Early and Late Fusion of Classifiers for the MediaEval Medico Task

M. Taschwer¹, M. J. Primus¹, K. Schoeffman¹, O. Marques²
¹ ITEC, Klagenfurt University, Austria
² Florida Atlantic University, USA

MediaEval’18, Sophia Antipolis, France
October 30, 2018
Approach

LIRE features

early fusion

GoogLeNet features

feature extraction

linear SVM

kernel SVM

random forest

logistic regression

16 class probabilities

multi-class classifiers

late fusion

single class prediction

run 1

run 2

run 3

run 4

run 5
Results

<table>
<thead>
<tr>
<th>Run</th>
<th>accuracy</th>
<th>F1</th>
<th>MCC</th>
<th>T / ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR</td>
<td>0.9873</td>
<td>0.8986</td>
<td>0.8919</td>
<td>0.119</td>
</tr>
<tr>
<td>LSVM</td>
<td><strong>0.9876</strong></td>
<td><strong>0.9008</strong></td>
<td><strong>0.8942</strong></td>
<td><strong>0.103</strong></td>
</tr>
<tr>
<td>KSVM</td>
<td>0.9865</td>
<td>0.8921</td>
<td>0.8849</td>
<td>25.808</td>
</tr>
<tr>
<td>RF</td>
<td>0.9843</td>
<td>0.8747</td>
<td>0.8664</td>
<td>0.828</td>
</tr>
<tr>
<td>RF-KSVM-LR</td>
<td>0.9875</td>
<td>0.9002</td>
<td>0.8936</td>
<td>26.783</td>
</tr>
</tbody>
</table>

T – mean prediction time per image w/o feature extraction
Results

<table>
<thead>
<tr>
<th>Run</th>
<th>accuracy</th>
<th>F1</th>
<th>MCC</th>
<th>$T$ / ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR</td>
<td>0.9873</td>
<td>0.8986</td>
<td>0.8919</td>
<td>0.119</td>
</tr>
<tr>
<td>LSVM</td>
<td>0.9876</td>
<td>0.9008</td>
<td>0.8942</td>
<td>0.103</td>
</tr>
<tr>
<td>KSVM</td>
<td>0.9865</td>
<td>0.8921</td>
<td>0.8849</td>
<td>25.808</td>
</tr>
<tr>
<td>RF</td>
<td>0.9843</td>
<td>0.8747</td>
<td>0.8664</td>
<td>0.828</td>
</tr>
<tr>
<td>RF-KSVM-LR</td>
<td>0.9875</td>
<td>0.9002</td>
<td>0.8936</td>
<td>26.783</td>
</tr>
</tbody>
</table>

$T$ – mean prediction time per image w/o feature extraction
Findings

• Unsuccessful design options:
  – GoogLeNet features trained for surgical actions on a different endoscopic video dataset
  – Ensemble of nested dichotomies (binary classifiers arranged in a tree)

• Successful techniques:
  – Early fusion of traditional and CNN-based features
  – Traditional linear classifiers are both effective and efficient on this dataset