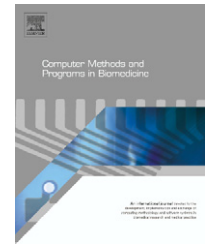




ELSEVIER

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

A multi-classifier system for the characterization of normal, infectious, and cancerous prostate tissues employing transrectal ultrasound images

Dimitris Glotsos^{a,*}, Ioannis Kalatzis^a, Pantelis Theocharakis^b, Pantelis Georgiadis^b, Antonis Daskalakis^b, Kostas Ninos^a, Pavlos Zoumboulis^c, Anna Filippidou^c, Dionisis Cavouras^a

^a Department of Medical Instruments Technology, Technological Educational Institute of Athens, Ag. Spyridonos, Aigaleo, Athens 12210, Greece

^b Medical Image Processing and Analysis Laboratory, Department of Medical Physics, University of Patras, Rio-Patras 26500, Greece

^c Ultrasound Department, Echonet, 319 Kifissias Avenue, Kifissia, Athens 14561, Greece

ARTICLE INFO

Article history:

Received 1 September 2008

Received in revised form

30 June 2009

Accepted 6 July 2009

Keywords:

Prostate

Transrectal ultrasound

Textural features

Pattern classification

Multi-classifier systems

ABSTRACT

A computer-aided diagnostic system has been developed for the discrimination of normal, infectious and cancer prostate tissues based on texture analysis of transrectal ultrasound images. The proposed system has been designed using a panel of three classifiers, which have been evaluated individually or as a multi-classifier scheme, using the external cross-validation procedure. Clinical data consisted of 165 transrectal ultrasound images, characterized by an experienced physician as normal (55/165), cancerous (55/165), and infectious (55/165) prostate cases. From each image, the physician delineated the most representative regions of interest, from which, 23 textural features were extracted. Classification was seen as a two level hierarchical decision tree. Normal from infectious and infectious from cancer cases were discriminated at the 1st and 2nd level of the decision tree, respectively. The best classification results for the 1st level were 89.5%, whereas for the 2nd level 90.1%. The utilization of multi-classifier system improved the discrimination of prostate pathologies as compared to individual classifiers; for infectious prostate cases improvement was from 87.3% to 88.7% and for cancer prostate cases improvement was from 84.1% to 91.4%. In terms of overall system performance (the decision tree's node propagating error taken into account), best classification accuracies were 89.5%, 79.6% and 82.7% for the recognition of normal, infectious and cancer cases, respectively. The proposed system might be used as a second opinion tool for assisting diagnosis of different prostate pathologies.

© 2009 Elsevier Ireland Ltd. All rights reserved.

* Corresponding author at: Medical Image and Signal Processing (medisp) Lab., Department of Medical Instruments Technology, Technological Educational Institute of Athens, Ag. Spyridonos Street, Egaleo 12210, Greece. Tel.: +30 210 5385 375.

E-mail addresses: dimglo@teiath.gr, dimglo@yahoo.com (D. Glotsos).

URLs: <http://www.teiath.gr/stef/tio/medisp/index.htm> (D. Glotsos), <http://medisp.bme.teiath.gr/> (D. Glotsos).

0169-2607/\$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.

doi:10.1016/j.cmpb.2009.07.003

1. Introduction

The three most common and significant pathologies of the prostate gland are infection (which refers to prostate inflammation), hyperplasia (that is, abnormal growth of benign prostate cells) and cancer. Standard diagnostic tests for discriminating these pathologies are digital rectal examination and transrectal ultrasound imaging (TRUS) [1,2]. Other diagnostic tests comprise power Doppler imaging [3] and combined ultrasonic and magnetic-resonance methods [4]. TRUS imaging is performed for assessing the size of the prostate gland and assisting the visual identification of tumors [5,6]. Although TRUS is among the most important tools for diagnosing prostate pathologies, its accuracy is limited by means of poor inter- and intra-observer reproducibility of the observing physicians [7,8]. To ameliorate for this effect, several computer-aided diagnostic (CAD) systems have been proposed as second opinion tools [9–16]. CAD systems comprise two important stages: (i) the feature extraction stage, in which the morphological and textural patterns of the prostate gland in ultrasound images are encoded in the form of numerical features (patterns) and (ii) the classification stage, in which these patterns are interpreted into diagnostic conclusions, using pattern classification methods.

In general, all the developers of proposed CAD systems agree upon the characteristics (morphological and textural) that seem to be important for the evaluation of prostate pathologies in ultrasound images. However, the existence of multiple classification algorithms implies that there is a lack of consensus among experts as to which is the best single classification approach. Additionally, the reported specificities, sensitivities and overall accuracies in discriminating normal from cancerous prostate tissues vary significantly from 78% to 85.7%. The latter variation might be possibly due to the utilization of different (a) classification methodologies [10,12,15,17], (b) diagnostic features, (c) ROI (region of interest) selection, (d) pixel resolution sizes and (e) ultrasound imaging equipment. An important aspect that remains problematic in CAD systems is that the evaluation of the classifiers' performance has been applied mostly internally to the feature selection process [12,18]. The latter has been identified as generating optimistic estimates of the classifier's performance [19–22]. Thus, features selected using internal validation processes (like cross-validation) are most likely to generalize poorly to unseen data. The latter effect has not been investigated by previous studies making difficult the extraction of conclusions regarding generalization. Moreover, there are few studies that have used wrapper feature selection methods [13] (i.e. features are selected independently of the classifier construction stage [19–22]), but these studies, have reported relatively low accuracies (area under the receiver operating characteristic (ROC) curve 61.6%). Moreover, most of previous studies have explored the discrimination of normal from cancer prostate cases. However, there are other important subcategories deserving special attention, such as hypertrophy and infectious prostates. The integration of hypertrophy in a CAD system has been examined by [15] with accuracy 78%. However, infectious cases, possibly representing early stages of cancer development [23,24],

have not been yet considered in the construction of CAD systems.

In this study a CAD system has been developed for discrimination of normal, infectious and cancer prostate tissues based on texture analysis of transrectal ultrasound images. The proposed CAD system presents three unique characteristics compared to other studies, attempting to overcome limitations in existing literature, mentioned in previous paragraph: (a) in contrast to previous studies that have used internal validation procedures, an external cross-validation procedure has been used in this study for evaluating the performance of the CAD system. In this way, the generalization capacity of the system to unseen data could be assessed [19–22]. (b) The CAD system has been designed using a multi-classifier system in contrast to previous studies that have used single classifiers. In this way, we have tried to investigate whether suitable combinations of classifiers might improve success rates in predicting different prostate pathologies. (c) Finally, to the best of our knowledge, the proposed CAD system comprises the first attempt to examine and to put emphasis in discriminating infectious from normal and cancerous cases. The latter discrimination is of major significance since infectious cases have been reported as possible early stage symptoms of emerging prostate cancer [23,24].

2. Methods

2.1. Region of interest (ROI) acquisition

A total number of 165 TRUS prostate images from 165 patients were obtained by an experienced physician (P.Z.), using an HDI-3000 ATL digital ultrasound system with a wide band (5–12 MHz) probe. The image dataset comprised 55 Prostate Normal cases (PRNO) with no pathological findings, 55 prostate infection cases (PRIN) and 55 prostate cancer cases (PRCA). All pathological cases were subjected to ultrasound guided biopsy by an experienced pathologist (A.F.) and were histological confirmed. A video grabber was used for image digitization at $320 \times 256 \times 8$ bit. In each image, the physician, with the help of a custom made program (Fig. 1), chose a 30×30 pixel size region of interest, which included part of the corresponding region of interest (ROI) for the PRIN, PRCA and PRNO. Considering that our system's resolution was 16.66 pixels/mm, each ROI's area was 3.24 mm^2 .

2.2. Feature extraction

A series of features were computed from each ROI. Four features were computed from the ROI's histogram (mean value, standard deviation, skewness, and kurtosis), 14 from the co-occurrence matrices [25] and 5 from the run-length matrices [26]. In this way, 23 features represented each case (patient) since feature extraction was based on 30×30 ROI extracted by the participating radiologist from most representative region of the patient's image. A more detailed description of how these features are computed can be found in Appendix A. All 23 features were normalized to zero mean and unit standard

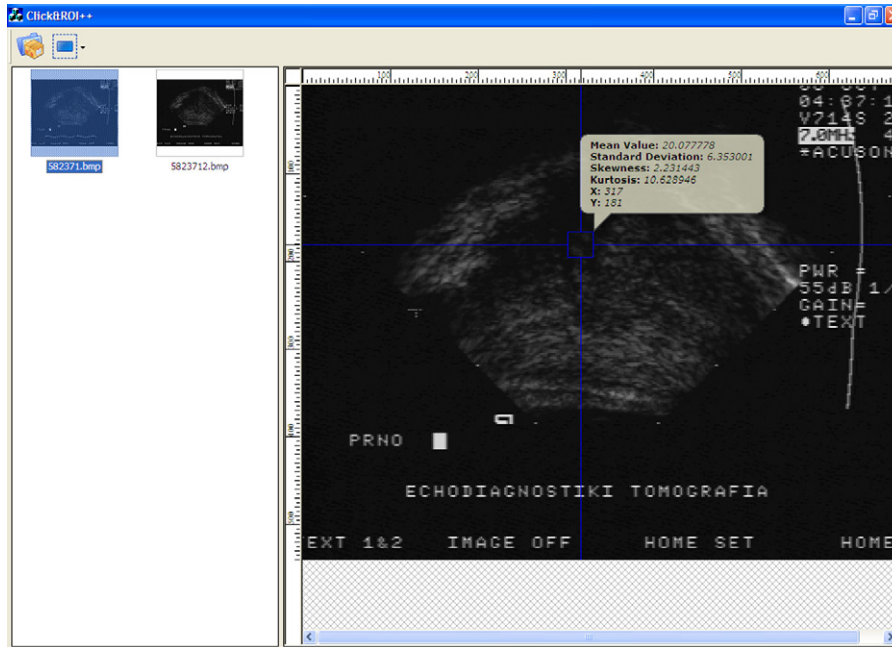


Fig. 1 – Custom made program for ROI acquisition.

deviation [27], according to the following equation

$$\tilde{x}_i = \frac{x_i - \mu}{\sigma} \quad (1)$$

where x_i is the i th feature, \tilde{x}_i the normalized i th feature, μ the mean value and σ is the standard deviation of the i th feature of all patterns in all classes considered.

2.3. Design of the classification scheme

A two level hierarchical decision tree was designed to discriminate the normal prostate cases from the infectious and cancerous cases. At the 1st level, the cancer and the prostatitis cases were grouped into the prostate abnormal (PRAB) class and were classified against the normal prostate (PRNO) cases. At the 2nd level, the abnormal cases were further classified into cases with cancer and infection. The selection of the two level hierarchical decision tree structure was motivated by the fact that this structure resembles the diagnostic procedure followed by expert physicians in diagnosis of prostate cancer based on TRUS images.

At each level, classification was performed using three different classifiers either individually or combined in a multi-classifier (MC) system. Classifiers employed were the Cubic Least Square Mapping Probabilistic Neural Network (LSM³PNN) [28,29], the Quadratic Bayesian (BAYES) [27] and the Support Vector Machines (SVM) employing a radial-basis function kernel [30]. Many classifier evaluation methods exist in literature, such as the leave-one-out, the k -fold [27], or the external cross-validation [19–22]. The selection of the validation process is an important task; there are few validation processes, which allow the extraction of conclusions regarding the generalization of the methods used. In two previous studies [19,31], it has been demonstrated that cross-validation

samples should be kept external to the feature selection process. Moreover in [19], researchers have also assessed several techniques for estimating the generalization error. They showed that the external 10-fold cross-validation error is one of the most unbiased estimators of the generalization error. Hence, we have used the external cross-validation. Accordingly, data were randomly split into two main subsets: a training subset comprising 2/3 of available samples (training data) and a test subset consisting of the remaining 1/3 of available samples (test data). Training data were used for feature selection and training of each classifier using the exhaustive search (combinations up to five features) and the leave-one-out methods [32]. Following feature selection, the performance of each classifier and the MC-system (described in the following paragraph) was then evaluated on the remaining test data. This process was repeated 10 times for 10 different random splits of all available samples into training and test data. The average accuracies and ranges were computed. In this way the external cross-validation estimate of the classifiers' performance was determined. Thus, with the external cross-validation process, at each stage of the cross-validation process, the same feature selection method is implemented only to the training data, external to the test subset, correcting in this way for the feature selection bias. Fig. 2 illustrates the design of the system using individual classifiers and internal validation method (exhaustive search and leave-one-out). The construction of the system using external validation (10-fold external cross-validation) and individual classifiers is depicted in Fig. 3. Finally, Fig. 4 demonstrates the structure of the MC-system under external validation.

The MC-system was constructed using as ensemble combination rule the majority vote (MV) [33,34]. The MV MC-system output was formulated according to the decision of the majority of classifiers (in our case, at least more than half of the

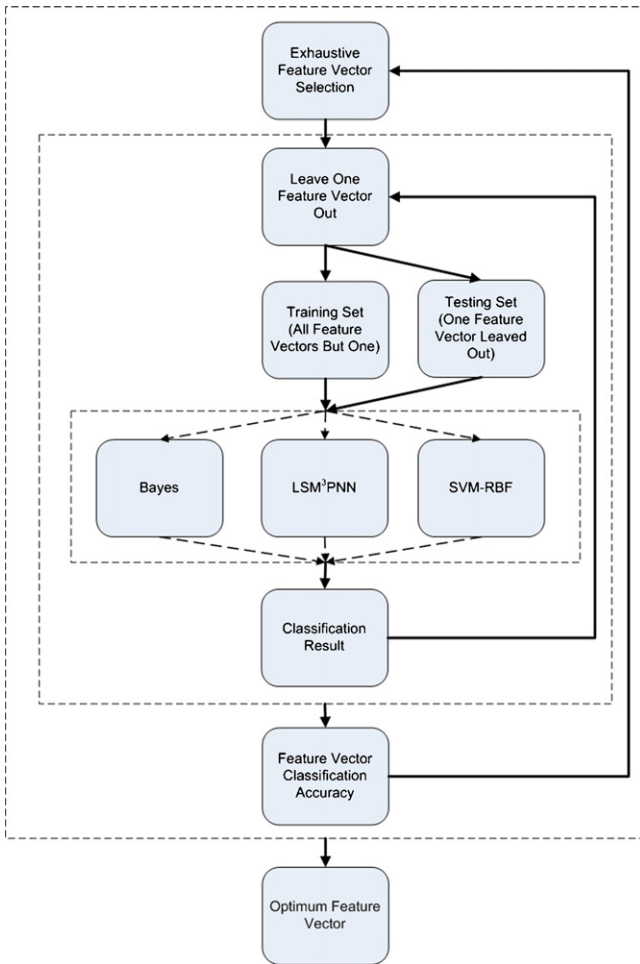


Fig. 2 – Design of the system using individual classifiers and internal validation method (exhaustive search and leave-one-out).

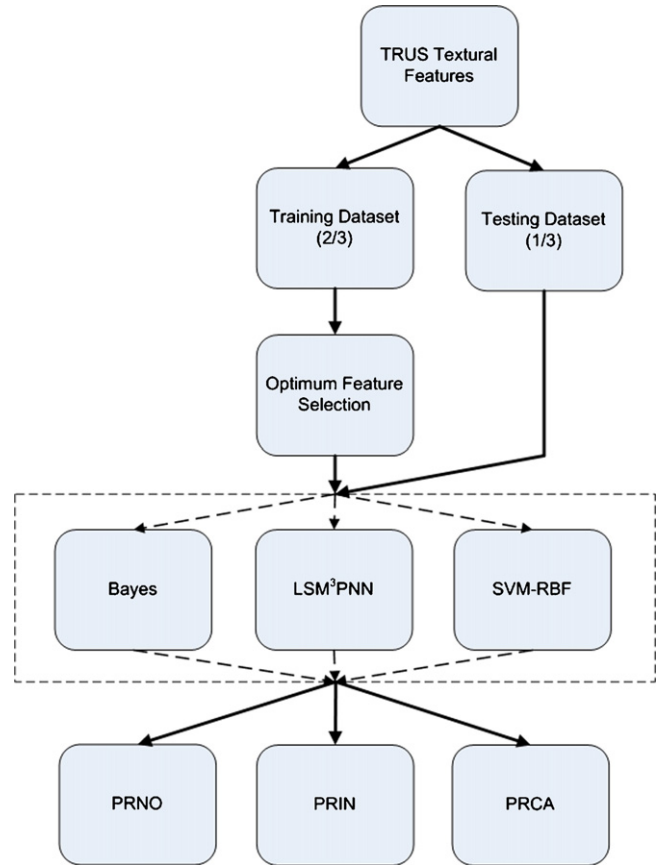


Fig. 3 – The construction of the system using external validation (10-fold external cross-validation) and individual classifiers.

Table 1 – Performance of individual classifiers using internal validation (internal leave-one-out) against external validation (10-fold external cross-validation).

	Internal validation			External validation		
	BAYES	LSM ³ PNN	SVM-RBF	BAYES	LSM ³ PNN	SVM-RBF
1st level						
PRNO	96.4	94.5	76.4	94.7	89.9	85.0
PRAB	87.3	91.8	90.0	85.3	89.2	90.9
Overall	90.3	92.7	85.5	88.6	89.5	88.8
2nd level						
PRIN	89.1	94.5	92.7	90.9	87.3	86.6
PRCA	98.2	90.9	80.0	79.7	84.1	83.7
Overall	93.6	92.7	86.4	85.4	85.9	85.2
Overall system performance						
PRNO	90.3	92.7	85.5	88.6	89.5	88.8
PRIN	77.8	86.8	83.4	77.5	77.9	78.7
PRCA	85.7	83.4	72.0	68.0	75.0	76.1

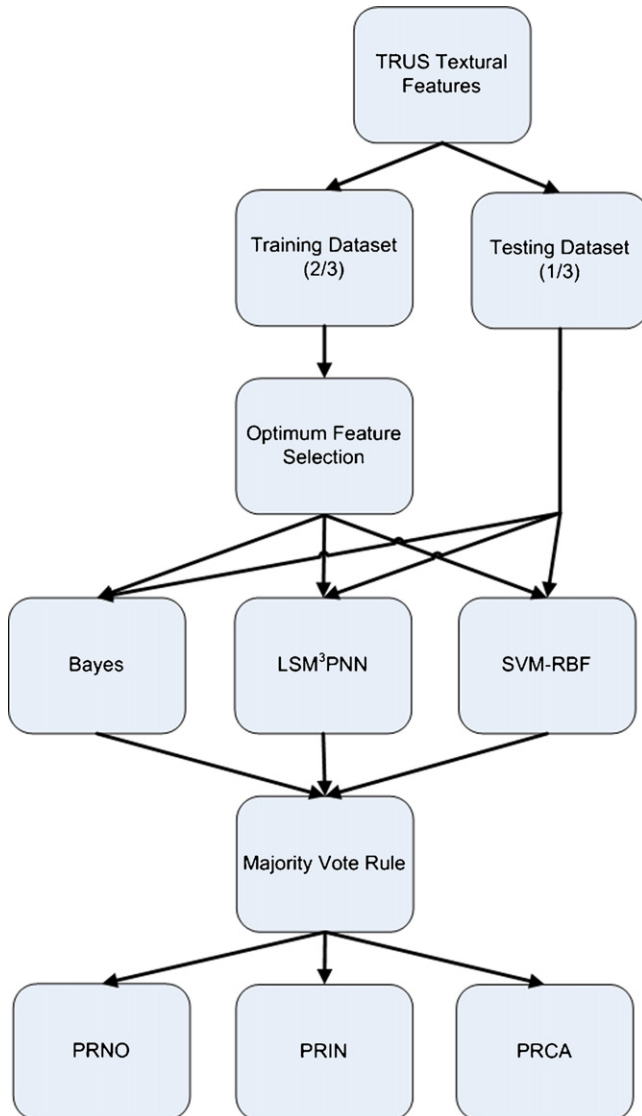


Fig. 4 – The structure of the MC-system under external validation.

classifiers should agree on that decision):

$$C_{MV} = \arg \max_k \sum_{i=1}^r d_{ki} \quad (2)$$

where C_{MV} is the output class of the MV MC-system, $d_{ki} \in \{0, 1\}$ is a binary decision value of k th class for the i th classifier (1 corresponds to classification to class k and 0 to classification in any other class) and r is the number of classifiers.

All algorithms have been developed in custom MATLAB® code.

3. Results

Table 1 depicts the performance of individual classifiers using internal validation (internal leave-one-out) against external validation (10-fold external cross-validation) for both levels

Table 2 – 1st level of the decision tree: sensitivity, specificity and overall accuracy of the MC-system (SVM, BAYES and LSM³PNN) in discriminating PRNO from PRAB cases.

Average success rates (%)	
PRNO	86.7
PRAB	89.8
Overall accuracy	88.7

Table 3 – 2nd level of the decision tree: sensitivity, specificity and overall accuracy of the MC-system (SVM, BAYES and LSM³PNN) in discriminating PRIN from PRCA cases.

Average success rates (%)	
PRCA	91.4
PRIN	88.7
Overall accuracy	90.1

Table 4 – Performance of individual classifiers using 10-fold external cross-validation against the MC-system for both levels of the decision tree. Bold numbers indicate highest accuracies achieved.

	BAYES	LSM³PNN	SVM-RBF	MC-system
Overall system performance				
PRNO	88.6	89.5	88.8	88.7
PRIN	77.5	77.9	78.7	79.6
PRCA	68.0	75.0	76.1	82.7

of the decision tree. The overall system accuracy for each classifier (last three rows of Table 1) was determined by multiplication of the each classifier's performance at each level [27]. According to the results obtained from internal validation, best classifier for the 1st level of the decision tree was the LSM³PNN with overall accuracy 92.7%, whereas for the 2nd level the Bayesian attained the maximum accuracy with 93.6%. These high classification rates reduced notably when external validation was applied. For both the 1st and 2nd levels of the decision tree the LSM³PNN resulted to the highest accuracies, with 89.5% and 85.9%, respectively. In terms of overall system performance (the decision tree's node propagating error taken into account), the proposed methodology led to the more accurate results for the discrimination of PRNO, PRIN and PRCA cases using the LSM³PNN with 89.5%, 77.9% and 75.0%, respectively.

Tables 2 and 3 presents average classification accuracies attained using an ensemble classifier scheme comprising the SVM, BAYES and LSM³PNN classifiers for both levels of the decision tree. The MC-system's sensitivity, specificity and overall accuracy in discriminating normal from abnormal prostate cases were 86.7%, 89.8% and 88.7%, respectively. For the 2nd level, sensitivity, specificity and overall accuracy reached 91.4%, 88.7% and 90.1%, respectively. Table 4 illustrates comparative results for the individual classifiers versus the MC-system in terms of overall system performance (the decision tree's node propagating error taken into account). The latter results were obtained using a 10-fold external cross-validation process. Fig. 5 illustrates the best design of the CAD system, for both levels of the decision tree, according to the 10-

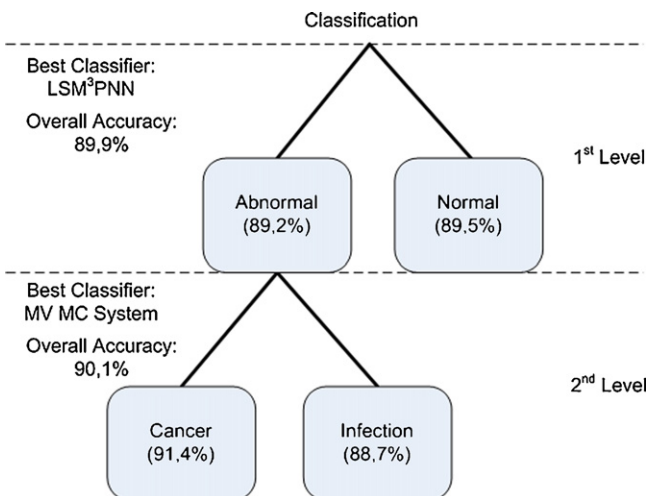


Fig. 5 – Best structure of the CAD system. According to the 10-fold external cross-validation results, optimum performance at the 1st level of the decision tree has been obtained using the LSM³PNN. For the 2nd level, the MC-system gave the highest results.

fold external cross-validation results. Finally, Table 5 presents the external cross-validation performance (in terms of classification accuracy and area under Receiver Operation Curve (ROC)) of individual classifiers using the sequential forward floating selection method at the feature selection stage. All classifiers present high values, from 0.941 to 0.947 for the 1st level and from 0.956 to 0.971 for the 2nd level of the classification tree.

4. Discussion

There has been intensive research concerning pattern recognition methods for automating prostate diagnosis on ultrasound images, due to the poor depiction of different prostate tissues in these images. Excellent literature surveys concerning computer-aided detection and classification of prostate tissues may be found in [7,18]. Basset et al. [15] has developed a rule based expert system that discriminated normal, benign

Table 5 – External validation performance (in terms of classification accuracy and area under ROC curve) of individual classifiers using the sequential forward floating selection method at the feature selection stage.

	BAYES	LSM ³ PNN	SVM-RBF
1st level			
PRNO	93.0	86.5	77.3
PRAB	79.8	83.1	90.8
Overall	84.5	84.3	86.1
AUROC	0.947	0.941	0.946
2nd level			
PRIN	95.2	88.4	91.2
PRCA	87.8	86.5	87.5
Overall	91.2	87.4	89.2
AUROC	0.971	0.956	0.969

hypertrophy, and cancer prostate cases with 78% accuracy. Lorenz et al. [12] have utilized pattern recognition methods with neuro-fuzzy classifiers for the separation of benign from malignant prostate tissues with success rates up to 86%. Huynen et al. [9] explored classification trees as possible tools for identifying normal and malignant prostate tissues with an accuracy of 85.7%. Moreover, in a recent study, hidden markov models and k-NN classification have been compared for discrimination of prostate tissues into cancerous and non-cancerous, with results, in terms of area under the receiver operating characteristic curve, up to 61.4% [14]. Direct comparison of the results presented in our study with those reported in literature is very difficult to perform due to the different experimental set up of each study (i.e. number of cases, ultrasound units, image resolution, selected features, ROIs sizes [5,9,10,12,15,16]). Reported accuracies range from 78% to 85.7%. There are also some studies that have reported accuracies up to 93.7% [13], using, however, extremely small datasets (16 ROIs in total), subject to important statistical variations. Results presented in this study are in line with those reported in literature, however, with a significant difference: Results have been obtained using an external cross-validation process, in contrast to most previous studies [12,18] that have used internal validation processes. Internal validation processes endanger the generation of optimistically biased estimates of the classification method's performance, whose occurrence has been noted in our experiments. For the best individual classifier (LSM³PNN) classification rates reduced drastically when the external cross-validation process was used (see Tables 1–4). The latter confirms the argument that if internal validation processes are applied, the resulting estimates of the classifier's performance are considerably biased. Thus, up to this point, we have developed a CAD system that ensures to a certain extent that its success rates in discriminating normal, infectious and cancer prostate tissues will remain the same even when evaluating new prostate cases.

According to [35], the exhaustive search and the branch and bound methods are the only feature selection methods that guarantee optimal feature subset selection. Exhaustive search is computationally intensive and cannot be used for large number of features combinations. However, due to sparsely of data (to use a larger number of features it is required a larger number of cases), in our case combinations up to five features might be regarded as adequate considering the high classification results that we have obtained using these features. Other methods such as sequential floating forward search, 'plus l-take away r', and best individual features are suboptimal. We have selected to use exhaustive search for optimal feature selection. For comparative purposes we have also used sequential floating selection (see Table 5).

Following, we have investigated a classifier combination scheme as an alternative to using individual classifiers. The motive was to test whether the utilization of an effective combination of a panel of classifier's, which explore data from a different perspective (geometrical criteria for the SVM, statistical criteria for the Bayesian and augmentation in conjunction with non-parametric density estimation for the LSM³PNN), might enhance sensitivity and specificity in discriminating different prostate pathologies. The latter has been verified to a certain extent. Regarding the discrimination of PRNO, the

MC-system gave slightly inferior results compared to the best individual classifier (89.5% and 88.7%, respectively). However, classification of PRIN and PRCA enhanced. For PRIN cases, classification accuracy notably enhanced from 77.9% to 79.6%, whereas for PRCA case, success rates increased from 76.1% to 82.7%. These results lead to two important conclusions: (a) the potential of MC-system to boost accuracy in prostate pathologies classification and (b) the potential to sufficient accurately discrimination between PRCA from PRIN cases. The latter discrimination has gained wide attention recently [23,24], since infectious prostate tissues may represent one mechanism through which prostate cancer develops, thus, their accurate recognition is of increasing merit. To the best of our knowledge, neither MC-systems nor the discrimination of PRIN from PRCA cases has been reported in literature.

In this study, a CAD system has been developed to investigate whether accurate separation of normal, infectious and cancerous prostate tissues might be feasible by analysing textural properties of TRUS images. Three important conclusions have risen: (a) Results are in line with those presented in literature, however, it has to be stressed that the proposed system has been evaluated using an external cross-validation process. Thus, we may argue that results are indicative of the generalization of the method to new prostate cases. (b) The utilization of MC-system improves the discrimination of prostate pathologies (infectious prostate cases: improvement from 77.9% to 79.6%; cancer prostate cases: improvement from 76.1% to 82.7%). Finally, (c) the discrimination of infectious and cancerous prostate cases is feasible with reasonably high accuracies, reaching 79.6% and 82.7%, respectively.

5. Conclusions

As a conclusion, it can be considered that the combination of the complementary information from many classifiers in the form of the proposed multi-classifier system is an effective alternative to single classifier systems for prostate diagnosis. Additionally, the proposed system can be used as a second opinion tool for assisting physicians in prostate tissues characterization when analysing transrectal ultrasound prostate images.

Conflict of interest

None declared.

Appendix A. A short description of the textural features employed in this paper

A.1. Features derived from the image gray-level histogram [27]

1. Mean value

$$m = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} I(x, y) \quad (A.1)$$

where N_x and N_y are the image dimensions and $I(x,y)$ is the gray-level of the image pixel with coordinates (x,y) .

2. Standard deviation

$$s = \left[\frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} (I(x, y) - m)^2 \right]^{1/2} \quad (A.2)$$

3. Skewness

$$sk = \frac{1}{N_x N_y s^3} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} [I(x, y) - m]^3 \quad (A.3)$$

4. Kurtosis

$$ku = \frac{1}{N_x N_y s^4} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} [I(x, y) - m]^4 \quad (A.4)$$

A.2. Features derived from the image gray-level co-occurrence matrix

The (i,j) -element of the co-occurrence matrix [25] for angle θ of an image matrix I is defined as

$$P_\theta(i, j) = \#\{(x_1, y_1), (x_2, y_2) : f(x_1, y_1) = i, f(x_2, y_2) = j, \\ x_2 = x_1 + \cos \theta, y_2 = y_1 + \sin \theta\}$$

where (x_1, y_1) and (x_2, y_2) are two consecutive pixels of the image at angle θ .

In the present work, the co-occurrence matrices for angles $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ were calculated, and their mean value was used for independency over rotation.

From the co-occurrence matrix the following features can be defined:

5. Angular second moment

$$ASM = \sum_i \sum_j p(i, j)^2 \quad (A.5)$$

where $p(i,j)$ is the (i,j) -element of the normalized co-occurrence matrix and N_g is the number of gray-levels in the image.

6. Contrast

$$CON = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_i \sum_j p(i, j) \right\} \quad (A.6)$$

$$\left. \begin{array}{l} \\ \\ |i-j| = n \end{array} \right\}$$

7. Correlation

$$COR = \frac{\sum_i \sum_j [ij p(i, j)] - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (A.7)$$

where μ_x, μ_y are the mean values of P_x, P_y correspondingly and σ_x, σ_y their standard deviations, where $P_x(i) = \sum_j^{N_y} p(i, j)$ and $P_y(i) = \sum_i^{N_x} p(i, j)$.

8. Auto-correlation

$$ACO = \sum_i^{N_g} \sum_j^{N_g} i j p(i, j) \quad (A.8)$$

9. Sum of squares

$$SSQ = \sum_i^{N_g} \sum_j^{N_g} (i - \mu)^2 p(i, j) \quad (A.9)$$

10. Inverse difference moment

$$IDF = \sum_i^{N_g} \sum_j^{N_g} \frac{p(i, j)}{1 + (i - j)^2} \quad (A.10)$$

11. Entropy

$$ENT = - \sum_i^{N_g} \sum_j^{N_g} p(i, j) \ln(p(i, j)) \quad (A.11)$$

12. Sum entropy

$$SEN = - \sum_{i=2}^{i=2N_g} P_{x+y}(i) \ln(P_{x+y}(i)) \quad (A.12)$$

where $P_{x+y}(k) = \sum_i^{N_g} \sum_j^{N_g} p(i, j)$, $k = 2, 3, \dots, 2N_g$ and

$$P_{x-y}(k) = \sum_i^{N_g} \sum_{\substack{i+j=k \\ |i-j|=k}}^{N_g} p(i, j), \quad k = 0, 1, \dots, N_g.$$

13. Sum average

$$SAV = \sum_{i=2}^{i=2N_g} i P_{x+y}(i) \quad (A.13)$$

14. Sum variance

$$SVA = \sum_{i=2}^{i=2N_g} (i - SEN)^2 P_{x+y}(i) \quad (A.14)$$

15. Difference variance

$$DVA = \text{VARIANCE}(P_{x-y}(i)) \quad (A.15)$$

16. Difference entropy

$$DEN = - \sum_{i=0}^{N_g-1} P_{x-y}(i) \ln(P_{x-y}(i)) \quad (A.16)$$

17. Information measure of correlation 1

$$IMC1 = \frac{ENT - ENTXY1}{\max\{ENTX, ENTY\}} \quad (A.17)$$

where $ENTX = - \sum_i^{N_g} \sum_j^{N_g} p_x(i, j) \ln(p_x(i, j))$, $ENTY = - \sum_i^{N_g} \sum_j^{N_g} p_y(i, j) \ln(p_y(i, j))$, and $ENTXY1 = - \sum_i^{N_g} \sum_j^{N_g} [p(i, j) \ln(p_x(i, j)p_y(i, j))]$.

18. Information measure of correlation 2

$$IMC2 = [1 - e^{-2(ENTXY2-ENT)}]^{1/2} \quad (A.18)$$

where $ENTXY2 = - \sum_i^{N_g} \sum_j^{N_g} [p_x(i, j)p_y(i, j) \ln(p_x(i, j)p_y(i, j))]$.

A.3. Features derived from the image gray-level run-length matrix:

Let $Q_{RL}(i, j)$ the (i, j) -element of the run-length matrix [26]. Then, i is the gray-level and j is the maximum number of consecutive pixel with gray-level equal to i at an angle θ .

In the present work, the run-length matrices for angles $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ where calculated, and their mean value was used for independency over rotation.

From the run-length matrix the following features can be defined:

19. Short-run emphasis

$$SRE = \frac{\sum_i^{N_g} \sum_j^{N_r} Q_{RL}(i, j)/j^2}{\sum_i^{N_g} \sum_j^{N_r} Q_{RL}(i, j)} \quad (A.19)$$

20. Long-run emphasis

$$LRE = \frac{\sum_i^{N_g} \sum_j^{N_r} Q_{RL}(i, j)/j^2}{\sum_i^{N_g} \sum_j^{N_r} Q_{RL}(i, j)} \quad (A.20)$$

21. Gray-level non-uniformity

$$GLNU = \frac{\sum_i^{N_g} \left[\sum_j^{N_r} Q_{RL}(i, j) \right]^2}{\sum_i^{N_g} \sum_j^{N_r} Q_{RL}(i, j)} \quad (A.21)$$

22. Run-length non-uniformity

$$RLNU = \frac{\sum_j^{N_r} \left[\sum_i^{N_g} Q_{RL}(i, j) \right]^2}{\sum_i^{N_g} \sum_j^{N_r} Q_{RL}(i, j)} \quad (A.22)$$

23. Run-percentage

$$RP = \frac{\sum_i^{N_g} \sum_j^{N_r} Q_{RL}(i, j)}{P} \quad (A.23)$$

where P the maximum number of run-lengths N_r .

REFERENCES

- [1] W.H. Cooner, B.R. Mosley, C.L. Rutherford Jr., et al., Prostate cancer detection in a clinical urological practice by ultrasonography, digital rectal examination and prostate specific antigen, *J. Urol.* 143 (1990) 1146-1152.
- [2] K. Shinohara, T.M. Wheeler, P.T. Scardino, The appearance of prostate cancer on transrectal ultrasonography: correlation of imaging and pathological examinations, *J. Urol.* 142 (1989) 76-82.

- [3] M. Inahara, H. Suzuki, H. Nakamachi, et al., Clinical evaluation of transrectal power doppler imaging in the detection of prostate cancer, *Int. Urol. Nephrol.* 36 (2004) 175–180.
- [4] E.J. Feleppa, S. Dasgupta, J.A. Ketterling, et al., Tissue-type imaging (TTI) of prostate cancer based on combined ultrasonic and magnetic-resonance methods: Latest developments, in: *IEEE Ultrasonics Symposium*, 2006.
- [5] T. Loch, Computerized supported transrectal ultrasound (C-TRUS) in the diagnosis of prostate cancer, *Urologe A* 43 (2004) 1377–1384.
- [6] E.J. Feleppa, A. Kalisz, J.B. Sokil-Melgar, et al., Typing of prostate tissue by ultrasonic spectrum analysis, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* 43 (1996) 609–619.
- [7] Y. Zhu, S. Williams, R. Zwiggelaar, Computer technology in detection and staging of prostate carcinoma: a review, *Med. Image Anal.* 10 (2006) 178–199.
- [8] S. Tong, H.N. Cardinal, R.F. McLoughlin, D.B. Downey, A. Fenster, Intra- and inter-observer variability and reliability of prostate volume measurement via two-dimensional and three-dimensional ultrasound imaging, *Ultrasound Med. Biol.* 24 (1998) 673–681.
- [9] A.L. Huynen, R.J.B. Giesen, J.J.M.C.H. Delarosette, et al., Analysis of ultrasonographic prostate images for the detection of prostatic-carcinoma—the automated urologic diagnostic expert-system, *Ultrasound Med. Biol.* 20 (1994) 1–10.
- [10] A.L. Ronco, R. Fernandez, Improving ultrasonographic diagnosis of prostate cancer with neural networks, *Ultrasound Med. Biol.* 25 (1999) 729–733.
- [11] U. Scheipers, H. Ermert, H.J. Sommerfeld, et al., Ultrasonic multifeature tissue characterization for prostate diagnostics, *Ultrasound Med. Biol.* 29 (2003) 1137–1149.
- [12] A. Lorenz, M. Blüm, H. Ermert, Th. Senge, Comparison of different neuro-fuzzy classification systems for the detection of prostate cancer in ultrasonic images, in: *IEEE Ultrasonics Symposium*, Toronto, Ont., Canada, 1997.
- [13] M.M.A. Salama, S.S. Mohamed, Computer-aided diagnosis for prostate cancer using support vector machine, in: *SPIE Conference on Medical Imaging: Visualization, Image-Guided Procedures, and Display*, 2005.
- [14] R. Llobet, J.C. Perez-Cortes, A.H. Toselli, A. Juan, Computer-aided detection of prostate cancer, *Int. J. Med. Inform.* 76 (2007) 547–556.
- [15] O. Basset, Z. Sun, J.L. Mestas, G. Gimenez, Texture analysis of ultrasonic images of the prostate by means of co-occurrence matrices, *Ultrasonic Imaging* 15 (1993) 218–237.
- [16] R.J.B. Giesen, A.L. Huynen, R.G. Aarnink, et al., Construction and application of hierarchical decision tree for classification of ultrasonographic prostate images, *Med. Biol. Eng. Comput.* 34 (1996) 105–109.
- [17] R.J.B. Giesen, A.L. Huynen, R. Aarnink, et al., Computer analysis of transrectal ultrasound images of the prostate for the detection of carcinoma: a prospective study in radical prostatectomy specimens, *J. Urol.* 154 (1995) 1397–1400.
- [18] M. Moradi, P. Mousavi, P. Abolmaesumi, Computer-aided diagnosis of prostate cancer with emphasis on ultrasound-based approaches: a review, *Ultrasound Med. Biol.* 33 (2007) 1010–1028.
- [19] C. Ambroise, G.J. McLachlan, Selection bias in gene extraction on the basis of microarray gene-expression data, *Proc. Natl. Acad. Sci. U.S.A.* 99 (2002) 6562–6566.
- [20] B. Efron, Bootstrap methods: another look at the jackknife, *Ann. Statist.* 7 (1979) 1–26.
- [21] B. Efron, How biased is the apparent error rate of a prediction rule? *J. Am. Stat. Assoc.* 81 (461) (1986) 470.
- [22] G.J. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*, Wiley, New York, 1992.
- [23] L.K. Dennis, C.F. Lynch, J.C. Torner, Epidemiologic association between prostatitis and prostate cancer, *Adult Urol.* 60 (2002) 78–83.
- [24] W.G. Nelson, A.M. DeMarzo, T.L. DeWeese, W.B. Isaacs, The role of inflammation in the pathogenesis of prostate cancer, *J. Urol.* 172 (2004) S6–S12.
- [25] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, *IEEE Trans. Syst. Man Cybernet.* 3 (1973) 610–621.
- [26] M.M. Galloway, Texture analysis using gray-level run lengths, *Comput. Graphics Image Process.* 4 (1975) 172–179.
- [27] S. Theodoridis, K. Koutroubas, *Pattern Recognition*, Academic Press, 1999.
- [28] D.F. Specht, Probabilistic neural networks, *Neural Netw.* 3 (1990) 109–118.
- [29] N. Piliouras, I. Kalatzis, N. Dimitropoulos, D. Cavouras, Development of the cubic least squares mapping linear-kernel support vector machine classifier for improving the characterization of breast lesions on ultrasound, *Comput. Med. Imag. Graphics* 28 (2004) 247–255.
- [30] C. Burges, A tutorial on support vector machines for pattern recognition, *Data Min. Knowl. Discov.* 2 (1998) 121–167.
- [31] R. Simon, M.D. Radmacher, K. Dobbin, L.M. McShane, Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification, *J. Natl. Cancer Instit.* 95 (2003) 14–18.
- [32] A.K. Jain, R.P.W. Duin, J.C. Mao, Statistical pattern recognition: a review, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (2000) 4–37.
- [33] L.A. Alexandre, A.C. Campilho, M. Kamel, On combining classifiers using sum and product rules, *Pattern Recog. Lett.* 22 (2001) 1283–1289.
- [34] J. Kittler, M. Hatef, R.W. Duin, J. Matas, On combining classifiers, *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (1998) 226–239.
- [35] K.J. Anil, R.P.W. Duin, J. Mao, Statistical pattern recognition: a review, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (2000) 4–37.