Optimized data analysis pipeline for MALDI MSI based tumor typing from FFPE tissue samples evaluated on six benchmark classification tasks

D. Lachmund\(^1\), J. von Schröder\(^1\), T. Boskamp\(^1,2\), L. Hauberg-Lotte\(^1\), J.H. Kobarg\(^3\), S.O. Deininger\(^3\), K. Kriegsmann\(^4\), M. Kriegsmann\(^4\), R. Casadonte\(^5\), J. Kriegsmann\(^5\), P. Maaß\(^1,2\)

\(^1\)University of Bremen, Bremen, Germany \(^2\)SCiLS, Bremen, Germany \(^3\)Bruker Daltonik, Bremen, Germany \(^4\)University of Heidelberg, Heidelberg, Germany \(^5\)Proteopath, Trier, Germany

**Goals**
- Develop optimized pre-processing pipeline for MALDI MSI based tumor typing
- Consider different clinical tumor typing and subtyping tasks
- Consider intra- and inter-lab scenarios and different instrument types

**Tasks**

<table>
<thead>
<tr>
<th>Task</th>
<th>Instrument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Antalya autoflex</td>
<td>Four tumor entities, 8 TMAs</td>
<td>Lung, pancreas, colon, breast</td>
</tr>
<tr>
<td>2. Bruker MTR rapiflex</td>
<td>Six tumor entities on one TMA</td>
<td>Lung, pancreas, colon, breast</td>
</tr>
<tr>
<td>3. Lx ADSQ autoflex</td>
<td>Eight TMAs with mix of adenocarcinoma and squamous cell carcinoma, ax</td>
<td>Lung, pancreas, colon, breast</td>
</tr>
<tr>
<td>4. MDP axf intrax autoflex</td>
<td>Breast, ovary tumors, 5 TMAs</td>
<td>Measured in two labs</td>
</tr>
<tr>
<td>5. MDP axf intra cross-validation</td>
<td>Same as above, but intra-lab cross-validation</td>
<td></td>
</tr>
<tr>
<td>6. MDP axf rapiflex</td>
<td>Breast, ovary tumors, 5 TMAs</td>
<td>Single lab</td>
</tr>
</tbody>
</table>

**Normalization**
- Non-linear intensity profile normalization (IPN, bottom row) improves comparability across different acquisitions (Boskamp et al, ASMS 2018)

**Alignment**
- Mild Gaussian kernel spatial denoising (right) increases signal-to-noise ratio as compared to original data (left)

**Downsampling**
- Dimensionality reduction by downsampling to peak areas over 0.4 Da intervals (Boskamp et al, ASMS 2018)

**Data**
- Benchmark panel acquired from 25 TMAs, 2031 cores and 1410 patients total

**Transformation**
- Mass calibration based on statistical peptide mass model reduces misalignment (Boskamp et al, ASMS 2018)

**Classification**
- Balanced accuracy 82% and 92%
- Performance gain over baseline (TIC only) 9.5 ... 39.8% pts. for five of six tasks
- Mass alignment / downsampling alone yields 5 ... 18.5% pts. for five of six tasks

**Conclusion**
- Systematic investigation of six benchmark problems yields an optimized pre-processing pipeline for MALDI MSI tumor typing applications
- Significant performance gains achieved in intra- and inter-lab scenarios
- Improved robustness towards SOP variations and technical variability