Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/14/20230783

A comparison of various deep learning models on blood glucose management of type 1 diabetes: A brief review

Z. Jiang[1,3], N. Yang[2]

[1]Boston University, Boston MA 02215, USA
[2]Purdue University, West Lafayette IN 47906, USA

[3]Corresponding Author: Zhuojian Jiang, zjiang2@bidmc.harvard.edu

Abstract. Type 1 diabetes has become one of the most common chronic diseases nowadays because patients' pancreas cannot produce sufficient insulin, which helps the blood sugar to enter the cells, which will cause them to build up in the bloodstream and leads to complications and diseases. Therefore, a basal-bolus insulin therapy that contains daily insulin injections has become routine for patients with type 1 diabetes to help regulate the blood sugar level. To better monitor the blood glucose level, digital health monitors are becoming the trending method for type 1 diabetes patients. Meanwhile, all the data generated by the digital monitoring devices made researchers realize that deep learning algorithms could be implanted to help the device better predict a patient's blood sugar levels. In this paper, we aim to present a review of testing several state-of-the-art deep-learning models on blood glucose prediction. We have identified a literature search and focused on the deep-learning algorithms for glucose management. After detailed explanations of each model, we employ them on one mutual dataset to identify the direct prediction results of each model and compare the pros and cons of these models according to the results report. While all these models have the most advanced frameworks, the lack of feature varieties and data availability becomes their limitations. However, followed by the increased focus on the digital health field, these challenges might soon get resolved, which leads to more comprehensive models that could be further deployed in clinical conditions.

Keywords: type 1 diabetes, deep learning, deep neural networks.

1. Introduction

Diabetes mellitus is a chronic disease affecting the patient's metabolic process, where their pancreas cannot produce sufficient insulin or does not work at total capacity (citation). Studies indicate that about 463 million people have diabetes, which is still increasing and is expected to be doubled in 20 years [1]. Based on the etiology of diabetes, three main clinical categories are listed here: type 1 diabetes (T1D), type 2 diabetes (T2D), and gestational diabetes mellitus (GDM) [2]. This article mainly focuses on type 1 diabetes (T1D). From previous CDC statistical studies on diabetes, approximately 28.7 million people are diagnosed with diabetes in the U.S. Among them, there are 1.6 million adults diagnosed with type 1 diabetes, which is expected to be doubled by 2040 [2]. Type 1 diabetes is a chronic disease caused by pancreas dysfunction, resulting in the patient's inability to produce sufficient insulin to deliver glucose into cells. T1D occurs when the immune system destroys the insulin-secreting beta cells of the pancreas [3]. Nevertheless, insufficient insulins could severely influence patients' life quality, and it may also

DOI: 10.54254/2755-2721/14/20230783

cause complications involving but not limited to kidney damage, heart and blood vessel disease, and eye damage [3]. Therefore, the patients must take insulin to help the insufficiency within their bodies. With the development of the digital health field, continuous blood glucose monitoring (CGM) has become one of the effective methods of glucose management. Its ability to measure real-time blood glucose data demonstrates the massive potential in the glucose management field. Nevertheless, researchers further combine the glucose data with conventional machine learning algorithms to predict short-term blood glucose levels change in patients to help regulate blood glucose level.

Furthermore, a more empowered method, deep learning, stepped onto the stage of digital healthcare. Due to the complexity of the pathology of type 1 diabetes, many features, such as the number of meals, need to be accounted for to predict accurate readings. With convolutional neural networks, large-scale raw data can be quickly processed without complicated preprocessing, and the higher dimensional models could handle more complex datasets with various features. While the other review articles recently demonstrated comprehensive studies on the digital healthcare field and presented the most advanced deep neural networks (DNNs), they still need to provide a more detailed and direct comparison of different neural network models. In the previous review on deep learning algorithms in T1DM, distinct datasets are employed to train the model. However, these datasets sometimes do not share mutual features, and lack of data availability caused indirect comparison between these models. Due to the increase in the diabetic population recently, the focus of this paper will help on deciding the most suitable deep learning models for daily glucose management, which will further benefit the life quality of the patients. Therefore, this paper will mainly focus on comparing various models with identical datasets with rich feature amounts to determine the advantage and disadvantages of each model.

2. Deep learning review

Deep learning is commonly employed in healthcare, particularly glucose management. Artificial neuron networks (ANNs) initially improved deep learning methods. ANNs are developed by imitating the brain neurons, which grants the ability to handle complex calculation processes [4]. ANNs are built with three major components, input layers, hidden layers, and output layers. As the number of hidden layers increases, the ANN can analyze more data features and simulate human brain behaviors. However, as the model becomes more complex and new frameworks are being developed, the hardware cannot follow the steps of software, which becomes the limitation of ANNs. Fortunately, the potential of ANNs made people believe in its future in data analysis and many other fields, such as automation. Followed by the rapid evolution of computer hardware like graphic processing units (GPUs) from decades ago, the capability of deep learning also gets further developed. Nowadays, many developers have built generations of supporting software frameworks like Theano [5], Tensorflow [6], and Pytorch [7] to help people quickly build in-depth deep neuron networks (DNNs).

Although various DNN models are built by different frameworks, they could be generally divided into three categories: supervised learning, unsupervised learning, and reinforcement learning. Unsupervised learning uses an unlabeled dataset for model training, meaning the algorithms will learn limited or no information from the dataset. They are typically involved in tasks like grouping and dimensionality reduction. On the other hand, reinforcement learning also involves using unlabeled data. However, the model will explore new strategies to complete the assigned tasks and try to balance the exploitation and exploitation of the strategies. Unsupervised and reinforcement learning generally has unique characteristics mainly used as feature detections and approximations. However, the mainstream method for glucose management region is supervised learning due to its capability on network weights and potential in large-scale handling features during model training [8] [9].

From the literature we reviewed, there are four types of supervised learning algorithms used in glucose management: deep multilayer perceptrons (DMLPs), convolutional neural networks (CNNs), and recurrent neural networks (RNNs). In the healthcare region, the purpose of DMLPs is the function of fully connected (FC) layers. Typically, the FC layers are combined with other supervised learning algorithms, such as CNN, to optimize objectives. Due to their ability to process multi-dimensional data sets and remarkable performance, CNNs are often used in automation or imaging processing tasks [10].

Although CNN is suited for multi-dimensional data analysis, the original time cost and calculation power requirement are tremendous. However, large-scale data analysis becomes available with the parallelized operations of GPU and tensor processing units (TPUs) [11]. Furthermore, with the combination of FC layers and other sub-sampling layers, time cost could be significantly saved while accuracy and precision remain ensured.

On the other hand, RNNs can analyze sequential data containing time stamps, which is suited for glucose management because all the CGM data are timestamped with temporal features. However, the challenge for traditional RNN cells is that they need help handling back-propagation training, gradient vanishing, and exploding problems like other fully connected networks [12]. However, with the technique of long short-term memory (LSTM) [13] and gated recurrent units (GRUs) [14], RNNs can preserve long-term information and overcome the problems mentioned above. These advanced RNNs allow algorithms to handle various prediction, classification, and regression tasks. RNNs are the perfect short-term glucose prediction method by employing their ability to analyze sequential data.

3. Methodology

To identify and analyze the progress and new challenges brought by machine learning, we conduct a review focusing on the 60 minutes blood glucose level prediction by focusing on the performance of different models based on the Ohio T1DM database. Ohio T1DM includes 5 minutes of blood glucose level with other daily activities features of patients: insulin dose, time of physical activities, and other physical features including skin temperature and galvanic skin response [15]. Given these data, building blood glucose level prediction models becomes possible. In this study, we will compare the feature selection and model structures.

3.1. Search strategies

In our paper search, the keywords "OhioT1DM", "glucose", and "prediction" were combined using Boolean operators AND/OR. The detailed query was: (T1DM AND glucose AND prediction)

3.2. Inclusion and exclusion criteria

This review selected full-text studies focusing on machine learning method application on glucose level prediction based on the T1DM data set. after removing the unrelated and repeated works, the final collection of paper was organized into two categories: features extraction and glucose level prediction.
The expectation of the included studies should have the following features: contain details of data processing, present the structure of models applied, evaluate model prediction performance by standard metrics. Abstracts, posters, technique reports, and reviews were excluded.

3.3. Information Extraction

From the selected paper collection.The following pre-defined categories were used to present the selected studies.

1) Model:

This category first summarizes the model architectures of prediction models. The details of the hybrid model were also included.

2) Features:

We present the features that have been used to predict glucose levels. The input dimensions are distinct from different models, and we will study the potential effect of input dimensions on the performance of the deep learning prediction model.

2) Development process:

The category summarizes the process of building prediction models, which includes data processing, training, testing, and validations. For hybrid models, the roles of each deep learning will be specified.

3) Main outcome:

The result of performance evaluation. Most of the evaluations were presented as RMSE calculations.

4) Prediction period:

DOI: 10.54254/2755-2721/14/20230783

The period of prediction. Most studies have 30 min prediction period.

5) Limitation:

As a review of the performance of deep learning models based on the Ohio T1DM data set, this category describes the limitations specified in the selected studies and future work inspirations and improvements.

Table 1. Summary of selected articles focused on glucose management [16]-[25].

Ref num	Model	Feature	Development process	Main outcomes	Prediction period	Limitation
1	Casual CNN (Wave Net method)	CGM data, Insulin event; Carbohydrate intake; Time index	Pre-processing; calculate changes of glucose level in PH; train model;	RMSE for 30-min PH: 21.73 mg/dL	30 min	the model not able to forecast the fluctuate after insulin event and carbohydrate intake
2	NPE(decomposed convolution) + LSTM	19 dimension(all fields in Ohio T1DM data set)	Data preprocessing; separate model test; combine model test	RMSE for 30-min PH: 17.80 mg/dL	30 min	NPE do not require a feature engine
3	Stack LSTM, Vanilla RNN	CGM values, carbohydrate intake from the meal, insulin dose as a bolus, and 5-min aggregation of step count	Selecting prediction features; Kalman smoothing; training; test; validating	RMSE for 30, 60 min: 6.45 17.24 mg/dL	30 min, 60 min	The prediction accuracy of glucose over traditional range decrease
4	Multi-layer CNN	N/A	Preprocessing; dilated CNN; post processing	RMSE for 30, 60 min: 19.28, 31.83 mg/dL	30 min, 60 min	The training quality highly relies on data set quality
5	Bi-directional LSTM	Top 5 most relevant and common features	Preprocessing; feature selection; evaluation	RMSE for 30 min: 20.8 mg/dL	30 min	Performance were affected by fluctuation in glucose values and missing data
6	RNN	CGM values alone	Preprocessing; training; testing; evaluation	RMSE for 30 min: 18.867 mg/dL	30 min, 60 min	There is potential relevance between glucose level and other features while only CGM was involved in the study

Table 2. (continued).

7	RNN with LSTM cells	CGM value alone	Prepro-cessing; training; evaluation	RMSE for 30 min: 20.1 mg/dL	30 min	The model is hard to predict hypoglycemic event
8	Dilated RNN	sampling time, CGM values, meal intake and insulin dose	Filling miss-ing data; training;	RMSE for 30 min: 18.9 mg/dL	30 min	the accuracy of database still needs to be im-proved. the prediction does not consider the potential ac-tivities changes
9	Hybrid LSTM + WaveNet + GRU	N/A	Training three blocks separately; decision level fusion	RMSE for 30, 45, 60: 21.9, 29.12, 35.10 mg/dL	30,45,60 min	The fusion stage can involve other patients' history to improve the prediction preci-sion
10	MS-LSTM model: multi lag structure	BG value, basal insu-lin dosage, bolus insu-lin dosage, carbohy-drate intake, and timestamp	Remove out-liers; filling missing data; training-vali-dation ratio is 9/1	RMSE for 30, 60 min: 19.048, 32.029 mg/dL	30 min, 60 min	The prediction accuracy was in-fluenced by miss-ing data and rapid fluctuation in blood glucose level

4. Result

The initial search generated 202 results, and then the initial collection was filtered by our inclusion and exclusion criteria. We manually selected ten papers from the remaining papers by full-text inspection. The earliest publication was published in 2018, and the rest of the papers were published from 2020 to 2022. This trend is a brave new study field. The detailed description for each category is presented in Table I.

4.1. Prediction strategies

The prediction model will take patients' historical data as input to predict the blood glucose level in a 30 to 60 minutes period. Showing in table I, when selecting features as input to predict the glucose level, there are two preferences among those papers: using only CGM values and combining other features. Based on the evaluation result, there is no significant difference between the prediction accuracy around those two preferences. Martinsson mentioned that the relationship between other features and glucose levels could be complex. Therefore, they decided not to include features rather than CGM value to implement the prediction [21]. While Zhu selected insulin event and carbohydrate intake with CGM value to involve the prediction. In Zhu's study, the model did not respond to the blood glucose level fluctuation caused by those events sensitively [16].

DOI: 10.54254/2755-2721/14/20230783

4.2. Potential factors that can influence the accuracy

From the reported deep learning prediction model performance, the RMSE for 30 min varies from 6.45 to 21.9. Except for the minimum RMSE given by the stack LSTM model, other RMSEs are around 19 to 22. We observed that the RMSEs decrease as more researchers participate in this study field, indicating that prediction accuracy increases as new models are developed. As mentioned in all of the papers, the missing data in Ohio state T1DM is considered a main factor that can influence the effectiveness of prediction models. To decrease the effect of missing data, some researchers choose to fill missing data by interpolation and extrapolation, like Zhu [16]. In contrast, other studies skip the missing data to avoid the uncertain effect of these added data. The fluctuation of blood glucose level itself is also a factor that can influence the precision of prediction [16][20][25]. The turnover change in blood glucose level is usually related to carbohydrate intake events and insulin events. However, the earliest CNN model of Zhu needs to respond to these two events with corresponding blood glucose change fast enough. The consequence is a lag between the prediction value and the actual value. The following study introduces NPE into the LSTM model to process the influence of physiological events on blood glucose events. The RMSE of 30 min improved to 17.8 mg/dL, which is the second best in our collection [17].

5. Discussion

5.1. Limitation and challenge

While the prediction precisions have improved over the last three years, more than precision is needed to provide a reliable glucose prediction service. As mentioned in the table, prediction accuracy relies on training data quality [19][25]. Depending on the data preprocessing method, the prediction accuracy of a specific range can have poor performance. Zhu’s DRNN model is hard to predict the hypoglycemic event [22]. By applying Kalmon smoothing in the preprocessing stage, the model's performance on abnormal blood glucose levels was worse than other levels [18]. Therefore, data preprocessing is a field that needs more attention. Then, the delay between the prediction and actual value is a universal limitation among all selected studies.

5.2. Future opportunities

First, introducing NPE into traditional machine learning models can improve accuracy by strengthening the model's response to special physiological events [17]. The existing deep learning models rely on feature selection by deep learning itself. Second, as revealed by studies, the performance of models highly depends on the dataset's quality and data preprocessing method. Thus, how to improve the dataset quality in practice and covering the dataset's deficiency is a field worth more study.

6. Conclusions

In this review, we present a brief review and comparison of the most advanced deep-learning algorithms in the glucose management region. Within the literature we reviewed, we mainly focused on comparing various DNNs and testing their performances on the same dataset to provide the most direct analysis of each architecture. Nevertheless, we also identified the existing challenges these pieces of literature are facing, which include different methods of preprocessing that can ultimately influence the accuracy of the model and the limitation of specific models that heavily rely on the quality of the training dataset. However, we believe that in the future, there will be more comprehensive datasets in clinical settings that could provide more features and high-quality data points for more robust model training. This review will provide guidelines and standards for future research on glucose management models, especially on model selection and customization for individual patients. Furthermore, the direct comparison between selected models in the review also exploits further research on parameters optimization, neuron network framework, and customized feature selections, which leads to more accurate and stable predictions. We also expect that shortly, deep learning algorithms can accurately and safely detect blood glucose variations and help diabetes patients improve their life quality.

References

[1] American Diabetes Association and others. Classification and diagnosis of diabetes. Diabetes care, 40 (no. Supplement 1): 11-24, 2017
[2] Divers J, Mayer-Davis EJ, Lawrence JM, et al., “Trends in Incidence of Type 1 and Type 2 Diabetes Among Youths — Selected Counties and Indian Reservations”,. MMWR Morb Mortal Wkly Rep;69:161–165, 2020.
[3] Centers for Disease Control and Prevention and others. National diabetes fact sheet: national estimates and general information on diabetes and prediabetes in the united states. US department of health and human services, centers for disease control and prevention, 201(1): 2568-2569,2011.
[4] I. Goodfellow, Y. Bengio and A. Courville, Deep Learn., Cambridge, MA, USA:MIT press, 2016
[5] J. Bergstra et al., "Theano: A CPU and GPU math expression compiler", Proc. Python Sci. Comput. Conf. (SciPy), vol. 4, no. 3, pp. 1-7, 2010.
[6] M. Abadi et al., "Tensorflow: A system for large-scale machine learning", Phroc.12th {USENIX} Symp. Operating Syst. Des. Implementation, pp. 265-283, 2016.
[7] A. Paszke et al., "Pytorch: An imperative style high-performance deep learning library", Proc. Adv. Neural Inf. Process. Syst., pp. 8024-8035, 2019.
[8] G. E. Hinton, S. Osindero and Y.-W. Teh, "A fast learning algorithm for deep belief nets", Neural Comput., vol. 18, no. 7, pp. 1527-1554, 2006.
[9] V. Mnih, K. Kavukcuoglu, D. Silver, et al. “Human-level control through deep reinforcement learning”. Nature 518, pp. 529–533, 2015.
[10] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition", Proc. IEEE, vol. 86, no. 11, pp. 2278-2324, 1998.
[11] D. C. Cireşan, U. Meier, L. M. Gambardella and J. Schmidhuber, "Deep big simple neural nets for handwritten digit recognition", Neural Comput., vol. 22, no. 12, pp. 3207-3220, 2010.
[12] Y. Bengio, P. Simard and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult", IEEE Trans. Neural Netw., vol. 5, no. 2, pp. 157-166, 1994.
[13] S. Hochreiter and J. Schmidhuber, "Long short-term memory", Neural Comput., vol. 9, no. 8, pp. 1735-1780, 1997.
[14] K. Cho, B. van Merriënboer, D. Bahdanau and Y. Bengio, "On the properties of neural machine translation: Encoder–decoder approaches", Proc. 8th Workshop Syntax Semantics Struct. Stat. Transl., pp. 103-111, 2014.
[15] C. Marling and R. Bunescu, "The OhioT1DM dataset for blood glucose level prediction", Proc. 3rd Int. Workshop Knowl. Discovery Healthcare Data, 2018.
[16] T. Zhu, K. Li, P. Herrero, J. Chen and P. Georgiou, "A deep learning algorithm for personalized blood glucose prediction", 3rd Int. Workshop Knowledge Discovery Healthcare Data IJCAI-ECAI 2018, pp. 64-78, 2018.
[17] K. Gu, R. Dang and T. Prioleau, "Neural Physiological Model: A Simple Module for Blood Glucose Prediction," 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 5476-5481, 2020.
[18] M.F. Rabby, Y. Tu, M.I. Hossen, I. Lee and A. Maida. “Stacked LSTM based deep recurrent neural network with kalman smoothing for blood glucose prediction.” BMC Med Inform Decis Mak 21, 101, 2021.
[19] K. Li, C. Liu, T. Zhu, P. Herrero and P. Georgiou, "GluNet: A deep learning framework for accurate glucose forecasting", IEEE J. Biomed. Health Infrom., vol. 24, no. 2, pp. 414-423, Jul. 2019.
[20] B. Ananth, et al. "Blood glucose level prediction as time-series modeling using sequence-to-sequence neural networks." CEUR Workshop Proceedings, 2020.
[21] J. Martinsson, A. Schliep, B. Eliasson and O. Mogren, "Blood glucose prediction with variance estimation using recurrent neural networks", J. Healthcare Informat. Res., vol. 4, no. 1, pp. 1-18, 2020.

DOI: 10.54254/2755-2721/14/20230783

[22] J. Martinsson, A. Schliep, B. Eliasson, C. Meijner, S. Persson, and O. Mogren. “Automatic Blood Glucose Prediction with Confidence Using Recurrent Neural Networks.” KHD@IJCAI, 2018.

[23] T. Zhu, K. Li, P. Herrero, J. Chen and P. Georgiou, "Dilated recurrent neural networks for glucose forecasting in type 1 diabetes", J. Healthcare Informat. Res., pp. 1-17, 2020.

[24] D. Hatice Vildan, M. Taskiran, and T. Yildirim. "Blood glucose prediction with deep neural networks using weighted decision level fusion." Biocybernetics and Biomedical Engineering 41.3 (2021): 1208-1223.

[25] V Doorn, William PTM, et al. "Machine learning-based glucose prediction with use of continuous glucose and physical activity monitoring data: The Maastricht Study." PloS one 16.6 (2021): e0253125.