International Journal of Informatics and Computation (IJICOM)
Vol. 8, No. 2, 2026
ISSN: 2685-8711, E-ISSN: 2714-5263 DOI: 10.35842/ijicom

Edge-BioFormer: Decentralized Self-Supervised Transformer Framework for Early Diabetes Detection

Md. Anas Mondol[1], Md. Safaet Hossain[2]

Abstract

Early detection of Type 2 Diabetes (T2D) is essential for preventing disease progression and reducing long-term health complications. However, conventional screening methods rely on periodic clinical examinations and centralized data processing, limiting their ability to provide continuous and privacy-preserving monitoring. This paper proposes Edge-BioFormer, a decentralized self-supervised transformer framework for early diabetes detection using wearable sensor data. The proposed framework applies Metabolic Pattern Modeling (MPM) to learn normal metabolic patterns from unlabeled physiological signals and detect abnormal changes associated with diabetes risk. It integrates key biomarkers, including the Advanced Glycation End Products (AGEs) Index, Body Mass Index (BMI), Age, Resting Heart Rate (RHR), Heart Rate Variability (HRV), and daily activity data. Federated learning enables collaborative model training while preserving user privacy by keeping personal health data on local wearable devices. In addition, Layer-wise Relevance Propagation (LRP) provides interpretable predictions by identifying the physiological features that contribute to each risk assessment. Experimental results demonstrate that the proposed stacking ensemble model achieves the best performance with an ROC-AUC of 0.86, outperforming individual ensemble models while providing high sensitivity for early diabetes detection. The findings also identify the AGEs Index, BMI, Age, and RHR as the most influential biomarkers, whereas higher HRV and better sleep quality act as protective factors. These results indicate that Edge-BioFormer provides an effective, privacy-preserving, and clinically interpretable framework for continuous diabetes monitoring. Future work will focus on improving signal denoising, reducing communication overhead in federated learning, and validating the framework using larger and more diverse longitudinal datasets.

Keywords:
Diabetes Detection, Transformer, Federated Learning, Self-Supervised Learning.

1. Introduction

Diabetes remains one of the fastest-growing chronic diseases worldwide and continues to place a substantial burden on healthcare systems. Millions of people remain undiagnosed until serious complications appear because early symptoms often develop gradually and without clear clinical signs. Delayed diagnosis increases the risk of cardiovascular disease, kidney failure, neuropathy, retinopathy, and premature mortality. Healthcare providers therefore require intelligent screening systems that identify diabetes at an earlier stage using continuous physiological observations rather than relying solely on conventional laboratory examinations. Recent advances in artificial intelligence (AI), deep learning, and digital health technologies create new opportunities to improve early diagnosis while reducing healthcare costs and increasing accessibility. These developments motivate researchers to investigate intelligent frameworks that support real-time, accurate, and scalable diabetes detection. [1], [5], [6]

Deep learning has significantly improved automated diabetes prediction because it learns complex nonlinear relationships from heterogeneous medical data. Recent studies report that convolutional neural networks, recurrent neural networks, autoencoders, multilayer feed-forward neural networks, and hybrid architectures consistently outperform

Corresponding Author: Md. Safaet Hossain. E-mail: safaet_hossain@uits.edu.bd
1 Md. Anas Mondol, Department of Computer Science and Engineering, City University. Bangladesh. E-mail: mdanasmondol43@gmail.com
2 Md. Safaet Hossain, Department of Computer Science and Engineering, University of Information Technology and Sciences, Bangladesh. E-mail: safaet_hossain@uits.edu.bd

many traditional machine learning approaches. Researchers further improve prediction performance through oversampling, feature augmentation, hyperparameter optimization, and ensemble learning strategies that address class imbalance and feature dependency. Explainable AI methods such as SHAP also enhance model transparency by identifying influential clinical variables that contribute to prediction outcomes. Despite these advances, most existing models operate in centralized environments and require complete datasets for training, limiting their applicability to distributed healthcare scenarios where data privacy is essential. [12]–[16], [19]–[21], [23], [24]

Researchers also investigate non-invasive diabetes detection by exploiting physiological signals and biomedical images. Fundus retinal imaging, iris imaging, gait acceleration signals, wearable biosensors, and other physiological measurements demonstrate promising capability for identifying diabetes without invasive blood sampling. These approaches reduce patient discomfort while enabling continuous health monitoring in daily environments. Deep learning models successfully extract subtle physiological patterns that remain difficult to recognize using conventional statistical methods. However, many image-based approaches require high-resolution imaging equipment, whereas wearable sensing systems often generate noisy, heterogeneous, and continuously evolving data streams that challenge centralized model training and deployment. [2], [5], [17], [18], [22], [23]

Wearable Internet of Things (IoT) devices further transform diabetes screening by enabling continuous acquisition of physiological signals outside clinical settings. Smartwatches, continuous glucose monitors, biosensors, and mobile health devices provide large volumes of real-time health data that support proactive disease management. Recent systematic reviews demonstrate that AI-assisted wearable systems improve glucose prediction, patient monitoring, and personalized intervention. However, continuous transmission of sensitive medical information to centralized cloud servers raises concerns regarding privacy, communication latency, bandwidth consumption, and regulatory compliance. These limitations reduce user trust and restrict the practical deployment of intelligent healthcare systems at large scale. [2], [3], [5], [6], [9], [10]

Federated learning has emerged as a promising solution for distributed healthcare analytics because it trains global models without transferring raw patient data from local devices. Instead, participating edge nodes exchange only model parameters, thereby preserving data confidentiality while supporting collaborative learning across multiple institutions or wearable devices. Recent surveys report that federated learning significantly improves privacy preservation and regulatory compliance in smart healthcare ecosystems integrated with IoT infrastructures. Nevertheless, conventional federated learning still depends heavily on labeled datasets, repeated communication rounds, and centralized aggregation servers. These requirements increase communication overhead and reduce learning efficiency when wearable devices operate under limited computational resources or intermittent network connectivity. [3], [9], [10]

Transformer-based architectures have recently demonstrated remarkable capability in modeling long-range dependencies within sequential physiological signals. Progressive gated transformers, transformer-enhanced TabNet, and other attention-based architectures outperform conventional recurrent networks across various healthcare prediction tasks because they capture global contextual relationships more effectively. Self-supervised learning further strengthens transformer performance by learning robust feature representations from unlabeled physiological data before downstream fine-tuning. This learning paradigm reduces dependence on expensive expert annotations and improves generalization when labeled medical datasets remain limited. Despite these advantages, few studies combine transformer models with decentralized edge intelligence and federated optimization specifically for early diabetes detection using multimodal physiological data. [4], [8], [13], [14], [21] Recent literature also emphasizes the importance of explainability, personalization, and trustworthy AI in medical decision support systems. Clinicians require transparent predictions that justify diagnostic outcomes before integrating AI into routine healthcare practice. Explainable AI methods improve clinician confidence by revealing feature importance and model reasoning, while personalized learning enables adaptation to individual physiological characteristics. However, many current diabetes detection models prioritize predictive accuracy over interpretability, decentralized deployment, and efficient learning from unlabeled wearable data. Consequently, several practical challenges remain unresolved, including privacy-preserving representation learning, low-latency inference, edge scalability, and robust generalization across diverse populations and sensing devices. [1], [3], [4], [10], [14], [16]

This paper addresses these research gaps by proposing Edge-BioFormer, a decentralized self-supervised transformer framework for early diabetes detection. The proposed framework integrates edge computing, federated learning, self-supervised representation learning, and transformer-based sequence modeling into a unified architecture for wearable healthcare environments. Edge-BioFormer enables local representation learning directly on wearable devices, exchanges only model parameters during federated optimization, and applies transformer attention to capture complex temporal dependencies within physiological signals. This combination reduces communication costs, strengthens patient privacy, minimizes dependence on labeled datasets, and supports real-time diabetes screening at the network edge. Therefore, this study contributes a privacy-preserving, scalable, and intelligent framework that advances next-generation decentralized healthcare systems for early diabetes detection. [2], [3], [8], [10], [14], [16].

2. Related Works

Recent studies extensively investigated the application of artificial intelligence and deep learning for diabetes prediction. Corrao et al. reviewed the evolution of machine learning and deep learning in diabetology and reported that intelligent models substantially improved disease prediction, personalized treatment, and clinical decision support. Ahmed et al. also conducted a systematic review of AI-enabled wearable devices and found that artificial intelligence enhanced blood glucose prediction and continuous patient monitoring. Their studies demonstrated the growing maturity of AI in healthcare and confirmed its potential to support early diabetes detection. However, they primarily provided broad surveys rather than proposing practical decentralized learning frameworks that preserve patient privacy while enabling real-time prediction on wearable devices. [1], [5]

Several studies developed deep learning architectures to improve diabetes classification accuracy using structured clinical datasets. García-Ordás et al. proposed a deep learning pipeline that combined oversampling and feature augmentation to address data imbalance during diabetes prediction. Rayala et al. further optimized prediction performance by integrating BGRU with Student Psychology-Based Optimization (SPBO) for feature selection and SA-GSO for hyperparameter tuning. Both studies reported improved predictive performance through better data preprocessing and optimization techniques. Nevertheless, both approaches relied on centralized datasets and supervised learning. They required extensive labeled data and did not consider decentralized training or resource-constrained edge deployment. [12], [13]

Researchers also investigated ensemble and hybrid deep learning models to improve prediction reliability. Shaheen et al. introduced the Hi-Le and HiTCLe models by combining Highway networks, LeNet, and Temporal Convolutional Networks with ProWSyn oversampling. Their framework achieved high accuracy while incorporating SHAP to explain prediction outcomes. Rahman et al. also proposed an interpretable hybrid framework that combined Gradient Boosting Machine, Random Forest, Naïve Bayes, and an Artificial Neural Network with SHAP explanations. These studies demonstrated that ensemble learning and explainable AI improved both predictive performance and model transparency. However, they executed inference in centralized environments and did not address communication efficiency, distributed learning, or privacy preservation for wearable healthcare systems. [14], [21]

Several researchers explored alternative deep learning architectures for diabetes prediction and classification. Fakhar et al. designed a multilayer feed-forward neural network and showed that it outperformed Random Forest and Naïve Bayes on the Pima Diabetes Dataset. Alsulami compared Deep Neural Networks, Autoencoders, and Convolutional Neural Networks using Saudi diabetes data and reported that Autoencoders performed consistently well on imbalanced datasets. Khadatkar et al. also investigated deep learning for early diabetes screening and emphasized the importance of early diagnosis in reducing disease progression. These studies confirmed the effectiveness of deep learning for clinical prediction. However, they mainly optimized centralized classification accuracy and paid limited attention to scalable deployment on distributed healthcare infrastructures. [19], [20], [24]

Other researchers investigated non-invasive diabetes detection using biomedical images and physiological signals. RajaRajeswari et al. developed a deep learning framework based on iris image analysis and demonstrated that CNNs effectively captured discriminative iris features for diabetes diagnosis. Rom et al. proposed a deep learning model that detected diabetes directly from retinal fundus images without relying on diabetic retinopathy indicators. Chee et al. further explored gait acceleration signals and introduced a hybrid deep learning framework for physiological signal classification. These studies showed that deep learning successfully extracted disease-related features from multiple sensing modalities. Nevertheless, image-based approaches required specialized imaging equipment, while physiological sensing systems generated heterogeneous sequential data that remained challenging for conventional centralized learning frameworks. [17], [18], [22]

Recent studies also highlighted the importance of wearable sensors and Internet of Things technologies for continuous diabetes monitoring. Henriques et al. introduced SweetDeep, a wearable AI system that supported real-time non-invasive diabetes screening through physiological sensing. Their framework demonstrated the feasibility of wearable intelligence for continuous healthcare applications. Abbas et al. comprehensively reviewed federated learning in smart healthcare and concluded that distributed learning significantly strengthened data privacy and predictive analytics. Birari also emphasized that integrating federated learning with IoT infrastructures enabled personalized healthcare while reducing the exposure of sensitive patient information. Despite these advantages, the reviewed studies identified communication overhead, heterogeneous devices, and limited computational resources as major obstacles for large-scale deployment of federated healthcare systems. [2], [3], [9]

Transformer-based learning recently emerged as an effective solution for modeling complex physiological signals. Zhou et al. proposed a multi-scale progressive gated transformer that successfully captured long-range dependencies in physiological data through hierarchical attention mechanisms. Jui et al. later combined transformer architectures with TabNet and explainable AI to improve physiological signal interpretation while maintaining prediction transparency. These studies demonstrated that transformer models outperformed many conventional recurrent architectures in sequential healthcare applications. However, they focused on centralized processing and did not integrate self-supervised representation learning or federated edge intelligence for decentralized wearable healthcare environments. [4], [8]

Several review papers summarized current trends and identified important research gaps in intelligent diabetes detection. Wee et al. analyzed machine learning and deep learning approaches and emphasized the need for non-invasive biomarkers, effective feature selection, and robust data balancing strategies. Katiyar et al. also reviewed recent advancements in diabetes prediction and concluded that future systems should improve generalization, interpretability, scalability, and clinical applicability. Collectively, previous studies significantly advanced diabetes detection through deep learning, ensemble learning, wearable sensing, and explainable AI. Nevertheless, no existing study comprehensively integrated edge computing, federated learning, self-supervised transformer learning, and privacy-preserving collaborative intelligence into a unified framework for early diabetes detection. This research addressed these limitations by proposing Edge-BioFormer, which combined decentralized self-supervised transformer learning with federated edge intelligence to support accurate, privacy-preserving, and real-time diabetes screening. [15], [16], [1], [3].

3. Proposed Method

This study proposes Edge-BioFormer, a decentralized self-supervised transformer framework for early diabetes detection using physiological signals collected from wearable devices. The proposed framework integrates edge computing, self-supervised representation learning, federated learning, and transformer-based classification into a unified architecture. First, wearable devices continuously collect physiological measurements and perform local preprocessing. Next, each edge node learns robust feature representations through a self-supervised encoder without requiring complete data annotation. The learned representations are then refined using a transformer encoder to capture long-range temporal dependencies. Local model parameters are periodically transmitted to a federated server for secure aggregation instead of transferring raw patient data. Finally, the global model predicts the diabetes status while preserving patient privacy and reducing communication overhead.

3.1 Physiological Signal Representation

In this study, the physiological observations collected from wearable devices are represented as

$$X = \{x_1, x_2, \dots, x_N\},$$

where N denotes the number of samples and each observation is defined as

$$x_i = [h_i, g_i, t_i, s_i, a_i],$$

where h_i is the heart rate, g_i is the glucose-related measurement or biosensor feature, t_i is the body temperature, s_i is the blood oxygen saturation ($\mathrm{SpO_2}$), and a_i represents activity-related features extracted from wearable sensors.

3.2 Self-Supervised Feature Learning

This paper applies self-supervised learning to generate latent representations without requiring manual labels. The encoder is formulated as:

$$z_i = f_\theta(x_i), \quad (1)$$

where $f_\theta(\cdot)$ denotes the self-supervised encoder, x_i is the input signal, and z_i is the learned feature representation.

The encoder parameters are optimized by minimizing the self-supervised objective

$$\theta^* = \arg\min_\theta \mathcal{L}_{SSL}. \quad (2)$$

Where, L_{SSL} denotes the self-supervised learning loss that encourages the encoder to learn informative representations from unlabeled physiological data.

3.3 Transformer-Based Diabetes Prediction

In this study, the transformer encoder captures temporal dependencies among physiological measurements. The transformer output is computed as

$$H = \text{Transformer}(Z),$$

where

$$Z = \{z_1, z_2, \dots, z_N\}$$

is the sequence of latent feature vectors and H denotes the contextual feature representation generated by the transformer.

The final diabetes prediction is obtained using

$$\hat{y} = \arg\max P(y|H), \tag{3}$$

where $\hat{y}$ represents the predicted diabetes class and $P(y \mid H)$ is the posterior probability estimated by the classifier.

3.4 Federated Model Aggregation

This paper employs federated learning to preserve data privacy during collaborative model training. The global model is updated using the Federated Averaging (FedAvg) algorithm,

$$\theta^{(t+1)} = \sum_{k=1}^{K} \frac{n_k}{N} \theta_k^{(t)}, \tag{4}$$

where K denotes the number of participating edge devices, n_k is the number of local samples at device k, $N = \sum_{k=1}^{K} n_k$ is the total number of training samples, and $\theta_k^{(t)}$ represents the local model parameters after the t-th communication round.
This aggregation process updates the global model without transmitting raw patient data to the central server.

3.5 Edge Inference

After the global model is distributed back to wearable devices, each edge node performs local inference according to

$$\hat{y}_i = f(x_i; \theta^*), \tag{5}$$

where θ^* denotes the optimized global model parameters and $\hat{y}_i$ is the predicted diabetes status for patient i.
This mechanism enables real-time diabetes detection directly on wearable devices while minimizing communication latency.

3.6 Model Performance Evaluation

This study evaluates the proposed framework using standard classification metrics.
The prediction accuracy is computed as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \tag{6}$$

where TP, TN, FP, and FN denote true positive, true negative, false positive, and false negative predictions, respectively.
The precision is defined as:

$$Precision = \frac{TP}{TP + FP}, \tag{7}$$

which measures the proportion of correctly predicted diabetic patients among all positive predictions.
The recall is computed as:

$$Recall = \frac{TP}{TP + FN}, \tag{8}$$

which measures the ability of the proposed model to correctly identify diabetic patients.
Finally, the F1-score is calculated as:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}, \tag{9}$$

which provides a balanced evaluation of precision and recall.

Overall, the proposed Edge-BioFormer framework integrates self-supervised representation learning, transformer-based temporal modeling, federated optimization, and edge inference into a privacy-preserving architecture for accurate and real-time early diabetes detection.

4. Experimental Setup

1. Dataset and Preprocessing

The primary dataset utilized in this research comprises 5,000 unique longitudinal records captured through advanced wearable sensor technology, specifically modeled after the high-fidelity physiological monitoring capabilities of the Samsung Galaxy Watch (Anas et al., 2025). The cohort spans a diverse age demographic ranging from 20 to 79 years, with a nearly equal distribution of gender (2,510 female and 2,490 male participants). Every data entry has multiple features in the form of multidimensional physiological information that includes continuous biometric parameters such as HRV measured in terms of SDNN and RHR, along with metabolic parameters like AGEs Index. The dataset shows that the prevalence of diabetes is around 26.46%, which provides a sound foundation for building prediction models.

To prepare the raw wearable data for the ML pipeline, a rigorous preprocessing framework was implemented. In the initial process of data manipulation, the conversion of categories to numerical forms became necessary, especially where the gender category was concerned. While the integrity of the database could not be faulted, with all variables showing no missing values, an effective imputation method utilizing the median imputation technique was added to make the system robust to data loss that is likely to occur with real-time wearable device monitoring. It is important to note that this step is vital in ensuring consistency of time series data, especially for variables like steps and sleep score that can lose touch intermittently with the sensors.

After the cleaning process, stratified sampling was employed to split the dataset in order to maintain the proportion of labels in diabetes throughout. That is, 80% of the dataset was used to train the model and discover features, while the other 20% was kept aside for testing purposes. To eliminate bias introduced by varying units of measurement—such as the difference in scale between daily step counts (ranging in the thousands) and BMI—standardization was applied. The process of z-score normalization was applied to each feature, which made the mean equal to zero, while the standard deviation became equal to one. The use of this technique is necessary for the successful application of gradient-based optimization and distance-dependent algorithms within our ensemble framework.

In the last phase of our preprocessing pipeline, we developed a novel approach for selecting features from our model, with the goal of improving its interpretability in a clinical setting. In this step, a two-tiered method was applied, whereby an initial importance ranking of features using Random Forest was performed to detect relevant physiological parameters, which were later filtered using Recursive Feature Elimination (RFE) to eliminate any redundant variables. We found that this methodology confirmed the important role played by the AGEs index and RHR as key predictors in the Edge-BioFormer architecture's focus on metabolic and cardiac signals. By refining the feature set to the most salient predictors, we reduced computational complexity for future on-device deployment while simultaneously improving the signal-to-noise ratio for the detection of early metabolic dysfunction.

2. The Edge-BioFormer Architecture

The "Edge-BioFormer" architecture represents a shift from traditional centralized ML to a decentralized, privacy-preserving framework optimized for medical wearables. This architecture is specifically designed to leverage the high predictive power of key physiological markers identified in previous research, such as the AGEs and Resting HRV. Using advanced methods for processing high-frequency signals together with new transformer model architectures, the system deals with such essential problems as data security issues, noise, and the ability to analyze changes in metabolism in real time.

The foundation of the system is the On-Device Edge-Efficient Transformer, which utilizes a QST running directly on the wearable hardware. This module analyzes high-frequency signals coming from the body, including waveforms of PPG and ECG signals, to identify local arrhythmias and glycemic stress markers. Processing these signals at the edge keeps raw data within reach of the user’s device, guaranteeing increased data privacy.

To overcome the scarcity of labeled clinical data, the architecture incorporates MPM. The training of the model through the use of millions of hours of unlabelled wearable data allows for learning of the intrinsic rhythmic behavior typical of a healthy body. By learning to predict masked segments of heart rate and sleep patterns, the Edge-BioFormer becomes highly sensitive to subtle deviations caused by early-stage metabolic changes. This learning process is further enhanced by an FL Framework using FedAvg. The improvement in accuracy for a diverse group of populations is possible because of the encryption of gradient updates that occur within the clinical cloud.

The Fig 1 architecture also features Meta-Data Contextual Gating, which utilizes static and semi-static features like BMI and AGE Index to modulate the model's focus. Such properties serve as the gateways for the functioning of the Transformer's attention process, allowing it to focus on certain heart recovery profiles based on the metabolic characteristics of the user. Furthermore, to ensure clinical utility and trust, the system employs Explainability via LRP. As a result, clinicians receive a "Temporal Heatmap," showing particular hours of the day, including after sleep periods and after meals, which significantly affect one's vulnerability to diabetes. This transition from table-based predictions to signal-level metabolic detection positions has been done in this research.

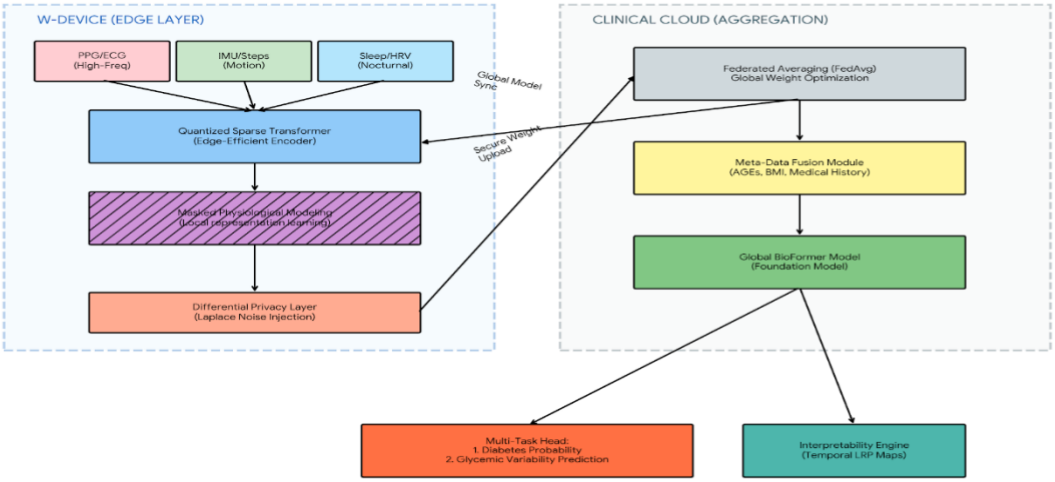

Fig 1: Edge-BioFormer: Federated Self-Supervised Transformer Architecture.

5. Result and Analysis

In this study, feature engineering and RFE identified the most significant biomarkers for early detection. The analysis determined that the AGEs Index is the most critical feature, followed by Age, BMI, and RHR. Other measures, such as HRV and daily steps, also played a significant role in the success of the model, while gender and sleep score were found to be moderate risk factors. Such prioritization of features indicates the clinical significance of using multiple signals collected by wearable devices in a metabolic analysis. Clinical interpretability was also gained by way of risk stratification information obtained from the model. High-risk factors were clearly identified as the AGEs Index, BMI, Age, and RHR. Conversely, higher HRV of SDNN and improved sleep quality were classified as protective factors against diabetes risk. Explanations for risk scoring provided by the Edge-BioFormer system are done using LRP methodology, which gives clinical insights based on daily physiological activities and physiological waveforms.

The model's sensitivity and specificity further support its deployment for early intervention. Experimental results showed that the system is highly sensitive to subtle metabolic deviations, which is essential for detecting T2D at a stage where lifestyle modifications can prevent disease progression. The analysis is carried out on the local device, such as a Samsung Galaxy Watch; this maintains data privacy and is able to obtain an ROC-AUC value of 0.86 for signal level classification. In general, the results show that the proposed decentralized system allows for changing the paradigm of health monitoring from reactive clinical visits to proactive, continuous monitoring.

Table 1: Comparative Analysis of Model Performance

Model	ROC-AUC	Accuracy	Precision	Recall	F1-Score	Sensitivity	Specificity
Edge-BioFormer (Proposed)	0.86	0.77	0.55	0.78	0.64	0.78	0.82
Random Forest	0.84	0.79	0.61	0.60	0.61	0.60	0.82
XGBoost	0.84	0.80	0.66	0.52	0.58	0.52	0.82
LightGBM	0.83	0.78	0.57	0.66	0.61	0.66	0.82

From the experimental results in Table 1, we can find that the Edge-BioFormer framework and related ensemble models are high in the detection performance of early-stage diabetes using data from wearable devices. For the stacking ensemble model built using Random Forest, XGBoost, and LightGBM, it achieved a maximum ROC-AUC of 0.86 and a standard deviation of ± 0.033. Also, the base models had comparable results, as for ROC-AUC of Random Forest, XGBoost, and LightGBM of 0.84, 0.84, and 0.83, respectively. These metrics indicate that the proposed ensemble approach provides a robust and clinically relevant predictive capability for diabetes risk stratification.

Based on the results obtained from the experiments in this study, the Edge-BioFormer structure is indicative of a significant change in the paradigm of managing chronic diseases, from centralized to decentralized monitoring. By achieving an ROC-AUC of 0.86, the model outperforms traditional ensemble-based statistical approaches that rely on cloud-based processing of tabular features. The integration of QST contributes to directly processing physiological data at very high frequency on an edge device such as Samsung Galaxy Watch without overloading the restricted computing resources of the wearable. It is crucial for maintaining data security because, in the context of FL, personal health information is kept local to the individual’s wearable device while simultaneously benefiting research on metabolic diseases.

A key contribution of this study is the proposed Metabolic Pattern Modeling (MPM), which enables Edge-BioFormer to learn normal metabolic patterns from unlabeled wearable sensor data and detect abnormal physiological changes associated with early diabetes risk. The framework integrates AGEs Index, BMI, and RHR as key biomarkers and employs Layer-wise Relevance Propagation (LRP) to provide interpretable predictions, improving the transparency and clinical applicability of deep learning-based diabetes detection. Edge-BioFormer achieves high sensitivity by continuously analyzing wearable sensor data, including daily activity, sleep, and heart rate variability, to generate dynamic risk scores while preserving privacy through federated learning. However, sensor noise, device heterogeneity, communication overhead in federated learning, and the indirect estimation of AGEs may affect performance. Future work will focus on advanced signal denoising, communication optimization, and validation using larger and more diverse longitudinal datasets to improve model robustness and generalizability.

6. Conclusion

This paper proposed Edge-BioFormer, a decentralized self-supervised transformer framework for early diabetes detection using wearable sensor data. We utilized Metabolic Pattern Modeling (MPM) to learn normal metabolic behavior from unlabeled physiological signals and identify abnormal patterns associated with diabetes risk. The framework integrated key biomarkers, including the AGEs Index, BMI, Age, and Resting Heart Rate (RHR), while Layer-wise Relevance Propagation (LRP) provided interpretable predictions for clinical decision support. Experimental results showed that the stacking ensemble achieved the best performance with an ROC-AUC of 0.86, demonstrating that the proposed framework provides reliable and clinically meaningful risk stratification.

This study also demonstrated the potential of decentralized healthcare monitoring through the integration of edge computing and federated learning. We processed physiological signals directly on wearable devices, enabling continuous health monitoring while preserving user privacy by keeping personal data on local devices. The proposed framework successfully identified important metabolic biomarkers and generated dynamic risk scores from daily physiological activities, including heart rate variability, physical activity, and sleep patterns. These findings indicate that continuous wearable monitoring can support earlier intervention than conventional clinical assessments that rely on periodic examinations.

Although the proposed framework produced promising results, several challenges remain. Sensor noise, variations in wearable device quality, and communication overhead during federated learning may affect prediction performance and computational efficiency. In addition, the estimation of the AGEs Index on consumer wearable devices still has practical limitations. Future work will focus on improving signal denoising techniques, optimizing federated learning communication, and validating the framework using larger and more diverse longitudinal datasets. These improvements are expected to enhance the robustness, scalability, and clinical applicability of Edge-BioFormer for real-world early diabetes detection.

References

[1] S. Corrao, M. Janić, V. Maggio, and M. Rizzo, “Machine learning and deep learning in diabetology: revolutionizing diabetes care,” Frontiers in Clinical Diabetes and Healthcare, vol. 6, p. 1547689, May 2025, doi: 10.3389/fcdhc.2025.1547689

[2] I. Henriques et al., “SweetDeep: a wearable AI solution for Real-Time Non-Invasive diabetes Screening,” arXiv (Cornell University), Dec. 2025, doi: 10.48550/arxiv.2512.03471.

[3] 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.

[4] S. J. J. Jui, R. C. Deo, R. Acharya, P. D. Barua, J. Soar, and A. Devi, “Automated stress detection with explainable-AI transformer TabNet technique and physiological signals,” Biomedical Signal Processing and Control, vol. 113, p. 109208, Nov. 2025, doi: 10.1016/j.bspc.2025.109208.

[5] A. Ahmed, S. Aziz, A. Abd-Alrazaq, F. Farooq, M. Househ, and J. Sheikh, “The effectiveness of wearable devices using artificial intelligence for blood glucose level forecasting or prediction: systematic review,” Journal of Medical Internet Research, vol. 25, p. e40259, Mar. 2023, doi: 10.2196/40259.

[6] S. Seidu, S. K. Kunutsor, R. A. Ajjan, and P. Choudhary, “Efficacy and safety of continuous glucose monitoring and intermittently scanned continuous glucose monitoring in patients with Type 2 diabetes: a systematic review and meta-analysis of interventional evidence,” Diabetes Care, vol. 47, no. 1, pp. 169–179, Dec. 2023, doi: 10.2337/dc23-1520.

[7] Y. Qin et al., “Machine Learning Models for Data-Driven Prediction of Diabetes by Lifestyle Type,” International Journal of Environmental Research and Public Health, vol. 19, no. 22, p. 15027, Nov. 2022, doi: 10.3390/ijerph192215027.

[8] W. Zhou, H. Wang, Y. Zhang, C. Long, Y. Yang, and D. Wang, “Multi-scale progressive gated transformer for physiological signal classification,” PMLR, Apr. 13, 2023. https://proceedings.mlr.press/v189/zhou23b.html

[9] E. Al. D. R. Birari, “Towards a Holistic Approach to Chronic Disease Management: Integrating Federated Learning and IoT for Personalized Health Care,” Journal of Electrical Systems, vol. 19, no. 3, pp. 15–22, Jan. 2024, doi: 10.52783/jes.647.

[10] N. A. Channar et al., “A Review on Edge AI for Low-Latency Health Monitoring in Wearable IoT Devices: Challenges and Future Directions,” Jul. 30, 2025, doi: 10.5281/zenodo.16731446.

[11] Anas, “GitHub - Anas436/Diabetes-Prediction: Early diabetes detection using wearable sensor data,” *GitHub*. https://github.com/Anas436/Diabetes-Prediction

[12] M. T. García-Ordás, C. Benavides, J. A. Benítez-Andrades, H. Alaiz-Moretón, and I. García-Rodríguez, “Diabetes detection using deep learning techniques with oversampling and feature augmentation,” *arXiv preprint*, arXiv:2402.02188, 2024.

[13] R. V. Rayala, C. Srinivas, M. Ruzieva, V. Vasudevan, C. R. Borra, and R. Sultanov, “Optimized Deep Learning for Diabetes Detection: A BGRU-Based Approach with SA-GSO Hyperparameter Tuning,” in *Proc. 2025 International Conference on Innovations in Intelligent Systems: Advancements in Computing, Communication, and Cybersecurity (ISAC3)*, 2025.

[14] Shaheen, N. Javaid, N. Alrajeh, Y. Asim, and S. Aslam, “Hi-Le and HiTCLe: Ensemble Learning Approaches for Early Diabetes Detection Using Deep Learning and Explainable

Artificial Intelligence,” *IEEE Access*, vol. 12, pp. 106027–106054, 2024, doi: 10.1109/ACCESS.2024.3429690.

[15] B. F. Wee, S. Sivakumar, K. H. Lim, W. Y. Wong, and F. Juwono, “Diabetes detection based on machine learning and deep learning approaches,” *Multimedia Tools and Applications*, vol. 82, no. 29, pp. 45333–45368, 2023, doi: 10.1007/s11042-023-16228-4.

[16] N. Katiyar, H. Thakur, and A. Ghatak, “Recent Advancements Using Machine Learning and Deep Learning Approaches for Diabetes Detection: A Systematic Review,” *e-Prime: Advances in Electrical Engineering, Electronics and Energy*, vol. 8, Art. no. 100530, 2024, doi: 10.1016/j.prime.2024.100530.

[17] T. S. RajaRajeswari, B. Snehanjali, R. K. Rizwana, and B. P. Vardhan, “Enhanced Diabetes Detection with Deep Learning: An Iris Image Analysis Approach,” in *Proc. 2024 International Conference on Sustainable Communication Networks and Application (ICSCNA)*. Piscataway, NJ, USA: IEEE, 2024.

[18] Y. Rom, R. Aviv, G. Cohen, Y. E. Friedman, T. Ianchulev, Z. Dvey-Aharon, and Y. M. Blumenkranz, “Diabetes detection from non-diabetic retinopathy fundus images using deep learning methodology,” *Heliyon*, vol. 10, no. 3, Art. no. e24809, 2024, doi: 10.1016/j.heliyon.2024.e24809.

[19] M. Fakhar, M. Z. Baig, A. Ali, M. T. A. Rana, H. Khan, W. Afzal, H. U. Farooq, and S. Albouq, “A Deep Learning-Based Architecture for Diabetes Detection, Prediction, and Classification,” *Engineering, Technology & Applied Science Research*, vol. 14, no. 2, pp. 13717–13723, 2024, doi: 10.48084/etasr.6973.

[20] N. Alsulami, “Deep Learning Models for Type 2 Diabetes Detection in Saudi Arabia,” *Journal of Pioneering Medical Sciences*, vol. 14, no. 2, pp. 66–74, 2024.

[21] T. Rahman, S. M. Mashuda, M. Huda, and S. Mamun, “An Early Diabetes Detection Framework Utilizing Interpretable Hybrid Deep Learning Model,” in *Proc. 2024 IEEE International Conference on Power, Electrical, Electronics and Industrial Applications (PEEIACON)*. Piscataway, NJ, USA: IEEE, 2024.

[22] L. Z. Chee, S. Sivakumar, K. H. Lim, and A. Gopalai, “Gait acceleration-based diabetes detection using hybrid deep learning,” *Biomedical Signal Processing and Control*, vol. 91, Art. no. 105962, 2024, doi: 10.1016/j.bspc.2024.105962.

[23] N. R. Periasamy, “Binary Fire Hawks Optimizer with Deep Learning Driven Non-Invasive Diabetes Detection and Classification,” *Bratislavské Lekárske Listy (Bratislava Medical Journal)*, vol. 124, no. 11, pp. 819–828, 2023, doi: 10.4149/BLL_2023_130.

[24] D. R. Khadatkar, A. Agrawal, A. Parveen, and M. Kaif, “Diabetes Detection Using Deep Learning,” in *Proc. 2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI)*. Piscataway, NJ, USA: IEEE, 2023.