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

### Methods

#### Dish Detection

#### Segmentation

1. **Image acquisition:** An image with an elliptical dish is captured by the user.
2. **Edge detection:** The image is downsized and its grey-level histogram equalized. The equalized image is fed to the Canny filter to extract edges.
3. **Edge preprocessing:** Captured by the user.
4. **Robust fitting:** Edge junctions and sharp angles are removed to get curvilinear edge components. Small components are filtered out.
5. **Seed generation:** Groups of components are randomly sampled in increasing size and tested of the edge-set in a RANSAC-like paradigm.
6. **Region growing:** The plate-seed is created automatically as a band inside the ellipse border, followed by:
   - **Automatic:** Seeds are generated on a regular grid.
   - **Semi-auto:** Users generate seeds by swiping the smartphone screen.
7. **Region merging:** The two semi-auto grown regions are merged using the statistical region merging principle, and the ratio of color distance to edge size.

#### Results

The three proposed methods were evaluated on a dataset of 1600 manually annotated images. The chosen evaluation metric is based on the average overlap between the result and the ground truth regions for each image:

\[
\text{Fsum}(G \rightarrow R) = \frac{\sum_i \max_j \left| R_i \cap G_j \right|}{\sum_j |R_j|}
\]

The harmonic mean, \( F_{\text{sum}} \), of \( F_{\text{sum}}(G \rightarrow R) \) and \( F_{\text{sum}}(R \rightarrow G) \) is used as overall score, averaged over all images.

For the ellipse detection, the average score was **99.1%**, with 99.3% of samples reaching accuracies above 98%. Table 1 presents the results of the proposed segmentation methods compared to the literature.

#### Conclusions

The proposed detection and segmentation methods showed high efficiency, outperforming state of the art. These results indicate they are viable solutions for convenient diet assessment from images on mobile devices. For future work, we will investigate the use of texture features and 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]. The consortium would like to thank Swisscom for generously providing the mobile devices and services.

---

**Table 1 – Performance comparison for segmentation methods**

<table>
<thead>
<tr>
<th>Segmentation Method</th>
<th>Average F_{\text{sum}}</th>
<th>Time (sec/image)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed</td>
<td>88.2</td>
<td>0.45</td>
</tr>
<tr>
<td>Meanshift</td>
<td>87.5</td>
<td>2.1</td>
</tr>
<tr>
<td>Local variation</td>
<td>82.6</td>
<td>2.8</td>
</tr>
<tr>
<td>Ultrametric contours</td>
<td>69.2</td>
<td>19</td>
</tr>
<tr>
<td>Floodfill</td>
<td>89.9</td>
<td>0.52</td>
</tr>
<tr>
<td>Proposed</td>
<td>90.8</td>
<td>0.49</td>
</tr>
</tbody>
</table>

