Automated food log

Baptiste NOGARET

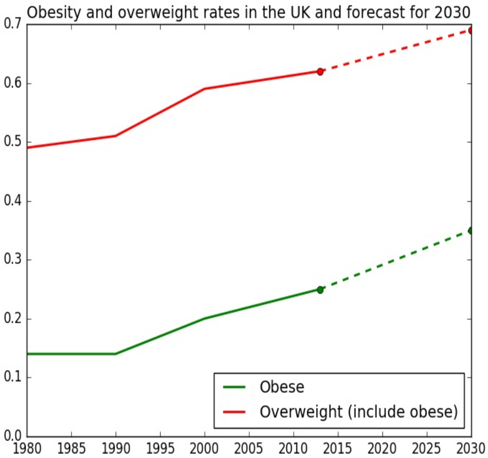
Supervised by Dr Stefan RÜGER

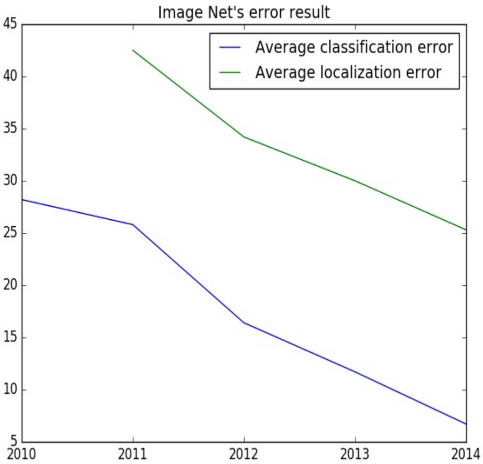
21st July 2016

www.cranfield.ac.uk

Table of contents

Why a food log analysis system?





Overall process

Roughly copying the procedure of FoodLog [1, 2, 3, 4, 5]:

- Generate a relevant dataset
- Extract characteristic
- Machine learning
- For a new picture from a user, classify and estimate intake

Focus on localization and classification

Dataset

Name	Release date	Number of pictures	Type of food	Number of classes	Multiple food items
PFID [6]	2009	4545	American fast-food	101	No
UEC FOOD 100 [7]	2012	14361	Japanese	100	Yes
FIDS 30 [8]	2013	971	Fruit	30	No
ETHZ Food-101 [9]	2014	101 000	European	100	No
FooDD [10]	2015	3000	Fruit	23	Yes
UEC FOOD 256 [11]	**2015**	**31395**	**World**	**256**	**Yes**

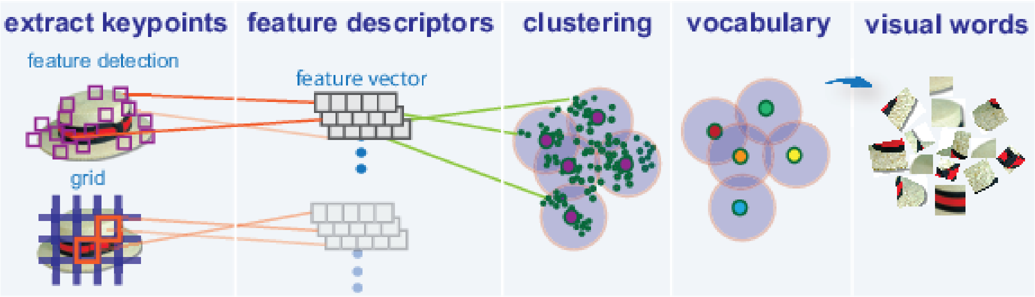
Dataset

Example of multi-items

Feature description

Bag of visual words

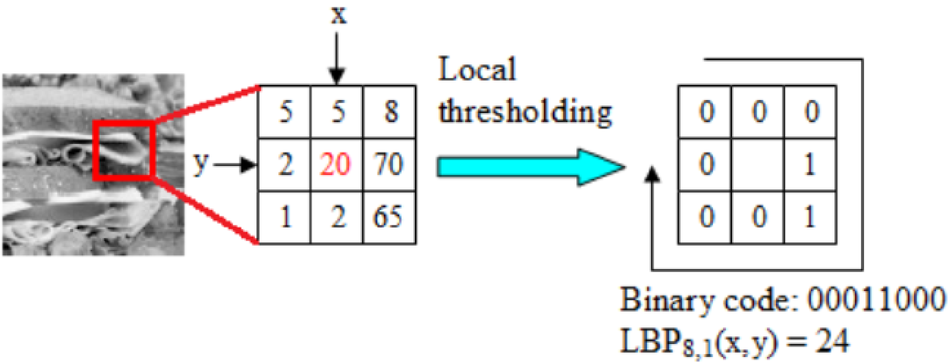
Common feature descriptor, use in [6, 12, 13]



Feature description

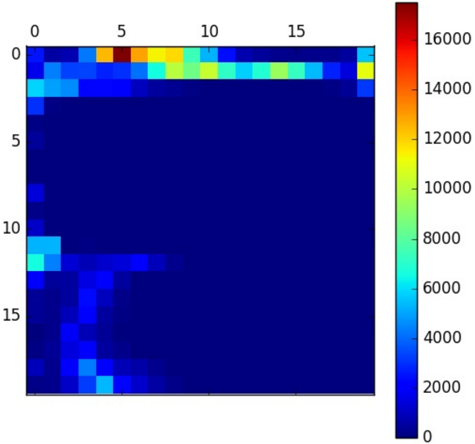
Local binary pattern

Use in [14, 15] for texture description



Feature description

Color moments and histograms



Mean:

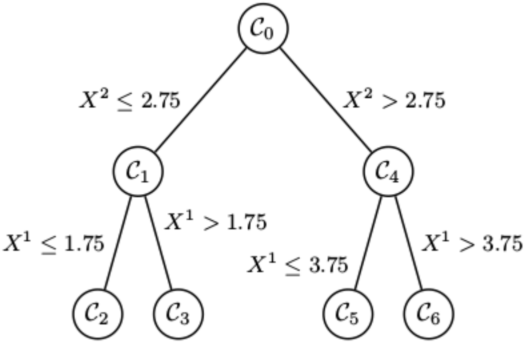
$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

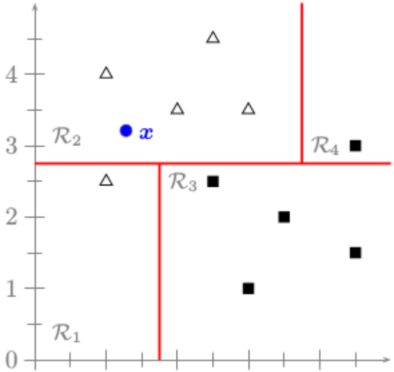
Variance:

$$\mathrm{Var}(X) = \sum_{i=1}^{n} p_i \cdot (x_i - \mu)^2$$

Classifiers

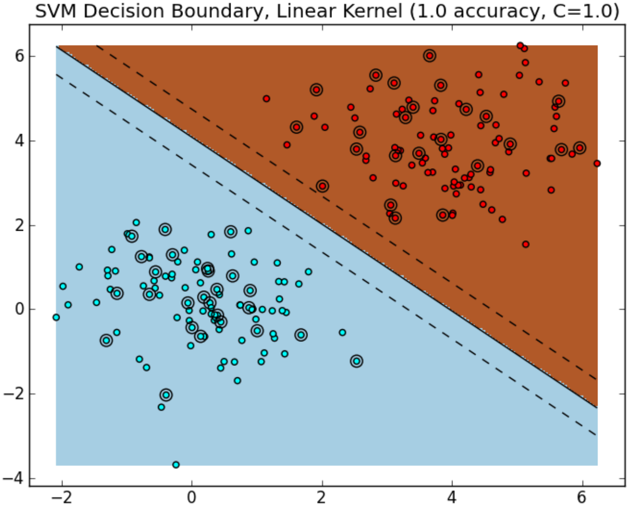
Decision tree and random forest





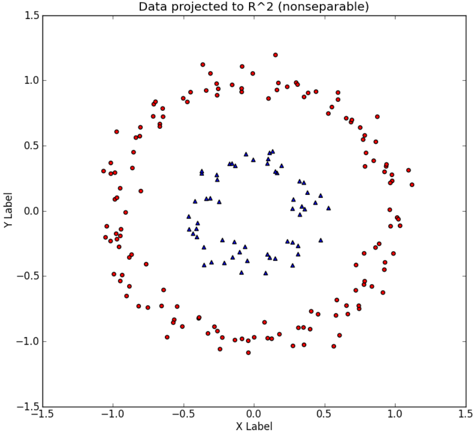
Classifiers

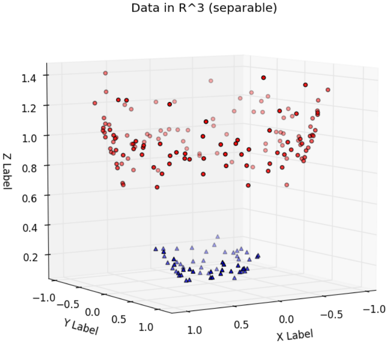
SVM



Classifiers

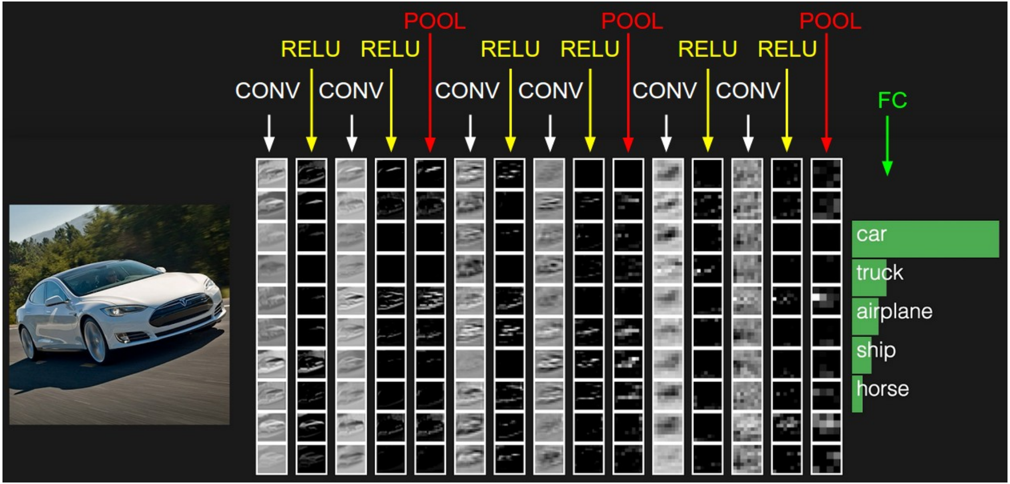
SVM and kernel trick



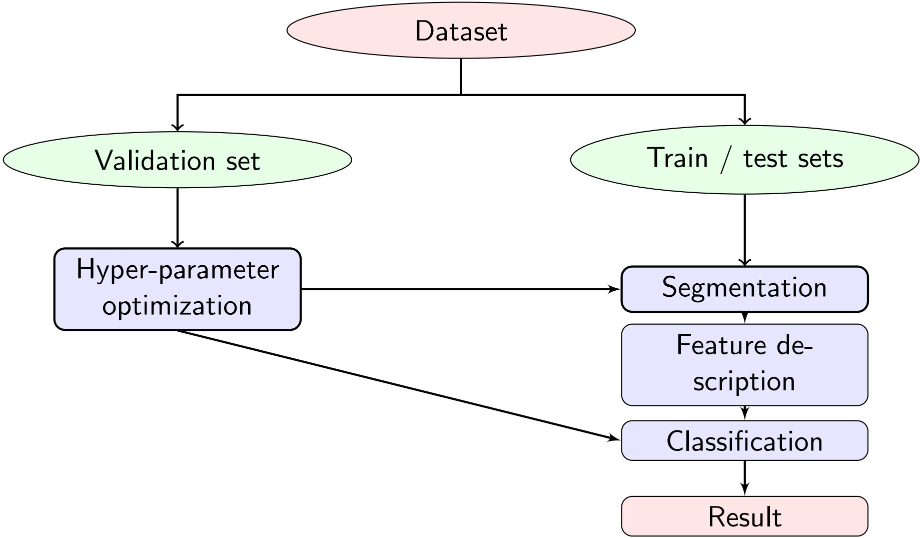


Classifiers

CNN



Structure


Dataset
Validation set
Train / test sets
Hyper-parameter optimization
Segmentation
Feature de-scription
Classification
Result

Results

Localization

Using a DCNN pre-trained on [16] to detect saliency object

Correcteness metric: As describe in [17], must have an intersection over union greater than 50 %

$$IoU = \frac{area(B_p \cap B_{gt})}{area(B_p \cup B_{gt})}$$

Metric	My method	DCNN from [18]
Accuracy	**73 %**	60 %
Recall	**74 %**	80 %
Precision	**79 %**	70 %

Results

Classification

Method	Average accuracy
CNN as descriptor + RF	**40 %**
BoW (1000 words)+ SVM with χ^2	10 %
LBP + color historams and moments + Decision tree	5 %
LBP + color historams and moments + SVM	11 %
LBP + color historams and moments + RF	16 %
DCNN from [18]	63 %
DCNN from [19]	67 %

Results

Segmentation and classification

Segmentation: DCNN followed by the classification: CNN as a descriptor and RF

UEC FOOD 256

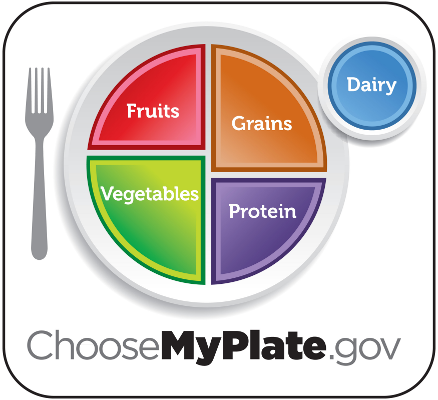
Accuracy	My method	DCNN from [18]
Overall	**28 %**	36 %
Segmentation	**74 %**	60 %
Classification	**38 %**	60 %

UEC FOOD 100

Accuracy	My method	[20]	[21]
Overall	**33 %**	-	-
Segmentation	**67 %**	60 %	-
Classification	**50 %**	-	72 %

Future work and comment

- Use a better feature descriptor / classifier
- Regroup the classification in 5 big categories for food intake as in [4]

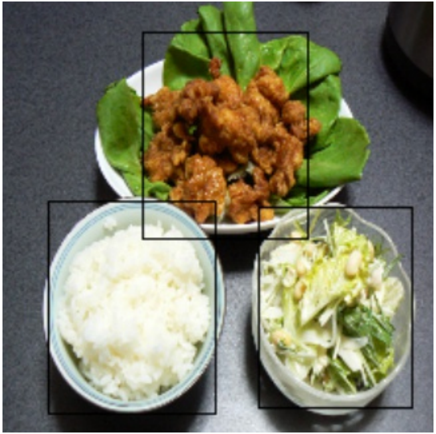


Automated food log - Baptiste NOGARET

Segmentation and classfication applied on two reference datasets

Obtain great localization results

Overall accuracy of 28 % (to date, the best result is 36 % in [18]).



References I

- Keigo Kitamura, Toshihiko Yamasaki, and Kiyoharu Aizawa. “FoodLog: Capture, Analysis and Retrieval of Personal Food Images via Web”. In: **Proceedings of the ACM multimedia 2009 workshop on Multimedia for cooking and eating activities - CEA ’09** (2009), p. 23.
- Keigo Kitamura, Toshihiko Yamasaki, and Kiyoharu Aizawa. “Food log by analyzing food images”. In: **ACM international conference on Multimedia** (2008), p. 999.
- Gamhewage C. De Silva and Kiyoharu Aizawa. “Clustering meal images in a web-based dietary management system”. In: **Proceedings - IEEE International Conference on Multimedia and Expo**. IEEE, 2011, pp. 1–6.

References II

- Kiyoharu Aizawa et al. "Food balance estimation by using personal dietary tendencies in a multimedia food log". In: **IEEE Transactions on Multimedia** 15.8 (2013), pp. 2176–2185.
- Hokuto Kagaya, Kiyoharu Aizawa, and Makoto Ogawa. "Food Detection and Recognition Using Convolutional Neural Network". In: **ACM Multimedia**. 2. 2014, pp. 1085–1088.
- Mei Chen et al. "PFID: Pittsburgh Fast-food Image Dataset". In: **Proceedings - International Conference on Image Processing, ICIP** (2009), pp. 289–292.
- Yuji Matsuda, Hajime Hoashi, and Keiji Yanai. "Recognition of multiple-food images by detecting candidate regions". In: **Proceedings - IEEE International Conference on Multimedia and Expo**. IEEE, 2012, pp. 25–30.

References III

- Škrjanec Marko. "Automatic fruit recognition using computer vision". Mentor: Matej Kristan. Bsc thesis. Faculty of Computer and Information Science, University of Ljubljana, 2013.
- Lukas Bossard, Matthieu Guillaumin, and Luc Van Gool. "Food-101 - Mining discriminative components with random forests". In: **Lecture Notes in Computer Science**. Vol. 8694 LNCS. PART 6. 2014, pp. 446–461.
- Parisa Pouladzadeh Abdulsalam Yassine and Shervin Shirmohammadi. "FooDD: Food Detection Dataset for Calorie Measurement Using Food Images". In: **New Trends in Image Analysis and Processing – ICIAP 2015 Workshops** 9281 (2015), pp. 441–448.

References IV

Yoshiyuki Kawano and Keiji Yanai. "Automatic expansion of a food image dataset leveraging existing categories with domain adaptation". In: **Lecture Notes in Computer Science** 8927 (2015), pp. 3–17.

Hajime Hoashi, Taichi Joutou, and Keiji Yanai. "Image recognition of 85 food categories by feature fusion". In: **Proceedings - 2010 IEEE International Symposium on Multimedia, ISM 2010**. IEEE, 2010, pp. 296–301.

Vinay Bettadapura et al. "Leveraging context to support automated food recognition in restaurants". In: **Proceedings - 2015 IEEE Winter Conference on Applications of Computer Vision, WACV 2015**. 2015, pp. 580–587.

References V

- Zhimin Zong et al. "On the combination of local texture and global structure for food classification". In: **Proceedings - 2010 IEEE International Symposium on Multimedia, ISM 2010**. IEEE, 2010, pp. 204–211.
- Duc Thanh Nguyen et al. "Food image classification using local appearance and global structural information". In: **Neurocomputing** 140 (2014), pp. 242–251.
- Jianming Zhang et al. "Unconstrained Salient Object Detection via Proposal Subset Optimization". In: **IEEE Conference on Computer Vision and Pattern Recognition(CVPR)** (2016).
- M. Everingham et al. **The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results**.

References VI

- Marc Bolaños and Petia Radeva. "Simultaneous Food Localization and Recognition". In: (2016), pp. 2–7.
- Keiji Yanai and Yoshiyuki Kawano. "Food image recognition using deep convolutional network with pre-training and fine-tuning". In: **2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)**. IEEE, 2015, pp. 1–6.
- Wataru Shimoda and Keiji Yanai. "CNN-based food image segmentation without pixel-wise annotation". In: **New Trends in Image Analysis and Processing – ICIAP 2015 Workshops**. Vol. 9281. 2015, pp. 449–457.
- Yoshiyuki Kawano and Keiji Yanai. "Food Image Recognition with Deep Convolutional Features". In: **ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)** (2014), pp. 589–593.