A 3D voxel neighborhood classification approach within a multiparametric MRI classifier for prostate cancer detection

A 3D Voxel Neighborhood Classification Approach within a Multiparametric MRI Classifier for Prostate Cancer Detection

Francesco Rossi[1], Alessandro Savino[1], Valentina Giannini[2], Anna Vignati[2], Simone Mazzetti[2], Alfredo Benso[1], Stefano Di Carlo[1], Gianfranco Politano[1], and Daniele Regge[2]

[1] Politecnico di Torino, Control and Comp. Engineering Department, Torino, Italy
E-mail: <firstname>.<lastname>@polito.it
Home page: http://www.sysbio.polito.it
[2] Radiology Unit, Candiolo Cancer Institute FPO, IRCCS, Candiolo (Torino), Italy
E-mail: <firstname>.<lastname>@ircc.it

Abstract. Prostate Magnetic Resonance Imaging (MRI) is one of the most promising approaches to facilitate prostate cancer diagnosis. The effort of research community is focused on classification techniques of MR images in order to predict the cancer position and its aggressiveness. The reduction of False Negatives (FNs) is a key aspect to reduce mispredictions and to increase sensitivity. In order to deal with this issue, the most common approaches add extra filtering algorithms after the classification step; unfortunately, this solution increases the prediction time and it may introduce errors. The aim of this study is to present a methodology implementing a 3D voxel-wise neighborhood features evaluation within a Support Vector Machine (SVM) classification model. When compared with a common single-voxel-wise classification, the presented technique increases both specificity and sensitivity of the classifier, without impacting on its performances. Different neighborhood sizes have been tested to prove the overall good performance of the classification.

Keywords: Prostate cancer; magnetic resonance imaging; support vector machine; MRI classification.

1 Introduction

Prostate Cancer (PCa) is one of the most frequent cancer in males, and it is the third leading cause of cancer-related death among European men [1]. According to clinical guidelines, one of the most commonly used methods to detect prostate cancer is a Transrectal Ultrasound (TRUS) guided biopsy that, unfortunately, has been proven to provide limited efficacy to differentiate malignant from benign tissues [2]. Another accepted screening method is the antigen (PSA) blood test, which has been linked to over diagnosis and over treatments [3]. Recently, diagnostic improvements have been made by evaluating the information extracted from magnetic resonance image (MRI) sequences such as conventional

morphological T1-weighted (T1-w) and T2-weighted (T2-w) imaging, diffusion-weighted MRI (DW-MRI) and dynamic contrast-enhanced MRI (DCE-MRI). These Multiparametric MRI techniques are promising alternatives for the detection of prostate cancer, as well as the evaluation of its aggressiveness [4–6].

Research studies have shown that Support Vector Machine (SVM) classifiers provide good results for classification [7–9]. Nevertheless, they have to cope with False Positives (FPs) and False Negatives (FNs) that affect the final results. From a clinical point of view FNs may lead to underestimating the cancer by detecting only portions of it, whereas FPs may lead to an extra care. Both scenarios are not acceptable and should be avoided. The reduction of FPs and FNs is therefore still a challenge that needs to be solved to effectively use Computer Aided Detection (CAD) tools or Decision Support System (DSS) for PCa. In particular, physicians look at FNs reduction (and related sensitivity increment) to avoid misprediction by detecting regions that cannot be easily seen with the naked eye.

On the computational side, these aspects are usually addressed either by implementing extra filter steps, trained to increase the specificity by reducing FPs, or by modifying the classification method. Sometimes both approaches are taken, negligently forgetting that procedures adding post-prediction filters may also decrement True Positives (TPs) or may negatively impact on future improvements of the classification step.

In this article, we propose a new methodology that relies on a 3D voxel-wise neighborhood features evaluation instead of single voxel one. We implemented the classification pipeline on top of a SVM supervised machine learning classification technique. The tool is mainly written resorting to Insight Toolkit (ITK) libraries [10] to provide a modular and cross-platform implementation of the flow. Moreover, actual implementation based ITK algorithms may take advantage of the multiple processors present in most common systems and ensure faster classification time.

In terms of classification, preliminary results show interesting classification improvements w.r.t. the single voxel classification process.

2 Materials and Methods

In this section, we first describe the available multiparametric MRI dataset, then we present the idea of the 3D-voxel neighborhoods classifier approach, together with its implementation based on ITK and LIBSVM tool [11].

2.1 Dataset

The available dataset consists of 28 patients, who underwent MRI before prostatectomy. The mean age is 64 years and they were selected among patient of the same hospital. Since personal data are confidential and removed from the MRI sequences, we have no other information about them. A pathologist contoured tumors on histological sections and a radiologist outlined regions of interest

(ROI) on the T2-w images in areas corresponding to each PCa. Non tumor ROI were also outlined in each patient to define a balanced dataset useful in the training stage. A total of 28 tumors with size bigger than 0.5 cc (median: 1,64 cc; 1st-3rd quartile: 0,75-2,25 cc) were included in the dataset. DW-MRI and DCE-MRI sequences were automatically aligned to T2-w images and finally each pixel, belonging to the prostate automatically segmented, was represented as a vector of quantitative parameters (i.e. T2-w signal intensity, the ADC value and K trans)[12, 13].

2.2 Classifier

Before going into methods and implementation details, we briefly describe the architecture of the classification tool. It consists of two different stages: a training and a testing stage. Both stages are summarized in Figure 1. Since the ITK libraries follows the Objected Oriented Programming (OOP) paradigm, it can be noticed that the architecture is fully modular. Together with the usage of the LIBSVM library as a reliable implementation for a SVM classifier [11], this choice aims at guaranteeing compatibility and flexibility towards further functionality improvements.

The very first step of both stages is a data adaptation provided by means of an ITK filter (called DataProcess in the schema of Figure 1). This step is merely required by the discrepancy between the actual dataset format and the one imposed by LIBSVM. It consists on vector permutations and aggregations; actual values are not modified during this process. If the DataProcess filter, due to further changes on the dataset format, will be no longer necessary, it will be avoided.

Once adapted, all the data are normalized and standardized as required by the SVM classification methodology [14]. The Normalization class filter implements this step. It provides both the normalized data and the normalization parameters. Normalization parameters are necessary to apply the normalization process to new data, i.e., when any further prediction is needed.

The normalized data are then forwarded to the training class, where the final SVM classifier model is trained and created. The Training class filter is also able to output a final model description that is very useful to set up the classifier whenever needed, without training the model from scratch again. Avoiding re-training time waste hugely impacts to the classifier timing performances.

When the training is completed and we need to employ it for prediction, the testing flow is very similar to the training one: data coming from the feature extraction are formatted and subsequently normalized resorting to the parameters saved during the training phase. The prediction is also performed restoring the model previously saved during the training phase. The outputs is an ITK compliant image, representing a probability map where, for each pixel of the original morphological sequence, a cancer probability value is computed.

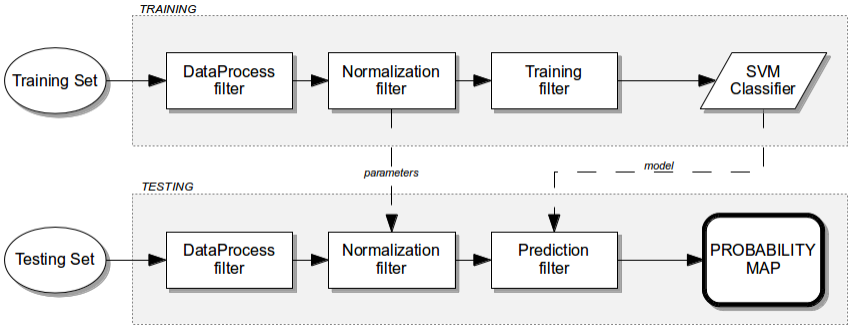

Fig. 1. The Classifier Training and Testing Flows

2.3 The 3D Voxel Neighborhood Approach

At the morphological level, tumors are clusters of voxels, spreading along three dimensions. Even on MRI sequences, all of them start as a two-dimensions artifact on one slice and then progress slice by slice until the end of their mass. Thus, it makes sense to consider at the same time not only on the set of features extracted for the voxel under-evaluation but also on the pixels surrounding it, both on the same slice and on the previous and following slices. Under these assumptions, when the classifier deals with potential FP voxels surrounded by True Negative (TN) voxels, we may increase the chance of classifying them as TN. The same way potential FN voxels within a TP voxels neighborhood likely results in a TP classification.

To formalize this idea, we first define a radius concept as the distance from a central voxel (see Figure 2). This distance can be evaluated as a voxel-wise one (R_{vx}) or using a common length metrics, i.e. millimeters (R_{mm}), resorting to the spacing information ($S_{x,y,z}$) available within the header of any MRI sequence compliant with the DICOM standard [15]. This last type of evaluation should allow better results since spacing is known to be bigger on Z-axis than on X- and Y-axis (e.g., 3 millimeters from one slice to another versus 0.3 millimeters between two adjacent voxels within a slice). The radius identifies a 3D box where all belonging voxels can be processed together.

Assuming that for each axial direction (x, y, z) we can relate a voxel-wise radius with common length metrics radius (millimeters in the equation) as:

$$R_{vx} = \left\lceil \frac{R_{mm}}{S_{(x,y,z)}} \right\rceil \tag{1}$$

we are able to express the number of voxel selected with respect to each direction as:

$$\begin{cases} NumVox_x = ((2R_{vx}) + 1) \\ NumVox_y = ((2R_{vx}) + 1) \\ NumVox_z = ((2R_{vx}) + 1) \end{cases} \tag{2}$$

Eventually, equation 2 allows to evaluate the total number of voxels processed in the 3D neighborhood as:

$$\begin{aligned} NumVox_{tot} &= (NumVox_x NumVox_y NumVox_z) \\ &= ((2R_{vx}) + 1)^3 \end{aligned} \tag{3}$$

Figure 2 shows quantitatively how many voxel will be considered when the radius is set to 1 voxel. Both the Training and the Prediction class filters are able to process 3D voxel-wise neighborhoods as well as single voxels.

Once a radius is defined, each feature belonging to the voxels included in the 3D box is evaluated by averaging its value among all single voxel values. This averaged value is employed as final value of the feature. Currently, studies on different evaluation approaches are under analysis.

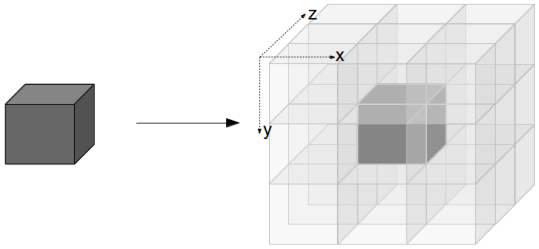

Fig. 2. 3D Voxel Neighborhood with a $R_{vx} = 1$. A total of 27 voxels are then considered.

It has to be emphasized in here that, during the developed supervised training, where the dataset is originated by user-selected malignant and benignant ROI, the Training class filter averages the feature value only of voxels of the same ROI target type. This strategy avoids the training class filter to break the supervised training rules. Moreover, this way, volumetric information could be weighted in the malignant voxels set described by $NumVox_{tot}$ neighborhood voxels; thereby FP isolated voxels could be filtered and the borders of malignant ROI may result more accurate. In line with these assumption physicians consider and highlight tumors with a minimum volume size that have diagnostic relevance (i.e. 0,5cc). It is also important to mention that features are collected voxel by voxel. Any other form of feature extraction introduces losses in the original data.

Experiments were performed with different radius values in order to investigate the classifier performances against single voxel-wise classification and will

be provided on Section 3. In technical terms, the neighborhood selection is implemented making use of *itkConstNeighborhoodIterator* and *itkImageRegionConstIterator* ITK class templates.

3 Results

In this section we present the results for 3D voxel neighborhood classifier and we discuss about obtained performances.

3.1 Data Evaluation

The training set was built upon 28 patients and a leave-one-out cross-validation (LOOCV) has been implemented to validate the model. We performed different experiments varying the radius (R_{vx}): six different radius (R_i) size were tested: $R_{i=\{0,1,2,3,4,5\}}$, where $R_{i=0}$ means a standard single voxel-wise classification. Figure 3 compares results by means of the following statistical functions:

- Area Under the ROC Curve (AUC) [16]
- Sensitivity (Se): $Se = \frac{TP}{TP+FN}$
- Specificity (Sp): $Sp = \frac{TN}{TN+FP}$
- Accuracy (Acc): $Acc = \frac{TP+TN}{TP+TN+FP+FN}$

In particular, we present arithmetic mean (AM) and standard deviation (SD) to compare classifier prediction performances.

Since the final decision has to be binary (tumor vs healthy tissue), a thresholds cutoff was set to 0.5 value during the experimental setup. This is a very weak threshold, which commonly leads to worst results comparing with best cutoff search algorithms, such as the Yoden one [17, 18]. The meaning of our choice is strongly related to the way physicians exploit the classification results. Within a diagnosis path, the CAD software is usually able to show the physician the classification results as a colored overlay map on the T2-w sequence [19]. Colors help focusing on the tumor areas but their relation to the classifier outcome is changed as physician commodity.

3.2 Discussion

Figure 3 reveals that our 3D voxel-wise neighborhood classifier provide an improvement of the classification performances. We highlight the AM improvement in term of AUC, Se, Sp and Acc. In particular, when comparing R_0 to R_5, the results seem very promising.

Generally speaking, the progressive increase of the radius relates to a continuous improvement in results. The Se reveals a significant improvement ranging its average from 0.71 to 0.78 (R_0 to R_5), as well as the improvement of the AUC value is significant when single voxel-wise classifier performances are compared with R_5 neighborhood classifier. In this case the average rises from 0.76 to 0.81.

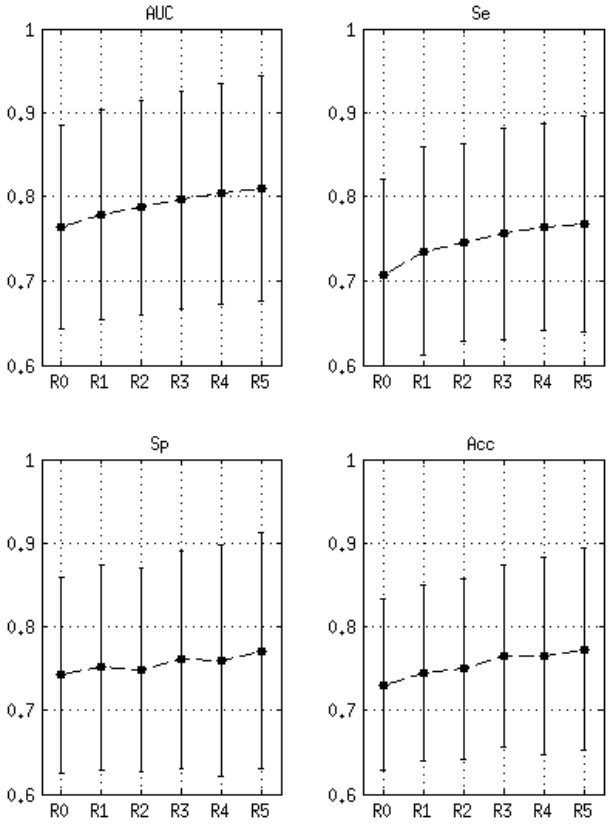

Fig. 3. AUC, Se, Sp and Acc average (dot) and standard deviation (whiskers) values, with relation to radius size (R_i), for testing set.

Nevertheless, the Sp seems to do not benefit from our proposed technique (R_0 to R_5 delta is less than 0.03), suggesting that the way features on different voxel are considered may need further investigation to refine the methodology. In particular, the AM variation expressed along the radius, may suggest that more considerations on the morphological characteristic of tumors among the dataset are needed. We expected such kind of impact on final results thus further investigations will be planned on that.

Eventually, SD values are generally not negligible but this may depend on the actual available dataset size.

4 Conclusion

We present a 3D voxel neighborhood SVM classifier methodology, implemented using ITK libraries, based on MRI sequences to discriminate prostate cancer lesion from healthy tissue.

Obtained results indicate improvements if compared against traditional single voxel-wise classifier; especially FNs take advantage from the proposed approach. Some minor drawbacks suggest a further analysis involving an extended dataset to confirm the validity volumetric neighborhood approach performance.

References

1. European Cancer Observatory (ECO), http://eu-cancer.iarc.fr/
2. Hegde, J.V., Mulkern, R.V., Panych, L.P., Fennessy, F.M., Fedorov, A., Maier, S.E., Tempany, C.M.C.: Multiparametric MRI of Prostate Cancer: An Update on State-Of-The-Art Techniques and their Performance in Detecting and Localizing Prostate Cancer. Journal of Magnetic Resonance Imaging 37, 1035-1054 (2013)
3. Ongun, S., Celik, S., Gul-Niflioglu, G., Aslan, G., Tuna, B., Mungan, U., Uner, S., Yorukoglu, K.: Are Active Surveillance Criteria Sufficient for Predicting Advanced Wtage Prostate Cancer Patients? Actas Urologicas Espanolas 38, 499-505 (2014)
4. Tan, C.H., Wang, J.H., Kundra, V.: Diffusion Weighted Imaging in Prostate Cancer. European Radiology 21, 593-603 (2011)
5. Desouza, N.M., Riches, S.F., VanAs, N.J., Morgan, V.A., Ashley, S.A., Fisher, C., Payne, G.S., Parker, C.: Diffusion-Weighted Magnetic Resonance Imaging: a Potential Non-Invasive Marker of Tumor Aggressiveness in Localized Prostate Cancer. Clinical Radiology 63, 774-782 (2008)
6. Turkbey, B., Bernardo, M., Merino, M.J., Wood, B.J., Pinto, P.A., Choyke, P.L.: MRI of Localized Prostate Cancer: Coming of Age in the PSA Era. Diagnostic and Interventional Radiology 18, 34-45 (2012)
7. Artan, Y., Haider, M.A., Langer, D.L., van der Kwast, T.H., Evans, A.J., Yang, Y.Y., Wernick, M.N., Trachtenberg, J., Yetik, I.S.: Prostate Cancer Localization With Multispectral MRI Using Cost-Sensitive Support Vector Machines and Conditional Random Fields. Ieee Transactions on Image Processing 19, 2444-2455 (2010)
8. Zhou, T., Lu, H.L.: Multi-Features Prostate Tumor Aided Diagnoses Based on Ensemble-SVM. 2013 Ieee International Conference on Granular Computing (Grc) 297-302 (2013)
9. Shah, V., Turkbey, B., Mani, H., Pang, Y.X., Pohida, T., Merino, M.J., Pinto, P.A., Choyke, P.L., Bernardo, M.: Decision Support System for Localizing Prostate Cancer Based on Multiparametric Magnetic Resonance Imaging. Medical Physics 39, 4093-4103 (2012)
10. The Insight Segmentation and Registration Toolkit, http://www.itk.org
11. Chang, C.C., Lin, C.J.: LIBSVM: A Library for Support Vector Machines. Acm Transactions on Intelligent Systems and Technology 2, 27 (2011), Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm
12. Peng, Y.H., Jiang, Y.L., Yang, C., Brown, J.B., Antic, T., Sethi, I., Schmid-Tannwald, C., Giger, M.L., Eggener, S.E., Oto, A.: Quantitative Analysis of Multiparametric Prostate MR Images: Differentiation between Prostate Cancer and Normal Tissue and Correlation with Gleason Score-A Computer-aided Diagnosis Development Study. Radiology 267, 787-796 (2013)

13. Tamada, T., Sone, T., Jo, Y., Yamamoto, A., Ito, K.: Diffusion-Weighted MRI and its Role in Prostate Cancer. Nmr in Biomedicine 27, 25-38 (2014)
14. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification, ISBN:0-471-05669-3. Wiley, New York (2001)
15. Pianykh, O.S.: Digital Imaging and Communications in Medicine (DICOM), ISBN:978-3-642-10849-5. Springer-Verlag Berlin Heidelberg (2012)
16. Hanley, J.A., McNeil, B.J.: The Meaning and Use of the Area Under a Receiver Operating Characteristic (ROC) Curve. Radiology 143, 29-36 (1982)
17. Martnez-Camblor, Pablo. Nonparametric Cutoff Point Estimation for Diagnostic Decisions with Weighted Errors, Revista Colombiana de Estadstica, 34(1), 133-146 (2011)
18. Fluss, Ronen; Faraggi, David; FAU - Reiser, Benjamin, Estimation of the Youden Index and its associated cutoff point., Biometrical journal. Biometrische Zeitschrift, Aug;47(4):458-72. (2005)
19. Savino A., Benso A., Di Carlo S., Giannini V., Vignati A., Politano G., Mazzetti S., Regge D., A Prostate Cancer Computer Aided Diagnosis Software including Malignancy Tumor Probabilistic Classification. In: International Conference on Bioimaging (BIOIMAGING), Eseo, Angers, FR, 3-6 March. pp. 49-54 (2014)