Food Recognition and Leftover Estimation for **Daily Diet Monitoring**

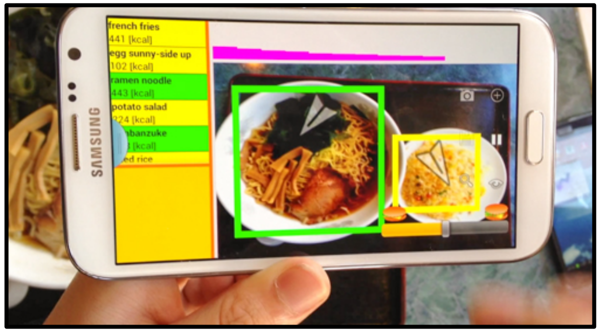
Gianluigi Ciocca, **Paolo Napoletano**, Raimondo Schettini

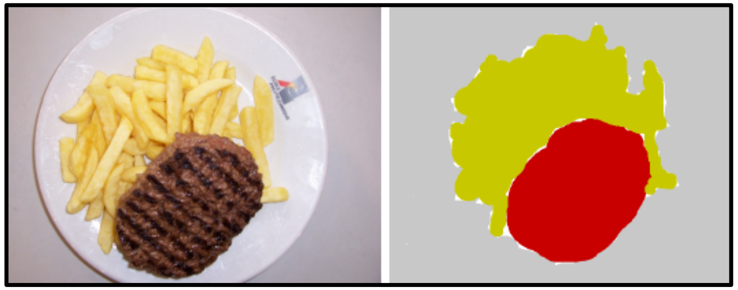
@MADiMa 2015, Genova, Italy, September 8th 2015

Automatic dietary monitoring

Nowadays, **technology** can support the users in keep tracks of their **food consumption** in a more user friendly way allowing for a more comprehensive **daily dietary monitoring**.;

Recent findings showed that **computer vision** techniques can help to automatically **recognize food** and **estimate its quantity**

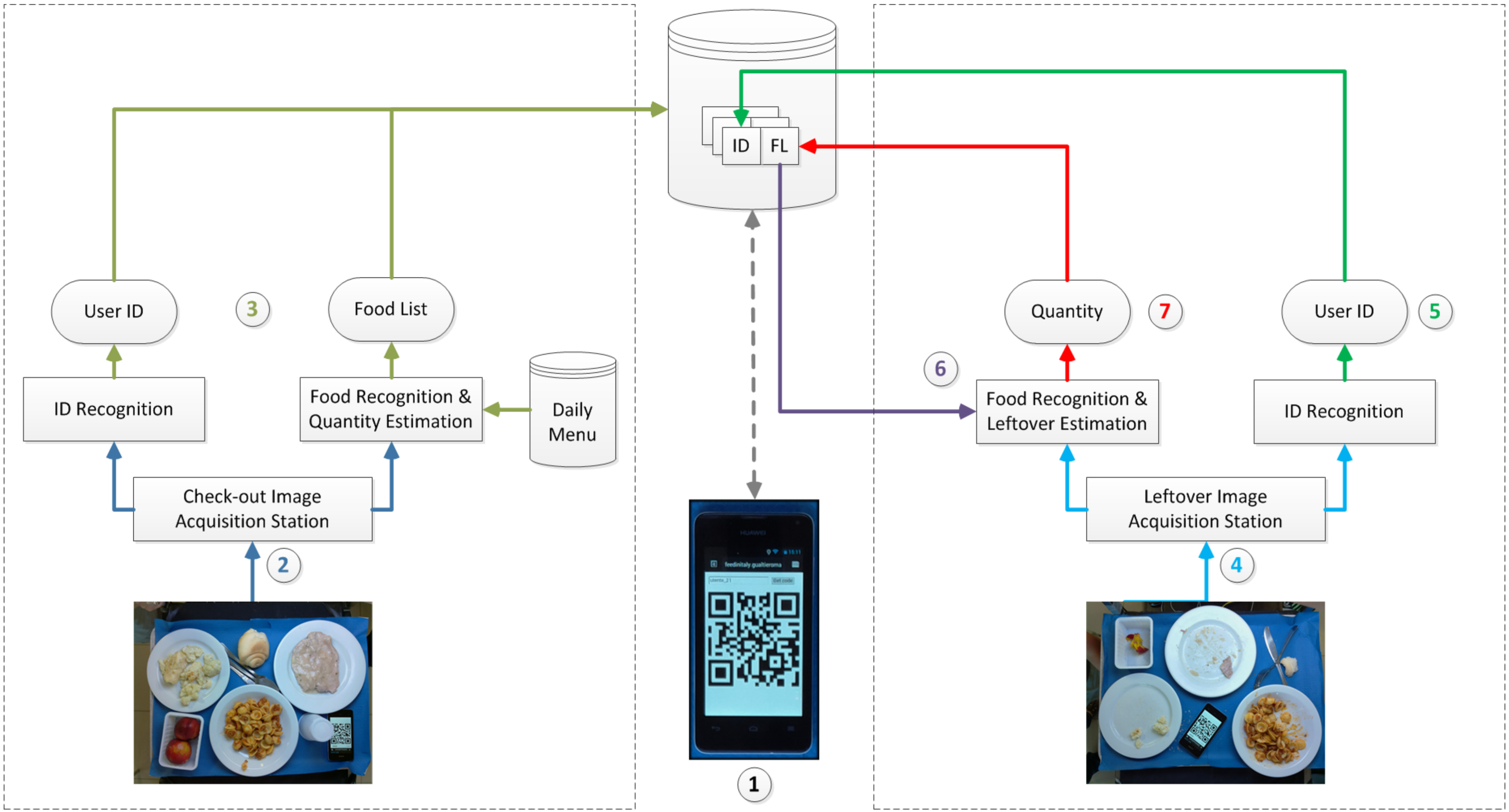




Proposed System // Workflow

Automatic dietary monitoring of canteen customers

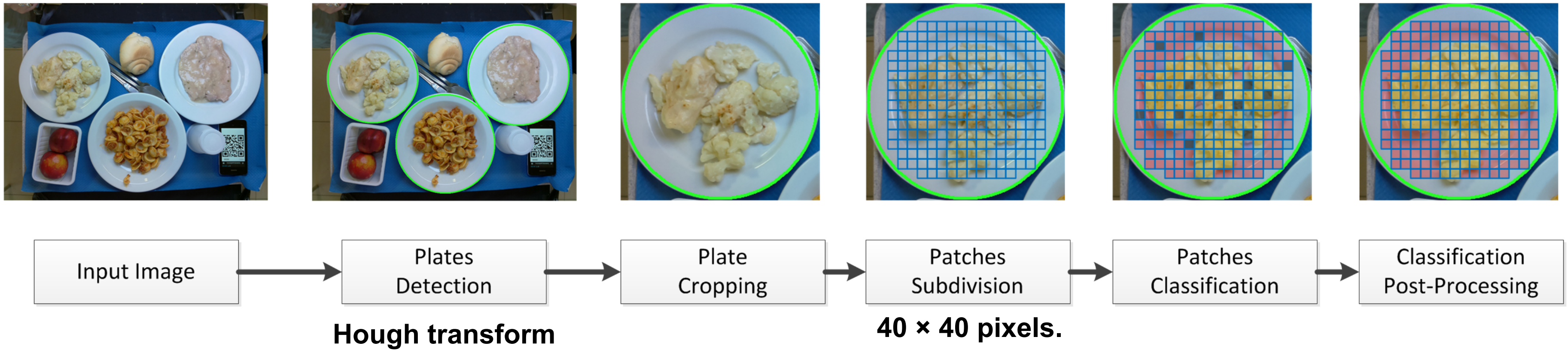
(based on robust computer vision techniques for food recognition and leftover estimation)



Proposed System // Food recognition

Automatic dietary monitoring of canteen customers

(based on robust computer vision techniques for food recognition and leftover estimation)



From each patch, a visual descriptor is extracted and submitted to a **pre-trained k-NN** classifier in order to receive a **classification label**.

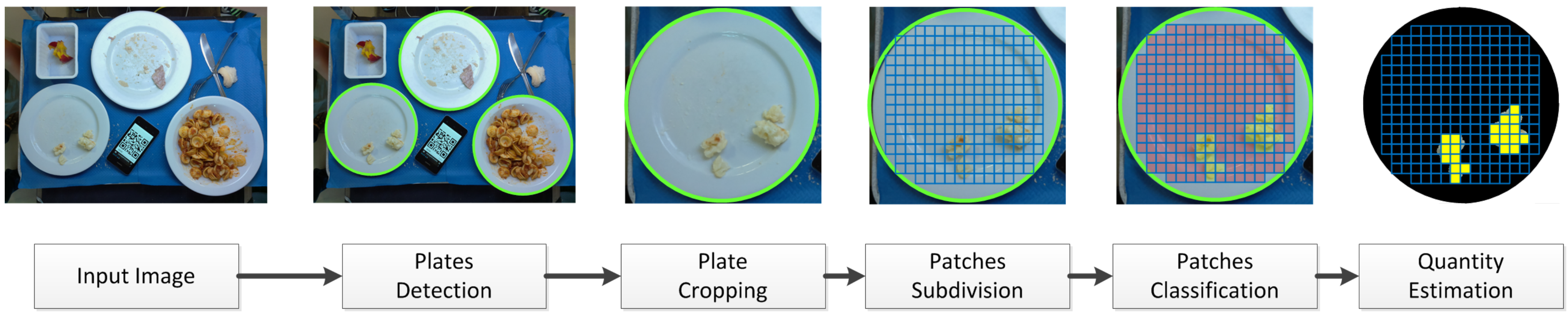
The labels of the the patches are then **post-processed** to remove spurious labels in order to have more homogeneous groups of labels that correspond to the **food regions**.

Name	Description	Length
CEDD	Color and Edge Directivity Descriptor	144
Gabor	Gabor features. Mean and st.dev. of RGB DFT at $(\theta, f)=(4,4)$	96
OG	Opponent Gabor. Gabor on iter-intra channel combinations	264
LBP	Non-uniform, invariant Local Binary Pattern with (r,n)=(1,8)	54
LCC	Local Color Contrast	499
CM	Two sets of five normalized Chromaticity Moments	10
CWT	Complex Wavelet features. RGB mean and st.dev. at three scales	18

Proposed System // Leftover estimation

Automatic dietary monitoring of canteen customers

(based on robust computer vision techniques for food recognition and leftover estimation)



The personnel of the canteen is bounded to follow the regulations provided by nutritionists in the form of nutritional tables, and to serve a specific amount of food (that depends on its calories and nutrients). This somewhat simplifies the problem of the estimation of the food quantity.

$$r_{ic}^{est} = \frac{\#Patches\ leftover}{\#Patches\ before}$$

Once we have identified this ratio, the corresponding amount of calories is deduced by the precompiled nutritional tables.

Experiments // Canteen scenario

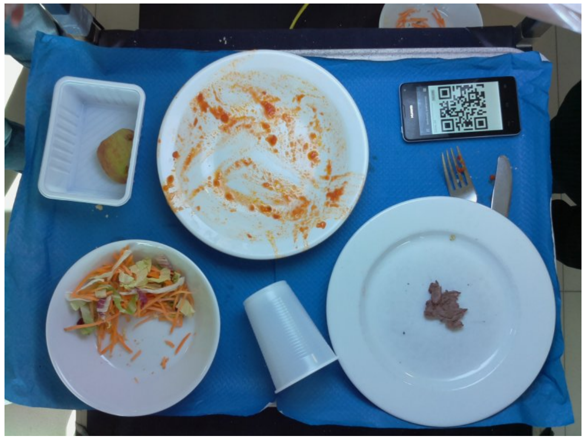
Monitored and recorded the meal of 1000 customers of a real canteen
(1000 customers correspond to 2000 tray images, 1000 before and 1000 after the meal)

- **Customer behavior:** each customer selected 3 dishes from the daily menu that included 15 different dishes;

- **Tray acquisition:** automatic photographic system that includes a raspberry motherboard, an embedded camera and a motion sensor. The system automatically detects when the tray has to be acquired.

- **Ground-truth:** The annotations have been created using the IAT – image annotation tool [2], that permitted to draw a polygon around the food: r_{ic}^{gt}

before



after

Experiments

2 metrics for food recognition:

Standard Accuracy (**SA**) and the **Macro Average Accuracy** (**MAA**)

$$SA = \frac{\sum_{c=1}^{C} TP_c}{\sum_{c=1}^{C} NP_c}; \qquad MAA = \frac{1}{C}\sum_{c=1}^{C} A_c = \frac{1}{C}\sum_{c=1}^{C} \frac{TP_c}{NP_c}$$

Denoting ***NPc*** the number of positives, i.e., the number of times the class ***c*** occurs in the dataset; ***TPc*** the number of true positives for class ***c***, i.e., the number of times that the system recognizes the dish ***c***; ***C*** the number of classes,

leftover estimation:

Error

$$Error = \sum_{c=1}^{C} w_c \sum_{i=1}^{I} |r_{ic}^{gt} - r_{ic}^{est}|$$

where w_c is the class weight and ***I*** is the number of test customers. The class weight is defined as the number of elements of the class divided by the total number of elements.

Experiments

Automatic dietary monitoring of canteen customers

(based on robust computer vision techniques for food recognition and leftover estimation)

Classes	w_c(%)		CEDD	OG	Gabor	LBP	LLC	CM	CWT
			visual descriptors						
bistecca	(3.8%)		100.00	100.00	100.00	27.50	97.50	91.25	80.00
carote	(7.6%)		100.00	100.00	100.00	100.00	100.00	98.75	100.00
cavolfiore	(8.6%)		100.00	100.00	98.89	97.22	98.33	97.22	98.33
fagiolini	(7.6%)		100.00	100.00	100.00	99.38	100.00	100.00	96.25
frittata	(7.6%)		100.00	100.00	100.00	81.25	93.75	83.12	100.00
fusilli ragu	(8.6%)		100.00	100.00	100.00	85.56	100.00	97.22	100.00
insalata mista	(2.4%)		100.00	92.00	42.00	58.00	100.00	90.00	32.00
lenticchie	(7.1%)		98.67	99.33	96.67	68.00	94.67	28.67	57.33
minestra	(6.7%)		100.00	100.00	97.86	99.29	97.86	93.57	100.00
pasta cime rapa	(8.6%)		100.00	100.00	100.00	100.00	100.00	100.00	100.00
pasta sugo	(2.4%)		100.00	100.00	100.00	28.00	76.00	100.00	98.00
piselli	(7.1%)		99.33	100.00	98.67	94.67	100.00	88.00	98.00
pollo ferri	(7.6%)		96.86	97.48	67.30	62.26	76.10	93.71	69.18
scaloppina	(8.6%)		98.90	99.45	99.45	13.81	98.34	97.79	98.90
tortino	(5.7%)		91.67	90.83	79.17	22.50	79.17	83.33	80.00
		SA	**99.05**	99.00	94.33	74.14	95.05	89.57	90.38
		MAA	**99.03**	98.61	92.00	69.16	94.11	89.51	87.20

The system is capable of estimating the **relative quantity of eaten food** with an average error of about **15 percentage points**, with the **best** and **worst** cases being **7** and **34 percentage points** respectively.

Conclusions

Results achieved on a real canteen scenario are promising with an average accuracy in recognition of about 99%, and and average error in food estimation of 15 percentage points.

The proposed food recognition and leftover estimation system can serve multiple purposes:

1. at the check-out station, the food recognition allows to keep track the eaten food and the user's dietary habits;
2. using the list of recognized foods, an automatic billing procedure can be activated speeding up the check- out;
3. by evaluation the leftovers, we can better estimate the food intakes in terms of calories ingested.

As a future work,

1. we will experiment the system on large scale food datasets
2. we will experiment CNN based features