

# Food Image Segmentation for Dietary Assessment

Joachim Dehais<sup>1</sup>, Marios Anthimopoulos<sup>1,2</sup>, Stavroula Mougiakakou<sup>1,3</sup>

<sup>1</sup> ARTORG Center for Biomedical Engineering Research, University of Bern, Switzerland

<sup>2</sup> Department of Emergency Medicine, Bern University Hospital, Switzerland

<sup>3</sup> Department of Endocrinology, Diabetes and Clinical Nutrition, Bern University Hospital, Switzerland

## Introduction

Diet management is a key factor for the prevention and treatment of diet-related chronic diseases, but to date lacks scalable solutions. To solve this problem, computer vision systems aim to provide automated food intake assessment using meal images. We propose new and efficient methods for dish segmentation as first steps of a dietary assessment system. An interactive method is proposed along with an automatic one, greatly improving reliability even in difficult cases.

## Methods

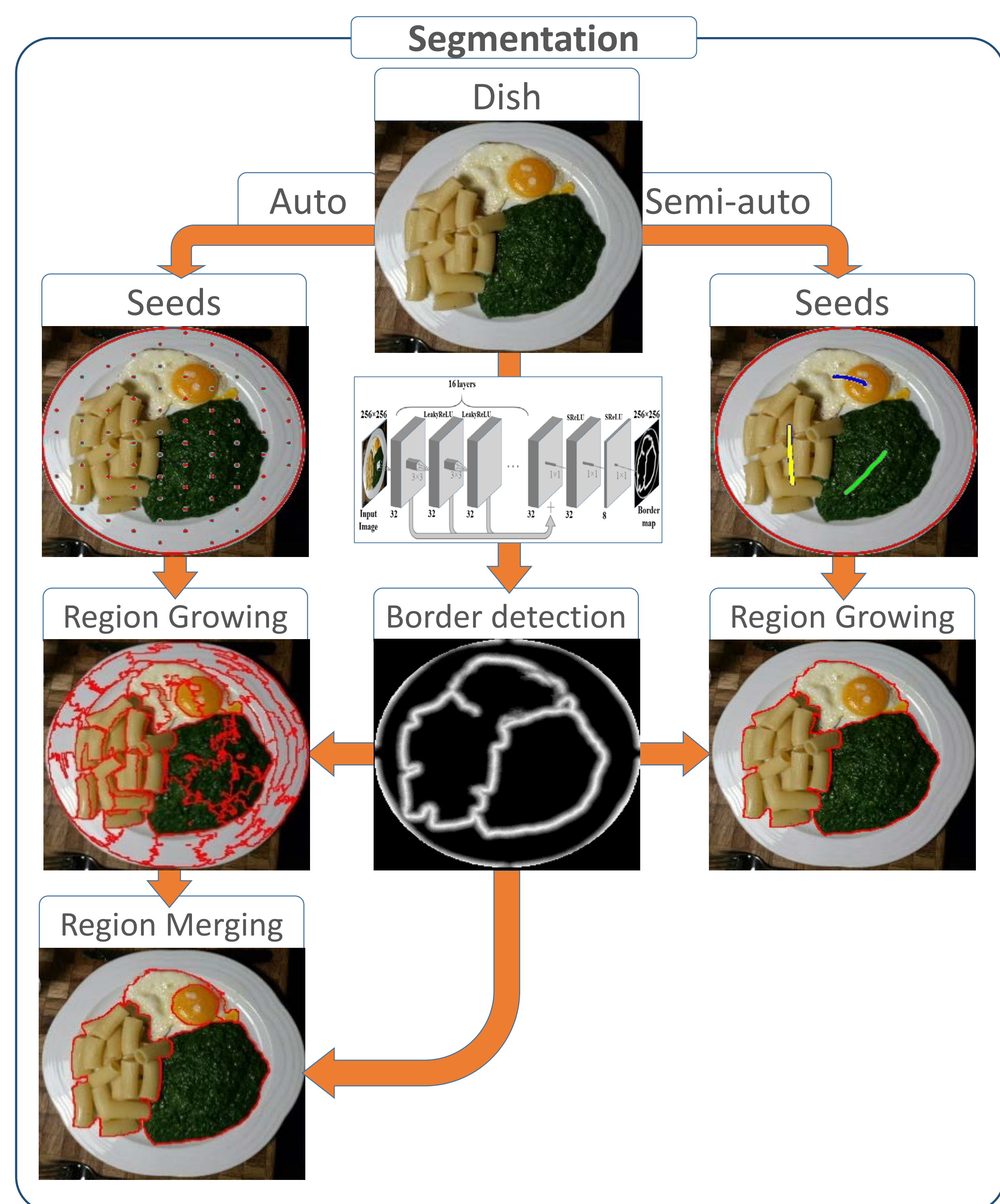


Figure 1 – Overview of the proposed method

**Dish:** Input of the method is an image of a dish already detected with methods such as [1].

**Border detection:** A deep CNN is trained on manually annotated images. The CNN has 16 convolutional layers with 32 kernels of size 3×3. The activations from all 16 layers are summed and passed through three convolutional layers with 32, 8 and 1 kernels of size 1×1.

**Seed generation:** The plate-seed is created automatically as a band inside the ellipse border, followed by:

**Automatic:** Seed generation on a regular grid

**Semi-auto:** User-generated seeds by swiping the smartphone screen.

**Region growing:** Seeds are grown into full regions using the Seeded Region Growing paradigm with a distance (*Dist*) combining: (i) a luminosity dampened colour distance, (ii) the computed border magnitude, and (iii) the distance from the seed.

**Region merging:** In the automatic case, the regions grown in the previous step are merged based using the Statistical Region Merging principle, and the ratio of the distance *Dist* to edge size.

## Results

### Experimental setup:

- The used dataset consists of 821 meal images with one round dish and large variety of foods and conditions.
- 30% of the images were taken by the authors and the rest by end-users
- Food and dish locations were manually annotated.
- 70% of the dataset was randomly chosen for training and the rest 30% for testing.
- The chosen evaluation metric is based on the average overlap between the resulted and the ground truth regions for each image.

Table 1 – Results and comparison with previous works

	Segmentation Method	Performance	Time (sec/image)
Automatic	Proposed	87.6	0.5
	Previous work [1]	85.6	0.45
	Mean-shift [2]	82.6	1.4
	Local Variation [3]	72.0	1.9
Semi-auto	Proposed	92.2	0.43
	Previous [1]	91.3	0.41
	Flood fill	85.7	0.6

[1] J. Dehais, M. Anthimopoulos, S. Mougiakakou, "Dish Detection and Segmentation for Dietary Assessment on Mobile Phones", 1st International Workshop on Multimedia Assisted Dietary Management, ICIAP Workshops, Genova, Italy, Sept. 2015  
[2] Anthimopoulos, M.; Dehais, J.; Diem, P.; Mougiakakou, S., "Segmentation and recognition of multi-food meal images for carbohydrate counting," IEEE 13th International Conference on Bioinformatics and Bioengineering (BIBE), (2013)  
[3] Felzenszwalb, P. F., Huttenlocher, D. P.: Image segmentation using local variation. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 98-104 (1998)

## Conclusions

The proposed segmentation methods outperformed the state of the art. The use of a CNN for border detection permits the gradual improvement of the automatic method towards the semi-automatic by retraining on the results of the latter. For future work, we will investigate the use of depth information to enhance the segmentation results.

## Acknowledgements

The work was funded in part by the Bern University Hospital "InselSpital" and the European Union Seventh Framework Programme (FP7-PEOPLE-2011-IAPP) under grant agreement n° 286408 [www.gocarb.eu].

