Large-Scale Twitter Food Photo Mining and Its Applications

The Fifth IEEE International Conference on Multimedia Big Data (BIGMM)
Sep. 11th 2019

Keiji Yanai, Kaimu Okamoto,
Tetsuya Nagano and Daichi Horita

Artificial Intelligence eXploration Research Center
The Univ. of Electro-Communications, Tokyo, Japan

Overview of this talk

- **Introducing our Twitter photo mining works since Feb. 2011** (one month before the big earthquake)
  - **Geotagged tweet photo analysis**
    - Real-time geo-tweet photo mapping system [2012]
    - Event photo mining from geo-tweet photos [2012-2016]
    - Finding regional tendency on Twitter photos [2019]
  - **Twitter food photo mining**
    - Statistics on food image collection for 8 years [2012-]
    - Regional tendency on Twitter food photos [2019]
    - Applications of a large-scale Twitter food photoDB [2018-]
      - Food image translation by GAN, mobile app., food VR

Bac

- Various kinds of photos are posted to SNSs such as twitter Instagram facebook every seconds.
- Photos on SNSs are posted with text messages and meta data such as geotags.

- SNSs can be regarded as useful data sources for multimedia research.

twitter Instagram facebook

Why twitter?

- twitter provides the API to watch the Tweet stream in the real time way.
  - Twitter API statuses/filter
    (formerly TW Streaming API (～2018/8))

- Instagram facebook do not provide real-time API.
  - Most of the Facebook msg. are not public.
  - Instagram msg. are public, but its API is highly restricted (mainly designed for mobile apps.)

Google images vs twitter

Google images flickr — Image Search eng.

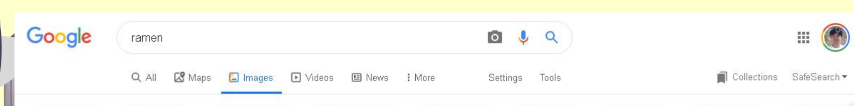
- Typical, very relevant
- Distributed over the Web
  - Various purpose

twitter — Micro-blog

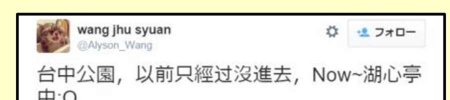
- Everyday life
- w/ tweet messages
  - Not describing

The biggest difference is instant or not.

- Appropriate for training CNN



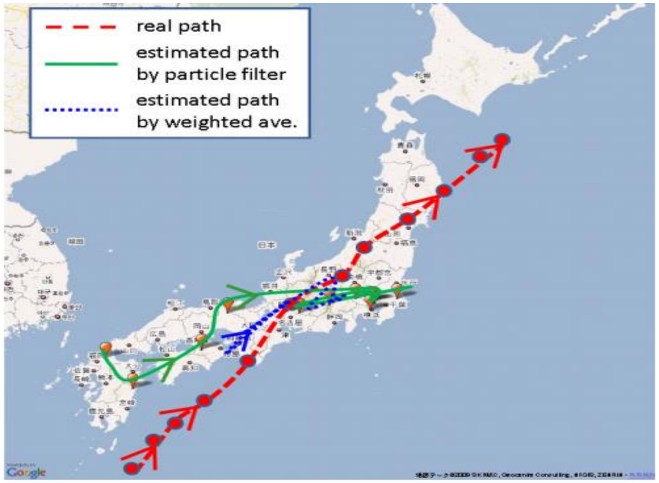
- 30 million/day



Twitter photos are more helpful to understand the current state/trend of the worlds .

Many text mining works using twitter

- Text analysis ⇒ so many
  - Event detection
  - Trend mining
  - Positive/Negative reputation



Typhoon trajectory estimated by tweets [WWW 2010]

- Photo analysis ⇒ limited before, but recently increasing due to CNN
  - Evaluation of relatedness between msg. and img.
  - Brand image mining, event photo detection
  - Fake News detection

Characteristic of the Twitter photos

- Normal condition : everday life
  - Food
  - Scene
  - People
- Something special: event photos
  - Artificial public events sport games
  - Natural phenomena earthquake, typhoon
  - Personal events
    - go hiking, travel, birthday

Mining two types of photos

- Event photo : special

- Food photo: everyday-life

Twitter photos: special event

Special big event photos on March 11th 2011 around Tokyo area

Twitter photos: normal event

Everyday-life photos on March 11th 2011 in the western part of Japan

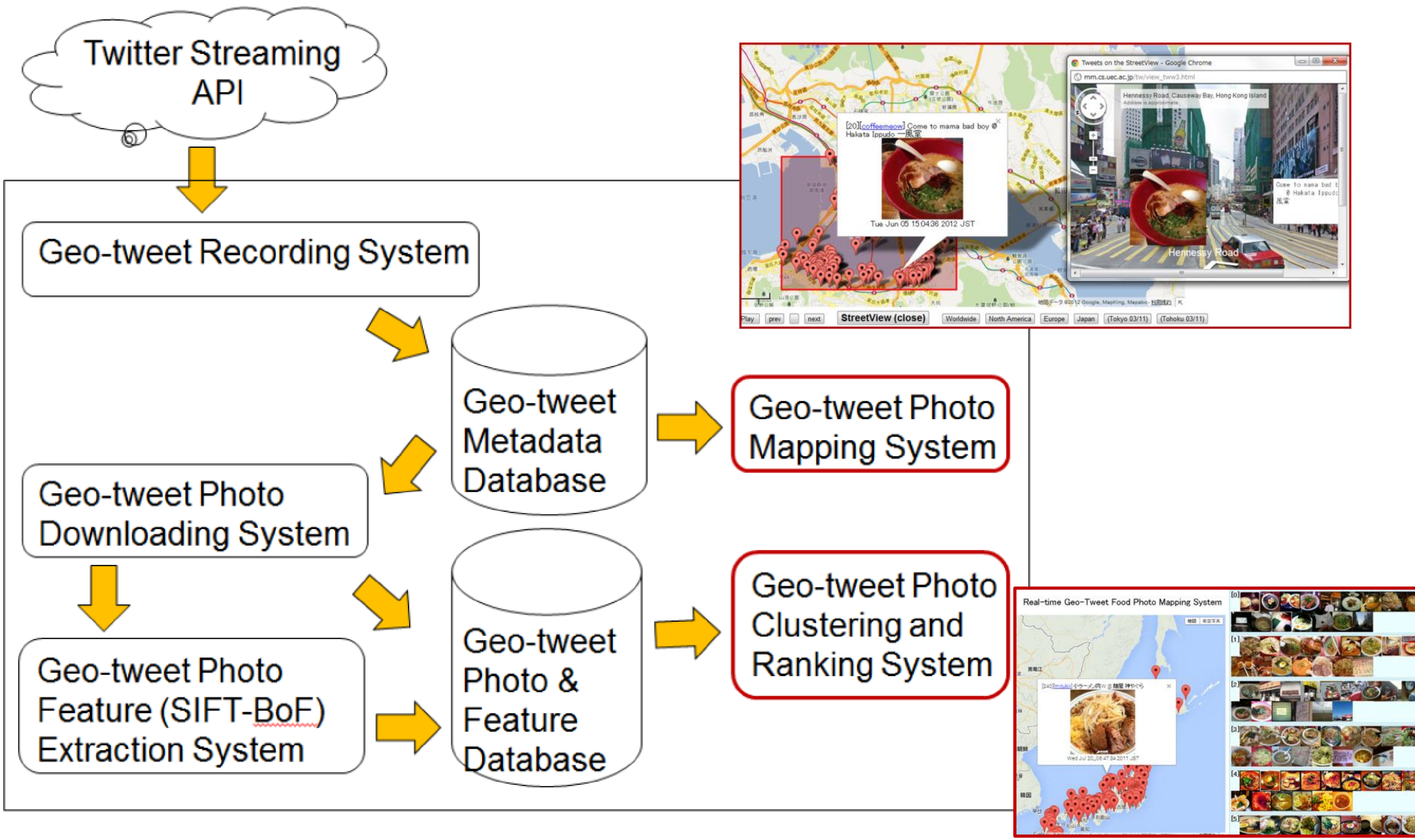
World Seer: A Real-time Geo-Tweet Photo Mapping System

ICMR 2012

World Seer: Real-time Twitter Photo Mapping System [ICMR 2012]

Build geo-photo tweet database for research

Monitoring the TW stream & Recording Geo-Photo Tweets



demo

- http://mm.cs.uec.ac.jp/tw/

News: TW St API discon.

- Twitter Streaming API was discontinued officially Aug. 2018.
  But it was still available until the end of July. 2019.
  (When we wrote this review paper, It was still available.)
- Aug 1st 2019, a new connection to the Twitter Streaming API was not accepted anymore.

Still OK : realtime tweet API

- An alternative method is provided.
  - Twitter API statuses/filter

Unfortunately in our system (in fact **MY system**) API is not update. So currently it does not work.

I will update it soon
after completing reviewing five MMM papers.

Tweet photo database 2011/2～2019/7

- Since Feb. 2011, we have collected
  - several billion photo tweets
  - 321million geo-photo tweets

  (5M geo-tweets/month before May 2015,
  0.5M geo-tweets/month after May 2015)
- We used this data for
  - Event Photo Mining
  - Food Photo Mining
  - Visual Topic Tendency Analysis
  - Training of GANs for food image translation

Twitter Event Photo Mining

Yusuke Nakaji and Keiji Yanai: Visualization of Real World Events with Geotagged Tweet Photos, IEEE ICME Workshop on Social Media Computing (SMC), (2012).

Takamu Kaneko and Keiji Yanai: **Visual Event Mining from Geo-tweet Photos**, IEEE ICME Workshop on Social Multimedia Research (SMMR), (2013).

Takamu Kaneko and Keiji Yanai: Event Photo Mining from Twitter Using Keyword Bursts and Image Clustering, Neurocomputing, Elsevier, Vol.172, pp.143–158 (2016).

Demo

- http://mm.cs.uec.ac.jp/kaneko-t/tw/jp/index.html

- http://mm.cs.uec.ac.jp/kaneko-t/tw/us/index.html

Twitter Event Photo Mining

- Mine the photos related to the events happened in the specific areas and times

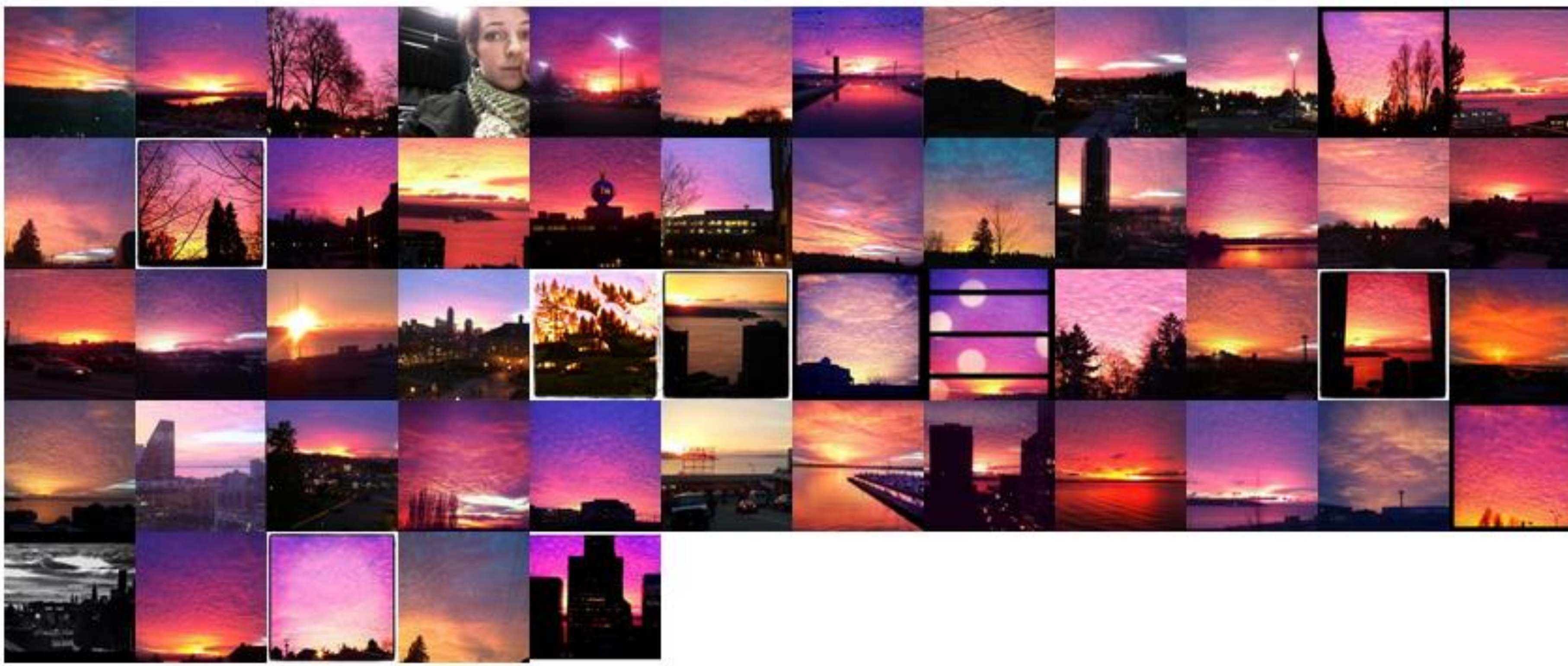
The results of detected event photos in 2012

Twitter Event Photo Mining

• sunset January 13, 2012 ド

sunset January 13, 2012

Cluster No.1 num="53" bof="156.684" color="336.837" weight="10.757" score="61.224"



Objective

- Detect events from Twitter stream
  - Weather, natural events
  - Festivals, sport games
- Understand events visually
  - Select representative photos
  - Place event photos on a map

Mapping events with the photo

Processing flow

Textual analysis

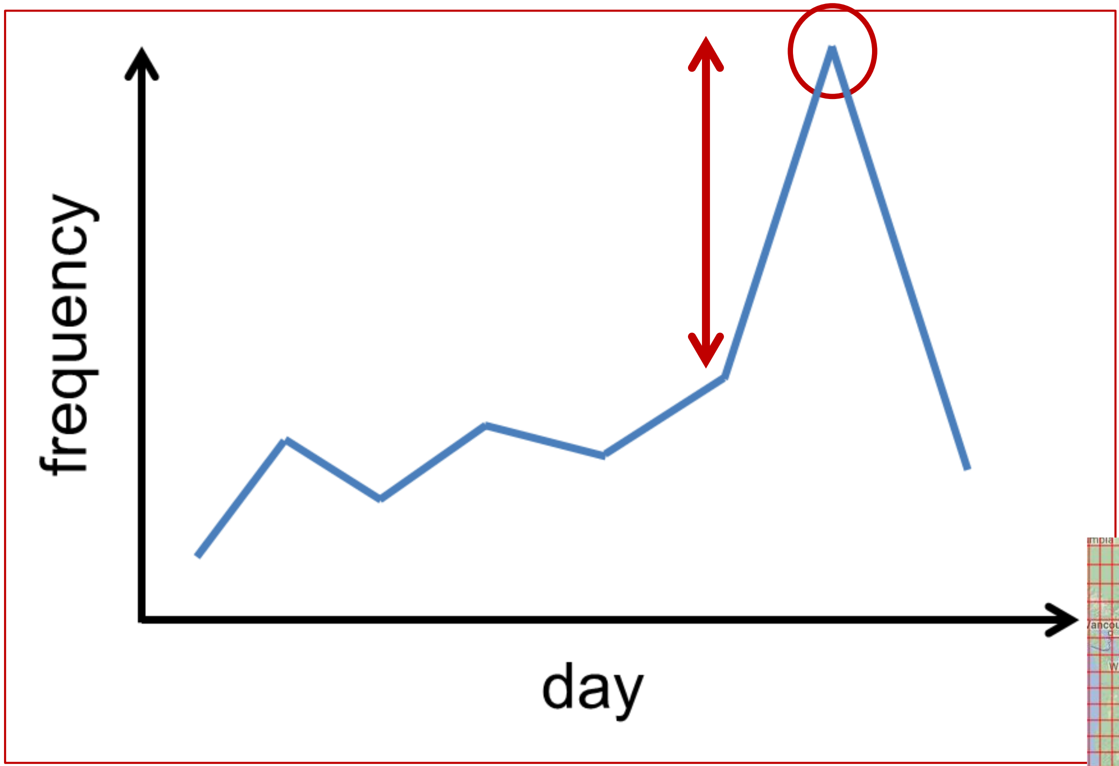
1. Event keyword detection
2. Keyword unification

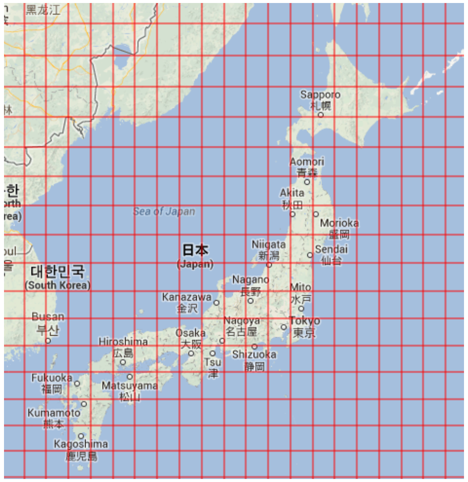
Visual feature analysis

3. Event photo clustering
4. Mapping event with photos

Event Keyword Burst Detection

- Examine change of daily frequency





Event Photo Clustering

- Hand-crafted image features (not CNN !)
  - Bag-of-Features with SURF
  - Color histograms
- Ward clustering method
  - a hierarchical clustering method
  - threshold is 300 (both)

$$E(C) = \sum_{x \in C} ((x_{BoF} - \overline{x_{BoF}})^2 w_{BoF} + (x_{RGB} - \overline{x_{RGB}})^2 w_{RGB})$$

Experiments

- Japan Dataset
  - Feb 10th, 2011 to Sep 30th, 2012
  - about 3 million geo-tweet photos

- US Dataset
  - Jan 1st, 2012 to Dec 31st, 2012
  - about 17 million geo-tweet photos

Results of Keyword Detection

国立大学法人 UEC TOKYO 電気通信大学

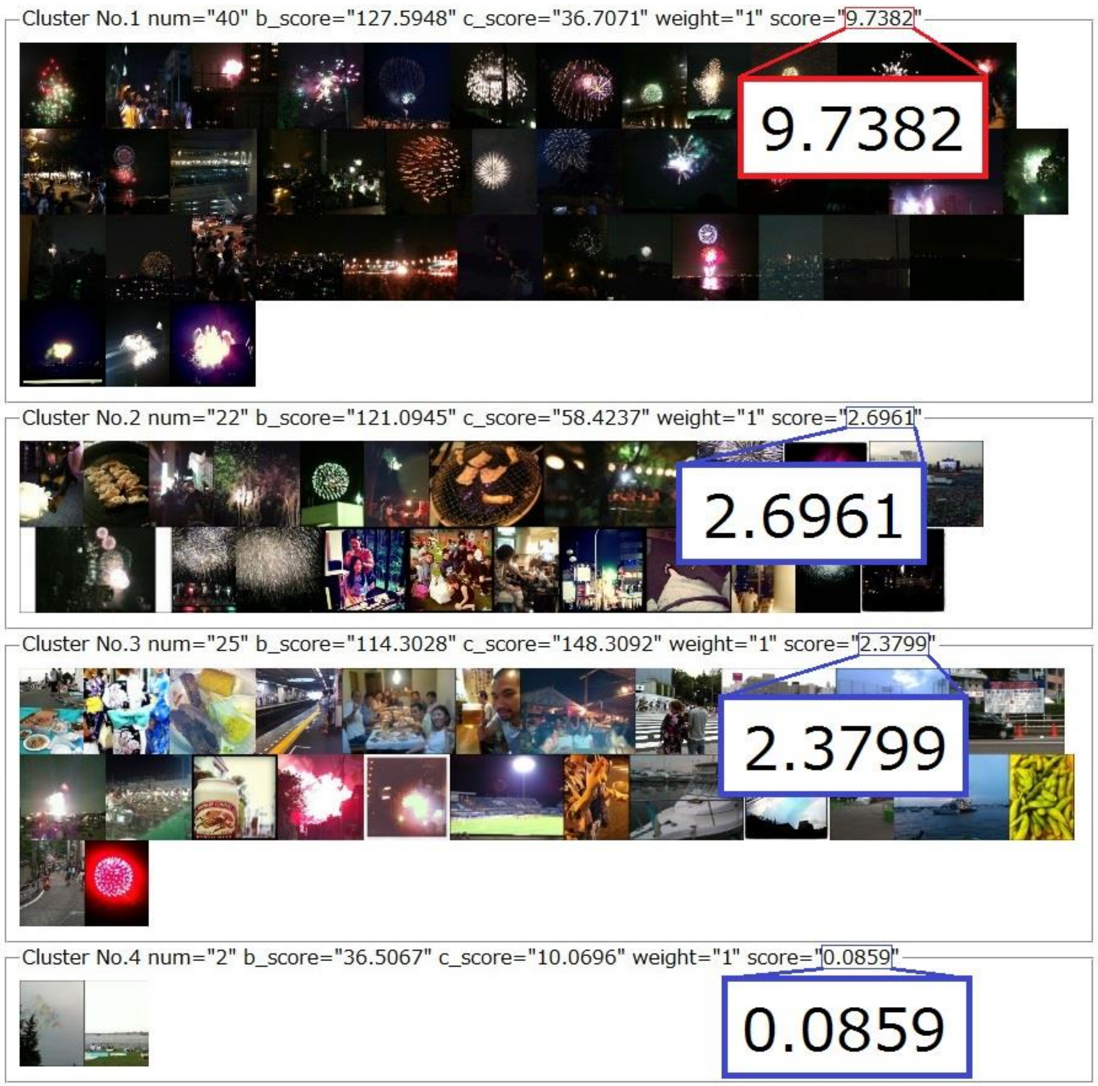
Japan

Keyword	Date
snow	11/02/2011
earthquake	11/03/2011
fireworks	06/08/2011
typhoon	21/09/2011
Mt. Fuji	24/09/2011
Apple	06/10/2011
eclipse	10/12/2011
illumination	10/12/2011
Christmas	24/12/2011
New years eve	31/12/2011
sunrise	01/01/2012
firefly	06/05/2012

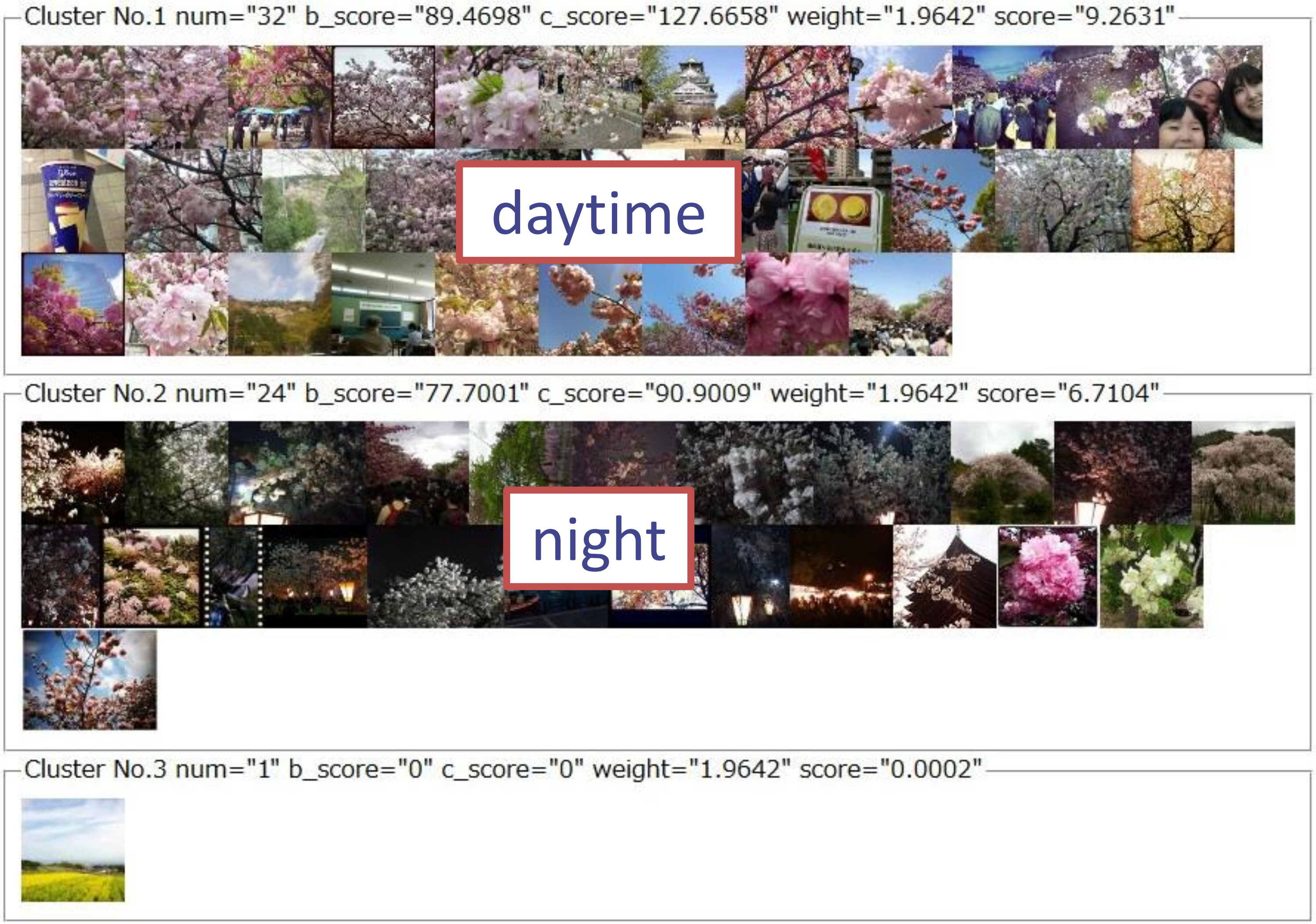
USA

Keyword	Date
snow	09/01/2012
sunset	13/01/2012
Grammy	12/02/2012
Valentines	14/02/2012
SXSW	09/03/2012
Easter	08/04/2012
shuttle	17/04/2012
WWDC	10/06/2012
hurricane	26/08/2012
rainbow	05/09/2012
49ers	18/10/2012
NYE	31/12/2012

“fireworks” photo clusters



“cherry blossoms” photo clusters

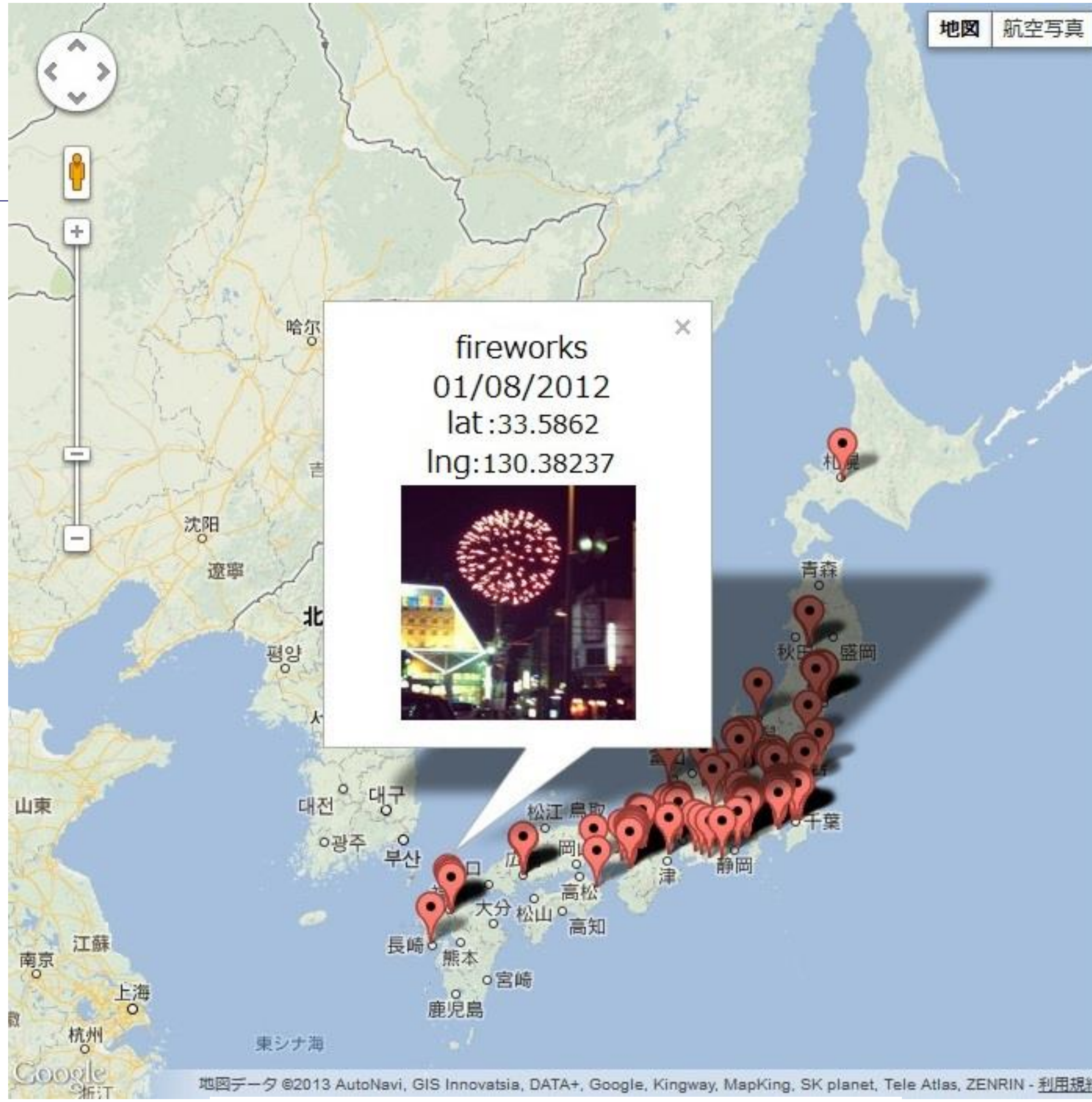


国立大学法人 電気通信大学

Mapping Results

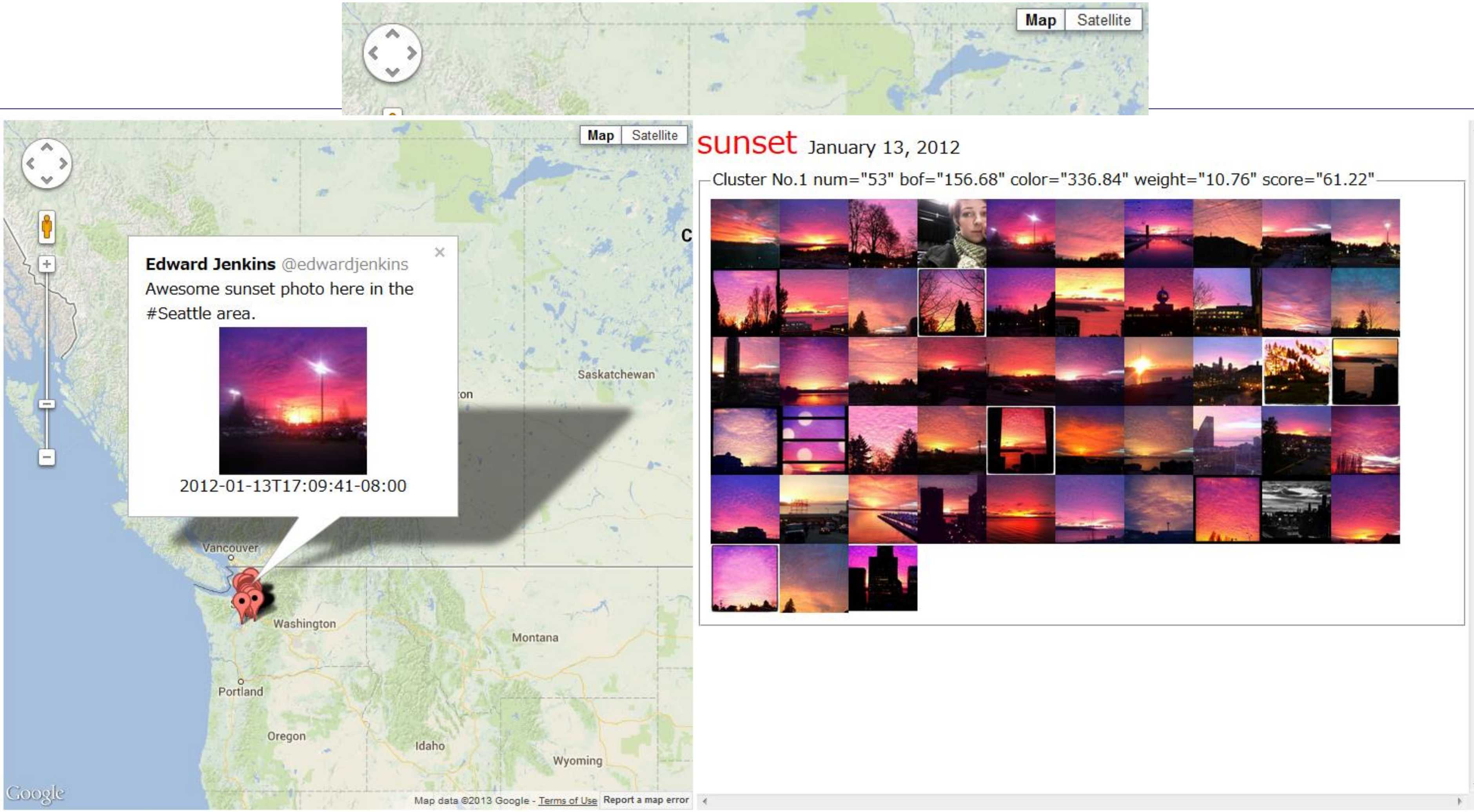
- Map event in a map
  - Calculate coordinates of event
  - Correspond information and the photo
- Summary of results

	Japan	USA
# events	258	1676
accuracy	65.5%	72.5%



Research dataset

“River’s Edge Music Festival”


Map
Satellite
Edward Jenkins @edwardjenkins
Awesome sunset photo here in the #Seattle area.
2012-01-13T17:09:41-08:00
Saskatchewan
Vancouver
Washington
Montana
Portland
Oregon
Idaho
Wyoming
Map data ©2013 Google - Terms of Use
Report a map error
sunset January 13, 2012
Cluster No.1 num="53" bof="156.68" color="336.84" weight="10.76" score="61.22"

“sunset”

Visual Topic Analysis of Twitter Photo Analysis

Unpublished.

2million photo clustering using only visual features (no text)

- http://mm.cs.uec.ac.jp/twimg/
  (BOF features)

- http://mm.cs.uec.ac.jp/twimg/dcnn.cgi
  (CNN features)

Most of the tweet texts do not explain the attached images directly. So text-based analysis might restrict target images too much. ⇒ **Twitter images with only visual analysis**

Food is one of the major topics of Twitter photos

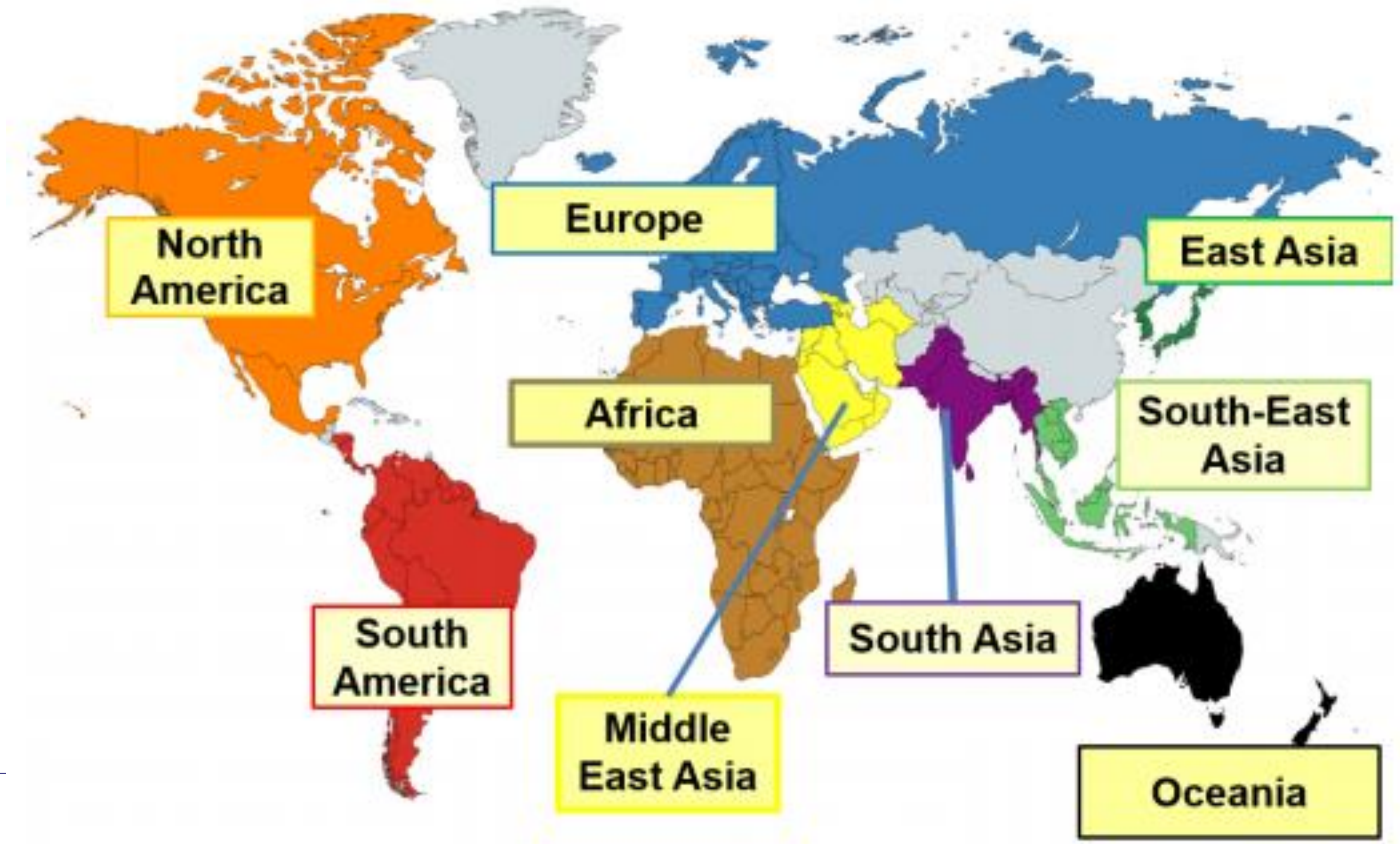
- Visual topic analysis with half-million Twitter photos employing DCNN feature.

Finding regional tendency on Twitter photos using only image features

Tetsuya Nagano, Takumi Ege, Wataru Shimoda and Keiji Yanai: A Large-scale Analysis of Regional Tendency of Twitter Photos Using Only Image Features, Proc. of IEEE International Conference on Multimedia Information Processing and Retrieval (MIPR), (2019).

Regional tendency analysis on Twitter geotagged photos

- Apply visual clustering based photo topic analysis on each of the regions over the world.

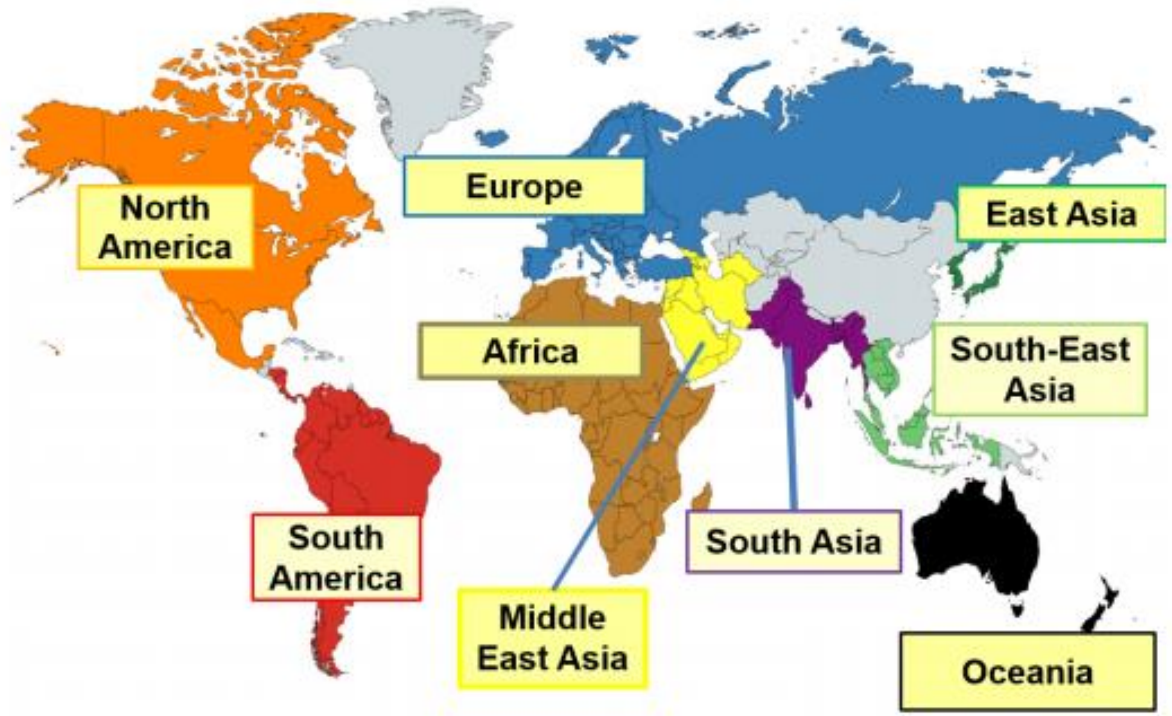


Experiments

- **Dataset**
  - Collected from January to June in 2016
  - 2,161,000 geotagged Twitter images
- **Feature**
  - CNN (128-d compressed by PCA)
- **K-means**
  - K-means with one-tenth images
    - Assigned rest of images into the nearest clusters
  - K=100

Regions

- East Asia, North America, South America, Europe, Africa, Middle East, South Asia and South-East Asia, Oceania



Clustering results (CNN features)

Select five representative topics from observation of clusters

- pre-selected photo genres.
  - "people"
  - "building"
  - "document"
  - "scene"
  - "food"

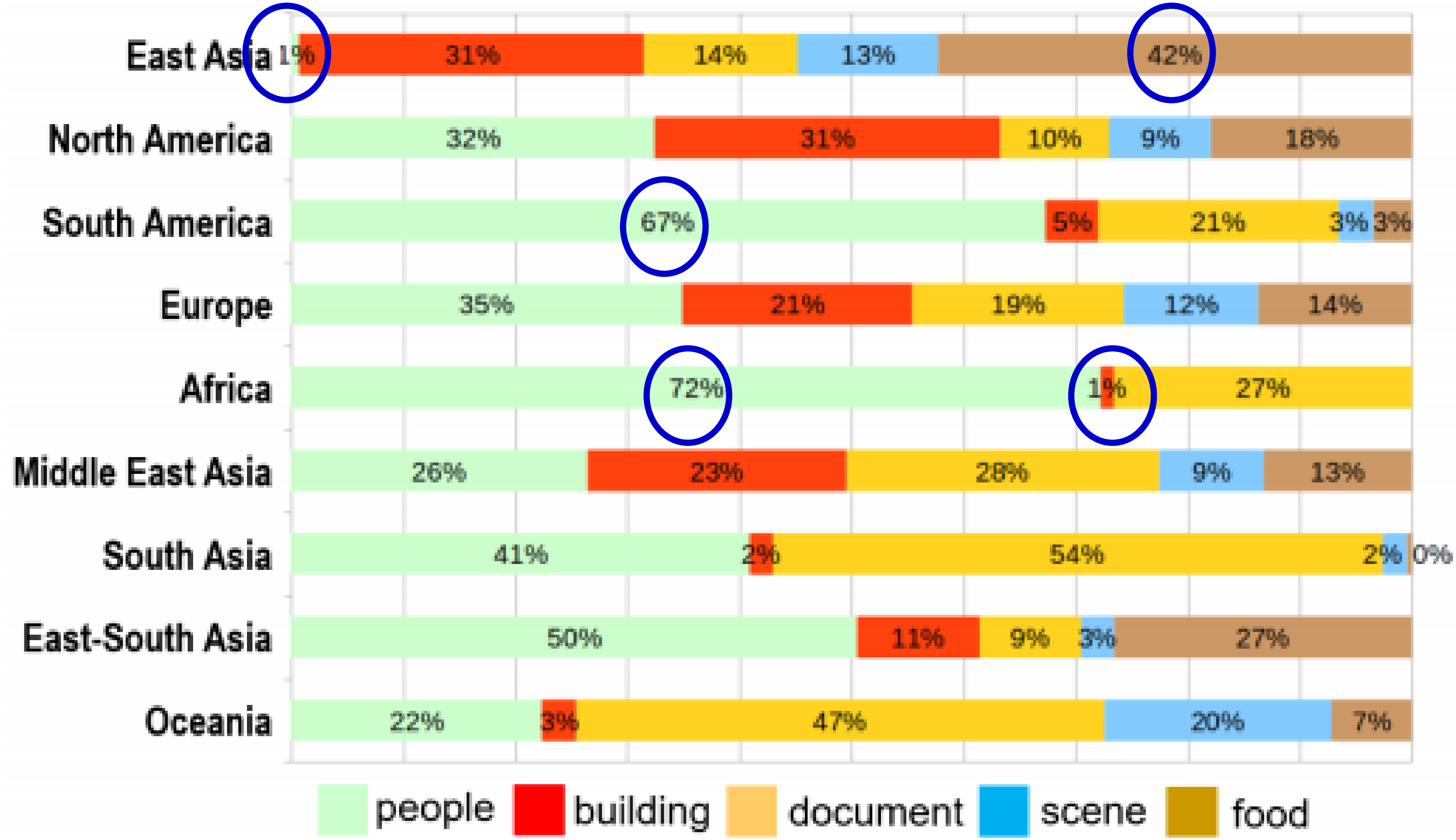
Figure 5. "Food" in East Asia.

Figure 6. "Building" in North America.

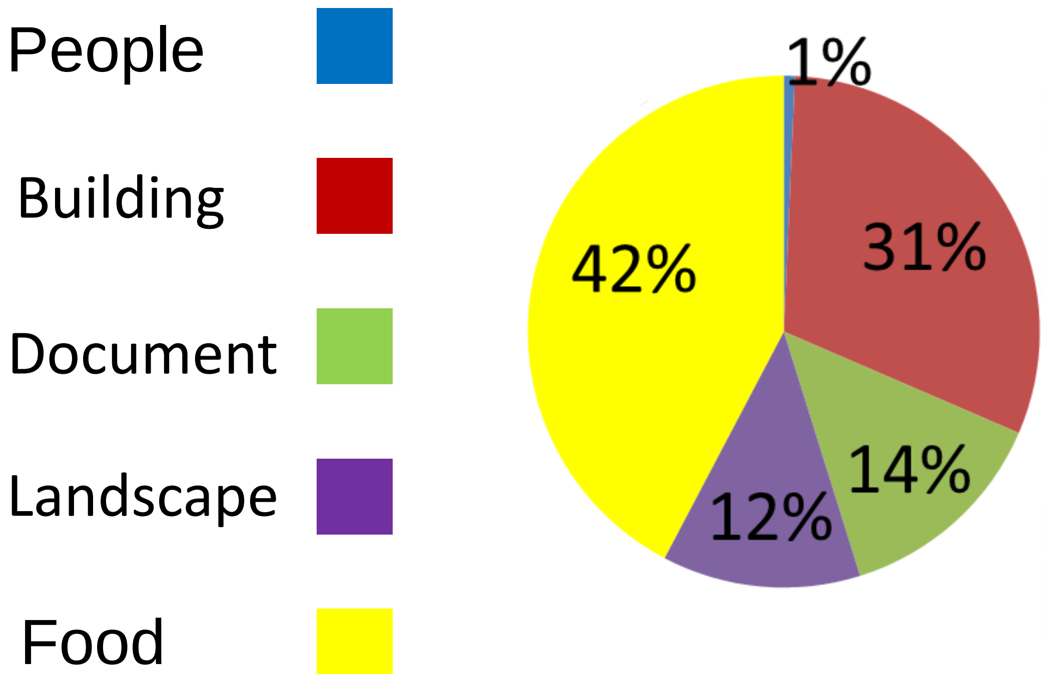
Figure 7. "People" in South America.

Figure 8. "Document" in Middle East.

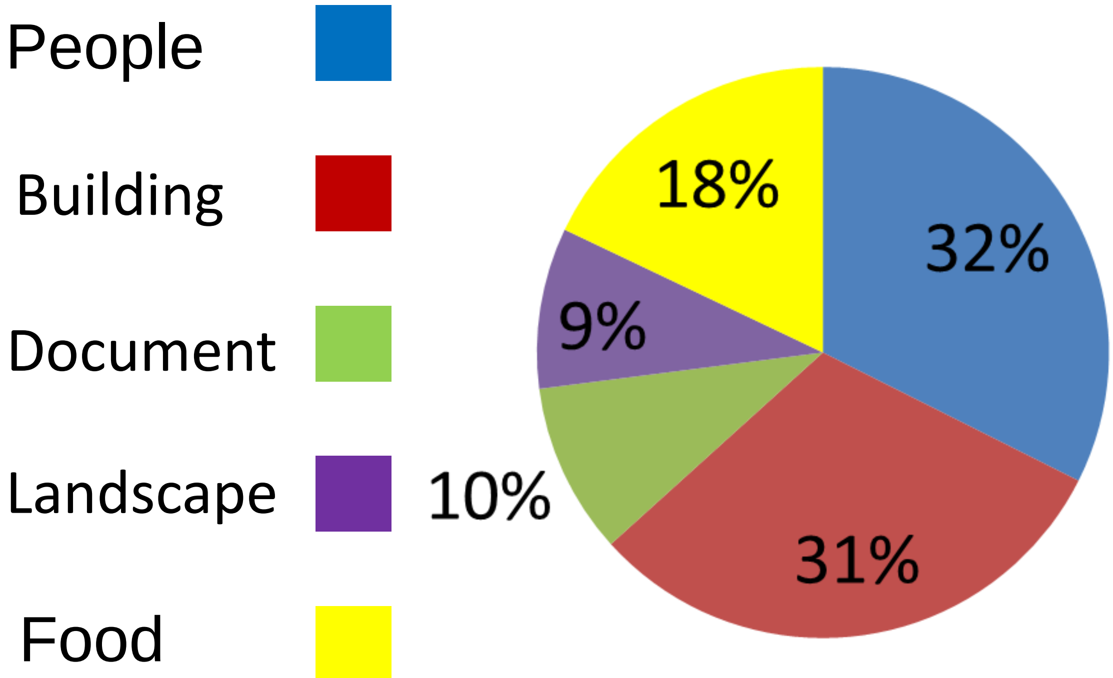
Ratio of five topics on each region



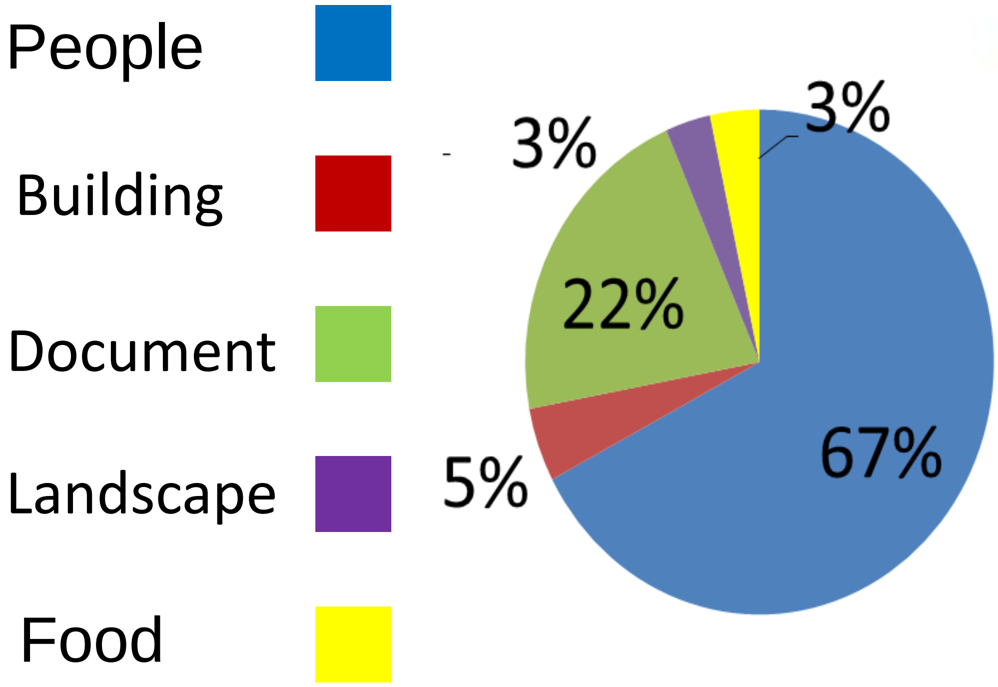
- **East Asia**
  - No people photos
  - Many building and food photos
  - The total ratio of building and food photos were more than 70%



- **North America**
  - The ratio of people and building were high more than 60%.

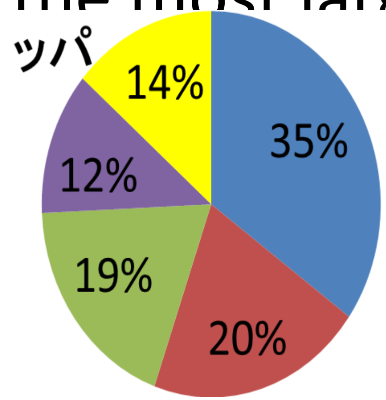


- **South America**
  - People photos are the most popular genre (67%)



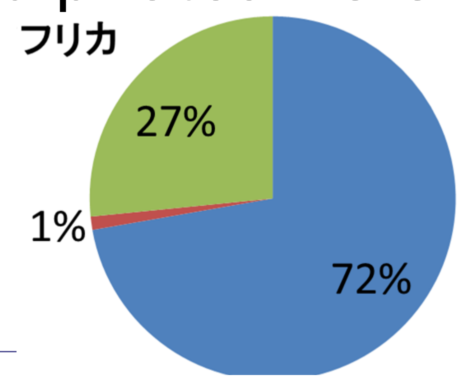
Analysis of Regional Tendency of Photo Genres

- **Europe**
  - The number of posted photos was the most large
  - The genres were well balanced.



People Document Food

- **Africa**
  - Almost no building, scene and food photos were posted
  - People photos occupied 70%.



Building Landscape

• Middle East

– Although the number of posts were fewer than Europe, all the five genres were balanced as well

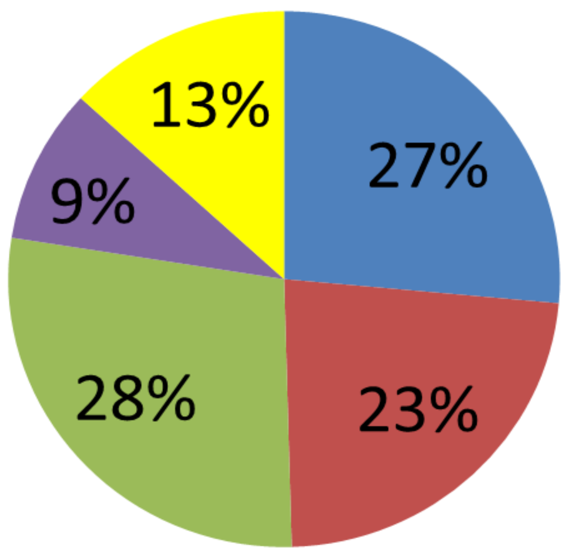
People

Building

Document

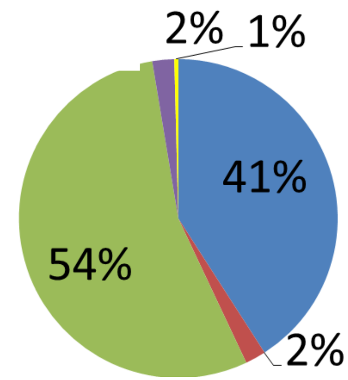
Landscape

Food



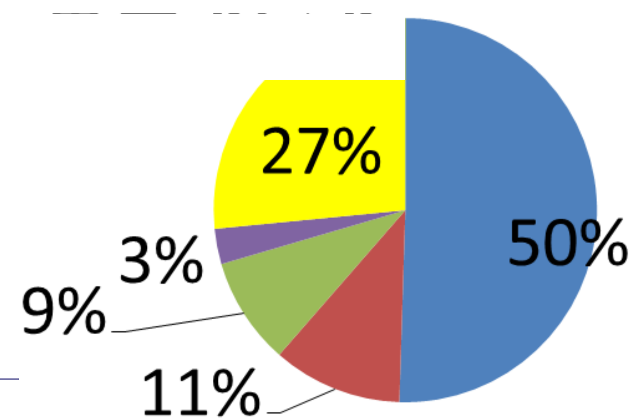
Analysis of Regional Tendency regarding Photo Topics

- **South Asia**
  - More than half of the photos were document photos.
  - This tendency was not observed in other regions



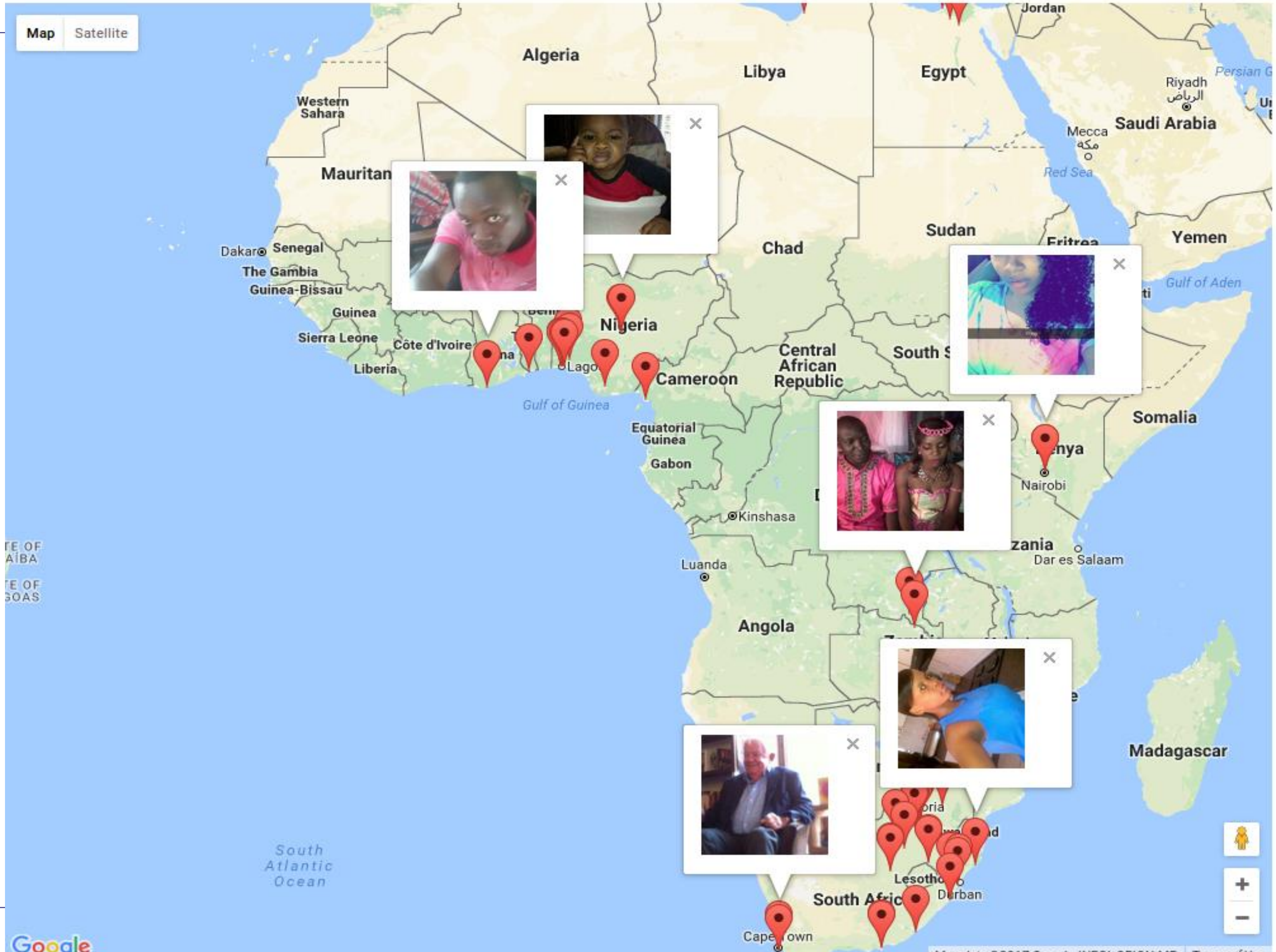
- **SouthEast Asia**
  - People photos are the most and in addition food photos was the second most

People | Document | Food
Building | Landscape



Sourth America : People

Africa : People



East Asia : Food

Sourth East Asia : Food

Findings

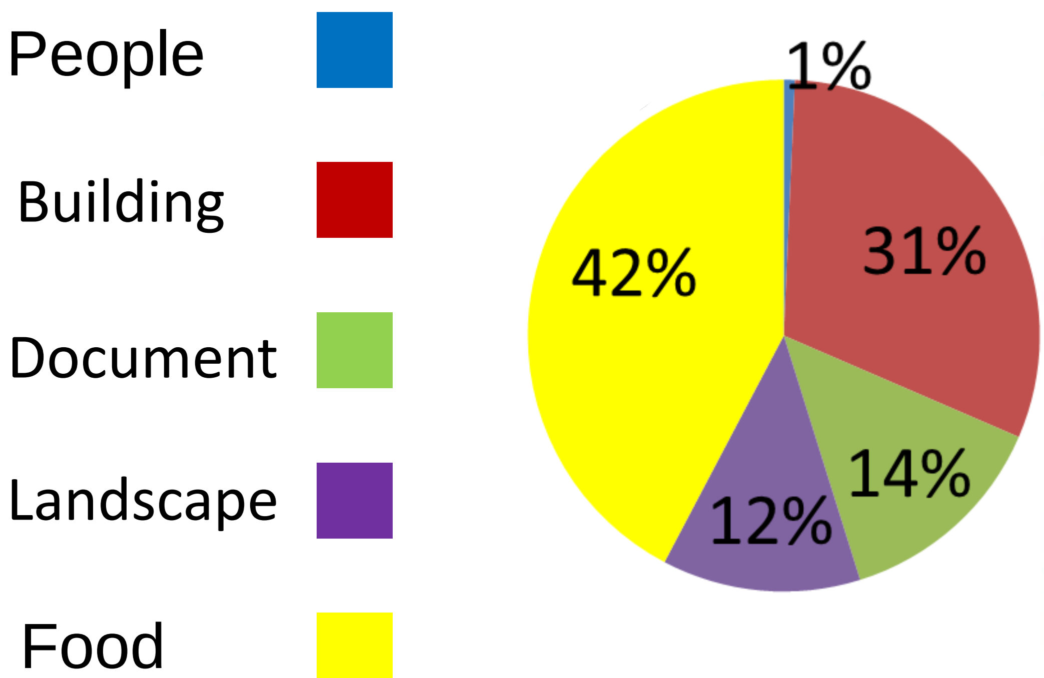
- Tendency
  - East Asia and East-South Asia,
    - Food photos are relatively high
  - South America, South Asia and East-South Asia
    - people photos are exceptionally high.
  - Europe and MiddleEast
    - well balanced.
- East Asia enjoys posting food photos
- South America, South Asia and EastSouth Asia like to post people photos without caring privacy issue.

Food Twitter Photo mining

Why food ?

Food is one of the major topics in Twitter photos.

Especially in East Asia



We are actively working on food images.

FoodCam : [Kawano et al. MTA13]

- Real-time mobile food recognition
  Android application

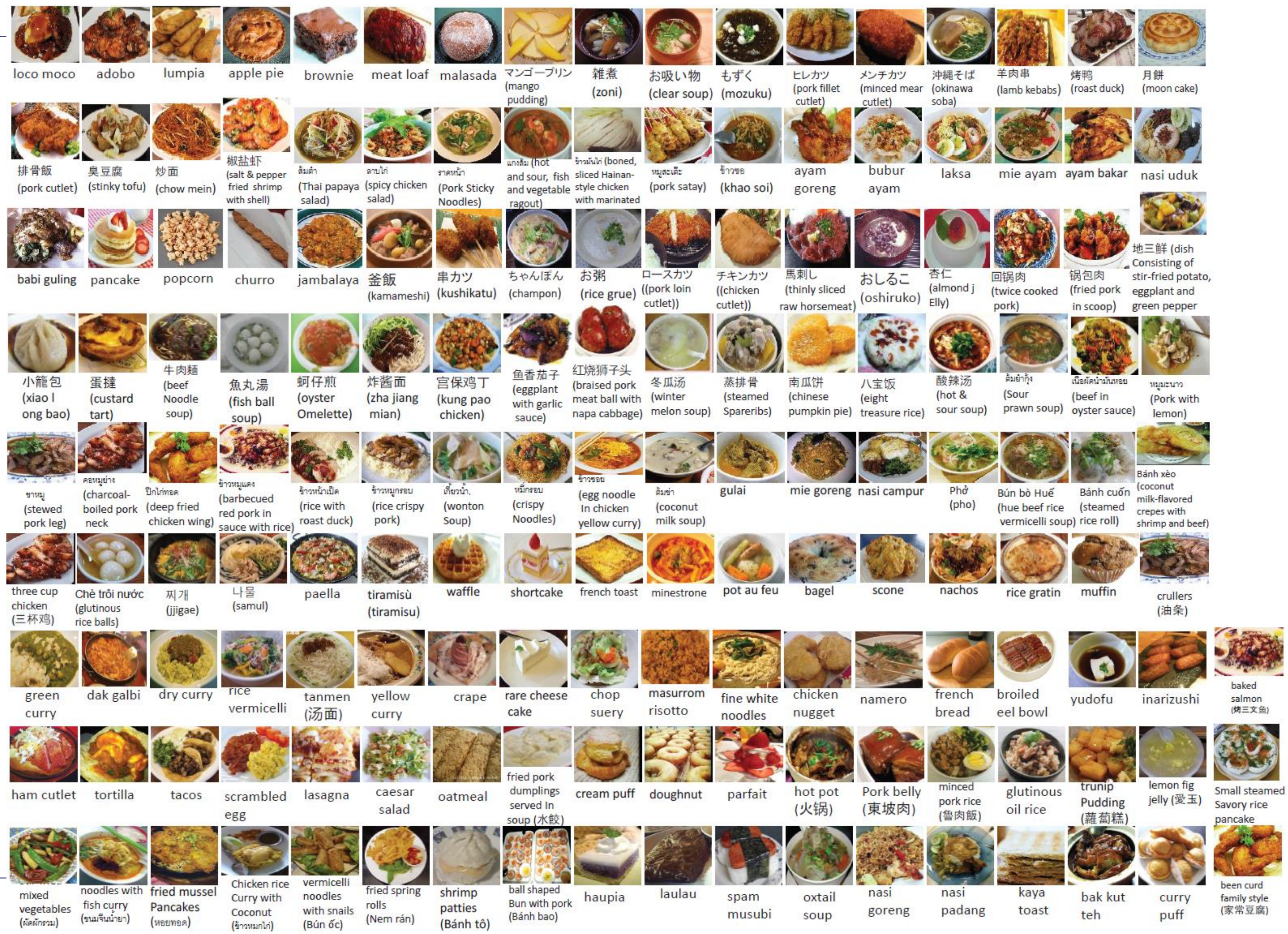
http://foodcam.mobi/

UEC-FOOD 100


rice
eels on rice
pilaf
chicken-'n'-egg on rice
port cutlet on rice
beef curry
sushi
chicken rice
fried rice
tempura bowl
bibimbap
toast
croissant
roll bread
raisin bread
chip butty
hamburger
pizza
sandwiches
udon noodle
tempura udon
soba noodle
ramen noodle
beef noodle
tensin noodle
fried noodle
spaghetti
Japanese-style pancake
takoyaki
gratin
sauteed vegetables
croquette
grilled eggplant
sauteed spinach
vegetable tempura
miso soup
potage
sausage
oden
omelet
ganmodoki
jiaozi
stew
teriyaki grilled fish
fried fish
grilled salmon
salmon meuniere
sashimi
grilled pacific saury
sukiyaki
sweet and sour pork
lightly roasted fish
steamed egg hotchpotch
tempura
fried chicken
sirloin cutlet
nanbanzuke
boiled fish
seasoned beef with potatoes
hambarg steak
steak
dried fish
ginger pork saute
spicy chili-flavored tofu
yakitori
cabbage roll
omelet
egg sunny-side up
natto
cold tofu
egg roll
chilled noodle
stir-fried beef and peppers
simmered pork
boiled chicken and vegetables
sashimi bowl
sushi bowl
fish-shaped pancake with bean jam
shrimp with chill source
roast chicken
steamed meat dumpling
omelet with fried rice
cutlet curry
spaghetti meat sauce
fried shrimp
potato salad
green salad
macaroni salad
Japanese tofu and vegetable chowder
pork miso soup
chinese soup
beef bowl
rice ball
pizza toast
dipping noodles
french fries
goya chanpuru
kinpira-style sauteed burdock
hot dog
mixed rice

国立大学法人 電気通信大学

UEC-FOOD 256



国立大学法人 電気通信大学

UEC-FOOD 256

FoodRec: foodrec app with UECFOOD100

by Hamlyn Centre-Imperial College(UK)

UEC-FOOD as a Fine-Grained Image Classification Dataset

arXiv.org > cs > arXiv:1502.07802

Search or Article-id

Computer Science > Computer Vision and Pattern Recognition

Modelling Local Deep Convolutional Neural Network Features to Improve Fine-Grained Image Classification

ZongYuan Ge, Chris McCool, Conrad Sanderson, Peter Corke

(Submitted on 27 Feb 2015)

We propose a local modelling approach using deep convolutional neural networks (CNNs) for fine-grained image classification. Recently, deep CNNs trained from large datasets have considerably improved the performance of object recognition. However, to date there has been limited work using these deep CNNs as local feature extractors. This partly stems from CNNs having internal representations which are high dimensional, thereby making such representations difficult to model using stochastic models. To overcome this issue, we propose to reduce the dimensionality of one of the internal fully connected layers, in conjunction with layer-restricted retraining to avoid retraining the entire network. The distribution of low-dimensional features obtained from the modified layer is then modelled using a Gaussian mixture model. Comparative experiments show that considerable performance improvements can be achieved on the challenging Fish and UEC FOOD-100 datasets.

UECFOOD-256 is used in NVIDIA ICCV2019 paper

- Ming-Yu Liu et al: Few-Shot Unsupervised Image-to-Image Translation, ICCV2019.

x → style (y1,y2) → x'
food image translation

We are organizing food WS.

MADiMa 2019

ACM MM Workshop related to "food multimedia"

5th International Workshop on Multimedia Assisted Dietary Management

In conjuction with the 27th ACM International Conference on Multimedia (ACMMM2019), Nice, France

Organization

Workshop chairs



Stavroula Mougiakakou, University of Bern, Switzerland

Giovanni Maria Farinella, University of Catania, Italy

Keiji Yanai, The University of Electro-Communications, Tokyo, Japan

Paper submission deadline: July 8th, 2019

Notification of acceptance: August 5th, 2019

Camera-ready deadline: August 12th, 2019

Workshop date: **October 21th**

Twitter Real-time Food Photo Mining

Keiji Yanai and Yoshiyuki Kawano: Twitter Food Image Mining and Analysis for One Hundred Kinds of Foods , Pacifit-Rim Conference on Multimedia (PCM), (2014).

Keiji Yanai, Kaimu Okamoto, Tetsuya Nagano and Daichi Horita: Large-Scale Twitter Food Photo Mining and Its Applications, IEEE International Conf. on Big Multimedia (BIGMM), (2019)

Twitter Real-time Food Photo Mining System (mm.cs.uec.ac.jp/tw/)

UEC TOKYO 国立大学法人 電気通信大学

- What kinds of foods are being eaten in Japan ?

Objective

VS

Ramen　　Curry

- Which food is the most popular in Japan?
  - "Ramen vs Curry" problem ⇒ very controversial
  - I would like to put a period to this controversy by Twitter food photo mining !!!

Approach for food photo mining

- Two-step food photo selection
  - [1] Keyword-based tweet selection
  - [2] Image-based photo selection
    - 101-class (100 foods + non-food) classification (fine-tuned AlexNet)

Keyword-based selection → 100 Food/non-food classification

Experiments

- Collect photo tweets via Twitter Streaming API
  - From 2011/5 to 2019/07/16 (8year 2month)
  - About several billion photo tweets

- Search for the tweets including any of 100-food names (in Japanese) and apply a food CNN
  - 16,044,090 images ⇐ Apply the 101-food CNN

- 2,308,988 food photos (14.4%)
  13,735,102 non-food photos (85.6%)

Twitter food photo ranking

rank	foods	#photos








Ramen noodle is the most popular food in Japan.
I have solved "ramen vs curry" problem !!!
(And I got a "ramen vs curry" expert !!)

Precision of the top 5 foods (May 2011-Aug. 2013)

国立大学法人 UEC TOKYO 電気通信大学

Food	(1) KW	(2) f/n	(3) spec.	(4) ALL
ramen	275,652 **72.0%**	200,173 92.7%	84,189 95.0%	80,021 **99.7%**
curry	224,685 **75.0%**	163,047 95.0%	62,824 97.0%	59,264 **99.3%**
sushi	86,509 **69.0%**	43,536 86.0%	48,019 72.3%	25,898 **92.7%**
tsukemen	33,165 **88.7%**	24,896 96.3%	28,846 93.7%	22,158 **99.0%**
omelet	34,125 **90.0%**	28,887 96.3%	18,370 98.0%	17,520 **99.0%**

Only keyword search (Ramen noodle) (72.0%)

After applying 100-class food classifier (final)(99.7%)

Only keyword search (curry) (75.0%)

Final results (curry) (99.3%)

Some interesting findings

- Letters or drawings are sometimes drawn on omelets with ketchup

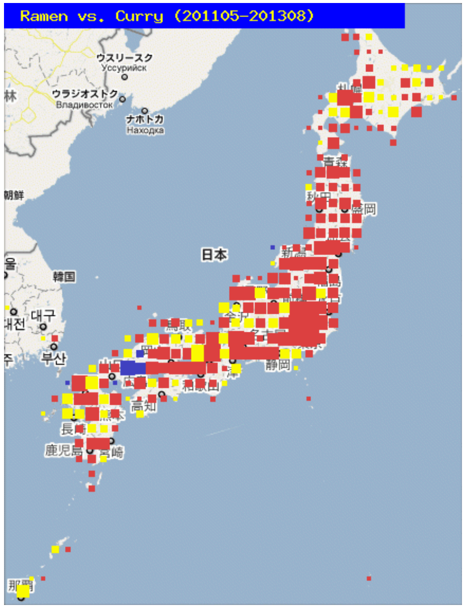
- Fast-foods such as humberger (rank 30^{th}) and beef bowl (rank 27^{th}) are ranked lower, since their appearance is always the same.

Not worth posting fastfood photos to Twitter

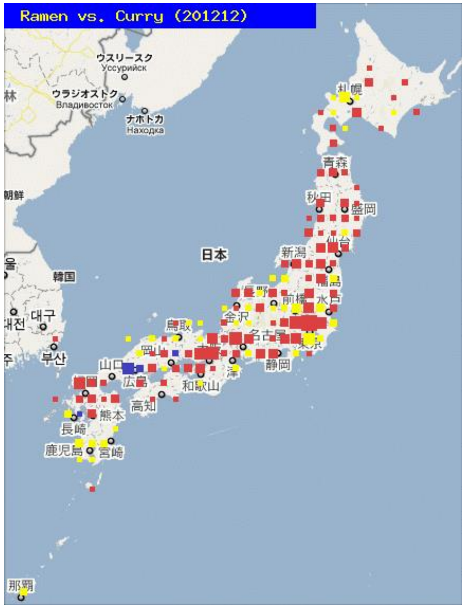
Omlet wall paper

Geographical-Temporal analysis on ramen vs curry

12.6% of the obtained food photos have geotag.

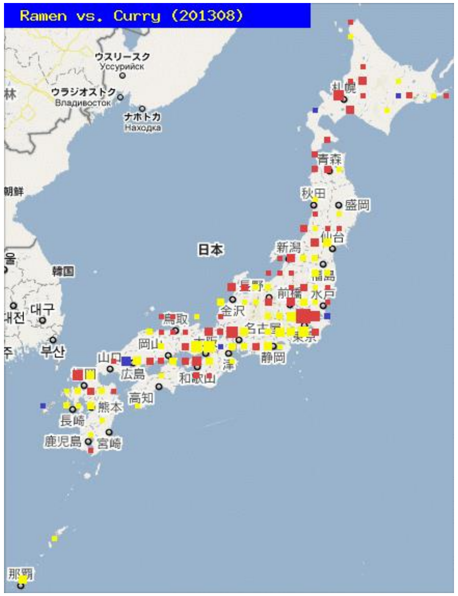


Whole year



Dec. (winter)

Ramen is popular.



Aug. (summer)

Curry gets more popular than ramen only in summer.

Ramen **Curry**

I appeared TV program in Japan (2018/11/1)

I commented on "ramen vs curry" problem

as a "ramen vs curry" expert.


Ramen won !!!
twitter

Curry lost !

But the rank is second.

Regional Tendency Analysis on Twitter Food Photo

K. Okamoto and K. Yanai, “Analyzing regional food trends with geo-tagged twitter food photos,” in Proc. of International Conference on Content-Based Multimedia Indexing (CBMI), 2019.

Method overview

The method is almost the same as the work on regional tendency analysis on generic images. The difference is using a food/non-food classifier at first.

[1] Classifying food and non-food photos

[2] Extracting food CNN features

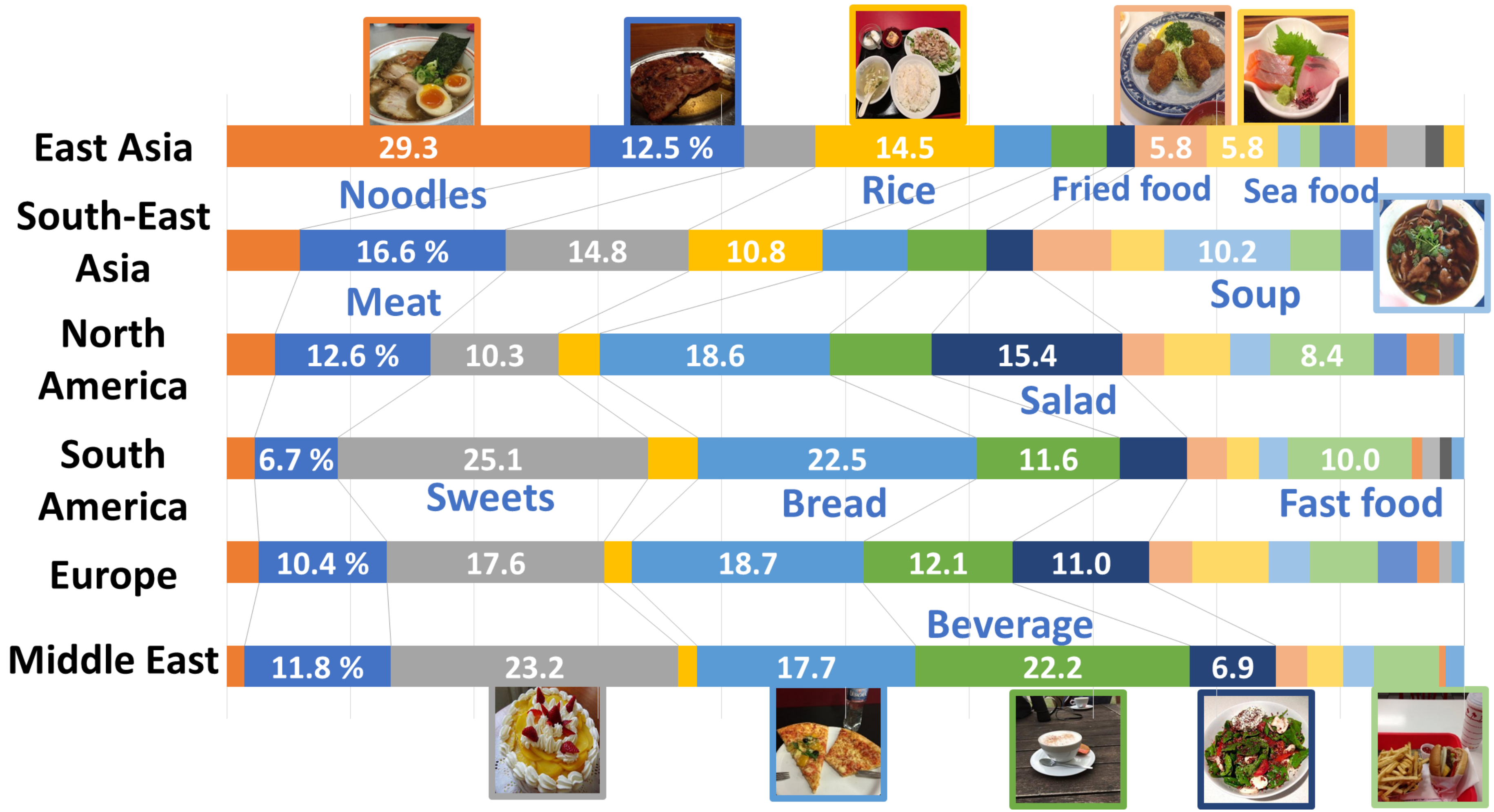
[3] Clustering and analyzing of regional tendency

Experiments

- Selecting food images from Twitter images in 2016 for whole a year with the food/non-food classifier

- 190,000 food images from 3.78 million raw Twitter images

Results



East Asia
South-East Asia
North America
South America
Europe
Middle East
29.3
12.5 %
14.5
5.8
5.8
Noodles
Rice
Fried food
Sea food
16.6 %
14.8
10.8
10.2
Meat
Soup
12.6 %
10.3
18.6
15.4
8.4
Salad
6.7 %
25.1
22.5
11.6
10.0
Sweets
Bread
Fast food
10.4 %
17.6
18.7
12.1
11.0
Beverage
11.8 %
23.2
17.7
22.2
6.9

Tendency analysis

- **East Asia** : "**Noodles**" and "**Rice**", which are relatively rare in the other regions, are included at the top.
- **South-East Asia** : ***"Soup"*** is ranked at the top, and many dishes are made of vegetables and meats in soup.
- **North America, South America, Europe** : It turned out that 4 items of the top 5 items of the categories are the same
- **Middle East** : The top five food categories are the same as Europe. However there are many brown coffee photos which are a unique type to Middle East.

Applications of Large-Scal Twitter Food Photo DB

[food image translation (food GAN)]

R. Tanno, D. Horita, W. Shimoda, and K. Yanai, “Magical rice bowl: Real-time food category changer,” ACM Multimedia, demo, 2018.

D. Horita, R. Tanno, W. Shimoda, and K. Yanai, “Food category transfer with conditional cycle gan and a largescale food image dataset,” in Proc. of International Workshop on Multimedia Assisted Dietary Management (MADIMA), 2018.

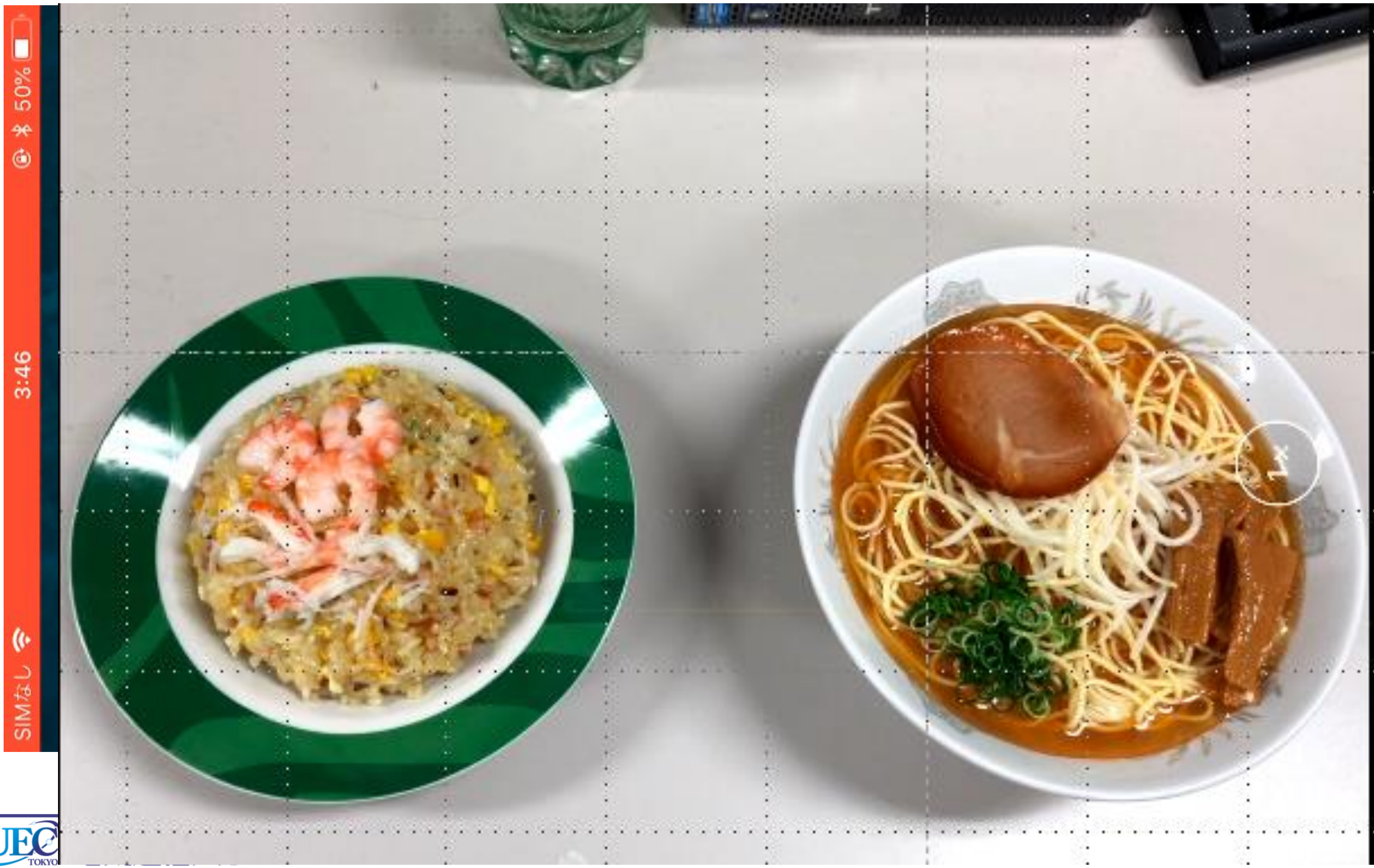
[Food VR]

S. Naritmo, R. Tanno, T. Ege, and A. K. Yanai, “Cnnbased food transformation on hololens,” in Proc. of International Workshop on Interface and Experience Design with AI for VR/AR (DAIVAR), 2018.

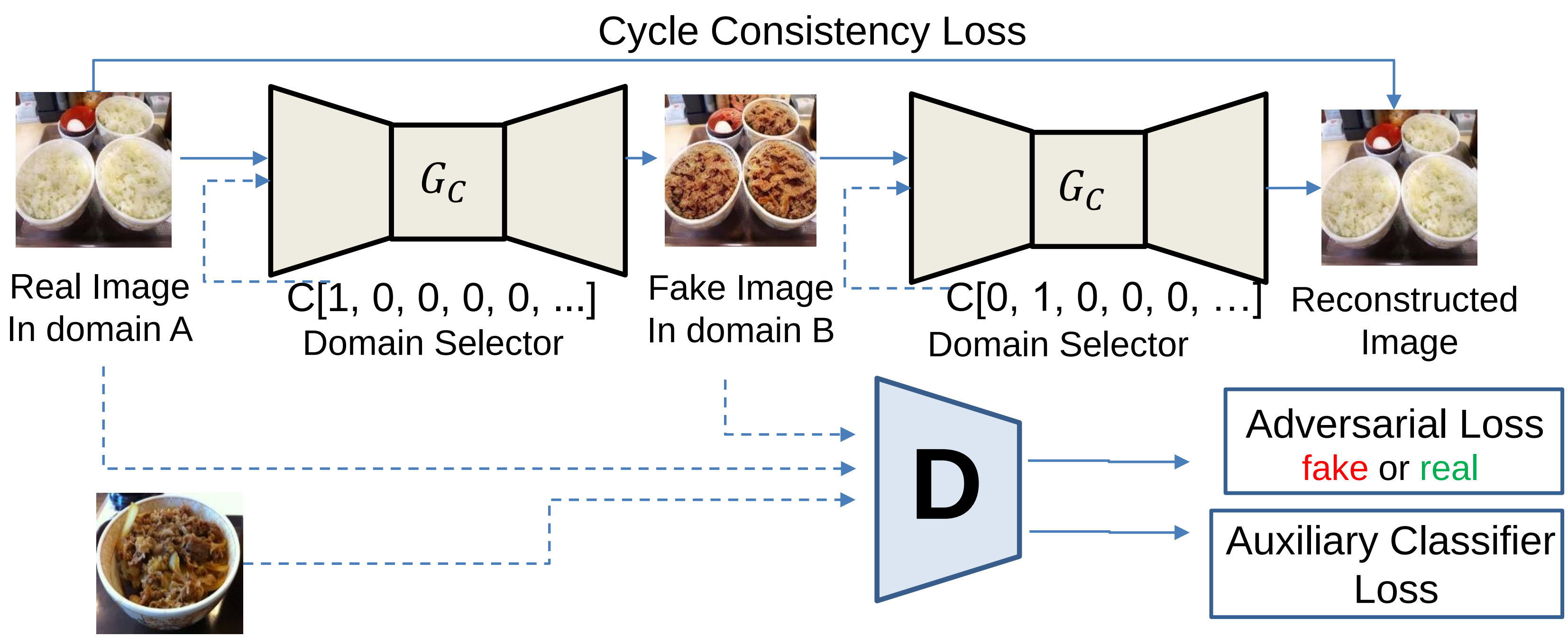
Food Image Translation

(FoodGAN)

Real-time Food Translation “MagicalRiceBowl”



Original StarGAN



Experiments

- We use foods, which have similar dish plates as target food category for simplification.

Selected 10 kinds of food category.


Curry

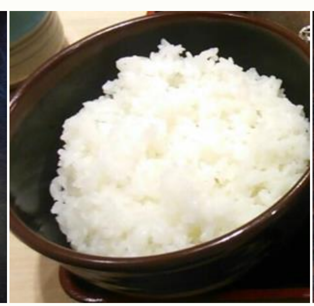

Fried rice


Beef bowl


Chilled noodle


Meat spaghetti


Ramen


Rice


Buckwheat noodle


Eel bowl


Fried noodle

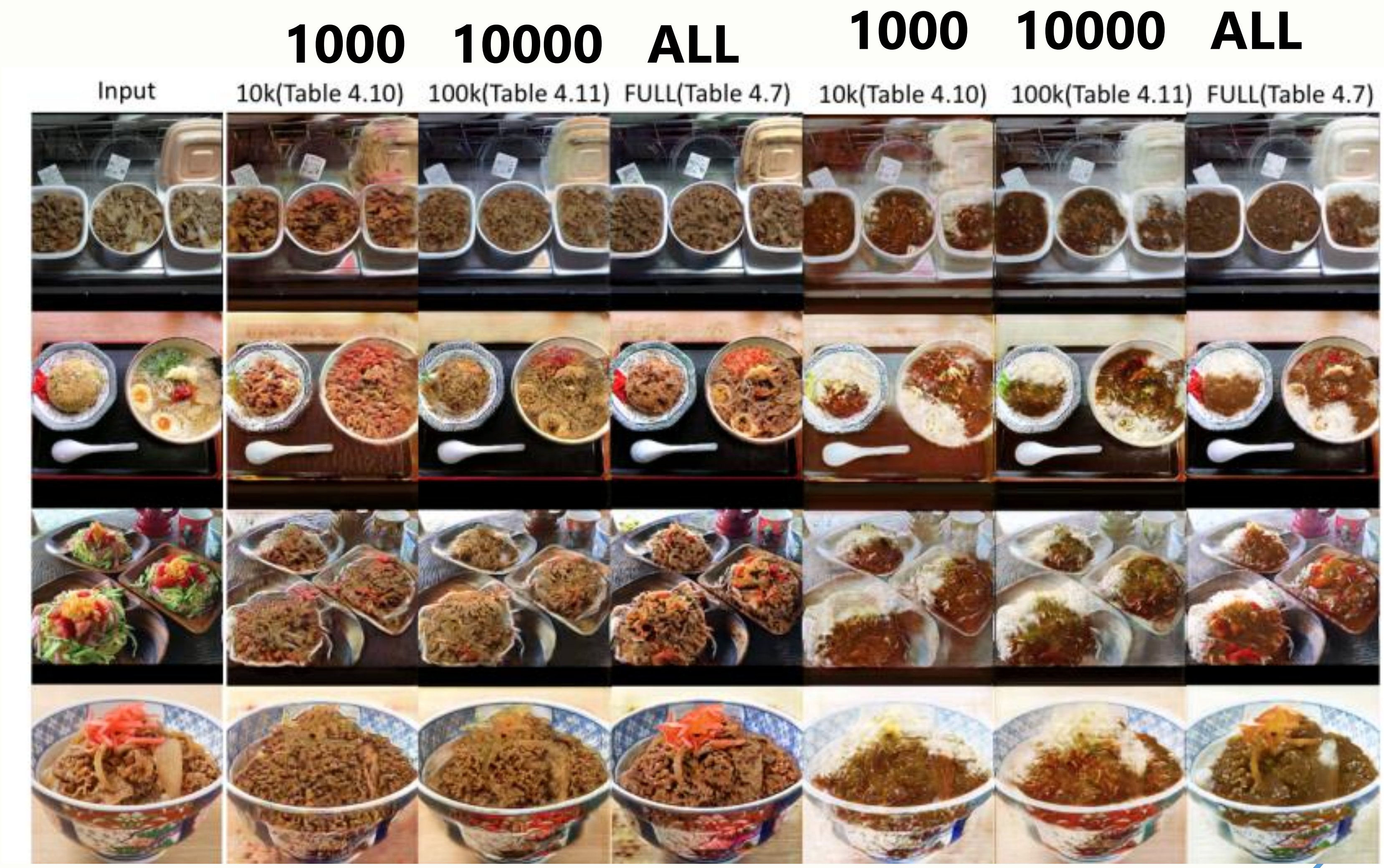
Experimanetal Data

- Total amount:
  - 230k images
  - Traing：0.9
  - Testing：0.1

Twitter Stream から収集

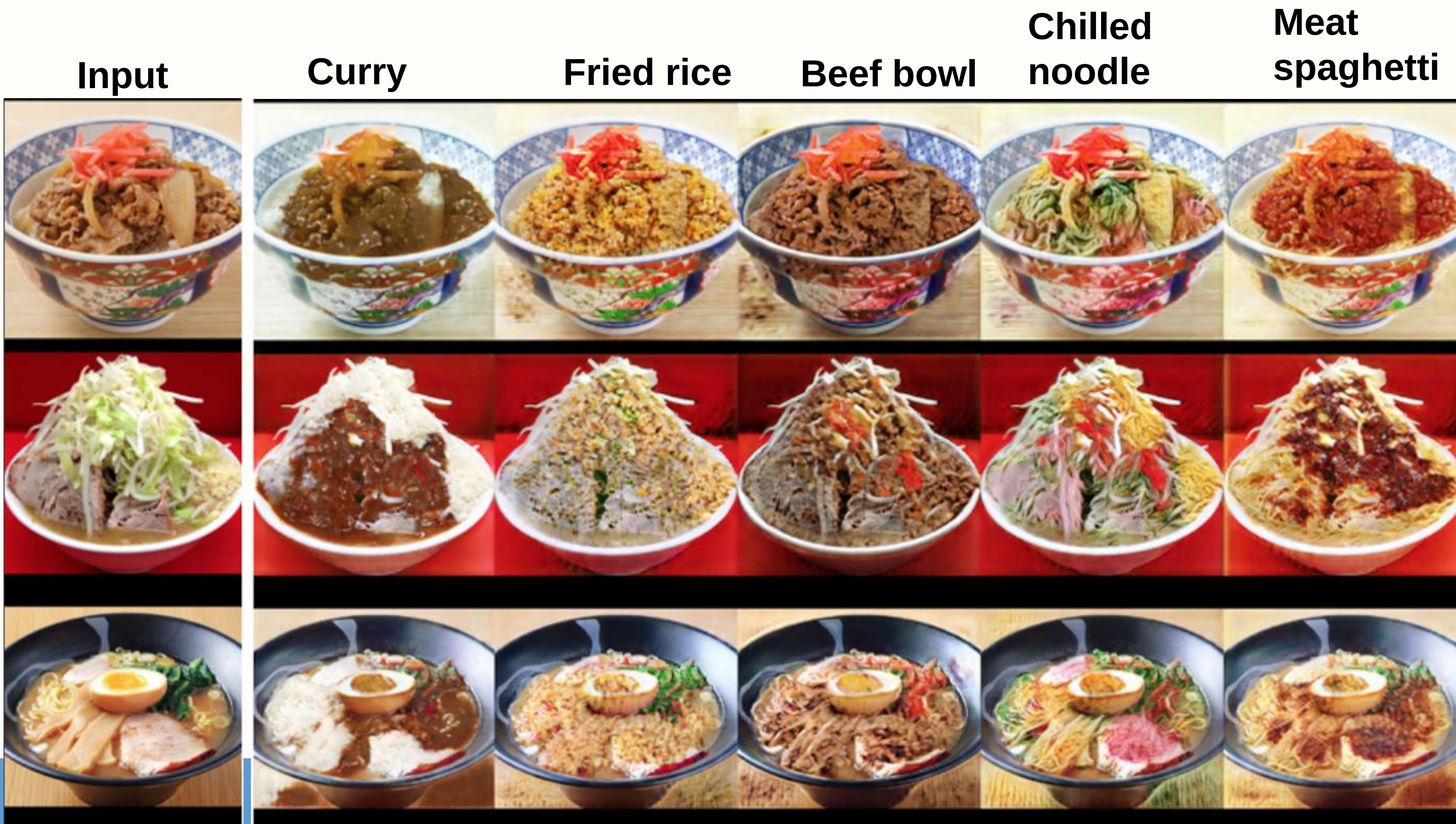
Target category	Image number
Chilled noodle	13,499
Meat spaghetti	7,138
Buckwheat noodle	3,530
Ramen	74,007
Fried noodle	24,760
Rice	21,324
Curry rice	34,216
Beef bowl	18,396
Eel bowl	5,329
Fried rice	27,854
total	230,053

The number of training images affects quality



Experimental results

- In case of one food included in an image



Experimental results

- In case of one food included in an image

Experimental results

- In case of multiple foods included in an image

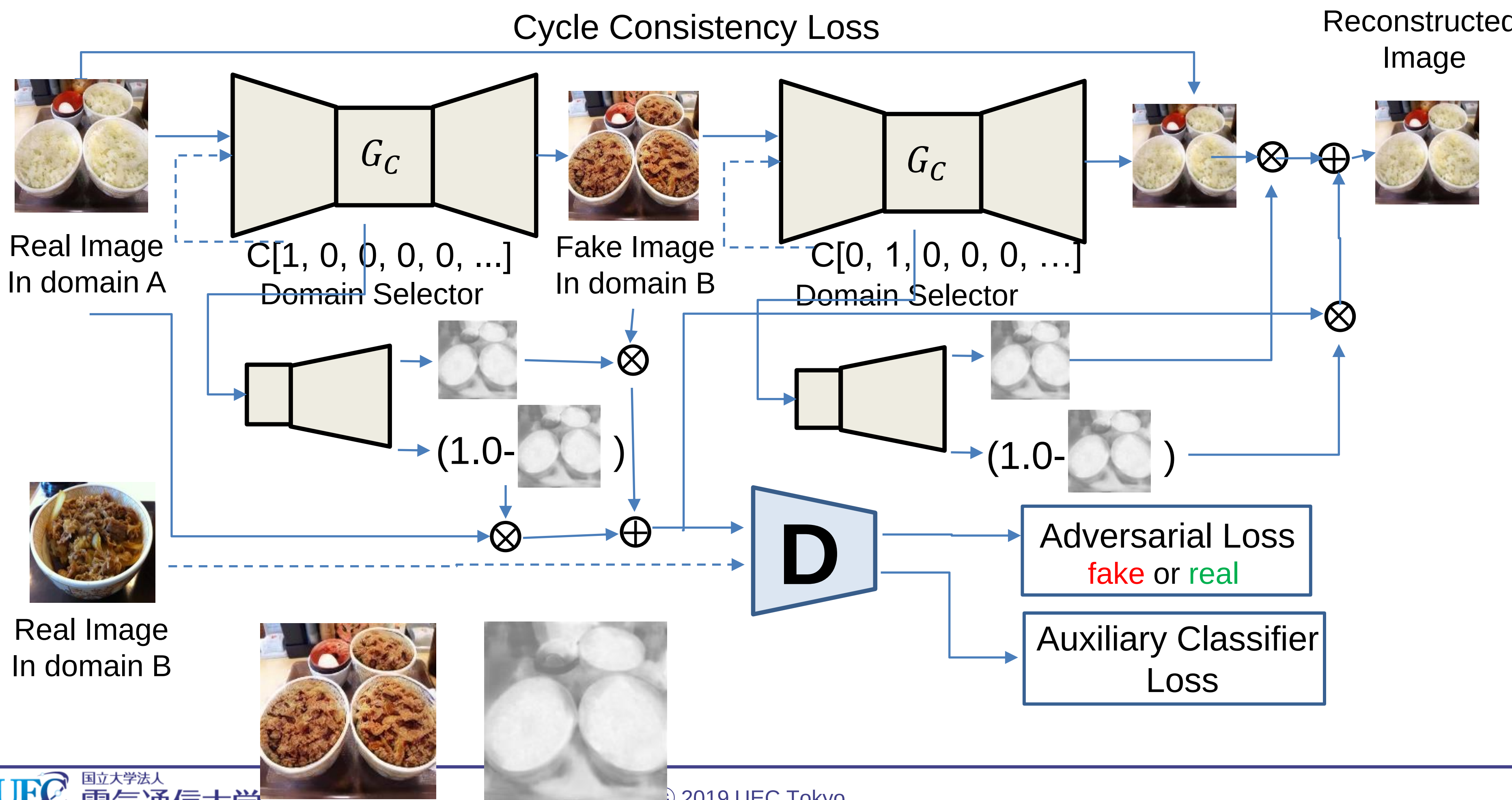
Input Curry Fried rice Beef bowl Chilled noodle Meat spaghet

Experimental results

- In case of multiple foods included in an image

input | ramen | rice | Buckwheat noodle | Eel bowl | Fried noodle

[extention] Attentional StarGAN (network of “MagicalRiceBowl”)



Attentional StarGAN

- w/o attention

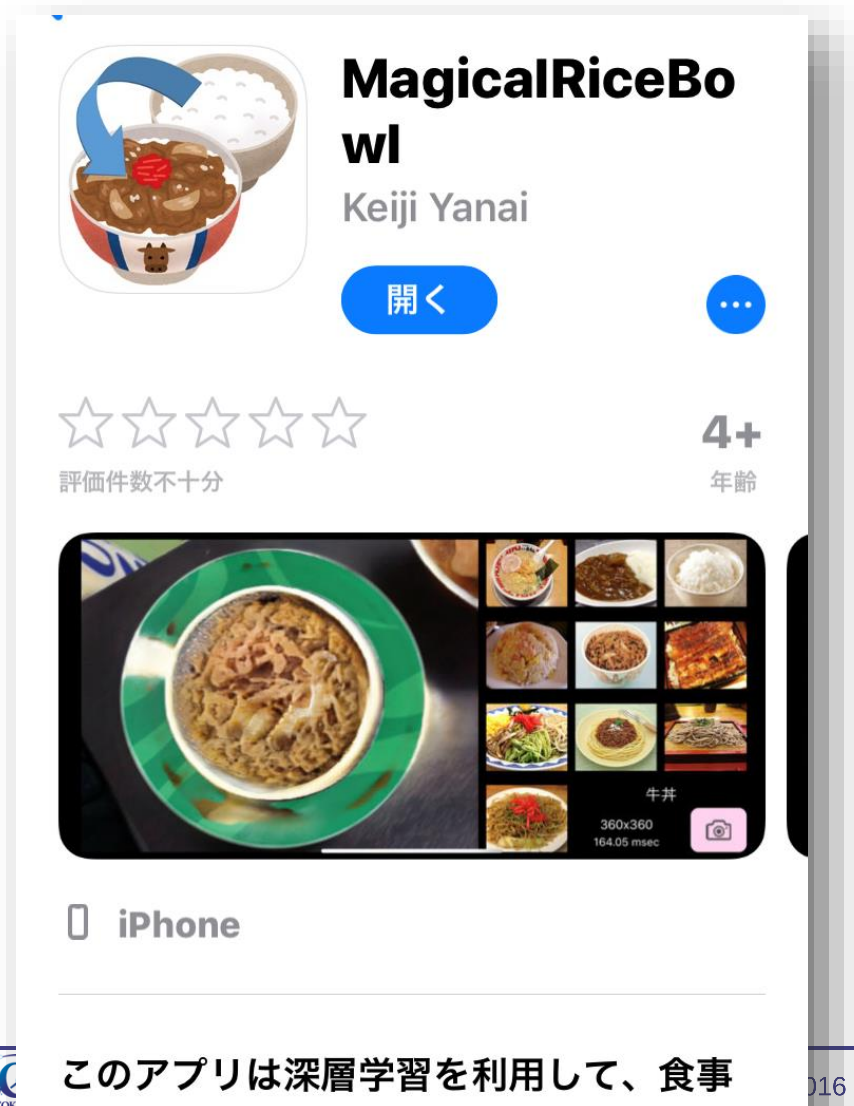
Background is also translated.

- w/ attention

Background is almost unchanged.

MagicalRiceBowl with the model with attention

MagicalRiceBowl on iOS App Store

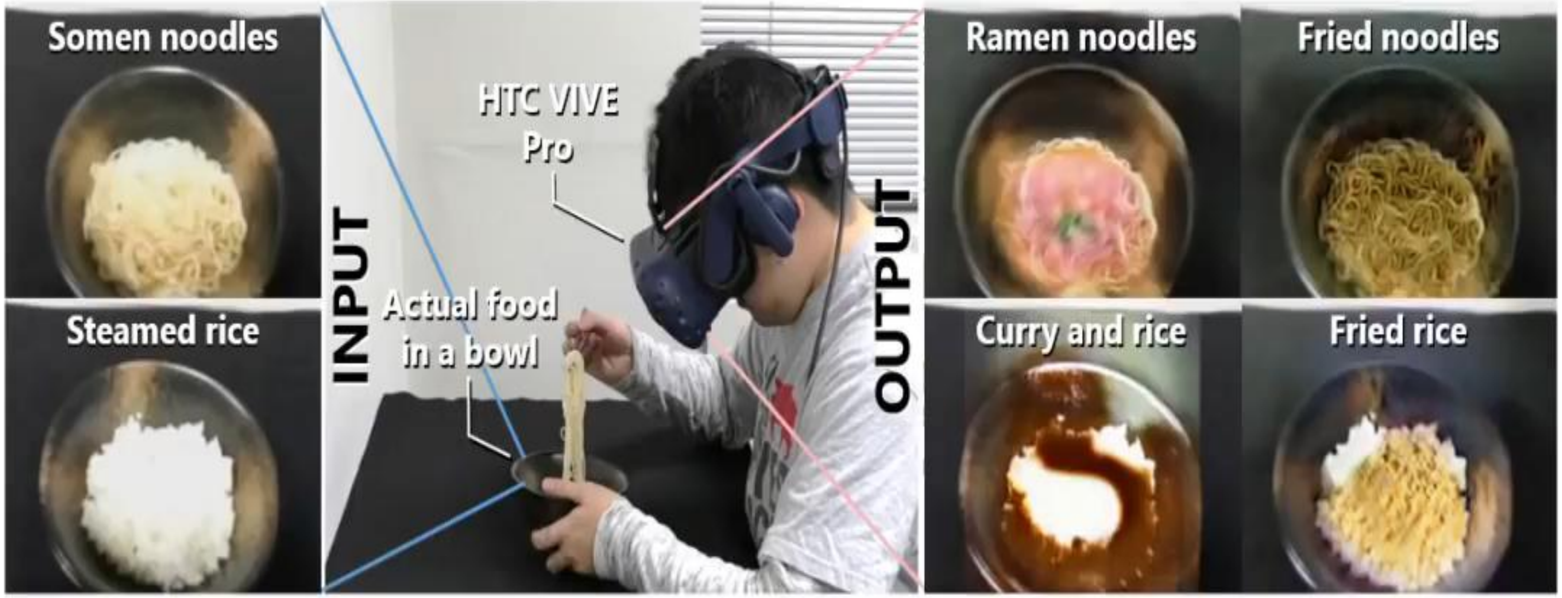


Only working
on iPhone 8 or more
(8/X/Xs/XR/11/11pro)

Using Neural Engine.

Food VR

Enchanting Your Noodles: GAN-based Real-time Food-to-Food Translation and Its Impact on Vision-induced Gustatory Manipulation



Kizashi Nakano[1] Kiyoshi Kiyokawa[1]
Daichi Horita[2] Keiji Yanai[2]
Nobuchika Sakata[1] Takuji Narumi[3]

[1] Nara Institute of Science and Technology
[2] The University of Electro-Communications
[3] The University of Tokyo

Cybernetics & Reality Engineering
Cyber Interface Lab.

https://www.youtube.com/watch?v=BJaQ5IF6iEI

食事変換＋HoloLens
(VR Restaurant)

newszero 2019/05/24

Conclusions

Conclusions

- **Introduced our Twitter photo mining works since Feb. 2011** (one month before the big earthquake)
  - **Geotagged tweet photo analysis**
    - Real-time geo-tweet photo mapping system [2012]
    - Event photo mining from geo-tweet photos [2012-2016]
    - Finding regional tendency on Twitter photos [2019]
  - **Twitter food photo mining**
    - Statistics on food image collection for 8 years [2012-]
    - Regional tendency on Twitter food photos [2019]
    - Applications of a large-scale Twitter food photoDB [2018-]
      - Food image translation by GAN, mobile app., food VR

“Thank you” by fake ramen images