On the Exploitation of One-Class Classification to Distinguish Food vs Non-Food Images

G. M. Farinella, D. Allegra, F. Stanco, S. Battiato.

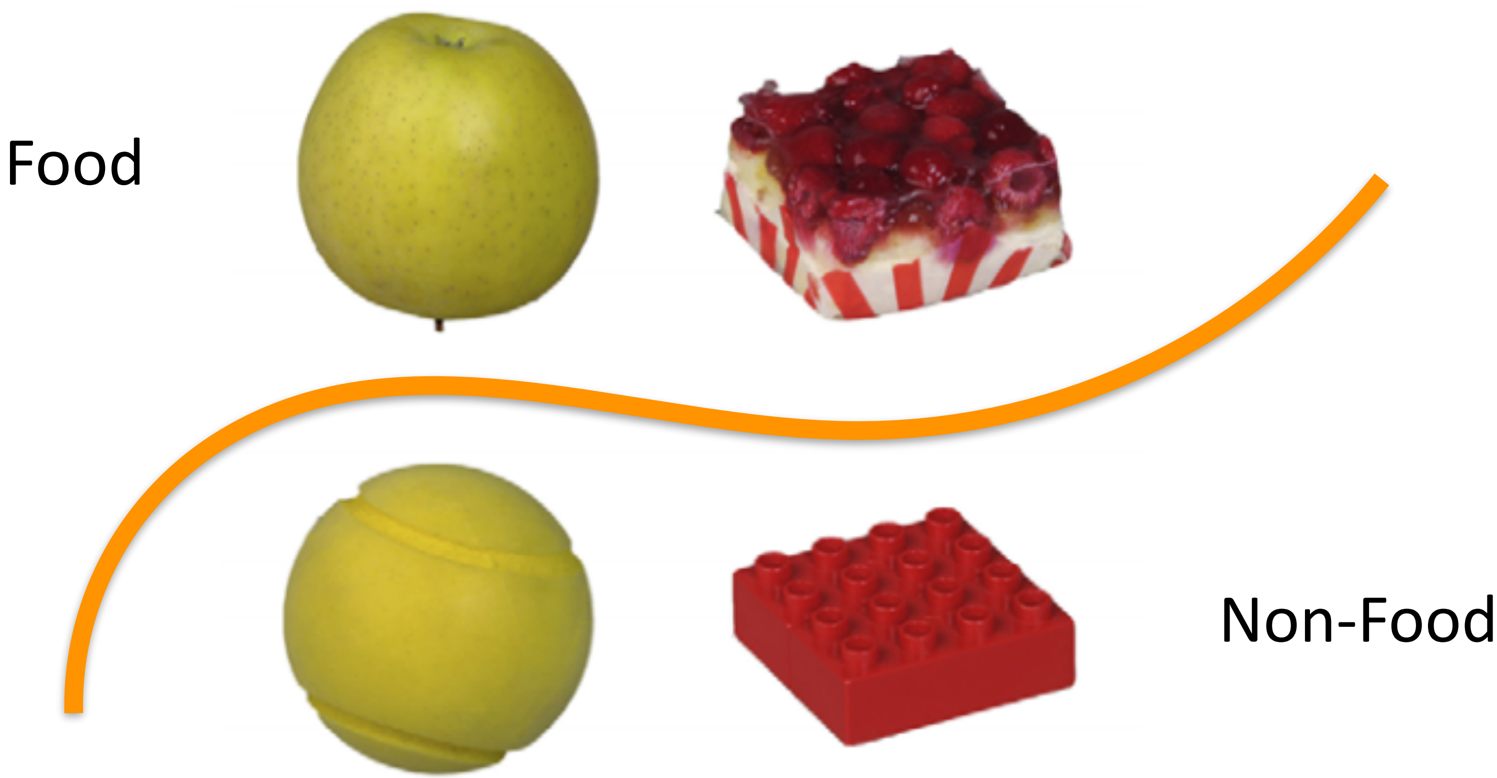
1st International Workshop on Multimedia Assisted Dietary Management

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

Genova, Italy, September 8th 2015

MADiMa2015

Is it Food?



Very Simple for humans, challenging for a machine…

Is it Food?

Art-Inspired Food: Salad with a taste of Kandinsky

Chef: Charles Michel, University of Oxford

Not so simple for humans, very challenge for a machine…

One-Class Classification

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

MADiMa2015

Classic Approach: Two-Class Classification

Kagaya, H., Aizawa, K.: Highly Accurate Food/Non-Food Image Classification Based on a Deep Convolutional Neural Network. In: Proceedings of the International Workshop on Multimedia Assisted Dietary Management, (2015)

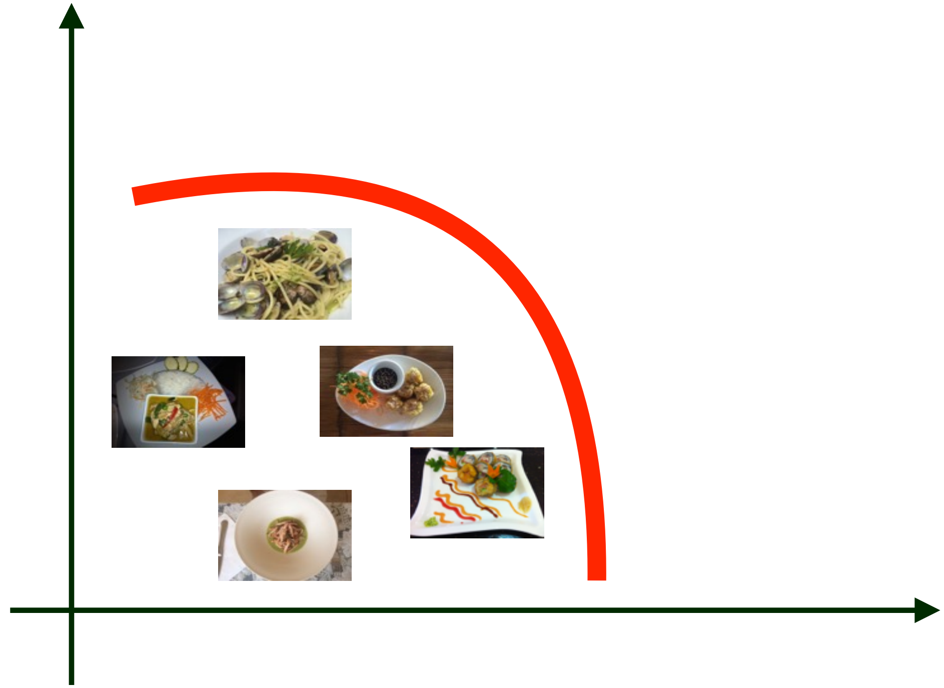
Kitamura, K., Yamasaki, T., Aizawa, K.: Foodlog: capture, analysis and retrieval of personal food images via web. In: Proceedings of the Workshop on Multimedia for Cooking and Eating Activities, pp. 23–30 (2009)

Kagaya, H., Aizawa, K., Ogawa, M.: Food detection and recognition using convolutional neural network. In: Proceedings of the ACM International Conference on Multimedia, pp. 1085–1088 (2014)

We want to learn a discriminant boundary

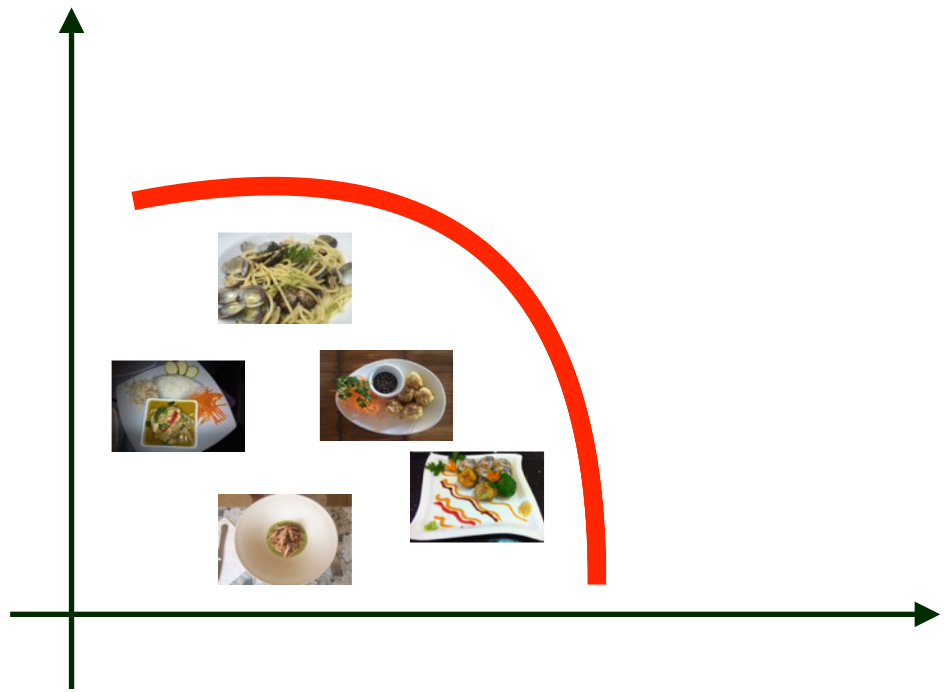
Food vs. non-Food classification problem requires the ability to decide whether a new observation belongs to the same distribution as existing food observations (it is an inlier), or should be considered as different than food (it is an outlier).

We want to learn a discriminant boundary



Can we learn the discriminant boundary using only the food images?

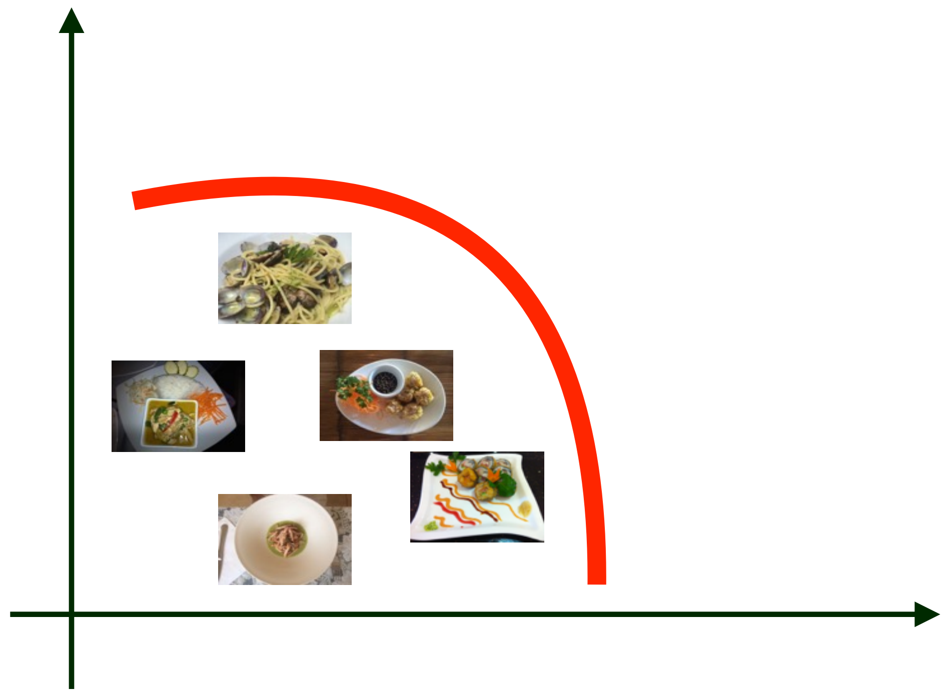
We want to learn a discriminant boundary



Can we learn the discriminant boundary using only the food images?

It is “simple” to collect images of food.

We want to learn a discriminant boundary

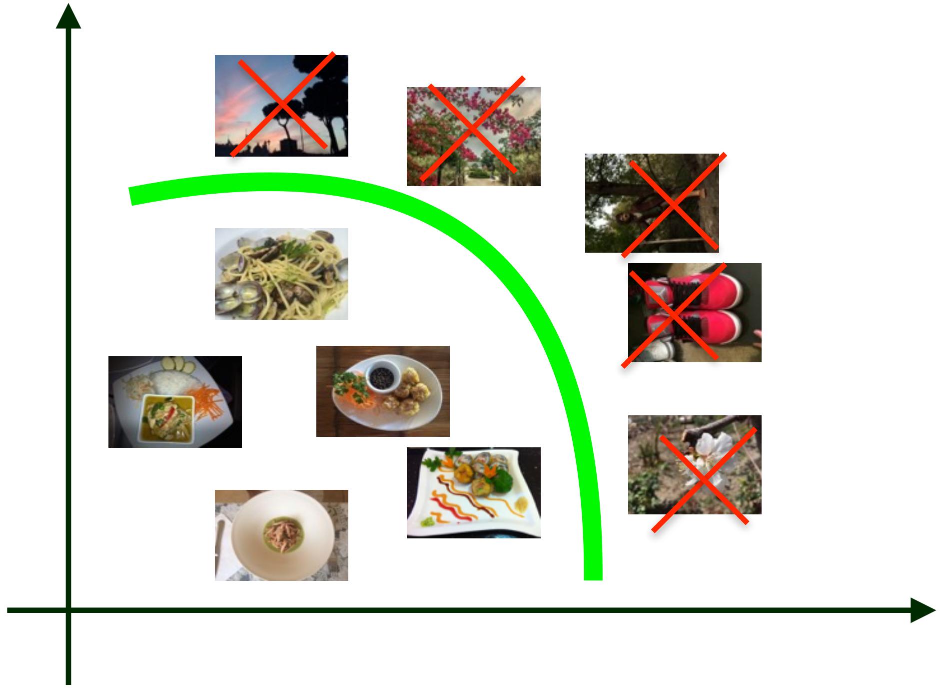


Can we learn the discriminant boundary using only the food images?

It is “simple” to collect images of food.

It is not well defined how to gather the Non-Food classes for the training set for the classic food vs non-food supervised binary classification.

Unsupervised Learning Problem



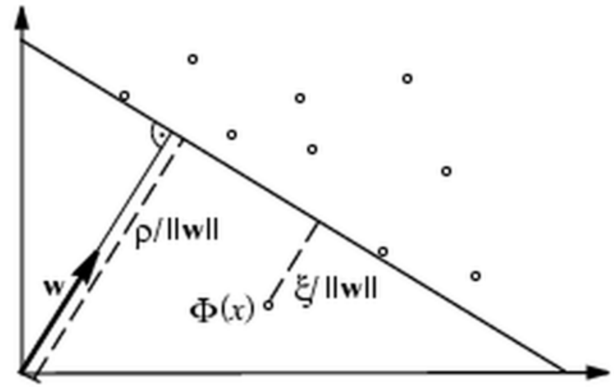
We consider Food vs non-Food classification problem as a **novelty detection problem** where the training data is not polluted by outliers (i.e., we have only food images), and we are interested in detecting anomalies in new observations (i.e., non-food images).

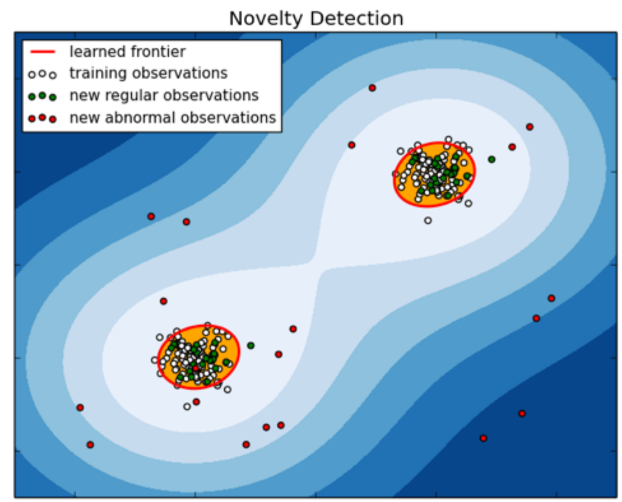
One-Class SVM

$$\min_{\mathbf{w},\boldsymbol{\xi},\rho} \ \frac{1}{2}||\mathbf{w}||^2 + \frac{1}{\nu m}\sum_{i=1}^{m}\xi_i - \rho,$$

$$\text{subject to } \langle \mathbf{w}, \boldsymbol{\Phi}(x_i)\rangle \geq \rho - \xi_i, \ \ \xi_i \geq 0,$$

$$f^m_{\mathbf{w},\rho}(x) = \operatorname{sgn}\left(\langle \mathbf{w}, \boldsymbol{\Phi}(x)\rangle - \rho\right)$$





- x_i is the data *i;*
- $\mathbf{w}$ is a weight vector;
- ξ_i is a slack variable to penalize the outliers;
- $\Phi(\cdot)$ is a map function;
- $\rho = \langle \mathbf{x}, \Phi(x)\rangle$ is an offset hyperplane.
- $\nu \in (0, 1]$ is an upper bound on the fraction of data that may be outliers

Image Representation

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

Image Representation

The image representation plays a key role in the image classification problems:

To test our approach we choose to use three different kinds of handcrafted features: Texton, SIFT and PRICoLBP (see the following link for references: **http://iplab.dmi.unict.it/UNICT-FD889**)

- Bag of Textons has been tested with MR8 filters bank and the Schmid one and 2200 visual words, in both gray and L*a*b* domain.
- Bag of SIFT is considered in RGB domain with 2200 visual words and dense sampling.
- PRICoLBP is considered in gray and color domain.

Also a combination of the aforementioned features has been considered.

We have also tests considered the image representation obtained with CNN (AlexNet Model).

Experimental Datasets

1st International Workshop on Multimedia Assisted Dietary Management

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

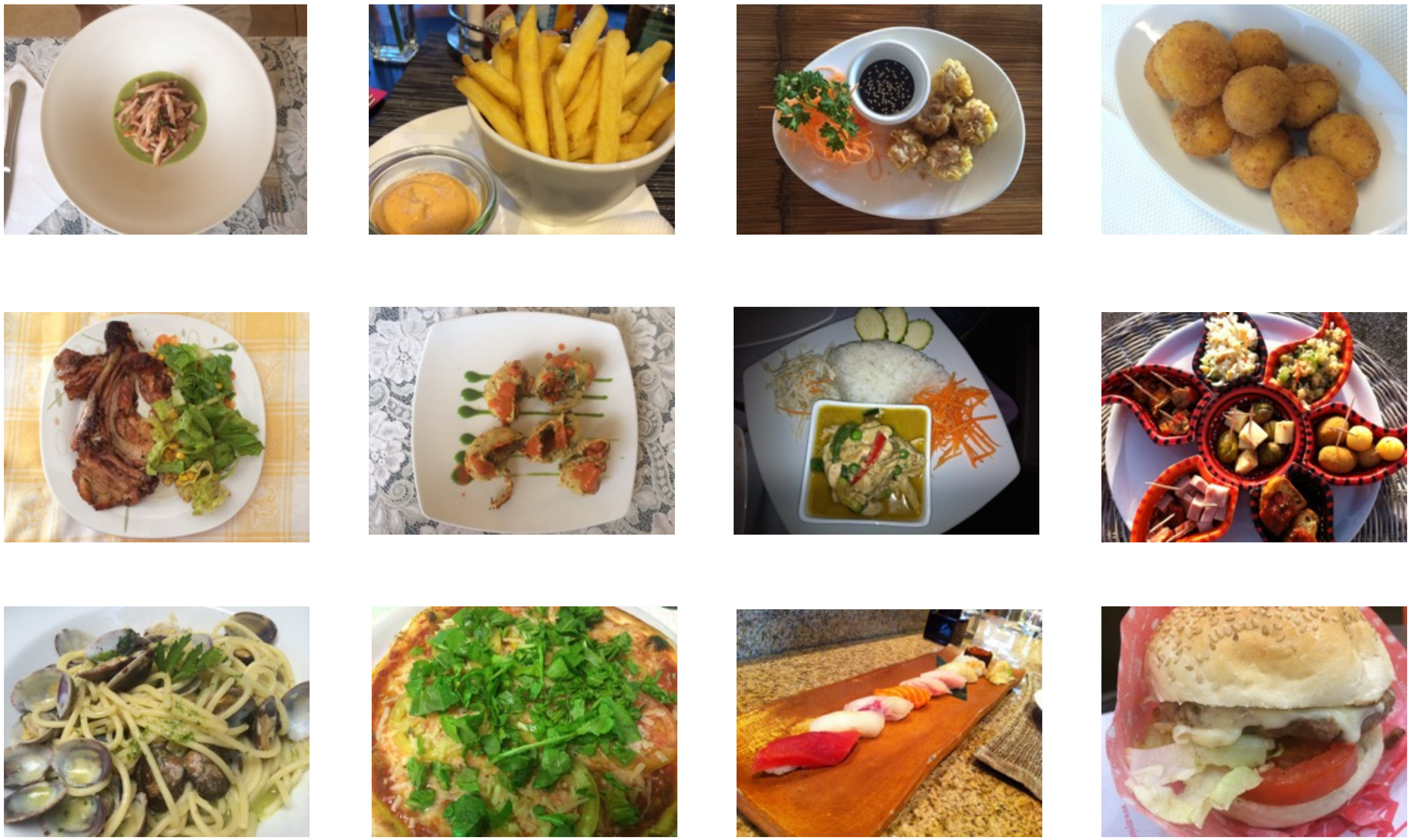
Genova, Italy, September 8th 2015

UNICT-FD889 includes different variabilities

- Scale
- Depth of the plate
- Background
- Pose
- Position
- Orientation
- Illumination
- Nationality/Type/Class
- Appearance
- Shape
- Self-Occlusion
- Shadows
- Camera Effects

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

UNICT-FD889 - Plate Shape Variability



UNICT-FD889 - Class Variability

High Intra-Class Variability

Low Between-Class Variability

UNICT-FD889 - Resolution and Complexity

All the images have been acquired with an iPhone smartphone (minimum resolution is 2592 × 1936 pixels). Images are rescaled to 320x240 pixels in the experiments.

Level of Description

Homogeneous → **Heterogeneous**

Food and Non-Food Images from Flickr

We also collected image form flicker to build two more dataset for testing purpose.

To give possibility to the community to work on the problem all the datasets used to perform the experiments, as well as the experimental protocol, are available at **http://iplab.dmi.unict.it/madima2015/**

Experimental Settings and Results

1st International Workshop on Multimedia Assisted Dietary Management

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

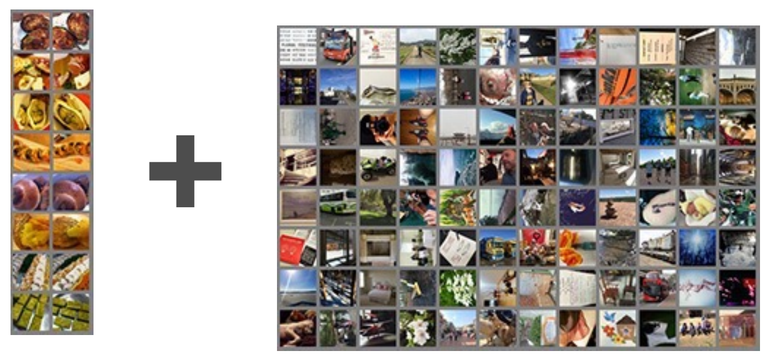
Genova, Italy, September 8th 2015

First Experiment

In the first experiment UNICT-FD889 dataset has been divided in two subsets for training (2855 images) and testing (728 images). Tests have been repeated three times.

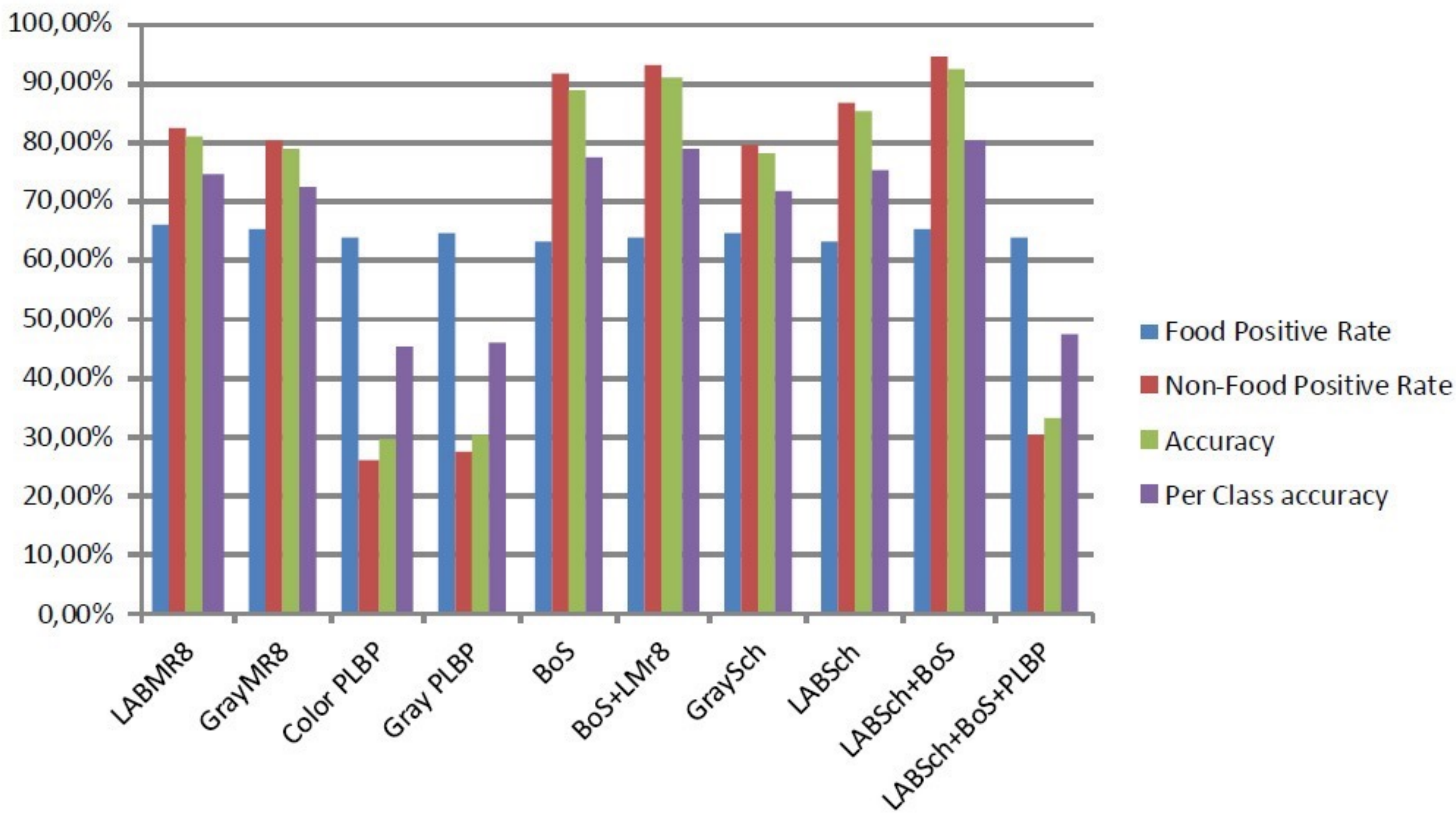
The 8005 non-food images downloaded from Flickr have been used for testing.

Training Set



Test Set

First Experiment - Results



Misclassified Images - First Experiment

Non-Food classified as Food

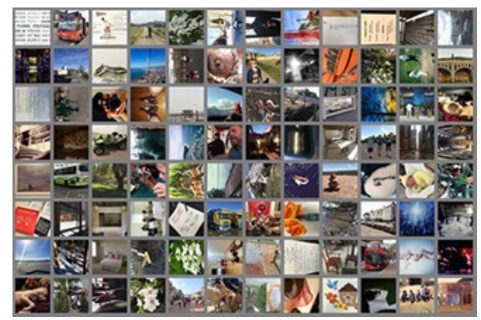
Food classified as Non-Food

Second Experiment

In the second experiment the entire UNICT-FD889 dataset is used for training the classifier.

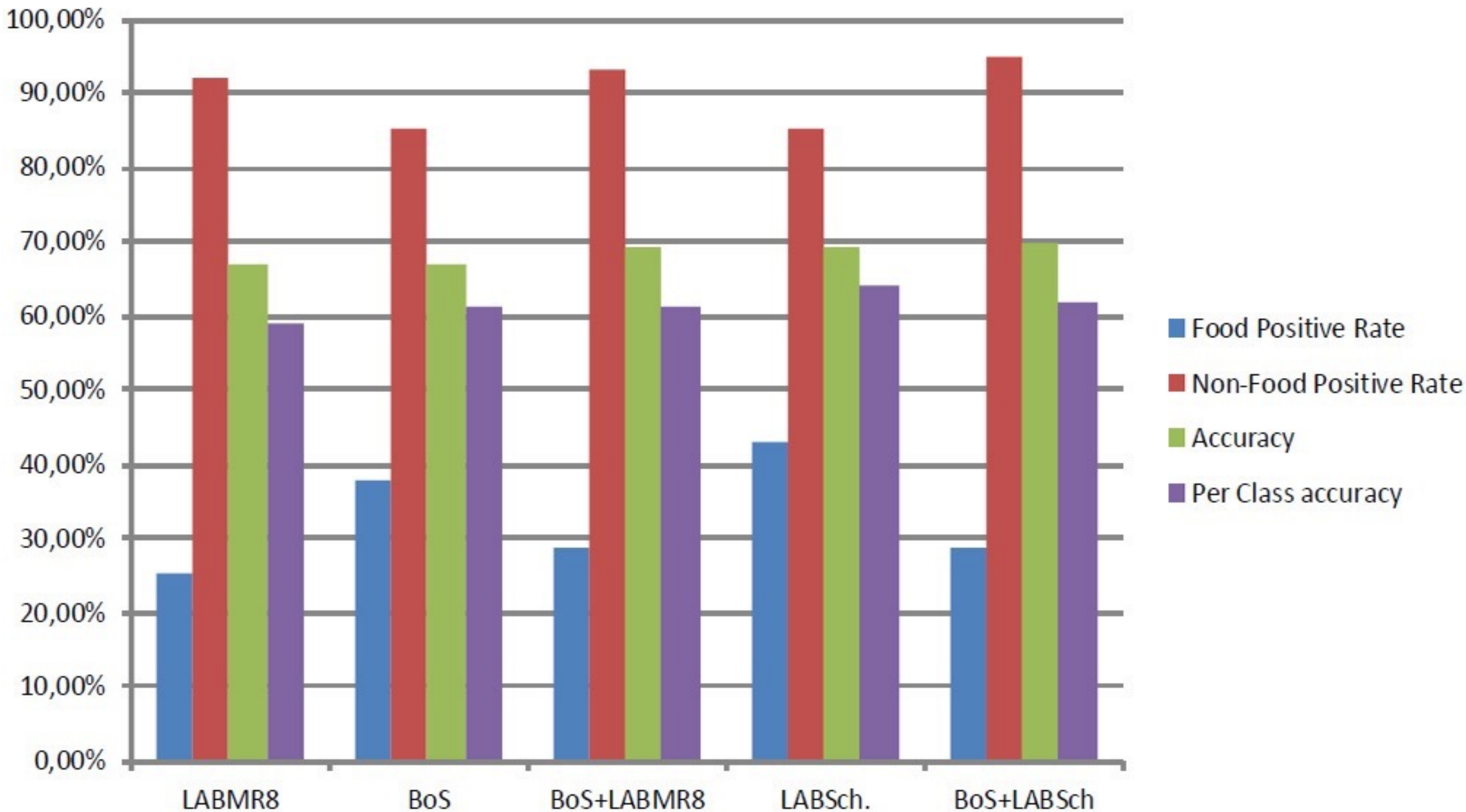
The test set consists of 8005 non-food images and 4805 food images downloaded from Flickr.

Training Set



Test Set

Second Experiment - results

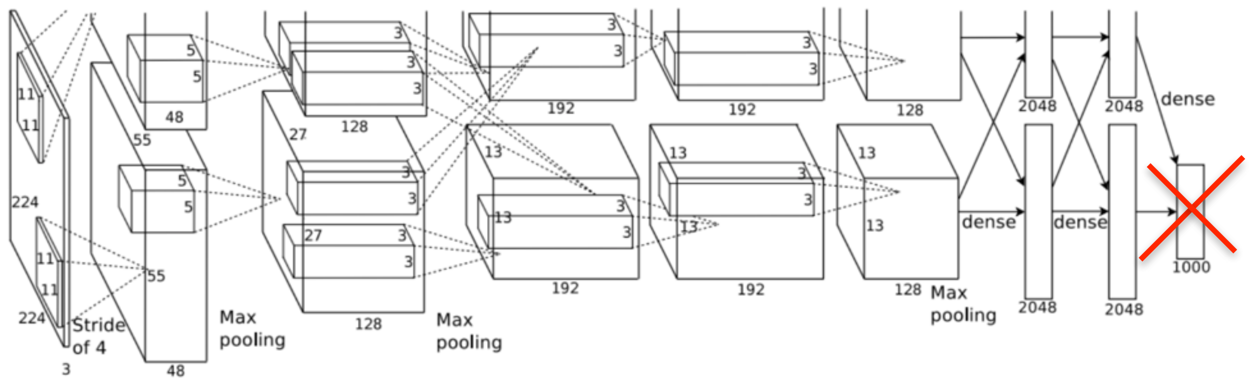


Misclassified Images - Second Experiment

Non-Food classified as Food

Food classified as Non-Food

Results considering CNN Features



OCC Approach	LABSch + BoS	CNN Features FC7
Food	65.43%	52.33%
Non-Food	94.44%	62.53%
Accuracy	92.02%	61.68%

Classic Two Class Approach	CNN Features FC7	CNN Features FC6
Accuracy	77.57%	84.01%

Summary

- <u>**One-Class Classification Approach**</u>

- Considering the results of the two experiments, it is clear that by learning from the food class only it is possible to achieve <u>**low false positive**</u> rate for food vs non-food classification already with simple image representations.

- On the other hand, the <u>**classification accuracy of the food class is still to low**</u> to be considered useful to monitor the food intake and the behavior of a person.

- By considering the two experiments the main observation is that, <u>**when the food class to be recognized is represented in the training, the food classification performance is higher**</u> (i.e., when images of food used for testing are visually similar to the once in the training, despite high geometric and photometric variabilities).

- The <u>**combination of different features achieve a better discrimination**</u>.

- **Datasets and Experimental protocols are available at** http://iplab.dmi.unict.it/madima2015/

On the Exploitation of One-Class Classification to Distinguish Food vs Non-Food Images

G. M. Farinella, D. Allegra, F. Stanco, S. Battiato.

Thank you for your attention.

Datasets and experimental protocol for comparison are available at

http://iplab.dmi.unict.it/madima2015/

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