

1st International Workshop on  
Multimedia Assisted Dietary Management (MADiMa)

Highly Accurate Food/Non-Food Image Classification  
based on a Deep Convolutional Neural Network

---

6<sup>th</sup>, September 2015

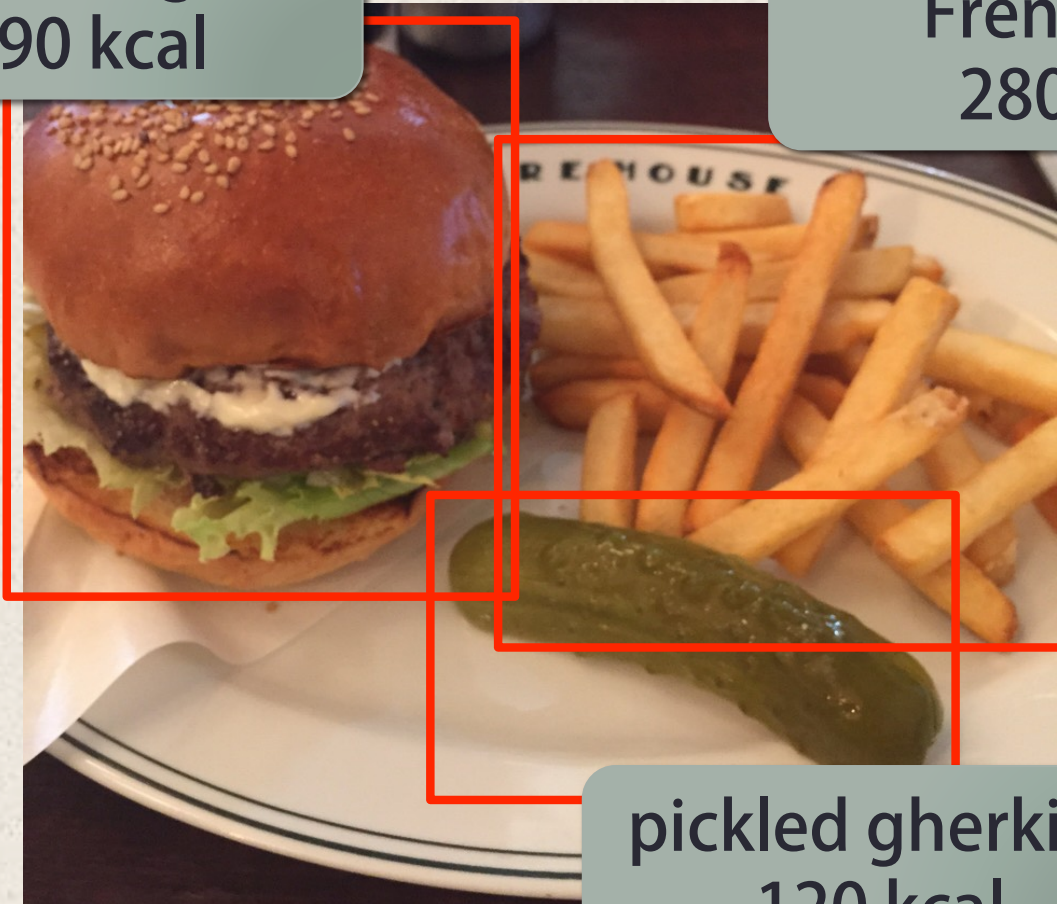
Hokuto Kagaya and Kiyoharu Aizawa

The University of Tokyo

# Food Image Analysis

Hamburger  
290 kcal

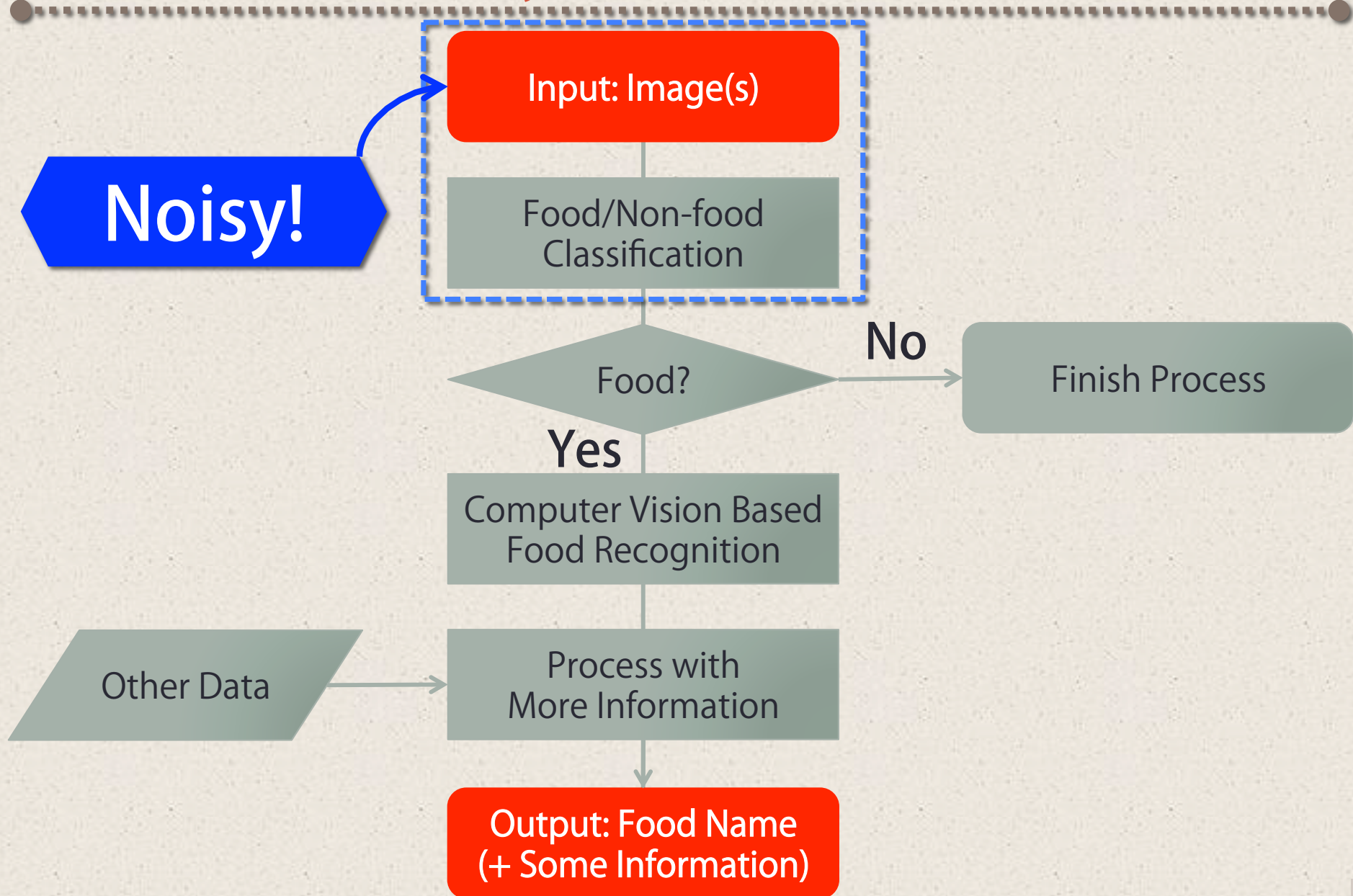
French Fry  
280 kcal



pickled gherkins  
120 kcal

\* Displayed calories are not real values.

# Food Analysis Framework

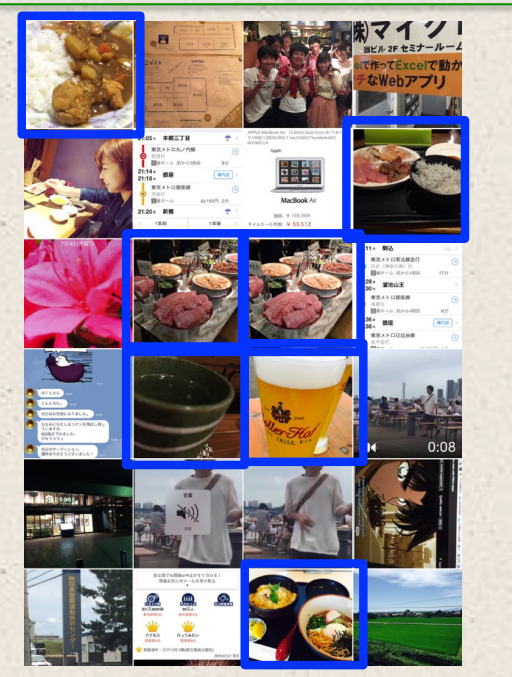




# Noisy Input Images

- Case 1: Personal photo albums
  - Personal photos is usually not arranged
- We want our photos in personal albums to be classified food or non-food automatically.

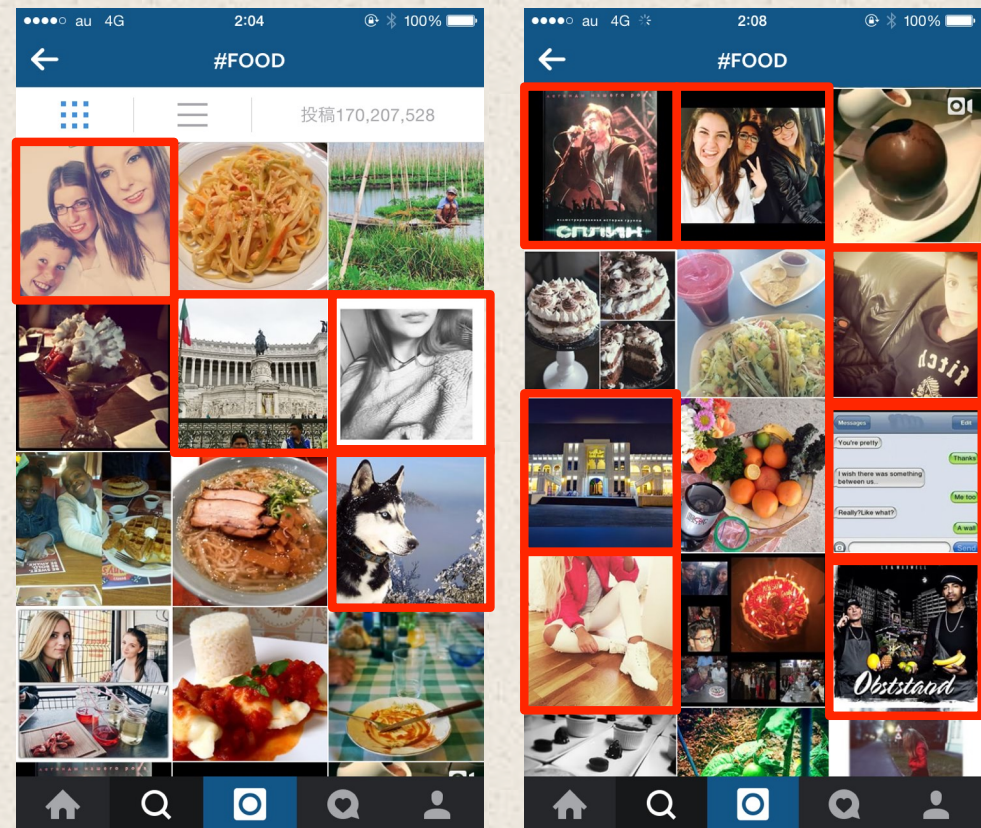
My photo album in iPhone



# Noisy Input Images

- Case 2: SNS photos
  - Hash tags are very confusing

The search results of "#food" in Instagram





# Food Image Analysis

- Health Care, Lifelog, Entertainment (e.g. SNS)
- FoodLog App<sup>[1]</sup>
  - Assist food recording using image processing



[1] <http://app.foodlog.jp/>

# Food Image Analysis

- Google made a statement about AI to count calories in food porno pictures (May, 2015)

The image shows a collage of news articles and advertisements. At the top left is a CNN Money header with navigation links for Business, Markets, Tech, and Luxury, and a stock tickers search bar. Below it is a sub-header 'Innovate' and the headline 'Google says it can count calories in a picture'. To the right is a CNET article titled 'Google cooking up AI to count calories in your food photos' by Danny Gallagher, dated June 1, 2015. The article includes a grid of food photos (burgers, Swirlers ice cream, cookies, peach cobbler, and a dessert with sticks) and a social media sharing bar. Below the article is a GMC advertisement with the slogan 'PRECISION MATTERS' and 'WE ARE PROFESSIONAL GRADE'. At the bottom right is a 'THIS WEEK'S MUST READS' section featuring the article 'Google cooking up AI to count calories in your food photos' as item 1.

CNN Money Business Markets Tech Luxury stock tickers

Innovate

Google says it can count calories in a picture

CNET · Sci-Tech · Google cooking up AI to count calories in your food photos

## Google cooking up AI to count calories in your food photos

A Google research scientist spills the beans on an artificial-intelligence project that tries to calculate the calories in your food porn pictures – and learns from its mistakes.

by **Danny Gallagher** @thisisdannyg / June 1, 2015 1:54 PM PDT

PRECISION MATTERS EXPAND

GMC WE ARE PROFESSIONAL GRADE

THIS WEEK'S MUST READS /

- 1 Google cooking up AI to count calories in your food photos

<http://money.cnn.com/2015/06/03/technology/google-calories-food/>

<http://www.cnet.com/news/google-working-on-ai-that-can-count-calories-in-your-food-photos/>



# Related Work

---

- Food item recognition from one image
  - Fusing general appearance features<sup>[Hoashi+, 2010]</sup>
    - Dataset : Japanese foods (collected from web)
    - Used SIFT, Gabor feature, Color feature (MKL-SVM)
  - Original feature specific to food images<sup>[Yang+, 2010]</sup>
    - Dataset : General USA fast foods (PFID)
    - “pair-wise local features”
  - Deep Learning <sup>[Kagaya+, 2014](ours)</sup>
    - Dataset: FoodLog Dataset
    - Compared to existing techniques, 10-15% better performance.



# Related Work

---

- Food item recognition from one image
  - Fusing general appearance features<sup>[Hoashi+, 2010]</sup>
    - Dataset : Japanese foods (collected from web)
    - Used SIFT, Gabor feature, Color feature (MKL-SVM)
  - Original feature specific to food images<sup>[Yang+, 2010]</sup>
    - Dataset : General USA fast foods (PFID)
    - “pair-wise local features”
  - Deep Learning <sup>[Kagaya+, 2014](ours)</sup>
    - Dataset: FoodLog Dataset
    - Compared to existing techniques, 10-15% better performance.

[2] H. Hoashi et al., “Image recognition of 85 food categories by feature fusion”, IEEE ISM, 2010.

[3] S. Yang et al., “Food recognition using statistics of pairwise local features”, CVPR, 2010.

[4] H. Kagaya et al., “Food Detection and Recognition Using Convolutional Neural Network”, ACM MM 2014.

# Related Work

---

- More and more food image dataset
  - PFID [Chen+, 2009]
  - Food-101 [Bossard+, 2014]
    - 101 classes \* 1,000 images from web
  - UEC Food-256 [Kawano+, 2014]
    - 256 classes \* 101~729 images from web
  - FoodLog Dataset (ours) [Kagaya+, 2014]



# FoodLog App

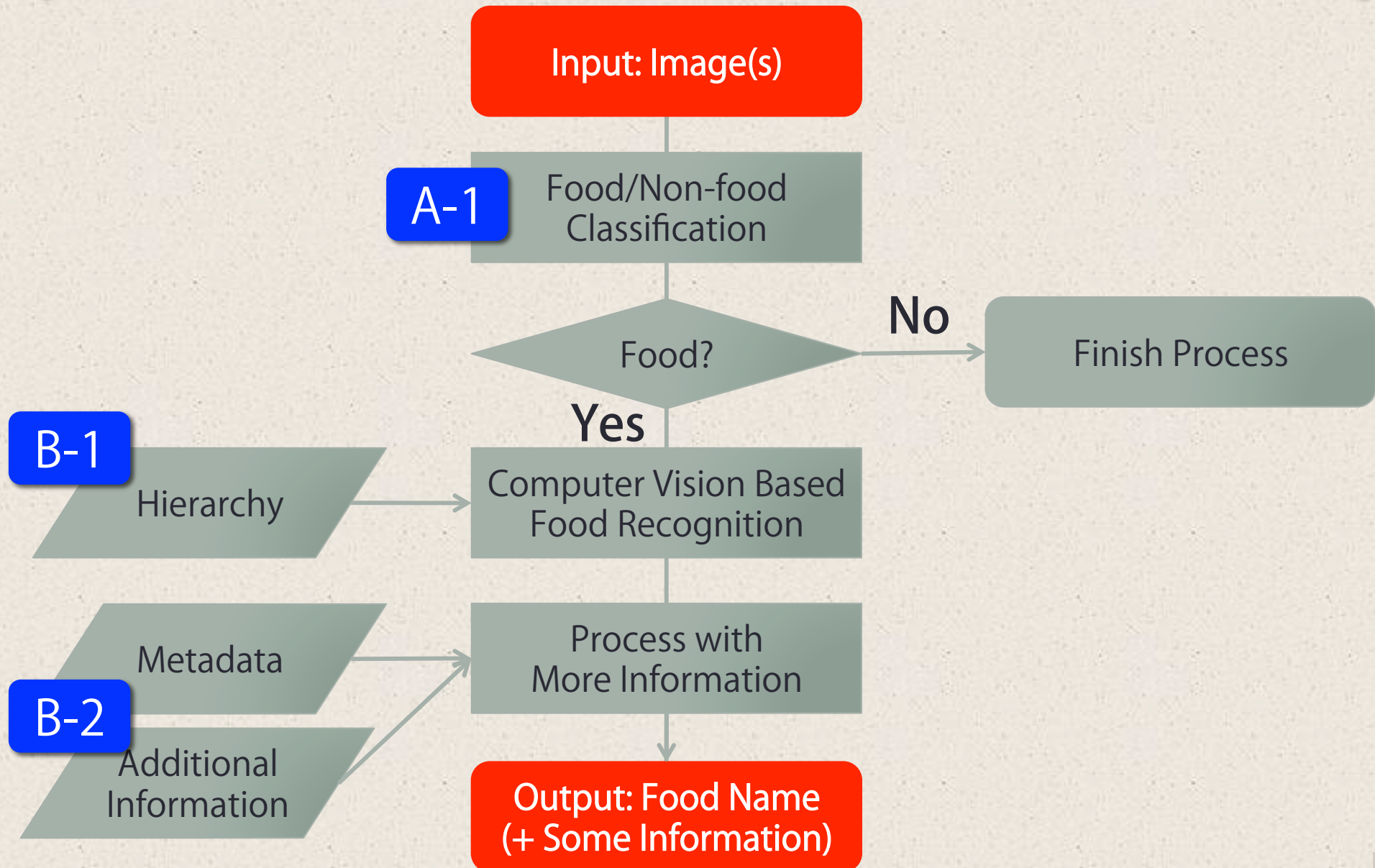
- Smartphone application for food recording helped using image processing<sup>[1]</sup>
  - Available in iOS and Android
  - Image retrieval from user's past food record using color feature with spatial pyramid



- ✓ Personal trends is very important
- ✗ Many manual processes



# Proposed Framework





# Problem and Contribution

---

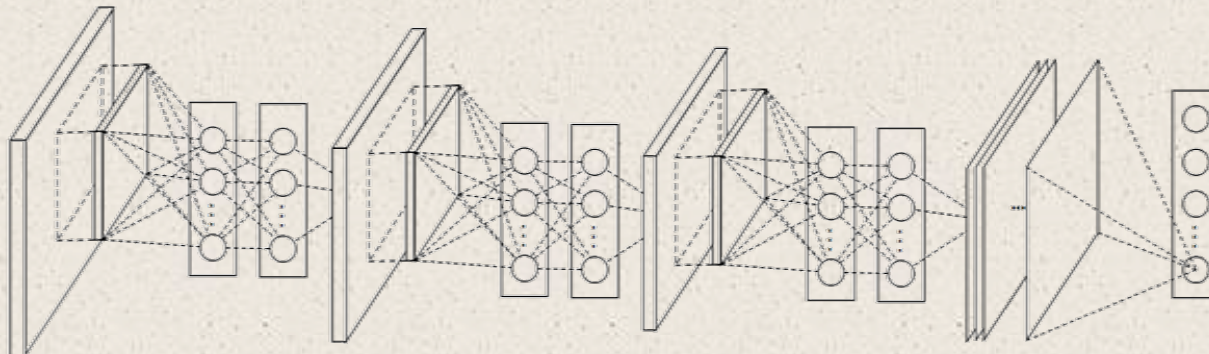
- Food/Non-Food Classification
  - judges whether an image is a food or not

## Contribution

- Build two novel datasets
- Evaluate the performance of a deep CNN for food/non-food recognition in the experiment within and across the same dataset.

# Classification with CNN-NIN

- Convolutional Neural Network (CNN)<sup>[14]</sup>
  - state-of-the-art for general image classification
  - Also effective for food recognition<sup>[4]</sup>
- Network in Network (NIN)<sup>[15]</sup>
  - We employ NIN as the model because it performs better than usual AlexNet and it is memory efficient
  - mlpconv layers and global average pooling



[14] Y. LeCun et al., "Gradient-based learning applied to document recognition", Proc. IEEE, 1998.

[15] M. Lin et al., "Network in Network", ICLR, 2014.



# Datasets

- Build two datasets for food/non-food classification
  - ① Instagram Food/Non-Food Dataset (Insta.)
    - Collect images tagged with “#food” from Instagram<sup>[12]</sup>
    - Motivated by another application: filtering SNS
  - ② Food-101/Caltech-256 Dataset (F/C)
    - Collect images in Food-101<sup>[10]</sup> as food class and Caltech-256<sup>[13]</sup> (except for some food images) as non-food class

Classess	#images (Insta.)	#images (F/C)
Food	4,230	25,250
Non-Food	5,428	28,322
Unspecified	342	-
<b>Total</b>	<b>10,000</b>	<b>53,572</b>

[12] <http://instagram.com/>

[13] G. Griffin et al., “Caltech-256 object category dataset”, Caltech Technical Report, 2006.

# Classification Protocol

No	Training	Testing
1	Instagram	Instagram
2	Food-101/Caltech-256	Food-101/Caltech-256
3	Food-101/Caltech-256	Instagram
4	Instagram	Food-101/Caltech-256

- No. 1 + 2 (Within a dataset)
  - change train/test ratio from 0.5 to 0.9 and average the result of five trial for each ratio.

- No. 3 + 4 (Across datasets)
  - is the model general?

# Within one dataset

Train: Insta. / Test: Insta.

		Predicted class	
		non-food	food
Actual class	non-food	<b>0.95</b>	<b>0.05</b>
	food	<b>0.06</b>	<b>0.94</b>

Accuracy: 94.8%

Train: F/C / Test: F/C

		Predicted class	
		non-food	food
Actual class	non-food	<b>0.97</b>	<b>0.03</b>
	food	<b>0.04</b>	<b>0.96</b>

Accuracy: 96.4%

Train / Test ratio is 0.8 / 0.2



# Wrong Cases (Instagram)

Ground Truth: Food



Package

Ground Truth: Non-food



Flower

# Wrong Cases (Food/Caltech)

## Ground Truth: Food



Package



Food area is small

## Ground Truth: Non-food



Living thing





# Across datasets

Train: F/C / Test: Insta.

		Predicted class	
		non-food	food
Actual class	non-food	<b>0.95</b>	0.05
	food	0.13	<b>0.87</b>

Accuracy: 91.5%

Train: Insta. / Test: F/C

		Predicted class	
		non-food	food
Actual class	non-food	<b>0.86</b>	0.14
	food	0.05	<b>0.95</b>

Accuracy: 90.6%



# Wrong Cases (across)

\*correct in "within" but wrong in "across"

## Ground Truth: Food



## Ground Truth: Non-food



# Comparison with The Previous Methods

- For comparison, we utilized the dataset in [Kagaya+, 2014]
  - Baseline: SIFT/Circle feature/Color feature + SVM<sup>[19]</sup>
    - It is used the web version of FoodLog

Method	Accuracy
Baseline <sup>[19]</sup>	89.7 ± 0.73%
Kagaya+, 2014	93.8 ± 1.39%
<b>This study (CNN-NIN)</b>	<b>99.1 ± 0.81%</b>



# Conclusion & Future Work

---

- Highly accurate food/non-food image classification for pre-process of food image analysis
  - Built novel datasets (we will get them open publicly)
  - Used a deep convnet as a state-of-the-art and conducted some experiments
  - Evaluated the performance using two datasets
    - within scheme: 95-96%
    - across scheme: 90-91%
- Future: we really introduce this study to the current FoodLog system



Thank you for your kind attention!

---

Any question?