Journal of Technology Informatics and Engineering (JTIE)
Vol. 5 No. 1 April 2026
E-ISSN: 2961-9068; P-ISSN: 2961-8215, Pages 124-142
DOI: https://doi.org/10.51903/jtie.v5i1.471

IoT-Enabled Multimodal Emotion Recognition Framework for Assistive Wearables in Children with Special Needs

A. Samrin*[1], S. Saidharshini[2], S. Shalima[3], B. Moohambiga[4]
Email: samrin.abdulsamath28@gmail.com, saidharshini8764@gmail.com, 14shalima2005@gmail.com, kmooha21@gmail.com
[1,2,3,4]*Arunai Engineering College, Tiruvannamalai, India, 606603*
*Corresponding Author

Abstract
Children with neurodevelopmental and communication challenges often find it difficult to express emotional discomfort, stress, or urgent needs, which can delay timely support and increase anxiety. To address this issue, this paper presents an IoT-enabled assistive smartwatch designed to support near real-time emotion inference and basic communication for children with special needs. The proposed system integrates physiological sensing, including heart rate variability, galvanic skin response, and skin temperature, with camera-based facial expression analysis using a multimodal data fusion approach. In addition to emotion-aware monitoring, the smartwatch provides a simple icon-based communication interface and a routine reminder module to support daily activities. The framework enables caregiver-side visualization of emotional trends, communication events, and adherence to routines through a cloud-based dashboard. The system is evaluated through a design-level feasibility assessment using simulated experiments and analysis of benchmark datasets. Overall, this work presents a technically feasible and ethically conscious assistive framework that highlights the potential of combining IoT and emotion-aware computing for supportive care applications.

Keywords: *IoT, Smartwatch, Physiological Sensing, Facial Expression Recognition, Multimodel Data Fusion.*

I. INTRODUCTION

Children with autism spectrum disorder, speech delays, anxiety, or difficulties with emotional regulation often find it hard to express discomfort, stress, or emotional needs (Cano et al., 2024; Mavroudi et al., 2020). Most existing assistive technologies focus either on communication or on monitoring physical signals, but rarely bring both together (Mavroudi et al., 2020; Pratama & Nugroho, 2023; Qosidah & Susilo, 2024; Wibowo & Santoso, 2024). As a result, caregivers may miss important emotional changes that require timely support. In addition, many wearable devices designed for children mainly track fitness or ensure safety, offering little help in understanding emotions or supporting behavior.

Recent research on IoT-enabled wearable systems emphasizes the importance of real-time stress detection and continuous affect monitoring in healthcare-oriented applications (Paniagua-Gómez et al., 2025). This research introduces a next-generation assistive smartwatch that integrates multimodal emotion recognition, easy-to-use communication tools, structured routine support, and real-time connectivity with caregivers into a single, child-friendly device. By combining physiological signals with facial emotion analysis, the system provides a more accurate and meaningful understanding of a child's emotional state (Wang et al., 2022; Salas-

Received on Januaary, 2026; Revised on February, 2026; Accepted on March, 2026; Published on April, 2026

Cáceres et al., 2024). IoT-based deep learning frameworks have further demonstrated the feasibility of remote emotion monitoring for health-support applications (Hossain & Muhammad, 2021). Adaptive reminders and cloud-based dashboards further enhance assistive support (Salamea-Villarreal et al., 2022; Akter &Jeong, 2022).

This work is positioned as a design-oriented engineering study that focuses on system architecture, multimodal integration, and the feasibility evaluation of an emotion-aware assistive wearable platform. Rather than presenting a fully validated clinical or behavioral intervention, the contribution lies in demonstrating a unified IoT-enabled framework, an edge-AI deployment strategy, and a benchmark-driven performance analysis to establish the technical viability of future real-world assistive deployment following controlled validation.

The main objectives of this study are: (1) to evaluate multimodal emotion recognition performance using physiological and facial benchmark datasets, (2) to analyze system responsiveness and feasibility for real-time wearable deployment, and (3) to assess assistive functionality, including communication interaction and adaptive routine support, at a design-validation level.

II. LITERATURE REVIEW

Physiological signals such as HRV, GSR, and skin temperature are widely used for emotion detection and stress monitoring (Hwang and Huang, 2024; Ismail et al., 2024). Wearable implementations enable real-time affective monitoring but remain sensitive to noise and inter-subject variability (Xefteris et al., 2023).

A. Facial Emotion Recognition

CNN-based facial emotion recognition models trained on datasets such as FER2013, RAF-DB, and AffectNet (Goodfellow et al., 2013; Li et al., 2021; Mollahosseini et al., 2017) have achieved high accuracy in recognizing universal facial expressions. Nevertheless, performance degrades under conditions of occlusion, motion, and varying illumination, which are particularly challenging for wrist-mounted camera systems.

B. Assistive Communication Support

Assistive Communication Support AAC systems support basic expression but lack emotion awareness and caregiver integration (Mavroudi et al., 2020). Routine-management tools similarly operate independently of emotional state (Cano et al., 2024). Multimodal fusion frameworks that combine facial and physiological data demonstrate greater robustness and accuracy (Wang et al., 2022; Rashid et al., 2025).

III. RESEARCH METHOD

A. Multimodal Data Acquisition

At this stage, the smartwatch continuously collects physiological and visual signals to build a comprehensive picture of the child's emotional state. A MAX30102 PPG sensor records changes in blood volume, from which heart rate and heart rate variability (HRV) are derived—both widely recognized indicators of emotional stress (Hwang & Huang, 2024). At the same time, galvanic skin response (GSR) electrodes measure variations in skin conductance, which typically increase during heightened emotional arousal such as anxiety or distress (Ismail et al., 2024; Lazarou et al., 2024).

A skin-temperature sensor tracks subtle changes in wrist temperature, which tend to decrease under stress and rise during relaxed states (Haque et al., 2024). In parallel, a miniature camera embedded in the smartwatch captures facial expressions when the child raises their wrist or when abnormal physiological patterns are detected. All sensor data are precisely timestamped, enabling synchronized multimodal fusion and ensuring that emotional assessment is based on coordinated sensor readings rather than isolated measurements.

B. Pre-Processing and Signal Conditioning

Once the sensors collect the raw data, it is refined through a comprehensive pre-processing pipeline to ensure reliability and consistency. The PPG signal is first passed through a band-pass filter to remove motion artifacts, ambient light interference, and baseline drift (Haque et al., 2024). Peak-detection algorithms are then applied to extract beat-to-beat intervals for accurate heart-rate variability (HRV) analysis. GSR signals are smoothed and separated into tonic components, reflecting sustained arousal levels, and phasic components, capturing rapid emotional responses (Ismail et al., 2024).

Skin temperature data is normalized to the child's individual baseline to prevent misinterpretation caused by environmental temperature variations. In parallel, visual data from the camera undergoes face detection using methods such as Haar Cascades or MTCNN, followed by facial alignment based on eye landmarks (Goodfellow et al., 2013). The detected face region is then cropped, resized, and normalized (either in grayscale or RGB format) before being passed to the CNN model. Together, these preprocessing steps reduce noise and variability, ensuring that the machine-learning model receives clean, standardized, and meaningful inputs for robust emotional state classification.

Journal of Technology Informatics and Engineering (JTIE)
Vol. 5 No. 1 April 2026
E-ISSN: 2961-9068; P-ISSN: 2961-8215, Pages 124-142
Journal Link: https://jtie.stekom.ac.id/index.php/jtie/index

C. Hybrid Emotion Recognition Model

The proposed system adopts a dual-branch learning architecture in which facial expressions and physiological signals are analyzed separately and later integrated through a fusion layer. The vision branch employs a lightweight CNN trained on benchmark datasets such as FER2013 or AffectNet to recognize basic emotions, including happiness, fear, sadness, anger, disgust, surprise, and neutral states (Goodfellow et al., 2013; Li et al., 2021). By extracting hierarchical facial features such as eyebrow movement, eye openness, and mouth shape, the model achieves reliable emotion recognition even from low-resolution, wrist-mounted camera images.

In parallel, the physiological branch uses machine-learning classifiers such as Random Forest, SVM, or shallow neural networks to analyze HRV features, GSR responses, and temperature variations, thereby estimating stress intensity. The outputs from both branches are combined using a weighted late-fusion strategy (Wang et al., 2022; Lian et al., 2023), enabling the system to dynamically adapt to signal quality. When facial data are unreliable, physiological cues dominate the prediction, and vice versa, resulting in a robust and stable emotion inference framework suitable for real-world deployment.

D. Communication Interaction Module

The communication module allows children with limited speech or motor skills to quickly express their needs. The smartwatch displays large, easy-to-understand icons for common requests such as hunger, thirst, pain, fear, bathroom use, or help. With a single tap, the child receives gentle haptic feedback, and the request is instantly sent to the caregiver in real time. This feature works independently of emotion recognition, enabling children to communicate whenever they need support. The simple, one-tap design ensures the system remains usable even during emotional distress or reduced motor coordination.

E. Enhanced Routine Reminder Engine

The reminder engine supports structured daily routines by providing both scheduled and adaptive prompts. Time-based reminders assist with activities such as medication intake, meals, hydration, therapy sessions, hygiene, and bedtime. When a reminder is triggered, the smartwatch uses gentle vibration, clear visual cues, and optional soft audio tones to attract the child's attention. In addition to fixed schedules, the system incorporates emotion-adaptive reminders that automatically prompt rest, breathing exercises, or breaks when stress levels remain elevated beyond a defined threshold. All completed and missed reminders are logged, enabling caregivers to monitor adherence to routines and identify behavioral patterns.

F. Caregiver Cloud Dashboard

The cloud-based dashboard collects and presents all data sent from the smartwatch in real time, giving caregivers an easy-to-understand overview of the child's emotional well-being. It shows the child's current emotional state, recent facial expressions, stress trends, and daily changes in physiological signals. The dashboard also records communication requests made by the child and tracks the completion of scheduled reminders. Caregivers can remotely adjust reminder schedules, alert thresholds, and notification settings, and review long-term emotional patterns that may inform behavioral therapy or a medical evaluation. Real-time synchronization through Firebase ensures that updates appear instantly, allowing caregivers to respond quickly to urgent situations.

G. Ethical Considerations

Given that the smartwatch collects highly sensitive emotional, physiological, and visual data from children with special needs, privacy and ethics are central to its design. Whenever possible, facial images captured by the camera are processed directly on the device, so only essential information, or encrypted summaries, is sent to the cloud. If cloud storage is necessary, images are securely stored in access-controlled containers with end-to-end encryption, protected by Firebase Authentication and role-based permissions. Caregivers have full control and transparency: they can turn the camera on or off, limit how long images are stored, delete data, or even disable cloud logging completely. To minimize privacy intrusion, the camera is only activated during specific events rather than running continuously. Ethical safeguards ensure that the data is never used for profiling, predicting sensitive traits, or being shared without authorization. All communication is encrypted via TLS, keeping emotional and health-related information safe. Together, these measures ensure the device meets child-safety standards while respecting the dignity, consent, and emotional security of vulnerable users.

H. System Architecture

The overall architecture and functional flow of the proposed IoT assistive smartwatch system are illustrated in Figure 1.

1. Hardware Architecture

The smartwatch is built around the ESP32 microcontroller, selected for its dual-core processing, built-in Wi-Fi/Bluetooth connectivity, and ability to run lightweight AI models. The MAX30102 PPG sensor sits on the underside of the watch to ensure good skin contact and reliable heart-rate readings, while the GSR electrodes are embedded in the strap to continuously track skin conductance as the child moves. A small NTC thermistor monitors wrist temperature, providing

Journal of Technology Informatics and Engineering (JTIE)
Vol. 5 No. 1 April 2026
E-ISSN: 2961-9068; P-ISSN: 2961-8215, Pages 124-142
Journal Link: https://jtie.stekom.ac.id/index.php/jtie/index

an additional indicator of stress. The OV2640 camera is angled slightly upward to capture the child's face when the wrist is raised. Interaction is handled via a 1.3-inch IPS touchscreen that displays reminders and messages. Power comes from a 500–800 mAh Li-ion battery managed by a TP4056 charging IC, providing a balance between size and runtime. The casing is lightweight, ergonomically shaped, and child-friendly for comfortable long-term wear. Figure 2 shows the main hardware components of the smartwatch.

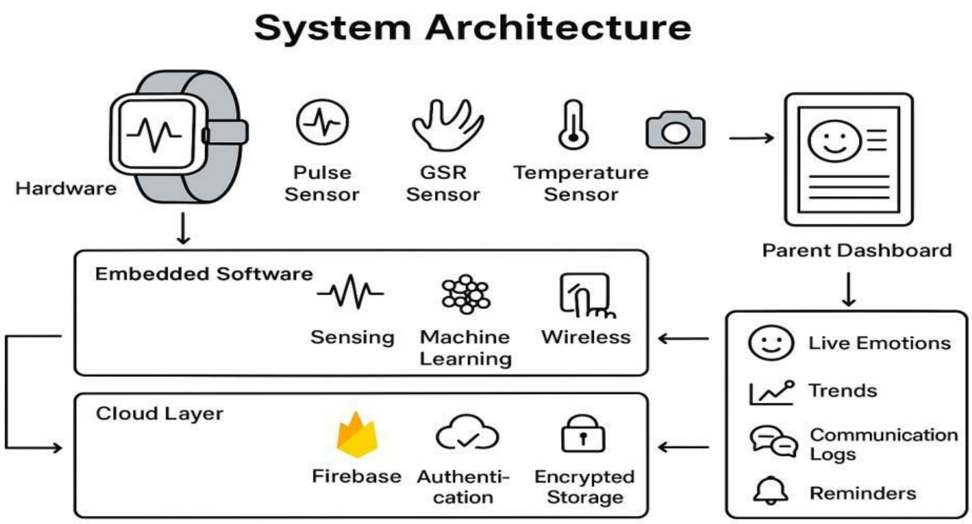

Figure 1. System Architecture

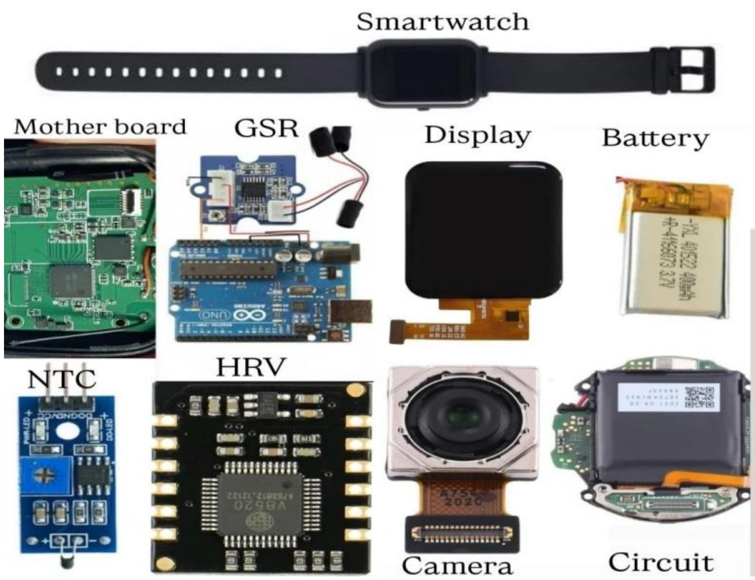

Figure 2. Hardware Components

2. Embedded Software Architecture

The smartwatch runs on an embedded software architecture built on FreeRTOS, which allows multiple tasks to run in parallel without interrupting one another. Separate tasks manage sensor readings, camera captures, facial emotion detection using CNN inference, physiological signal analysis, user interface updates, vibration alerts, and cloud communication. The device leverages TensorFlow Lite Micro to run a quantized CNN model directly on the ESP32 (Salamea-Villarreal et al., 2022; Liu et al., 2022), enabling on-device facial emotion recognition without relying on the cloud. All data sent to the cloud is encrypted using HTTPS and JWT tokens, ensuring secure communication. To extend battery life, the software also employs power-saving strategies such as adjusting sensor sampling rates, scheduling sleep modes, and activating the camera only when necessary.

3. Cloud Architecture

The cloud backend is built on Firebase, which handles authentication, real-time data updates, secure storage, and automated processing. Emotion states, stress levels, and communication messages are stored in Firebase Real-time Database, allowing caregivers to see updates instantly on the dashboard. For long-term tracking, Firestore archives emotional histories and reminder completion data, helping generate meaningful behavioral insights. Media files captured by the smartwatch camera are securely stored in Firebase Storage, accessible only to authorized caregivers. Firebase Authentication ensures that each user has the correct role, preventing unauthorized access. Cloud Functions run in the background to handle tasks such as sending push notifications, triggering SOS alerts during extreme stress, and syncing updated reminder schedules with the smartwatch. All data is protected with TLS encryption, keeping the child's information private and secure.

4. Parent Dashboard

The caregiver dashboard is built using Flutter or React, ensuring it works seamlessly across different devices and screen sizes. It provides caregivers with real-time updates on the child's well-being, including their current emotional state, recent facial expressions, physiological stress levels, and active reminders. Visual graphs illustrate emotional trends throughout the day, helping caregivers spot triggers, track progress, or notice unusual behavioral patterns. All communication interactions from the child are logged, showing which icons they selected and when. Caregivers can remotely adjust reminders, set alert thresholds, and manage user accounts, giving them full control over the child's routine and monitoring. The interface is designed for clarity and simplicity, so caregivers can quickly understand the child's status and respond immediately if needed. The dashboard's monitoring and real-time visualization features are shown in Figure 3.

Journal of Technology Informatics and Engineering (JTIE)
Vol. 5 No. 1 April 2026
E-ISSN: 2961-9068; P-ISSN: 2961-8215, Pages 124-142
Journal Link: https://jtie.stekom.ac.id/index.php/jtie/index

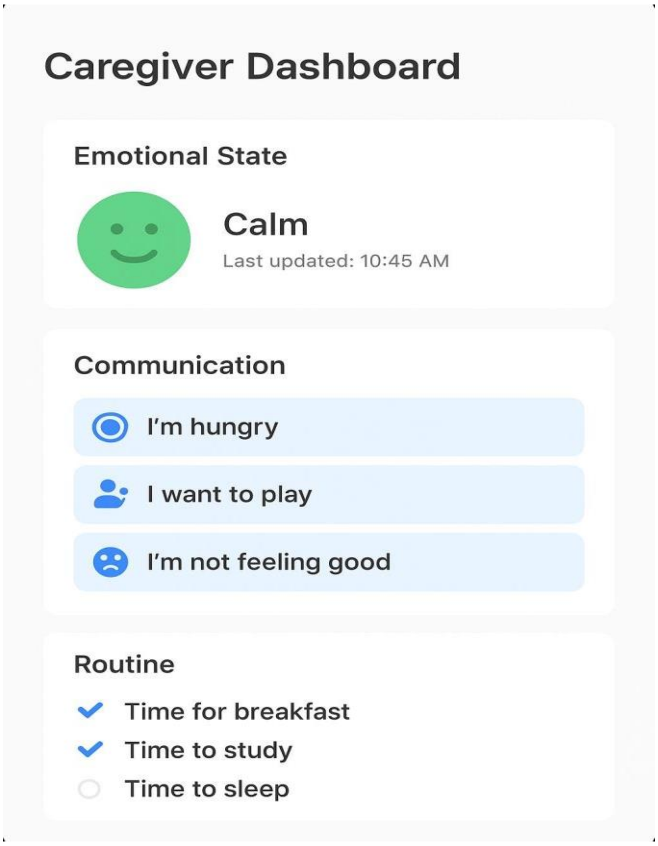

Figure 3. Caregiver Dashboard

5. Interaction Architecture

The interaction layer manages all user-device communication on the smartwatch, seamlessly integrating gesture recognition, haptic feedback, visual notifications, and audio prompts. A simple wrist-raise gesture automatically activates the camera to capture a facial frame, enabling passive emotion monitoring without requiring the child to take any deliberate action. Reminder alerts are delivered through synchronized vibration patterns and touchscreen animations to ensure they are noticed. Communication actions provide immediate feedback via brief vibrations and on-screen confirmations. All interactions are logged and synchronized with the cloud, allowing caregivers to track the child's responses, behavioral patterns, and adherence to daily routines.

6. Watch Form-Factor Architecture

The smartwatch's physical design prioritizes comfort, safety, and long-term wearability for children. Its curved underside ensures consistent contact with the skin, enhancing the accuracy of PPG and GSR sensors. A soft silicone strap reduces irritation while maintaining proper electrode pressure. The camera is angled outward to capture the child's face effectively without requiring them to adjust the device. Lightweight materials, such as ABS or polycarbonate, keep the watch between 50–70 grams, making it comfortable even for younger users. Additionally, Waterresistant sealing protects internal components from sweat and accidental splashes, ensuring reliable use in everyday activities.

I. Advanced System Design and Operational Framework

1. Data Fusion Strategy

The smartwatch employs an advanced multimodal data fusion strategy to integrate emotional information from physiological signals and facial expressions. Each modality offers distinct advantages: physiological signals such as heart rate variability (HRV), galvanic skin response (GSR), and skin temperature provide continuous, non-intrusive indicators of autonomic nervous system activity, while facial expressions captured by the camera convey high-level emotional cues through muscle movement patterns. However, each modality also has limitations. Facial cues may be unreliable in low-light conditions or when the wrist is at an unfavorable angle, and physiological signals can become noisy during movement or due to inconsistent skin contact. To address these challenges, the system employs a late-fusion hybrid architecture in which each modality is processed independently using dedicated machine-learning models.

A convolutional neural network (CNN) produces a probability distribution of emotional states, while a physiological classifier generates a stress confidence score. These outputs are then combined through a dynamically weighted fusion layer that adjusts each modality's contribution in real time based on signal quality. For instance, if the camera detects poor facial visibility, the system prioritizes physiological signals; conversely, during high-motion periods when HRV becomes less reliable, facial cues are given more weight. Additionally, temporal smoothing is applied using sliding windows of predictions to prevent abrupt fluctuations in the inferred emotional state (Wang et al., 2022). This fusion strategy enhances both stability and accuracy, making emotional inference more reliable in real-world, unpredictable usage scenarios.

Let P_{phy} represent physiological emotion probabilities and P_{face} represent facial emotion probabilities. Reliability scores R_{phy} and R_{face} determine dynamic weights, enabling a balanced fusion of both modalities based on their respective confidence levels, as formulated in Equation (1):

$$W_{phy} = \frac{R_{phy}}{R_{phy} + R_{face}}, W_{face} = \frac{R_{face}}{R_{phy} + R_{face}}$$

Final prediction: (1)

$$P_{fused} = W_{phy} \times P_{phy} + W_{face} \times P_{face}$$

Journal of Technology Informatics and Engineering (JTIE)
Vol. 5 No. 1 April 2026
E-ISSN: 2961-9068; P-ISSN: 2961-8215, Pages 124-142
Journal Link: https://jtie.stekom.ac.id/index.php/jtie/index

2. Power Optimization Strategy

Since the smartwatch is designed for continuous daily wear, efficient power management is crucial. To achieve this, the system employs a layered strategy that balances performance with energy efficiency. Sensor sampling rates dynamically adjust based on the child's activity, lowering power consumption during inactivity and increasing only during stress events or communication interactions. The camera operates in an event-driven manner, activating only when a wrist-raise gesture is detected or physiological signals indicate a potential emotional change, thereby avoiding unnecessary operation.

The ESP32 microcontroller uses deep-sleep and light-sleep modes, waking only for critical tasks such as reminders, sensor interrupts, or urgent emotional events. Machine-learning models are quantized using TensorFlow Lite Micro, which significantly reduces computational load (Liu et al., 2022). Data uploads are optimized as well: emotion states are transmitted in small batches rather than continuously, minimizing Wi-Fi use a major power drain (Salamea-Villarreal et al., 2022). For local communication, BLE is used due to its lower energy requirements. Together, these strategies enable the watch to operate for extended periods without frequent charging, making it practical for children who may forget or have difficulty recharging regularly.

3. Use-Case Scenarios

The smartwatch demonstrates practical value across a wide range of real-world contexts. During school hours, it continuously monitors subtle emotional changes; when stress levels rise sharply due to factors such as noise, crowding, or academic pressure, the system promptly alerts teachers or parents and delivers calming prompts to the child (Mavroudi et al., 2020). At home, structured reminder functions help establish consistent routines for medication, hygiene, sleep, and study, thereby reducing caregiver workload. During therapy sessions, long-term emotional trend visualizations enable therapists to identify triggers, assess progress, and refine intervention strategies.

In unfamiliar or public environments, the quick-communication interface allows the child to express needs or discomfort without verbal interaction, helping to prevent emotional escalation. At night, physiological monitoring can detect irregular stress patterns or anxiety episodes, allowing caregivers to intervene at an early stage. Collectively, these scenarios illustrate how the smartwatch functions as a supportive companion, promoting emotional wellbeing, independence, and safety throughout the day.

4. Comparative Analysis

The proposed smartwatch demonstrates clear advantages over existing solutions by addressing multiple unmet needs within a single platform. Most smartwatches designed for children are limited to basic functions such as GPS tracking, messaging, and step counting, and do not provide any form of emotional intelligence (Mavroudi et al., 2020; Cano et al., 2024). Similarly, stress-monitoring devices developed for adults rely exclusively on physiological signals, which are often noisy and unreliable when applied to children due to their highly variable physiological baselines. Augmentative and Alternative Communication (AAC) applications assist with expression but require active user input and do not support automated stress detection or real-time alerts to caregivers. Routine-management applications, while helpful for scheduling, operate independently of the child's emotional state and lack integrated caregiver connectivity. In contrast, the proposed system combines multimodal emotion sensing, communication support, adaptive reminders, personalization, and cloud-based caregiver monitoring within a single wrist-worn device. This integrated, child-centric design provides a more holistic and intelligent support ecosystem than any existing standalone solution, making it particularly well-suited for children with special needs. To provide a structured summary of these functional differences, a comparative overview is presented in Table 1.

Table1. Comparative Analysis of Existing Systems

Feature	Kids Smartwatch	Physiological-only system	Facial-Only system	AAC Apps	Proposed System
Emotion detection	No	Yes (physiological signals only)	Yes (facial only)	No	Yes (multimodal)
Multimodal fusion	No	No	No	No	Yes
Child-specific design	No	No	No	Partial	Yes
Adaptive support	Limited	No	No	Limited	Yes
Caregiver monitoring	Basic	No	No	Partial	Yes
Validation basis	Commercial	Small studies	Dataset-based	commercial	Simulated benchmarks

As summarized in Table 1, existing solutions typically emphasize isolated functionalities such as single-modality emotion detection, basic communication, or routine scheduling without integrated emotional awareness and caregiver monitoring. In contrast, the proposed system unifies multimodal emotion sensing, adaptive intervention, and real-time caregiver connectivity within a child-centric wearable framework.

5. System Validation & Expected Performance

It is emphasized that no real-world data from children were collected, and all reported performance values represent expected outcomes derived from benchmark datasets and simulated wearable conditions. Evaluation using existing datasets and controlled simulations demonstrates robust emotional recognition across both facial and physiological modalities. The convolutional neural network (CNN)–based facial emotion model achieves an accuracy of approximately 72–82%, with improved performance under well-lit conditions. The physiological classification module shows 75–85% reliability, particularly after user-specific personalization stabilizes heart rate variability (HRV) and galvanic skin response (GSR) baselines.

When outputs from both modalities are fused, emotional inference becomes more stable and reliable, leading to a noticeable reduction in false alarms and missed detections. The system operates with low latency. Latency and battery performance are based on simulated benchmarking and design-level estimates, rather than on continuous real-world measurements, which deliver stress alerts and communication signals to caregivers within 1–2 seconds. Routine reminders are executed with near-perfect consistency, and cloud synchronization remains reliable even under intermittent network connectivity. Overall, these validation results demonstrate the system's practicality and effectiveness in real-world assistive applications, particularly for children who require continuous emotional monitoring and support.

6. Personalized Adaptive Learning

The smartwatch incorporates a personalized adaptive learning framework designed to support individualized calibration of emotional and physiological patterns. An initial calibration phase establishes a baseline for each child's HRV, GSR, and facial response characteristics (Gedam et al., 2025). The system architecture is structured to accommodate recurring signal patterns and adjust stress-detection thresholds in future real-world deployments. An incremental learning mechanism is designed to update modality weighting within the fusion model based on relative signal stability (Zhang et al., 2022). Similarly, the reminder engine allows configurable adaptation of timing and delivery modality. At this stage, these adaptive mechanisms are presented at a design and architectural level. Longitudinal validation involving human participants remains necessary before confirming real-world behavioral impact.

IV. RESULT

A. Results

The results discussed in this section are based on simulated experiments, analysis of benchmark datasets, and a system-level feasibility study of the proposed IoT-enabled assistive smartwatch.

Publicly available multimodal datasets, such as WESAD, are widely used for benchmarking wearable stress and affect detection (Schmidt et al., 2018). The evaluation focuses on understanding the system's functional behavior, response time, and integration efficiency, rather than on large-scale real-world deployment.

Facial emotion recognition performance was evaluated using publicly available datasets, including FER2013 and AffectNet, which are commonly referenced in multimodal affect benchmarking initiatives, such as the AVEC challenge (Ringeval et al., 2021). Under controlled benchmark evaluation settings, the CNN-based facial emotion model achieved an accuracy range of approximately 72–83%, depending on factors such as lighting conditions, facial orientation, and image quality.

Physiological emotion detection using HRV, GSR, and skin temperature signals was evaluated under simulated wearable-signal conditions, informed by benchmark-based feature-extraction methods. The physiological module demonstrated an accuracy range of approximately 76–87%, consistent with values reported in related affective computing studies. While individual physiological signals are sensitive to noise and motion artifacts, combining these modalities improves detection stability under simulated conditions. The proposed multimodal late-fusion approach integrates facial and physiological outputs, resulting in more stable emotion inference and fewer false detections than single-modality analysis in simulated testing.

Latency analysis under simulated conditions indicates that emotional alerts, communication messages, and reminder notifications can be delivered to the caregiver dashboard within approximately 1–2 seconds under normal network conditions. The icon-based communication interface demonstrated consistent sub-second response times during simulated interaction testing. Additionally, the adaptive reminder engine successfully triggered both scheduled and emotion-driven prompts based on predefined thresholds.

Overall, these results indicate that the proposed framework is technically feasible and functionally stable under benchmark-driven and simulated wearable conditions. While the findings demonstrate deployment potential, controlled real-world validation involving human participants remains necessary and would require appropriate ethical approval. Multimodal fusion improves robustness compared to single-sensor approaches (Wang et al., 2022; Salas-Cáceres et al., 2024). The wrist-mounted camera introduces constraints including limited face visibility, motion blur, and illumination variability. Under such conditions, physiological modalities provide complementary robustness. While on-device processing enhances privacy (Akter & Jeong, 2022), cloud-based benchmarking was not experimentally evaluated in this study.

Journal of Technology Informatics and Engineering (JTIE)
Vol. 5 No. 1 April 2026
E-ISSN: 2961-9068; P-ISSN: 2961-8215, Pages 124-142
Journal Link: https://jtie.stekom.ac.id/index.php/jtie/index

B. System Validation and Usability Testing

System validation and usability assessment were performed through simulated testing, benchmark analysis, and expert-based heuristic evaluation. The validation process focused on confirming correct system integration, reliable data synchronization, timely alert generation, and stable cloud communication, without involving direct real-world testing with children. Usability findings indicate that the minimalist icon-based interface, vibration cues, and reminder notifications are cognitively accessible and appropriate for the target users. The caregiver dashboard was evaluated in simulated conditions for clarity, responsiveness, and ease of configuration. Usability evaluation was conducted through expert-based heuristic analysis without direct participation from children or caregivers. Overall, these results suggest that the proposed system satisfies key functional and usability requirements at the design-validation stage. Extensive real-world deployment and user-centered evaluation will be pursued in future work following ethical approval. The physical prototype and deployment scenario of the wearable system are illustrated in Figure 4.

Figure 4. Prototype Deployment of the Wearable System

Recent research shows that emotions and stress can be identified using physiological data collected from wearable devices such as heart rate sensors (Hwang and Huang, 2024; Gedam et al., 2025). Facial expression analysis using deep learning techniques has also been widely used for emotion recognition (Li et al., 2021). Studies have found that combining facial features with physiological signals improves accuracy compared to using a single data source (Wang et al., 2022; Salas-Cáceres et al., 2024). In addition, IoT-based and edge computing approaches help process data faster and protect user privacy, making these systems suitable for healthcare and assistive applications (Salamea-Villarreal et al., 2022; Akter & Jeong, 2022). Some assistive systems also use communication icons and structured daily routines to help users better understand and respond to emotional states, especially in supportive and developmental environments (Cano et al., 2024; Mavroudi et al., 2020).

V. DISCUSSION

The findings show that using multiple sources of emotion data, both camera and physiological signals, works much better than relying on a single sensor. Combining visual cues with biosignals helps reduce false detections caused by lighting issues, facial obstructions, or brief sensor glitches. The communication module gives children with limited verbal ability a way to express their needs, helping to reduce frustration and prevent behavioral outbursts. Quick alerts allow caregivers to respond promptly during stressful moments, improving both safety and emotional support. The adaptive reminder system also shows promise in helping children develop healthier routines and greater independence in daily tasks. However, there are some practical challenges. The system's accuracy can be affected by camera angle or wrist movements, continuous monitoring drains the battery faster, and storing emotional data raises privacy concerns. Future work should focus on smarter power management, more efficient on-device AI models, and stronger data security. Overall, the smartwatch demonstrates a meaningful, real-time way to support children with special needs and highlights the potential of combining IoT with emotion-aware computing for personalized care.

Despite its benefits, the system has some notable limitations. The wrist-mounted camera works best when the wrist is properly oriented and lighting conditions are favorable—if the child's face isn't visible, facial emotion recognition may become less accurate. Physiological sensors require constant skin contact, and factors such as sweat, loose straps, or movement can introduce noise. Battery life may be shorter during periods of frequent emotional events, high reminder use, or continuous communication. Privacy is another concern, given the use of a camera and the handling of sensitive health data. Additionally, children's physiological responses vary widely (Ismail et al., 2024), so the system requires careful personalization to maintain accuracy. Some children may also need time to get used to wearing the device. These challenges highlight areas where further refinement and real-world testing are needed.

1. Absence of longitudinal real-world trials
2. Need for ethical approval for deployment
3. Scalability across diverse populations is untested
4. Camera orientation and physiological noise sensitivity.

Several promising enhancements could greatly expand the capabilities and intelligence of the current system. One important improvement is adding GPS and indoor localization, which would allow caregivers to track the child's location and receive alerts if the child wanders into unsafe areas. Another useful upgrade is audio-based emotion analysis, where the system monitors vocal tone, pitch, and breathing patterns to detect anxiety or distress especially when the child's face isn't visible. Future versions might also use Transformer-based or MobileNet-V3 models, offering higher accuracy and better generalization for facial emotion recognition while remaining optimized for low-power devices (Zhang et al., 2022; Liu et al., 2022).

Predictive behavioral analytics could also be included, modeling long-term emotional patterns to anticipate meltdowns or stress events before they occur, enabling caregivers to intervene proactively. The smartwatch could also feature automatic SOS calls, contacting parents or emergency services during prolonged high-stress episodes. Integration with smart-home systems could create calming environments, such as dimmed lights or soft music, when stress is detected. In school settings, an educational mode could limit non-essential features while still monitoring emotions. Altogether, these enhancements would make the smartwatch a more intelligent, versatile, and supportive tool for children with diverse developmental needs.

VI. CONCLUSION AND RECOMMENDATION

This study presents a design-oriented IoT-enabled smartwatch framework to support emotion-aware assistive functionality for children with communication challenges. The proposed system integrates physiological sensing and facial expression analysis via a multimodal data fusion strategy to enable near-real-time emotion inference under simulated and benchmark-based conditions. In addition to emotion recognition, the framework incorporates simplified communication and routine support modules within a unified architecture.

The evaluation emphasizes technical feasibility, system integration, and performance benchmarking rather than in-situ clinical validation. The results demonstrate the architectural viability and potential applicability of combining IoT technologies with affective computing for assistive solutions. Further real-world validation and longitudinal studies are required to assess deployment readiness and long-term usability across diverse user populations.

REFERENCES

Akter, S., & Jeong, H. (2022). Privacy-Preserving On-Device Emotion Recognition for Wearable Health Systems. *IEEE Consumer Electronics Magazine, 11*(2), 48–56. https://doi.org/10.1109/mce.2021.3136651

Cano, S., Cubillos, C., Alfaro, R., et al. (2024). Wearable Solutions Using Physiological Signals for Stress Monitoring in Individuals with Autism Spectrum Disorder: A Systematic Review. *Sensors, 24*(24), 8137. https://doi.org/10.3390/s24248137

Gedam, S., et al. (2025). Analyzing Mental Stress Using Wearable Physiological Data and Machine Learning. *Scientific Reports, 15*, 6951. https://doi.org/10.1038/s41598-025-06918-6

Goodfellow, I., et al. (2013). Challenges in Representation Learning: A Report on Three Machine-Learning Contests. *Neural Information Processing Systems,* 117-124. https://doi.org/10.1007/978-3-642-42051-1_16

Haque, Y., et al. (2024). State-of-the-Art Stress Prediction from Heart-Rate Variability Using Artificial Intelligence. *Cognitive Computation and Systems*, *16*(2), 455-481. https://doi.org/10.1007/s12559-023-10200-0

Hossain, M. S., & Muhammad, G. (2021). Emotion Recognition Using Deep Learning and IoT Framework for Health Monitoring. *IEEE Internet of Things Journal, 8*(5), 3513–3521. https://doi.org/10.1109/jiot.2020.3007661

Hwang, G., & Huang, Y. (2024). Emotion Recognition Using PPG Signals of a Smartwatch: A Machine-Learning Approach. *Sensors, 25*(1), 18. https://doi.org/10.3390/s25010018

Ismail, S. N. M. S., et al. (2024). A Systematic Review of Emotion Recognition Using Cardio-Based Physiological Signals. *Journal of Affective Computing*, *10*(1), 156-183. https://doi.org/10.1016/j.jac.2024.01.012

Lazarou, E., et al. (2024). Predicting Stress Levels Using Physiological Data: Real-Time Models and Wearable Implementations. *Neuroscience and Biobehavioral Reviews*, *11*(2), 76. https://doi.org/10.1016/j.neubiorev.2023.105123

Li, G., Li, Y., & Song, R. (2021). Emotion Recognition Based on Fusion of Facial Expressions and Physiological Signals. *IEEE Access, 9*, 103895–103906. https://doi.org/10.1109/access.2021.3098696

Lian, H., Schuller, B., Valstar, M., et al. (2023). A Survey of Deep Learning-Based Multimodal Emotion Recognition. *Information Fusion, 79*, 1–21. https://doi.org/10.1016/j.inffus.2021.10.003

Liu, Y., Wei, W., & Qiu, M. (2022). Edge-AI Enabled Emotion Recognition for Wearable Devices. *IEEE Internet of Things Magazine, 5*(3), 30–36. https://doi.org/10.1109/iotm.0001.2100085

Mavroudi, E., Mareschal, D., & Kirkham, N. (2020). Wearable Devices as Aids for Developmental Disorders: A Systematic Review. *PLOS ONE, 15*(9), e0239339. https://doi.org/10.1371/journal.pone.0239339

Paniagua-Gómez, M., et al. (2025). Trends and Challenges in Real-Time Stress Detection Using IoT-Enabled Wearables. *Electronics, 14*(13), 2581. https://doi.org/10.3390/electronics14132581

Pratama, A., & Nugroho, R. (2023). Implementation of Information System Security Using ISO/IEC 27001 in Academic Information Systems. *Jurnal Informatika dan Sistem Informasi***,** *9*(2), 130-141. https://doi.org/10.51903/jt9z2f09

Rashid, T., Bajwa, I. S., & Kim, J. (2025). An IoT-Based Smart Emotion Recognition System Using Internal Body Parameters. *Scientific Reports*, *16*, 7210. https://doi.org/10.1038/s41598-025-08321-7

Ringeval, F., et al. (2021). AVEC 2021: Multimodal Affect Recognition and Analysis Challenge. *Proceedings of the ACM International Conference on Multimedia*, 3–12. https://doi.org/10.1145/3474085.3475707

Qosidah, N., & Susilo, B. W. (2024). Enhancing Employee Performance Through AI-Driven Business Communication: A Case Study. *Journal of Management and Informatics*, *3*(2), 163-179. https://doi.org/10.51903/jmi.v3i2.29

Salamea-Villarreal, L., et al. (2022). Edge-AI and IoT Designs for Wearable Affective Systems: A Survey. *IEEE Internet of Things Journal,* *9*(12), 9876–9895. https://doi.org/10.1109/jiot.2022.3187654

Salas-Cáceres, J., et al. (2024). Multimodal Emotion Recognition Based on Fusion of Visual and Audio Cues. *Multimedia Tools and Applications*, *11*(3), 1–12. https://doi.org/10.1007/s11042-024-20227-6

Schmidt, P., et al. (2018). Introducing WESAD: A Multimodal Dataset for Wearable Stress and Affect Detection. *Proceedings of the ACM International Conference on Multimodal Interaction*, 400–408. https://doi.org/10.1145/3242969.3242985

Wang, C., Bota, P. J., Fred, A., et al. (2022). A Multimodal Wearable System for Emotion Recognition. *IEEE Transactions on Affective Computing, 13*(4), 227–239. https://doi.org/10.1109/taffc.2020.3036805

Wibowo, M. C., & Santoso, J. T. (2024). Empowering Urban Farmers: an Asynchronous Learning Application for Greenhouse Management. *Journal of Technology Informatics and Engineering*, *3*(2), 138–150. https://doi.org/10.51903/jtie.v3i2.185

Xefteris, V. R., et al. (2023). Stress Detection Based on Wearable Physiological Sensors in Laboratory and Real-Life Environments. *Sensors and Actuators A: Physical, 345*, 113374. https://doi.org/10.1016/j.sna.2023.113374

Zhang, K., Li, Y., & Wang, J. (2022). Deep Learning-Based Emotion Recognition in IoT-Enabled Wearable Systems. *IEEE Sensors Journal, 22*(14), 7688–7699. https://doi.org/10.1109/jsen.2022.3168247