International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Design and Development of an IOT-Based Babycry Detection System Using ESP32-CAM with Telegram Notification Support

G.SANDHYA[1], Dr.ENNAM GOVINDA[2], K.DHAKSHAYANI[3], G.JYOSHNA[4], D.HARSHAVARDHAN[5]

[1]Assistant professor, Department of ECE, Avanthi Institute of Engineering &Technology
[2]Professor, Department of ECE, Avanthi Institute of Engineering &Technology
[2,3,4,5,6] Student, Department of ECE, Avanthi Institute of Engineering &Technology
[2]Email Id: ennamgovinda19@gmail.com

Abstract — This project presents an ESP32CAM based Baby Cry Detection System designed to assist parents and givers in continuously monitoring infants and receiving instant alerts when a baby cries. The system utilises the ESP32-CAM module integrated with an audio sensing mechanism to detect cry sound based on predefined sound intensity or pattern thresholds. When baby cry is detected, the system automatically triggers an alert mechanism without requiring constant human supervision, making it highly suitable for home and daycare environments. To enhance remote monitoring, a Telegram Bot is integrated with the system. Upon cry detection, the ESP32CAM sends a real-time notification message to the registered Telegram account, ensuring immediate awareness regardless of the caregiver's location. The ESP32-CAM also supports visual monitoring by capturing images or live video streams if required, thereby improving reliability and safety. The proposed system is compact, low-cost, energy-efficient, and easy to deploy, making it an effective IOT based solution for infant safety, parental assistance, and smart childcare applications.

Keywords: ESP32 CAM, Telegram Bot, Internet of Things (IoT), ESP32-CAM, Baby Cry Detection, Deep Learning, Convolutional Neural Network (CNN), Telegram Bot Notification, Infant Monitoring System, Machine Learning, Real-Time Monitoring, Smart Healthcare.

1. INTRODUCTION

Infant monitoring is an essential aspect of childcare, especially during the early stages of a baby's life when constant supervision is required. Babies communicate their needs primarily through crying, which may indicate hunger, discomfort, illness, or the need for attention. In

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

many households, parents or caregivers may not always be physically present near the baby, making it difficult to respond immediately when the baby cries, With the advancement of Internet of Things (IOT) technology, intelligent monitoring systems can be developed to assist caregivers by automatically detecting important events and sending alerts[1].

The Baby Cry Detection System using ESP32CAM is designed to monitor a baby in real time and notify caregivers when a crying event is detected. The system combines both audio and visual inputs to improve detection accuracy. An ESP32-CAM module captures images and streams video, while an external microphone collects sound signals from the environment. These signals are analyzed using machine learning techniques to determine whether the baby is crying[3][4]..

When a cry is detected, the system sends an instant notification to the caregiver through the Telegram messaging platform [9]. This allows parents to receive alerts on their smart phones from any where with internet access. The system also displays status information on a 16*2 LCD module for local monitoring.

The proposed solution is cost-effective, compact, and suitable for home automation and smart healthcare environments. By integrating IOT communication, machine learning models, and embedded hardware, the system provides a reliable and intelligent baby monitoring solution that improves safety and responsiveness in child care environments[5][6].

2. LITERATURE SURVEY

The field of baby health monitoring systems has witnessed significant advancements from traditional manual supervision to modern IOT enabled smart monitoring solutions. A literature survey of existing research in this domain helps in understanding the evolution of sensor-based monitoring, remote alert mechanisms, and intelligent health tracking systems. It also assists in identifying current limitations, such as lack f advanced analytics and security concerns, while exploring opportunities for innovation through improved integration of IOT, automation, and intelligent data processing techniques.

CONTINUOUS REAL TIME MONITORING:

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

The core concept behind the 10T-based Baby Health Monitoring System is continuous realtime monitoring of a baby's health and activities using connected sensors and internet communication. This ensures that the baby's condition is observed consistently without requiring constant physical supervision.

The collected sensor data is processed using programmed threshold logic, and whenever any abnormal condition is detected, instant alerts are transmitted to parents through Wi-Fi using a mobile application or Telegram chatbot. This automation enables timely response, enhances baby safety, reduces parental burden, and allows remote monitoring from anywhere, ensuring reliable and smart baby care.

INFANT CRY SIGNAL DIAGNOSTIC SYSTEM USING DEEP LEARNING AND FUSED FEATURES

The paper presented by Yara Zayed, Ahmad Hasasneh, and Chakib Tadj proposes a non-invasive deep learning-based system to diagnose neonatal diseases using infant cry audio signals. Their proposed fused deep learning model achieved an accuracy of 97.50%, demonstrating the effectiveness of CNN-based feature extraction and classification techniques [7].

Feature extraction methods such as Mel Frequency Cepstral Coefficients (MFCC) are widely used in speech and audio recognition systems due to their effectiveness in representing acoustic patterns [2]. Deep learning architectures, particularly CNNs, have significantly improved classification accuracy in audio and image processing applications [3][4].

The process begins with collecting cry recordings from hospitals, followed by preprocessing steps such as noise removal, segmentation into expiratory (EXP) cry segments, and normalization to handle data imbalance. Feature extraction is then performed across three audio domains: prosodic features using Harmonic Ratio (HR),cepstral features using Gammatone Frequency Cepstral Coefficients (GFCC), and image based features using spectrograms processed through a pertained VGG16 Convolution neural network for deep feature extraction. Instead of simple concatenation alone, the study emphasizes feature fusion during the deep learning process to enhance discriminative power. The fused features are fed into machine learning models including Support Vector Machine (SVM), Random Forest (RF), and Deep Neural Networks (DNN), with hyper parameter tuning via Research and Kera ' s Tuner. Performance is evaluated using accuracy, precision, recall, F I-score, confusion matrix, and ROC curves. The proposed fused deep learning model achieved a highest accuracy of 97.50%, outperforming previous benchmark studies and successfully extending the task from binary to multiclass classification (RDS, sepsis, healthy), demonstrating that combining spectrogram, HR, and GFCC features through deep learning significantly improves early medical diagnosis of infants using cry signals alone.

3. PROPOSED SYSTEM:

The proposed system is a Deep Learning-based 10T Baby Cry Detection and Monitoring System designed to intelligently detect. The system captures audio signals through a microphone sensor interfaced with the ESP32 microcontroller. The collected audio is processed using feature extraction techniques such as MFCC or spectrogram conversion and then analysed using a trained Convolution Neural Network (CNN) model for accurate cry detection. Upon detecting a cry, the system instantly sends notifications to caregivers through an internet-based alert mechanism. The proposed solution ensures improved accuracy, real-time monitoring, hardware implementation feasibility ,compared to existing traditional and basic 10T-based systems.Key improvements in the proposed system includes:

- **Automatic Feature Learning**:

Instead of relying only on manual feature extraction, the Deep Learning (CNN) model automatically learns complex acoustic patterns from audio data, improving detection performance.

- **IOT-Based Instant Alert System:**

The system provides real-time notifications to caregivers through internet connectivity, enabling remote monitoring.

- **Real-Time Embedded Implementation:**

Your model is integrated with hardware using the ESP32, making it practically deployable rather than just software-based research.

- **Improved Noise Robustness:**

The use of spectrogram/MFCC features with a trained neural network increases accuracy even in real-world noisy environments.

Advantages of Proposed System

- Uses machine learning for accurate cry detection
- Provides real-time notifications through Telegram
- Allows live video monitoring via ESP32 camera streaming
- Reduces false alarms through intelligent audio analysis
- Low-cost and compact hardware implementation
- Suitable for smart home and healthcare applications
- Easy integration with 10T platforms and mobile devices

3.1 BLOCK DIAGRAM

Input design refers to how data enters the system and how various components interpret and use this data. In the context of the 10Tbased Baby Cry Detection System, input design involves real-time audio acquisition, optional environmental monitoring, and network configuration inputs.

Key Inputs in the Project:

- **Infant Audio Signal:**

The primary input is the baby's cry captured using a microphone sensor. The sound waves are converted into electrical signals and sent to the ESP32CAM for processing and analysis.

- **Environmental Sensor Data (Optional):**

Temperature and humidity sensors may provide environmental data to monitor surrounding conditions affecting infant comfort.

- **Wi-Fi Configuration Input**:

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Network credentials are provided to connect the system to the internet for 10T-based notifications.

- **Power Input:**

Power is supplied through an adapter or battery module to ensure stable operation of the microcontroller and sensors.

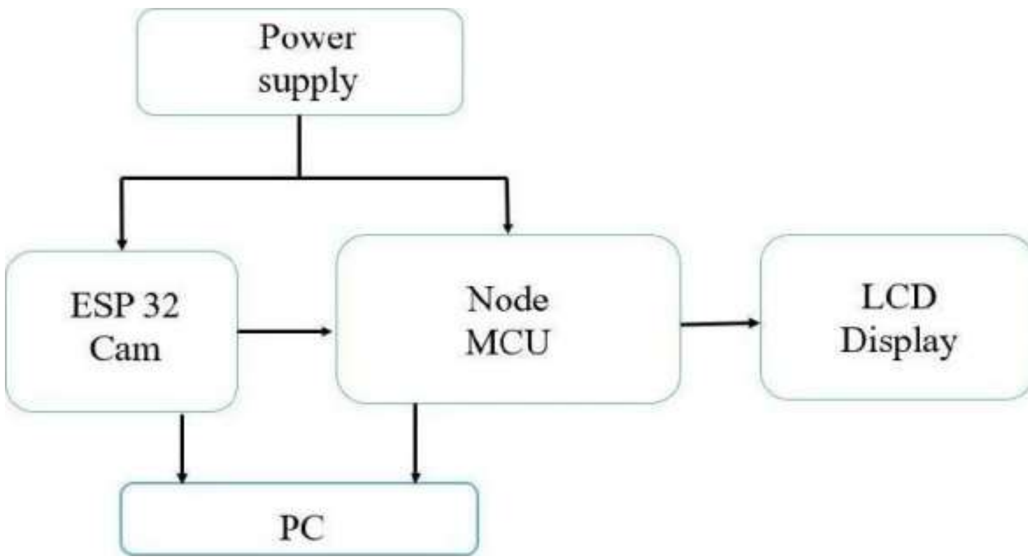

Figure 1:Block diagram

Design Consideration:

Audio signals must be filtered to reduce back ground noise.

Sampling rate must be sufficient for accurate feature extraction.

Wi-Fi communication must be secure and low latency.

Stable power supply is necessary to prevent system resets.

3.2 OUTPUT DESIGN

Output design focuses on how the system presents processed information to the caregiver. In this project, the key output is intelligent cry detection alerts and system status indications.

Key Outputs in the Project:

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

- **Cry Detection Notification:**

When a cry is detected and classified, the system sends a real-time alert through an 10T platform such as Telegram or a mobile application.

- **Cry Type Classification Display:**

The detected cry category (hunger, pain, discomfort) may be shown on an OLED/LCD display for local monitoring.

- **Buzzer Alert:**

A buzzer provides immediate audible notification when crying is detected.

- **LED Indicators:**

LEDs indicate system status such as Power ON, Wi-Fi Connected, and Cry Detected.

Design Consideration:

- **Real-Time Alert Generation:**

The system must provide immediate notifications after detecting a baby's cry to ensure quick caregiver response.

- **High Detection Accuracy**:

Outputs should be triggered only after proper cry classification to minimize false alarms.

- **Clear and Readable Indications**:

Display messages and LED indicators must be simple, visible, and easy to understand.

- **Reliable IOT Communication:**

Notifications sent through the ESP32-CAM should be stable and delivered without delay.

- **Power Efficiency:**

Output components like buzzers and displays should consume minimal power to ensure longer system operation.

4.RESULTS:

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

The proposed system was successfully implemented using ESP32-CAM hardware integrated with a microphone sensor and Telegram Bot notification support. The system effectively detected baby cries and generated real-time alerts.

The LCD module displayed the cry detection status locally, while Telegram notifications enabled remote monitoring through smart phones. Experimental results demonstrated reliable operation in real- time

Project Kit

Input

LCD Output

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

environments with acceptable accuracy and low latency

Figure 2: Input Baby cry signal and output detected signal in LCD

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Figure 3:Telegram BOT Output

5. CONCLUSION

The proposed 10T-based Baby Cry Detection System successfully integrates deep learning techniques with embedded hardware to create an efficient real-time infant monitoring solution. The system utilizes audio signal acquisition, feature extraction, and a trained Convolutional Neural Network (CNN) model to accurately detect baby cries. By combining intelligent audio analysis with IOT communication, the project ensures reliable detection even in the presence of background noise.

The implementation using ESP32-CAM makes the system cost-effective, compact, and suitable for practical deployment[8]. The hardware module continuously monitors the surrounding audio, processes the signal, and triggers alerts when a baby cry is detected. The integration with a Telegram BOT enables instant notification to caregivers, ensuring quick response and improved child safety. This real-time alert mechanism enhances the overall efficiency and usability of the system.

The deep learning model plays a crucial role in improving detection accuracy. By training the CNN model on cry and non-cry audio samples, the system learns distinctive acoustic patterns, resulting in better performance compared to traditional threshold-based or rule-based methods. The use of spectrogram-based features further strengthens the model's ability to handle variations in pitch, intensity, and environmental noise.

Overall, the project demonstrates how Artificial Intelligence and 10T can be effectively combined to address real-world healthcare and safety challenges. The system provides a scalable and reliable solution for smart baby monitoring, reducing caregiver stress and ensuring timely intervention. With further optimization and multimodal integration, the proposed model can be extended into a more advanced smart infant care system in the future.

6. REFERENCES

[1] M. Shafi, S. Ahmad, M. Burhan, and I. Haq, "IoT-Based Smart Baby Monitoring System with Emotion Recognition Using Machine Learning," *Wireless Communications and Mobile Computing*, vol. 2023, pp. 1–12, 2023.

[2] S. Davis and P. Mermelstein, "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 28, no. 4, pp. 357–366, 1980.

[3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[4] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Proceedings of the Advances in Neural Information Processing Systems (NIPS)*, 2012.

[5] F. Chollet, *Deep Learning with Python*. Manning Publications, 2017.

[6] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[7] Y. Zayed, A. Hasasneh, and C. Tadj, "Infant Cry Signal Diagnostic System Using Deep Learning and Fused Features," *Diagnostics*, vol. 13, no. 4, pp. 1–20, 2023.

[8] Espressif Systems, *ESP32-CAM Technical Reference Manual*, 2023.

[9] Telegram, *Telegram Bot API Documentation*, 2023.

[10] J. Zhang et al., "Deep Learning-Based Infant Cry Classification," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 5, pp. 120–126, 2018.