### Background & Objective

Some meal recording app can estimate food calories. But they ...

- Need user’s manual input of food categories and volumes.
- Estimate food calories for each dish by one.
- Are paid service to hire nutritionists who estimate food calories.

**Purpose : Image-based food calorie estimation**

### Method

**[ Flow of Dish Detection and Calorie Estimation ]**

1. Extract bounding boxes of food dishes by Faster R-CNN.
2. Crop image corresponding to each bounding box.
3. Give food calorie estimation network these cropped images one by one.

**1. Faster R-CNN : S. Ren et al. [1] 2015**

- High-speed and highly accurate detection system.
- End-to-end learning of the whole system is possible.
- Consists of two modules.

**Fast R-CNN detector**

Simultaneous estimation of classification and bounding boxes. Convolution once for the entire image.

- Region Proposal Network (RPN)
- CNN-based region proposal method
  - Share a conv layers with Fast R-CNN detector.

**We use Faster R-CNN as a food detector to detect each dish in a food image.**


- Image-based food calorie estimation with CNN.
- Regression-based method.
- Output food calories directly from single-dish food photos.

- We denote \( l_{re} \) as an relative error and \( l_{ab} \) as a absolute error, \( L_{cal} \) is defined as follows.

\[
L_{cal} = \lambda_{re} l_{re} + \lambda_{ab} l_{ab}
\]

\( l_{re} = \frac{| y - g |}{g} \quad l_{ab} = | y - g | \)

\( y \) is an estimated food calorie.

\( g \) is ground-truth.

- The food calorie of each detected dish are estimated by image-based food calorie estimation.

### Experiment 1 : Food detection

**DATASET**

- Two kinds of bounding box annotated food photo datasets.
  - UEC Food-100 [4]
  - Japanese school lunch photos collected from online school lunch sites.

**[ Food Detection 1 (on UEC Food-100) ]**

- Comparison to an exist work of Shimoda et al. [5] .
- We use mean Average Precision (mAP) of PASCAL VOC detection task for the evaluation.
- 11,566 single-dish photos for training, 1,174 multiple-dish photos for evaluation.

<table>
<thead>
<tr>
<th>Method</th>
<th>mAP (%)</th>
<th>10 class (all)</th>
<th>15 class (all)</th>
<th>15 class (test images 1:10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-CNN</td>
<td>26.0</td>
<td>21.8</td>
<td>25.7</td>
<td></td>
</tr>
<tr>
<td>[5]'s method</td>
<td>49.9</td>
<td>55.3</td>
<td>55.4</td>
<td></td>
</tr>
<tr>
<td>Faster R-CNN</td>
<td>42.0</td>
<td>46.3</td>
<td>57.9</td>
<td></td>
</tr>
</tbody>
</table>

### Experiment 2 : Food calorie estimation of multiple dishes

**DATASET**

- Bounding box annotated japanese school lunch photos for training of Faster R-CNN [1].
- Total food calorie annotated Japanese school lunch photos for the evaluation.

We also collected estimated school lunch photos for the evaluation.

Each image has a total calorie value of all the dishes.

- We fixed the calorie of "Milk" detected by Faster R-CNN to 134 kcal.
- 690 total food calorie annotated Japanese school lunch photos for the evaluation.

- The numbers in bounding boxes are estimated food calories of foods in each bounding box (kcal).
- ES : the estimated total food calorie (kcal).
- GT : the ground-truth of total food calorie (kcal).

### Conclusion & Future work

- Food detection by Faster R-CNN.
- We collected school lunch photo dataset by Web image mining.
- We estimate food calories from multiple-dishes food photos.
- Multi-task learning of food detection and food calorie estimation.
- Construction of large-scale food photos dataset.