Mobile Nutrition Monitoring System: Qualitative and Quantitative Monitoring

Nina Reščič
nina.rescic@ijs.si
Department of Intelligent Systems,
Jožef Stefan Institute
International Postgraduate School
Jozef Stefan
Ljubljana, Slovenia

Marko Jordan
Department of Intelligent Systems,
Jožef Stefan Institute
Ljubljana, Slovenia

Jasmijn de Boer
ConnectedCare
Nijmegen, Netherlands

Ilse Bierhoff
ConnectedCare
Nijmegen, Netherlands

Mitja Luštrek
mitja.lustrek@ijs.si
Department of Intelligent Systems,
Jožef Stefan Institute
Ljubljana, Slovenia

ABSTRACT

The WellCo project[1] aims to provide a mobile application featuring a virtual coach for behaviour changes aiming to achieve for healthier lifestyle. The nutrition monitoring module consists of two main parts - qualitative (Food Frequency Questionnaire) and quantitative (eating detection and bite counting). In this paper we present the nutrition monitoring module that connects both monitoring aspects as implemented in the virtual coach (mobile application).

KEYWORDS

nutrition monitoring, eating detection, FFQ

1 INTRODUCTION

Proper nutrition habits are beneficial for healthy lifestyle and help to prevent many chronic diseases, such as cancer, diabetes and hypertension. Automated monitoring has become really important i nutrition monitoring, but in only gives quantitative information (when is the user eating, how much did he eat...), while qualitative information (what is the user eating) is acquired by using 24 hour food recall diaries or by using Food Frequency Questionnaires (FFQs). In the WellCo project we aimed to developed a user friendly nutrition module, which monitors qualitative and quantitative aspects of users' nutrition. We combined the self-reported FFQ, Extended Short Form Food Frequency Questionnaire (ESFFFQ), developed and validated in the project project [5], with automated monitoring by using a commercially available wearable smartwatch. This paper describes the developed module and the improvements we made since our previous papers [5, 2, 7].

By using wrist-worn devices to collect data, it is possible to recognize eating gestures [4] or even count 'bites' or assess caloric intake [10]. Mirtchou et al. [3] explored eating detection by using several sensors and combining real-life and laboratory data. Edison et al. [8] proposed a method that recognizes each intake gesture separately and later the intake gestures within 60 minutes interval are clustered.

For qualitative monitoring we evaluated both dietary recalls and FFQs as self-reporting methods. However, dietary recalls require typing or complex food item selection which can be cumbersome on mobile devices, so we opted for FFQ. FFQs are the most commonly selected tools in nutrition monitoring as they are efficient, cost-effective and non-invasive [9, 6].The developed FFQ covers all key aspects of healthy diet, and is modular, so that only questions pertaining to certain aspects can be asked. This is important in ubiquitous settings where one wishes to minimize the required inputs from the user.

To our knowledge the developed application module is the first one to combine qualitative (validated FFQ) and quantitative monitoring (bite counting method) and to provide recommendations based on data gathered by monitoring.

2 METHOD

2.1 Method Overview

The paper describes the nutrition monitoring module developed in the Wellco project.

The **qualitative monitoring** starts with a five-question questionnaire that provides essential information about the user's diet. Based on this, some goals to improve the user's nutrition can already be recommended. However, the users are invited to answer a more extensive questionnaire that paints a more complete picture and allows recommending more goals. This questionnaire is an extended version of a validated questionnaire, and the extension was validated by us [5]. How successful the users are at achieving their goals is monitored with goal-specific questions on a bi-weekly basis.

The **quantitative monitoring** uses the accelerometer and gyroscope in a smartwatch to detect micromovements related to eating (e.g., picking up food, putting it into the mouth). From a sequence of such micromovement, we then recognise whether the user has made one "bite" (taken the food to the mouth). The improved method uses a Convolutional neural network to recognise the micromovements and a LSTM neural network to recognise bites. The latter achieved higher accuracy so it was the one selected to be integrated into the WellCo system.

[1] http://wellco-project.eu

Information Society 2020, 5–9 October, 2020, Ljubljana, Slovenia

2.2 FFQ - Qualitative Monitoring

When choosing goals that would help users of the WellCo virtual coach towards behavioural changes for healthier lifestyle, we were leaning on national dietary recommendation and dietary recommendations for elderly, combined with expert knowledge by the nutritionist involved in the project. A summary of national dietary recommendations is presented in Table 1.

Guidelines specifically for the elderly are very similar to national dietary recommendations for all three countries involved in pilots (Italy, Spain and Denmark), but they put additional emphasis on dairy consumption, as this is a good source of proteins and calcium, which are beneficial and often under-consumed; drinking enough water, as dehydration is often a problem with elderly; and leucine consumption (in milk, peanuts, oatmeal, peanuts, fish, poultry, egg white, wheat sprouts, etc). Given these recommendations, we chose goals we will suggest WellCo users to follow and use in order to improve their diet: *fruit consumption*, *vegetable consumption*, *salt consumption*, *fat consumption*, *fibre consumption*, *protein consumption*, *salt consumption*, *fish consumption* and *water consumption*.

In our search for a comprehensive but still short FFQ we found a validated questionnaire named Short Food Frequency Questionnaire (SFFQ)[1], which consists of 23 questions and fully covers five of our chosen goals – *fruit* and *vegetable consumption*, *sugar consumption*, *fat consumption* and *fish consumption*. To cover the four missing goals (protein, fibre, salt and water consumption) we added additional 8 questions, turning the SFFQ into the so-called Extended Short Food Frequency Questionnaire (ESFFQ). The validation of the questionnaire is described in our previous paper [5].

2.3 Quantitative Monitoring

The main objective of the smartwatch-based nutrition monitoring is bite counting (counting the number of time the user takes food to the mouth).

The bite-counting algorithm described in [2] was used as the base for all of the following work. When deciding how to present the results of the developed algorithm to the users in the mobile application, we had to make some improvements to our model. As the number of bites does not really give much useful information to the users, we decided to join individual bites into meals and to recognize meals as *snack*, *small meal* or *big meal*.

2.3.1 Datasets. To construct the bite detection algorithm, we created the Wild Meals Dataset (WMD). It includes 51 sessions and 99 meals, with known starting and ending time points, belonging to 11 unique subjects, recorded 'in-wild'. For 68 of those meals we have also obtained the approximate number of the corresponding bites, since the subjects were asked to count them while eating. Additionally we used the publicly available The Food Intake Cycle (FIC) dataset and The Free Food Intake Cycle (FreeFIC). All datasets contains tri-axial signals from accelerometers and gyroscopes in wrist devices with the sampling frequency of 100 Hz.

2.3.2 Meal detection method. The algorithm for meal detection was comprised of two parts: in the first part probabilities that given time periods are part of eating were assigned, whereas in the second part these probabilities were grouped together to form a meal.

First we linearly interpolated all accelerometer and gyroscope measurements as well as the probabilities of bites to 4Hz frequency. Next, the normalization was applied to interpolated accelerometer and gyroscope data. We constructed 90 s long sliding windows with a 2.5s step. Each window contained 360 of the previously obtained accelerometer, gyroscope and bite probability values (obtained with CNN and LSTM networks as described in [2]). 4Hz frequency was used to achieve faster training and predicting, while also enabling us to construct longer windows. A window was labelled as a positive instance, if the majority of the window belonged inside a meal.

To solve this machine learning task, an inception-type neural network was constructed, with the added GRU layers at the end. The inception part of the network is mainly made of two types of inception blocks. Both types consist of convolutional layers and end with a filter concatenation. The B block includes also a max pooling operation. Each block in the network is succeeded by a max pooling layer. The entire architecture is presented in Table 1. The inputs were transformed in the (batch size,timestamps,1,7) shape. "Prep" (preparation) in Table 1 refers to the yellow convolutional layers in Figure 5, whereas "Pool proj" refers to 1x1 convolutional layer after 4x1 max pooling layer. The final model used approximately 130 K parameters.

With the intention of smoother and better learning, the ratio between positive and negative instances was fixed to 1:2. During the sampling, we actually focused more on problematic areas, by first predicting with the network and then selecting problematic instances to train on. Learning rate was set to keep decreasing every few epochs. Certain hyper-parameters were subject to optimization during cross-validation, with the help of `hyperopt` library. The function to minimize was categorical cross entropy.

In the next part, the outputs $\in [0,1]$ of the neural network, which represent the probabilities that the given windows are eating instances, are taken to form possible/candidate meals. This is done in the following manner:

- **Round 1**: Find all probabilities, denoted as beacons, that are higher than a p1 threshold. Include also all probabilities that are closer than t1 seconds to any of the beacons. Set all the other probabilities temporarily to 0.
- **Round 2**: Find all probabilities that are higher than a p2 threshold and group them together, if they are immediately next to each other. For each group find the time distance to its nearest group. Finally remove all groups that have either 1 or 2 members and are more than t2 seconds away from the corresponding nearest group.
- **Round 3**: If there exist any two groups of the form [A,B] and [C,D], where $0 \leq C - B \leq t3$ (all in seconds), combine these two groups together to form a new group, [A,D]. This means that indices in [A,D] can now represent the probabilities of zero as well.
- **Round 4**: Similar as Round 3, but with a t4 parameter in place of t3.

At this point the probabilities of windows, previously temporarily set to zero, are switched back to their original values. For the final model, we obtained the following values of the above hyperparameters:

Since p2 > p1, this means that Round 1 in this particular case was not necessary, although in some other cases it could have been. Once the candidate meals have been obtained, the features are constructed for the ensemble of random forest, support vector machine, knn and gradient boosting algorithms. The ensemble

Table 1: Architecture of the network

Type	Units/Nodes	Kernel/stride	Output	1x1	4x1 prep	4x1	6x1 prep	6x1	Pool
Inception-A			360x1x128	32		64		32	
Max pool		3x1/2	180x1x128						
Inception-B			180x1x128	32	64	64	16	16	16
Max pool		3x1/2	90x1x128						
Inception-B			90x1x128	32	64	64	16	16	16
Max pool		3x1/2	45x1x128						
Inception-B			45x1x128	32	64	64	16	16	16
Max pool		3x1/2	23x1x128						
GRU			23x32						
GRU			32						
Dense	64		64						
Dropout(0.36)			64						
Dense	2		2						

Table 2: Hyperparameters.

p1	t1(sec)	p2	t2(sec)	t3(sec)	t4(sec)
0.46	61	0.87	120	63	61

makes the final decision whether a candidate meal is in fact a meal or not. The following features are created for each candidate meal:

- The mean, standard deviation, the 25th, 50th and 75th percentile of all the probabilities inside a given candidate meal.
- The mean and standard deviation of the first and second half of a potential meal, separately.
- The mass of all the future probabilities inside all the potential meals closer than 3 hours to a given candidate meal, divided by their time centre.
- The mass of all the past probabilities inside all the potential meals closer than 3 hours to a given candidate meal, divided by their time centre.

Hyper-parameters for each model in the ensemble, as well as p1, t1, p2 t2, t3 and t4 values, were calculated with a cross-validation, with the help of hyperopt library. The function to minimize was negative F1-score.

3 RESULTS

3.1 Bite Counting

In Table 4 we present the results of evaluation of our work. The analysis of the entire pipeline is based on Leave-One-Subject-Out double cross-validation. For calculation of the above statistics the following definitions were used:

- True positive prediction of a meal: any prediction of the respective meal for which the majority of the prediction laid inside the ground truth meal. If there was more than one prediction of eating for a certain meal, only one prediction is actually counted as a true positive, whereas all the others are not regarded as a false positive.. This is due to the possibility that the subjects didn't eat their entire recording time; as such it did not seem reasonable to penalize the pipeline for predicting more than one meal, however, only one true positive is counted in order not to encourage the algorithm to predict a bundle of eating instances.

Table 3: Results of bite recognition and meal detection algorithm.

	F1-score	precision	recall	cov_area	outside_area
Avg.	0.76	0.88	0.72	0.81	0.03

Table 4: Example of recommendations for qualitative monitoring (*goal_sugar*) and quantitative monitoring (*nutrition_number_of_meal*).

goal_sugar	It seems you don't eat enough vegetables. Vegetables are important sources of many nutrients, such as vitamins, minerals and dietary fibre. Try to eat 2 servings of vegetables per day. Serving is 1 cup of fresh or half cup of cooked vegetables.
nutrition_number_of_meal	Try to eat 3–5 meals per day (e.g. 3 bigger, 2 smaller). Avoid snacking between meals.

- For F1-score, precision and recall, def A was used, while cov_area and outside_area used def B. However, double cross-validation results show that all ground truth meals, with one exception, had at most one corresponding, true positive predicted meal.
- Covered area (cov_area): for a given ground truth meal, the length of the areas, which laid inside the ground truth meal, of the corresponding true positive meals, divided by the length of the ground truth meal.
- Outside area (outside_area): for a given predicted, true positive meal, the length of the area that laid outside the corresponding ground truth meal, divided by the length of the predicted meal.

3.2 Application Implementation

The application shows users the detected meals, number of bites and score quality for the chosen goals (see Figure 1). Based on the results we additionally show the user recommendations to follow in order to improve their nutrition. Example for recommendations for both, qualitative and quantitative monitoring is shown in table.

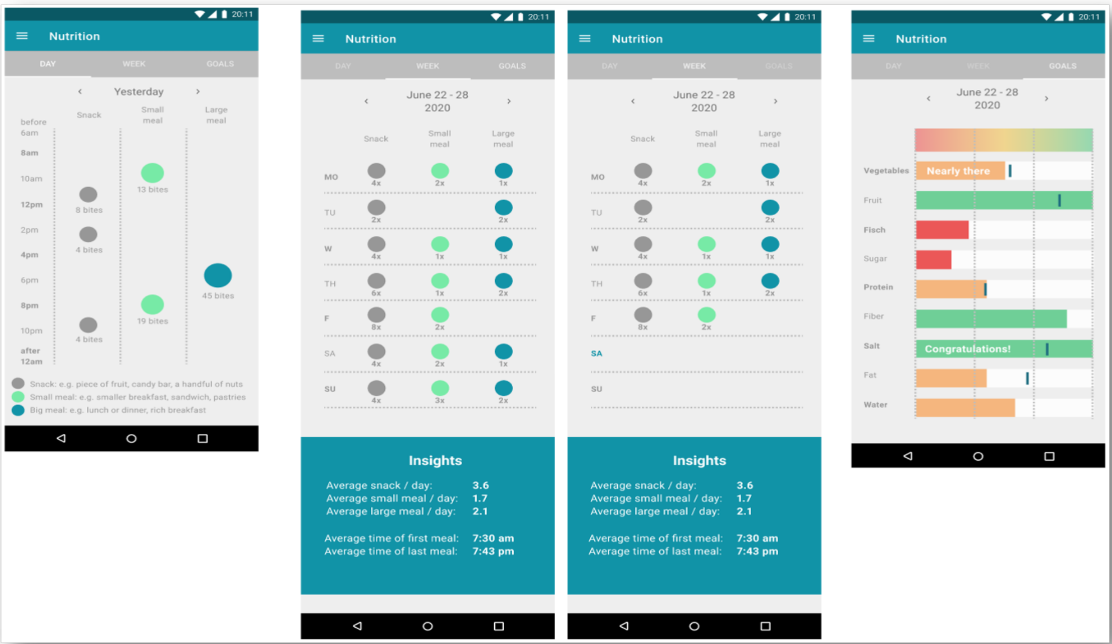

Figure 1: Application view for both monitoring tasks.

4 CONCLUSION

The developed nutrition monitoring module consists of two parts - qualitative monitoring and quantitative monitoring. Both of the developed modules are implemented in a mobile application. In our future work we would like to improve the developed eating detection and bite counting algorithms.

The developed FFQ (ESFFFQ) can be used to support a wide range of nutrition goals and minimizes the number of questions asked, so it is suitable for mobile nutrition monitoring. To make the application user friendly the questions from the FFQ will not be asked all at the same time, but separately during a course of fortnight. This means that some of the questions won't be asked, hence it is really important to ask the right questions. In our future work we will try to explore the problem of question ranking. With this we would be able to ask the questions in a specific order and loose as few information as possible.

5 ACKNOWLEDGMENTS

WellCo Project has received funding from the European Union's Horizon2020 research and innovation program under grant agreement No 769765.

REFERENCES

[1] Christine L Cleghorn, Roger A Harrison, Joan K Ransley, Shan Wilkinson, James Thomas, and Janet E Cade. 2016. Can a dietary quality score derived from a short-form ffq assess dietary quality in uk adult population surveys? *Public Health Nutrition*, 19, 16, 2915–2923. DOI: 10.1017/S1368980016001099.

[2] 2019. Counting bites with a smart watch. In *Slovenian Conference on Artificial Intelligence : proceedings of the 22nd International Multiconference Information Society*. Volume A, 49–52.

[3] Mark Mirtchouk, Drew Lustig, Alexandra Smith, Ivan Ching, Min Zheng, and Samantha Kleinberg. 2017. Recognizing eating from body-worn sensors: combining free-living and laboratory data. 1, 3. DOI: 10.1145/3131894.

[4] Raul I. Ramos-Garcia, Eric R. Muth, John N. Gowdy, and Adam W. Hoover. 2014. Improving the recognition of eating gestures using intergesture sequential dependencies. *IEEE Journal of Biomedical and Health Informatics*, 19, 3, 825–831.

[5] Nina Reščič, Eva Valenčič, Enej Mlinarič, Barbara Koroušić Seljak, and Mitja Luštrek. 2019. Mobile nutrition monitoring for well-being. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers* (UbiComp/ISWC '19 Adjunct). London, United Kingdom, 1194–1197.

[6] JS Shim, K Oh, and HC Kim. 2014. Dietary assessment methods in epidemiologic studies. *Epidemiol Health*, 36. DOI: 10.4178/epih/e2014009.

[7] Simon Stankoski, Nina Reščič, Grega Mežič, and Mitja Luštrek. 2020. Real-time eating detection using a smartwatch. In Junction Publishing, USA.

[8] Edison Thomaz, Irfan Essa, and Gregory D. Abowd. 2015. A practical approach for recognizing eating moments with wrist-mounted inertial sensing. In Association for Computing Machinery, New York, NY, USA. ISBN: 9781450335744. DOI: 10.1145/2750858.2807545.

[9] Frances Thompson and T Byers. 1994. Dietary assessment resource manual. *The Journal of nutrition*, 124, (December 1994), 2245S–2317S. DOI: 10.1093/jn/124.suppl_11.2245s.

[10] Shibo Zhang, William Stogin, and Nabil Alshurafa. 2018. I sense overeating. *Inf. Fusion*, 41, C, (May 2018), 37–47. DOI: 10.1016/j.inffus.2017.08.003.