Impact of Timestamp and Segmentation Map Selection for Cancerous Prostate Regions in DCE MRI Classification

Aleksas Vaitulevičius[1], Povilas Treigys[1], Jolita Bernatavičienė[1], Jurgita Markevičiūtė[1], Ieva Naruševičiūtė[2], Mantas Trakymas[2]

[1]Vilnius University, institute of data science and digital technologies

[2]National Cancer Instityte

aleksas.vaitulevicius@mif.stud.vu.lt, povilas.treigys@mif.vu.lt, jolita.bernataviciene@mif.vu.lt, jurgita.markeviciute@mif.vu.lt, ieva.naruseviciute@nvi.com, mantas.trakymas@nvi.com

Introduction

Dynamic contrast enhancement image sequence is one of the multiparametric magnetic resonance imaging modalities used to detect cancerous regions in the prostate. This image sequence is acquired by capturing the prostate region several times, resulting in prostate region images acquired in different timestamps with an interval of several seconds.

By investigating the previous research more thoroughly, this study aims to inspect whether the timestamp selection influences classification performance and to test segmentation map selection on functional data using the nearest centroid classifier.

Experimental investigation terms

- **Segmentation** is performed on a single timestamp by using **Simple Linear Iterative Clustering**. Two approaches are tested - **fixed** and **proportionate** number of segments to prostate size in slice.
- **Curve construction** consists of projecting segmentation map on all timestamps, calculating median on each projected and original segment and performing Bspline basis function smoothing on resulting timeseries with **landmark registration** step being tested.
- Experiments are performed on 13 patients separately and then aggregated by calculating medians.

Results

1st Experiment

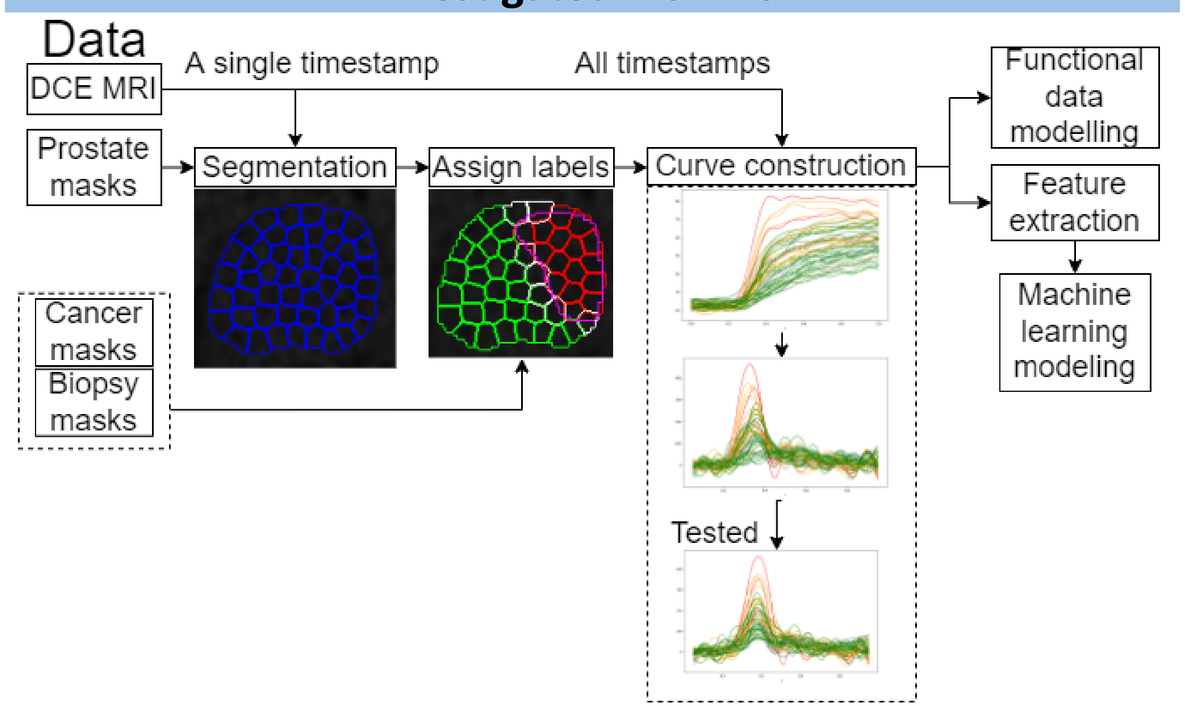
- Experiments of segmentation map selection on functional data using the nearest centroid classifier. Metrics are calculated on training set.
- Wilcoxon tests are used to determine statistical significance.

Medians on single combination of configurations basis						
registered	**fixed number of SLIC regions**	**precision**	**recall**	**F1**	**Balanced accuracy**	**specificity**
unregistered	Proportionate SLIC	**0.127**	**0.718**	**0.225**	**0.772**	**0.816**
	fixed SLIC	**0.136**	**0.748**	**0.239**	**0.801**	**0.834**
registered	Proportionate SLIC	**0.122**	**0.714**	**0.214**	**0.761**	**0.8**
	fixed SLIC	**0.127**	**0.727**	**0.225**	**0.784**	**0.828**

Medians on single configurations basis					
configuration	**precision**	**recall**	**F1**	**Balanced accuracy**	**specificity**
unregistered	**0.132**	**0.736**	**0.233**	**0.787**	**0.827**
registered	**0.123**	**0.714**	**0.218**	**0.77**	**0.825**
proportionate number of SLIC regions	**0.123**	**0.714**	**0.218**	**0.765**	**0.809**
fixed number of SLIC regions	**0.132**	**0.741**	**0.231**	**0.796**	**0.832**

Configurations		P-value
Registered	Unregistered	8.76×10^{-36}
Proportionate SLIC	Fixed SLIC	7.896×10^{-14}

Investigated workflow



2nd Experiment

- Experiments of timestamp selection influence on classification performance. **XGBoost** classification model is used.
- Features are scaled by normalizing them to interval [0, 1]. Training set is 70% of dataset and validation set is 30%. Balancing weights are used in training. Early stopping is used if validation set's F1 score does not improve after iterating through training set 10 times.
- **Hyper parameter tuning** is performed with **Tree of Parzen Estimators** algorithm and **maximizing F1 score**. Single training is performed with at least 400 trials by minimizing **logistic regression**. Number of gradient boosted trees is 180. Tuned hyper parameters are (with search space interval):
  - Subsample ratio of columns when constructing each tree [0.4, 0.8].
  - Min loss reduction required to make a further partition on a leaf node of the tree [0, 1].
  - Max depth of the XGBoost ensemble tree .[3, 18]
  - Min sum of instance weight (Hessian) needed in a child [0, 10].
  - L1 regularization term on weights [0, 1].
- Statistical significance is calculated on validation set's balanced accuracies with Friendman's tests.

only discrete	fixed number of slic regions	P-value
not only discrete	proportionate SLIC	**0.433**
	fixed SLIC	**0.35**
only discrete	proportionate SLIC	**0.778**
	fixed SLIC	**0.663**

Conclusions

- Modeling with functional data:
  - Unregistered functional data gives significantly more accurate results than unregistered and the comparison is statistically significant.
  - Fixed number of SLIC zones gives significantly more accurate results than proportionate number of SLIC zones to prostate size and the comparison is statistically significant.
- Modeling with extracted features: there is no significant difference between timestamps for segmentation.
- Further investigations:
  - The experiments should be repeated on higher data variability from more patients. The data variability could be used to explain the proportionate number SLIC zones performance with flat neural networks.
  - The search of ensemble classifier that merge the proposed scheme of processing DCE modality with processing other prostate MRI modalities could improve the results.