Computer-aided classification of the clinical significance of the prostate lesions in MRI images*

*Note: Sub-titles are not captured in Xplore and should not be used

1st Yesid Gutiérrez
dept. name of organization (of Aff.)
Universidad Industrial de Santander
Bucaramanga, Colombia
yesid.gutierrez@correo.uis.edu.co

2nd Fabio Martínez
dept. name of organization (of Aff.)
Universidad Industrial de Santander
Bucaramanga, Colombia
famarcar@saber.uis.edu.co

***Abstract*—This document is a model and instructions for LaTeX. This and the IEEEtran.cls file define the components of your paper [title, text, heads, etc.]. *CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

This document is a model and instructions for LaTeX. Please observe the conference page limits.

II. Proposed Method

A. MRI Data

The data used in this work was obtained from The Cancer Imaging Archive (TCIA) sponsored by the SPIE, NCI/NIH, AAPM, and Radboud University. The data-set contains 204 MRI cases for training and 142 MRI cases for testing, each case consists of four sets of MRI scan data: two sets of T2-weighted images (trans-axial and sagittal; DICOM format), Ktrans images (computed from dynamic contrast-enhanced (DCE) images; mhd format), and apparent diffusion coefficient (ADC) images (computed from diffusion-weighted (DWI) imaging; DICOM format)[1].

The data-set contains at least one prostate lesion with biopsy-proven malignancy status or with imaging findings with sufficiently low suspicion of clinical significance per patient and the position (i,j,k) of the lesion. Despite of the data-set provided, this work was done using only the subset of Ktrans images[1].

B. Ktrans Images

Ktrans is a key pharmacokinetic parameter computed from the available Dynamic contrast-enhanced T1-weighted series [1], Ktrans depends on three main factors plasma blood flow (F), vascular permeability (P), and capillary surface area (S) per unit mass.[2]

Ktrans images come in MHD format with a shape of (16 slices, 128 rows and 128 columns), only 1 Ktrans image is provided per finding.

Identify applicable funding agency here. If none, delete this.

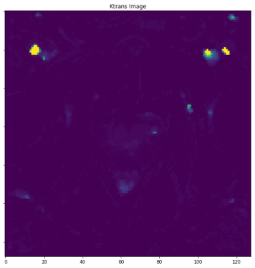

Fig. 1. Krans image of the Patient ProstateX-0143 at the slice 10.

C. Ktrans Images normalization

Ktrans images intensities may differ due to the lack of a standard image intensity scale in MRI [3]. Due this, a linear image normalization was proposed. [4]

$$I_N = (I - min) * \frac{newMax - newMin}{max - min} + newMin \quad (1)$$

where "I" represents the intensity of the image at the x,y,z position, "max" represents the maximum value of intensity at the slice z, "min" represents the minimum value of intensity at the slice z [4]. The "newMax" and newMin value were selected as 1 and -1 respectly.

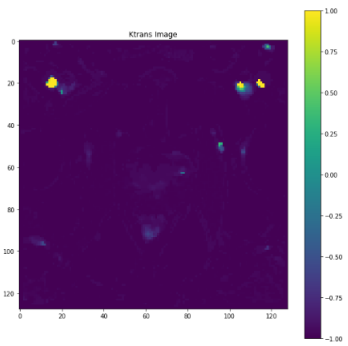

Fig. 2. Krans image normalization at the slice 10.

D. 3D patch extraction

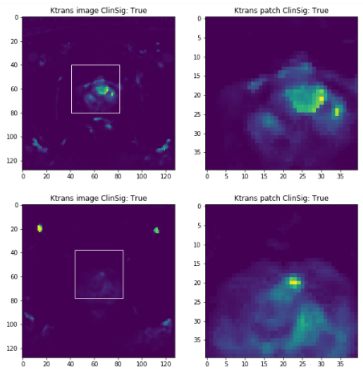

Fig. 3. Sample of a patch extraction in a Ktrans image with clinical significance.

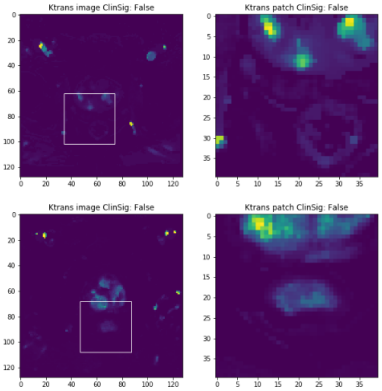

Fig. 4. Sample of a patch extraction in a Ktrans image without clinical significance.

A 3D patch extraction process was performed in order to avoid models learning the shape of the different zones of the prostate instead of the features of the lesion of study itself. This process was made with a window of 40 rows, 40 columns and 3 slices centrated at the ijk position provided at the data-set.

E. Clinical significance prediction of the prostate lesion

The main goal of this work is to predict the value of the clinical significance to a given prostate lesion, Clinical significance's value has two possibles status (non-malignancy status and malignancy status) that were represented as 0 and 1 respectively.

The training data-set was divided into 3 sub-data-sets, (peripheral zone data-set, transitional zone data-set and anterior fibromuscular stroma). A final experiment with all the zones of the data-set is also performed, additionally 4 different machine learning models were selected for the prediction task, these models were a random forest classifier with 20 trees [5], a K nearest neighbor classifier with k = 20 [5] a Support vector machine classifier with optimized hyper-parameters [5] and a Support Vector Machine classifier from LIBSVM [6]. Finally, a K-fold cross validation with 8 folds of data was performed with each one of the previous models and data-sets described.

III. Results

The results were summarized in the following 5 tables, each table contains the results of the 4 machine learning models (LIBSVM SVM, Random Forest, K nearest neighbor and SVM with the hyper-parameters provided by LIBSVM.) using a k-fold cross validation at different prostate zones.

TABLE I
LIBSVM RESULTS

Prostate Zone	***Scores's mean***
PZ	82.5
TZ	89.2857
AS	69.0909
ALL	80.236

TABLE II
PERIPHERAL ZONE RESULTS TABLE

Classifiers	**Metrics**	
	Scores mean	***Scores std***
Random Forest 20	77.512821	3.878353
K Nearest Neighbors 20	82.038462	1.494401
SVM optimized PZ	81.557692	2.252026

TABLE III
TRANSITIONAL ZONE RESULTS TABLE

Classifiers	**Metrics**	
	Scores's mean	***Scores's std***
Random Forest 20	89.507576	5.126807
K Nearest Neighbors 20	89.393939	2.322683
SVM optimized PZ	89.393939	2.322683

TABLE IV
ANTERIOR FIBROMUSCULAR STROMA RESULTS TABLE

Classifiers	**Metrics**	
	Scores's mean	***Scores's std***
Random Forest 20	56.547619	15.510493
K Nearest Neighbors 20	63.392857	16.535946
SVM optimized PZ	63.690476	21.287565

References

[1] Litjens G, Debats O, Barentsz J, Karssemeijer N, Huisman H. Computer-aided detection of prostate cancer in MRI. IEEE Transactions on Medical Imaging, (33)1083-92, 2014.
[2] Cuenod CA, Balvay D. Perfusion and vascular permeability: basic concepts and measurement in DCE-CT and DCE-MRI. Diagn Interv Imaging 2013; 94:1187-1204.
[3] Nyl, L. G., Udupa, J. K. (1999). On standardizing the MR image intensity scale. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 42(6), 1072-1081.
[4] Rafael C. Gonzlez, Richard Eugene Woods (2007). Digital Image Processing. Prentice Hall. p. 85. ISBN 0-13-168728-X.
[5] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

TABLE V
ALL ZONES RESULTS TABLE

Classifiers	Metrics	
	Scores's mean	***Scores's std***
Random Forest 20	75.880871	5.201011
K Nearest Neighbors 20	77.887057	2.201882
SVM optimized PZ	78.780589	3.170147

[6] Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1–27:27, 2011.

[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove the template text from your paper may result in your paper not being published.