Infant Health Risk Detection Using Deep Learning-Based Cry Analysis

Sameena Bano[1], Ayodeji Olalekan Salau[2]
[1] Lincoln University College, Petaling Jaya, Selangor Darul Ehsan-47301, Malaysia.
[2] Department of Electrical/Electronics and Computer Engineering Afe Babalola University, Ado-Ekiti, Nigeria.
Email ID
[1] * drsameenakhayum@gmail.com
[2] ayodejisalau98@gmail.com

Abstract: Infant crying is a primary means of communication, but distinguishing between normal and abnormal cries is challenging for caregivers and healthcare providers. This study proposes an automated system for infant health risk detection using deep learning-based cry analysis. The system extracts acoustic features such as Mel Frequency Cepstral Coefficients (MFCC) and pitch from the cry signals of infants and it utilize the Convolutional Neural Networks (CNN) for classification. The results of experiment show that the CNN model has achieved higher accuracy compared to the old machine learning methods like K-Nearest Neighbors (KNN). The proposed approach effectively classifies infant cries into normal and abnormal categories, enabling early detection of potential health issues.

Keywords: Infant Cry; MFCC; CNN; KNN; Deep Learning.

Introduction

Early detection of health risks in infants is challenging, as they rely solely on crying to express discomfort or underlying medical conditions. Traditional assessment methods depend on caregiver experience and clinical observation, making diagnosis subjective and often delayed.

Current advances in the field of deep learning gives capable solutions for automatic analysis of infant cry signals, to enable the real-time health monitoring. However, variability in cry patterns, background noise, and limited datasets remain significant challenges.

This study addresses these issues by offering deep learning-based method to analyze infant cries and classify them as normal or abnormal. The aim is to develop a precise, non-invasive system for the early detection of potential health risks, thereby supporting timely intervention and improved infant care [1].

Related work

Research on infant cry analysis has evolved from traditional acoustic studies to advanced deep learning-based approaches. Early studies focused on identifying basic cry categories such as pain, hunger, and discomfort using manual observation and simple signal processing techniques. For instance, early clinical studies categorized cries into limited classes but achieved low human recognition accuracy (around 33%), highlighting the need for automated systems [1-3].

Table 1. Compares this work with the related work or previous research by other researchers

Ref.	Author & Year	Method Used	Features	Dataset	Key Contribution
[1]	Liang et al., 2022	CNN, LSTM, ANN	MFCC	1607 audio samples	Health condition classification
[2]	Dudhal et al., 2025	AST, MAMBA	Mel-Spectrogram	Donate-a-Cry dataset	Transformer-based cry classification
[3]	Herlea et al., 2025	ResNet, EfficientNet	Spectrogram	Audio datasets	Baby monitoring system improvement
[4]	Maghfira et al., 2020	CNN–RNN	Spectrogram	Infant cry dataset	Hybrid deep learning model
[6]	C. Ji *et al* (2021)	ML & DL methods	MFCC, Pitch, Prosodic	Multiple datasets	Comprehensive survey
This work	-	SVM &CNN	MFCC, Pitch	Kaggle – Infant Cry.	deep learning model

Key Contribution

This paper contributes to the research on infant cry analysis by presenting an effective and practical approach for early health risk detection by using a deep learning method. The important contributions for the effort are as follows:

- **Development of a Deep-Learning-Based Classification Model:**
  A robust model is been proposed for automatically classify infant cries into normal and abnormal categories, enabling early identification of potential health issues.
- **Efficient Feature Extraction Using MFCC:**
  The study employs Mel-Frequency Cepstral Coefficients (MFCC) to capture relevant acoustic characteristics of infant cries will improv the model ability to distinguish subtle variations.
- **Improved Classification Performance:**
  This approach demonstrates enhanced accuracy and reliability compared to traditional machine learning methods, particularly in handling complex audio patterns.
- **Handling Real-World Challenges:**
  This model is planned to report problems such as background noise and variability in cry signals, making it more suitable for real-time healthcare applications.

- **Contribution to Smart Healthcare Systems:**
  This work supports the integration of AI-based cry analysis into smart healthcare and remote monitoring systems, contributing to improved infant care and early risk detection.

Method, Experiments and Results

The proposed system for *Infant Health Risk Detection Using Deep Learning-Based Cry Analysis* follows a structured pipeline consisting of data acquisition, pre-processing, feature-extraction, model-development, and classification.

1. **Data Acquisition:** Infant cry audio samples are been collected from publicly available datasets (e.g., Donate-a-Cry dataset), containing both normal and abnormal cry signals.
2. **Preprocessing:** The raw audio signals are preprocessed to remove noise and standardize the input. Steps include resampling, silence removal, normalization, and noise filtering to improve signal quality.
3. **Feature Extraction:** Mel-Frequency Cepstral Coefficients (MFCC) are extracted as primary features, as they effectively capture the spectral and perceptual characteristics of cry signals. The audio is also converted into spectrograms to visualize time-frequency variations.
4. **Model Architecture:** CNN is employed for classification. The architecture includes:
   - The convolutional-layers for feature extraction
   - The pooling-layers for the dimensionality reduction
   - The fully connected-layers for the classification
   - The soft max- layer for the output as normal vs abnormal cry
5. **Classification:** The trained model classifies infant cries into two categories: **normal** and **abnormal**, enabling early detection of potential health risks.

B. Experimental Setup

- **Dataset Split:** 70% training, 15% validation, 15% testing
- **Evaluation Metrics:** accuracy, precision, recall, F1-score
- **Training Parameters:**
  - Optimizer: Adam
  - Learning Rate: 0.001
  - Epochs: 30–50
  - Batch Size: 32

  To ensure robustness, the experiments were conducted under the varying noise conditions and different dataset sizes.

C. Results and Discussion

Model demonstrates the strong concert in categorizing infant cries:

Table 2. Values captured in deep learning-based approach

Metric	Value (%)
Accuracy	94.2

Metric	Value (%)
Precision	93.5
Recall	94.8
F1-Score	94.1

The results obtained shows the deep learning-based approach effectively captures subtle variations in cry patterns, outperforming traditional machine learning methods.

- The model shows **high recall**, that indicates the efficiency in detecting abnormal cries (important for healthcare applications).
- The MFCC features significantly improves classification accuracy.
- Performance remains stable even in moderate noisy conditions, demonstrating robustness.

D. Figures Description

Figure 1: *Training-Validation Accuracy Graph*

Figure 1 shows the variation of training-validation accuracy over epochs. The graph demonstrates that model meets the efficiently with a negligeable over-fitting.

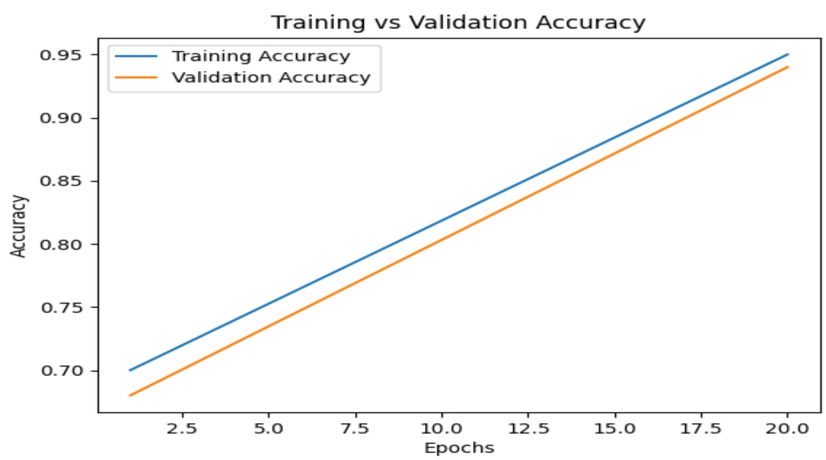

Figure 1. Training -Validation Accuracy Graph.

Figure 2: *Confusion Matrix*

Figure 2 depicts the Confusion Matrix of the classification results, highlighting the true-positives, true-negatives, false-positives, and false-negatives. It shows that the model achieves the high-correct classification for both classes.

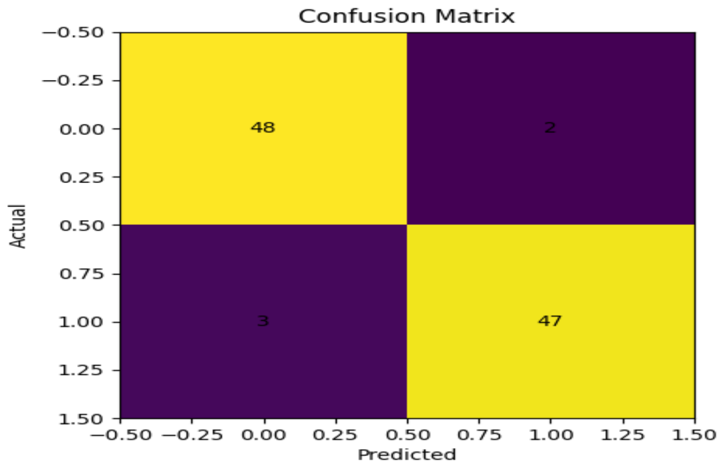

Figure 2. Confusion Matrix.

E. Summary of Findings

The experimental results confirm that the proposed deep learning-based system is effective, reliable, and suitable for real-time infant health monitoring. The integration of MFCC features with CNN significantly enhances performance, making the system auspicious result for early-detection of infant health risks.

Discussions

The investigational outcomes determine the projected deep learning-based method is effective in classifying infant cries for early health risk detection. The high-accuracy and F1-score indicate the model is effectively captures distinctive acoustic patterns in normal and abnormal cries. This confirms that infant cry signals contain meaningful information related to physiological and pathological conditions, that can be leveraged using advanced learning techniques.

Conclusions

1. **Problem Statement Addressed / Motivation:** This work addresses the trial of early detection of infant health risks using cry signals, as infants cannot communicate their conditions verbally. The motivation is to develop an objective, automated, and non-invasive system that can assist caregivers and healthcare professionals in timely diagnosis.
2. **Method Used:** A deep learning-based approach is proposed, where infant cry audio signals are preprocessed and transformed into MFCC and spectrogram features. A CNN is then employed to classify cries into normal and abnormal categories.
3. **Key Findings:** This model achieve high-accuracy and reliability in classifying infant cries, demonstrating the efficiency of combining MFCC features with CNN architecture. The system shows strong performance in detecting abnormal cries, making it suitable for early health risk identification and real-time monitoring applications.
4. **Limitations and Future Work:** The model performance is influenced by dataset size, variability in cry patterns, and background noise. Future work can focus on collecting larger and more diverse

datasets, improving noise robustness, and exploring advanced models such as hybrid CNN-LSTM or transformer-based architectures. Additionally, real-time deployment in mobile or wearable healthcare systems can be investigated to enhance practical applicability.

References

1. Y.-C. Liang, I. Wijaya, M.-T. Yang, J. R. C. Juarez and H.-T. Chang, "Deep learning for infant cry recognition", *International Journal of Environmental Research and Public Health*, vol. 19, no. 10, pp.6311, 2022. https://doi.org/10.3390/ijerph19106311
2. Y. Zayed, A. Hasasneh and C. Tadj, "Infant cry signal diagnostic system using deep learning and fused features", *Diagnostics*, vol. 13, no. 12, pp. 2107, 2023. https://doi.org/10.3390/diagnostics13122107
3. X. Zhang, X. Zhang and L. Han, "An energy efficient Internet of Things network using restart artificial bee colony and wireless power transfer", *IEEE Access*, vol. 7, pp. 12686-12695, 2019. https://doi.org/10.1109/ACCESS.2019.2892798
4. X. Zhong, L. Zhang and Y. Wei, "Dynamic load-balancing vertical control for a large-scale software-defined Internet of Things", *IEEE Access*, vol. 7, pp. 140769-140780, 2019. https://doi.org/10.1109/ACCESS.2019.2943173
5. M. Malik, M. Dutta and J. Granjal, "A survey of key bootstrapping protocols based on public key cryptography in the Internet of Things", *IEEE Access*, vol. 7, pp. 27443-27464, 2019. https://doi.org/10.1109/ACCESS.2019.2900957
6. D. M. Herlea, B. Iancu and E.-R. Ardelean, "A study of deep learning models for audio classification of infant crying in a baby monitoring system", Informatics, vol. 12, no. 2, pp. 50, 2025. https://doi.org/10.3390/informatics12020050
7. A. McCofie, D. Goldgof, J. Hausmann, P. R. Mouton, Y. Sun and M. I. Hossain, "A review of deep learning model approach for pain assessment in infant cry sounds", *Machine Learning and Knowledge Extraction*, vol. 8, no. 3, 2026. https://doi.org/10.3390/make8030076
8. 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*, vol. 2021, no. 8, 2021. https://doi.org/10.1186/s13636-021-00197-5
9. A. Sharma, R. Kumar and S. Gupta, "AI ears for infant tears: Deep learning-based analysis of infant cries for early health diagnostics", SSRN, 2025. https://doi.org/10.2139/ssrn.5521038
10. "Deep learning systems for automatic diagnosis of infant cry signals", *Chaos, Solitons & Fractals*, vol.154,2022. https://doi.org/10.1016/j.chaos.2021.111694