DeepFoodCam: A DCNN-based Real-time Mobile Food Recognition System

Ryosuke Tanno, Koichi Okamoto, Keiji Yanai
Department of Informatics, The University of Electro-Communications, Tokyo
{tanno-r,okamoto-k,yanai}@mm.inf.uec.ac.jp

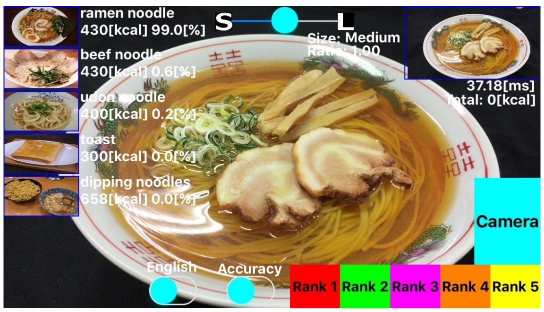

Figure 1: A screen-shot of "DeepFoodCam" for iOS.

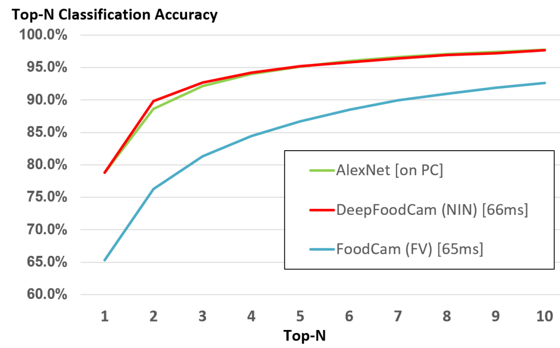

Figure 2: Performance comparison among NIN-based "DeepFoodCam", standard AlexNet and FV-based "FoodCam" on the UEC-FOOD100 dataset regarding the top-N classification accuracy.

Due to the recent progress of the studies on deep learning, deep convolutional neural network (DCNN) based methods have outperformed conventional methods with a large margin. Therefore, DCNN-based recognition should be introduced into mobile object recognition. However, since DCNN computation is usually performed on GPU-equipped PCs, it is not easy for mobile devices where memory and computational power is limited.

In this demo, we show the possibility of DCNN-based object recognition on mobile devices, especially on iOS and Android devices including smartphones and tablets in terms of processing speed and required memory. As an example of DCNN-based mobile applications, we show a food image recognition app, "DeepFoodCam". Previously we have proposed "FoodCam" [1] which is a mobile app of food image recognition employing linear SVM and Fisher Vector (FV) of HoG and color patches. In this demo, we show DCNN-based "DeepFoodCam" outperformed FV-based "FoodCam" greatly in terms of recognition accuracy as shown in Figure 2, although the processing time was kept almost the same.

We use multi-scale network-in-networks (NIN) [2] in which users can adjust the trade-off between recognition time and accuracy. We implemented multi-threaded mobile applications on both iOS and Android employing either NEON SIMD instructions or the BLAS library for fast computation of convolutional layers, and compared them in terms of recognition time on mobile devices. As results as shown in Table 1, it has been revealed that BLAS is better for iOS, while NEON is better for Android, and that reducing the size of an input image by resizing is very effective for speedup of DCNN-based recognition. In case of using iPad Pro with BLAS, we achieved 66.0ms as the processing time for one time recognition on iPad Pro. Please refer to [3] about the detail.

We have released the iOS-version food recognition app on Apple App Store. Please search for "DeepFoodCam". The Android-version app can be download from our project page, http://foodcam.mobi/.

MADiMa'16 October 16-16 2016, Amsterdam, Netherlands

ACM ISBN 978-1-4503-4520-0/16/10.
DOI: http://dx.doi.org/10.1145/2986035.2986044

Table 1: Recognition time [ms] on mobile devices.

	NIN(BLAS)	NIN(NEON)
iPad Pro	**66.0**	221.4
iPhone SE	**79.9**	251.8
Galaxy Note 3	1652	**251.1**

1. REFERENCES

[1] Y. Kawano and K. Yanai, "Foodcam: A real-time food recognition system on a smartphone," *Multimedia Tools and Applications*, vol. 74, no. 14, pp. 5263–5287, 2015.

[2] M. Lin, Q. Chen, and S. Yan, "Network in network," in *Proc. of International Conference on Learning Represenation Conference Track*, 2014.

[3] K. Yanai, R. Tanno, and K. Okamoto, "Efficient mobile implementation of a cnn-based object recognition system," in *Proc. of ACM International Conference Multimedia*, 2016.