



IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2320-6710



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 7, July 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.569



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com

A Review on Image-Based Estimation of Real Food Size for Accurate Food Calorie Estimation

Nidhi Meshram, Prof. Rupesh Sushir

Department of Electronics and Telecommunication Engineering, P.R. Pote Patil Education & Welfare Trust's Group of Institutions, College of Engineering & Management, Amravati, Sant Gadge Baba Amravati University, Maharashtra, India

ABSTRACT: In this paper, we review our works on image-based estimation of real size of foods for accurate food calorie estimation which including three existing works and two new works: (1)“CalorieCam” which is a system to estimate real food size based on a reference object, (2) Region segmentation based food calorie estimation, (3) “AR Deep CalorieCam V2” which is based on visual inertial odometry built in the iOSARKit library, (4) “Depth CalorieCam” which employs stereo cameras on iPhone X/XS, and (5)“RiceCalorieCam” which exploits rice grains as reference objects. Especially, the last two new methods achieved 10% or less estimation error, which was enough for robust food calorie estimation.

I. INTRODUCTION

In food image recognition, CNN-based methods have achieved great improvement in recent years and some smartphone applications employ them. However, in most of the calorie estimation, the estimated calories are just associated with the estimated food categories and these applications often require users to enter information such as size or volume, there are problems that it is a troublesome and subjective evaluation. Currently, no applications which can estimate food calories automatically exist. Although most of the image recognition tasks including food category recognition have been almost solved due to great progress of CNN based image recognition methods, fully-automatic food calorie estimation from a food photo has still remained an unsolved problem. We think that food calorie estimation not only helps people's health a lot, but also is promising as a new problem of image recognition studies. In this paper, we review three of our works related to food size estimation, and propose two new methods for food size estimation.

- (1) “CalorieCam” which is a reference-object-based food calorie estimation system.
- (2) Weakly-supervised segmentation based food calorie estimation
- (3) “AR Deep CalorieCam V2” which is a real food size and calorie estimation system based on visual inertial odometry implemented in iOSARKit.
- (4) “Depth CalorieCam” (new system) which is a food calorie estimation system exploiting iPhone stereo cameras.
- (5) Rice grain based size estimation (new method) which uses rice grains the size of which are usually almost the same as a reference object.

II. CALORIECAM

“CalorieCam” [1] is an image-based calorie estimation system which estimates food calories automatically by simply taking a food photo from the top with a pre-registered reference object. Taking into account usability and mobility of a system, we implemented the system as a stand-alone mobile application running on a normal Android smartphone.

To estimate food calorie from a single image, a user needs to register a size-known reference object in advance and to take a food photo with the registered reference object. As a reference object, we assume a personal belonging which we are always carrying such as a wallet and a credit card-size card in a wallet (except a smartphone for taking a meal photo). After taking a meal photo with a reference object, the system carries out segmentation of food items and the pre-registered reference object. Because the real size of the reference object is known (e.g. In case of a credit-card-size object, the size is 85.6mm x 54mm.), the system can estimate the real size of each detected food item by comparing the

number of pixels of the reference objects and the detected food items. By using the estimated real size and the equations to calculate food calorie from their size, the system finally estimates the calorie of the food items in the real photo. Figure 1 shows an example of usage and a screen-shot of the proposed system.



Figure 1. (Upper) A user taking a food photo with the reference object. (Lower) A screen shot

III. WEAKLY-SUPERVISED SEGMENTATION BASED CALORIE ESTIMATION

In the previous work, we assumed that one meal photo contains only one dish. This means we have to recognize dishes one by one for the case of multiple dishes. Then, in this work we introduced CNN-based region proposal to cope with multiple-dish meals. We estimate calories from segmentation results by taking account of area ratios of multiple foods. With our method, we can estimate food calories for a photo without multiple-view photos and specific reference objects such as wallets and cards.

The proposed method on CNN-based region detection consists of the following steps as shown in Fig.2:

- Apply selective search [6] and obtain 2000 bounding box proposals at most.
- Group them and select bounding boxes.
- Perform back propagation over the pre-trained CNN regarding all the selected bounding boxes.
- Obtain saliency maps by averaging BP outputs within each group
- Extract segments based on the saliency maps with GrabCut
- Apply non-maximum suppression (NMS) to obtain final region results.

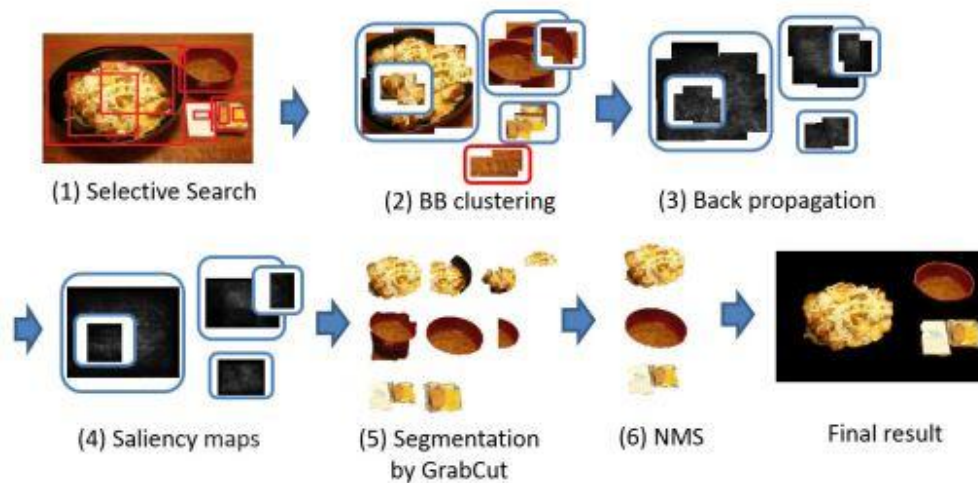


Figure 2. The processing flow of the proposed method.

After region segmentation, we choose a base food region the real size of which is expected to be the same as the standard dish in the database from the segmentation results of multiple foods. We do not fix the category of a base food so that there are a lot of patterns in food combination. Therefore, we define the priorities of food items for choosing a base food class. We decide the priorities based on a tendency of unchanging food volumes. Some food volumes change frequently, while some foods' volume rarely change. For example, in "Teishoku" which is Japanese traditional food combo menu, we can often change the volume of "rice" as options, while we cannot change "miso-soup" volume in general. In this case, "miso-soup" is more appropriate for a base food. In this way, there are differences in tendency of unchanging food volumes and we defined food ranking the volume of which are relatively stable as a high priority class for choosing a base food class. After selecting a base food the actual size and calorie is assumed to be known, we calculate calories of other dishes than the dish of the base food by comparing the number of pixels between each dish and the dishes of the base food taking account of calorie densities which depends on food categories.

IV. RICE GRAIN BASED SIZE ESTIMATION

In this work, we propose CNN-based real size estimation from boiled rice grains images. We use grains of boiled rice as a reference object and construct a CNN-based method that input a patch image of boiled rice grains images and output real size of a side of an input patch image. Figure 8 shows examples of real size annotated boiled rice photos, and Figure 3 shows the architecture of our network. The architecture of our network is based on VGG16. As shown in the Figure 3, it has an output layer composed of a single unit that outputs a real size. An input image is a patch image cropped from boiled rice grains image and the output is a real size of a side of the input patch image. In this work, since the input patch image is 244×244 , the output the real size for 224 pixels. The real size estimation is treated as a regression problem. Generally, in the regression problem, a mean square error is used as the loss function. In this work, we use the mean square error shown in the Equation (1).

$$L = \frac{1}{n} \sum_{k=1}^n (x_k - y_k)^2$$

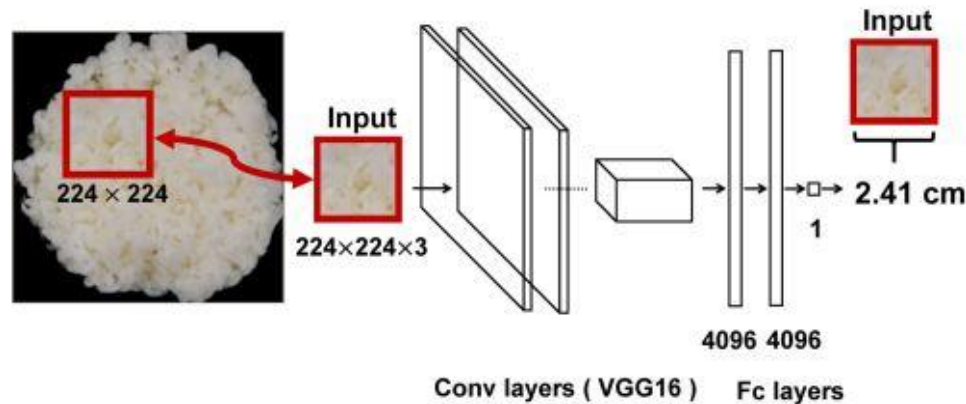


Figure 3. Our network that estimate a real size from boiled rice grains images.

Table 1. The results of the real size estimation from boiled rice grains images

Evaluation data.	abs. err.(cm/224pixels)	rel. err.(%)	corr.	≤ 5% err.(%)	≤ 10% err.(%)	≤ 20% err.(%)
Camera:COOLPIX, Small amount of water	0.212	7.182	0.958	41.667	75.000	91.667
Camera:COOLPIX, Medium amount of water	0.178	6.550	0.973	43.333	76.667	93.333
Camera:COOLPIX, Large amount of water	0.197	6.668	0.962	48.333	78.333	90.000
Camera:iPhone8 Plus, Small amount of water	0.127	5.652	0.945	50.000	75.000	98.333
Camera:iPhone8 Plus, Medium amount of water	0.170	7.512	0.903	43.333	68.333	88.333
Camera:iPhone8 Plus, Large amount of water	0.105	4.800	0.967	58.333	88.333	98.333

V.CONCLUSION

In this paper, we reviewed our three previous works and proposed two new works. Currently, DepthCalorieCam is the most promising approach. However, large-scale calorie annotated 3D food volume data is needed to extend the system into large-scale categories, which is very costly and time-consuming. In addition, the rice grain based method is also promising for meals containing white steamed rice. Especially it is appropriate for Japanese foods.

REFERENCES

- [1] K. Okamoto and K. Yanai. An automatic calorie estimation system of food images on a smartphone. In *Proc. of ACM MM Workshop on Multimedia Assisted Dietary Management*, 2016.
- [2] W. Shimoda and K. Yanai. CNN-based food image segmentation without pixel-wise annotation. In *Proc. of IAPR International Conference on Image Analysis and Processing*, 2015.
- [3] R. Tanno, T. Ege, and K. Yanai. AR DeepCalorieCam V2: Food calorie estimation with cnn and ar-based actual size estimation. In *Proc. Of ACM Symposium on Virtual Reality Software and Technology (VRST)*, 2018.
- [4] C. Rother, V. Kolmogorov, and A. Blake. GrabCut: Interactive foreground extraction using iterated graph cuts. *ACM Trans. on Graphics*, 23(3):309–314, 2004.
- [5] R. Tanno, K. Okamoto, and K. Yanai. DeepFoodCam: A DCNN based real-time mobile food recognition system. In *Proc. of ACM MM Workshop on Multimedia Assisted Dietary Management*, 2016.
- [6] J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, and A. W. M. Smeulders. Selective search for object recognition. *International Journal of Computer Vision*, 104(2):154–171, 2013.
- [7] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. *Springer*, pages 234–241, 2015.
- [8] Y. Matsuda, H. Hajime, and K. Yanai. Recognition of multiple-food images by detecting candidate regions. In *Proc. of IEEE International Conference on Multimedia and Expo*, pages 25–30, 2012.
- [9] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In *arXiv preprint arXiv:1409.1556*, 2014.



**Impact Factor:
7.569**



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details