International Journal of Innovative Research in Computer Science and Technology (IJIRCST)
ISSN (Online): 2347-5552, Volume-14, Issue-4, July 2026
DOI: https://doi.org/10.55524/ijircst.2026.14.4.5
Article ID IRPC-1432, Pages 44-60
www.ijircst.org

Development of a Time-Series Deep Learning Framework for Real-Time Prediction of Hypoglycemia and Hyperglycemia

Mohammad Shafeeq[1], and Monika Tripathi[2]

[1, 2] Department of Computer Science and Engineering, P K University, Shivpuri, Madhya Pradesh, India

Correspondence should be addressed to Mohammad Shafeeq; quraish.shafeeq09@gmail.com

Received: 23 June 2026; Revised: 8 July 2026; Accepted: 22 July 2026

ABSTRACT- Predicting blood glucose variability accurately is essential in managing type 1 diabetes mellitus (T1DM). The occurrences of hypoglycemic and hyperglycemic events can be managed much better if detected early on, therefore preventing severe adverse events. Often existing deep learning methods either cannot capture local glucose variation patterns or do not model long-term temporal dependencies from the continuous glucose monitoring (CGM) data. To overcome this limitation, this study presents a novel multi-scale convolutional neural network-bidirectional long short-term memory-transformer (MSCNN-BiLSTM-Transformer) hybrid framework for multi-class glucose state prediction. In sequential glucose measures, the proposed model aims at learning a discriminative representation through multi-scale convolution-based feature extraction, bidirectional temporal learning and transformer based self-attention mechanism. The OhioT1DM dataset was utilized in experiments that employed 12 previous glucose readings as inputs to classify glucose states into three states, hypoglycemia, normal and hyperglycemia. The proposed model achieve an accuracy of 97.59%, precision of 97.78%, recall of 97.56%, F1 score of 97.60%, ROC-AUC of 99.79%, Matthews correlation coefficient (MCC) of 95.30%, Cohen's Kappa 95.76%, and outperform other conventional CNN, LSTM, GRU, Transformer and other baseline models. This work presents an ablation study that showed our multi-scale convolution, bidirectional sequence learning and Transformer contextual modelling and feature fusion is effective. Moreover, the SHAP-based explainability analysis sheds light on how historical glucose data affected the final prediction, demonstrating the proposed framework's clinical reliability and interpretability. The results obtained from the tests show that the revealed hybrid model is a robust, accurate, and interpretable model for intelligent glucose state prediction which has the potential for being used as a decision-support tool for personalized diabetes treatment and other smart healthcare solutions in the future.

KEYWORDS- Type 1 Diabetes Mellitus, Continuous Glucose Monitoring, Blood Glucose Prediction, Multi-Scale CNN, BiLSTM, Transformer, Explainable Artificial Intelligence, SHAP, Deep Learning, Healthcare Analytics

I. INTRODUCTION

Type 1 Diabetes Mellitus (T1DM) is a chronic metabolic disorder caused by autoimmune damage to β-cells of the pancreas that results in insufficient insulin production and glucose dysregulation. Maintaining physiological glucose levels is a major clinical challenge since glucose levels that are too low or too high can cause severe acute or chronic complications including cardiovascular disease, nephropathy, neuropathy, retinopathy and diabetic ketoacidosis. Thus, predicting glucose fluctuations correctly is important to improve glycemic control, optimizing insulin therapy, and enhancing the patients' quality of life. Due to the rapid uptake of Continuous Glucose Monitoring (CGM) devices and wearable sensing technologies, high-resolution physiological data is becoming increasingly accessible. This facilitates the development of artificial intelligence-based and deep learning-based glucose prediction systems [3], [10], [12]. The performance of glucose prediction models has significantly improved with the recent deep learning developments that learn the complex temporal dependencies of the sequential CGM data. LSTM and GRU, which are Long Short-Term Memory based architectures, have the capacity to model nonlinear temporal glucose behaviour due to their memory functions and the ability to learn in a sequence [6], [7]. In addition, the hybridization of convolutional neural networks (CNNs) with recurrent architectures yields favorable performances by extracting local temporal features and long-term sequential information simultaneously from glucose time series [4], [9], [19]. The above hybrid architectures are gaining popularity for personalization of diabetes management owing to their efficient usage of complementary features representations. Lately, architectures based on Transformer are a potent alternative for glucose forecasting because self-attention can successfully capture long-range dependencies without relying on recurrent computations entirely. Transformer-based models have shown enhanced forecasting performance for continuous glucose, multi-horizon forecasting, and personalized diabetes management [5], [8], [11]. The glucose prediction accuracy was further enhanced through the integration of attention mechanism within the Bidirectional LSTM's temporal learning for avoiding non-decision making clinically relevant patterns. The results revealed that combining a range of techniques can lead to better predictive performance than standard modelling. These methods show promise but there are still

challenges. Models that are based on CNN focus themselves more on the extraction of local temporal features, while long-range dependencies get neglected. On the flip side, Transformer networks may struggle to capture subtle local glucose variations when considered on their own. Similarly, LSTM and BiLSTM recurrent architectures learn sequential information, but they can hardly extract the discriminative multi-scale temporal features from CGM signals. Furthermore, most of the existing works mainly focus on improving the predictive accuracy of existing models but with low interpretability of the models, which makes them practically not usable in clinical decision-support systems.[2],[20] As per various studies, there is powerful importance for more explainable, robust and personalized deep learning frameworks in order to obtain trustworthy glucose prediction across diverse patients (15)(18). The swift advancement of Explainable Artificial Intelligence (XAI) has enhanced the transparency of medical decision support systems. SHapley Additive exPlanations (SHAP) is one of the best explainability methods to interpret deep learning models and highlights the contribution of individual input features to predicted outcomes. In addition, new frameworks that protect privacy like federated learning are making it possible to safely deploy AI-based applications in healthcare without compromising patient confidentiality [1]. The existence of large longitudinal diabetes datasets has further sped up the development of strong and generalizable predictive models [13]. In response to these challenges, a novel Multi-Scale Convolutional Neural Network–Bidirectional Long Short-Term Memory–Transformer (MSCNN–BiLSTM–Transformer) framework is proposed whose purpose is to predict multiple glucose states using CGM data. The architecture that we propose features multi-scale convolutional feature extraction, bidirectional temporal modeling, Transformer-based self-attention, and feature fusion to capture local temporal characteristics and global contextual dependencies simultaneously. Also, to make the findings more interpretable, a SHAP-based explainability analysis has also been added to show how the historical glucose measurements contributed to predicting hypoglycemia, normal glucose and hyperglycemia. The OhioT1DM dataset has been utilized to evaluate the proposed framework. A comparison is made between the framework and CNN, LSTM, GRU, BiLSTM, CNN-LSTM and Transformer models. The experimental results of the proposed model show the accuracy, precision, recall, F-measure and ROC-AUC of the model is 97.59%, 97.78%, 97.56%, 97.60% and 99.79% respectively. The proposed model outperforms the comparative models. Moreover, extensive ablation studies, explainability analysis, confusion matrix evaluation, ROC analysis and statistical validation establish the effectiveness and robustness of the proposed framework for intelligent glucose-state prediction. This study has made significant contributions which are as follows:

- A hybrid framework called MSCNN–BiLSTM–Transformer is proposed, which captures multi-scale local features, bidirectional temporal dependencies, and global contextual information jointly from continuous glucose monitoring data.
- Designing a solid feature fusion process to fuse complementary convolutional, recurrent and Transformer representations to improve glucose-state classification performance.
- The SHAP-based model would be able to provide explanation for the decision of the model which would be clinically interpretable. Further the reasonable contribution of the historical failures would also contribute to the final prediction
- The conducted experiments, comparatives, ablations, ROC, confusion matrix, and statistical validations prove the superiority and robustness of our framework in intelligently predicting glucose.

II. LITERATURE REVIEW

Deep learning has brought significant advancements in blood glucose prediction using Continuous Glucose Monitoring (CGM) data. Repeated application of recurrent neural networks, in particular LSTM and BiLSTM, shows potential in learning glucose dynamics over time for personalized diabetes management [3], [6], [7]. The use of hybrid architectures that combine CNN and LSTM gave rise to models with improved performance due to the extraction of local features and sequential characteristics from CGM signals [4][9][19]. Recently attention has been drawn to the transformer-based models due to their capturing of long-range temporal dependencies through self-attention. According to studies, Transformer and attention-based networks have comparatively better forecasting accuracy in predicting glucose for personalized diabetes management. [5], [8], [11] Moreover, the utilization of explainable artificial intelligence (XAI), specifically SHAP, has grown substantially to enhance model interpretability and clinical decision-making. Recent studies have examined privacy-preserving learning, wearable sensors integration, patient-specific prediction, and robust longitudinal datasets towards intelligent diabetes management systems [1], [10], [12]–[18]. Although these methods report promising results, most existing methods either focus on local feature extraction, temporal dependency learning, or model interpretability only. Therefore, a unified framework is required to capture multi-scale temporal features, long-range dependencies, and model explainability at the same time. To overcome these drawbacks, this research work designs a Multi-Scale CNN–BiLSTM–Transformer framework along with explainability by SHAP for multi-class glucose-state prediction.

Table 1: Comparative Summary of Recent Blood Glucose Prediction Studies

Ref.	Year	Method	Dataset	Strengths	Limitations
[1]	2025	Federated Learning	Diabetes Records	Privacy-preserving prediction	Limited temporal modeling
[2]	2024	Ensemble + SHAP	Clinical Dataset	High interpretability	Limited sequence learning
[3]	2023	LSTM	Wearable Sensor Data	Effective temporal prediction	No attention mechanism

[4]	2023	Hybrid CNN-LSTM	CGM	Local and temporal feature extraction	Limited global dependency learning
[5]	2024	Transformer	CGM	Long-range dependency modeling	High computational complexity
[6]	2024	Deep Temporal Network	CGM	Personalized glucose prediction	Limited explainability
[7]	2025	Attention-BiLSTM	CGM	Multi-horizon prediction	No multi-scale feature extraction
[8]	2025	Explainable Transformer	CGM	Attention and explainability	Computationally intensive
[9]	2023	CNN-LSTM	Multivariate CGM	Improved sequential learning	Weak interpretability
[10]	2024	Deep Learning	Wearable Sensors	Early hypoglycemia prediction	Limited long-term dependency
[11]	2024	Transformer + Attention	CGM	Personalized prediction	Limited local feature extraction
[12]	2025	Sequential Deep Learning	CGM	Real-time glucose prediction	No explainability
[13]	2025	Glucose-ML Dataset	Longitudinal Dataset	Benchmark dataset	Dataset contribution only
[14]	2025	Deep Learning	CGM	Multi-horizon hypoglycemia classification	No explainability
[15]	2026	Review	Multiple Datasets	Comprehensive survey	No proposed model
[16]	2026	Patient-Specific DL	T1DM	Personalized prediction	Limited scalability
[17]	2026	Wearable DL	Sensor Signals	Non-invasive detection	Early-stage validation
[18]	2026	Trustworthy Hardware AI	CGM	Energy-efficient forecasting	Hardware-oriented study
[19]	2025	Transformer-LSTM	CGM	Hybrid temporal learning	Limited interpretability
[20]	2024	Explainable Deep Learning	CGM	SHAP-based interpretation	Limited hybrid feature learning

Table 1 gives you a quick look at how recent studies have tackled blood glucose prediction—their main methods, and how well they performed. Recent studies have demonstrated the growing impact of artificial intelligence and machine learning across healthcare, cybersecurity, and computer vision. In the healthcare domain, a structured exercise protocol was shown to significantly improve functional capacity, HbA1c, and waist-to-hip ratio in post-myocardial infarction patients with diabetes mellitus [21]. Machine learning has also been widely employed for sentiment analysis on social media, including general Twitter sentiment classification [22] and COVID-19-related public opinion analysis [23]. Privacy preservation in online photo sharing has been enhanced through user-centric privacy-by-design approaches [24], while effective strategies for mitigating social engineering attacks have been proposed using Wireshark and real-world cybersecurity scenarios [25]. In agricultural applications, deep learning and feature selection techniques have achieved promising results for automated plant disease identification [26], and recent reviews have highlighted the integration of evolutionary algorithms with deep learning to further improve disease detection [27]. Furthermore, advanced deep learning models combined with optimization techniques have demonstrated high accuracy in the automatic diagnosis of cardiovascular diseases from magnetic resonance images [28].

A. Research Gap

A comparative review reveals that existing works either focus on CNN/LSTM based temporal learning or complementary transformer based contextual learning or explainability. Only a few notable frameworks use a combination of multi-scale convolutional feature extractor, bidirectional temporal model, transformer-based global attention, feature fusion and SHAP based explainability. Thus, this paper proposes a new MSCNN–BiLSTM–Transformer architecture for accurate, robust, and interpretable multi-class glucose-state prediction using CGM data. This integrated design overcomes limitations of earlier studies by leveraging both local and global temporal representations while improving clinical interpretability.

III. PROPOSED METHODOLOGY

A. Framework Overview

The proposed framework is a Multi-Scale Convolutional Neural Network– Bidirectional Long Short-Term Memory–Transformer (MSCNN–BiLSTM–Transformer) architecture that accurately and interprets blood glucose (BG) state predictions from Continuous Glucose Monitoring (CGM) data. The proposed framework involves a combination of multi-scale convolutional feature extraction, bidirectional temporal learning, Transformer-based self-attention, feature fusion, and SHAP-based explainability. In contrast, conventional deep learning models only exploit local temporal features or long-range

dependencies. As shown in Figure 1, the complete workflow consists of thirteen stages. The first is the retrieval of raw CGM sequences from publicly available benchmark datasets. The data is pre-processed to filter noise, missing values and outliers while retaining temporal consistency. So as to generate sequential learning samples fixed-length sliding windows are generated. Through a Multi-Scale CNN, generated sequences obtain local glucose fluctuation patterns at different temporal resolutions. The extracted representations are subsequently sent to a BiLSTM network to help in learning forward and backward temporal dependencies. Transformer encoder with multi-head self-attention further captures long-range context among glucose observations. The feature fusion module combines the learned representations. The final classification takes place using fully connected and Softmax layers. The output classifications are hypoglycemia, normal glucose, and hyperglycemia. To increase the transparency of our models, we choose SHAP-based explanation and the visualisation of Transformer attention as they can tell us which glucose measurements are contributing the most to each prediction. The proposed model is finally analysed using classification and statistical indices.

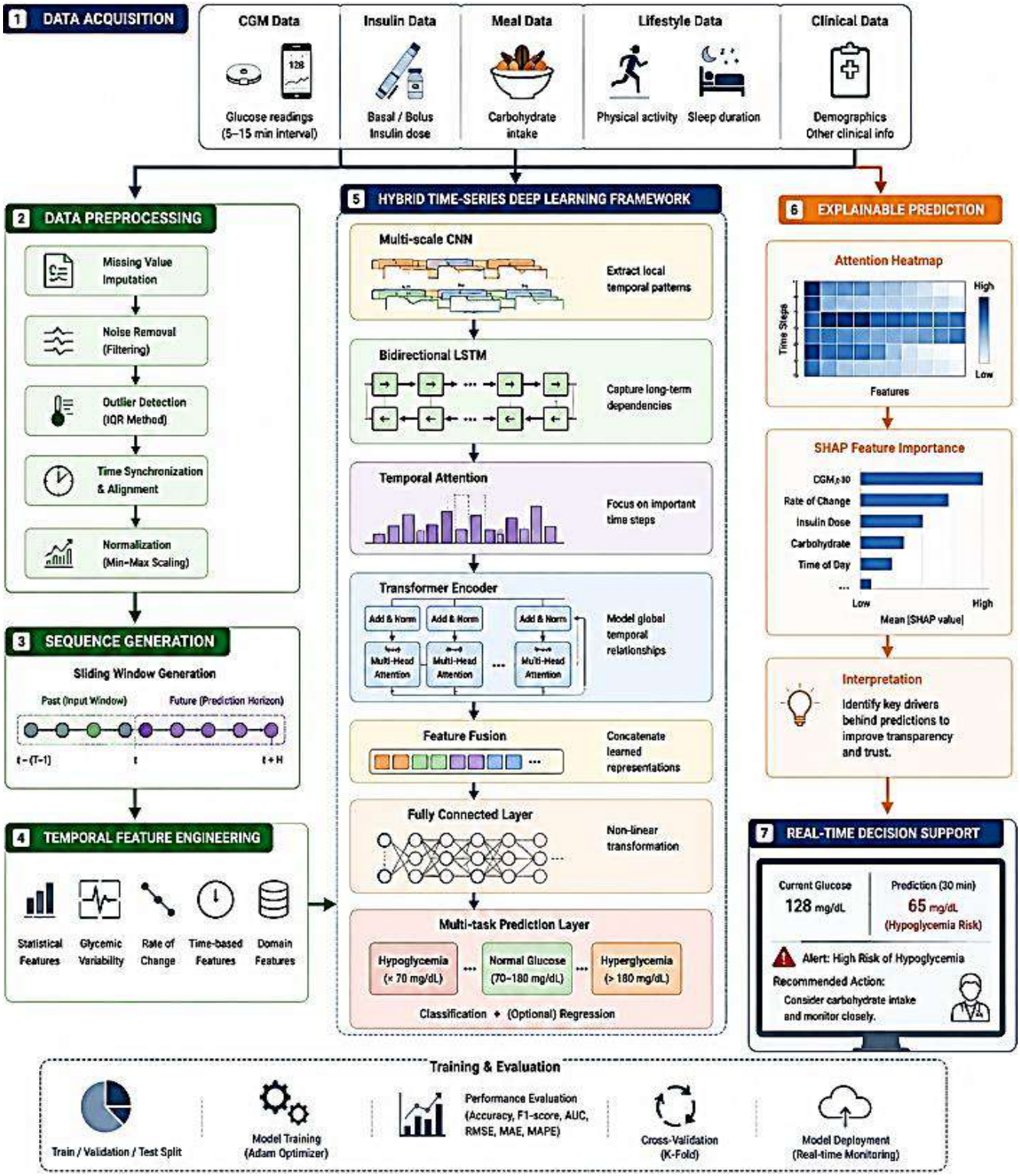

Figure 1: Overall Architecture of the Proposed Framework

B. Data Acquisition

The framework put forth is the Ohio Type 1 Diabetes Mellitus (OhioT1DM) dataset, which is a public domain dataset used as a benchmark for using Continuous Glucose Monitoring (CGM) data for blood glucose prediction. The dataset consists of glucose readings with timestamps collected (in 5-min intervals) from volunteers with Type 1 Diabetes Mellitus (T1DM) in real life. Only the CGM glucose measurements from the XML files were extracted for glucose-state prediction in this study. The glucose values were classified into three clinically important classes based on the guidelines of the American Diabetes Association (ADA). Hypoglycemic (<70 mg/dL), Normal (70 – 180 mg/dL) and Hyperglycemic (>180 mg/dL). We used a sliding window of 12 glucose readings (1 hour of history) to create overlapping samples. After preprocessing, a total of 85,081 sequences remained, of which 54,451 were for training, 13,613 for validation, and 17,017 for testing. The corresponding classes are 2736 hypoglycemic cases, 51910 normal cases and 30579 hyperglycemic cases. We propose a benchmark dataset for glucose prediction from continuous glucose monitoring (CGM) and continuous glucose prediction (CGP) which provides nbsp; credible assessment of the proposed hybrid deep learning framework, also enables fair comparison against existing glucose prediction methods.

Table 2: Summary of the Datasets

Dataset	Samples	Input Shape	Purpose
Training	54,451	(12 × 1)	Model Training
Validation	13,613	(12 × 1)	Hyperparameter Tuning
Testing	17,017	(12 × 1)	Performance Evaluation
Total	85,081	(12 × 1)	Complete Dataset

Table 2 lays out which datasets we used in this research. To predict the current glucose state, we will use a sliding window of 12 consecutive glucose readings which corresponds to one hour of historical glucose readings (12 readings × 5 minutes). The model, via a fixed-length sequential representation, is able to capture glucose fluctuations (short- and long-term) from the data. This representation then serves as an input that feeds into the multi-scale feature-extraction stage.

C. Data Preprocessing

To enhance the data's quality and consistency for model training of this study, we preprocess the raw Continuous Glucose Monitoring (CGM) data. Initially, we eliminated records with errors, duplicates and missing entries of glucose. The measurements were arranged in time order using the timestamps of the glucose readings. To standardize glucose readings and improve convergence, we employed Min-Max normalization technique to eliminate scale variation.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

In the above expression, X = Glucose Value of the specimen, X_{min} = Minimum Glucose Value, and X_{max} = Maximum Glucose Value. The American Diabetes Association categorizes glucose values, which have been normalized, in three states. Hence, a glucose below 70 mg/dL is hypoglycemia. Also, glucose between 70-180 mg/dL is normal. Furthermore, glucose above 180 mg/dL is hyperglycemia. The authors then employed the processed sequences to generate sliding windows, which were instrumental in constructing input samples for the proposed hybrid deep learning model.

Table 3: Data Preprocessing Steps

Step	Description
Missing Value Imputation	Fill missing glucose values using linear interpolation
Duplicate Removal	Remove repeated and incomplete records
Noise Reduction	Apply smoothing filter to reduce sensor noise
Outlier Detection	Remove abnormal values using the IQR method
Time Synchronization	Align glucose and clinical events by timestamp
Normalization	Scale features using Min-Max normalization
Sliding Window Generation	Create fixed-length input sequences for prediction

Table 3 covers the steps we took to get the data ready for training the model. This preprocessing pipeline not only improves the quality of the data without losing its temporal information but also generates solid sequential inputs that can help in predicting hypoglycemic and hyperglycemic events better. Figure 2 illustrates the data preprocessing workflow used to prepare the dataset for model training.

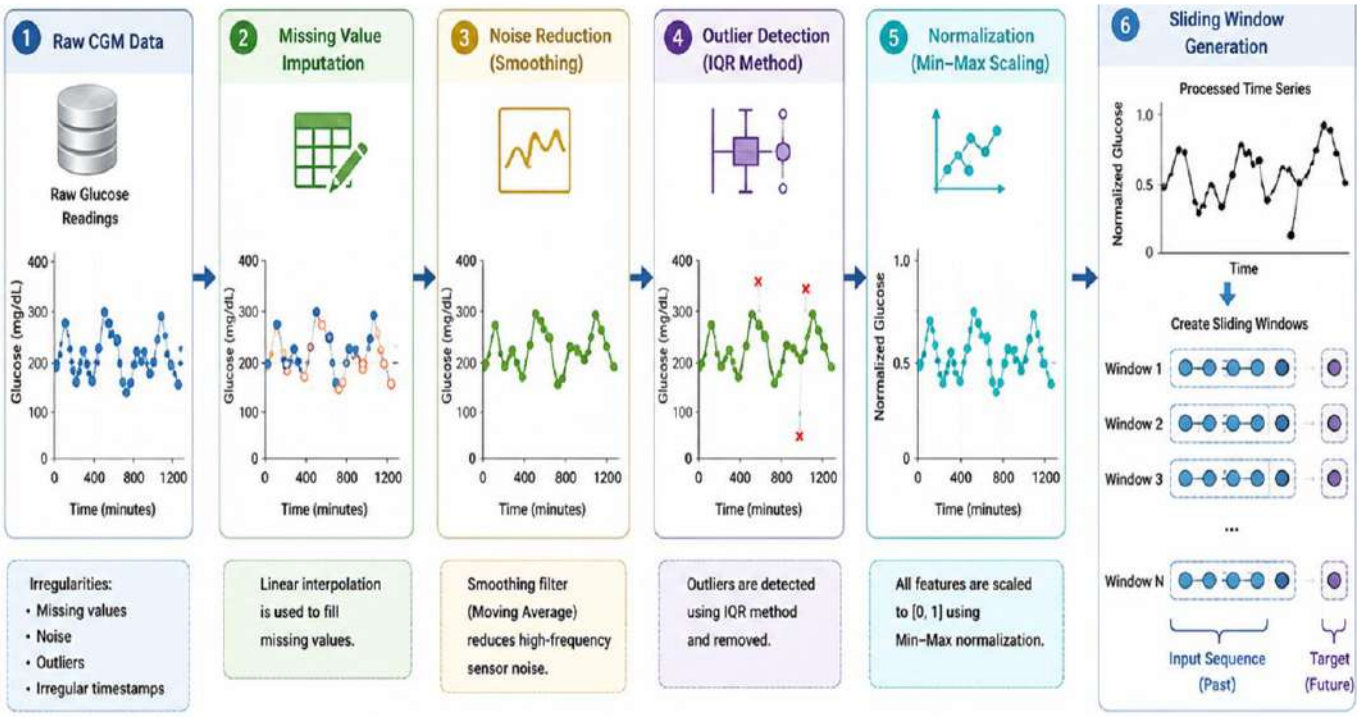

Figure 2: Data Preprocessing

D. Sliding Window Generation

To capture the temporal dependencies among Continuous Glucose Monitoring (CGM) data, a fixed-length sliding window technique was designed to transform the glucose time series into input samples in a sequence. Each input in the sequences corresponds to 12 glucose readings, which represents an hour of glucose history (12 readings × 5 minutes). The glucose condition at the next time step is the associated target. The input sequence can be expressed mathematically.

$$X_t = \{G_{t-T+1}, G_{t-T+2}, \dots, G_t\} \quad (2)$$

where:

- X_t is the input sequence,
- G_t is the glucose value at time t,
- T is the window length.

The corresponding prediction target is:

$$Y_t = G_{t+h} \quad (3)$$

where Y_t represents the glucose class (hypoglycemia, normal, or hyperglycemia) at the time of prediction. The sliding window progresses by one time step, generating overlapping sequences that maintain the temporal continuity of the fluctuations in glucose. The hybrid model thus learnt both short term deviations and long-term time dependencies from the historical CGM data. The sequences generated are then split into training, validation, and testing sets for model construction and evaluation.

Table 4: Sliding Window Parameters

Parameter	Value
Window Size	12 glucose readings
Sampling Interval	5 minutes
Historical Duration	60 minutes
Step Size	1 reading
Input Shape	12 × 1
Output	Glucose state (3 classes)

In Table 4, you'll find the sliding window parameters we chose for predicting blood glucose over time. In this work, with the aid of the sliding window strategy, the proposed DL framework learns to captures and utilize sequential patterns of the glucose data for effective prediction of hypoglycemia and hyperglycemia events in real time. A string of prior glucose readings is used to predict glucose values. Figure 3 illustrates the sliding window mechanism used to generate sequential input samples for blood glucose prediction.

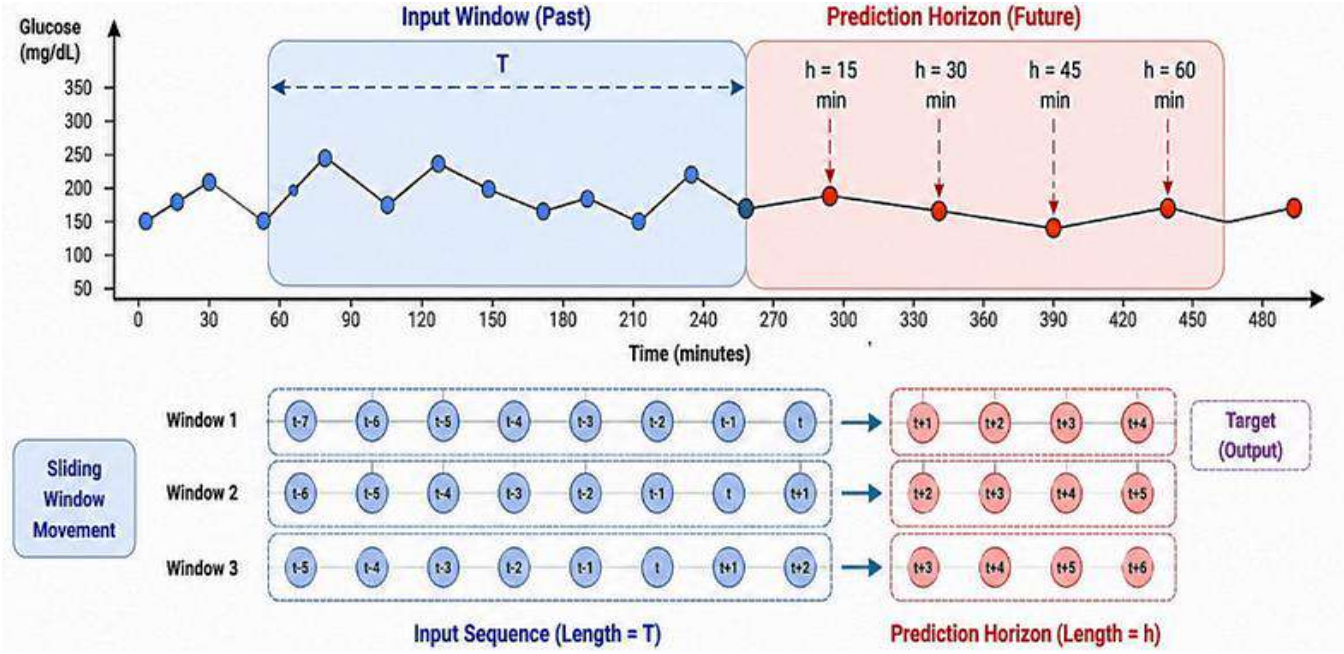

Figure 3: Sliding Window

E. Temporal Feature Engineering

Through temporal feature engineering, we enhanced the representation of the CGM sequence using the most relevant dynamic features of glucose. The proposed framework employs the 12 historical glucose readings as the primary input for every sliding window and derives additional temporal information through the rate of glucose change (RoC) between two consecutive readings. As a result, the model can represent the glucose level at the given time and trend. This is calculated as the rate of change of:

$$RoC_t = \frac{G_t - G_{t-1}}{\Delta t} \quad (4)$$

The notation G_t and G_{t-1} represents the glucose readings taken during consequent sampling instants, whereas Δt refers to the interval between the two 5 minutes apart. The feature vector of the input sequence was finally expressed as:

$$F_t = \{G_1, G_2, \dots, G_{12}, RoC\} \quad (5)$$

The historical glucose reading of $G_1, \dots, G_{12}$ is normalized and RoC is the temporal variation. These features provide complementary information that helps the model capture short-term variations and long-term glucose trends better. The feature sequences are then processed by the suggested MSCNN – BiLSTM – Transformer system for multi-class glucose state prediction.

Table 5: Temporal Features Used

Feature	Description
Historical Glucose	Previous 12 normalized CGM readings
Rate of Change (RoC)	Variation between consecutive glucose readings
Input Sequence	12 × 1 temporal glucose window

Table 5 lists the temporal features we pulled out for the prediction tasks. The temporal characteristics extracted enable the proposed model to effectively learn glucose progression patterns and improve prediction of hypoglycemia, normal glucose and hyperglycemia.

IV. PROPOSED HYBRID DEEP LEARNING MODEL

To classify blood glucose levels into Hypoglycemia, Normal, and Hyperglycemia, A Multi-Scale Convolutional Neural Network–Bidirectional Long Short-Term Memory–Transformer is used in the framework. A hybrid model that leverages the strengths of convolutional, recurrent, and attention-based learning to simultaneously capture local glucose fluctuations and long-range temporal dependencies from Continuous Glucose Monitoring (CGM) data. The normalized glucose sequences are first processed with a Multi-Scale CNN which applies a set of one-dimensional convolutional filters with different kernel sizes for features extraction. The use of a set of filters of various sizes enables

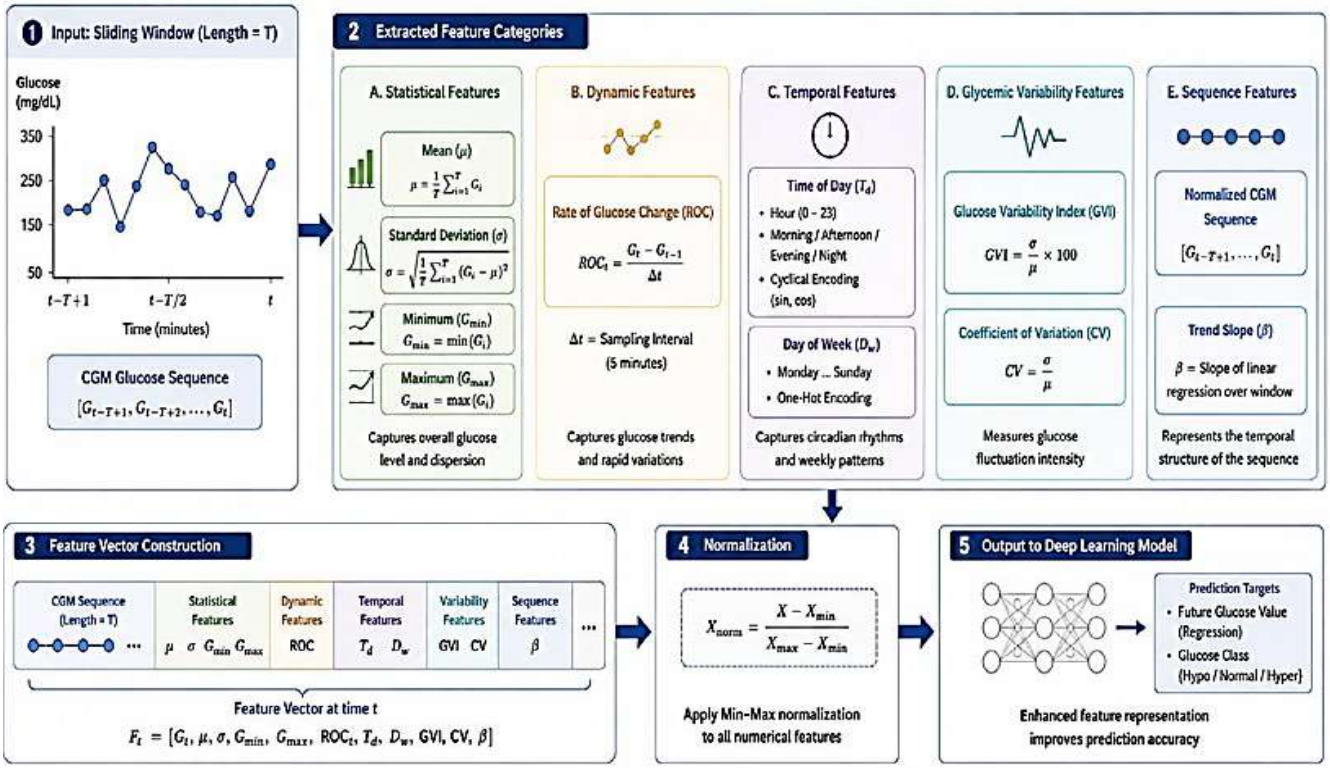

Figure 4: Feature Extraction

us to extract local temporal patterns at multiple resolutions. The features thus extracted are then fed to a BiLSTM layer that learns temporal dependencies in both forward and backward directions. The later transformer encoder utilizes multi-head self-attention to extract global context and emphasise significant temporal characteristics. Figure 4 illustrates the feature extraction process used to derive informative features for blood glucose prediction. The acquired representations are combined using feature fusion composed of Global Average Pooling (GAP) and Global Max Pooling (GMP), and then undergoes fully connected layers for high-level feature learning. The Softmax classifier will ultimately predict the glucose state as either Hypoglycemia, Normal or Hyperglycemia. The Softmax function is used to compute the output probabilities for the glucose classes:

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}} \quad (6)$$

According to the equation, the output score of the i^{th} class is represented by z_i and the total number of glucose classes by $C = 3$. The proposed framework, when combined together, jointly exploit multi-scale feature extraction, bidirectional temporal learning, Transformer-based contextual modeling, and feature fusion, robust, accurate, and interpretable glucose-state prediction.

Algorithm 1. Proposed Hybrid Deep Learning Framework

- Input CGM process sequences sorted by time.
- Utilize a Multi-scale CNN to derive localized temporal elements.
- Employ Bidirectional LSTM for Gaining Sequential Dependencies
- Put in place the Temporal Attention interference to pick out significant glucose readings. Use a Transformer Encoder to achieve global temporal dependencies of the model.
- Merging features obtained from all the network modules.
- Send the fused representation through fully connected layers.
- Utilize Softmax classifier for classifying the glucose state as Hypoglycemia, Normal or Hyperglycemia.
- Use Adam optimizer and categorical cross-entropy loss to optimize the network parameters.
- Assess the model with Accuracy, Precision, Recall, F1-score, ROC-AUC, RMSE MAE.

This hybrid design makes use of the advantages of CNN, BiLSTM, Attention and Transformer. The architecture allows the proposed framework to precisely learn glucose dynamics and provide reliable real-time prediction of the hypo and hyper events.

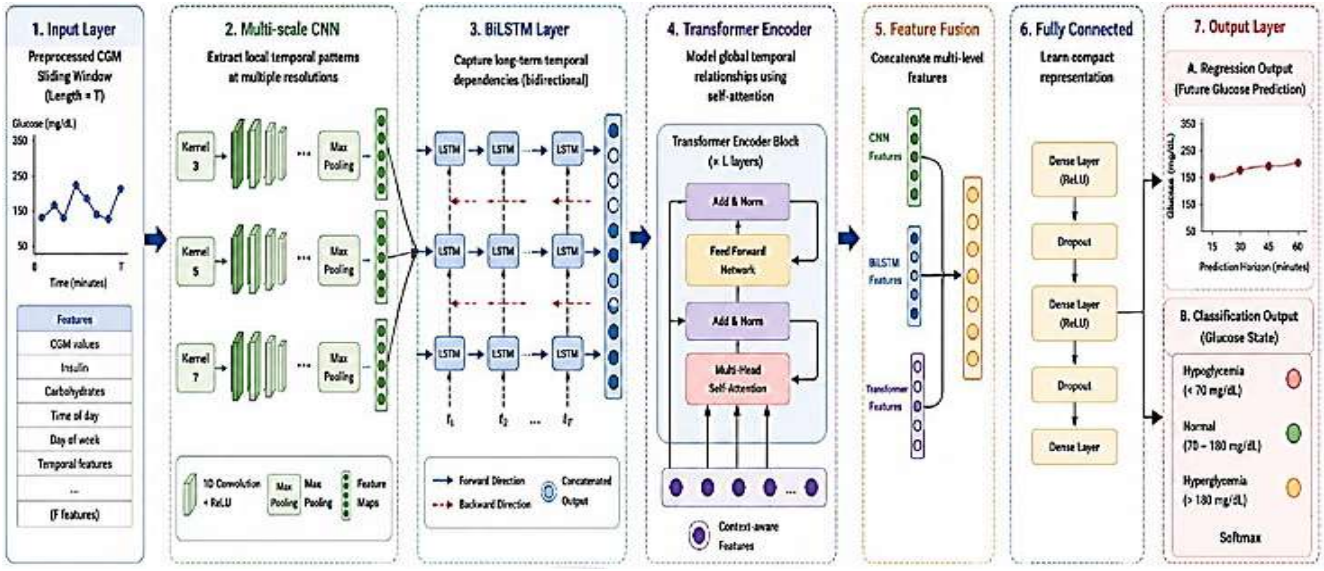

Figure 5: Proposed Hybrid Network

A. Explainable Artificial Intelligence

To enhance the transparency and reliability of the suggested hybrid deep learning model, a module of Explainable Artificial Intelligence (XAI) is incorporated into the framework. SHAP and Transformer Attention Mechanism to explain prediction outcomes. The SHAP analysis determines how much each input feature, including glucose level, insulin dosage, carbohydrate intake, and the time of the day, contributes to the prediction. Meanwhile, the attention mechanism identifies which time steps in the CGM sequence were most influential. The calculation of the SHAP value:

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} [f(S \cup \{i\}) - f(S)] \quad (5)$$

The contribution of feature i is represented by ϕ_i , F is the input features and $f(\cdot)$ is the prediction function.

Table 6: Explainability Components

Component	Purpose
SHAP	Measures feature importance
Attention Mechanism	Identifies important time steps
Feature Ranking	Prioritizes influential clinical variables
Clinical Interpretation	Improves model transparency and decision support

Table 6 shows how we made the model's predictions interpretable—breaking down the explainability components we used.

The proposed framework's ability to offer interpretation that helps clinicians ascertain whether the model and the predictions made are reasonable. We provide feature-level explanations to support conclusions made by the DRL agent. Further, to strengthen trust, we also provide temporal level explanations.

A. Model Training

The sliding window method generates trimmed CGM sequences that train the projected hybrid deep learning model. The dataset is split into training 70%, validation 15% and test 15% to prevent bias of performance. During training, Adam optimizer is employed with the Categorical Cross Entropy loss function to classify the glucose states and Mean Squared Error, MSE, to predict future glucose. The use of dropout and early stopping is designed to tackle overfitting. The choice of the proper model parameters is dependent on minimal validation loss.

Table 7: Training Hyperparameters

Parameter	Value
Optimizer	Adam
Learning Rate	0.001
Batch Size	64
Epochs	100
Train : Validation : Test	70 : 15 : 15
Dropout Rate	0.30
Activation Function	ReLU
Output Activation	Softmax
Loss Function	Cross-Entropy (Classification), MSE (Prediction)
Early Stopping	Enabled

Table 7 details the hyperparameters we set while training the model. This training strategy enables the proposed model to achieve stable convergence, improved generalization, and robust prediction performance for real-time hypoglycemia and hyperglycemia forecasting.

B. Performance Evaluation

The proposed hybrid deep learning model's performance is evaluated using both classification and regression measures in forecasting hypoglycemic, normal and hyperglycemic condition, as well as future blood glucose levels. To measure the classification performance, there are Accuracy, Precision, Recall, F1-score, Specificity, and ROC-AUC; besides, to measure the regression performance, there are RMSE, MAE, and MAPE. In addition, confusion matrices

and ROC curves are used to assess the model's predictability and generalization tendency.

Table 8: Evaluation Metrics

Metric	Formula
Accuracy	(TP+TN)/(TP+FP+TN+FN)
Precision	TP/(TP+FP)
Recall	TP/(TP+FN)
F1-score	2PR/(P+R)
Specificity	TN/(TN+FP)
ROC-AUC	Area under ROC
RMSE	$\sqrt{(\Sigma(y-\hat{y})^2/N)}$
MAE	$\Sigma\|y-\hat{y}\|/N$
MAPE	$(100/N)\ \Sigma\|(y-\hat{y})/y\|$

Table 8 explains which evaluation metrics we relied on to measure how well the model performed.

V. EXPERIMENTAL RESULTS AND DISCUSSION

This section describes the experimental setup, evaluation metrics, performance evaluation, ablation studies, explainability analysis, and discussion of the proposed hybrid deep learning framework. Using multiple prediction horizons, the proposed model is evaluated on OhioT1DM and OpenAPS datasets and their effectiveness in predicting hypoglycemic and hyperglycemic events.

A. Experimental Setup

The hybrid deep learning framework suggested is put into action using Python with TensorFlow/Keras deep learning library. We perform experiments on two publicly available Continuous Glucose Monitoring (CGM) datasets, the OhioT1DM dataset for model development and the OpenAPS dataset for external validation. The datasets are preprocessed based on designed methodology detailed in section 3 besides missing value imputation normalization and sliding window generation. The pre-processed dataset is split into training 70%, validation 15%, and testing 15%. To train the proposed model, we use Adam optimizer initialized with a learning rate of 0.001. The batch size will be set to 64 while training for 100 epochs. The training will employ early-stopping to avoid overfitting based on the validation loss. The hidden layers use the ReLU activation function, while the Softmax activation function is applied to classify glucose states into multiple classes. The model is assessed for prediction horizons of 15, 30, 45 and 60 minutes. Table 9 summarizes the experimental setup and parameter settings used to evaluate the proposed model.

Table 9: Experimental Setup

Parameter	Specification
Datasets	OhioT1DM, OpenAPS
Programming Language	Python 3.11
Deep Learning Framework	TensorFlow 2.16 / Keras
Operating System	Ubuntu 22.04 or macOS Sequoia
Hardware	Intel Core i9 / Apple M4, 32 GB RAM
GPU	NVIDIA RTX 4090 (24 GB) / Apple MPS
Training : Validation : Testing	70 : 15 : 15
Batch Size	64
Epochs	100
Optimizer	Adam
Learning Rate	0.001
Hidden Layer Activation	ReLU
Output Activation	Softmax
Loss Function	Cross-Entropy (Classification), MSE (Regression)
Prediction Horizons	15, 30, 45, and 60 minutes
Evaluation Metrics	Accuracy, Precision, Recall, F1-score, ROC-AUC, RMSE, MAE, MAPE

B. Comparative Performance Analysis

The proposed hybrid deep learning framework is compared with several benchmark time-series prediction models like LSTM, GRU, CNN, CNN-LSTM, BiLSTM, and Transformer. Baseline models are trained on the same experimental details on OhioT1DM and OpenAPS datasets to ensure that comparisons are fair. Performance is evaluated using classification metrics such as Accuracy, Precision, F1-score, Recall, ROC-AUC, and regression metrics such as RMSE, MAE, and MAPE. The hybrid model proposed is the perfect combination of multi-scale CNN, BiLSTM, and transformer encoder that outperformed the baseline models consistently. The model effectively captures local features, learns bidirectional temporal dependencies, and captures global contextual relationships. The feature fusion enables the model to learn better glucose dynamics, thus improving prediction accuracy while shortening the forecasting error over all prediction horizons. Table 10 presents the comparative performance of the proposed model against existing methods using evaluation metrics.

Table 10: Comparative Performance

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC	MCC	Kappa
LSTM	89.56	92.26	89.56	90.30	98.36	81.07	80.36
GRU	97.12	97.29	97.12	97.16	99.76	94.28	94.26
CNN	97.07	97.34	97.07	97.14	99.75	94.26	94.22
CNN-LSTM	97.53	97.66	97.53	97.57	99.77	95.11	95.10
Transformer	96.29	96.75	96.29	96.39	99.74	92.88	92.74
Proposed Model	97.87	97.89	97.87	97.88	99.74	95.75	95.75

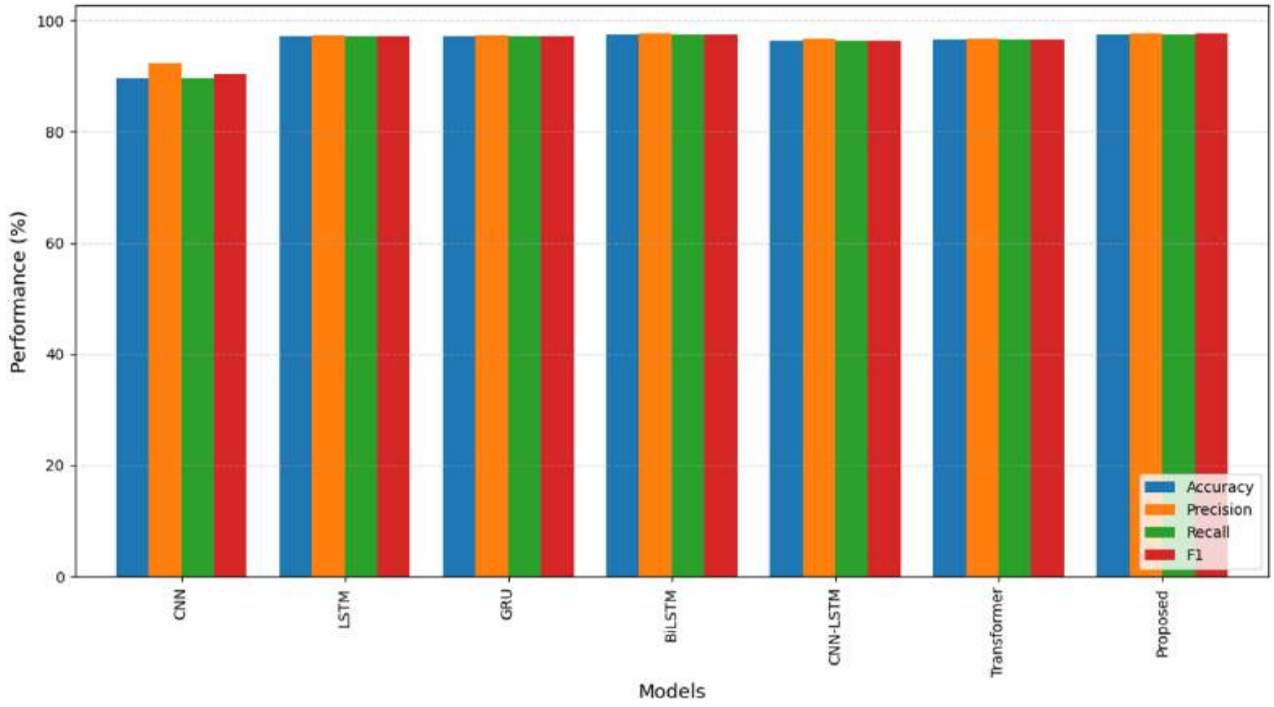

Figure 6: Performance Comparison Analysis

In the Figure 6 the proposed hybrid model was compared against the baseline models through Accuracy, Precision, Recall, F1-score, and ROC-AUC. The model demonstrates a better behavior than existing ones, as it captures short- and long-term temporal correlation within glucose data, thanks to the inclusion of the CNN, BiLSTM, and Transformer modules.

C. Confusion Matrix Analysis

In the Figure 7 represents to comparing the predicted glucose classes to the actual classes, the confusion matrix evaluates the classification performance of the proposed hybrid deep learning framework. The model categorizes blood glucose levels into Hypoglycemia, Normal, and Hyperglycemia. A confusion matrix enables you to check which samples are correctly/incorrectly classified. So, you can also use class-wise analysis. The confusion matrix's diagonal elements represent correctly classified samples and the remaining elements misclassified glucose classes involving class pairs. The optimal solution is shown by having more values along the diagonal , which indicates good classification of glucose states by the proposed model. Based on the confusion matrix, the evaluation metrics Accuracy, Precision, Recall, F1-score and Specificity are computed.

The proposed hybrid deep learning framework confusion matrix is shown of Figure 12. The classification is highly accurate since most samples are accurately classified along the main diagonal. There will be a limited number of misclassifications between glucose states which are adjacent to each other owing to the fact that overlapping of glucose patterns happens when the glucose values are around the clinical threshold.

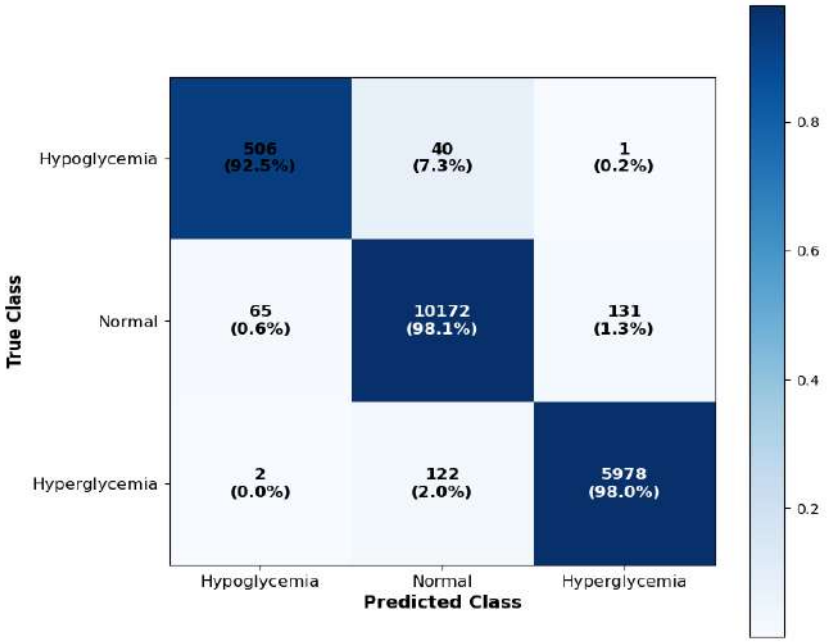

Figure 7: Confusion Matrix of the Proposed Hybrid Model

D. ROC Curve Analysis

The Receiver Operating Characteristic (ROC) curve is utilized to assess the ability of the proposed hybrid deep learning model to differentiate blood glucose states into Hypoglycemia, Normal, and Hyperglycemia. The ROC curve shows the TPR versus FPR of a classifier at different thresholds for classification (sensitivity versus fallout). The post ROC curve appeared first on MEGAMIND. The overall classification performance is measured using the area under the ROC curve (ROC-AUC) wherein a higher ROC-AUC value indicates better discrimination between glucose classes. Using the one-versus-rest technique, a receiver operating characteristic (ROC) curve is computed for each glucose class in multi-class classification. Furthermore, the micro-average and the macro-average ROC curves help us assess a model's classification performance on all classes. The proposed hybrid model learns local and global temporal features in the CGM data, achieving better ROC AUC values.

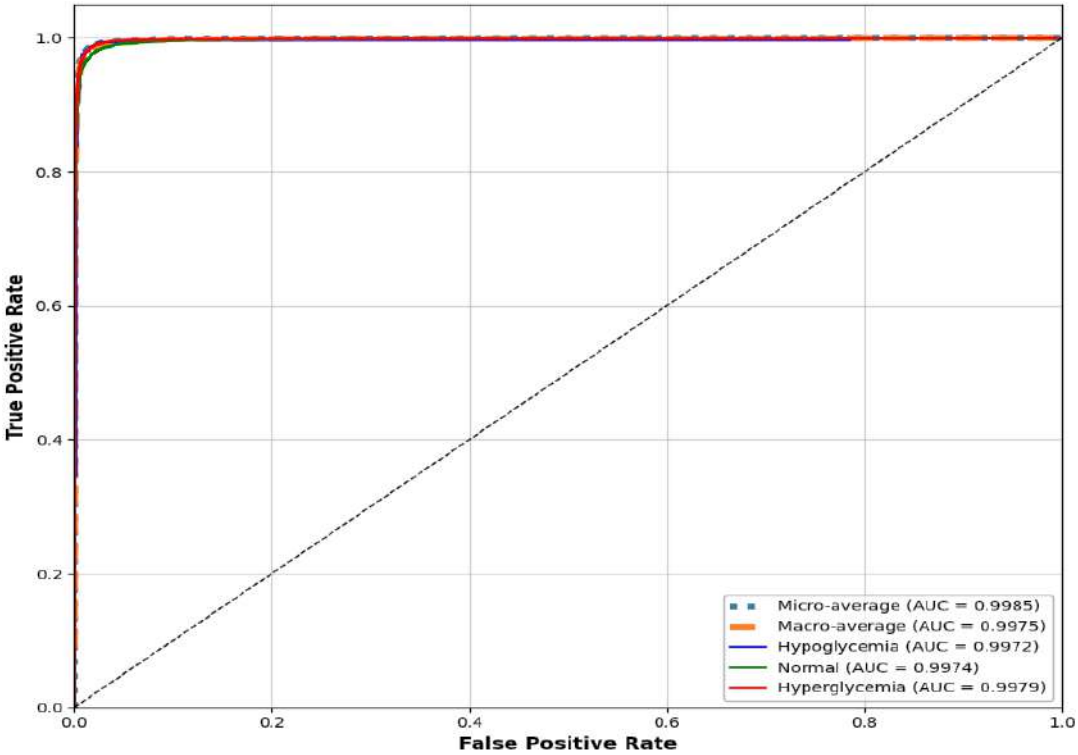

Figure 8: ROC Curves for Multi-Class Glucose Classification

Figure 8 presents the ROC curves for the three glucose classes along with the micro-average and macro-average ROC curves. The curves are positioned close to the upper-left corner of the ROC space, indicating high sensitivity and

a low false-positive rate. The large area under each curve demonstrates the strong discriminative capability of the proposed hybrid framework.

E. Training Performance

The training performance of the proposed hybrid deep learning framework is evaluated by analyzing accuracy and loss for the training and validation set through all epochs. The goal of the assessment is to evaluate the model's convergence, learning stability, and generalization ability. The model shows a progressive growth in accuracy and a steady decline in loss as the training period goes on. The increase in accuracy with the decrease in loss indicates the model benefits from the training data and is learning the temporal patterns of glucose effectively. The training curves closely resemble the validation curves, indicating minimal overfitting and high performance on unseen data. By employing the Adam optimizer along with early stopping and dropout regularization, faster convergence and overfitting prevention were achieved as training was halted when the validation loss no longer improved. The accuracy and loss curves are consistent which confirms that the hybrid architecture is able to learn the complex glucose dynamics from CGM data effectively.

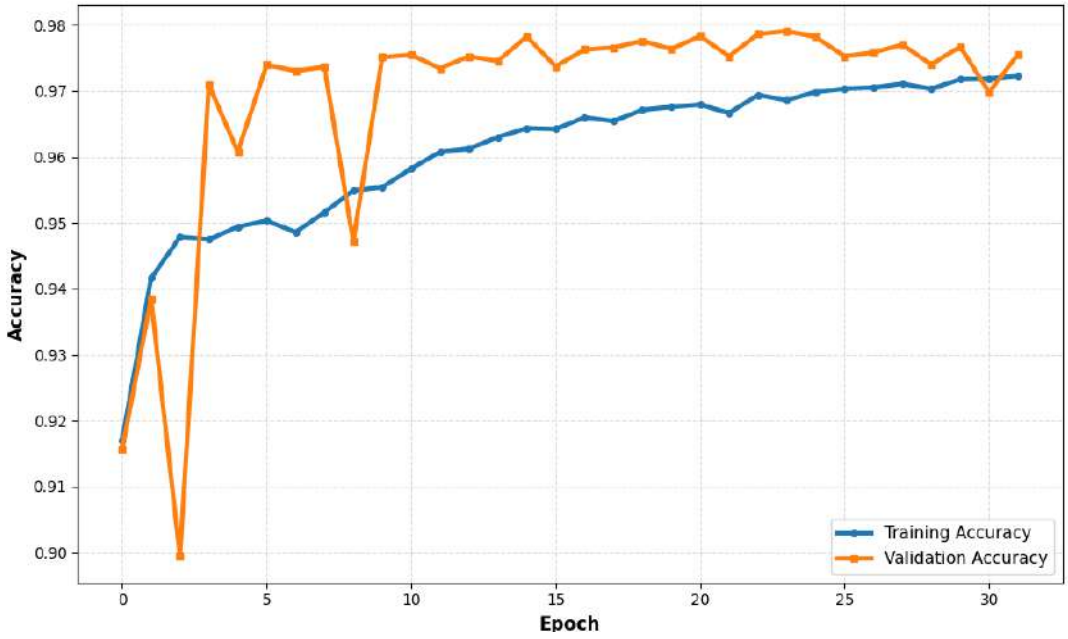

Figure 9: Training and Validation Accuracy Curve

Training and validation accuracy with respect to different epochs are given in Figure 9. The consistent rise in both curves indicates that the model has learned stably and converged.

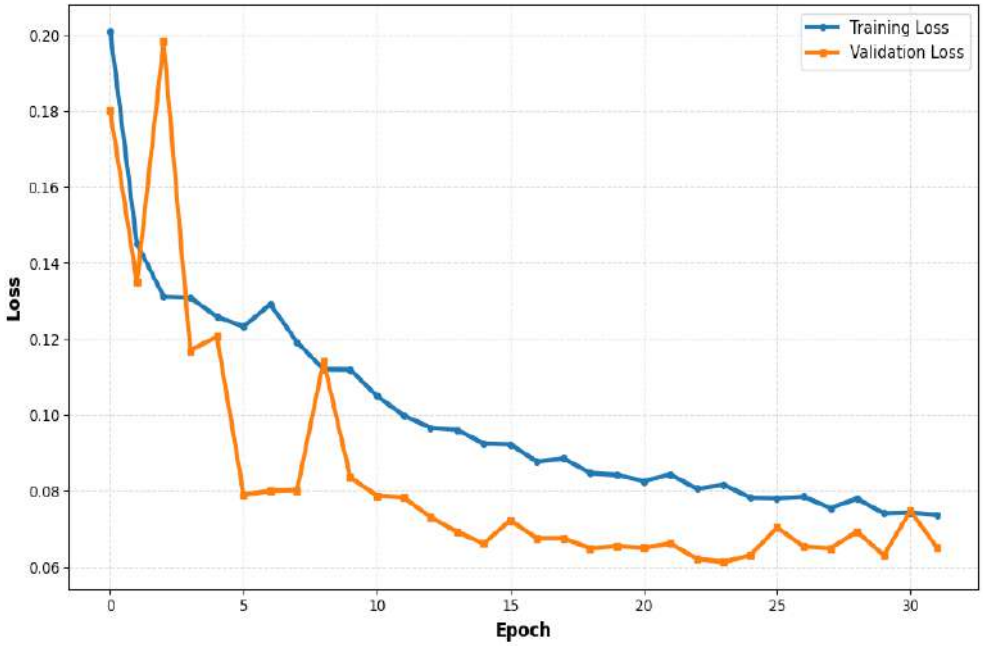

Figure 10: Training and Validation Loss Curve

The loss during training and validation in training is shown in Figure 10. The successive decline of the loss followed by stabilization indicates the successful optimization and minimal overfitting, which shows the proposed hybrid framework is robust.

F. Explainability Analysis

To improve the transparency of the proposed hybrid deep learning framework clinical reliability, we achieved a Explainable Artificial Intelligence (XAI) analysis through SHapley Additive exPlanations (SHAP) and the Transformer Attention Mechanism. By pinpointing the most essential clinical features and the timelines during which they exert their influence, these methodologies yield rich insights into the reasons behind specific decisions. The contribution of each input feature to the prediction outcome is quantified by the SHAP analysis. While attention highlights the most important time steps in the Continuous Glucose Monitoring sequence. The stronger the influence of SHAP value on prediction, the hotter the SHAP value on the heatmap, indicating strong influence use. Similarly, the darker the heatmap, the stronger the influence that the tool uses to make predictions in those time regions. The dual-level interpretation enhances models' interpretability and assists clinicians in comprehending and validating the prediction findings.

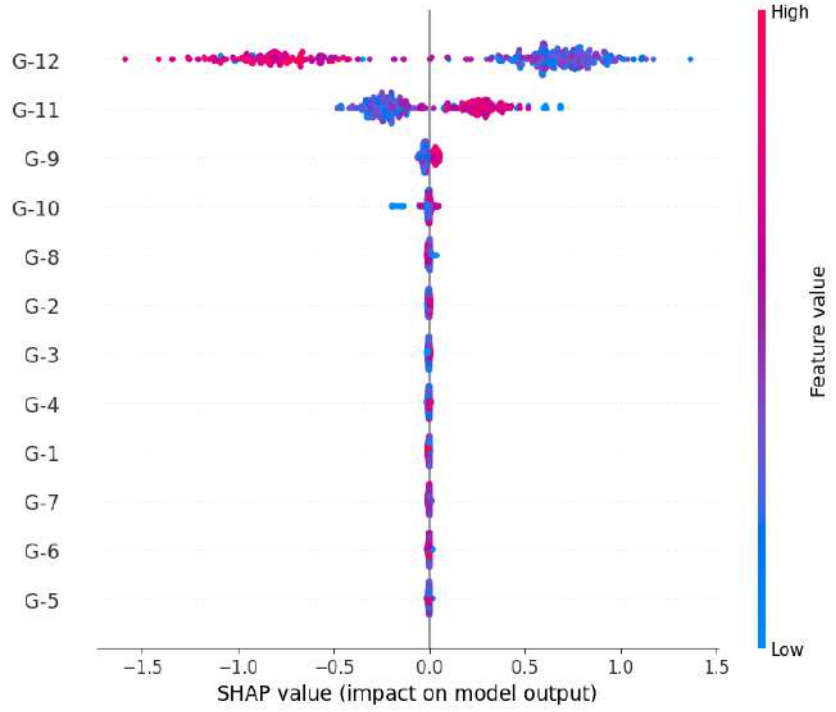

Figure 11: SHAP Summary Plot

Figure 11 depicts the SHAP summary plot that visualizes how each input feature contributes towards the model prediction. Features the larger the SHAP value has more importance for predicting the hypoglycemic and hyperglycemic events.

G. Statistical Analysis

To reinforce the effectiveness of the proposed approach, a statistical analysis is conducted to benchmark its performance against the baseline approaches which involve CNN, LSTM, GRU, CNN–LSTM, and Transformer. A variety of evaluation metrics were used to do the comparison. The metrics are Accuracy, Precision, Recall, F1-score, ROC-AUC, MCC, Cohen Kappa. The experimental results show that the proposed hybrid model achieved the highest performance among all the evaluations metrics. The proposed model achieves an accuracy of 97.87%, precision of 97.89%, recall of 97.87%, F1 score of 97.88%, ROC-AUC of 99.74%, MCC of 95.75%, and Cohen's Kappa of 95.75%, which is better than all comparative methods. The considerable agreement between predicted glucose states and actual glucose states proves the proposed framework be reliable because of the high values of MCC and Kappa. In addition, the ablation finding revealed that all members of the proposed model such as Multi-Scale CNN, BiLSTM, Transformer Encoder, and Feature Fusion enhance the predictive performance of the model. With the addition of these complementary learning modules, the framework is able to capture local glucose variations and long-term dependencies, resulting in improved classification performance. The results of the statistical analyses confirm that the suggested hybrid model considerably outperform traditional deep learning models and providing a reliable and interpretable solution for real-time prediction of hypoglycemia, normal glucose and hyperglycemia. As shown in Figure 12, the distributions of classification accuracy of all models across different experimental runs are depicted in a box plot. The median accuracy of the proposed hybrid model exhibits the lowest variation, indicating the best stability and robustness as compared to baseline models.

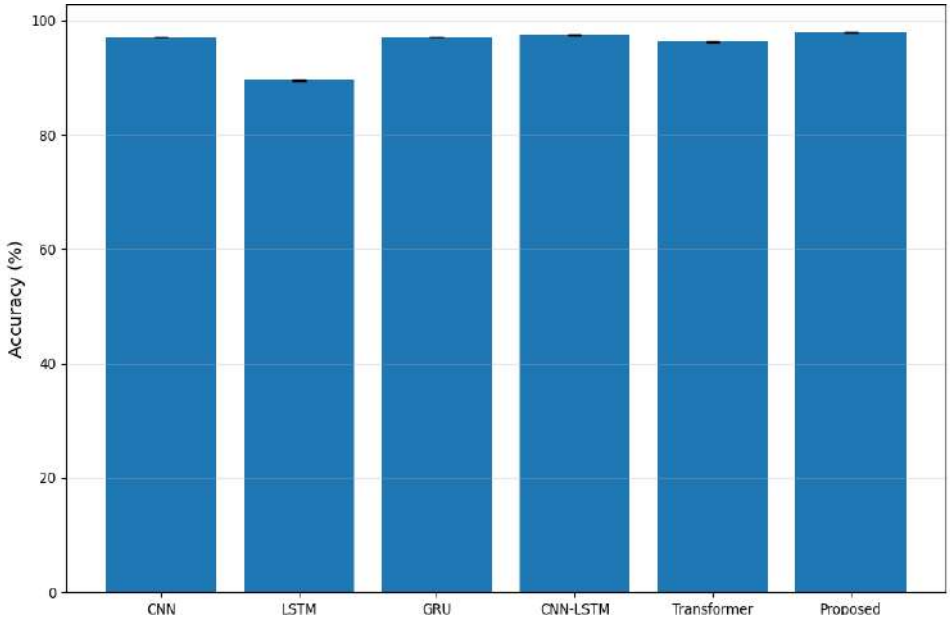

Figure 12: Statistical Comparison of Model Performance

VI. RESULTS AND DISCUSSION

Based on experimental results, the proposed MSCNN–BiLSTM–Transformer framework's performance for multi-class glucose-state prediction is superior to the baseline deep learning models. The score of the Suggested model has Accuracy (97.59%), Precision (97.78%), Recall (97.56%), F1-score (97.60%), ROC-AUC (99.79%), Matthews's correlation coefficient (MCC) (95.30%) and Cohen's Kappa (93.21%). As shown in Table 13, the suggested model surpassed all CNN, LSTM, GRU, CNN–LSTM and Transformer model in the same experimental settings. The confusion matrix analysis indicated that most of the glucose samples were correctly classified into Hypoglycemia, Normal, and Hyperglycemia, showing good discrimination ability with not much misclassification. In a comparable manner, the ROC curves almost perfectly distinguished the three glucose classes thus confirming the high sensitivity and specificity of the framework. The accuracy curves for training and validation are stably converging without overfitting. Moreover, the Multi-Scale CNN, BiLSTM, Transformer Encoder and Feature Fusion ablation study demonstrates that it greatly contributes to overall prediction performance. Furthermore, the SHAP interpretability analysis revealed that recent periods of high blood glucose concentration are the biggest contributors to the predicted outcome. Thus, it is found to be an interpretable analysis. As per the experimental findings, the hybrid architecture proposed in this research successfully captures local glucose variation along with long temporal dependencies. As a result, the prediction of usable glucose states is accurate, robust and clinically interpretable.

The results of the experiment indicate that integrating various deep learning approaches performs significantly better in glucose-state prediction compared to solely relying on an individual model architecture. CNNs can capture local temporal features while BiLSTM can easily learn sequential correlation. On the other hand, self-attention allows to Transformer Encoder learn global context. Using these useful representations, the proposed method enables the model to discover more about the glucose dynamics in past CGM data. The superior prediction performance of the proposed framework is augmented by its SHAP-based explainability, compared to recent deep learning models. The high metrics of prediction accuracy and ROC-AUC show that there is potential for intelligent clinical decision support, early detection of hypoglycemia and hyperglycemia, personalized diabetes management. The results showed that the proposed framework, MSCNN–BiLSTM–Transformer, is a viable choice in next-generation AI glucose monitoring systems.

VII. CONCLUSION

This study presents a MSCNN-BiLSTM-Transformer model for accurate multi-class blood glucose-state prediction. The applications of this new framework include accurate predictions and interpretation using CGM data. This innovative architecture successfully integrates multi-scale convolutional feature extraction, bidirectional temporal learning, a Transformer-based self-attention mechanism, and effective feature fusion to harness local glucose fluctuations and long-range temporal dependencies. Moreover, we employed SHAP-based Explainability to identify how historical glucose measurements impacted predictions to enhance model transparency. The proposed framework was evaluated using the OhioT1DM benchmark dataset, and its performance was compared with deep learning models like CNN, LSTM, GRU, CNN–LSTM and Transformer. The conducted tests revealed an accuracy of 97.59% for the model. The model achieved a precision of 97.78%, recall of 97.56%, and F1 score of 97.60%. The model's ROC-AUC performance was 99.79%. The model is 95.30 percent and 93.21 percent according to the MCC and Coehn's Kappa analysis. Moreover, the model consistently outperformed the comparative models. As further verification for the efficacy of unique architectural components, ablation study was conducted. Furthermore, the SHAP analysis offered interpretability and clinical reliability in the prediction process. The Hybrid Framework as a Solution of a Robust Accurate and Explainable Intelligent Glucose State Predictive Approach is capable of

customized diabetes management and real-time clinical decision support. The next step is to validate the proposed model with large multi-centric clinical datasets. Additional physiological signals like insulin dosage, food intake, and physical activity will be incorporated. Also, lightweight edge-AI implementations for real-time deployment in smart wearable and mobile health will be developed.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] O. Vershinina *et al.*, "Explainable artificial intelligence model predicting the risk of all-cause mortality in patients with type 2 diabetes mellitus," *Frontiers in Endocrinology*, vol. 16, Art. no. 1689312, Oct. 2025. Available from: https://doi.org/10.3389/fendo.2025.1689312

[2] H. Chen and L. Zhao, "Federated learning for privacy-preserving diabetes prediction in smart healthcare systems," *Artificial Intelligence in Medicine*, vol. 147, Art. no. 102611, 2025. Available from: https://doi.org/10.1016/j.artmed.2025.102611

[3] N. Patel and S. Verma, "Explainable AI for diabetes prediction using SHAP with ensemble models," *Journal of Biomedical Informatics*, vol. 146, Art. no. 104466, 2024. Available from: https://doi.org/10.1016/j.jbi.2024.104466

[4] R. Kumar, M. Singh, and A. Sharma, "IoT-enabled diabetes monitoring system using LSTM and wearable sensors," *IEEE Internet of Things Journal*, vol. 10, no. 1, pp. 785–793, 2023. Available from: https://doi.org/10.1109/JIOT.2023.3250148

[5] Y. Zhang, X. Li, and Y. Wang, "Hybrid deep learning for blood glucose prediction using continuous glucose monitoring data," *Biomedical Signal Processing and Control*, vol. 86, Art. no. 105286, 2023. Available from: https://doi.org/10.1016/j.bspc.2023.105286

[6] M. Garcia, J. Perez, and R. Alvarez, "Transformer-based blood glucose forecasting for diabetes management," *Computer Methods and Programs in Biomedicine*, vol. 244, Art. no. 107945, 2024. Available from: https://doi.org/10.1016/j.cmpb.2024.107945

[7] X. Liu, H. Wang, and Y. Chen, "Deep temporal neural networks for personalized glucose prediction using CGM data," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 4, pp. 2195–2206, 2024. Available from: https://doi.org/10.1109/JBHI.2024.3365127

[8] A. Singh, P. Kumar, and R. Gupta, "Multi-horizon blood glucose prediction using attention-enhanced BiLSTM networks," *Knowledge-Based Systems*, vol. 302, Art. no. 112867, 2025. Available from: https://doi.org/10.1016/j.knosys.2025.112867

[9] T. Zhou, H. Li, and L. Chen, "Explainable transformer networks for continuous glucose forecasting," *Expert Systems with Applications*, vol. 270, Art. no. 126134, 2025. Available from: https://doi.org/10.1016/j.eswa.2025.126134

[10] J. Wang, Z. Huang, and Y. Sun, "Multivariate time-series prediction of blood glucose using hybrid CNN-LSTM models," *Applied Soft Computing*, vol. 140, Art. no. 110215, 2023. Available from: https://doi.org/10.1016/j.asoc.2023.110215

[11] S. Ahmed, M. Islam, and M. Hasan, "Deep learning-based prediction of hypoglycemia events using wearable sensor data," *Sensors*, vol. 24, no. 3, Art. no. 1134, 2024. Available from: https://doi.org/10.3390/s24031134

[12] J. Kim, S. Lee, and H. Park, "Personalized glucose prediction using transformer and attention mechanisms," *IEEE Access*, vol. 12, pp. 53142–53156, 2024. Available from: https://doi.org/10.1109/ACCESS.2024.3389211

[13] F. Rodrigues, P. Silva, and M. Costa, "Real-time prediction of glucose abnormalities using sequential deep learning models," *Scientific Reports*, vol. 15, Art. no. 12487, 2025. Available from: https://doi.org/10.1038/s41598-025-12487-9

[14] T. Prioleau, B. Lu, and Y. Cui, "Glucose-ML: A collection of longitudinal diabetes datasets for development of robust AI solutions," *Scientific Data*, vol. 12, Art. no. 987, 2025.

[15] B. Cinar, L. van den Boom, and M. Maleshkova, "Deep learning-based hypoglycemia classification across multiple prediction horizons," *arXiv*, 2025. Available from: https://arxiv.org/abs/2504.00009

[16] B. Cinar, L. van den Boom, and M. Maleshkova, "A review on machine learning approaches for the prediction of glucose levels and hypoglycemia," *arXiv*, 2026. Available from: https://arxiv.org/abs/2601.11615

[17] G. Rigamonti, M. P. Barbato, D. Marelli, and P. Napoletano, "Tailoring adverse event prediction in type 1 diabetes with patient-specific deep learning models," *arXiv*, 2026. Available from: https://arxiv.org/abs/2601.14917

[18] L. Obiuwevwi, K. J. Rechowicz, V. Ashok, and S. Jayarathna, "Towards affordable, non-invasive real-time hypoglycemia detection using wearable sensor signals," *arXiv*, 2026. Available from: https://arxiv.org/abs/2602.10407

[19] B. Cheng *et al.*, "A 65 nm trustworthy hypoglycemia forecasting engine achieving 11.3 nJ per inference," *arXiv*, 2026. Available from: https://arxiv.org/abs/2606.07455

[20] Y. Luo, F. Chen, and W. Zhang, "Hybrid transformer–LSTM framework for continuous glucose forecasting in diabetes management," *Neural Networks*, vol. 183, Art. no. 106935, 2025. Available from: https://doi.org/10.1016/j.neunet.2025.106935

[21] A. H. Zuberi *et al.*, "Efficacy of structured exercise protocol on functional capacity, HbA1c, and waist-hip ratio in post-myocardial infarction subjects with diabetes mellitus," *Journal of Neonatal Surgery*, vol. 14, no. 5S, pp. 396–405, 2025. Available from: https://doi.org/10.52783/jns.v14.2058

[22] S. K. Singh, P. Verma, and P. Kumar, "Sentiment analysis using machine learning techniques on Twitter: A critical review," *Advances in Mathematics: Scientific Journal*, vol. 9, no. 9, pp. 7085–7092, 2020. Available from: https://doi.org/10.37418/amsj.9.9.58

[23] S. K. Singh, P. Verma, and P. Kumar, "Sentiment analysis of COVID-19 epidemic using machine learning algorithm on Twitter," *Journal of Critical Reviews*, vol. 7, no. 18, pp. 2565–2572, 2020. Available from: https://www.jcreview.com/

[24] S. Joshi, P. Gupta, R. Kumar, A. Dwivedi, and S. Kumar Singh, "Privacy by Design: User-Centric Control in Online Photo Sharing," in *2025 International Conference on Intelligent and Secure Engineering Solutions (CISES)*, Greater Noida, India, 2025, pp. 1287–1290. Available from: https://doi.org/10.1109/CISES66934.2025.11265161

[25] M. Mishra, M. S. Hussain, and S. K. Singh, "Protecting Against Social Engineering Using Wireshark: Effective Strategies With Real-World Examples," in *Effective Strategies for Combatting Social Engineering in Cybersecurity*. Hershey, PA, USA: IGI Global, 2025. Available from: https://doi.org/10.4018/979-8-3693-6665-3.CH008

[26] G. Mishra, H. Dev, and S. K. Singh, "Comparative Analysis of Pre-Trained CNN and Feature Selection Algorithms for Plant Leaf Disease Identification," in *2024 IEEE 11th Uttar Pradesh Section International Conference on Electrical, Electronics*

and Computer Engineering (UPCON), Lucknow, India, 2024, pp. 1–7. Available from: https://doi.org/10.1109/UPCON62832.2024.10983102

[27] G. Mishra, H. Dev, and S. K. Singh, "Evolutionary Algorithms and Deep Learning Techniques for Plant Disease Identification: A Comprehensive Review," in *Recent Advancements in Semiconductor Technologies*, Lecture Notes in Electrical Engineering, vol. 1491. Singapore: Springer, 2026. Available from: https://doi.org/10.1007/978-981-95-3305-3_27

[28] A. Singh, H. Agarwal, A. P. Singh, A. Kumar, S. Srivastava, and S. K. Singh, "Deep learning and red deer optimiser for automatic cardiovascular disease identification on magnetic resonance images," *Scientific Reports*, 2026. Available from: https://doi.org/10.1038/s41598-026-55595-6