SeeFood - Calories Prediction from Food Images

Gokberk Sahin [1] **Okan Alan** [1] **Emre Yazici** [1]
Department of Computer Science Hacetepe University Ankara,Turkey

Abstract

In this paper we introduce a method for calorie calculation. We employ a multi-view image recognition and R-CNN for object detection. We experiment with a dataset (1) of 3000 images. We aim to predict the calorie of the given food correctly.

1. Introduction

This projects aim is to find the food in the given image and predict the calorie of the food correctly. This algorithm takes two pictures of the food with a measure coin (one from side and one from above) and recognize the food in the picture. After finding the food, it calculates the calorie of the recognized food by finding the volume and mass of the food with the help of the coin in the picture.

2. Related Work

In recent years, some vision based methods emerged for food volume estimation. According to inputs, they can be grouped into two classes: single-view or multi-view.

The single view technique requires only single image and estimates food volume by using a reference for camera calibration such as dining plate, playing card, thumb, block or coin after food portion segmentation and identification. However, the use of reference objects is inconvenient in real application due to difficulty of fetching(2).

The multi-view technique requires at least two image for reconstructing three-dimensional (3D) models of food and calculating food portion size. In multi-view cases either a coin or a playing card is used as reference for camera calibration in three-dimensional (3D) reconstruction(2).

There is a study that calculates volume of fruits which are extracted trough Faster R-CNN (3) with the help of the coin in the picture as a reference. This study uses Faster R-CNN to get series of bounding boxes after the images inputted as RGB channels. Then using an image prossesing aproach to segment each bounding box and stores in matrix after replacing the background pixels with zeros. Using a 1 CNY coin(2.5 cm diameter) as a reference it calculates the volume.

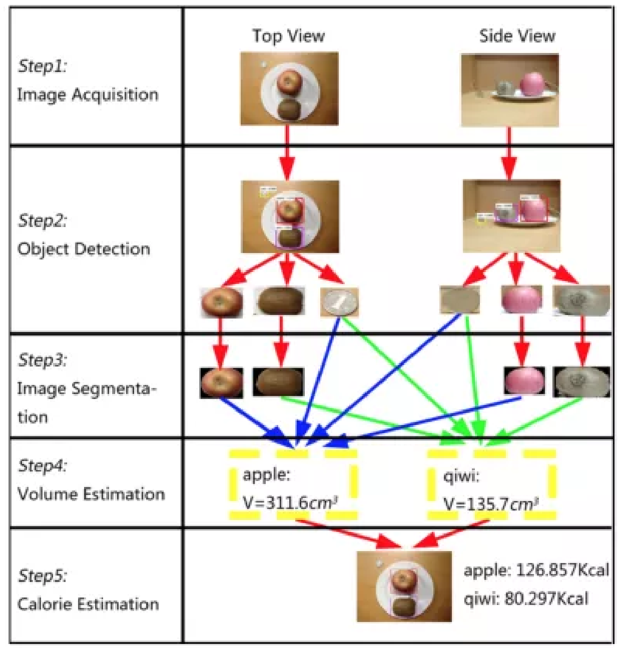

Figure 1. Calorie Estimation Flowchart. All pictures taken from: (3)

$$\alpha_S = \frac{2.5}{(W_S + W_S)/2}$$

In this equation W_S and H_S are width and height of the bounding box in side view picture.

$$\alpha_T = \frac{2.5}{(W_T + H_T)/2}$$

In this equation W_t and H_t are width and height of the bounding box in top view picture.

*Equal contribution [1]Hacettepe University, Department of Computer Engineering. Correspondence to: Machine Learning Class <Students>.

BBM406 Fundamentals of Machine Learning, Fall 2018. Copyright 2018 by the author(s).

Authors of this study categorized foods based on the shape and select volume estimation formula accordingly.

$$v = \begin{cases} \beta \times \frac{\pi}{4} \times \sum_{k=1}^{H_S} (L_S^k)^2 \times \alpha_S^3 & \text{if} \quad \textit{the shape is ellipsoid} \\ \beta \times (s_T \times \alpha_T^2) \times (H_S \times \alpha_S) & \text{if} \quad \textit{the shape is column} \\ \beta \times (s_T \times \alpha_T^2) \times \sum_{k=1}^{H_S} (\frac{L_S^k}{L_S^{MAX}})^2 \times \alpha_S & \text{if} \quad \textit{the shape irregular} \end{cases}$$

In this equation L_S^k is the number of foreground pixels in row k (k $\in$ 1,2,...,H_s). $L_{MAX} = max(Lk, ..., Lk)$ records maximum number of foreground pixels. β is compensation factor (default = 1.0).

For the volume estimation mass is needed. It can be calculated with $m = \rho \times v$ where $\rho(cm^3)$ represents volume and $v(g/cm^3)$ represents density.

Finally calorie can be obtained with $C = c \times m$ where $m(g)$ represents mass and $c(Kcal/g)$ represents calorie.

This study also provides a dataset of top and side view of the fruits that includes ground truth volume of fruits.

The main purpose of SeeFood is to efficiently and correctly predict the calorie of the food. SeeFood can be extended to other food types given that pictures of the foods in the dataset.

3. Methodology

3.1. Image Recognition

If we divide the problem to sub-problems we can see that the real problem is estimating the volume of a food. The accuracy of this estimation will directly influence the estimation of calories. As mentioned in related work, we will use a 5 step algorithm. After feeding the top and side view of food image we will detect the foods in the picture and the coin which will be used for scale. For detection we'll use the Tensorflow's object detection API (4) which outperforms YOLO (5), Faster-RCNN (6) and it's variants. In our data set it recognizes all the apples but since it is a pre-trained model it doesn't recognize the coins. We will train our own custom model for coin recognition as well. Then we will use image segmentation to get their bounding boxes and measures. Lastly, we will use these measures to actually obtain the volume of our food relative to the coin which should give a estimation about the real volume of the food.

3.2. Calorie and Volume Estimation

Foods in the data set will be divided to different shapes as ellipsoid, column and irregular. We'll use different formulas for different shapes which are mentioned in related work. Obviously instead of 1 CNY we can use 1TL or 0.5 TL coins for scale. After estimating the volume we will multiply the value with food's density to obtain mass. Also we know how many kcals per grams for each food so we can easily obtain the total calorie of the food multiplying by this value denoted as c. The mean error between estimated and real volume does not exceed $\pm 20\%$ in the previous work. Given that data set is relatively small we can expect to decrease this error by expanding the data set.

4. Experimental Evaluation

In this section, we introduce the experimental evaluations for three learning tasks.These are food classification, volume estimation and calories prediction. Results created by using ECUSTFD(ECUST Food Dataset) (7). This dataset divided into train and test sets.

4.1. Food Classification

We have 19 different type foods. These are apple, banana, bread, bun, doughnut, egg, fired dough twist, grape, lemon, litchi, mango, mooncake, orange, peach, pear, plum, qiwi, sachima, tomato. We have 2978 images in dataset. This is not big dataset. Therefor we split our dataset into three folds: training(%70 of the data), validation(%15 of the data), test(%15 of the data).

All images goes to object detection model for classification. We will use TensorFlow's object detection API for detection and classification.(8)

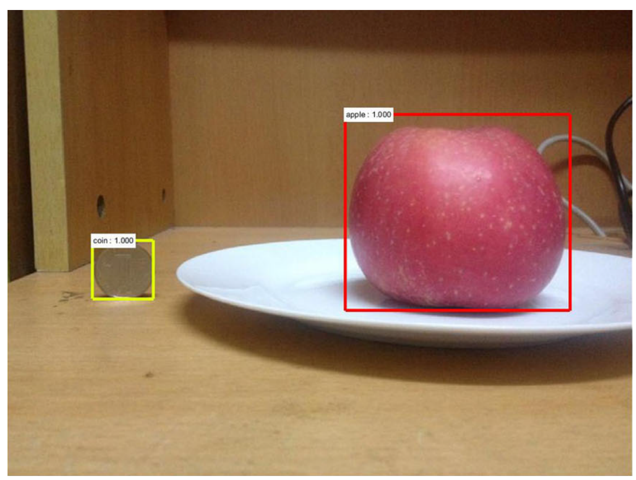

Figure 2.Detected Coin and Food

4.2. Volume Estimation

We know actual values of food in given images. We will use coin to reference point to find volume. We will use mean error to evaluate estimation results. Mean error ME is defined as:

$$ME_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \frac{v_j - V_j}{V_j}$$

4.3. Calorie Estimation

This is our main task in the project. We will use things what we learned in the previous two learning tasks which are detection/classification and volume estimation.

We know how many actual values of calorie per gram of the food. Using formulas that mentioned above and our previous estimates from the previous task, we can estimate calories.

5. Future Work

We have covered the methodologies and algorithm for detecting the foods and estimating the calories. First step for our future work is to detect the food and the coin in the image and get their bounding boxes. We will further analyze the methods for isolating these detecting images and getting a accurate values of their sizes in image. After this we can simply use the formulas to calculate the actual volume of the food. Since the precision of this estimation will greatly influence the estimation of calories we will mostly focus on improving this method and try to get better results so we can make better calorie estimations. Also we will expand the data set by taking new photos of existing foods. This way we can have chance to better understand how light and shooting angle conditions will affect the estimation results.

References

[1] J. L. Yanchao Liang. Computer vision-based food calorie estimation: Dataset, method, and experiment. [Online]. Available: https://arxiv.org/pdf/1705.07632.pdf

[2] B. H. L. I. I. M. Vera Kurkova, Yannis Manolopoulos, *Artificial Neural Networks and Machine Learning – ICANN 2018*, ser. 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings. Springer, 2018.

[3] SYNCED. Deep learning-based food calorie estimation method in dietary assessment. [Online]. Available: https://bit.ly/2PZZ1xD

[4] Z. Lu. Tensorflow object detection api. [Online]. Available: https://github.com/tensorflow/models/tree/master/research/object_detection

[5] J. Redmon. Yolo : Real-time object detection. [Online]. Available: https://pjreddie.com/darknet/yolo/

[6] R. G. J. S. Shaoqing Ren, Kaiming He. Faster r-cnn(object detection) towards data science. [Online]. Available: https://towardsdatascience.com/review-faster-r-cnn-object-detection-f5685cb30202

[7] Liang-yc. Ecust food dataset resized version. [Online]. Available: https://github.com/Liang-yc/ECUSTFD-resized

[8] Sentdex. [Online]. Available: https://www.youtube.com/watch?v=COlbP62-B-U&list=PLQVvvaa0QuDcNK5GeCQnxYnSSaar2tpku