An LSTM-based Approach Towards Automated Meal Detection from Continuous Glucose Monitoring in Type 1 Diabetes Mellitus

Maria Athanasiou
School of Electrical and Computer Engineering
National Technical University of Athens
Athens, Greece
mathanasiou@biosim.ntua.gr

Konstantia Zarkogianni
School of Electrical and Computer Engineering
National Technical University of Athens
Athens, Greece
kzarkog@biosim.ntua.gr

Konstantinos Karytsas
School of Electrical and Computer Engineering
National Technical University of Athens
Athens, Greece
konkarytsas@gmail.com

Konstantina S. Nikita
School of Electrical and Computer Engineering
National Technical University of Athens
Athens, Greece
knikita@ece.ntua.gr

2021 IEEE 21st International Conference on Bioinformatics and Bioengineering (BIBE) | 978-1-6654-4261-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/BIBE52308.2021.9635246

***Abstract*— Technological advancements in glucose sensing, insulin pumps, and closed-loop glucose control algorithms open new opportunities towards the realization of Artificial Pancreas (AP). However, the effective management of meal disturbances in these systems still remains a challenge. Meal detection algorithms eliminate the need for meal announcements and enable the shift to more automated and reliable AP systems. The aim of the present study is to develop and evaluate a personalized approach for the detection of meal disturbances in patients with Type 1 Diabetes Mellitus (T1DM). Long Short Term Memory Neural Networks (LSTM)'s inherent ability to efficiently handle sequential data is leveraged within an ensemble learning strategy towards the development of different versions of ensemble models. The models receive as input sequences of Continuous Glucose Monitoring (CGM) measurements (glucose profiles) of a 120-min duration and classify them as positive or negative for the onset of an ingested meal. *In silico* evaluation is performed using the UVA-PADOVA T1DM Simulator. All ensembles achieve acceptable discriminative performance (mean c-statistic: 75.12%-79.52%) and are able to detect meals in a timely manner (mean detection time: 7.08–12.84 min). Statistical analysis demonstrates the superiority of the simple averaging combination scheme over the other schemes in terms of the c-statistic.**

Keywords— Type 1 Diabetes Mellitus, artificial pancreas, meal detection, continuous glucose monitoring, machine learning, deep learning, ensemble learning

I. Introduction

Type 1 Diabetes Mellitus (T1DM) is a metabolic disorder that results from a chronic autoimmune destruction of the insulin-producing pancreatic beta cells, and is characterized by elevated blood glucose levels. Long-term hyperglycaemia due to the absence of insulin secretion is associated with the onset of macrovascular (coronary artery disease, peripheral arterial disease, and stroke) and microvascular (diabetic nephropathy, neuropathy, and retinopathy) complications [1]. The injurious effects of hyperglycaemia can be prevented through optimal glycaemic control, which involves regular glucose measurements and exogenous insulin administration [2].

Technological advances in glucose sensors and insulin pumps, along with algorithmic progress towards the automated estimation of appropriate insulin infusion rates, have brought forward the development of wearable Artificial Pancreas (AP), with the ultimate goal to enable effective T1DM management [3]. Despite the promising performance of closed-loop glucose control algorithms that has been reported in a wide range of studies, AP systems are still challenged by the inter- and intra-patient variability of glucose metabolism, the delays of subcutaneous glucose sensing and insulin infusion, and the presence of multiple disturbances such as meals and exercise [4]. When it comes to meal ingestion, in particular, postprandial glucose regulation constitutes one of the most arduous aspects of glycaemic control. The intra-day variations of insulin sensitivity along with the delay in subcutaneous insulin absorption with respect to meal-related glucose absorption, hamper the timely and efficient manifestation of the rapid-acting insulin's metabolic effect [5].

Along these lines, the majority of closed-loop glucose control systems have utilized a "semi-closed-loop" control strategy in order to handle meal disturbances, by incorporating meal announcements and generating advice on prandial insulin based on information about the meal size and the time of ingestion. In this context, patients are required to estimate the carbohydrate (CHO) content of the ingested meal, which is the main determinant of postprandial glucose rise [6][7]. However, CHO counting is a strenuous task for patients, limiting their degree of freedom throughout disease management, while also involving an estimated average error of 20%. Moreover, although the calculation of prandial insulin boluses solely considers CHO content, a meal's glycemic index and nutritional

Part of this research was supported within the framework of ENDORSE project, which is funded by the NSRF. Grant agreement: T1EΔK-03695.

TABLE I
VARIATION OF MEAL PROFILES IN THE SIMULATED SCENARIO

	Breakfast	Lunch	Snack	Dinner	Snack
CHO content (g)	[30,70]	[50,80]	[5,15]	[50,80]	[5,15]
Meal time (h)	[7:00, 8:00]	[12:00, 14:00]	[16:00, 17:00]	[18:00, 20:00]	[21:00, 23:00]

composition in terms of protein and lipids also have a significant impact on postprandial glucose concentrations [5].

Meal detection algorithms are viewed as a critical enabling component that addresses the above-mentioned challenges and eliminates the need for patient-inserted meal-related information, opening up the road towards more automated, reliable and robust AP systems. In this regard, several meal detection algorithms have been proposed, using measurements from continuous glucose monitoring (CGM) systems. Heuristic approaches, incorporating rules for threshold violations of glucose measurements and their rate of change have been widely investigated [8]. Noise rejection techniques, such as Finite Impulse Response (FIR) Filters and Kalman Filtering, are usually applied in these approaches for the reduction of glucose sensor noise [9][10]. Additionally, physiological models for the simulation of the glucose-insulin metabolism, including Bergman's minimal model and the Hovorka model, have been combined with the Unscented Kalman Filter (UKF) in order to enable the detection of threshold violations by the UKF states, indicating meal ingestion [11][12]. Such mathematical models have also been deployed in hybrid approaches, featuring probabilistic methods for the estimation of the glucose rate of appearance based on pre-defined meal shapes [13], as well as statistical methods for the identification of meal presence based on the comparison with the induced effect of previous meals [14]. Only a few studies have so far explored machine learning techniques towards meal detection. Fuzzy logic has been leveraged in a meal detection algorithm, incorporating a meal size estimation technique [7]. Linear Discriminant Analysis (LDA) has been recently utilized in a study, investigating the use of longer horizons of CGM measurements with different meal detection techniques [15].

In this work, a deep learning approach based on Long Short Term Memory Neural Networks (LSTM) is deployed towards the development of a personalized model for the detection of meal disturbances in T1DM patients. LSTM's inherent ability to handle sequential data is harnessed in order to identify complex patterns, associated with the onset of ingested meals, in 120-min sequences of CGM measurements (glucose profiles). An ensemble learning strategy is adopted by training differently configured LSTM-based individual models and investigating different combination schemes towards the creation of corresponding ensembles. To the best of the authors' knowledge, this is the first work proposing an LSTM-based meal detection algorithm.

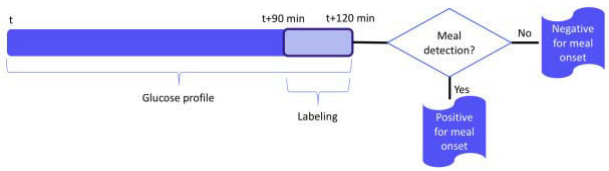

Figure 1. Glucose profiles' generation and labeling strategy.

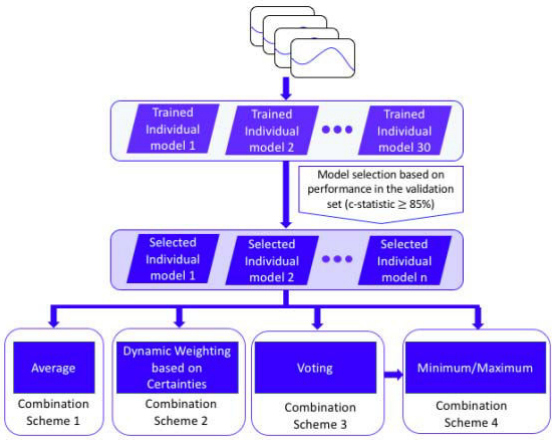

Figure 2. The applied ensemble learning approach. Four combination schemes were investigated, leading to four corresponding ensembles.

II. DATASET

The development and evaluation of the proposed personalized meal detection model was based on data generated by the UVA-PADOVA T1DM Simulator [16]. In particular, a seven-day scenario was simulated for the *in silico* populations of 10 adults, 10 adolescents, and 10 children, including varying meal ingestion timings and CHO meal content, as depicted in Table I.

III. METHODS

A. Data preparation: Generation of glucose profiles

Sequences of CGM measurements (glucose profiles) were used in order to compose the proposed model's input space [15]. More specifically, a sliding window of 120 min was applied to the simulated CGM data of the seven-day scenario, leading to the generation of glucose profiles of a 120-min duration. The 120-min window size was selected in order to enable the identification of glycaemic patterns across the entire spectrum of postprandial glucose response, that is typically characterized by a 2-3-hour duration [5]. Each glucose profile comprised 25 CGM measurements (recorded by a CGM system with a 5-min sampling rate), corresponding to a 120-min horizon. Subsequent glucose profiles overlapped by 24 measurements.

The generated glucose profiles were then assigned a label according to the presence, or not, of an ingested meal. Label assignment was based on prior knowledge about CGM and postprandial glucose response, according to which the time lag between changes in CGM and plasma glucose corresponds to roughly 15 minutes, while blood glucose levels rise approximately 10-15 minutes after the a meal ingestion [5]. Taking this into account, the last 7 CGM measurements of the generated glucose profiles, corresponding to the most recent 30 min of each profile, were considered towards the assignment of appropriate labels. Glucose profiles that included at least one CGM measurement, related to the onset of a meal, in their most recent 30-min period, were labeled as positive for a meal onset, otherwise they were labeled as negative for a meal onset. Eventually, 23% of the generated glucose profiles were

assigned to the "positive for a meal onset" class. The glucose profiles' generation and labeling strategy is shown in Fig. 1.

B. Ensemble learning approach

Fig. 2 depicts the conceptual framework of the proposed LSTM-based approach. In order to increase the model's discriminative ability, an ensemble learning method was applied. In particular, thirty LSTM-based individual models with different configurations in terms of hyperparameters' combinations (described in Section III.C) were trained on the same dataset. In this manner, a heterogeneous set of mapping functions was obtained, characterized by partially independent errors. The individual models which achieved a c-statistic score ≥ 85% in the validation set, were selected as the final components of the ensembles. Four schemes were investigated for combining the outputs of the selected individual models, resulting in four ensembles.

1) *Combination scheme 1:* The outputs of the individual models were simply averaged (*ensemble 1*).
2) *Combination scheme 2:* Dynamic weighted averaging based on certainties was deployed in order to obtain the final probability *(ensemble 2)* [17] .
3) *Combination scheme 3:* A voting scheme was applied to the outputs of the individual models. The optimal decision threshold of each individual model, based on the c-statistic, was used for the generation of the models' individual votes. In the case of split vote, glucose profiles were classified as negative for a meal onset with the aim of avoiding false detections (*ensemble 3)*.
4) *Combination scheme 4:* The maximum or minimum estimated probability generated by the individual models was applied, depending on whether the glucose profile was classified as positive or negative for a meal onset, based on combination scheme 3 (*ensemble 4*).

C. Long Short Term Memory Neural Networks

LSTM is a variant of Recurrent Neural Networks (RNNs), capable of learning order dependence in sequence prediction problems [18]. They provide a solution to the problem of vanishing gradients, which in the case of RNNs hampers learning of long data sequences. Given their inherent ability to handle sequential data efficiently, LSTMs have shown promising results in a wide range of classification and regression tasks. In contrast to the RNN cell, that is characterized by a simple structure, such as a single tanh layer, the structure of the LSTM cell includes four layers, which are called gates and enable the control of the information that is

TABLE II
VARIATION OF THE LSTM-BASED INDIVIDUAL MODELS' HYPERPARAMETERS

Hyperparameter	Value range
Learning rate	[0.0001, 0.01]
Number of hidden units (LSTM layer)	[100, 1000]
Maximum number of epochs	[50, 1000]
L2 Regularization	[0.0001, 0.01]
Dropout rate	[0, 1]

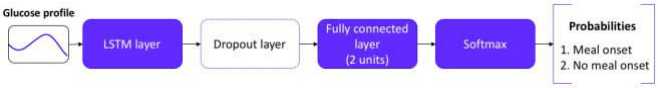

Figure 3. Architecture of the LSTM-based individual models.

added to or removed from the cell state C [19]. At first, the forget gate f determines which information from the input should be neglected by the cell state:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f). \quad (1)$$

where x_t and h_{t-1} represent the current input vector and the hidden vector of the previous state, respectively. Then, the input gate i decides which information can be used to update the cell state,

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (2)$$

and the cell input activation vector, including a set of new values that could be used to update the state, is created:

$$\tilde{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C). \quad (3)$$

These values are used to update the cell state from C_{t-1} to C_t based on the equation:

$$C_t = f_t \circ C_{t-1} + i_t \circ \tilde{C}_t . \quad (4)$$

The output gate o generates the output of the LSTM cell and updates the hidden vector from h_{t-1} to h_t, according to the following equations:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \circ \tanh(C_t). \quad (6)$$

In equations (1) to (6), σ is the sigmoid activation function, and W_x and b_x ($x = f, i, C$) are weight matrices and bias vector parameters, respectively.

In the present study, a many-to-one LSTM architecture was utilized, consisting of one LSTM layer, one fully connected layer, and a softmax layer, as shown in Fig. 3. A dropout layer was applied during training to prevent overfitting. Thirty different combinations of hyperparameters were generated, and were deployed within this architecture towards training 30 LSTM-based individual models. Table II summarizes the hyperparameters that were differentiated among the individual models during training. Bayesian optimization was applied in order to identify the hyperparameters' combinations that led to models with a c-statistic score ≥ 85% in the validation set [20]. These combinations were eventually used for training the LSTM-based individual models, that comprised the ensembles 1 to 4.

IV. RESULTS AND DISCUSSION

For each of the *in silico* patients, data corresponding to 70% of the simulated days were used for training purposes, while the remaining 30% were used for testing. During the training phase, a validation set corresponding to 15% of the training data was utilized. The ensembles' discriminative performance was assessed in terms of the c-statistic, Accuracy (ACC), Sensitivity (SENS), and Specificity (SPEC). ACC, SENS and SPEC were calculated by setting the probability threshold equal to the optimal threshold obtained by the c-statistic.

Table III summarizes the results obtained by ensembles 1 to 4 for the *in silico* populations of adults, adolescents and children. The combination schemes 1, 2, and 4 achieved acceptable

TABLE III
PERFORMANCE METRICS FOR ENSEMBLES 1 TO 4 IN THE *IN SILICO* PATIENTS' POPULATION OF UVA-PADOVA T1DM SIMULATOR

		Mean±SD					
		c-statistic (%)	**ACC (%)**	**SENS (%)**	**SPEC (%)**	**Mean meal detection time (min)**	**Mean percentage of detected meals (%)**
Ensemble 1	**Adults**	75.29±6.71	69.56±5.18	68.95±8.12	69.76±6.62	11.51±3.89	97.08±4.78
	Adolescents	79.52±4.01	71.27±5.66	77.69±8.11	69.25±8.39	7.99±4.67	99.01±3.13
	Children	79.05±4.76	70.84±5.87	76.29±8.44	69.12±9.01	8.09±5.71	98.0±4.18
	All	**77.95±5.45**	70.56±5.43	74.31±8.84	69.38±7.79	9.20±4.94	98.02±4.03
Ensemble 2	**Adults**	75.19±6.74	69.51±5.11	68.32±7.95	69.89±6.47	12.84±3.22	99.00±3.13
	Adolescents	79.49±4.01	71.16±5.43	77.62±8.28	69.12±8.00	8.08±4.65	99.00±3.13
	Children	78.95±4.81	70.77±6.26	77.06±10.25	68.79±10.27	8.21±5.50	98.02±4.18
	All	77.88±5.48	70.48±5.47	74.34±9.60	69.27±8.11	9.71±4.94	**98.68±3.43**
Ensemble 3	**Adults**	-	70.70±3.88	68.18±9.93	71.50±4.96	10.98±5.43	96.20±6.52
	Adolescents	-	71.46±5.27	78.46±5.67	69.25±7.41	7.08±2.92	99.01±3.13
	Children	-	72.09±4.12	74.76±6.22	71.26±5.70	9.42±3.52	98.02±4.18
	All	-	71.42±4.35	73.80±8.44	70.67±5.99	9.16±4.28	97.74.±4.80
Ensemble 4	**Adults**	75.12±7.12	71.06±3.20	71.05±12.50	71.06±5.20	10.21±5.48	96.20±6.52
	Adolescents	77.80±4.09	72.96±5.88	76.99±6.56	71.70±8.50	7.72±3.42	99.01±3.13
	Children	78.24±4.34	72.70±5.55	75.80±6.25	71.72±8.29	8.92±3.12	98.02±4.18
	All	77.06±5.36	**72.24±4.92**	**74.62±8.99**	**71.49±7.23**	**8.95±4.13**	97.74±4.80

discriminative performance (mean c-statistic: 75.12%-79.52%), while the application of different combination schemes led to varying levels of sensitivity and specificity. Moreover, all ensembles were able to detect high percentages of meals (mean percentage of detected meals: 96.20%-99.01%) in a timely manner (mean detection time: 7.08-12.84 min). All ensembles showed higher discriminative ability in children and adolescents compared to the adult population. This is justified by the higher glucose fluctuations observed in the populations of children and adolescents in the presence of meal disturbances.

The pairwise t-test was applied in order to compare the performance obtained by the different ensembles on the whole population, including adults, adolescents and children. Statistically significant differences were revealed in terms of the c-statistic: ensemble 1 vs ensemble 2 (p-value = 0.001), ensemble 1 vs ensemble 4 (p-value = 0.004) and ensemble 2 vs ensemble 4 (p-value = 0.007). Moreover, statistically significant differences were obtained with respect to accuracy: ensemble 1 vs ensemble 4 (p-value = 0.005), ensemble 2 vs ensemble 4 (p-value = 0.004) and ensemble 3 vs ensemble 4 (p-value = 0.027). In terms of specificity, statistically significant differences were observed between ensemble 1 and ensemble 4 (p-value = 0.044). There were no statistically significant differences in terms of sensitivity, mean meal detection time, and mean percentage of detected meals. Taking into consideration the unbalanced nature of the dataset (23% positive for meal onset glucose profiles), the statistically significant differences in terms of the c-statistic demonstrated the superiority of ensemble 1 over the other ensembles.

V. CONCLUSION

In this study, an LSTM-based approach towards the development of a meal detection algorithm was presented. An appropriate data preparation strategy was applied towards the generation and labeling of glucose profiles which composed the models' input space. Ensemble learning techniques were implemented in order to increase the models' discriminative ability, while different combination schemes were investigated and comparatively assessed. The models' performance was *in silico* evaluated. The obtained results indicated the potential of the ensembles to perform accurate and timely detection of meal disturbances. Future work concerns the validation of the proposed ensembles on data gathered from T1DM patients under real-life conditions.

REFERENCES

[1] M. J. Fowler, "Microvascular and macrovascular complications of diabetes," *Clinical Diabetes*, vol. 26, no. 2. American Diabetes Association, pp. 77–82, Apr. 2008.

[2] "IDF Atlas 9th edition and other resources." [Online]. Available: https://www.diabetesatlas.org/en/resources/. [Accessed: 03-May-2021].

[3] K. Zarkogianni *et al.*, "A Review of Emerging Technologies for the Management of Diabetes Mellitus," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 12, pp. 2735–2749, Dec. 2015.

[4] S. Kapil, R. Saini, S. Wangnoo, and S. Dhir, "Artificial Pancreas System for Type 1 Diabetes—Challenges and Advancements," *Explor. Res. Hypothesis Med.*, vol. 000, no. 000, pp. 1–11, Jul. 2020.

[5] A. El Fathi, M. Raef Smaoui, V. Gingras, B. Boulet, and A. Haidar, "The Artificial Pancreas and Meal Control: An Overview of Postprandial Glucose Regulation in Type 1 Diabetes," *IEEE Control Syst.*, vol. 38, no. 1, pp. 67–85, Feb. 2018.

[6] K. Zarkogianni, A. Vazeou, S. G. Mougiakakou, A. Prountzou, and K. S. Nikita, "An insulin infusion advisory system based on autotuning nonlinear model-predictive control," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 9, pp. 2467–2477, Sep. 2011.

[7] M. Zheng, B. Ni, and S. Kleinberg, "Automated meal detection from continuous glucose monitor data through simulation and explanation," *J. Am. Med. Informatics Assoc.*, vol. 26, no. 12, pp. 1592–1599, Dec. 2019.

[8] E. Dassau, B. W. Bequette, B. A. Buckingham, and F. J. Doyle, "Detection of a meal using continuous glucose monitoring: Implications for an artificial β-cell," *Diabetes Care*, vol. 31, no. 2, pp. 295–300, Feb. 2008.

[9] H. Lee and B. W. Bequette, "A Closed-loop Artificial Pancreas based on MPC: human-friendly identification and automatic meal disturbance rejection," *IFAC Proc. Vol.*, vol. 41, no. 2, pp. 4252–4257, Jan. 2008.

[10] R. A. Harvey, E. Dassau, H. Zisser, D. E. Seborg, and F. J. Doyle, "Design of the glucose rate increase detector: A meal detection module for the health monitoring system," *J. Diabetes Sci. Technol.*, vol. 8, no. 2, pp. 307–320, 2014.

[11] Z. Mahmoudi, K. Nørgaard, N. K. Poulsen, H. Madsen, and J. B. Jørgensen, "Fault and meal detection by redundant continuous glucose monitors and the unscented Kalman filter," *Biomed. Signal Process. Control*, vol. 38, pp. 86–99, Sep. 2017.

[12] K. Turksoy, S. Samadi, J. Feng, E. Littlejohn, L. Quinn, and A. Cinar, "Meal detection in patients with type 1 diabetes: A new module for the

multivariable adaptive artificial pancreas control system," *IEEE J. Biomed. Heal. Informatics*, vol. 20, no. 1, pp. 47–54, Jan. 2016.

[13]F. Cameron and G. Niemeyer, "Predicting blood glucose levels around meals for patients with type i diabetes," in *ASME 2010 Dynamic Systems and Control Conference, DSCC2010*, 2010, vol. 1, pp. 289–296.

[14]J. Weimer, S. Chen, A. Peleckis, M. R. Rickels, and I. Lee, "Physiology-Invariant Meal Detection for Type 1 Diabetes," *Diabetes Technol. Ther.*, vol. 18, no. 10, pp. 616–624, Oct. 2016.

[15]K. Kölle, T. Biester, S. Christiansen, A. L. Fougner, and Ø. Stavdahl, "Pattern Recognition Reveals Characteristic Postprandial Glucose Changes: Non-Individualized Meal Detection in Diabetes Mellitus Type 1," *IEEE J. Biomed. Heal. Informatics*, vol. 24, no. 2, pp. 594–602, Feb. 2020.

[16]C. Dalla Man, 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.

[17]K. Zarkogianni, M. Athanasiou, and A. C. Thanopoulou, "Comparison of Machine Learning Approaches Toward Assessing the Risk of Developing Cardiovascular Disease as a Long-Term Diabetes Complication," *IEEE J. Biomed. Heal. Informatics*, vol. 22, no. 5, pp. 1637–1647, Sep. 2018.

[18]S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[19]"Understanding LSTM Networks -- colah's blog." [Online]. Available: https://colah.github.io/posts/2015-08-Understanding-LSTMs/. [Accessed: 03-May-2021].

[20]P. I. Frazier, "A Tutorial on Bayesian Optimization," *arXiv*, Jul. 2018.