2023 IEEE International Conference on Digital Health (ICDH) | 979-8-3503-4103-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICDH60066.2023.00042

Blood Glucose Prediction for Type-1 Diabetics using Deep Reinforcement Learning

Peter Domanski*[1], Aritra Ray*[2], Farshad Firouzi[2], Kyle Lafata[2], Krishnendu Chakrabarty[3], and Dirk Pflüger[1]

[1]University of Stuttgart, *Stuttgart, Germany*
[2]Duke University, *NC, USA*
[3]Arizona State University, *AZ, USA*
[1]{*peter.domanski, dirk.pflueger*}*@ipvs.uni-stuttgart.de*
[2]{*aritra.ray, farshad.firouzi, kyle.lafata*}*@duke.edu*
[3]{*krishnendu.chakrabarty*}*@asu.edu*

***Abstract*—An accurate prediction of blood glucose levels for individuals affected with type-1 diabetes mellitus helps to regulate blood glucose through specific insulin delivery. In our work, we propose the design of a densely-connected encoder-decoder network in conjunction with Long-Short Term Memory networks. We formulate the blood glucose prediction as a deep reinforcement learning problem and evaluate our results on the OhioT1DM dataset. The OhioT1DM dataset contains blood glucose monitoring records in intervals of 5 minutes over 8 weeks for 12 patients affected with type-1 diabetes mellitus. Prior works aim to predict the blood glucose levels in prediction horizons of 30 and 45 minutes, corresponding to 6 and 9 data points, respectively. Compared to prior work with the best prediction accuracy so far with respect to the mean absolute error, we improve by 18.4% and 22.5% in 30-minute and 45-minute prediction horizons, respectively. Furthermore, for risk assessment in our predictions, we visualize the error and evaluate clinical risk through a surveillance error grid approach.**

***Index Terms*—Deep Reinforcement Learning, Long-Short Term Memory, Blood Glucose Predictions, Type-1 Diabetes**

I. Introduction

Diabetes mellitus is a chronic medical condition characterized by high glucose levels in the blood [1]. Without clinical interventions, individuals with high uncontrolled levels of blood glucose (BG) develop serious health complications, increasing the probability of heart disease, diabetic retinopathy, or kidney damage [2]. According to the International Diabetes Federation (IDF) [3], it is estimated that 537 million adults between the age of 20 to 79 years are affected with diabetes. Moreover, in 2021 alone, the IDF estimates that diabetes was responsible for 6.7 million deaths. The variants include type 1, type 2, gestational, pre-diabetes, monogenic, and cystic fibrosis-related diabetes [4]. In our paper, we focus on type 1 diabetes, an organ-specific autoimmune disease of pancreatic β cells [1], wherein the β cells eventually lose the ability of insulin synthesis.

In our paper, we propose the design of a Long-Short Term Memory Network (LSTM) encoding network in conjunction with Deep Reinforcement Learning (DRL). The input encoding network consists of densely-connected layers which exploit the representational power to extract informative features for the subsequent LSTM network, followed by an output encoding network. The output of the encoding network, intertwined with LSTM, is fed into a projection layer which generates a probability distribution across prediction values with a non-zero mean and bounded output to promote numerical stability. We formulate the BG prediction as a DRL learning task and integrate the Soft Actor-Critic (SAC) DRL algorithm in our study. We evaluate our proposed method on the OhioT1DM dataset [5], which contains CGM of 12 patients affected with type-1 diabetes mellitus, in intervals of 5 minutes, for 8 weeks.

We compare our results to prior work by Hatice et al. [6], which achieved the best results using Deep Neural Networks (DNN) with a weighted decision-level fusion of LSTM, GRU, and Wavelets on the OhioT1DM dataset. Hatice et al. defined the Prediction Horizon (PH) as the consecutive number of BG levels to be predicted. A 45-minute PH involves 9 consecutive prediction BG levels (as the OhioT1DM dataset considers a 5-minute interval). Our proposed method significantly outperforms the current best BG prediction approach with respect to both Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) in the 30-minute and 45-minute PH. The key contributions of our work are as follows:

- We propose the design of a densely-connected encoding-decoding network in conjunction with an LSTM model. In contrast to other works, we formulate the BG prediction task as a DRL problem and train our network using the SAC method.
- We train our proposed method on the OhioT1DM dataset. In comparison to the baseline, our proposed method outperforms their best results by 18.5% on MAE and 16.3% on RMSE for 30-minute PH, and by 22.5% MAE and 19.5% on RMSE for 45-minute PH.

*Equal Contribution

This research was supported by Advantest as part of the Graduate School "Intelligent Methods for Test and Reliability" (GS-IMTR) at the University of Stuttgart

979-8-3503-4103-4/23/$31.00 ©2023 IEEE
DOI 10.1109/ICDH60066.2023.00042

The rest of the paper is organized as follows: In Section III, we discuss our proposed methodology involving details of the LSTM encoding network and DRL algorithm. We delve into our experimental setup in Section IV, followed by the results in Section V. Section VI concludes our paper. The implementation of our method is publicly available on GitHub [7].

II. Related Work

Accurately predicting BG levels for individuals affected with type 1 diabetes plays an important role in informed insulin injection. Various external factors like carbohydrate intake, physical exhaustion induced by exercising, or stress contribute to the variation of BG levels in type-1 diabetics. Traditional Machine Learning (ML) approaches like random forests [8], Autoregressive Integrated Moving Average (ARIMA) [9], [10], to name a few, have been explored in the space of continuous glucose monitoring (CGM) and prediction. Deep learning-based popular approaches include LSTM [10], Gated Recurrent Units (GRU) [6], Convolutional Neural Networks (CNN) [11], and Convolutional Recurrent Neural Networks (CRNN) [12]. Variants of the DRL algorithm have also been used for Type-1 diabetes prediction [13]. In recent years, the algorithms have been popularly evaluated on simulator environments like Diabetes Mellitus Metabolic Simulator for Research (DMMS.R) [14], Padova T1D Simulator [15], Web-based Simulation Tool [16], Simglucose [17], GluCoEnv [18], to name a few. Some works also include evaluations on data collected from subjects affected with Type-1 diabetes [19]–[21].

In regards to blood glucose monitoring for Type-1 diabetics on the OhioT1DM [5] dataset, several methodologies have been proposed. Harry et. al. [22] proposed additional supervision on top of a neural network architecture, stacked with fully connected layers and residual backcasting, primarily by replacing the fully connected block structure with a recurrent neural network. Hadia et. al. [23] investigated themes of knowledge distillation and transfer learning over a student-teacher approach across the patients of OhioT1DM dataset. Other approaches include generative adversarial networks [24], multi-task networks [25], and sequence-to-sequence neural networks [26]. The one we consider as our baseline involves weighted decision level fusion over LSTM, WaveNet, and GRU [6].

III. Methodology

In recent years, approaches to blood glucose prediction increasingly use data-driven methods. Especially since 2017, deep learning-based approaches have gained attention and seem to be the most successful methods (see, e.g., [27]–[29]). Similar to many other studies, we use a patient's past BG values as input to predict future BG values. In training, we use 30 min of a patient's BG history to forecast 5, 30, and 45 min prediction horizons. After the training, we evaluate the models on unseen testing data. Fig. 1 shows an overview of the proposed reinforcement learning-based model. Detailed information about the network architecture is given in the subsequent sections.

A. Deep neural networks

Previous studies revealed deep neural networks as the most promising models for blood glucose prediction. Many different networks have been proposed, for example, convolutional neural networks. In this study, we use recurrent neural networks (RNN) and LSTM. In contrast to other works, we propose a DRL approach instead of supervised learning and adjust the architecture and training process to the requirements of our proposed DRL.

1) Densely-connected networks: A type of neural network where each neuron in one layer is connected to every neuron in the next layer. Among several others, the high representational power, flexibility, and effective feature learning (multiple levels of abstraction) are reasons to use densely-connected networks. In medical image analysis, they are successfully used with convolutional networks, e.g., in DenseNet [30].

Encoding networks: The input encoding network is a three-layer densely-connected neural network (consisting of 256, 512, and 256 neurons per layer with ReLU activation) that exploits the representational power to extract informative features for the subsequent LSTM network. Similarly, the output encoding networks extract features of the LSTM for the final blood glucose prediction of the projection layer.

Projection layer: Performs the blood glucose value prediction. The size of the (densely connected) layer matches the number of time steps in the prediction horizon. The projection layer uses a tanh normal projection to transform the output and generate a normal distribution with a mean and standard deviation which is necessary to represent the probability distribution across different prediction values (so-called actions) in DRL, see Section III-B. We use tanh normal projection to efficiently get a distribution with a non-zero mean, which helps to ensure that the output is bounded and increases numerical stability.

2) LSTM: A type of RNN explicitly designed for processing sequential data, e.g., time-series data. In addition to the input data, RNNs have a feedback mechanism (so-called memory) to use previous network outputs to make the current prediction. The structure of an LSTM in Fig. 1 shows multiple gates (input, forget, and output gate) that control the flow of information through time in the memory [31]. The LSTM cell processes the input x_t at time step t, the previous cell memory c_{t-1}, and the previous output h_{t-1} of the cell to generate the outputs c_t respectively h_t. The input gate in (1) and the forget gate in (2) control the cell state c_t, see (5). Together with the result of the output gate in (3), the hidden state h_t of the cell is given in (6).

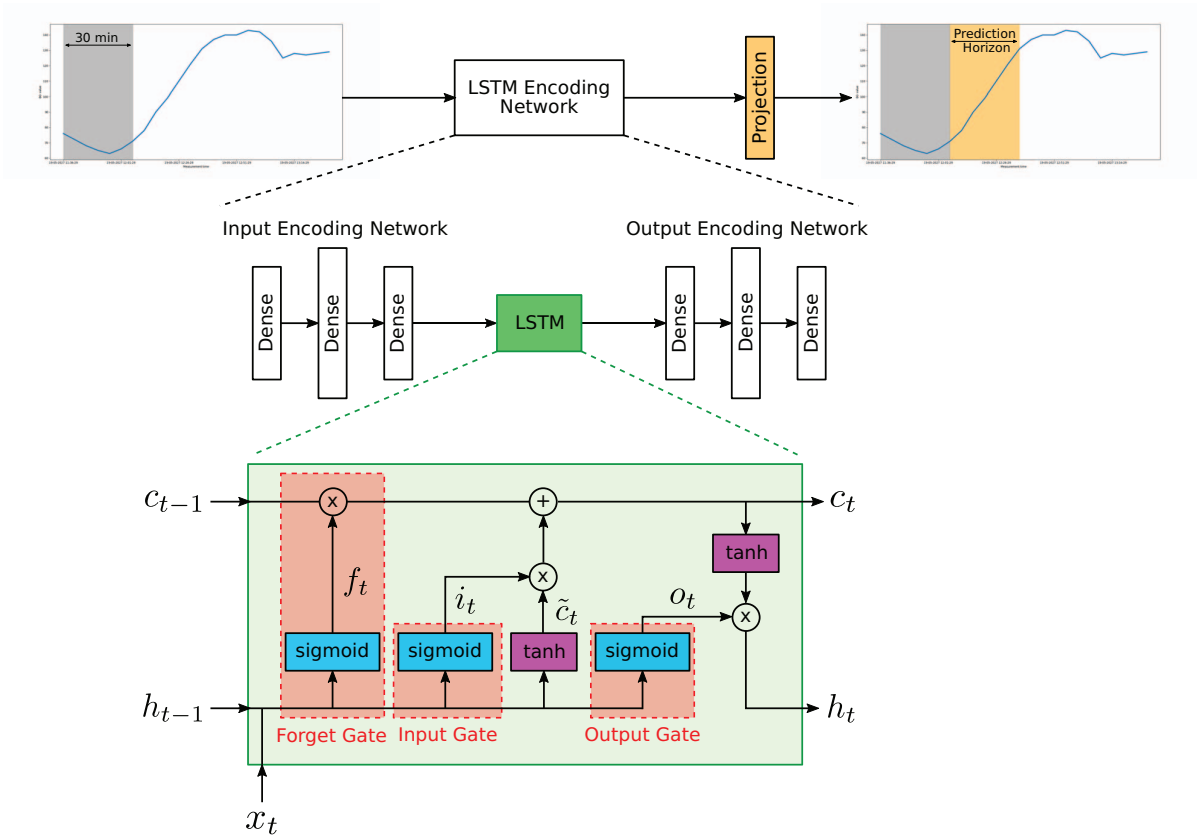

Fig. 1: Overview of the proposed network architecture including encoder networks, LSTM, and the final projection layer.

$$i_t = \text{sigmoid}\left(W_i x_t + U_i h_{t-1} + b_i\right) \quad (1)$$
$$f_t = \text{sigmoid}\left(W_f x_t + U_f h_{t-1} + b_f\right) \quad (2)$$
$$o_t = \text{sigmoid}\left(W_o x_t + U_o h_{t-1} + b_o\right) \quad (3)$$
$$\tilde{c}_t = \tanh\left(W_c x_t + U_t h_{t-1} + b_c\right) \quad (4)$$
$$c_t = f_t \odot c_{t-1} \odot i_t \odot \tilde{c}_{t-1} \quad (5)$$
$$h_t = o_t \odot \tanh\left(c_t\right) \quad (6)$$

These gates and the recurrent structure of the network allow us to address the vanishing gradients problem and to improve the learning of long-term dependencies [32]. In our proposed design, we use an LSTM with a cell size of 256.

B. Deep reinforcement learning

DRL combines the representational power of deep learning and the decision-making ability of reinforcement learning. DRL is a sub-field of machine learning for complex, high-dimensional decision-making tasks in which an agent learns an optimal behavior by taking action in its environment and receiving rewards or punishments with respect to the outcome of the actions [33]. DRL tasks can be described as Markov Decision Processes, which determine the space of environment states $\mathcal{S}$, the space of agent actions $\mathcal{A}$, and a scalar training signal $\mathcal{R}$ (reward). At each time step t, the agent gets a state $s_t \in \mathcal{S}$ and selects an action $a_t \in \mathcal{A}$ following a policy $\pi\left(a_t|s_t\right)$. Consequently, the agent receives a reward $r_t \in \mathcal{R}$ and the next state $s_{t+1} \in \mathcal{S}$. We parameterize the policy π with a deep neural network and maximize the cumulative reward (return) in training. Analog to the study in [34], we use Actor-Critic training algorithms. Actor-Critic algorithms use a so-called value function to predict future rewards. The value function in Eq. 7 denotes the expected total (discounted) reward starting from state s and γ the discount factor.

$$V(s) = \mathrm{E}_\pi\left[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s\right] \quad (7)$$

In the following, we focus on the Soft Actor-Critic (SAC) algorithm [35], which was the most successful in our study among the algorithms in [34].

Soft Actor-Critic: DRL algorithm for continuous control tasks following the Actor-Critic framework [36]. It uses two networks, the actor and the critic network, to learn the policy respectively to estimate a soft value function similar to Eq. 7, which is used to update the policy using Temporal Difference (TD) error. SAC's soft value function updates and entropy regularization improve exploration and prevent the policy from becoming too deterministic. In addition, the automatic temperature tuning simplifies the hyperparameter

tuning process, increasing stability and robustness in training [35]. However, training will require more computational effort, possibly increasing the training time compared to common supervised approaches.

In the following experiments, we use an actor and a critic network with a structure shown in Fig. 1. To formulate the BG prediction as a DRL task, we define the environment state at time t as a vector $[BG_{t-5}, BG_{t-4}, BG_{t-3}, BG_{t-2}, BG_{t-1}, BG_t]^T$, the reward as in Eq. 8, and actions as continuous BG values in the range from 35 to 500 mg/dl for a given prediction horizon (single-step ahead or multi-step ahead). During training, the agent outputs continuous BG values from 35 to 500, given the observed state of the environment.

$$R_{t+1} = -|BG_t - a_t| \tag{8}$$

The reward definition is such that the reward is close to zero if the output action a_t is close to the real BG value B_t and far away from zero (in the negative direction) otherwise. Formulating the BG prediction task as a decision-making task allows us to apply different DRL techniques.

IV. Experimental setup

A. Dataset

In this paper, we use the OhioT1DM dataset [5], widely used in the research of BG level prediction. The most recent release for the second Blood Glucose Level Prediction (BGLP) Challenge (2020) contains data from 12 subjects with type 1 diabetes on insulin pump therapy. It includes continuous glucose monitoring (CGM), insulin, physiological sensor, and self-reported life-event data throughout an eight-week data collection period for each of the 12 people. Throughout the eight-week period, they wore Medtronic 530G or 630G insulin pumps and used Medtronic Enlinte CGM sensors for data collection. The data is entirely anonymized, e.g., by randomly assigning ID numbers for each patient. The dataset includes 19 different features, especially the CGM blood glucose level given in 5 minutes intervals. The data is from 7 male and 5 female subjects aged from 20 to 80 years. Detailed information on the dataset is given in Table I.

B. Preprocessing

We normalize the data using z-score standardization widely used in ML algorithms. It scales the values of each feature in the data to have zero mean and unit variance. The feature-wise calculation of the z-score normalization is given in Eq. 9 and requires determining the distribution mean μ_x and standard deviation σ_x for each feature x.

$$x' = \frac{x - \mu_x}{\sigma_x} \tag{9}$$

The intuition using z-score normalization is that average BG values correlate with A1C and other relevant characteristics, e.g., hyperglycemia. Moreover, the normalization focuses on potentially dangerous changes in BG values and neglects statistics in the data that do not reveal information for accurate CGM [37].

C. Hyperparameter selection

Applying deep neural networks successfully to a given problem requires a suitable selection of hyperparameters that control the structure and properties of architectures and algorithms. In this work, we use random search to determine the most successful hyperparameters with respect to achievable RMSE values [38]. The sets of hyperparameter values were chosen in alignment with commonly used values in the literature.

D. Evaluation metrics

Evaluation metrics are measures to quantify the accuracy of a system. Many metrics exist to evaluate the accuracy of BG prediction tasks. In the following, we introduce the most common metrics to evaluate analytical and clinical accuracy. Evaluation of analytical accuracy includes quantitative methods for describing how closely the predictions match the ground truth measurements. In contrast, clinical accuracy is a qualitative measure to evaluate the clinical outcome of different treatment decisions. Therefore, the definition includes statistical metrics and the expert knowledge of clinicians.

Analytical accuracy: The most commonly used numerical metric is the Root Mean Squared Error (RMSE) given in Eq. 10. In addition, we use the Mean Absolute Error (MAE) in Eq. 11 to evaluate the analytical BG prediction accuracy in this study.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{\mathrm{predicted}} - y_{\mathrm{measured}}\right)^2} \tag{10}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_{\mathrm{predicted}} - y_{\mathrm{measured}}\right| \tag{11}$$

Clinical accuracy: A surveillance Error Grid (SEG) is the most recently used grid-based visualization method to evaluate clinical accuracy. It is a metric for error- and (clinical) risk assessment of BG measurements [39]. In general, the error grid shows a set of risk zones with scores (risk levels) for the clinical impact ranging from 0 (none) to 4 (extreme). Fig. 2 shows a simplified, discrete SEG with limits from 0 to 600 ml/dl and risk zones separated by 120 mg/dl intervals. The error grid on which our predictions are superimposed in Fig. 4 is continuously color-coded. The color represents the average risk rating of clinician respondents of a survey [39]. It represents the mean group decision in accordance with the consensus of experts. Consequently, the SEG is not symmetrical with respect to the identity line.

V. Results

We train our proposed DRL algorithm on the OhioT1DM dataset [5]. The OhioT1DM dataset contains continuous blood glucose level monitoring in intervals of 5 minutes for 12 patients. Our models are trained on the blood glucose recordings of each patient available in the OhioT1DM dataset in an 80-20

TABLE I: Description of the OhioT1DM dataset properties [5].

Patient ID	Gender	Age Range	Pump Model	Sensor Band	Training Samples	Test Samples
540	male	20-40	630G	Empetica	11947	2884
544	male	40-60	630G	Empetica	10623	2704
552	male	20-40	630G	Empetica	9080	2352
567	female	20-40	630G	Empetica	10858	2377
584	male	40-60	530G	Empetica	12150	2653
596	male	60-80	530G	Empetica	10877	2731
559	female	40-60	530G	Basis	10796	2514
563	male	40-60	530G	Basis	12124	2570
570	male	40-60	530G	Basis	10982	2745
575	female	40-60	530G	Basis	11866	2590
588	female	40-60	530G	Basis	12640	2791
591	female	40-60	530G	Basis	1ß847	2760

TABLE II: Multi Step Prediction Results (SAC, Batch size=1024, PH=30 mins)

Patient ID	Root Mean Square Error			Mean Absolute Error		
	Our Proposed Method	Baseline [6] (LSTM + WaveNet + GRU)	% Improvement Over Baseline	Our Proposed Method	Baseline [6] (LSTM + WaveNet + GRU)	% Improvement Over Baseline
540	19.254	25.28	23.837	13.807	18.77	12.996
544	15.287	19.76	22.637	11.348	14.36	12.996
552	14.499	19.43	25.378	10.580	14.66	27.831
559	19.372	21.78	19.779	13.473	15.35	12.228
563	17.779	20.43	12.976	12.308	14.39	14.468
567	22.206	23.96	7.321	14.682	17.41	12.625
570	15.803	18.06	12.497	11.715	12.85	8.833
575	20.370	25.02	18.585	14.011	16.77	16.452
584	24.306	24.84	3.237	16.606	18.57	10.576
588	15.024	21.26	29.332	11.040	15.55	29.003
591	18.096	23.76	23.838	12.996	18.61	28.436
596	17.319	19.23	9.938	12.625	13.58	7.032

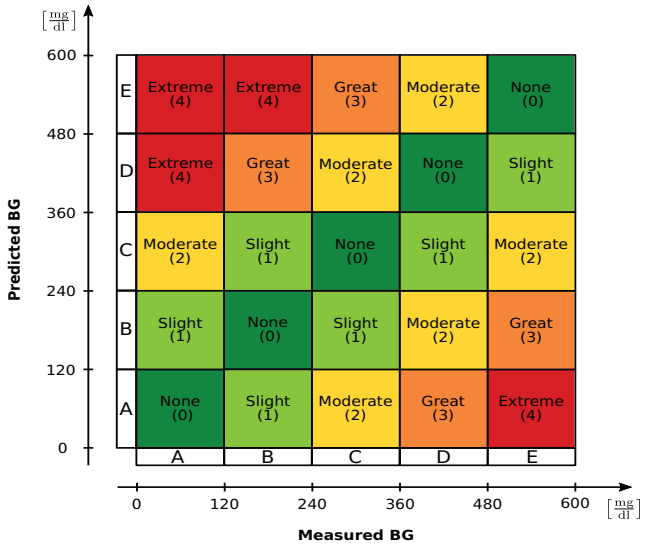

Fig. 2: Visualization of SEG risk zones.

TABLE III: Comparing Average in Multi-Step Prediction Results (SAC, Batch size=1024, PH=30 mins)

Method	Root Mean Square Error	Mean Absolute Error
LSTM [6]	22.13	16.02
Wavelet [6]	22.49	16.47
GRU [6]	22.00	15.91
WaveNet + LSTM [6]	22.35	16.29
WaveNet + GRU [6]	22.21	16.15
LSTM + GRU [6]	21.98	15.86
LSTM + WaveNet + GRU [6]	21.90	15.87
Deep RL (**Proposed Method**)	**18.32**	**12.93**

train-test split ratio. In the prediction phase, for every patient, we perform training on 30 minutes intervals (corresponding to 6 consecutive blood glucose level recordings) and perform prediction on either a single step (corresponding to the blood glucose level in the next 5 minutes) or multi-step prediction (corresponding to the blood glucose level in next 30 or 45 minutes). The evaluation results are visualized using SEG and compared to our baseline [6] with respect to RMSE and MAE. We discuss the prediction results for the entire dataset of 250 hours with regard to single-step and multi-step predictions as follows.

Multi-Step Predictions for 30 minutes: Our proposed SAC model predicts the blood glucose level for 6 consecutive steps (corresponding to the next 30 minutes), with a batch size of

TABLE IV: Multi Step Prediction Results (SAC, Batch size=1024, PH=45 mins)

Patient ID	Root Mean Square Error Our Proposed Method	Mean Absolute Error Our Proposed Method
540	24.995	18.025
544	20.527	14.849
552	19.826	14.426
559	24.833	17.606
563	20.888	14.356
567	27.873	18.492
570	19.322	14.158
575	25.151	17.249
584	31.512	22.687
588	20.452	14.701
591	23.084	16.382
596	22.723	16.945

TABLE V: Comparing Average in Multi-Step Prediction Results (SAC, Batch size=1024, PH=45 mins)

Method	Root Mean Square Error	Mean Absolute Error
LSTM [6]	12.93	21.61
Wavelet [6]	29.68	22.19
GRU [6]	29.22	21.50
WaveNet + LSTM [6]	29.46	21.87
WaveNet + GRU [6]	29.44	21.83
LSTM + GRU [6]	29.26	21.56
LSTM + WaveNet + GRU [6]	29.12	21.52
Deep RL (**Proposed Method**)	**23.43**	**16.66**

1024, and 50000 training steps while considering input data for 30 minutes. For all 12 patients, we evaluate the RMSE and MAE, as presented in Table II. Table III features the comparative results of our proposed method with respect to LSTM, Wavelet, GRU, and a possible combination of these methods. We observe that for all the patients, our proposed method can significantly improve over the baseline [6], with respect to both metrics. On average for all 12 patients, our proposed method improves RMSE by 16.34% and MAE by 18.4%. Our proposed method outperforms each of them.

Multi-Step Predictions for 45 minutes: Our proposed SAC model predicts the blood glucose level for 9 consecutive steps (corresponding to the next 45 minutes), with a batch size of 1024, and 50000 training steps while considering input data for 30 minutes. For all 12 patients, we evaluate the RMSE and MAE, as presented in Table IV. Table V features the comparative results of our proposed method with respect to LSTM, Wavelet, GRU, and a possible combination of these methods. We observe that for all the patients, our proposed method can significantly improve over the baseline [6], with respect to both metrics. On average for all 12 patients, our proposed method improves the RMSE by 19.53% and MAE by 22.5%. Our method outperforms each of them. To visualize how well our model fits the data, in Fig. 3, we present the ground truth in comparison to the prediction results for a representative sample of the testing dataset of 30 hours for patient ID 570 and 584.

TABLE VI: Single-Step Prediction Results (SAC, Batch size=64, PH=5 mins)

Patient ID	Root Mean Square Error	Mean Absolute Error
540	12.270	9.441
544	8.806	7.087
552	8.336	6.573
559	10.920	8.293
563	9.017	6.907
567	9.309	7.293
570	10.940	8.626
575	11.640	8.914
584	11.630	8.745
588	8.542	6.894
591	9.840	7.225
596	8.07	6.360
Average	9.943	7.697

Single Step Prediction for 5 minutes: Our proposed SAC model predicts the blood glucose level for one step (corresponding to the next 5 minutes), with a batch size of 64, and 50000 training steps while considering input data for 30 minutes. For all 12 patients, we evaluate the RMSE and MAE, as presented in Table VI. The average RMSE and MAE are 9.943 and 7.767. Owing to the lack of single-step prediction results in our baseline models, we refrain from any comparative results.

Surveillance Error Grid for Multi-Step Prediction: Through SEG, we assess the risk between predicted and ground truth blood glucose levels for all 12 patients. SEG helps us to estimate the percent predictions that lie within the risk zones, as defined by clinical practices. We present the average SEG in the multi-step prediction of 30 minutes and 45 minutes, for all 12 patients, in Table VII and Table VIII respectively. Tables IX and X present a comparison among all the prior works compared to our proposed method, for multi-step prediction of 30 minutes and 45 minutes respectively. We present the SEG visualization for the patients with the best and worst scores respectively, referenced by patient ID 570 and 584 using our proposed method, in Fig. 4. From the tables and the visualization of SEG as outlined in Fig. 4, we can conclude that a significant proportion of the prediction results lie in the no risk to slight risk zone, excluding a few outliers. Our prediction either outperforms or remains consistent with our baseline. This further bolsters the confidence in our prediction results.

VI. Conclusion

In our paper, we have proposed a novel methodology for time series modeling of BG levels in type-1 diabetes using a densely-connected encoder-decoder network and LSTM formulated as a DRL problem. We have evaluated our results for the OhioT1DM dataset benchmark. Compared to the prior

TABLE VII: Average SEG in Multi-Step Prediction Results (SAC, Batch size=1024, PH=30 mins)

Patient ID	None (0)	Slight (1)	Moderate (2)	Great (3)	Extreme (4)
540	85.104	14.896	0.000	0.000	0.000
544	90.815	9.185	0.000	0.000	0.000
552	89.541	10.459	0.000	0.000	0.000
559	89.928	9.952	0.120	0.000	0.000
563	89.085	10.915	0.000	0.000	0.000
567	86.869	12.963	0.168	0.000	0.000
570	90.367	9.560	0.073	0.000	0.000
575	89.161	10.801	0.039	0.000	0.000
584	85.332	14.517	0.075	0.075	0.000
588	91.361	8.531	0.108	0.000	0.000
591	89.592	10.408	0.000	0.000	0.000
596	89.011	10.989	0.000	0.000	0.000
Average	88.847	11.098	0.048	0.000	0.000

TABLE VIII: Average SEG in Multi-Step Prediction Results (SAC, Batch size=1024, PH=45 mins)

Patient ID	None (0)	Slight (1)	Moderate (2)	Great (3)	Extreme (4)
540	86.343	13.542	0.116	0.000	0.000
544	89.605	10.395	0.000	0.000	0.000
552	89.344	10.656	0.000	0.000	0.000
559	89.022	10.951	0.027	0.000	0.000
563	91.111	8.863	0.0261	0.000	0.000
567	85.035	14.515	0.366	0.084	0.000
570	92.015	7.790	0.195	0.000	0.000
575	86.817	13.157	0.0259	0.000	0.000
584	82.212	17.611	0.151	0.025	0.000
586	90.380	9.620	0.000	0.000	0.000
591	86.847	13.153	0.000	0.000	0.000
596	87.151	12.849	0.000	0.000	0.000

TABLE IX: Comparing Average SEG in Multi-Step Prediction Results (SAC, Batch size=1024, PH=30 mins)

Method	None (0)	Slight (1)	Moderate (2)	Great (3)	Extreme (4)
LSTM [6]	86.42	13.56	0.02	0.00	0.00
Wavelet [6]	85.91	14.06	0.03	0.00	0.00
GRU [6]	86.41	13.55	0.04	0.00	0.00
WaveNet + LSTM [6]	86.43	13.54	0.03	0.00	0.00
WaveNet + GRU [6]	86.39	13.57	0.04	0.00	0.00
LSTM + GRU [6]	86.52	13.44	0.04	0.00	0.00
LSTM + WaveNet+ GRU [6]	86.53	13.45	**0.02**	0.00	0.00
Deep RL (**Proposed Method**)	**88.85**	**11.10**	0.05	**0.00**	**0.00**

TABLE X: Comparing Average SEG in Multi-Step Prediction Results (SAC, Batch size=1024, PH=45 mins)

Method	None (0)	Slight (1)	Moderate (2)	Great (3)	Extreme (4)
LSTM [6]	80.95	18.98	0.07	0.00	0.00
Wavelet [6]	79.88	20.07	0.05	0.00	0.00
GRU [6]	81.08	18.85	0.07	0.00	0.00
WaveNet + LSTM [6]	81.04	18.90	0.06	0.00	0.00
WaveNet + GRU [6]	81.01	18.92	0.07	0.00	0.00
LSTM + GRU [6]	81.10	18.83	0.07	0.00	0.00
LSTM + WaveNet+ GRU [6]	81.14	18.80	**0.06**	**0.00**	0.00
Deep RL (**Proposed Method**)	**87.99**	**11.93**	0.08	0.01	**0.00**

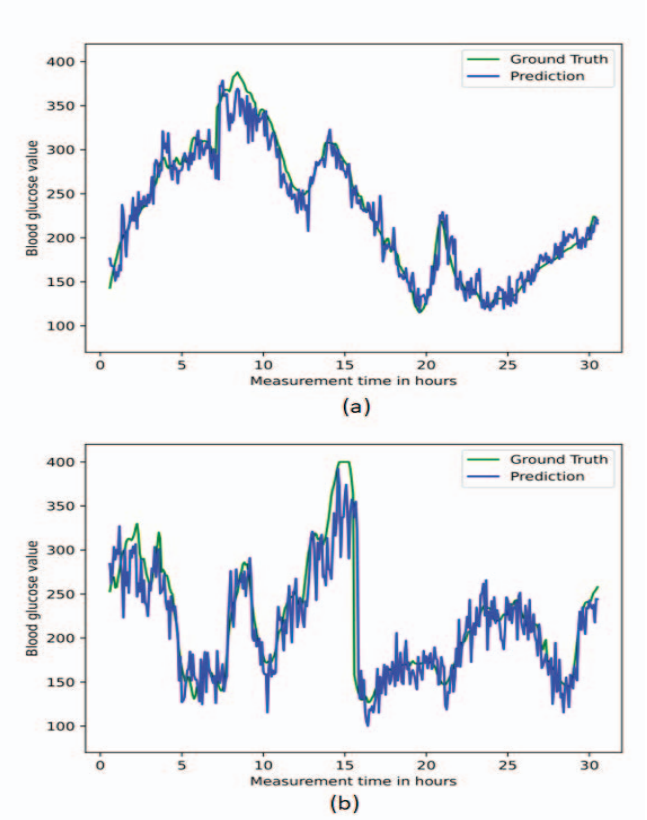

Fig. 3: Blood glucose ground truth in comparison to the prediction for a representative sample of the testing dataset of 30 hours for patient ID 570 (a) and patient ID 584 (b)

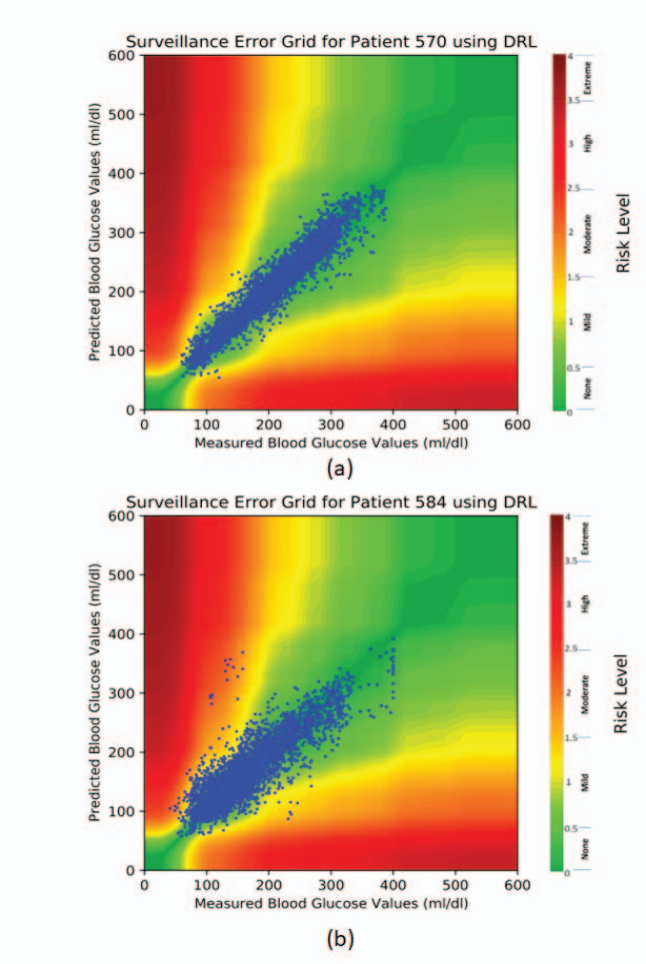

Fig. 4: SEG for Blood glucose predictions in comparison to the prediction for a representative sample of the testing dataset of 30 hours for patient ID 570 (a) and patient ID 584 (b)

work that achieved the best prediction accuracy, on average, and with respect to the mean absolute error, we have improved by 18.4% and 22.5% in 30-minute and 45-minute prediction horizons, respectively. Furthermore, for risk assessment in our predictions, we have visualized the error and evaluated clinical risk through a surveillance error grid approach. For future work, we aim to evaluate our proposed methodology of time-series modeling for other BG prediction datasets and for different categories of diabetes.

REFERENCES

[1] A. D. Association, "Diagnosis and classification of diabetes mellitus," *Diabetes care*, vol. 37, no. Supplement_1, pp. S81–S90, 2014.

[2] A. D. Deshpande, M. Harris-Hayes, and M. Schootman, "Epidemiology of diabetes and diabetes-related complications," *Physical therapy*, vol. 88, no. 11, pp. 1254–1264, 2008.

[3] K. Ogurtsova, L. Guariguata, N. C. Barengo, P. L.-D. Ruiz, J. W. Sacre, S. Karuranga, H. Sun, E. J. Boyko, and D. J. Magliano, "Idf diabetes atlas: Global estimates of undiagnosed diabetes in adults for 2021," *Diabetes research and clinical practice*, vol. 183, p. 109118, 2022.

[4] N. Elsayed, Z. ElSayed, and M. Ozer, "Early stage diabetes prediction via extreme learning machine," in *SoutheastCon 2022*. IEEE, 2022, pp. 374–379.

[5] C. Marling and R. Bunescu, "The ohiot1dm dataset for blood glucose level prediction: Update 2020," in *CEUR workshop proceedings*, vol. 2675. NIH Public Access, 2020, p. 71.

[6] "Blood glucose prediction with deep neural networks using weighted decision level fusion," *Biocybernetics and Biomedical Engineering*, vol. 41, no. 3, pp. 1208–1223, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0208521621001005

[7] 2023. [Online]. Available: https://github.com/PeterDomanski/BG-Prediction-for-Type-1-Diabetics-using-DRL

[8] K. VijiyaKumar, B. Lavanya, I. Nirmala, and S. S. Caroline, "Random forest algorithm for the prediction of diabetes," in *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)*, 2019, pp. 1–5.

[9] J. Yang, L. Li, Y. Shi, and X. Xie, "An arima model with adaptive orders for predicting blood glucose concentrations and hypoglycemia," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 3, pp. 1251–1260, 2019.

[10] D. N. Katsarou, E. I. Georga, M. Christou, S. Tigas, C. Papaloukas, and D. I. Fotiadis, "Short term glucose prediction in patients with type 1 diabetes mellitus," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2022, pp. 329–332.

[11] M. Jaloli and M. Cescon, "Long-term prediction of blood glucose levels in type 1 diabetes using a cnn-lstm-based deep neural network," *Journal of Diabetes Science and Technology*, p. 19322968221092785, 2022.

[12] K. Li, J. Daniels, C. Liu, P. Herrero, and P. Georgiou, "Convolutional recurrent neural networks for glucose prediction," *IEEE journal of*

biomedical and health informatics, vol. 24, no. 2, pp. 603–613, 2019.
[13] T. Zhu, K. Li, L. Kuang, P. Herrero, and P. Georgiou, "An insulin bolus advisor for type 1 diabetes using deep reinforcement learning," *Sensors*, vol. 20, no. 18, p. 5058, 2020.
[14] M. Ubl, T. Koutny, A. Della Cioppa, I. De Falco, E. Tarantino, and U. Scafuri, "Distributed assessment of virtual insulin-pump settings using smartcgms and dmms. r for diabetes treatment," *Sensors*, vol. 22, no. 23, p. 9445, 2022.
[15] C. D. Man, F. Micheletto, D. Lv, M. Breton, B. Kovatchev, and C. Cobelli, "The uva/padova type 1 diabetes simulator: new features," *Journal of diabetes science and technology*, vol. 8, no. 1, pp. 26–34, 2014.
[16] P. Colmegna, A. Bisio, R. McFadden, C. Wakeman, M. C. Oliveri, R. Nass, and M. Breton, "Evaluation of a web-based simulation tool for self-management support in type 1 diabetes: A pilot study," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 1, pp. 515–525, 2023.
[17] J. Xie, "Simglucose v0. 2.1," *Avaible: https://github. com/jxx123/simglucose. Accessed on: Jan-20-2020*, 2018.
[18] C. Hettiarachchi, "Glucoenv v0.1.0(2022)," https://github.com/chirathyh/GluCoEnv, 2022.
[19] P. P. San, S. H. Ling, and H. T. Nguyen, "Deep learning framework for detection of hypoglycemic episodes in children with type 1 diabetes," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, pp. 3503–3506.
[20] A. Bhimireddy, P. Sinha, B. Oluwalade, J. W. Gichoya, and S. Purkayastha, "Blood glucose level prediction as time-series modeling using sequence-to-sequence neural networks." CEUR Workshop Proceedings, 2020.
[21] T. Zhu, K. Li, P. Herrero, J. Chen, and P. Georgiou, "A deep learning algorithm for personalized blood glucose prediction." in *KHD@ IJCAI*, 2018, pp. 64–78.
[22] H. Rubin-Falcone, I. Fox, and J. Wiens, "Deep residual time-series forecasting: Application to blood glucose prediction." in *KDH@ ECAI*, 2020, pp. 105–109.
[23] H. Hameed and S. Kleinberg, "Investigating potentials and pitfalls of knowledge distillation across datasets for blood glucose forecasting," in *Proceedings of the 5th Annual Workshop on Knowledge Discovery in Healthcare Data*, 2020.
[24] T. Zhu, X. Yao, K. Li, P. Herrero, and P. Georgiou, "Blood glucose prediction for type 1 diabetes using generative adversarial networks," in *CEUR Workshop Proceedings*, vol. 2675, 2020, pp. 90–94.
[25] J. Daniels, P. Herrero, and P. Georgiou, "Personalised glucose prediction via deep multitask networks." in *KDH@ ECAI*, 2020, pp. 110–114.
[26] A. Bhimireddy, P. Sinha, B. Oluwalade, J. W. Gichoya, and S. Purkayastha, "Blood glucose level prediction as time-series modeling using sequence-to-sequence neural networks." CEUR Workshop Proceedings, 2020.
[27] S. Oviedo, J. Vehí, R. Calm, and J. Armengol, "A review of personalized blood glucose prediction strategies for t1dm patients," *International journal for numerical methods in biomedical engineering*, vol. 33, no. 6, p. e2833, 2017.
[28] J. Xie and Q. Wang, "Benchmarking machine learning algorithms on blood glucose prediction for type i diabetes in comparison with classical time-series models," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 11, pp. 3101–3124, 2020.
[29] M. De Bois, M. A. E. Yacoubi, and M. Ammi, "Glyfe: review and benchmark of personalized glucose predictive models in type 1 diabetes," *Medical & Biological Engineering & Computing*, vol. 60, no. 1, pp. 1–17, 2022.
[30] T. Zhou, X. Ye, H. Lu, X. Zheng, S. Qiu, and Y. Liu, "Dense convolutional network and its application in medical image analysis," *BioMed Research International*, vol. 2022, 2022.
[31] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[32] S. Hochreiter, Y. Bengio, P. Frasconi, J. Schmidhuber *et al.*, "Gradient flow in recurrent nets: the difficulty of learning long-term dependencies," 2001.
[33] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
[34] T. Liu, Z. Tan, C. Xu, H. Chen, and Z. Li, "Study on deep reinforcement learning techniques for building energy consumption forecasting," *Energy and Buildings*, vol. 208, p. 109675, 2020.
[35] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *International conference on machine learning*. PMLR, 2018, pp. 1861–1870.
[36] V. Konda and J. Tsitsiklis, "Actor-critic algorithms," *Advances in neural information processing systems*, vol. 12, 1999.
[37] I. B. Hirsch, T. Battelino, A. L. Peters, J. J. Chamberlain, G. Aleppo, and R. M. Bergenstal, *Role of continuous glucose monitoring in diabetes treatment*. American Diabetes Association Arlington, VA, 2018.
[38] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization." *Journal of machine learning research*, vol. 13, no. 2, 2012.
[39] D. C. Klonoff, C. Lias, R. Vigersky, W. Clarke, J. L. Parkes, D. B. Sacks, M. S. Kirkman, B. Kovatchev, and E. G. Panel, "The surveillance error grid," *Journal of diabetes science and technology*, vol. 8, no. 4, pp. 658–672, 2014.