Food Detection Using Faster R-CNN

Benny Mai
Simon Fraser University
bmai@sfu.ca

Jiali Cai
Simon Fraser University
jialic@sfu.ca

Ji xin Luo
Simon Fraser University
jxluo@sfu.ca

Junjie Liao
Simon Fraser University
jla597@sfu.ca

Wendy Huang
Simon Fraser University
wha57@sfu.ca

1 Introduction

Food makes the essentials of our daily life, and wouldn't it be interesting to be able to detect different types of food through visual media? In our project, we will study modern object detection technologies. Specifically, given an image of foods, we are aiming at detecting the existence of foods and the bounding boxes containing the food.

2 Method

There has been some excellent previous research on object detection, which is useful for the purpose of our project. For example, Faster R-CNN [3] is an end-to-end model for object detection. In our project, we are going to finetune a pre-trained Faster R-CNN model, which is already embedded in PyTorch. The dataset we use for fine-tuning the model is UEC FOOD [1, 2], which contains photos of different kinds of Japanese food and each food photo has a bounding box indicating the location of the food item. UEC FOOD has two versions, one of which comprises 100 categories and another of which comprises 256 categories. However, each category comprises very limited images (roughly 100), which are likely to cause overfitting. Thus, we will augment the dataset to allow better training by flipping, adding random noises, or changing the illumination. Given the time limit for this project, we will train 10 most representative foods. We will also compare the training results with different parameters, such as feature extractor, learning rate, and momentum.

3 Additional Work

If time permits, we will detect foods as much as possible. Meanwhile, if time permits, we are going to deploy our model in production. For example, we are going to probably build an AI cashier based on our model.

References

[1] Kawano, Y., & Yanai, K. (2014, September). Automatic expansion of a food image dataset leveraging existing categories with domain adaptation. *European Conference on Computer Vision*. pp. 91-99. Springer, Cham.

[2] Matsuda, Y., Hoashi, H., & Yanai, K. (2012, July). Automatic Expansion of a Food Image Dataset Leveraging Existing Categories with Domain Adaptation *Proc. of ECCV Workshop on Transferring and Adapting Source Knowledge in Computer Vision*. pp. 25-30. IEEE.

[3] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Advances in neural information processing systems*. pp. 91-99.