

## Data-Driven Classifiers for Early Meal Detection Using ECG

Muhammad A. Cheema<sup>1</sup>, Pallavi Patil<sup>2</sup>, Salman I. Siddiqui<sup>1</sup>, Pierluigi Salvo Rossi<sup>1,\*</sup>, Øyvind Stavadahl<sup>2,\*\*</sup>, and Anders Lyngvi Fougner<sup>2,\*\*</sup>

<sup>1</sup>Department of Electronic Systems, Norwegian University of Science and Technology, Trondheim, Norway

<sup>2</sup>Department of Engineering Cybernetics, Norwegian University of Science and Technology, Trondheim, Norway

\* Senior Member, IEEE; \*\* Member, IEEE

**Abstract**—This study investigates the potential of the electrocardiogram (ECG) to perform early meal detection, which is critical for developing a fully-functional automatic artificial pancreas. The study was conducted in a group of healthy subjects with different ages and genders. Two classifiers were trained: one based on neural networks (NNs) and working on features extracted from the signals; one based on convolutional NNs (CNNs) and working directly on raw data. During the test phase, both classifiers correctly detected all the meals, with the CNN outperforming the NN in terms of misdetected meals (MMs) and detection time (DT). Reliable meal onset detection with short detection time has significant practical implications: it reduces the risk of postprandial hyperglycemia and hypoglycemia, it reduces the mental burden of meal documentation for patients and related stress.

**Index Terms**—Classification, meal detection, neural networks, diabetes mellitus type 1, electrocardiogram.

### I. INTRODUCTION

Type 1 diabetes mellitus (DM1) is a chronic medical condition that affects millions of individuals worldwide. It is characterized by insufficient production of insulin by the pancreas, leading to high blood glucose level (BGL) which can cause sustained impairments if untreated [1]. Maintaining BGL within normal range is imperative for the proper functioning of the human body, thus DM1 patients require external insulin. Administering insulin manually can be a meticulous and time-intensive procedure. Moreover, manual insulin dosing may not be a practical solution for children and elderly patients, who may struggle to either remember or self-administer the insulin injections [2]. Continuous subcutaneous insulin injection (CSII) needs frequent manual interventions which can add to stress of DM1 patients.

A hybrid artificial pancreas (HAP) can control BGL and comprises of three key components: (i) continuous glucose monitoring (CGM) system measuring BGL; (ii) insulin pump to deliver insulin; and (iii) control system implemented in the insulin pump [3]. CGM-based HAPs measure BGL in interstitial fluid, not capillary glucose, resulting in a physiologic lag ranging from 5 min to 15 min, depending on glucose-change rate [4]. To reduce postprandial glucose excursions, meal bolus insulin should be infused prior to meal start. The HAP depends on manual input (by the user) for meal-related information (e.g. meal start and content), increasing management burden.

Automatic meal detection in an artificial pancreas enables more precise and timely insulin dosing, ideally by reducing/eliminating

manual meal-related input for patients and addressing challenges related to latency and variability in glucose response. Several approaches for automatic meal detection have been recently investigated: incorporating meal estimation based on CGM systems [5], using abdominal-sound based methods to predict meals [6], [7] and using hybrid models that combine model-based and data-driven approaches [8], [9], [10]. As for detection time, the hybrid model requires approximately 10 min whereas abdominal-sound methods ranges within 5 min and 10 min [6], [7].

ECG-based meal-onset detection represents a robust alternative or supplement to the approaches based on abdominal sounds. Unlike abdominal sounds, ECG signals are unaffected by ambient sound, thus offering better reliability. Meal intake evokes the sympathetic activation of the cardiovascular system when preparing the body to digest the incoming meal [11]. As a result, the heart rate and the amount of blood pumped from the heart to the body increase [12]. Alterations are also seen in ECG parameters (e.g. shape and size of T-wave, duration of QT interval, spectral powers associated with sympathetic-parasympathetic activity) [11], [13], [14].

In this study, classifiers based on neural networks (NNs) and on convolutional NNs (CNNs) are trained, tested and compared using ECG signals. The reason for selecting NNs and CNNs is that the former are among the most simple and basic machine-learning method (thus representing a natural baseline), while the latter have shown excellent combination in terms of performance, complexity and training stability for time-series processing. A relevant overview on recent data-driven methodologies for time series processing is found in [15]. The paper is organized as follows: Sec. II describes the hardware and the procedure for collecting the ECG signals; Sec. III presents the data pre-processing for feature extraction from the ECG signals and the classifiers used for meal detection; Sec. IV illustrates the results; and finally Sec. V concludes the paper.

### II. DATA ACQUISITION

The study is based on 24 recordings from 12 healthy volunteers (at most 4 recordings per subject) without any reported history of cardiovascular or gastrointestinal disorders. The participants were instructed to consume breakfast 4 hours before the experiment and abstain from consuming caffeine 1 hour before the start of the recording. The study was approved by the Regional Committee for Medical and Health Research Ethics (REK-midt approval no. 84374) and was conducted at the Norwegian University of Science and Technology, Trondheim, Norway.

During the experiment, a three-lead bipolar ECG was recorded by placing a yellow electrode and a black electrode on the right and left collar bone of the chest, respectively, while a blue electrode was positioned above the umbilicus, as shown in Fig. 1. The ECG

Corresponding author: M.A. Cheema (e-mail: asaad.cheema@ntnu.no).

Associate Editor: F. Falcone.

This work was partially supported by the Research Council of Norway under the project ML4ITS within the IKTPLUSS framework and under grant no. 294828 through the Centre for Digital Life Norway (digitallifenorway.org).

Digital Object Identifier 10.1109/LENS.2023.3307106

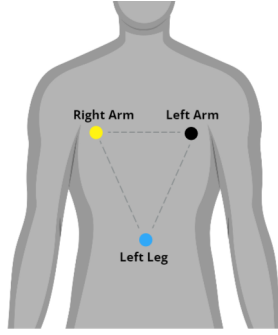


Fig. 1: The electrode placement on body.

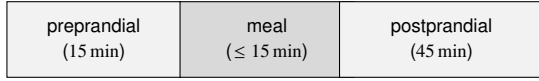


Fig. 2: Timeline for the recording session.

signal was recorded with a sampling frequency of 2048 Hz using software BioGraph Infiniti T7500M and ProComp Infiniti SA7500 encoder (ThoughtTech, Montreal, Quebec, Canada). The participants remained seated during the entire procedure, and ECG signals were recorded starting 15 min before meal intake and ending 45 min after the meal was finished. The duration of the meal was approximately 15 min. The protocol is depicted in Fig. 2.

### III. DATA PROCESSING

A pre-processing phase mitigates noise in the ECG signal by using a 4th-order low-pass Butterworth filter with cutoff frequency 40 Hz.

#### A. NN-Based Classifier

The first classifier is based on a fully-connected (FC) NN architecture processing features extracted from ECG signals.

After low-pass filtering, the initial part (approximately 30 min) of each recording is considered, with the start time of the meal occurring after about 15 min (in all cases between 12 min and 21 min). The truncated signals are downsampled to 256 Hz and both original and downsampled versions are segmented into  $S = 359$  consecutive *frames* with 10 s duration and 50% overlap for features extraction. A total of  $N = 27$  features per frame are extracted including higher order statistics, temporal-domain and spectral-domain representations. Statistical features like kurtosis, skewness, energy and variance are calculated from the frames of the original signal. ECG signals contain relevant information related to five peaks called *fiducial points* (namely P, Q, R, S, and T), and several features are related to the amplitude and the (relative) temporal location of the fiducial points. The R wave was detected using the Pan-Tompkins algorithm [16]. Spectral features related to energy and energy ratios in various frequency bands (very low: 0.003–0.04 Hz, low: 0.04–0.15 Hz and high: 0.15–0.5 Hz) of heart rate variability are also computed from the frames of the downsampled version.

A *feature matrix*  $\mathbf{X} \in \mathbb{R}^{S \times N}$  is created for each recording to be used as the input to the classifier. When preparing the feature matrix for the NN-based classifier, min-max scaling is applied to each feature column for data normalization followed by median filtering for outliers removal. A binary *response vector*  $\mathbf{y} \in \mathbb{R}^{S \times 1}$  is introduced

as the output of the classifier for training, where the  $s$ th entry is 0 (resp. 1) if the  $s$ th frame happens before (resp. after) the meal start.

The NN-based classifier is made of:

- Input layer with  $N$  nodes;
- FC layer with 100 nodes, followed by Leaky Re-Lu activation function;
- FC layer with 50 nodes, followed by Leaky Re-Lu activation function and a dropout layer (to prevent over-fitting);
- FC layer with only one output, followed by Sigmoid activation function.

For training purposes, we utilized the ADAM optimizer [17] with a learning rate of  $10^{-5}$  and employed the mean squared error (MSE) as the objective function.

#### B. CNN-based Classifier

The second classifier is based on a 1D CNN architecture directly processing raw data<sup>1</sup>.

After low-pass filtering, min-max scaling is introduced for data normalization, and finally downsampling by a factor of 8 is performed to match the frequency of 256 Hz and speed up successive computations. Frames with 10 s duration and 50% overlap for building feature matrices are identified and analogous response vector as in the case of the NN classifier are introduced. In this case, each column of the feature matrix contains raw signal data.

The overall structural details of the considered CNN are:

- 1D Conv with 16 output channels, a kernel size equal to 512 and stride 2, followed by Leaky Re-Lu activation function;
- 1D Conv with 8 output channels, a kernel size equal to 512 with stride 2 followed by Leaky Re-Lu activation function;
- Max pooling with a window of size equal to 5 and stride 2;
- FC layer with output size equal to 100, followed by Leaky Re-Lu activation function, and a dropout layer;
- FC layer with only one output, followed by Sigmoid activation function.

For training purposes, we utilized the ADAM optimizer with a learning rate of  $10^{-4}$  and employed MSE as the objective function.

#### C. Classification

In this work, 24 recordings from 12 volunteers are utilized for the purposes of training, validation, and testing: 16 of these are designated for training and validation, employing the leave-one-out cross-validation (LOOCV) method, while the remaining 8 are reserved for testing (3 subjects were in common in all sets). Selecting 1 meal from the 16 options results in 16 distinct combinations or folds, which are employed to train and validate the classification models more effectively. Also, for improved robustness, we selected 5 trained models (based on their performance in terms of detection time, we excluded the best and the worst models) and perform detection based on a fusion strategy among those models.

More specifically, each classifier (NN-based and CNN-based) provides a soft decision  $d_s \in [0, 1]$  about the  $s$ th frame. The soft output from the classifier is filtered according to an Exponentially Weighted Moving Average (EWMA) for an individual (hopefully more reliable) decision. EWMA [20] is a sequential change detection

<sup>1</sup>Over the past two decades, CNNs have demonstrated their effectiveness in extracting features from raw data in several applications [18] [19].

procedure that exploits past observations and is applied to reduce the number of false alarms. EWMA relies on the following equation:

$$z_s = \alpha d_s + (1 - \alpha)z_{s-1}, \quad (1)$$

where  $\alpha$  is a parameter determining a trade-off between current and past values from the classifier. The outputs ( $z_s$ ) from the 5 classifiers are then averaged and the result is converted to a final binary decision based on a threshold mechanism.

#### D. Performance Metrics

The performance of two classifiers is assessed in terms of: true positive rate (TPR) and false positive rate (FPR), from a sample-by-sample perspective; number of misdetected meals (MMs), number of undetected meals (UMs), and detection time (DT), from a meal-event perspective. DT is computed as the difference in the meal start between the actual meal given by the response vector and the predicted vector from the trained model. The meal start is defined as the time instant when the label vector transits from 0 to 1. UMs happen when the actual label from the response vector shows the presence of the meal, but the classifier fails to detect it. MMs happen when the meal is predicted by the classifier, while the actual labels from the response vector shows no meal.

### IV. NUMERICAL RESULTS

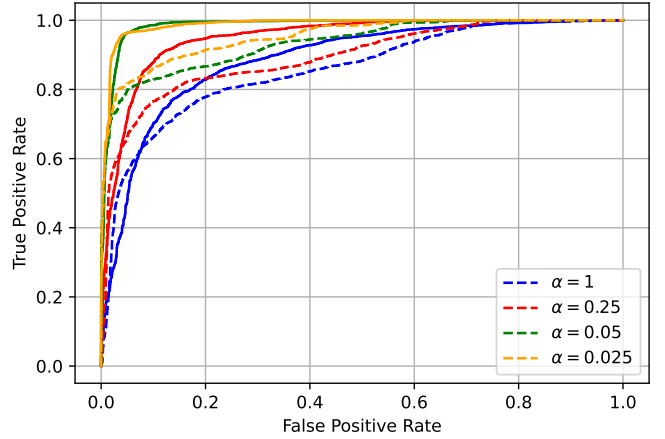
This section presents the performance results of the two classifiers, both implemented in python using Pytorch package. The Receiver Operating Characteristic (ROC) curve is computed for both the LOOCV and test sets. The parameter  $\alpha$  in Eq. (1) impacts the overall performance such that both the TPR and the DT decrease with it, for a fixed FPR, until saturation. Negligible FPR is prioritized to avoid hypoglycemia, thus comparisons with a meal-event perspective are made with a threshold providing FPR approximately equal to 0.015.

Fig. 3 shows that both classifiers exhibit interesting performance from a sample-by-sample perspective. Fig. 3(a) shows the performance during validation for different values of  $\alpha$ , with the CNN outperforming the NN when operating at low FPR. Assuming a probability of false detection equal to 0.1 (and excluding the case with  $\alpha = 1$ ), the CNN-based classifier achieves probabilities of detection in the range (0.9, 1), while the NN-based classifier in the range (0.7, 0.9). Fig. 3(b) shows the performance during testing with  $\alpha = 0.05$ . Apparently, the CNN-based classifier is significantly better than the NN-based classifier on the test set: when operating at FPR up to 0.2, the range for TPR increases from (0.3, 0.8) to (0.6, 0.9).

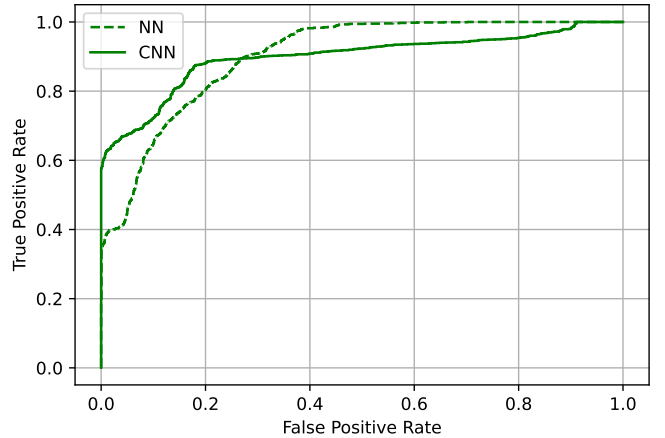
The impact of the parameter  $\alpha$  is also shown in Fig. 4 where the meal-event perspective is considered. Performance metrics are shown for different values of  $\alpha$ , with  $\alpha = 0.05$  being a desirable solution with minimized numbers of MM and UM, and reduced DT. It is worth noticing that the calculation of the DT only includes correctly-detected meals.

Fig. 5 shows the performance in terms of DT on the test set of the selected classifiers (with  $\alpha = 0.05$  and  $FPR \approx 0.015$ ). Again, the CNN-based classifier outperforms the NN-based classifier on the test set: the DT for every test meal is kept below 4 min only for the CNN case. The negative bars represent MM, showing how early they are declared before the meal actually starts.

Table 1 summarizes all the experimental results. Both classifiers perform similarly, but the CNN-based classifier exhibits better



(a) Validation



(b) Testing ( $\alpha = 0.05$ )

Fig. 3: ROCs for NN (dashed lines) and CNN (solid lines) classifiers.

Table 1: Meal-Detection Result Summary

Classifiers	CNN	NN
# Training Meals (Each Training Term)	15	15
# Validation Meals (Each Training Term)	1	1
# Test Meals	8	8
# Misdetected Meals (MMs) - (LOOCV)	1	3
# Undetected Meals (UMs) - (LOOCV)	0	2
Average Detection Time (DT) in min - (LOOCV)	1.87	1.51
# Misdetected Meals (MMs) - (Test)	0	3
# Undetected Meals (UMs) - (Test)	0	0
Average Detection Time (DT) in min - (Test)	1.83	2.63

generalization capabilities than the NN-based classifier (both from sample-by-sample and meal-event perspectives) as shown during the performance comparison on the test set. Also, CNNs operate directly on raw data, thus not requiring features extraction and selection. We remark that the overall achieved performance is better than previously-published solutions based on CGM or sound, which exhibit lower detection capabilities ( $TPR < 0.5$ ) and longer DT in the order of 10 – 40 min [8], [9] and 4 – 10 min [6], [7], [21], respectively.

### V. CONCLUSION

Automated early meal detection is crucial for insulin infusion in an artificial pancreas. We compared two data-driven classifiers, based on

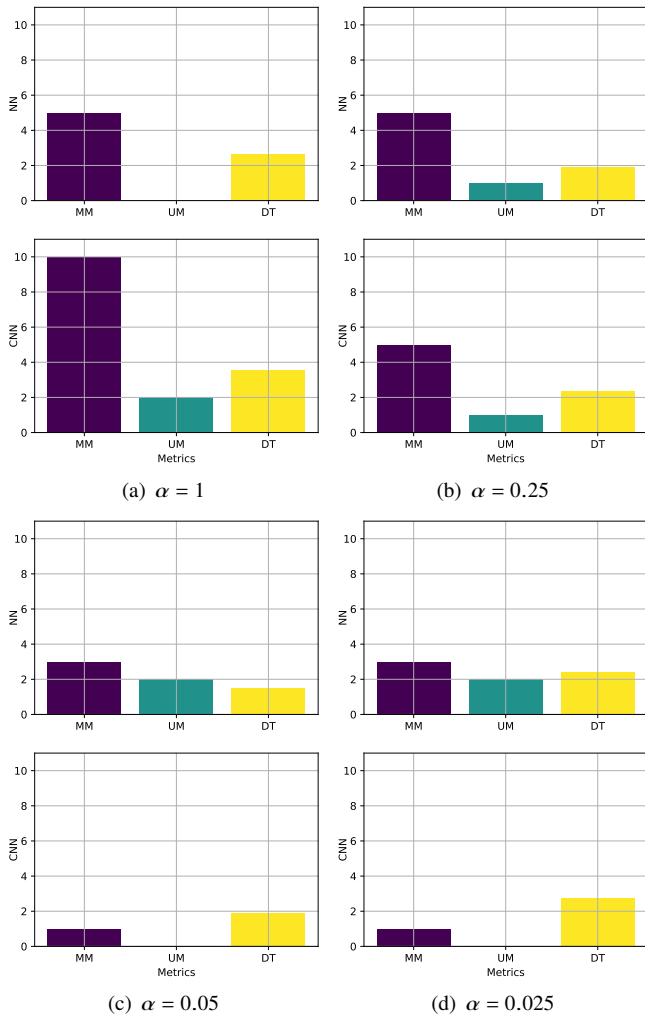


Fig. 4: Impact of  $\alpha$  on MMs, UMs and DT [min] in validation meals.

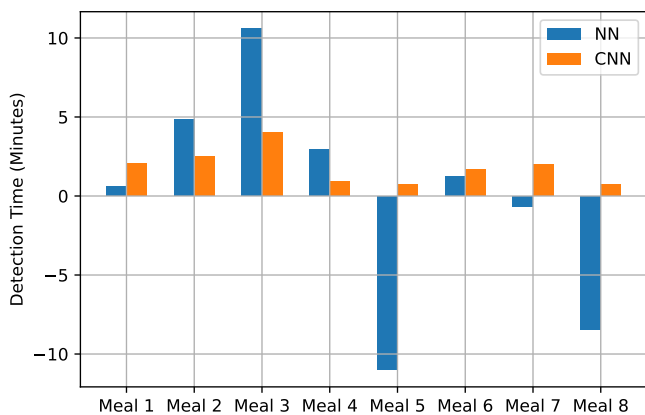


Fig. 5: DT performance on the test set ( $\alpha = 0.05$  and  $FPR \approx 0.015$ ).

NNs and on CNNs, for early meal detection using ECG signals from real-world experiments. EWMA is introduced as a tuning mechanism to reduce false alarms and detection delay. Compared with existing algorithms, based on CGM and sound, the proposed approach appears faster ( $DT < 2$  min vs.  $> 4$  or 10 min) and safer ( $TPR \approx 0.7$  vs. 0.5), thus potentially providing better glucose control in DM1 patients.

## REFERENCES

- [1] S. Yagihashi, H. Mizukami, and K. Sugimoto. Mechanism of diabetic neuropathy: Where are we now and where to go? *J. Diabetes Investigation*, 2:18–32, 2011.
- [2] J. Burdick, H. P. Chase, R. H. Slover, Kerry Kniewel, Laura Scrimgeour, Aristides K Maniatis, and Georgeanna J Klingensmith. Missed insulin meal boluses and elevated hemoglobin A1c levels in children receiving insulin pump therapy. *Pediatrics*, 113:e221–e224, 2004.
- [3] A. Saunders, L. H. Messer, and G. P. Forlenza. Minimed 670G hybrid closed loop artificial pancreas system for the treatment of type 1 diabetes mellitus: Overview of its safety and efficacy. *Expert Review of Medical Devices*, 16(10):845–853, 2019.
- [4] Howard A Wolpert. Use of continuous glucose monitoring in the detection and prevention of hypoglycemia, 2007.
- [5] M. Zheng, B. Ni, and S. Kleinberg. Automated meal detection from continuous glucose monitor data through simulation and explanation. *J. American Medical Informatics Association*, 26(12):1592–1599, 2019.
- [6] K. Kölle, A. L. Fougner, R. Ellingsen, S. M. Carlsen, and Ø. Stavdahl. Feasibility of early meal detection based on abdominal sound. *IEEE J. Translational Engineering in Health and Medicine*, 7:1–12, 2019.
- [7] M. A. Cheema, S. I. Siddiqui, and P. Salvo Rossi. Comparison of different classifiers for early meal detection using abdominal sounds. In *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pages 420–424, 2022.
- [8] 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. Biomedical and Health Informatics*, pages 594–602, 2019.
- [9] 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. Biomedical and Health Informatics*, 20:47–54, 2015.
- [10] R. A. Harvey, E. Dassau, H. Zisser, D. E. Seborg, and F. J. Doyle III. Design of the glucose rate increase detector: A meal detection module for the health monitoring system. *J. Diabetes Science and Technology*, 8(2):307–320, 2014.
- [11] C.-L. Lu, X. Zou, W. C. Orr, and J. D. Z. Chen. Postprandial changes of sympathovagal balance measured by heart rate variability. *Digestive Diseases and Sciences*, 44(4):857–861, 1999.
- [12] E. Widerlöv, K.-G. Jostell, L. Claesson, B. Odlind, M. Keisu, and U. Freyschuss. Influence of food intake on electrocardiograms of healthy male volunteers. *European J. Clinical Pharmacology*, 55(9):619–624, 1999.
- [13] K. Hnatkova, D. Kowalski, J. J. Keirns, E. M. van Gelderen, and M. Malik. Qtc changes after meal intake: Sex differences and correlates. *J. Electrocardiology*, 47(6):856–862, 2014.
- [14] J. Täubel, G. Ferber, L. Van Langenhoven, T. Del Bianco, S. Fernandes, D. Djumanov, J. K. Kanters, C. Graff, and A. J. Camm. The cardiovascular effects of a meal: J-Tpeak and Tpeak-Tend assessment and further insights into the physiological effects. *J. of Clinical Pharmacology*, 59(6):799–810, 2019.
- [15] M. A. Belay, S. S. Blakseth, A. Rasheed, and P. Salvo Rossi. Unsupervised anomaly detection for IoT-based multivariate time series: Existing solutions, performance analysis, and future directions. *Sensors*, 23(5):1–24, 2023.
- [16] J. Pan and W. J. Tompkins. A real-time QRS detection algorithm. *IEEE Trans. Biomedical Eng.*, 32(3):230–236, 1985.
- [17] D. P. Kingma and J. Ba. ADAM: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [18] J. Brownlee. 1D convolutional neural network models for human activity recognition. *Machine Learning Mastery. Retrieved July*, 26:2021, 2018.
- [19] Q. Shen, H. Qin, K. Wei, and G. Liu. Multiscale deep neural network for obstructive sleep apnea detection using RR interval from single-lead ECG signal. *IEEE Trans. Instrum. Meas.*, 70:1–13, 2021.
- [20] L. Xie, S. Zou, Y. Xie, and V. V. Veeravalli. Sequential (quickest) change detection: Classical results and new directions. *IEEE J. Sel. Areas Inf. Theory*, 2(2):494–514, 2021.
- [21] T. S. Sunilkumar, E. Söiland, Ø. Stavdahl, and A. L. Fougner. Pilot study of early meal onset detection from abdominal sounds. In *2019 E-Health and Bioengineering Conf. (EHB)*, pages 1–4, 2019.