Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 53-68 e-ISSN: 2656-8632

RESEARCH ARTICLE OPEN ACCESS

DCRNet: Hybrid Deep Learning Architecture for Forecasting of Blood Glucose

Ketan Lad [1], and **Maulin Joshi** [2]

[1] Department of Electronics and Communication Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India
[2] Department of Electronics and Communication, Sarvajanik College of Engineering and Technology, Gujarat, India

Corresponding author: Maulin Joshi (e-mail: maulin.joshi@scet.ac.in), **Author(s) Email**: Ketan Lad (e-mail: ketanlad_18@yahoo.co.in)

Abstract **Maintaining blood glucose (BG) levels within the euglycemic range is essential for patients with type 1 diabetes (T1D) to prevent both hypoglycemia and hyperglycemia. Often, BG concentration changes due to unannounced carbohydrate intake during meals or an inappropriate amount of insulin dosage. Timely forecasting of BG can help take appropriate actions in advance to keep BG within the euglycemic range. Recent studies indicate that deep learning techniques have demonstrated improved performance in this field. Deep learning approaches often struggle to precisely predict future BG levels. To address these challenges, this paper introduces a novel hybrid deep learning architecture called DCRNet. This architecture incorporates a Dilated Convolution layer that effectively detects multi-scale patterns while minimizing parameter count. Additionally, it utilizes Long Short-Term Memory (LSTM) to handle contextual dependencies and maintain the temporal order of the extracted features. DCRNet predicts future BG levels for short-term durations (15, 30, and 60 minutes) using information on glucose, meals, and insulin dosages. The proposed architecture's performance is evaluated on 11 simulated subjects from the UVA/Padova T1D Mellitus simulator and 12 actual subjects from the OhioT1DM dataset. In contrast to previous works, the proposed architecture achieves root mean square errors (RMSEs) of 3.42, 6.45, and 17.73 mg/dL for simulated subjects and 12.57, 20.72, and 34.41mg/dL for actual subjects, for prediction horizons (PH) of 15-, 30-, and 60-minute, respectively. The proposed architecture is also evaluated using the mean absolute error (MAE), which is 2.11, 4.47, and 11.78 mg/dL for simulated subjects and 7.9, 14.13, and 25.5 mg/dL for actual subjects, for 15-, 30-, and 60-minute PH. The experimental findings validate that the proposed architecture, which uses a dilated convolutional LSTM, outperforms other recent state-of-the-art models.**

Keywords **Dilated convolution, forecasting, Long Short-Term Memory, Blood Glucose, deep neural network, continuous glucose monitoring**

I. Introduction

Diabetes has become a serious health issue that has reached concerning levels. Currently, approximately 589 million people all over the world are living with diabetes. By 2030 and 2045, the total number of people who have diabetes will rise to 643 and 783 million, respectively [1]. The β cells that create insulin are damaged in people with Type 1 diabetes (T1D), preventing the pancreas from generating sufficient insulin [2]. Since insulin helps control the storage and use of glucose from carbohydrate consumption, its regulation can be disrupted, resulting in disease. It is challenging to control standard glucose-insulin transactions, as insufficient insulin or no insulin creates long-term complications that result in permanent damage to organs and nerves [3]. Worldwide, millions of people have already lost their lives due to diabetes and face high medical costs, which significantly highlight the clinical and financial consequences of inadequate glycaemic control. Therefore, T1D individuals must depend on external insulin to maintain their blood glucose levels (BG) within the targeted range of 70 to 180 mg/dL [4] to avoid hypoglycaemic or hyperglycaemic events.

Recently, the development of closed-loop mechanisms such as the artificial pancreas (AP) is an excellent example of significant technological advancements in this field [5]. These cutting-edge technologies mimic natural insulin release by using automated, glucose-responsive insulin delivery [6]. To maximize glycemic control, the main objective of AP systems is to reduce both hyperglycemia and hypoglycemia [7]. Although automated basal insulin delivery has been successfully achieved by AP systems, postprandial glucose control (PPGC) remains a challenge. For those with type 1 diabetes who have

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

to manually determine their pre-prandial insulin dosage based on their meal information, this presents a serious problem [8]. To address this challenge, adding an advanced BG forecasting model for T1D management can help emphasize PPGC [9]. Recent studies in artificial intelligence (AI), mainly in machine learning (ML) techniques, offer better glycaemic control by predicting future BG levels [10], [11]. Artificial neural networks (ANN) and other machine-domain learning (ML)-based methods are highlighted in the current research as having the ability to predict future BG levels, optimize insulin delivery, and identify hypoglycemia and hyperglycemic episodes early [12], [13], [14]. However, inaccurate or delayed glucose forecasts may result in incorrect insulin dosage, raising the possibility of hypoglycaemic and hyperglycaemic episodes that could result in hospitalization or permanent organ damage.

Recent studies have utilized recurrent and convolutional architectures to develop clinically useful BG forecasting models. Compared with conventional techniques, these models have demonstrated better performance on clinically important error metrics and root mean square error (RMSE) [15], [16], [17], [18], [19], [20]. Furthermore, studies on neural architecture search (NAS) and optimization indicate that modifying the architecture and hyperparameters can enhance prediction accuracy. Nevertheless, instead of radically altering the convolutional receptive field to successfully capture multi-scale glycaemic patterns, these approaches frequently focus on improving already-existing kernel designs [21]. GLYFE performs a thorough analysis of several customized glucose prediction models over a range of time periods. According to the results, kernel-based nonlinear models perform better than linear models in terms of clinical acceptability and accuracy. It's crucial to remember that not all models exhibit the same level of performance across all measures and prediction horizons (PHs) [22]. Bidirectional Long Short-Term Memory (Bi-LSTM) approaches frequently surpass traditional approaches across several non-invasive activities. However, by adding front-end encoders that efficiently capture both spectral and temporal multi-scale patterns, these models could be enhanced [23]. Hyperparameters of the LSTM are optimized using an advanced particle swarm optimization (PSO) approach, and variational modal decomposition (VMD) is utilized to decompose Continuous Glucose Monitoring (CGM) data into intrinsic modal functions that improve RMSEs for short-term predictions. The ensemble decomposition and PSO approaches can reduce rapid changes in BG. However, real-time or low-resource implementations limit the model's efficacy [24]. Recent studies have also used the Grey Wolf optimization method to enhance a stacking ensemble, thereby improving model performance. However, the absence of multi-scale feature extraction limits its accuracy in predicting rapid changes, reducing its suitability for lightweight, real-time deployment [25]. Recent efforts have focused more on personalized strategies, such as hybrid multimodal networks that use attention-based deep learning (DL) frameworks to combine glucose, insulin, meal, and physical activity data, and incrementally retrained LSTM models that adapt to each patient's CGM history [26]. Due to several factors that influence actual glucose levels beyond insulin infusions, accurately predicting BG levels remains challenging. The role of these parameters as input components in BG forecasting across various prediction horizons (PHs) has been investigated in earlier research, leading to increased accuracy [27], [28]. However, the findings show that this approach fails to learn both short-term and long-term patterns [29], [30], [31]. It has been demonstrated that feature engineering techniques that convert event data into continuous inputs can enhance forecast accuracy. Small sample sizes and restricted generalizability, however, remain significant limitations [32]. For PHs of 60 to 90 minutes, Lara-Abelenda et al. show that customized models based on Large Language Models (LLMs) can achieve competitive predicting accuracy and decrease variability among patients. However, this approach may not perform effectively for short-term PHs of 15 to 30 minutes in individuals with limited data due to the complexity offered by the model, and it still struggles with abrupt changes in BG levels [33].

Transformer topologies consistently surpass conventional recurrent models in representing long-term glucose dynamics across various PHs. However, their approach is computationally inefficient and may struggle in real-time deployment scenarios or in environments with limited data availability [34]. Factors such as sensor noise, patient variability, and external influences complicate accurate forecasting. These factors highlight the importance of having models that can accurately represent both long-term dynamics and local patterns [35]. Although lightweight explainable predictors are attractive for applications with limited resources, they frequently fall short of hybrid deep models in terms of accuracy across longer time horizons [36]. For instance, by including features such as food intake, insulin dosage, and step count, Kalman-smoothed stacked LSTM forecasts for T1D patients exhibit noticeably low RMSEs at 30- and 60-minute time frames. However, due to smoothing, these models typically perform poorly at extreme glucose levels and during rapid glucose fluctuations [37]. Evolving risk-aware ensembles enhance safety-critical forecasting, although they compromise on computational simplicity [38]. Hybrid architectures that combine convolutional feature extractors with sequence models have repeatedly shown superior

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

short and mid-horizon glucose-forecasting accuracy [39].

The Convolutional Recurrent Neural Network (CRNN) improves short-term glucose prediction accuracy, but its limited capacity to model very long-term dependencies and its high computational cost limit its performance [40]. Despite the hybrid CNN-GRU model outperforming single-model baselines in real-time blood glucose forecasting accuracy, its generalizability is limited due to its dependency on a single-patient dataset, lack of validation across a variety of populations, and shorter prediction horizons [41]. Hybrid transformer-LSTM models are efficient at capturing long-range dependencies, but they require substantial data and computational resources [42]. However, unless very deep stacks or large kernels are used, these hybrid models may not be able to capture long-range, multiscale temporal patterns; they are also computationally expensive for real-time edge deployment, remain relatively black-box, and are sensitive to CGM noise. These limitations motivate an approach that provides an LSTM with compact, informative sequences, enabling reliable sequential learning by employing dilated convolutions to effectively broaden the receptive field and capture multiscale temporal patterns. Compared to conventional CNN-LSTM stacks, this combination is expected to enhance long-horizon accuracy with fewer parameters and improved appropriateness for real-time applications.

In order to overcome the limitations noted in recent studies, the proposed hybrid dilated convolutional recurrent network (DCRNet) combines multi-scale spatial feature extraction through dilated convolutions with temporal dependency modeling via LSTM. This allows for enhanced prediction accuracy, robustness, and real-time applicability for next-generation AP systems while preserving optimal computational cost. The key contributions of this paper are as follows:

1. The proposed DCRNet is a compact hybrid architecture that combines a dilated convolutional layer with an LSTM network. The dilated convolutional neural network (DCNN) benefits from a broader receptive field, which enhances its ability to process multidimensional long signals, while the LSTM effectively captures non-linear stochastic relationships among multiple inputs.
2. An in-silico dataset is created that contains time series BG data points based on 1,344 meal scenarios from 11 in-silico adult subjects with T1D using the UVA/Padova T1DM simulator [43], [44].
3. DCRNet is validated on this in-silico dataset and demonstrates superior performance, achieving RMSEs of 3.42, 6.45, and 17.73 mg/dL, and mean absolute errors (MAEs) of 2.11, 4.47, and 11.78 mg/dL for PH of 15, 30, and 60 minutes, respectively.
4. DCRNet is also validated on 12 clinical T1D subjects from the OhioT1DM Dataset [45], where it outperforms the baseline, achieving RMSEs of 12.57, 20.72, and 34.41 mg/dL, along with MAEs of 7.9, 14.13, and 25.5 mg/dL for the same PH of 15, 30, and 60 minutes, respectively.

This paper is structured as follows: Part II details the datasets used, their preparation, and the proposed hybrid DL architecture. Part III evaluates the performance of this architecture and compares it with previous work. Part IV examines the findings and effectiveness of the proposed architecture, and Part V provides a conclusion.

II. Method

A. Datasets

In-silico and clinical datasets are used to train and validate the proposed architecture. Both datasets (the virtual cohort and the OhioT1DM dataset) are standard reference datasets, and the relevant data for all subjects have been captured to adequately account for inter-subject variability. The details of both datasets are as follows:

1. In-silico data

Eleven adult simulated subjects are selected from the UVA/Padova T1D Mellitus (T1DM) simulator. Four daily meals are planned: breakfast, lunch, snacks, and dinner. Various meal scenarios are generated by adjusting the minimum and maximum carbohydrate values for each meal. For breakfast, lunch, snacks, and dinner, the base carbohydrate amounts considered are 35, 45, 0, and 50, respectively. To create new meal scenarios, 10g is added to each meal based on the base values. The maximum carbohydrate intake for each meal is set at 85, 105, 30, and 120 grams. This process resulted in the creation and simulation of 1,344 meal scenarios for the 11 in silico adult subjects over the course of a day. The generated dataset consists of a time series containing 1,440 BG points recorded throughout the day for one adult, for each meal scenario. BG points are recorded every five minutes, along with corresponding carbohydrate and insulin data, to train the models. Data from eight subjects are used for training, while the remaining subjects are used for testing.

2. Clinical data

The OhioT1DM dataset was developed for the BG Level Prediction Challenge (BGLPC). It involved 12 individuals with T1D who participated for eight weeks. During this period, BG data were collected using CGM devices. Daily data on various factors were also gathered using a fitness band and a smartphone app. The dataset includes five females and seven males,

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

with ages ranging from 20 to 60. Participants recorded their carbohydrate intake, stress levels, sleep quality, exercise duration, and any illness symptoms using their smartphones. Additionally, the fitness bands collected data on heart rate, movement acceleration, Galvanic Skin Response (GSR), skin temperature, and step count. BG levels were recorded every five minutes, and both bolus and basal insulin doses were administered and noted. The dataset is provided in two separate sets: one for training and one for testing.

B. Data preprocessing and data preparation

1. Feature normalization

In data preprocessing, time-series data is time-aligned and normalized for DCNN. This modification prevents large-magnitude features from dominating by ensuring that each feature contributes proportionately within the same numerical range.

Let:

$$\boldsymbol{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T], \mathbf{x}_t \in \mathbb{R}^n \quad (1)$$

$\boldsymbol{X}$ denotes a multivariate time series having n features at each time step of length T as shown in Eq. (1) [20]. All input variables (blood glucose (BG), carbohydrate consumption (CHO), and insulin (BOLUS)) are adjusted using Min-Max normalization as per Eq. (2) [46] in order to ensure consistent feature scaling and enhance convergence stability.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (2)$$

where, X denotes the original value, $X_{\min}$ and $X_{\max}$ represent the minimum and maximum values of the corresponding feature in the training dataset, and $X' \in [0,1]$ is the normalized output used for model training.

2. Windowing and Segmentation

The time series normalized dataset is transformed into a collection of matched inputs and their corresponding intended outputs. To achieve this, various segments are created from the time-series data along with the desired outputs. While the dataset consists of continuous time-series samples taken every five minutes, the last 30 minutes of data (6 prior samples) are used as model input for prediction horizons of 15, 30, and 60 minutes. Let t represents the recent time index and y_t be the BG value at time t. The structure of every training instance, $\mathbf{X}_{\mathrm{t}}$, is represented by Eq. (3) [21].

$$\mathbf{X}_{\mathrm{t}} = [x_{t-5}, x_{t-10}, x_{t-15}, x_{t-20}, x_{t-25}, x_{t-30}] \quad (3)$$

where $x_{t-i} = [\mathrm{BG}_{t-i}, \mathrm{CHO}_{t-i}, \mathrm{BOLUS}_{t-i}]$ represents the feature vector at time $t - i$.

The forecasting horizon (*h*) affects the corresponding target output, as represented by Eq. (4) [14]:

$$y_{t+h} = \mathrm{BG}_{t+h} \quad (4)$$

Where $h \in \{15,30,60\}$ minutes. y_{t+h} and BG_{t+h} represent predicted BG value for horizon *h*.

Thus, the model learns a mapping for every horizon, as represented by Eq. (5) [14]:

$$f: \mathbf{X}_t \rightarrow y_{t+h} \quad (5)$$

A sliding window technique with a stride of s=1 (overlapping windows) is used to segment the dataset, guaranteeing optimal use of temporal data.

The sliding window approach, as illustrated in Fig. 1 can be used to segment the time-series dataset and prepare it for deep convolutional neural networks (DCNN). In this method, the input data is generated by collecting six samples (with a window size of w = 6) for learning purposes. Targeted BG is selected using the PH of the 3rd, 6th, or 12th position BG sample from time series data after a specified window for PH 15, 30, or 60 minutes, respectively. Since BG readings are taken every five minutes in this study, the prediction model forecasts the BG level after 15 minutes (as represented by the blue box) or 30 minutes (as shown by the green box) by utilizing the previous 30 minutes of historical data (indicated by the red boxes). The portion for the 60-minute PH is not shown in Fig. 1.

C. Proposed DCRNet architecture

A new hybrid DL architecture, DCRNet, is proposed for predicting BG levels. The proposed hybrid architecture sequentially combines dilated convolutional layers with long short-term memory (LSTM) layers to efficiently capture both local patterns and global temporal dependencies. DCRNet is trained and validated using in-silico data generated from the UVA/Padova T1DM simulator, as well as clinical data sourced from the OhioT1DM dataset. DCRNet comprises three components: Preprocessing and data preparation, DCNN, and LSTM. Fig. 2 shows the detailed architecture of DCRNet. C, I, and G represent carbohydrates in the meals, insulin dosage, and BG.

1. Dilated CNN

In CNN, convolution refers to the computation in each layer. The limited receptive field of CNN is one of its issues. The number of layers and their receptive field are comparable. To solve the problem, we employ dilated convolutions, which increase the receptive field without appreciably increasing computational cost. DCNN expands the receptive field without lowering spatial resolution or increasing parameters by utilizing dilated convolutions. The dilated 1D convolutional layers extract local contextual features from adjacent and distant time steps. The feature extraction approach for the dilated convolutional layers is described as Eq. (6) [46]:

$$\mathrm{h}_t^{(l)} = \emptyset\left(\sum_{k=0}^{K-1}\left(W_k^{(l)}\mathrm{W}_{l,k}\,\mathrm{X}_{t-d\cdot k}^{(l-1)} + b^{(l)}\right)\right) \quad (6)$$

where, $\mathrm{h}_t^{(l)}$represents the output feature map at layer l, $W_k^{(l)}$and $b^{(l)}$denote the kernel weights and biases, $\mathrm{W}_{l,k}$

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

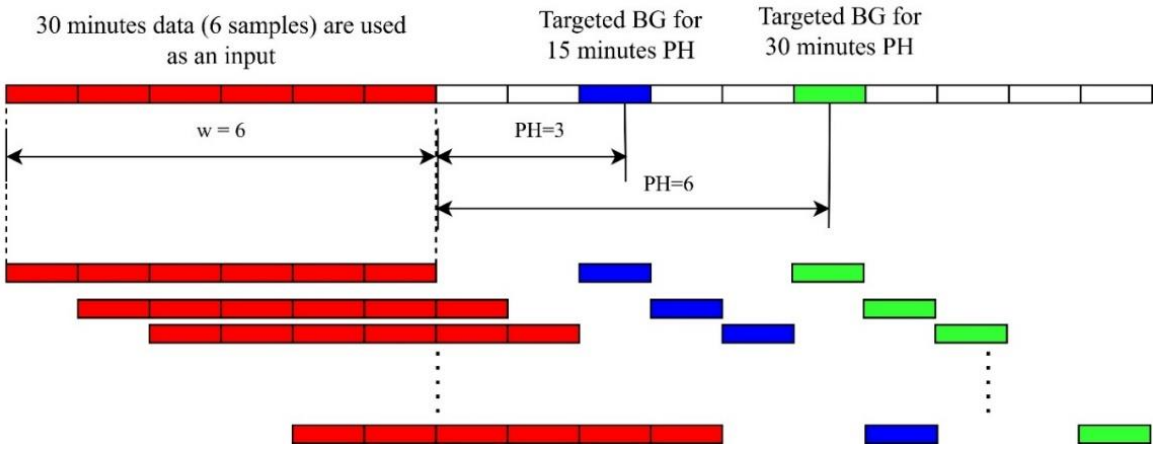


Fig. 1 Time-series data segmentation to create train and test dataset for PH= 15 and 30 minutes

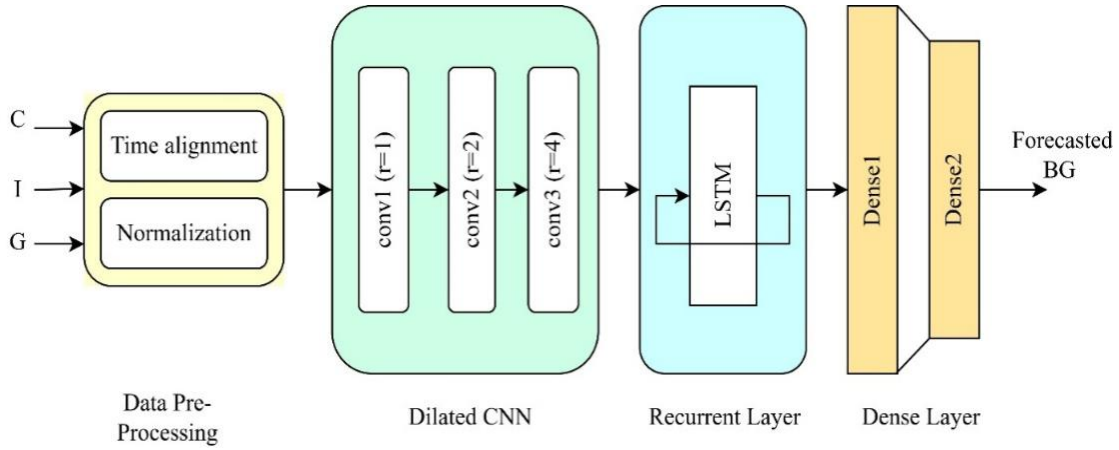


Fig. 2 Detailed architecture of proposed DCRNet: Data preprocessing, three Dilated CNN layers, one LSTM and two dense layers

denotes convolution kernel weights, r is the dilation rate (1, 2, or 4), K is the kernel size, and $\emptyset(\cdot)$ is the ReLU activation function.

ReLU is represented as Eq. (7) [47]:

$$\emptyset(x) = ReLU\ (x) = \max\ (0, x) \quad (7)$$

Each output channel j in $\mathbf{h}_t^{(l)}$ implements a learned linear filter over a small temporal stencil of the input features (with spacing d_l). With d_l increasing across layers (e.g., 1,2,4), the effective receptive field at layer L becomes R, as represented by Eq. (8) [48].

$$R = 1 + (K-1) \cdot \left(\sum_{l=0}^{L-1} \prod_{j=0}^{l} d_j\right) \quad (8)$$

where, K represents a constant kernel size and d_j represents an increasing dilation rate across layers. So local filters at earlier layers capture rapid changes (high-frequency), while deeper dilated layers capture mid-range structures. A standard convolution slides a small kernel (such as 3x3) over the input. In a DCNN, the receptive field is increased without raising the computational cost by inserting gaps (zeros) between kernel elements. The gap's spacing is determined by the dilation rate (*d*). The dilation rate remains the same in each layer, like {1,1,1,1…} in standard convolution, but is increasing exponentially in DCNN, such as {1,2,4,8,}.

Fig. 3 illustrates the 3 × 3 convolutional processes with d of 1 and 2 [20]. As illustrated in Fig. 3, the dilated convolution becomes a conventional convolution when r = 1. Fig. 3 (A) displays a 3 × 3 receptive field (d = 1), whereas Fig. 3 (B) displays a 5 × 5 receptive field (d = 2).

2. LSTM

Dilated convolutions are unable to explicitly simulate long-term sequential dynamics, yet they effectively local extract spatial-temporal context. To imitate global temporal dependencies, an LSTM is fed the feature sequence from the last dilated convolutional layer. An LSTM [24] consists of three main components: the input gate, the output gate, and the forget gate. The long-term memory cell's information storage is managed by the

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

input gate. The forget gate controls which data is erased from the LSTM cell. Information from the forget gate, current input, and input gate is combined and extracted by the output gate. This processed information is then used in the next time step. Fig. 4 shows LSTM Architecture [38]. $h_t^{(h)}$ represents the input from the hidden state. The forget gate creates f_t by adding $h_t^{(h)}$ and $h_{t-1}^{(h)}$, then passing the result via the sigmoid (σ) function. The forget gate regulates which LSTM cell data is discarded. g_t and $\overline{S_t}$ are utilized in the input gate to generate new information stored in the LSTM cell. To create a new cell state, S_t, the product of S_{t-1}, is added to the product of g_t and ($\overline{S_t}$). $h_t^{(h)}$ is the product of y_t and *Tanh* (S_t).

The multi-scale feature map derived from the dilated convolutions is fed into the LSTM. The output sequence from the last convolutional layer, $\{h_t^{(L)}\}_{t=1}^{T}$, is fed into an

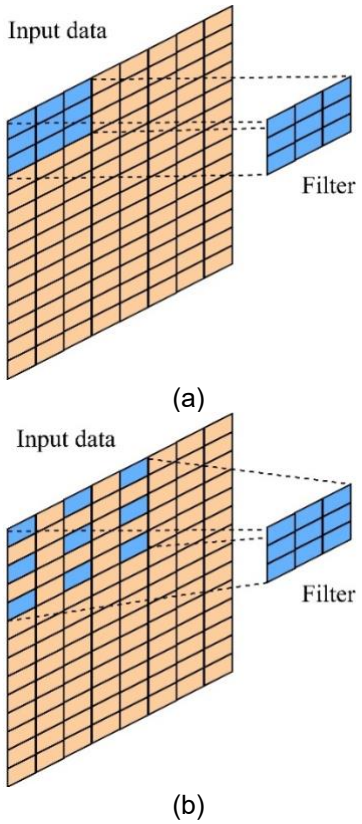


Fig. 3 Receptive field expansion in DCNN (A) DCNN with d = 1 (B) DCNN with d = 2 [20]

LSTM to model temporal dependencies. Long-term dependencies in the temporal sequence are captured at this level. The LSTM cell's output is determined using Eq. (9) to Eq. (14) [24] for each time step *t*.

$$S_t = g_t * \overline{S_t} + f_t * S_{t-1} \quad (9)$$

$$g_t = \sigma\left(W_h\left[h_{t-1}^{(h)} + h_t^{(h)}\right] + b_h\right) \quad (10)$$

$$f_t = \sigma\left(W_f\left[h_{t-1}^{(h)} + h_t^{(h)}\right] + b_f\right) \quad (11)$$

$$\overline{S_t} = tanh\left(W_S\left[h_{t-1}^{(h)} + h_t^{(h)}\right] + b_c\right) \quad (12)$$

$$h_t^{(h)} = y_t * (tanh\, S_t) \quad (13)$$

$$y_t = \sigma\left(W_y\left[h_{t-1}^{(h)} + h_t^{(h)}\right] + b_o\right) \quad (14)$$

b_h, b_f, b_s, b_y represents bias, W_h, W_f, W_S, W_y represents weights, * represents the product of Hadamard. The final LSTM output (e.g., the last hidden state $h_T^{(h)}$) is passed through fully connected layers for feature fusion and task-specific output.

3. Dense (fully connected) layer

In order to provide a final representation appropriate for

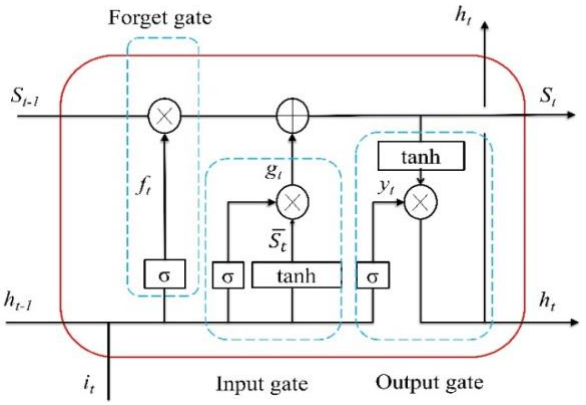


Fig. 4 An LSTM's architecture for identifying long-range temporal relationships in time-series information [24]

the output task, dense layers include temporal features. This permits the LSTM feature space to be mapped nonlinearly to the output space. The generalized dense layer, z_n, and the last dense layer for the regression task, $\hat{y}$ are represented by Eqs. (15) [14] and (16) [14], respectively.

$$z_n = \phi\left(W_n h_T^{(h)} + b_n\right) \quad (15)$$

$$\hat{y} = \psi(W_{n+1} z_{n+1} + b_{n+1}) \quad (16)$$

where, $\phi(\cdot)$ is the ReLU activation function, $\psi(\cdot)$ is the linear activation function for regression used at the last dense layer, $h_T^{(h)}$ represents the combined DCNN and LSTM outputs, and $W_n, W_{n+1}, b_n, b_{n+1}$ are trainable weights/biases.

The overall hybrid network output ($\hat{y}$) is represented as Eq. (17):

$$\hat{y} = f(X'; \Theta) = \psi(W_n\, \phi(W_{n+1}\, LSTM(Conv_{d_L}^{(L)} \circ \cdots \circ \mathrm{Conv}_{d_1}^{(1)}(X')) + b_n) + b_{n+1}), \quad (17)$$

where, Θ denotes the set of all trainable parameters for the BG forecasting model, X' denotes the preprocessed multivariate input sequence, and $W_n, W_{n+1}, b_n, b_{n+1}$ are trainable weights/biases of dense layers. Convolutional layers learn high-frequency, short-term temporal patterns (local dependencies), while the LSTM learns

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

long-range, low-frequency temporal associations (global dependencies) by integrating these extracted features over time.

D. Optimization and training

The Mean Squared Error (MSE) loss function, which is described as Eq. (18) [10], is minimized in order to optimize the model parameters $\Theta = \{\mathrm{W}, \mathrm{b}\}$:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (18)$$

where, y_i represents true target value for the i^{th}sample, $\hat{y}_i$ represents the model-predicted value for the i^{th}sample and N represents the number of samples in the dataset.

The Adam optimizer, which adaptively modifies learning rates for each parameter, is used to carry out the optimization. The Eq. (19) to Eq. (22) [49] are the expressions for the parameter update rule:

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \quad (19)$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \quad (20)$$

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t}, \hat{v}_t = \frac{v_t}{1-\beta_2^t} \quad (21)$$

$$\Theta_{t+1} = \Theta_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t}+\epsilon} \quad (22)$$

where, $g_t = \nabla_{\Theta_t} MSE$ is the gradient at iteration t, η is the learning rate, and β_1, β_2, and ϵ are Adam hyperparameters. The network was trained end-to-end using this objective until convergence.

E. Modeling Local and Global Dependencies in Glucose Dynamics

The DCRNet architecture integrates features from the dilated convolutional (CNN) and recurrent (LSTM) modules to explicitly characterize local and global temporal dependencies. Long-term trends controlled by physiological patterns are reflected in global dependencies, whereas local dependencies show transient glucose variations induced by meals or insulin events. Dilated convolutional layers use constrained receptive fields centered on time t to extract local features that result in localized feature maps $\mathrm{h}_t^{(L)}$, as shown in Eq. (23):

$$h_t^{(L)} = f_{CNN}(x_{t-R:t}) \quad (23)$$

where, the dilation rate d_l controls the effective temporal receptive field, denoted by R. These feature maps capture multi-scale temporal fluctuations in the input signal and quantify short-term associations. The CNN feature sequence $\{h_1^{(L)}, \ldots, h_T^{(L)}\}$ is fed into the LSTM, which integrates it over time to simulate long-term dependencies, as shown in Eq. (24):

$$z_t = f_{LSTM}(h_{1:t}^{(L)}) \quad (24)$$

where, z_t represents the LSTM output that captures global temporal context, preserving information across distant time steps.

F. Summary of proposed algorithm for BG forecasting architecture

Eq. (25) demonstrates the complete algorithm's processes, from input multivariate features to final prediction:

$$X_{1:t} \xrightarrow{f_{Pre}} X'_{1:t} \xrightarrow{f_{CNN}} H_{1:t}^{(L)} \xrightarrow{f_{LSTM}} L_{1:t} \xrightarrow{f_{Dense}} \hat{y}_t \quad (25)$$

where, $\mathrm{X}_{1:t} \in \mathbb{R}^{T\times n}$denotes the multivariate time-series input with n features (glucose, insulin, carbohydrate intake), $X'_{1:t}$ represents normalized input, $H_{1:t}^{(L)}$represents output of last DCNN layer, $L_{1:t}$ represents output of LSTM and represents $\hat{y}_t$ final output of proposed DCRNet. This representation illustrates the hierarchical flow from normalized multivariate input data, through CNN-based short-term feature extraction and LSTM-based long-term modeling, to final glucose prediction via dense layers, optimized end-to-end. The following steps summarize the proposed work for BG forecasting, explained in earlier sections:

Step 1:

The preprocessing function $f_{\mathrm{Pre}}(\cdot)$ standardizes and normalizes input data, $X_{1:t}$, to ensure numerical stability and give $X'_{1:t}$, as shown in Eq. (26) [46]:

$$X'_{1:t} = f_{Pre}(X_{1:t}) = \frac{X_{1:t}-X_{min}}{X_{max}-X_{min}} \quad (26)$$

Step 2:

The pre-processed multivariate features are applied to DCNN. It extracts short-term temporal features using dilated convolutions, as shown in Eq. (27):

$$H_{1:t}^{(L)} = f_{CNN}(X'_{1:t}) \quad (27)$$

where, $H_{1:t}^{(L)}$ represents the output of the last layer of the dilated convolution layer.

Step 3:

The last layer output of the DCNN is applied to LSTM, which captures long-term dependencies and gives $L_{1:t}$, as shown in Eq. (28):

$$L_{1:t} = f_{LSTM}(H_{1:t}^{(L)}) \quad (28)$$

Step 4:

The output of LSTM is fed to dense layers that perform regression to estimate future BG values, as shown in Eq. (29):

$$\hat{y}_t = f_{\mathrm{Dense}}(L_{1:t}) \quad (29)$$

Step 5:

Model training is formulated as minimizing the mean squared error (MSE) loss and is optimized using the Adam algorithm.

III. Result

A. Architecture design and Hyperparameter settings

In the proposed architecture, we utilize three layers of DCNN, one LSTM layer, and two dense layers.

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

The first three DCNN layers consist of 8, 16, and 32 neurons, respectively, with the parameters *r* set to 1, 2, and 4 for each layer. The LSTM layer contains 64 cells. The two dense layers consist of 128 and 64 neurons, respectively, with 30% dropout applied after each. Forecasting BG is conducted using six samples. For a fair comparison with current models, the proposed DCRNet uses continuous blood glucose levels, carbohydrate intake during meals, and insulin as inputs. The results are evaluated against various models, including Multi-Linear Regression (MLR), Random Forest (RF), Decision Trees (DT), AdaBoost, XGBoost, Fully Connected networks (FC), Convolutional Neural Network (CNN), Dilated CNN (DCNN), Simple Recurrent Neural Network (RNN), and LSTM.

Table 1. Hyperparameter settings for MLR, DT, RF and AdaBoost ML models

Model	Hyperparameter setting
MLR	-
DT	Gini impurity is used to assess the quality of a data split.
RF	Ten estimators are considered, and Gini impurity is used to measure the quality of the split.
AdaBoost	Learning rate = 1 and Maximum estimators = 50

Table 1 presents the hyperparameter settings for the ML models used during implementation. The FC model consists of six hidden layers, each containing 10 neurons. The CNN includes two convolutional layers with filter sizes of 6 and 16, followed by a max-pooling layer and two dense layers with 128 and 64 neurons, respectively. The Simple RNN is designed with two recurrent layers containing 64 and 32 neurons, respectively, along with two dense layers with 16 and 8 neurons, respectively. The LSTM model has three LSTM recurrent layers with 64, 32, and 16 neurons, followed by four dense layers with 64, 32, 16, and 8 neurons. The DCNN is structured with three convolutional layers having filter sizes of 8, 16, and 32, along with two FC (dense) layers containing 128 and 64 neurons. The Rectified Linear Unit (ReLU) serves as the activation function, and the Adam optimizer is utilized in all implementations.

B. Evaluation metrics

The performance of implemented models is evaluated using root mean square error (RMSE) and mean absolute error (MAE) between actual observed and forecasted BG values. RMSE is calculated according to Eq. (30) [20],

$$RMSE = \sqrt{\frac{1}{m}\sum_{k=1}^{m}\left(F_k - \widehat{F_k}\right)^2} \quad (30)$$

where, F_k = Actual observed output, $\widehat{F_k}$ = forecasted output, m = number of predictions. MAE is a metric that quantifies the mean scope of the absolute errors between Forecasted and actual observed outputs. It is the mean of the absolute differences between the Forecasted and the actual observed BG values. MAE is calculated as per Eq. (31) [10],

$$MAE = \frac{1}{m}\sum_{k=1}^{m}\left|F_k - \widehat{F_k}\right| \quad (31)$$

where, F_k = Actual output, $\widehat{F_k}$ = Predicted output, m = number of predictions.

C. Performance evaluation on in-silico subjects using UVA/Padova T1DM simulator

To evaluate the robustness and generalizability of the proposed DCRNet architecture, the model is validated using data from 11 virtual adult subjects of the UVA/Padova T1DM simulator. The model is trained on data from eight subjects simultaneously to capture shared glucose dynamics and inter-individual variability, while the remaining three subjects are reserved for independent testing, thereby representing population-level generalization rather than subject-specific tuning.

Table 2 compares the performance of various models, including MLR, DT, RF, AdaBoost, XGBoost, FC, CNN, DCNN, Simple RNN, LSTM, and the proposed DCRNet. The evaluation is based on RMSE (mg/dL) and MAE (mg/dL) for three different prediction horizons: 15, 30, and 60 minutes. The metrics are computed over the entire test dataset to capture overall predictive performance. For the 15-minute PH, MLR, RF, and DT demonstrated relatively low error rates, outperforming AdaBoost and FC. MLR recorded RMSE values of 3.06 mg/dL and 1.57 mg/dL, while AdaBoost showed significantly higher errors, with values exceeding 11.09 mg/dL and 9.85 mg/dL. The DL models, such as LSTM and Simple RNN, achieved moderate performance, with RMSE values of 7.37 mg/dL and 7.25 mg/dL, respectively, compared to FC and CNN. Among the existing models, DCNN achieved the best performance, with a 2.87 mg/dL RMSE value and a 2.11 mg/dL MAE. The proposed DCRNet surpassed the benchmark models, achieving RMSE and MAE values of 3.42 mg/dL and 2.11 mg/dL, respectively. However, MLR and DCNN still provided slightly better performance for the 15-minute PH. The models DT, RF, and XGBoost achieve lower error rates for 30 minutes of PH, with RMSE values of 8.19, 7.16, and 6.95 mg/dL, respectively. In comparison, the baseline model, MLR, demonstrates impressive performance, with RMSE and MAE values of 6.24 and 4.01 mg/dL, respectively, surpassing those of other baseline and ensemble models. Among the baseline DL models, CNN and Simple RNN show decent performance for the same 30-minute pH period,

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

Table 2 Performance comparison of BG forecasting models for in-silico subjects of UVA/Padova T1DM simulator

Metrics	Model name	Future Prediction Horizon (PH) (in minutes)		
		15	**30**	**60**
RMSE (mg/dL)	MLR	3.06	6.24	17.55
	DT	4.39	8.19	21.5
	RF	3.84	7.16	19.19
	AdaBoost	11.09	17.19	31.05
	XGBoost	4.07	6.95	18.11
	FC	10.29	11.00	20.14
	CNN	9.82	8.36	17.88
	Simple RNN	7.25	8.87	23.27
	LSTM	7.37	7.76	17.56
	DCNN	2.87	5.54	17.88
	DCRNet (proposed)	**3.42**	**6.45**	**17.73**
MAE (mg/dL)	MLR	1.57	4.01	12.25
	DT	2.12	4.86	12.59
	RF	1.76	4.26	11.29
	AdaBoost	9.85	15.74	27.07
	XGBoost	2.72	4.64	12.93
	FC	7.93	7.8	12.8
	CNN	7.04	5.34	10.73
	Simple RNN	6.21	6.77	17.24
	LSTM	5.63	5.29	11.15
	DCNN	2.11	3.19	11.72
	DCRNet (proposed)	**2.11**	**4.47**	**11.78**

recording RMSE values of 8.36 and 8.87 mg/dL, respectively. However, LSTM outperforms them with RMSE and MAE values of 7.76 and 5.29 mg/dL, respectively. DCNN exhibits superior performance among the existing models, achieving RMSE and MAE values of 5.54 mg/dL and 3.19 mg/dL, respectively, for the 30-minute PH.

The proposed DCRNet outperforms the benchmark models, achieving RMSE and MAE values of 3.42 and 2.11 mg/dL, respectively, except for MLR and DCNN.

At a PH = 60 minutes, MLR demonstrates relatively lower errors compared to DT, RF, AdaBoost, and XGBoost, achieving RMSE and MAE values of 17.55 and 12.25 mg/dL, respectively. In contrast, AdaBoost shows significantly higher errors, with RMSE and MAE values exceeding 31.05 and 27.07 mg/dL, respectively. DL models, including CNN and LSTM, performed well, recording RMSE values of 17.88 and 17.56 mg/dL at the 60-minute PH, respectively. However, the performance of FC and Simple RNN notably declines as PH increases, with RMSE values reaching 20.14 and 23.27 mg/dL, respectively. Among the existing baseline and ensemble models, DCNN achieves superior performance, with an RMSE of 17.88 mg/dL and an MAE of 11.72 mg/dL. The proposed DCRNet consistently outperforms the benchmark models, except for LSTM, achieving RMSE and MAE values of 17.73 and 11.78 mg/dL, respectively. Overall, when compared to the ML and DL baselines for three PH, the proposed DCRNet performs consistently. Although this study does not specifically measure inter-individual variability, the continuous performance trend across horizons suggests that the proposed hybrid architecture generalizes well to individuals who are not visible.

D. Performance evaluation on actual subjects using the OhioT1DM dataset

Table 3 presents the experimental outcomes for three PH (15, 30, and 60 minutes) across 12 subjects from the OhioT1DM dataset, in terms of RMSE and MAE. Consequently, the training samples from all 12 subjects are combined to form the training set, and the test samples are combined to form the test set for assessment. With this method, the estimate of generalization across diverse subjects is more accurate.

For a PH of 15 minutes, traditional ML approaches, such as MLR, DT, and RF, demonstrate moderate

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

performance with RMSE values of 13.1, 18.64, and 13.1 mg/dL, respectively. Conversely, ensemble-based approaches such as AdaBoost have suboptimal predicting accuracy, evidenced by an RMSE of 32.05 mg/dL. The proposed DCRNet attains optimal accuracy, with an RMSE of 12.57 mg/dL and an MAE of 7.9 mg/dL. These figures exceed all baseline models, with the nearest competitor, a DCNN, attaining an RMSE of 12.80 mg/dL and an MAE of 8.13 mg/dL. After 30 minutes, the baseline DL models, LSTM and Simple RNN, exhibit reasonable performance, with RMSEs of 21.21 and 23.6 mg/dL and MAEs of 14.46 and 16.88 mg/dL, respectively. DCRNet achieves a minimum RMSE of 20.72 mg/dL and an MAE of 14.13 mg/dL, underscoring the efficacy of the proposed architecture. Although DCNN maintains strong performance, evidenced by an RMSE of 20.83 mg/dL and an MAE of 14.28 mg/dL, the suggested model regularly surpasses alternative baseline models in optimizing both error measures.

For a PH of 60 minutes, DCRNet demonstrates stable performance, with RMSE and MAE values of 34.41 mg/dL and 25.50 mg/dL, respectively. In comparison, DCNN achieves slightly better results, reporting an RMSE of 34.30 mg/dL and an MAE of 25.36 mg/dL. The conventional and ensemble models, such as MLR, DT, RF, AdaBoost, and XGBoost show substantially higher error rates. Meanwhile, the performance of DL baselines, including FC, CNN, simple RNN, and LSTM, significantly declines at this time horizon. Their RMSE values are 37.49, 40.08, 40.16, and 35.60 mg/dL, with corresponding MAE values of 28.56, 31.59, 29.87, and 26.31 mg/dL, respectively. Consistent error trends across both simulated and real datasets imply that DCRNet successfully learns subject-independent temporal and physiological patterns, despite the fact that individual-level findings were not presented separately. Inter-individual factors such as insulin sensitivity, absorption of carbohydrates, and sensor noise are likely to cause minor variances in performance; however, the hybrid dilated-convolutional and LSTM structure reduces these effects by simulating both local fluctuations and long-term glucose trends simultaneously.

IV. Discussion

A. Performance evaluation on in silico and actual subjects

The findings indicate a number of significant discoveries about how well BG prediction performed. Traditional machine learning methods, such as MLR, DT, and RF, perform well enough for short-term predictions but have trouble generalizing as the prediction window gets wider. The persistent underperformance of ensemble-based techniques, especially AdaBoost, raises the possibility that boosting strategies are not appropriate for the extremely dynamic nature of physiological data, like glucose variations.

In contrast, DL models exhibit stronger predictive power, with CNN, LSTM, and DCNN outperforming most conventional models. Their efficacy, however, differs depending on the time frame. The limitations of LSTM and Simple RNNs in modeling complex, long-range temporal relationships are highlighted by their performance: they perform well for short-term predictions but significantly deteriorate for long-term ones. Even though CNN-based methods are good at identifying small temporal patterns, their accuracy also decreases as the PH rises. DCNN performs competitively, achieving the lowest MAE of all baseline models for in-silico participants, especially at the 15- and 30-minute PH. However, DCRNet consistently achieves reduced Root Mean Square Error (RMSE) values across all PHs and maintains competitive MAE values for the real patient dataset, outperforming DCNN and all other models. This implies that the suggested architecture provides a superior trade-off between long-term robustness and short-term sensitivity.

DCRNet's ability to capture both local and global dependencies in glucose dynamics is demonstrated by its persistent superiority in both RMSE and MAE measurements. Its resistance to the degradation commonly observed in other models over extended pHs is probably a result of its hybrid design, which combines convolutional and recurrent components. Additionally, the slight variations between DCRNet and DCNN at the 60-minute PH show that convolutional-based methods remain reliable baselines, while DCRNet's improvements yield measurable gains in prediction reliability.

B. Comparison with state-of-the-art models

Table 4 compares state-of-the-art models with the proposed DCRNet architecture for PH at 15, 30, and 60 minutes, regarding RMSE (mg/dL) and MAE (mg/dL). For a fair comparison, we focus on papers that test the forecasting model using six or twelve patients from the Ohio T1DM dataset, reporting results for PHs of 30 and 60 minutes. Table 4 shows that the proposed model significantly outperforms existing approaches, especially in short-term glucose forecasting. At a PH of 15 minutes, where previous studies did not provide results, the model achieves impressively low errors (RMSE=12.57 mg/dL, MAE=7.90 mg/dL), demonstrating its ability to accurately capture rapid glucose fluctuations. At 30 minutes PH, the model surpasses most previous studies, boasting the lowest MAE (14.13 mg/dL) and a competitive RMSE (20.72 mg/dL). This reflects both improved average prediction accuracy and more reliable short-term forecasts. Compared with the work of Dudukchu et al. [16] and Cioppa et al. [38], the reductions in RMSE and MAE are significant.

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

Table 3. Comparison of Performance among BG Forecasting Models for actual subjects of OhioT1DM dataset

Metrics	Model name	Future Prediction Horizon (PH) (in minutes)		
		15	**30**	**60**
RMSE (mg/dL)	MLR	13.1	22.19	35.84
	DT	18.64	31.14	49.31
	RF	13.1	21.82	35.51
	AdaBoost	32.05	37.5	44.75
	XGBoost	13.38	22.24	35.36
	FC	26.07	36.56	37.49
	CNN	38.94	33.22	40.08
	Simple RNN	16.6	23.6	40.16
	LSTM	14.17	21.21	35.6
	DCNN	12.80	20.83	34.30
	DCRNet (proposed)	**12.57**	**20.72**	**34.41**
MAE (mg/dL)	MLR	8.42	15.60	27.03
	DT	11.68	21.33	36.25
	RF	8.3	15.11	26.36
	AdaBoost	24.85	29.71	36.08
	XGBoost	8.72	15.7	26.53
	FC	19.80	26.91	28.56
	CNN	37.77	25.85	31.59
	Simple RNN	11.94	16.88	29.87
	LSTM	9.50	14.46	26.31
	DCNN	8.13	14.28	25.36
	DCRNet (proposed)	**7.9**	**14.13**	**25.50**

Furthermore, the performance remains comparable to the best values reported by Giancotti et al. [30] and Nemat et al. [27]. At the 60-minute PH, the differences between the models are minor, with the proposed method performing similarly to or slightly better than existing approaches, indicating stable long-term predictive capacity. This demonstrates that DCRNet outperforms recent hybrid models and maintains its computational efficiency. The comparison emphasizes the effectiveness of dilated convolutional layers in capturing multi-scale features. It discusses the significance of LSTM in modelling long-term glucose trends, which enhances efficacy across different prediction horizons. These findings prove the robustness and clinical applicability of the proposed architecture in the domain of BG forecasting models. In summary, DCRNet shows its potential to accurately predict BG for the PH of 15, 30, and 60 minutes in both in-silico and real subjects. This justifies the role of DCRNet in real-world applications where accurate forecasting is essential for proactive diabetes management.

C. Statistical Analysis of baseline and proposed models and Its Significant Implications

A detailed statistical analysis between baseline and proposed models is carried out for the 60-minute PH (Table 5). This highlights the computational challenge and clinical applicability of BG forecasting models, serving to support the evaluation of model performance. For all the implemented models, $p < 0.001$ except DT ($p>0.005$). With a mean prediction error of 1.87 mg/dL ($t = 9.694$, $p < 0.001$), an inference latency of 0.054 ± 0.004 s, and 83,817 trainable parameters, the proposed DCRNet proved to be both computationally efficient and statistically robust. While most baseline models showed significant bias, DCRNet maintained a tiny but persistent variance (< 2 mg/dL), which was statistically detectable but clinically insignificant, according to confidence tests. Real-time viability for closed-loop insulin delivery is confirmed by its low computing requirement and smalllatency (< 0.02% of the 5-min sampling window). Additionally, the model's sequence-based training facilitates adaptation to patient-specific glucose dynamics, and its convolutional and recurrent architecture intuitively reduces the influence of transient

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

Table 4. Comparison of state-of-art models and proposed DCRNet architecture for PH of 15, 30 and 60 minutes in terms of RMSE (mg/dL) and MAE (mg/dL)

Research works	Input feature used	No. of subjects	PH=15		PH=30		PH=60	
			RMSE	MAE	RMSE	MAE	RMSE	MAE
Dudukchu et al. [16]	CGM	12	-	-	21.9	15.86	35.10	-
Giancotti et al. [30]	CGM	6	-	-	20.48	14.59	34.16	25.54
Nemat et al. [27]	CGM	12			20.21	14.36	34.66	25.89
Nemat et al. [27]	CGM, Carbohydrate intake, insulin, Physical activity	12	-	-	20.76	14.56	34.85	25.60
Cioppa et al. [38]	CGM, Carbohydrate intake, insulin	6	-	-	21.34	-	-	-
Dudukchu et al. [50]	CGM	12	-	-	21.54	15.39	-	-
Proposed DCRNet	CGM, Carbohydrate intake, insulin	12	12.57	7.9	20.72	14.13	34.41	25.50

Table 5 Comparative performance and statistical analysis of baseline and proposed models for 60-minute PH, highlighting learnable parameters, inference latency, mean prediction error, and significance testing results (t-statistic and p-value)

Models	Learnable parameters	Inference (s)	Mean prediction error (mg/dL)	t-statistic	p-value
MLR	19	0.0001 ± 0.00002	0.86	4.372	< 0.001
DT	9,85,271	0.0007 ± 0.0002	0.43	1.57	> 0.05
RF	3,50,45,656	0.0041 ± 0.0004	0.73	3.643	< 0.001
AdaBoost	750	0.0036 ± 0.0006	-15.576	-66.132	< 0.001
XGBoost	1600	0.0004 ±0.0001	1.08	5.454	< 0.001
FC	871	0.048 ±0.003	9.43	46.319	< 0.001
CNN	10,835	0.32 ±0.06	23.36	110.270	< 0.001
Simple RNN	8,129	0.3 ± 0.04	23.64	114.12	< 0.001
LSTM	37,249	0.3 ± 0.05	14.29	68.903	< 0.001
DCNN	34,409	0.049 ± 0.007	2.435	12.675	< 0.001
DCRNet (proposed)	83,817	0.054 ± 0.004	1.87	9.694	< 0.001

sensor noise. All of these characteristics indicate that DCRNet is clinically and practically ready for integration into wearable or embedded insulin administration systems, in addition to achieving statistically proven accuracy.The DCRNet model is appropriate for real-time and embedded implementations in CGM systems because it provides a balanced trade-off between accuracy and computing efficiency from the perspective of deployment. On typical GPU hardware, the model training procedure takes a few hours and requires substantial processing power. The proposed model predicts each test sample in milliseconds during inference, enabling it for real-time BG forecasting and making it suitable for implementation in wearable or embedded devices. These features indicate the suitability of practical deployment of the proposed model on low-power hardware platforms that are utilized in closed-loop insulin delivery systems and CGMs.

D. Limitations of DCRNet and future scope

This study acknowledges several significant limitations despite the exceptional performance of the proposed DCRNet architecture. The study mainly focuses on in-silico and the OhioT1DM dataset, which lack the complexity and diversity of glucose dynamics that are

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

present in larger, more heterogeneous populations. The proposed architecture also demonstrates strong predictive performance for short- and medium-term BG levels. However, sparse data, sensor noise, and variations in BG regulation among different people may affect the performance of the proposed architecture.

Therefore, future research will focus on validating the architecture using large, more diverse clinical datasets collected across various activity levels, demographic conditions, and treatment methods to enhance generalizability and ensure reliability in real-world clinical settings. The architecture forecasting accuracy may further increase utilizing the data of individuals that have distinct glucose dynamics and subject-specific adaptive mechanisms, including transfer learning or personalized fine-tuning. To determine whether the architecture can be used in closed-loop insulin delivery systems, real-time implementation studies must also be done. These studies must include an assessment of latency and computational efficacy. To mitigate the sensor artifacts and emphasize architecture compression, further research will concentrate on adaptive filtering approaches and create a robust protocol for clinical validation. Further adding redundant features and uncertainty estimations to the forecasting architecture will emphasize the reliability and safety of the proposed architecture.

E. Implications of DCRNet

This study has major implications for predictive modelling and diabetes control. The results highlight DCRNet's promising potential in decision-support systems, closed-loop insulin delivery systems, and early warning systems for both hypoglycaemia and hyperglycaemia. It demonstrates a reliable ability to produce steady and accurate forecasts across various PHs. DCRNet is an excellent alternative for efficient and adaptive diabetes management due to its capacity to balance short-term sensitivity with long-term robustness. Additionally, the observed enhancements over state-of-the-art models underscore the crucial role of hybrid deep learning architectures in developing data-driven, personalized healthcare solutions.

V. Conclusion

This work aims to develop a novel hybrid deep learning architecture called DCRNet. This architecture increases prediction accuracy for BG forecasting by utilizing a multilayer dilated convolutional layer with an LSTM network, allowing the CNN to identify patterns or features within multidimensional time series data. Dilation convolution increases the receptive field without adding additional parameters or increasing computational costs, effectively capturing dependencies over longer time series. The proposed architecture achieves the lowest RMSEs of 3.42, 6.45, and 17.73 mg/dL for simulated subjects and 12.57, 20.72, and 34.41 mg/dL for actual subjects with prediction horizons (PH) of 15, 30, and 60 minutes, respectively. The proposed architecture also achieves the lowest MAE: 2.11, 4.47, and 11.78 mg/dL for simulated subjects and 7.9, 14.13, and 25.5 mg/dL for actual subjects, for 15-, 30-, and 60-minute PH. This architecture helps predict BG levels in individuals with T1D, allowing prompt corrective actions to regulate BG levels and prevent health complications arising from hyperglycaemia and hypoglycaemia. However, if model complexity is a concern, MLR is a simpler alternative to DCRNet. It can be used for forecasting, as its performance is comparable to that of the proposed architecture.

The proposed DCRNet is capable of reliably predicting both short- and long-term glucose concentrations from multivariate physiological inputs. The model achieves better predictive accuracy than existing methods by balancing short-term sensitivity and long-term robustness through the integration of dilated convolutions for local feature extraction, LSTM-based long-term modeling, and adaptive fusion. These results demonstrate that it can improve individualized therapies and real-time glucose management. Generalization across many demographics and physiological situations is one of the remaining unknowns. Future work will concentrate on validating DCRNet on larger heterogeneous datasets, refining the architecture for real-time deployment, lowering computing costs through hardware-efficient implementations, and investigating adaptive personalization techniques to further improve clinical applicability.

Acknowledgment

The authors are thankful to Ohio University for providing the OhioT1DM dataset. It is helpful in the performance evaluation of the proposed architecture.

Funding

No external funding was used during this research.

Data Availability

Not applicable.

Author Contribution

Mr. Ketan Lad is responsible for the conceptualization, analysis, manuscript writing, and programming, while Dr. Maulin Joshi oversaw the supervision and manuscript editing.

Declarations

Ethical Approval

This research does not involve human or animal participants; therefore, ethics approval was not required.

Consent for Publication Participants.

Consent for publication was given by all participants

Competing Interests

The authors declare that they have no competing interests or conflicts of interest to disclose.

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

References

[1] "IDF Diabetes atlas 11th Edition," 2025.

[2] E. A. James, A. V. Joglekar, A. K. Linnemann, H. A. Russ, and S. C. Kent, "The beta cell-immune cell interface in type 1 diabetes (T1D)," Dec. 01, 2023, *Elsevier GmbH*. doi: 10.1016/j.molmet.2023.101809.

[3] S. A. Antar *et al.*, "Diabetes mellitus: Classification, mediators, and complications; A gate to identify potential targets for the development of new effective treatments," Dec. 01, 2023, *Elsevier Masson s.r.l.* doi: 10.1016/j.biopha.2023.115734.

[4] M. J. Schoelwer *et al.*, "Predictors of Time-in-Range (70-180 mg/dL) Achieved Using a Closed-Loop Control System," *Diabetes Technol Ther*, vol. 23, no. 7, pp. 475–481, Jul. 2021, doi: 10.1089/dia.2020.0646.

[5] C. Toffanin, L. Magni, and C. Cobelli, "Artificial Pancreas: In Silico Study Shows No Need of Meal Announcement and Improved Time in Range of Glucose with Intraperitoneal vs. Subcutaneous Insulin Delivery," *IEEE Trans Med Robot Bionics*, vol. 3, no. 2, pp. 306–314, May 2021, doi: 10.1109/TMRB.2021.3075775.

[6] M. N. Pelkey, M. E. Boyle, A. Long, J. C. Castro, C. B. Cook, and B. Thompson, "Hybrid Closed-Loop Insulin Pump Technology Can Be Safely Used in the Inpatient Setting," *Endocrine Practice*, vol. 29, no. 1, pp. 24–28, Jan. 2023, doi: 10.1016/j.eprac.2022.11.006.

[7] M. Nwokolo and R. Hovorka, "The Artificial Pancreas and Type 1 Diabetes," *J Clin Endocrinol Metab*, vol. 108, no. 7, pp. 1614–1623, Jun. 2023, doi: 10.1210/clinem/dgad068.

[8] D. Pollakova, A. Lauria Pantano, U. Di Folco, M. R. Nardone, and C. Tubili, "Advanced hybrid closed loop (artificial pancreas) and carbohydrate count in type 1 diabetes," *Med J Nutrition Metab*, vol. 17, no. 3, pp. 219–227, Sep. 2024, doi: 10.3233/MNM-230125.

[9] B. M. Ahmed, M. E. Ali, M. M. Masud, M. R. Azad, and M. Naznin, "After-meal blood glucose level prediction for type-2 diabetic patients," *Heliyon*, vol. 10, no. 7, Apr. 2024, doi: 10.1016/j.heliyon. 2024.e28855.

[10] M. Zhang, K. B. Flores, and H. T. Tran, "Deep learning and regression approaches to forecasting blood glucose levels for type 1 diabetes," *Biomed Signal Process Control*, vol. 69, Aug. 2021, doi: 10.1016/j.bspc.2021.102923.

[11] C. Ji, T. Jiang, L. Liu, J. Zhang, and L. You, "Continuous glucose monitoring combined with artificial intelligence: redefining the pathway for prediabetes management," *Front Endocrinol (Lausanne)*, vol. 16, 2025, doi: 10.3389/fendo.2025.1571362.

[12] S.-M. Lee, D.-Y. Kim, and J. Woo, "Glucose Transformer: Forecasting Glucose Level and Events of Hyperglycemia and Hypoglycemia," *IEEE J Biomed Health Inform*, vol. 27, no. 3, pp. 1600–1611, Mar. 2023, doi: 10.1109/JBHI.2023.3236822.

[13] J. Alvarado, J. M. Velasco, F. Chavez, F. Fernández-de-Vega, and J. I. Hidalgo, "Combining wavelet transform with convolutional neural networks for hypoglycemia events prediction from CGM data," *Chemometrics and Intelligent Laboratory Systems*, vol. 243, Dec. 2023, doi: 10.1016/j.chemolab.2023.105017.

[14] G. Alfian *et al.*, "Blood glucose prediction model for type 1 diabetes based on artificial neural network with time-domain features," *Biocybern Biomed Eng*, vol. 40, no. 4, pp. 1586–1599, Oct. 2020, doi: 10.1016/j.bbe.2020.10.004.

[15] W. Seo, S.-W. Park, N. Kim, S.-M. Jin, and S.-M. Park, "A personalized blood glucose level prediction model with a fine-tuning strategy: A proof-of-concept study," *Comput Methods Programs Biomed*, vol. 211, p. 106424, Nov. 2021, doi: 10.1016/j.cmpb.2021.106424.

[16] H. V. Dudukcu, M. Taskiran, and T. Yildirim, "Blood glucose prediction with deep neural networks using weighted decision level fusion," *Biocybern Biomed Eng*, vol. 41, no. 3, pp. 1208–1223, Jul. 2021, doi: 10.1016/j.bbe.2021.08.007.

[17] D. Kalita, H. Sharma, J. K. Panda, and K. B. Mirza, "Platform for precise, personalised glucose forecasting through continuous glucose and physical activity monitoring and deep learning," *Med Eng Phys*, vol. 132, p. 104241, Oct. 2024, doi: 10.1016/j.medengphy.2024.104241.

[18] J. Shao *et al.*, "Generalization of a Deep Learning Model for Continuous Glucose Monitoring–Based Hypoglycemia Prediction: Algorithm Development and Validation Study," *JMIR Med Inform*, vol. 12, pp. e56909–e56909, May 2024, doi: 10.2196/56909.

[19] U. Krishnamoorthy, V. Karthika, M. K. Mathumitha, H. Panchal, V. K. S. Jatti, and A. Kumar, "Learned prediction of cholesterol and glucose using ARIMA and LSTM models – A comparison," *Results in Control and Optimization*, vol. 14, Mar. 2024, doi: 10.1016/j.rico.2023.100362.

[20] Li Kezhi and Liu C., T. Zhu, P. Herrero, and P. Georgiou, "GluNet: A Deep Learning Framework for Accurate Glucose Forecasting,"

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

IEEE J Biomed Health Inform, vol. 24, no. 2, pp. 414–423, Feb. 2020, doi: 10.1109/JBHI.2019.2931842.

[21] P. Domanski, A. Ray, K. Lafata, F. Firouzi, K. Chakrabarty, and D. Pflüger, "Advancing blood glucose prediction with neural architecture search and deep reinforcement learning for type 1 diabetics," *Biocybern Biomed Eng*, vol. 44, no. 3, pp. 481–500, Jul. 2024, doi: 10.1016/j.bbe.2024.07.006.

[22] M. De Bois, M. A. El Yacoubi, and M. Ammi, "GLYFE: review and benchmark of personalized glucose predictive models in type 1 diabetes," *Med Biol Eng Comput*, vol. 60, no. 1, pp. 1–17, Jan. 2022, doi: 10.1007/s11517-021-02437-4.

[23] X. Huang *et al.*, "Comparison of feature learning methods for non-invasive interstitial glucose prediction using wearable sensors in healthy cohorts: a pilot study," *Intelligent Medicine*, vol. 4, no. 4, pp. 226–238, Nov. 2024, doi: 10.1016/j.imed.2024.05.002.

[24] W. Wang, M. Tong, and M. Yu, "Blood Glucose Prediction with VMD and LSTM Optimized by Improved Particle Swarm Optimization," *IEEE Access*, vol. 8, pp. 217908–217916, 2020, doi: 10.1109/ACCESS.2020.3041355.

[25] V. B. Liu, L. Y. Sue, O. M. Padilla, and Y. Wu, "Optimizing blood glucose predictions in type 1 diabetes patients using a stacking ensemble approach," *Endocrine and Metabolic Science*, vol. 18, Jun. 2025, doi: 10.1016/j.endmts.2025.100253.

[26] Y. Shen and S. Kleinberg, "Personalized Blood Glucose Forecasting from Limited CGM Data Using Incrementally Retrained LSTM," *IEEE Trans Biomed Eng*, vol. 72, no. 4, pp. 1266–1277, Apr. 2025, doi: 10.1109/TBME.2024.3494732.

[27] H. Nemat, H. Khadem, J. Elliott, and M. Benaissa, "Data-driven blood glucose level prediction in type 1 diabetes: a comprehensive comparative analysis," *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-70277-x.

[28] A. Neumann, Y. Zghal, M. A. Cremona, A. Hajji, M. Morin, and M. Rekik, "A data-driven personalized approach to predict blood glucose levels in type-1 diabetes patients exercising in free-living conditions," *Comput Biol Med*, vol. 190, May 2025, doi: 10.1016/j.compbiomed.2025.110015.

[29] M. He, W. Gu, Y. Kong, L. Zhang, C. J. Spanos, and K. M. Mosalam, "CausalBG: Causal Recurrent Neural Network for the Blood Glucose Inference with IoT Platform," *IEEE Internet Things J*, vol. 7, no. 1, pp. 598–610, 2020, doi: 10.1109/JIOT.2019.2946693.

[30] R. Giancotti *et al.*, "Forecasting glucose values for patients with type 1 diabetes using heart rate data," *Comput Methods Programs Biomed*, vol. 257, Dec. 2024, doi: 10.1016/j.cmpb.2024.108438.

[31] G. Cappon, F. Prendin, A. Facchinetti, G. Sparacino, and S. Del Favero, "Individualized Models for Glucose Prediction in Type 1 Diabetes: Comparing Black-Box Approaches to a Physiological White-Box One," *IEEE Trans Biomed Eng*, vol. 70, no. 11, pp. 3105–3115, Nov. 2023, doi: 10.1109/TBME.2023.3276193.

[32] H. Butt, I. Khosa, and M. A. Iftikhar, "Feature Transformation for Efficient Blood Glucose Prediction in Type 1 Diabetes Mellitus Patients," *Diagnostics*, vol. 13, no. 3, p. 340, Jan. 2023, doi: 10.3390/diagnostics13030340.

[33] F. J. Lara-Abelenda, D. Chushig-Muzo, P. Peiro-Corbacho, A. M. Wägner, C. Granja, and C. Soguero-Ruiz, "Personalized glucose forecasting for people with type 1 diabetes using large language models," *Comput Methods Programs Biomed*, vol. 265, Jun. 2025, doi: 10.1016/j.cmpb.2025.108737.

[34] M. A. Karagoz, M. D. Breton, and A. El Fathi, "A Comparative study of transformer-based models for multi-horizon blood glucose prediction," in *IFAC-PapersOnLine*, Elsevier B.V., May 2025, pp. 83–88. doi: 10.1016/j.ifacol.2025.06.027.

[35] S. J. Kim, J. S. Moon, and S. Y. Jung, "Long-term blood glucose prediction using deep learning-based noise reduction," *ICT Express*, Aug. 2025, doi: 10.1016/j.icte.2025.05.009.

[36] T. Koutny and M. Mayo, "Predicting glucose level with an adapted branch predictor," *Comput Biol Med*, vol. 145, Jun. 2022, doi: 10.1016/j.compbiomed.2022.105388.

[37] M. F. Rabby, Y. Tu, M. I. Hossen, I. Lee, A. S. Maida, and X. Hei, "Stacked LSTM based deep recurrent neural network with kalman smoothing for blood glucose prediction," *BMC Med Inform Decis Mak*, vol. 21, no. 1, Dec. 2021, doi: 10.1186/s12911-021-01462-5.

[38] A. Della Cioppa, I. De Falco, T. Koutny, U. Scafuri, M. Ubl, and E. Tarantino, "Reducing high-risk glucose forecasting errors by evolving interpretable models for Type 1 diabetes," *Appl Soft Comput*, vol. 134, Feb. 2023, doi: 10.1016/j.asoc.2023.110012.

[39] K. Li, J. Daniels, C. Liu, P. Herrero, and P. Georgiou, "Convolutional Recurrent Neural Networks for Glucose Prediction," *IEEE J Biomed Health Inform*, vol. 24, no. 2, pp. 603–613, 2020, doi: 10.1109/JBHI.2019.2908488.

[40] M. Jaloli and M. Cescon, "Long-Term Prediction of Blood Glucose Levels in Type 1 Diabetes

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245

Using a CNN-LSTM-Based Deep Neural Network," *J Diabetes Sci Technol*, vol. 17, no. 6, pp. 1590–1601, Nov. 2023, doi: 10.1177/19322968221092785.

[41] R. I. Alkanhel *et al.*, "Hybrid CNN-GRU Model for Real-Time Blood Glucose Forecasting: Enhancing IoT-Based Diabetes Management with AI," *Sensors*, vol. 24, no. 23, Dec. 2024, doi: 10.3390/s24237670.

[42] X. Xiong, X. L. Yang, Y. Cai, Y. Xue, J. F. He, and H. Su, "Exploring the potential of deep learning models integrating transformer and LSTM in predicting blood glucose levels for T1D patients," *Digit Health*, vol. 11, Jan. 2025, doi: 10.1177/20552076251328980.

[43] D. M. Chiara, F. Micheletto, D. Lv, M. Breton, B. Kovatchev, and C. Cobelli, "The UVA/PADOVA type 1 diabetes simulator: New features," *J Diabetes Sci Technol*, vol. 8, no. 1, pp. 26–34, 2014, doi: 10.1177/1932296813514502.

[44] R. Visentin *et al.*, "The UVA/Padova Type 1 Diabetes Simulator Goes from Single Meal to Single Day," *J Diabetes Sci Technol*, vol. 12, no. 2, pp. 273–281, Mar. 2018, doi: 10.1177/1932296818757747.

[45] M. Cindy and B. Razvan, "The ohioT1DM dataset for blood glucose level prediction: Update 2020," *CEUR Workshop Proc*, vol. 2675, pp. 71–74, 2020, [Online]. Available: http://smarthealth.cs.ohio.edu/bglp/OhioT1DM-dataset-paper.pdf

[46] A. K. Singh and D. Koundal, "A Temporal Convolutional Network for modeling raw 3D sequences and air-writing recognition," *Decision Analytics Journal*, vol. 10, Mar. 2024, doi: 10.1016/j.dajour.2023.100373.

[47] A. Zaitcev, M. R. Eissa, Z. Hui, T. Good, J. Elliott, and M. Benaissa, "A Deep Neural Network Application for Improved Prediction of HbA 1c in Type 1 Diabetes," *IEEE J Biomed Health Inform*, vol. 24, no. 10, pp. 2932–2941, 2020, doi: 10.1109/JBHI.2020.2967546.

[48] C. Xu, X. Wang, and S. Zhang, "Dilated convolution capsule network for apple leaf disease identification," *Front Plant Sci*, vol. 13, Nov. 2022, doi: 10.3389/fpls.2022.1002312.

[49] A. Ghosh, A. Sufian, F. Sultana, A. Chakrabarti, and D. De, "Fundamental concepts of convolutional neural network," in *Recent Trends and Advances in Artificial Intelligence and Internet of Things. Intelligent Systems Reference Library*, vol. 172, Springer, Cham, 2020, pp. 519–567. doi: 10.1007/978-3-030-32644-9_36.

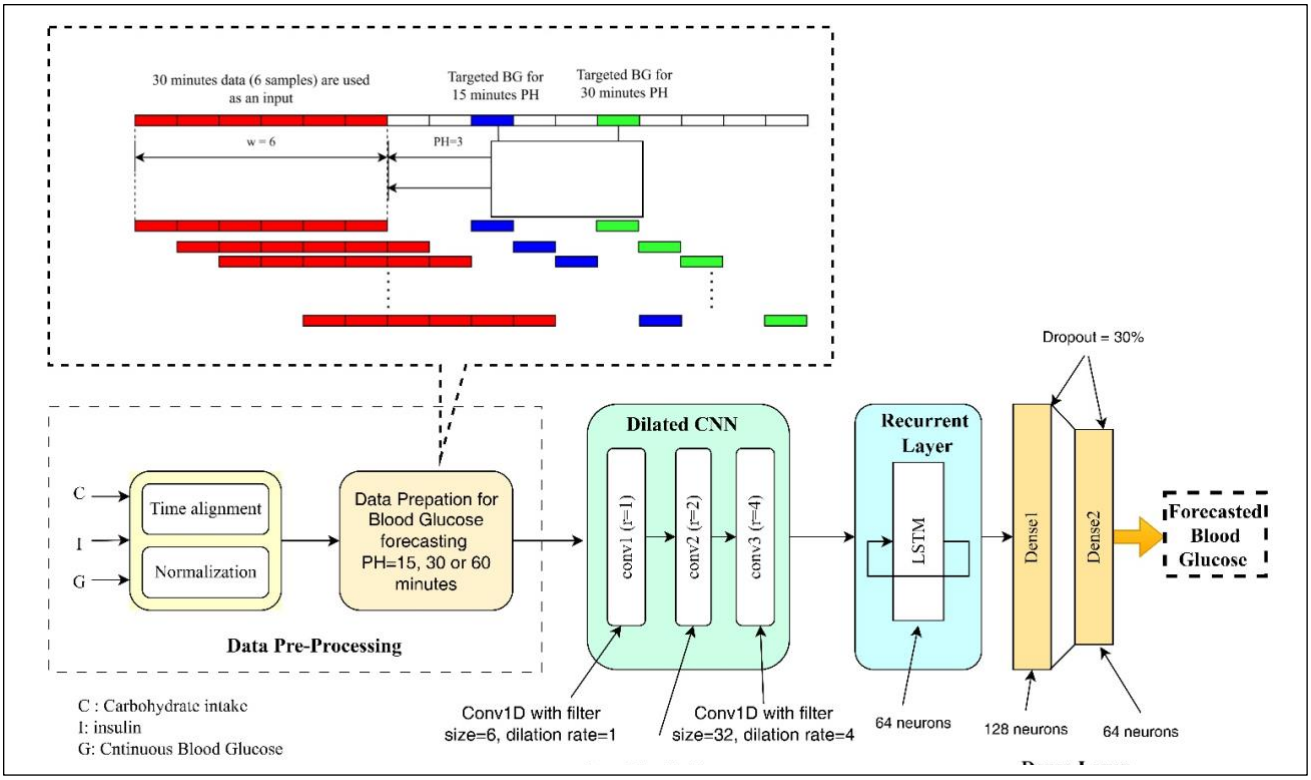
[50] H. V. Dudukcu, M. Taskiran, and T. Yildirim, "Consolidated or individual training: Which one is better for blood glucose prediction?" in *2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, IEEE, Aug. 2021, pp. 1–6. doi: 10.1109/INISTA52262.2021.9548612.

Author Biography

Mr. Ketan K. Lad earned a Bachelor of Engineering degree in Electronics and Communication Engineering from Bhavnagar University in Gujarat, India, in 2008. He then completed his M.Eng. degree in the same field at Gujarat Technological University in 2012. Currently, he is a PhD scholar at Gujarat Technological University. Additionally, he has been working as an Assistant Professor in the Electronics and Communication Department at L.D. College of Engineering in Ahmedabad, Gujarat, for the past two years. With a total of thirteen years of teaching experience, his research focuses on artificial intelligence, machine learning, embedded system design, and biomedical applications within electronics and communication engineering.

Dr. Maulin Joshi is a distinguished academic and faculty member at Sarvajanik College of Engineering and Technology (SCET), known for his expertise in Electronics and Communication Engineering. With over 26 years of experience in academia, he holds the position of professor in the department and serves as the Dean of Academics. Dr. Joshi's academic qualifications include a Ph.D., and he specializes in teaching subjects such as Signal Processing, Mobile Communication, Wireless Communication, and Soft Computing Techniques. In addition to his teaching responsibilities, Dr. Joshi has contributed significantly to research, published numerous papers, and delivered expert lectures in his field. His involvement in administrative duties and active participation in professional bodies demonstrate his commitment to academic and institutional excellence.

Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245



Manuscript Received 20 September 2025; Revised 25 November 2025; Accepted 5 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1245