Demo Abstract: Speech2Health: Tracking Calorie Intake with Unstructured Spoken Data on Mobile Devices

Niloofar Hezarjaribi
Washington State University
nhezarja@eecs.wsu.edu

Seyed Ali Rokni
Washington State University
alirokni@eecs.wsu.edu

Ramin Fallahzadeh
Washington State University
rfallah@eecs.wsu.edu

Hassan Ghasemzadeh
Washington State University
hassan@eecs.wsu.edu

ABSTRACT

Diet and physical activity are important factors in self-management and prevention of many chronic diseases. The always-carried nature of smart-phones makes them advantageous for eating behavior administration. We introduce a voice-based system, called Speech2Health, which devises speech processing, natural language processing (NLP), and text mining techniques. In this demo, we first present the design and implementation of our platform. We present a mobile app, web portal, and back-end storage for nutrition database. Secondly, we evaluate utility of Speech2Health for calorie intake monitoring in real-time. Furthermore, we performed a user study comparing our system with two other nutrition monitoring approaches.

CCS CONCEPTS

• **Computer systems organization** → **Embedded systems**; *Redundancy*; Robotics; • **Networks** → Network reliability;

KEYWORDS

Speech Recognition, String Matching, Machine Learning, Natural Language Processing

ACM Reference format:
Niloofar Hezarjaribi, Seyed Ali Rokni, Ramin Fallahzadeh, and Hassan Ghasemzadeh. 2017. Demo Abstract: Speech2Health: Tracking Calorie Intake with Unstructured Spoken Data on Mobile Devices. In *Proceedings of The 16th ACM/IEEE International Conference on Information Processing in Sensor Networks, Pittsburgh, PA USA, April 2017 (IPSN 2017),* 2 pages.
DOI: http://dx.doi.org/10.1145/3055031.3055044

IPSN 2017, Pittsburgh, PA USA

DOI: http://dx.doi.org/10.1145/3055031.3055044

1 INTRODUCTION

Research in nutrition monitoring and physical activity aims to enhance lifestyle behaviors. Mobile health is becoming a popular topic as people become increasingly aware about their health conditions. Studies show that 45% of adults use smart-phones [4]; moreover, using mobile devices is advantageous due to their various embedded sensors and features such as GPS, camera, and wireless connectivity. Current technologies using smart-phones either require users to enter food intake information in form of text [3, 5, 6] or image [1, 2]. Text-based approaches require end-user to enter diet data manually. Image-based approached are time sensitive; moreover, visual information by itself may not be enough to accurately compute calorie intake.

Previous research studies suggest that voice input is more beneficial when the user is disabled, keyboard of the device is small, spelling is important, and user's hands and eyes are busy. We introduce a new framework for voice-based nutrition monitoring and propose a hierarchy of data processing modules including speech to text conversion, nutrient attribute extraction, and string matching in order to extract nutrient information from spoken language.We devise a nutrient attribute extraction by utilizing Natural Language Processing (NLP) and developing a novel pattern mapping approach for tagging nutrient data. To this end, we utilize NLP for tokenizing the sentences and tagging the words. Utilizing the tokenized words, their associated tags, and a dataset of extracted patterns, we assign nutrient-specific tags to the words. Furthermore, we develop a 2-tier string search algorithm including an exact matching and an edit-distance based approximation matching to search food items from a nutrition database and to compute nutrient values such as calorie intake, protein, etc.

2 SYSTEM DESIGN

The GUI of Speech2Health consists of four pages: 1) homepage, which shows the nutrient intake of the user in a progress chart; 2) diary page, wherein nutrition intake of the user is recorded using a push button or by entering text; 3) progress page, shows a summary of user's nutrition intake; 4) message page, where user can send/receive messages to/from clinician. A screen-shot of the mobile app is shown in Figure 1. Once the user's data is recorded, the mobile app sends the data to a web host. In case the web host is not available, the data is saved in a local database. This is done through an

API that allows the data to be sent from the phone to the web host with an Internet connection. We utilized Google's

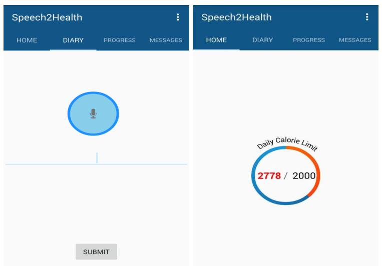

Figure 1: An example of the pages in mobile app

open source speech recognition API for Android platform. This way, the spoken data is converted into natural language text. We implemented our nutrient attribute extraction and string matching modules in Python. For nutrient attribute extraction, we employed Natural Language Toolkit (NLTK). In order to perform approximate matching, we utilized Levenshtein function. Furthermore, the nutrition dataset was obtained from the USDA database and was embedded in hash-map to allow fast lookup with a time complexity of O(1).

3 CASE STUDY

The system is evaluated with real data collected with 20 subjects in an experimental setting that mimics noise-free as well as realistic noisy environments where the spoken data are entered to the system. A script was provided to the subjects incorporating variety of food names. In order to increase the nutrient calculation accuracy a similarity algorithm was embedded in the system. The algorithm aimed to fix error of the application and compensate incompleteness of the database. After audio signal is converted to text, the output is fed into the Nutrient Attribute Extraction module. This module is responsible for tokenizing and parsing the sentence. After tagging the tokens utilizing POS, nutrition specific data are located based on predefined set of patterns obtained from the database. The sentences are tokenized which breaks them to stream of words. Next, each word is read by the POS tagging module and predefined tags are assigned to them. This module detects food name, portion size, and time-stamp based on predefined patterns. An example of this procedure is shown in Figure 2.

The accuracy of calculated calorie using Nutrient Attribute Extraction in presence of error-free text is calculated using equation 1, wherein Cal_{NAE} and Cal_{Actual} are the amount of calculated and actual calorie of the script respectively. The accuracy of the system is 97.69% on average.

$$ACC_{CalorieCalculation} = (\frac{Cal_{NAE}}{Cal_{Actual}}) \times 100 \quad (1)$$

We assess user acceptance of our nutrition monitoring approach with 10 additional participants to compare voice-based nutrition recording with two other diet recording systems including text-based and image-based methods in a week-long experiment. The result is show in Table 1.

Table 1: Questionnaire results analysis, 1 the worst and 5 is the best numbers for each category

Categories	Voice	Text	Image
Attention			
Frequency of apps' usage	3.89	3.13	2.75
Interaction			
User satisfaction of apps' ease of use	4.25	3.13	2.25
Privacy			
User satisfaction of apps' privacy	3.56	3.75	2.38
Impact			
Rating of each app	2.44	2.00	1.56
Rating			
likliness of Speech2Health be the primary app	3.25	N/A	N/A

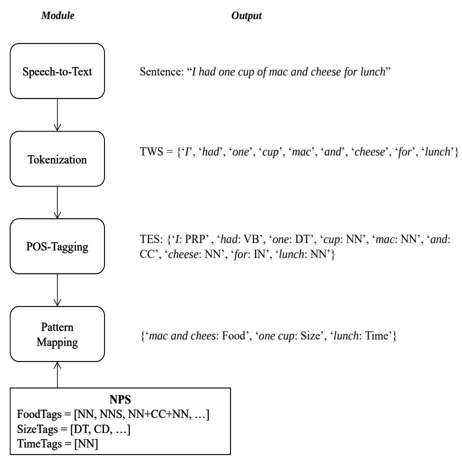

Figure 2: An example of Nutrient Attribute Extraction

4 FUTURE WORK

Our goal is to provide user with other nutrient values (e.g. protein, fiber, etc.) in addition to calorie intake. Moreover, we plan on using our system in conjunction with other sensor modalities to develop a hybrid nutrition monitoring platform.

REFERENCES

[1] G. Ciocca, P. Napoletano, and R. Schettini. Food recognition and leftover estimation for daily diet monitoring. In *International Conference on Image Analysis and Processing*, pages 334–341. Springer, 2015.

[2] H. He, F. Kong, and J. Tan. Dietcam: Multiview food recognition using a multikernel svm. *IEEE journal of biomedical and health informatics*, 20(3):848–855, 2016.

[3] S. S. Intille, C. Kukla, R. Farzanfar, and W. Bakr. Just-in-time technology to encourage incremental, dietary behavior change. In *AMIA*. Citeseer, 2003.

[4] L. Rainie. Smartphone ownership update: September 2012. *Pew Internet & American Life Project*, 2012.

[5] J. J. Rodrigues, I. M. Lopes, B. M. Silva, and I. d. L. Torre. A new mobile ubiquitous computing application to control obesity: Sapofit. *Informatics for Health and Social Care*, 38(1):37–53, 2013.

[6] C. C. Tsai, G. Lee, F. Raab, G. J. Norman, T. Sohn, W. G. Griswold, and K. Patrick. Usability and feasibility of pmeb: a mobile phone application for monitoring real time caloric balance. *Mobile networks and applications*, 12(2-3):173–184, 2007.