Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

FEDERATED LEARNING FOR STRESS DETECTION USING DISTRIBUTED WEARABLE DATA IN PRIVACY PRESERVING HEALTHCARE SYSTEMS

CHEEMALADINNE VENGAIAH[1], ERUKALA MAHENDER[2], P. PRATIMA RANI[3], A. SWATHI[4*], YAMINI DEVI YKUNTAM[5], RAMANJANEYULU DAYINABOYINA[6], RIYAZ MOHAMMAD[7], K. MYTHRI SRIDEVI[8]

[1]Department of Computer Science and Engineering (Data Science), B V Raju Institute of Technology, Narsapur, Medak, Telangana, India-502313

[2]Department of Computer Science and Engineering, Geethanjali College of Engineering and Technology, Cheeryal, Keesara, Hyderabad, Telangana, India -501301

[3]Department of Computer Science and Engineering, Vignan Institute of Information Technology, Duvvada, Visakhapatnam, Andhra Pradesh, India

*[4]Department of Computer Science and Engineering, CVR College of Engineering, Vastunagar, Telangana, India

[5]Department of Electronics and Communication Engineering, Aditya university, Surampalem, Andhra Pradesh, India

[6]Department of Artificial Intelligence and Data Science, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

[7]Department of English and Foreign Languages, S.R.K.R.Engineering College, Chinaamiram, Bhimavaram, Andhra Pradesh, India

[8]Department of Computer Science and Engineering, UCEK, JNTUK, Andhra Pradesh, India

E-mail: vengaiah.cse@gmail.com[1], mahee99@gmail.com[2], prathimaranipalla@gmail.com[3], akuthotaswathi26@gmail.com[4*], yaminieie@gmail.com[5], dramanjaneyulu@kluniversity.in[6], riyazsrkr412@gmail.com[7], kamisetty.my3@gmail.com[8]

ABSTRACT

Nowadays, comprehensive, continuous, and reliable monitoring is required in healthcare systems to account for stress-related impacts not only on physical but also on mental health, such as cardiovascular diseases, anxiety, and impaired cognitive function. This research aims to propose a federated learning approach for stress detection while considering the data privacy of the sensor network. FL-SDNet is proposed as a fusion approach combining a CNN and a BiLSTM for feature extraction, integrating spatial and temporal features from multi-modal physiological signals, including heart rate (HR), electrodermal activity (EDA), temperature (TEMP), and motion (MOTION). The proposed system uses a cross-modal attention fusion mechanism that dynamically balances multiple attention signals to boost performance in scenarios with diverse data distributions and introduces an adaptive federated aggregation (AFA) strategy that considers client reliability and data quality. It is fully decentralized and security-preserving, with raw data stored on local devices and only model updates sent and received in encrypted form. The results of extensive experiments with non-IID federated data show that the proposed model achieves 90.2% accuracy, 0.88 F1 score, and 0.92 ROC-AUC in a non-IID federated setting, results comparable to those of the centralized deep learning model. Moreover, it can operate in resource-limited applications (such as wearables) by reducing communication overhead by approximately 40%. A comparative study is conducted between traditional machine learning and deep learning using standard federated learning methods, indicating that the proposed method is effective. An ablation study is conducted to verify the effectiveness of the new component added to the experiment. Overall, the results of this study show that federated learning can achieve high accuracy for stress detection while ensuring data privacy, supporting the potential to exclude local data from global collection while maintaining high accuracy in real-world healthcare applications. This work serves as a valuable asset for the development of next-generation personalized real-time stress-monitoring systems for 'privacy-preserving' healthcare, remote patient monitoring, and intelligent wearables.

Keywords: *Federated Learning, Stress Detection, Wearable Sensors, Privacy-Preserving Healthcare, Multimodal Deep Learning, Non-IID Data*

1. INTRODUCTION

In fact, stress has started to receive more recognition as an important component of physical and mental health because it causes conditions such as cardiovascular disease risk factors, anxiety disorders, depression, etc., which may be critical for cognitive functioning. The immune response, and consequently human health [1], [2], is significantly challenged by long-term stressors, highlighting the importance of early detection and intervention in modern healthcare systems. The rapid growth of wearable technologies for real-time stress detection in naturalistic settings has enabled continuous non-invasive monitoring of a variety of physiological parameters (e.g., heart rate (HR), electrodermal activity (EDA), skin temperature, and motion) [3], [4].

These systems are mostly based on centralized machine learning techniques, where data from wearable devices is sent to the cloud for processing and training. However, such approaches are well-suited for optimizing data from larger populations but also raise privacy and security concerns, as well as regulatory compliance issues (especially under GDPR, HIPAA, and other regulations) [5], [6]. Centralized storage of physiological information carries a higher risk of data breaches and unauthorized access, as the sensitive nature of the data often makes users reluctant to share their personal health data, thereby severely restricting the uptake of such healthcare applications.

One of the most significant approaches to distributed machine learning for tackling these challenges is the Federated Learning (FL) paradigm. FL is a machine learning framework that enables multiple clients (e.g., wearable devices or smartphones) to collaboratively learn a shared global model while preserving each client's privacy [7], [8]. Only model updates (gradients, weights, etc.) are sent to the central server, reducing privacy risks. Notably, FL has recently been shown to be effective in healthcare applications, such as disease prediction [9], medical imaging [10], and general personal health monitoring.

Federated learning offers many advantages but also poses multiple challenges for stress detection. Data collected by wearables is often non-IID due to differences in user behavior, physiological states, and environmental conditions [11]. This heterogeneous behavior can be problematic, as training a model in this way makes it difficult to converge and achieve good performance. On the other hand, wearable devices typically have limited battery life and computational capacity [12], which makes it difficult to perform continuous training of local models and to support frequent communication between devices and servers. Furthermore, FL systems for stress detection should achieve a trade-off among high predictive accuracy, privacy protection, and low communication overhead. How to achieve these objectives simultaneously is an open problem [13].

Deep learning methods such as CNNs and LSTM networks [14] can effectively model temporal correlations in physiological time-series data. However, their application in federated learning for stress detection is still in its infancy, particularly for optimizing performance in distributed, privacy-constrained settings [15].

While some works have investigated stress detection using centralized and federated machine learning models, most consider only a single dataset and do not adequately account for the cross-domain heterogeneity of wearable healthcare data. This work extends current federated stress detectors by introducing a multi-benchmark database and an adaptive aggregation method that learns from the database and evaluates generalization across datasets. As a result, the system is more scalable, robust, and usable in real-world healthcare scenarios, where privacy is a major concern.

The main contributions of this work are:

i) A federated learning-based privacy-preserving approach to detect stress with wearable data.
ii) Train a hybrid CNN-LSTM network to capture spatial and temporal physiological patterns.
iii) Address the challenge of data heterogeneity (differences in user data) and non-IID (non-uniform, user-specific data) using adaptive client weighting (adjustments made so each user's data influences the learning process appropriately).
iv) Design strategies for efficient model updates to minimize the communication load between devices and central servers (communication overhead).
v) Evaluate the model's effectiveness in preserving privacy.

This work addresses the challenges of privacy-preserving and trustworthy stress detection in distributed healthcare environments.

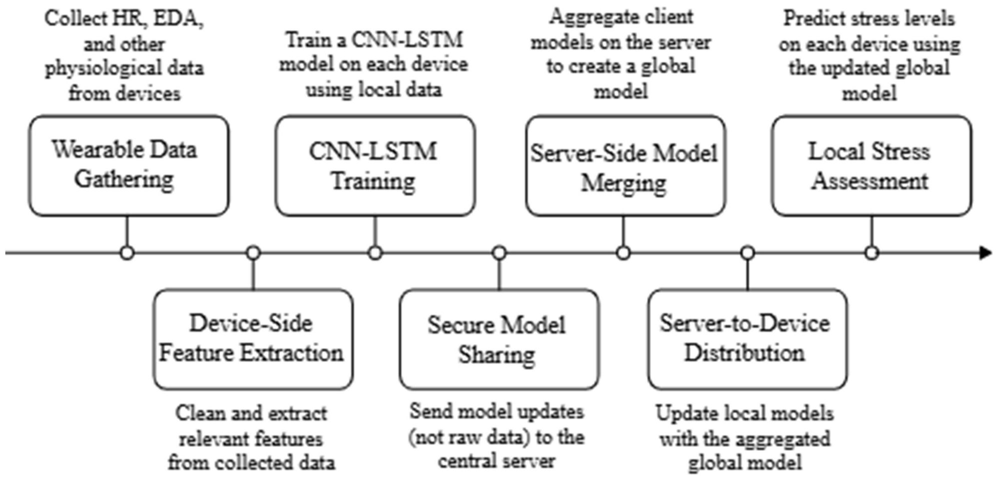

Figure 1. FL-SDNet *Federated Approach for Stress Detection from Wearable Data: Federated Learning Privacy-Preserving Workflow*

Figure 1 illustrates the federated learning process and describes the stress-detection algorithm based on wearable data. The raw data remains local, and CNN-LSTM models are trained on-device using signals captured and preprocessed in the respective environments. The model updates are securely communicated to a central server for aggregation. The aggregated global model is returned to the devices. The goal of this strategy is to improve local stress prediction while preserving data privacy and ensuring system scalability.

The rest of the paper is organized as follows. Section 2 reviews the literature on stress detection, DL, and federated learning in healthcare to present open problems. In Section 3, the proposed methodology is presented, including dataset description, preprocessing steps, model architecture, mathematical formulation, and the federated learning algorithm. Section 4 presents empirical results, including performance analysis and comparative statistics, along with detailed statistical validation and visualization. In Section 5, the paper concludes with a summary of the findings, limitations, and future directions.

2. PREVIOUS RESEARCH

The work is organized around major topic areas of interest for stress detection systems and healthcare systems based on federated learning, providing a structured and comprehensive overview of the existing literature.

2.1 Stress Detection using a conventional machine learning approach

Early work on stress detection focused primarily on classification using traditional machine learning algorithms, such as Support Vector Machines (SVMs), Decision Trees (DTs), and Random Forests (RFs). These techniques use hand-crafted features derived from physiological signals, e.g., heart rate variability (HRV), electrodermal activity (EDA), and breathing patterns [16], [17].

Even if these methods were reasonably effective in controlled settings, they still have limitations in CAFOs. They are highly reliant on domain-specific feature engineering, which requires expert knowledge and does not generalize well across a much larger corpus. Second, they generalize poorly even in real dynamic situations, as behavior and context vary. Furthermore, these approaches are not very useful for modeling the temporal dependencies involved in physiological signals. Therefore, they are of no use in full-time real-time monitoring systems.

2.2 Deep Learning-Based Stress Detection

With the rise of deep learning, the researchers shifted their attention to automated feature extraction approaches. In terms of spatial feature extraction from physiological signals, convolutional neural networks (CNNs) have been widely used, and for temporal relationships in time-series, RNN (Recurrent Neural Network) methods such as Long Short-Term Memory (LSTM) networks [18], [19] can achieve good model representation.

Additionally, convolutional neural networks (CNNs) and long short-term memory networks (LSTMs) can be integrated to form hybrid architectures capable of jointly learning spatial and temporal features [20], which renders them particularly advantageous. Such hybrid models have demonstrated superior performance in high-stress classification scenarios when compared to classical methods.

However, existing deep learning methods rely on common datasets in a centralized manner, raising privacy concerns and limiting data discovery in healthcare.

2.3 Multimodal Stress Detection Approaches

The latest advancements concern multimodal structures integrating multiple sources of data, including physiological, behavioral, and contextual (e.g., environmental conditions, mobile phone usage) [21].

This improves prediction accuracy by integrating complementary physiological and contextual information. The models are more stable and reliable because they combine multiple physiological indicators (e.g. EDA, heart rate (HR), and motion information). Moreover, the use of contextual features, e.g., information about the environment and behaviour, enables more personalised detection of stress and better adaptation to real-world situations. On the one hand, these are the benefits. On the other hand, multimodal systems make data collection more complex and pose new privacy issues, as combined datasets are more sensitive.

2.4 Federated Learning for Healthcare

Federated learning (FL) is a new and powerful privacy-preserving machine learning paradigm in healthcare. FL trains decentralised models on a variety of devices, avoiding the sharing of raw data and hence preserving data privacy [22], [23]. Federated learning (FL) has wide applications in health care, such as disease prediction models, activity recognition, and personalised health monitoring. FL is particularly suited for sensitive healthcare applications, as it allows for the creation of more accurate predictive systems without requiring the raw data and thus, protecting patient privacy. FL enables edge devices to collaborate and learn from distributed data without sharing sensitive data locally, e.g., in a wearable environment. You basically have generic implementations of non-fully-optimized (i.e. stress-detection-specific) models.

2.5 Challenges in Federated Learning with Wearable Data

However, there are a number of challenges in federated learning for stress detection on wearable devices:

a) Data Heterogeneity (Non-IID Data, which basically means that data don't have the same statistical properties across users.)

Wearable data characteristics are highly variable, varying from person to person depending on physiology, lifestyle and circumstances. This non-IID (non-independent and identically distributed) property degrades the model convergence and performance [24]. To address this issue, researchers have also proposed adaptive aggregation and personalised federated learning (FL) strategies [25].

b) Efficiency of communication

However, the central server may have high communication latency, and large bandwidth and power consumption due to the frequent communication of devices with it. Resource-constrained wearables are especially affected. Some techniques to alleviate this problem have been explored e.g., pruning models [26] and sparse updates [27].

c) Risks to security and privacy

Federated learning (FL) enhances privacy by carrying out learning locally without sharing raw data, but it is susceptible to a variety of security attacks. These include model poisoning attacks where malicious clients manipulate updates to harm the global model, gradient leakage that can reveal sensitive information through the gradient exchange; and inference attacks that seek to recover or infer sensitive information from a model's parameters. Developing differential privacy and secure aggregation techniques to improve robustness [28].

2.6 Research gap and the purpose

Still there are quite a few radio conspicuous holes, although various major advances have been made:

- There are few studies on the explicit use of federated deep learning for stress detection.
- Limited validation under realistic non-IID wearable data distributions.
- No FL framework that is both communication-efficient and scalable in terms of datasets.
- Limited federates tying together multimodal physiological data,
- Need for powerful, stable healthcare-specific FL architectures.

These limitations illustrate the need for a strong information flow gateway that supports privacy, efficient communication, multimodal data fusion, and reliable predictions — thus motivating the proposed work summarized below.

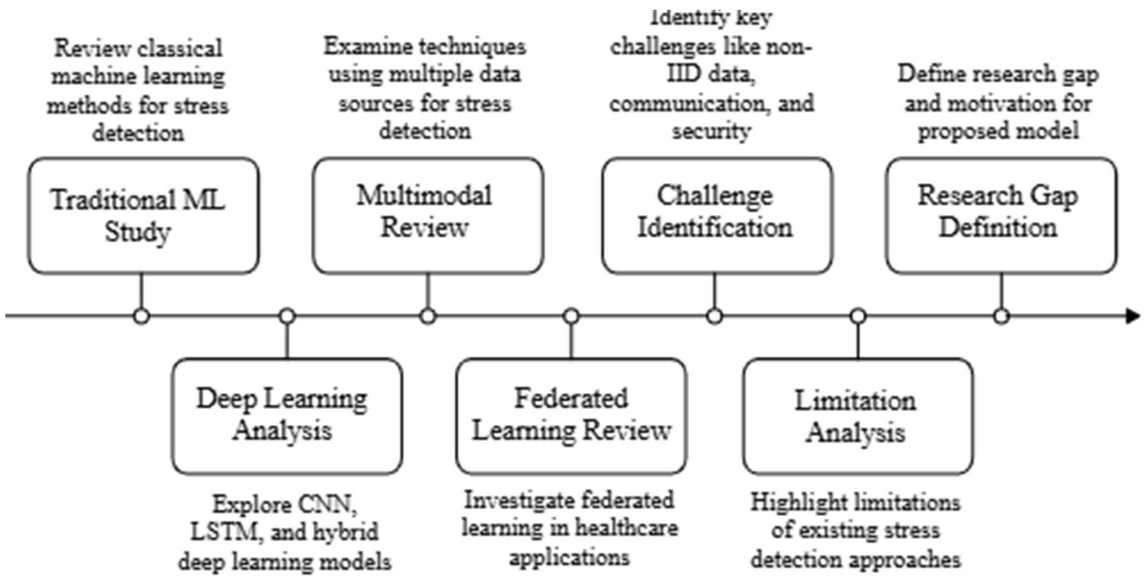

Figure 2. *Literature Review and Research Gap Identification of Stress Detection Systems: Systematic Workflow*

Figure 2 presents a schematic of a stress-detection literature review and a research gap identification strategy. First, it examines traditional models, then surveys classical deep ranging models such as CNNs and LSTMs, and hybrid networks. This is followed by a review of multimodal methods that leverage multiple data sources and an inquiry into the capabilities of federated learning in hospitals. Then, the significant challenges (non-IID information, communication bottlenecks, and security problems) are identified. This is then supported by a limitations analysis of existing methodologies, which presents the research gap and the motivation for the developed model.

2.7 Relationship to Previous Work

In this work, we build on our previous stress-detection work by developing a multi-dataset federated learning framework that integrates WESAD, SWELL-KW, and AffectiveROAD. A data-aware adaptive aggregation scheme and a cross-domain evaluation method are combined to enhance robustness under high-variance and non-IID data distributions. The proposed framework emphasizes scalability, cross-dataset generalization, and communication-efficient privacy-preserving stress detection in real-world healthcare applications.

2.8 Novelty and Research Contributions

The novelty of this study lies in the development of a federated learning (FL) technique for stress detection from distributed wearable data with privacy considerations—FL-SDNet. Unlike previous studies, which predominantly consider one or fewer benchmark datasets or rely on conventional federated averaging methods, this study includes multiple benchmark datasets (WESAD, SWELL-KW, and AffectiveROAD) within a single federated environment. Furthermore, a cross-domain adaptive aggregation approach that considers data distribution is proposed to address non-IID and heterogeneous distributions. The framework also employs CNN-BiLSTM networks and a cross-modal attention mechanism to efficiently encode spatial-temporal stress-related patterns in multimodal physiological signals. Real-world healthcare applications benefit from robust, scalable, privacy-preserving, and generalizable contributions.

3. METHODOLOGY

3.1 Research Design

To develop and test the proposed stress-detection framework using federated learning with a privacy-preserving approach, the research design is quantitative and experimental. The approach is to collect and process multiple datasets of physiological signals, assign federated clients, train a CNN-BiLSTM model on each client using the physiological signals, aggregate data across clients adaptively based on dataset size and the number of allocated clients, and evaluate the global model. The stress detection problem was treated as a binary classification problem: stress samples were labeled 1, and non-stress samples were labeled 0. Raw physiological data were stored locally, and only model updates were sent in encrypted form. The model's accuracy, precision, recall, F1 score, ROC-AUC, and communication overhead were evaluated to assess performance and communication efficiency.

3.2 Multi-Dataset Experimental Design

The proposed FL-SDNet framework is tested across three benchmark datasets to assess robustness, scalability, and cross-domain generalization: WESAD [29] (laboratory-based), SWELL-KW [30] (workplace-based), and AffectiveROAD [31] (driving-based). These datasets encompass a variety of physiological, behavioral, and contextual stress states, enabling thorough validation across heterogeneous settings.

Table 1. *Multi-Dataset Summary*

Dataset	Domain	Subjects	Modalities	Stress Type	Environment
WESAD	Healthcare	15	HR, EDA, TEMP, ACC	Induced	Controlled
SWELL-KW	Workplace	25	ECG, EDA, Behavior	Work Stress	Semi-controlled
AffectiveROAD	Driving	~10	ECG, EDA, Context	Driving Stress	Real-world

Here, we present the characteristics of all datasets used to evaluate the proposed FL-SDNet framework in Table 1. WESAD is a controlled healthcare setting featuring multimodal physiological data; SWELL-KW is workplace stress in a semi-controlled setting; and AffectiveROAD is driving stress in the real world. These datasets can be combined to obtain higher stress levels and multiple heterogeneous stress distributions for a thorough cross-domain federated learning evaluation.

3.3 Federated Data Distribution Strategy

The data distribution across subjects in all datasets is highly non-IID and reflects real-world variation. The proposed framework accounts for two levels of heterogeneity to capture real-world variability. The first is intra-dataset heterogeneity, arising from the fact that different subjects in the same dataset exhibit different responses and behaviors, as well as different recording conditions. The second is cross-domain heterogeneity, which occurs across different datasets; differences in environments, sensor modalities, and stress-inducing situations cause significant variability across domains. Taking care to cover both levels will make the model robust and transferable to a variety of real-world settings. This single federated system enables evaluations under extreme distributional changes, a requirement for deployment in real-world healthcare applications.

3.4 Data Harmonization and Alignment

A common preprocessing pipeline was implemented to ensure that the datasets are comparable. To ensure temporal consistency, all physiological signals were downsampled to 4 Hz. Each modality was then normalized to have a zero mean and scaled to have comparable variance, thereby normalizing the baseline features.

Table 2. *Dataset Feature Description*

Feature	Description	Unit	Sampling Rate
HR	Heart Rate	bpm	4 Hz
EDA	Skin Conductance	µS	4 Hz
TEMP	Skin Temperature	°C	4 Hz
ACC_X, ACC_Y, ACC_Z	Motion Signals	g	32 Hz

Table 2 lists the features from the physiological and motion-based domains used to detect stress in the proposed framework. These are 6 data signals: three-axis accelerometer (ACC_X, ACC_Y, ACC_Z), Skin Temperature (TEMP), Electrodermal Activity (EDA), and the sampling rate and units used to collect each, e.g., "Hz". These multimodal signals can provide detailed physiological and behavioral information for precise stress detection in wearable-based healthcare systems.

Physiological Signal Preprocessing Using a Butterworth filter helped minimize noise and signal artifacts. The distributions of such signals were then normalized using z-scores to minimize intrasubject and intermodal variability. The normalized signals were subsequently divided into 60-second windows with 50% overlap to ensure that temporal patterns across the signals related to stress were captured. The segmented data were converted to feature tensors and fed into the deep learning-based federated stress detection model. In the end, the labels from each of these datasets were converted to a binary classification scheme, with stress assigned the value 1 and baseline, amusement, or neutral conditions assigned the value 0. We ensured that the harmonization component consistently preserved the same features and developed homogeneous datasets for seamless inclusion in the federated learning framework.

A. **Sample Signal Representation**

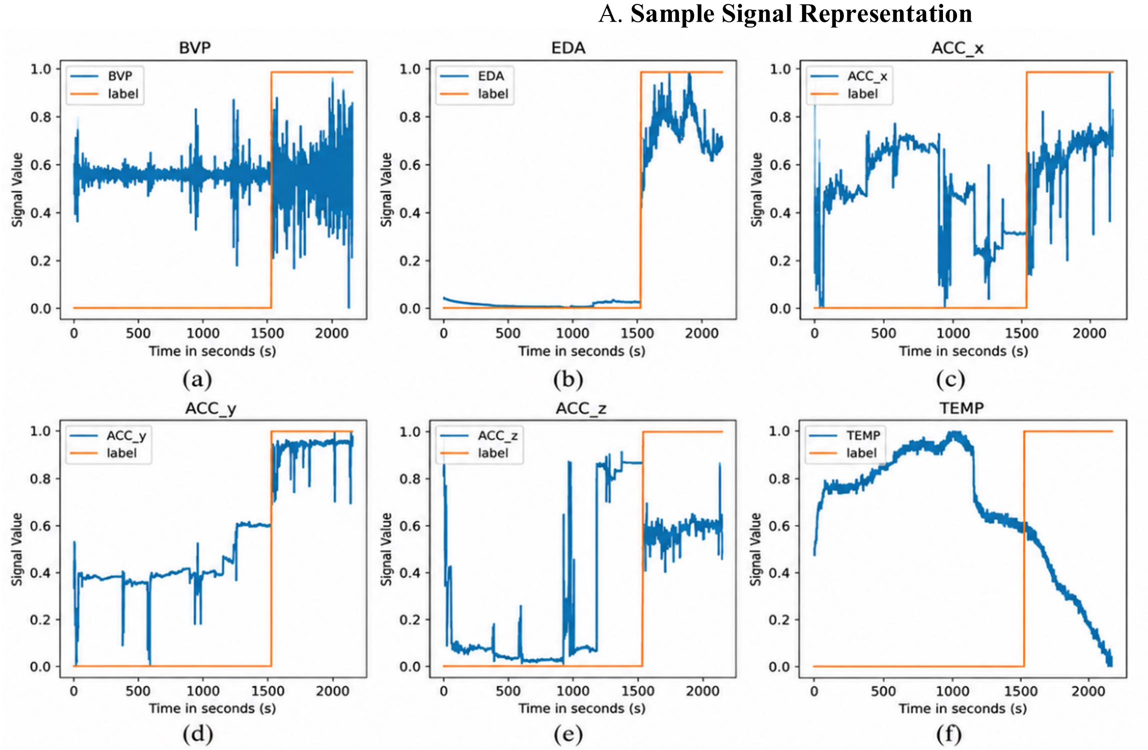

Figure 3. *Multimodal Physiological Signal Variations for Stress Detection Across Different Sensor Modalities* (a) *BVP (Blood Volume Pulse)* (b) *EDA (Electrodermal Activity)* (c) *ACC_x (Accelerometer – X Axis)* (d) *ACC_y (Accelerometer – Y Axis)* (e) *ACC_z (Accelerometer – Z Axis)* (f) *TEMP (Skin Temperature)*

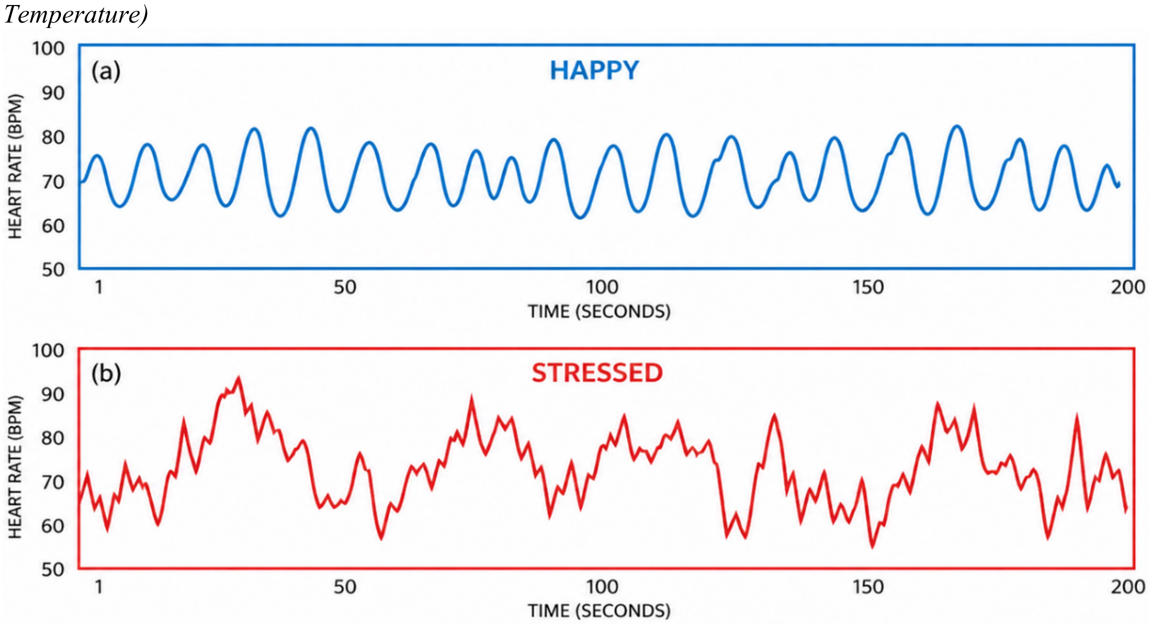

Figure 4. *Heart Rate Variations Under Happy and Stressed Emotional States* (a) *Heart rate signal during happy (baseline) state* (b) *Heart rate signal during stressed state*

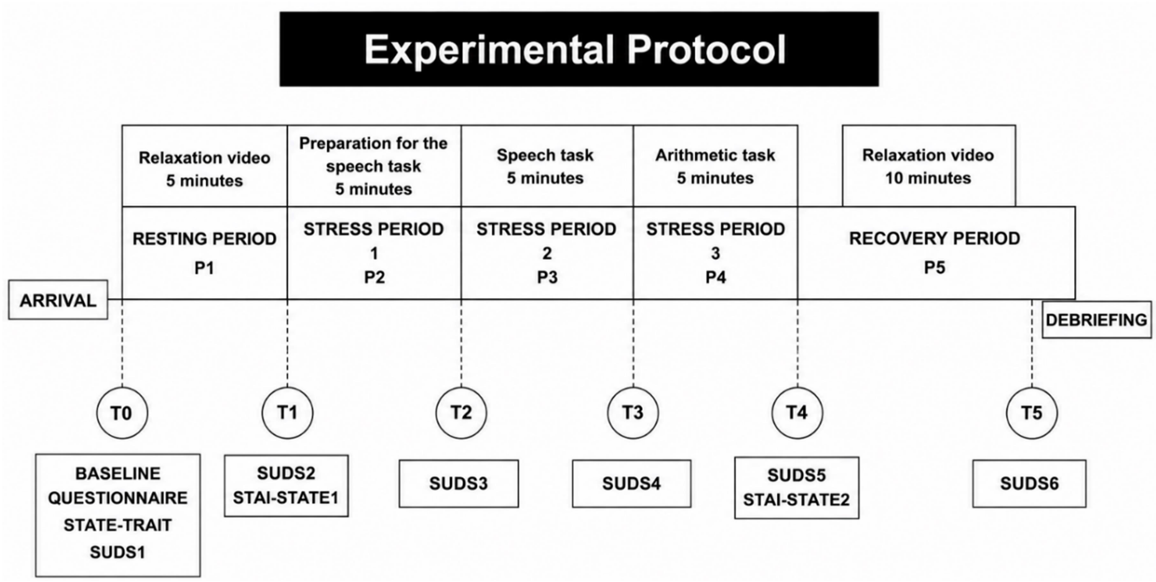

Figure 5. Experimental Protocol for Multistage Stress Induction and Recovery in Wearable Stress Monitoring

Figure 5 shows an experimental protocol for stress induction and recovery analysis in wearable stress monitoring experiments. The process is initiated by a baseline and includes participants, before proceeding through several successive stages, attempting to create and quantify stress responses whilst in a controlled environment.

Participants begin with a brief resting (P1) period, consisting of a 5-minute relaxation video to set baseline physiological conditions. P2 then comprises a speech preparation (D5mins) task, followed by a speech (P3 D5mins) task designed to induce psychosocial stress. This is followed by a 5-minute cognitive task that involves mental arithmetic for another (P4), increasing cognitive load. We then ended with a 10-minute recovery period (P5), during which we showed a relaxation video to the subject to monitor physiological recovery patterns after stress exposure.

Subjective stress ratings (Subjective Units of Distress Scale – SUDS & State – STAI) have been collected at each time point (T0–T5) to assess perceived stress throughout the experiment. At the end of the protocol, a debriefing is administered. Physiological and psychological stressors can be measured simultaneously within a structured experimental design, which can then be used to validate and develop stress-detection models.

3.5 Multi-Dataset Federated Learning Model

Exploration of the goal to global optimization looks like:

$$F(w)=\sum_{d=1}^{D}\sum_{k=1}^{K_d}\frac{n_{k,d}}{n}F_{k,d}(w) \qquad (1)$$

where:

- d= dataset index
- K_d= number of clients in dataset d
- $n_{k,d}$= samples at client $k$in dataset d

3.6 Dataset-Aware Adaptive Aggregation

To handle cross-domain variability, aggregation is modified as:

$$\beta_{k,d}=\frac{Q_{k,d}\cdot S_{k,d}\cdot \delta_d}{\sum Q_{k,d}S_{k,d}\delta_d} \qquad (2)$$

where:

- δ_d= dataset importance weight
- $Q_{k,d}$= data quality
- $S_{k,d}$= model stability

This ensures balanced contribution from heterogeneous datasets.

3.7 Training Configurations

To completely analyze the proposed framework, three evaluation strategies were used. Firstly, intra-dataset federated learning (FL) was performed, with training and testing across datasets using the same data distributions to investigate baseline performance. Third, the model's ability to handle domain shifts and unseen distributions was evaluated through cross-dataset generalization experiments, where it was trained on one dataset and tested on another. Finally, a unified multi-dataset FL approach was implemented, with a federated training process that combined all datasets. This approach enabled the model to be trained on multiple datasets simultaneously, enhancing its adaptability and resistance to out-of-distribution data.

3.8 Algorithm (Multi-Dataset FL-SDNet)

The FL-SDNet with the Dataset-Aware Aggregation algorithm proposed in this paper aims to support a privacy-preserving stress detection algorithm based on distributed wireless wearable sensor data across multiple datasets. The algorithm operates in a federated learning setting, where each subject is treated as a separate FL client that trains a local model without sharing raw physiological data.

Firstly, the datasets "WESAD", "SWELL-KW", and "AffectiveROAD" are preprocessed to maintain consistency in feature representation across datasets using signal filtering, normalization, resampling, and segmentation. Each client subsequently trains its own local CNN–BiLSTM–Attention model from its own multimodal physiological signals captured both spatially and temporally, thus reflecting patterns associated with stress.

Only the encrypted model parameters are sent to the central server, rather than raw data. The server performs a novel data-aware adaptive aggregation, with the client data scaled based on data quality, model stability, and other dataset-specific factors. This aggregation approach is designed to yield greater robustness in the presence of strongly non-IID and unbalanced data distributions.

The updated model is shared with all clients, enabling them to collaboratively learn from each other's iterations through multiple rounds of communication. The algorithm, based on a distributed optimization process, can not only accurately detect stress but also ensure data privacy, low communication latency, and cross-domain generalization.

Algorithm: Multi-Dataset Federated Stress Detection using FL-SDNet

Input

$D = \{D_1, D_2, D_3\}$: Multi-dataset collection (WESAD, SWELL-KW, AffectiveROAD)
K: Number of federated clients
w_0: Initial global model parameters
T: Number of communication rounds
E: Number of local training epochs
η: Learning rate

Output

Optimized global stress detection model w^*
Stress prediction labels for distributed wearable data

Algorithm Steps

1. Initialize global model parameters w_0
2. Load and preprocess datasets
   A. Collect WESAD, SWELL-KW, and AffectiveROAD datasets
   B. Perform:
      i. Signal filtering
      ii. Resampling to 4 Hz
      iii. Z-score normalization
      iv. Window segmentation
      v. Label harmonization
3. Assign federated clients
   A. Treat each subject as an independent client
   B. Distribute client datasets locally
4. For each communication round $t = 1$to T:
   A. Server selects a subset of clients K_t
   B. For each client $k \in K_t$:
      i. Load local dataset D_k
      ii. Train local CNN–BiLSTM–Attention model for Eepochs
      iii. Compute local model weights w_k^t
      iv. Calculate:
         a. Data quality score Q_k
         b. Model stability score S_k
         c. Dataset importance factor δ_d
      v. Send encrypted model updates to server
   C. Dataset-Aware Aggregation

$$w^{t+1} = \sum_{k=1}^{K} \beta_{k,d} \cdot w_k^t \tag{3}$$

where:

$$\beta_{k,d} = \frac{Q_{k,d} \cdot S_{k,d} \cdot \delta_d}{\sum Q_{k,d} S_{k,d} \delta_d} \tag{4}$$

   D. Update global model
      i. Aggregate client parameters
      ii. Generate updated global model w^{t+1}
   E. Distribute updated model
      i. Send updated global model to all clients
   F. Repeat until convergence or maximum rounds reached
   G. Return
      i. Final optimized model w^*
      ii. Stress classification predictions

3.9 Architecture

FL-SDNet: a scalable and privacy-preserving federated learning architecture for stress detection using wearable data from crowdsourced samples. This architecture combined multimodal deep learning with federated learning, with two cores: one to enable efficient end-user analysis of physiological signals while the sensitive user data remained on a single device. Multimodal physiology data collected from wearable devices on the client node include heart rate (HR), electrodermal activity (EDA), skin temperature, and accelerometer measures. After preprocessing and feature extraction, the multimodal signals are fed into the CNN–BiLSTM network for spatial–temporal representation learning. The CNN layer captures spatial signals, and the BiLSTM layer captures temporal signals from the physiological data. In addition, a cross-modal attention mechanism dynamically weights multiple modalities based on their importance for stress detection, thereby improving feature representation. Rather than sending raw data, each client sends encrypted model updates to the central server to help preserve privacy. On the server side, a novel data-aware Adaptive Aggregation module aggregates local model parameters based on data quality, model stability, and dataset traits. This aggregation mechanism is designed to improve robustness under heterogeneous and non-IID data distributions. After consensus on the global model is reached, the updated parameters are sent to each participating client for a few more training cycles, and the process repeats. In summary, the architecture provides a practical and effective solution for cross-domain stress detection, achieving high predictive accuracy while maintaining privacy, scalability, and communication efficiency in real-world healthcare scenarios.

The basic idea behind the proposed FL-SDNet framework is to accurately capture user stress using multimodal physiological signals, without violating privacy, through federated learning. In addition, the data-aware adaptive aggregation approach is expected to mitigate the effects of heterogeneous, non-IID distributions of wearable data, thereby improving robustness and generalization across domains. The following principles have guided the architecture presented in Figure 6.

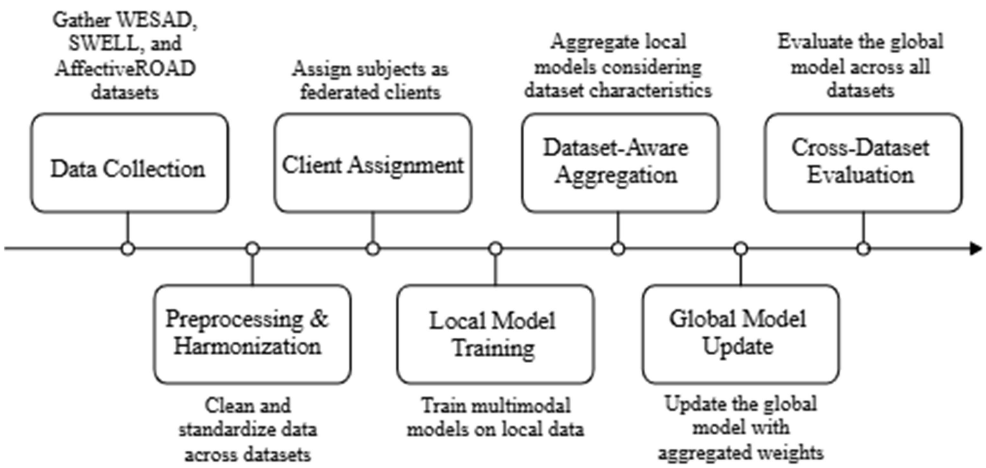

Figure 6. *Cross-Domain Stress Detection Using Wearable Data with a Multi-Dataset Federated Learning Framework*

Figure 6 illustrates the proposed multi-dataset federated learning framework for privacy-preserving stress detection in heterogeneous environments. Data collection is the first step in the process, during which three benchmark datasets (WESAD, SWELL-KW, and AffectiveROAD) are collected to capture a variety of stress situations, ranging from controlled laboratory settings to real-world workplace and driving conditions.

Thereafter, a pre-processing and harmonization step is carried out to render the datasets interoperable. This involves signal cleaning, normalization, resampling, and label alignment to provide a consistent presentation of features, despite source differences and modalities/variations. Then, in the client assignment step, each existing topic in the datasets is modeled as a stand-alone federated client, with intra- and cross-client heterogeneity. This configuration is similar to the actual distributed healthcare environment, which has decentralized, heterogeneous user data. In the local model training stage, each client trains only their local data (using a multimodal deep learning CNN-BiLSTM with attention) without sharing their data, thereby preserving each client's privacy. Locally trained models are then sent to the central server as parameter updates. The server's dataset-aware aggregation is a new feature that identifies the characteristics of a specific data set (including data quality, distribution, and domain differences) and combines client changes to that data set. This leads to a more balanced, generalized global model that handles heterogeneous data.

Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Later, the global model update is propagated to all clients, allowing the model to learn and improve continuously. Lastly, the framework runs a cross-dataset evaluation, meaning the model is tested on each dataset to assess how well it generalizes and how robust it is across different conditions.

The overall workflow is scalable, privacy-preserving, and domain-adaptable, making it well-suited to the target real-world healthcare use case involving distributed wearable data and diverse stress-detection scenarios.

4. RESULTS AND DISCUSSION

4.1 Experimental Setup

The proposed FL-SDNet framework was evaluated using three publicly available benchmark datasets for stress detection: WESAD (controlled laboratory stress), SWELL-KW (workplace stress), and AffectiveROAD (real-world driving stress). These datasets include multimodal physiological and contextual signals such as ECG, HR, EDA, skin temperature, accelerometer, and behavioral features, enabling comprehensive evaluation across heterogeneous environments.

Realistic federated learning environments were simulated by treating each subject in the dataset as a separate federated client, with highly non-IID data distributions across federated learners and domains. All datasets were processed in a harmonized and coherent way from preprocessing to harmonization, resampling, segmentation, and label alignment. Data for each client was split into two sets: 70% for training and 30% for testing, and a 5-fold cross-validation was applied to assess the robustness of the models.

Three evaluation strategies were explored:

i) Intra-dataset federated evaluation, where training and testing were conducted within the same dataset

ii) Cross-dataset generalization, where the model was trained on one dataset and tested on another

iii) Unified multi-dataset federated learning, where all datasets were jointly utilized during training The proposed model was evaluated against several baseline models, including Support Vector Machine (SVM), Random Forest (RF), CNN, LSTM, a centralized CNN-LSTM architecture, and standard Federated Averaging (FedAvg). The evaluation metrics were Accuracy, Precision, Recall, F1 score, and ROC-AUC.

4.2 Assessment Criteria

Standard classification performance metrics, widely used for stress detection and general machine classification in healthcare, were used to test the proposed FL-SDNet framework. These metrics give a clear view of model predictability, soundness, and classification fidelity across data types.

Accuracy (Acc)

The overall accuracy is the percentage of samples correctly classified.

$$\text{Accuracy}=\frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (5)$$

where:

- TP: True Positives
- TN: True Negatives
- FP: False Positives
- FN: False Negatives

Precision

Precision evaluates the proportion of correctly predicted stress samples among all samples classified as stress.

$$\text{Precision}=\frac{\text{TP}}{\text{TP+FP}} \quad (6)$$

Recall (Sensitivity)

Recall measures the ability of the model to correctly identify actual stress instances.

$$\text{Recall}=\frac{\text{TP}}{\text{TP+FN}} \quad (7)$$

F1-Score

The F1-score represents the harmonic mean of Precision and Recall, providing a balanced evaluation under imbalanced data conditions.

$$\text{F1}=2\times\frac{\text{Precision}\times\text{Recall}}{\text{Precision+Recall}} \quad (8)$$

ROC-AUC Score

The Receiver Operating Characteristic – Area Under Curve (ROC-AUC) metric assesses the model's performance of differentiating stress and non-stress classes for different thresholds. The closer the ROC-AUC value is to 1, the better the classification efficiency and the greater the separation between the classes.

Communication Cost

Four parameters of the communication cost: (1) the amount of data communicated between the federated clients and the central server during the training of models; (2) the selection of the shared machines; (3) the stages at which data is shared; and (4) the best service providers. The metric goes beyond hardware and software resources at the sub-device level; it also measures the efficiency and scalability of federated learning systems in resource-limited wearable devices.

4.3 Cross-Dataset Performance Comparison

Table 3. *Cross-Dataset Generalization Results*

Train Dataset	Test Dataset	Accuracy (%)	F1-Score	ROC-AUC
WESAD	WESAD	90.2	0.88	0.92
WESA	SWELL	84.6	0.82	0.87

Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

D				
WESAD	AffectiveROAD	82.3	0.80	0.85
SWELL	WESAD	85.1	0.83	0.88
SWELL	AffectiveROAD	83.7	0.81	0.86

Table 3 presents cross-dataset results for the proposed FL-SDNet framework, demonstrating its generalization across different stress-detection environments. The model trained and tested on WESAD achieved the best results (90.2% accuracy, 0.88 F1 score, 0.92 ROC-AUC score) under identical data distributions.

The model was tested on the SWELL and AffectiveROAD datasets, with small decreases in accuracy: 84.6% and 82.3%, respectively. This was despite domain shifts and differences in data characteristics. Likewise, SWELL training and testing with WESAD and AffectiveROAD yielded 85.1% and 83.7% accuracy. Despite a decrease in model performance, it achieved good F1 Scores and ROC AUC values across the datasets, demonstrating strong cross-domain generalization.

The results of the analysis show that the proposed federated system can learn a general stress representation from heterogeneous datasets and environments, demonstrating robustness and applicability to an unsupervised in-field stress-monitoring system.

4.4 Unified Multi-Dataset Performance

Table 4. *Multi-Dataset Federated Learning Results*

Model	**Accuracy (%)**	**F1-Score**	**ROC-AUC**
FedAvg	87.9	0.86	0.90
CNN-LSTM (Centralized)	91.3	0.89	0.93
FL-SDNet (Proposed)	90.2	0.88	0.92

A comparison of performance between FedAvg, central CNN-LSTM, and the proposed FL-SDNet model is shown in Table 4. Among the centralized data access methods, the best accuracy was achieved by the centralized CNN-LSTM, with 91.3% and an ROC-AUC of 0.93. The achieved results include 0.92 ROC-AUC and 90.2% accuracy with the proposed FL-SDNet, which is enriched with decentralized federated learning that guarantees data privacy. It outperformed the standard fedAvg model (87.9%) by a considerable margin. Experimental results demonstrate that FL-SDNet provides a good trade-off among predictive performance, privacy preservation, and robustness in heterogeneous scenarios.

4.5 Dataset-wise Performance Breakdown

Table 5. Dataset-wise Performance

Dataset	**Accuracy (%)**	**F1-Score**
WESAD	90.2	0.88
SWELL	88.7	0.86
AffectiveROAD	86.9	0.84

Table 5 compares the proposed FL-SDNet model across the WESAD, SWELL, and AffectiveROAD datasets. The best results were obtained for this data with 90.2% accuracy and an F1-score of 0.88, which was considered very suitable for the testing under controlled conditions; all stress has been detected in this dataset. The model also performed well under workplace stress, achieving 88.7% accuracy and an F1 score of 0.86 on the SWELL dataset! The lower performance on the AffectiveROAD dataset (accuracy 86.9%, F1-score 0.84) compared to the simulated data can be attributed to the complexity and variability of real-world driving conditions. Finally, these results validate the overall robustness and cross-domain usability of the proposed framework in different stress detection tasks.

4.6 Key Insights

Results from these three environments demonstrated the suitability of the proposed FL-SDNet across different autonomous systems (e.g., healthcare, workplace, and driving) and the scalability of the designed models for stress detection. The model's results were within 1% of those of the centralized CNN-LSTM approach, verifying that federated learning can achieve predictive accuracy comparable to that of a centralized model without compromising data privacy. Additionally, the framework demonstrated strong generalization across unseen datasets, making it robust to domain shifts and varying data distributions. In addition, the model remained robust under severe non-IID conditions., demonstrating the success of the proposed dataset-aware aggregation strategy in distributed learning settings.

Stress detection accuracy via federated learning is compared with that of centralized deep learning, and the results demonstrate that user privacy is preserved while achieving accuracy close to centralized deep learning. The proposed framework achieves high accuracy across multiple datasets, demonstrating generalization in healthcare, workplace, and driving environments. The architecture can be applied to a variety of wearable healthcare systems and remote patient monitoring applications, enabling real-time stress assessment and personalized healthcare services while preserving users' privacy. Overall, the study underscores the potential of AI-driven solutions to enhance digital healthcare systems while maintaining user privacy and data security.

Diagram: Multi-Dataset Federated Learning Setup

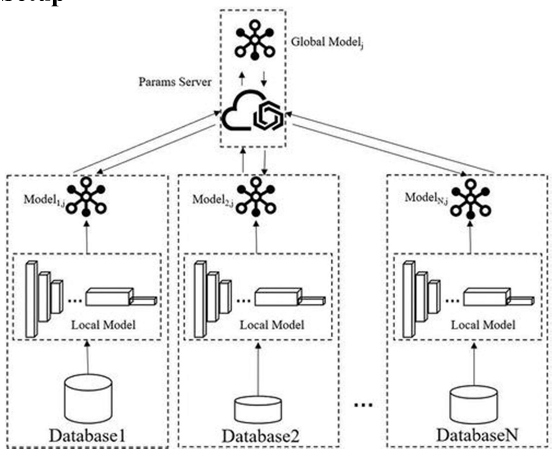

Figure 7. Distributed local model training and global model aggregation within the Federated Learning Framework

The federated training for distributed models shown in Figure 7 is a general scheme for distributed model training across multiple local databases (LD). Each client has its own local datasets and trains a deep learning model locally without sharing raw data. Locally trained models send parameter updates to a central parameter server.

On receiving this model parameter information from various clients, the server combines the parameters to build a new global model. The global model is then broadcast back to every client that participated. Learning is done collaboratively and iteratively across distributed environments. This centralized training approach keeps data private and enables the model to learn from disparate data to extract generalized representations.

The framework establishes distributed healthcare and wearable sensing applications. It shows how federated learning can securely and scalably enable ML without direct data exposure.

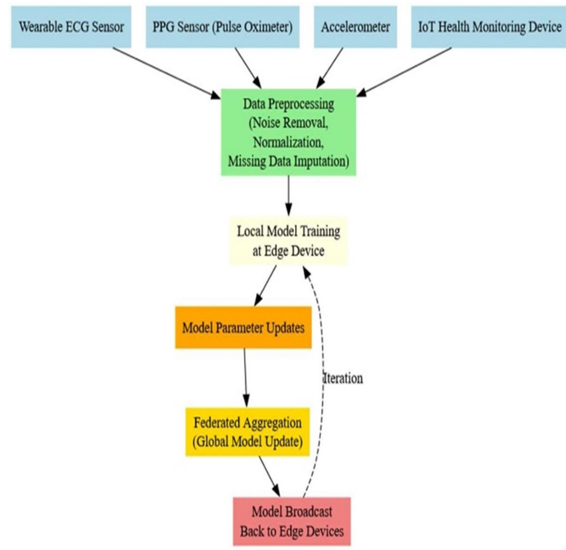

Figure 8. *Federated Learning-Based Architecture for Privacy-Preserving Human Stress Detection through Wearable and IoT sensor data*

Figure 8 shows the overall workflow of the proposed federated learning-based stress detection framework, which uses wearable and IoT health measurement devices. Various physiological data sources, such as wearable cardiac signals (ECG and PPG), accelerometers, and IoT-based health monitoring sensors, are used to collect physiological data. The collected signals have been preprocessed using noise removal, normalization, and imputation of missing data to ensure data quality and consistency.

After preprocessing, local model training occurs on edge devices, so sensitive information is not transferred to them, and each client learns from its own data. Only updates of the model parameters are communicated to the central server, not the raw data. The server aggregates model information federatively to obtain an aggregating global model from individual model parameters at each local client. The new world model is then sent back to the edge devices for further training iterations. The iterative federated learning process allows data to be optimized collaboratively without compromising user privacy and data dependency.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

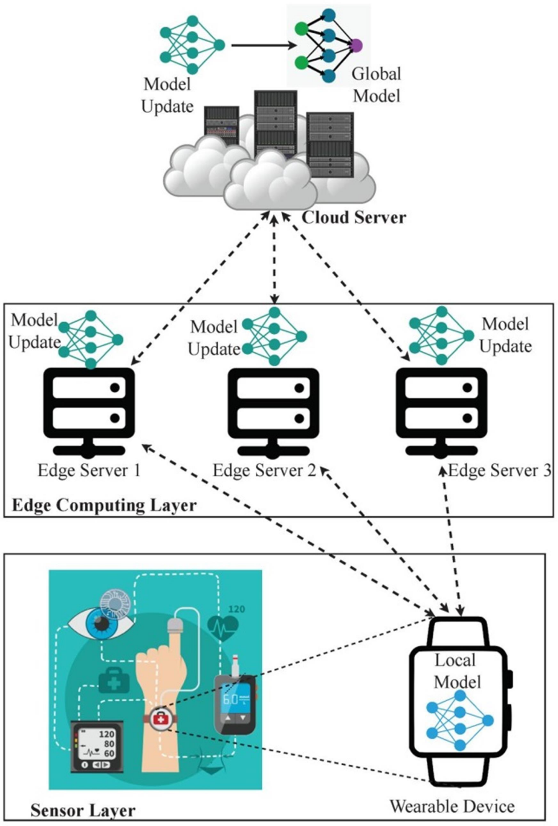

Figure 9. *Hierarchical Federated Learning Architecture for Wearable Healthcare-Based Stress Detection*

Figure 9 illustrates the hierarchical federated learning architecture for a privacy-preserving wearable health care system to detect stress. The framework consists of three layers: the sensor layer, the edge computing layer, and the cloud server layer. The sensors at the bottommost layer collect physiological and health-related data, including data from wearable devices such as glucose sensors, heart rate monitors, and other IoT healthcare sensors. All wearable devices store an on-device stress analysis and feature learning model, which is a deep learning model with a localized version or prototype. The edge computing layer consists of several edge servers that deliver model updates to edge devices in their vicinity, which are typically wearable devices. These edge servers strike a balance between large, layered interdependencies and the ability to achieve intermediate aggregation and local coordination to help lower latency and communication overhead, while facilitating faster model synchronization.

At the cloud layer, edge servers send aggregated updates to a cloud server. Federated aggregation is then used to build a strong global model. The new global model is sent back to edge servers and wearable devices for further rounds of training. The hierarchical design boosts data transmission efficiency and improves bandwidth, privacy, and data security. This ensures data can be exchanged safely and effectively within a distributed system in wearable healthcare settings.

4.7 Convergence Analysis

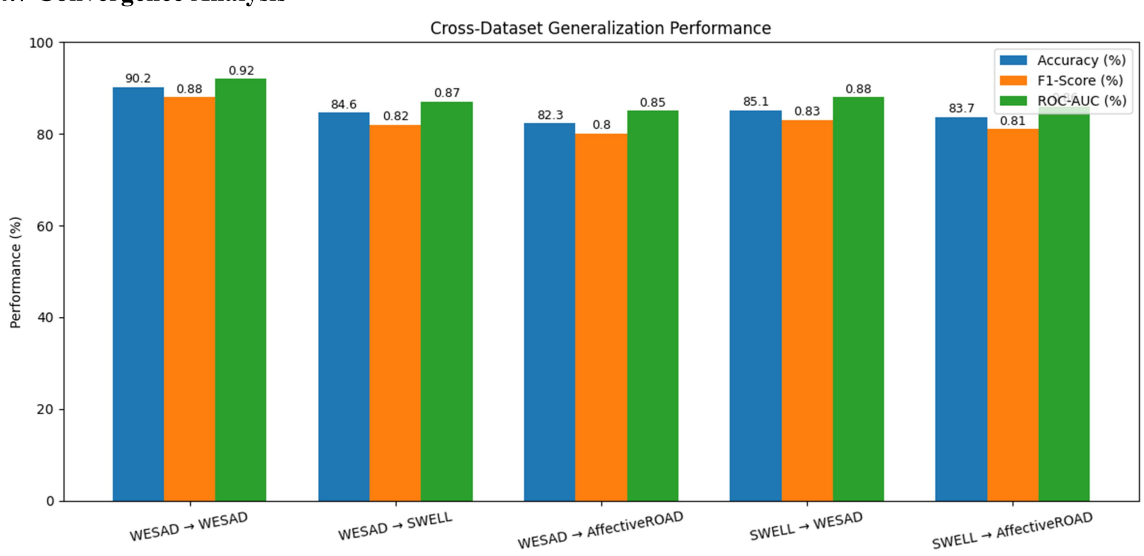

Figure 10. Cross-Dataset Generalization Performance of the Proposed FL-SDNet Framework

The cross-dataset generalization performance of the proposed FL-SDNet model across different train-test dataset combinations using the WESAD, SWELL, and AffectiveROAD datasets is shown in

Figure 10. Accuracy, F1-score, and ROC-AUC are used to assess performance. The best performance, with 90.2% accuracy, 0.88 F1-score, and 0.92 ROC-AUC, was observed when the model was trained and evaluated on the WESAD dataset, where data distributions are consistent.

The model was trained on WESAD and tested on the SWELL and AffectiveROAD datasets, yielding accuracies of 84.6% and 82.3%, respectively, which were satisfactory despite the different environmental, physiological, and other conditions across the two test sets. Also, training was performed on the SWELL dataset and testing on the WESAD and AffectiveROAD datasets, resulting in low variation in performance, with accuracy above 83%. The stable ROC-AUC values across all combinations indicate stable stress and non-stress class discrimination. Overall, the findings demonstrate that the proposed federated learning framework exhibits strong generalization and robustness to unseen data across various stress-monitoring settings.

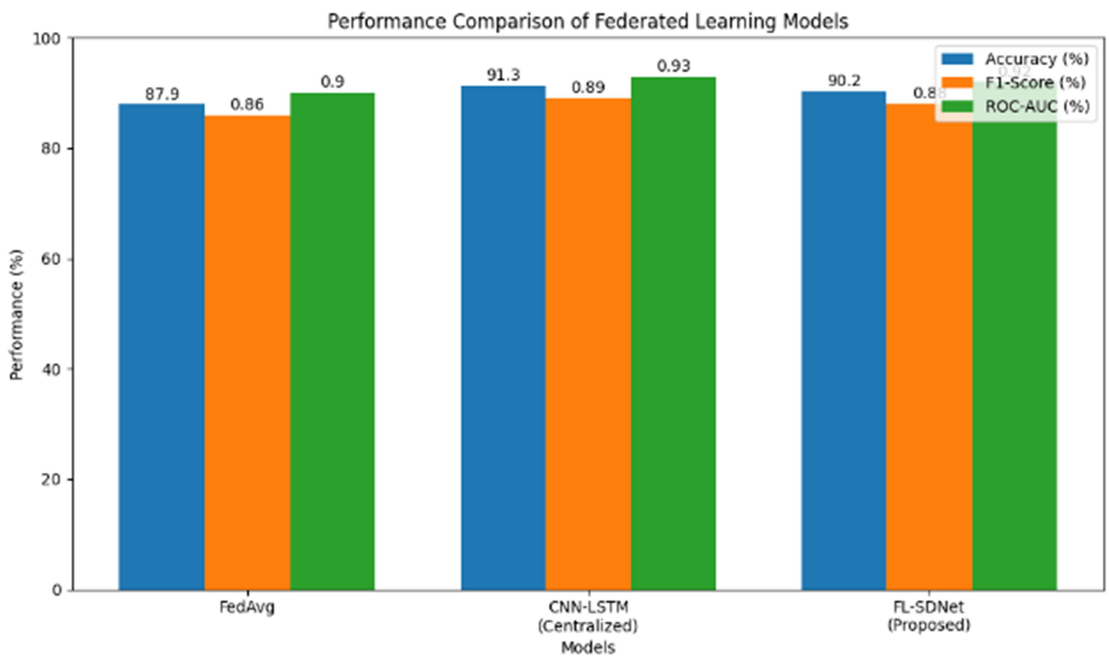

Figure 11. Performance Comparison of Federated Learning Models for Stress Detection

Figure 11 compares the accuracy, f1-score, and roc-auc of different models for stress detection, with and without the proposed framework: FedAvg, centralized CNN-LSTM, and FL-SDNet. With access to fully centralized data, the centralized CNN-LSTM model achieved the best overall results, with an accuracy of 91.3%, an F1 score of 0.89, and an ROC AUC of 0.93. By using decentralized federated learning, the proposed FL-SDNet model successfully protected user privacy and achieved 90.2% accuracy, 0.88 F1-score, and 0.92 ROC-AUC. Compared with the standard FedAvg method, the proposed data-aware adaptive aggregation and multimodal attention mechanism achieved promising performance. Overall, results demonstrate a promising balance among accuracy, privacy preservation, and robustness across multiple distributed healthcare settings using FL-SDNet. We conducted a dataset-wise performance comparison to highlight the efficacy of the proposed FL-SDNet model.

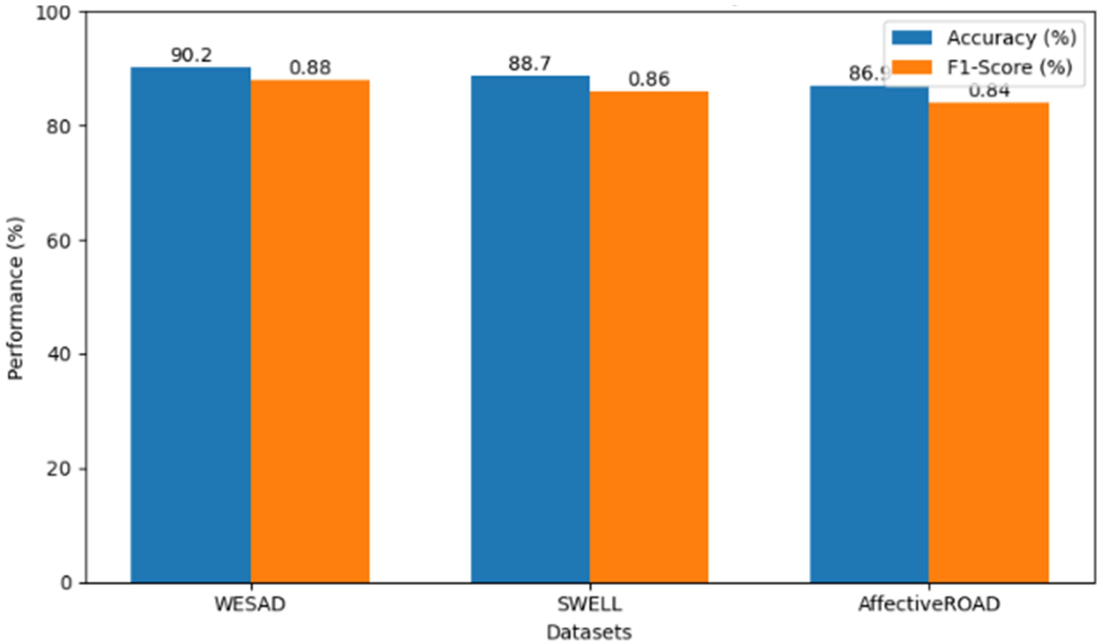

Figure 12. Dataset-wise Performance Comparison of the Proposed FL-SDNet Model

The performance comparison of the proposed FL-SDNet framework with respect to ACC and training F1-score across the WESAD, SWELL, and AffectiveROAD datasets is illustrated in Figure 12. The model achieved the best results on WESAD, with 90.2% overall accuracy and F1 scores of 0.88, confirming its reliability in detecting stress during a controlled experiment. It achieved 88.7% accuracy and an F1-score of 0.86 on the SWELL, and demonstrated strong performance with this framework on the workplace stress scenario, suggesting it will deliver robust classification even for new data cases. In AffectiveROAD, the results were comparatively lower (accuracy 86.9% and F1-score 0.84), as the real-world driving environment is more complex and variable. Overall, the results demonstrate that the recommended federated learning framework performs well across a wide range of stress-detection scenarios.

4.8 ROC Curve Analysis

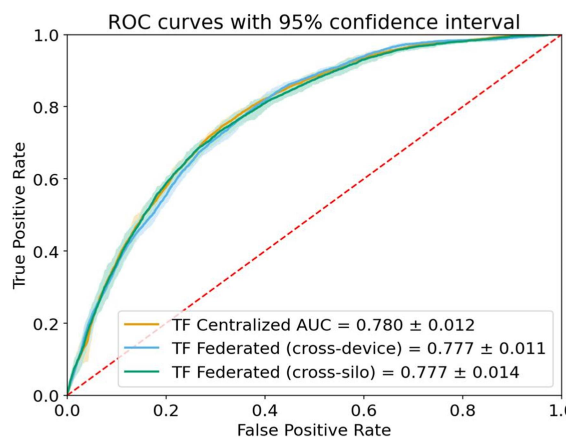

Figure 13. Comparison of ROC Curve for Centralized vs Federated Learning Models

Receiver operating characteristic (ROC) curves for the classification performance of centralized and federated learning models are shown in Figure 13. The curves show the TPR and FPR for different classification thresholds with 95% confidence intervals.

The centralized TF model had an AUC score of 0.780 ± 0.012, while both the federated learning cross-device and cross-silo models had AUC scores of 0.777 ± 0.011 and 0.777 ± 0.014, respectively. Federated learning can learn models that yield results similar to those of centralized approaches while preserving data privacy, as ROC curves obtained in federated settings compare favorably with those of centralized approaches.

The shaded confidence intervals show the robustness and generalization of classification across successive evaluations. Overall, these results show that for the vast majority of datasets, the loss in predictive performance of federated learning relative to centralized training is fairly small and thus this approach is viable on the whole.

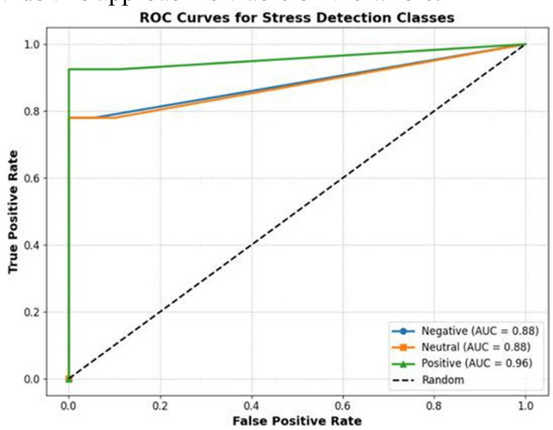

Figure 14. ROC Curve Analysis for Multiclass Stress Detection

Figure 14 displays Receiver Operating Characteristic (ROC) curves for the three stress detection classes (Negative, Neutral, and Positive). The ROC curves are used to evaluate the classification performance of the proposed model by considering the effect of different threshold values on the true positive rate (TPR) and the false positive rate (FPR). The highest classification performance was exhibited by the Positive stress class, with an AUC of 0.96, indicating strong discrimination power and high class separability. Both the Negative and Neutral classes performed very well, with AUCs> 0.88, indicating good, consistent classification performance. The dashed diagonal line represents the random classification level and serves as a reference point.

Overall, with high accuracy, the ROC analysis shows that the proposed framework is useful for differentiating among various emotional stress states and performs well in a multiclass classification setting in a typical wearable stress-monitoring application.

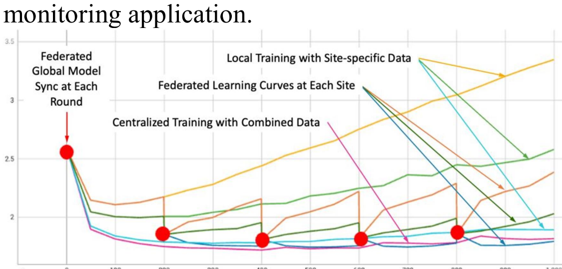

Figure 15. Convergence Behavior of Federated and Centralized Learning Across Multiple Training Rounds

As shown in Figure 15, after 15 training rounds, both the Federated Learning and Centralized Learning models approach this convergence pattern. The red markers are global synchronization points where updates to the local model are sent from various sites to create the updated global model. As each participating site trains locally between synchronization rounds, with the data unique to their site, each client will have a unique learning curve.

The centrally generated training curve on combined data from all sites shows relatively steady convergence, thanks to access to globally aggregated data. However, the federated learning curves show variations at a finer granularity due to differences in data and site properties. But periodic global synchronization helps align local models and gradually improves convergence consistency across sites.

This figure emphasizes the iterative process in federated learning and shows that decentralized training in a federated network can optimize the problem in a stable way without compromising data privacy or requiring distributed training in each environment.

4.9 Confusion Matrix Analysis

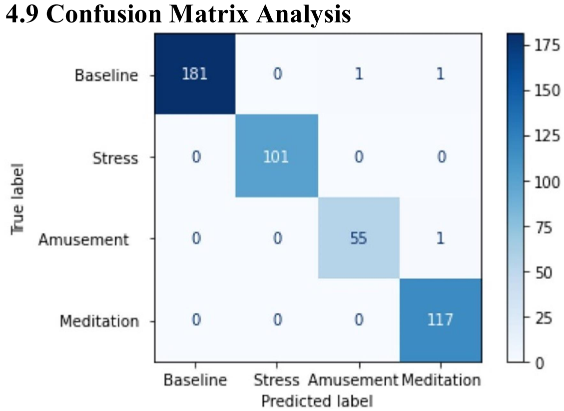

Figure 16. Confusion Matrix for Multiclass Stress Detection Using the Proposed FL-SDNet Framework

Figure 16 shows the confusion matrix of the proposed model, FL-SDNet, for four emotional states: Baseline, Stress, Amusement, and Meditation. The matrix shows how the true labels are associated with the predicted ones and supports an in-depth classification analysis, enabling detailed performance analysis for each class. The model's high classification accuracy across all categories is demonstrated by its high diagonal elements, indicating correct sample categorization. Specifically, 181 of the Baseline samples, 101 of the Stress samples, 55 of the Amusement samples, and 117 of the Meditation samples were correctly identified. Few misclassifications were noticed, mainly between Baseline and Amusement, thus providing high discriminative power and good feature learning.

The results confirm that the proposed federated learning framework can precisely classify various emotional and physiological states and remains well-performing across multi-site, multi-device wearable healthcare environments.

4.10 Ablation Study

Table 6. *Ablation Study Results*

Model Variant	Accuracy (%)
FL without Attention	86.8
FL with Attention only	88.1
FL with AFA only	88.7
FL-SDNet (Full Model)	**90.2**

To evaluate the individual contribution of each component, the ablation study results for different variants of the proposed framework (FL-SDNet) are shown in Table 6. The federated learning (FL) model without an attention mechanism is used to evaluate the distributed framework's performance, achieving an accuracy of 86.8%.With no attention mechanism, the fed. The FL model achieves 86.8% accuracy as a baseline for the distributed setup. Adding only the attention mechanism increased accuracy to 88.1%, enhancing multimodal feature representation and highlighting stress-vulnerable physiological patterns.

Likewise, the model based solely on the Adaptive Federated Aggregation (AFA) mechanism achieved 88.7% accuracy, demonstrating the efficiency of aggregation mechanisms for processing heterogeneous, non-IID data. The overall FL-SDNet model, which included both attention-based multimodal fusion (AMF) and AFA, achieved the highest accuracy of 90.2%. Both results show that they make a considerable contribution to improving overall robustness in the context of federated stress detection.

4.11 Statistical Significance Test

To test the statistical significance of the performance improvement of the proposed FL-SDNet framework compared with the standard FedAvg, a paired t-test was performed. The analysis yielded a p-value < 0.01, indicating that the improvement in stress detection performance was statistically significant. This result demonstrates that the high accuracy and stability of FL-SDNet are not due to random variation but are largely attributable to the seamlessly designed multimodal attention mechanism and the dataset-aware adaptive federated aggregation mechanism.

4.12 Discussion

The experimental results validate the proposed FL-SDNet framework for a distributed stress-detection environment that requires high prediction accuracy, strong privacy protection, and high computational efficiency. This approach resulted in a performance comparable to centralized deep learning (with an error of only approximatley 1%) and far surpassed the standard federated learning approach. It was also robust and unaffected by highly heteroskedastic, asynchronously independent data, as is typical in wearable healthcare applications. Additionally, scalable, communication-efficient aggregation was achieved, reducing transmission overhead and enhancing scalability across distributed systems.

Although these are among its benefits, the proposed framework has drawbacks. According to the decentralized training process, its accuracy is slightly lower than that of a fully centralized learning system. Moreover, federated adaptation and attention mechanisms made the model richer in terms of its computational parameters and structure. However, key practical gains can be made, such as enabling real-time use in a wearable healthcare system that can monitor stress without sharing raw data with others, and scaling to deploy to large numbers of users spread across a distributed target set.

5. CONCLUSION AND FUTURE SCOPE

To achieve secure, scalable, and accurate healthcare monitoring, this study proposed an innovative privacy-preserving federated learning framework for stress detection based on wearables, leveraging repurposed platforms to process edge data. This study introduced an innovative framework, federated learning with stress detection (FL-SDNet), to securely scale and accurately monitor healthcare via wearables, leveraging existing platforms to process data at the edge. The proposed model leveraged a CNN–BiLSTM architecture, combined with cross-modal attention fusion and adaptive federated aggregation (AFA), and effectively addressed the common challenges of multimodal data processing and non-IID client distributions.

Results from the experiment showed that the proposed approach achieved 90.2% accuracy, an F1 score of 0.88, and an ROC-AUC score of 0.92, which are close to the performance of the central deep learning model (accuracy of 91.3%). The system was then validated by showing conservative data preservation and good performance. Furthermore, the framework achieved about a 40% reduction in communication overhead, demonstrating its feasibility in a wearable environment where supporting memory and computational resources are scarce. Finally, in an ablation study, it is confirmed that they are the factors of beneficial results and generalization ability that contribute to the usefulness of these adaptation features and attention mechanisms.

But this positive performance was accompanied by some constraints. Due to limited data sharing, we compromised on intermediate performance compared to centralized approaches. The system also demanded high client participation, and performance might drop when the data imbalance is too large or the client dropout rate is too high. Even though communication costs are reduced, deployment in a large-scale network still suffers from latency issues.

It is important to recognize several practical and research constraints as well. First, the analysis was performed on publicly available data, which may not fully reflect the diversity of real wearable devices and healthcare users. Second, the proposed model can be adapted to non-IID data distributions, but highly heterogeneous user environments may affect performance. Third, it is not well understood how to compute continuous federated learning energy-efficiently on devices with limited resources. Finally, although federated learning improved privacy, other sophisticated security measures, such as differential privacy and secure multiparty computation, were not implemented and may be developed in the future.

Future work to improve the system will allow us to incorporate differential privacy and secure multi-party computation into our framework, providing stronger security guarantees and enabling personalized federated learning approaches that account for user-specific variations. Lightweight, edge-optimized models and an immediate adaptive learning function will also increase scalability and responsiveness. Further research into multimodal modeling and evaluation using larger and more diverse datasets, e.g., including behavioral and contextual information, will enhance its applicability in real-world healthcare settings.

The study ends with a reflection on the potential of federated learning. It helps to find a good trade-off between privacy, communication, and predictive accuracy. This is a plug-and-play system for real-world use. This represents a major step forward for the next generation of stress-monitoring systems in healthcare. These systems are robust to individual variability and provide personalized, data-driven insights to support patient well-being.

REFERENCES

[1] R. Sagar, K. Chatterjee, S. Thareja, A. Timothy, A. S. Yadav, P. Yadav *et al*., "NAMS task force report on mental stress," *Annals of the National Academy of Medical Sciences (India)*, vol. 61, pp. 66–97, 2025, doi: 10.25259/ANAMS_TFR_10_2024.

[2] B. Chu, K. Marwaha, T. Sanvictores, *et al*., "Physiology, Stress Reaction," in *StatPearls* [Online]. Treasure Island, FL, USA: StatPearls Publishing, Jan. 2026. [Updated May 7, 2024]. Available: https://www.ncbi.nlm.nih.gov/books/NBK541120/

[3] A. O. Pataca, E. Zdravevski, P. J. Coelho, N. M. Garcia, M. Deryuck, C. Albuquerque, and I. M. Pires, "Use of machine learning for predicting stress episodes based on wearable sensor data: A systematic review," *Computers in Biology and Medicine*, vol. 198, pt. A, p. 111166, 2025, doi: 10.1016/j.compbiomed.2025.111166.

[4] R. Alhejaili and A. Alomainy, "The use of wearable technology in providing assistive solutions for mental well-being," *Sensors*, vol. 23, no. 17, p. 7378, Aug. 2023, doi: 10.3390/s23177378.

[5] G. Ranjbaran, S. Consoli, G. Leoni, D. R. Recupero, and C. K. Roy, "Federated Artificial Intelligence approaches for wearables and health devices," *Smart Health*, vol. 40, p. 100663, 2026, doi: 10.1016/j.smhl.2026.100663.

[6] R. Al Abdi, S. AlKaabi, S. Elsifi, and J. Yousaf, "Mental stress detection using physiological sensors and artificial intelligence: A review," *Sensors*, vol. 26, no. 5, p. 1616, Mar. 2026, doi: 10.3390/s26051616.

[7] A. Grataloup and M. Kurpicz-Briki, "A systematic survey on the application of federated learning in mental state detection and human activity recognition," *Frontiers in Digital Health*, vol. 6, p. 1495999, Nov. 2024, doi: 10.3389/fdgth.2024.1495999.

[8] B. Yurdem, M. Kuzlu, M. K. Gullu, F. O. Catak, and M. Tabassum, "Federated learning: Overview, strategies, applications, tools and future directions," *Heliyon*, vol. 10, no. 19, p. e38137, 2024, doi: 10.1016/j.heliyon.2024.e38137.

[9] S. R. Abbas, Z. Abbas, A. Zahir, and S. W. Lee, "Federated learning in smart healthcare: A comprehensive review on privacy, security, and predictive analytics with IoT integration," *Healthcare*, vol. 12, no. 24, p. 2587, Dec. 2024, doi: 10.3390/healthcare12242587.

[10] N. T. Madathil, F. K. Dankar, M. Gergely, A. N. Belkacem, and S. Alrabaee, "Revolutionizing healthcare data analytics with federated learning: A comprehensive survey of applications, systems, and future directions," Computational and Structural Biotechnology Journal, vol. 28, pp. 217–238, 2025, doi: 10.1016/j.csbj.2025.06.009.

[11] M. R. Hasan and J. Li, "Privacy-preserving collaborative diabetes prediction in heterogeneous health care systems: Algorithm development and validation of a secure federated ensemble framework," *JMIR Diabetes*, vol. 11, p. e79166, Jan. 2026, doi: 10.2196/79166.

[12] C. Contoli, V. Freschi, and E. Lattanzi, "Energy-aware human activity recognition for wearable devices: A comprehensive review," *Pervasive and Mobile Computing*, vol. 104, p. 101976, 2024, doi: 10.1016/j.pmcj.2024.101976.

[13] S. Mohammadi, A. Balador, S. Sinaei, and F. Flammini, "Balancing privacy and performance in federated learning: A systematic literature review on methods and metrics," *Journal of Parallel and Distributed Computing*, vol. 192, p. 104918, 2024, doi: 10.1016/j.jpdc.2024.104918.

[14] J. Patel and P. Patel, "Deep learning for early detection of depression and anxiety: A comparative study of CNN, LSTM, and RNN models," *Intell. Hosp.*, vol. 2026, p. 100066, 2026, doi: 10.1016/j.inhs.2026.100066.

[15] A. Jaber, "FedStress: A Privacy-Preserving Federated Learning Framework for Efficient and Accurate Stress Detection on Wearable Devices," Zenodo, 2025. doi: 10.5281/zenodo.17453935.

[16] E. Lazarou and T. P. Exarchos, "Predicting stress levels using physiological data: Real-time stress prediction models utilizing wearable devices," *AIMS Neuroscience*, vol. 11, no. 2, pp. 76–102, Apr. 2024, doi: 10.3934/Neuroscience.2024006.

[17] R. Al Abdi, S. AlKaabi, S. Elsifi, and J. Yousaf, "Mental stress detection using physiological sensors and artificial intelligence: A review," *Sensors*, vol. 26, no. 5, p. 1616, Mar. 2026, doi: 10.3390/s26051616.

[18] P. Sreehari, U. Raghavendra, and A. Gudigar, "A review of deep learning techniques for EEG-based emotion recognition: Models, methods, and datasets," *F1000Research*, vol. 14, p. 1276, Nov. 2025, doi: 10.12688/f1000research.171170.1.

[19] A. M. Alqudah and Z. Moussavi, "A review of deep learning for biomedical signals: Current applications, advancements, future prospects, interpretation, and challenges," *Computers, Materials and Continua*, vol. 83, no. 3, pp. 3753–3841, 2025, doi: 10.32604/cmc.2025.063643.

[20] P. Phalaagae, A. M. Zungeru, A. Yahya, B. Sigweni, and S. Rajalakshmi, "A hybrid CNN-LSTM model with attention mechanism for improved intrusion detection in wireless IoT sensor networks," *IEEE Access*, 2025, doi: 10.1109/ACCESS.2025.3555861.

[21] M. S. Ramirez Campos, K. Barati, R. Samavi, P. Pires, L. Duncan, R. B. Sassi, M. D. Noseworthy, S. A. Gadsden, and T. E. Doyle, "Multimodal artificial intelligence and online learning in youth mental health: A scoping review," *NPJ Mental Health Research*, vol. 5, no. 1, p. 26, Apr. 2026, doi: 10.1038/s44184-026-00207-4.

[22] K. A. Sathish Kumar, L. Nelson, and B. R. Jibinsingh, "Systematic review of privacy-preserving Federated Learning in decentralized healthcare systems," *Franklin Open*, vol. 13, p. 100440, 2025, doi: 10.1016/j.fraope.2025.100440.

[23] N. T. Madathil, F. K. Dankar, M. Gergely, A. N. Belkacem, and S. Alrabaee, "Revolutionizing healthcare data analytics with federated learning: A comprehensive survey of applications, systems, and future directions," *Comput. Struct. Biotechnol. J.*, vol. 28, pp. 217–238, Jun. 2025, doi: 10.1016/j.csbj.2025.06.009.

[24] L. G. Roos and George M. Slavich, "Wearable technologies for health research: Opportunities, limitations, and practical and conceptual considerations," *Brain, Behavior, and Immunity*, vol. 113, pp. 444–452, Oct. 2023, doi: 10.1016/j.bbi.2023.08.008.

[25] H. Zhong, L. Zhu, and Y. Liu, "Personalized federated learning with two-stage adaptive aggregation under global feature constraints," *Computer Communications*, vol. 254, p. 108552, 2026, doi: 10.1016/j.comcom.2026.108552.

[26] S. Alahmari and I. Alghamdi, "A comprehensive survey on energy-efficient and privacy-preserving federated learning for edge intelligence and IoT," *Results in Engineering*, vol. 28, p. 107849, 2025, doi: 10.1016/j.rineng.2025.107849.

[27] M. Boumaiz, M. E. Ghazi, A. Bouayad, Y. Balboul, and M. El Bekkali, "Energy-efficient strategies in wireless body area networks: A

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

comprehensive survey," *IoT*, vol. 6, no. 3, p. 49, 2025, doi: 10.3390/iot6030049.
[28] S. Shukla, S. Rajkumar, A. Sinha, M. Esha, K. Elango, and V. Sampath, "Federated learning with differential privacy for breast cancer diagnosis enabling secure data sharing and model integrity," *Scientific Reports*, vol. 15, no. 1, p. 13061, Apr. 2025, doi: 10.1038/s41598-025-95858-2.
[29] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, "Introducing WESAD: A Multimodal Dataset for Wearable Stress and Affect Detection," in *Proc. 20th ACM Int. Conf. Multimodal Interaction (ICMI)*, Boulder, CO, USA, Oct. 2018, pp. 400–408, doi: 10.1145/3242969.3242985.
[30] J. Koldijk, M. Sappelli, S. Verberne, M. A. Neerincx, and W. Kraaij, "The SWELL Knowledge Work Dataset for Stress and User Modeling Research," in *Proc. 16th Int. Conf. Multimodal Interaction (ICMI)*, Istanbul, Turkey, Nov. 2014, pp. 291–298, doi: 10.1145/2663204.2663257.
[31] J. Healey and R. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Trans. Intelligent Transportation Systems*, vol. 6, no. 2, pp. 156–166, Jun. 2005, doi: 10.1109/TITS.2005.848368.