Intelligent Real-Time Hypoglycemia Prediction for Type 1 Diabetes

Nevien A. Mahdy
Electrical Engineering Department
Benha Faculty of Engineering, Benha University
Benha, Egypt
niven.mahdy@bhit.bu.edu.eg

Ahmed F. Elnokrashy
[1]*Electrical Engineering Department*
Benha Faculty of Engineering, Benha University
Benha, Egypt
[2]*Computer Science Department*
Faculty of Information Technology and Computer Science, Nile University
Giza, Egypt
ahmed.elnokrashy@bhit.bu.edu.eg

D. A. Hammad
Basic Engineering Sciences Department
Benha Faculty of Engineering, Benha University
Benha, Egypt
doaa.hammad@bhit.bu.edu.eg

Wael A. Mohamed
Electrical Engineering Department
Benha Faculty of Engineering, Benha University
Benha, Egypt
Wael.ahmed@bhit.bu.edu.eg

***Abstract*—Hypoglycemia in Type 1 Diabetes (T1D) refers to a condition where blood glucose (BG) levels drop to abnormally low levels, typically below 70 mg/dL. This can occur when there is an excessive amount of insulin relative to the blood glucose level, leading to an imbalance that can be dangerous and potentially life-threatening if not promptly treated. The availability of large amounts of data from continuous glucose monitoring (CGM), insulin doses, carbohydrate intake, and additional vital signs, together with deep learning (DL) techniques, has revolutionized algorithmic approaches for BG prediction in T1D, achieving superior performance. In our study, we employed a Long Short-Term Memory (LSTM) neural network architecture to predict hypoglycemia events in patients with T1D. For the training and testing, we utilized the OhioT1DM (2018) dataset. In addition, real-time data collected from an individual patient for the evaluation. This patient utilized the CGM FreeStyle Libre (FSL) system, along with a smartwatch to monitor step count. The LSTM model exhibited performance demonstrating exceptional levels of sensitivity, specificity, and accuracy scores of 97.09%, 94.17%, and 95.63%, respectively, when assessed using the Ohio test dataset. Our research provides strong evidence supporting the system's efficacy in managing hypoglycemia events in individuals diagnosed with T1D.**

***Index Terms*—Hypoglycemia, Deep learning, Type 1 diabetes, Glucose prediction, Classification**

I. Introduction

Type 1 diabetes (T1D) is a chronic medical condition characterized by pancreatic insulin deficiency, leading to impaired regulation of blood glucose levels. T1D is an autoimmune disease where the immune system mistakenly targets and destroys the insulin-producing β cells in the pancreas [1], [2], [3]. The management of T1D requires the adoption of intensive insulin therapies, which represent a lifelong commitment. These therapies enable patients to effectively control elevated blood glucose (BG) levels, thereby preventing hyperglycemia [4], they also pose the risk of hypoglycemia. Individuals with T1D face the complex challenge of achieving glycemic control within a normal range, striking a balance between preventing and avoiding hyperglycemia events [5].

The scientific integration of continuous glucose monitoring (CGM) with multiple daily injections (MDI) therapy results in the acquisition of a substantial dataset, which presents promising prospects for advancing diabetes management [6]. Expanding on this framework, machine learning (ML) methodologies have been utilized to forecast BG levels continuously. Continuous predictive modeling plays a crucial role in empowering patients to take proactive measures to prevent the occurrence of hypoglycemia or hyperglycemia. Extensive research efforts are dedicated to this topic, as evidenced by numerous investigations in the field [7], [8], [9], [10]. In a comprehensive study conducted by Oviedo and collaborators on blood glucose prediction techniques [11], only a minor portion, comprising 13% of the studies, focused on predicting adverse glycemia events, while the vast majority, accounting for 87%, concentrated on the continuous forecasting of BG levels.

The primary focus of most studies is centered around predicting hypoglycemia [12], [13], [14], [15], [16]. For instance, Reddy et al. [13] applied ML to predict the onset of hypoglycemia events during aerobic exercise in individuals with T1D. In a distinct study, Oviedo et al. [12] utilized support vector machines (SVM) with historical data from 10 adults with T1D undergoing Sensor-Augmented Pump (SAP) therapy to forecast postprandial hypoglycemia occurrences.

979-8-3503-6263-3/24/$31.00 ©2024 IEEE

Other studies have introduced the prediction of hypoglycemia using personalized models, as demonstrated in the cited references [1], [17]. Current methodologies have refocused their efforts on predicting a spectrum of adverse events in T1D, encompassing both hypoglycemia and hyperglycemia [18], [19]. Their approach entails the classification of adverse glycemia events rather than employing regression, resulting in enhanced performance outcomes.

In this study, we present an approach for predicting hypoglycemia events in patients with T1D. Initially, we conducted data preprocessing on data collected from a volunteer and the OhioT1DM dataset. Subsequently, we applied a Long Short-Term Memory (LSTM) neural network to forecast glucose levels using a diverse set of features. Furthermore, we employed a classification framework to assess the predictive performance of the model in anticipating hypoglycemia events. Finally, we introduced a comprehensive system for predicting and managing hypoglycemia events, which we subsequently validated.

II. Materials and Methods

The proposed hypoglycemia management technique includes thorough preprocessing, data splitting, model construction, hyperparameter tuning, and comprehensive evaluation using two distinct test datasets to assess the effectiveness of our hypoglycemia prediction strategy. The workflow of our approach is described below in Fig. 1.

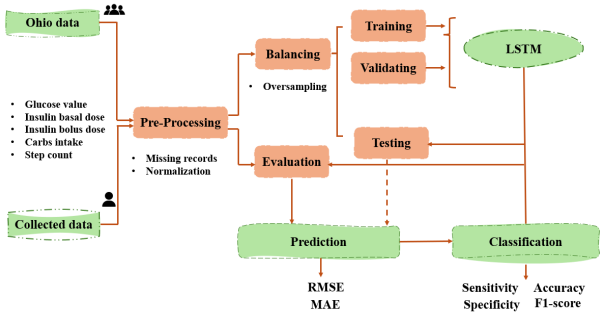

Fig. 1. The framework of our proposed methodology.

In this study, we utilized the OhioT1DM dataset [20], which consists of comprehensive data from 12 patients diagnosed with T1D undergoing insulin pump therapy. The dataset includes data collected from CGM devices and activity bands worn by the patients, who diligently recorded their daily activities using a designated smartphone application over an approximately eight-week period. The Blood Glucose Level Prediction (BGLP) Challenge, initiated in 2018, provided data from six patients (referred to as Group 1) from this dataset. Subsequently, in the 2020 iteration of the challenge, an additional dataset comprising six more patients (designated as Group 2) was introduced. For our study, we specifically focused on using the Group 1 dataset due to the presence of specific data features that were deemed essential for our research objectives and were not available in the Group 2 dataset. A detailed description of patient devices, data format, and dataset characteristics can be found in the OhioT1DM dataset publication [20].

Additionally, authentic patient data obtained during a case study was utilized for validating the model. A 17-year-old male patient diagnosed with T1D utilized the CGM FreeStyle Libre (FSL) system for continuous monitoring of BG levels. The patient also employed the TicWatch Pro 3 Ultra smartwatch to collect data on physical activity. The patient received comprehensive training on operating these devices, including data extraction from the smart-watch via the Google Fit platform. Detailed records were maintained for each administered insulin dose, covering both rapid-acting and long-acting insulin, as well as estimates for the carbohydrate content of every meal consumed. Participation in the study was formalized through the execution of a consent form, signifying explicit agreement to participate in the research.

A. Data Pre-Processing

In our study, both the OhioT1DM and Collected datasets were utilized, incorporating a common set of five essential recorded features for experimental investigations. These features included CGM values, insulin basal rate (indicating long-acting insulin administration), bolus amount (reflecting short-acting insulin doses), carbohydrate intake (indicating meal consumption), and step count. To align with the Collected dataset, we resampled the 5-minute interval CGM values to 15-minute intervals using the mean method. Missing data, particularly regarding glucose values, were addressed using forward and backward filling techniques for data imputation. Bolus amounts, step counts, and carbohydrate intake values were set to zero when absent or unrecorded. Insulin basal rate values were imputed using forward filling, given that changes were only recorded when adjustments occurred; missing data indicated continuity of the most recent value. To standardize feature values and ensure consistency across the datasets, normalization was applied, scaling all features to a uniform range from zero to one.

B. Data Balancing

The OhioT1DM and Collected datasets displayed a significant imbalance in the number of CGM intervals containing reported hypoglycemia events compared to those without such events. To address this imbalance and achieve a more equitable data distribution, we implemented an oversampling technique. we partitioned the data into discrete units, each consisting of a specific number of sequential samples. We then classified each unit based on the blood glucose level of its final sample. Units were categorized as hypoglycemia units if the last sample exhibited blood sugar levels below 70 mg/dL [2], [21], and as normal units if the last sample showed blood sugar levels at or above 70 mg/dL. Following classification, we identified the number of hypoglycemic units and normal units to guide the oversampling procedure for the training dataset. This involved duplicating the number of units within the minority class

(hypoglycemia units) to achieve a balanced representation equivalent to that of the majority class (normal units).

The OhioT1DM dataset was divided into separate subsets for training, validation, and testing. Specifically, 70% of the hypoglycemia units were allocated to the training subset, while 20% were designated for validation, and the remaining 10% were reserved for testing. Additionally, the Collected dataset was exclusively used for testing purposes. We established a scenario for the number of samples per unit. The scenario comprised 13 samples per unit, covering a total temporal span of three hours of data, as illustrated in Fig. 2.

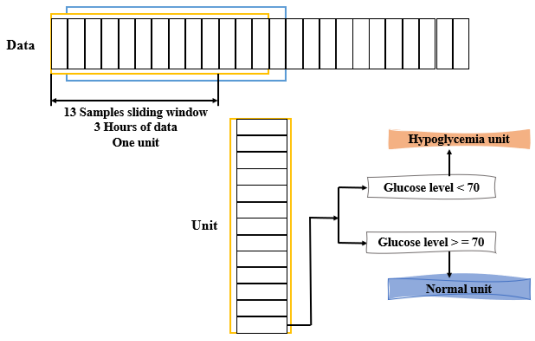

Fig. 2. The methodology employed for establishing the unit entailed iteratively adjusting the sampling window by one sample.

C. Long Short-Term Memory (LSTM) Networks

LSTMs were initially proposed by Hochreiter and Schmidhuber [22]. LSTMs are a specialized category of Recurrent Neural Networks (RNNs) designed to capture dependencies across extended temporal sequences effectively. This unique capability enables LSTMs to retain crucial sequential patterns, enhancing parameter learning efficiency, particularly for time series data analysis. In our study, the LSTM model incorporates historical information spanning a specified number of time steps as input to generate single-step prediction. This approach involves leveraging past data to inform current predictions, enabling the model to capture temporal dependencies and make accurate forecasts based on historical context. we implemented it using the Keras platform. The model architecture included an LSTM layer followed by a dropout layer, a densely connected layer, and a final output layer with a single unit for prediction.

During training, we employed a batch size of 32 and trained the model for 200 epochs. To monitor performance and ensure model reliability, we implemented checkpointing, saving model weights every 20 epochs and selecting the best-performing model checkpoint from epoch 80.

D. Hyperparameter Tuning

It is essential to recognize that default hyperparameters may not be optimal for specific datasets, necessitating the fine-tuning of these parameters to achieve substantial improvements in model performance. In our endeavor to optimize hyperparameters, we employed Bayesian optimization [23], which has proven effective in systematically improving model configurations. A comprehensive description of the model's hyperparameters, along with their corresponding optimized values, is provided below in Table I.

TABLE I
HYPERPARAMETER CONFIGURATIONS.

Hyperparameter	LSTM model
No. of units for the first layer	200
Dropout rate	0.1
No. of units for the dense layer	150
Learning rate	0.01

E. Classification for Hypoglycemia Events Prediction

Following the prediction phase, a classification threshold was applied. Specifically, the threshold was set at 70 mg/dL for glucose levels. Glucose values above this threshold were categorized as "Normal events," while values below were classified as "Hypoglycemia events".

F. Performance Metrics

This approach employed a diverse set of metrics to evaluate the performance of the implemented methodology.

1. Prediction

To evaluate the performance of the model following the prediction of BG levels, we employed the root mean squared error (RMSE) as computed by (1) and mean absolute error (MAE) as expressed in (2).

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{n=1}^{n}(y_i - \hat{y}_i)^2} \tag{1}$$

$$\text{MAE} = \frac{1}{n}\sum_{n=1}^{n}|y_i - \hat{y}_i| \tag{2}$$

where y_i is actual glucose level and $\hat{y}_i$ is predicted glucose level, both measured in mg/dl.

2. Classification

To evaluate the classification performance, metrics derived from the confusion matrix were utilized. Table II presents a concise summary of primary evaluation metrics, including sensitivity (SE), specificity (SP), accuracy, and the F1 score, utilized to assess the model's performance.

III. RESULTS

A. Prediction Performance

Table III presents the RMSE and MAE outcomes for the Ohio test dataset. Table IV provides a detailed summary of the RMSE and MAE outcomes, with specific emphasis on BG values below the critical threshold of 70 mg/dL.

TABLE II
METRICS TO EVALUATE THE PERFORMANCE OF THE MODEL.

Performance Metrics
SE = $\frac{TP}{TP+FN}$
Sp= $\frac{TN}{TN+FP}$
Accuracy = $\frac{TP+TN}{TP+FN+TN+FP}$
F1score = $\frac{2TP}{2TP+FP+FN}$

TABLE III
TOTAL RMSE AND MAE RESULTS FOR THE TOTAL TEST DATASET.

Test dataset	RMSE	MAE
Ohio	21.998	9.64
Collected	41.75	20.571

B. Classification

Table V presents a detailed synthesis of the BG classification outcomes. Following BG prediction, each sample undergoes systematic classification as either indicating a hypoglycemia event or representing a normal event.

IV. DISCUSSION

The results reveal a notable increase in both the total RMSE and MAE for the Collected test dataset in comparison to the Ohio test dataset. Specifically, when focusing on hypoglycemia cases, there is a substantial escalation in the RMSE and MAE error rates, underscoring the challenge of accurately identifying hypoglycemia events within the Collected test dataset. Additionally, The LSTM model exhibits significantly higher accuracy when evaluated on the Ohio test dataset compared to the Collected test dataset.

The difference in performance when using real-time data compared to the OhioT1DM dataset is attributed to the fact that patients in the OhioT1DM dataset used insulin pumps, whereas the real-time data did not. This discrepancy in insulin delivery methods introduces variability that affects the model's ability to generalize and maintain its predictive accuracy across different datasets.

For instance, in the evaluation of the LSTM model using the Ohio test dataset, which included 103 hypoglycemia events and 103 normal events, it identified 106 cases as hypoglycemia events and 100 cases as normal events. However, when applied to the Collected dataset, which contained 46 cases of hypoglycemia events and 46 cases of normal events, it classified 7 cases as hypoglycemia events and 85 cases as normal events.

TABLE IV
RMSE AND MAE RESULTS FOR BG VALUES LESS THAN 70 MG/DL.

Test dataset	RMSE	MAE
Ohio	9.049	6.124
Collected	49.934	41.142

TABLE V
SE, SP, ACCURACY, AND F1-SCORE FOR THE LSTM MODEL.

Test dataset	LSTM model	
Ohio	SE %	97.09
	SP %	94.17
	Accuracy %	95.63
	F1-score %	95.69
Collected	SE %	13.04
	SP %	97.83
	Accuracy %	55.43
	F1-score %	22.64

V. CONCLUSION

In this paper, we propose an advanced approach to predict hypoglycemia events in patients with T1D using a comprehensive feature set that includes glucose levels, insulin doses, carbohydrate intake, and step count. Accurately predicting hypoglycemia events is crucial as it empowers patients to take proactive preventive measures, thereby enhancing their overall health and safety.

Our findings indicate that the LSTM model demonstrates robust performance with the Ohio test dataset. However, its performance was less satisfactory when applied to real-time data, highlighting a limitation in the model's ability to generalize to real-world scenarios.

To address these limitations, future works should explore the implementation of more advanced DL models, which could potentially offer improved performance and greater robustness. Additionally, expanding the range of features incorporated into the analysis could enhance the model's effectiveness, ensuring it captures a more comprehensive set of factors influencing hypoglycemia events. By advancing these methodologies, we can move closer to developing reliable, real-time prediction tools that significantly benefit patients with T1D.

REFERENCES

[1] A. Bertachi, L. Biagi, I. Contreras, N. Luo, and J. Vehí, "Prediction of blood glucose levels and nocturnal hypoglycemia using physiological models and artificial neural networks."

[2] A. Bertachi et al., "Prediction of nocturnal hypoglycemia in adults with type 1 diabetes under multiple daily injections using continuous glucose monitoring and physical activity monitor," Sensors (Switzerland), vol. 20, no. 6, Mar. 2020, doi: 10.3390/s20061705.

[3] A. Parcerisas et al., "A Machine Learning Approach to Minimize Nocturnal Hypoglycemic Events in Type 1 Diabetic Patients under Multiple Doses of Insulin," Sensors, vol. 22, no. 4, Feb. 2022, doi: 10.3390/s22041665.

[4] A. Aliberti et al., "A Multi-Patient Data-Driven Approach to Blood Glucose Prediction," IEEE Access, vol. 7, pp. 69311–69325, 2019, doi: 10.1109/ACCESS.2019.2919184.

[5] E. Chen, F. King, M. A. Kohn, E. K. Spanakis, M. Breton, and D. C. Klonoff, "A Review of predictive low glucose suspend and its effectiveness in preventing nocturnal hypoglycemia," Diabetes Technology and Therapeutics, vol. 21, no. 10. Mary Ann Liebert Inc., pp. 589–601, Oct. 01, 2019. doi: 10.1089/dia.2019.0119.

[6] I. Contreras and J. Vehi, "Artificial intelligence for diabetes management and decision support: Literature review," Journal of Medical Internet Research, vol. 20, no. 5. JMIR Publications Inc., 2018. doi: 10.2196/10775.

[7] I. Contreras, S. Oviedo, M. Vettoretti, R. Visentin, and J. Vehí, "Personalized blood glucose prediction: A hybrid approach using grammatical evolution and physiological models," PLoS One, vol. 12, no. 11, Nov. 2017, doi: 10.1371/journal.pone.0187754.

[8] IEEE Engineering in Medicine and Biology Society, Annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2012 San Diego, California, USA, Aug. 28 - Sept. 1, 2012.

[9] C. Zecchin, A. Facchinetti, G. Sparacino, G. De Nicolao, and C. Cobelli, "Neural network incorporating meal information improves accuracy of short-time prediction of glucose concentration," IEEE Trans Biomed Eng, vol. 59, no. 6, pp. 1550–1560, 2012, doi: 10.1109/TBME.2012.2188893.

[10] R. Bunescu, N. Struble, C. Marling, J. Shubrook, and F. Schwartz, "Blood glucose level prediction using physiological models and support vector regression," in Proceedings - 2013 12th International Conference on Machine Learning and Applications, ICMLA 2013, IEEE Computer Society, 2013, pp. 135–140. doi: 10.1109/ICMLA.2013.30.

[11] S. Oviedo, J. Vehí, R. Calm, and J. Armengol, "A review of personalized blood glucose prediction strategies for T1DM patients," International Journal for Numerical Methods in Biomedical Engineering, vol. 33, no. 6. Wiley-Blackwell, Jun. 01, 2017. doi: 10.1002/cnm.2833.

[12] S. Oviedo, I. Contreras, C. Quirós, M. Giménez, I. Conget, and J. Vehi, "Risk-based postprandial hypoglycemia forecasting using supervised learning," Int J Med Inform, vol. 126, pp. 1–8, Jun. 2019, doi: 10.1016/j.ijmedinf.2019.03.008.

[13] R. Reddy, N. Resalat, L. M. Wilson, J. R. Castle, J. El Youssef, and P. G. Jacobs, "Prediction of Hypoglycemia During Aerobic Exercise in Adults With Type 1 Diabetes," J Diabetes Sci Technol, vol. 13, no. 5, pp. 919–927, Sep. 2019, doi: 10.1177/1932296818823792.

[14] O. Mujahid, I. Contreras, and J. Vehi, "Machine learning techniques for hypoglycemia prediction: Trends and challenges," Sensors (Switzerland), vol. 21, no. 2. MDPI AG, pp. 1–21, Jan. 02, 2021. doi: 10.3390/s21020546.

[15] D. Dave et al., "Feature-Based Machine Learning Model for Real-Time Hypoglycemia Prediction," J Diabetes Sci Technol, vol. 15, no. 4, pp. 842–855, 2020, doi: 10.1177/1932296820922622.

[16] V. Felizardo, D. MacHado, N. M. Garcia, N. Pombo, and P. Brandao, "Hypoglycaemia Prediction Models With Auto Explanation," IEEE Access, vol. 10, pp. 57930–57941, 2022, doi: 10.1109/ACCESS.2021.3117340.

[17] A. Bhimireddy, P. Sinha, B. Oluwalade, J. W. Gichoya, and S. Purkayastha, "Blood Glucose Level Prediction as Time-Series Modeling using Sequence-to-Sequence Neural Networks." [Online]. Available: http://ceur-ws.org/Vol-2148/

[18] F. Iacono, L. Magni, and C. Toffanin, "Personalized LSTM-based alarm systems for hypoglycemia and hyperglycemia prevention," Biomed Signal Process Control, vol. 86, Sep. 2023, doi: 10.1016/j.bspc.2023.105167.

[19] F. D'Antoni et al., "Layered Meta-Learning Algorithm for Predicting Adverse Events in Type 1 Diabetes," IEEE Access, vol. 11, pp. 9074–9094, 2023, doi: 10.1109/ACCESS.2023.3237992.

[20] C. Marling and R. Bunescu, "The OhioT1DM Dataset for Blood Glucose Level Prediction: Update 2020," 2020. [Online]. Available: http://www.jdrf.org/impact/research/artificial-

[21] C. Duckworth et al., "Explainable Machine Learning for Real-Time Hypoglycemia and Hyperglycemia Prediction and Personalized Control Recommendations," J Diabetes Sci Technol, 2022, doi: 10.1177/19322968221103561.

[22] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Comput, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[23] P. K. P, M. A. B. V, and G. G. Nair, "An efficient classification framework for breast cancer using hyper parameter tuned Random Decision Forest Classifier and Bayesian Optimization," Biomed Signal Process Control, vol. 68, Jul. 2021, doi: 10.1016/j.bspc.2021.102682.