FooDD: Food Detection Dataset for Calorie Measurement Using Food Images

Parisa Pouladzadeh[1(✉)], Abdulsalam Yassine[1], and Shervin Shirmohammadi[1,2]

[1] Distributed and Collaborative Virtual Environments Research Laboratory, University of Ottawa, Ottawa, Canada
{ppouladzadeh,ayassine,shervin}@discover.uottawa.ca
[2] Colleges of Engineering and Natural Sciences, Istanbul Şehir University, Istanbul, Turkey

Abstract. Food detection, classification, and analysis have been the topic of in-depth studies for a variety of applications related to eating habits and dietary assessment. For the specific topic of calorie measurement of food portions with single and mixed food items, the research community needs a dataset of images for testing and training. In this paper we introduce FooDD: a Food Detection Dataset of 3000 images that offer variety of food photos taken from different cameras with different illuminations. We also provide examples of food detection using graph cut segmentation and deep learning algorithms.

Keywords: Food image dataset · Calorie measurement · Food detection

1 Introduction

Food images, taken by people using their smartphones, are used in many proposed systems for food recognition, detection, and classification. Detection of food ingredients from their image is a key process in calorie measurement systems used for treatment of chronical illness such as diabetes, blood pressure, obesity, etc. However, for the specific topic of calorie measurement of food portions with single and mixed food items, the research community is lacking a public and free dataset of images for testing and training, making comparison across different food recognition methods more challenging. For this purpose, in this paper we introduce a dataset of 3000 images, offering a variety of food poses taken from different cameras with different illuminations. Acquisition of an accurate dataset will, at the end, support the realization of effective treatment programs for patients. In our previous work [1][2][3], we proposed a system using Vision-Based Measurement (VBM) [6] to improve the accuracy of food intake reporting. Our system runs on smartphones and allows the user to take a picture of the food and measure the calorie intake automatically. This paper presents the food image dataset. Furthermore, we provide examples of food detection using graph cut segmentation [4] and deep learning algorithms [5]. Our dataset can aid further research on different types of food recognition and learning algorithms. The rest of the paper is organized as follows: In section 2 we evaluate existing food datasets and explain the novelty and contribution of our dataset. In section 3, we briefly explain our food

V. Murino et al. (Eds.): ICIAP 2015 Workshops, LNCS 9281, pp. 441–448, 2015.
DOI: 10.1007/978-3-319-23222-5_54

recognition system, while in section 4 we present the dataset collection methodology. Section 5 describes our proposed system. Finally, in section 6, we conclude the paper.

2 Related Work

In [12], the authors have collected 101 different fast food images such as burgers, pizza, salads, etc. But the collection has only fast food images. The dataset in [13] introduces 101 food categories, with 101,000 images, which are mostly mixed food. They sampled 750 training images automatically. Additionally, 250 test images were collected for each class, and were manually cleaned. However, the training images were not cleaned, and thus contain some amount of noise. In [14], the proposed framework consists of images from the Web with the category of names. The noise is filtered out using a "foodness" classifier and an adaptive SVM. Also in [15], the authors investigate features and their combinations for food image analysis and a classification approach based on k-nearest neighbors and vocabulary trees. The system is evaluated on a food image dataset consisting of 1453 images in 42 food categories acquired by 45 participants in natural eating conditions. In [16], the authors collected the dataset from different restaurant food images which contains 61 different categories of food items. Also in [17] and [18], the authors collected a dataset of different food images with different smart phones. In our dataset, we collected around 3000 different images with different categorization as we are going to discuss in next sections. We considered both single and mixed food portions. By so doing, our system is trained to achieve higher accuracy in mixed and non-mixed food. Also, by having the user's finger in the image, we can easily calculate the size of each food in order to have better calorie estimation. Compared to other food datasets, our dataset has a number of advantages and is in fact the only dataset that provides all of the following features:

1- We include both single and mixed food portions. By so doing, the food recognition system trained with our dataset can achieve higher accuracy in mixed and non-mixed food.
2- We use multiple brands of cameras to capture the image of the same food item, providing an opportunity for the food recognition system to become more robust with changes in camera brands.
3- We provide multiple lighting conditions, again for the same food item, allowing more research in developing systems that are more resilient to lighting changes.
4- We provide multiple shooting angles for the same food item, allowing the development of more accurate food recognition methods.
5- We provide a calibration reference (user's thumb) in our images, allowing more accurate measurement of the size of food ingredients, leading to higher accuracy in calories measurement.

In our dataset we have collected around 3000 different images with different categorization as we are going to discuss in next sections.

3 Food Recognition

In this section, we briefly explain how our food recognition system works. For a complete description with details about specific image processing and machine learning techniques, design choices, experiments, and the required accuracy, we refer the readers to [1] [2], and [3]. To measure food calorie, we use a mobile device with camera that supports wireless connection, such as any of today's smartphones. The system will enable the mobile device to take pictures of the food for analysis and immediate response to the user. The overall system design is shown in Figure 1.

Fig. 1. Overall System Design

First, in order to have accurate results for our segmentation, a simple transformation must be performed on the image to change the image size into standard format. To do so, the size of each image will be compared with standard size categorizes. If the image size is not compatible with any size category, some cropping or padding techniques will be applied to the image. We have defined one size category, i.e. 970×720 for simplicity. Larger images will be adjusted to this size, before performing any image processing technique. In next step, at the segmentation step, each image is analyzed to extract various segments of the food portion. We paid significant attention to the segmentation mechanism design to ensure that images are processed appropriately. Particularly, we have used color segmentation, k-mean clustering, and texture segmentation tools. Furthermore, in our classification and food recognition analysis, we have nominated Cloud SVM and deep neural network method to increase the accuracy of the recognition system. Finally we have measure the calorie of the food.

4 Dataset Collection

For the collection of the food images in our dataset, we divided the food images into two different collections; single food portions and mixed food portions. We took into consideration important factors that affect the accuracy of our results. Specifically, we used a variety of the following components: Camera, Lighting, Shooting Angle, White Plate, Thumb, Single and Mixed Food. Each component will be described in details in the following subsections.

4.1 Camera

The camera itself will have an effect on the results in terms of its lens, hardware, and software. As such, we used three different cameras for our experiments, consisting of Canon SD1400 (14.1-megapixel resolution, 2.7-inch Pure Color System LCD, 28mm wide-angle lens; 4x optical zoom and Optical Image Stabilizer) iPhone 4(5 megapixel resolution, LED flash, VGA-quality photos and video at up to 30 frames per second with the front camera), and Samsung S4 cameras (13 megapixel resolution, 41, auto-focus, LED flash, Dual shot).

4.2 Lighting

Lighting and illumination is one of the important parameters which affect the system outcome because illumination directly affects the image segmentation algorithm, which in turn affects the rest of the algorithms. To take this into account, we put the same plate in three different locations with different illuminations (sunlight) and we took pictures.

4.3 Shooting Angle

Another effective parameter is the angle of photography; we have chosen three different angles which are approximately 30, 90, and 150 degrees from the plate of food for all pictures. This means that for each plate in 3 different lighting locations we have also taken 3 pictures from different angles.

4.4 White Plate

For all images we have considered a white plate to ignore the background of the images. By using white plate, food segmentation and food recognition will be easier to perform.

4.5 Thumb

The thumb of the user and its placement on the plate are also shown in Figure 2. There is a one-time calibration process for the thumb, which is used as a size reference to measure the real-life size of food portions in the picture [1]. An example of food picture capturing and thumb isolation and measurement are shown in Figure 1. Compared to the calibration methods of similar systems, using the thumb is more flexible, controllable, and reliable. For users with thumb disability or amputated thumbs, another finger or a coin can be used instead, the latter still more ubiquitous than special plates or cards used in other systems.

(a) (b) (c)

Fig. 2. (a, b) Test images with thumb (c) Calculation of the thumb dimensions [1]

4.6 Single and Mixed Food

We have divided our food into two different groups: single food and fruits such as apple, orange, and bread, and mixed food which includes different food portions in a plate of food such as salad, pizza, and kebab with rice.

4.7 Food Item Types

The dataset contains images taken with different cameras, illuminations, and angles. Having a wide variety of food and fruits gives a better and more reliable dataset in order to increase the accuracy of calorie food measurement systems. The name and number of single food images which are included in the dataset are shown in Figure 3. In the dataset, the images are divided into 6 categories considering the capturing device, background, and lighting condition. For example, images in category 1 are captured with a Samsung camera, within a light environment with a white background, and from different shooting angles.

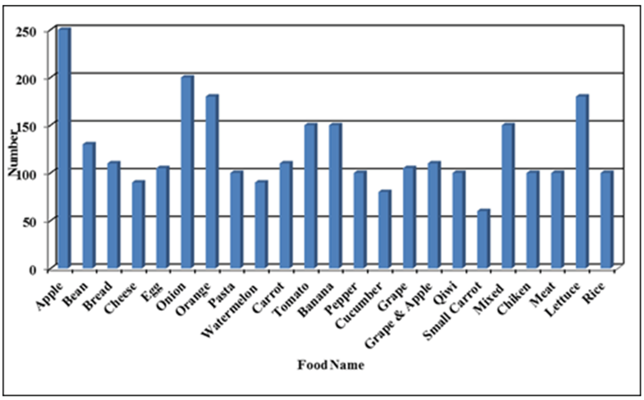

Fig. 3. Food Items in the Dataset

The categories in the dataset are shown in Table I. Each category contains more than 100 images, including various food items from Figure 3. Figure 4 shows sample images takes by the user. Note that the thumb is used as a calibration means to determine the size of the food items in the image [1].

Table 1. I Different Food Categories

Category	Camera	Lighting
1	Samsung-S4	Light Environment
2	Samsung-S4	Dark Environment
3	IOS-4	Light Environment
4	IOS-4	Dark Environment
5	CanonSD1400	Light Environment
6	CanonSD1400	Dark Environment

Fig. 4. Sample food images taken by the user

5 Experimental Results

This section presents the experimental results of our system using the food images dataset. In this work, we have combined Graph cut segmentation and deep neural network. The dataset is used in the learning process of these two methods, which allow us to improve the accuracy of our food classification and recognition significant compared to our previous work [1][3]. By recognizing the food portions and also by having the size and shape of the food portions from graph cut algorithm, we can calculate the calorie of the whole food portions.

In this experiment, our dataset comprises of 30 different categories of food and fruits. These food and fruit images are divided into training and testing sets, where around 50% of the images from each group are used to train the system and the remaining images serve as the testing set.

The results are shown in Figure 5. We can see that graph cut segmentation outperforms normal segmentation but is outperformed by Deep Neural Network algorithm which has 100% accuracy in our dataset. In addition, our system recognized food portions very accurately in about 3 seconds, on average.

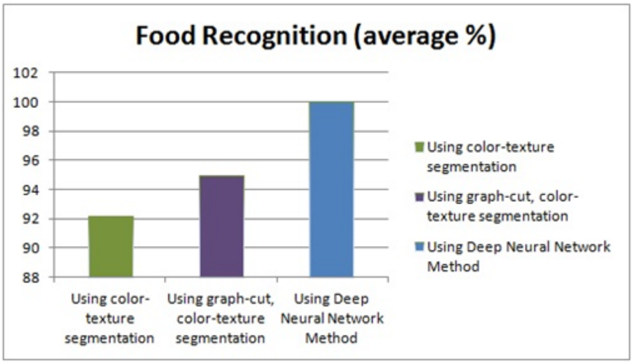

Fig. 5. Food Recognition Accuracy for Single Food

6 Dataset Availability and Format

The dataset is available for free as long as it is used for research purposes. Instructions on how to obtain it are posted at our website[1].

7 Conclusion

Food detection and classification is a problem gaining much importance in health-related applications. Since algorithms to accomplish this task are currently being developed and refined, diverse and complementary datasets for evaluation are not only helpful but also necessary to aid research. In this paper, we provided a dataset comprising 3000 food images. In this dataset, we placed careful attention in generating the food images characteristics pertained to camera type, shooting angle, and illumination variations. A strong feature of our dataset is the good distribution of single and mixed food images. Thus, it can be used to facilitate testing and benchmarking of various food detection algorithms. We provided experimental results using color-texture segmentation, graph cut segmentation, and deep neural network algorithms on this dataset, and invite researchers to devise suitable benchmarks and share with the research community.

References

1. Pouladzadeh, P., Shirmohammadi, S., Almaghrabi, R.: Measuring Calorie and Nutrition from Food Image. IEEE Transactions on Instrumentation & Measurement **63**(8), 1947–1956 (2014)
2. Pouladzadeh, P., Shirmohammadi, S., Bakirov, A., Bulut, A., Yassine, A.: Cloud-Based SVM for Food Categorization. Multimedia Tools and Applications, p. 18. Springer, June 3, 2014. doi: 10.1007/s11042-014-2116-x

3. Pouladzadeh, P., Shirmohammadi, Yassine, A.: Using graph cut segmentation for food calorie measurement. In: IEEE International Symposium on Medical Measurements and Applications, pp. 1–6, June 2014
4. Yuri, Y.B., Lea, G.F.: Graph Cuts and Efficient N-D Image Segmentation. International Journal of Computer Vision **70**(2), 109–131 (2006)
5. Krizhevsky, A., Sutskever, I., and Hinton, G.: ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems (NIPS) (2012)
6. Shirmohammadi, S., Ferrero, A.: Camera as the Instrument: The Rising Trend of Vision Based Measurement. IEEE Instrumentation and Measurement Magazine **17**(3), 41–47 (2014)
7. Sun, M., et al.: Determination of food portion size by image processing. Engineering in Medicine and Biology Society, 871–874, August 2008
8. Schölkopf, B., Smola, A., Williamson, R., Bartlett, P.L.: New support vector algorithms. Neural Computation **12**(5), 1207–1245 (2000)
9. Burke, L.E., et al.: Self-monitoring dietary intake: current andfuture practices. Journal of renal nutrition the official journal of the Council on Renal Nutrition of the National Kidney Foundation **15**(3), 281–290 (2005)
10. Beasley, J.: The pros and cons of using pdas for dietary self-monitoring. J. Am. Diet Assoc. **107**(5), 739 (2007)
11. Gao, C., Kong, F., Tan, J.: Healthaware: tackling obesity with health aware smart phone systems. In: IEEE International Conference on Robotics and Biometics, pp. 1549–1554 (2009)
12. Chen, M., Dhingra, K., Wu, W., Yang, L., Sukthankar, R., Yang, J.: PFID: pittsburgh fast-food image dataset. In: International Conference on Image Processing, pp. 289–292 (2009)
13. Bossard, L., Guillaumin, M., Van Gool, L.: Food-101 – mining discriminative components with random forests. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014, Part VI. LNCS, vol. 8694, pp. 446–461. Springer, Heidelberg (2014)
14. Kawano, Y., Yanai, K.: Automatic expansion of a food image dataset leveraging existing categories with domain adaptation. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) ECCV 2014 Workshops. LNCS, vol. 8927, pp. 3–17. Springer, Heidelberg (2015)
15. He, Y., Xu, C., Khanna, N, Boushey, C.J., Delp, E.J.: Analysis of food images: features and classification. In: IEEE International Conference on Image Processing (ICIP), pp. 2744–2748 (2014)
16. Kong, F., Tan, J.: DietCam: regular shape food recognition with a camera phone. In: International Conference on Body Sensor Networks (BSN), pp. 127–132 (2011)
17. Farinella, G.M., Allegra, D., Stanco, F.: A benchmark dataset to study the representation of food images. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) ECCV 2014 Workshops. LNCS, vol. 8927, pp. 584–599. Springer, Heidelberg (2015)
18. Kawano, Y., Yanai, K.: FoodCam: A real-time food recognition system on a smartphone. Multimedia Tools and Applications. Springer, April 12, 2014