AI-Driven Infant Voice Analysis for Early Diagnosis of Health Issues

Ayushi Kaushik
Apex Institute of Technology (CSE)
Chandigarh University
Punjab, India
ayushilenin23@gmail.com

Aditya
Apex Institute of Technology (CSE)
Chandigarh University
Punjab, India
jangraaditya8@gmail.com

Merry K P
Apex Institute of Technology (CSE)
Chandigarh University
Punjab, India
merrypaulose@gmail.com

I. Abstract

The well-being and health of infants have long been a primary concern for caregivers and healthcare professionals. The inability of infants to communicate effectively through conventional means often leads to challenges in early detection of discomfort, health issues, or emotional states. This paper introduces an innovative voice-based monitoring system for infants, empowered by cutting-edge artificial intelligence (AI) technologies, aimed at revolutionizing infant care. The voice-based monitoring system utilizes AI algorithms to analyse and interpret the vocalizations and sounds emitted by infants, providing real-time insights into their emotional and physical states. Through this advanced analysis, caregivers can promptly respond to the infant's needs, whether related to hunger, discomfort, or potential health concerns. This system offers a novel and non-intrusive approach to infant monitoring, reducing the stress and uncertainty that often accompany parenthood.[1] We delve into the technical intricacies of this system, discussing the AI algorithms used for voice analysis and pattern recognition. Furthermore, we present the results of a comprehensive study conducted to evaluate the system's effectiveness in diverse real-world scenarios, highlighting its accuracy in detecting and classifying infant vocalizations. The findings demonstrate the potential of this technology to enhance the quality of infant care significantly. Ethical considerations are a central focus of this research, emphasizing the importance of data privacy, security, and responsible AI development. We explore the ethical dimensions of AI-powered infant monitoring, addressing concerns related to data collection and potential misuse of sensitive information. We advocate for a balanced approach that combines the benefits of technology with the warmth and care of human interaction.

Keywords — Voice Based Monitoring System, Artificial Intelligence, Voice Analysis, Voice Pattern Recognition.

II. Introduction

Worldwide, about 130 million babies are born each year. Properly meeting the needs of newborns is a huge challenge, especially for first-time parents , and nurses are able to recognize crying based on their collective experiences." system, but new parents face difficulties calming their babies due to the apparent similarity of crying symptoms It began in the 19th century, when the Wasz–Hawkert research group, Trained nurses facilitated by, identified crying four distinct types as sounds—*Belly Pain, Burping, Discomfort, Hungry and Tired [1].* Preliminary studies confirmed that trained adult listeners can discriminate different types of infant features in pitch but that training human perception of infant features proves more difficult compared to machine learning object models for training a. Mukhopadhyay's study showed that, in contrast to the highest classification accuracy of 33.09% obtained when a group of individuals were trained to recognize specific scream sounds, machine learning algorithms using spectral and prosodic features showed a much higher accuracy of 83.62% on the same data [2]. The development of intelligent devices capable of recognizing babies' behavior not only fulfills emergency care needs, but also paves the way for intelligent robotic caregivers in the future If babies' daily life needs will ever be understood beside, another important aspect of infant temperament research is disease prognosis. Certain disorders affect infants' vocal and respiratory systems, causing crying symptoms with symptoms typical of healthy infants. Examination of behavioral symptoms represents a noninvasive and rapid method of diagnosis, particularly useful in areas of limited access to medical equipment and knowledge In the early years of infant behavioral research was focused on classifying common symptoms. Saraswati's review [3] includes 34 papers published between 2019 and 2023, focusing primarily on the classification of normal and pathological tear signs. These projects were established to identify various diseases, such as hypo-acoustic conditions, asphyxia, hypothyroidism, hyperbilirubinemia, cleft palate, etc. Objective To use tear signal analysis as a means of rapid diagnosis of health issues and effective in infants, which can save lives, especially in areas where medical resources are scarce.[7] Research on infant characteristics takes a comprehensive approach including data collection, rain model development, product extraction and selection, and classification Because of the sensitivity of procurement data, researchers face challenges with data achieving the necessary. Data collection methods require researchers to capture cryclips themselves or to obtain access to data sets from other authors. Data are typically recorded in various settings such as hospitals, neonatal intensive care units (NICUs), homes, and hospitals, either in real time or through the installation of electronic recorders infant bedside over time, making signal processing necessary to eliminate background noise and crack cracks form the basis for spit databases. Once the database is established, the next step is feature extraction, which involves extracting features from different domains of screaming signals, such as time domain, cepstral domain, or prosodic domain and then using feature selection to de select the most suitable features for effective classification models . The use of appropriate machine learning models tailored to specific spit features is essential for accurate classification or detection. The resurgence of artificial intelligence (AI) in the 1990s brought neural networks to the forefront of infant behavior research. Neural networks induced by biological brain systems include neuronal connections, weights, activating functions, and outputs. 2000s Scaled conjugate gradient neural networks, multi-level perceptrons,[9] general regression neural networks, evolutionary neural networks, probabilistic neural networks, neuro-fuzzy networks, and time -Neural methods such as delay neural networks were widely adopted Hidden Markov Models and Support Vector Machines (SVM) also found application during this period.

Sr. No.	Name of Dataset	Source	Number of Samples	Genders	Number of Classes	Age Group
1.)	Donate-a-cry-corpus	*Donate-a-cry mobile application*	457 Voice Samples	2	5	0-24 Months

Table 1 – overview of the dataset

Concurrently, novel neural network architectures like Convolutional Neural Network (CNN), Long short-term memory (LSTM) CNN, CNN+LSTM, and ResNet50+LSTM opened new avenues in infant cry research. This survey primarily reviews signal processing techniques and machine learning methods developed in the past decade for infant cry research. It covers typical databases, pre-processing approaches, diverse features in time and frequency domains, and suprasegmental features of infant cry signals. [4] The focus is on state-of-the-art methods utilizing CNN, LSTM, and ResNet50 algorithms for classification and detection. The survey concludes with a list of resources for researchers interested in this domain and outlines potential future directions for research in this area.

Sr. No.	Classes	Number of Samples
1.)	*Belly Pain*	16
2.)	*Burping*	8
3.)	*Discomfort*	27
4.)	*Hungry*	382
5.)	*Tired*	24

Table 2 – Number of samples obtained

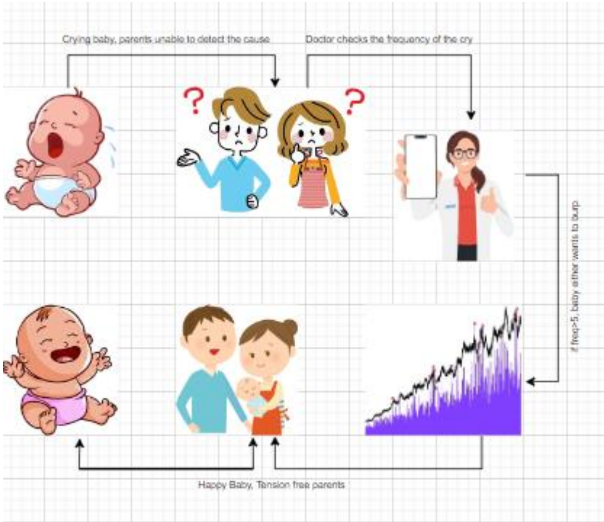

Figure 1 – flowing showing the working model

III. LITERATURE REVIEW

DATASET DESCRIPTION

In order to construct a resilient model for infant cry detection, an assemblage of infant auditory recordings gathered in authentic environmental contexts were acquired and subsequently subjected to annotation. The aggregation encompassed audio records from the Donate-a-cry mobile application, featuring a high likelihood of cry occurrences.[8] This dataset encompasses vocal samples from infants in various states, categorized into five classes denoting the causes of crying: abdominal discomfort, burping, general unease, hunger, and fatigue. The dataset comprises vocal recordings from two distinct genders, male and female, spanning the age spectrum from newborn to 24 months. The dataset has been taken from a GitHub repository: https://github.com/gveres/donateacry-corpus/tree/master/donateacry_corpus_cleaned_and _updated_data

Size of the Dataset

The dataset comprises a total of 457 vocal samples, distributed across specific categories as follows: 16 samples pertaining to instances of abdominal pain, 8 samples corresponding to burping, 27 samples associated with discomfort, 382 samples indicative of hunger, and 24 samples representative of fatigue. To facilitate model development and evaluation, the dataset was partitioned into training and validation subsets. This partitioning was executed in an 8:2 ratio, where the training dataset encompasses 80% of the total data, containing a sum of 366 samples, while the validation dataset encompasses the remaining 20%, comprising 91 samples in total.

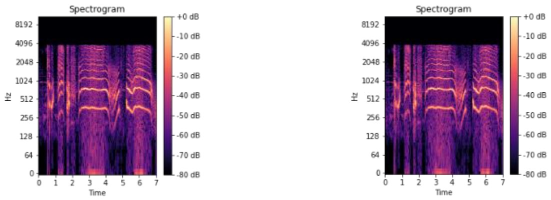

Figure 2 – Training and Validation spectrograph for Belly Pain

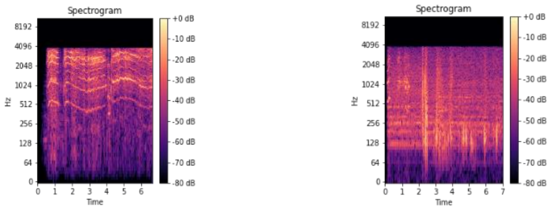

Figure 3– Training and Validation spectograph for Burping

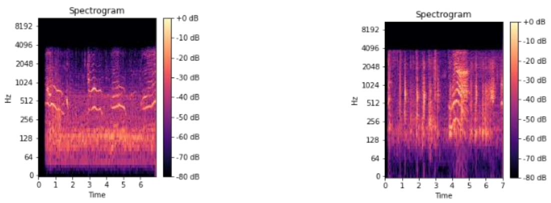

Figure 4– Training and Validation spectograph for Discomfort

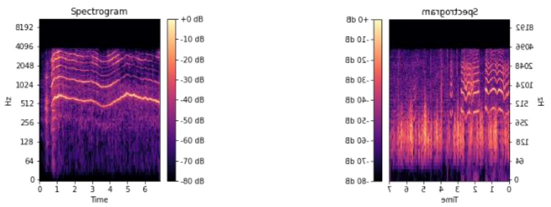

Figure 5 – Training and Validation spectograph for Hungry

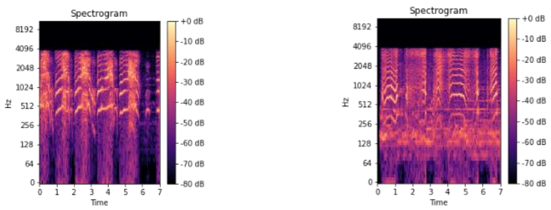

Figure 6 – Training and Validation spectograph for Tired

Sr. No.	Data Split	Number of Samples
1.)	*Training Data*	366
2.)	*Validation Data*	91

Table 3 – Number of samples in training and validation data

TYPES OF BABY CRY

Infant vocal expressions serve as fundamental communication tools, facilitating the conveyance of needs and emotional states. [6] The categorization of these cries delineates distinct types: Hunger Cry, marked by rhythmic persistence denoting the requirement for sustenance; Tiredness Cry, characterized by whiny tones and accompanied by behaviors such as eye rubbing, indicative of fatigue or excessive stimulation; Discomfort Cry, intermittently expressed and displaying fussiness, arising from issues like wet diapers or irritations; Pain Cry, a sudden and intense expression signaling substantial distress, potentially attributable to colic or injuries; and the Burping Cry, emerging post-feeding, manifesting as strained cries and squirming, denoting trapped air discomfort necessitating burping for gas pain alleviation.[1] Comprehensive understanding of these cries is pivotal for responsive caregiving, fostering secure attachments, and augmenting infant well-being.

DEEP LEARNING

Deep learning constitutes a subset of machine learning methodologies that emulate the intricate operations of neural networks within the human brain. [5]These methodologies encompass sophisticated algorithms and neural network architectures meticulously crafted to comprehend and derive hierarchical learning from extensive datasets. Employing layered neural nodes, these systems traverse copious data layers, progressively extracting abstract features.

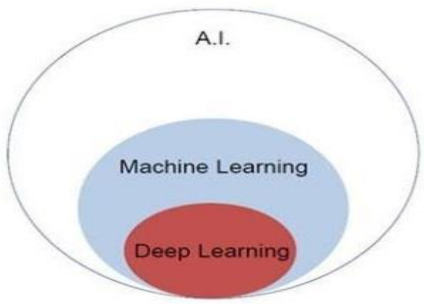

Figure 7 – Artificial Intelligence, Machine Learning, and Deep Learning venn diagram

The pivotal prowess of deep learning resides in its autonomous discernment of intricate data patterns, correlations, and representations, thereby facilitating tasks inclusive of image and speech recognition, natural language processing, and complex decision-making processes. Notably, models within deep learning, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have brought about transformative shifts across diverse domains, attaining cutting-edge performance levels in domains historically reliant on human-like cognitive capacities.[11] The adaptability of this technology and its adeptness in assimilating diverse data types continue to steer pioneering advancements spanning industries ranging from healthcare to autonomous vehicular technologies and beyond. [9]

TITLE	YEAR	TOOLS AND METHOD	RESULTS	ACCURACY
Is Speech the New Blood? Recent Progress in AI- Based Disease Detection from Audio in a Nutshell	2022	The paper discusses the use of deep learning methods, expert-designed feature extractors, and classical machine learning methodologies for disease detection from audio data, CNN	The paper discusses recent advances in the use of speech data for the automatic audio-based detection of diseases, including respiratory diseases, psychiatric disorders, developmental disorders, and neurodegenerative disorders	69%
Automated Speech Recognition System to Detect Babies' Feelings through Feature Analysis	2022	Machine learning and artificial intelligence were used to distinguish cry tones in real time through feature analysis, HMM	The study used machine learning and artificial intelligence to distinguish cry tones in real-time through feature analysis, producing real-time results after recognizing a child's cry sounds.	80%
A Review of Infant cry Analysis and Classification	2018	automatic infant cry research are data acquisition, pre-processing, feature extraction, feature selection, and classification, ANN	been achieved in various areas of infant cry research, including pathological cry identification, cry reason classification, and cry sound detection	80.56%
Infant Crying Detection in Real World Enviornments	2022	LENA (Language ENvironment Analysis) is a commercial product used by developmental psychologists to capture and process relevant acoustic events, including infant crying, in everyday environments, CNN	The model trained on in-lab data underperformed when presented with real-world data, highlighting the need for real-world datasets	61.30%
Deep Learning for Infant Cry Recognition	2022	Deep learning algorithms such as Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) were used for recognizing infants' necessities and differentiating healthy and sick infants.	They used feature extraction and a combination of CNN and a stacked restricted Boltzmann machine (RBM) to classify infant cry signals based on the health status of the baby. The developed model classified pathological conditions such as hunger, need for diaper changing, emotional needs, and pain caused by medical treatment.	60%

Figure 8 – Table of previous paper analysis

REMOTE MONITORING FOR BABIES

Remote monitoring systems for infants utilize technological apparatuses to remotely observe and track infant behavior and physiological parameters, contributing valuable insights into their overall health and welfare. These systems integrate an array of sensors, cameras, and wearable devices to meticulously monitor vital signs, sleep patterns, and environmental variables. Through intricate data analysis, these systems facilitate the identification of potential triggers or factors associated with instances of infant crying.[15] Anomalies detected in sleep patterns or abrupt fluctuations in physiological markers, such as heart rate or body temperature, remotely signal potential discomfort or underlying illnesses. Furthermore, audio-visual inputs captured by surveillance mechanisms assist in deciphering contextual stimuli prompting infant distress, enabling caregivers to swiftly intervene by adjusting environmental conditions or attending to the infant's immediate needs.[17] This technological paradigm supports a proactive caregiving approach, furnishing real-time insights into infant requirements, thereby augmenting parental responsiveness and fostering optimal infant well-being.

IV. PROPSED SYSTEM

The proposed framework presents a deep learning neural network model specifically tailored for the comprehensive analysis of audio data, employing spectrogram conversion as a fundamental preprocessing stage. This pivotal process involves the transformation of raw audio data into spectrographs,[7] enabling the representation of sound frequencies across temporal dimensions. Subsequently, the model integrates diverse neural network architectures, encompassing Convolutional Neural Networks (CNNs), CNNs augmented with Long Short-Term Memory (LSTM) networks, and a hybridized approach amalgamating ResNet50 with LSTM.[9]

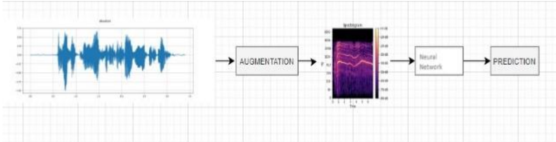

Figure 9 – Flow chart of proposed system

Initiating with the conversion of audio into spectrograms, this preparatory step enables the extraction of intricate temporal and frequency-based features. These spectrographs serve as primary inputs for the distinct neural network configurations. [11] The CNN architecture specializes in extracting salient features from spectrograph images, while the CNN+LSTM model capitalizes on the amalgamation of CNNs' image processing prowess with LSTM's sequence modeling capabilities. Furthermore, the ResNet50+LSTM hybrid architecture integrates ResNet50's robust feature extraction capabilities with LSTM's proficiency in capturing temporal dependencies within the spectrogram data.[14]

The comparative evaluation of these models aims to discern their respective efficacies in capturing nuanced audio patterns, emphasizing their performance in tasks such as sound classification or identification.[12] Evaluation metrics, encompassing accuracy, precision, recall, and F1-score, serve to provide comprehensive insights into the efficacy of each architectural configuration in discerning complex audio features. This research endeavor stands to significantly contribute to advancements in audio-based deep learning applications.

V. METHODOLOGY

The methodology employed herein represents a structured and methodical approach to appraise the proposed deep learning neural network model tailored for audio analysis of infant crying. The procedural sequence commences with meticulous data augmentation, entailing the transformation of raw audio files into spectrographs, effectively delineating sound frequencies temporally.

Following the derivation of spectrograms, a suite of distinct neural network architectures is implemented, comprising Convolutional Neural Networks (CNNs), CNNs integrated with Long Short-Term Memory (LSTM) networks, and a composite model merging ResNet50 with LSTM.[13] These architectures are primed to accept spectrograph representations as inputs, facilitating the extraction of salient features and discernment of intricate patterns.

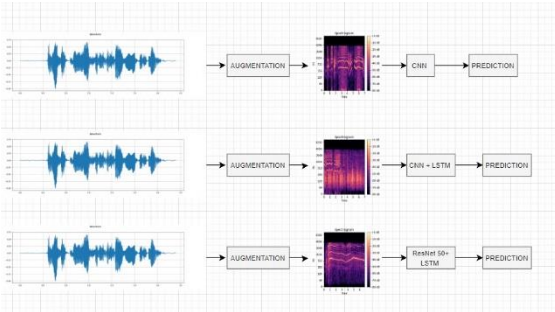

Figure 10 – Flow chart of all three neural networks

To ensure a comprehensive evaluation, a structured comparative analysis is systematically executed across diverse neural network configurations.[5] The models undergo rigorous training, and validation utilizing pertinent datasets, underpinned by the utilization of standardized performance metrics encompassing accuracy, precision, recall, and F1-score.

This rigorous and methodologically sound approach fosters a systematic exploration into the strengths and constraints inherent in diverse neural network architectures for audio data analysis, thereby furnishing invaluable insights into their effectiveness and practical applicability in real-world scenarios.[1][2]

V. RESULTS

The study's results encapsulate the findings stemming from the comprehensive implementation and assessment of diverse neural network architectures within the proposed deep learning framework dedicated to audio analysis. Through the training and subsequent evaluation of these models utilizing spectrograph representations derived from the audio dataset, a spectrum of performance metrics, including accuracy.

Nevertheless, the limited dataset size imposed inherent constraints, leading to a discernible convergence among all models. The restricted dataset substantially influenced the models' performances, yielding akin predictive capabilities and outcomes. Despite the variances in architectural compositions—comprising CNN, CNN+LSTM, and ResNet50+LSTM—the dataset's limitations curtailed the models' ability to exhibit substantial divergence in performance. This observed uniformity underscores the profound impact of dataset constraints on model differentiation, accentuating the challenge of attaining diverse performance outcomes within a confined dataset context.

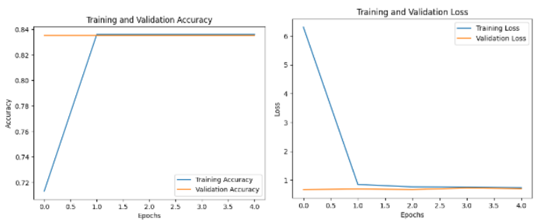

Figure 11 – Training and Validation Accuracy, Training and Validation Loss Of CNN model

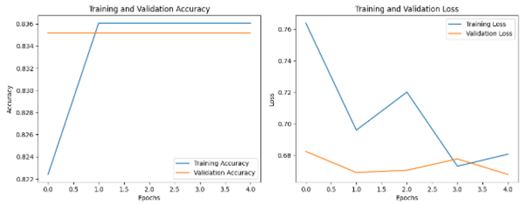

Figure 12 – Training and Validation Accuracy, Training and Validation Loss Of CNN+LSTM model

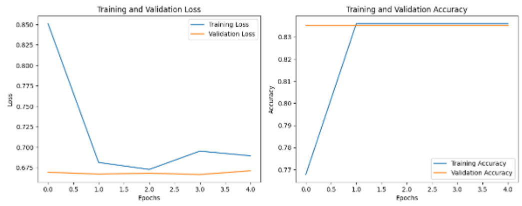

Figure 13 – Training and Validation Accuracy, Training and Validation Loss Of ResNet50+LSTM model

Sr. No.	Model	Training Accuracy	Validation Accuracy	Test Accuracy
1.	CNN	83.61%	83.52%	83.52%
2.	CNN+ LSTM	83.61%	83.52%	83.52%
3.	ResNet 50+ LSTM	83.61%	83.52%	83.52%

Table 4 – Accuracy achieved by different models

VI. CONCLUSION

The paper "AI-Driven Infant Voice Analysis for Early Diagnosis of Health Problems" proposes the use of CNN, CNN+LSTM, ResNet50+LSTM model for early diagnosis of health problems in infants The study investigates the performance of these models work in health information discovery by analyzing infant temperament cues. The results show that the CNN+LSTM and ResNet50+LSTM models have the same accuracy as the end-to-end CNN model, which is 83.62% in the training phase, 83.52% in the validation phase, and 83.52% in the testing phase.

Recommendations of Future Research -

Suggestions for future research on infant voice analysis using AI using CNN, CNN+LSTM, and ResNet50+LSTM models include:

1. Hybrid CNN-LSTM models with efficient over-parameter tuning can be used to predict other diseases in infants, such as autism, cerebral palsy, and hearing loss [4]

2. The proposed method for rapid application of CNN-based infant characteristic features can be extended to detect other diseases affecting infant vocal cord development[5]

3. Explanation with artificial intelligence (XAI) techniques can be used to interpret the results of AI-driven infant voice analysis models.[8]

4. Analysis of acoustic and vocal quality features can be used to classify infant and other maternal voices.[9]

5. Another research framework of CNN structures can be explored to develop new efficient block architectures.[14]

6. Long-term and short-term memory networks (LSTM) can be used to model spatio-temporal patterns for micro-expression recognition (MER).

In conclusion, future research could investigate the use of AI-based infant vocal evaluation to predict other diseases, detect other abnormalities in infant vocal cord development, interpret the results using XAI techniques [7] , and other infant and maternal voices are segmented. Furthermore, new block architectures can be developed that are more efficient for new research paradigms of CNN architecture, and spatio-temporal patterns can be modeled using LSTM networks for micro-expression recognition

VII. REFRENCES

1. Towards using cough for respiratory disease diagnosis by leveraging Artificial Intelligence: A survey - Ijaz, Aneeqa, Muhammad Nabeel, Usama Masood, Tahir Mahmood, Mydah Sajid Hashmi, Iryna Posokhova, Ali Rizwan, and Ali Imran. "Towards using cough for respiratory disease diagnosis by leveraging Artificial Intelligence: A survey." Informatics in Medicine Unlocked 29 (2022): 100832.

2. Embedded AI-based digi-healthcare - Ashfaq, Zarlish, Rafia Mumtaz, Abdur Rafay, Syed Mohammad Hassan Zaidi, Hadia Saleem, Sadaf Mumtaz, Adnan Shahid, Eli De Poorter, and Ingrid Moerman. "Embedded AI-based digi-healthcare." Applied Sciences 12, no. 1 (2022): 519.

3. AI-based monitoring of retinal fluid in disease activity and under therapy - Schmidt-Erfurth, Ursula, Gregor S. Reiter, Sophie Riedl, Philipp Seeböck, Wolf-Dieter Vogl, Barbara A. Blodi, Amitha Domalpally et al. "AI-based monitoring of retinal fluid in disease activity and under therapy." Progress in retinal and eye research 86 (2022): 100972.

4. Precision Medicine, AI, and the Future of Personalized Health Care - Johnson, Kevin B., Wei-Qi Wei, Dilhan Weeraratne, Mark E. Frisse, Karl Misulis, Kyu Rhee, Juan Zhao, and Jane L. Snowdon. "Precision medicine, AI, and the future of personalized health care." Clinical and translational science 14, no. 1 (2021): 86-93.

5. Applications of Artificial Intelligence and Big Data Analytics in m-Health: A Healthcare System Perspective - Khan, Z. Faizal, and Sultan Refa Alotaibi. "Applications of artificial intelligence and big data analytics in m-health: a healthcare system perspective." Journal of healthcare engineering 2020 (2020): 1-15.

6. INFANT CRYING DETECTION IN REAL-WORLD ENVIRONMENTS - Yao, Xuewen, Megan Micheletti, Mckensey Johnson, Edison Thomaz, and Kaya de Barbaro. "Infant crying detection in real-world environments." In ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 131-135. IEEE, 2022

7. A REVIEW OF INFANT CRY ANALYSIS AND CLASSIFICATION - Ji, Chunyan, Thosini Bamunu Mudiyanselage, Yutong Gao, and Yi Pan. "A review of infant cry analysis and classification." EURASIP Journal on Audio, Speech, and Music Processing 2021, no. 1 (2021): 1-17.

8. Automated Speech Recognition System to Detect Babies' Feelings through Feature Analysis – Yasin, Sana, Umar Draz, Tariq Ali, Kashaf Shahid, Amna Abid, Rukhsana Bibi, Muhammad Irfan et al. "Automated Speech Recognition System to Detect Babies' Feelings through Feature Analysis." Computers, Materials & Continua 73, no. 2 (2022).

9. ChatGPT for healthcare services: An emerging stage for an innovative perspective - Javaid, Mohd, Abid Haleem, and Ravi Pratap Singh. "ChatGPT for healthcare services: An emerging stage for an innovative perspective." BenchCouncil Transactions on Benchmarks, Standards and Evaluations 3, no. 1 (2023): 100105

10. GitHub - gveres/donateacry-corpus: an infant cry audio corpus that's being built through the Donate-a-cry campaign - see http://donateacry. com. https://github.com/gveres/donateacry-corpus. Accessed 07 Aug 2020

11. M. Severini, D. Ferretti, E. Principi, S. Squartini, Automatic detection of cry sounds in neonatal intensive care units by using deep learning and acoustic scene simulation. IEEE Access. 7, 51982–5199 (2019). https://doi.org/10.1109/ACCESS.2019.2911427

12. X. Zhang, Y. Zou, Y. Liu, in Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). AICDS: an infant crying detection system based on lightweight convolutional neural

13. L. Liu, Y. Li, K. Kuo, in 2018 International Conference on Information and Computer Technologies, ICICT 2018. Infant cry signal detection, pattern extraction and recognition,

14. S. Sharma, P. R. Myakala, R. Nalumachu, S. V. Gangashetty, V. K. Mittal, in 2017 7th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos, ACIIW 2017. Acoustic analysis of infant cry signal towards automatic detection of the cause of crying, (2018), pp. 117–122. https://doi.org/10.1109/ACIIW.2017.8272600

15. C. Ji, X. Xiao, S. Basodi, Y. Pan, in Proceedings - 2019 IEEE International Congress on Cybermatics: 12th IEEE International Conference on Internet of Things, 15th IEEE International Conference on Green Computing and Communications, 12th IEEE International Conference on Cyber, Physical and So. Deep learning for asphyxiated infant cry classification based on acoustic features and weighted prosodic features, (2019). https://doi.org/10.1109/iThings/GreenCom/CPSCom/SmartData.2019.00206

16. G. Gu, X. Shen, P. Xu, in Proceedings of 2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference, IMCEC 2018. A set of DSP system to detect baby crying, (2018), pp. 411–415. https://doi.org/10.1109/IMCEC.2018.8469246

17. Y. Lavner, R. Cohen, D. Ruinskiy, H. Ijzerman, in 2016 IEEE International Conference on the Science of Electrical Engineering, ICSEE 2016. Baby cry detection in domestic environment using deep learning, (2017). https://doi.org/10.1109/ICSEE.2016.7806117

18. D. Ferretti, M. Severini, E. Principi, A. Cenci, S. Squartini, in 2018 26th European Signal Processing Conference (EUSIPCO). Infant cry detection in adverse acoustic environments by using deep neural networks, (2018), pp. 992–996. https://doi.org/10.23919/EUSIPCO.2018.8553135

19. A. Chittora, H. A. Patil, in International Conference on Text, Speech, and Dialogue. Significance of unvoiced segments and fundamental frequency in infant cry analysis, vol. 9302, (2015), pp. 273–281. https:// doi.org/10.1007/978-3-319-24033-6_31

20. S. Bano, K. M. Ravikumar, in Proceedings of the IEEE International Conference on Soft-Computing and Network Security, ICSNS 2015. Decoding baby talk: a novel approach for normal infant cry signal classification, (2015), pp. 24–26. https://doi.org/10.1109/ICSNS.2015.7292392

21. S. Orlandi, C. A. Reyes Garcia, A. Bandini, G. Donzelli, C. Manfredi, Application of pattern recognition techniques to the classification of full-term and preterm infant cry. J. Voice. 30(6), 656–663 (2016). https:// doi.org/10.1016/j.jvoice.2015.08.007

22. M. V. Varsharani Bhagatpatil, An automatic infant's cry detection using linear frequency cepstrum coefficients (LFCC). Int. J. Sci. Eng. Res. 5(12), 1379–1383 (2014)

23. S. Yamamoto, Y. Yoshitomi, M. Tabuse, K. Kushida, T. Asada, Recognition of a baby's emotional cry towards robotics baby caregiver. Int. J. Adv. Robot. Syst. 10 (2013). https://doi.org/10.5772/55406

24. A. K. Singh, J. Mukhopadhyay, K. S. Rao, in 2013 Indian Conference on Medical Informatics and Telemedicine, ICMIT 2013. Classification of infant cries using source, system and supra-segmental features, (2013), pp. 58–63. https://doi.org/10.1109/IndianCMIT.2013.6529409

25. K. Manikanta, K. P. Soman, M. Sabarimalai Manikandan, in 2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS). Deep learning based effective baby crying recognition method under indoor background sound environments, vol. 4, (2019), pp. 1–6. https://doi.org/10.1109/ CSITSS47250.2019.9031058

26. G. Joshi, C. Dandvate, H. Tiwari, A. Mundhare, in Proceedings - 2017 International Conference on Vision, Image and Signal Processing, ICVISP 2017. Prediction of probability of crying of a child and system formation for cry detection and financial viability of the system, (2017), pp. 134–141. https://doi.org/10.1109/ICVISP.2017.33

27. R. Torres, D. Battaglino, L. Lepauloux, in International Conference on Engineering Applications of Neural Networks. Baby cry sound detection: a comparison of hand crafted features and deep learning approach, (2017). https://doi.org/10.1007/978-3-319-65172-9_15

28. M. Moharir, M. U. Sachin, R. Nagaraj, M. Samiksha, S. Rao, Identification of asphyxia in newborns using GPU for deep learning, (2017), pp. 236–239. https://doi.org/10.1109/I2CT.2017.8226127

29. C. C. Onu, I. Udeogu, E. Ndiomu, U. Kengni, D. Precup, G. M. Sant'anna, E. Alikor, P. Opara, Ubenwa: cry-based diagnosis of birth asphyxia. Nips, 2–5 (2017). https://doi.org/1711.06405

30. M. U. Sachin, R. Nagaraj, M. Samiksha, S. Rao, M. Moharir, GPU based deep learning to detect asphyxia in neonates. Indian J. Sci. Technol. 10, 3 (2017). https://doi.org/10.17485/ijst/2017/v10i3/110617

31. O. M. Badreldine, N. A. Elbeheiry, A. N. M. Haroon, S. Elshehaby, E. M. Marzook, in ICENCO 2018 - 14th International Computer Engineering Conference: Secure Smart Societies. Automatic diagnosis of asphyxia infant cry signals using wavelet based mel frequency cepstrum features, (2019), pp. 96–100. https://doi.org/10.1109/ICENCO.2018.8636151

32. H. B. Sailor, H. A. Patil, in Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH. Auditory filterbank learning using ConvRBM for infant cry classification, (2018), pp. 706–710. https://doi.org/10.21437/Interspeech.2018-1536

33. J. Saraswathy, M. Hariharan, V. Vijean, S. Yaacob, W. Khairunizam, in Proceedings - 2012 IEEE 8th International Colloquium on Signal Processing and Its Applications, CSPA 2012. Performance comparison of Daubechies wavelet family in infant cry classification, (2012), pp. 451–455. https:// doi.org/10.1109/CSPA.2012.6194767

34. M. Hariharan, L. S. Chee, S. Yaacob, Analysis of infant cry through weighted linear prediction cepstral coefficients and probabilistic neural network. J. Med. Syst. 36(3), 1309–1315 (2012). https://doi.org/10.1007/ s10916-010-9591-z

35. L. Le, A. N. M. H. Kabir, C. Ji, S. Basodi, Y. Pan, in Proceedings - 2019 IEEE 16th International Conference on Mobile Ad Hoc and Smart Systems Workshops, MASSW 2019. Using transfer learning, SVM, and ensemble classification to classify baby cries based on their spectrogram images, (2019). https://doi.org/10.1109/MASSW.2019.00028

36. T. Nadia Maghfira, T. Basaruddin, A. Krisnadhi, Infant cry classification using CNN - RNN. J. Phys. Conf. Ser. 1528(1), 012019 (2020). https://doi. org/10.1088/1742-6596/1528/1/012019

37. S. P. Dewi, A. L. Prasasti, B. Irawan, in Proceedings - 2019 IEEE International Conference on Signals and Systems, ICSigSys 2019. The study of baby crying analysis using MFCC and LFCC in different classification methods, (2019), pp. 18–23. https://doi.org/10.1109/ICSIGSYS.2019.8811070

38. I. A. Banica, H. Cucu, A. Buzo, D. Burileanu, C. Burileanu, in 2016 International Conference on Communications (COMM). Automatic methods for infant cry classification, (2016), pp. 51–54. https://doi.org/ 10.1109/ICComm.2016.7528261

39. K. Sharma, C. Gupta, S. Gupta, in 2019 10th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2019. Infant weeping calls decoder using statistical feature extraction and Gaussian mixture models, (2019), pp. 1–6. https://doi.org/10.1109/ ICCCNT45670.2019.8944527

40. M. Huckvale, in Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH. Neural network architecture that combines temporal and summative features for infant cry classification in the Interspeech 2018 Computational Paralinguistics Challenge, (2018