Comparative Study of Four Widely used Classifiers for Prostate Cancer Detection with Multi-parametric MRI

Chengyan Wang[1], Shuangjuan Cheng[2], Juan Hu[2], He Wang[2], Xuedong Yang[2], Jue Zhang[1,3], Xiaoying Wang[1,2], and Jing Fang[1,3]

[1]*Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, Beijing, China,* [2]*Department of Radiology, Peking University First Hospital, Beijing, Beijing, China,* [3]*College of Enigneering, Peking University, Beijing, Beijing, China*

Purpose:

Prostate cancer (PCa) is the second most frequently diagnosed cancer and the sixth leading cause of cancer death among men worldwide [1]. Several studies [2-4] have proven that the diagnostic accuracy of PCa detection can be significantly improved by combining different MR sequences, and several computer-aided diagnosis (CAD) systems have been proposed to integrate the MR information. However, no comparison has been made to find out which system performs better. In this study, we evaluate the performance of four widely used classifiers using leave-one-out (LOO) method.

Materials and Methods:

Study Population:

Between January 2006 and January 2010, totally 256 consecutive patients who were suspected of PCa by urologist in the prostate MR database were selected into our retrospective study, with Clinical Trial Ethic Committee permission. Among these patients, 126 cases (mean age 70 ± 9 years, range 37-86 years) were confirmed to be PCa by biopsy, the other 130 cases (mean age 67 ± 7 years, range 37-85 years) were not detected of PCa by serial biopsy and long-term follow-up (16 to 63 months, mean 36 months). All MR examinations were carried out on a 1.5 Tesla clinical MR scanner (Signa TM; GE Medical Systems, Milwaukee, WI), using a pelvic phased-array coil. The multi-parametric MR images of a typical example of PCa was shown in Fig.1.

Image Preprocessing:

The general data flow diagram of this CAD system was presented in Fig.1. Before analysis, T2-weighted images were standardized to correct for background and nonlinearity of the MR image intensity scale. apparent diffusion coefficient (ADC) maps were calculated based on the diffusion-weighted (DW) images (b = 0 and 800 s/mm^2). For the dynamic contrast enhanced (DCE) MR series, the relative signal enhancement (SI(t)-SI(t=0))/SI(t=0) (where 't' stands for time) was preferred to the raw signal. A linear registration scheme was used to spatially align the corresponding MRI protocols.

Features Extraction&Selection :

Regions of interest (ROIs) with 8×8 pixels were manually sketched from the cancerous regions and randomly selected from the normal regions in the prostate MR images by two experienced radiologists (7 and 9 years of prostate MR experience). The feature set was derived from T2 value, gray-level histogram and co-occurrence matrix (GLCM) of T2-weighted image, ADC value, and semi-quantitative features extracted from DCE curves. After extracting these features above, we used the Sequential Forward Selection (SFS) feature selection scheme to identify an ensemble of features that will allow for optimal identification of PCa.

Classifier Design:

Four kinds of classifiers were trained and compared: multi-layer perceptron (MLP), support vector machine (SVM), logistic regression (LR) and linear discriminant analysis (LDA) classifiers. All of the methods were implemented using MATLAB® (2010a, The Mathworks Inc., Natick, MA). Receiver operating characteristic (ROC) analysis was performed for all the patients, and area under the ROC curve (AUC) was used to evaluate the performance of different classifiers

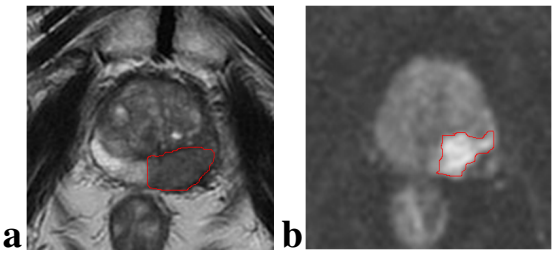

Fig.1. Multi-parametric MR images of a typical PCa patient, where the red outlines are the most suspicious regions of PCa.
a) T2WI; b) DWI (b = 800 s/mm^2).

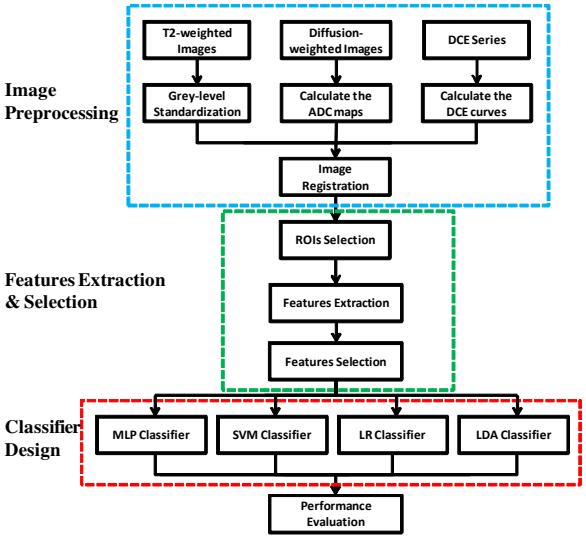

Fig.2. General data flow diagram of this CAD system.

Results:

Totally, 238 ROIs were selected from the peripheral zone (PZ) (146 control and 92 cancerous ROIs), and 188 ROIs were selected from the central gland (CG) (136 control and 52 cancerous ROIs).

The AUCs were 0.821 ± 0.038 for the LDA classifier, 0.891 ± 0.031 for the SVM classifier, 0.838 ± 0.031 for the LR classifier, and 0.909 ± 0.029 for the MLP classifier (see Fig.3). The accuracies were 0.809, 0.888, 0.840 and 0.915 for these four classifiers, respectively.

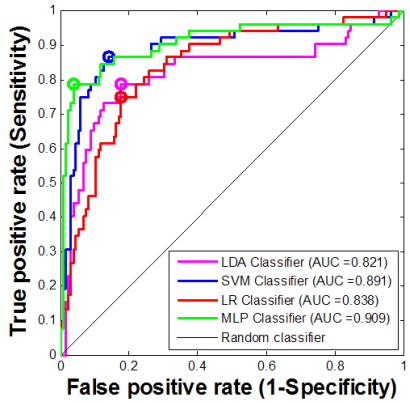

Fig.3. ROC curves of the results calculated by four different classifiers of **a)** LDA classifier; **b)** SVM classifier; **c)** LR classifier ; **d)** MLP classifier. Cut-offs marked with circles were selected to reach the best accuracy.

Discussion:

LDA classifier seems to achieve the lowest AUC comparing with the others. Though this method is popular because it's not computationally demanding and easy to interpret, but it makes certain assumption that the features used have linear relations, whereas the other classifiers do not. SVM and MLP classifiers achieve high AUC of 0.891 and 0.909 respectively, which proves the complex non-linear relationships between the input features.

Conclusions:

The study's main contribution is to evaluate and compare four widely used classifiers (LDA, LR, SVM and MLP classifier). MLP and SVM classifiers which have been successfully applied in many branches of medical diagnostics seem promising here. Because of their abilities to resolve nonlinear complex relations among input variables, without the need for any prior assumptions about these relations, MLP and SVM models are more advisable for PCa detection.

References:

[1] Jemal, et al. Ca Cancer J Clin 2011;61:69–90.
[2] Kozlowski, et al. J. Magn. Reson. Imaging 2006; 24:108–13.
[3] Girouin, et al. Eur. Radiol 2007; 17: 1498–509.
[4] Cheikh, et al. Eur. Radiol 2009; 19: 770–8.