# An Automatic Calorie Estimation System of Food Images on a Smartphone

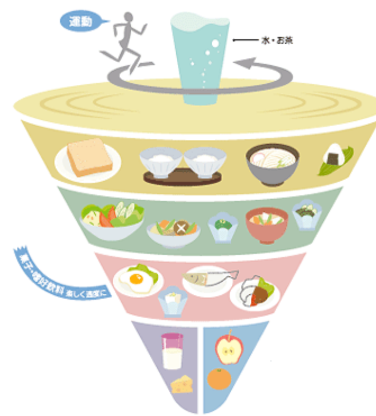
Koichi Okamoto and Keiji Yanai

The University of Electro-Communications, Tokyo

# Introduction (1)

- Recording everyday meals is important.

| Dietary health control | Nutritional study |
| --- | --- |
| e.g. Foodlog | |


# Introduction (2)

- Smartphone-based Food logging services

## Foodlog


Need to teach the amount of foods by selecting items

Food category recognition and manual vol. estimation

## CaloNavi


Dietary advices by professional nutrition

human cost → Pay service

# Objective

- Automatic calorie estimation system:

  **New system: CalorieCam**

- Standalone mobile system

  - very rapid recognition (less than one second)

  - all the processing inside a smartphone

We focus on calorie intakes rather than carbohydrate .

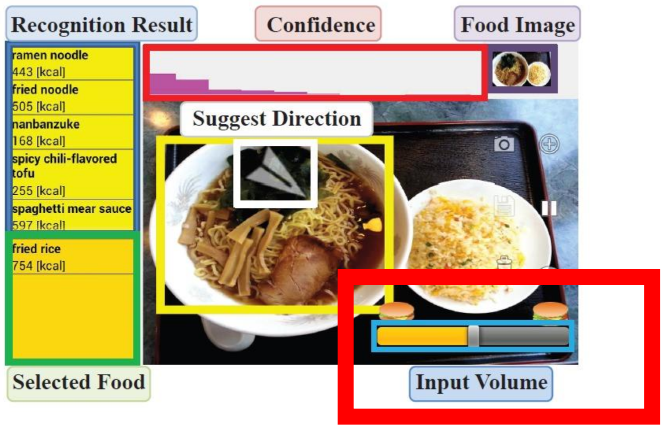
**For better usability**

*e.g. user can give up eating*

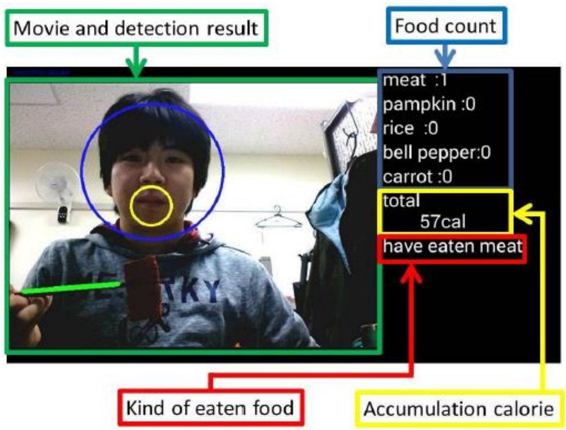
| dishes | Calorie Value |
|---|---|
| Pork cutlet | 528kcal |
| Rice | 172kcal |
| Miso soup | 43kcal |
| Salad | 27kcal |

# Related works (1)

- Some food recognition systems can estimate food calories in the simplified way.
  - Indicating the amount of food manually
  - Counting the number of the food pieces




**FoodCam**
**(Kawano et al. 2014 )**




**GrillCam**
**(Okamoto et al. 2015)**

# Related works (2)

- Several work on automatic food volume estimation
  - 3D volume reconstruction by multiple images
  - 3D volume estimation by CNN
  - Using the size-known reference objects

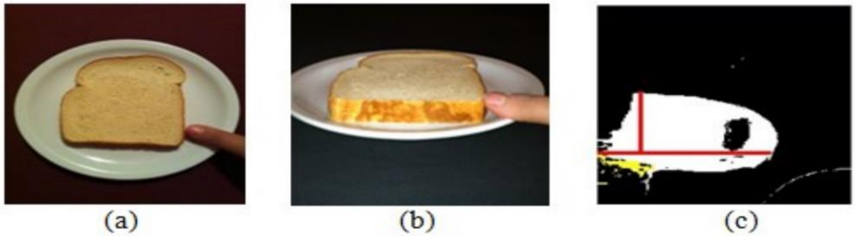
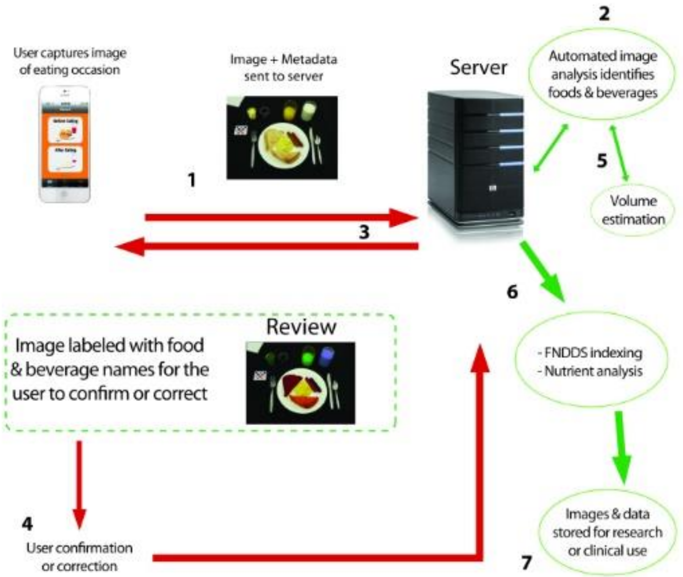


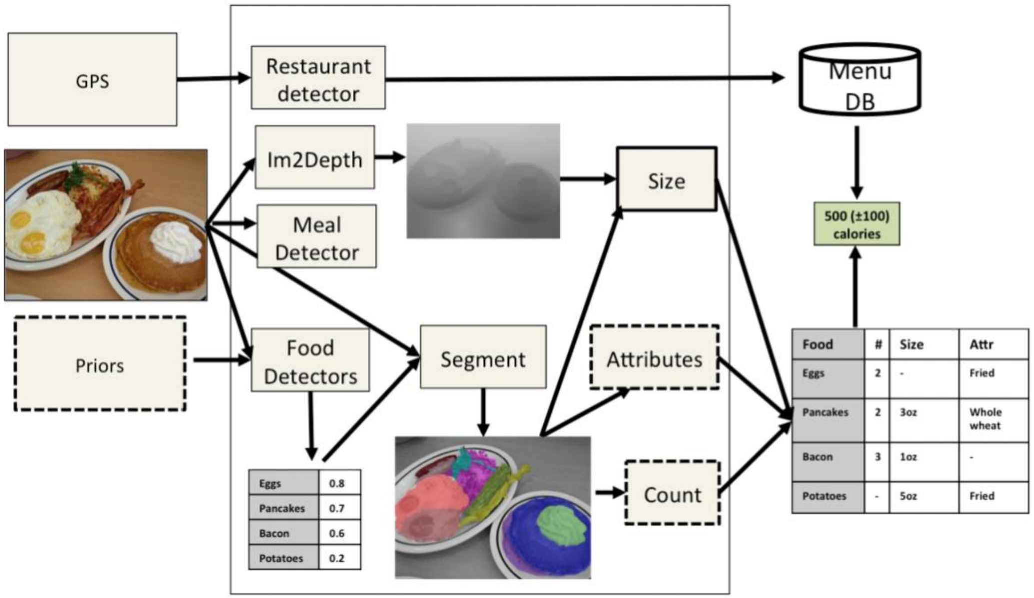

**Figure 2 (a, b) Test images with food and thumb (c) Calculation of the thumb dimensions**




*1 P. Pouladzadeh, S. Shirmohammadi, and R. Almaghrabi. Measuring calorie and nutrition from food image. IEEE Transactions. on Instrumentation and Measurement, Vol. 63, No. 8, pp. 1947–1956, 2014.
*2 The TADAProject www.tadaproject.org

# Related works (3)

- Im2Calories [Myers et.al. 2015]
  - All the processing inside a smartphone
  - CNN-based methods (categorization, 3D vol. est.)
  - Not yet released as a mobile application
    - Just presented some ideas for image-based calorie estimation.



# Design of the proposed system

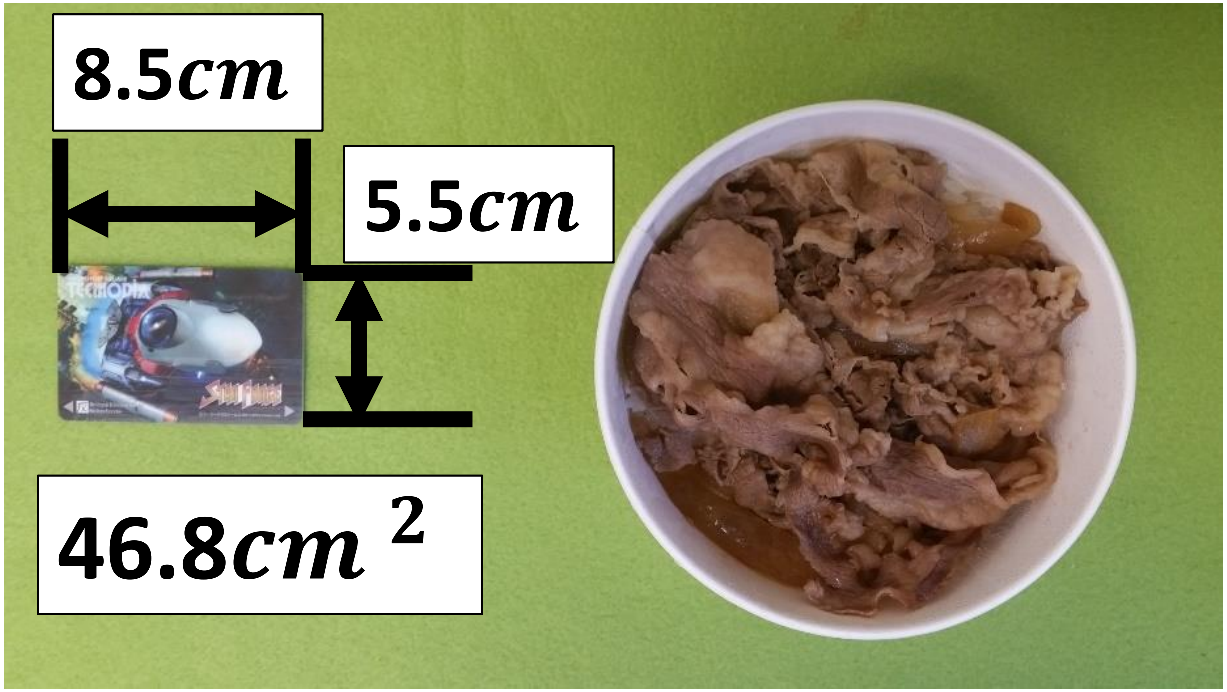
- All the processing inside a smartphone
  - Very quick calorie estimation from a food image

- Simple 2D-size-based calorie estimation

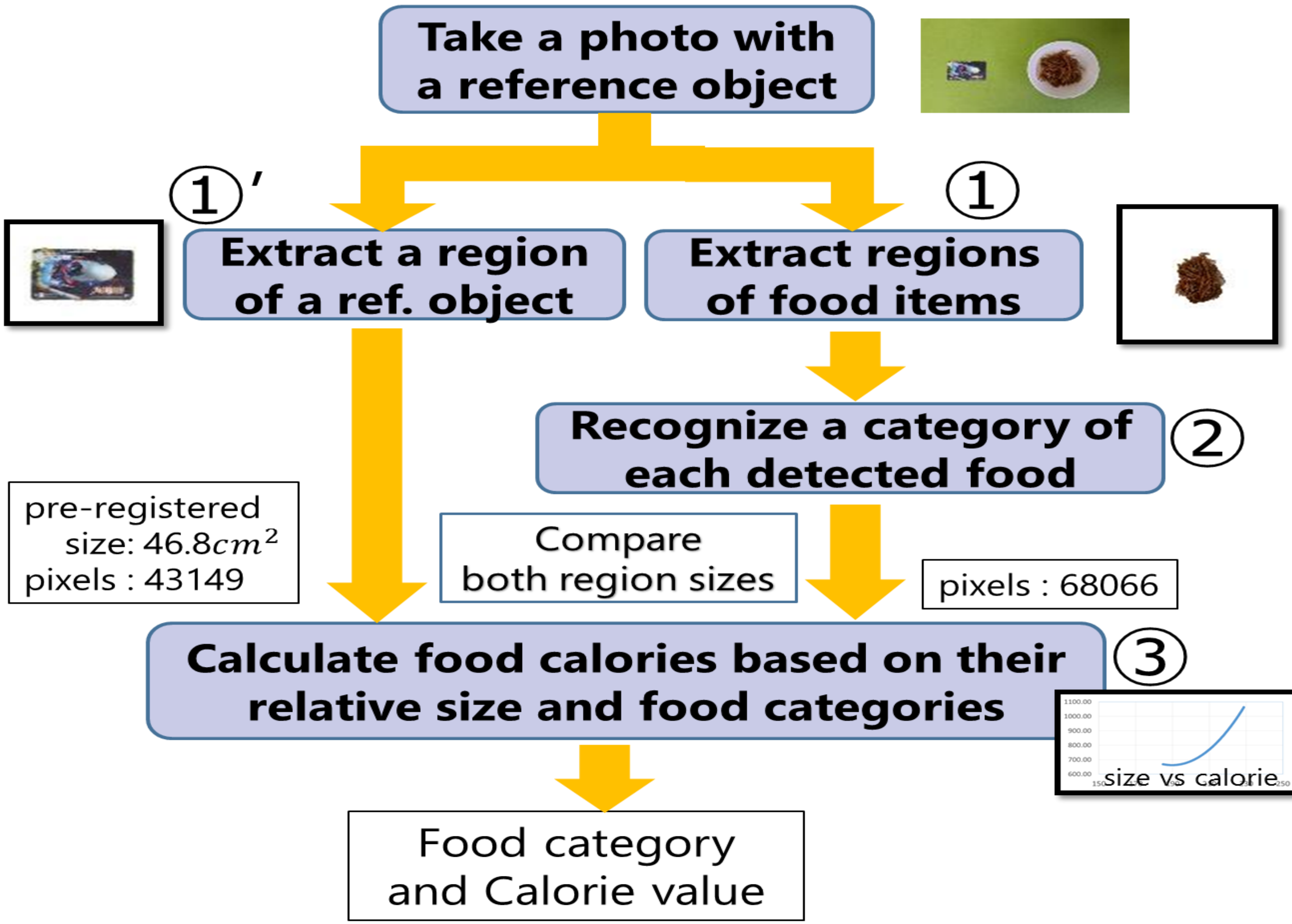
- Use CNN only for food categorization
  - Use the “DeepFoodCam” food recognition engine, while use conventional methods for segmentation.

# Conditions to be assumed

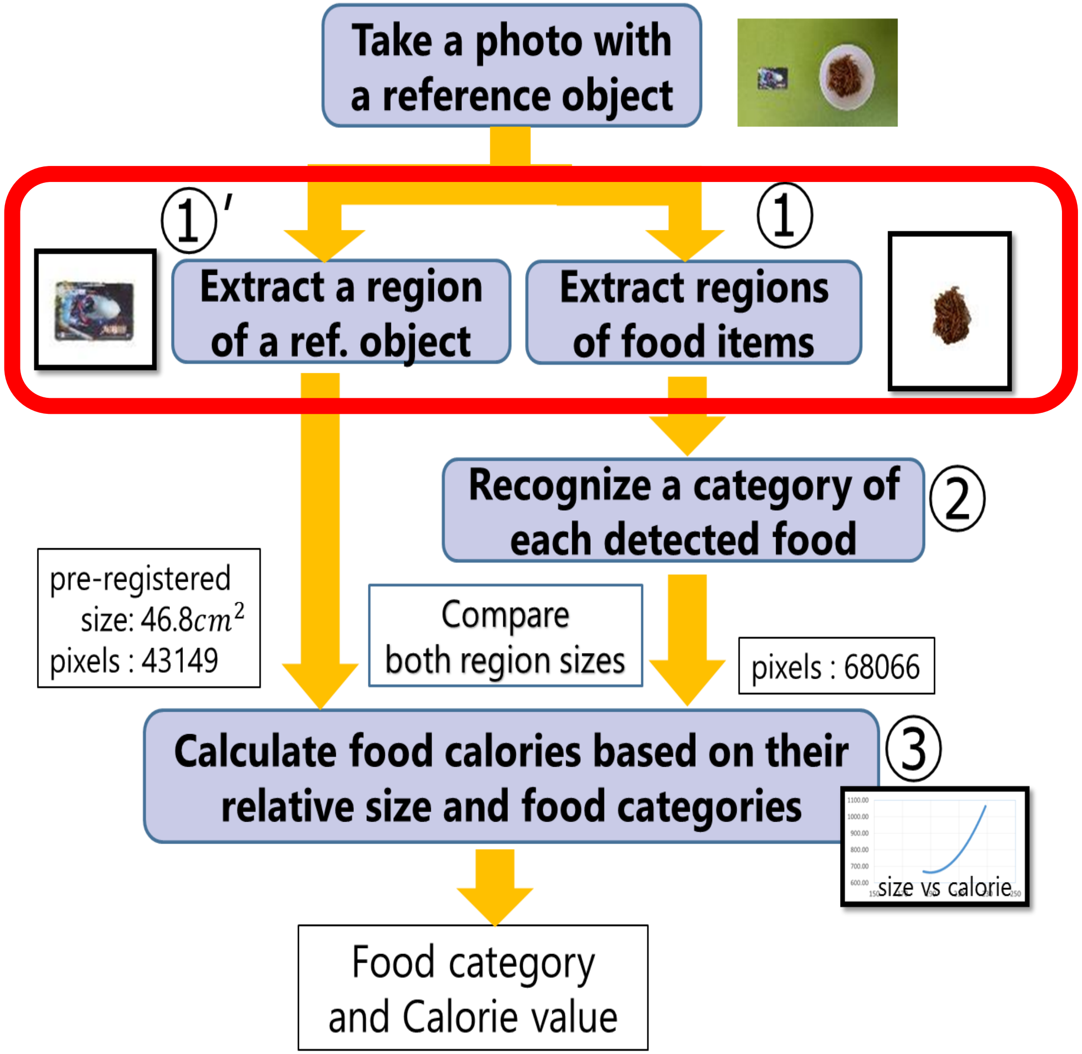
- Uniform back-ground
- Size-known reference obj.
- Taken from top (top-view photo)



# Processing flow



# Step 1: region extraction



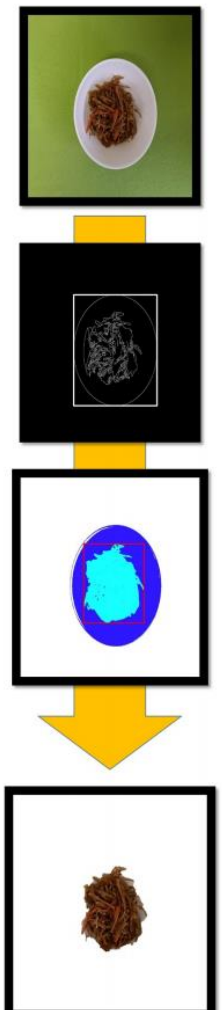

# Dish region detection

Provide a meal photo

Detect a bounding box of
a dish region based on edges

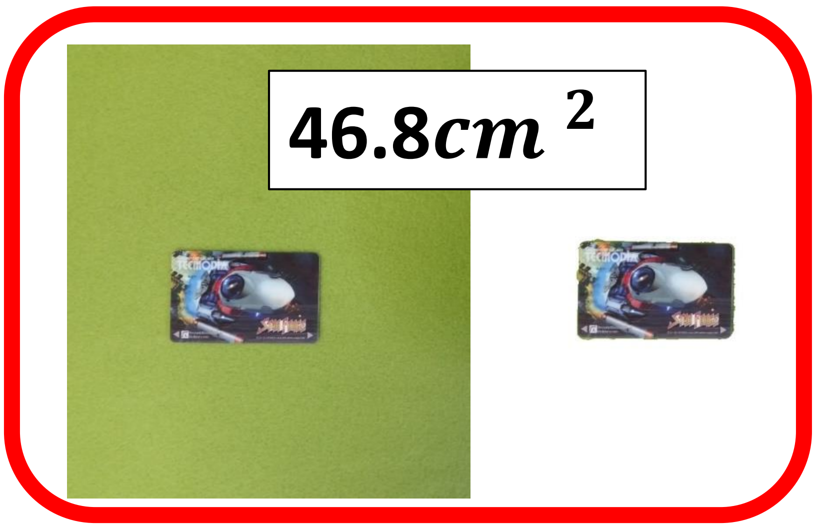
Detect a bounding box of
a food region by k-means

Detect an accurate food
region by GrubCut

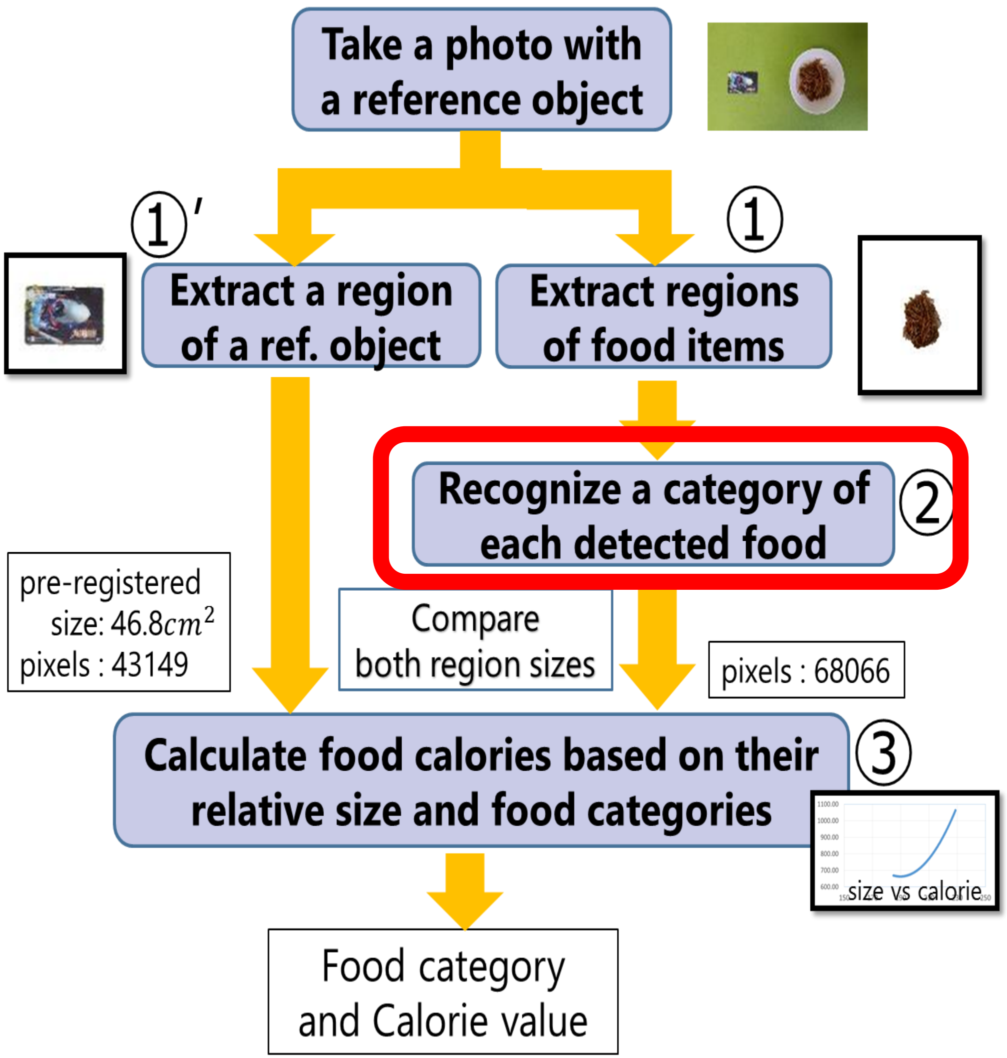


# Extraction of a given reference

- In the same way as food region extraction
- Any shape is possible, since only the size is important for estimating actual size of foods.




# Step 2: food classification

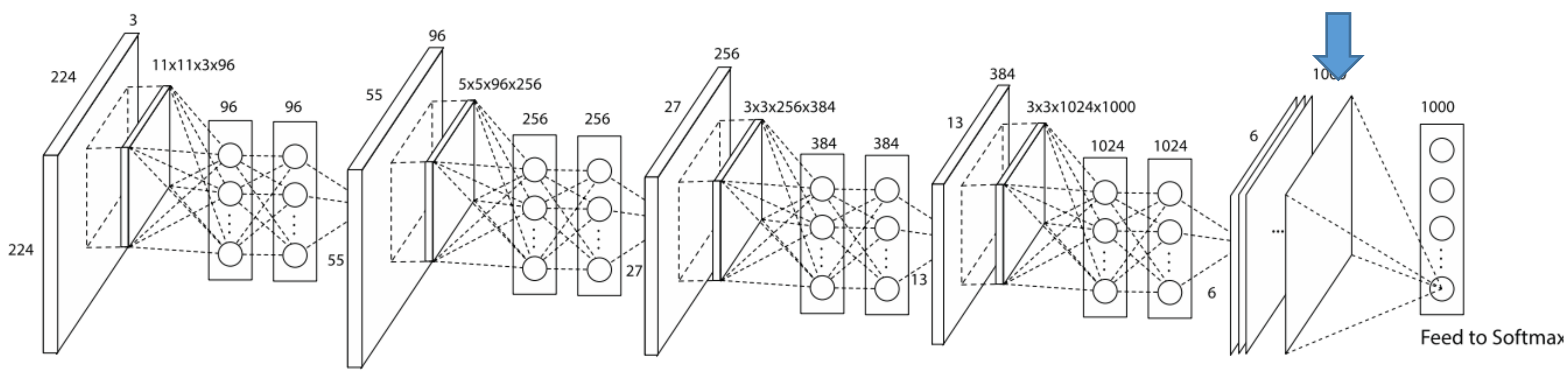



``Kinpira Burdock''

# CNN-based mobile food recognition engine (1)

- Use the “DeepFoodCam” mobile food rec. engine
- CNN: Network in Network(NIN)
  - No fully connected layer
  - AlexNet 60million ⇔ NIN 7.5million

**Global average pooling**




*1 L. Min, C. Qiang, and Y. Shuicheng. Network in network.
In Proc. of International Conference on Learning Represenation Conference Track, 2013.

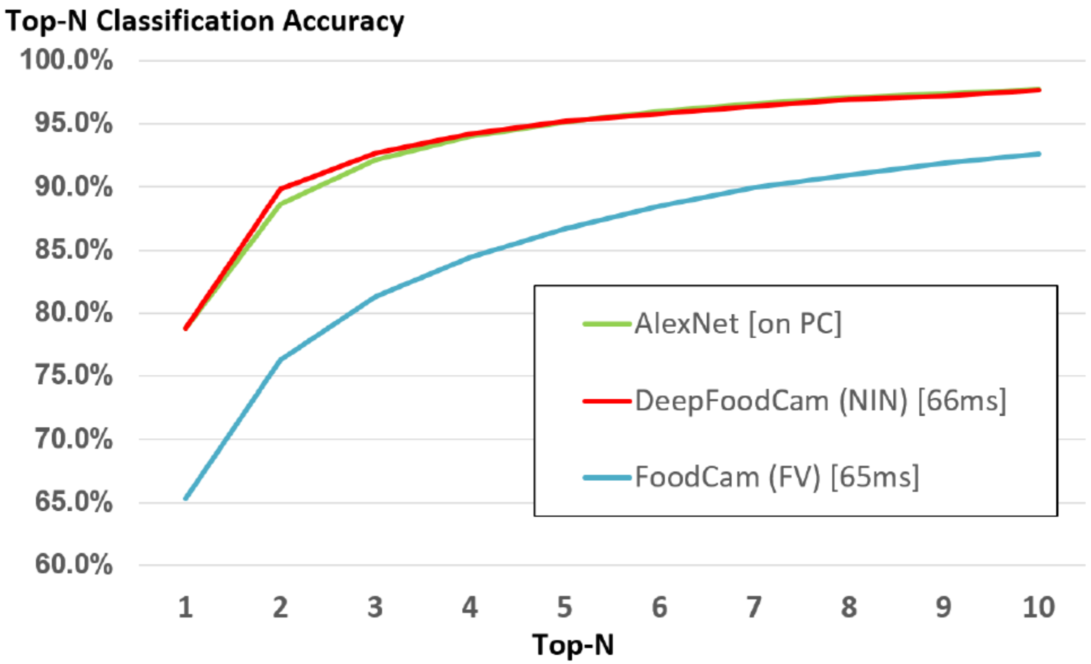
# CNN-based standalone mobile food recognition engine (2)

- Pre-training: ILSVRC1000 classes + 1000 food-related classes in ImageNet
- Finetuning: UECFOOD-100
  (100 kinds of Japanese foods)

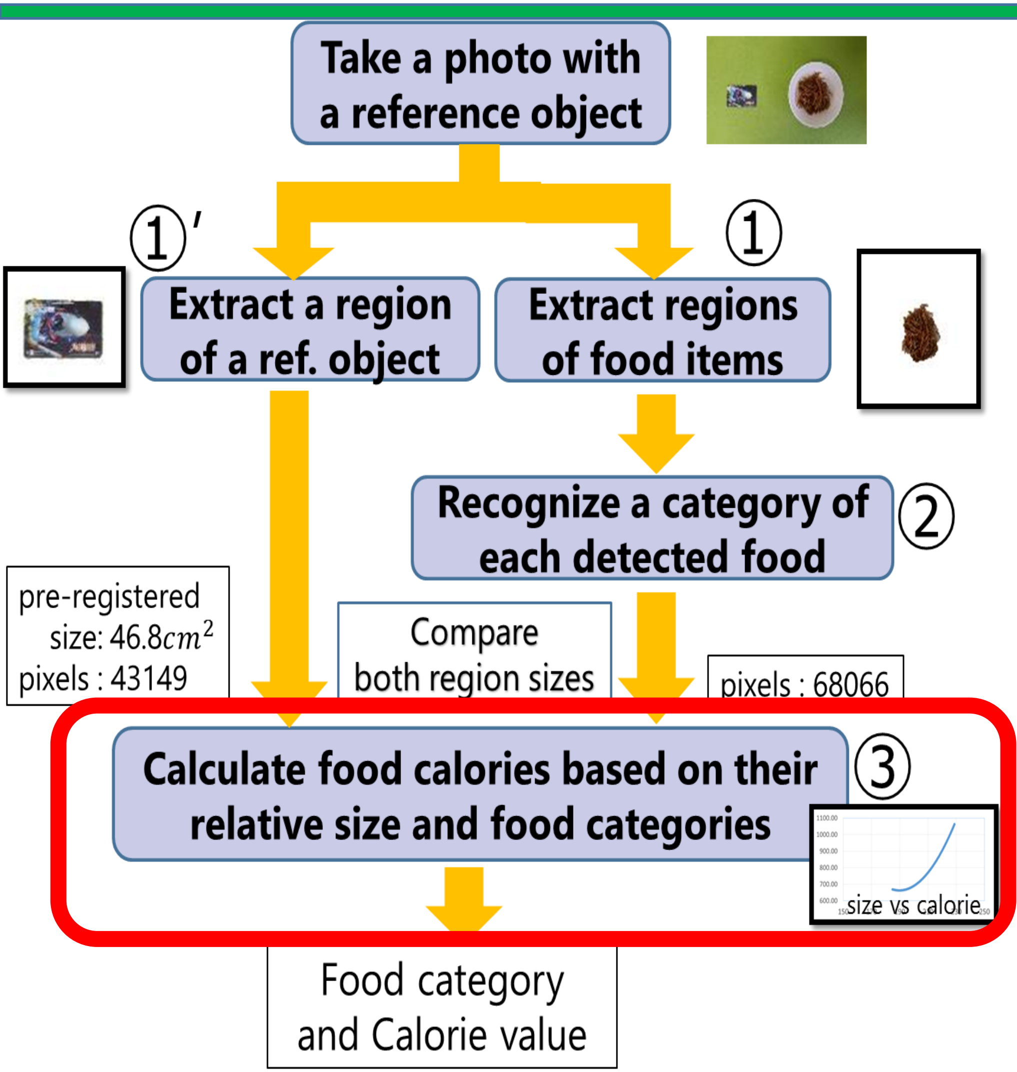
- Performance:
  78.8% (top1)
  95.2% (top5)
- Time:

  55.7ms / img
  (227x227) (iPhone7+)

  26.2ms / img
  (160x160) (iPhone7+)



# Step 3: food calories estimation




# Estimation of real size of foods

- 2D-Size-based calorie estimation


**Reference Object**
**#pixel: 43149**
**Real size: 46.8$cm^2$(known)**

**"Kinpira burdock"**
**#pixel: 68066**
**Real size: ?? $cm^2$**

$$size_{food} = size_{ref}/\#pix_{ref} * \#pix_{food}$$

→ **78.82 $cm^2$**

# Quadratic curve based calorie estimation

“Kinpira burdock”
#pixel: 68066
Real size: **78.8 $cm^2$**
**a=0.001,b=0.5,c=37**

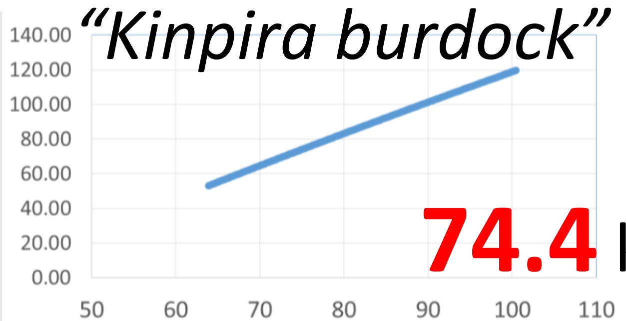
``Beef bowl’’
#pixel: 21043
Real size : **208$cm^2$**
**a=1.8,b=0.4,c=190**

Use quadratic curve based estimation

$$cal = a_i * size_{food}^2 + b_i * size_{food} + c_i$$

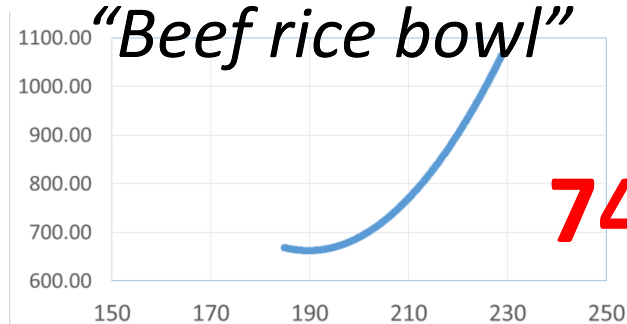
where $a_i, b_i, c_i$ are trained params for $i$-th categories.

*“Kinpira burdock”*




**74.4** kcal

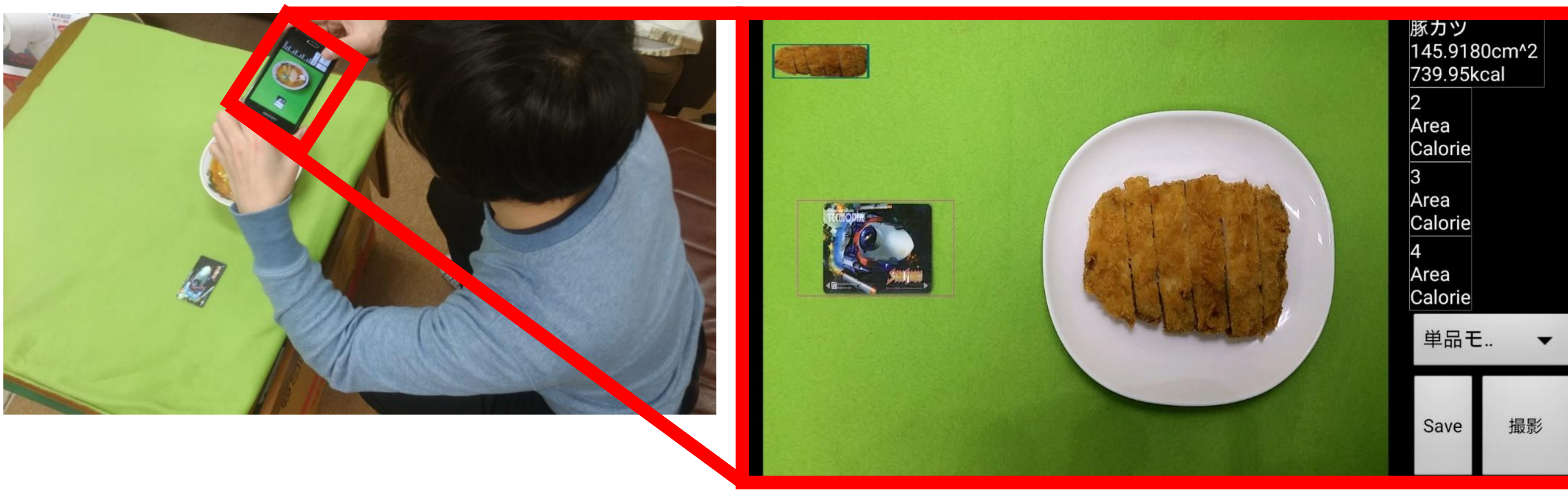
*“Beef rice bowl”*




**744.0** kcal

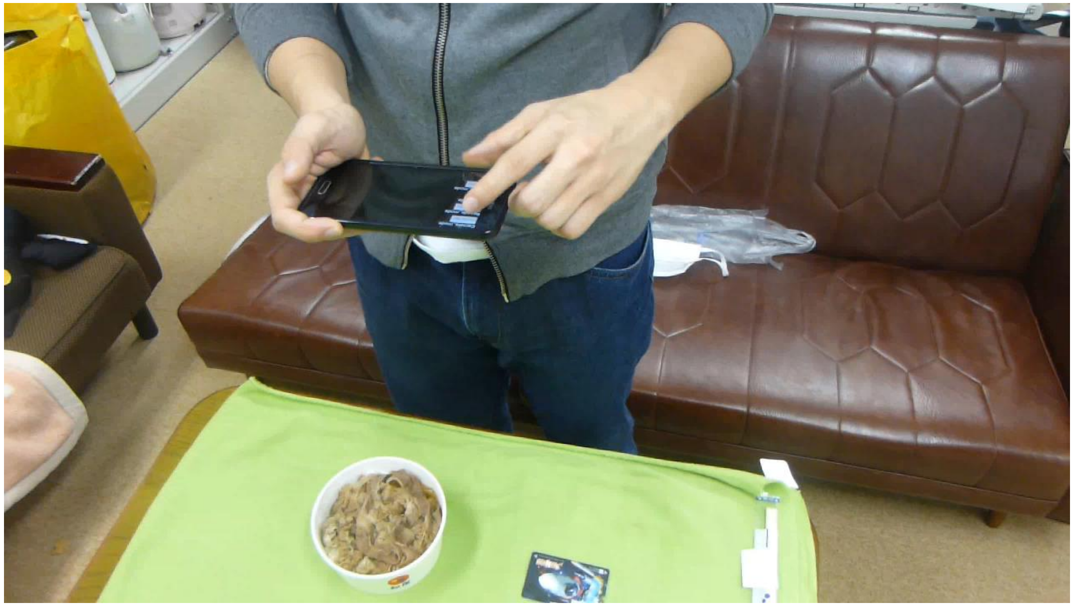
# User Interface

- Implemented the proposed system as Android app.
- Took only less than 1 seconds for one food image.




**Download (only Japanese version)**
**http://foodcam.mobi/calorie/**

# How to use the system



**The processing time is less than 1 second !**

# Experiments

- Evaluation on calorie estimation accuracy
  - on PC

- User Study
  - using Android app.

# Calorie estimation

- Dataset : 120 images (60 for training, 60 for eval)
  - 20 kinds of Japanese dishes with 3 different sizes
  - Prepare all the dishes and take photos in our lab
  - Use for training of calorie estimation parameters
    ($a_i$, $b_i$ , $c_i$ : quadratic curve parameters)


Chill shrimp 268kcal


French fries 454kcal

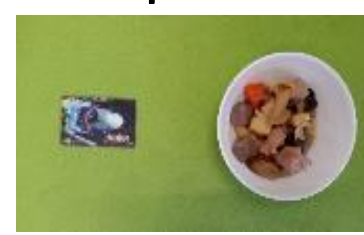

Spring rolls 428kcal


Sweet pork 292kcal


Takoyaki 241kcal


Gyoza 246kcal


Yakitori 165kcal


Boiled chikuzen 85kcal

# Evaluation on Calorie Estimation

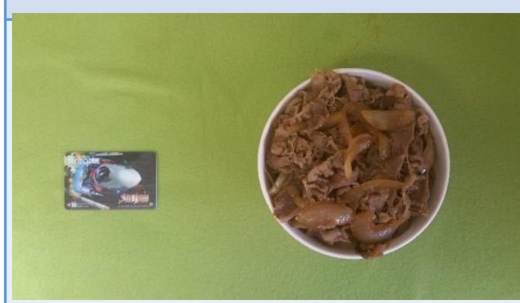
20 dishes for 3 images for each （totally 60 images）

| Mean avg err. | Mean SD | Mean relative err. | Mean relative SD |
|---|---|---|---|
| **52.2** kcal | ± **40.4** kcal | **21.3%** | ± **0.82** |

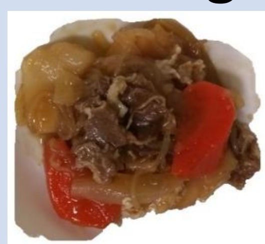
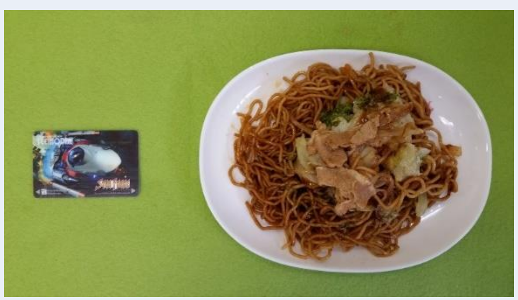
Error: average absolute value of the difference

Relative error : average relative value of the difference

# The cases of good estimation

| Input image | dish region | GT | Estimat ion. | Error | Ralative err. |
|---|---|---|---|---|---|
|  | Pork cutlet  | 586 kcal | 559 kcal | 27 kcal | 0.05 |
|  | Beef bowl  | 1322 kcal | 1417 kcal | 95 kcal | 0.07 |

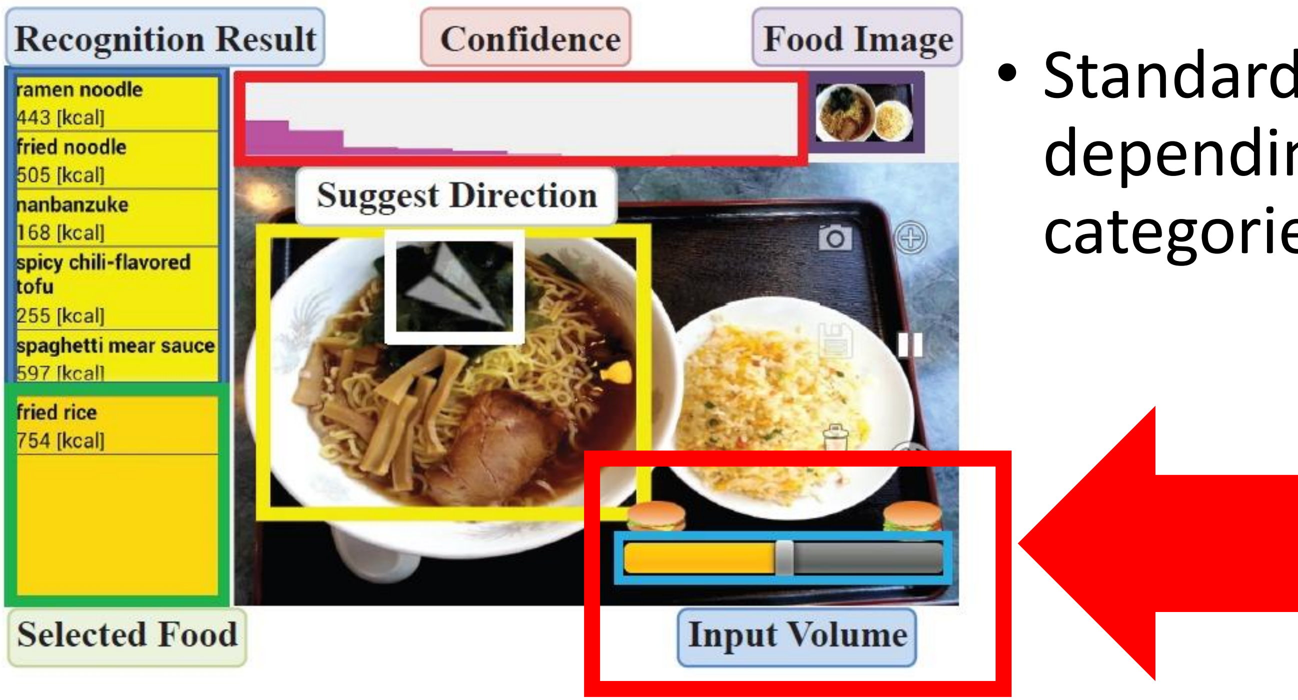
# The cases of bad estimation

| Input image | dish region | GT | Estimat ion. | Error | Ralative err. |
|---|---|---|---|---|---|
|  | **Niku Jaga**  | **170** kcal | **122** kcal | **48** kcal | **0.28** |
|  | **Yakisoba**  | **425** kcal | **519** kcal | **94** kcal | **0.22** |

# User Study

- Subjects: **12** students who have no knowledge on nutrition
- Task: estimate food calories with two systems:
  FoodCam (baseline) and CalorieCam (proposed)
- Target foods : beef rice bowl, croquette, salad


# Baseline system:
# FoodCam (Kawano et al. 2015)




- Standard calorie amount depending on the dish categories

**Manual volume selector**

# Estimation by users

| | | FoodCam | | Proposed | |
|---|---|---|---|---|---|
| Dish | GT | Avg. err | Avg.SD | Avg. err | Avg.SD |
| **Beef bowl** | **962** | **-53.25** | **±209.79** | **-242** | **±55.10** |
| **Croquette** | **552** | **-242** | **±91.26** | **-47.08** | **±52.52** |
| **Salad** | **14** | **54.83** | **±36.28** | **4.86** | **±11.87** |

The average errors were reduced except for a beef bowl.

The standard diviation (SD) were also reduced, which means the proposed app achived more stable estimation than FC.

# Evaluation of usability by users

- 5-step evaluation on usability of the system

| | Usability (5-step) |
|---|---|
| FoodCam | **2.83 ± 0.80** |
| Proposed (CalorieCam) | **4.25 ± 0.72** |

- CalorieCam was much simpler, since it is an automatic calorie estimation system. (Foodcam is a manual system).

# Conclusions

- We proposed an automatic calorie estimation mobile system implemented as an Android app which enable
  calorie estimation within one second.
- The system achieved automatic region extraction of dishes and a reference object.
- Avg. error **52**kcal, relative error **20%** for 20 kinds of foods
- Higher usability than the baseline manual system.

# Future work

- CNN-based region segmentation (+3D volume estimation)
  - For complicated background
- Add more food category
  - It is not practical to prepare and take photos of hundreds kinds of foods.
  - Use calorie-annotated recipe data on the Web