AR DeepCalorieCam V2: Food Calorie Estimation with CNN and AR-based Actual Size Estimation

Ryosuke Tanno*

Department of Informatics, The University of Electro-Communications, Tokyo
tanno-r@mm.inf.uec.ac.jp

Takumi Ege, Keiji Yanai

Department of Informatics, The University of Electro-Communications, Tokyo
{ege-t,yanai}@mm.inf.uec.ac.jp

ABSTRACT

In most of the cases, the estimated calories are just associated with the estimated food categories, or the relative size compared to the standard size of each food category which are usually provided by a user manually. In addition, in the case of calorie estimation based on the amount of meal, a user conventionally needs to register a size-known reference object in advance and to take a food photo with the registered reference object. In this demo, we propose a new approach for food calorie estimation with CNN and Augmented Reality (AR)-based actual size estimation. By using Apple ARKit framework, we can measure the actual size of the meal area by acquiring the coordinates on the real world as a three-dimensional vector, we implemented this demo app. As a result, it is possible to calculate the size more accurately than in the previous method by measuring the meal area directly, the calorie estimation accuracy has improved.

KEYWORDS

Food Calorie Estimation, Augmented Reality, Deep Learning, iOS, Application

ACM Reference Format:
Ryosuke Tanno and Takumi Ege, Keiji Yanai. 2018. AR DeepCalorieCam V2: Food Calorie Estimation with CNN and AR-based Actual Size Estimation. In *VRST 2018: 24th ACM Symposium on Virtual Reality Software and Technology (VRST '18), November 28-December 1, 2018, Tokyo, Japan.* ACM, New York, NY, USA, Article 4, 2 pages. https://doi.org/10.1145/3281505.3281580

1 INTRODUCTION

In recent years, due to a rise in healthy thinking on eating, many people take care of their eating habits, and some people record daily diet regularly. To assist them, many mobile applications for recording everyday meals have been released so far. Some of them employ food image recognition which can estimate not only food names but also food calories. However, in most of the cases, the estimated calories are just associated with the estimated food categories, or the relative size compared to the standard size of each food category which are usually provided by a user manually. Most of the applications does not estimate calories based on the amount of foods.

On the other hand, there are some studies on calorie estimation considering the amount of meals. For example, in [3], 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, they assume a personal belonging which they are always carrying such as a credit-card-size card. After taking a meal photo with a reference object, the system carries out segmentation of food items and the pre-registered reference object. Hence, 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 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. However, with the above method, it is absolutely necessary to have a pre-registered reference object to estimate the real size of meal. In addition, the dispersion of calorie estimation accuracy is large, since indirectly obtaining the meal area from the reference object depends on the segmentation accuracy of the meal area.

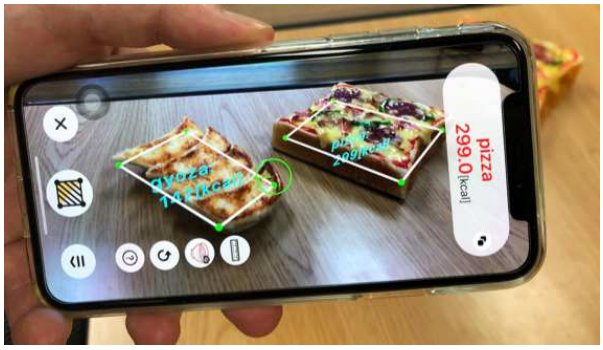

Figure 1: Example of the use of the proposed iOS App "AR DeepCalorieCam V2"

In this demo, we propose a new approach for food calorie estimation with CNN and AR-based actual size estimation. By using Apple ARKit framework, it is possible to measure the actual size of the meal area by acquiring the coordinates on the real world as a three-dimensional vector. Figure 1 shows an example of usage of the proposed "AR DeepCalorieCam V2".

Our contributions are as follows.

- The reference object, which was conventionally necessary to obtain the actual meal area, became unnecessary by using AR technology on mobile devices such as iPhone.
- By measuring the meal area directly, it is possible to calculate the size more accurately than in the previous method. (Please refer to the Table 1)

2 TECHNICAL DESCRIPTION

The demo system outline is as shown in the Figure 2. The processing of the proposed system consists of the following two steps:

- Recognize a category of each food item.
- Directly calculate the size of the region of the food item of food items using AR and then calculate food calories based on their actual size and food categories.

*Currently he is working for NTT Communications.

VRST '18, November 28-December 1, 2018, Tokyo, Japan

ACM ISBN 978-1-4503-6086-9/18/11.
https://doi.org/10.1145/3281505.3281580

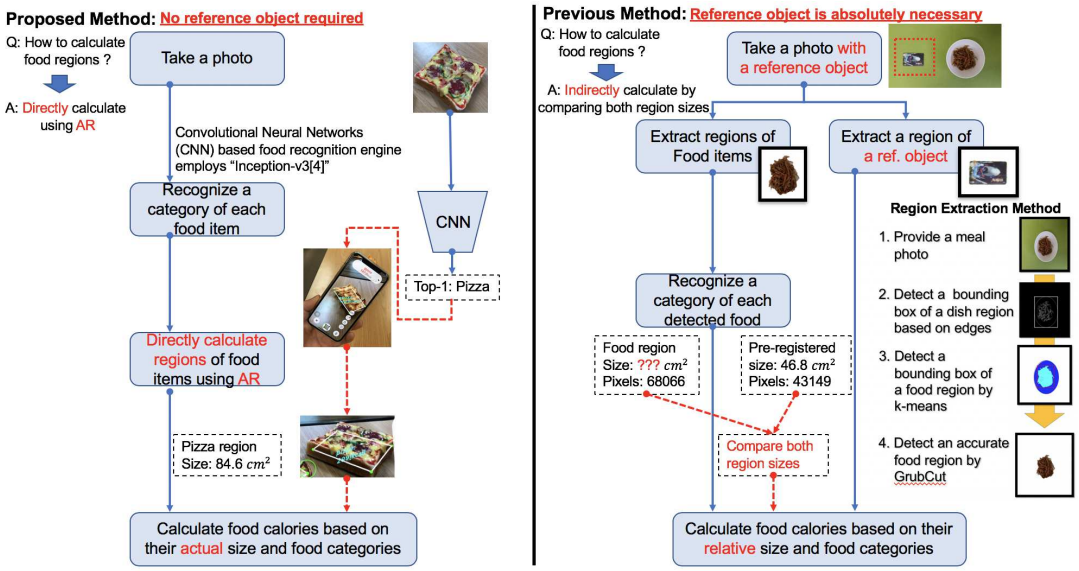

Figure 2: An overview of the demo system(Left: Proposed Method, Right: Previous Method)

Table 1: Estimation errors by the baseline [1] and the baseline [3] and the proposed system.

		baseline[1]		baseline[3]		proposed	
food name	real value(kcal)	avg. err.	SD	avg. err.	SD	avg. err.	SD
beef rice bowl	962	-53.25	209.79	-242	55.1	-67.14	18.8
croquette	552	-242	91.26	-47.08	52.52	-127.0	9.0
salad	14	54.83	36.28	4.86	11.87	-0.95	0.16

2.1 Food Category Recognition

For mobile implementation, the memory capacity and processing time of the device is an important factor in implementing deep learning. For this time, we decided to use Inception-v3 [4] which has a light memory, faster inference and high classification accuracy. We fine-tune a pre-trained ImageNet model in Keras Deep Learning framework with UEC-FOOD100 dataset [2].

2.2 Calorie Estimation

We assume that the height of food portion correlates with the size of foods and food categories, and we estimate calories of food items directly from the food sizes estimated from the top-view image. To do that, we use not simple linear estimation but quadratic curve estimation from the 2D size of foods to their calories. The quadratic curve of each food category is estimated based on the training data annotated with real food calories independently. We use a quadratic curve estimation from the 2D size of foods to their calories.

In the case of the proposed method, we can calculate the real size of foods area directly using AR technology on mobile devices such as iPhone. Therefore, the reference object, which was conventionally necessary to obtain the actual meal area, became unnecessary. By using Apple ARKit framework, we can measure the actual size of the meal area by acquiring the coordinates on the real world as a three-dimensional vector and then calculate food calories based on their actual size and food categories.

3 DEMO OVERVIEW

This demo is performed by the step on the left side of Figure 2. First of all, we recognize meal categories by taking food samples on iPhone. Next, we directly calculate regions of food items using AR. Finally, we calculate food calories based on their actual size and food categories.

4 CONCLUSION

In this demo, we propose a new approach for food calorie estimation with CNN and AR-based actual size estimation. In the future, we plan to implement an automatic calorie estimation system that combines the proposed method and segmentation, taking into account the size of the meal area.

ACKNOWLEDGMENTS: This work was supported by JSPS KAKENHI Grant Number 15H05915, 17H01745, 17H05972, 17H06026 and 17H06100.

REFERENCES

[1] Y. Kawano and K. Yanai. 2013. Real-time Mobile Food Recognition System. In *Proc. of CVPR International Workshop on Mobile Vision (IWMV).*
[2] Y. Matsuda, H. Hoashi, and K. Yanai. 2012. Recognition of Multiple-Food Images by Detecting Candidate Regions. In *Proc. of IEEE International Conference on Multimedia and Expo (ICME).*
[3] K. Okamoto and K. Yanai. 2016. An Automatic Calorie Estimation System of Food Images on a Smartphone. In *Proc. of ACM MM Workshop on Multimedia Assisted Dietary Management (MADiMa).*
[4] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. 2015. Rethinking the Inception Architecture for Computer Vision. In *Proc. of arXiv:1512.00567.*