Support Vector Neural Networks versus Logistic Regression MR based diagnostic model for classification of transition zone prostate cancer

Nikolaos Dikaios[1,2], Jokha Alkalbani[2], Alex Kirkham[3], Clare Allen[3], Hashim Ahmed[4], Mark Emberton[4], Alex Freeman[5], Steve Halligan[2], Stuart Taylor[2], David Atkinson[2], and Shonit Punwani[2]

[1]Medical Physics, UCL, London, Greater London, United Kingdom, [2]Centre of Medical Imaging, UCL, Greater London, United Kingdom, [3]Radiology, UCL, Greater London, United Kingdom, [4]Urology, UCL, Greater London, United Kingdom, [5]Histopathology, UCL, Greater London, United Kingdom

Purpose: Multi-parametric MRI (mp-MRI) facilitates identification of transition zone cancers, yet its overall diagnostic accuracy is likely lower in this part of the prostate compared with the peripheral zone. Benign hyperplastic nodules within the transition zone likely make the localisation of cancer difficult. Logistic regression (LR) models[1] for classifying transition zone (TZ) prostate cancer (PCa) on mp-MRI were previously derived and validated. Here we explore whether the application of support vector machine (SVM) neural network (SVNN) algorithms can improve classification accuracy. The proposed SVNN algorithm is trained on 70 patients and temporally validated on a second independent cohort of 85 patients.

SVNN: Logistic regression (LR) analysis is commonly used in diagnostic modelling, because it provides a deterministic model yielding weighting factors for each contributing feature. SVM with a non linear kernel classify data into a richer feature space, improving the classification accuracy but increasing the computational cost. SVNN is a SVM-like maximal margin training algorithm for NN that solves the following optimization

$$\min_{w_1,w_2,b_1,b_2}\left(e_{\min}+e_{\max}+\frac{C}{N}\sum_{i=1}^{N}H(z_i y_i)\right)$$

Where z is the target class, y is the model output i.e. $y=w_2 \cdot Sigmoid(w_1 \cdot x+b_1)+b_2$, w_1/w_2 are the synaptic weights of the hidden/output layer, b_1/b_2 are the bias vectors of the hidden/output layer, x are the input features, C is a regularization term, and H is the Hinge loss. The regularization scheme used, called eigenvalue decay, penalize the minimum and maximum eigenvalues of $(w_1 w_1^T)$ in order to improve the classification margin and the generalization ability if the algorithm.

Methods: A total of 155 patients (training cohort 70; temporal validation cohort 85 patients) underwent mp-MRI and transperineal-template-prostate-mapping (TPM) biopsy. 28/70 patients from the training cohort, and 25/85 patients from the temporal validation cohort had significant cancer on TPM. Positive cores were classified by three cancer definitions: (i) anycancer; (ii) definition-1 (≥Gleason 4+3 or ≥6mm cancer core length (CCL)) [high risk-significant]; and (iii) definition-2 (≥Gleason 3+4 or ≥4mm CCL) cancer [intermediate-high risk-significant]. Previous work[1] has showed that the three mp-MRI variables that best classified TZ PCa are the Apparent Diffusion Coefficient (ADC), normalized T2 signal intensity and Maximum Enhancements (ME). A LR and SVNN model were trained based on these parameters and validated on the temporal cohort using receiver operating characteristic (ROC) analysis for each cancer definition. SVNN labels training datasets by $y_i \epsilon \{+1,-1\}$, probabilistic output is generated from the binary y_i using the method of Lin et al (2007)[3]. This is necessary to generate receiver operating characteristic (ROC) and calculate the area under curve (AUC). Probabilistic maps of a non cancer patient and a patient with significant cancer were calculated.

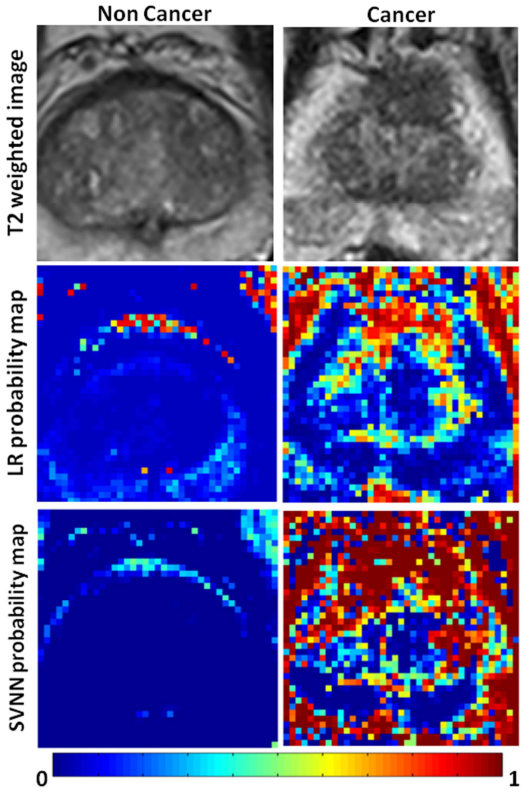

Figure1: Probability maps generated by the LR and SVNN models for a non-cancer and a def 1cancer prostate patient

Table 1: Evaluation of the LR and the SVNN model on the training and the temporal validation cohort using ROC analysis. AUC stands for area under curve, LB and UB are the lower and upper bound of the confidence interval (CI)

		Training Cohort				Temporal Cohort			
				Asymp. 95% CI				Asymp. 95% CI	
		AUC	Std	LB	UB	AUC	Std	LB	UB
anyCa	LR	0.78	0.06	0.67	0.89	0.76	0.05	0.66	0.87
	SVNN	0.85	0.05	0.76	0.94	0.80	0.05	0.70	0.89
def1	LR	0.80	0.05	0.70	0.90	0.67	0.06	0.55	0.79
	SVNN	0.89	0.04	0.81	0.96	0.75	0.06	0.64	0.86
def2	LR	0.79	0.05	0.68	0.89	0.70	0.08	0.55	0.85
	SVNN	0.86	0.05	0.77	0.94	0.75	0.07	0.62	0.87

Results: Table 1 illustrates that SVNN had higher ROC-AUC than LR on the training cohort for all cancer definitions. There was a drop in performance of both algortihms when applied to the temporal validation cohort. However, the SVNN ROC-AUC was consistently higher than for the LR model for all cancer definitions. Figure 1 demonstrates an example of the probability maps generated from each model when applied to patients with and without histologically confirmed cancer. The LR localisation of cancer was a closer approximation of the cancer localised by radiologists (shown on T2 weighted images), whereas SVNN is over-sensitive. Where no cancer was present the SVNN probabilities were consistently lower than LR.

Conclusions: SVNN[1] exploits new training methods, in explicit SVM, for multilayer perceptron (MLP) NN. This avoids the overtraining of NN and is faster than usual SVM. This study suggests that more complex training methods like SVNN do not have a clear advantage over simpler methods like LR. Further LR is less computationally demanding and easy to interpret.

Acknowledgements: This work was supported by the CRUK/EPSRC KCL/UCL comprehensive cancer imaging centre. ND was supported by UK EPSRC grants EP/I018700/1 and EP/H046410/1.

References: [1] Dikaios N et al, Eur Rad 2014; 32,229-35. [2] O. Ludwig, PhD Thesis, University of Coimbra, 2012. [3] Lin HT et al, Mach Learn 2007; 68:267-76.