InfaSafe: A Comprehensive, Non-invasive Infant Monitoring System

Lukas Klicker, Alex Maliwat, Joanna Findura, Xinrui Yu, Mikhail Gromov and Jafar Saniie

Embedded Computing and Signal Processing (ECASP) Research Laboratory (http://ecasp.ece.iit.edu)
Department of Electrical and Computer Engineering
Illinois Institute of Technology, Chicago, IL, USA.

Abstract— InfaSafe emerges as a novel approach to infant health monitoring, uniquely positioned at the convergence of advanced artificial intelligence and edge computing. This system is designed not as a definitive solution but as an advanced platform for comprehensive data archiving, offering valuable insights into the complex and elusive nature of Sudden Unexpected Infant Death (SUID). InfaSafe utilizes AI algorithms for real-time pose estimation, breathing surveillance, and cry analysis, all within an edge computing framework that facilitates prompt and efficient data handling. This paper explores the development and capabilities of InfaSafe, underscoring its role in providing crucial, real-time insights and alerts for caregivers and its potential to contribute significantly to our understanding of neonatal health and SUID. The focus is on leveraging technological advancements to gather comprehensive data, which can be instrumental in shaping future research and interventions in neonatal care.

Keywords—infant monitoring, SIDS, feature extraction, deep learning

I. Introduction

Each year, the United States grapples with the tragic reality of approximately 3,400 sudden unexpected infant deaths (SUID), a term that encompasses fatalities among infants less than one year old where the cause is not immediately apparent. The Centers for Disease Control and Prevention (CDC) categorizes these deaths into three primary types: Sudden Infant Death Syndrome (SIDS), with 1,389 cases; unknown causes, accounting for 1,062 cases; and accidental suffocation and strangulation in bed (ASSB), contributing to 905 cases in 2020 alone.[1] This distribution underscores the multifaceted nature of SUID and highlights the critical need for comprehensive preventive strategies.

Despite significant advancements in infant care and monitoring technologies, the decline in SUID (Sudden Unexpected Infant Death) rates, which began in the early 1990s following the American Academy of Pediatrics safe sleep recommendations and the initiation of the Safe to Sleep® campaign, has plateaued since 1999 [1]. The CDC's longitudinal data reveals a concerning trend: a shift in classification from SIDS to either ASSB or unknown causes, suggesting evolving patterns in the risk factors associated with these deaths. This shift, coupled with the findings from recent studies, points to the complex interplay of factors such as sleep position, ambient temperature, and bedding, which can significantly influence SUID risk [2-3].

The InfaSafe project innovates by introducing a comprehensive, non-invasive baby monitoring system. This system takes advantage of advances in computing technology, including Artificial Intelligence for infant pose analysis, audio monitoring, thermal imaging to monitor infant breathing, and environmental sensors, all seamlessly integrated into an edge computing device, empowering parents and caregivers with interactive, real-time insights. Infasafe's comprehensive visual, audio and environmental information provides a holistic view of the baby's well-being. Infasafe is designed to be scalable with multiple infant monitoring for use in hospitals and nurseries.

The primary objective of InfaSafe is to significantly reduce the incidence of SIDS, unknown causes, and ASSB by providing caregivers and medical professionals with a comprehensive, real-time monitoring and data analysis tool. This tool adheres to and advances the American Academy of Pediatrics' recommendations for a safe sleeping environment, incorporating the latest research findings into its design and functionality [4].

II. Related Works

In recent years, rapid advancements in infant monitoring systems have been driven by an increasing demand for reliable, non-invasive, and affordable solutions. These systems are crucial for reducing risks associated with SUID. Traditionally, many of these systems have relied on wearable technologies, such as smart clothing, bracelets, or patches. While effective in monitoring basic vital signs like heart rate and oxygen saturation, these wearables often fall short in comfort and convenience and may even pose risks to infant safety. Additionally, they do not provide a holistic view of an infant's well-being.

In contrast, computer vision and artificial intelligence (AI) present promising alternatives that could overcome the limitations of wearables. These technologies utilize advanced sensors and sophisticated algorithms to capture and analyze a wide array of data concerning the infant's behavior, physiology, and environment—all without direct physical contact. Implementing these cutting-edge approaches in infant monitoring, however, introduces several technical challenges. These include managing occlusions, optimizing performance under low lighting conditions, reducing the impact of noise and motion blur, and ensuring the monitoring systems' accuracy, robustness, scalability, and interpretability.

A. Pose Estimation

Recent advancements in pose estimation have significantly impacted infant monitoring technologies. Studies like "Simple Baselines for Human Pose Estimation and Tracking" by Bin Xiao et al. [5] and "Real-time Multi-Person 2D Pose Estimation using Part Affinity Fields" by Zhe Cao et al. [6] have simplified the complexity of pose estimation algorithms achieving state-of-the-art results. These methodologies, particularly deep learning models for accurate and real-time pose detection, offer a solid foundation for applications in non-invasive infant monitoring systems. However, adapting these technologies for the specific context of infant monitoring, where accuracy and non-invasiveness are paramount, presents unique challenges. InfaSafe leverages these advancements, tailoring pose estimation techniques to accurately monitor infants' movements and positions without physical contact, addressing the critical need for safe and effective monitoring solutions.

B. Respiration Rate Monitoring

Monitoring respiration rates in infants has seen promising developments through non-contact methods, notably thermal imaging. Research by Lalit Maurya et al. [7] and Carina Barbosa Pereira et al. [8] has demonstrated the efficacy of thermal and visible imaging in accurately detecting neonatal respiratory rates. These studies underscore the potential of thermal imaging as a non-invasive technique capable of overcoming the limitations posed by traditional contact-based sensors, such as skin damage in preterm neonates. Despite these advancements, ensuring the accuracy and reliability of respiration rate monitoring in various environmental conditions remains a challenge. InfaSafe incorporates these non-contact thermal imaging techniques, enhancing the system's ability to monitor respiration rates accurately and safely, even in challenging conditions.

C. Cry Analysis

The analysis of infant cries as a means to diagnose and understand infants' needs has advanced with the application of methodologies from automatic speech recognition. The work of Liu et al. [9] in classifying cry signals based on audio features has opened new avenues for non-invasive diagnostics. By employing techniques such as linear predictive coding and Mel frequency cepstral coefficients [9], researchers have developed systems capable of interpreting a baby's needs from their cries, potentially reducing caregiver stress and aiding in the early diagnosis of conditions. However, distinguishing between normal and abnormal cries, especially in noisy environments, poses significant challenges. InfaSafe aims to build upon these cry analysis techniques, integrating advanced audio processing algorithms to enhance the system's ability to provide actionable insights into infants' well-being based on cry analysis.

III. SYSTEM DESIGN

A. System Description

InfaSafe is architected as a comprehensive IoT device to enhance infant care by monitoring critical health parameters in real time (see Fig. 1). It integrates advanced hardware and software to provide a holistic view of an infant's well-being.

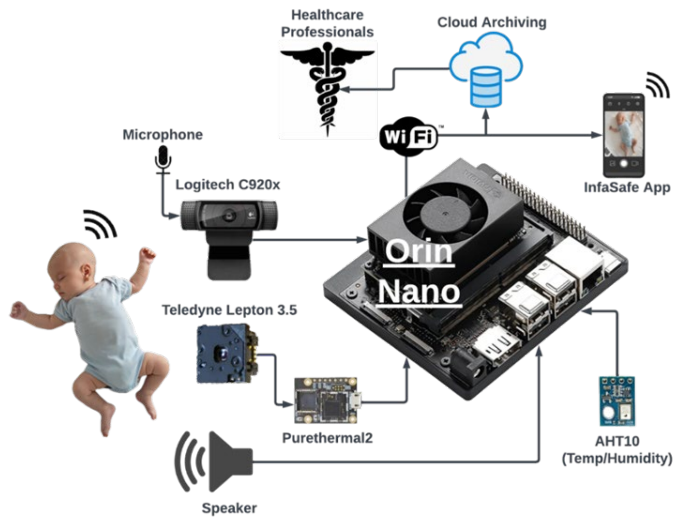

Fig. 1. InfaSafe System Map

InfaSafe employs known mitigation techniques against Sudden Unexpected Infant Death (SUID) by monitoring sleep ecology, environment, body temperature, sound, and respiration rate, leveraging AI to analyze and interpret complex data streams. InfaSafe Research and Data Archiving promotes pediatric research by archiving data related to SUID risk factors and events that could lead to unexplained infant deaths, offering valuable insights for future studies. User Interface and Accessibility designed with caregivers in mind, InfaSafe features an intuitive UI that delivers real-time information processed on the edge. It ensures safety, affordability, and ease of use without requiring subscriptions. An encrypted cloud infrastructure facilitates secure sharing with healthcare professionals for enhanced pediatric care.

A. Hardware Components

Jetson Nano serves as the core processing unit, equipped with an NVIDIA 128 CUDA core Maxwell GPU for efficient parallel computations. The system is designed for AI edge applications, capable of handling real-time AI processing tasks. The connected WiFi adapter ensures wireless connectivity, enabling seamless data transmission and system integration with the cloud and mobile devices.

Webcam/Microphone (Logitech C920X Pro Webcam): provides high-resolution video capture with low-light features. The webcam's stereo microphone is crucial for the cry analysis feature, capturing high-quality audio data for precise sound analysis.

Thermal Imaging Camera Module (FLIR Teledyne Lepton 3.5) features allow for high-accuracy measurements with minimal calibration. It produces frames at nine fps, allowing for accurate respiration rate frequency sampling and providing accurate real-time data. Its less than 50 millikelvins thermal sensitivity enables it to detect minute temperature variations, a key aspect for accurately monitoring an infant's thermography. This high level of sensitivity is crucial for precise temperature measurements and for detecting changes essential in calculating infant body temperature variations and respiration rates. Breakout Board Thermal Module (PureThermal2 GroupGets):

This breakout board from GroupGets for the Lepton camera module provides onboard image processing.

The AHT10 temperature and humidity sensor provides accurate environmental readings. These measurements are vital for maintaining a safe infant sleep environment and calibrating the thermal frames.

B. Software Libraries

The InfaSafe system's design integrates a carefully curated selection of software libraries and frameworks, each chosen for its specific capabilities to optimize performance, ensure reliability, and enhance user experience. Below, we detail the components in the logical order of their roles and interactions within the system.

1) *NVIDIA JetPack (Operating System):* NVIDIA JetPack, the foundational operating system, is crucial for managing the Jetson Nano's capabilities. It enables hardware acceleration, which is essential for the processing-intensive tasks that InfaSafe requires, such as real-time pose estimation and thermal image analysis. JetPack's comprehensive suite of tools and libraries ensures the system can efficiently manage its computational resources.

2) *Jetson Inference:* Following the operating system, Jetson Inference [10] is a deep learning library that harnesses the power of JetPack to provide AI-driven functionalities, including human pose estimation through PoseNet. This library is instrumental in enabling efficient inference on the Jetson Nano, utilizing the hardware acceleration provided by JetPack for real-time data processing.

3) Python Programming Language: Python is chosen for its readability, simplicity, and extensive ecosystem, which includes libraries for scientific computing, machine learning, and data visualization. InfaSafe leverages Python for its core programming due to these strengths. Specifically, the Jetson Utils included in Jetson Inference are utilized for seamless hardware communication, facilitating direct interaction with the system's AI models and sensor data.

4) Sensor Integration Libraries: The Adafruit CircuitPython library integrates the AHT10 temperature and humidity sensor, showcasing Python's versatility in handling hardware communication. CircuitPython offers a user-friendly approach to accessing sensor data, which is crucial for monitoring the environmental conditions around the infant. The GroupGets UVC Library is essential for interfacing with the Lepton 3.5 thermal imaging camera, enabling high-precision thermal data capture.

5) Image and Signal Processing: OpenCV: For image processing tasks, OpenCV is employed to interpolate and align IR and RGB frames, a crucial step for accurate thermal readings. Its robust functionalities support the system's need for precise image manipulation. NumPy and SciPy: These libraries are fundamental for manipulating IR frame arrays and performing signal processing, especially for calculating the respiration rate. NumPy's efficient handling of large, multi-dimensional arrays and matrices, combined with SciPy's signal processing modules, enables real-time sensor data analysis.

6) Audio Processing Libraries: Librosa: Utilized for detailed audio analysis, Librosa aids in extracting features from the baby's sounds, which is vital for identifying different states or needs based on audio cues. PyAudio: This library captures audio data, facilitating the monitoring of sounds within the baby's environment.

7) *User Interface (UI) Frameworks: Flutter:* For the mobile application, Flutter is chosen for its ability to create natively compiled applications for multiple platforms from a single codebase. This cross-platform UI framework allows parents to receive real-time data visualization and alerts on their devices, enhancing the system's user engagement and interaction. Flask: It serves as the foundational framework for our web-based user interface, enabling caregivers to monitor their infants remotely with ease and security. This lightweight yet powerful Python framework is at the heart of our application's backend, supporting the development of responsive web applications capable of efficiently handling Hypertext Transfer Protocol Secure HTTPS requests. A key aspect of Flask's utility in our system architecture is its seamless integration with Application Programming Interfaces (APIs), which facilitate communication between the web interface, the application's backend, and external services.

IV. IMPLEMENTATION

A. AI-Driven Features

The section on AI-driven features in the InfaSafe system introduces the core technological innovations that underpin its monitoring capabilities. InfaSafe incorporates sophisticated artificial intelligence to enhance infant safety comprehensively. This segment explicitly outlines the integration of two key components: pose estimation and audio analysis.

1) *PoseNet*: *PoseNet*'s application within InfaSafe is pivotal for assessing the infant's sleeping position. It is precisely engineered to detect the positioning of the nose and eye key points, enabling the system to evaluate whether an infant is in a safe sleeping posture. An 'unsafe sleeping position' event is triggered if the nose keypoint is not detected within the camera frame, signaling potential risk and alerting caregivers. Our implementation leverages the convolutional backbone of DenseNet121, selected for its depth and compatibility with infrared frame interpolation. This choice balances the need for a comprehensive depth of analysis with the requirement for maintaining sufficient frames per second (fps) for accurate pose estimation and seamless integration with thermal imaging data. DenseNet121 and ResNet18 represent two pre-trained models available within Jetson Inference for PoseNet deployment. These models, distinguished by their unique architectural designs—ResNet18's efficient residual learning approach and DenseNet121's dense connectivity pattern—offer a spectrum of capabilities. ResNet18 is optimized for speed, making it ideal for real-time applications with limited computational resources. DenseNet121, though requiring more computational power, excels in accuracy due to its complex structure and extensive layering. The choice between ResNet18 and DenseNet121 for PoseNet's deployment hinges on the specific demands of the

monitoring scenario—balancing between the necessity for real-time performance and the imperative for precision. In our application, DenseNet121's superior depth facilitates a more detailed analysis, which is crucial for infant monitoring.

2) *Cry Analysis:* For the separate aspects of cry analysis for baby monitoring, we tackle identifying the presence of a cry and the classification of one. To enable this, we require a dataset combining infant crying and other ambient sounds familiar in a household. We use the Donateacry dataset [11] and the ESC-50 audio dataset [12] for this. The Donateacry corpus was part of a campaign for users to upload their infant's crying along with metadata. There are five types of cries in the cleaned data: hunger, need for burping, belly pain, discomfort, and tiredness. For this database, the clips are tagged by the users themselves. Also, more than half of the labels are 'hungry.' This poses an obvious limitation of the efficacy of the labeling, but our considerations were put to the side due to the scarcity of data. There is a total of 457 5-second clips after data processing. ESC-50 is a collection of 2000, 5-second-long recordings of environmental Audio divided into 50 semantical categories. One of these categories is removed in feature extraction because it is 'crying baby.' The feature selected for classification is the first ten the Mel frequency Cepstral Coefficients (mfccs) averaged over the 5-second clip.

For identification of the crying, we use a support vector machine (SVM) with an Rbf kernel [13]. The SVM is trained on a combined Donateacry and ESC-50 dataset. The ten mics are extracted from the audio clips and averaged to be the feature extracted from the clip. We will use a k nearest neighbors model (KNN) with three neighbors for the classification. Here, we use the KNN model because as we only use the Donateacry dataset, the feature space cannot be segmented to produce sufficient results (a valid positive classification rate > 20% is better than guessing randomly uniformly). Both models use sci-kit-learn as their backing implementation.

The real-time audio analysis works by gathering the Audio through an input stream opened by Pyaudio. The handler uses a callback function that operates on chunks defined by the class initialization. In our case, there are 2048 audio frames per chunk. These audio frames are collected inside the class and every five seconds, the mfccs for each frame is computed and averaged, and then a prediction is generated for the SVM and KNN models.

B. *Thermal Imaging & Sensor Integration*

1) Body Temperature: The body temperature monitoring module is intricately designed to work in tandem with the PoseNet model to accurately determine the infant's body temperature through facial feature recognition. This system employs the Lepton 3.5 thermal imaging camera, leveraging the camera's precision in detecting thermal variances. A transformation matrix converts RGB frame key points specifically those identifying the infant's facial features into corresponding locations within the IR frame. Special attention is given to the regions around the right and left eyes, from which the highest temperature values are extracted. This selection is based on research indicating the eyes as reliable indicators of core body temperature. By mapping thermal data onto these keypoints, the module ensures precise and targeted temperature measurements.

2) Respiration Rate: The monitoring of the respiration rate adopts a similar approach to body temperature measurement by creating a specific region based on the nose keypoints within the IR frame. This region encompasses the nostrils and the area directly below them, capturing the temperature fluctuations that occur with each breath. Unlike the method for body temperature, which seeks the maximum temperature, the average temperature within this region is calculated and analyzed over time to determine the respiration rate.

The development of the respiration rate analysis algorithm draws upon critical findings from leading research [7,8], focusing on the area beneath the nose where temperature changes due to breathing are most pronounced. The algorithm identifies the respiration rate by analyzing temperature fluctuations in the frequency domain, effectively isolating the breathing frequency from unrelated noise.

To refine the accuracy of this analysis, advanced digital signal processing techniques are applied:

1) Initial Processing: A Hamming window is applied to the data of the nose region, emphasizing the gradient of temperature changes within this area. The average temperature of this region across sequential frames forms the basis of our respiration signal.

2) Signal Processing: The signal, consisting of 270 points (equivalent to 30 seconds at nine frames per second), begins processing once it accumulates 120 points. This setup includes a moving average component to mitigate any displacement of the region being monitored. Using the most recent 50 points, a Hampel filter identifies and excludes outliers, enhancing signal clarity.

3) *Filtering and FFT:* A Butterworth filter removes frequencies unrelated to respiration after normalization. Subsequently, the Fast Fourier Transform (FFT) magnitude is computed. The dominant frequency identified in this spectrum corresponds to the infant's respiration rate, providing a precise measurement.

4) *Data Continuity:* To maintain the integrity of the analysis, the data buffer is reset if the nose point becomes undetectable, ensuring only relevant and continuous data is processed.

This section demonstrates the integration of advanced thermal imaging techniques with sophisticated signal processing algorithms to monitor vital health parameters non-invasively, enhancing the InfaSafe system's capability to provide real-time, accurate health monitoring of infants.

C. *Software Architecture*

Integrating diverse technologies into a cohesive IoT device, our project, InfaSafe, presents a sophisticated software architecture designed to monitor infant health in real-time. The system's operation is orchestrated through six key functional threads, each responsible for a distinct aspect of data acquisition and processing, ensuring comprehensive monitoring and analysis.

1) *Vision and Position Monitoring:* Utilizing a Logitech Camera, the system employs computer vision techniques through the RGB Camera Thread. This thread leverages PoseNet for continuously monitoring the infant's position, capturing real-time video input, and analyzing it to detect the infant's posture and movements.

2) *Thermal Analysis:* The IR Camera Thread takes on the task of processing infrared (IR) frames to identify regions of interest. These regions are determined based on keypoints detected by PoseNet, with coordinates transformed for thermal analysis. This is instrumental in assessing body temperature and identifying critical points for respiration rate monitoring.

3) *Environmental Monitoring:* The Environmental Sensor Thread captures ambient temperature and humidity to ensure a comprehensive health monitoring environment. This data is periodically updated, allowing cross-thread access to conditions that may impact the infant's well-being.

4) *Audio Surveillance:* Capturing environmental sounds, the Audio Thread processes audio inputs from the Logitech webcam. Audio data is segmented for detailed analysis, providing insights into potential distress signals or environmental conditions.

5) *Data Aggregation and Analysis:* The Main Processing Thread acts as the system's nucleus, integrating data from all threads. It evaluates this information against set thresholds, applying advanced signal processing techniques to respiration and temperature data. This thread is pivotal in identifying deviations from normal parameters triggering system alerts as necessary.

6) Alert Generation and Event Management: Operational intelligence culminates in the Event Thread, which processes system flags to generate alerts. These alerts are crafted based on the frequency and nature of the flagged events, ensuring timely and accurate notification of potential issues.

In addition to real-time monitoring and alert generation, the system emphasizes data archiving. Events are meticulously logged with corresponding flags and are supplemented with saved frames. The software architecture embodies a holistic approach to infant health monitoring. By seamlessly integrating vision, thermal imaging, environmental sensing, and audio analysis within a structured multi-threaded framework, InfaSafe ensures a comprehensive, real-time surveillance system designed to safeguard infant health and well-being.

D. User Interface

In the design phase of our application, we conceptualized the user interface with a focus on simplicity, functionality, and user engagement. Fig. 2. illustrates the initial sketches that laid the groundwork for our application's interface. These sketches represent the envisioned layout and functionality of the main screens within the app, including the Login, Live, Recordings, and Graphs screens. This preliminary design phase was crucial for establishing a coherent vision for the application's user interface, emphasizing intuitive navigation and seamless user interaction. Using Firebase for authentication and integrating Google/Gmail account linkage, the Login screen sketch was designed to offer a straightforward entry point for users, prioritizing security and convenience. The Live, Recordings, and Graph screens' sketches further detail our approach to providing real-time monitoring capabilities, easy access to recorded content, and insightful data visualizations, respectively.

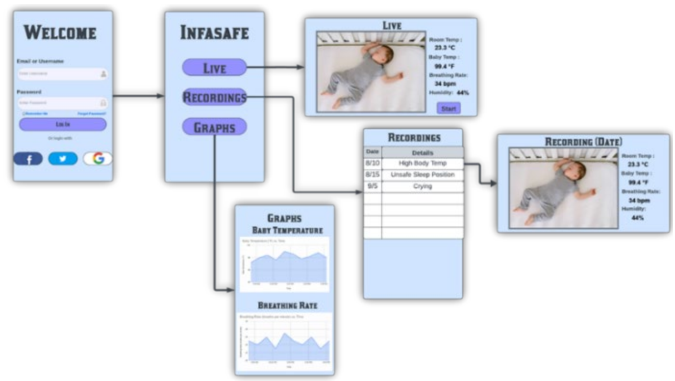

Fig. 2. InfaSafe Application User Interface Map

V. Results

The results section presents the empirical outcomes of deploying the InfaSafe system.

A. Pose Analysis

PoseNet proficiently delivered keypoints at an impressive rate of approximately ten frames per second. This capability was crucial for the system's ability to track movements and maintain continuous monitoring, which is essential for the accurate signal processing of body temperature and respiration rate. During our tests, the DenseNet121 model outperformed the ResNet model in terms of the consistency and reliability of keypoint detection. Fig. 3. Shows an example of the DenseNet121performance.

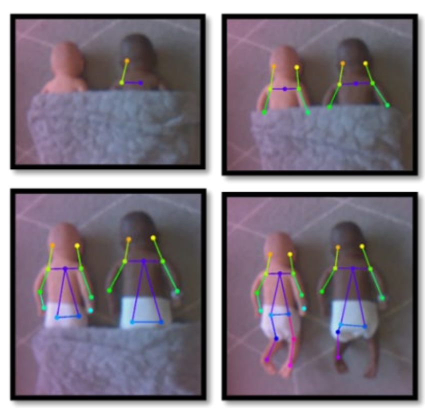

Fig. 3. Pose estimation on infant doll with PoseNet Utilizing Densenet121 showing various unsafe sleeping positions and blanket cover.

B. Body Temperature

The module was tested against standard medical thermometers, focusing on understanding the differences in readings. While some variances were noted, these were primarily accounted for by considering the skin's emissivity rate of 98% and that medical thermometers measure internal (mouth) temperature. Although the absolute temperature readings were reasonably accurate, we found that the changes in temperature readings over time were even more precise. This aspect is particularly crucial for monitoring and detecting rapid changes in an infant's condition, making our system highly effective for real-time health monitoring. Fig. 4. shows pose estimation and thermal monitoring. Fig. 5. Shows the extracted temperature values.

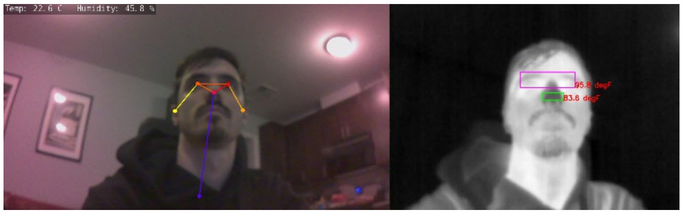

Fig. 4. Display of Pose Estimation and Thermal Monitoring: The left panel illustrates real-time pose estimation with keypoints connected to a subject's body. The right panel shows corresponding thermal imagery with transformed regions of interest—purple indicates the body temperature region, and green denotes the respiration rate area.

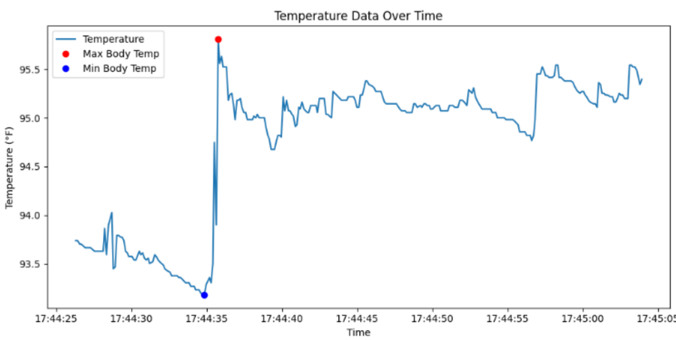

Fig. 5. The panel graphs the extracted temperature values from the designated body temperature region over time, providing a continuous and non-invasive monitoring tool.

C. *Respiration Rate*

In the evaluation of respiration rate determination, our results affirm the system's adeptness in calculating respiration rates with a high degree of accuracy when the nose keypoint is visible. The incorporation of various signal processing techniques, such as the application of the Hampel filter and a moving average calculation, has yielded stable and reliable respiration rate measurements (see Fig. 6).

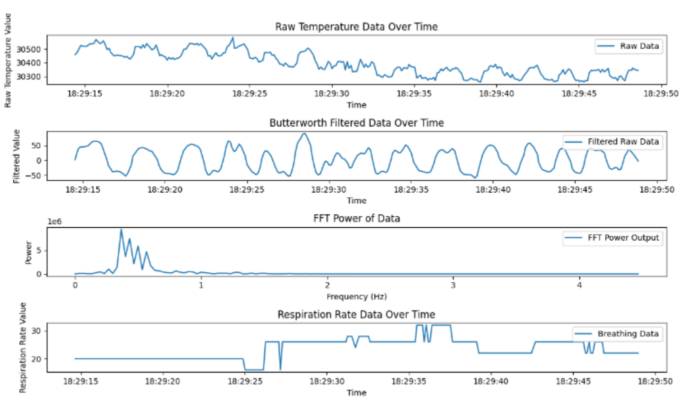

Fig. 6. Respiration rate signal extraction stages raw data, then filtered data, which is then FFT and analyzed for the most common frequency or respiration rate, which is then tracked in real-time as long as the nose is present.

VI. Conclusion

This project has successfully engineered a prototype for an IoT-based health monitoring system aimed at mitigating the risks associated with Sudden Unexpected Infant Death (SUID). Leveraging advanced computer vision, thermal imaging, and integrated sensor data, the InfaSafe monitors and analyzes key parameters such as an infant's pose, body temperature, breathing, and sleep environment to ensure safety and well-being. This platform lays a robust foundation for research and data collection, which are imperative for drawing meaningful decisions. InfaSafe can function as a platform tailored for medical professionals, enabling the real-time observation of infant health with a specific focus on preventing SUID. While the current prototype marks a significant step forward, further development is essential to solidify our understanding, particularly concerning infant health outcomes.

References

[1] "Data and Statistics for SIDS and SUID," Centers for Disease Control and Prevention. Accessed: Jan. 25, 2024. [Online]. Available: https://www.cdc.gov/sids/data.htm

[2] S. E. Parks, A. B. Erck Lambert, F. R. Hauck, C. R. Cottengim, M. Faulkner, and C. K. Shapiro-Mendoza, "Explaining Sudden Unexpected Infant Deaths, 2011–2017," Pediatrics, vol. 147, no. 5, p. e2020035873, May 2021.

[3] S. E. Parks, C. L. DeSisto, K. Kortsmit, J. M. Bombard, and C. K. Shapiro-Mendoza, "Risk Factors for Suffocation and Unexplained Causes of Infant Deaths," Pediatrics, vol. 151, no. 1, p. e2022057771, Dec. 2022.

[4] R. Y. Moon, R. F. Carlin, I. Hand, and THE TASK FORCE ON SUDDEN INFANT DEATH SYNDROME and THE COMMITTEE ON FETUS AND NEWBORN, "Evidence Base for 2022 Updated Recommendations for a Safe Infant Sleeping Environment to Reduce the Risk of Sleep-Related Infant Deaths," Pediatrics, vol. 150, no. 1, p. e2022057991, Jun. 2022.

[5] B. Xiao, H. Wu, and Y. Wei, "Simple Baselines for Human Pose Estimation and Tracking," Apr. 2018.

[6] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Real-time Multi-person 2D Pose Estimation Using Part Affinity Fields," Jul. 2017, pp. 1302–1310.

[7] L. Maurya, R. Zwiggelaar, D. Chawla, and P. Mahapatra, "Non-contact respiratory rate monitoring using thermal and visible imaging: a pilot study on neonates," J Clin Monit Comput, vol. 37, no. 3, pp. 815–828, Jun. 2023.

[8] C. B. Pereira, X. Yu, M. Czaplik, R. Rossaint, V. Blazek, and S. Leonhardt, "Remote monitoring of breathing dynamics using infrared thermography," Biomed Opt Express, vol. 6, no. 11, pp. 4378–4394, Oct. 2015.

[9] L. Liu, W. Li, X. Wu, and B. X. Zhou, "Infant cry language analysis and recognition: an experimental approach," IEEE/CAA J. Autom. Sinica, vol. 6, no. 3, pp. 778–788, May 2019.

[10] D. Franklin, "Deploying Deep Learning." Jul. 01, 2023. Accessed: Jul. 01, 2023. [Online]. Available: https://github.com/dusty-nv/jetson-inference

[11] G. Veres, "gveres/donateacry-corpus." Feb. 15, 2024. Accessed: Feb. 24, 2024. [Online]. Available: https://github.com/gveres/donateacry-corpus

[12] K. J. Piczak, "ESC: Dataset for Environmental Sound Classification," in Proceedings of the 23rd ACM International Conference on Multimedia, in MM '15. New York, NY, USA: Association for Computing Machinery, Oct. 2015, pp. 1015–1018.

[13] B. Schölkopf, A. J. Smola, R. C. Williamson, and P. L. Bartlett, "New Support Vector Algorithms," Neural Computation, vol. 12, no. 5, pp. 1207–1245, May 2000.