Edge AI Empowered Personalized Privacy-Preserving Blood Glucose Prediction with Federated Deep Learning

Xinyi Yang
Department of Computer Science
North Dakota State University
Fargo, USA
xinyi.yang@ndsu.edu

Juan Li
Department of Computer Science
North Dakota State University
Fargo, USA
j.li@ndsu.ed

***Abstract*— Glucose prediction can greatly benefit people with diabetes by allowing them to anticipate and proactively manage changes in their glucose levels. In this paper, we propose a novel glucose prediction mechanism that works with wearable devices to accurately predict a person's glucose levels in real-time without sending sensitive personal glucose data to a third-party cloud. This distributed, lightweight, personalized glucose prediction mechanism works with IoT devices, such as continuous glucose monitoring system, to analyze patterns in glucose levels, insulin doses, and food intake to provide predictions of future glucose levels. Specifically, we applied a personalized federated deep learning algorithm that can train the model on multiple IoT devices, while keeping personal and sensitive glucose data on each device and avoiding the centralization of sensitive data. Moreover, a personalization component is integrated into the federated learning model to allow for the creation of personalized models for each individual user. This can be particularly useful in the context of diabetes management, where individual differences in insulin sensitivity, food preferences, and physical activity can impact glucose levels in unique ways. Finally, the algorithm is optimized for IoT devices that have limited processing power, memory, and battery life. Experimental results on simulated data justify the performance of the proposed system.**

Keywords—edge AI, artificial neural network, recurrent neural network, deep learning, federated learning, diabetes management, glucose prediction

I. INTRODUCTION

Diabetes is a chronic health condition characterized by impaired insulin production or utilization, leading to elevated blood glucose levels. Diabetes affects people all over the world, including in the United States. According to the Centers for Disease Control and Prevention, in the US, approximately 37.3 million people, or 11.3% of the population, have diabetes [1]. This number has been steadily increasing over time and is projected to continue to rise in the coming years.

There are two main types of diabetes: type 1 and type 2. Type 1 diabetes is an autoimmune condition in which the body's immune system attacks and destroys the cells that produce insulin. People with type 1 diabetes require regular insulin injections to regulate their glucose levels. Type 2 diabetes is the most common form of diabetes and is characterized by insulin resistance, meaning that the body's cells are not able to effectively use the insulin that is produced. This can lead to elevated levels of glucose in the blood. High levels of glucose in the blood over an prolonged period can lead to serious health complications, such as heart disease, stroke, kidney failure, blindness, and amputations [2]. Therefore, it is important to maintain tight control of glucose levels for people with both type 1 and type 2 diabetes. Deep learning (DL) has become a popular choice in research; however, it often requires large amounts of data to achieve optimal performance. Nonetheless, the privacy and safety of patient data are paramount, as sharing sensitive information can lead to serious safety and security concerns, especially in healthcare.

In this context, Edge AI, which refers to the use of artificial intelligence algorithms on the edge of a network, rather than relying solely on the cloud or a central server. It offers a promising solution to address the challenge of data dependency in DL models while also enhancing data privacy and enabling personalized models. However, one of the major challenges of edge AI is the limited computing power of IoT devices, which can affect the performance and accuracy of machine learning algorithms. This can make it difficult to process large amounts of data and make accurate predictions in real-time. In addition, since IoT devices are often battery-powered their ability to analyze and make predictions in real-time using complex energy-intensive algorithms is restricted.

To address these problems, we propose a personalized federated learning (FL) model that allows people's devices, such as continuous glucose monitoring (CGM) or smartwatch, to collaboratively make glucose predictions and improve the accuracy of machine learning models. Individual devices collect and analyze data locally, and the data remains on each device without being shared or combined. Instead, findings (in the form of model parameters) are sent to the cloud, where they are aggregated to update and improve the global machine learning model. In this way, we can address privacy concerns while still

This work was supported by the National Science Foundation (NSF) with award numbers: 1722913 and 2218046.

allowing for the development of powerful and effective machine learning models.

The federated prediction model exploits the computational power of all users and is trained over a larger set of data points. However, this model generates a common prediction for all the users without adapting to different individuals. This can be a problem in the context of diabetes management, as individual differences in insulin sensitivity, food preferences, and physical activity can impact glucose levels in unique ways. To solve this problem, we introduce a personalized learning component to the FL model, so that individual device can adapt their prediction model based on their local dataset. This personalized federated predictor keeps all the benefits of the FL while leading to a more personalized model for each individual user.

The rest of the paper is organized as follows. In Section II, we review research and applications related to our study. In Section III, we explain the detailed methodology of the proposed approach. We present the evaluation results in Section IV and conclude the paper in Section V.

II. Related Work

Machine learning algorithms have been used for glucose prediction for several decades. These algorithms are based on mathematical models trained using historical glucose data to make predictions about future glucose levels. A simple approach is using linear regression for glucose prediction. For example, Li et al. used a simple mathematical formula to fit a line to the data, which can then be used to make predictions about future glucose levels [3]. In another work, Wang et al. used multiple linear regression to make predictions [4]. Another common algorithm is decision trees, which involves creating a tree-like model to represent the relationships between various factors that impact glucose levels, such as food intake, physical activity, and insulin dose. For instance, a recent research combines decision trees with ensemble machine learning algorithm to predict blood glucose levels [5]. There are many other machine learning-based glucose predicting systems, such as support vector machines [3]. These traditional machine learning algorithms have been shown to be effective in making glucose predictions but have limitations, including handling complex data patterns, dealing with missing data, and adapting to changes over time.

In recent years, there has been a growing interest in utilizing DL techniques for glucose prediction in diabetes management. Researchers used artificial neural networks (ANN) [6]–[8], convolutional neural networks [9]–[11], and neural networks (RNN) [12]–[14] to analyze CGM data and predict glucose levels. Zhu et al. [11] addressed the glucose prediction problem as a classification task and developed a model using casual dilated convolutional neural network layers, employing fast WaveNet algorithms. In another work, RNN was applied on pre-clustered data to predict glucose levels [12].

FL enables decentralized training of models across multiple devices. Researchers have started to apply FL in the health domain, as it has the potential for training machine learning models on large-scale datasets while preserving the privacy of individuals' data. This is particularly important in healthcare, where privacy concerns are high and sensitive personal health information needs to be protected. For example, FL has been proposed to connect electronic health record data from medical institutions, allowing them to share their experiences but not their data [15]–[18]. In addition, FL can be used to train models on a large number of devices, such as wearable devices and smartphones, that collect data about various health parameters. This can result in more accurate and robust models, as the models can be trained on a diverse range of data from a large number of individuals. For example, researchers proposed to use FL and IoT devices for remote health monitoring [19], [20].

FL has been extensively employed in various diabetes research studies. It has been prominently used in analyzing diabetic retinopathy severity based on images [21], predicting the risk of diabetes mellitus [22], and forecasting the onset of diabetes [23]. Additionally, researchers have explored the application of FL in glucose prediction, achieving an average accuracy of approximately 0.65 across six patients [24]. To the best of our knowledge, only one study has ventured to combine DL techniques, specifically convolutional neural network, with the FL approach in the domain of glucose prediction. The study conducted training and testing using a dataset spanning a duration of five days. The results showed Root Mean Square Error (RMSE) values of 27.45, 28.58, 51.79, and 51.79 mg/dL, respectively, when predicting glucose levels based on the last 9, 15, 21, and 27 minutes of preceding data from a cohort of six patients [25].

Despite progress in DL for glucose prediction, several research gaps persist:

1. Limited access to comprehensive patient data due to privacy concerns hinders the availability of large-scale datasets for training DL models.
2. Patients with abundant local personal device data face challenges in performing independent analysis due to computational and technical barriers.
3. Previous studies on federated deep learning (FDL) for glucose prediction often focused on a small number of patients, potentially missing valuable insights from a more diverse patient population. Additionally, these studies reported relatively high RMSE values, highlighting the need for improved prediction model accuracy.

Therefore, this study aims to address these limitations by utilizing a personalized FDL approaches to analyze the data of 30 patients without sharing their sensitive information with a central repository. Each patient will train a personalized model using their own devices, leveraging the computational power available on their end. By focusing on a limited number of days of data, this approach ensures efficient computation while maintaining privacy.

III. Methodology

A. System Architecture

The proposed privacy-preserving and personalized FDL glucose prediction system comprises two essential components: edge devices and a centralized server. Each edge device (such as a wearable device (CGM) or smartphone) hosts its local FDL model, using time-series data (e.g., glucose levels, physical activity, heartbeat) collected from the user's wearable devices to make predictions. The local models include a personalization component, tailoring the model to each device's unique data

patterns. The centralized server aggregates local models from edge devices and updates the global model, which is a DL model trained using parameters from all the local models across the edge devices.

The architecture of personalized federated learning (FL) is depicted in Figure 1, illustrating the system's components and interactions. The system operates as follows:

- Data Collection: Edge devices collect data and use their local models to make predictions based on the collected data.
- Model Parameters Transfer: Edge devices transmit their local model parameters to the centralized server.
- Global Model Training: Centralized server aggregates the received local model parameters to update the global model.
- Model Parameters Update: Centralized server sends the updated global model parameters back to the edge devices .
- Personalization: edge devices use the updated global model parameters to update their local models, incorporating the personalized component.

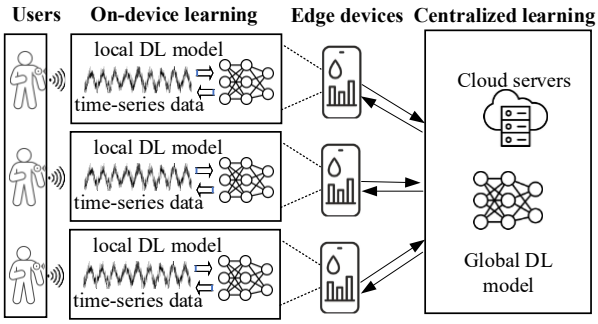

Fig. 1. System architecture

This decentralized and scalable approach enables continuous updating of the global model based on data collected by numerous edge devices . Importantly, it preserves individual data privacy and allows for personalized predictions.

B. Learning Models

This study explores two popular DL models, ANN and RNN, for glucose prediction, lerveraging their distinct strengths in handling time-series data. The decision to use either ANN or RNN depends on the characteristics of the data and the complexity of temporal dependencies. If the temporal dependencies are relatively short-term and the sequence length is not excessively long, ANN might be sufficient and computationally efficient. On the other hand, if the data exhibits long-term temporal dependencies and requires modeling complex sequences, RNN would be more appropriate due to their ability to capture and utilize information from past observations. After using the ANN or RNN model, we employ Federated Learning (FL) to facilitate collaborative and privacy-preserving model training across multiple devices.

Artificial Neural Network

ANN has been widely used to solve the regression problems through a learning process by adjusting the values of the connection weights among the neurons [26]. This study used the multilayer perceptron, which is the most widely adopted ANN model in different industries [27]. A multilayer perceptron may have one or more hidden layers and finally an output layer. It is described as being fully connected, with each node connected to every node in the next and previous layer [28]. After conducting multiple tests, two hidden layers were considered in this study, as this configuration led to efficient processing and satisfactory performance for the ANN.

Recurrent Neural Network

RNN can effectively model sequential data by capturing temporal dependencies through recurrent connections [29]. This unique architecture allows RNN to maintain an internal state or memory of previous inputs, enabling them to consider the context and history of the input sequence when making predictions [30]. This makes RNN well-suited for tasks such as time series analysis, where the current input relies on the information from previous inputs. In this study,we utilized an RNN with two hidden layers as shown in Figure 2. Again, the choice of two hidden layers for RNN was based on extensive testingas it strikes a balance between complexity and efficiency, making it the most suitable choice for the task.

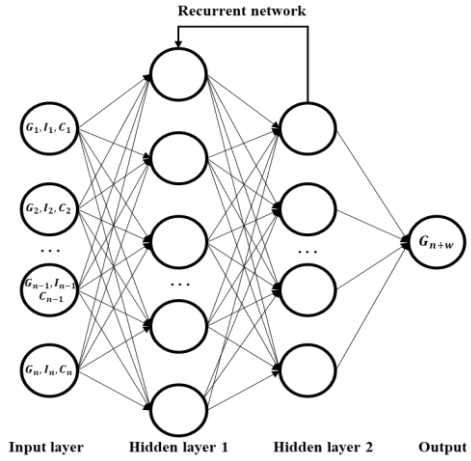

Fig. 2. RNN architecture with two hidden layers.

Federated learning

To achieve collaborative learning while preserving data privacy, the proposed personalized FDL glucose prediction system uses FL with the standard Federated Averaging (FedAvg) algorithm, as proposed by McMahan et al. [31]. The FL process involves two key components: 1) local model updates on edge devices with local model personalization; and 2) model aggregation on the centralized server using the FedAvg algorithm.

Local Model Personalization: After receiving the global model parameters from the centralized server , each of the K edge device maintains its local FDL model, allowing for personalized predictions. The personalization component tailors the local models during training, capturing individual data patterns and characteristics specific to each edge device's data through fine-tuning. This fine-tuning process enables the local models to adapt to the nuances and specific patterns in the local data, enhancing prediction accuracy for individual patients and accommodating variations in glucose dynamics.

FedAvg Algorithm: The FedAvg algorithm orchestrates the collaborative learning process between the edge devices and the centralized server. At round *t* of training:

a. Local Model Updates: Each edge device k (k = 0 to K) conducts local stochastic gradient $g_{t,k}$ with respect to their model parameters ${w_t}^k$.

b. Model Parameter Transfer: After local updates, each edge device transmits the latest model parameters ${w_{t+1}}^k$ to the centralized server. In which,

$$ {w_{t+1}}^k = {w_t}^k + \eta\, g_{t,k} \tag{1}$$

where ${w_t}^k$ is the local model parameter maintained in the k^{th} device at the t^{th} round, η is the learning rate.

c. Model Aggregation: The centralized server receives the updated model parameters from active devices and aggregates them using a weighted average. The weight for each edge device's model parameters is determined by the ratio of its data points n_k to the total (n) across all devices (K) as equation 2:

$$ w_{t+1} = \sum_{k=1}^{K} \frac{n_k}{n} w_{t+1}^k \tag{2}$$

d. Broadcast and Global Update: The averaged global model parameters w_{t+1} are broadcasted back to all edge devices, serving as the updated global model.

The FL approach enables continuous updating of the global model by collecting local model parameters from multiple edge devices. Each device uses its local data to generate a personalized model, ensuring data privacy without centralizing sensitive patient data. This privacy-preserving framework enhances glucose prediction accuracy for individual patients by leveraging collective knowledge.

Performance Measurement

To assess the predictive ability of the proposed personalized FDL models, this study used RMSE as a metric. RMSE quantifies the average magnitude of the differences between predicted and actual values, providing a measure of prediction accuracy. It considers both the direction and magnitude of errors, allowing for meaningful comparisons between different models. The RMSE value is produced by [32]:

$$ RMSE = \frac{\sum_{m=1}^{M} (\widehat{Y_m} - Y_m)^2}{n} \tag{3}$$

where $\widehat{Y_m}$ and Y_m are the predicted and actual values for observation m, with a total of M oberseataions.

IV. Evaluation

We have performed extensive experiments to evaluate the performance of the proposed approach. In this part, we demonstrate our experimental results.

A. Data Generation

Due to the limited patient data availability, many studies on glucose prediction have utilized simulators like the UVA/Padova Type 1 diabetes simulator, approved by the Food and Drug Administration [33]. In this study, we used a Python implementation of the UVa/Padova Simulator developed by Xie [34]. This custom simulator, based on the UVa/Padova Simulator (2008 version), allowed us to conduct experiments and analysis effectively. The simulator consists of thirty virtual patients, with ten patients in each of the three age groups: adults, adolescents, and children. Each virtual patient is equipped with a unique glucose-insulin response model, basal-bolus control model, and glucose sensor model. While the simulator allowed for automatic generation of food intake and insulin amounts, researchers can also customize parameters to simulate various scenarios and adapt treatment to indivisual virtual patients. However, the default scenario of random meals, food amount, and timing generation may not accurately represent real patients' typical eating habits and variations.

To address this limitation and create a more realistic simulation, this study defined food intake parameters, includeing number of meals, food amount, and timing. This customization allowed for better representation of real-world dietary patterns and individual patient habits. The predefined meal schedule consisted of breakfast, lunch, dinner, and snacks, accounting for 30%, 45%, 25%, and 5% of the total daily food intake, respectively. Time and food amount variations, ranging from 1 to 30 minutes and 5%, respectively, were introduced to accurately simulate real-world scenarios and individual patient habits.

For each patient, this study generated 5 days of data, capturing glucose, insulin, and food intake (represented as carbohydrates (CHO)) values at 3-minute intervals. CHO values remained zero except during designated time slots for food intake, and insulin values remained constant until 3 minutes after CHO intake, when additional insulin was administered to counteract the anticipated increase in glucose levels. Figure 3 presents the glucose, CHO, and insulin values of patient adult#001 simulated over a 5 days using self-designed food intake and time inputs. The first day's data is displayed within a 24-hour timeframeThe y-axis on the left represents the scale for glucose, while the y-axis on the right represents the scales for insulin and CHO. The RNN employed a window slide of 3 hours to predict glucose levels 30 minutes in advance, while the ANN used a single data point for the same prediction timeframe.

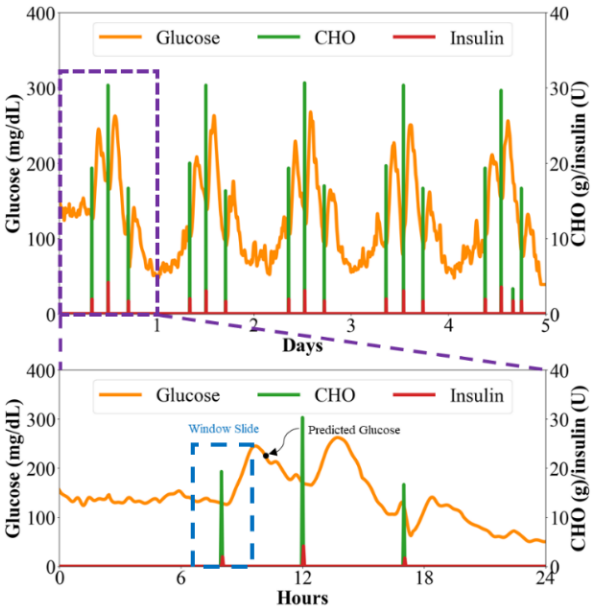

Fig. 3. Glucose, CHO, and insulin profiles - 5 Days & Day 1 Overview

B. Model Setup

Our model setup involves the following steps: Initially, we use a small dataset consisting of 5 patients' blood glucose data spanning 5 days each to train the initial global model. Once the initial global model is trained, it is distributed to 25 individual patients participating in the FL process. These individual patient then train their local models using the initial global model as a starting point and adapt them to their personalized data and device. Through FL, their models continuously improve and adapt based on the unique characteristics of their individual blood glucose data. The improvements from the individual models are periodically aggregated to update and enhance the global model. The iterative process of local model training, aggregation, and global model updates is repeated to refine the predictions over multiple rounds.

ANN and RNN- based DL models were utilized for glucose prediction 30 minutes ahead. The ANN model considers glucose, CHO, and insulin values at a single time point to predict the glucose value 30 minutes later. In contrast, the RNN model used a 3-hour window slide of time series data to predict the glucose value. This allows the RNN model to capture temporal dependencies for improved accuracy. Both models have two hidden layers with 64 and 32 neurons, respectively, and a batch size of 32. The models were trained using 80% of the data, while the remaining 20% was used for testing.

C. Results

In this section, we present the outcomes of our FL approach focusing on fine-tuning individual patient prediction models and observing improvement over time. Due to limited data availability, we used 10 epochs for each local and global model. To ensure model stability and prevent overfitting, we set the patience to 5, allowing sufficient epochs before considering early stopping. However, we noticed instances where the RMSE increased with more epochs, indicating possible overfitting. To address this concern, we implemented model checkpointing, saving the best-performing local models at different epochs. These best localmodels are then used to create a new global model, leveraging the best-performing models from all local devices, thereby improving the overall performance.

The initial global models (ANN and RNN-based) were created using data from 5 patients. Each of the 25 edge devices received the global model and fine-tuned it with 1 to 10 epochs to generate their first local model. Figure 4 displays the average performance of patients in each group (adolescent, adult, child) and the overall average RMSE for glucose prediction 30 minutes ahead for both ANN and RNN-based DL models. The results indicate that the ANN-based model had minor improvements in the first 2 epochs, whereas the RNN-based models consistently decreased in RMSE from epoch 1 to 10.

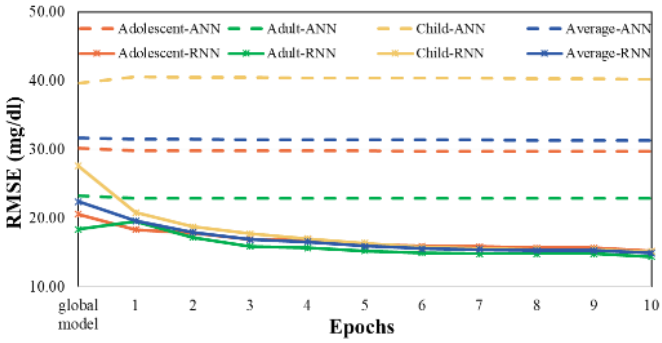

Fig. 4. RMSE values for glucose predictions using ANN model and RNN model.

Table I presents the RMSE values for the ANN and RNN-based DL models with 10 epochs for each patient (local model) and initial global models. The percentage change (**% Δ**) in RMSE is calculated by comparing the local models with the global models. Through FL, substantial improvements were observed in the prediction models for each patient. The initial global model, distributed to the patients, served as a strong foundation for refining their local models. As patients continued training their models with their personalized data, the models adapted to their unique physiological variations, leading to more accurate and tailored predictions. For the ANN-based global model, the percentage change ranges from 0.2% to 4.71% when fine-tuning the first local model across the 25 devices. In contrast, the percentage change for the RNN-based global model varies from 21.97% to 70.73% upon fine-tuning the first local model.

TABLE I. COMPARATIVE ANALYSIS OF RMSE IN ANN AND RNN-BASED DL MODELS: LOCAL MODEL, GLOBAL MODEL, AND PERCENTAGE CHANGE.

Patients	Local Model (ANN)	Global Model (ANN)	% Δ (ANN)	Local Model (RNN)	Global Model (RNN)	% Δ (RNN)
adolescent#003	23.29	23.72	-1.82	13.94	17.87	-21.97
adolescent#004	27.64	27.99	-1.25	13.70	19.44	-29.51
adolescent#005	24.01	24.70	-2.77	13.53	17.86	-24.23
adolescent#006	26.79	27.79	-3.59	11.91	18.32	-34.99
adolescent#007	46.87	47.35	-1.02	14.53	28.19	-48.45
adolescent#008	44.70	45.30	-1.31	14.23	26.75	-46.82
adolescent#009	26.58	27.11	-1.96	13.67	18.55	-26.31
adolescent#010	16.67	17.19	-3.03	13.61	17.79	-23.50
adult#003	28.30	28.89	-2.02	13.02	18.59	-29.93
adult#004	29.07	29.74	-2.26	13.09	18.55	-29.43
adult#005	20.04	20.50	-2.22	13.20	19.14	-31.00
adult#006	19.33	19.76	-2.17	13.37	17.96	-25.55
adult#007	18.56	19.11	-2.84	12.53	17.24	-27.32
adult#008	16.83	17.16	-1.96	13.28	17.67	-24.81
adult#009	20.43	20.98	-2.61	13.36	18.82	-29.04
adult#010	29.34	29.82	-1.59	12.53	18.98	-33.96
child#002	20.26	20.51	-1.25	14.34	19.28	-25.62
child#003	45.87	46.25	-0.81	10.44	22.51	-53.64
child#004	24.14	24.63	-1.98	12.15	24.20	-49.78
child#005	12.40	12.42	-0.20	12.21	21.91	-44.28
child#006	41.67	42.12	-1.08	13.04	23.45	-44.40
child#007	32.58	33.11	-1.61	13.06	24.75	-47.24
child#008	86.36	87.29	-1.07	15.18	51.86	-70.73
child#009	49.72	52.18	-4.71	13.79	34.23	-59.71
child#010	46.49	47.49	-2.11	13.07	25.99	-49.71

*unit of RMSE is mg/dL

After completing the FL approach, we compared it with centralized learning (CL) methods. Table II presents a contrast between the ANN and RNN-based CL (ANN-CL and RNN-CL) and their FL counterparts (ANN-FL and RNN-FL). The table displays the average RMSE for different patient groups using ANN-CL, RNN-CL, ANN-FL, and RNN-FL models. or instance, the RNN-FL model achieved an RMSE of 13.03 for the child group, while the RNN-CL model resulted in an RMSE of 22.77. Overall, the FDL models exhibited lower RMSE compared to the CL models. The percentage change was -4.21% to -11.97% for ANN-CL versus ANN-FL, and -17.72% to -42.77% for RNN-CL versus RNN-FL.

The comparison reveals that our FL approach demonstrated results comparable to the CL model while maintaining decentralized patient data and ensuring data privacy. Unlike CL methods that rely on sharing sensitive data with a central server, our approach leveraged the knowledge from diverse individual models without compromising patient privacy.

TABLE II. RMSE COMPARISON: ANN AND RNN-BASED CL VS. FL MODELS.

Patients	Adolescent	Adult	Child	Average
ANN-CL	31.88	23.74	45.37	35.17
RNN-CL	18.75	15.86	22.77	19.83
ANN-FL	29.57	22.74	39.94	31.12
RNN-FL	13.64	13.05	13.03	13.23
ANN-CL vs ANN-FL (%)	-7.25	-4.21	-11.97	-11.52
RNN-CL vs RNN-FL (%)	-27.25	-17.72	-42.77	-33.27

*unit of RMSE is mg/dL

We not only consider the CL approach, where all 25 patients send their five days of data to a central server to generate one model for all the data. We also explore the scenario where patients prioritize privacy and prefer not to share their data. In this case, each of the 25 devices trains an ANN and RNN-based DL model with their 5 days of data.

Table III compares the results of the ANN and RNN models alone to their performance with FL. Using the ANN model alone with ANN-FL showed no significant difference, likely due to limited data available for effective ANN training. However, the RNN- FL demonstrates remarkable performance. Despite the challenge of achieving low RMSE values with only 5 days of patient data, the RNN-FL achieved substantial percentage decreases in RMSE 39.61%, 38.83%, 56.32%, and 46.63% for the adolescent, adult, child, and average groups, respectively, compared to using the RNN model alone. These findings underscore the effectiveness of FL in significantly improving the RNN model's performance for glucose prediction tasks.

TABLE III.RMSE COMPARISON: ANN AND RNN-BASED DL VS. FDL MODELS.

	Adolescent	Adult	Child	Average
ANN	30.23	23.25	41.16	31.93
RNN	22.59	21.33	29.83	24.79
ANN-FL	29.57	22.74	39.94	31.12
RNN-FL	13.64	13.05	13.03	13.23
ANN vs. ANN-FL (%)	-2.19	-2.22	-2.95	-2.55
RNN vs. RNN-FL (%)	-39.61	-38.83	-56.32	-46.63

*unit of RMSE is mg/dL

We compared the performance of FDL for each patient and selected the best improvement among the three groups compared to using DL alone. The percentage decrease in RMSE between the predicted values using the RNN model alone and the predicted values using the RNN- FL is 49.58%, 50.91%, and 72.44% for patients adolescent#007, adult#010, and child#008, respectively. This substantial improvement highlights the effectiveness of using FL in conjunction with the RNN model for glucose prediction. Figure 5 displays the 12 hours real data, the predicted glucose values from the RNN model, and the predicted glucose values from the RNN model with FL for patients adolescent#007, adult#010, and child#008.

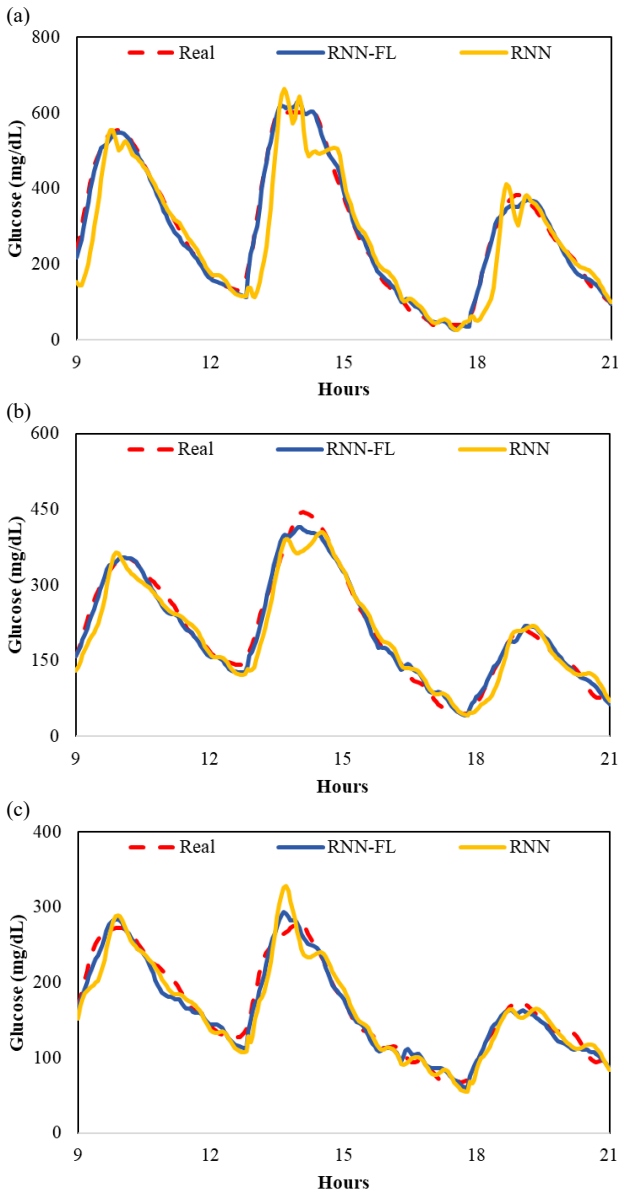

Fig. 5. Comparison of real glucose and predicted glucose (RNN and RNN-FL) for patient (a) adolescent#007, (b) adult#010, and (c) child#008.

V. Conclusions

In conclusion, we have developed a system that utilizes edge devices like CGM and smartwatches to help patients manage diabetes effectively. Our approach employs personalized FDL, allowing model training on multiple IoT devices while maintaining data privacy. The integration of personalization components tailors models to individual users, considering unique factors affecting glucose levels. The system is optimized for IoT devices with limited resources, and experimental results on simulated datasets validate its performance in diabetes management. This privacy-preserving and personalized approach empowers patients to make informed decisions for better health outcomes. Future work includes expanding the research to larger patient populations and incorporating additional edge devices to enhance diabetes management capabilities.

References

[1] "National Diabetes Statistics Report 2020. Estimates of diabetes and its burden in the United States.," 2020.

[2] D. K. Wukich *et al.*, "Amputation and infection are the greatest fears in patients with diabetes foot complications," *Journal of Diabetes and its Complications*, vol. 36, no. 7, p. 108222, 2022.

[3] J. Li and C. Fernando, "Smartphone-based personalized blood glucose prediction," *ICT Express*, vol. 2, no. 4, pp. 150–154, 2016.

[4] J. Wang *et al.*, "Multiple linear regression and artificial neural network to predict blood glucose in overweight patients," *Experimental and clinical endocrinology & diabetes*, vol. 124, no. 01, pp. 34–38, 2016.

[5] Aashima, S. Bhargav, S. Kaushik, and V. Dutt, "A combination of decision trees with machine learning ensembles for blood glucose level predictions," in *Proceedings of International Conference on Data Science and Applications: ICDSA 2021, Volume 2*, Springer, 2022, pp. 533–548.

[6] G. Alfian *et al.*, "Blood glucose prediction model for type 1 diabetes based on artificial neural network with time-domain features," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 4, pp. 1586–1599, 2020.

[7] T. Hamdi *et al.*, "Artificial neural network for blood glucose level prediction," in *2017 International Conference on Smart, Monitored and Controlled Cities (SM2C)*, IEEE, 2017, pp. 91–95.

[8] J. B. Ali, T. Hamdi, N. Fnaiech, V. Di Costanzo, F. Fnaiech, and J.-M. Ginoux, "Continuous blood glucose level prediction of type 1 diabetes based on artificial neural network," *Biocybernetics and Biomedical Engineering*, vol. 38, no. 4, pp. 828–840, 2018.

[9] K. Li, J. Daniels, C. Liu, P. Herrero, and P. Georgiou, "Convolutional recurrent neural networks for glucose prediction," *IEEE journal of biomedical and health informatics*, vol. 24, no. 2, pp. 603–613, 2019.

[10] F. D'Antoni *et al.*, "Prediction of glucose concentration in children with type 1 diabetes using neural networks: An edge computing application," *Bioengineering*, vol. 9, no. 5, p. 183, 2022.

[11] T. Zhu, K. Li, P. Herrero, J. Chen, and P. Georgiou, "A Deep Learning Algorithm for Personalized Blood Glucose Prediction.," in *KHD@ IJCAI*, 2018, pp. 64–78.

[12] Y. Dong, R. Wen, Z. Li, K. Zhang, and L. Zhang, "Clu-RNN: A new RNN based approach to diabetic blood glucose prediction," in *2019 IEEE 7th International Conference on Bioinformatics and Computational Biology (ICBCB)*, IEEE, 2019, pp. 50–55.

[13] D.-Y. Kim *et al.*, "Developing an individual glucose prediction model using recurrent neural network," *Sensors*, vol. 20, no. 22, p. 6460, 2020.

[14] M. Munoz-Organero, "Deep physiological model for blood glucose prediction in T1DM patients," *Sensors*, vol. 20, no. 14, p. 3896, 2020.

[15] S. Boughorbel, F. Jarray, N. Venugopal, S. Moosa, H. Elhadi, and M. Makhlouf, "Federated uncertainty-aware learning for distributed hospital ehr data," *arXiv preprint arXiv:1910.12191*, 2019.

[16] R. Duan *et al.*, "Learning from electronic health records across multiple sites: A communication-efficient and privacy-preserving distributed algorithm," *Journal of the American Medical Informatics Association*, vol. 27, no. 3, pp. 376–385, 2020.

[17] L. Huang, A. L. Shea, H. Qian, A. Masurkar, H. Deng, and D. Liu, "Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records," *Journal of biomedical informatics*, vol. 99, p. 103291, 2019.

[18] Z. Li, K. Roberts, X. Jiang, and Q. Long, "Distributed learning from multiple EHR databases: contextual embedding models for medical events," *Journal of biomedical informatics*, vol. 92, p. 103138, 2019.

[19] Q. Wu, K. He, and X. Chen, "Personalized federated learning for intelligent IoT applications: A cloud-edge based framework," *IEEE Open Journal of the Computer Society*, vol. 1, pp. 35–44, 2020.

[20] Y. Zhao, H. Haddadi, S. Skillman, S. Enshaeifar, and P. Barnaghi, "Privacy-preserving activity and health monitoring on databox," in *Proceedings of the Third ACM International Workshop on Edge Systems, Analytics and Networking*, 2020, pp. 49–54.

[21] J. Lo *et al.*, "Federated learning for microvasculature segmentation and diabetic retinopathy classification of OCT data," *Ophthalmology Science*, vol. 1, no. 4, p. 100069, 2021.

[22] Y. Su, C. Huang, W. Zhu, X. Lyu, and F. Ji, "Multi-party diabetes mellitus risk prediction based on secure federated learning," *Biomedical Signal Processing and Control*, vol. 85, p. 104881, 2023.

[23] M. Lincy and A. M. Kowshalya, "Early detection of type-2 diabetes using federated learning," *International Journal of Scientific Research in Science, Engineering and Technology*, vol. 12, pp. 257–267, 2020.

[24] I. De Falco *et al.*, "A Federated Learning-Inspired Evolutionary Algorithm: Application to Glucose Prediction," *Sensors*, vol. 23, no. 6, p. 2957, 2023.

[25] P. V. Astillo, D. G. Duguma, H. Park, J. Kim, B. Kim, and I. You, "Federated intelligence of anomaly detection agent in IoTMD-enabled Diabetes Management Control System," *Future Generation Computer Systems*, vol. 128, pp. 395–405, 2022.

[26] C. Y. Tang, K. Fung, E. W. Lee, G. T. Ho, K. W. Siu, and W. Mou, "Product form design using customer perception evaluation by a combined superellipse fitting and ANN approach," *Advanced Engineering Informatics*, vol. 27, no. 3, pp. 386–394, 2013.

[27] F. Rosenblatt, "Principles of neurodynamics. perceptrons and the theory of brain mechanisms," Cornell Aeronautical Lab Inc Buffalo NY, 1961.

[28] M. W. Gardner and S. Dorling, "Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences," *Atmospheric environment*, vol. 32, no. 14–15, pp. 2627–2636, 1998.

[29] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[30] A. Fan, T. Lavril, E. Grave, A. Joulin, and S. Sukhbaatar, "Addressing some limitations of transformers with feedback memory," *arXiv preprint arXiv:2002.09402*, 2020.

[31] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*, PMLR, 2017, pp. 1273–1282.

[32] H. Michael, J. Christopher, N. John, and L. William, "Applied linear statistical models," *5th ed. Boston: McGraw-Hil*, 2005.

[33] C. D. Man, F. Micheletto, D. Lv, M. Breton, B. Kovatchev, and C. Cobelli, "The UVA/PADOVA type 1 diabetes simulator: new features," *Journal of diabetes science and technology*, vol. 8, no. 1, pp. 26–34, 2014.

[34] J. Xie, "Simglucose v0. 2.1 (2018)," *URL https://github. com/jxx123/simglucose*, 2018.