



Search



Write

Get unlimited access to the best of Medium for less than \$1/week. [Become a member](#)

Deep Learning-Based Food Calorie Estimation Method in Dietary Assessment



Synced · Follow

Published in SyncedReview · 8 min read · Aug 3, 2017

203

1



Due to the improvement in people's standards of living, obesity rates are increasing at an alarming speed, and this is reflective to the risks in people's health. People need to control their daily calorie intake by eating healthier foods, which is the most basic method to avoid obesity. However, although food packaging comes with nutrition (and calorie) labels, it's still not very convenient for people to reference. Thus, scientists started to use machine learning algorithms in computer vision to help people determine the caloric value in the food they eat. During the 2015 Rework Deep Learning Summit in Boston, Google scientist Kevin Murphy presented a deep learning algorithm that was used to analyze static food image. By analyzing the composition of the food in picture, the algorithm can calculate how much calories the dish has.

This paper is trying to provide a more efficient way of estimating calories. First, it needs the top view and side view images of the food being analyzed. Then, it will use Faster R-CNN to detect the food and calibration object, after which, a GrabCur algorithm is used to determine the food's contour. After estimating the volume of food, the authors can finally estimate the amount of calories.

Introduction

When people's Body Mass Index (BMI) is over 30 (kg/m²), they are generally considered to be obese. High BMI can increase the risk of illnesses like heart disease [1]. The main reason of obesity is due to the imbalance between the amount of caloric intake (consumption) and energy output (expenditure). Because of unwillingness to record and track, lack of related nutritional information or other reasons, patients often experience trouble in controlling the amount of calories they consume. There are lots of proposed methods to estimate calories based on computer vision [2, 3, 4, 5], but after the authors' analysis, the accuracy of detection and volume estimation still need to be improved. In this paper, the main difference from other similar approaches is that it requires an input of two images, and the use Faster R-CNN to detect the object and GrabCut algorithm to obtain each food's contour. After that, the authors can estimate each food's volume and calories.

Material and Methods

A. Calorie Estimation Method Based On Deep Learning

This method is shown in Figure 1. As mentioned before, the process of estimating calories requires two images from top and side, and each image should include the calibration object. Here, the authors choose Faster Region-based Convolutional Neural Networks (Faster R-CNN) [5] to detect objects, and GrabCut algorithm [6] as the segmentation algorithm.

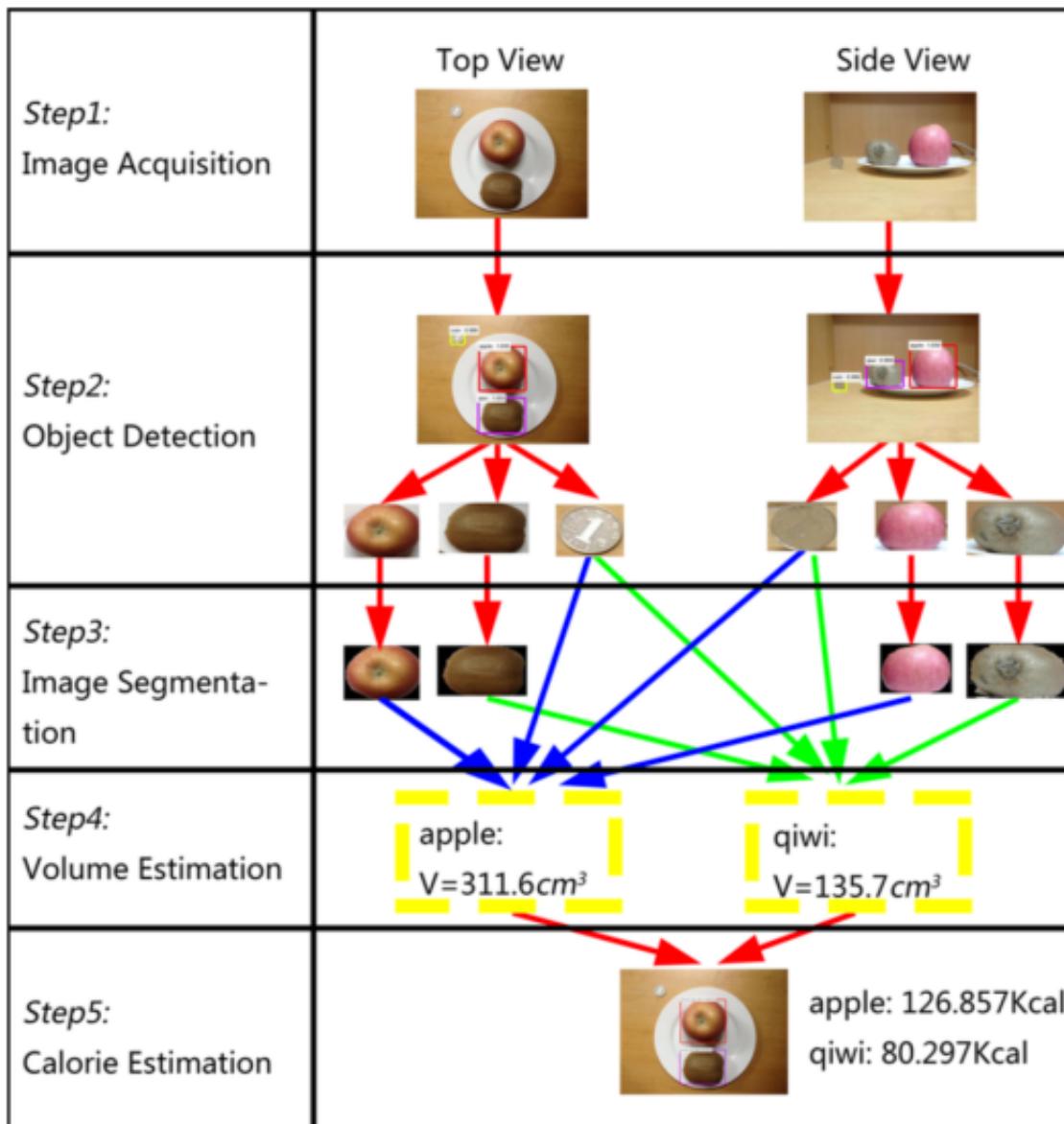


Figure 1: Calorie Estimation Flowchart

B. Deep Learning Based Objection Detection

The authors chose Faster R-CNN instead of using semantic segmentation method such as Fully Convolutional Networks (FCN). Here, after the images are inputted as RGB channels, the authors can get a series of bounding boxes, which means the class if judged.

C. Image Segmentation

This process uses an image processing approach to segment each bounding box. As mentioned above, the bounding boxes around the object that GrabCut needs can be provided by Faster R-CNN. After segmentation, we can get a series of food images stored in matrix, but with the the values of the background pixels being replaced by zeros. This will leave only the foreground pixels.

D. Volume Estimation

To estimate the volume, the authors calculate the scale factors based on calibration objects. The authors use a 1 CNY coin to show the specific process of calculating the volume. The diameter of the coin is 2.5 cm, and the side view's scale factor was calculated with Equation 1.

$$\alpha_s = \frac{2.5}{(W_s + H_s)/2} \quad (1)$$

In this equation, W_s is the width of the bounding box, H_s is the height of the bounding box. Similarly, the top view's scale can be calculated with Equation 2.

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

After, the authors divide the foods into three categories based on shape: ellipsoid, column, irregular. Different volume estimation formula will be selected for different types of food, according to Equation 3. HS is the height of side view PS and LkS is the number of foreground pixels in row k ($k \in 1, 2, \dots, HS$). LMAX = max(Lk ,...,Lk), it records the maximum number of foreground pixels in PS. β is a compensation factor (default value = 1.0). After that, for each food type there will be a unique value.

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

E. Calorie Estimation

After estimating the volume, the next step is to estimate each food's mass. It can be calculated in Equation 4, Where v (cm^3) represents the volume of current food, and ρ (g/cm^3) represents its density value

$$m = \rho \times v \quad (4)$$

Then the calorie of the food can be obtained with Equation 5.

$$C = c \times m \quad (5)$$

Where $m(g)$ represents the mass of current food and $c(Kcal/g)$ represents its calories per gram.

Results and Discussion

A. Dataset Description

In this paper, the authors use their own food dataset name ECUSTFD (downloadable on this [website](#)) ECUSTFD contains 19 kinds of food. They use a smart phone to take the required images, and each pair of images contains a top view and a side view. A 1 CNY coin is used as the calibration object. Additionally, for each image in ECUSTFD, they provide annotations, volume and mass records.

B. Object Detection Experiment

The author use a comparison experiment to choose the object detection algorithm. The numbers of training images and testing images are shown in Figure 2. Average Precision was used to evaluate the object detection results. In the test set, Faster R-CNN achieves 93.0% while Exemplar SVM achieves 75.9%.

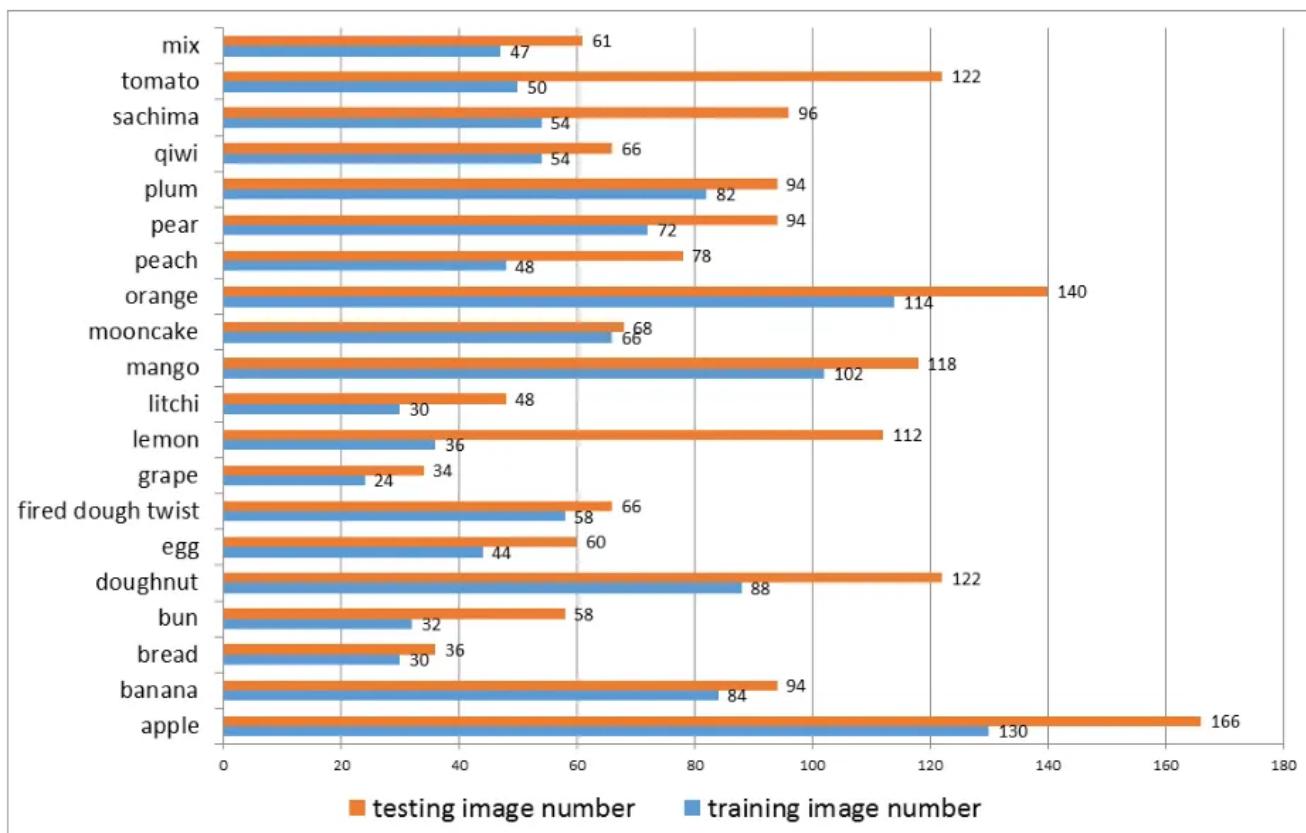


Figure 2: training and testing image numbers

C. Food Calorie Estimation Experiment

β (compensation factor) in Equation 3 can be calculated with Equation 6, where k is the food type, and N is the number of volume estimation.

$$\beta_k = \left(\frac{\sum_{i=1}^N V_k^i}{\sum_{i=1}^N v_k^i} \right) \quad (6)$$

p in Equation 4 can be calculated with Equation 7.

$$\rho_k = \left(\frac{\sum_{i=1}^N M_k^i}{\sum_{i=1}^N v_k^i} \right) \quad (7)$$

After that, the authors give the shape definition, estimation images number, β , ρ of each food in Table 1.

Table 1: Shape Definition and Parameters in Our Experiments

Food Type	shape	estimation image number	β_k	ρ_k
apple	ellipsoid	122	1.13	0.88
banana	irregular	82	0.64	0.59
bread	column	26	0.70	0.13
bun	irregular	32	1.14	0.43
doughnut	irregular	42	1.38	0.42
egg	ellipsoid	30	1.00	1.16
fired dough twist	irregular	48	1.30	0.77
grape	column	24	0.25	0.25
lemon	ellipsoid	34	1.11	1.04
litchi	irregular	30	0.98	0.95
mango	irregular	20	1.24	1.33
mooncake	column	64	1.04	1.24
orange	ellipsoid	110	1.12	0.99
peach	ellipsoid	48	1.07	1.09
pear	irregular	72	1.12	1.09
plum	ellipsoid	82	1.22	1.19
qiwi	ellipsoid	54	1.11	1.09
sachima	column	54	1.13	0.25
tomato	ellipsoid	46	1.20	1.07

Then, by using the images from the test set, the results can be shown in

Table 2.

Table 2: Volume and Mass Estimation Experiment Results

Food Type	estimation	mean	mean	volume	mean	mean	mass
	image number	volume	estimation	error	mass	estimation	error
apple	158	332.78	320.65	-3.65	263.60	250.02	-5.15
banana	92	162.17	128.13	-20.99	146.61	116.92	-20.25
bread	20	155.00	102.03	-34.17	29.04	18.71	-35.57
bun	56	247.50	237.02	-4.23	78.19	89.91	15.00
doughnut	10	166.00	197.58	19.03	63.44	59.44	-6.31
egg	42	52.38	56.40	7.67	61.20	65.80	7.51
fired dough twist	38	64.74	70.73	9.26	40.60	41.72	2.76
grape	26	240.00	203.98	-15.01	219.50	203.07	-7.48
lemon	112	96.79	100.54	3.88	94.24	94.71	0.49
litchi	46	43.48	45.35	4.30	44.01	44.38	0.84
mango	38	80.53	91.89	14.12	90.58	98.66	8.92
mooncake	68	70.00	55.04	-21.37	64.01	65.81	2.82
orange	100	240.00	266.37	10.99	218.44	234.69	7.44
peach	72	106.67	116.53	9.25	111.29	117.81	5.86
pear	70	266.86	265.57	-0.48	256.43	257.12	0.27
plum	94	100.00	115.98	15.98	105.14	112.92	7.41
qiwi	60	122.00	127.19	4.25	122.30	124.68	1.95
sachima	96	147.29	136.85	-7.09	31.89	30.13	-5.50
tomato	120	171.00	199.40	16.61	180.55	179.21	-0.74

The authors use mean volume error to evaluate volume estimation results. The definition of mean volume error is as shown in Equation 8, where food type is i , $2N_i$ is the number of images Faster R-CNN recognizes correctly.

$$ME_V^i = \frac{1}{N_i} \sum_{j=1}^{N_i} \frac{v_j - V_j}{V_j} \quad (8)$$

The definition of mean mass error is in Equation 9.

$$ME_M^i = \frac{1}{N_i} \sum_{j=1}^{N_i} \frac{m_j - M_j}{M_j} \quad (9)$$

For the results in Table 2, we see that most types of food's estimation results are closer to reference real values. Other than banana, bread, and mooncake, the mean error between estimation volume and true volume does not exceed $\pm 20\%$. Even if the drainage method is not that accurate, but the estimation method can be accepted.

Conclusion

This paper gives us a calorie estimation method, and the results of the experiments show promise.

Since the images are taken from smartphones, and the image processing methods used here are well-developed, this proposed method can be

easily integrated into health apps as an engineering solution. Nevertheless, from a research perspective, I think this paper has two limitations. First, there is no comparison with prior work. The authors did provide a literature review in the introduction, but I think they should have compared their results with the results in those prior work. If this approach can achieve a better performance, then we can say that this paper provides a more effective way. Unfortunately we can't say that, because the author didn't provide a series of comparison experiment. Secondly, I am not sure if the dataset is accurate or big enough. The authors just say that they take the images from a smartphone, but they didn't tell whether there is a standard to collect the images. Like the light intensity, and the number of pixels. Besides, in Table 2 we can see that the mean error is still large, which indicates that there is some room to make the mean error much smaller.

Reference

- [1] W. Zheng, D. F. McLerran, B. Rolland, X. Zhang, M. Inoue, K. Matsuo, J. He, P. C. Gupta, K. Ramadas, S. Tsugane, Association between body- mass index and risk of death in more than 1 million asians, *New England Journal of Medicine* 364 (8) (2011) 719–29.
- [2] W. Jia, H. C. Chen, Y. Yue, Z. Li, J. Fernstrom, Y. Bai, C. Li, M. Sun, Accuracy of food portion size estimation from digital pictures acquired by a chest-worn camera., *Public Health Nutrition* 17 (8) (2014) 1671–81.
- [3] Z. Guodong, Q. Longhua, Z. Qiaoming, Determination of food portion size by image processing, 2008, pp. 119–128.
- [4] Y. Bai, C. Li, Y. Yue, W. Jia, J. Li, Z. H. Mao, M. Sun, Designing a wearable computer for lifestyle evaluation., in: *Bioengineering Conference*, 2012, pp. 93–94.

- [5] P. Pouladzadeh, P. Kuhad, S. V. B. Peddi, A. Yassine, S. Shirmohammadi, Mobile cloud based food calorie measurement (2014) 1–6.
- [6] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks, in: Advances in neural information processing systems, 2015, pp. 91–99.
- [7] C. Rother, V. Kolmogorov, A. Blake, Grabcut: Interactive foreground extraction using iterated graph cuts, in: ACM transactions on graphics (TOG), Vol. 23, ACM, 2004, pp. 309–314.

Author: Shixin Gu | **Editor:** Joni Chung | **Localized by Synced Global Team:** Xiang Chen

Machine Learning

Nutrition

Technology

Food

Computer Science



Written by Synced

36K Followers · Editor for SyncedReview

AI Technology & Industry Review — syncedreview.com | Newsletter:
<http://bit.ly/2IYL6Y2> | Share My Research <http://bit.ly/2TrUPMI> | Twitter:
@Synced_Global

Follow

More from Synced and SyncedReview



 Synced in SyncedReview

Standford U's MAPTree: Redefining Decision Trees—...

Decision trees have emerged as one of the cornerstones of modern machine learning...

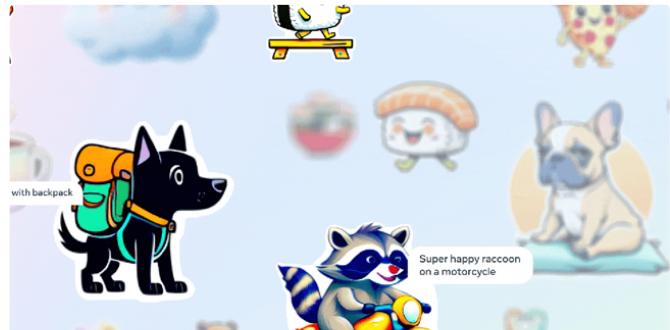
★ · 4 min read · Oct 3

48

1



...



 Synced in SyncedReview

Meta AI's Long-Context LLMs: Redefining the Landscape of...

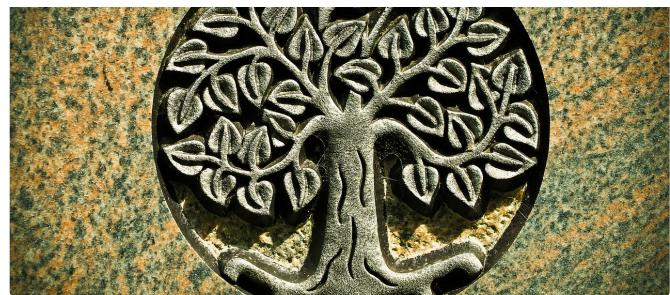
★ · 3 min read · Sep 28

76

1



...



 Synced in SyncedReview

NNAISENSE's New Class of Generative Model: Bayesian Flo...

In recent years, large-scale neural networks have ushered in a transformative era in...

 · 4 min read · Oct 4

 15 

 Synced in SyncedReview

Google Unveils the Enigma of Memorization and Generalizati...

The success of modern, overparameterized neural models has sparked significant...

 · 4 min read · Oct 17

 21 

[See all from Synced](#)

[See all from SyncedReview](#)

Recommended from Medium



 Barack Obama 

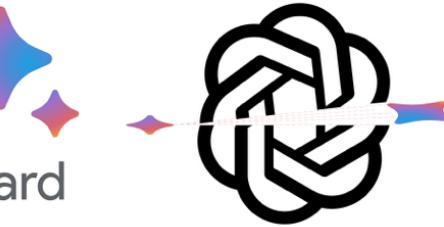
Thoughts on Israel and Gaza

It's been 17 days since Hamas launched its horrific attack against Israel, killing over...

5 min read · 6 days ago

 32K  816

+ ...



 AL Anany 

The ChatGPT Hype Is Over—Now Watch How Google Will Kill...

It never happens instantly. The business game is longer than you know.

 · 6 min read · Sep 1

 16.9K  503

+ ...

Lists



Predictive Modeling w/ Python

20 stories · 535 saves



AI Regulation

6 stories · 165 saves



ChatGPT prompts

27 stories · 565 saves



Practical Guides to Machine Learning

10 stories · 618 saves

Bye Bye, Spotify

Marco Peixeiro  in Towards Data Science

TimeGPT: The First Foundation Model for Time Series...

Explore the first generative pre-trained forecasting model and apply it in a project...

★ · 12 min read · 5 days ago

 1.1K

 14



Scott-Ryan Abt in Pitfall

Bye Bye, Spotify

And see ya later, all you subscription services in my little empire

★ · 4 min read · Aug 20

 14.3K

 342



Dr. Ashish Bamania  in Level Up Coding

Google Rejected Max Howell(Creator Of Homebrew)...

Can you solve this Google interview question?

★ · 4 min read · Oct 3

 6.4K

 71



Maximilian Vogel in MLearning.ai

The ChatGPT list of lists: A collection of 3000+ prompts,...

Updated Oct-18, 2023. New developer resources, marketing & SEO prompts, new...

11 min read · Feb 8

 10.3K

 125



See more recommendations

[Help](#) [Status](#) [About](#) [Careers](#) [Blog](#) [Privacy](#) [Terms](#) [Text to speech](#) [Teams](#)