Obesity Surgery

Artificial neural network-based automatic detection of food intake for neuromodulation in treating obesity and diabetes --Manuscript Draft--

Manuscript Number:	OBSU-D-19-01170R1	
Full Title:	Artificial neural network-based automatic detection of food intake for neuromodulation in treating obesity and diabetes	
Article Type:	Original Contribution	
Keywords:	Food intake detection, Obesity, Diabetes, Neuromodulation, Machine Learning	
Corresponding Author:	Jiande Chen, Ph.D Johns Hopkins University Baltimore, Maryland UNITED STATES	
Corresponding Author Secondary Information:		
Corresponding Author's Institution:	Johns Hopkins University	
Corresponding Author's Secondary Institution:		
First Author:	Elizabeth R.M. Heremans	
First Author Secondary Information:		
Order of Authors:	Elizabeth R.M. Heremans	
	Amy Chen	
	Ximeng Wang	
	Jiafei Cheng	
	Feng Xu	
	Agustin Martinez	
	Georgios Lazaridis	
	Sabine Van Huffel	
	Jiande Chen, Ph.D	
Order of Authors Secondary Information:		
Funding Information:	Foundation for the National Institutes of Health (R01DK107754)	Professor Jiande Chen
Abstract:	<p>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). 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. Conclusion: 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.</p>	

Artificial neural network-based automatic detection of food intake for neuromodulation in treating obesity and diabetes

Elisabeth R. M. Heremans^{1,2}, Amy S. Chen³, Ximeng Wang¹, Jiafei Cheng¹, Feng Xu⁴,
Agustin E. Martinez², Georgios Lazaridis², Sabine Van Huffel², and Jiande D Z Chen¹

1. Departments of Medicine and Biomedical Engineering, Johns Hopkins University,
School of Medicine, Baltimore, Maryland, USA
2. Department of Electrical Engineering-ESAT, STADIUS Center for Dynamical
Systems, Signal Processing and Data Analytics, KU Leuven, Leuven B-3001, Belgium.
3. College of Letters and Science, University of California, Berkeley, USA
4. Yinzhou Hospital Affiliated with Ningbo University School of Medicine

Running title: Automatic detection of food intake for neuromodulation

Correspondence:

Jiande Chen, PhD

Departments of Medicine and Biomedical Engineering

Johns Hopkins University School of Medicine

Tel: 410-550-6853, fax: 410-550-7861

Email: jchen184@jhu.edu

Abstract:

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). *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. *Conclusion:* 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

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

1
2
3
4 automated food intake detection methods. Without an automatic food intake detection algorithm,
5 the neuromodulation therapy would have to be applied continuously or depend on the action of
6 patients (manually turn on the device) that imposes a serious compliance issue [19]. Therefore,
7 automatic food intake detection methods are becoming increasingly crucial to obesity, overweight,
8 diabetes, and FGID treatment. Wearable sensors for real-time detection can utilize acoustics [20],
9 gyroscopic sensors [21], electroglottography [22], piezoelectricity [23], and respiratory signals
10 [24]. However, most of these approaches are obtrusive or do not meet expectations for required
11 accuracy. Furthermore, if these sensors were to report food intake to an implanted device, they
12 would additionally require some form of wireless communication as they could not be directly
13 attached to the implanted device. Because of these limitations, there is a need to develop new
14 approaches to detecting food intake that can be incorporated into an implantable stimulator.

15
16 Food intake is highly regulated by the central nervous system [25]. The autonomic nervous
17 system also plays a key role in the control of energy homeostasis and body weight [26]. The
18 autonomic system consists of a sympathetic and a parasympathetic branch, whose activities are
19 affected differently by periods of fasting and food intake. Food intake has been reported to
20 consistently increase sympathetic activity and decrease vagal activity [27], [28].

21
22 The cardiac autonomic function can be assessed by the spectral analysis of heart rate variability
23 (HRV) [28], [29]. The spectrum of HRV displays two major spectral components, i.e., a low-
24 frequency (LF) component (0.04-0.15 Hz) and a high-frequency (HF) component (0.15-0.40 Hz).
25 The LF reflects mainly sympathetic activity with some parasympathetic input, while the HF
26 represents solely parasympathetic activity. In addition, the LF/HF ratio reflects the sympathovagal
27 balance.

28
29 Since food intake alters cardiac autonomic functions that can in turn be measured by changes in
30 the HRV, we hypothesize that the HRV can be used to detect food intake. The HRV signal can be
31 assessed from the electrocardiogram (ECG), whereas the ECG can be recorded noninvasively and
32 repetitively using surface electrodes. The ECG signals obtained from surface electrodes can be
33 used to train an artificial neural network (ANN) and in real neuromodulation applications, the ECG
34 signal can be easily acquired using the implantable neuromodulation device via one electrode
35 placed at its stimulation lead and the case of the stimulator (serving as another electrode) (Review
36 #1, comment #1). Lu et al already assessed the change in LF, HF, and LF/HF ratio after eating a
37 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 (Reviewer #2, comment #2).

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

A. 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.

B. Experiment 1: A Small Sample Series with a Standard Test Meal

1) Experimental Protocol

After 6 hours 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: 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: 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. (Reviewer #2, comment #3) The recording consisted of 60 minutes in the fasting state and 10 minutes during eating. The meal given to the subjects consisted of around 450 calories. 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.

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

3) *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

1
2
3
4 feeding stage. The Levenberg-Marquardt backpropagation function was selected because its
5 performance was among the best. The mean squared normalized error was chosen as the
6 performance measure. During training, the samples were randomly divided into 80% training
7 samples and 20% validation samples.
8
9

10
11 This study considered 3 different epoch lengths: 2, 5, and 8 minutes. In the time-frequency
12 analysis, decreasing the epoch length increases the time resolution but leads to a loss in frequency
13 resolution [33]. Thus, there is a trade-off between frequency resolution and time resolution when
14 the ANN tries to differentiate between samples of the fasting and feeding stage. The minimum 2
15 minute epoch length was selected based on the fact that in practice, occurrences of meals shorter
16 than 2 minutes are quite rare. Moreover, detecting such short timescale changes in the HRV would
17 in practice result in a high false positive rate, as transient changes in the HRV on such a small
18 timescale can be related to numerous other factors than feeding. The upper limit of 8 minutes was
19 determined based on a similar reasoning: further increasing the epoch length would mean that
20 short, small meals would go undetected too easily, as the features are averaged measures over the
21 whole epoch length. This would in turn result in a high false negative rate (Reviewer #2, comment
22 #6).
23
24

25
26 Three numbers of neurons in the hidden layer were compared: 3, 5, and 10. The information in
27 the following paragraph is based on [34]. If the number of neurons in the hidden layer is too low,
28 the classification problem will be underfitted, i.e., the neurons will not be able to adequately solve
29 the classification problem. However, having too many neurons may result in overfitting and an
30 increase in training time. Unfortunately, there is no simple way to select the perfect amount of
31 neurons. There are, however, some rule-of-thumb methods, one of which suggests that the size of
32 the hidden layer should be between the size of the input layer and the size of the output layer. With
33 this in mind, we tested the performance of the neural network with 3, 5 and 10 neurons in the
34 hidden layer. (Reviewer #2, comment #6)
35
36

37
38 The training and testing of the ANNs was performed in several ways, which are explained and
39 discussed in the following paragraphs. First, leave-one-subject-out (LOSO) cross-validation was
40 used. This way of testing and training evaluated how well the classifier performed on unseen, new
41 subjects. Then, leave-one-subject-out and leave-one-out (LOSO-LOO) cross-validation were used,
42 to evaluate how adding samples of specific patients to the training set would improve the accuracy.
43 Finally, the performance of the ANN was tested in a more realistic setting, in which the ANN was
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 first trained on all 16 subjects and then retrained on some recordings of a new, unseen patient. The
5 testing was performed on unseen recordings of the same patient. This simulated the way the ANN
6 would be trained and retrained in its application in future implant devices.
7

8
9
10 LOSO cross-validation was the first method used for training and testing. It is similar to K-fold
11 cross-validation, but the partitioning of the dataset is different. Each partition contains all the
12 samples corresponding to one subject [35]. LOSO cross-validation thus consisted of training the
13 ANN with the samples of all except one subject and then testing on the samples of that one subject.
14 All subjects were used as a test set exactly once. This cross-validation method was used to test
15 how well the designed classifier generalized to new, unseen subjects. In the context of an
16 implantable stimulating device, LOSO would be the equivalent of training the classifier
17 beforehand on a database of different people and using the algorithm to detect food intake in a new
18 subject without retraining it.
19

20
21 The second way of training and testing evaluated the ANN's performance with retraining on a
22 particular subject. We used a combination of leave-one-subject-out and leave-one-out (LOSO-
23 LOO) cross-validation. The first part of the method consisted of training the ANN with LOSO,
24 i.e., training on all but one subject. In the second part, the ANN was retrained on all but one sample
25 of the previously excluded subject. This part consisted of applying LOO cross-validation to the
26 samples of one particular subject, with each sample of the subject being used as a test set exactly
27 once. The performance was expected to improve with respect to LOSO cross-validation because
28 of the subject-retraining.
29

30
31 The classifier's performance was represented as follows: the sensitivity, specificity, accuracy,
32 and area under the Receiver Operating Characteristic curve (AUC) [33]. Due to the random
33 initialization of some of the parameters of the ANN, the LOSO procedure did not lead to
34 reproducible results. Therefore, it was performed ten times and the performance values were each
35 averaged over the ten iterations.
36

37 38 39 40 41 **C. Experiment 2: A Large Sample Series with Varying Meal Size**

42 43 44 45 46 47 48 49 50 51 52 53 **1) Experimental protocol**

54
55 The second dataset contained ECG recordings from 37 healthy control subjects and 73 patients
56 with functional dyspepsia. Of the total of 110 subjects, 68 were female and 42 were male. The
57 inclusion and exclusion criteria for the healthy control were the same as in Experiment 1. The
58 inclusion criteria for patients were: 1) age: 18-65 years; 2) met criteria for functional dyspepsia
59
60
61
62
63
64
65

1
2
3
4 according to Rome IV criteria; 3) willing to sign the consent form. Exclusion criteria for patients:
5
6 1) pregnant or lactating, 2) a history of gastrointestinal surgery, 3) any other diseases that may
7
8 explain the symptom of functional dyspepsia, 4) severe psychological or mental diseases, or 5)
9
10 allergic to adhesive ECG electrodes. Similar to Experiment 1, no limits were set on the body weight
11
12 and height. (Reviewer #2, comment #3). All of the recordings were made using the same recording
13
14 technology and electrode placement as in experiment 1. This time however, no recordings were
15
16 obtained during the feeding stage. The recordings consisted of 30 minutes in the fasting stage
17
18 (before eating) and 30 minutes in the postprandial state (after eating).

19
20 Instead of giving a fixed amount of a solid meal, the subjects in this experiment were asked to
21
22 consume a liquid meal at a rate of 60 mL/min until complete fullness. The meal was prepared by
23
24 dissolving 100g of Nestle Full Cream Milk Powder and 50g of Cola powder in 1120 mL of water.
25
26 The tolerated volume averaged 855 ± 294 mL. The data used to compare the fasting data was the
27
28 data obtained after the maximum intake of the nutrient liquid meal instead of the data during eating
29
30 or drinking. All postprandial data were acquired immediately after termination of eating/drinking.
31
32 (Reviewer #2, comment #4).

32 2) *Signal processing and feature extraction*

33
34 Each ECG signal was processed with the same methods as in experiment 1. This processing
35
36 consisted of R-peak detection, tachogram computation, and resampling. The same time domain
37
38 and frequency domain analyses were performed to extract the eleven features from tachogram
39
40 segments.

41 3) *Artificial neural network*

42
43 We should note that the classification problem to be solved in experiment 2 was different from
44
45 the problem in experiment 1. In experiment 1, the classes to be distinguished were the fasting and
46
47 the feeding stage. In this second experiment, an ANN was designed to classify samples as
48
49 belonging to either the fasting stage (before eating) or the postprandial stage (after eating). The
50
51 classification was performed similarly as in experiment 1. The training examples that were fed to
52
53 the ANN consisted of samples with eleven entries corresponding to features of the HRV. Each
54
55 sample corresponded to a label indicating whether it belonged to the fasting or the postprandial
56
57 stage. The design parameters (training function, performance function, training/validation ratio
58
59 and number of layers) chosen for the ANN were the same as in experiment 1. Again, the three
60
61 different ways of training and testing were performed. The LOSO and LOSO-LOO cross-
62
63
64
65

validation methods were applied with window lengths of 2, 5 and 8 minutes, 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 minutes of fasting data and 30 minutes of postprandial data were used here.

RESULTS

1) Features of HRV

Fig. 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.

Fig. 3 shows some of the time- and frequency-domain features of the HRV of a healthy subject using an epoch length of 8 minutes. The figure shows a complete time window from the fasting stage to the postprandial stage. The fasting stage stopped at 32 minutes, after which the feeding stage started (indicated by the dashed line in Fig. 3). At 48 minutes, 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 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 to 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 (Reviewer #2, comment 5).

2) 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 number of neurons didn't 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 minute 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 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 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 minutes and the number of neurons in the neural network was 10.

3) 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 minute epochs (Table 5).

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 minutes 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 [37-41]. 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 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

1
2
3
4 the feeding stage of experiment 1 lasted only 10 minutes, the epoch length had a major impact on
5 the amount of samples available for training. For example, each subject had just 1 sample in the
6 feeding stage when an 8 minute epoch was used. The tests with 2 and 5 minute epoch lengths had
7 more training samples in the feeding stage. The shorter the epoch length, the more training samples
8 in the feeding stage and the better the ANN was able to identify the feeding stage. This effect was
9 not due to the intrinsic quality of features obtained with certain window lengths, but due to the fact
10 that the ANN had not enough training data per subject. It primarily had a large impact on the
11 sensitivity, and consequently also influenced the accuracy and AUC. Table 3 shows that the
12 highest sensitivity obtained was 0.79 (2 minute epoch) and the lowest was 0.13 (8 minute epoch).
13 Overall, using LOSO-LOO cross-validation improved the accuracy compared to LOSO cross-
14 validation: the mean accuracy increased to 0.85 and the maximum accuracy to 0.88. Again, the
15 number of neurons appeared to have no effect on the performance.

16
17 Experiment 2 was designed to mimic real-world situations by including a large number of
18 subjects, mixing healthy controls and patients, and providing a meal with different size. In practical
19 applications, a large number of subjects can be used to train the ANN since the ECG measurement
20 is noninvasive. Patients with functional dyspepsia were included in the experiment to represent
21 diversity of subjects as they may have different postprandial changes in autonomic functions.
22 Different meal sizes (varied from 400 mL to 1200 mL) were allowed in this experiment to reflect
23 future clinical application. In addition, the ECG data during feeding were excluded to represent
24 future applications that do not need stimulation during feeding. Nevertheless, experiment 2 yielded
25 similar results; the LOSO method results showed no influence from epoch length or number of
26 neurons (Table 4) while the LOSO-LOO method showed higher performance values for shorter
27 epoch lengths (Table 5). Unlike experiment 1, which had a 10 minute feeding period, this
28 experiment had 30 minute fasting and postprandial periods. Therefore, longer epoch lengths did
29 not decrease performance values as drastically as in experiment 1. Performance values still
30 increased with shorter epochs because of their positive effect on the amount of training samples.
31 Table 5 shows that the highest sensitivity was 0.87 (2 minute epoch) and the lowest was 0.80 (8
32 minute epoch) for the LOSO-LOO method. This was an improvement with comparison to the
33 LOSO method, yielding a mean accuracy of 0.64 (Table 4). Neither method showed impact from
34 the number of neurons.

1
2
3
4 It is also important to note that the specificity was similarly influenced by epoch length in
5 experiment 2, whereas experiment 1 showed no correlation between the two. In all experiment 1
6 tests, specificity was considerably higher than sensitivity. In other words, the ANN was capable
7 of detecting fasting better than feeding, a result of the unbalanced sample proportion between the
8 two stages. During the experiments, the ECG was recorded for 1 hour during the fasting stage and
9 only 10 minutes during feeding. As a result, the ANN was trained with more samples from the
10 fasting stage, leading to better performance in detecting this stage than the feeding stage. In
11 general, experiment 2 avoided this issue by having equal fasting and postprandial periods
12 (Reviewer #2, comment #7). This led to more balanced sensitivity and specificity. Lowering the
13 decision threshold can improve the sensitivity at the expense of decreasing the specificity. A trade-
14 off can be made by selecting the desired sensitivity and specificity in the ROC curve and applying
15 the corresponding threshold to the algorithm.

16
17 In both experiments 1 and 2, the final part was to train the ANN on the original subjects and to
18 test it on a new, unseen subject. First, the ANN was not retrained on the new subject; then, it was
19 retrained on one recording of this subject; finally, it was retrained on two recordings of the subject.
20 The purpose of this was to mimic the ANN's training process in a real-life application. If an
21 implantable stimulator were to be used in a patient, it would first need to acquire a sufficient
22 amount of patient-specific ECG to retrain the ANN, which can be done easily before the treatment
23 starts. In this way, the classifier would be specified to match the patient's own characteristics.

24
25 The results of the subject-specific training and testing are reported in Fig. 4 and 5 for experiment
26 1 and 2, respectively. They showed that the performance was, in fact, significantly improved after
27 retraining on subject-specific recordings. In experiment 1 (Fig. 4), we see an exceptionally high
28 specificity of 1 in all three cases. The other performance values were greatly improved when the
29 ANN was retrained on one recording of the new subject. Retraining on one recording was enough
30 to obtain almost maximal performance values when the ANN was tested on the second and third
31 recording. When the network was retrained on the first and the second recording and tested on the
32 third one, there was no more significant improvement. The accuracy, sensitivity, specificity and
33 AUC were 0.993, 0.979, 1 and 1. The perfect amount of data to retrain on should be investigated
34 in further studies that include more than three subject-specific data recordings. In experiment 2
35 (Fig. 5), all performance measures except specificity were greatly improved when subject-specific
36 retraining was included. The specificity was decreased from 0.95 to 0.92 in exchange for a huge
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 improvement in sensitivity from 0.05 to 0.83. Moreover, all the performance values were
5 additionally increased by retraining on two subject-specific recordings rather than one. After
6 retraining on two recordings, the values for accuracy, sensitivity, specificity and AUC were,
7 respectively, 0.95, 0.94, 0.96 and 0.98. We expect these values to further improve with more
8 subject-specific retraining. Comparing the subject-specific retraining in experiment 1 and 2 shows
9 that the dataset on which the ANN is trained beforehand plays a large role in how well it performs
10 initially in detecting food intake in a new subject. Fortunately, subject-specific retraining of the
11 ANN can easily be used to improve this performance up to the desired accuracy.
12
13
14
15
16
17
18

19 The promising results in this study suggest that the HRV can be used to detect food intake
20 dynamically. However, this study has limitations, and a few more issues need to be addressed
21 before the ANN can be used in an implantable stimulation device. Further studies should include
22 more data, particularly in the fasting stage under different situations rather than pure laboratory
23 setting. Conditions other than eating that may alter autonomic functions should be included in
24 future studies, such as exercise, walking, and running. Future work should also consider different
25 types of meals, including snacks, to see how the changes in features (between fasting and feeding)
26 depend on the type and amount of food. Finally, the decision threshold in the ANN should be tuned
27 depending on the desired sensitivity and specificity.
28
29
30
31
32
33
34

35 We would conclude this discussion with some general considerations on the clinical application
36 of ANNs in future neuromodulation therapies for treating obesity, diabetes, and FGIDs. An ANN
37 can first be trained on a general dataset. Then, using the ECG acquired from a candidate patient,
38 the same ANN can be retrained to achieve a higher accuracy. After the retraining, the weights of
39 the ANN are fixed. The pacemaker can then be programmed to include a simple, static, input-
40 output function determined by the final weights of the ANN. We should note that the nature of
41 these treatments does not require perfect detection of fasting and feeding. For example, the
42 decision threshold could be adapted to trigger stimulation when eating a full meal, but not when
43 eating a snack. Changing this threshold is the equivalent of making a trade-off between sensitivity
44 and specificity. With these considerations in mind, ANNs trained for food intake detection have
45 high potential for clinical applications.
46
47
48
49
50
51
52
53
54

55 In conclusion, this study investigates a promising approach to detect food intake using the ECG
56 as primary signal. It uses time- and frequency-domain features derived from the dynamic analysis
57 of the HRV. The features suggest that the sympathovagal balance increases upon food intake. The
58
59
60
61
62
63
64
65

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.

ACKNOWLEDGEMENT

This work was supported by a grant from National Institutes of Health (R01DK107754).

CONFLICT OF INTEREST

There is no conflict of interest to describe.

REFERENCES

- [1] Sallam HS, Chen JDZ. Colon Electrical Stimulation: Potential Use for Treatment of Obesity. *Obesity*. 2011 Sep;19(9):1761-67.
- [2] Arrone LJ. Epidemiology, morbidity, and treatment of overweight and obesity. *J of Clin Psychiatry*. 2001;62(23):13-22.
- [3] Spieker EA, Pyzocha N. Economic impact of obesity. *Prim Care*. 2016 March;43:83–95.
- [4] Department of Health and Human Services [Internet]. 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: Aug. 2, 2018].
- [5] American Diabetes Association [Internet]. 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 Sep;30(9):23-40.

- 1
2
3
4 [7] Talley N. Functional gastrointestinal disorders as a public health problem.
5
6 Neurogastroenterol Motil. 2008 May;20(1):121-9.
7
8
9 [8] Cordera R, Adami GF. From bariatric to metabolic surgery: looking for a “disease modifier”
10
11 surgery for type 2 diabetes. World J Diabetes. 2016 Jan;7(2):27-33.
12
13
14 [9] Department of Health and Human Services, USA [Internet]. Insulin, Medicines, & Other
15
16 Diabetes Treatments. Bethesda, MD: National Institute of Diabetes and Digestive and Kidney
17
18 Diseases; c1995-2019 [cited 2018 Aug 2]. Available from: [https://www.niddk.nih.gov/health-](https://www.niddk.nih.gov/health-information/diabetes/overview/insulin-medicines-treatments)
19
20 [information/diabetes/overview/insulin-medicines-treatments](https://www.niddk.nih.gov/health-information/diabetes/overview/insulin-medicines-treatments).
21
22
23 [10] Huertas-Ceballos A, Macarthur C, Logan S. Dietary interventions for recurrent abdominal
24
25 pain (RAP) in childhood. Cochrane Database Syst Rev. 2002; CD003019.
26
27
28 [11] Lebenthal E, Rossi TM, Nord KS, Branski D. Recurrent abdominal pain and lactose
29
30 absorption in children. Pediatrics. 1981;67:828–832.
31
32
33 [12] Grover M, Drossman DA. Psychotropic agents in function gastrointestinal disorders. Curr
34
35 Opin Pharmacol. 2008 Dec;8(6):715-723.
36
37
38 [13] Bursch B. Psychological/cognitive behavioral treatment of childhood functional abdominal
39
40 pain and irritable bowel syndrome. J Pediatr Gastroenterol Nutr. 2008 Nov;47:706–707.
41
42
43 [14] Whitfield KL, Shulman RJ. Treatment options for functional gastrointestinal disorders: from
44
45 empiric to complementary approaches. Pediatr Ann. 2009 May;38(5):288-294.
46
47
48 [15] Yin J, Abell TD, McCallum RW, Chen JDZ. Gastric neuromodulation with Enterra system
49
50 for nausea and vomiting in patients with gastroparesis. Neuromodulation. 2012 May;15(3):224–
51
52 231.
53
54
55 [16] Greenway F, Zheng J. Electrical stimulation as treatment for obesity and diabetes. J
56
57 Diabetes Sci Technol. 2007 Mar;1(2):251–259.
58
59
60
61
62
63
64
65

- 1
2
3
4 [17] Chen JDZ, Yin J, Wei W. Electrical therapies for gastrointestinal motility disorders. *Expert*
5
6 *Rev Gastroenterol Hepatol*. 2017 May;11(4):407–418.
7
8
9 [18] Apovian CM, Shah SN, Wolfe BM, Ikramuddin S, Miller CJ, Tweden KS, et al. Two-Year
10
11 Outcomes of Vagal Nerve Blocking (vBloc) for the Treatment of Obesity in the ReCharge Trial.
12
13 *Obes Surg*. 2016 Aug;27:169–176.
14
15
16 [19] Vu T, Lin F, Alshurafa N, Xu W. Wearable Food Intake Monitoring Technologies: A
17
18 Comprehensive Review. *Computers*. 2017 Jan;6(1):4.
19
20
21 [20] Amft O. A wearable ear pad sensor for chewing monitoring. *Proceedings of the IEEE*
22
23 *Sensors 2010 Conference*; 2010 Nov 1-4; Waikoloa, USA. IEEE Xplore, 2010.
24
25
26 doi:10.1109/icsens.2010.5690449
27
28 [21] Dong Y, Hoover A, Scisco J, Muth E. A new method for measuring meal intake in humans
29
30 via automated wrist motion tracking. *Appl Psychophysiol Biofeedback*. 2012 Sep;37(3):205–
31
32 215.
33
34
35 [22] Farooq M, Fontana JM, Sazonov E. A novel approach for food intake detection using
36
37 electroglottography. *Physiol Meas*. 2014 May;35(5):739–751.
38
39
40 [23] Farooq M, Sazonov E. Comparative testing of piezoelectric and printed strain sensors in
41
42 characterization of chewing. *Proceedings of the 37th Annual International Conference of the*
43
44 *IEEE Engineering in Medicine and Biology Society*; 2015 Aug 25-29; Milano, Italy. IEEE
45
46 Xplore; 2015:7538-7541
47
48
49 [24] Dong B and Biswas S. Wearable diet monitoring through breathing signal analysis.
50
51 *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine*
52
53 *and Biology Society*; 2013 Jul 3-7; Osaka, Japan. IEEE Xplore; 2013:1186-1189.
54
55
56 [25] Schwartz MW, Woods SC, Porte D, Seeley RJ, Baskin DG. Central nervous system control
57
58
59
60
61
62
63
64
65

1
2
3
4 of food intake. *Nature*. 2000 Apr;404(6778):661–671.

5
6 [26] Messina G, De Luca V, Viggiano A, Ascione A, Iannaccone T, Chieffi S, et al. Autonomic
7 nervous system in the control of energy balance and body weight: personal contributions. *Neurol*
8
9 *Res Int*. 2013 Apr;2013:639280.

10
11
12 [27] Sakaguchi T, Arase K, Fisler JS, Bray GA. Effect of starvation and food intake on
13
14 sympathetic activity. *Am J Physiol*. 1988 Aug;255(2):R284–8.

15
16
17 [28] Lu CL, Zou X, Orr WC, Chen JD. Postprandial changes of sympathovagal balance
18
19 measured by heart rate variability. *Dig Dis Sci*. 1999 Apr;44(4):857–861.

20
21
22 [29] Stuckey MI, Tulppo MP, Kiviniemi AM, Petrella RJ. Heart rate variability and the
23
24 metabolic syndrome: a systematic review of the literature. *Diabetes Metab Res Rev*. 2014
25
26 Nov;30(8):784–793.

27
28
29 [30] Blum A. *Neural Networks in C++: An Object-Oriented Framework for Building*
30
31 *Connectionist Systems*. Wiley; 1992.

32
33
34 [31] Pan J, Tompkins WJ. A real-time QRS detection algorithm. *IEEE Trans Biomed Eng*. 1985
35
36 Mar;32(3): 230–236.

37
38
39 [32] Welch P. The use of fast Fourier transform for the estimation of power spectra: A method
40
41 based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and*
42
43 *Electroacoustics*. 1967;15(2):70–73.

44
45
46 [33] Rangayyan BRM. *Biomedical Signal Analysis: A Case-Study Approach*. IEEE Press Wiley-
47
48 Interscience; 2002.

49
50
51 [34] Heaton J. The Number of Hidden Layers [Internet]. Heaton Research; 2017 Jun 1 [cited
52
53 2018 Aug 1]. Available from: <https://www.heatonresearch.com/2017/06/01/hidden-layers.html>.

54
55
56 [35] Hegenbart S, Uhl A, Vécsei A. Systematic assessment of performance prediction techniques
57
58
59
60
61
62
63
64
65

1
2
3
4 in medical image classification: a case study on celiac disease. *Inf Process Med Imaging*.
5
6 2011;22:498–509.
7
8
9 [36] Chen JD, Lin Z, McCallum RW. Noninvasive feature-based detection of delayed gastric
10 emptying in humans using neural networks. *IEEE Trans Biomed Eng*. 2000 Mar;47(3):409–412.
11
12 [37] Aberle J, Busch P, Veigel J, Duprée A, Roesch T, zu Eulenburg C, et al. Duodenal Electric
13 Stimulation: Results of a First-in-Man Study. *Obes Surg*. 2016;26:369–375.
14
15 [38] Khawaled R, Blumen G, Fabricant G, Ben-Arie J, Shikora S. Intestinal electrical stimulation
16 decreases postprandial blood glucose levels in rats. *Surg Obes Relat Dis*. 2009;5:692–697.
17
18 [39] Ye F, Liu Y, Li S, Chen JDZ. Hypoglycemic Effects of Intestinal Electrical Stimulation by
19 Enhancing Nutrient-Stimulated Secretion of GLP-1 in Rats. *Obes Surg*. 2018;28:2829–2835.
20
21 [40] Li S, Chen JD. Pulse Width-Dependent Effects of Intestinal Electrical Stimulation for
22 Obesity: Role of Gastrointestinal Motility and Hormones. *Obes Surg*. 2017;27:70-77.
23
24 [41] Liu J, Xiang Y, Qiao X, Dai Y, Chen JD. Hypoglycemic effects of intraluminal intestinal
25 electrical stimulation in healthy volunteers. *Obes Surg*. 2011;21:224-30.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figure

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.

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

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 minutes.

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 minutes and the number of neurons is 10.

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 minutes and the number of neurons is 10.

TABLE 1: TIME- AND FREQUENCY-DOMAIN FEATURES

Feature	Description
HR	Mean heart rate of each epoch (in bpm)
mean	
HR SD	Standard deviation of the heart rate in each epoch (in bpm)
NN	Mean normal-to-normal interval, i.e., Mean time interval between consecutive heart beats (R-R interval) in each epoch (in s)
mean	
NN	Difference between longest and shortest R-R interval in each epoch (in s)
Max-Min	
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., $\frac{LF}{LF+HF}$

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

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

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

1

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

Abstract:

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). *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. *Conclusion:* 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

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

1
2
3
4 automated food intake detection methods. Without an automatic food intake detection algorithm,
5 the neuromodulation therapy would have to be applied continuously or depend on the action of
6 patients (manually turn on the device) that imposes a serious compliance issue [19]. Therefore,
7 automatic food intake detection methods are becoming increasingly crucial to obesity, overweight,
8 diabetes, and FGID treatment. Wearable sensors for real-time detection can utilize acoustics [20],
9 gyroscopic sensors [21], electroglottography [22], piezoelectricity [23], and respiratory signals
10 [24]. However, most of these approaches are obtrusive or do not meet expectations for required
11 accuracy. Furthermore, if these sensors were to report food intake to an implanted device, they
12 would additionally require some form of wireless communication as they could not be directly
13 attached to the implanted device. Because of these limitations, there is a need to develop new
14 approaches to detecting food intake that can be incorporated into an implantable stimulator.

15
16 Food intake is highly regulated by the central nervous system [25]. The autonomic nervous
17 system also plays a key role in the control of energy homeostasis and body weight [26]. The
18 autonomic system consists of a sympathetic and a parasympathetic branch, whose activities are
19 affected differently by periods of fasting and food intake. Food intake has been reported to
20 consistently increase sympathetic activity and decrease vagal activity [27], [28].

21
22 The cardiac autonomic function can be assessed by the spectral analysis of heart rate variability
23 (HRV) [28], [29]. The spectrum of HRV displays two major spectral components, i.e., a low-
24 frequency (LF) component (0.04-0.15 Hz) and a high-frequency (HF) component (0.15-0.40 Hz).
25 The LF reflects mainly sympathetic activity with some parasympathetic input, while the HF
26 represents solely parasympathetic activity. In addition, the LF/HF ratio reflects the sympathovagal
27 balance.

28
29 Since food intake alters cardiac autonomic functions that can in turn be measured by changes in
30 the HRV, we hypothesize that the HRV can be used to detect food intake. The HRV signal can be
31 assessed from the electrocardiogram (ECG), whereas the ECG can be recorded noninvasively and
32 repetitively using surface electrodes. The ECG signals obtained from surface electrodes can be
33 used to train an artificial neural network (ANN) and in real neuromodulation applications, the ECG
34 signal can be easily acquired using the implantable neuromodulation device via one electrode
35 placed at its stimulation lead and the case of the stimulator (serving as another electrode) (Review
36 #1, comment #1). Lu et al already assessed the change in LF, HF, and LF/HF ratio after eating a
37 standardized meal; however, they did not consider the dynamic nature of the HRV [28]. Although
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

the HRV is also affected by factors other 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

A. 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.

B. Experiment 1: A Small Sample Series with a Standard Test Meal

1) Experimental Protocol

After 6 hours 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: 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: 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 minutes in the fasting state and 10 minutes during eating. The meal given to the subjects consisted of around 450 calories. 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.

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

3) *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

1
2
3
4 feeding stage. The Levenberg-Marquardt backpropagation function was selected because its
5 performance was among the best. The mean squared normalized error was chosen as the
6 performance measure. During training, the samples were randomly divided into 80% training
7 samples and 20% validation samples.
8
9

10
11 This study considered 3 different epoch lengths: 2, 5, and 8 minutes. In the time-frequency
12 analysis, decreasing the epoch length increases the time resolution but leads to a loss in frequency
13 resolution [33]. Thus, there is a trade-off between frequency resolution and time resolution when
14 the ANN tries to differentiate between samples of the fasting and feeding stage. The minimum 2
15 minute epoch length was selected based on the fact that in practice, occurrences of meals shorter
16 than 2 minutes are quite rare. Moreover, detecting such short timescale changes in the HRV would
17 in practice result in a high false positive rate, as transient changes in the HRV on such a small
18 timescale can be related to numerous other factors than feeding. The upper limit of 8 minutes was
19 determined based on a similar reasoning: further increasing the epoch length would mean that
20 short, small meals would go undetected too easily, as the features are averaged measures over the
21 whole epoch length. This would in turn result in a high false negative rate.
22
23

24
25 Three numbers of neurons in the hidden layer were compared: 3, 5, and 10. The information in
26 the following paragraph is based on [34]. If the number of neurons in the hidden layer is too low,
27 the classification problem will be underfitted, i.e., the neurons will not be able to adequately solve
28 the classification problem. However, having too many neurons may result in overfitting and an
29 increase in training time. Unfortunately, there is no simple way to select the perfect amount of
30 neurons. There are, however, some rule-of-thumb methods, one of which suggests that the size of
31 the hidden layer should be between the size of the input layer and the size of the output layer. With
32 this in mind, we tested the performance of the neural network with 3, 5 and 10 neurons in the
33 hidden layer.
34
35

36
37 The training and testing of the ANNs was performed in several ways, which are explained and
38 discussed in the following paragraphs. First, leave-one-subject-out (LOSO) cross-validation was
39 used. This way of testing and training evaluated how well the classifier performed on unseen, new
40 subjects. Then, leave-one-subject-out and leave-one-out (LOSO-LOO) cross-validation were used,
41 to evaluate how adding samples of specific patients to the training set would improve the accuracy.
42 Finally, the performance of the ANN was tested in a more realistic setting, in which the ANN was
43 first trained on all 16 subjects and then retrained on some recordings of a new, unseen patient. The
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

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.

C. Experiment 2: A Large Sample Series with Varying Meal Size

1) 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: 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:

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 minutes in the fasting stage (before eating) and 30 minutes 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 100g of Nestle Full Cream Milk Powder and 50g 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.

2) *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.

3) *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 minutes, 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 minutes of fasting data and 30 minutes of postprandial data were used here.

RESULTS

1) Features of HRV

Fig. 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.

Fig. 3 shows some of the time- and frequency-domain features of the HRV of a healthy subject using an epoch length of 8 minutes. The figure shows a complete time window from the fasting stage to the postprandial stage. The fasting stage stopped at 32 minutes, after which the feeding stage started (indicated by the dashed line in Fig. 3). At 48 minutes, 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 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 to 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.

2) 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 number of neurons didn't 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-minute 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 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 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 minutes and the number of neurons in the neural network was 10.

3) 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-minute epochs (Table 5).

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 minutes 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 [37-41]. 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 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 minutes, 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-minute epoch was used. The tests with 2- and 5-minute 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 was 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-minute epoch) and the lowest was 0.13 (8 minute epoch). Overall, using LOSO-LOO cross-validation improved the accuracy compared to 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 size. 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 mL to 1200 mL) were allowed in this experiment to reflect 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-minute feeding period, this experiment had 30-minute 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-minute epoch) and the lowest was 0.80 (8-minute epoch) for the LOSO-LOO method. This was an improvement with comparison to the LOSO method, yielding a mean accuracy of 0.64 (Table 4). Neither method showed impact from the number of neurons.

1
2
3
4 It is also important to note that the specificity was similarly influenced by epoch length in
5 experiment 2, whereas experiment 1 showed no correlation between the two. In all experiment 1
6 tests, specificity was considerably higher than sensitivity. In other words, the ANN was capable
7 of detecting fasting better than feeding, a result of the unbalanced sample proportion between the
8 two stages. During the experiments, the ECG was recorded for 1 hour during the fasting stage and
9 only 10 minutes during feeding. As a result, the ANN was trained with more samples from the
10 fasting stage, leading to better performance in detecting this stage than the feeding stage. In
11 general, experiment 2 avoided this issue by having equal fasting and postprandial periods. This led
12 to more balanced sensitivity and specificity. Lowering the decision threshold can improve the
13 sensitivity at the expense of decreasing the specificity. A trade-off can be made by selecting the
14 desired sensitivity and specificity in the ROC curve and applying the corresponding threshold to
15 the algorithm.

16
17 In both experiments 1 and 2, the final part was to train the ANN on the original subjects and to
18 test it on a new, unseen subject. First, the ANN was not retrained on the new subject; then, it was
19 retrained on one recording of this subject; finally, it was retrained on two recordings of the subject.
20 The purpose of this was to mimic the ANN's training process in a real-life application. If an
21 implantable stimulator were to be used in a patient, it would first need to acquire a sufficient
22 amount of patient-specific ECG to retrain the ANN, which can be done easily before the treatment
23 starts. In this way, the classifier would be specified to match the patient's own characteristics.

24
25 The results of the subject-specific training and testing are reported in Fig. 4 and 5 for experiment
26 1 and 2, respectively. They showed that the performance was, in fact, significantly improved after
27 retraining on subject-specific recordings. In experiment 1 (Fig. 4), we see an exceptionally high
28 specificity of 1 in all three cases. The other performance values were greatly improved when the
29 ANN was retrained on one recording of the new subject. Retraining on one recording was enough
30 to obtain almost maximal performance values when the ANN was tested on the second and third
31 recording. When the network was retrained on the first and the second recording and tested on the
32 third one, there was no more significant improvement. The accuracy, sensitivity, specificity and
33 AUC were 0.993, 0.979, 1 and 1. The perfect amount of data to retrain on should be investigated
34 in further studies that include more than three subject-specific data recordings. In experiment 2
35 (Fig. 5), all performance measures except specificity were greatly improved when subject-specific
36 retraining was included. The specificity was decreased from 0.95 to 0.92 in exchange for a huge
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 improvement in sensitivity from 0.05 to 0.83. Moreover, all the performance values were
5 additionally increased by retraining on two subject-specific recordings rather than one. After
6 retraining on two recordings, the values for accuracy, sensitivity, specificity and AUC were,
7 respectively, 0.95, 0.94, 0.96 and 0.98. We expect these values to further improve with more
8 subject-specific retraining. Comparing the subject-specific retraining in experiment 1 and 2 shows
9 that the dataset on which the ANN is trained beforehand plays a large role in how well it performs
10 initially in detecting food intake in a new subject. Fortunately, subject-specific retraining of the
11 ANN can easily be used to improve this performance up to the desired accuracy.
12
13
14
15
16
17
18

19 The promising results in this study suggest that the HRV can be used to detect food intake
20 dynamically. However, this study has limitations, and a few more issues need to be addressed
21 before the ANN can be used in an implantable stimulation device. Further studies should include
22 more data, particularly in the fasting stage under different situations rather than pure laboratory
23 setting. Conditions other than eating that may alter autonomic functions should be included in
24 future studies, such as exercise, walking, and running. Future work should also consider different
25 types of meals, including snacks, to see how the changes in features (between fasting and feeding)
26 depend on the type and amount of food. Finally, the decision threshold in the ANN should be tuned
27 depending on the desired sensitivity and specificity.
28
29
30
31
32
33
34

35 We would conclude this discussion with some general considerations on the clinical application
36 of ANNs in future neuromodulation therapies for treating obesity, diabetes, and FGIDs. An ANN
37 can first be trained on a general dataset. Then, using the ECG acquired from a candidate patient,
38 the same ANN can be retrained to achieve a higher accuracy. After the retraining, the weights of
39 the ANN are fixed. The pacemaker can then be programmed to include a simple, static, input-
40 output function determined by the final weights of the ANN. We should note that the nature of
41 these treatments does not require perfect detection of fasting and feeding. For example, the
42 decision threshold could be adapted to trigger stimulation when eating a full meal, but not when
43 eating a snack. Changing this threshold is the equivalent of making a trade-off between sensitivity
44 and specificity. With these considerations in mind, ANNs trained for food intake detection have
45 high potential for clinical applications.
46
47
48
49
50
51
52
53
54

55 In conclusion, this study investigates a promising approach to detect food intake using the ECG
56 as primary signal. It uses time- and frequency-domain features derived from the dynamic analysis
57 of the HRV. The features suggest that the sympathovagal balance increases upon food intake. The
58
59
60
61
62
63
64
65

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.

ACKNOWLEDGEMENT

This work was supported by a grant from National Institutes of Health (R01DK107754).

CONFLICT OF INTEREST

There is no conflict of interest to describe.

REFERENCES

- [1] Sallam HS, Chen JDZ. Colon Electrical Stimulation: Potential Use for Treatment of Obesity. *Obesity*. 2011 Sep;19(9):1761-67.
- [2] Arrone LJ. Epidemiology, morbidity, and treatment of overweight and obesity. *J of Clin Psychiatry*. 2001;62(23):13-22.
- [3] Spieker EA, Pyzocha N. Economic impact of obesity. *Prim Care*. 2016 March;43:83–95.
- [4] Department of Health and Human Services [Internet]. 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: Aug. 2, 2018].
- [5] American Diabetes Association [Internet]. 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 Sep;30(9):23-40.

- 1
2
3
4 [7] Talley N. Functional gastrointestinal disorders as a public health problem.
5
6 Neurogastroenterol Motil. 2008 May;20(1):121-9.
7
8
9 [8] Cordera R, Adami GF. From bariatric to metabolic surgery: looking for a “disease modifier”
10
11 surgery for type 2 diabetes. World J Diabetes. 2016 Jan;7(2):27-33.
12
13
14 [9] Department of Health and Human Services, USA [Internet]. Insulin, Medicines, & Other
15
16 Diabetes Treatments. Bethesda, MD: National Institute of Diabetes and Digestive and Kidney
17
18 Diseases; c1995-2019 [cited 2018 Aug 2]. Available from: [https://www.niddk.nih.gov/health-](https://www.niddk.nih.gov/health-information/diabetes/overview/insulin-medicines-treatments)
19
20 [information/diabetes/overview/insulin-medicines-treatments](https://www.niddk.nih.gov/health-information/diabetes/overview/insulin-medicines-treatments).
21
22
23 [10] Huertas-Ceballos A, Macarthur C, Logan S. Dietary interventions for recurrent abdominal
24
25 pain (RAP) in childhood. Cochrane Database Syst Rev. 2002; CD003019.
26
27
28 [11] Lebenthal E, Rossi TM, Nord KS, Branski D. Recurrent abdominal pain and lactose
29
30 absorption in children. Pediatrics. 1981;67:828–832.
31
32
33 [12] Grover M, Drossman DA. Psychotropic agents in function gastrointestinal disorders. Curr
34
35 Opin Pharmacol. 2008 Dec;8(6):715-723.
36
37
38 [13] Bursch B. Psychological/cognitive behavioral treatment of childhood functional abdominal
39
40 pain and irritable bowel syndrome. J Pediatr Gastroenterol Nutr. 2008 Nov;47:706–707.
41
42
43 [14] Whitfield KL, Shulman RJ. Treatment options for functional gastrointestinal disorders: from
44
45 empiric to complementary approaches. Pediatr Ann. 2009 May;38(5):288-294.
46
47
48 [15] Yin J, Abell TD, McCallum RW, Chen JDZ. Gastric neuromodulation with Enterra system
49
50 for nausea and vomiting in patients with gastroparesis. Neuromodulation. 2012 May;15(3):224–
51
52 231.
53
54
55 [16] Greenway F, Zheng J. Electrical stimulation as treatment for obesity and diabetes. J
56
57 Diabetes Sci Technol. 2007 Mar;1(2):251–259.
58
59
60
61
62
63
64
65

- 1
2
3
4 [17] Chen JDZ, Yin J, Wei W. Electrical therapies for gastrointestinal motility disorders. *Expert*
5
6 *Rev Gastroenterol Hepatol*. 2017 May;11(4):407–418.
7
8
9 [18] Apovian CM, Shah SN, Wolfe BM, Ikramuddin S, Miller CJ, Tweden KS, et al. Two-Year
10
11 Outcomes of Vagal Nerve Blocking (vBloc) for the Treatment of Obesity in the ReCharge Trial.
12
13 *Obes Surg*. 2016 Aug;27:169–176.
14
15
16 [19] Vu T, Lin F, Alshurafa N, Xu W. Wearable Food Intake Monitoring Technologies: A
17
18 Comprehensive Review. *Computers*. 2017 Jan;6(1):4.
19
20
21 [20] Amft O. A wearable ear pad sensor for chewing monitoring. *Proceedings of the IEEE*
22
23 *Sensors 2010 Conference*; 2010 Nov 1-4; Waikoloa, USA. IEEE Xplore, 2010.
24
25
26 doi:10.1109/icsens.2010.5690449
27
28 [21] Dong Y, Hoover A, Scisco J, Muth E. A new method for measuring meal intake in humans
29
30 via automated wrist motion tracking. *Appl Psychophysiol Biofeedback*. 2012 Sep;37(3):205–
31
32 215.
33
34
35 [22] Farooq M, Fontana JM, Sazonov E. A novel approach for food intake detection using
36
37 electroglottography. *Physiol Meas*. 2014 May;35(5):739–751.
38
39
40 [23] Farooq M, Sazonov E. Comparative testing of piezoelectric and printed strain sensors in
41
42 characterization of chewing. *Proceedings of the 37th Annual International Conference of the*
43
44 *IEEE Engineering in Medicine and Biology Society*; 2015 Aug 25-29; Milano, Italy. IEEE
45
46 Xplore; 2015:7538-7541
47
48
49 [24] Dong B and Biswas S. Wearable diet monitoring through breathing signal analysis.
50
51 *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine*
52
53 *and Biology Society*; 2013 Jul 3-7; Osaka, Japan. IEEE Xplore; 2013:1186-1189.
54
55
56 [25] Schwartz MW, Woods SC, Porte D, Seeley RJ, Baskin DG. Central nervous system control
57
58
59
60
61
62
63
64
65

1
2
3
4 of food intake. *Nature*. 2000 Apr;404(6778):661–671.

5
6 [26] Messina G, De Luca V, Viggiano A, Ascione A, Iannaccone T, Chieffi S, et al. Autonomic
7 nervous system in the control of energy balance and body weight: personal contributions. *Neurol*
8
9 *Res Int*. 2013 Apr;2013:639280.

10
11
12 [27] Sakaguchi T, Arase K, Fisler JS, Bray GA. Effect of starvation and food intake on
13
14 sympathetic activity. *Am J Physiol*. 1988 Aug;255(2):R284–8.

15
16
17 [28] Lu CL, Zou X, Orr WC, Chen JD. Postprandial changes of sympathovagal balance
18
19 measured by heart rate variability. *Dig Dis Sci*. 1999 Apr;44(4):857–861.

20
21
22 [29] Stuckey MI, Tulppo MP, Kiviniemi AM, Petrella RJ. Heart rate variability and the
23
24 metabolic syndrome: a systematic review of the literature. *Diabetes Metab Res Rev*. 2014
25
26 Nov;30(8):784–793.

27
28
29 [30] Blum A. *Neural Networks in C++: An Object-Oriented Framework for Building*
30
31 *Connectionist Systems*. Wiley; 1992.

32
33
34 [31] Pan J, Tompkins WJ. A real-time QRS detection algorithm. *IEEE Trans Biomed Eng*. 1985
35
36 Mar;32(3): 230–236.

37
38
39 [32] Welch P. The use of fast Fourier transform for the estimation of power spectra: A method
40
41 based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and*
42
43 *Electroacoustics*. 1967;15(2):70–73.

44
45
46 [33] Rangayyan BRM. *Biomedical Signal Analysis: A Case-Study Approach*. IEEE Press Wiley-
47
48 Interscience; 2002.

49
50
51 [34] Heaton J. The Number of Hidden Layers [Internet]. Heaton Research; 2017 Jun 1 [cited
52
53 2018 Aug 1]. Available from: <https://www.heatonresearch.com/2017/06/01/hidden-layers.html>.

54
55
56 [35] Hegenbart S, Uhl A, Vécsei A. Systematic assessment of performance prediction techniques
57
58
59
60
61
62
63
64
65

1
2
3
4 in medical image classification: a case study on celiac disease. *Inf Process Med Imaging*.
5
6 2011;22:498–509.
7
8
9 [36] Chen JD, Lin Z, McCallum RW. Noninvasive feature-based detection of delayed gastric
10 emptying in humans using neural networks. *IEEE Trans Biomed Eng*. 2000 Mar;47(3):409–412.
11
12 [37] Aberle J, Busch P, Veigel J, Duprée A, Roesch T, zu Eulenburg C, et al. Duodenal Electric
13 Stimulation: Results of a First-in-Man Study. *Obes Surg*. 2016;26:369–375.
14
15 [38] Khawaled R, Blumen G, Fabricant G, Ben-Arie J, Shikora S. Intestinal electrical stimulation
16 decreases postprandial blood glucose levels in rats. *Surg Obes Relat Dis*. 2009;5:692–697.
17
18 [39] Ye F, Liu Y, Li S, Chen JDZ. Hypoglycemic Effects of Intestinal Electrical Stimulation by
19 Enhancing Nutrient-Stimulated Secretion of GLP-1 in Rats. *Obes Surg*. 2018;28:2829–2835.
20
21 [40] Li S, Chen JD. Pulse Width-Dependent Effects of Intestinal Electrical Stimulation for
22 Obesity: Role of Gastrointestinal Motility and Hormones. *Obes Surg*. 2017;27:70-77.
23
24 [41] Liu J, Xiang Y, Qiao X, Dai Y, Chen JD. Hypoglycemic effects of intraluminal intestinal
25 electrical stimulation in healthy volunteers. *Obes Surg*. 2011;21:224-30.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit <http://www.adobe.com/go/acrreader>.

Windows is either a registered trademark or a trademark of Microsoft Corporation in the United States and/or other countries. Mac is a trademark of Apple Inc., registered in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit <http://www.adobe.com/go/acrreader>.

Windows is either a registered trademark or a trademark of Microsoft Corporation in the United States and/or other countries. Mac is a trademark of Apple Inc., registered in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit <http://www.adobe.com/go/acrreader>.

Windows is either a registered trademark or a trademark of Microsoft Corporation in the United States and/or other countries. Mac is a trademark of Apple Inc., registered in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit <http://www.adobe.com/go/acrreader>.

Windows is either a registered trademark or a trademark of Microsoft Corporation in the United States and/or other countries. Mac is a trademark of Apple Inc., registered in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit <http://www.adobe.com/go/acrreader>.

Windows is either a registered trademark or a trademark of Microsoft Corporation in the United States and/or other countries. Mac is a trademark of Apple Inc., registered in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit <http://www.adobe.com/go/acrreader>.

Windows is either a registered trademark or a trademark of Microsoft Corporation in the United States and/or other countries. Mac is a trademark of Apple Inc., registered in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit <http://www.adobe.com/go/acrreader>.

Windows is either a registered trademark or a trademark of Microsoft Corporation in the United States and/or other countries. Mac is a trademark of Apple Inc., registered in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit <http://www.adobe.com/go/acrreader>.

Windows is either a registered trademark or a trademark of Microsoft Corporation in the United States and/or other countries. Mac is a trademark of Apple Inc., registered in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit <http://www.adobe.com/go/acrreader>.

Windows is either a registered trademark or a trademark of Microsoft Corporation in the United States and/or other countries. Mac is a trademark of Apple Inc., registered in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.