

# A Mobile-Based Diet Monitoring System for Obesity Management

Bruno Vieira Resende e Silva<sup>1</sup>, Milad Ghiasi Rad<sup>1</sup>, Juan Cui<sup>1\*</sup>, Megan McCabe<sup>2</sup> and Kaiyue Pan<sup>3</sup>

<sup>1</sup>Department of Computer Science and Engineering at University of Nebraska, Lincoln, USA

<sup>2</sup>Department of Complex Bio Systems at UNL, Lincoln, USA

<sup>3</sup>Department of Computer Science at McGill University, Canada

## Abstract

Personal diet management is key to fighting the obesity epidemic. Recent advances in smartphones and wearable sensor technologies have empowered automated food monitoring through food image processing and eating episode detection, with the goal to conquer drawbacks of traditional food journaling that is labour intensive, inaccurate, and low adherent. In this paper, we present a new interactive mobile system that enables automated food recognition and assessment based on user food images and provides dietary intervention while tracking users' dietary and physical activities. In addition to using techniques in computer vision and machine learning, one unique feature of this system is the realization of real-time energy balance monitoring through metabolic network simulation. As a proof of concept, we have demonstrated the use of this system through an Android application.

**Keywords:** Food classification; Food image processing; Food volume estimation; Healthy eating; Metabolic network simulation; Activity tracking; Wearable device

## Introduction

Healthy diet with balanced nutrition is key to the prevention of overweight and obesity, cardiovascular disease, as well as other life-threatening metabolic comorbidities such as type 2 diabetes, and cancer [1], which warrants personal diet monitoring. In contrast to the traditional manual food logging that is time consuming and hard to sustain [2], smartphones applications such as MyFitnessPal, LoseIt and Fooducate, have demonstrated high level of usability [3,4] by providing effective dietary feedback [5]. However, many of these applications require significant amount of manual input from users and poorly perform in assessing the exact ingredients and food portion of a meal [6], which has hindered users' experience in a long run. In order to make food journaling easier and more accurate, we proposed to develop a novel automated system that integrates diet recoding via interactive food recognition and assessment through smartphone apps, exercise detection via wearable devices, and personalized energy balance monitoring through metabolic network modeling, and just-in-time dietary intervention. For instance, a user can take photos of his/her meal using smartphone and within seconds, receive nutritional information about the underlying food items. With more food logging activities, the system is capable of identifying individuals' eating patterns and rendering interventions, e.g., recommending healthier food or providing warnings when detecting bad eating habits. To accomplish this, we first explored new methodologies in Computer Vision and Machine Learning to address key issues in each of the following components: 1) a comprehensive food image database that contains diverse and abundant images from a large number of food classes, in order to avoid the food discrepancy when training a food-image classifier [7]; 2) a food segmentation strategy that can correctly identify all items in an image from the background regardless the lighting conditions or if the foods are mixed or not [8]; 3) a Machine Learning model to be trained for classifying each segmented item; 4) volume and weight estimation to be performed on each food item, followed by the nutrient analysis [9,10]. In addition, one unique feature included in this system is a metabolic network simulation that takes into consideration individual's basal metabolism and monitors the real-time energy production in the presence of nutrients available in the meal.

The rationale behind the modeling is that, with different respective metabolic baselines, individuals may respond differently in terms of energy production to the same meal or similar combination of nutrients. This paper is organized as follows: it starts with a general review of the related work in food image processing and classification, followed by an overview of the entire workflow of this project. We then present the details of our methodologies and results, followed by the discussion on remaining challenges and future outlook to close the paper.

## Related Work

As mentioned above, a complete automated food monitoring and dietary management system should be composed of a comprehensive food image database, robust food segmentation and food classification, accurate food volume estimation, and insightful dietary feedback and advices. It is notable that every step involves technical challenges, which has been documented in the related research. Current food image datasets vary in many aspects, e.g., type of cuisine, number of food groups, and total images per food class. For instance, Menu-Match dataset [11] contains 41 food classes and a total of 646 images captured in 3 distinct restaurants while PFID [12] has 61 classes with a total of 1098 pictures captured in fast food restaurants and laboratory. There is no default food image database for general classification purpose since most databases archive specific food type. For examples, the UNIMIB2016 database has Italian food images from a campus dining hall and the UEC Food-100 [8] consists of items from Chinese culinary. Chen [13] and PFID consist of images from traditional Japanese dishes and American fast food, respectively while Food-101 [14] and UEC Food-256 [15] contain a mix of eastern and western food. Except for

\*Corresponding author: Juan Cui, Department of Computer Science and Engineering at University of Nebraska, Lincoln, NE 68588 USA, Tel: 4024725023; E-mail: [jcui@unl.edu](mailto:jcui@unl.edu)

Received March 26, 2018; Accepted April 02, 2018; Published April 06, 2018

Citation: Silva BVRE, Rad MG, McCabe M, Pan K, Cui J (2018) A Mobile-Based Diet Monitoring System for Obesity Management. J Health Med Informat 9: 307. doi: 10.4172/2157-7420.1000307

Copyright: © 2018 Silva BVRE, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

food type, factors such as if the picture was obtained in the wild or in a controlled environment, or whether the images is segmented or not have been taken into consideration when developing those databases. The objective of segmentation, when dealing with food, is to localize and extract food items from the image. For examples, one approach asks user to draw bounding boxes over food items on the smartphone screen, and performs segmentation using GrabCut algorithm over selected areas [16]. Another strategy segments items by integrating four methods to detect candidate region, including the whole image (assuming each image has one food), Deformable Part Model (DPM), a method utilizing sliding windows to detect object regions), circle detector (detecting circular in an image), and JSEG segmentation to segment regions. In addition, the work presented in [17] tries to segment food by its ingredients and their spatial relationship applying Semantic Texton Forest (STF). Segmentation of food images is often challenging due to the following reasons: 1) The food image may not present specific attributes such as edges and defined contour [17]; 2) one food item can be underneath the other, being obstructed and hidden in the given image [17]. Furthermore, external factors such as illumination can also interfere negatively in this aspect, where shadows can be identified as part of the food or even a new food item [18]. Currently, there are two major classification strategies for food image recognition, traditional Machine Learning approach using handcrafted features and Deep Learning approach. The former usually start with a set of visual features extracted from the food image and use them to train a prediction model based on Machine Learning algorithms such as Support Vector Machine (SVM), Bag of Features, or K Nearest Neighbors. For example, one study uses features of SIFT (Scale-invariant feature transform), LBP (Local Binary Pattern), color and Gabor filter, with a multiclass Adaboost. Menu-Match extracts features of SIFT, LBP, color, HOG (Histogram of Oriented Gradient) and MR8 to train a SVM classifier. However, there is a common concern that general image features, as listed above, may not be descriptive enough to distinguish foods since the properties of the same food may change when the food is prepared in different ways [19]. For example, Penne and Spaghetti have same color and texture but distinct shape. On the other hand, it has been recently shown that the deep learning classification often outperforms traditional Machine Learning approaches. For example, in [20], color and HOG features are integrated to a strategy similar to Bag of Features, called Fisher Vectors, which achieved accuracy of 65.3% on UEC Food-100. Based on the same database, the Deep Learning architecture DCNN-FOOD [21] showed an improvement of 13.5% over the handcrafted method. Next, estimation of food volume and nutrient content represents another challenge in automatic food analysis. In fact, not even an expert dietitian can estimate the total calories without a precise instrument, e.g., a scale. Crowdsourcing [21] and a depth sensor camera have been applied for food volume estimation and nutrition assessment. In addition, user's finger was also used as reference while taking a picture of the plate to estimate food volume. Similarly, another study used a checkerboard to help obtaining depth information alongside camera calibration [22]. Last, users can receive feedbacks based on the detected food habit. For examples, one introduces a Semantic Healthcare Assistant for Diet and Exercise (SHADE) that can identify user habits and generate suggestions not only for diet, but also for exercise for diabetic control [23]. Similarly, Lee et al. presents a personal food recommendation agent that can creates a meal plan according to a person's lifestyle and particular health needs towards a certain health goal [24].

### Overview of the System and Methods

In this study, we propose a new system that comprises the following functional modules (Figure 1): 1) a food image and nutrition fact

database used for training a food classification model and assessing nutrients of each classified food; 2) an interactive smartphone system for image-based food recognition; 3) a metabolic network simulation for monitoring the real-time energy balance, which was integrated in such automated system for the first time; 4) an intervention module which identifies individual's eating patterns based on logged meals and activity information and provides feedback. To access this system, users interact with a web-based system through a smartphone application that captures images from meals and exercise activities from a wearable device (currently using Fitbit). The output consists of classified food and nutritional information corresponding to the meal. By monitoring the individual's energy production from food and energy expenditure from exercise, our system is able to provide users their real-time energy balance and recommendations about the best eating time and food portion. After detecting user' eating habits, the intervention module can also generates timely warnings and feedback, as recommended in Kerr DA [5]. Below we will briefly describe the design of each component and focus on the technical implementation mainly on food recognition and assessment and energy minoring in this study.

### Food recognition and assessment

The food recognition system is designed to deal with the challenges in food image segmentation, classification, and volume and nutrient estimation. Figure 1 shows the workflow of this module. The smartphone app requires users to take four pictures of meals, one from the top of

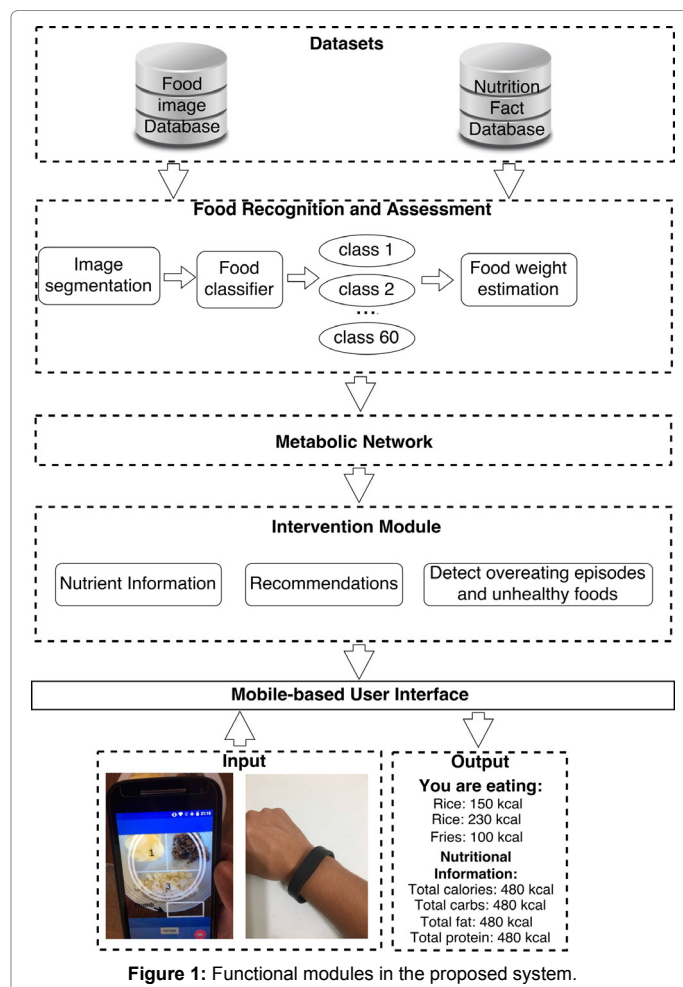


Figure 1: Functional modules in the proposed system.

the plate and one from each of the three sides of the plate. The plate image will be partitioned into 3 fixed areas in order to improve food localization and segmentation. The user will be asked to place each food item in one of the partitions if possible, as seen on Figure 2. We adopt the user's fingertip as a reference to ease the volume and nutrients estimation. The length and width of the selected finger will be asked at the first registration of the app.

### Food segmentation

To facility the segmentation, we first partition the image view into three sections for users to place their plate on the center of the camera screen (Figure 2) and cover each food item on each section. Once the image center is determined, Otsu's segmentation [25] was first performed to extract food items based on a threshold value that separates background to foreground, assuming that the image has only food (foreground) and plate (background). Subsequently, a color-image clustering [26] is performed to separate the finger from all three side views (Figure 2).

### Feature extraction and food classification

Visual information from food images are extracted using feature extractors such as Local Binary Patterns (LBP), color information, texture and Scale-Invariant Feature Transform (SIFT), known as handcrafted features. Specifically, we divided the cropped image center into four quadrants, and from each extracted color, the Histogram of Oriented Gradients (HOG) was extracted from 32 x 32 and 16 x 16 grids on the cropped area. The LBP feature was extracted following a modified approach: instead of 1 feature vector [27], we combined features from radius 1, 3, 5, thresholding the center pixel against its 8, 16 and 24 adjacent neighbours (horizontal, vertical, and diagonal), respectively. The resulted vector has more features than the original LBP that encodes more information from the image. Particularly, it captures larger scale structure information that is important to differentiate one (food) image from another. Speeded Up Robust Features (SURF) [28] and Gabor features [29] were also extracted. Based on the reported performance on food image classification using SVM, Bag of Features [30], and K Nearest Neighbors, we decided to train our system using quadratic SVM [31] based on the collected features on each image. To train an effective food-classification model and benchmark the

prediction performance, we selected the most comprehensive and widely-used database, Food-101, in this study, along with our locally compiled new food classes from the internet. The new expanded database contains 60 food classes, with 1,000 images each. All newly-collected images have been resized to 299 x 299 pixels. For classification purpose, we first divided all images into training set (70%), testing set (20%), and independent validation set (10%). Meanwhile, 10-fold cross-validation was performed for feature evaluation. In contrast to the traditional handcrafted approach, we also applied the Deep Learning method in this study that automatically extracts the features from each image before classification. The pre-trained model Inception V3 available on Google's Tensor Flow framework [32] was used after we modified its final layer to classify classes in our food image dataset.

### Weight estimation

After identifying all food items from an image, it is important to assess the nutrient content, e.g., amount of carbohydrates, sugar, proteins, lipids, and calorie, which will require weight estimation, another major challenge. Specifically, each classified food item will have its area (by the top picture) and height (based on side images) estimated using user's fingertip as a reference object. As illustrated in Figure 2, the fingertip and the food items need to be positioned in specific areas when taking the pictures. After segmentation of the top image, we have the total pixels that form the finger area and the food items, respectively, ready for the estimation of the size of each food area. Based on the side pictures. One can estimate the food height after segmentation. The final step is to multiply the area by height, which results in food volume. As mentioned in the previous section, our system contains a database of food items with correlated volume/weight and nutrient information, which was collected, based on a nutritional facts table from the USDA Food Composition Database [33] and can be directly used for the calculation of nutrient content based on the estimated volume and weight.

### Metabolic network modeling

After discovering what is on the plate, nutrition, and the energy consumption, all information will be sent to the metabolic network to monitor individual's energy balance. The metabolic modeling was performed based on the major ATP-production related metabolic pathways, as demonstrated in Figure 3. Specifically, the nutrients identified from each meal, including mainly proteins, carbohydrates, sugar, and fats, are input to the system to be broken down through three main interconnected pathways participating in the metabolism of proteins, polysaccharides, and lipids. The carbohydrate pathways are the backbone of the energy production network where the breakdown of big carbohydrate molecules is the first step of a series of reactions that produces alpha-DGlucose as the input for the glycolysis pathway. This pathway alongside with pyruvate oxidation, TCA cycle, oxidative phosphorylation, amino acid, and fatty acid pathways complete the model as central energy-related metabolism. Each of the twenty amino acids may participate in one or more metabolic pathways, including carbohydrates and lipids pathways. On the other hand, lipids will be broken down into smaller fatty acid molecules, which is easier for ATP production. Combining these set of reactions helps to build a comprehensive network model to be used in the proposed tool.

The first step in metabolic modeling is to build an independent database achieving parameters regarding participating metabolites and chemical reactions in each selected metabolic pathway. Pathway information was collected from the KEGG database [34] while the initial concentrations were collected from HMDB database [35] and

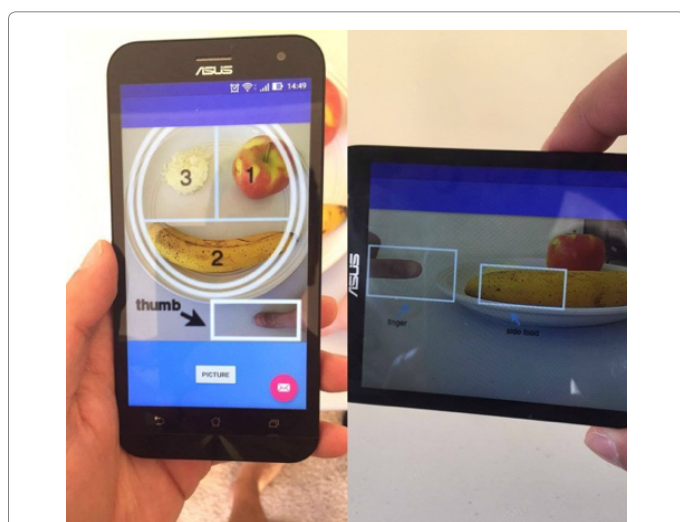
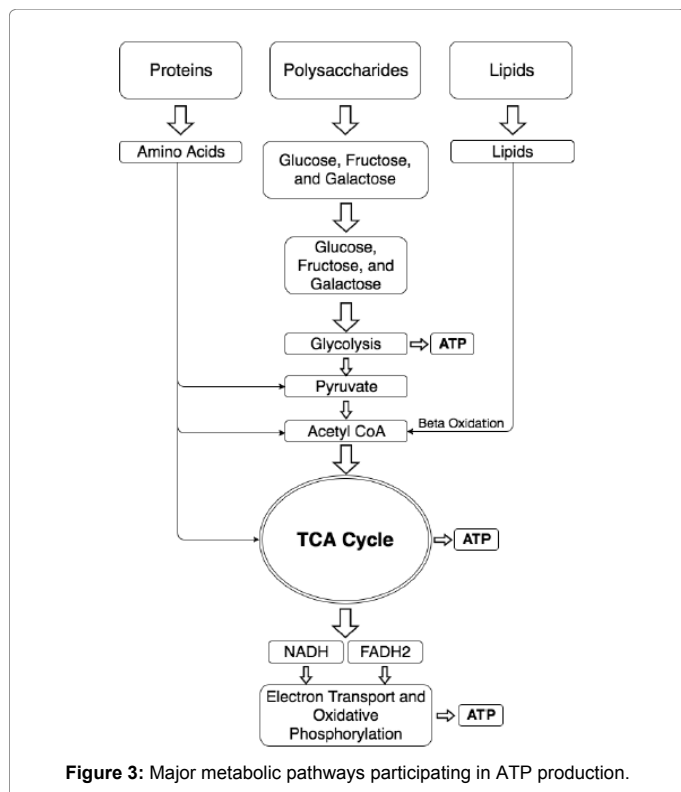


Figure 2: Pictures taken from the top and side of the plate for classification and weight estimation (with one finger as the reference).





literature, e.g., [36]. Using the Michaelis-Menten equations, ODEs (Ordinary Differential Equations) were derived for each metabolite and Runge-Kutta method was used to calculate the concentration of each metabolite for each time step [37]. As this study focuses on the energy level, the ATP and glucose level were monitored. The real-time energy balance between energy produced by food intake and energy expenditure from exercise will determine the overall energy gain or loss of each individual. More detailed description of the metabolic system was presented in our recent work [38].

### Intervention module

The intervention module is responsible for generating insightful information about each meal. It is basically a junction of three interventional functions, which are: 1) creating a comprehensive report that contains all nutrient information about the meal, nutrition report about each meal; 2) detecting if users are overeating or consuming unhealthy food, and recommending them to reduce quantity or take alternative healthier food; 3) sending alert to users when they tend to overeat according to the detected eating habit. The comprehensive report consists of complete nutrition information, such as calories, protein, and carbohydrate, based on the size of each classified food item in the meal. The recommendations are generated at each meal of a day (breakfast, lunch, and dinner) based on the energy balance of the user. Alongside with exercise data from activity trackers, which indicates the total of calories burned throughout the day, it will verify if the user consumes more energy than needed that possibly leads to weight gain. If that happens, smaller food amount will be recommended, or healthier food will be suggested to substitute unhealthy ones. For the latter, foods are labelled according to their levels of fat, saturated fats, sugar, and salt, where low calorie/healthy foods are marked as green, medium level ones as yellow, and high calorie/unhealthy ones are red according to [39]. This traffic light like system has shown that people

start eating healthier when they see such a color-based label [40]. A similar system has utilized additional nutrients such as cholesterol, sodium and dietary fibre to classify American [41]. We labelled the food items in our nutrition database based on this recent approach. Advisory messages are displayed to warn users of unhealthy eating or alter them not to overeat during the next meal, if the previous one was skipped.

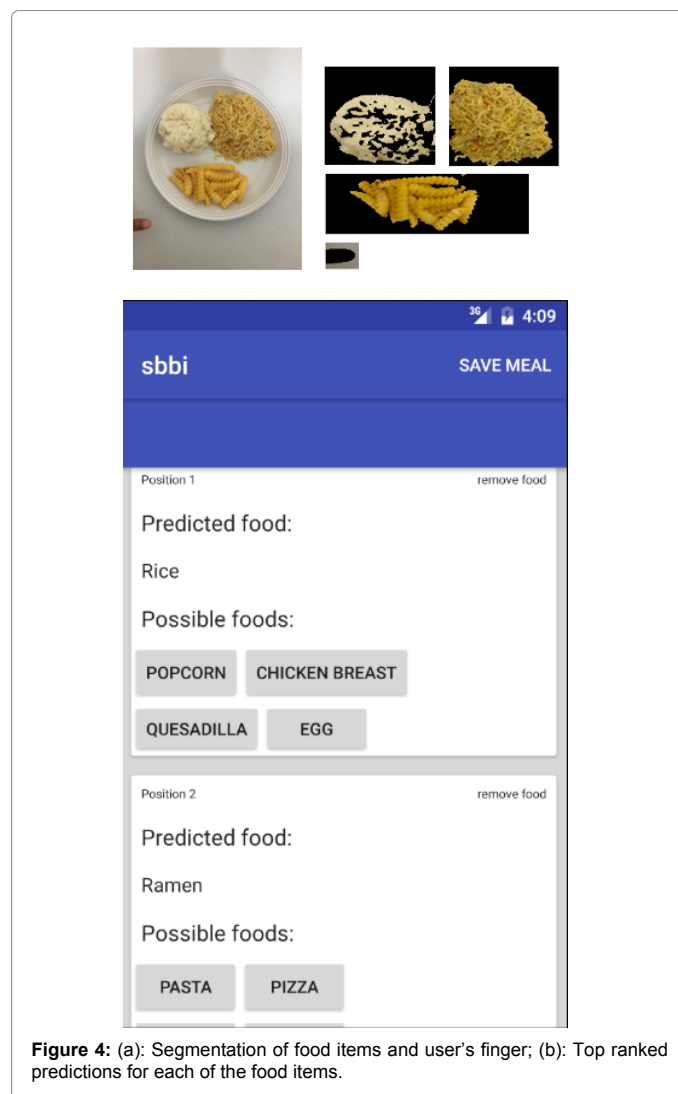
### Results

This section contains the results derived from each of aforementioned analysis.

#### Image segmentation and classification

After receiving the uploaded meal images, the system performs segmentation on the top image to extract each food items, along with user's finger (Figure 4a). We did the same segmentation for every image in our database before training the classification system.

After each item has been identified from the image, it is subject to aforementioned classifiers with the compiled image features. We first evaluated the performance of our modified LBP approach. In this case, 5 vectors were extracted from different radiuses (from 1 to 5) instead of one vector (as described in methods). As shown in Table 1, the new



LBP feature outperforms existing one on food image classification when we compared to other three studies [42]. Therefore, we included modified LBP and performed feature selection through combining with other feature groups. As a result, different feature sets have led to somehow different performance, as shown in Table 2. Our SVM model (using color, HOG, modified LBP, Gabor, and SURF) outperforms the standard one (51.1% versus 43.0%) (Table 1), when validated on Food-101 dataset.

More importantly, when trained based on the extended dataset where the Food-101 dataset was artificially extended by applying random distortions in the training images, such as cropping images, distorting brightness, contrast, saturation and hue, instead of including more food class, our model can achieve higher accuracy of 65.5% (with 59.0% sensitivity and 72.0% specificity). Furthermore, our approach using Deep Learning strategy shows much more improved performance than the traditional handcrafted approach, achieving overall performance as 87.2% (with 90.0% sensitivity and 84.4% specificity) on the expanded dataset. The main reason behind is that Deep Learning strategies learn relevant features automatically through convolutional layers, compared to the pre-defined features that might not be effective enough to distinguish images. When compared to similar models presented in [43], our model shows comparable result on an expanded database. After classification, our system returned the identified items for each food (Figure 4b).

### Food weight estimation and nutrient analysis

After correctly classify each food item, volume and weight estimation was performed (as described in Methods). Here we used an example of a one-day diet (including breakfast, lunch and dinner, in Table 3 to showcase the analysis. With identification of what is on the plate and the corresponding weight, our system outputs the total nutrient and calorie intake for each logged meal, calculated based on the aforementioned nutrient database. Table 3 shows the identified items, estimated weight and calories, compared to the ground truth, which is within  $\pm 5\%$  and  $\pm 8\%$  variation, respectively. Such information will be displayed to the user and serves as input to the metabolic analysis.

### Metabolic analysis

Through metabolic network modeling, ATP production was calculated to help understanding the body response in presence of different nutrients (or metabolites) derived from the food. COPASI [44,45] was first used to derive the ODEs of the model and then MATLAB was used for simulation. Figure 5 illustrated the body's response to starvation and to the intake of three meals by the user. In this example, the user has consumed 95 grams of banana, an apple weighted 140 grams, and 50 grams of cookies as the breakfast at 8am, 180 grams of cooked rice, 150 grams of ramen, and 60 grams of French fries for the lunch at 12pm, and 240 grams of cooked rice, 55 grams of chicken breast, and 100 grams of French fries at 6pm as dinner.

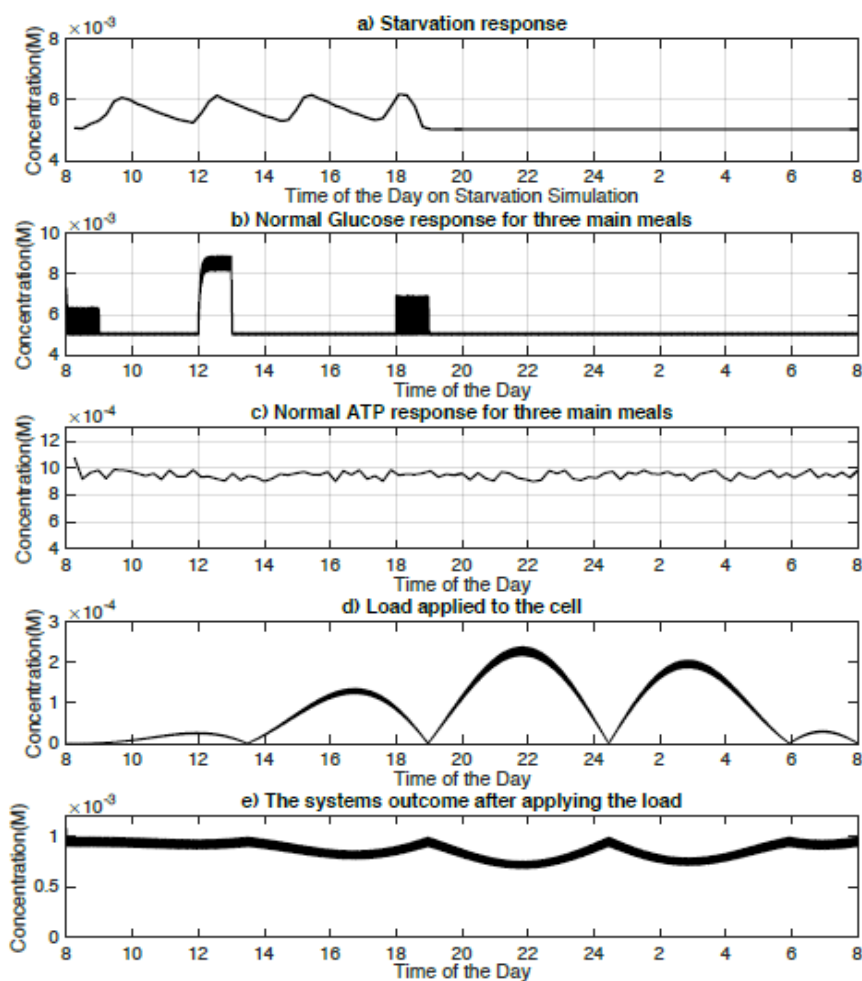
Study	Method	Database	Classification Accuracy
Extract LBP Feature [47]	Standard LBP+ SVM	Chen et al., 2012	32.80%
Chen, et al. [19]			45.90%
Bejjbom, et al. [11]			43.60%
Our implementation	Modified LBP+SVM		48.00%

Table 1: Performance comparison on food image classification using different local binary pattern approaches.

Study	Approach		Database	Classification accuracy
	Features	Methods		
Our study	Color, HOG, traditional LBP	SVM	Food-101	43.00%
	Color, HOG and modified LBP	SVM		47.10%
	Color, HOG, modified LBP, Gabor and SURF	SVM		51.1%*
	SURF and L*a*b color values	RFDC		50.80%
Our study	Color, HOG, modified LBP and Gabor	SVM	Expanded database	65.50%
<b>Deep Learning methods</b>				
Study	Architecture		Database	Classification accuracy
Bossard et al., 2014 [14]	Provided by Caffe library		Food-101	56.40%
Our study	Inception v3			69.40%
Hassannejad et al., 2016 [44]				88.30%
Our study	Inception v3		Expanded database	87.20%

(\*The indicated classification accuracy is evaluated based on the testing set (Methods) and the 10-fold cross validation using the selected features is 47.2%)

Table 2: The classification performance on food-101 and our food database.



**Figure 5:** (a) Glucose concentration in starvation state, (b) Glucose concentration for a normal diet of three main meals within a day, (c) ATP concentration of the normal diet, (d) the load applied to mimic the activity (the ATP consumption) captured by exercise tracker and (e) the ATP concentration for one day after applying the given load.

Here, we set 100 seconds as interval in a period of 1 hour for each small portion of the meal to be [44] absorbed.

After running the simulations, the ATP and glucose concentrations are shown starting from 8am within a day. In Figure 5a, the glucose concentration is shown for the entire day in the starvation state. Once starvation starts, the free glucose is consumed at first; then, glycogen starts to be consumed for the next 8 to 10 hours to maintain the glucose level in normal condition. If glycogen is completely burned, the third phase will start by consuming the fat resources. Figure 5b and 5c show the trend of glucose and ATP concentration with three different meals during a day, respectively. In Figure 5d, an arbitrary periodic load was applied to the system to mimic the ATP consumption captured by the exercise tracker. At the end in Figure 5e, the overall ATP balance after applying the load has been demonstrated where the ATP concentration goes to the same level at 8 am of the previous day. This information is helpful to alert the user to adjust his/her diet and exercise to manipulate the energy balance. In addition, using the similar metabolic system, a proof of concept study on glucose monitoring and a new standalone tool developed by our group was presented in a recent work [45].

### Intervention module

Based on the nutritional information obtained from the three meals above, it is possible to verify if user has consumed more calories than needed and if all food items are healthy. Nutritional information from the meals is presented on Table 3.

Data about energy intake estimated from previous step, along with energy consumption from the activity tracer, was further processed by the intervention module. For example, considering that a user burned 1,700 calories throughout the day but took 2,154 calories from the diet, characterizing an overeating episode. In this case, the intervention module identifies that fries are unhealthy according to local database and suggests up to three healthy food items that are absent in any of the three meals, e.g., to substitute fries for salad, tomato or fish. Meanwhile, this module also identifies that ramen and rice contribute to the high calorie within the day, being responsible for 30.5% and 25.3% for the total calories intake. However, they are not unhealthy as fries; therefore, user will receive a suggestion to reduce quantities of these food. For example, user will be informed that by reducing rice to 92 grams or ramen to 204

Meal	Food item	Weight (in grams)		Calorie (in kcal)		Protein (in grams)		Lipid (in grams)		Carbohydrate (in grams)		Fiber (in grams)		Sugar (in grams)	
		est	grt	est	grt	est	grt	est	grt	est	grt	est	grt	est	grt
Breakfast	Banana	89	95	79	85	0.97	1.04	0.29	0.3	20	21.7	2.3	2.5	11	12
	Apple	134	140	70	73	0.34	0.36	0.23	0.2	19	19.3	3.25	3.4	14	15
	Cookies	52	50	255	245	2.96	2.85	11.44	11	35	33.8	0.83	0.8	17	16
Lunch	Rice	190	180	247	234	4.52	4.28	0.4	0.4	54	51.5	0.53	0.5	0	0
	Ramen	130	150	572	660	13.22	15.3	22.87	26	78	90.4	3.8	4.3	2.6	3
	French fries	46	60	133	174	1.52	1.99	6.05	7.9	18	23.7	1.76	2.3	0	0
Dinner	Rice	246	240	320	312	5.85	5.71	0.51	0.5	70	68.6	0.71	0.7	0	0
	Chicken breast	60	55	90.5	83	18.32	16.8	1.89	1.7	0	0	0	0	0	0
	French fries	80	100	232	290	2.65	3.31	10.53	13	32	39.5	3.12	3.9	0	0
TOTAL		1027	1070	1998	2156	50.35	51.6	54.21	62	327	349	16.3	18	44	46

**Table 3:** Illustration of the estimated weight (est.) versus the ground truth (grt.) based on the prediction of three sample meals.

grams, he/she will obtain a healthy balance between calories intake and burned.

## Conclusion

This paper presents a proof-of-concept study of an image based classification system using smartphone applications specifically designed for automated food recognition and dietary intervention. Particularly, the entire framework can be broken down into four major parts that involve new strategies for comprehensive food image databases, classifiers capable for food item recognition, food volume estimation, and nutrient analysis that provide information for diet intervention. In addition, an energy-related metabolic model was implemented including all the chemical reactions participating in the main ATP-producing metabolic pathways. The results of a meaningful trend for ATP concentration in presence of regular meals or a random load have demonstrated the feasibility of the given model. Worth mentioning is the increasingly growing application of Deep Learning methods in image-based food recognition, which outperformed traditional approaches using handcrafted features.

## Future Work

Even though improved performance has been demonstrated, challenging issues still remain and novel algorithms and techniques in image segmentation, classification, and food weight estimation are highly desired. In addition, enormous food diversity around the world has posed extreme challenges for building a versatile food classification system; therefore we will continue efforts in scaling up the food image databases. It would be interesting to study if new 3D cameras embedded in devices can help in segmentation and food volume estimation. Furthermore, by adding more related signaling pathways into the metabolic model, e.g., insulin signaling, the ATP production and glucose level can be estimated with higher accuracy. In parallel, the metabolic model will be transferred into a more flexible environment like MATLAB, through which we will build a new standalone package that enable an ease and reliable application for the similar research. Additionally, we believe the increased application of wearable sensor devices, especially those can be integrated into smartphone, will revolutionize this line of research and as a whole the food monitoring system will be useful for effective health promotion and disease prevention. For example, eating episodes detected by several wearable devices, such as glasses with load cells [46], glasses connected to sensors on temporalis muscle and accelerometer [47], and wrist motion track [48], can provide more food intake information in addition to the image-based strategy. We believe that such information collected

by multi-monitoring technologies [49], pertinent to users' diet habit pattern, can serve as starting point for more precise food consumption analysis and diet interventions.

## Acknowledgment

The authors would like to thank all SBBI members who have been involved in this work for providing helpful discussions and technical assistance. No conflict of interest declared.

This work was supported by the National Institutes of Health funded COBRE grant [1P20GM104320], UNL Food for Health seed grant and the Tobacco Settlement Fund as part of Cui's start up grant support.

## References

- Allen JK, Stephens J, Dennison CR, Stewart KJ, Hauck S (2013) Randomized controlled pilot study testing use of smartphone technology for obesity treatment. J Obes vol. 2013: 1-7.
- Champagne CM, Bray GA, Kurtz AA, Monteiro JB, Tucker E (2002) Energy Intake and Energy Expenditure: A Controlled Study Comparing Dietitians and Non-dietitians. J Am Diet Assoc 102: 1428-1432.
- Carter MC, Burley VJ, Nykjaer C, Cade JE (2013) Adherence to a Smartphone Application for Weight Loss Compared to Website and Paper Diary: Pilot Randomized Controlled Trial. J Med Internet Res 15: 32.
- Boushey CJ, Harray AJ, Kerr DA, Schap TE, Paterson S (2015) How willing are adolescents to record their dietary intake? The mobile food record. JMIR MHealth UHealth 3: e47.
- Kerr DA, Harray AJ, Pollard CM, Dhaliwal SS, Delp EJ, et al. (2016) The connecting health and technology study: a 6-month randomized controlled trial to improve nutrition behaviours using a mobile food record and text messaging support in young adults. Int J Behav Nutr Phys Act 13: 52.
- Cordeiro F, Epstein DA, Thomaz E, Bales E, Jagannathan AK, et al. Barriers and Negative Nudges: Exploring Challenges in Food Journaling. Proc SIGCHI Conf Hum Factor Comput Syst 2015: 1159-1162.
- Ciocca G, Napoletano P, Schettini R (2016) Food recognition: a new dataset, experiments and results IEEE J Biomed Health Inform 99: 1-1.
- Matsuda Y, Hoashi H, Yanai K (2012) Recognition of Multiple-Food Images by Detecting Candidate Regions. IEEE Int Conf Multimed Expo 2012: 25-30.
- Meyers A (2015) Im2Calories: Towards an Automated Mobile Vision Food Diary. IEEE Int Conf Compu Vis 2015: 1233-1241.
- Pouladzadeh P, Shirmohammadi S, Yassine A (2014) Using graph cut segmentation for food calorie measurement. IEEE Int Sympos Med Meas App 2014: 1-6.
- Beijbom O, Joshi N, Morris D, Saponas S, Khullar S (2015) Menu- Match: Restaurant-Specific Food Logging from Images. IEEE 2015: 844-851.
- Chen M, Dhingra K, Wu W, Yang L, Sukthankar R, et al. (2009) PFID: Pittsburgh fast-food image dataset. IEEE Int Conf Imag Process 2009: 289-292.
- Chen MY (2012) Automatic Chinese Food Identification and Quantity Estimation. SIGGRAPH 29: 1-29.

14. Bossard L, Guillaumin M, Gool LV (2014) Food-101-Mining Discriminative Components with Random Forests. *ECCV 2014*: 446-461.
15. Kawano Y, Yanai K (2014) Automatic Expansion of a Food Image Dataset Leveraging Existing Categories with Domain Adaptation. *ECCV 2014*: 3-17.
16. Kawano Y, Yanai K (2013) Real-Time Mobile Food Recognition System. *IEEE Conf Compu Vis Pat Recog Workshop 2013*: 1-7.
17. Yang S, Chen M, Pomerleau D, Sukthankar R (2010) Food recognition using statistics of pairwise local features. *IEEE Comp Soc Conf Compu Vis Patt Recog 2010*: 2249-2256.
18. He Y, Xu C, Khanna N, Boushey CJ, Delp EJ (2014) Analysis of food images: Features and classification. *Proc Int Conf Image Proc 2014*: 2744-2748.
19. Chen J, Ngo C (2016) Deep-based Ingredient Recognition for Cooking Recipe Retrieval. *ACM 2016*: 32-41.
20. Yanai K, Kawano Y (2015) Food image recognition using deep convolutional network with pre-training and fine-tuning. *IEEE Int Conf Multim Expo Worksh 2015*: 1-6.
21. Noronha J, Hysen E, Zhang H, Gajos KZ (2011) Platemate: Crowdsourcing Nutritional Analysis from Food Photographs. *ACM 2011*: 1-12.
22. Zhu F, Bosch M, Khanna N, Boushey CJ, Delp EJ (2015) Multiple Hypotheses Image Segmentation and Classification With Application to Dietary Assessment. *IEEE J Biomed Health Inform* 19: 377-388.
23. Faiz I, Mukhtar H, Khan S (2014) An integrated approach of diet and exercise recommendations for diabetes patients. *IEEE Int Conf e-Health Net App Serv 2014*: 537-542.
24. Lee CS, Wang MH, Li HC, Chen WH (2008) Intelligent ontological agent for diabetic food recommendation. *IEEE Int Conf Fuzz Syst 2008*: 1803-1810.
25. Otsu N (1979) A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern* 9: 62-66.
26. Mignotte M (2008) Segmentation by Fusion of Histogram-Based  $K$ -Means Clusters in Different Color Spaces. *IEEE Trans Image Process* 17: 780-787.
27. Timo O, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE 2002*: 971-987.
28. Bay H, Tuytelaars T, Gool LV (2006) SURF: Speeded Up Robust Features. *ECCV 2006*: 404-417.
29. Mehrotra R, Namuduri KR, Ranganathan N (1992) Gabor filter-based edge detection. *Pattern Recognit* 25: 1479-1494.
30. Anthimopoulos MM, Gianola L, Scarnato L, Diem P, Mougiakakou SG (2014) A Food Recognition System for Diabetic Patients Based on an Optimized Bag-of-Features Model. *IEEE J Biomed Health Inform* 18: 1261-1271.
31. Wan V, Campbell WM (2000) Support vector machines for speaker verification and identification. *IEEE 2*: 775-784.
32. Tensor F (2016) An open source machine learning framework for everyone.
33. Ogata H (2016) USDA Food Composition Databases.
34. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, et al. (1999). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 27: 29-34.
35. Wishart S (2007) HMDB: The human metabolome database. *Nucleic Acids Res* 35: 1.
36. BD Bennett, EH Kimball, Gao M, Osterhout R, Van Dien SJ (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat Chem Biol* 5: 593-599.
37. Fehlbeg E (1969) Low-order classical Runge-Kutta formulas with stepsize control and their application to some heat transfer problems. *NASA Tech Rep 1969*: 315.
38. Ghiasi RM (2017) A simulation model of glucose-insulin metabolism and implementation on OSG. *IEEE 2017*: 1-6.
39. Lobstein T, Davies S (2009) Defining and labelling 'healthy' and 'unhealthy' food. *Public Health Nutr* 12: 331-340.
40. Thorndike AN, Riis J, Sonnenberg LM, Levy DE (2014) Traffic-Light Labels and Choice Architecture: Promoting Healthy Food Choices. *Am J Prev Med* 46: 143-149.
41. Larrivee S, Greenway FL, Johnson WD (2015) A Statistical Analysis of a Traffic-Light Food Rating System to Promote Healthy Nutrition and Body Weight. *J Diabetes Sci Technol* 9: 1336-1341.
42. <https://www.mathworks.com/help/vision/ref/extractlbpfeatures.html>
43. Hassannejad H, Matrella G, Ciampolini P, Munari ID, Mordonini M, et al. (2016) Food Image Recognition Using Very Deep Convolutional Networks. *Multim Ass Diet Manag* 2016: 41-49.
44. Hoops S (2006) COPASI-A Complex Pathway Simulator. *Bioinformatics* 22: 3067-3074.
45. Silva BVR, Cui J (2017) A Survey on Automated Food Monitoring and Dietary Management Systems. *J Health Med Informat* 8: 272.
46. Chung J, Chung J, Oh W, Yoo Y, Lee WG, et al. (2017) A glasses-type wearable device for monitoring the patterns of food intake and facial activity. *Sci Rep* 7: 41690.
47. Farooq M, Sazonov E (2016) A Novel Wearable Device for Food Intake and Physical Activity Recognition. *Sensors (Basel)* 16: 7.
48. Dong Y, Hoover A, Scisco J, Muth E (2012) A new method for measuring meal intake in humans via automated wrist motion tracking. *Appl Psychophysiol Biofeedback* 37: 205-215.
49. Oliver A (2011) Ambient, on-body, and implantable monitoring technologies to assess dietary behavior. *Food Nut* 2011: 3507-3526.