A Digital Technology Framework to Optimise the Self-Management of Obesity

Patrick McAllister
School of Computing and Mathematics
Ulster University
Newtownabbey, Northern Ireland
Mcallister-p2@email.ulster.ac.uk

Anne Moorhead
School of Communication
Ulster University
Newtownabbey, Northern Ireland
a.moorhead@ulster.ac.uk

Huiru Zheng
School of Computing and Mathematics
Ulster University
Newtownabbey, Northern Ireland
h.zheng@ulster.ac.uk

Raymond Bond
School of Computing and Mathematics
Ulster University
Newtownabbey, Northern Ireland
rb.bond@ulster.ac.uk

 ISBN 978-1-4503-4462-3/16/09…$15.00 DOI: http://dx.doi.org/10.1145/2968219.2978096

Abstract
Obesity is increasing globally and can cause major chronic conditions. Much research has been completed in utilising digital technologies to optimise the self-management of obesity. This research proposes an obesity management framework which highlights digital technologies to promote self-management of obesity. This work discusses preliminary research using image classification to promote food logging and crowdsourcing to determine calorie content of food images through aggregating the predictions of experts and non-experts. Preliminary results from image classification show SMO classifier achieved 73.87% accuracy in classifying 15 food items, which is promising as computer vision methods could be incorporated into food logging methods. Crowdsourcing results show that aggregated expert group mode percentage error was +2.60% (SD 3.87) in predicting calories in meals and non-expert group mode percentage error was +29.07% (SD 20.48). Further analysis on the crowdsourcing dataset will be completed to ascertain how many experts or non-experts is needed to get the most accurate calorie prediction.

Author Keywords
Obesity, management, image, classification, crowdsourcing, food, calories, classifier

Introduction

Obesity is a major health concern in the UK, Ireland and internationally [1,5,4]. Obesity is used to described an individual who is excessively overweight. Being obese can have a detrimental effect on an individuals' health as it can contribute to chronic conditions such as diabetes, bowel cancer, heart disease, and hypertension [2]. The main cause of obesity is attributed to individuals who consume high amount of calories but do not burn of the energy through physical activity or exercise, excess energy is then stored as fat in the body. Adults with a Body Mass Index (BMI) greater than 30 are likely to be obese. The World Health Organisation (WHO) states that 41 million children under the age of 5 years were obese worldwide in 2014 [3]. In the UK, research revealed that 67.1% of men and 57.2% of women aged 16 years and over were overweight or obese in 2013 [4]. The Northern Ireland Health Survey 2014/15 stated that 60% of adults were overweight or obese; and 28% of children were overweight or obese [4]. The Foresight Report 2012 suggested that a 15% of males and 15% of females under 20 years of age will be classed as obese by 2025 [5].

Related Work

Much work has been completed in using digital technologies to promote management of obesity. For example, there has been an increase in the use of obesity management applications. Research has shown that keeping a logbook can help promote and maintain weightloss [6]. Obesity management services and applications are available across many devices; web applications offer the ability to document energy intake. Other research use images as a means to document energy intake. In [7], image analysis and crowdsourcing was combined to develop a system that determines the nutritional content from an image. The system crowd sources nutritional analysis from the Amazon Mechanical Turk Platform. The image is analysed to determine food type and portion size by the number of users on this platform and then calorie content is extracted from an API. Other research uses machine learning techniques to identify food types. In [8] research was conducted to predict what food item is present in images. This research [8] uses SURF (Speeded-Up-Robust-Features) with bag-of-features (BoG) and achieved 81.55% in classifying food images correctly. Research in [9] dataset achieved 50.76% accuracy in classifying food images through incorporating Random Forest classification with SURF and Colour features.

Research Framework

Figure 1 defines the areas in the proposed obesity management framework. Each component is associated with an area related to obesity management. These components have been informed by the literature review into current methods and technologies that have been used to promote the management of obesity [7,8,9,10]. This framework seeks to combine different methods to inform the development of an application to help promote the management of obesity and research has been completed in component 2, energy intake.

Methodology

This paper presents the work in Energy Intake monitoring. Two studies have been carried out: food image classification and crowdsourcing.

Energy Intake Input – Food Image Classification

This component will be concerned with using machine learning to promote food logging by classifying food items within an image to make it convenient for the user to document their daily intake. Research has shown that self-monitoring is crucial to weight loss [6]. Experiments were completed using a food image dataset [11]. A combination of local and global feature types was extracted from the images. Fifteen food types were selected with 100 images in each folder. Images were manually segmented to remove non-food items in images. Features extracted are Speeded-Up-Robust-Features (SURF), LAB colour features, local binary patterns (LBP) features, and SFTA (Segmented Fractal Texture Analysis) textual features. SURF features and colour features were extracted using a Bag-of-Features (BoF) model [12] as described in Figure 2.

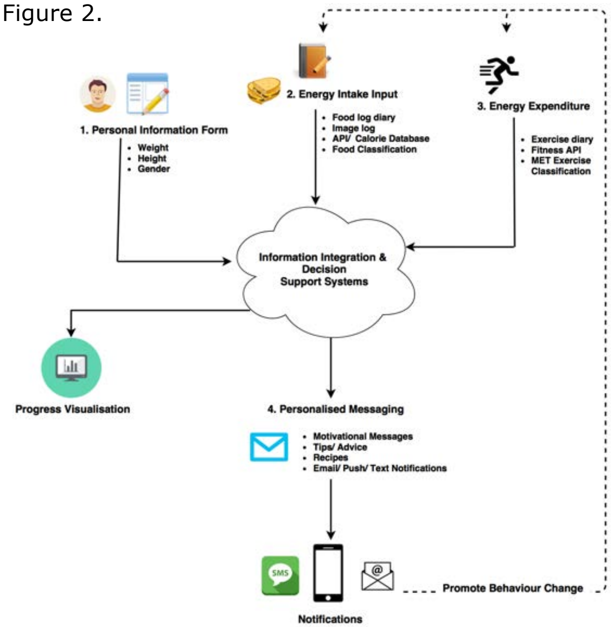

Figure 1. Proposed Obesity Management Framework

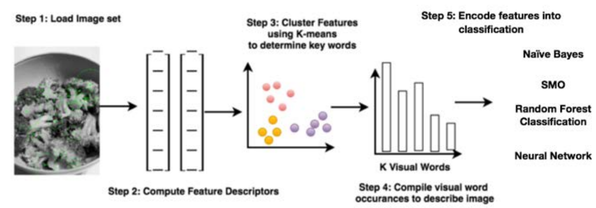

Figure 2. Illustration of the BoF process.

Four Machine learning classifiers, Naïve Bayes, Sequential minimal optimization (SMO), Random Forest and Neural Network were used in the study for comparison. Table 1 lists the classifier parameters that were used in the experiments. Matlab (vR2016a) was used to extract features from the image training set and Weka (v3.7.13) was used for the classification. Preliminary evaluation metrics such as accuracy were used to measure the performance using 10-fold cross validation.

Table 1: Table showing classifier parameters used.

Classifier	**Parameters**
Naïve Bayes	Weka Default
SMO	Kernel: PolyKernel
Random Forest	300 Trees
Neural Network	2 layers, 100 nodes in each layer.

Energy Intake Input – Crowdsourcing

The second research study involved in researching the use of crowdsourcing to determine the calorie content in meals. The use of crowdsourcing to predict calories in meals has been explored with promising results [7]. The main aim of this study was to determine the feasibility of identifying calories of meals in photographs. This work is a continuation of research completed in [10]. The objective of this research is to determine the accuracy of individuals estimating calories of meals using an online survey and to determine differences in accuracy between experts and non-experts. Two user groups were recruited, experts and non-experts. The experts group consisted of dieticians and nutritionists. The non-expert group consists of adults aged 18-65 who are self-reported as healthy and are not dieticians or nutritionists. Both groups completed the same online survey which consists of a series of demographic questions and also 15 photographs of meals. They were asked to estimate the calorie content in each photograph. The survey also consisted of a section that asked participants about their confidence levels in predicting calorie content. Statistical analysis was applied to measure percentage agreement e.g. percentage error between experts and non-experts, determining difference between mode and mean of experts and non-experts. When preparing the data for analysis, subjects that partially completed the survey were disregarded. Data was analysed using Microsoft Excel (v2016).

Results

Food Image Classification

Figure 3 shows the accuracy results of the classification experiments.

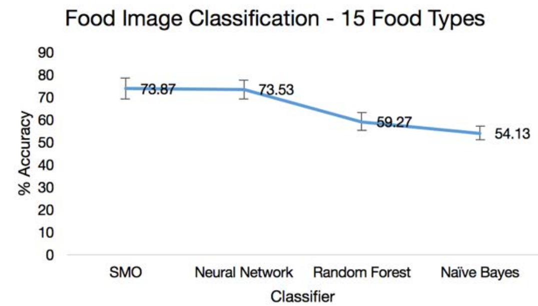

Figure 3: Results of 10-Fold Cross Validation using 4 classifiers.

Results from experiments using a combination of features extracted from the image dataset (SURF, Colour, LBP, SFTA) show that the Sequential Minimal Optimisation (SMO) classifier achieved the highest accuracy with 73.87% (SD 4.56).

Crowdsourcing

Results revealed that the average percentage difference between the non-expert group predictions and the actual calorie content was 40.62%(SD 29.39) (n=126) and 9.25% (SD 9.24) for expert (n=22). Percentage error was calculated to measure how accurate the aggregated calorie predictions for each group compared to the actual content of each meal. For non-expert group the average mode percentage error was 29.07% (SD 20.48) and for expert group the average mode percentage error was 2.60% (SD 3.87). Table 2 contains a list of performance metrics that were used to determine which method provides the closest match to the actual calorie content. This was completed by listing the average absolute calorie difference for each meal for each group. An overall average was computed and This was then compared it to the true calorie content.

Table 2: Calorie difference between non-experts and experts identifying calories in meals.

Group & Metric	Calorie difference (Kcal)
All Mean	124.13
Experts Mean	31.58
Non-Experts Mean	141.54
All Mode	48.73
Experts Mode	8.66
Non-Experts Mode	107.26

Discussion & Future Work

In regards to classifying food images experiments, SMO achieved the highest accuracy in predicting food images with 73.87% (SD 4.56). Future work will include expanding the amount of food types used. Research will be completed in changing parameters of each classifier to obtain optimal results. From the results the 'expert mode' result achieves the lowest calorie difference compared to other group metrics used. The results also show the expert group was predicted calories with a 9.25% difference compared to the true calorie content. Non-expert groups were not as successful, with a 40.62% percentage difference. In regards to crowdsourcing, further evaluation analysis will be completed in comparing individual participant predictions against aggregated group predictions. More participants will be invited to complete the survey to gather more responses.

References

1.WHO | Obesity. Who.int, 2016. http://www.who.int/topics/obesity/en/..

2. NHLBI. 1998. NHLBI Obesity Education Initiative Expert Panel on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults. NIH publication no. 98-4083.

3. Childhood overweight and obesity. World Health Organization, 2016. http://www.who.int/dietphysicalactivity/childhood/en/.

4. Health-ni.gov.uk, 2016. https://www.health-ni.gov.uk/sites/default/files/publications/dhssps/hsni-first-results-14-15.pdf.

5. Gov.uk, 2016. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/287937/07-1184x-tackling-obesities-future-choices-report.pdf.

6. National Heart Lung and Blood Institute and Nhlbi National Institutes of Health (NIH) National Heart, Lung, and Blood Institute. 1998. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults. The Evidence Report, NIH Publication No. 98-4083. Archives of Internal Medicine 158, Suppl 2: 51S–209S. http://doi.org/10.1001/jama.2012.39

7. Jon Noronha, Eric Hysen, Haoqi Zhang, and Krzysztof Z Gajos. 2011. Platemate: crowdsourcing nutritional analysis from food photographs. Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11: 1. http://doi.org/10.1145/2047196.2047198

8. Y Kawano and K Yanai. 2013. Real-Time Mobile Food Recognition System. Computer Vision and Pattern Recognition Workshops (CVPRW), 2013 IEEE Conference on: 1–7. http://doi.org/10.1109/CVPRW.2013.5

9. Lukas Bossard, Matthieu Guillaumin, and Luc Van Gool. 2014. Food-101 - Mining discriminative components with random forests.

Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 446–461. http://doi.org/10.1007/978-3-319-10599-4_29

10. Moorhead, Anne, Raymond Bond, and Huri Zheng. Smart Food: Crowdsourcing of experts in nutrition and non-experts in identifying calories of meals using smartphone as a potential tool contributing to obesity prevention and management. Bioinformatics and Biomedicine (BIBM), 2015 IEEE International Conference on. IEEE, 2015..

11. 1. Mei-Yun Chen, Yung-Hsiang Yang, Chia-Ju Ho, et al. 2012. Automatic Chinese food identification and quantity estimation. SIGGRAPH Asia: 1–4. http://doi.org/10.1145/2407746.2407775

12. Image Category Classification Using Bag of Features - MATLAB & Simulink Example. Uk.mathworks.com, 2016. http://uk.mathworks.com/help/vision/examples/image-category-classification-using-bag-of-features.html?requestedDomain=uk.mathworks.com.