

Food Understanding from Digital Images

Giovanni Maria Farinella

gfarinella@dmf.unict.it

www.dmf.unict.it/farinella

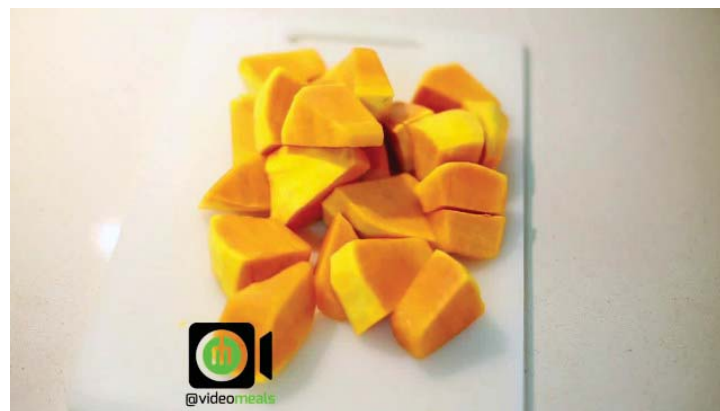
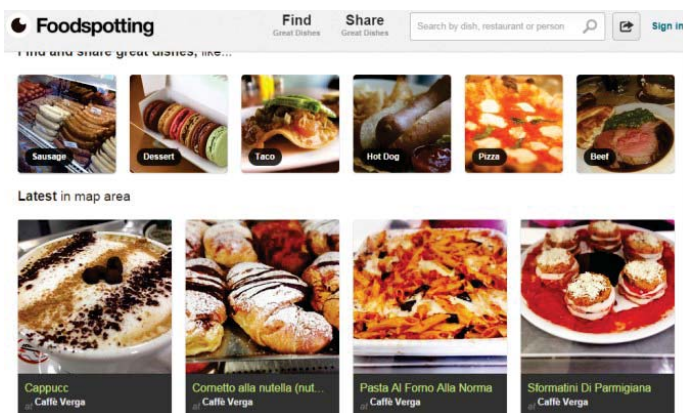
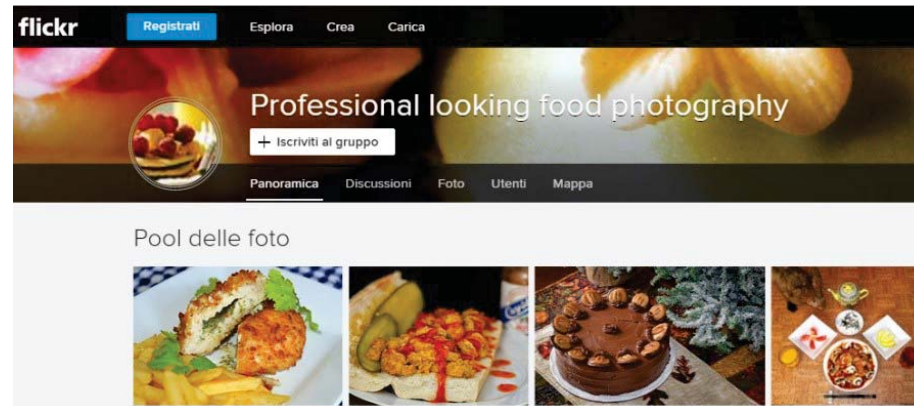
People Love Food!

Food is nowadays one of the most photographed objects.



<http://iplab.dmi.unict.it/DCT-GIST/>

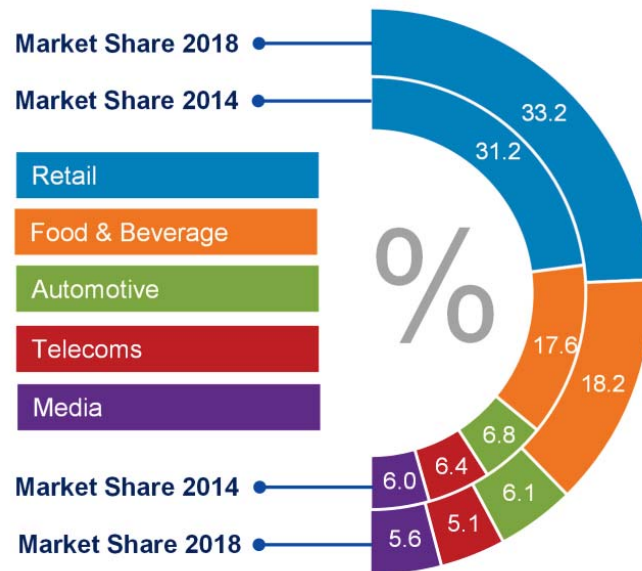
People Love to Share Food Images/Videos!



The number of food images and videos on the web is increasing and novel social networks for food lovers (e.g., foodspotting) are more and more popular.

Food and Social

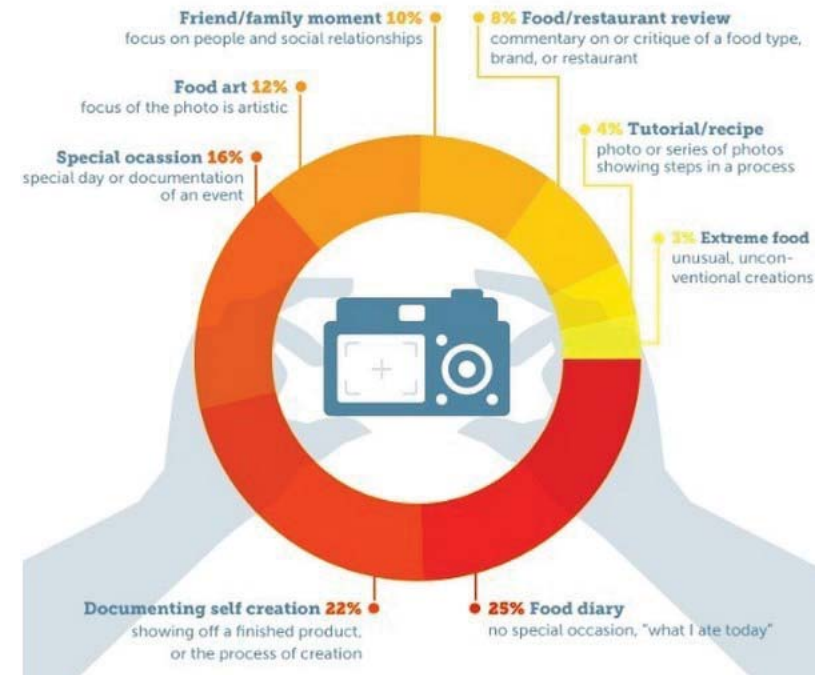
US Social Media Advertising Spending, by Five Largest Industries, 2014–2018



Source: Socintel360

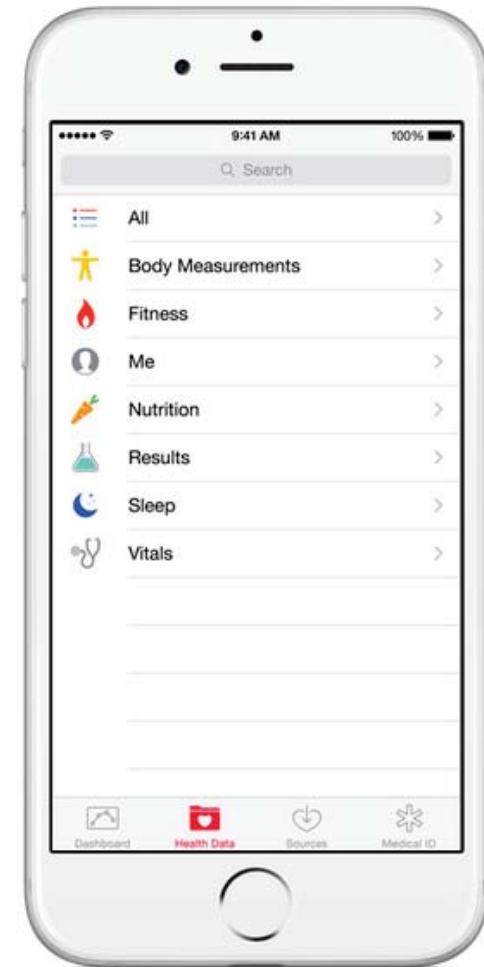
© Socintel360

Why do people share food photos?



Food Images and Industry

- **Roche:** Diabetes Self-Management and Carbohydrate Counting
- **Shazam:** recently announced food recognition
- **SRI International - Meal Snap:** Meal Snap lets you take pictures of the meal you eat, and then magically tells you what food was in your meal. Oh yeah, we estimate how many calories you ate too. Food tracking has never been easier.
- **Ellips** - Vision systems for inspecting and grading fruits and vegetables
- **Montrose Technology:** Vision systems for the baked goods industry. Systems monitor bake color, shape, and size of bread, cookies, tortillas, etc.
- **Jawbone:** tools to manage weight, sleep better and be more fit.
- **SCiO:** near-infrared spectroscopy to analyze objects. Current applications for SCiO include analyzing the caloric and nutritional content of food.
- **TellSpecopedia:** evidence-based information on thousands of global food ingredients and their impact on health.



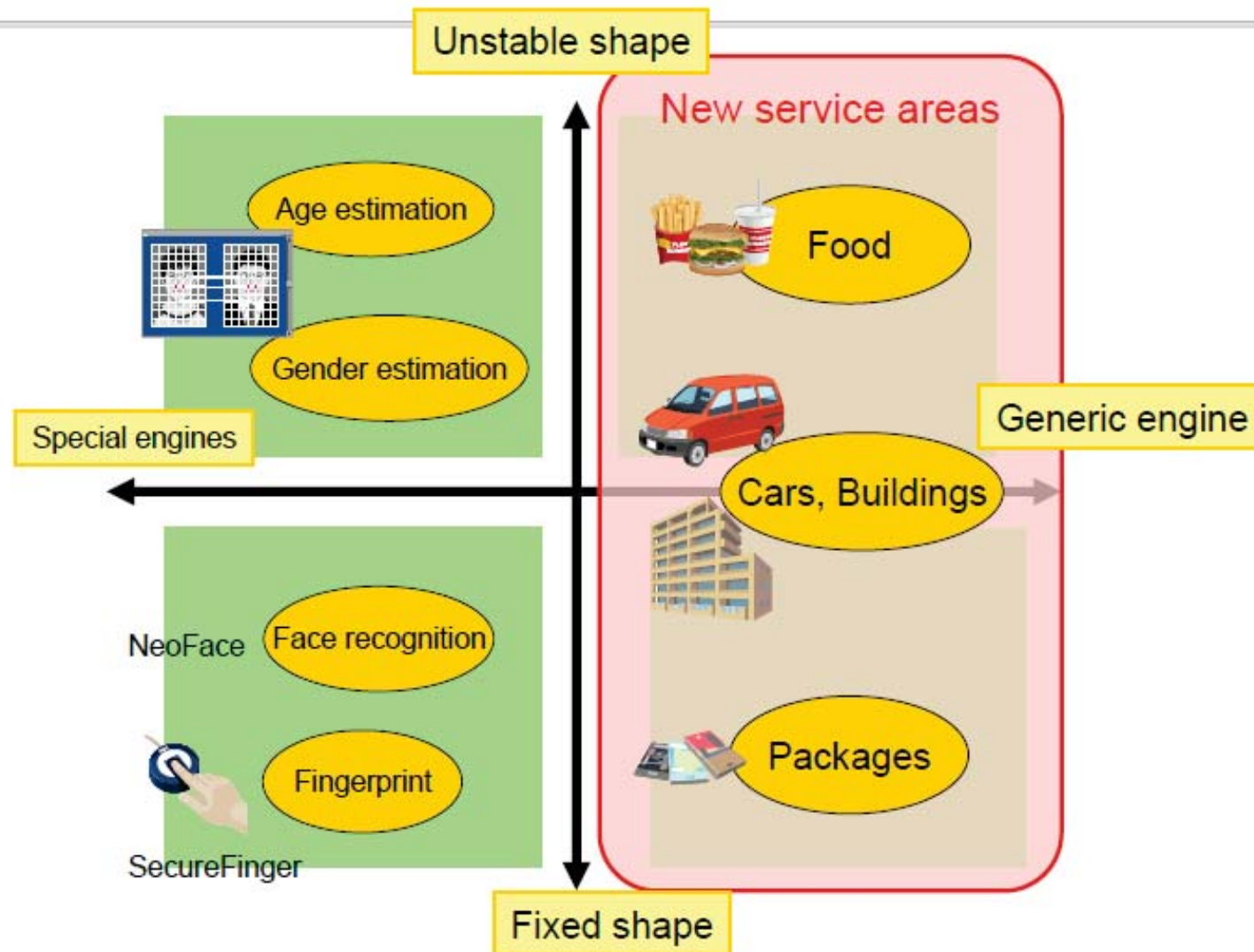
Veggie Vision – IBM (1996)

Veggie Vision: A Produce Recognition System

R. M. Bolle, J. Connell, N. Haas, R. Mohan, G. Taubin
Proc. of the Third IEEE Workshop on Applications of
Computer Vision (WACV-96),
Sarasota FL, pp. 224-251, December 1996.



New service areas







RESEARCH & INNOVATION

Horizon Prizes

[European Commission](#) > [Research & Innovation](#) > [Horizon Prizes](#) > [The Prizes](#) > [Food Scanner](#)

Food Scanner - € 1 Million

[Introduction](#)

[Rules & Guidance](#)

Given the increase in food-related health problems, the challenge set for this €1 million prize is to develop an affordable and non-invasive mobile solution that will enable users to measure and analyse their food intake. This solution will especially benefit people with conditions such as obesity, allergies or food intolerance.

Challenge

The **Horizon Prize for a food scanner** will be awarded to a mobile solution that analyses precisely, quickly and efficiently food composition, nutrition facts and potentially harmful ingredients such as allergens. It should also be able to provide feedback to users regarding their health and lifestyle.

MADIMA2015

Full paper submission
May 18th, 2015

1st International Workshop on Multimedia Assisted Dietary Management

In conjunction with the 18th International Conference on Image Analysis and Processing

Genova, Italy, September 8th 2015



Topics

Topics of interest include (but are not limited to) the following:

- Ubiquitous and mobile computing for dietary assessment
- Computer vision for food detection, segmentation and recognition
- 3D reconstruction for food portion estimation
- Augmented reality for food portion estimation
- Wearable sensors for food intake detection
- Computerized food composition (nutrients, allergens) analysis
- Multimedia technologies for eating monitoring
- Smartphone technologies for dietary behavioral patterns
- Food multimedia databases
- Evaluation protocols of dietary management systems
- Multimedia assisted self-management of health and disease

Workshop chairs

Stavroula Mougialakou, University of Bern, Switzerland

Giovani Maria Farinella, University of Catania, Italy

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

INVITED SPEAKERS

*Prof. Edward Sazonov, University of Alabama, USA

- Talk Title: **Wearable solutions for detection and characterization of food intake**

* Dr. Gerald Cultot, Research Programme Officer – European Commission

- Talk Title: **The role of food scanning in the portfolio of the EU well-being**

www.madima.org

Food Understanding from Images

Automatic answering questions about food from images

- Is there any food in the image?
- Where is the food in the image?
- Which class of food is in the image?
- Which are the ingredients of the detected food?
- Does it contain allergic ingredients (e.g. nuts)?
- Which is the volume of the detected food?
- How many calories?
- Can you recommend me food based my preferences?

:

:

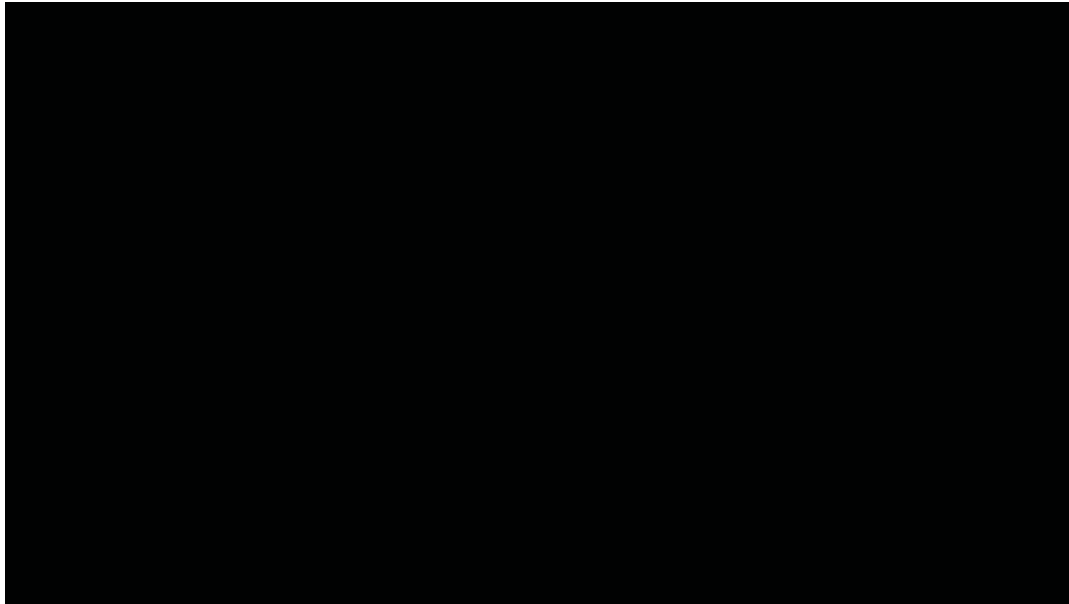


Packed Food

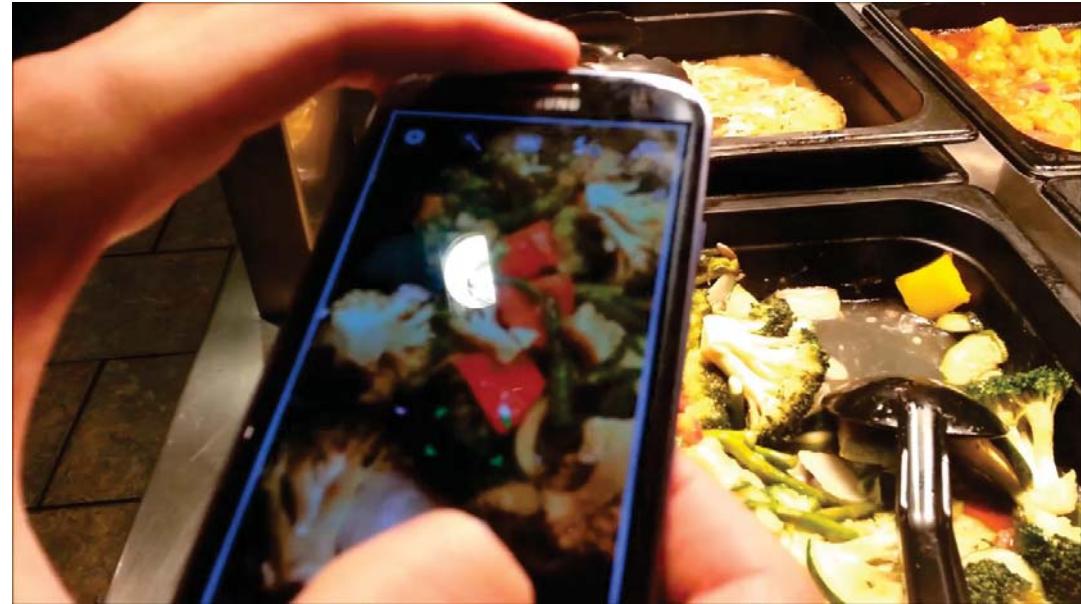


Unpacked Food

Food/Beverage Recognition



Packed Food/Beverage



Unpacked Food

Degree of Challenge

PFID Pittsburgh Fast-Food Image Dataset

Carnegie Mellon



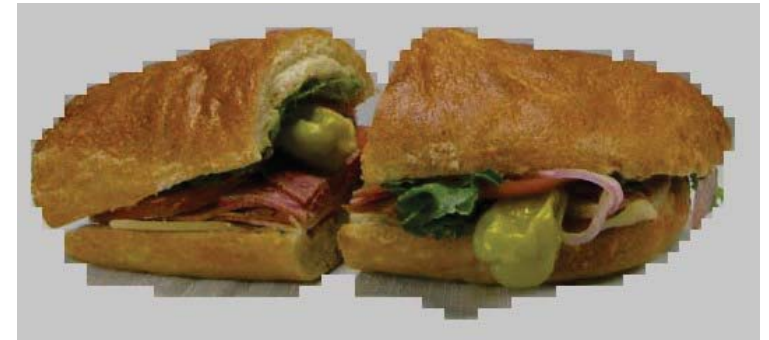
This dataset contains 1098 images belonging to 61 categories (can be grouped in 7 major categories). There are 3 instances of each food item. Each instance is taken by 6 point of view. Clear testing protocol.



Sandwich	Salad & Sides	Bagel	Donut	Chicken	Taco	Bread & Pastry	Average
228	36	24	24	24	12	18	52.3

Drawbacks

- Small Dataset (only 1098 images)
- Only fast-food images
- Lab settings
- Low resolution images
- Not suitable for the study of food representation



UNICT-FD889

A Benchmark Dataset to Study the Representation of Food Images

We introduce a new food images dataset composed by 3583 images related to 889 distinct dishes of food of different nationalities (e.g., Italian, English, Thai, Indian, Japanese, etc.).

Food images have been acquired by users during meals with a smartphone (e.g., iPhone 3GS or iPhone 4) in unconstrained settings (e.g., different backgrounds and light environmental conditions).

Images acquired by users during meals with iPhone

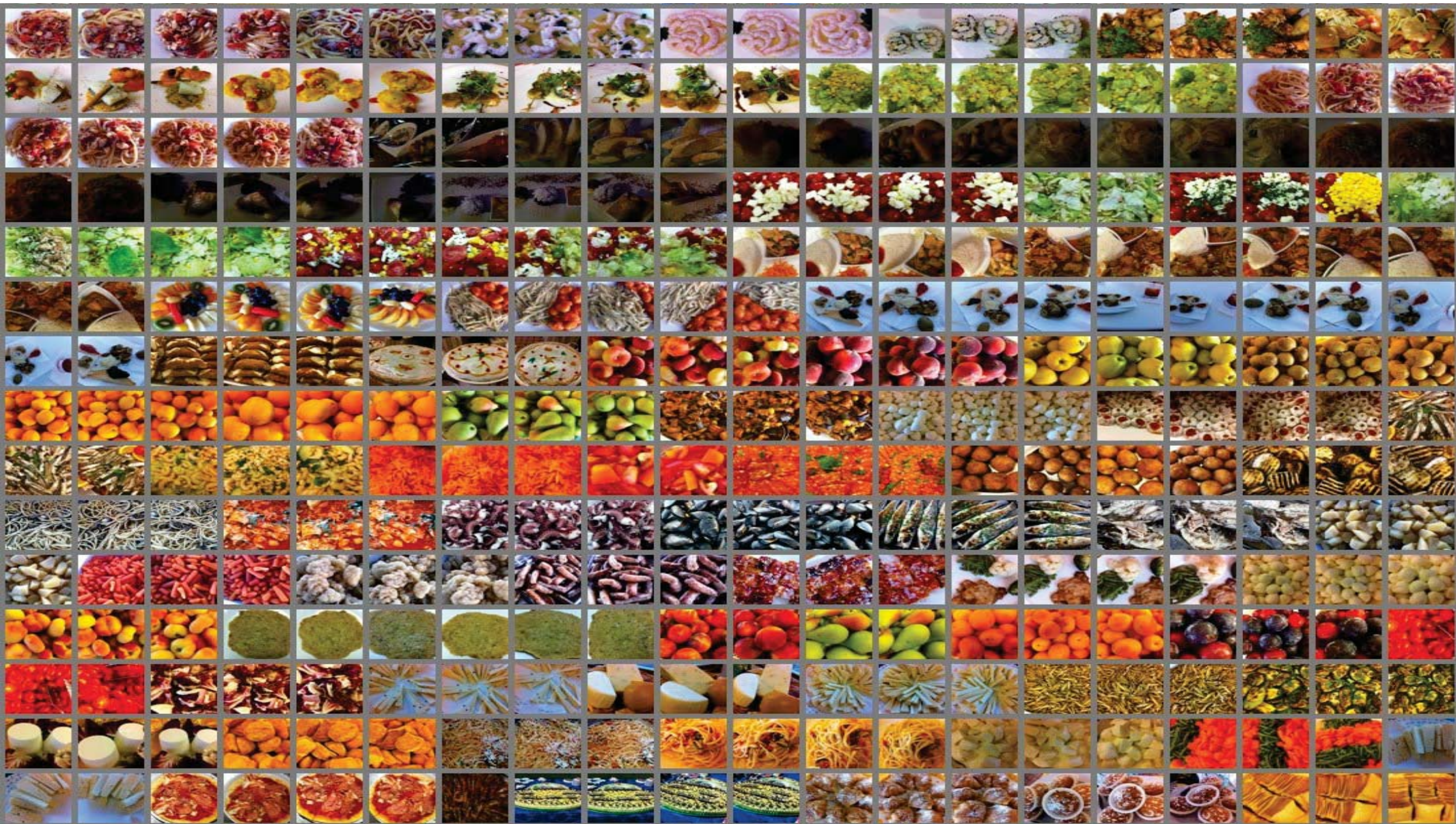


December 2010 - September 2014

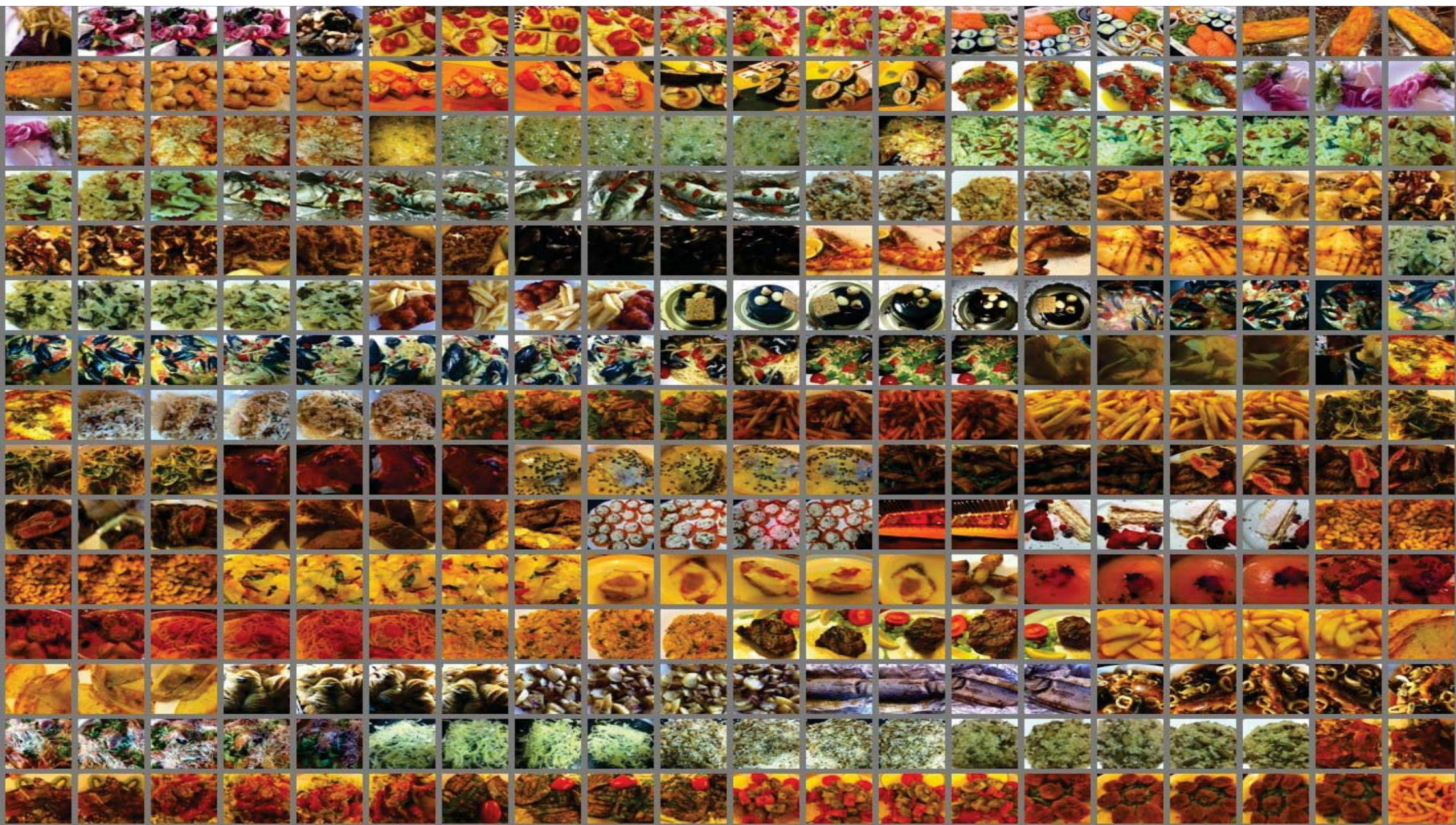
The Dataset

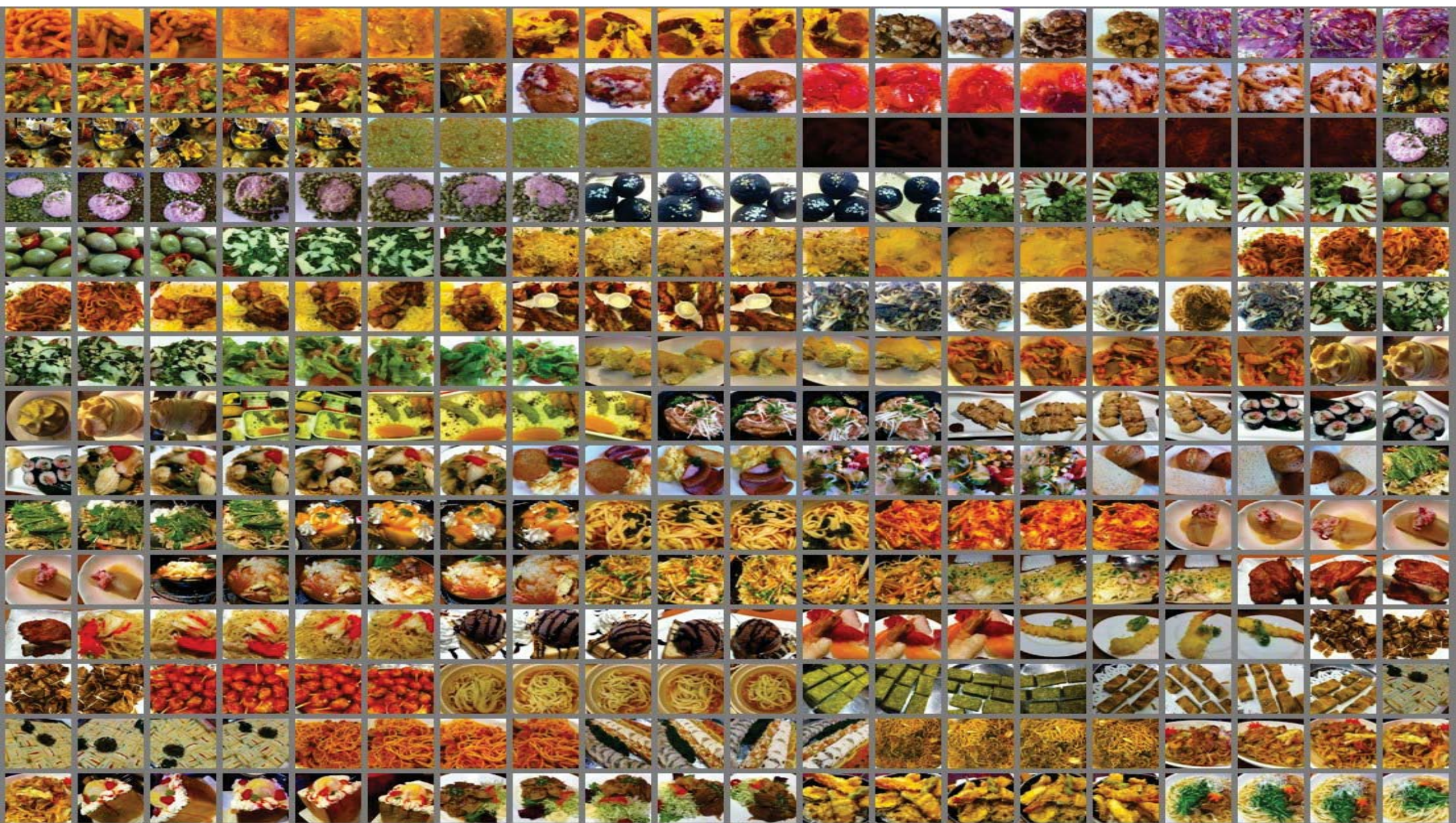
<http://iplab.dmi.unict.it/UNICT-FD889/>



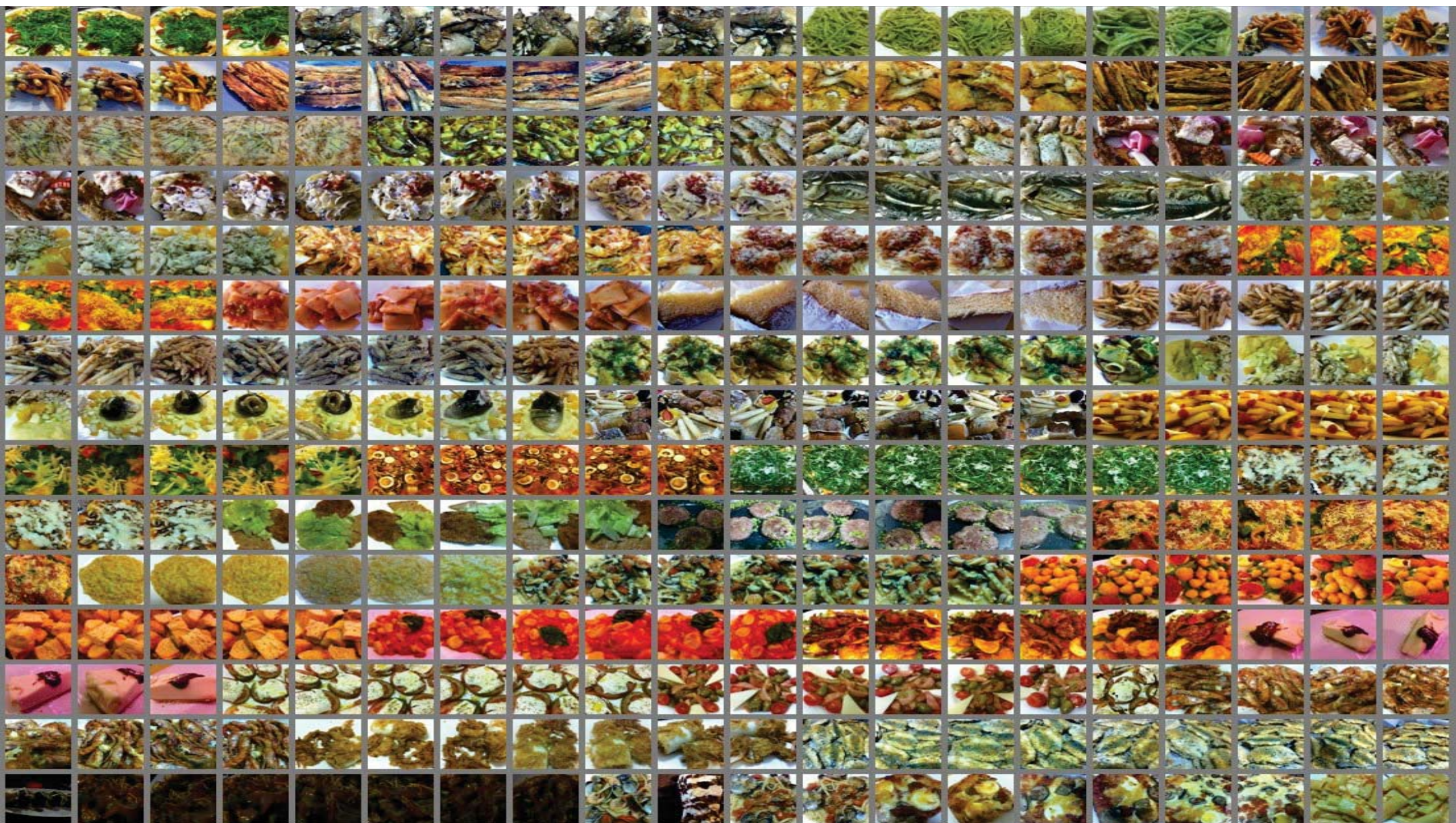






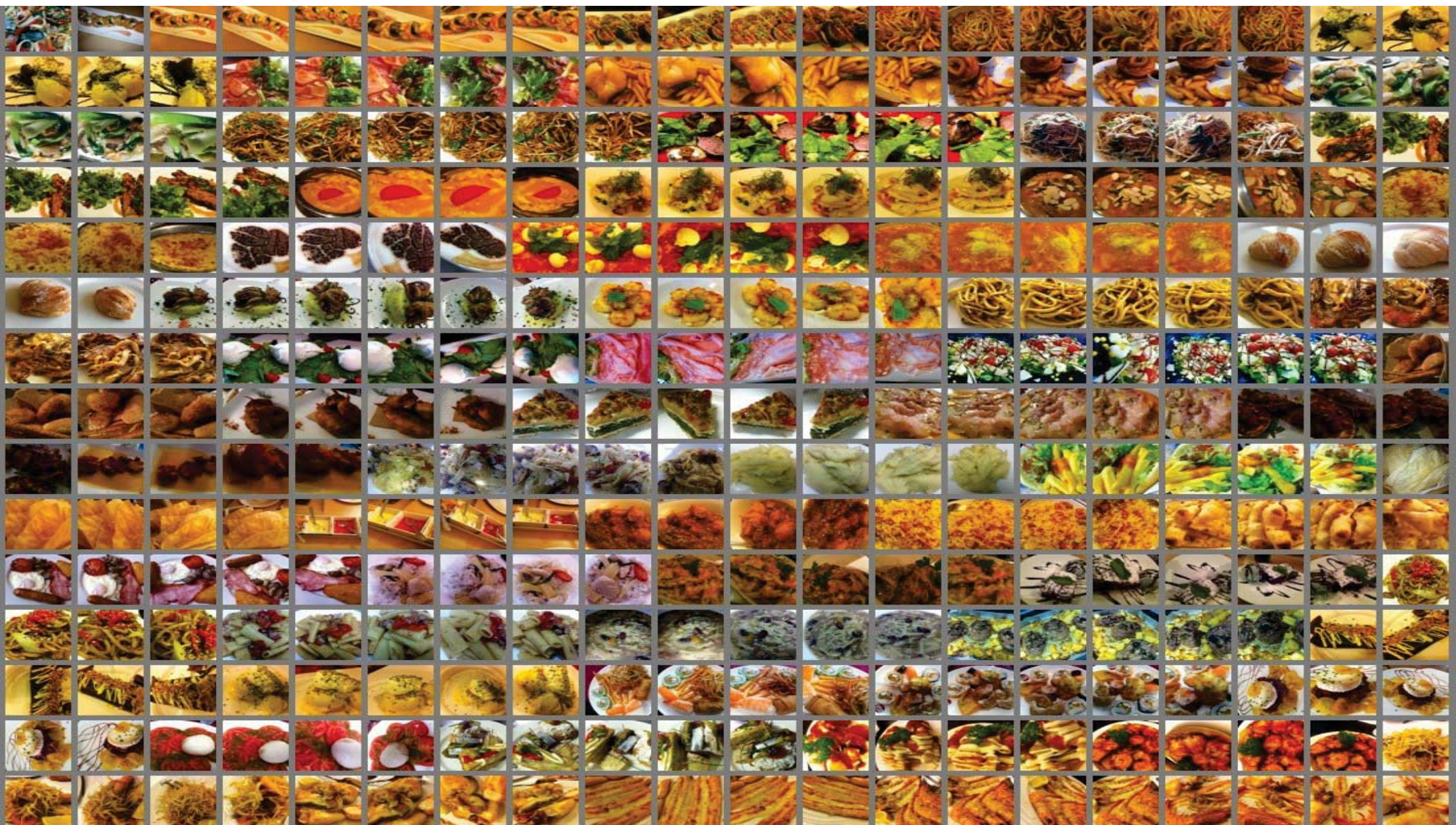




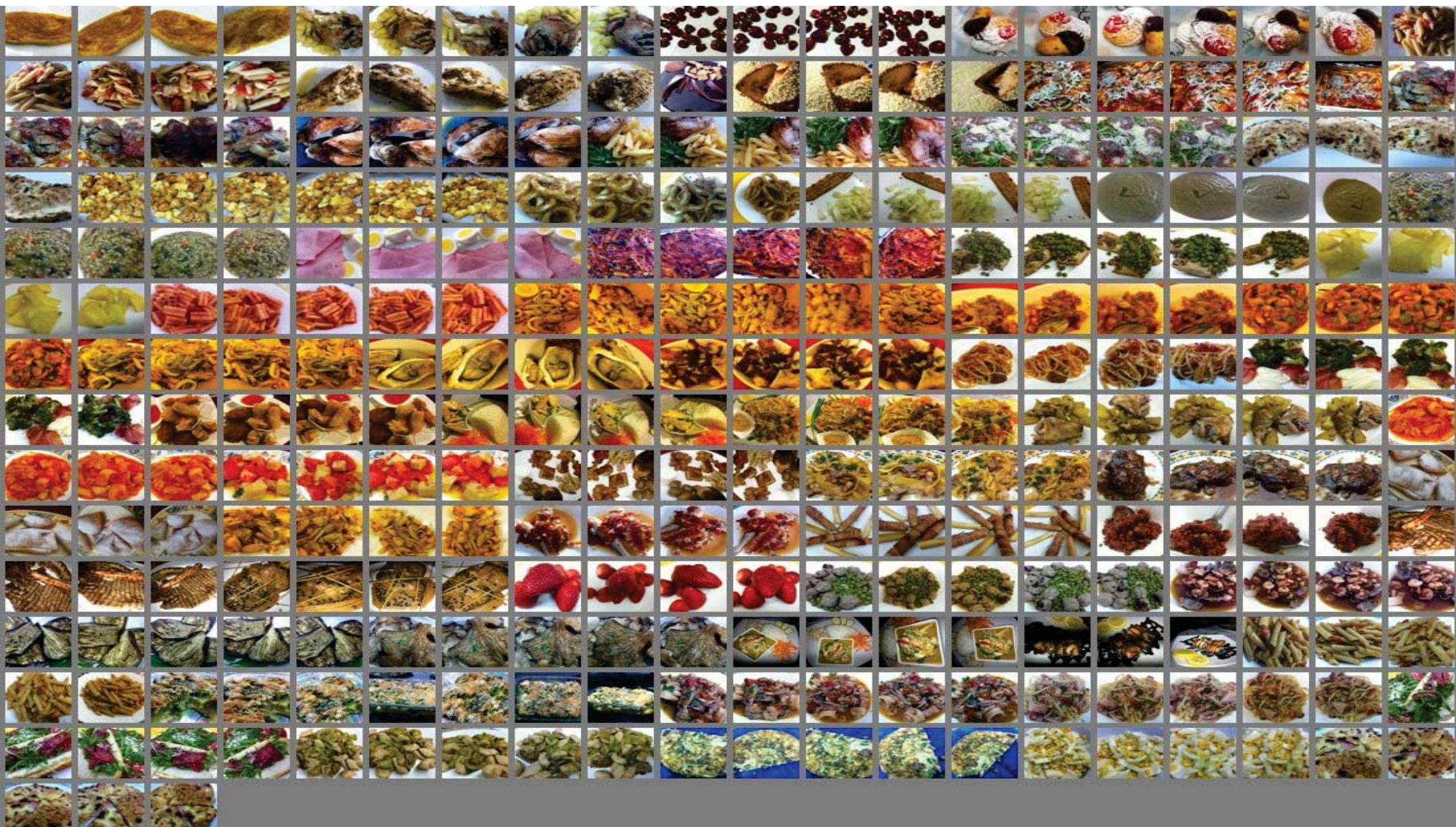












How to Represent Food Images?

- Bags of Textons [1,2,3]
- PRICoLBP [4]
- SIFT [5]

[1] G. M. Farinella et al., *Representing Scenes for Real-Time Context Classification on Mobile Devices*, Pattern Recognition, Vol. 48, Issue 4, 2015

[2] G. M. Farinella et al., *Classifying food images represented as bag of textons*, IEEE International Conference on Image Processing, 2014

[3] G. M. Farinella, D. Allegra, F. Stanco, "A Benchmark Dataset to Study the Representation of Food Images", International Workshop on Assistive Computer Vision and Robotics, Zürich, Switzerland, 12 September 2014

[4] Qi X. et al., *Pairwise rotation invariant co-occurrence local binary pattern*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014

[5] Lowe D.G., *Distinctive Image Features from Scale-Invariant Keypoints*, International Journal of Computer Vision, 60(2), 91-110, 2004

Bag of Words Representation

Medicine

Of all the sensory impressions proceeding to the brain, the visual experiences are the dominant ones. Our perception of the world around us is based essentially on the messages that reach our eyes. For a long time, the retinal image was considered as a simple projection on a screen. As a movie screen receives a visual image, the retina receives a visual image. It was not until the discovery of the visual pathway, the optic nerve, that we began to know the more complex nature of the visual perception. Hubel and Wiesel, following the work of the various centers of the visual cortex, demonstrated that the message about the image falling on the retina undergoes a complex analysis in a system of nerve cells stored in columns. In this system each cell has its specific function and is responsible for a specific detail in the pattern of the retinal image.

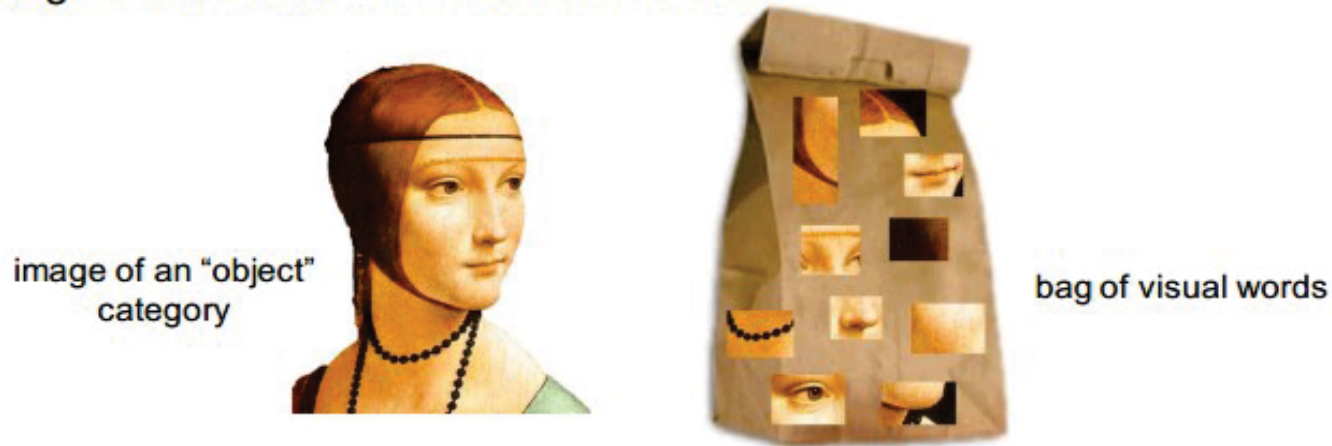
Business

China is forecasting a trade surplus of \$90bn (£51bn) to \$100bn this year, a threefold increase on 2004's \$32bn. The Commerce Ministry said the surplus would be created by a predicted 30% increase in exports to \$750bn, compared with \$575bn in 2004. Imports are expected to be \$660bn. The increase in the trade surplus will annoy the US, which has long complained about China's deliberate export-led growth strategy. China's deliberate export-led growth strategy has annoyed the US, which has long complained about China's deliberate export-led growth strategy. China's deliberate export-led growth strategy has annoyed the US, which has long complained about China's deliberate export-led growth strategy.

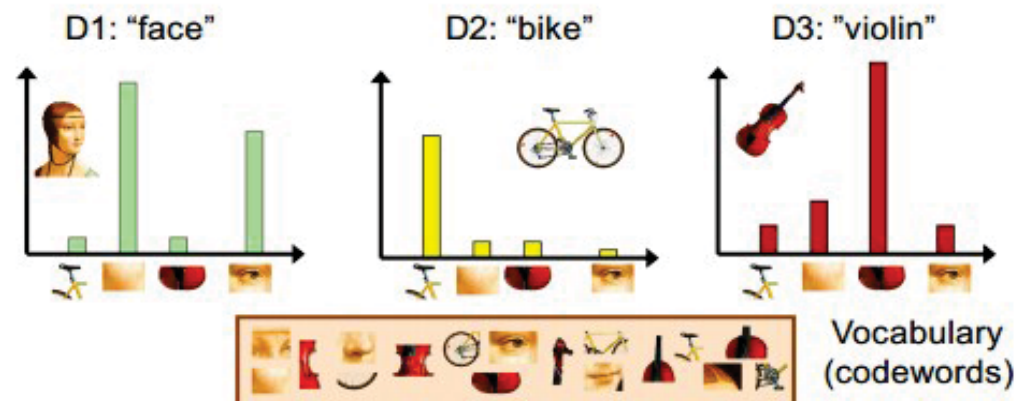
A text - such a sentence or a document - is represented as an unordered collection of words, (belonging to a vocabulary) disregarding grammar and even word order.

Bag of Words Representation

- An image can be treated as a document, and features extracted from the image are considered as the "visual words"...



Bag of (visual) Words: an image is represented as an unordered collection of visual words



Textons Vocabulary

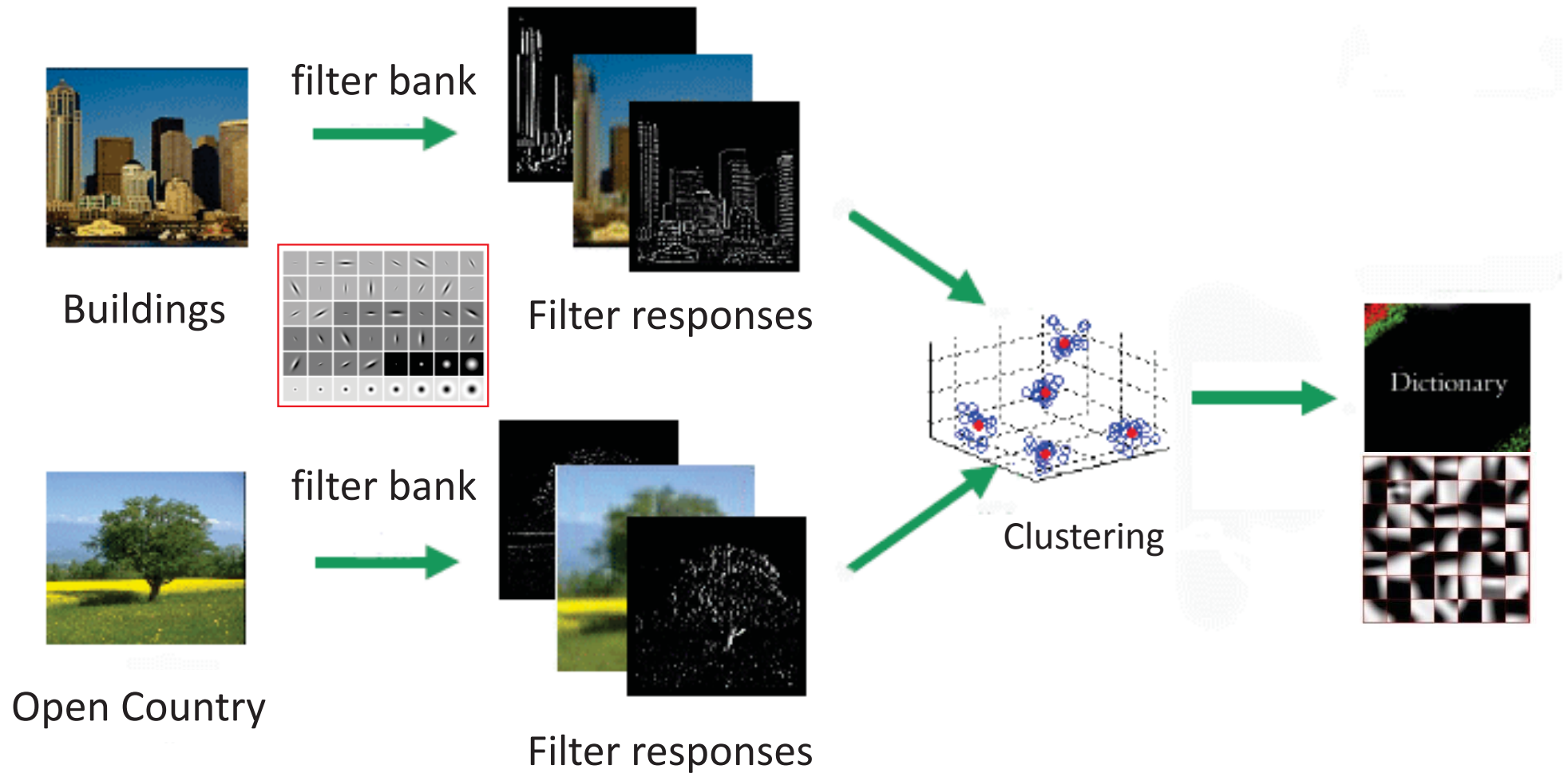
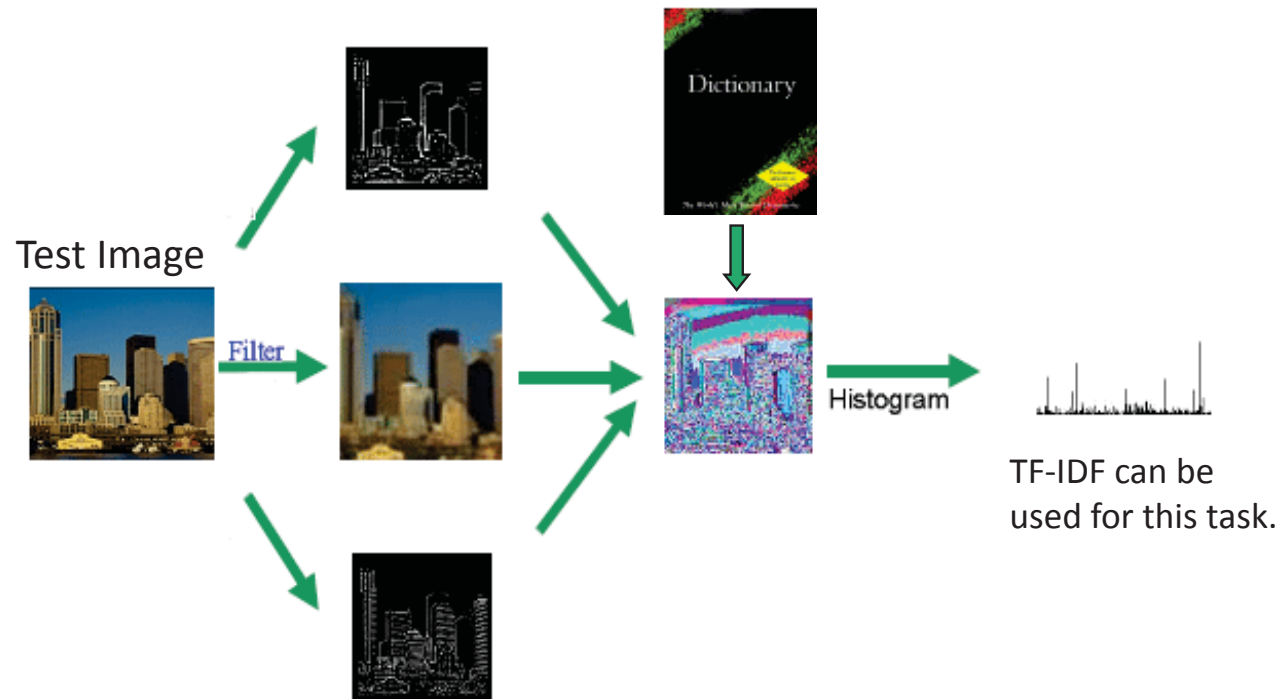


Image Representation



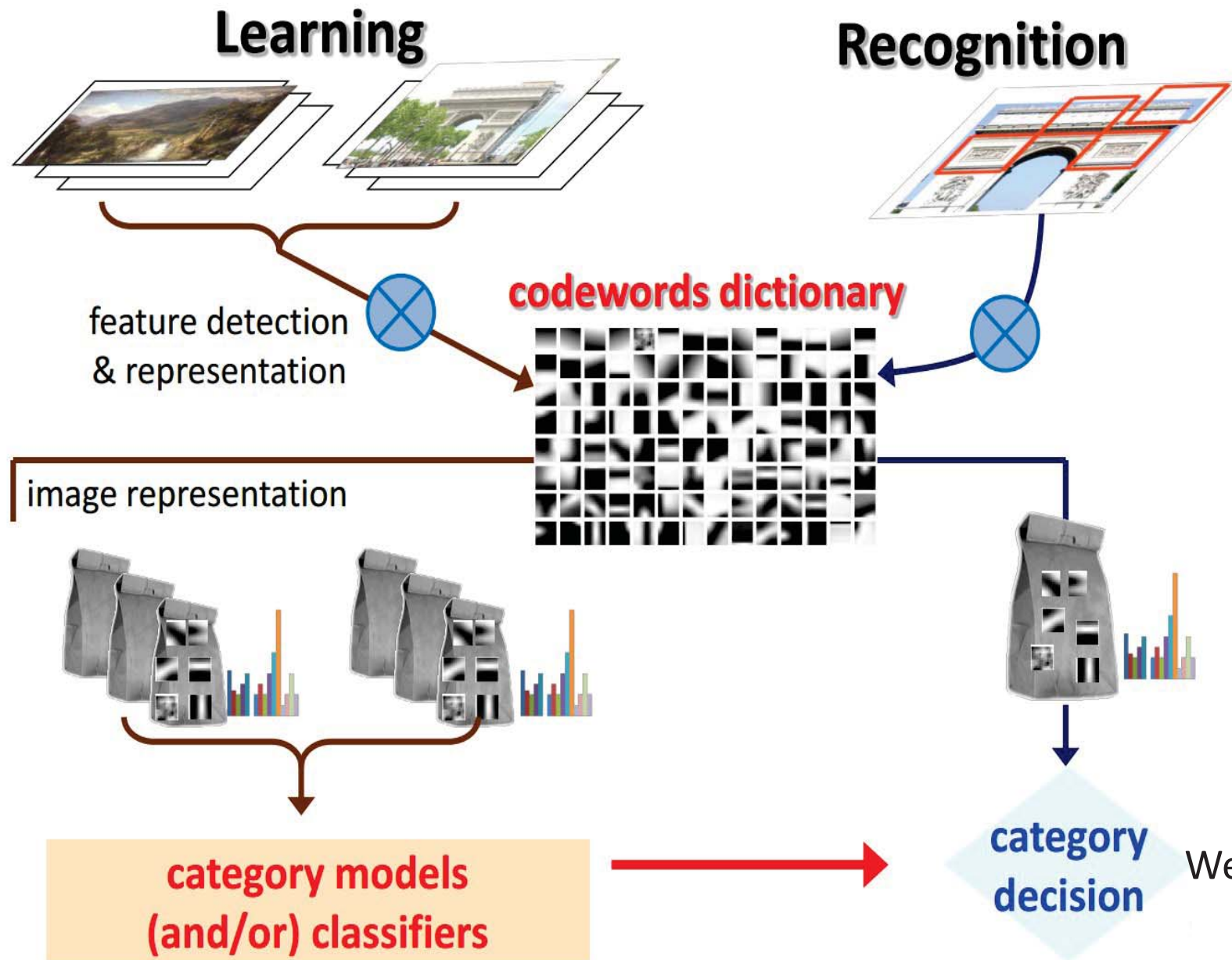
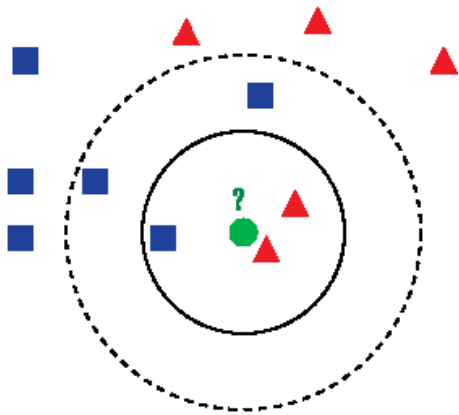


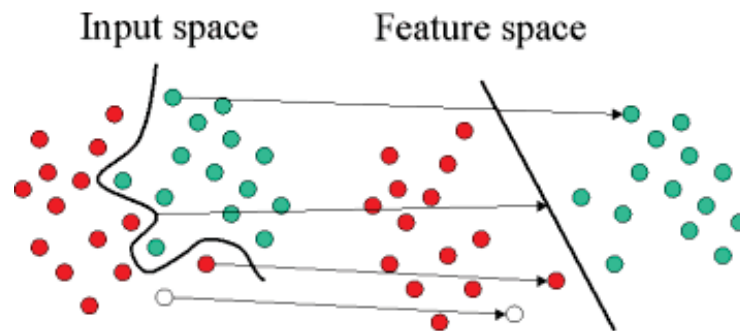
Image Representation Parameters

- Impact of vocabulary modality (Global vs Class Based)
- Impact of vocabulary size (Small vs Large)
- Impact of texture information (Textons/PriCoLBP vs SIFT)
- Impact of color information (Gray vs Color)

Classification and Evaluation



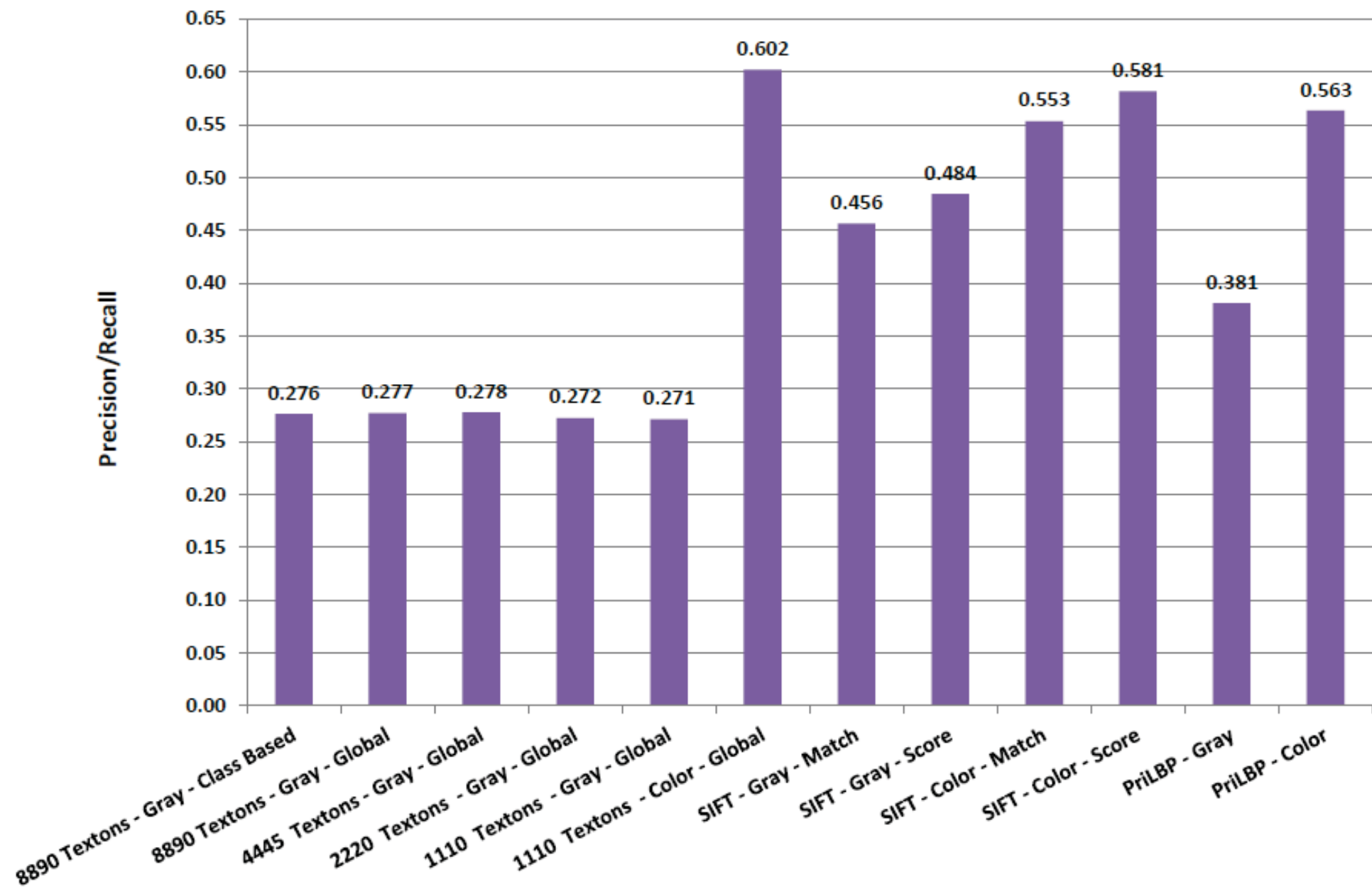
K-Nearest Neighbors

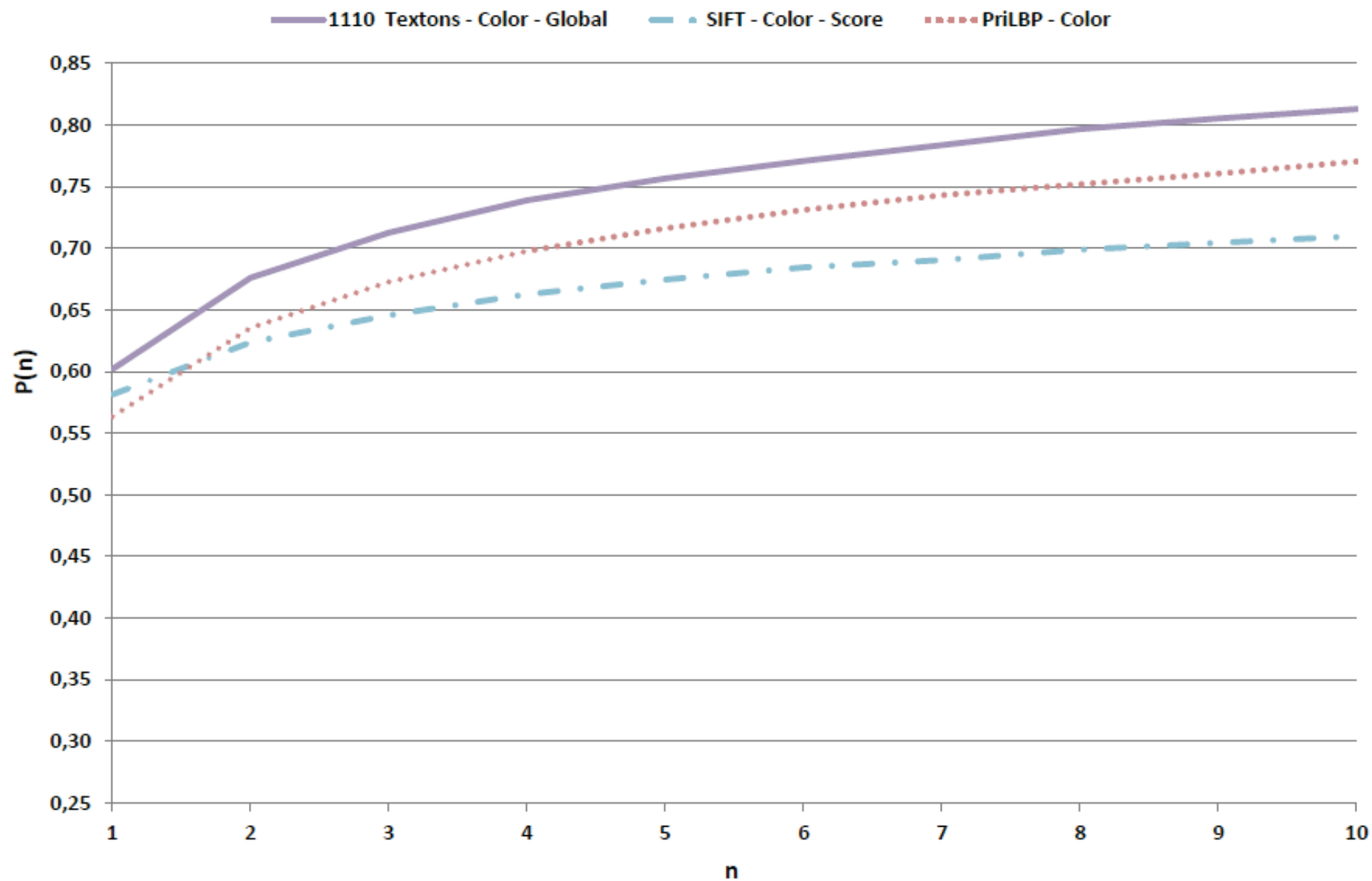


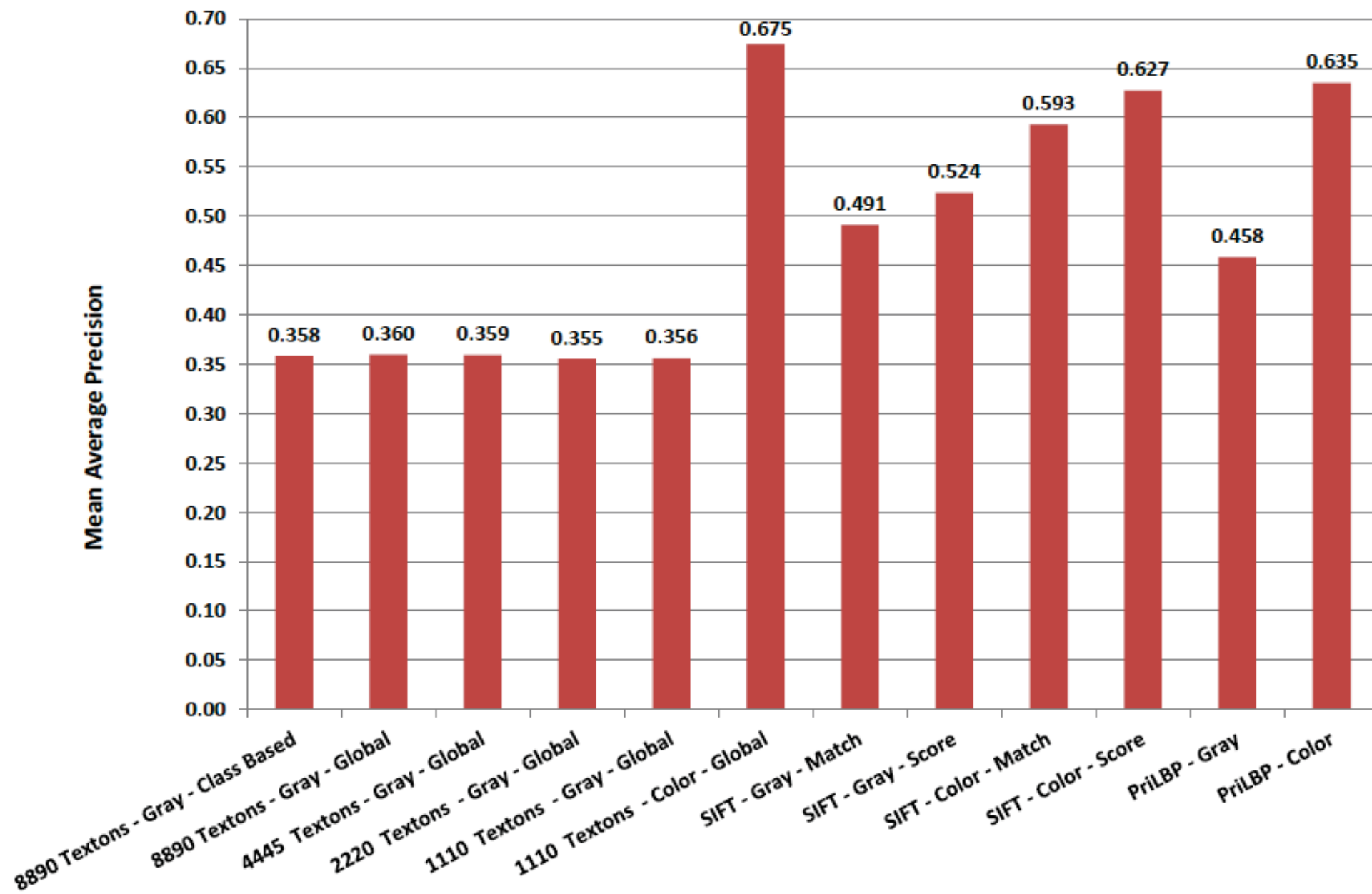
Support Vector Machine

$$P(n) = \frac{Q_n}{Q}$$

$$\text{MAP}(Q) = \frac{1}{|Q|} \sum_{j=1}^{|Q|} \frac{1}{m_j} \sum_{k=1}^{m_j} \text{Precision}(R_{jk})$$







Current Investigations @ IPLAB

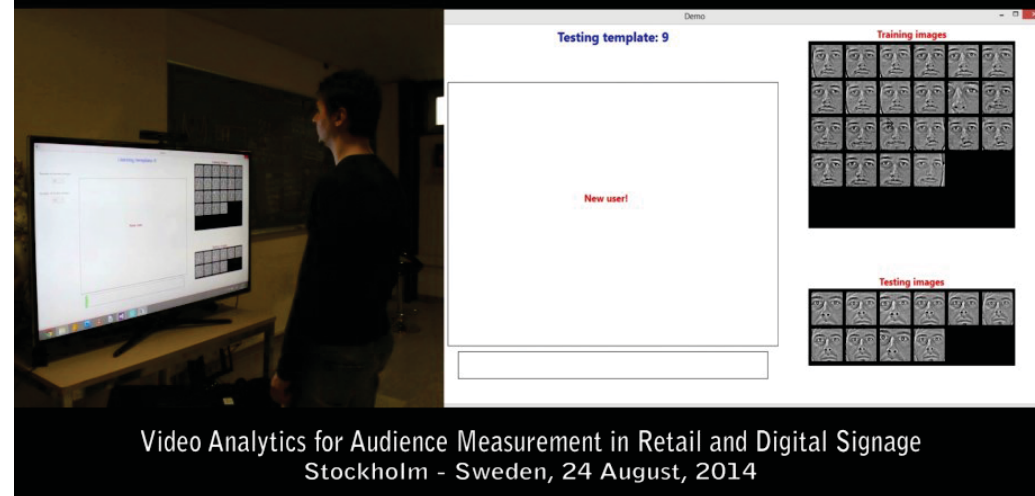
- Dataset extension (UNICT-FD1500)
- Wearable technologies
(e.g., dietary monitoring)
- 3D of Food
(e.g., quantity estimation)
- Augmented Reality
(e.g., to display information on food)
- Quality of food
(e.g., pizza segmentation)
- Person Re-Identification
(e.g., Food/Beverage Machines)
- Soft Biometrics for marketing/retail
(e.g., gender/age recognition)

Some of the above topics we are in collaboration with:



Face Re-Identification for Digital Signage Applications

G. M. Farinella, G. Farioli, S. Battiato, S. Leonardi, G. Gallo



Video Analytics for Audience Measurement in Retail and Digital Signage
Stockholm - Sweden, 24 August, 2014

Conclusions

- Food Images and Social
- Food Images and Industry
- Food Images and Research Open Challenges
- Opportunities in the context of health and marketing/retail



UNIVERSITÀ
degli STUDI
di CATANIA

DIPARTIMENTO DI
MATEMATICA E INFORMATICA

Thank you for your attention

Giovanni Maria Farinella

CONTACT

gfarinella@dmf.unict.it

www.dmf.unict.it/farinella

