Fruit Calories Estimation Using 3D Measurements

Pyone Thant Win*
pwin17@earlham.edu
Earlham College
Richmond, Indiana

ABSTRACT

Calories estimation of food from images is a well-researched field. Many image-based food calorie estimation systems use a reference object to measure the volume of food and the calories of food. However, it is inconvenient for users as the system requires them to have the reference object next to the food every time a user wants to measure food calories. This project will solve this problem by using a virtual reference object. In this project, calories of fruits are measured using two input images from the user. Volume and calories estimation will be made in an entirely image-based method. The fruit calorie estimation system will be developed as an Android application as this will be the most convenient way for users to apply this method.

KEYWORDS

Volume Estimation, Calories Estimation, Android Development

ACM Reference Format:
Pyone Thant Win. 2020. Fruit Calories Estimation Using 3D Measurements. In ,. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Food is essential for human, and it is important that people consume a healthy number of calories from nutritious foods. Calories management softwares and applications are used by wide variety of people. According to Statistica, there are 27.4 million unique users of fitbit, the health and fitness app [2]. Volume estimation and calories estimation are fairly studied areas of research. The state-of-art image-based food calories estimation system requires users to take a picture of the food with reference object. This can be inconvenient for users as they need to have a physical reference object with them at all time they want to measure food volume or calories. This project serves as the very first step to an attempt to create a new volume and calories estimation system with solely image-based calculations. The physical reference object is replaced with an image ruler as a reference for users to compare. Image recognition and image processing techniques are not used in this project, but they are included in the future works. In order to convert from volume to mass for calories estimation a fruit density is required. However, there is no open-source database of fruits' densities. Therefore, a fruit density database is contributed in this study. This project will aim for a more user-friendly and accurate system for calorie estimation. The final product of this project is an android application. All calculations in this project are based on the two input images from the user. As a beginning step of this project, I will be focusing on fruits only.

*CS488 Senior research

Win '20, Earlham College Computer Science Senior Capstone, January–May, 2020, Richmond, IN

ACM ISBN XXX-X-XXX-XXXX-X/XX/XX…$0.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

2 RELATED WORK

2.1 Food Detection and Recognition Techniques

Although image recognition is not integrated in this study, it is important to notice the works from past studies since this project is only the first step to the bigger project and image recognition will be applied in the future projects. Different technologies for food recognition are used in the past studies. The following are the most commonly used methods.

2.1.1 Analysis of color and texture features. Detecting and recognizing fruits from their color and texture is one approach of image recognition. Arivazhagan et al. used HSV (hue, saturation and value) model for fruit recognition [12]. In their research, the input images in the RGB format are converted to the HSV format. The hue is used to detect the object through orientation, lighting and camera. The saturation is used to examine the texture of the object and the value is used to retrieve color features of the object. With this approach, the freshness of a fruit can be determined. This approach is applicable to various agricultural uses. They have achieved an accuracy rate of 92.6% [13].

2.1.2 Linear kernel SVM. Another method of ingredient recognition is the use of linear kernel SVM. It was commonly used in studies related to food ingredient recognition that were conducted before 2012. Maruyama et al. implemented a mobile cooking recipe recommendation system using linear SVM with one-vs-rest strategy [9]. The researchers integrated bag-of-features with SURF and linear SVM. They chose linear kernel because it was less expensive compared to non-linear kernel in spite of the high classification performance of non-linear kernel. The linear kernel is more practical because it has faster run-time which is suitable for real-time recipe recommendation. The result of this research revealed that the classification rate is 83.93% for 30 kinds of food.

Kawano and Yanai also applied color histogram, SURF bag-of-features and linear SVM and fast x¬² kernel to their study [7]. Users draw square boxes around the dishes which were cut and extracted using GrubCut. Then, the cut image is inserted as input to the linear SVM model. The linear SVM model implemented for this study has

50 linear SVM classifiers one for each category. The researchers added Efficient Sub-window Search (ESS) as an input of SVM model to the make the linear SVM model run more efficiently with O(1) run time. The total running time for recognition and direction estimation is 0.34 seconds and about 3 bounding boxes of dishes can be processed at the same time with 1 second run time in total.

2.1.3 Convolutional Neural Network (CNN). CNN is the state-of-art algorithm of deep learning in analyzing images. Kayaga et al. conducted a research on food detection and recognition using CNN and compared its performance and SVM model's performance [6]. The researchers contributed on improving CNN's hyper parameters. It is found that the accuracy rate of food recognition using CNN is better than the one that uses SVM. Chen and Ngo implemented a program that retrieves cooking recipe from a picture of the dish [1]. In their experiment, they developed a multi-task learning model using deep convolutional neural network (DCNN) to solve zero-shot retrieval and content-based retrieval. The reason is that the program needs to detect and recognize ingredients of a dish and find the recipe at the same time. As a conclusion, they had found that the performance of ingredient recognition is better than normal method They had also found that this model shows promising performance for zero-shot recipe retrieval. Pouladzadeh et al. implemented a deep CNN and compared the recognition rate of high resolution food portion images with color-texture segmentation system [11]. The results of deep CNN's recognition rate was quite promising as the maximum failure to detect rate was 2%. Similarly, Singla et al. ran a CNN model on datasets contributed by them for food detection and food/non-food classification [14]. They achieved the overall accuracy of 99.2% on food/non-food image classification and 83.6% on food categorization.

2.2 Volume Estimation Techniques

Gingold et al. explained the details of 2D to 3D modeling in his paper [4]. The author went into details on how to create a cylinder from 2D spine, tilting the image and changes in angles associated with it, mirroring the image, generating the side-view of a drawing etc. Another research that integrates volume estimation to food image is the most similar study to this project. Hassannejad et al. implemented a three-step approach to food volume estimation [5]. In their study, they created a mobile application that scans the food. Then, user mark the food and non-food parts of the picture which goes through Gaussian Mixture Model (GMM) algorithm and Graph-cut algorithm to segment the food image from the picture. The segmented food image goes through 3D modeling algorithm that takes original image, segmented image and distance between phone camera and food. A checked cardboard is placed next to the food as a size reference. 6 images of food was taken as it was determined the optimized result. This study had achieved an average accuracy rate of 92%.

Martin et al. did the similar study, but they calculated the estimated area of food using reference card [8]. Similarly, Chang et al. attempted volume and mass estimation of food in their research [16]. In their study, food image was taken to create a 3D model with the help of reference card and several equations. From the volume, density and mass were calculated and the result was compared with an actual density and mass of food. The error rate of 10% was resulted in this study. This study was about multi-view volume estimation of food. Xu et al. also developed equations for 3D modelling of food images from single view in their study [15].

2.3 Calorie Estimation Techniques

One approach to the food calorie estimation is called multi-task CNN, where the neural network is trained simultaneously to perform multiple tasks. Ege and Yanai implemented and used this method in their research with which estimate food calories, food category, food ingredient and cooking direction [3]. They ran the model on a Japanese food dataset where the food calories are displayed on the website. They went through the same process with an American food dataset. 13% improvement in accuracy is found when compared to calorie estimation directly from picture of food. In this method, a picture of food is taken together with an object of known size. Okomoto and Yanai also used this method for food calorie estimation [10]. In their experiment, a picture of meal is taken together with a standard object. Then, the image of food and standard object are extracted from the picture and compared to get the size of a food. The detected food leads to which food category it is. Based on food size and category, the calorie estimation can be extracted. The absolute error resulted from this experiment was 50kcal and the relative error was 20%.

3 DESIGN

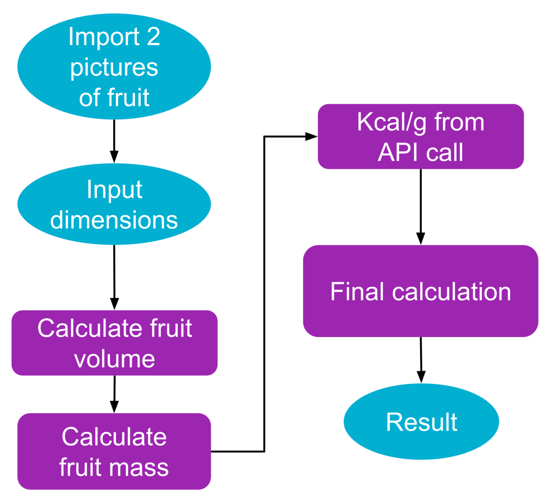

Figure 1: Project Design

The blue segments are the ones users will be seeing and experiencing. The purple are the back-end part of the process. In this application, the density database of fruits is built-in. When the user opens the application, the name of a fruit will be asked. This will be followed by the user importing two pictures and inputting the dimensions of the fruit using the ruler reference shown in the interface with imported pictures. Once all the values are given, the application will calculate the volume of the fruit, which will be used in calculating the mass of the fruit. The resulted mass and the kilocalories per 100 grams of the fruit from API will be used to return the total kilocalories of the fruit.

4 METHOD

4.1 Density Database

Currently there are 39 fruits and vegetables that can be consumed raw in the database. The database is currently a comma-separated values file. The intersection of a list of fruits from fruits360 dataset and raw fruits' information of USDA Food Central Database is used to get a list of fruits to be put in the database. Density is calculated using the equation:

$$density = \frac{mass}{volume}$$

In the USDA database, raw fruits' weight for 1 cup in grams is available. 1 cup is equivalent to 236.588 cubic centimeters. When looking for the mass of 1 cup of fruits, the fruits in chopped, diced or sliced conditions are used to the closest accuracy possible. The database is available at https://docs.google.com/spreadsheets/d/1jDqD9IfIDZZxEZ65345FQg2RRMgWfvh4kH176UefY4E/edit?usp=sharing. Since the size of the density database is relatively small for now, it is included as a raw .csv file.

4.2 Android Application

As shown in Figure 1, the blue, or oval-shaped, portions are what users will experience. This section will unravel the front-end part in detail. When a user opens the application, they are required to input the name of a fruit they would like to measure. The, the user will have to import the first picture, which is the side view of the fruit. After the user has imported the first picture, the picture will be shown together with the centimeter ruler picture with actual length as shown in Figure 2. Please note that the image size has been adjusted to fit in so the picture will not portray the real length ruler. The ruler is obtained from "https://catchydesk.com/wp-content/uploads/2019/04/Inch-ruler-actual-size.png".

Next, the user will be asked the width and the height of the fruit in the image in centimeters. After the user had put both values, they will be asked to import the top view of the fruit. Similarly, they will also be asked to put the width and height of the apple in the top-view image as well. By doing these simple steps, the users can obtain the estimated calories of fruit within a short amount of time. It is very important to recognize that the pictures are expected to be taken 15 centimeters away from the fruit for the maximum accuracy.

4.3 Volume and Calories Estimation

This section will explain the back-end work of the application. All four values that the user has put will be converted into real world values by using the equation below.

$$\text{Object size} = \frac{\text{Object size in image * Object distance}}{\text{Focal Length}}$$

At least two out of the four values should match or be very close to one of the two values put for the side-view dimensions of fruit because there can only be 3 dimensions to calculate volume. The

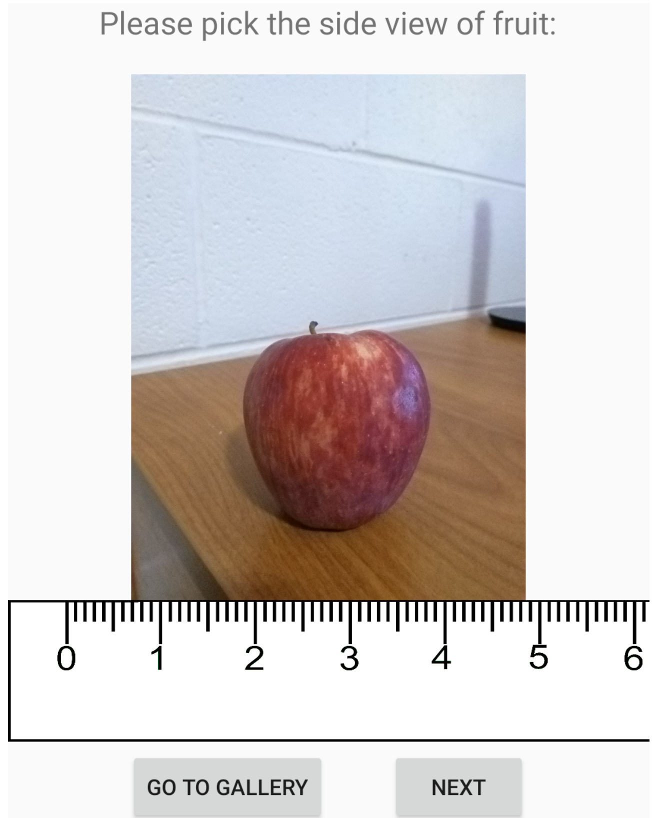

Figure 2: Image with ruler as appeared in the application

similar or the same one will be omitted. The other three values will be used to calculate volume of fruit in cuboid model (length * width * height). The volume with the density from database will be used to calculate mass of the fruit. The kilocalories of fruit per 100 grams will be retrieved from API using the input fruit name. This value will be converted to become kilocalories per 1 gram, and this value will be used to calculate and return the final kilocalories to the user. Users can repeat this process as many times as they please.

4.4 Accuracy testing

Testing is an essential part in every study. This application depends mainly on the user inputs. Therefore, the accuracy rate will be very dependent on the user input. In order to check the accuracy rate of this app, the following procedure was used.

(1) The kilocalories estimations of six apples were obtained from the application.
(2) The volumes of the six apples were measured using submerging in the water method as shown in Figure 3.
(3) The measured volumes are multiplied to the density of an apple from the database to obtain mass.
(4) The mass with kilocalories per 100 grams from the API are used to calculate the total kilocalories of all apples.
(5) The values are compared to calculate best, worst and average accuracy rates

Figure 3: Volume calculation of an apple

5 RESULTS

On average, the volume of an apple is about 200 cubic centimeters. The kilocalories of six apple ranged from 85.488 to 87.308. The average kilocalories of 6 apples from manual calculation is 86.11, and the average from the application is 74.89. It can be noted that in the worst case at Apple F, the system can return about 378% less than the real value. However, in the best case at Apple B, the accuracy rate of 95.67% is achieved. The average accuracy rate achieved from this project is 86.98%. The estimations of the lengths, the widths and the heights of apples in images ranged from 1.5 centimeters to 2.3 centimeters.

	Apple A	Apple B	Apple C	Apple D	Apple E	Apple F
Real Calories (in kcal)	85.49	85.72	86.98	87.31	86.39	84.76
Estimated Calories (in kcal)	62.48	82.01	115.5	58.58	78.1	52.72
Difference (in kcal)	23.01	3.71	-28.52	28.73	8.29	32.04

Table 1: Difference between real and estimation in each trial

Average Real Calories (in kcal)	86.11
Average Estimated Calories (in kcal)	74.89
Difference (in kcal)	11.21
Average Accuracy Percentage	86.98%

Table 2: Accuracy percentage on average

6 CONCLUSION

This project has proved that it is possible to implement an entirely image-based volume and calories estimation system. Without any image processing techniques, the results already look promising. When such techniques are included, the accuracy rate can only be better.

7 FUTURE DIRECTION

The final goal of this project is to achieve the results from the past studies by using solely image processing with the help of deep learning. In order to reach the goal, the accuracy rate must be above 90% on average and the error rate must not be higher than 20%. If the size of fruit in an image can be obtained through image processing, the accuracy rate will be better as well. The next immediate step for this project is to integrate image recognition system into the application.

ACKNOWLEDGMENTS

I would like to thank Dr. Becky Thomas and Dr. Charlie Peck for every help I have received throughout the process. I would also like to think Computer Science Department at Earlham College for the support.

REFERENCES

[1] Jingjing Chen and Chong-wah Ngo. 2016. Deep-Based Ingredient Recognition for Cooking Recipe Retrieval. In *Proceedings of the 24th ACM International Conference on Multimedia* (Amsterdam, The Netherlands) *(MM '16)*. Association for Computing Machinery, New York, NY, USA, 32–41. https://doi.org/10.1145/2964284.2964315

[2] J. Clement. 2019 (accessed May 10, 2020). *Leading health and fitness apps in the U.S. 2018, by users*. https://www.statista.com/statistics/650748/health-fitness-app-usage-usa/.

[3] Takumi Ege and Keiji Yanai. 2017. Image-Based Food Calorie Estimation Using Knowledge on Food Categories, Ingredients and Cooking Directions. In *Proceedings of the on Thematic Workshops of ACM Multimedia 2017* (Mountain View, California, USA) *(Thematic Workshops '17)*. Association for Computing Machinery, New York, NY, USA, 367–375. https://doi.org/10.1145/3126686.3126742

[4] Yotam Gingold, Takeo Igarashi, and Denis Zorin. 2009. Structured Annotations for 2D-to-3D Modeling. *ACM Transactions on Graphics (TOG)* 28, 5 (2009), 148. https://doi.org/10.1145/1618452.1618494

[5] Hamid Hassannejad, Guido Matrella, Paolo Ciampolini, Ilaria De Munari, Monica Mordonini, and Stefano Cagnoni. 2017. A New Approach to Image-Based Estimation of Food Volume. *Algorithms* 10 (06 2017), 66. https://doi.org/10.3390/a10020066

[6] Hokuto Kagaya, Kiyoharu Aizawa, and Makoto Ogawa. 2014. Food Detection and Recognition Using Convolutional Neural Network. In *Proceedings of the 22nd ACM International Conference on Multimedia* (Orlando, Florida, USA) *(MM '14)*. Association for Computing Machinery, New York, NY, USA, 1085–1088. https://doi.org/10.1145/2647868.2654970

[7] Y. Kawano and K. Yanai. 2013. Real-Time Mobile Food Recognition System. In *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 1–7. https://doi.org/10.1109/CVPRW.2013.5

[8] Corby Martin, Sertan Kaya, and Bahadir Gunturk. 2009. Quantification of Food Intake Using Food Image Analysis. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* 2009 (09 2009), 6869–72. https://doi.org/10.1109/IEMBS.2009.5333123

[9] Takuma Maruyama, Yoshiyuki Kawano, and Keiji Yanai. 2012. Real-Time Mobile Recipe Recommendation System Using Food Ingredient Recognition. In *Proceedings of the 2nd ACM International Workshop on Interactive Multimedia on Mobile and Portable Devices* (Nara, Japan) *(IMMPD '12)*. Association for Computing Machinery, New York, NY, USA, 27–34. https://doi.org/10.1145/2390821.2390830

[10] Koichi Okamoto and Keiji Yanai. 2016. An Automatic Calorie Estimation System of Food Images on a Smartphone. In *Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management* (Amsterdam, The Netherlands) *(MADiMa '16)*. Association for Computing Machinery, New York, NY, USA, 63–70. https://doi.org/10.1145/2986035.2986040

[11] P. Pouladzadeh, P. Kuhad, S. V. B. Peddi, A. Yassine, and S. Shirmohammadi. 2016. Food calorie measurement using deep learning neural network. In *2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings*. 1–6. https://doi.org/10.1109/I2MTC.2016.7520547

[12] Arivazhagan Selvaraj, Newlin Shebiah, Selva Nidhyananthan, and Lakshmanan Ganesan. 2010. Fruit Recognition using Color and Texture Features. *Journal of Emerging Trends in Computing and Information Sciences* 1 (10 2010), 90–94.

[13] Adnan Shehzad, Nauman Zafar, Mir Hassan, and Zhidong Shen. 2019. Food Item Recognition and Intake Measurement Techniques. In *Proceedings of the 2019 11th International Conference on Machine Learning and Computing* (Zhuhai, China) *(ICMLC '19)*. Association for Computing Machinery, New York, NY, USA, 405–411. https://doi.org/10.1145/3318299.3318379

[14] Ashutosh Singla, Lin Yuan, and Touradj Ebrahimi. 2016. Food/Non-Food Image Classification and Food Categorization Using Pre-Trained GoogLeNet Model. In *Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management* (Amsterdam, The Netherlands) *(MADiMa '16)*. Association for Computing Machinery, New York, NY, USA, 3–11. https://doi.org/10.1145/2986035.2986039

[15] C. Xu, Y. He, N. Khanna, C. J. Boushey, and E. J. Delp. 2013. Model-based food volume estimation using 3D pose. In *2013 IEEE International Conference on Image Processing*. 2534–2538. https://doi.org/10.1109/ICIP.2013.6738522

[16] Chang Xu, Ye He, Nitin Khannan, Albert Parra, Carol Boushey, and Edward Delp. 2013. Image-Based Food Volume Estimation. In *Proceedings of the 5th International Workshop on Multimedia for Cooking Eating Activities* (Barcelona, Spain) *(CEA '13)*. Association for Computing Machinery, New York, NY, USA, 75–80. https://doi.org/10.1145/2506023.2506037