JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES
www.journalimcms.org

ISSN (Online) : 2454 -7190 Vol.-19, No.-2, February (2024) pp 39-65 ISSN (Print) 0973-8975

AI FOR INFANT WELL-BEING: ADVANCED TECHNIQUES IN CRY INTERPRETATION AND MONITORING

Ananjan Maiti[1], Chiranjib Dutta[2], Jyoti Sekhar Banerjee[3] Panagiotis Sarigiannidis[4]

[1]Dept. of CSE, Guru Nanak Institute of Technology, Kolkata, INDIA-700114

[2]Dept. of CA, Guru Nanak Institute of Technology, Kolkata, INDIA-700114

[3]Dept. of CSE (AI & ML), Bengal Institute of Technology, Kolkata, INDIA-700150

[4]Dept. of Electrical and Computer Engineering, University of Western Macedonia, Kozani, Greece

Corresponding Author: **Jyoti Sekhar Banerjee**

Email: tojyoti2001@yahoo.co.in

https://doi.org/10.26782/jmcms.2024.02.00003

(Received: December 16, 2023; Revised: January 21, 2024; Accepted: February 7, 2024)

Abstract

In order to improve the welfare of newborns, this study investigates the use of sound-recognition-based artificial intelligence (AI) approaches to the interpretation and monitoring of infant screams. Crying has long been a problem because it is the primary means of communication between infants and caregivers. The limitations of conventional interpretation techniques are discussed. These limitations include the subjective nature of interpretation and the inability to detect subtle variations in crying patterns. The goal of the research is to categorize crying patterns based on the cries of male and female infants and identify noises that are a sign of distress. The study utilized the Mel Frequency Cepstral Coefficients (MFCC) method to extract features from internet-sourced MP3 and WAV audio data. The technique successfully captured the unique qualities of each crying sound using various machine-learning models, including Random Forest and XGBoost. These models outperformed others with accuracy rates of 94.5% and 94.2%, respectively. These findings show how well these algorithms perform in correctly categorizing various newborn cries. The findings of this study establish the platform for possible Internet of Things (IoT) and healthcare framework implementations targeted at supporting parents in caring for their newborns by offering an insightful understanding of the distinctive vocalizations connected with weeping.

Keywords: infant cry interpretation, machine learning, artificial intelligence, infant monitoring, real-time systems, privacy concerns, XGBoost.

I. Introduction

Infant crying is their main form of communication and frequently indicates discomfort, hunger, or a desire for attention. Caregivers may have trouble figuring out what their newborn is trying to say, leading to a sluggish or misguided response. Artificial intelligence (AI) and machine learning (ML) are two fields of technology that have demonstrated potential approaches to simplify the interpretation of newborn cries, allowing for quicker diagnosis and more effective intervention. Classifying infant screams is a crucial topic of research. Using spectral pictures and hand-crafted feature sets, Ozseven (2023) [LIX] successfully classified infant cries by utilizing deep neural network models and hand-crafted features. Similar to this, Srinivasa et al. (2022) [XXXV] used artificial neural networks (ANNs) to construct an autonomous infant cry speech recognition system with an average accuracy of 94.77%. By examining cry signals, Khalilzad et al. (2022) [LXIV] concentrated on discriminating between sepsis and respiratory distress syndrome (RDS) in neonates. They employed machine learning classifiers to achieve accuracy levels of over 92%. These improvements do not, however, come without difficulties. To improve comprehension and interpretation of infant cry signals, Ji, C., Mudiyanselage, T. B., Gao, & Pan (2021) [XIII] emphasized the need for additional study in data processing, feature extraction, and neural network classification. Due to the potential for medical and societal uses of infant screams, Liu et al. (2019) [XXXVIII] stressed the significance of additional research into child cries. Improved accuracy and dependability in automated pain evaluation in newborns are needed, according to Zamzmi et al. (2017) [XX]. AI is essential for deciphering and observing newborn screams. The accuracy of cry interpretation can be significantly enhanced through the implementation of AI adaptability and personalization, as stressed by Zhong et al. (2021) [XLII]. Similar ideas were expressed by Kim et al. (2018) [XL], who highlighted AI's potential for accurately and effectively evaluating huge amounts of data. AI-enabled behavioral analysis can offer insightful information on the infant's mental state, enabling early detection and intervention. The study uses machine learning and artificial intelligence to understand baby screams using Mel Frequency Cepstral Coefficients (MFCCs). It classifies data into unique groups of crying sounds using various classifiers, including SVM, random forests, gradient boosting, XGBoost, and neural networks. This device has some basic hardware requirements, which include a sound-sensitive microphone module, a Raspberry Pi, memory support for an SD card, a bank of a power source, an end-to-end internet connection, and a Python-based audio processing library. The result shows that the Random Forest and XGBoost models showed high classification accuracy, with accuracy of 94.5% and 94.2%, respectively. The major contributions of this study are as follows: (1) The study advances automated baby cry categorization by utilizing machine learning and artificial intelligence to develop precise and effective systems for deciphering and classifying newborn vocalizations. (2) The research aims to increase the well-being of newborns by understanding infant cries without any human intervention. By successfully categorizing extreme distress and caregiving reactions, this study lays the groundwork for better health outcomes for infants. (3) Here, this employs various machine learning models, including Random Forest, Gradient Boosting, Support Vector Machine, Nearest Neighbors (KNN), Linear Discriminant Analysis (LDA), Neural Networks, and XGBoost. The research demonstrates their capacity to make valuable contributions to the domain of infant cry interpretation and monitoring. (4) The study utilizes the MFCC feature extraction from cry sounds. This enables a thorough investigation of the auditory-related information linked to various styles of distress-related vocal responses. (5) Incorporating the concept of the Internet of Things (IoT) and healthcare frameworks could lead to improvements in neonatal care and a better comprehension of early childhood development.

The work provided in this paper shows how artificial intelligence, machine learning, and healthcare can be used to address neonates' well-being. A fundamental problem in caregiving and parenting is addressed by the research by concentrating on the interpretation and monitoring of infant screams. The well-being of newborns and a child's general development can be greatly impacted by the capacity to correctly classify and distinguish various sorts of screams. The proposed framework has the potential to have a substantial impact on infants and their carers' lives, which is made more noble by its unique approach, use of cutting-edge technology, and promising outcomes.

II. Background Studies on Infant Cries

Traditional methods for interpreting infant cries have been found to be often labor-intensive and unreliable, primarily depending on subjective interpretation (Kim et al., 2021 [XLVI]; Kumari et al., 2017 [XXI]; Chen et al., 2020 [XXII]). In contrast, AI-based approaches have paved the way for more objective, efficient, and precise analysis, enabling automated interpretation, time-saving, and increased support for caregivers. Such approaches facilitate early identification, reduce the chances of misinterpretation, and enhance the emotional bond between infants and their caregivers. They also offer contextual analysis and personalization to capture individual variations and external influences (Barr et al., 2021 [XXVII]; Srinivasa et al., 2018 [XXXV]). Among several deep learning methods employed for understanding baby cries, Convolutional Neural Networks (CNN) and Long-Short Term Memory (LSTM) have demonstrated significant success. They offer high recognition accuracy, providing reliable responses (Liang et al., 2022 [LXIII]; Lahmiri et al., 2022 [XLVII]). These studies have underscored the potential of AI in early health issue detection, specific infant needs recognition, and support for caregivers. AI and machine learning have been integrated into various devices and systems, such as baby monitors and infant incubators (Matikolaie et al., 2022

[XVIII]; Sutanto et al., 2020 [XVII]; Khan, 2021 [LX]). In these applications, AI proves to be instrumental in understanding newborn behaviors, needs, and health moving forward. Hence, the adoption of AI in interpreting infant cries is gaining traction due to its potential to enhance the understanding of infants' emotional and physical health, aid caregivers' responses, and foster more personalized and effective care.

Table. 1: Related studies regarding Infant Cry

Study	Citation	Method	Results
Kim et al., 2021 [XLVI]	Traditional methods of interpretation are limited due to their subjective nature.	AI-based approach for automated interpretation.	Improvements in terms of accuracy, efficiency, and time-saving.
Kumari et al., 2017 [XXI]	Traditional methods may be comparatively less accurate	AI-based approach for automated interpretation.	Time-saving, improved caregiver support, early detection, and intervention in cases of infant distress.
Chen et al., 2018 [XXII]	Traditional methods of cry interpretation are subjective and time-consuming.	AI-based approach using machine learning algorithms.	Improvements in accuracy, efficiency, time-saving, early detection, and intervention.
Liang et al., 2022 [LXIII]	Cries hold valuable insights into infants' needs.	Deep learning algorithms for cry recognition.	CNN and LSTM showed 95% accuracy, precision, and recall in differentiating healthy and sick infants.
Lahmiri et al., 2022 [XLVII]	Cry records provide valuable health insights.	Multiple deep learning systems to improve diagnosis.	CNN achieved the highest accuracy and sensitivity in differentiating between healthy and unhealthy cries.
Matikolaie and Tadj, 2022 [XVIII]	Need for accurate analysis of cry audio signals.	Machine learning approaches to identify septic newborns from healthy ones.	Achieved an 86% F-score for expiration and 83.90% for inspiration.

Infant vocalizations or crying serve as the primary communication mode for these pre-lingual organisms. A wealth of information about an infant's physical and emotional health could be unearthed by interpreting these cries, and this is an area of avid interest to global researchers. Traditional interpretive methods, however, fall prey to errors and inconsistency due to the constraints of subjective analysis and labor-intensive procedures. Understanding the influences on infant cry patterns can offer valuable insights into a child's well-being. AI-based monitoring and interpretation of these patterns can aid caretakers in timely identification and intervention of potential health issues (Ricci et al., 2020 [XV]; Barr et al., 2021 [XXVII]; Srinivasa et al., 2018 [XXXV]). Modern research is focusing on using advanced AI, particularly deep learning, to improve baby cry interpretation. Modern machine learning algorithms and signal processing methods can help figure out the meanings of different types of cries, providing a fair, useful, and objective platform (Liang et al., 2022 [LXIII]; Lahmiri et al., 2022 [XLVII]; Matikolaie and Tadj, 2022 [XVI]; Sutanto et al., 2021 [XVII]; Khan, 2021 [LX]; Bashiri and Hosseinkhani, 2020 [I]; Hussain et al., 2019 [LXI]). This method can be greatly applicable to wearable devices to enhance their accessibility and usability for parents and healthcare professionals alike.

Deep learning techniques like artificial neural networks, multi-layer perceptrons, convolutional neural networks, and LSTM have been used to categorize baby crying, with methods like Mel-Frequency Cepstral Coefficients and cepstrum analysis-based coefficients achieving high accuracy rates. The rise of IoT and smart devices has supported the development of efficient baby cry monitoring and recognition systems, offering parents a non-intrusive way to monitor their infants and reduce caregiver stress. These methods have achieved high accuracy, up to 95%, in distinguishing between healthy and sick infants and recognizing specific infant needs (Liang et al., 2022 [LXIII]; Lahmiri et al., 2022 [XLVII]; Matikolaie and Tadj, 2022 [XVI]; Bashiri and Hosseinkhani, 2020 [I]. The rise of IoT and smart devices has greatly supported the development of efficient baby cry monitoring and recognition systems (Sutanto et al., 2021 [XVII]; Khan, 2021 [LX]; Hussain et al., 2019 [LXI]).

Future research is called upon to improve data quality, manage the data collection environment, and explore other feature extraction methods for comparison with the current ones. Strategies such as the K-Nearest Neighbor (K-NN) algorithm, Linear Frequency Cepstral Coefficient (LFCC), and convolutional neural network (CNN) have also shown significant promise in detecting and classifying baby cries with high accuracy, thereby offering potential benefits to infant care (Dewi et al., 2019 [XLVIII]; Zhang et al., 2018 [LXII]). For instance, Matikolaie et al.'s, 2022 [XVI] method combined probabilistic neural networks and support vector machine algorithms, demonstrating the importance of machine learning in neonate care. Furthermore, addressing pediatric feeding issues is critical due to their potential severe and costly impacts, as underscored by Babbitt et al. [XLIV] AI-based methods for cry interpretation have been instrumental in enhancing accuracy and efficiency, thereby enabling early detection and intervention. This improvement aids in bolstering caregiver support while minimizing subjectivity and inconsistency.

The creation of a newborn cry corpus is vital for research as it can assist in identifying infant needs and potential health issues. The factors to consider while creating such a corpus include the origin of the cry, the infant's age, health, and weight, as discussed by Chittora et al. (2017) [II]. Lyså et al. (2021) [XXVI] provide an ethnographic study on frustration education in urban China amidst the socio-cultural shifts influenced by the one-child policy. Meanwhile, Rosales-Pérez et al. (2015) [III] delve into the rising importance of AI as a strategic factor for national development, stressing the need for responsible governance frameworks, especially in relation to the UN's Sustainable Development Goals (SDGs). For effective newborn cry interpretation and monitoring technology, hardware infrastructure plays a crucial role. Computational bases such as Raspberry Pi microcontrollers ensure efficient operations, including data management, running machine learning algorithms, and networking. Specialized equipment like the CSL 4400 used by Wermke et al. (2009) [XXXIII] facilitates the detailed analysis necessary to draw correlations between newborn cry melodies and language development. In addition, wearable devices and integrated smart monitoring tools stand as intriguing alternatives. Pan et al. (2018) [XII] successfully demonstrated the viability of integrating machine learning methods into wearable devices using technologies like Deep Neural Networks (DNN), Gaussian Mixture Models (GMM), Hidden Markov Models (HMM), and Digital Signal Processing (DSP). Notably, Tareq Khan's smart alert system showcases the use of NVIDIA's Jetson Nano microcontroller in a smart baby monitor capable of identifying potentially dangerous situations and alerting caregivers. This exemplifies the need for potent processors and essential peripherals for data collection, processing, and transmission in AI-based baby monitoring systems. The choice of hardware can vary based on the complexity and intended use of the system.

Table. 2. Hardware Requirements and Used in Infant Monitoring Systems

Study	Hardware Used	Applications
Wermke et al., 2009 [XXXIII]	CSL 4400	Analyzing infants' cry melody complexity
Pan et al., 2018 [XII]	STM32 microcontroller, DSP technology, Bluetooth 4.0BLE technology	Recognizing baby crying and implementing intelligent monitoring
Tareq Khan [LX]	NVIDIA's Jetson Nano mic controller, night vision camera, Wi-Fi connectivity	Detecting potentially dangerous situations and sending alerts to caregivers

III. Proposed Methodology

The purpose of the methodology section of this paper is to show how AI was used to recognize and categorize infant screams. The objective is to use machine learning approaches to distinguish between distinct infant vocalizations, with a focus on identifying various cries [IV-V]. Potential improvements in the healthcare and parental support systems are provided by this procedure. The study mainly makes use of a collection of baby cry sounds that were gathered from several web sources. A wide variety of physiological variants are represented by this thorough compilation, providing a more complete and representative set of data for the study. The average length of each sound clip in the database is 20 seconds, ensuring that there is sufficient sample size for analysis while also allowing for quick processing [VI-VII]. The study uses the MFCC feature extraction technique to analyze the cry sounds. This method makes it possible to recognize and record the unique qualities that each cry sound possesses [VIII]. These characteristics offer a “fingerprint” of the cry that machine learning algorithms can utilize for training and categorization. The study includes extensive use of several machine learning models, such as ensemble methods like gradient boosting, extreme gradient boosting (XGBoost), supervised random forest, support vector classification, the K-nearest neighbors algorithm, linear discriminant analysis (LDA), neural networks, etc. The effectiveness of these advanced machine learning models was measured with different levels of categorization for various criteria. It highlights the importance of assessing their effectiveness in real-world healthcare and parental aid frameworks.

III.i. Step-by-Step Approach of the Suggested Framework

In the investigation of infant cries, audio data was initially collected from online sources in MP3 and WAV formats. This data underwent preprocessing activities such as format standardization, volume normalization, and noise removal to make it suitable for further analysis [XXIII]. The Mel Frequency Cepstral Coefficients (MFCC) method was used to measure how well the cry sounds were broken down into meaningful parts, with each cry being represented by its own unique set of numbers. These features were used to train multiple machine learning models, such as ensemble methods like gradient boosting, extreme gradient boosting (XGBoost), supervised random forest, support vector classification, the K-nearest neighbors algorithm, linear discriminant analysis (LDA), neural networks, etc. The effectiveness of these advanced machine learning models was measured with different levels of categorization for various criteria, which have been aiming to classify the cries into various categories such as male versus female and normal versus extreme distress. Upon evaluating these models, Random Forest and XGBoost outperformed the rest. The most effective model(s) were then deployed in an Internet of Things (IoT) or healthcare framework, creating an interface for the model to receive new cry sounds and return predicted categories. This system (see Figure 1) is designed for end-user interpretation where parents or caregivers could interpret their baby’s cries via a user-friendly interface such as a smartphone app or a display on a baby monitor device.

III.ii. Cry Audio Dataset

In our AI-focused research on the interpretation and surveillance of infant sounds, a key aspect was gathering a reliable, relevant, and robust database of infant cries. The search for this data commenced with Google Scholar, a well-established academic search engine. Our hunt was focused within a particular time range and utilized precise keywords, such as “baby cry analysis”, “infant vocalization”, and “acoustic analysis of infant cries”, to sift through the abundance of literature available. The next step involved a scrutiny of the identified research papers. We examined each one, concentrating primarily on the methodologies used and the sources cited for infant cry data (see Table 3). This process unveiled a collection of online resources hosting a variety of baby-crying sounds, primarily https://www.freesound.org and https://www.soundsnap.com [XXIV]. These resources served as the raw material, enabling us to construct a diverse and representative dataset of infant cries to aid our AI-centric research.

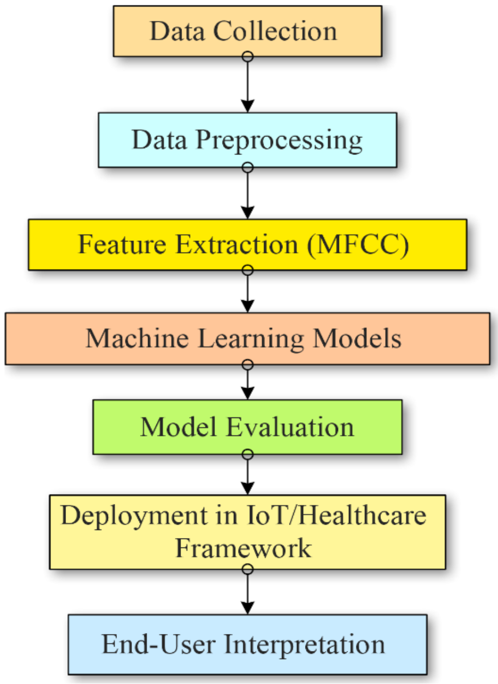

Fig. 1. Step-by-Step Approach Towards Proposed System

The collected audio files had an average duration of 20 seconds each. While accumulating the data was critical, it was equally vital to refine it to enhance its relevance and specificity. We used sound converters to modify the audio files into a format compatible with our analytical tools. Moreover, sound cutters were deployed to extract specific sections from longer recordings, focusing our data on distinct instances of infant crying. These technologies streamlined our data collection process and helped fine-tune the dataset to match our research needs. This thorough method ensured our data was ready for analysis and future AI modelling. We, thus, created a comprehensive dataset of infant cries, an invaluable asset for ongoing and future research in this field. Our refined dataset, filled with a broad spectrum of crying patterns and high-quality audio, paves the way for extensive studies on infant vocalizations and acoustic analysis. This systematic, meticulous approach highlights the crucial role of data collection in research and affirms the integrity and reliability of our dataset, specifically designed for AI-driven analysis of infant cries. Consequently, we have detailed a comprehensive and effective methodology for data collection and refinement in this critical research area.

Table. 3: Crying Sound Types with duration of the Sounds (in seconds)

Crying Sound Types	Duration of Sounds Present (in seconds)
Boy Crying sounds	109
Girl Crying sounds	128
Hard Crying sounds	120
Hiccups sound	40
Total	397

III.iii. Preprocessing and Feature Extraction

Cry audio data is processed to improve the extraction of MFCC features. This process involves adding external noise, using filters, and lowering noise. First, the original audio data is contaminated by outside noise [IX]. It is then added after being repeated until it is the same length as the original audio data [XLIX, XVI]. The following stage involves applying several noise-reduction strategies and filters to the noisy audio data:

The low-frequency noise is eliminated using a high-pass filter. Signals can only flow through this filter if their frequency is greater than a specific cutoff frequency. Half of the audio data's sampling rate, or the Nyquist frequency, is used to transform the cutoff frequency into this frequency. Then, this filtered data is kept in a different location. The data that has been high-pass filtered is then given a low-pass filter. In contrast to a high-pass filter, this filter only lets through sounds that have a frequency lower than a specific cutoff frequency. The cutoff frequency is likewise reduced to a portion of the Nyquist frequency in a similar manner. The data that goes through two filters is then kept apart. The high-pass filtered data is then effectively subtracted from the low-pass filtered data by a complementary filter. This method can be used to enhance specific frequency bands, such as vocals in audio data. The complementary filter's output is subjected to median filtering. This filtering method aids in removing brief noise bursts from audio data.

Finally, spectral subtraction is used to further reduce noise. To determine the strength and phase of the audio signal, one must compute the Short-Time Fourier Transform (STFT) of the audio data. The magnitude is then multiplied by a complex number produced by Euler's formula using the phase as the argument, which is then subtracted by the magnitude's median. The inverse STFT is then used to transform the outcome back to a time-domain signal. To enable more successful extraction of the MFCC characteristics, this audio preprocessing pipeline helps to refine the audio data by adding controlled noise and then using several filters and noise reduction algorithms. The steps are given below:

1. Initialization: Let LL be the set of required libraries: L={librosa,matplotlib.pyplot,numpy,pywt}

2. Load Audio: Given an MP3 audio file path PP, load the audio signal using the librosa function: Y,S=librosa.load(P) Where:

- YY is the audio time series.
- SS is the sampling rate.

3. Add Noise: Generate random noise NN of the same length as YY: N=numpy.random.randn(len(Y)) Scale the noise by a factor ff (e.g., 0.005) and add to the original audio signal to produce the noisy audio: Ynoisy=Y+f·N

4. Denoise Audio:

a. Decompose the noisy audio signal into wavelet coefficients: C=pywt.wavedec(Ynoisy,'db1',level=6)C=pywt.wavedec(Ynoisy,'db1',level=6) Where CC is a list of wavelet coefficients at different levels.
b. For the last two levels of the wavelet coefficients, determine a threshold value and threshold the coefficients:

Ti=0.5×numpy.std(C[i])Ti=0.5×numpy.std(C[i]) C[i]=pywt.threshold(C[i],Ti)

For i=len(C)−1i=len(C)−1 and i=len(C)−2i=len(C)−2.

c. Reconstruct the denoised audio signal from the thresholded coefficients: Ydenoised=pywt.waverec(C,'db1')

Noise contamination, especially from sources like baby cry sounds and environmental sounds, is a pervasive issue in audio signal processing, affecting clarity and causing potential misinterpretations. The proposed noise-removal method offers a systematic solution for addressing this problem, emphasizing the utilization of wavelet transformations. To begin, the necessary libraries for this process are identified, namely: librosa, for audio handling; matplotlib.pyplot, for visualization; numpy, for numerical operations; and pywt, vital for wavelet processing. The audio data, which could include recordings of baby cry sounds and environment sounds, is sourced from an MP3 file, and, using librosa, is loaded to retrieve the audio time series and its associated sampling rate. An inherent challenge in audio signal processing is simulating real-world noise conditions, such as the ambient noise from environmental sounds or the distinct interruption of a baby crying. In this procedure, artificial noise is added to the original audio signal, providing a controlled environment to test the denoising technique. This is achieved by generating random noise, scaling it down, and then combining it with the original signal, resulting in a deliberately 'contaminated' audio. The core of this method hinges on the denoising process. Wavelet transformation, a mathematical tool adept at capturing signal irregularities, is employed to decompose the noisy audio signal into wavelet coefficients. These coefficients offer multi-resolution representations of the audio signal, making them invaluable for detecting noise at different scales. The method then focuses on the coefficients from the last two levels, which are most likely to contain noise. By setting a threshold derived from the standard deviation of these coefficients, noise is identified and nullified. Subsequently, the cleaned audio signal is reconstructed from these processed coefficients. In essence, this technique provides a robust mechanism to purify audio signals, harnessing the mathematical prowess of wavelet transformations. Such methods are integral for enhancing audio clarity in various applications, from medical diagnostics to voice assistants, ensuring that disruptions like baby cries or environmental sounds are minimized.

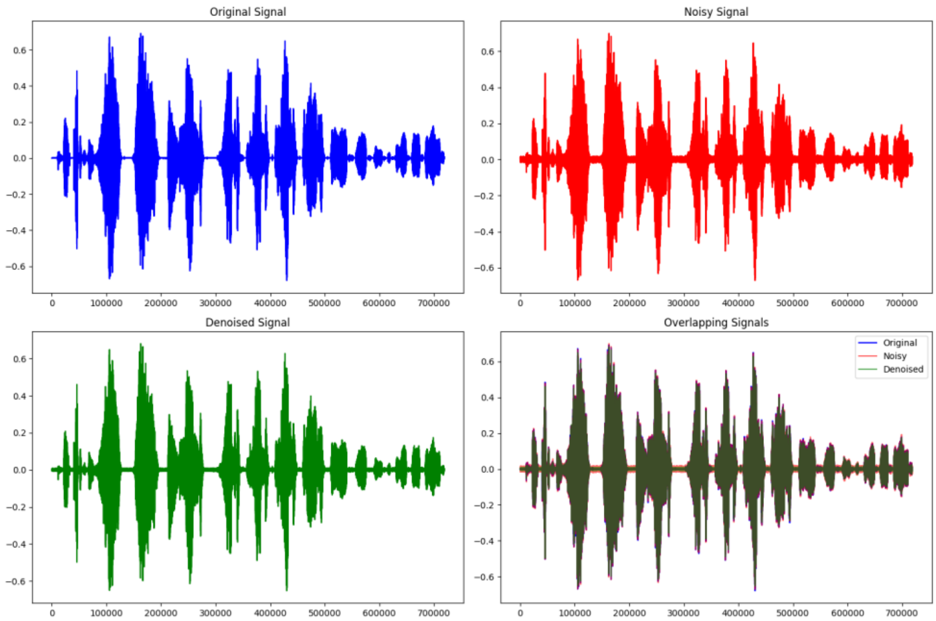

Fig. 2. Baby Cry Original Signal (blue), Noisy signal (red), Removed noise (green), and Overlapping signal (grey)

This study relies heavily on Mel Frequency Cepstral Coefficients (MFCCs), designed specifically for the analysis and interpretation of newborn cries using artificial intelligence (AI). Owing to their ability to capture unique sound characteristics and effectively mimic the human auditory system's perception and interpretation of audio signals, MFCCs are a well-established feature in speech and audio analysis. Using the MFCC feature extraction technique, the study can extract vital information from a vast database of infant cries. Infant cries are complex vocalizations conveying a spectrum of emotions and needs. Subtle variations in pitch, volume, and tone of these cries could potentially provide significant insights about the infant's state. In this context, MFCCs offer their unique value. MFCCs process these cries and extract useful features, highlighting the underlying patterns in the data. They accomplish this by focusing on the frequency bands most perceptible to the human ear, and the cepstral coefficients capture the spectral shape, which is crucial for differentiating between various sound sources. In this study, the utilization of MFCCs enables the transformation of raw audio data into a more comprehensible format that AI and machine learning models can understand and learn from. By converting the cries into a set of MFCCs, the study effectively creates a 'fingerprint' for each cry, encapsulating its unique attributes. These fingerprints are then used to train machine learning models, to distinguish between different types of cries, such as those indicative of hunger, pain, or discomfort. In essence, MFCCs provide a crucial bridge between the world of raw audio data and AI and machine learning. By leveraging these features, this study is able to delve into the rich, complex information contained within infant cries, translating it into actionable insights that could significantly contribute to infant welfare and parental assistance.

III.iv. Applied Machine Learning Algorithms

Classifiers, which sort data points into distinct groups or categories, are an essential component of Machine Learning models [XIV, XXXIV]. In essence, classification algorithms try to determine which class or category an input belongs to given the input.

1. SVM (Support Vector Machine): This machine learning model is strong and versatile, and it can do linear or non-linear classification. It looks for the best hyperplane to use for separating all of the data points from one class from those from the other. SVMs operate by raising the problem dimension to a point where hyperplanes can be used to divide the data into several categories.
2. Random Forests: This ensemble learning technique combines several decision trees, which are weak learners, to create a powerful prediction model. Random Forests tend to increase prediction accuracy and reduce over-fitting by merging many decision trees to provide predictions.
3. "Gradient Boosting" and "XGBoost" are two other ensemble machine learning techniques that provide reliable predictive models. The main principle of boosting is to sequentially add new models to the ensemble. Each additional model eventually reduces the system's overall loss function. Gradient boosting decision tree systems are noted for their performance and speed, and XGBoost is an improved and customized version of one such system.
4. K-Nearest Neighbors (KNN): This non-parametric technique is utilized for regression and classification. An object is allocated to the class most frequently chosen by its k closest neighbors based on a majority vote among the item's neighbors.

5. Linear Discriminant Analysis (LDA): This technique is used to identify a linear combination of features that distinguishes between two or more classes of objects or events. A linear classifier may be created using the resulting combination.
6. Neural Networks: The brains of deep learning algorithms, these are a subset of machine learning. They are made to recognize patterns and are loosely modelled after the human brain. They label or cluster raw input to interpret sensory data using a form of machine perception.

Each of these classifiers has benefits and drawbacks, and performance may vary significantly depending on the dataset and task at hand [XXXIX]. As a consequence, the project's specific needs should always be taken into consideration while choosing a classifier.

III.v. Proposed Hardware for AI-based Cry Detection

The design of a Raspberry Pi-based IoT device for infant cry detection would have several components, each with a specific function for recording, processing, and categorizing infant cries in accordance with the machine-learning approaches covered in the study [XIX, XXV] (see table 4). The input interface, the Microphone Module, would be in charge of precisely recording baby cries. It would catch the minute variations in cries, which may signal different requirements or potential health problems. The Raspberry Pi would serve as the system's central processing unit, with the Raspberry Pi 4 Model B being the most highly suggested model. Using machine learning models [X, XI, XIV, XXVIII-XXXII, XXXVI-XXXVII, XLI, XLIII, XLV, L-LVIII] like Random Forest or XGBoost, which have proven to be very accurate at classifying various sorts of infant cries, it would perform data processing chores. The operating system, Python scripts containing the machine learning models, and the gathered cry sound files would all need to be stored on an SD Card. Its storage space must be large enough to accommodate the projected volume of data. The power supply is yet another essential component that provides the Raspberry Pi with the energy it needs to function. A Micro USB power source with at least 2.5A of power is recommended to ensure reliable performance. An Internet connection would be necessary for the data transmission and result distribution. Depending on the Raspberry Pi model being utilized, either an Ethernet connection or a Wi-Fi module could be used to establish this. To process the audio data, an Audio Processing Library is required. The Mel Frequency Cepstral Coefficients (MFCC) features, which are essential for describing and categorizing the infant cries, can be extracted in Python using libraries like Librosa. As a result, each of these elements helps a Raspberry Pi-based IoT device for infant cry detection run efficiently while utilizing the power of machine learning to reveal important information about child wellbeing.

Table. 4: Hardware requirements for implementing the proposed model

Component	Description
Microphone Module	This is used to record the infant's cry sounds. It should be capable of capturing clear sound within the expected range of the infant cry.
Raspberry Pi	This is the central processing unit where data is processed and the machine learning models are run. A Raspberry Pi 4 Model B is recommended for its superior processing capabilities.
SD Card	This is used to store the operating system, Python scripts for the models, and the cry sound files. An SD card with high data write speed is recommended.
Power Supply	The Raspberry Pi requires a power supply to function. A Micro USB power supply with at least 2.5A power is needed.
Internet Connection	This is necessary for transmitting the data and results. This could be achieved through an Ethernet connection or a Wi-Fi module if it is not already built-in in your Raspberry Pi.
Audio Processing Library	A software library for processing audio data is required. For Python, libraries like Librosa can be used for MFCC feature extraction.

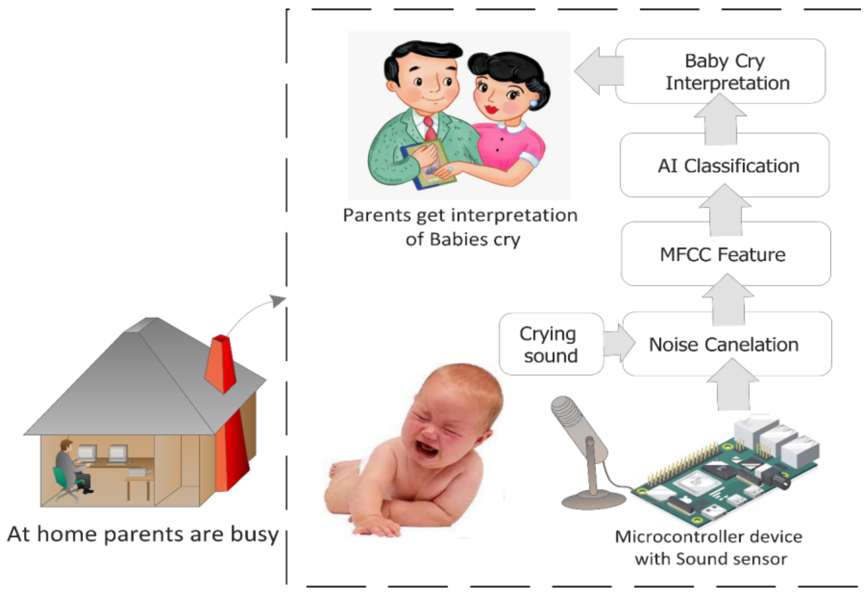

Fig. 3. System Landscape: How baby cry will be interpreted

In Fig. 3, it is seen that the proposed innovative cry interpretation device will function as a reliable tool for infant care. This advanced technology is designed with the capability to isolate the distinctive sound of a baby's cry, eliminating any background noise that might interfere.

(1) It is seen that parents often struggle to respond to their infant's needs due to some distractions from household duties at home.
(2) The baby’s cry pattern is a crucial signal that may indicate its needs or discomforts, such as hunger, fatigue, or illness.
(3) An IoT device with a sound sensor captures a baby’s cry, typically using an analog-to-digital converter for data conversion.
(4) This technique effectively cancels or reduces background noise, such as TV sounds or conversations, to enhance the distinctness and understanding of a baby's scream.
(5) MFCC features as a part of signal processing to extract valuable distinctive features from the sound of the baby’s cry. The pitch, volume, and timbre could provide clues about the baby’s emotional and mental state at that time.
(6) Advanced AI models are used to analyze the MFCC features and classify the baby’s cry into different categories. These categories include various needs or states like discomfort, hiccups, etc.
(7) A set of predefined rules that map the baby’s cry is then interpreted, with each category to a specific mental and physical state.
(8) Finally, the interpretation is relayed to the parents, perhaps through an app on their phone or a display on the device itself. The interpretation can then help the parents understand why their baby is crying and respond appropriately.

V. Results and Discussion

In this infant cry recognition study, various machine-learning techniques were used to categorize different types of baby cries. To train and evaluate these models, Mel Frequency Cepstral Coefficient (MFCC) features extracted from baby cry audio data were used.
The results show that the effectiveness of the different models varies greatly (see Figure 4). The Random Forest algorithm, the best method, achieved a remarkable accuracy of 94.5%. Like the XGBoost algorithm, Random Forest is known for its high accuracy and robustness when processing complex data sets. This algorithm is an ensemble learning algorithm that combines multiple decision trees to make more accurate predictions. The challenge was to achieve the complexity of the task without overfitting, which is sometimes a concern in machine learning tasks.

In second place in terms of accuracy rating is Extreme Gradient Boosting (XGBoost) with 94.2%. This shows that this model was best able to accurately classify the different types of screams. The Gradient Boosting framework used by XGBoost is responsible for the accuracy of the algorithm, which creates powerful predictive models by merging several weak models. Gradient Boosting achieved an accuracy of 78.2%, while the K-Nearest Neighbors model showed an increase. The performance of neural networks was 75.5%, while SVC and LDA models performed the worst due to their complexity and high dimensionality.

The results shed light on the possibility of machine learning to detect baby cries, with Random Forest and XGBoost showing particularly encouraging results. It also emphasizes the importance of selecting the appropriate model for the task at hand and the available data. Each classifier in the study would have a unique confusion matrix related to infant cry recognition (see Figures 5 to 10). This matrix could be divided into four types: True Positives (TP), where the model correctly predicted a particular type of positive class. True Negatives (TN), where the model correctly predicted the absence of a particular type of scream, namely the negative class. The model misclassified False Negatives (FN) as negative and False Positives (FP) as positive.

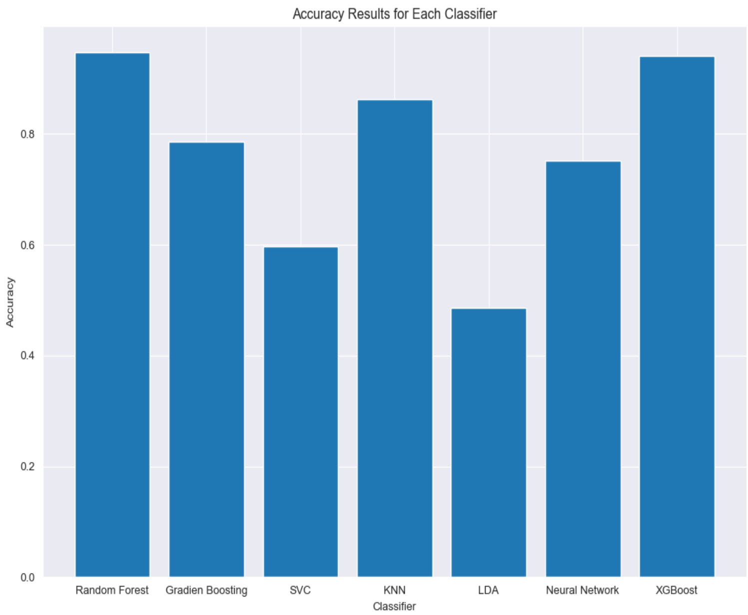

Fig. 4. Accuracy Results for each Applied Classifier

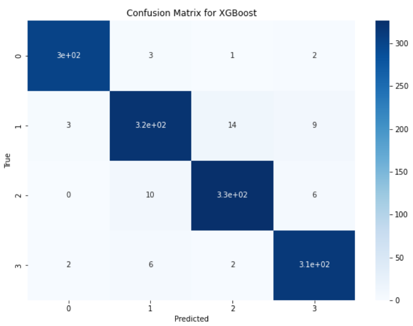

Fig. 5. Confusion matrix of XGBoost

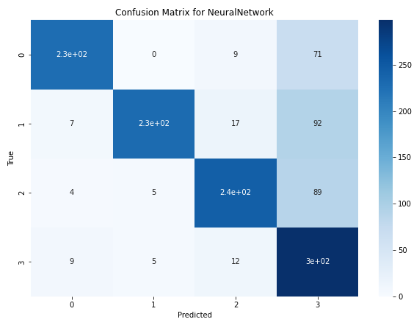

Fig. 6. Confusion matrix of Neural Network

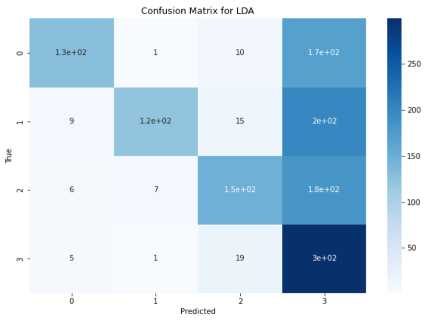

Fig. 7. Confusion Matrix of LDA

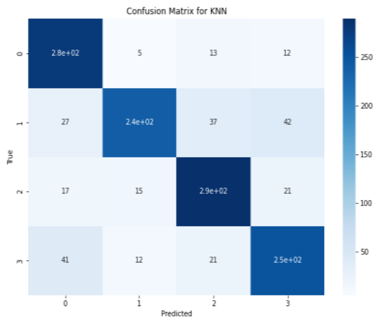

Fig. 8. Confusion Matrix of KNN

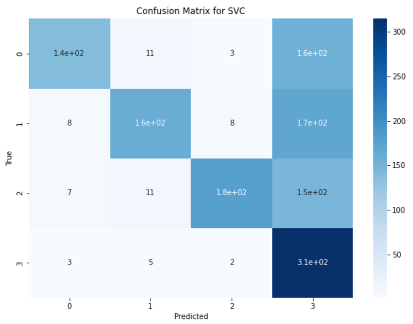

Fig. 9. Confusion Matrix of SVC

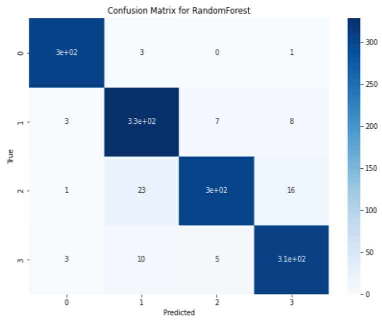

Fig. 10. Confusion Matrix of Random Forest

The results in Table 5 and Figures 5 to 11 show that Random Forest is the best algorithm in terms of precision, recall, F1, and AUC-ROC and has quite high precision rates (94.6%) and recall (94.5%) offers. This simply means that the model can correctly classify the output with good balance, true positives, and false positives. With a high F1 score (94.4%), the Random Forest also shows great skill in harmonizing precision and recall. In addition, it also shows high-class separation capabilities, as shown by the almost perfect AUC-ROC (99.5%). XGBoost comes as a reasonably competent second-best performer. It falls behind Random Forest, but still the efficacy it shows is respectably good, with precision, recall, and F1 score all around 94.2%, along with a strong AUC-ROC of 99.3%. This signals good classification altogether, including the especially good class separation ability. KNN is the third best by accuracy but still lower than the first two algorithms. The low value in terms of recall (86.3%) and F1 score (86.2%) is a bad signal that it may not have a very good ability in the capability to capture all positive cases and it may penalize its balanced classification capacity. Gradient Boosting achieved slightly lower scores than KNN with an accuracy of 78.2%, a precision of 78.5%, a recall of 78.2%, an F1 of 78.2%, and an AUC-ROC of 94.6%. Although these values are lower than KNN, they are still good, showing that gradient boosting was able to model the dataset effectively. Neural networks achieved 75.5% accuracy, 82.5% precision, 75.5% recall, 76.3% F1 score, and 93.9 AUC-ROC. The neural network had higher precision but lower recall compared to gradient boosting, suggesting that it was better at predicting the positive class but worse at detecting the negative class.

SVC achieved significantly worse results than the other algorithms. Accuracy was 59.6%, precision 73.6%, recall 59.6%, F1 60.3%, and AUC-ROC 82.7. The low values indicate that SVC had difficulty modelling the dataset effectively. The LDA algorithm had very poor precision and recall (68% and 48.3%), the lowest F1 score (47.8%) and AUC-ROC (77.4). In other words, there will be major deficiencies in the precise classification of cases and class separations, which is therefore highly likely. The form of this method does not seem suitable for scenarios such as this study.

Table 5: Summary of details of results of following matrices

Sl No.	Algorithm	Accuracy	Precision	Recall	F1 score	AUC-ROC
1	Random Forest	94.5	94.6	94.5	94.4	99.5
2	XGBoost	94.2	94.2	94.2	94.2	99.3
3	KNN	86.3	86.5	86.3	86.2	97.5
4	Gradient Boosting	78.2	78.5	78.2	78.2	94.6
5	Neural Network	75.5	82.5	75.5	76.3	93.9
6	SVC	59.6	73.6	59.6	60.3	82.7
7	LDA	48.3	68.0	48.3	47.8	77.4

In summary, Random Forest is the clear first choice, XGBoost, KNN and Gradient Boosting are reliable other options. Neural networks, SVC and LDA do not appear to be compatible with high-stakes classification. The results actually demonstrate that the selection of algorithms must be strongly focused on the precise requirements and constraints of the classification task. Random Forest is overwhelmingly preferred because of its robust, balanced classification and differentiation between classes.

The following procedures are used in this study to assess classifier performance using a macro-average Receiver Operating Characteristic (ROC) curve in Fig 11. Interpolation of ROC Curves: To provide consistency and to allow for average calculation, we first interpolate each ROC curve to a common grid of false positive rates, ranging from 0 to 1.

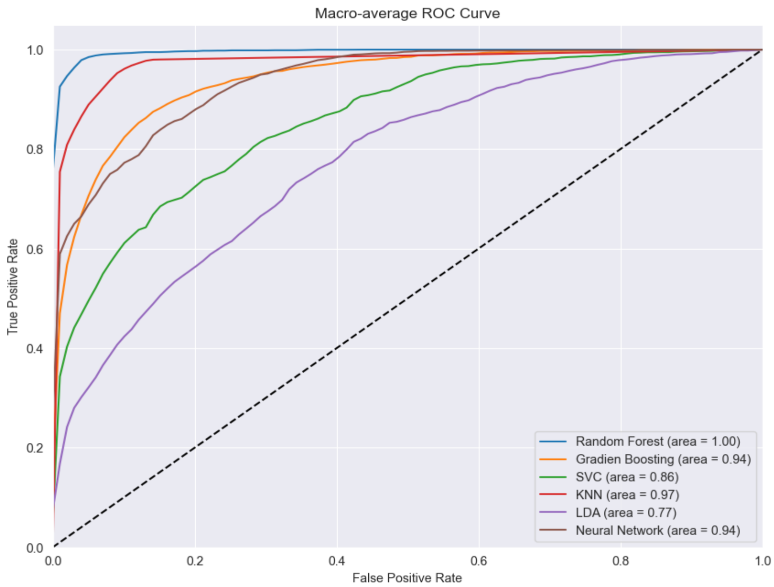

Fig. 11. ROC and AUC plot results to assess classifier performance

Calculation of Average True Positive Rate: We calculate the average true positive rate for all classes at each place on this common grid. This method makes sure that every class, regardless of size or distribution within the dataset, contributes equally to the final metric.

Plotting the Macro-average ROC Curve: We depict the macro-average ROC curve using the computed average true positive rates and the common grid of false positive rates. This curve offers a comprehensive assessment of the classifier's performance in multi-class settings by aggregating performance across all classes. This methodology is used in the study to give a thorough assessment of classifier performance that equally accounts for and represents each class in the dataset.

IV.i. Comparative Study

The field of infant cry interpretation has seen significant advancements. Dewi et al. (2019) [XLVIII] achieved 90% accuracy in cry detection using LFCC feature extraction and K-NN classification. Zhang et al. (2018) [LXII] differentiated between infant and non-infant crying with 86% accuracy using a lightweight CNN. Matikolaie et al. (2022) [XVI] used machine learning to identify unhealthy newborns, achieving an F-measure over 80%. Babbitt et al. (1994) [XLIV] successfully treated severe feeding disorders using behavioral analysis. The proposed method, which differentiates cries of male and female infants and identifies extreme distress screams, surpasses previous studies with a 94.5% accuracy using a Random Forest classifier. The outcome so far has been satisfactory as it has been seen which is better than previous studies (see table 6).

Table. 6: Comparing Efficiency Measures on Various Infant Cry Interpretation Methods

Sl No	Author	Features and AI/ML Algorithm Used	Outcome
1.	Srinivasa et al., 2018 [XXXV]	Artificial neural networks (ANNs).	autonomous infant cry speech recognition system with average accuracy of 94.77%.
2.	Khalilzad et al., 2022 [LXIV]	By examining cry signals concentrated on discriminating between sepsis and respiratory distress syndrome (RDS) in neonates.	Accuracy levels of over 92%.
3.	Liang et al., 2022 [LXIII]	Deep learning algorithms for cry recognition.	CNN and LSTM showed 95% accuracy, precision, and recall in differentiating healthy and sick infants.
4.	Lahmiri et al., 2022[XLVII]	Mel-Frequency Cepstral Coefficients (MFCC) and cepstrum analysis-based coefficients.	up to 95% and 64%, in distinguishing between healthy and sick infants, and recognising specific infant needs.
5.	Matikolaie et al., 2022 [XVIII]	MFCC and auditory-inspired amplitude modulation (AAM) features with traditional ML methods.	F-measure of over 80%.
6.	Dewi et al., 2019 [XLVIII]	LFCC feature extraction and K-NN classification.	90% accuracy in classifying whether a baby is crying or not.
7.	Zhang et al., 2018 [LXII]	Lightweight CNN in a client-server framework with a robot prototype.	86% detection accuracy rate in differentiating between infant crying and non-infant crying events.
8.	Proposed method	Detecting cries and separating the cries of male and female infants and identifying screams that are a sign of extreme distress.	Achieved 94.5 % accuracy using the Random Forest classifier.

V. Conclusion

The research presented in this study represents a pioneering advancement in the area of automated classification of infant cries. This work constitutes a noteworthy contribution to the development of healthcare and parenting support applications. A range of infant crying sounds, encompassing baby girl crying, baby boy crying, and hard, painful crying, were effectively categorized via a sequence of machine learning models. The aforementioned highlights the significance and variety of vocalizations associated with crying as a crucial component in the growth of children and the possibility for machine intelligence to comprehend and decipher these cues. The utilization of MP3 and WAV formats sourced from the internet yielded a strong foundation for analytical purposes. The process of transforming the sound involved the utilization of Mel Frequency Cepstral Coefficients (MFCC) feature extraction. This method is known for its effectiveness in capturing the distinct characteristics of individual crying sounds. The utilization of this methodology has facilitated a comprehensive analysis of the distinct auditory characteristics inherent in diverse forms of vocal expression associated with distress, thereby illuminating the intricacies involved in their comprehension. For the aim of categorization, several different machine learning models were used. Models included eXtreme Gradient Boosting (XGBoost), Linear Discriminant Analysis (LDA), Neural Networks, K-Nearest Neighbors (KNN), Support Vector Classification (SVC), Random Forest, and Gradient Boosting. The diverse range of models presented distinct viewpoints regarding the data, thereby showcasing the capacity of machine learning to make valuable contributions to this particular domain. The Random Forest and XGBoost algorithms demonstrated superior performance, attaining accuracy rates of 94.5% and 94.2% correspondingly. The aforementioned demonstration highlights the efficacy of ensemble learning techniques in addressing the complexities inherent in human auditory signals. The aforementioned models have demonstrated the capacity for reliable categorization of various types of infant cries, while also establishing a foundation for future progress in this domain. The implications of these findings are promising for future prospects. Enhanced comprehension of infant crying has the potential to facilitate more refined caregiving reactions, which may lead to better health outcomes for infants and alleviate the burden of caregivers. Moreover, the integration of these discoveries into practical implementations, such as medical facilities and resources for parental guidance, could potentially yield significant advancements in neonatal care and enhanced comprehension of initial childhood growth.

Conflict of Interest:

The author declares that there was no conflict of interest regarding this paper.

Reference

I. A. Bashiri and R. Hosseinkhani, : 'Infant crying classification by using genetic algorithm and artificial neural network.' Acta Medica Iranica, 531-539 (2020).

II. A. Chittora and H. A. Patil, : 'Data collection of infant cries for research and analysis.' Journal of Voice, 31(2), 252-e15 (2017).

III. A. Rosales-Pérez, C. A. Reyes-García, J. A. Gonzalez, O. F. Reyes-Galaviz, H. J. Escalante, and S. Orlandi, : 'Classifying infant cry patterns by the Genetic Selection of a Fuzzy Model.' Biomedical Signal Processing and Control, 17, 38-46 (2015).

IV. A. McStay and G. Rosner, : 'Emotional artificial intelligence in children's toys and devices: Ethics, governance and practical remedies.' Big Data & Society, 8(1), 2053951721994877 (2021).

V. A. Virani, L. Duffett-Leger, and N. Letourneau, : 'Parenting apps review: in search of good quality apps.' Mhealth, 5 (2019).

VI. A. Kachhi, S. Chaturvedi, H. A. Patil, and D. K. Singh, : 'Data Augmentation for Infant Cry Classification.' In 2022 13th International Symposium on Chinese Spoken Language Processing (ISCSLP), 433-437. IEEE (2022, December).

VII. A. Sharma, A. Thakur, M. Garg, and S. Sharma, : 'Early intervention saved the life of mother and baby: Management of a parturient with snake bite with neurotoxicity.' Indian Journal of Anaesthesia, 65(Suppl 3), S139 (2021).

VIII. A. K. Olsavsky, J. Stoddard, A. Erhart, R. Tribble, and P. Kim, : 'Reported maternal childhood maltreatment experiences, amygdala activation and functional connectivity to infant cry.' Social cognitive and affective neuroscience, 16(4), 418-427 (2021).

IX. A. Pothula, M. A. R. Mondol, S. Narasimhan, S. M. Islam, and D. Park, : 'SEDRo: A Simulated Environment for Developmental Robotics.' In 2020 Joint IEEE 10th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob), 1-2. IEEE (2020, October).

X. A. Chakraborty, B. Singh, A. Sau, D. Sanyal, B. Sarkar, S. Basu, & J. S. Banerjee: Intelligent vehicle accident detection and smart rescue system. In Applications of Machine Intelligence in Engineering (pp. 565-576). CRC Press (2022).

XI. A. Chakraborty, J. S. Banerjee, R. Bhadra, A. Dutta, S. Ganguly, D. Das, S. Kundu, M. Mahmud, & G. Saha: A framework of intelligent mental health monitoring in smart cities and societies. *IETE Journal of Research*, 1-14 (2023).

XII. C. Pan, W. Zhao, S. Deng, W. Wei, Y. Zhang, and Y. Xu, : 'The Methods of Realizing Baby Crying Recognition and Intelligent Monitoring Based on DNN-GMM-HMM.' In 2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 352-356. IEEE (2018, May).

XIII. C. Ji, T. B. Mudiyanselage, Y. Gao, and Y. Pan, : 'A review of infant cry analysis and classification.' EURASIP Journal on Audio, Speech, and Music Processing, 2021(1), 1–17 (2021).

XIV. C. Pan, J. S. Banerjee, D. De, P. Sarigiannidis, A. Chakraborty, and S. Bhattacharyya, : 'ChatGPT: A OpenAI Platform for Society 5.0.' In Doctoral Symposium on Human Centered Computing, 384-397. Singapore: Springer Nature Singapore (2023, February).

XV. D. M. Romeo, M. Ricci, M. Picilli, B. Foti, G. Cordaro, and E. Mercuri, : 'Early neurological assessment and long-term neuromotor outcomes in late preterm infants: a critical review.' Medicina, 56(9), 475 (2020).

XVI. F. S. Matikolaie, & C. Tadj: Machine learning-based cry diagnostic system for identifying septic newborns. Journal of Voice (2022).

XVII. F. Fahmi, W. Shalannanda, I. Zakia, and E. Sutanto, : 'Design of an IoT-based smart incubator that listens to the baby.' In IOP Conference Series: Materials Science and Engineering, Vol. 1003, No. 1, 012153. IOP Publishing (2020, December).

XVIII. F. S. Matikolaie, Y. Kheddache, and C. Tadj, : 'Automated newborn cry diagnostic system using machine learning approach.' Biomedical Signal Processing and Control, 73, 103434 (2022).

XIX. G. Esposito, J. Nakazawa, P. Venuti, and M. H. Bornstein, : 'Componential deconstruction of infant distress vocalizations via tree-based models: A study of cry in autism spectrum disorder and typical development.' Research in Developmental Disabilities, 34(9), 2717-2724 (2013).

XX. G. Zamzmi, R. Kasturi, D. Goldgof, R. Zhi, T. Ashmeade, and Y. Sun, : 'A review of automated pain assessment in infants: features, classification tasks, and databases.' IEEE reviews in biomedical engineering, 11, 77–96 (2017).

XXI. H. Kumari and M. Paliwal, : 'Concept of Trisutra Ayurveda: A Review.' International journal of research in ayurveda and pharmacy, 8, 13-16 (2017).

XXII. H. Xuan, Z. Zhang, S. Chen, J. Yang, and Y. Yan, : 'Cross-modal attention network for temporal inconsistent audio-visual event localization.' In Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, No. 01, 279-286 (2020, April).

XXIII. H. I. Salehin, Q. R. A. Joy, F. T. Z. Aparna, A. T. Ridwan, and R. Khan, : 'Development of an IoT based smart baby monitoring system with face recognition.' In 2021 IEEE World AI IoT Congress (AIIoT), 0292-0296.

XXIV. https://www.soundsnap.com/

XXV. H. Norholt, : 'Revisiting the roots of attachment: A review of the biological and psychological effects of maternal skin-to-skin contact and carrying of full-term infants.' Infant Behavior and Development, 60, 101441 (2020).

XXVI. I. M. Lyså, : 'Managing Risk and Balancing Minds Transforming the Next Generation through 'Frustration Education'.' In Childhood Cultures in Transformation 30 Years of the UN Convention on the Rights of the Child in Action towards Sustainability. Brill (2021).

XXVII. J. M. Leventhal, G. A. Edwards, and R. G. Barr, : 'Abusive Head Trauma and Infant Crying—A Flawed Analysis.' JAMA pediatrics, 175(11), 1181-1181 (2021).

XXVIII. J. S. Banerjee, A. Chakraborty, M. Mahmud, U. Kar, M. Lahby, G. Saha: Explainable Artificial Intelligence (XAI) Based Analysis of Stress Among Tech Workers Amidst COVID-19 Pandemic. In: Advanced AI and Internet of Health Things for Combating Pandemics. Springer (2023).

XXIX. J. K. Mandal, S. Misra, J. S. Banerjee, S. Nayak (eds.): Applications of machine intelligence in engineering. In: Proceedings of 2nd Global Conference on Artificial Intelligence and Applications (GCAIA, 2021), 8–10 September 2021, Jaipur, India. CRC Press (2022).

XXX. J. S. Banerjee, M. Mahmud, D. Brown: Heart rate variability-based mental stress detection: an explainable machine learning approach. SN Comput. Sci. 4(2), 176 (2023).

XXXI. J. S. Banerjee, S. Bhattacharyya, A. J. Obaid, & W. C. Yeh, (Eds.): Intelligent Cyber-Physical Systems Security for Industry 4.0: Applications, Challenges and Management. CRC Press, (2022).

XXXII. J. Chattopadhyay, S. Kundu, A. Chakraborty, & J. S. Banerjee: Facial expression recognition for human computer interaction. New Trends in Computational Vision and Bio-inspired Computing: Selected works presented at the ICCVBIC 2018, Coimbatore, India, 1181-1192, (2020).

XXXIII. K. Wermke and W. Mende, : 'Musical elements in human infants' cries: in the beginning is the melody.' Musicae Scientiae, 13(2_suppl), 151-175 (2009).

XXXIV. K. Wakunuma, G. Ogoh, D. O. Eke, and S. Akintoye, : 'Responsible AI, SDGs, and AI Governance in Africa.' In 2022 IST-Africa Conference (IST-Africa), 1-13. IEEE (2022, May).

XXXV. K. G. Srinivasa, B. J. Sowmya, A. Shikhar, R. Utkarsha, and A. Singh, : 'Data analytics assisted internet of things towards building intelligent healthcare monitoring systems: IoT for healthcare.' Journal of Organizational and End User Computing (JOEUC), 30(4), 83–103 (2018).

XXXVI. K. Das, & J. S. Banerjee: Green IoT for intelligent cyber-physical systems in industry 4.0: A review of enabling technologies, and solutions. In Applications of machine intelligence in engineering, 463-478, CRC Press (2022).

XXXVII. K. Das, & J. S. Banerjee: Cognitive Radio-Enabled Internet of Things (CR-IoT): An Integrated Approach towards Smarter World. In Applications of Machine intelligence in Engineering (pp. 541-555). CRC Press (2022).

XXXVIII. L. Liu, W. Li, X. Wu, and B. X. Zhou, : 'Infant cry language analysis and recognition: an experimental approach.' IEEE/CAA Journal of Automatica Sinica, 6(3), 778–788 (2019).

XXXIX. L. Provenzi, J. Lindstedt, K. De Coen, L. Gasparini, D. Peruzzo, S. Grumi, and S. Ahlqvist-Björkroth, : 'The paternal brain in action: a review of human fathers' fMRI brain responses to child-related stimuli.' Brain Sciences, 11(6), 816 (2021).

XL. M. K. Kim and D. J. Kim, : 'Effects of oral stimulation intervention in newborn babies with Cri du Chat syndrome: Single-Subject Research Design.' Occupational Therapy International, 2018.

XLI. M. Lahby, V. Pilloni, J. S. Banerjee, M. Mahmud (eds): Advanced AI and Internet of Health Things for Combating Pandemics. Springer (2023).

XLII. Q. Gao, L. Tong, L. Tang, W. Zhong, and H. Zhu, : 'Parental knowledge on infant crying and abusive head trauma and relevant shaking behaviors in China.' Child Abuse & Neglect, 115, 105025 (2021).

XLIII. R. Majumder, M. Dasgupta, A. Biswas, & J. S. Banerjee: IoT-Based Smart City for the Post COVID-19 World: A Child-Centric Implementation Emphasis on Social Distancing. In Applications of Machine intelligence in Engineering (pp. 599-612). CRC Press (2022).

XLIV. R. L. Babbitt, T. A. Hoch, D. A. Coe, M. F. Cataldo, K. J. Kelly, C. Stackhouse, and J. A. Perman, : 'Behavioral assessment and treatment of pediatric feeding disorders.' Journal of Developmental & Behavioral Pediatrics, 15(4), 278-291 (1994).

XLV. R. Gupta, M. H. Choudhury, M. Mahmud, & J. S. Banerjee: Patent Analysis on Artificial Intelligence in Food Industry: Worldwide Present Scenario. In Doctoral Symposium on Human Centered Computing (pp. 347-361). Singapore: Springer Nature Singapore (2023).

XLVI. S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, and E. E. Bolton, : 'PubChem in 2021: new data content and improved web interfaces.' Nucleic acids research, 49(D1), D1388-D1395 (2021).

XLVII. S. Lahmiri, C. Tadj, and C. Gargour, : 'Nonlinear statistical analysis of normal and pathological infant cry signals in cepstrum domain by multifractal wavelet leaders.' Entropy, 24(8), 1166 (2022).

XLVIII. S. P. Dewi, A. L. Prasasti, and B. Irawan, : 'The study of baby crying analysis using MFCC and LFCC in different classification methods.' In 2019 IEEE International Conference on Signals and Systems (ICSigSys), 18-23. IEEE (2019, July).

XLIX. S. P. Dewi, A. L. Prasasti, and B. Irawan, : 'Analysis of LFCC feature extraction in baby crying classification using KNN.' In 2019 IEEE International Conference on Internet of Things and Intelligence System (IoTaIS), 86-91. IEEE (2019, November).

L. S. Biswas, L. K. Sharma, R. Ranjan, S. Saha, A. Chakraborty, & J. S. Banerjee: Smart farming and water saving-based intelligent irrigation system implementation using the internet of things. In Recent Trends in Computational Intelligence Enabled Research (pp. 339-354). Academic Press (2021).

LI. S. Bhattacharyya, J. S. Banerjee, S. Gorbachev, K. Muhammad, & M. Koeppen (eds.): Computer Intelligence Against Pandemics. De Gruyter (2023).

LII. S. Bhattacharyya, J. S. Banerjee, & D. De (eds.): Confluence of Artificial Intelligence and Robotic Process Automation. Springer (2023).

LIII. S. Bhattacharyya, J. S. Banerjee, M. Köppen (eds.): Human-Centric Smart Computing: Proceedings of ICHCSC 2022, vol. 316. Springer Nature (2022).

LIV. S. Bhattacharyya, J. S. Banerjee, M. Köppen (eds.): Human-Centric Smart Computing: Proceedings of ICHCSC 2023, Springer Nature, (2024) (Press)

LV. S. Bhattacharyya, J. S. Banerjee, D. De, M. Mahmud (eds): Intelligent Human Centered Computing. Human 2023. Springer Tracts in Human-Centered Computing. Springer, Singapore (2023).

LVI. S. Guhathakurata, S. Saha, S. Kundu, A. Chakraborty, J. S. Banerjee: South Asian countries are less fatal concerning COVID-19: a fact-finding procedure integrating machine learning & multiple criteria decision-making (MCDM) technique. J. Inst. Eng. (India): Series B, 1–15 (2021).

LVII. S. Guhathakurata, S. Kundu, A. Chakraborty, J. S. Banerjee: A novel approach to predict COVID-19 using support vector machine. In: Data Science for COVID-19, pp. 351–364. Academic Press (2021).

LVIII. S. Guhathakurata, S. Saha, S. Kundu, A. Chakraborty, J. S. Banerjee: A new approach to predict COVID-19 using artificial neural networks. In: Cyber-Physical Systems, pp. 139–160. Academic Press (2022).

LIX. T. Ozseven, : 'Infant cry classification by using different deep neural network models and hand-crafted features.' Biomedical Signal Processing and Control, 83, 104648 (2023).

LX. T. Khan, : 'An intelligent baby monitor with automatic sleeping posture detection and notification.' AI, 2(2), 290-306 (2021).

LXI. T. Hussain, K. Muhammad, S. Khan, A. Ullah, M. Y. Lee, and S. W. Baik, : 'Intelligent baby behavior monitoring using embedded vision in IoT for smart healthcare centers.' Journal of Artificial Intelligence and Systems, 1(1), 110-124 (2019).

LXII. X. Zhang, Y. Zou, and Y. Liu, : 'AICDS: An infant crying detection system based on lightweight convolutional neural network.' In Artificial Intelligence and Mobile Services–AIMS 2018, 7, 185-196. Springer International Publishing (2018).

LXIII. Y. C. Liang, I. Wijaya, M. T. Yang, J. R. Cuevas Juarez, and H. T. Chang, : 'Deep learning for infant cry recognition.' International Journal of Environmental Research and Public Health, 19(10), 6311 (2022).

LXIV. Z. Khalilzad, Y. Kheddache, and C. Tadj,: 'An entropy-based architecture for detection of sepsis in newborn cry diagnostic systems.' Entropy, 24(9), 1194 (2022).