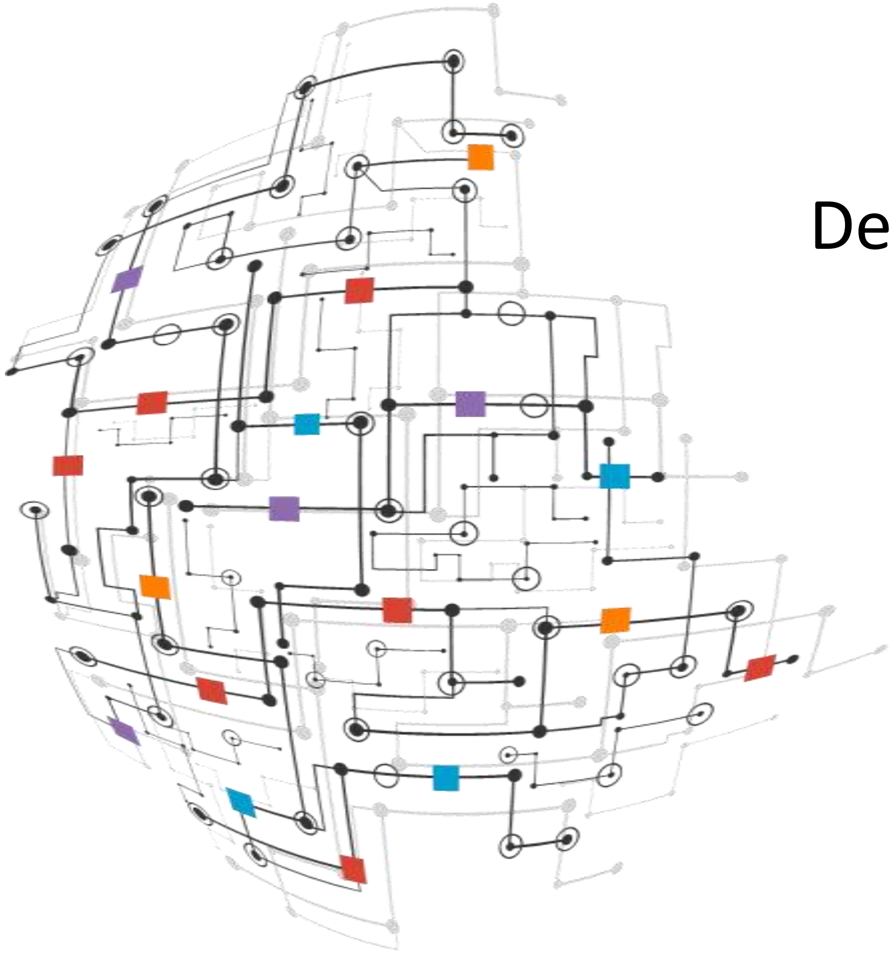
VISUAL 2022 - The Seventh International Conference on Applications and Systems of Visual Paradigms



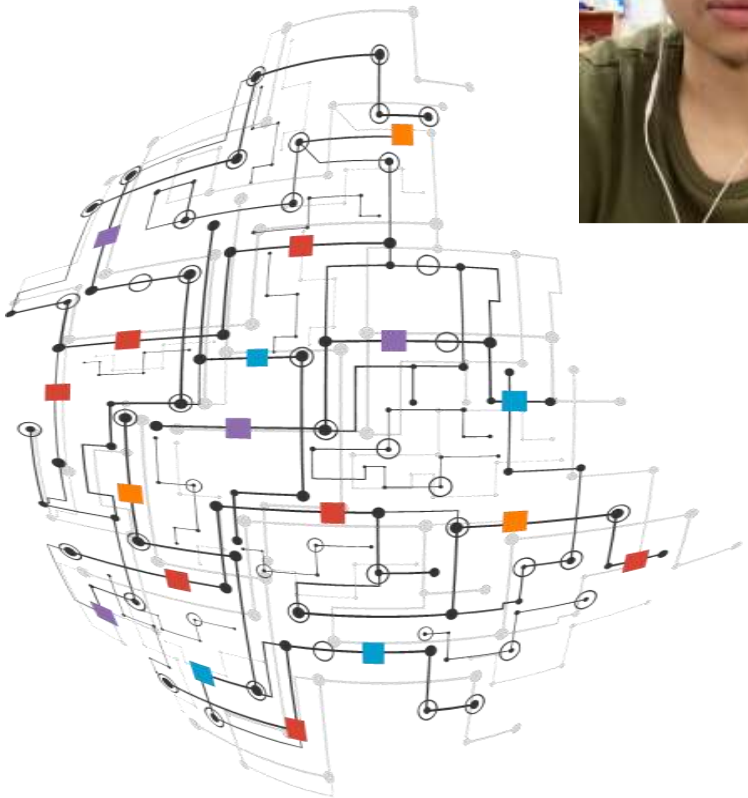
Deep Learning-Based Food Identification and Calorie Estimation System

Authors: Chen-Hao Wang, Jung-Tang Huang

Graduate student of National Taipei University of Technology
Presenter: Chen-Hao Wang
Email:kevin8901300179@gmail.com

國立臺北科技大學
National Taipei University of Technology

June 26, 2022 to June 30, 2022 - Porto, Portugal



Chen-Hao Wang

Email:kevin8901300179@gmail.com

Education

- M.S. Major in Graduate Institute of Automation Technology, National Taipei University of Technology, 2022

Research

- Network theory and application
- Algorithm
- Big data analysis
- Artificial intelligence(Deep Learning)

TAIPEI TECH

國立臺北科技大學

National Taipei University of Technology

IARIA

Introduction

Goals	Contributions
• Train a YOLO model for food. • Maximize usage data. • Calorie Estimation.	• More than 90% recognition rate can be obtained. • Calorie estimation error is less than 10%.

Hardware

Raspberry Pi	Item
CPU	ARM Cortex-A72 (ARMv8) 1.5GHz
GPU	H.265 (4Kp60), H.264 (1080p60 / 1080p30)，OpenGL ES 3.0
Memory	2/4/8 GB (LPDDR4-3200)
USB	USB 3.0 *2、USB 2.0 *2

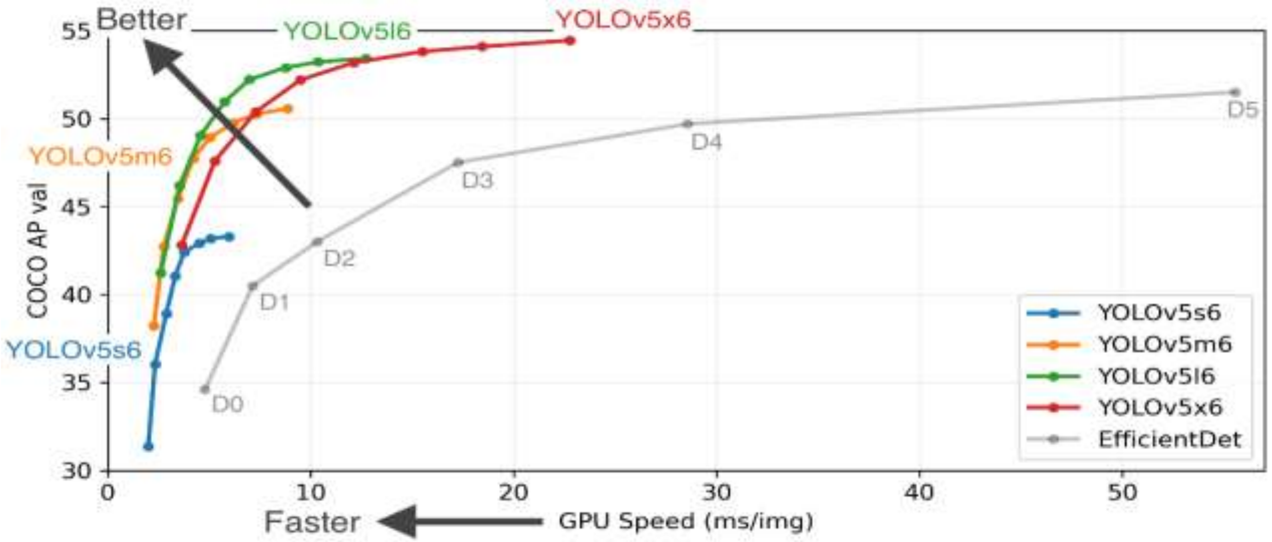
Depth Camera	Item
RGB FOV(H×V)	69°×42°(±1°)
Depth resolution	1280×720
RGB resolution	1920×1080
Min depth distance	0.105m
Max depth distance	10m

Data Acquisition

- In this work, a food dataset named Food-101 is used, which contains a large number of images.
- The pictures collected by myself, a total of 400 photos in 4 categories, including: rice, eggs, shrimp, broccoli, etc.
- Collect as many pictures of different angles and light as possible.

Yolo v5

- The YOLO (YouOnly Look Once) algorithm was originally proposed by Joseph Redmon.
- In YOLO v5, its author Glenn Jocher gave a total of four versions of the target detection network, divided into Yolov5s, Yolov5m, Yolov5l, Yolov5x. Compared with YOLO v4, YOLO v5 will reduce the size Reduced by 90%.
- On Tesla P100 YOLO v5 claims that it can achieve 140FPS fast detection, YOLO v4 The results are obtained at 50FPS, but the accuracy of the two is almost the same.



Volume Estimation

- Using yolov5 label the food type and record the coordinates.

- Depth images of pre- and post-meal food are recorded, and the depth difference for each pixel on the image is calculated. The depth difference of each point is the height of the actual height.

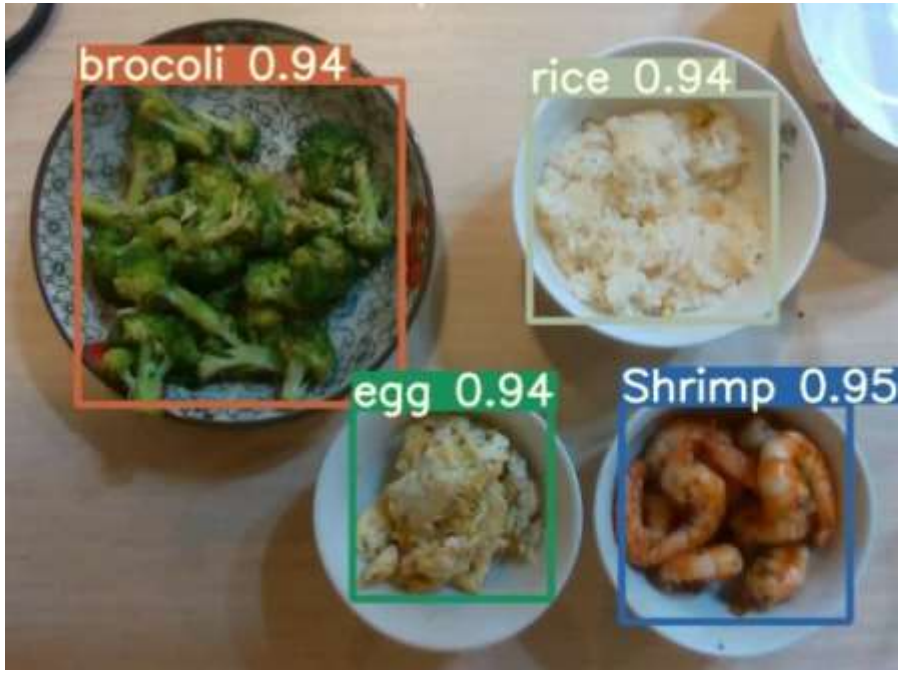
- The pixels go through linear regression to find the actual area of each pixel. Finally, sum and integrate all depth differences to get the true volume of the food.

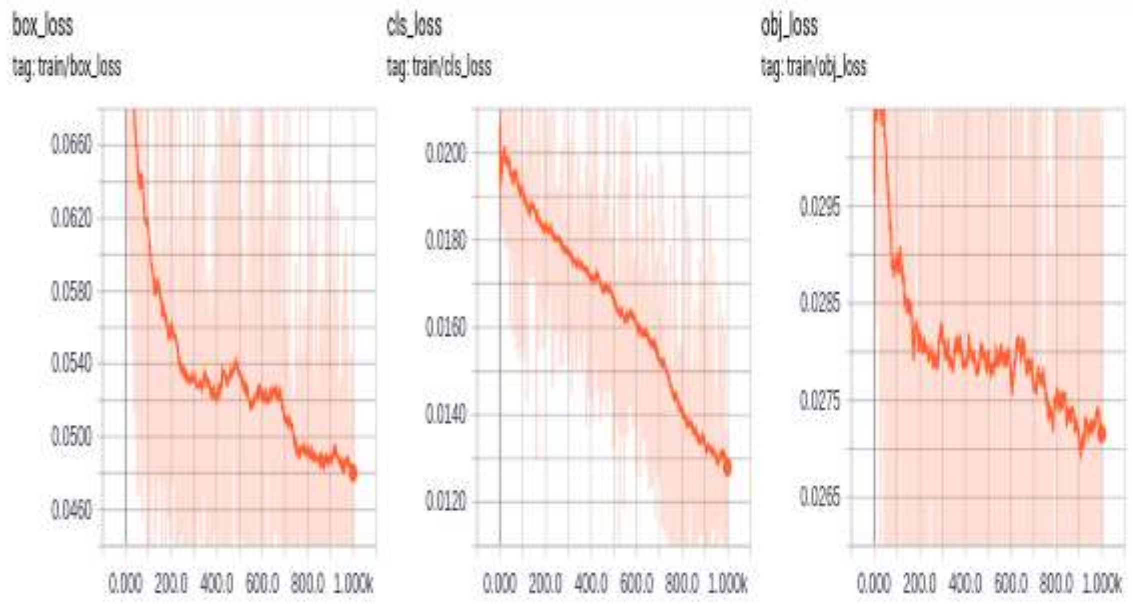
Calories Estimation

- For the convenience of calculation Calories of food, need to know the weight of the food, the density formula is used for Convert volume to weight, and density of each food obtained by the drainage method. Every food has one Density is different. Finally, convert calories and Get the three nutrients from food.

Image recognition Results

- We randomly select 10 images from the dataset and use our trained model to identify them with an average accuracy of over 90% for individual food items, The recognition rate of recognition is shown. The recognition rate is above 93%.
- The trained loss function is below 0.01.





Calorie Estimation Results

- This result is the same as the previous experiment, we randomly select 10 images from the dataset and use our model to estimate the calories of the food, record the real weight and the estimated weight, average all the recorded weights and calculate the error. has an error of less than 10%, which means that the calorie error will also be less than 10 calories.

Food	Data		
	Actual weight (g)	Estimated weight (g)	Error (%)
Rice	82.8	76.8	7.8
Egg	71	72.8	2.5
Broccoli	70.6	65.8	7.2
Shrimp	97.6	103.8	6.3

Conclusions

- Results show that it is feasible to use depth cameras for image recognition and calorie estimation.
- Uses a depth camera for image recognition and calorie estimation of the food, more than 90% recognition rate can be obtained.
- In most cases, while the heat estimation error is less than 10%.

Future Work

Experiment	Technology
• Increased experimental volume to increase experimental confidence.	• Increase the recognition rate and the types of recognition. • Reduce errors in calorie estimates.

References

[1] N. Dalal, and B. Triggs, “Histograms of oriented gradients for human detection,” 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), June 2005.
[2] H. Kaiming, G. Georgia, D. Piotr,and G. Ross, “Mask R-CNN,” Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 2961–2969, 2017.
[3] R. Shaoqing, H. Kaiming, G. Ross, and S. Jian, “Faster R-CNN: towards real-time object detection with region proposal networks,” NIPS 15: Proceedings of the 28th International Conference on Neural Information Processing Systems, Vol. 1, pp. 91–99, Dec 2015.
[4] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You only look once: Unified real-time object detection", Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779–788, 2016.
[5] G. Shroff, A Smailagic, and D. P. Siewiorek, “Wearable context-aware food recognition for calorie monitoring,” 2008 12th IEEE International Symposium on Wearable Computers, Oct. 2008.
[6] M. Anthimopoulos ,J Dehais, and P Diem, “Segmentation and recognition of multi-food meal images for carbohydrate counting,” in 13th IEEE International Conference on BioInformatics and BioEngineering, Nov. 2013.

References

[7] T. Ege, and K. Yanai “Estimating Food Calories for MultipleDish Food Photos,” 2017 4th IAPR Asian Conference on Pattern Recognition (ACPR), Nov. 2017.
[8] A. H. Abdulnabi, G. Wang,J .Lu, and K. Jia, “Multi-Task CNN Model for Attribute Prediction,” in IEEE Transactions on Multimedia, vol. 17, pp. 1949–1959, Nov. 2015.
[9] Y. Liang, and J. Li, “Deep Learning-Based Food Calorie Estimation Method in Dietary Assessment,” Cornell University, pp. 12, Jun 2017.
[10C. Rother, V. Kolmogorov, and A. Blake, “GrabCut: Interactive foreground extraction using iterated graph cuts,” ACM Trans. Graph., vol. 23, pp. 309–314, 2004.
[11] Y. Lu, D. Allegra, M. Antonopoulos , and F. Stanco, “A multi-task learning approach for meal assessment,” CEA/MADiMa '18: Proceedings of the Joint Workshop on Multimedia for Cooking and Eating Activities and Multimedia Assisted Dietary Management , pp. 46–52, July 2018.
[12] Y. Ando, T. Ege, J. Cho, and K. Yanai, “DepthCalorieCam: A Mobile Application for Volume-Based FoodCalorie Estimation using Depth Cameras,” MADiMa '19: Proceedings of the 5th International Workshop on Multimedia Assisted Dietary Management, pp. 76–81, Oct 2019.
[13] H. Raikwar, H. Jain, and A. Baghel, “Calorie Estimation from Fast Food Images Using Support Vector Machine,” International Journal on Future Revolution in Computer Science & Communication Engineering, vol. 4, pp. 98–102, April 2018.