Obesity Surgery (2020) 30:2547–2557
https://doi.org/10.1007/s11695-020-04511-6

ORIGINAL CONTRIBUTIONS

Artificial Neural Network-Based Automatic Detection of Food Intake for Neuromodulation in Treating Obesity and Diabetes

Elisabeth R. M. Heremans[1,2] · Amy S. Chen[3] · Ximeng Wang[1] · Jiafei Cheng[1] · Feng Xu[4] · Agustin E. Martinez[2] · Georgios Lazaridis[2] · Sabine Van Huffel[2] · Jiande D. Z. Chen[1]

Published online: 26 February 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Purpose Neuromodulation, such as vagal nerve stimulation and intestinal electrical stimulation, has been introduced for the treatment of obesity and diabetes. Ideally, neuromodulation should be applied automatically after food intake. The purpose of this study was to develop a method of automatic food intake detection through dynamic analysis of heart rate variability (HRV).

Materials and Methods Two experiments were conducted: (1) a small sample series with a standard test meal and (2) a large sample series with varying meal size. Electrocardiograms (ECGs) were collected in the fasting and postprandial states. Each ECG was processed to compute the HRV. For each HRV segment, time- and frequency-domain features were derived and used as inputs to train and test an artificial neural network (ANN). The ANN was trained and tested with different cross-validation methods.

Results The highest classification accuracy reached with leave-one-subject-out-leave-one-sample-out cross-validation was 0.93 in experiment 1 and 0.88 in experiment 2. Retraining the ANN on recordings of a subject drastically increased the achieved accuracy for that subject to values of 0.995 and 0.95 in experiments 1 and 2, respectively.

Conclusions Automatic food intake detection by ANNs, using features from the HRV, is feasible and may have a great potential for neuromodulation-based treatments of meal-related disorders.

Keywords Food intake detection · Obesity · Diabetes · Neuromodulation · Machine learning

Abbreviations

ANN	Artificial neural networks
FGIDs	Functional gastrointestinal disorders
HRV	Heart rate variability

✉ Jiande D. Z. Chen
jchen184@jhu.edu

[1] Departments of Medicine and Biomedical Engineering, Johns Hopkins University, School of Medicine, Baltimore, MD, USA

[2] Department of Electrical Engineering-ESAT, STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, KU Leuven, B-3001 Leuven, Belgium

[3] College of Letters and Science, University of California, Berkeley, CA, USA

[4] Yinzhou Hospital Affiliated with Ningbo University School of Medicine, Ningbo, China

Introduction

Many of the most prevalent health problems in the world are associated with food intake, digestion, and nutrient absorption. These problems include obesity, diabetes, and functional gastrointestinal disorders (FGIDs). Around 31% of adults in the United States (US) are obese; they are at a high risk of numerous chronic and sometimes fatal diseases, such as coronary heart diseases, type 2 diabetes, hypertension, stroke, and cancers [1, 2]. In addition, obesity represents a substantial economic impact in healthcare, with an estimated total annual cost over $275 billion in the US [3]. Diabetes affects 12.2% of US adults as of 2015, costing $245 billion in the United States in 2012 [4]. Diabetes increases the risk for numerous other conditions, such as cardiac diseases, stroke, kidney diseases, and neuropathy [5]. Finally, FGIDs affect up to 25% of Americans as of 2006 [6]. The combined direct and indirect cost of irritable bowel syndrome alone reached $21.9 billion in the US in 2002 [7]. Because FGIDs are chronic, patients experience persistent symptoms and have a decreased quality of life [6, 7].

Traditional treatments of obesity include behavior change, pharmacological treatments, and bariatric surgery. The only treatment with long-term effectiveness is bariatric surgery, a risky and invasive procedure [1, 8]. Diabetes is currently treated through insulin and oral medications, often in conjunction with lifestyle changes. Bariatric surgery and artificial pancreas technology are long-term possibilities, but they are also invasive [9]. FGID treatment frequently incorporates dietary changes [10, 11], medications [12], psychological intervention [13], and other alternative medicines [14].

Electrical stimulation or neuromodulation has recently been under intensive investigation as a potential treatment option for obesity, diabetes, and FGIDs. Gastrointestinal electrical stimulation could induce weight loss, as it alters gastrointestinal motility, increases satiety, and decreases food intake [1]. Gastric electrical stimulation with the Enterra system is used to treat gastroparesis [15]. Approaches for treatment of obesity and diabetes include stimulation to the gut, the subdiaphragmatic sympathetic, and the vagal nerve [16, 17]. For example, vBloc therapy blocks the vagus nerve to suppress hunger and has been approved for treating obesity [18].

Electrical stimulation or neuromodulation during or after eating would be most effective for most electrical stimulation-based treatments as it is used to suppress food intake and absorption by altering postprandial gastrointestinal motility and hormonal secretion. This explains the need for automated food intake detection methods. Without an automatic food intake detection algorithm, the neuromodulation therapy would have to be applied continuously or depend on the action of patients (manually turn on the device) that imposes a serious compliance issue [19]. Therefore, automatic food intake detection methods are becoming increasingly crucial to obesity, overweight, diabetes, and FGID treatment. Wearable sensors for real-time detection can utilize acoustics [20], gyroscopic sensors [21], electroglottography [22], piezoelectricity [23], and respiratory signals [24]. However, most of these approaches are obtrusive or do not meet expectations for required accuracy. Furthermore, if these sensors were to report food intake to an implanted device, they would additionally require some form of wireless communication as they could not be directly attached to the implanted device. Because of these limitations, there is a need to develop new approaches to detecting food intake that can be incorporated into an implantable stimulator.

Food intake is highly regulated by the central nervous system [25]. The autonomic nervous system also plays a key role in the control of energy homeostasis and body weight [26]. The autonomic system consists of a sympathetic and a parasympathetic branch, whose activities are affected differently by periods of fasting and food intake. Food intake has been reported to consistently increase sympathetic activity and decrease vagal activity [27, 28].

The cardiac autonomic function can be assessed by the spectral analysis of heart rate variability (HRV) [28, 29]. The spectrum of HRV displays two major spectral components, i.e., a low-frequency (LF) component (0.04–0.15 Hz) and a high-frequency (HF) component (0.15–0.40 Hz). The LF reflects mainly sympathetic activity with some parasympathetic input, while the HF represents solely parasympathetic activity. In addition, the LF/HF ratio reflects the sympathovagal balance.

Since food intake alters cardiac autonomic functions that can in turn be measured by changes in the HRV, we hypothesize that the HRV can be used to detect food intake. The HRV signal can be assessed from the electrocardiogram (ECG), whereas the ECG can be recorded noninvasively and repetitively using surface electrodes. The ECG signals obtained from surface electrodes can be used to train an artificial neural network (ANN) and in real neuromodulation applications, the ECG signal can be easily acquired using the implantable neuromodulation device via one electrode placed at its stimulation lead and the case of the stimulator (serving as another electrode). Lu et al. already assessed the change in LF, HF, and LF/HF ratio after eating a standardized meal; however, they did not consider the dynamic nature of the HRV [28]. Although the HRV is also affected by factors than food intake, we expect that food intake can be discriminated from other factors by considering different features of the HRV and by analyzing these in a subject-specific manner. This is because the food intake is a unique process that is different from other factors.

The objective of the present study was to perform a dynamic analysis of the HRV to detect food intake, with the goal of using the algorithm in an implantable electrical stimulator in a real-time fashion. The automatic food intake detection was achieved by training an artificial neural network (ANN) with HRV features as inputs, and then using this ANN to detect food intake.

Materials and Methods

Neural Network Method

The ANN is a machine learning technique inspired by the way neurons are connected in the brain [30]. It is composed of neurons, each of which can receive a number of inputs and process these to compute its activation, or output. ANNs are structured in layers, in which outputs from one layer are weighted and passed to the next layer as inputs (Fig. 1).

Since neural networks can process large amounts of data and can easily learn from examples, they are suitable for food intake detection, a two-class classification problem in which the samples are classified as belonging to the fasting or feeding stage. In this study, the training examples consisted of parameters derived from subject data as input and one of two states as output.

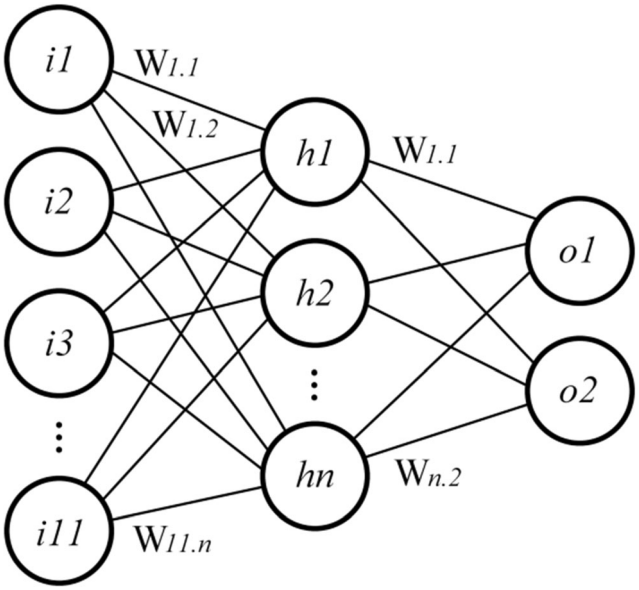


Fig. 1 Neural network with three layers used in this study. The input layer size is eleven, the hidden layer consists of a variable amount of neurons, and the output layer size is two

Experiment 1: a Small Sample Series with a Standard Test Meal

Experimental Protocol

After 6 h or more of fasting, the ECG was recorded from 16 subjects (all healthy adult volunteers) using an electrocardiogram (ECG) device (MedKinetic, Ningbo, China) via three surface electrodes. The inclusion criteria are as follows: (1) age, 18–65 years; (2) absence of any clinical symptoms; (3) taking no medications; (4) willingness to sign the consent form. The exclusion criteria are as follows: (1) any history of abdominal surgery; (2) any systemic diseases; (3) taking any medications during the 3 days before the test; (4) allergic to adhesive ECG electrodes. No limits were set on the body weight and height. The recording consisted of 60 min in the fasting state and 10 min during eating. The meal given to the subjects consisted of around 450 cal. It included a donut, a cookie, a yogurt, an egg, and 80 ml of water. The subjects were asked to remain still and avoid talking during the whole period of ECG measurements.

Signal Processing and Feature Extraction

The ECG signal was processed in MATLAB using a custom-designed software. In order to compute the HRV, first the R-peaks of the ECG signal were detected using the Pan-Tompkins algorithm [31] that finds QRS complexes based on the analysis of their slope, amplitude, and width. Once the R-peaks were detected, the R-R interval signal, or tachogram (which expresses the HRV), was derived from them. This beat-to-beat sampled signal was then resampled to 4 Hz using linear interpolation. We then analyzed the tachogram dynamically in the time domain and frequency domain. The total length of the signal was divided into non-overlapping segments, or epochs, of different lengths: 2, 5, and 8 min. For the time domain analysis of the HRV, different time-domain features were computed on each epoch of the non-resampled tachogram (Table 1). Each epoch was thus defined as one sample, with a number of corresponding features. Each of the samples was labeled as "fasting" or "feeding." The label "fasting/feeding" was accorded to samples corresponding to segments of the ECG during the fasting/feeding state.

For the dynamic analysis of the HRV in the frequency domain, frequency-domain features were computed for each quasi-stationary epoch of the resampled tachogram. Each epoch's power spectrum, or power spectral density (PSD), was estimated using Welch's method [32], by averaging the periodograms of half-overlapped Hanning windowed segments of each epoch.

After computing the power spectral density (PSD) of the HRV for every epoch, we extracted the frequency-domain features (Table 1). These features were all derived from the PSD in multiple frequency bands, namely the very low frequency band (VLF, 0.003–0.04 Hz), the low frequency band (LF, 0.04–0.15 Hz), and the high frequency band (HF, 0.15–0.4 Hz). The total power in these frequency bands was obtained by integrating the PSD over the respective frequency ranges.

Altogether, the feature extraction process results in a total of 6 time-domain and 5 frequency-domain features for each

Table 1 Time- and frequency-domain features

Feature	Description
HR mean	Mean heart rate of each epoch (in bpm)
HR SD	Standard deviation of the heart rate in each epoch (in bpm)
NN mean	Mean normal-to-normal interval, i.e., mean time interval between consecutive heart beats (R-R interval) in each epoch (in s)
NN max-min	Difference between longest and shortest R-R interval in each epoch (in s)
RMSSD	Square root of the mean-squared differences between successive R-R intervals (in s)
pNN50	Proportion of interval differences of successive R-R intervals greater than 50 ms in each epoch (in %)
VLF	Very low frequency spectral power (0.003 to 0.04 Hz)
LF	Low frequency spectral power (0.04 to 0.15 Hz)
HF	High frequency spectral power (0.15 to 0.4 Hz)
LF/HF	Ratio of low to high frequency power
nLF	Normalized low frequency power relative to HF + LF, i.e., $\frac{LF}{LF+HF}$

epoch and for each subject. Each feature is thus composed of numerous samples that correspond to the epochs into which the tachogram was divided.

Training and Testing of ANN

Detecting food intake is a two-class classification problem, in which the samples are classified as belonging to the fasting or feeding stage. We tackled this classification problem by training a pattern recognition network in MATLAB. The training examples that were fed to the ANN consisted of samples and their corresponding labels. Each sample was a vector with eleven entries corresponding to features of the HRV. Its label classified it as belonging to the fasting or the feeding stage. The Levenberg-Marquardt backpropagation function was selected because its performance was among the best. The mean squared normalized error was chosen as the performance measure. During training, the samples were randomly divided into 80% training samples and 20% validation samples.

This study considered 3 different epoch lengths: 2, 5, and 8 min. In the time-frequency analysis, decreasing the epoch length increases the time resolution but leads to a loss in frequency resolution [33]. Thus, there is a trade-off between frequency resolution and time resolution when the ANN tries to differentiate between samples of the fasting and feeding stage. The minimum 2 min epoch length was selected based on the fact that in practice, occurrences of meals shorter than 2 min are quite rare. Moreover, detecting such short timescale changes in the HRV would in practice result in a high false positive rate, as transient changes in the HRV on such a small timescale can be related to numerous other factors than feeding. The upper limit of 8 min was determined based on a similar reasoning: further increasing the epoch length would mean that short, small meals would go undetected too easily, as the features are averaged measures over the whole epoch length. This would in turn result in a high false negative rate.

Three numbers of neurons in the hidden layer were compared: 3, 5, and 10. The information in the following paragraph is based on [34]. If the number of neurons in the hidden layer is too low, the classification problem will be underfitted, i.e., the neurons will not be able to adequately solve the classification problem. However, having too many neurons may result in overfitting and an increase in training time. Unfortunately, there is no simple way to select the perfect amount of neurons. There are, however, some rule-of-thumb methods, one of which suggests that the size of the hidden layer should be between the size of the input layer and the size of the output layer. With this in mind, we tested the performance of the neural network with 3, 5, and 10 neurons in the hidden layer.

The training and testing of the ANNs was performed in several ways, which are explained and discussed in the following paragraphs. First, leave-one-subject-out (LOSO) cross-validation was used. This way of testing and training evaluated how well the classifier performed on unseen, new subjects. Then, leave-one-subject-out and leave-one-out (LOSO-LOO) cross-validation were used, to evaluate how adding samples of specific patients to the training set would improve the accuracy. Finally, the performance of the ANN was tested in a more realistic setting, in which the ANN was first trained on all 16 subjects and then retrained on some recordings of a new, unseen patient. The testing was performed on unseen recordings of the same patient. This simulated the way the ANN would be trained and retrained in its application in future implant devices.

LOSO cross-validation was the first method used for training and testing. It is similar to K-fold cross-validation, but the partitioning of the dataset is different. Each partition contains all the samples corresponding to one subject [35]. LOSO cross-validation thus consisted of training the ANN with the samples of all except one subject and then testing on the samples of that one subject. All subjects were used as a test set exactly once. This cross-validation method was used to test how well the designed classifier generalized to new, unseen subjects. In the context of an implantable stimulating device, LOSO would be the equivalent of training the classifier beforehand on a database of different people and using the algorithm to detect food intake in a new subject without retraining it.

The second way of training and testing evaluated the ANN's performance with retraining on a particular subject. We used a combination of leave-one-subject-out and leave-one-out (LOSO-LOO) cross-validation. The first part of the method consisted of training the ANN with LOSO, i.e., training on all but one subject. In the second part, the ANN was retrained on all but one sample of the previously excluded subject. This part consisted of applying LOO cross-validation to the samples of one particular subject, with each sample of the subject being used as a test set exactly once. The performance was expected to improve with respect to LOSO cross-validation because of the subject retraining.

The classifier's performance was represented as follows: the sensitivity, specificity, accuracy, and area under the receiver operating characteristic curve (AUC) [33]. Due to the random initialization of some of the parameters of the ANN, the LOSO procedure did not lead to reproducible results. Therefore, it was performed ten times and the performance values were each averaged over the ten iterations.

Experiment 2: a Large Sample Series with Varying Meal Size

Experimental Protocol

The second dataset contained ECG recordings from 37 healthy control subjects and 73 patients with functional dyspepsia. Of the total of 110 subjects, 68 were female and 42

were male. The inclusion and exclusion criteria for the healthy control were the same as in experiment 1. The inclusion criteria for patients were as follows: (1) age, 18–65 years; (2) met criteria for functional dyspepsia according to Rome IV criteria; (3) willing to sign the consent form. Exclusion criteria for patients are as follows: (1) pregnant or lactating, (2) a history of gastrointestinal surgery, (3) any other diseases that may explain the symptom of functional dyspepsia, (4) severe psychological or mental diseases, or (5) allergic to adhesive ECG electrodes. Similar to experiment 1, no limits were set on the body weight and height. All of the recordings were made using the same recording technology and electrode placement as in experiment 1. This time however, no recordings were obtained during the feeding stage. The recordings consisted of 30 min in the fasting stage (before eating) and 30 min in the postprandial state (after eating).

Instead of giving a fixed amount of a solid meal, the subjects in this experiment were asked to consume a liquid meal at a rate of 60 mL/min until complete fullness. The meal was prepared by dissolving 100 g of Nestle Full Cream Milk Powder and 50 g of cola powder in 1120 mL of water. The tolerated volume averaged 855 ± 294 mL. The data used to compare the fasting data was the data obtained after the maximum intake of the nutrient liquid meal instead of the data during eating or drinking. All postprandial data were acquired immediately after termination of eating/drinking.

Signal Processing and Feature Extraction

Each ECG signal was processed with the same methods as in experiment 1. This processing consisted of R-peak detection, tachogram computation, and resampling. The same time-domain and frequency-domain analyses were performed to extract the eleven features from tachogram segments.

Artificial Neural Network

We should note that the classification problem to be solved in experiment 2 was different from the problem in experiment 1. In experiment 1, the classes to be distinguished were the fasting and the feeding stage. In this second experiment, an ANN was designed to classify samples as belonging to either the fasting stage (before eating) or the postprandial stage (after eating). The classification was performed similarly as in experiment 1. The training examples that were fed to the ANN consisted of samples with eleven entries corresponding to features of the HRV. Each sample corresponded to a label indicating whether it belonged to the fasting or the postprandial stage. The design parameters (training function, performance function, training/validation ratio, and number of layers) chosen for the ANN were the same as in experiment 1. Again, the three different ways of training and testing were performed. The LOSO and LOSO-LOO cross-validation methods were applied with window lengths of 2, 5, and 8 min, and 3, 5, and 10 neurons in the hidden layer. The sensitivity, specificity, accuracy, and AUC were computed to evaluate the ANNs' performances.

Similarly as in experiment 1, we also tried the third way of training and testing, to quantify how much the ANN improves by going from general to subject-specific training. First, the ANN was trained on the original 110 subjects and tested on a new, unseen subject. Then, it was retrained on one and on two recordings of the new subject. Its performance was evaluated for all three situations. The new subject used in this experiment was the same one as the new subject in experiment 1, with three recordings of which 30 min of fasting data and 30 min of postprandial data were used here.

Results

Features of HRV

Figure 2 shows a typical example of a tachogram, in which the red-dashed line indicates the beginning of the feeding stage, defined as the moment the subject starts eating. The magnitude of the signal (R-R interval) clearly decreases after the subject starts eating, demonstrating an increase in heart rate.

Figure 3 shows some of the time- and frequency-domain features of the HRV of a healthy subject using an epoch length of 8 min. The figure shows a complete time window from the fasting stage to the postprandial stage. The fasting stage stopped at 32 min, after which the feeding stage started (indicated by the dashed line in Fig. 3). At 48 min, the subject stopped eating and the postprandial stage started (indicated by the dot-dashed line in Fig. 3. The figure shows an increase in the HR mean when the person started eating. The NN mean is reciprocal to the HR mean and decreased when food intake started. The RMSSD also decreased during the feeding stage, whereas the LF/HF and nLF (normalized low frequency) increased when food intake started. The postprandial stage was characterized by a decrease in the HR mean, LF/HF, and nLF, and an increase in the NN mean and RMSSD. The HR mean and NN mean were different in the postprandial stage compared with the fasting stage, while the other features in Fig. 3 had similar values in the postprandial stage as in the fasting stage. This figure represents typical dynamic changes of HRV parameters from fasting to eating to postprandial. The HRV data from the FD patients showed similar dynamic changes of these HRV parameters.

Automated Detection of Food Intake Based on Experiment 1

Classification results are shown in Tables 2 and 3. With the LOSO cross-validation (Table 2), the epoch length and

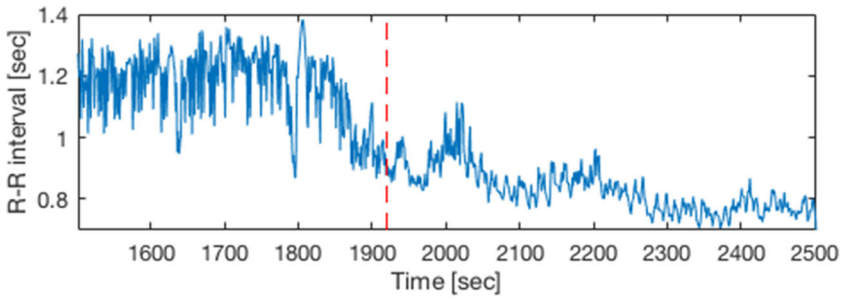


Fig. 2 Part of the tachogram of one subject. Dashed line indicates the moment when the subject starts eating

number of neurons did not have a large influence on the performance. An increasing epoch length did decrease the sensitivity a little. This effect was a lot stronger when LOSO-LOO cross-validation was used (Table 3). With LOSO, the mean accuracy was 0.83, the mean sensitivity was 0.51, and the mean specificity was 0.89. With LOSO-LOO, the ANN reached maximal accuracy (0.93) and sensitivity (0.79) values with 2 min epochs. The mean specificity increased to 0.97.

The last part of the experiment consisted of training the ANN with the data of the original 16 subjects and testing it on a new, unseen subject. First, the ANN was not retrained with the data of the new subject. Then, the ANN was retrained

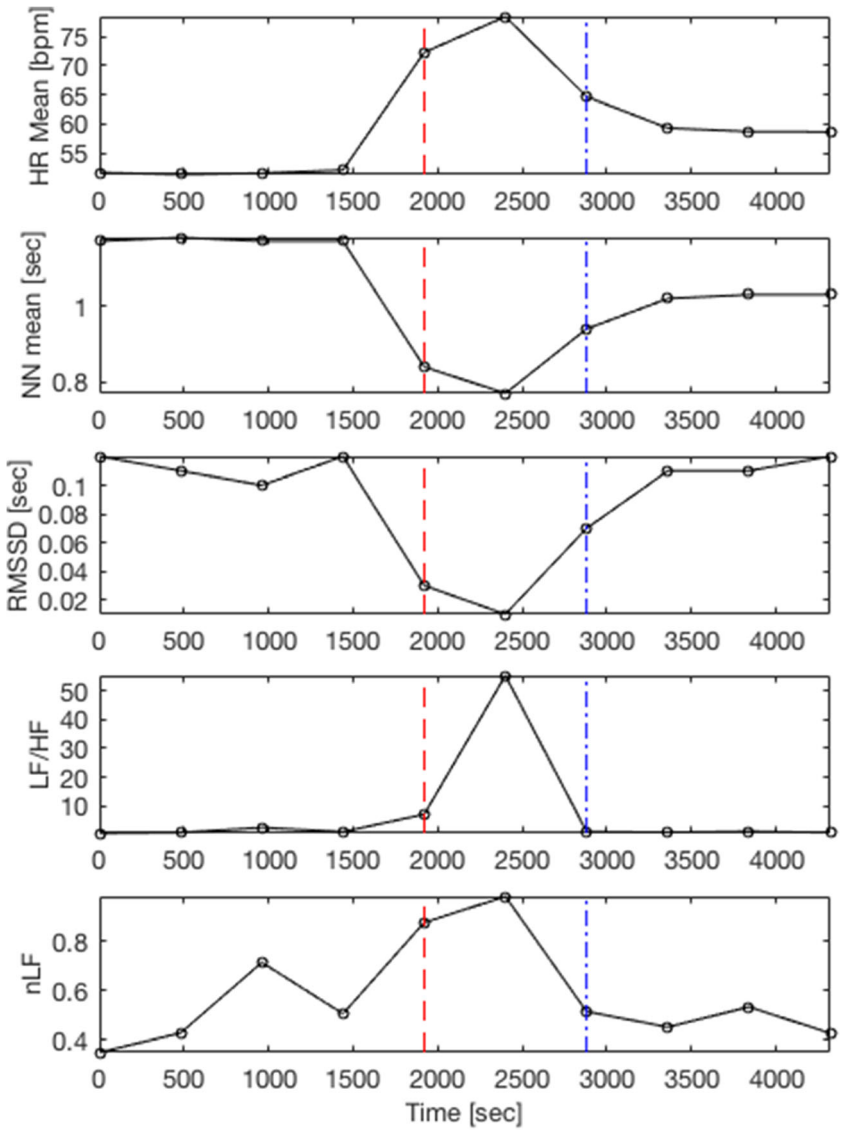


Fig. 3 Features for one subject. Dashed line indicates the moment when the subject starts eating. Dot-dashed line indicates the start of the postprandial stage. Epoch length, 8 min

Table 2 ANN performance in exp 1 with leave-one-subject-out (LOSO)

Epoch length (min)	Number of neurons	Accuracy	Sensitivity	Specificity	AUC
2	3	0.82	0.54	0.89	0.81
2	5	0.82	0.54	0.88	0.82
2	10	0.81	0.58	0.86	0.81
5	3	0.84	0.47	0.91	0.79
5	5	0.84	0.47	0.90	0.79
5	10	0.85	0.52	0.90	0.80
8	3	0.83	0.43	0.91	0.74
8	5	0.82	0.52	0.88	0.77
8	10	0.82	0.53	0.87	0.77

on one and two recordings of the new subject. The resulting accuracy, sensitivity, specificity, and AUC are shown in Fig. 4. The specificity was exceptionally high in this example: it was equal to 1 for all of the cases, meaning that all fasting samples were correctly classified for the new subject. All other performance values were improved by retraining the ANN with the data of the new subject. The performance values were already close to maximal when the ANN was retrained on the first recording of the new subject and tested on the second and third recording. Consequently, there was a little room for improvement when more recordings of the same subject were added to the retraining data. In this part of the experiment, the epoch length was 2 min and the number of neurons in the neural network was 10.

Automated Detection of Food Intake Based on Experiment 2

For the experiment with fasting and postprandial data from 110 subjects (with no feeding data and varying meal size), the performance is reported in Tables 4 and 5, showing sensitivity and specificity values that are more balanced than in Tables 2 and 3. The performance values of the two experiments cannot be compared exactly because the two experiments had very different meals (fixed solid meal vs liquid meal with varying size) and conditions (feeding vs postprandial). For the LOSO cross-validation method (Table 4), the sensitivity values were generally higher than the specificity values. The mean accuracy was 0.64. In Table 5, the reported performance of the ANN is better than in Table 4. Using LOSO-LOO as cross-validation, method increased the mean accuracy to 0.85 and the maximum accuracy to 0.88. The accuracy, sensitivity, and specificity reached a maximum for 2 min epochs (Table 5).

Table 3 ANN performance in exp 1 with leave-one-subject-out and leave-one-out

Epoch length (min)	Number of neurons	Accuracy	Sensitivity	Specificity	AUC
2	3	0.93	0.78	0.97	0.93
2	5	0.93	0.79	0.97	0.94
2	10	0.93	0.77	0.97	0.93
5	3	0.90	0.54	0.96	0.85
5	5	0.91	0.54	0.97	0.87
5	10	0.91	0.51	0.97	0.85
8	3	0.85	0.12	0.96	0.68
8	5	0.85	0.06	0.97	0.66
8	10	0.86	0.13	0.98	0.73

The last part of the experiment again consisted of training the ANN with the data of the original 110 subjects and testing it on a new, unseen subject. First, the ANN was not retrained with the data of the new subject. Then, the ANN was retrained on one and two recordings of the new subject. The resulting accuracy, sensitivity, specificity, and AUC are shown in Fig. 5. All performance values except for the specificity were improved when the ANN was retrained on the new subject. The figure shows an improvement in all performance values when the ANN was retrained with two rather than one recording of the new subject. In this part, the epoch length was 2 min and the number of neurons in the neural network was 10.

Discussion

In this study, the ECG was processed to derive the HRV, whose features were dynamically extracted and then used to train an ANN to detect food intake. The major findings from the experiments were as follows: (1) the LOSO-LOO cross-validation method yielded higher performance values than the LOSO method; (2) the highest accuracies obtained with LOSO-LOO cross-validation in experiments 1 and 2, respectively, were 0.93 and 0.88; (3) testing with subject-specific retraining resulted in better classification performance than testing without said retraining.

The automated food intake detection is required for the treatment of obesity, diabetes, and FGIDs using neuromodulation via an implantable stimulator, such as intestinal electrical stimulation for obesity [36–41]. The primary advantage of using the ECG as a signal to detect food intake is the fact that in future neuromodulation therapies using an implantable pulse generator (IPG), the stimulation lead and the IPG can be used as electrodes for detecting the ECG, eliminating the need for special sensors/electrodes. The advantage of using an ANN for the food intake detection is that ample data are available for training the ANN, and most importantly, the algorithm for the detection is simple once the ANN is trained, which is ideal for the implantable stimulator as the

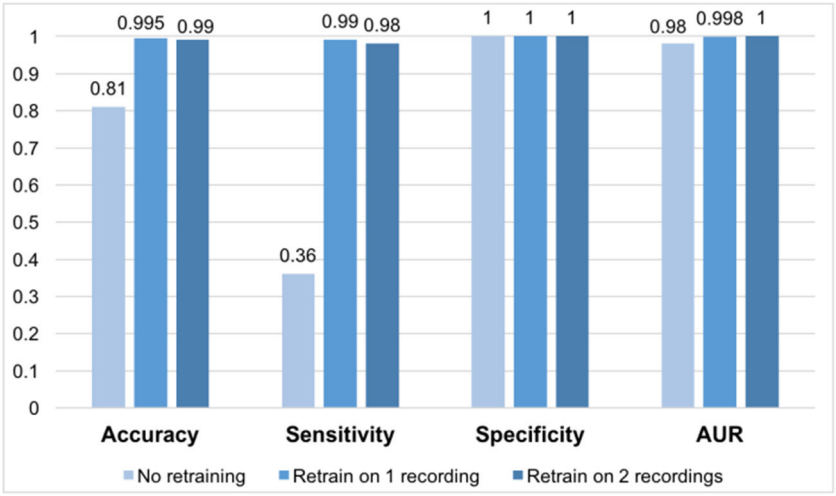


Fig. 4 The ANN's performance in experiment 1 (16 subjects, fasting and feeding stage) when the ANN is tested on recordings of a new, unseen subject, and when the ANN is tested on the same subject after retraining on 1 and 2 recordings of this subject. The epoch length is 2 min and the number of neurons is 10

training of the ANN can be done without the use of IPG, and once the ANN is trained, the weights of the ANN can be uploaded to the IPG for on-line detection of food intake.

With the LOSO cross-validation method in experiment 1 (small sample size but with feeding data and fixed meal), the mean accuracy was 0.83. The ANN demonstrated little to no change in performance when tested with different epoch lengths and numbers of neurons. As depicted in Table 2, the performance values showed only minor fluctuations. Conversely, Table 3 shows how the epoch length had a clear influence on results obtained with the LOSO-LOO method. Because the feeding stage of experiment 1 lasted only 10 min, the epoch length had a major impact on the amount of samples available for training. For example, each subject had just 1 sample in the feeding stage when an 8-min epoch was used. The tests with 2- and 5-min epoch lengths had more training samples in the feeding stage. The shorter the epoch length, the more training samples in the feeding stage and the better the ANN were able to identify the feeding stage. This effect was not due to the intrinsic quality of features obtained with certain window lengths, but due to the fact that the ANN had not enough training data per subject. It primarily had a large impact on the sensitivity, and consequently also influenced the accuracy and AUC. Table 3 shows that the highest sensitivity obtained was 0.79 (2-min epoch) and the lowest was 0.13 (8-min epoch). Overall, using LOSO-LOO cross-validation improved the accuracy compared with LOSO cross-validation: the mean accuracy increased to 0.85 and the maximum accuracy to 0.88. Again, the number of neurons appeared to have no effect on the performance.

Experiment 2 was designed to mimic real-world situations by including a large number of subjects, mixing healthy controls and patients, and providing a meal with different sizes. In practical applications, a large number of subjects can be used to train the ANN since the ECG measurement is noninvasive. Patients with functional dyspepsia were included in the experiment to represent diversity of subjects as they may have different postprandial changes in autonomic functions. Different meal sizes (varied from 400 to 1200 mL) were allowed in this experiment to reflect

Table 4 ANN performance in exp 2 with leave-one-subject-out (LOSO)

Epoch length (min)	Number of neurons	Accuracy	Sensitivity	Specificity	AUC
2	3	0.64	0.68	0.59	0.68
2	5	0.64	0.68	0.59	0.68
2	10	0.63	0.68	0.59	0.68
5	3	0.65	0.70	0.60	0.69
5	5	0.65	0.70	0.60	0.69
5	10	0.65	0.69	0.61	0.68
8	3	0.64	0.69	0.59	0.70
8	5	0.65	0.69	0.60	0.70
8	10	0.64	0.68	0.61	0.69

Table 5 ANN performance in exp 2 with leave-one-subject-out and leave-one-out

Epoch length (min)	Number of neurons	Accuracy	Sensitivity	Specificity	AUC
2	3	0.88	0.87	0.88	0.95
2	5	0.87	0.87	0.88	0.95
2	10	0.87	0.87	0.87	0.94
5	3	0.86	0.86	0.85	0.93
5	5	0.85	0.86	0.85	0.93
5	10	0.86	0.85	0.86	0.93
8	3	0.82	0.81	0.82	0.90
8	5	0.81	0.80	0.82	0.89
8	10	0.81	0.81	0.82	0.89

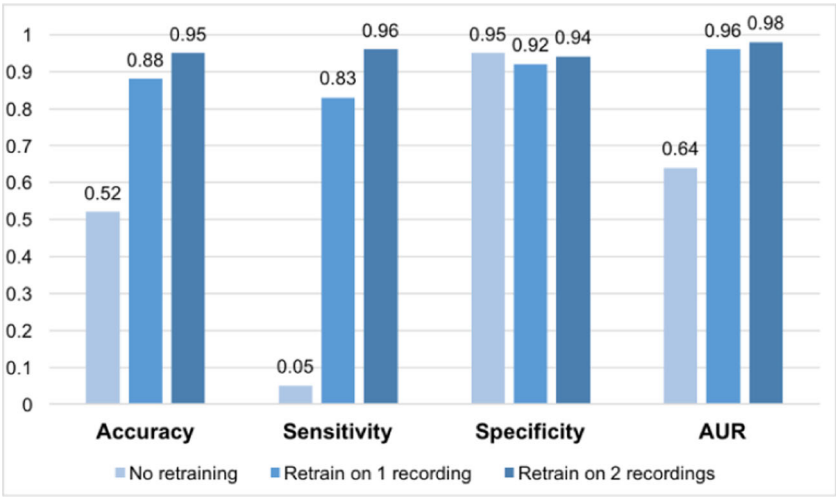


Fig. 5 The ANN's performance in experiment 2 (110 subjects, fasting and postprandial stage) when the ANN is tested on recordings of a new, unseen subject, and when the ANN is tested on the same subject after retraining on 1 and 2 recordings of this subject. The epoch length is 2 min and the number of neurons is 10

future clinical application. In addition, the ECG data during feeding were excluded to represent future applications that do not need stimulation during feeding. Nevertheless, experiment 2 yielded similar results; the LOSO method results showed no influence from epoch length or number of neurons (Table 4) while the LOSO-LOO method showed higher performance values for shorter epoch lengths (Table 5). Unlike experiment 1, which had a 10-min feeding period, this experiment had 30-min fasting and postprandial periods. Therefore, longer epoch lengths did not decrease performance values as drastically as in experiment 1. Performance values still increased with shorter epochs because of their positive effect on the amount of training samples. Table 5 shows that the highest sensitivity was 0.87 (2-min epoch), and the lowest was 0.80 (8-min epoch) for the LOSO-LOO method. This was an improvement with comparison with the LOSO method, yielding a mean accuracy of 0.64 (Table 4). Neither method showed impact from the number of neurons.

It is also important to note that the specificity was similarly influenced by epoch length in experiment 2, whereas experiment 1 showed no correlation between the two. In all experiment 1 tests, specificity was considerably higher than sensitivity. In other words, the ANN was capable of detecting fasting better than feeding, a result of the unbalanced sample proportion between the two stages. During the experiments, the ECG was recorded for 1 h during the fasting stage and only 10 min during feeding. As a result, the ANN was trained with more samples from the fasting stage, leading to better performance in detecting this stage than the feeding stage. In general, experiment 2 avoided this issue by having equal fasting and postprandial periods. This led to more balanced sensitivity and specificity. Lowering the decision threshold can improve the sensitivity at the expense of decreasing the specificity. A trade-off can be made by selecting the desired sensitivity and specificity in the ROC curve and applying the corresponding threshold to the algorithm.

In both experiments 1 and 2, the final part was to train the ANN on the original subjects and to test it on a new, unseen subject. First, the ANN was not retrained on the new subject; then, it was retrained on one recording of this subject; finally, it was retrained on two recordings of the subject. The purpose of this was to mimic the ANN's training process in a real-life application. If an implantable stimulator were to be used in a patient, it would first need to acquire a sufficient amount of patient-specific ECG to retrain the ANN, which can be done easily before the treatment starts. In this way, the classifier would be specified to match the patient's own characteristics.

The results of the subject-specific training and testing are reported in Figs. 4 and 5 for experiment 1 and 2, respectively. They showed that the performance was, in fact, significantly improved after retraining on subject-specific recordings. In experiment 1 (Fig. 4), we see an exceptionally high specificity of 1 in all three cases. The other performance values were greatly improved when the ANN was retrained on one recording of the new subject. Retraining on one recording was enough to obtain almost maximal performance values when the ANN was tested on the second and third recording. When the network was retrained on the first and the second recording and tested on the third one, there was no more significant improvement. The accuracy, sensitivity, specificity, and AUC were 0.993, 0.979, 1, and 1. The perfect amount of data to retrain on should be investigated in further studies that include more than three subject-specific data recordings. In experiment 2 (Fig. 5), all performance measures except specificity were greatly improved when subject-specific retraining was included. The specificity was decreased from 0.95 to 0.92 in exchange for a huge improvement in sensitivity from 0.05 to 0.83. Moreover, all the performance values were additionally increased by retraining on two subject-specific recordings rather than one. After retraining on two recordings, the values

for accuracy, sensitivity, specificity, and AUC were, respectively, 0.95, 0.94, 0.96, and 0.98. We expect these values to further improve with more subject-specific retraining. Comparing the subject-specific retraining in experiment 1 and 2 shows that the dataset on which the ANN is trained beforehand plays a large role in how well it performs initially in detecting food intake in a new subject. Fortunately, subject-specific retraining of the ANN can easily be used to improve this performance up to the desired accuracy.

The promising results in this study suggest that the HRV can be used to detect food intake dynamically. However, this study has limitations, and a few more issues need to be addressed before the ANN can be used in an implantable stimulation device. Further studies should include more data, particularly in the fasting stage under different situations rather than pure laboratory setting. Conditions other than eating that may alter autonomic functions should be included in future studies, such as exercise, walking, and running. Future work should also consider different types of meals, including snacks, to see how the changes in features (between fasting and feeding) depend on the type and amount of food. Finally, the decision threshold in the ANN should be tuned depending on the desired sensitivity and specificity.

We would conclude this discussion with some general considerations on the clinical application of ANNs in future neuromodulation therapies for treating obesity, diabetes, and FGIDs. An ANN can first be trained on a general dataset. Then, using the ECG acquired from a candidate patient, the same ANN can be retrained to achieve a higher accuracy. After the retraining, the weights of the ANN are fixed. The pacemaker can then be programmed to include a simple, static, input-output function determined by the final weights of the ANN. We should note that the nature of these treatments does not require perfect detection of fasting and feeding. For example, the decision threshold could be adapted to trigger stimulation when eating a full meal, but not when eating a snack. Changing this threshold is the equivalent of making a trade-off between sensitivity and specificity. With these considerations in mind, ANNs trained for food intake detection have high potential for clinical applications.

In conclusion, this study investigates a promising approach to detect food intake using the ECG as primary signal. It uses time- and frequency-domain features derived from the dynamic analysis of the HRV. The features suggest that the sympathovagal balance increases upon food intake. The parameters serve as inputs to train an ANN to detect food intake. When discriminating between fasting and feeding with LOSO cross-validation, the highest accuracy obtained was 0.85. Subject-specific retraining improves the classification accuracy to 0.995. Based on these promising results, we suspect that ANN-based food intake detection has high potential in minimally invasive electrical stimulation treatment for obesity, diabetes, and FGIDs.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval Statement All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

1. Sallam HS, Chen JDZ. Colon electrical stimulation: potential use for treatment of obesity. Obesity. 2011;19(9):1761–7.
2. Arrone LJ. Epidemiology, morbidity, and treatment of overweight and obesity. J Clin Psychiatry. 2001;62(23):13–22.
3. Spieker EA, Pyzocha N. Economic impact of obesity. Prim Care. 2016;43:83–95.
4. Department of Health and Human Services. Atlanta: Centers for Disease Control and Prevention; c1995-2019 [cited 2018 Aug 2]. Available from: https://www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf. (Accessed 2 Aug 2018).
5. American Diabetes Association. Complication; c1995-2019 [cited 2018 Aug 2]. Available from: http://diabetes.org/living-with-diabetes/complications/.
6. Parkman HP, Doma S. Importance of gastrointestinal motility disorders. Pract Gastroenterol. 2006;30(9):23–40.
7. Talley N. Functional gastrointestinal disorders as a public health problem. Neurogastroenterol Motil. 2008;20(1):121–9.
8. Cordera R, Adami GF. From bariatric to metabolic surgery: looking for a "disease modifier" surgery for type 2 diabetes. World J Diabetes. 2016;7(2):27–33.
9. Department of Health and Human Services, USA. Insulin, medicines, & other diabetes treatments. Bethesda: National Institute of Diabetes and Digestive and Kidney Diseases. c1995-2019 [cited 2018 Aug 2]. Available from: https://www.niddk.nih.gov/health-information/diabetes/overview/insulin-medicines-treatments
10. Huertas-Ceballos A, Macarthur C, Logan S. Dietary interventions for recurrent abdominal pain (RAP) in childhood. Cochrane Database Syst Rev. 2002:CD003019.
11. Lebenthal E, Rossi TM, Nord KS, et al. Recurrent abdominal pain and lactose absorption in children. Pediatrics. 1981;67:828–32.
12. Grover M, Drossman DA. Psychotropic agents in function gastrointestinal disorders. Curr Opin Pharmacol. 2008;8(6):715–23.
13. Bursch B. Psychological/cognitive behavioral treatment of childhood functional abdominal pain and irritable bowel syndrome. J Pediatr Gastroenterol Nutr. 2008;47:706–7.
14. Whitfield KL, Shulman RJ. Treatment options for functional gastrointestinal disorders: from empiric to complementary approaches. Pediatr Ann. 2009;38(5):288–94.
15. Yin J, Abell TD, McCallum RW, et al. Gastric neuromodulation with Enterra system for nausea and vomiting in patients with gastroparesis. Neuromodulation. 2012;15(3):224–31.
16. Greenway F, Zheng J. Electrical stimulation as treatment for obesity and diabetes. J Diabetes Sci Technol. 2007;1(2):251–9.
17. Chen JDZ, Yin J, Wei W. Electrical therapies for gastrointestinal motility disorders. Expert Rev Gastroenterol Hepatol. 2017;11(4):407–18.

18. Apovian CM, Shah SN, Wolfe BM, et al. Two-year outcomes of vagal nerve blocking (vBloc) for the treatment of obesity in the ReCharge trial. Obes Surg. 2016;27:169–76.
19. Vu T, Lin F, Alshurafa N, et al. Wearable food intake monitoring technologies: a comprehensive review. Computers. 2017;6(1):4.
20. Amft O. A wearable ear pad sensor for chewing monitoring. Proceedings of the IEEE Sensors 2010 Conference; 2010 Nov 1-4; Waikoloa, USA IEEE Xplore. 2010. https://doi.org/10.1109/icsens.2010.5690449
21. Dong Y, Hoover A, Scisco J, et al. A new method for measuring meal intake in humans via automated wrist motion tracking. Appl Psychophysiol Biofeedback. 2012;37(3):205–15.
22. Farooq M, Fontana JM, Sazonov E. A novel approach for food intake detection using electroglottography. Physiol Meas. 2014;35(5):739–51.
23. Farooq M, Sazonov E. Comparative testing of piezoelectric and printed strain sensors in characterization of chewing. In: Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2015 Aug 25–29; Milano, Italy. IEEE Xplore; 2015:7538–7541
24. Dong B, Biswas S. Wearable diet monitoring through breathing signal analysis. In: Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2013 Jul 3–7; Osaka, Japan. IEEE Xplore; 2013:1186–1189.
25. Schwartz MW, Woods SC, Porte D, et al. Central nervous system control of food intake. Nature. 2000;404(6778):661–71.
26. Messina G, De Luca V, Viggiano A, et al. Autonomic nervous system in the control of energy balance and body weight: personal contributions. Neurol Res Int. 2013;2013:639280.
27. Sakaguchi T, Arase K, Fisler JS, et al. Effect of starvation and food intake on sympathetic activity. Am J Phys. 1988;255(2):R284–8.
28. Lu CL, Zou X, Orr WC, et al. Postprandial changes of sympathovagal balance measured by heart rate variability. Dig Dis Sci. 1999;44(4):857–61.
29. Stuckey MI, Tulppo MP, Kiviniemi AM, et al. Heart rate variability and the metabolic syndrome: a systematic review of the literature. Diabetes Metab Res Rev. 2014;30(8):784–93.
30. Blum A. Neural networks in C++: an object-oriented framework for building connectionist systems: Wiley; 1992.
31. Pan J, Tompkins WJ. A real-time QRS detection algorithm. IEEE Trans Biomed Eng. 1985;32(3):230–6.
32. Welch P. The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Trans Audio Electroacoust. 1967;15(2):70–3.
33. Rangayyan BRM. Biomedical signal analysis: a case-study approach: IEEE Press Wiley-Interscience; 2002.
34. Heaton J. The Number of hidden layers [Internet]. Heaton Research; 2017 [cited 2018 Aug 1]. Available from: https://www.heatonresearch.com/2017/06/01/hidden-layers.html.
35. Hegenbart S, Uhl A, Vécsei A. Systematic assessment of performance prediction techniques in medical image classification: a case study on celiac disease. Inf Process Med Imaging. 2011;22:498–509.
36. Aberle J, Busch P, Veigel J, et al. Duodenal electric stimulation: results of a first-in-man study. Obes Surg. 2016;26:369–75.
37. Khawaled R, Blumen G, Fabricant G, et al. Intestinal electrical stimulation decreases postprandial blood glucose levels in rats. Surg Obes Relat Dis. 2009;5:692–7.
38. Ye F, Liu Y, Li S, et al. Hypoglycemic effects of intestinal electrical stimulation by enhancing nutrient-stimulated secretion of GLP-1 in rats. Obes Surg. 2018;28:2829–35.
39. Li S, Chen JD. Pulse width-dependent effects of intestinal electrical stimulation for obesity: role of gastrointestinal motility and hormones. Obes Surg. 2017;27:70–7.
40. Liu J, Xiang Y, Qiao X, et al. Hypoglycemic effects of intraluminal intestinal electrical stimulation in healthy volunteers. Obes Surg. 2011;21:224–30.
41. Chen JD, Lin Z, McCallum RW. Noninvasive feature-based detection of delayed gastric emptying in humans using neural networks. IEEE Trans Biomed Eng. 2000;47(3):409–12.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.