Foodness Proposal for Multiple Food Detection by Training with Single Food Images

Wataru Shimoda Keiji Yanai

The University of Electro-Communications, Tokyo, Japan

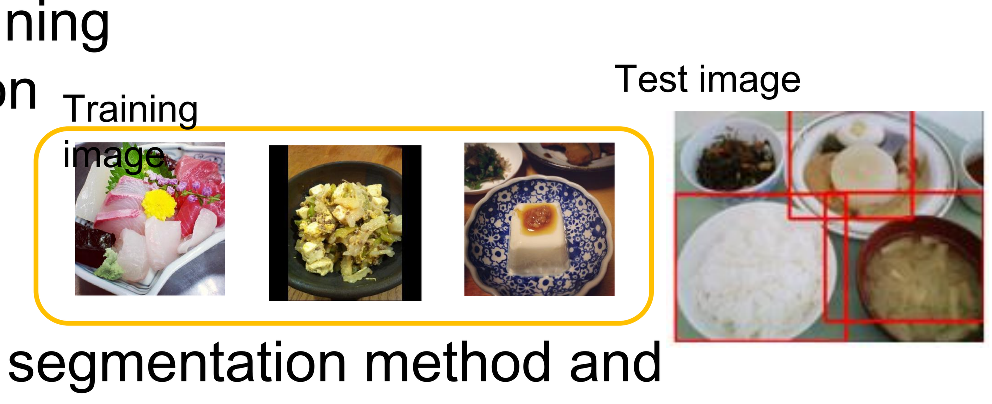
Objective

・ *Weakly supervised detection*

・ Use only image level annotation

・ Use only single label for training

・ Target is multi-food detection



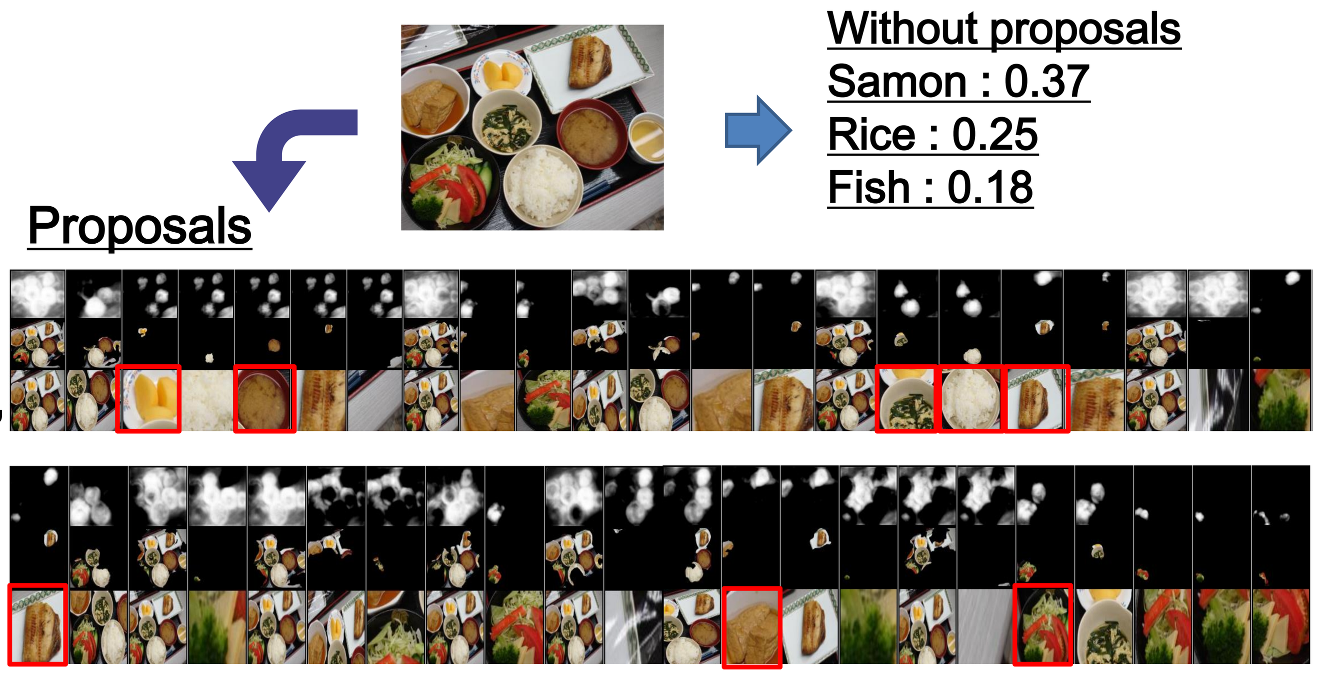
Contribution

- Combine weakly supervised segmentation method and proposal base approach
- Improve accuracy from weakly supervised segmentation results.
- improve computational cost from proposal base method.

Food region proposal

・ We regard estimated regions of upper rank classes as proposals.

・ If there are no target foods category in fact, the estimated food regions are belong to any food region.

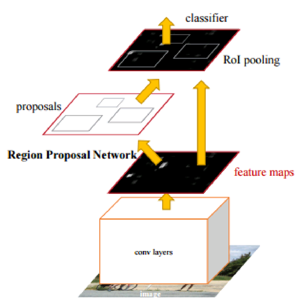


Related work

Fully supervised method

・ Faster RCNN

-Use bounding box annotation



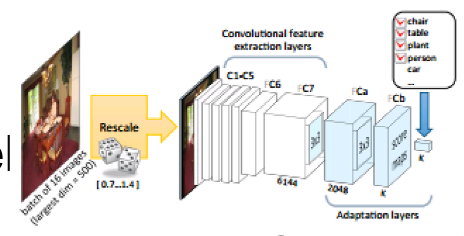
[Ren et al. NIPS 2015]

Weakly supervised method

・ Fully convolutional network

+ Global max pooling

-Train with single label and multi-label



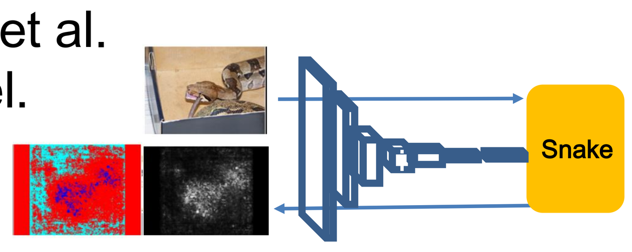
[Oquab et al. CVPR 2015]

・ Distinct class specific saliency maps

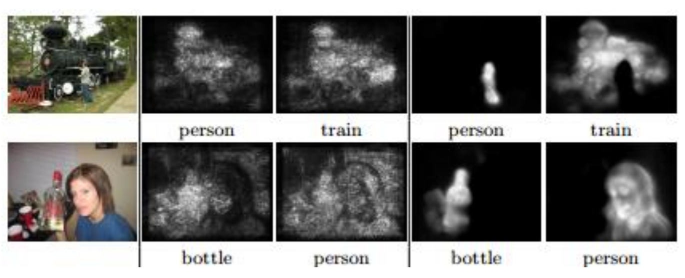
-Improved visualization of Simonyan et al.

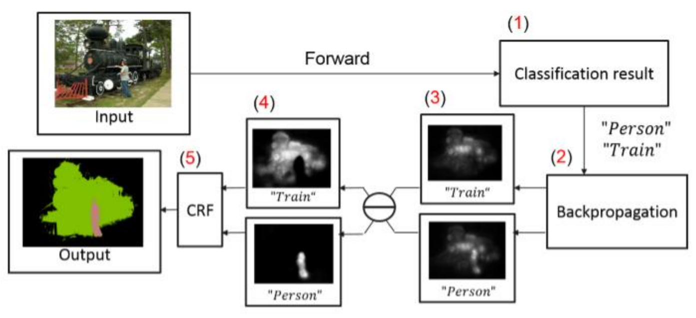
-Train with single label and multi label.

-Achieved state-of-the-art in weakly supervised segmentation.



[Simonyan et al. ICLR 2014]





[Shimoda et al. ECCV 2016]

Our method

・ Train with only single label

・ Test for multi object images.

・ Existence methods assume to train with Pascal VOC or MSCOCO which has multi label annotation.

-Most of existence datasets and web images have only single label.

Training images
-single label

Test images
-multi label

Background

Problems

・ Weakly supervised segmentation for containing multi label in training data

-Distinct class specific saliency maps

: Causes significant performance drop by training with only single label.

・ Detection and segmentation by proposal

-We can achieve detection in bottom-up approach by proposal base method.

-previous works: RCNN, SDS

: generates around 2000 candidates.

: Large computational cost.

Idea

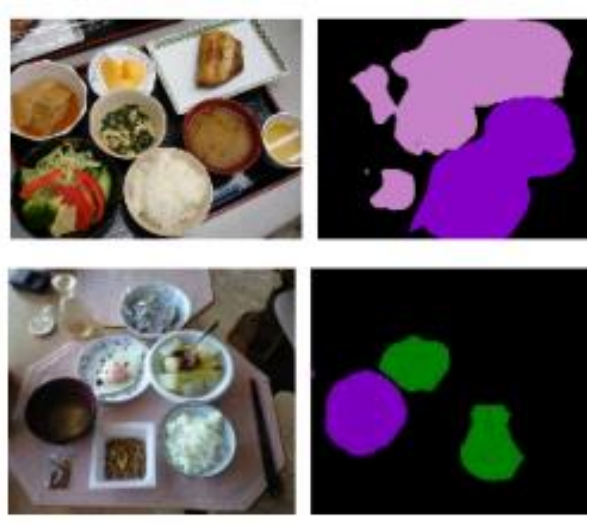
Weakly supervised segmentation results are low confidence.

However regions respond only food regions

We consider CNN could transfer only food concept.

Regard low confidence segmentation results as proposal candidates.

Combine weakly supervised segmentation and proposal base detection method.



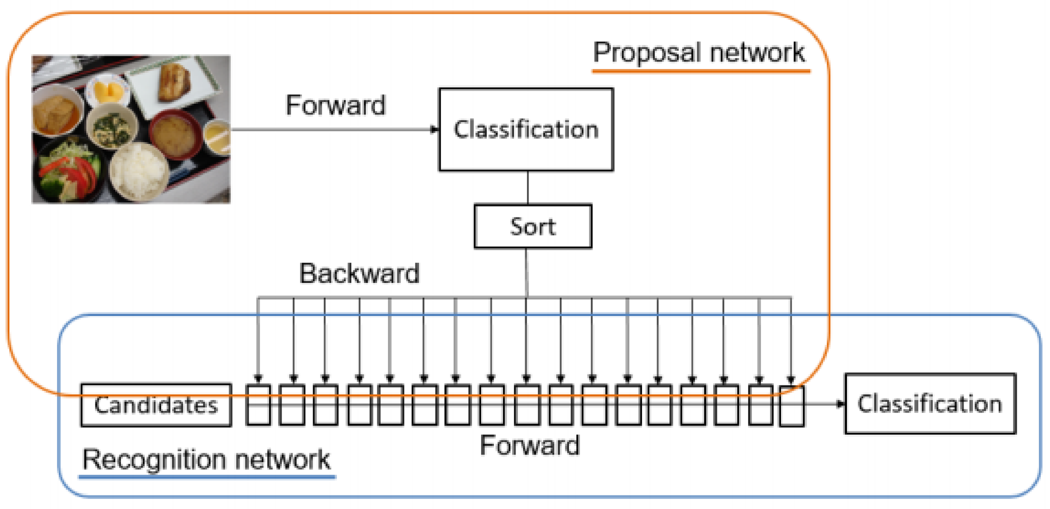
Overview

・ Sort recognition result

・ Estimate upper rank food region

・ Re-recognize estimated region

・ Unify recognition result by NMS

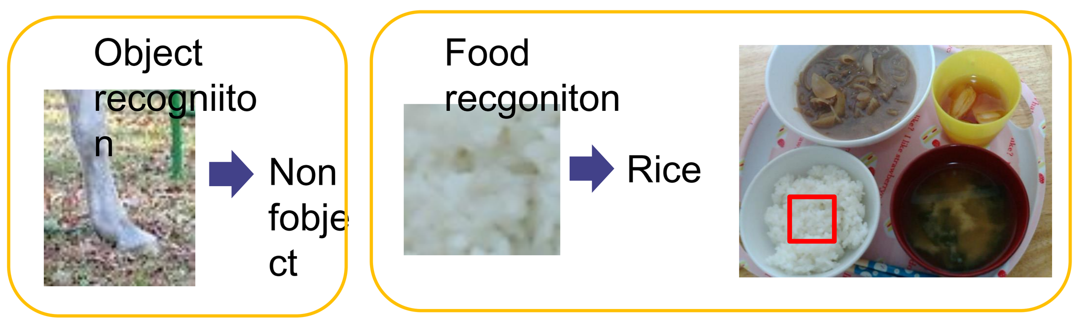


Recognition of food region candidates

Difference in object detection and food detection

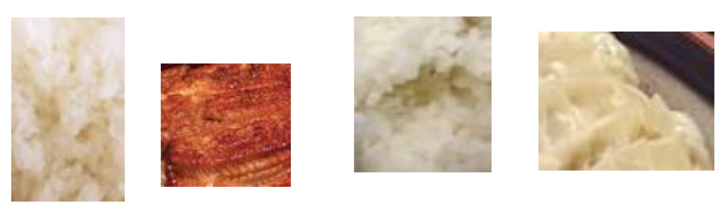
-Small region recognized as food

-Similar to texture recognition



We generate food patches by random cropping.

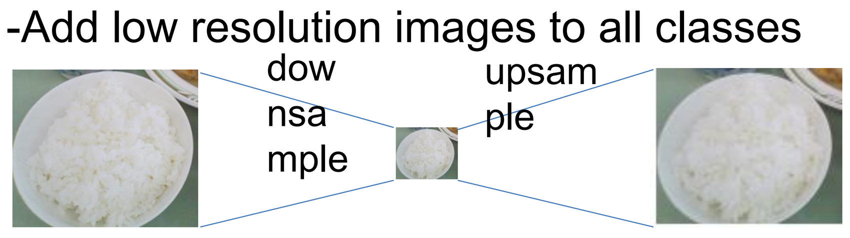
-separate food patches class from general food.



Low resolution images classified as patches

Since, We also generat low resolution images

-Add low resolution images to all classes



Experiments

Training :

・ UECFOOD 100 + Web images

-food 100 class : 1000 images for each category + non- food : 10000 images :

・ Training without bounding box and multi label.

Test

・ UECFOOD 100 multiple food dataset

- include at least one category of UECFOOD100
- Each class image number vary.
- We separate evaluation by image number.

Evaluation metric

Average precision for detection.

"100 class" is average of all class AP.

"53 class" is average of 53 class AP which class has at least 10 images.

"11 class" is average of 11 class AP which class has at least 50 images.

Detection results with different conditions

Patch images	Low resolution images	100 class	53 class	11 class
−	−	33.5	35.1	33.3
○	−	32.2	34.8	31.8
○	○	36.4	39.9	36.3

Comparison of global pooling methods

method	100 class	53 class	11 class
Average pooling	36.4	39.9	36.3
Max pooling	38.9	42.5	38.1

Comparison of other proposal methods

Method	100 class	53 class	11 class	Proposal speed [s]	recognition speed [s]
SS	38.3	39.1	35.7	7.6	35.0
MCG	33.9	43.7	33.4	2.5	35.0
Ours 10 class	33.1	33.0	33.2	0.5	1.1
Ours 20 class	36.5	40.1	37.7	1.0	2.6
Ours 30 class	38.9	42.5	38.1	1.4	3.8