DCE MRI Modality Investigation for Cancerous Prostate Region Detection: Case Analysis

Aleksas Vaitulevicius
Vilnius University
Institute of Data Science and Digital Technologies
Vilnius, Lithuania
aleksas.vaitulevicius@mif.stud.vu.lt

Povilas Treigys
Vilnius University
Institute of Data Science and Digital Technologies
Vilnius, Lithuania
povilas.treigys@mif.vu.lt

Jolita Bernataviciene
Vilnius University
Institute of Data Science and Digital Technologies
Vilnius, Lithuania
jolita.bernataviciene@mif.vu.lt

Roman Surkant
Vilnius University
Institute of Applied Mathematics
Vilnius, Lithuania
roman.surkant@gmail.com

Jurgita Markeviciute
Vilnius University
Institute of Applied Mathematics
Vilnius, Lithuania
jurgita.markeviciute@mif.vu.lt

Ieva Naruseviciute
National Cancer Institute
Vilnius, Lithuania
ieva.naruseviciute@nvi.lt

Mantas Trakymas
National Cancer Institute
Vilnius, Lithuania
mantas.trakymas@nvi.lt

ABSTRACT

Typically, prostate evaluation is done by using different imaging sequences of magnetic resonance imaging. Dynamic contrast enhancement, one of such scanning modalities, allow to spot higher vascular permeability and density caused by the malignant tissue. Authors of this paper investigate the ability to identify malignant prostate regions by the functional data analysis and standard machine learning techniques. The dynamic contrast enhanced images of the prostate are divided into the regions and based on those time-signal intensity curves are calculated. Two classification approaches: functional k-Nearest Neighbors and machine learning Support Vector Machine are used to model signal curve behavior on temporal variation matrix and timestamp based prostate region division of image data. Preliminary research shows that both functional data analysis and machine learning classification methods are able to identify highest saturation timestamp that gives best tissue classification results on timestamp based dynamic contrast enhanced region map obtained by Simple Linear Iterative Clustering algorithm. Cancer region classification results are better when the dynamic contrast enhanced images are subdivided into regions at each timestamp than when using a temporal variation matrix.

Keywords

prostate cancer, functional data analysis, machine learning, component, dynamic contrast-enhanced MRI

1 INTRODUCTION

Prostate cancer is one of the leading causes of cancer death worldwide. Among males, prostate cancer has second highest incidence rate after lung cancer according to the research given in paper [Bra20]. Although death rates have been decreasing in some countries, it remains a considerable disease affecting many patients. Due to nature of cancer, early diagnosis and treatment is critical. Preliminary identification of cancer involves Prostate-Specific Antigen (PSA) screening measuring concentration of a protein produced by the prostate, and the concentration is elevated inpatients with prostate cancer. Due to high level of false-negative and false-positive cases in PSA testing which lead to incorrect biopsies, a less invasive and more reliable procedure is needed. With the introduction of PI-RADS in the paper [Alq20], a structured reporting scheme for multi-parametric (mp) prostate Magnetic Resonance Imaging

(MRI) based on literature evidence given in the same paper and consensus expert opinion, the interpretation and performance of prostate cancer evaluation has considerably improved. Cancer evaluation is done by using different types of imaging (T2 Weighted images (T2W), Diffusion Weighted Images (DWI), Apparent Diffusion Coefficient map (ADCmap), Dynamic Contrast-Enhanced (DCE) images, etc.), each having own acquisition methods and purpose. Radiologists typically use at least several imaging sequences for more accurate diagnosis.

This research is focused solely on DCE sequence. Prostate DCE MRI data is gathered by capturing imaging sequences of the entire region of prostate during an intravenous injection of a contrast agent (typically gadolinium). Over a course of several minutes, a set of cross-sectional images is created at different time moments, usually every few seconds. The role of the contrast agent is to evaluate angiogenesis of tumor in DCE imaging. Since blood vessels are essential to cancer growth, tumors typically have higher vascular permeability and density which attracts higher amount of contrast medium as it is described in paper [Low11]. After acquiring such data, each cross-sectional image can be segmented into regions by using algorithm such as Simple Linear Iterative Clustering (SLIC) algorithm. Those regions can then be aggregated to single value by calculating mean, median or another metric of each region. Therefore, collected data can be fitted to functions $f_{xy} : T \rightarrow I$, where T is set of time points in which observations were made, I - set of aggregated intensity values and x, y are coordinates of the pixel. Functional Data Analysis (FDA) can be applied on these functions to detect, characterize, and monitor tumors together with machine learning methods.

A lot of work was already done in prostate cancer localization research by using T2W sequences for example in the paper [Juc16]. Moreover, DWI sequences were used to solve prostate cancer segmentation and severity evaluation problems. Examples of such papers are [Hot16], [WuC15] and [Bar15]. However, recent improvement in DCE MRI technology, described in paper [Cha18], create a motive to research DCE MRI sequences. The example of such paper is [Liu19]. However, in this paper only machine learning without FDA approach is tested. Therefore, the aim of this paper is to investigate a dynamic contrast imaging evaluation method for cancerous prostate zones localization in with focus to compare FDA and classical machine learning classification approaches while validating algorithms identifies cancerous zones with the ground truth samples obtained by histological tissue analysis after biopsy. The data for the investigation was provided by the Lithuanian National Cancer Institute (NVI) under the terms of bioethical agreement.

2 WORKFLOW

The structure of analysis workflow consists of data preparation, data preprocessing, segmentation, curve construction, data visualization, functional data analysis modeling and machine learning modeling steps, Fig. 1. The data used for investigation consists of four types: DCE MRI images (an example of single slice and 3 different timestamps is given in Fig. 2), prostate region masks (see Fig. 3), cancer region masks (see Fig. 4), and biopsy results.

The peculiarity of the MRI DCE image construction is that during the scan patient's prostate is scanned numerous times observing contrast agent saturation in tissue. Fast tissue region contrast saturation and fast agent washout may indicate malignant tissue region. As the maximum contrast saturation timestamp is not known prior, thus in this research experiments are conducted with two different data acquisition for magnetic resonance images approaches: segmenting temporal variance matrix calculated between all timestamps and segmenting each timestamp separately.

The block labeled as Data in Fig. 1 corresponds to data used in experiment. Examples of data types of data, used in experiment, are shown in Fig. 2, Fig. 3 and Fig. 4. In the Fig. 5 the overlay of these data types and biopsy masks is provided. More detailed characteristics of data types are explained in this section.

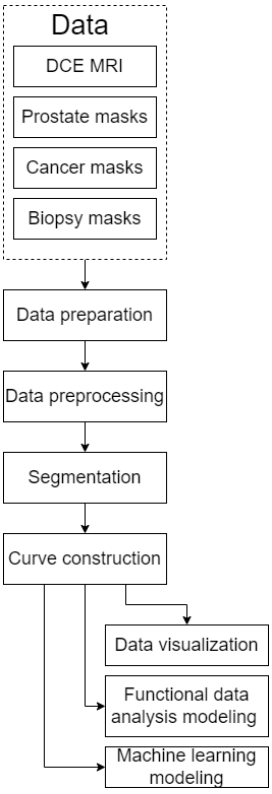

Figure 1: Analysis workflow structure.

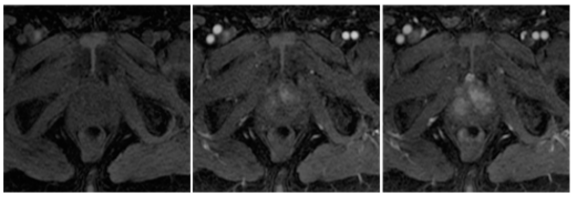

Figure 2: Example of DCE images of a single slice in different timestamps.

2.1 DCE MRI

DCE MRI data are grayscale images that are captured during an MRI procedure. Such images indicate the distribution of contrast agent in a particular cross-section of patient's pelvic region. The differentiating aspect of dynamic contrast imaging compared to other image sequence types is its bi-dimensional nature. DCE MRI sequence has a temporal dimension indicating the time moment when the DCE image was recorded ("timestamp"), and a spatial dimension linking the DCE image to a location of the cross-section in patient's body ("slice"). This allows to interpret the data in a 3D time series-based manner.

As stated earlier DCE MRI modality is the primary target of this research. This dataset contains 135660 anonymized images of 144 patients. Each patient having on average 41 slices (total range: 25 - 102) and 26 timestamps (total range: 5 - 55). For the case analysis and classification one patient data having explicit cancerous regions confirmed by biopsy were selected as not all of the patients have histological registration performed yet. This patient has 26 slices with prostate and 31 timestamps. Authors plan to generalize the investigation on all patients after finished histology registration.

2.2 Prostate region masks

Prostate masks are binary image-type data which indicates the region of prostate for each DCE MRI slice (all timestamps of the same slice have the same prostate region). Such masks were segmented manually by medical experts at NVI. Since MRI covers an area both above and below the prostate, some prostate masks are blank (i.e., are black).

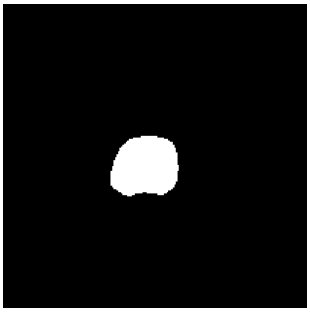

Figure 3: Prostate mask of a single slice.

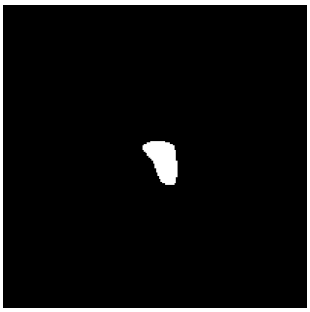

Figure 4: Cancerous region mask.

2.3 Cancer region masks

Similarly, to prostate region masks, cancer region masks are binary images showing the location of cancer tumor which are manually segmented by experts and may be blank. Tumors can be of three types: malignant, clinically insignificant, and benign. Cancer masks indicate only the suspected region of cancer according to medical experts. Factual tumor type is diagnosed through a biopsy.

2.4 Biopsy results

Each patient underwent on average 15 biopsies (total range: 3 - 25), each biopsy showing tumor type and tumor severity based on a Gleason score [Alq20]. The biopsy outcome data is split into two datasets of different format:

- Tabular: contains numerical identifier of patient, slice, biopsy, three Gleason scores (first, second, and combined).
- Mask: multi-label mask showing the location of biopsies; each biopsy has a numerical identifier linking to the tabular dataset.

While biopsies provide ground truth label for tumors, there are two flaws:

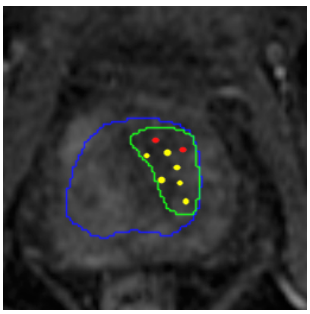

Figure 5: Overlay image combines middle image of Fig. 2, Fig. 3, Fig. 4 and biopsy mask images: prostate region (blue), cancer region (green), malignant biopsy (red), clinically insignificant or benign biopsy (yellow).

- Biopsies provide information only in a point-wise manner. The gaps between biopsies are essentially missing data where ground truth information is unavailable.
- To perform a biopsy, a medical expert inserts a needle into prostate and extracts a tissue sample for examination. The needle punctures several slices without knowing which slice exactly the tissue was taken from. Such process introduces a systematic uncertainty, for instance, a slice having no tumors may show positive diagnosis if cancerous tissue was extracted from a different slice above or below.

2.5 Data preparation

Before workflow data preparation is performed. The first step of data preparation is extracting metadata from each image object (MRI, cancer masks, biopsy masks). Extracted metadata parameters are patient ID, timestamp ID, slice ID, resolution, maximum pixel intensity, minimum pixel intensity, mean pixel intensity. This information allows to detect incompatible data (e.g. differing image resolution) or missing data (e.g. MRI images containing only pixels of identical value). The second step is biopsy aggregation which is the collection and aggregation of biopsy information from multiple CSV-type files into single source for more convenient processing.

2.6 Data preprocessing

The next step of workflow is data preprocessing. Data preprocessing consist of 2 steps. The first step is mask rescaling which is transformation of prostate and cancer mask files into binary format compatible with used algorithms and the other step is image rescaling which is transformation of MRI images into appropriate format with 8-bit pixel precision (256 grayscale values).

2.7 Segmentation

As mentioned before experiments in this research are done with two different approaches. The first approach is using Temporal variation matrix (TVM) calculated between all timestamps. The first step is calculating TVM. This step is a construction of a matrix with identical resolution as source MRI image which represents statistical variance of signal value of each pixel between all timestamps of a selected slice. TVM shows regions with high (bright) and low (dark) change in signal intensity over time which is used for region of interest segmentation. Example of a TVM is shown in Fig. 6 which is calculated from slice whose timestamp examples are displayed in the Fig. 2.

The next step is TVM segmentation by applying SLIC algorithm, introduced in paper [Ach12], to separate the prostate region into segments. Selected SLIC algorithm

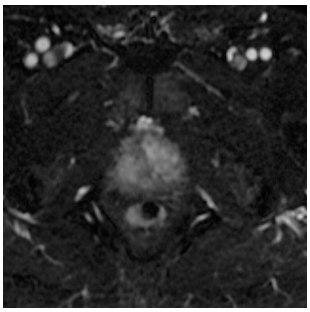

Figure 6: Temporal variation matrix calculated from all timestamps of slice whose timestamp examples are given in Fig. 2.

parameters are 50 and 7 for segment number and compactness respectively. Those parameters were selected using expert judgement based on the following criteria:

- Number of segments parameter should be high enough to have separation between cancerous and healthy tissue, but low enough to remain computationally relevant and keep segments visually distinguishable.
- Compactness parameter should capture similar intensities, but still prioritize color proximity in favor of maintaining circular shape because cancerous growth typically does not have irregular formation.

Acquired segment locations are projected back into MRI images that were used for TVM construction (example of TVM segments in Fig. 7). Lastly, segment-level aggregation is performed by calculating the means of pixel intensities inside each segment.

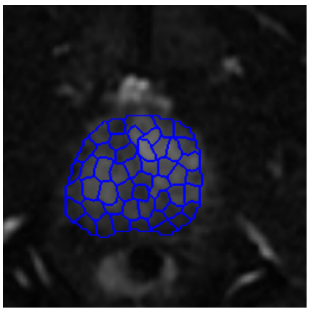

Figure 7: Segmented and zoomed TVM displayed in Fig. 6.

The second approach is segmenting each MRI image separately. The first step is segmentation of each MRI image by applying SLIC algorithm to separate the prostate region into segments in each timestamp separately. Resulting in the number of segmentations which is equal to number of timestamps. The parameters chosen for SLIC algorithm are the same as for the first approach. Then each segmentation is projected to all timestamps e.g.: patient, for which MRI scan

was applied 30 times, will have 30 timestamps. Each timestamp is segmented resulting in 30 segmentations. Each segmentation is projected to all 30 timestamps. Further workflow is applied for each segmentation separately. Lastly, segment-level aggregation is performed by calculating the means of pixel intensities inside each segment.

2.8 Curve construction

The resulting time series of each segment contains information of mean intensity of intensities within each segment at the given timestamp. The greater intensity indicates higher concentration of contrast agent. As explained in the introduction, the malignant tissue accumulates contrast agent faster than a healthy one. Therefore, these time series are used to construct functional data which later on is transformed to represent mean intensity growth speed within each region.

Each segment's aggregated intensity values are used to construct a single time series curve. Prior to this step, each segment is labeled into binary classes. Positive class 1 is assigned to segments which has an intersection with cancerous region mask $\geq 50\%$, overlaps with at least one malignant biopsy and does not overlap with any other biopsy. Negative class 0 is assigned to segments which does not overlap with malignant or clinically insignificant biopsies and does not have intersection with cancerous region mask. Meanwhile, the remaining segments are not used in the training or validation. These segments are either:

- have clinically insignificant or malignant biopsy while no overlap with cancerous region mask.
- have $< 50\%$ overlap with cancerous region mask.
- have $\geq 50\%$ overlap with cancerous region mask but no malignant biopsy.

The example of labeled segments is displayed in the Fig. 8. Segments marked with green contour have negative class, while red contour - positive class. Segments with orange contour have overlap $\geq 50\%$ with cancerous region mask and no overlap with malignant biopsy. Furthermore, segments with white contour either have intersection $< 50\%$ with cancerous region mask or no overlap with cancerous region mask but have other than benign biopsy. Segments with green and red contour are used in modeling while the rest are not. The curves of discrete points of this example are displayed in Fig. 9.

Afterwards, the time axis of those curves of discrete points are normalized to interval [0, 1]. The resulting curves of discrete points are then smoothed by using B-spline basis function whose calculation is presented in the paper [DeB72]. The parameters of this calculation are:

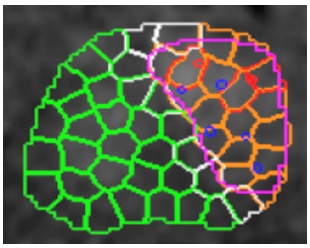

Figure 8: Example of segment classification where segments marked with red contour have class 1, segments with green contour - class 0 and segments with white or orange contours are not used in modeling.

- Order - the order of polynomial function, called B-splines. The used value for this parameter is 4.
- basis function number - number of basis functions to use for calculation. The used value for this parameter is 18.

These parameters were obtained by calculating the best Generalized Cross Validation (GCV) score on a single patient by using grid search-type algorithm. Tested combination of parameters are 2, 3, 4, 5 for order and 5, 6, ..., 29, 30, 35, 40 for basis function number. Functional data, related to Fig. 9, are displayed in Fig. 10.

Furthermore, in curve construction step, derivatives of first degree are calculated from functional data. These calculations are described in the paper [But76] and these derivatives are interpreted as velocity of intensity change over time. The first degree derivatives of functional data, related to Fig. 10, are displayed in Fig. 11.

Lastly, these 1st degree derivatives of functional data are registered by using landmark registration The chosen landmarks are the points in t axis in which functional data has a maximum value. The registrations are performed for each patient and segmentation separately. Registered functional data related to Fig. 11 are displayed in Fig. 12.

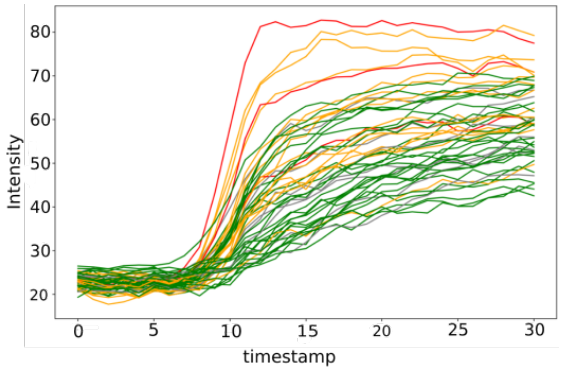

Figure 9: Example of curve created from discrete points where red curves are curves of class 1, green curves - of class 0 while orange is not used in modeling.

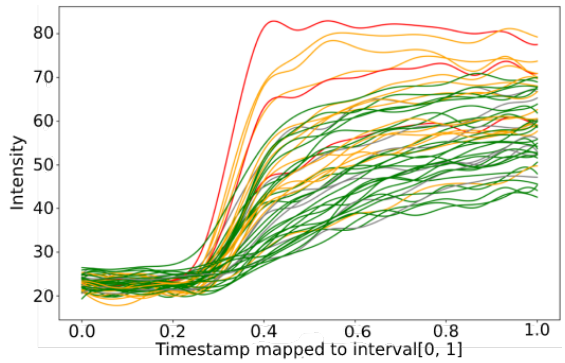

Figure 10: Functional data displayed in Fig. 9 and created by using B-spline.

2.9 Functional data analysis modeling

In this workflow, 2 different approaches of SLIC region classification are tested. The first is Functional Data Analysis approach. In this approach model is created by using k-Nearest Neighbors (kNN) algorithm and using functional data as training set. The neighboring parameter was obtained by the grid search. Seven neighbors were depicted as the optimal choice.

2.10 Machine learning modeling

The second approach of this workflow to classify SLIC regions is modeling by using Machine learning method - Support Vector Machine (SVM) algorithm. Training set used for training is extracted features from functional data derivatives: integrated depth, modified band depth, maximum intensity, time of maximum intensity and 10 uniformly spaced intensities from discretized curve derivative in interval [0.05, 0.95].

Due to the unbalanced class problem SVM model is validated by using cross validation method - stratified 5 k-folds.

3 RESULTS

In the Table 1, metrics of results of SLIC zone classification with different approaches are shown. For precision, recall and F1 metrics the positive class is cancerous region and referred as 1. Classes are highly unbalanced, the number of data points with class 1 for TVM approach is 63 while for single timestamp segmentation - 64. Meanwhile, the number of data points with class 0 for TVM approach is 1184 while for single timestamp segmentation - 1185.

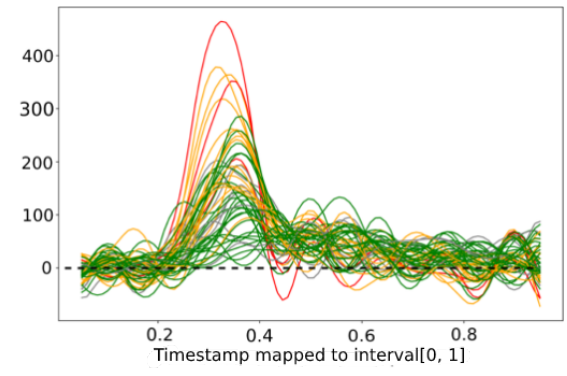

Figure 11: 1st degree derivatives of curves displayed in 10.

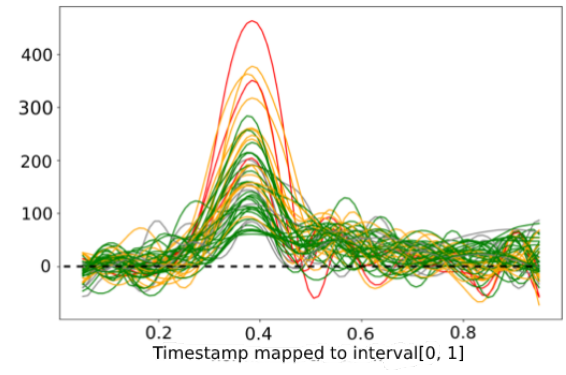

Figure 12: Registered functional data displayed in Fig. 11.

For segmenting in each image separately approach multiple different are obtained as number of different ways to segment the prostate to regions is equals to number of timestamps. In the Table I., the chosen timestamps for each FDA and modeling approach are the one which achieve the best results. For FDA modeling it was 10th timestamp and for ML modeling - 8th.

Number of SVM models are equal to number of folds used in stratified K-folds method which is 5. The metrics of these models are aggregated by calculating mean and standard deviation STD. The best results is produced by the kNN functional data classifier.

4 CONCLUSION

Preliminary research investigating MRI-DCE modality scans by applying functional data analysis and machine learning methods was presented in the paper. The results obtained by the comparison of the machine learning and FDA methods allows authors to conclude:

- Both FDA and ML based classification approaches gives best results at almost the same tissue saturation timestamp while using non-TVM intensity map.
- Grid search of neighboring parameter indicated that the seven neighbors is an optimal choice giving best classification results.
- Timestamp base SLIC application outperforms TVM intensity mapping giving FDA kNN classification precision of 1, recall 0.65, and F1 score - 0.71.

Metric	**FDA modeling**		**ML modeling**							
	TVM	***Best timestamp SLIC segmentation***	***TVM***				***Best timestamp SLIC segmentation***			
Set	**-**	**-**	***Training***		***Testing***		***Training***		***Testing***	
	-	**-**	***Mean***	***STD***	***Mean***	***STD***	***Mean***	***STD***	***Mean***	***STD***
Precision	0.929	1	0.969	0.016	0.8	0.267	0.960	0.08	0.7	0.4
Recall	0.413	0.625	0.504	0.036	0.303	0.125	0.719	0.109	0.5	0.316
F1	0.571	0.714	0.663	0.034	0.428	0.153	0.817	0.085	0.567	0.327
Balanced accuracy	0.706	0.778	0.752	0.018	0.649	0.062	0.859	0.055	0.75	0.158
Specificity	0.998	1	0.999	0.0004	0.996	0.005	0.9998	0.0004	0.999	0.002

Table 1: Modelling results

- SLIC algorithm applied to TVM intensity mapped images in ML modeling produces more stable testing results in terms of standard deviation than those obtained by SLIC applied best timestamp.

Obtained results shows potential points of further action for the modeling results improvement:

- Class discrimination can be improved by incorporating data from the functional boxplot analysis.
- More extensive search of new features, feature combinations, feature transformations may improve model learning.
- Extensive search of more suitable model parameters may improve model learning.
- Proposed workflow needs to be applied on all 144 patient data as currently not all of them has histological tissue analysis performed. Moreover, a lot of patients DCE image data results in unusual functional data. This may indicate faults in data or limitations of the proposed workflow. Therefore, those cases have to be examined.
- underrepresented classes strongly affect model's ability to segregate different classes. Thus resampling techniques such as over-sampling as well as configuring balancing class weights during model training may improve class discrimination and shall be used in further experiments with greater number of patients.
- Further experiments have to be conducted with greater number of patients to determine the timestamp for segmentation which achieves the most accurate results.

5 ACKNOWLEDGEMENT

The authors are thankful for the high performance computing resources provided by the Information Technology Research Center of Vilnius University.

6 REFERENCES

[Ach12] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Susstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE transactions on pattern analysis and machine intelligence 34, no. 11 (2012): 2274-2282.

[Alq20] Alqahtani, S., Wei, C., Zhang, Y., Szewczyk-Bieda, M., Wilson, J., Huang, Z., and Nabi, G. Prediction of prostate cancer Gleason score upgrading from biopsy to radical prostatectomy using pre-biopsy multiparametric MRI PIRADS scoring system. Scientific reports 10.1 (2020): 1-9.

[Bar15] Barrett, T., Priest, A.N., Lawrence, E.M., Goldman, D.A., Warren, A.Y., Gnanapragasam, V.J., Sala, E., and Gallagher, F.A. Ratio of Tumor to Normal Prostate Tissue Apparent Diffusion Coefficient as a Method for Quantifying DWI of the Prostate. American Journal of Roentgenology vol 205, (2015): 585-593. Doi: 10.2214/AJR.15.14338.

[Bra20] Bray, Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, and L.A. Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians 68.6 (2018): 394-424.

[But76] Butterfield, K.R. The computation of all the derivatives of a B-spline basis. IMA Journal of Applied Mathematics 17, no. 1 (1976): 15-2

[Cha18] Chatterjee, A., He, D., Fan, X., Wang, S., Szasz, T., Yousuf, A., Pineda, F., Antic, T., Mathew, M., Karczmar, G.S., and Oto, A. Performance of ultrafast DCE-MRI for diagnosis of prostate cancer. Academic radiology, 25(3) (2018): 349-358.

[DeB72] De Boor, C. On calculating with B-splines. Journal of Approximation theory 6, no. 1 (1972): 50-62.

[Hot16] Hotker, A.M., Mazaheri, Y., Aras, O., Zheng, J., Moskowitz, C.S., Gondo, T., Matsumoto,

K., Hricak, H., and Akin, O. Assessment of Prostate Cancer Aggressiveness by Use of the Combination of Quantitative DWI and Dynamic Contrast-Enhanced MRI. AJR. American journal of roentgenology vol. 206,4 (2016): 756-63. Doi:10.2214/AJR.15.14912.

[Juc16] Jucevicius, J., Treigys, P., Bernataviciene, J., Briediene, R., Naruseviciute, I., Dzemyda, G., and Medvedev, V. Automated 2D Segmentation of Prostate in T2-weighted MRI Scans. International journal of computers communication & control, [S.l.], v. 12, n. 1, p. 53-60, dec. 2016. ISSN 1841-9844.

[Liu19] Liu, B., Cheng, J., Guo, D.J., He, X.J., Luo, Y.D., Zeng, Y., and Li, C.M.. Prediction of prostate cancer aggressiveness with a combination of radiomics and machine learning-based analysis of dynamic contrast-enhanced MRI. Clinical radiology, 74(11) (2019): 896-e1.

[Low11] Low, R. N., Fuller, D. B., and Muradyan, N. Dynamic gadolinium-enhanced perfusion MRI of prostate cancer: assessment of response to hypofractionated robotic stereotactic body radiation therapy. American Journal of Roentgenology 197.4 (2011): 907-915.

[WuC15] Wu, C.J., Wang, Q., Li, H., Wang, X.N., Liu, X.S., Shi, H.B., and Zhang, Y.D. DWI-associated entire-tumor histogram analysis for the differentiation of low-grade prostate cancer from intermediate-high-grade prostate cancer. Abdom Imaging 40, 3214-3221 (2015).