Dish Detection and Segmentation for Dietary Assessment on Mobile Phones

J. Dehais\textsuperscript{1,2}, M. Anthimopoulos\textsuperscript{1}, S. Mougiakakou\textsuperscript{1,3}

\textsuperscript{1}Diabetes Technology Research Group, ARTORG Center for Biomedical Engineering Research, Bern, Switzerland
\textsuperscript{2}Graduate School for Cellular and Biomedical Sciences, University of Bern, Switzerland
\textsuperscript{3}Department of Endocrinology, Diabetes and Clinical Nutrition, Bern University Hospital – Inselspital, Bern, Switzerland
GoCARB System

Captured Image(s) → Dish detection → Segmentation → Nutrient Estimation

<table>
<thead>
<tr>
<th>CHO</th>
<th>Protein</th>
<th>Fat</th>
</tr>
</thead>
<tbody>
<tr>
<td>33</td>
<td>28</td>
<td>8</td>
</tr>
</tbody>
</table>
Dish detection

**Input**
An image of the meal inside an elliptic dish

**Process**
1. Preprocessing and Canny edge detection
2. Filtering of edge graph
3. Incremental RANSAC on edge segments

**Output**
A map with label codes for background and foreground
Ellipse equation
Segmentation
Seed region growing

**Input**
“Seed regions”: pixel sets

**Process**
1. Luminosity-dampened CIE94 colour distance
2. Iterative growth of the regions:
   Add the border pixel with the smallest distance to its region

**Output**
A map with label codes for background and food items
**Segmentation**

Statistical region merging

**Input**

Segmentation map

**Process**

1. Luminosity-dampened CIE94 colour distance

2. Iterative merging of the regions:
   Merge the two regions with the smallest distance between them

**Output**

A map with label codes for background and food items
Segmentation

1. Seeds

2. Region growing

3. Region Merging

Automatic

Semi-automatic
Evaluation - Procedure

**Dataset:**
- 1600 Manually annotated images
- Single large round plate
- Food from the local hospital restaurant

**Scoring:**
- Average overlap between ground truth and result

\[
N_{\text{sum}}(G \Rightarrow R) = \frac{\sum_{i} \max_{j}(|R \cap G|)}{\sum_{i} |R_{i}|}
\]

\[
F_{\text{sum}} = 2 \times N_{\text{sum}}(G \Rightarrow R) \times N_{\text{sum}}(R \Rightarrow G) / (N_{\text{sum}}(G \Rightarrow R) + N_{\text{sum}}(R \Rightarrow G))
\]
## Evaluation

<table>
<thead>
<tr>
<th></th>
<th>Automatic Average $F_{\text{sum}}$ (%)</th>
<th>Time (s/image)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed</td>
<td>88.2</td>
<td>0.45</td>
</tr>
<tr>
<td>Mean-shift [1]</td>
<td>87.5</td>
<td>2.1</td>
</tr>
<tr>
<td>Local Variation [2]</td>
<td>82.6</td>
<td>2.8</td>
</tr>
<tr>
<td>Ultrametric contours [3]</td>
<td>69.2</td>
<td>19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Semi-Automatic Average $F_{\text{sum}}$ (%)</th>
<th>Time (s/image)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed</td>
<td>90.8</td>
<td>0.49</td>
</tr>
<tr>
<td>Flood fill</td>
<td>89.9</td>
<td>0.52</td>
</tr>
</tbody>
</table>


Thank you for your attention.

(Photo credit: Kem Sypher)
Dish detection
Incremental RANSAC

1. Each segment contains s edge points, \( s \gg 5 \)
2. A conic can be generated from \( k \leq 5 \) segments

Incremental sampling:
1. Set \( k=1 \)
2. Apply RANSAC
   Max consensus set size: \( n_k \)
3. If \( n_k > n_{k-1} \) and \( k<5 \)
   1. Set \( k=k+1 \)
   2. Repeat 2-3

Average accuracy: 99.1%
Segmentation
Colour Distance

- CIE94 \((L,a,b,L',a',b') = \sqrt{\Delta L_N^2 + \Delta C_N^2 + \Delta H_N^2}\), with
- \(\Delta L_N = (L - L')\),
- \(\Delta C_N = \sqrt{a^2 + b^2} - \sqrt{a'^2 + b'^2})/(1 + 0.045 \sqrt{a^2 + b^2})\),
- \(\Delta H_N = \sqrt{(a-a')^2 + (b-b')^2 - \Delta C_N^2/(1 + 0.015 \sqrt{a^2 + b^2})}\)
- \(\text{dist}(L,a,b,L',a',b') = |\Delta L_N| + \Delta C_N^2 + \Delta H_N^2\)

<table>
<thead>
<tr>
<th></th>
<th>Av. (F_{sum}) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Automatic</strong></td>
<td></td>
</tr>
<tr>
<td>RGB - Euclidian</td>
<td>61.1</td>
</tr>
<tr>
<td>CIE94</td>
<td>80.2</td>
</tr>
<tr>
<td>Proposed</td>
<td>88.2</td>
</tr>
<tr>
<td><strong>Semi-Automatic</strong></td>
<td></td>
</tr>
<tr>
<td>RGB - Euclidian</td>
<td>72.3</td>
</tr>
<tr>
<td>CIE94</td>
<td>78.2</td>
</tr>
<tr>
<td>Proposed distance</td>
<td>90.8</td>
</tr>
</tbody>
</table>
Segmentation Merging Distance

\[ \text{Dist}(R_i, R_j) = \frac{\text{dist}(L_i, a_i, b_i, L_j, a_j, b_j)}{\sqrt{|\text{edge}_{i,j}|}} \], where

- \( L_i \): Median \( L_p, p \) in \( R_i \)
- \( a_i, b_i \): Average \( a_p, b_p, p \) in \( R_i \)

<table>
<thead>
<tr>
<th>Merging Cost</th>
<th>Av. ( F_{\text{sum}} ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color distance</td>
<td>85.8</td>
</tr>
<tr>
<td>Proposed</td>
<td>88.2</td>
</tr>
</tbody>
</table>