Published in the co-located with 24th European Conference on Artificial Intelligence (ECAI 2020), 29-30 August 2020, Santiago de Compostela, Spain, which should be cited to refer to this work.

A Deep Learning Approach for Blood Glucose Prediction of Type 1 Diabetes

Jonas Freiburghaus and **Aïcha Rizzotti-Kaddouri** and **Fabrizio Albertetti** [1]

Abstract. An essential part of this work is to provide a data-driven model for predicting blood glucose levels that will help to warn the person with type 1 diabetes about a potential hypo- or hyperglycemic event in an easy-to-manage and discreet way. In this work, we apply a convolutional recurrent neural network on a real dataset of 6 contributors, provided by the University of Ohio [5]. Our model is capable of predicting glucose levels with high precision with a 30-minute horizon (RMSE = 17.45 [mg/dL] and MAE = 11.22 [mg/dL]), and RMSE = 33.67 [mg/dL] and MAE = 23.25 [mg/dL] for the 60-minute horizon. We believe this precision can greatly impact the long-term health condition as well as the daily management of people with type 1 diabetes.

1 INTRODUCTION

Type 1 diabetes is a disease in which the cells responsible for insulin production are destroyed. Because insulin is the hormone that triggers absorption of glucose within the cells, people with diabetes need to monitor their glucose concentration in the blood and readjust it by frequent insulin injections, following a well-defined medical protocol (e.g., once during the day and once before each meal, to keep blood sugar levels within the normal range). The main challenge in handling diabetes is the optimization of insulin injections in order to avoid hypoglycemia and hyperglycemia. This is complicated by the fact that besides insulin intake and diet, glucose levels are also affected by several other factors such as physical activity, lifestyle, mental state, stress, etc. Despite the various accomplishments made in continuous diabetes monitoring (a.k.a. continuous glucose monitoring, CGM), such methods remain invasive. Furthermore, they are only able to provide the glycemic state at a given time, when the insulin level may already be unacceptable (too high or too low). A proactive detection could therefore dramatically improve the daily handling of diabetes by the patients themselves.

This work presents an approach based on deep learning algorithms for predicting glucose levels in the future (30-minute and 60-minute horizons). Our work is based on the architecture of a recurrent neural network (CRNN) from [3] and proposes certain variants, such as multi-step predictions, regression model using blood glucose level data for each person every 5 minutes, and the inclusion of other data such as basal insulin, bolus insulin, and meal values.

The goal of using a CRNN architecture is twofold. (1) Convolutional layers act as filters and automatically learn to detect the features of interest for prediction. They are also particularly convenient for analyzing time series with little signal processing required. And (2), recurrent neural network are well-known for the capacity to learn long-term relationships between the different values. For instance, it is necessary for the network to be able to capture a correlation between the ingestion of carbohydrates now and a change in the blood glucose level in a near future.

This paper is organized as follows: Section 2 presents related work for glucose prediction, Section 3 formulates the problem of glucose prediction and our contribution, Section 4 details our methodology and discuss experimental results. Finally, in Section 5 we summarize the importance of our contribution and suggest some future work.

2 RELATED WORK

Predicting blood glucose levels for diabetes (type 1 or type 2) using machine learning has gained a lot of attention and has resulted in several methods and applications being proposed recently. They are however either based on solely measuring the glucose levels or the resulting prediction accuracy is not yet high enough to be considered as a reliable predictor of a potential critical glycemic condition. Several types of regression algorithms can be used, including SVR, classic statistical methods such as ARIMA, deep learning neural networks, or even a naive persistence algorithm, to name a few.

Gu et al. [2] propose a personalized smartphone-based non-invasive blood glucose monitoring system that detects abnormal blood glucose levels events by jointly tracking meal, drug and insulin intake, as well as physical activity and sleep quality. It automatically collects daily exercise and sleep quality, and predicts the current blood glucose level of users, together with manual records of food, drug and insulin intake. It needs re-calibration using CGM devices once every three weeks and is based on multi-division deep dynamic recurrent neural network framework. Plis et al. [6] propose a solution that uses a generic physiological model of blood glucose dynamics to generate features for a SVR model that is trained on contributor specific data. It is shown that, the model could be used to anticipate almost a quarter of hypoglycemic events 30 minutes in advance, however the demonstrated corresponding precision is 42%. Contreras et al. [1] present an alternative approach to glucose levels prediction, based on previous studies that incorporated medical knowledge into a grammar aimed to build expression for glucose that considered previous glucose values, carbohydrate intake, and insulin administration. They extend the previous research to investigate a novel and complementary approach that uses symbolic regression through grammar evolution to determine an approximate glucose levels and fluctuations using personalized blood glucose predictive models. In the same order of topic, we note the work in [4]. It also contains a comparison with a prediction using latent variable with exogenous input (LVX [9]) model, in this model bolus insulin and meal are included in the predictor matrix X but in this work only meal information is

[1] HES-SO University of Applied Sciences and Arts of Western Switzerland, email: aicha.rizzotti@he-arc.ch

added . We can see that the RMSE for a prediction horizon of 60 min for clinical data, varies from 37.02 to 35.96 [mg/dL]. In [3], they use a deep learning architecture for predicting 30-minute and 60-minute horizons on both real and simulated patients. For real patients and a 30-minute horizon, they report an RMSE of 21.07 [mg/dL]. However, only a few approaches have used deep learning algorithms for CGM on clinical data and more specifically for this dataset provided by the University of Ohio [5].

3 Glucose Prediction

The aim of this work is to predict glucose levels in advance in order to avoid situations of hyperglycemia or hypoglycemia, as well as others negative effects on the health. For instance, chronic hyperglycemia may induce fatigue and vision problems among others. For that purpose, we created a model capable of predicting the glycemia of type 1 diabetes, where values must be as accurate as possible. In this context, the metrics are the RMSE and the MAE. The smaller these values, the more reliable the model. The real gain for a patient is to be able to make decisions at any given time considering the prediction of future values, and possibly avoid glycemia-related discomforts while minimizing intrusive methods.

4 METHODOLOGY

4.1 Approach

Our approach can be summarized by the following steps:

1. Data importation
2. Data preprocessing
3. Implementation of the CRNN prediction model with multi-step forecasting
4. Training, testing and tuning on selected features
5. Delivery of the forecasted blood glucose levels

Data importation involves loading, merging, and aligning values from multiple sources under the same time scale. The selected features are basal/bolus insulin, carbohydrates, and blood glucose levels.

In the preprocessing step, all variables must have measurements carried over at the same time. This requires the use of subsampling or oversampling methods and, corollary, defining imputing methods. Linear interpolation is used for the glucose level on the training set to resample the time series at a frequency of 5 minutes. The others features may be imputed with null values when required, as their nature is sparse. We can also use domain-specific functions, such as an equation describing the absorption rate over time for carbohydrates. The preprocessing steps are summarized by :

1. Save all the blood glucose timestamps
2. Resample the features to a time delta of 1 second
3. Forward fill the missing values by using the last available values
4. Fill the left missing values with 0
5. Resample the features to a time delta of 5 minutes
6. Smooth each feature with a 1D Gaussian filter over a window containing the past 2 hours of data

We use the saved timestamps at preprocessing step 1 to generate the results at the same timestamps as the measured values.

Linear interpolation is used during the training process to impute the missing values. This allows to have more data points for the model to be trained on. However, we should note that linear interpolation is not ideal for big gaps of missing values. The interpolation is not used at test time as it could lead to a data peek. Meaning the predictions would be contaminated by future values.

The CRNN model target values are based on the following equation:

$$y_{t+L} = bg_{t+L} - bg_t, \text{for } L = 1, 2, \ldots, 12 \tag{1}$$

where bg_t is the blood glucose value at time t, L is the lag value in timesteps for the horizon, and y the label to predict, that is the differentiated value of the blood glucose level.

For instance, if the blood glucose level is 80 mg/dL at the current time and 60 mg/dL 30 minutes later, the label for a prediction horizon of 30 minutes at the current time would be -20 mg/dL.

Respectively, as the model does not predict directly the blood glucose level but only the difference from the last known value. The predicted blood glucose level is obtained with the following equation:

$$\widehat{bg}_{t+L} = bg_t + \hat{y}_{t+L}, \text{for } L = 1, 2, \ldots, 12 \tag{2}$$

where bg_t is the blood glucose value at time t, $\widehat{bg}_{t+L}$ is the predicted blood glucose level at time $t + L$, $\hat{y}_{t+L}$ is the predicted blood glucose level difference at time t with lag L, L is the lag value in timesteps for the horizon, and y the label to predict.

It is important to note that the CRNN only outputs the values $\hat{y}_{t+L}$.

The model is capable of giving a prediction for each 5 minutes prediction horizon up to 60 minutes that is 5, 10, ..., 60 minutes. This feature may give valuable information to a user and thereby improve their blood glucose level control. An example of such a prediction is given in Fig. 5.

The overall architecture of the CRNN is based on [3] and described in Fig. 1. The input signals time series are fed into a CNN for extracting relevant features. The purpose of the pooling layers is to gradually reduce the spatial dimension while keeping only the highest values included in the pooling window. Then, these features are fed into an RNN layer to model the relationships over time. Finally, a dense neural network is used as a last layer for regressing the desired target.

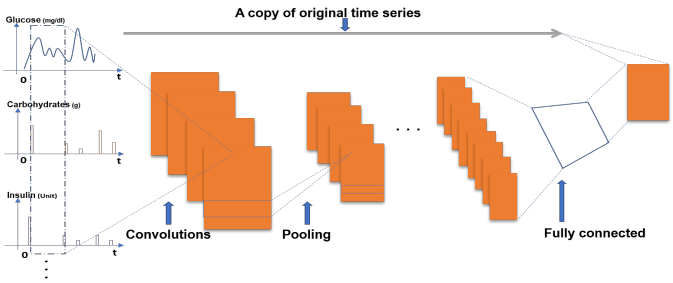

Figure 1. A multi-layer CRNN composed of convolutional layers, pooling layers, an RNN network, and a dense neural network

The training of a model is done on a contributor-basis, that is one model is trained per contributor. The reason is that each glucose response is individual, and a one-population model does not seem reasonable for CGM.

4.2 Experimental Results

4.2.1 Dataset

Our results reported in this work are based on the OhioT1DM dataset [5]. For each contributor a train set as well as test set are provided. One model only is pretrained on the data from the 6 contributors of 2018. Then for each 6 data contributor of 2020 transfer learning is applied, resulting in one trained model per contributor.

The reported results are based on the following signals: glucose level, basal insulin, bolus insulin, and meal for 6 data contributors (540, 544, 552, 567, 584, and 596). While more signals were available, we decided to use only these signals. Indeed, we performed several tests with the complete dataset and the preliminary results indicated better results with a limited set of features.

The use of a sliding window consisting of the last 2 hour data, resulting in 24 data points is based on [3]. Cross- and/or auto-correlation may provide a good starting point to find a reasonable sliding window size. Of course, the computing complexity must be taken into consideration depending on the targeted deployment hardware.

4.2.2 Architecture and learning process

The detailed architecture of the CRNN is presented in Table1.

Layer description	Output dimension
Convolution 1D	(Batch size, 24, 8)
Max pooling 1D	(Batch size, 12, 8)
Convolution 1D	(Batch size, 12, 16)
Max pooling 1D	(Batch size, 6, 16)
Convolution 1D	(Batch size, 6, 32)
Max pooling 1D	(Batch size, 3, 32)
LSTM	(Batch size, 64)
Dense	(Batch size, 256)
Dense	(Batch size, 32)
Dense	(Batch size, 12)

Table 1. Neural network layers and output shapes

The model is pretrained on batches of size 1024 over 1000 epochs, with an RMSProp optimizer. The learning rate is initially set to 0.001 and is reduced with a factor of 0.1 when the model does not progress after 3 epochs. Early stopping is used similarly with a patience of 50 epochs in order to regularize the model. The last model's weights with the lowest validation loss are then restored.

For each data contributor of 2020 the pretrained model is loaded and trained similarly as the pretraining stage. With the only difference that the learning rate is reduced with a patience of 15 epochs and that one model is saved for each contributor.

4.2.3 Results

The RMSE and the MAE are calculated for the six contributors using the following equations.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{3}$$

where $\hat{y}_i$ is the predicted value, y_i the truth value and n the number of observation.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{4}$$

The first observation is the errors systematically increase for each contributor over time. It is not surprising that the larger the prediction window, the larger the error in general grown up.

A comparison between contributors was also performed (Fig. 2 and Fig. 3). As we can see, for example with the contributor 596, the prediction curve for 30 minutes and 60 minutes follows the real curve with an RMSE = 13.34 and MAE = 9.08 [mg/dL], and RMSE = 27.74 and MAE = 19.13 [mg/dL]. These specific curves are also detailed in Fig. 4.

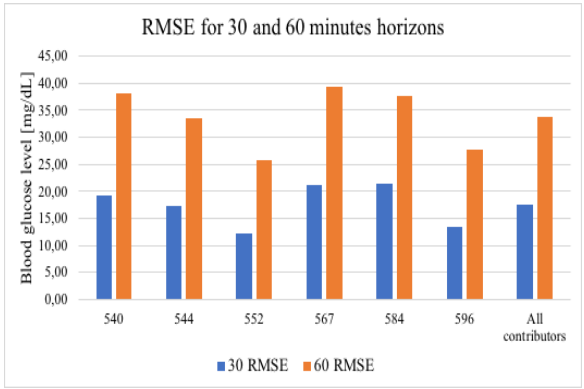

Figure 2. RMSE-30 vs RMSE-60 horizons for the 6 contributors of 2020

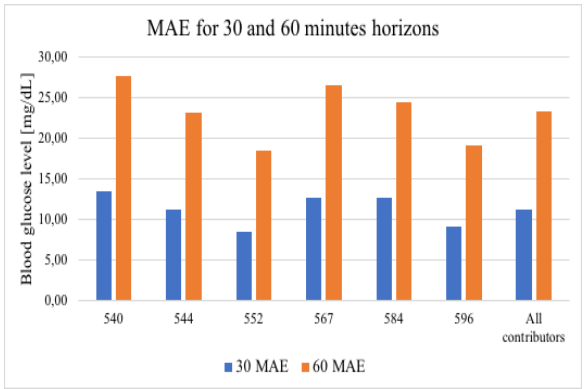

Figure 3. MAE-30 vs MAE-60 horizons for the 6 contributors of 2020

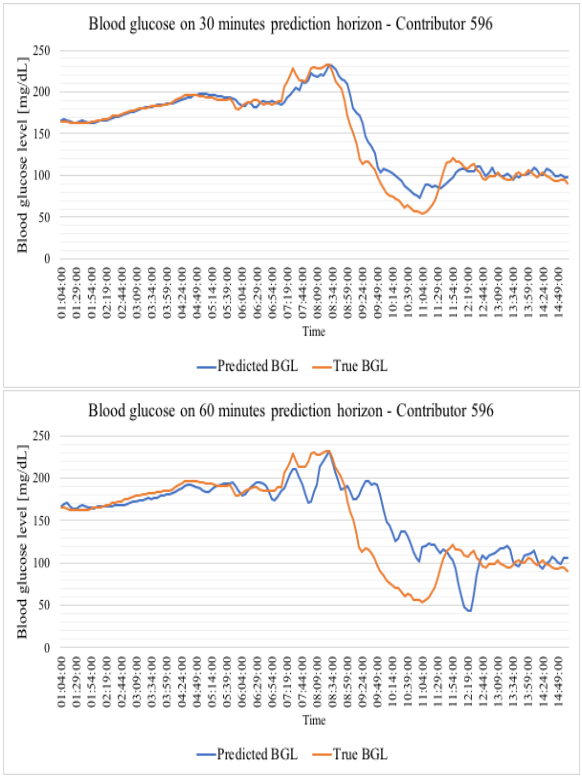

Figure 4. Example of prediction results of contributor 540, for 30 and 60 minutes

The evaluation of the prediction from the contributor-side can be performed by a multi-step prediction (as illustrated in Fig. 5). Prediction is made through one forward pass of a horizon of the last 2 hours data and outputs the horizon for the next hour represented by the orange curve. In the given example, the model seems to have predicted well the tendencies.

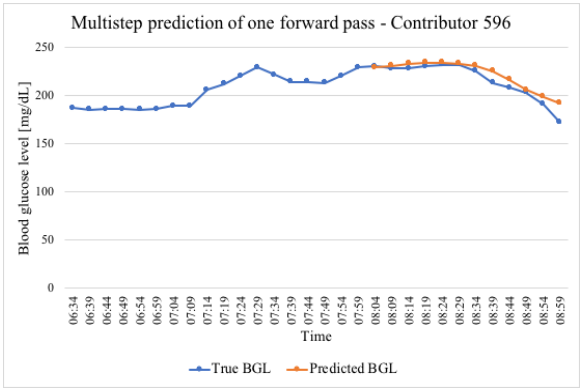

Figure 5. Multi-step prediction of one forward pass

A comparison of our approach with other algorithms is summarized in Table 2. We notice that the RMSE of our work at horizon 30 is smallest and at 60 minutes similar to [3]. Let us highlight that in (1) the database is different. The persistent algorithm is the baseline model. It forecasts the blood glucose level by using the last know value.

$$\hat{y}_{t+L} = bg_{t-1}, \text{for } L = 1, 2, \ldots, 12 \quad (5)$$

Prediction horizon	Metrics (mg/dL)	(1) Li's CRNN	(2) BASELINE	(3) CRNN
		Overall		
30	**MAE**	NA	18.13 ± 0.00	11.22
	RMSE	$21.0.7 \pm 2.35$	25.76 ± 0.00	17.45
60	**MAE**	NA	30.70 ± 0.00	23.25
	RMSE	33.27 ± 4.79	42.00 ± 0.00	33.67

Table 2. Comparison of different algorithms: (1) Li's CRNN from [3], (2) baseline with persistence forecast (where the previous value is predicted), and (3) the proposed CRNN. Let us note that (1) uses different data (not publicly available) than from (2) and (3).

With the hypothesis of a hypo-glycemia starting below 70 mg/dL and a hyper-glycemia starting above 150 mg/dL, we believe that our model delivers relevant and actionable results for real patients. In our opinion, a RMSE of 17.45 mg/dL for a 30-minute horizon indicates that data-driven decisions could be made in regard to avoiding hypo or hyper-glycemic related events.

5 CONCLUSION

In this paper we described a model for predicting future blood sugar levels of people with type 1 diabetes. A CRNN approach was proposed with the advantages of using only 4 different signals and very little signal processing. The evaluation was performed using RMSE and MAE metrics, with different horizons and on multiple contributors. The results were compared with different algorithms.The results report low error rates given the problematic of glucose prediction, and in our opinion could be considered for real-world implementation. Yet, several research tracks remain to be explored. For instance, testing additional features that can influence blood sugar levels such as stress or illness. We can also think of extracting manual features from the given signals with signal processing methods, and defining domain-specific imputation methods, such as for the absorption of carbohydrates over time. It would be also interesting to further personalize predictions as suggested in [7]. Another direction could be to use reinforcement learning approaches for the insulin recommendation, such as in [8].Those self-learning approaches are adaptable and personalize the daily insulin values to ensure glucose control, despite inter and intra-patient variability.

ERRATUM

During the making of the camera ready version, we found an error in the preprocessing stage thanks to the great reviews. The missing values were treated using a linear interpolation during the testing of the model. Thereby the predictions were contaminated by future values. This error was corrected by removing the interpolation and dealing with missing values as explained in this paper version.

Code source

It is available at https://github.com/JonasFreibur/BLGP-HES-SO

ACKNOWLEDGEMENTS

The authors would like to thank the HES-SO for funding this research.

REFERENCES

[1] Iván Contreras, Silvia Oviedo, Martina Vettoretti, Roberto Visentin, and Josep Vehí, 'Personalized blood glucose prediction: A hybrid approach using grammatical evolution and physiological models.', *PloS one 12.11*, (2017).

[2] Weixi Gu, Yuxun Zhou, Zimu Zhou, Xi Liu, Han Zou, Pei Zhang, Costas J Spanos, and Lin Zhang, 'Sugarmate: Non-intrusive blood glucose monitoring with smartphones.', *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol*, (2017).

[3] Kezhi Li, John Daniels, Chengyuan Liu, Pau Herrero-Vinas, and Panteli Georgiou, 'Convolutional recurrent neural networks for glucose prediction.', *EEE journal of biomedical and health informatics*, (2019).

[4] Chengyuan Liu, Josep Vehi, Nick Oliver, Pantelis Georgiou, and Pau Herrero, 'Enhancing blood glucose prediction with meal absorption and physical exercise information.', *arXiv preprint arXiv:1901.07467*, (2018).

[5] Razvan C. MARLING, Cindy et BUNESCU, 'The ohiot1dm dataset for blood glucose level prediction.', *KHD@ IJCAI*, (update 2020).

[6] Kevin Plis, Razvan Bunescu, Cindy Marling, Jay Shubrook, and Frank Schwartz, 'A machine learning approach to predicting blood glucose levels for diabetes management. in proceedings of aaai workshop on modern artificial intelligence for health analytics.', *AAAI Press*, 35–39, (2014).

[7] Qingnan Sun, Marko V Jankovic, João Budzinski, Brett Moore, Peter Diem, Christoph Stettler, and Stavroula G Mougiakakou, 'A dual mode adaptive basal-bolus advisor based on reinforcement learning.', *IEEE journal of biomedical and health informatics*, **23**(6), 2633–2641, (2018).

[8] Qingnan Sun, Marko V Jankovic, and Stavroula G Mougiakakou, 'Reinforcement learning-based adaptive insulin advisor for individuals with type 1 diabetes patients under multiple daily injections therapy', in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3609–3612. IEEE, (2019).

[9] Chunhui Zhao, Eyal Dassau, Lois Jovanovič, Howard C Zisser, Francis J Doyle III, and Dale E Seborg, 'Predicting subcutaneous glucose concentration using a latent-variable-based statistical method for type 1 diabetes mellitus', *Journal of diabetes science and technology*, **6**(3), 617–633, (2012).