Collaborative project with the Clatterbridge Cancer Centre, NHS Foundation Trust

Multi-parametric intra-subject classification of prostate cancer using Support Vector Machines

7030DATSCI – Data Science Project

Riley, Patrick
686073

Abstract

Prostate cancer is becoming increasingly prevalent with approximately 1 in 8 men being diagnosed, making it the most common cancer in men in the UK. It has been found that nearly two-fifths of cases are diagnosed at late stages. Multi-parametric Magnetic Resonance Imaging (mpMRI) is used widely to aid detection of prostate cancer and determining its aggressiveness, although it is not the primary tool. Machine learning methodologies can be applied to 'learn' and assist clinical decisions, aiding multidisciplinary healthcare teams in diagnosis and treatment stages of patient care. This report looks at applying Support Vector Machines to support the development of machine learning based diagnosis of prostate cancer on the appearance of tumours, through mpMRI to support mpMRI as a primary diagnosis tool. Model training is conducted based on the SPIE-AAPM-NCI PROSTATEx challenge data set and through *collaboration with the Clatterbridge Cancer Centre, NHS Foundation Trust*, the methodology can be applied onto anonymised cancer patient data, providing real results. However, reported performance evaluation is based on a 10-fold cross-validation split of the challenge data set, based on an outcome variable denoting the level of the Gleason score value of the tumour finding. Seven data combinations were tested including different scan parameters and patient level metadata. The best models achieved AUC measures of 0.746 and 0.796.

Acknowledgements

I would like to thank my family who have continually pushed me forward throughout my academic life. I am forever in debt to them for their support, their thoughts and contributions to the work. You have all helped me throughout my stress, and sarcasm by calming me down and joining me in the joyous times also. Good luck with me in the next stage!

I would also like to thank Steven, for putting up with a lot more of my sarcasm than he admittedly would like! Thank you for making me a very happy partner and long may the support continue, from me to you as well. Any stress I have had with this work, you have changed my outlook and used it to help me push on. You must put up with that a little longer!

I must thank Dr Sandra Ortega-Martorell, Dr Ivan Olier and Prof Paulo Lisboa for their constant support and guidance, intertwined with some amusing conversation, throughout this study. Listening to their experience with similar areas has been enriching for my academic development and I cannot thank them enough. I am excited to continue working with them throughout my PhD studies.

Thank you to Dr Marc Rea at the Clatterbridge Cancer Centre, NHS Foundation Trust, for his insight and support throughout the work.

I would also like to acknowledge the Department of Applied Mathematics and the Astrophysics Research Institute at LJMU for the inspiring delivery of both my undergraduate and postgraduate degrees.

Thank you to Steven & Ryan for proof-reading.

Contents

Abstract 1
Acknowledgements 2
List of figures 4
List of tables 4
List of equations 4
Introduction 5
Literature Review 6
Support Vector Machines 6
Classification of prostate cancer lesions 6
Segmentation of prostate cancers 7
Description of data source: PROSTATEx Challenge data set and Clatterbridge Cancer Centre data 9
PROSTATEx dataset 9
DICOM data 9
Metadata 10
Clatterbridge Cancer Centre data 13
Methodology 14
Implementation 14
Pre-processing: Image equalisation, ROI extraction, normalisation and class weightings 14
Equalisation of the images 14
Issues with dataset/metadata 15
Extracting the ROI 15
Normalisation 15
Class weighting 15
Methods: Support Vector Machines 15
Kernel selection 17
Hyperparameter selection 18
Methods: Logistic Regression 18
Validation: 10-fold cross validation 19
Results 20
Combinations of data used for models 20
Analysis of results 20
How the methodology may be applied to data from the Clatterbridge Cancer Centre 23
Conclusions 24
A Self Evaluation of Project Outcome 25
References 26

List of figures

Figure 1: The three anatomical planes studied upon each patient in this work. 11
Figure 2: ijk example. This is slice 12 with the box around the position given from 'ij' of 'ijk' - in this case, around column 139 and row 166. This shows how the patch is extracted 12
Figure 3: original image (left) and equalised image (right) through histogram equalisation 14
Figure 4: SVM decision boundaries 16
Figure 5: k-fold cross validation 19
Figure 6: ROC plot for model 2: T2 transverse, ADC transverse and lesion zone. AUC 0.796. 22
Figure 7: ROC plot for model 3: T2 transverse, ADC transverse, lesion zone and patient metadata (age and weight). AUC 0.746. 22
Figure 8: Application of the model onto an MRI image from the Clatterbridge Cancer Centre 23

List of tables

Table 1: ijk example 11
Table 2: Distribution of ages and weights of patients in the PROSTATEx data set 12
Table 3: Distribution of zones of lesion findings in the PROSTATEx data set 12
Table 4: Hyperparameter tuning values for C and γ 18
Table 5: Combinations of data and metadata for models fitted. Frequency of outcome class values and class weightings for SVM are also listed. 20
Table 6: SVM and Logistic Regression has been performed on each data set combination test as referred to in Table 5. Each test was run three times, and the best AUC is listed, along with the average AUC across all three runs. The best AUC score is 0.796, SVM test 2. 21

List of equations

Equation 1: Cumulative distribution function (cdf) of all probabilities lying in the domain 15
Equation 2: Formula for new probability distribution function of equalised image 15
Equation 3: Class Weighting equation 15
Equation 4: Linear model SVM. 16
Equation 5: Canonical representation of the decision hyperplane. 16
Equation 6: SVM optimisation problem: the factor of 12 is included for convenience. 16
Equation 7: constrained optimisation with Lagrange multipliers for finding the decision boundary 17
Equation 8 17
Equation 9 17
Equation 10: classifying new data points - evaluating the sign of $y(\boldsymbol{x})$ 17
Equation 11 17
Equation 12 17
Equation 13 17
Equation 14: Radial Basis Function kernel. 18
Equation 15: Fitting the log odds and explanatory variables to a linear model - Logistic Regression. Y = (0,1) is the binary variable, X = $(X1, \dots, Xn)$ are the n explanatory variables and $\beta = (\beta 0, \dots, \beta n)$ are the regression coefficients to be estimated based on the data. 18
Equation 16: Logistic function to predict probability of occurrence to assign a class label as a prediction. 19

Introduction

Prostate cancer is the most common cancer in men in the UK, with over 40,000 new cases diagnosed per year. [1] It develops slowly with no signs for many years and no apparent symptoms until the prostate is large enough to affect the urethra. The prostate is a small gland in the pelvis, only found in men, and its purpose is to help in the production of semen. Causes of prostate cancer are largely unknown and treatment may not be immediate – symptoms may not appear for decades and treatment may not be needed in the first instance. However, problems can develop physically and mentally if left untreated, such as urinary incontinence, anxiety or depression.

Early detection and diagnosis of prostate cancer is considered a key factor. Recently the charity, Orchid found that near two-fifths of prostate cancer cases were diagnosed at a later stage and that for the ageing population, prostate cancer provision is important. [2] [3]

Current practice for diagnosis of prostate cancer is through a 'trans-rectal ultrasound biopsy,' usually conducted after a positive blood screening test. However, this test has a relatively low specificity, leading to overdiagnosis and therefore overtreatment. Although Magnetic Resonance Imaging (MRI) diagnosis can overcome this, it requires specialist knowledge to review the prostate MRI, which is also time consuming. This leads to early and automatic detection and classification from MRI scans which this study looks to tackle, which reduces the stated problems and strengthens the use of prostate screening through MRI.

This work is an attempt to classify the aggressiveness of prostate cancers from MRI scans of the prostate, on a within subject analysis basis, through machine learning. This work is a collaborative project with The Clatterbridge Cancer Centre, NHS Foundation Trust. The methodology can be applied on anonymous patients' various MRI scans of the prostate which have been provided for use in this research.

This project will look at Support Vector Machines (SVM) as first described by Cortes & Vapnik [4] and Boser et al. [5], a method related to statistical learning theory and computational learning theory [6] which was a popularised method through success in handwritten digit recognition; from a study by Bottou et al. [7] the test error being the same of a carefully constructed neural network. The SVM is a supervised machine learning algorithm, in this study being used for classification. It is based on minimising an upper bound of the generalisation error through a hyper-plane separating classes as best as possible. The "support vectors" are the points on the edge of the group being separated. This work takes the ideas used by Kitchen & Seah [8].

The following section will focus on a literature survey into this field of work on SVMs and classification of prostate lesions, as well as segmentation as another area of prostate cancer studies followed by a description of the data source used in this report, the PROSTATEx challenge data set. The methodology section will show how the SVM is used against this classification problem, through its application onto various parameters of MRI scan. The results section will display the output of the method with AUC measures along with an analysis of the outputs. The conclusion section will study the findings of the work. Finally, a self-evaluation of the project outcome is presented, which considers how well the study went and what further work will take place.

Literature Review

Recently there has been work on classification of prostate cancers, although some problems remain such as the large intra-subject prostate shape variation which this work will attempt to overcome. It is known that different professionals may consider boundaries of a prostate cancer to be different [9] and so, accurate classification is an area with wide research. Previous work has been consulted and a study of these is presented here. As work on segmentation of prostate cancers is also a wide area of research with close links to classification, a study of these is also presented.

Support Vector Machines

Support Vector Machines (SVMs) were first presented by Cortes & Vapnik [4] as a learning machine for two-group classification problems, by mapping input vectors to a high dimensional feature space where a linear decision surface is constructed. The special properties of the decision surface ensure a strong generalisation ability of the learning machine. SVMs tend to perform better than other machine learning methodologies and statistical models that are developed. Implementation of SVMs have a broad application, such as model forecasting [10] and image retrieval [11].

Classification of prostate cancer lesions

Kitchen & Seah [8] present an application of SVM on the PROSTATEx data set upon three different modalities available, noting how the methodology is conceptually simple and provides competitive results in a low training time. Like this work, a $5mm * 5mm$ centred patch around the lesion over multiple modalities. SVMs are applied for this image patch data alone, another for the image patch data with information on the location of the lesion and another with a class weighting applied due to the imbalance of the outcome variable. After hyperparameter tuning and kernel selection, an AUC measure of 0.82 is achieved in the latter model. Although it was not the best performing model of the challenge, the simplicity and flexibility of the model is noted.

Deep learning methods can be applied on 3D multiparametric MRI. Liv et al. [12] faced the challenges of heterogeneous raw data and the relatively small sample size available in medical imaging through development of a deep convolutional neural network technique, performing end-to-end training of it on 3D multiparametric MRI images of prostates. In this paper, the PROSTATEx challenge data set was used where the pre-processing included data co-registration and slicing. The 'XmasNet' deep learning architecture was created through a 'greedy bagging' algorithm. It was compared to 'XGBoost,' created through gradient boosting. Comparisons were made to a previously tested mpMRI model, with the deep learning 'XmasNet' achieving the strongest AUC measure of 0.84 on the test dataset, whereas 'XGBoost' achieved 0.80. This shows strong potential for deep learning in prostate cancer imaging diagnosis.

Fehr et al. [13] studied automatic classification of prostate cancer Gleason scores. This work presents a machine learning-based automatic classification of prostate cancer aggressiveness by combining apparent diffusion coefficient (ADC) and T2-weighted (T2-w) MRI-based texture features. In this work, the images were acquired from 147 patients after data cleaning. After the tumours and normal structures were identified and segmented by research fellows, texture features were created for combinations of ADC and T2-w images. Three different methods were used for classifying the Gleason scores: *t test Support Vector Machine, Adaptive Boosting (AdaBoost) and Recursive Feature Selection Support Vector Machine.* Each method was evaluated using 10-fold cross-validation, and the hyperparameters for each classifier were selected separately through K-fold cross validation. The efficiency of classifying a Gleason score of 6(3+3) (n = 34) against a Gleason score of 7 or higher (n = 159), was assessed, as well as a Gleason score of 7(3+4) (n = 114) against a Gleason score of 7(4+3) (n = 26). This method distinguished between the Gleason scores of 6(3+3) and Gleason scores of 7 or

higher with 93% accuracy for cancers occurring in the peripheral and transition zones of the prostate, and 92% for cancers occurring in the peripheral zone alone.

Segmentation of prostate cancers

Ortega-Martorell et al. [14] utilised Non-Negative Matrix Factorisation (NMF) and one of its variants, Convex-NMF, on human brain data to obtain and interpret types of constituent tissue (infiltrating and necrotic tissues) and class prototypes (tumour types) with unsupervised and semi-supervised methods respectively. The two methods were applied for comparison, with Convex-NMF yielding better results. The steps taken were to first transform the data to maximise class separation using the Fisher Information matric, which amplified distances along the direction of the classes. Applying this mapped the data into a Riemannian space; through Multidimensional Scaling methods, the data was then mapped into a Euclidean space. Finally, Convex-NMF was applied for source extraction, obtaining class prototypes. The results of the paper shown that semi-supervised methods yielded better classification accuracies across training and test datasets, where two sources were employed; furthermore, a similar model also provided discrimination between two types of tumour which in general, are difficult to distinguish between.

Delimitation of brain tumours in mice in a fully unsupervised methodology, through Convex-NMF for source extraction, is presented by Ortega-Martorell et al. [15]. MRI and MRS (magnetic resonance spectroscopy) were acquired from both brain and tumour tissue as well as ex vivo post-mortem histology slides of the same animals, for evaluation. Two underlying source signals were calculated through NMF variants through the study of attained spectra within voxels of the data, with the signals resembling tumour and non-tumour capturing separation between the two main tissue types. Calculated correlations between spectra in each voxel led onto the sources yielding the highest correlation gaining its assigned label as per the sources. “Source-based label maps” were created representing both calculated sources, with strong rates of sensitivity and specificity in most cases against the supervised labelling, as well as reasonable results when borderline pixels were considered. Convex-NMF was used for source extraction as it performed consistently well with accuracies over 98% throughout for correlation results. This paper showed that the nature of an unsupervised approach for Convex-NMF; the algorithm itself can help radiologists tackle one of the main sources of uncertainty in the clinical management of brain tumours, which is the difficulty of appropriately delimiting the pathological area.

Reda et al. [16] used an evolving geometric deformable model which was guided by a speed function depending on intensity attributes extracted from diffusion weighted images (DWI) with non-negative matrix factorization to segment the prostate. The system in this work takes the DWI images of the prostate, segments the prostate by guiding a deformable model through NMF, constructs cumulative distribution functions (CDFs) of the estimated apparent diffusion coefficients as a global descriptor of water diffusion inside the prostate tissue, and classifies the extracted CDFs with an auto-encoder network, trained by deep learning. This network can separate benign and malignant prostates. The system was tested on DW-MRI data for 53 subjects and through a ‘leave-one-subject-out’ mode of experiment the overall diagnostic accuracy was 100% for the $b - value = 700 s/mm^2 only$. Where separate training and test sets were used, the accuracy was 92.3%.

Toth et al. [17] presented a scheme for automated prostate segmentation in vivo through multi-modal MRI, and MR spectroscopy. This was through a hierarchical unsupervised spectral clustering scheme using MRS data to isolate the region of interest (ROI) corresponding to the prostate and using an Active Shape Model (ASM) segmentation scheme where the ASM is initialized within the ROI previously obtained. The ASM here was trained by identifying user-selected landmarks on the prostate boundary on T2 MRI images. Over experiments of 150 prostate MRI slices through cross validation, gave an

average segmentation sensitivity, specificity, overlap and positive predictive value of 89%, 86%, 83% and 93% respectively.

Liu et al. [18] applied a deformable ellipse model to find an ellipse that best fits the prostate shape. Then, this ellipse was used to initiate the level set and constrain the level set evolution with a shape penalty term. Finally, certain post processing methods were applied to refine the prostate boundaries. This method was applied to diffusion-weighted MRI data, to test the performance. The Dice similarity coefficient (DSC) and its standard deviation was used for the calculation of the accuracy, was 0.91 ± 0.05. This work was later developed for 3D prostate segmentation [19] by describing the shape of the prostate by parametric deformable ellipses, which includes shape information in the 'guiding energy function.' Coarse segmentation to extract shape information is applied, then the shape prior is incorporated into the active contour model, leading to application of morphological operations to refine the segmentation results. Reflective of the previous work, the DSC was 0.86.

An older method for segmentation of prostate cancer relied on ultrasound images, as shown in [20]. The edge detection algorithm is three-fold in this work – contrast enhancement and speckle reduction in the transrectal ultrasound images, smoothing through an anisotropic diffusion filter and detection of the most probable edges describing the prostate. It was found that the accuracy of the prostate edges that were detected, were as good as those of the human observers.

Description of data source: PROSTATEx Challenge data set and Clatterbridge Cancer Centre data

The data analysed in this study are of DICOM format – Digital Imaging and Communication in Medicine. It is known as the international standard to: transmit, store, retrieve, print, process and display medical imaging information [21]. This includes its header information which contains data for each image as well as pixel level data of the images. A description of the data sets and the metadata is presented.

PROSTATEx dataset

The 2017 PROSTATEx challenge data set has been used to create the models in this study. [22]. The International Society for Optics and Photonics (SPIE), American Association of Physicists in Medicine (AAPM) and the National Cancer Institute (NCI) produced the dataset for the Prostate MR Classification Challenge, which was focussed on quantitative image analysis methods for the diagnostic classification of clinically significant prostate cancers [23]. The data were acquired for various parameters of MRI scan, including 'T1-weighted', 'T2-weighted' and 'diffusion-weighted' images which were available in most cases.

Because this dataset is a challenge dataset, the truths for the training dataset are known however, the truths for the test dataset as of the time of the study, are still unobtainable. The "PROSTATEX *training* dataset" is used in this study and is referred to as the PROSTATEx data set herein[1]. The data set contains 330 lesion findings over 204 anonymised patients; each patient has had prostate mpMRI scans performed on them and the data is held under various parameters and planes of MRI scan.

Each patient has had various parameters of scan taken for different planes. The three different planes are sagittal, coronal and transverse; the sagittal plane goes from side-to-side of the region of interest (ROI); the coronal plane goes front-to-back of the ROI; and the transverse plane goes from top-to-bottom of the ROI. These planes are represented in Figure 1. The scans hold a varying number of images – or, 'slices' – around the prostate, creating a 3D view around the scan area. Each case also has a variety of other parameters of data available, such as K-trans images, although these are not used in this work. The T2-weighted images are available for every patient, and the transversal plane for the apparent diffusion coefficient (ADC) scan parameter holds images for almost all cases and so are also used.

DICOM data

The data in the PROSTATEx data set is held as DICOM files. A DICOM file holds header and image data sets packed into a single file. [24] Various amounts of data can be held for each file depending on what is entered: in the header this includes demographic information on the patient (such as age, sex, weight and birthdate), acquisition parameters for the imaging study (such as MR modality), image dimensions, matrix size and more. Personal data can be withheld if required. The header is followed by a single attribute containing all the pixel intensity data for the image. This is held as a long series of 0s and 1s, which can be reconstructed as the image by using information from the header.

For this study, the age and weight of the patient is available from the DICOM header data and is applied to some of the models. It is stated in [8] that the extra metadata increases performance modestly and this is done in this work to study its effect.

[1] 10-fold cross validation is used to mitigate the use of just the training data and this and is described in the Methodology section of the report

Metadata

Various metadata is provided with the dataset for the provided DICOM images and the lesion findings as comma separated value (CSV) files. These have been used together to locate the lesion findings from the correct image.

1. Image metadata

   This metadata documents all the relevant DICOM files containing a lesion finding with various details provided, including:

   ProxID – patient identifier; ***Name*** – series description (scan parameter); ***fid*** – unique lesion finding ID; ***Pos*** – scanner coordinate position of finding; ***ijk*** – *i*th column, *j*th row and *k*th slice coordinates of the finding. Other details are available such as voxel spacing and values relating to locating the relevant files.

2. Lesion findings metadata

   The lesion findings are documented and relate to the image metadata. The data available here includes:

   ProxID, ***fid***, ***Pos*** (the same as in the "Image metadata"), ***zone*** – the zone of the prostate in which the lesion is located and the outcome variable (each zone is listed in Table 3) and,
   ClinSig – described as the 'truths' and takes the values of TRUE and FALSE.

   ClinSig is described in the metadata, as an "Identifier available in training set that identifies whether this is a clinically significant finding. Either the biopsy Gleason score was 7 or higher."

3. Screenshots around each lesion finding and are provided for reference to identify the correct slice. These are used as a reference point for ROI extraction.

At least one lesion was present in each patient and their locations are provided alongside their "*ClinSig*" truth. Over the 204 patients available in the PROSTATEx training data set, there were 330 lesion findings located in various prostate zones. Out of the 330 findings, 254 are not significant (*ClinSig* = FALSE, Gleason score of 6 or less) and 76 are significant (*ClinSig* = TRUE, Gleason score of 7 or higher). The range of ages and weights available in the PROSTATEx dataset are provided in Table 2, and the distribution of lesions found in each zone is provided in Table 3 as this data is also used in the study. Herein, where *ClinSig* is denoted as TRUE, it will be referred to as 'significant' or 'a significant tumour,' and vice versa.

The Gleason score system is based on histologic patterns at relatively low magnification by the extent of glandular differentiation and the pattern of growth of the tumour in the prostatic stroma. Multiple growth patterns were grouped into five grades. The original diagram of how the view of the pattern relates to which histologic grade is available in [25]. The Gleason score can range from 2 to 10, by adding the primary grade pattern – the pattern that is most predominant in the area - and the secondary grade pattern – the second most common pattern.

After data cleaning, based on what is available from each scan parameter, different numbers of lesion findings are used to create each model and therefore different amounts of non-significant and significant tumours are available. These are listed in the Results section of the paper.

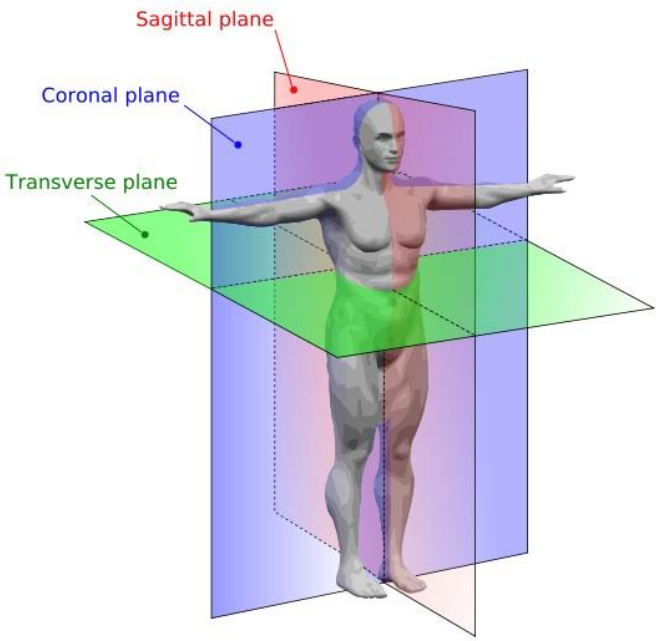

Figure 1: The three anatomical planes studied upon each patient in this work.

Image obtained from https://my-ms.org/mri_planes.htm

An example of the '**ijk**' value is provided. Table 1 shows an extract from the image metadata. Within the T2- weighted coronal ('t2_tse_cor0') parameter of MRI scan, for patient "ProstateX-0000," its first (and in this case, only) lesion is in column 139, row 166 of slice number 12. It is known from the lesion findings metadata if the finding is a clinically significant tumour or not. Figure 2 shows a representation of the area of the finding from the relevant slice in the MRI scan and how the ROI is extracted. The image/pixel level data of the box around the lesion will be taken. Similar to [8], a centred $5mm\ x\ 5mm$ patch is extracted at a resolution of 1px / mm, from the DICOM pixel level data. This is further described in the Methodology section of the report.

ProxID	**Name**	**fid**	**ijk**
ProstateX-0000	t2_tse_cor0	1	139 166 12

Table 1: ijk example

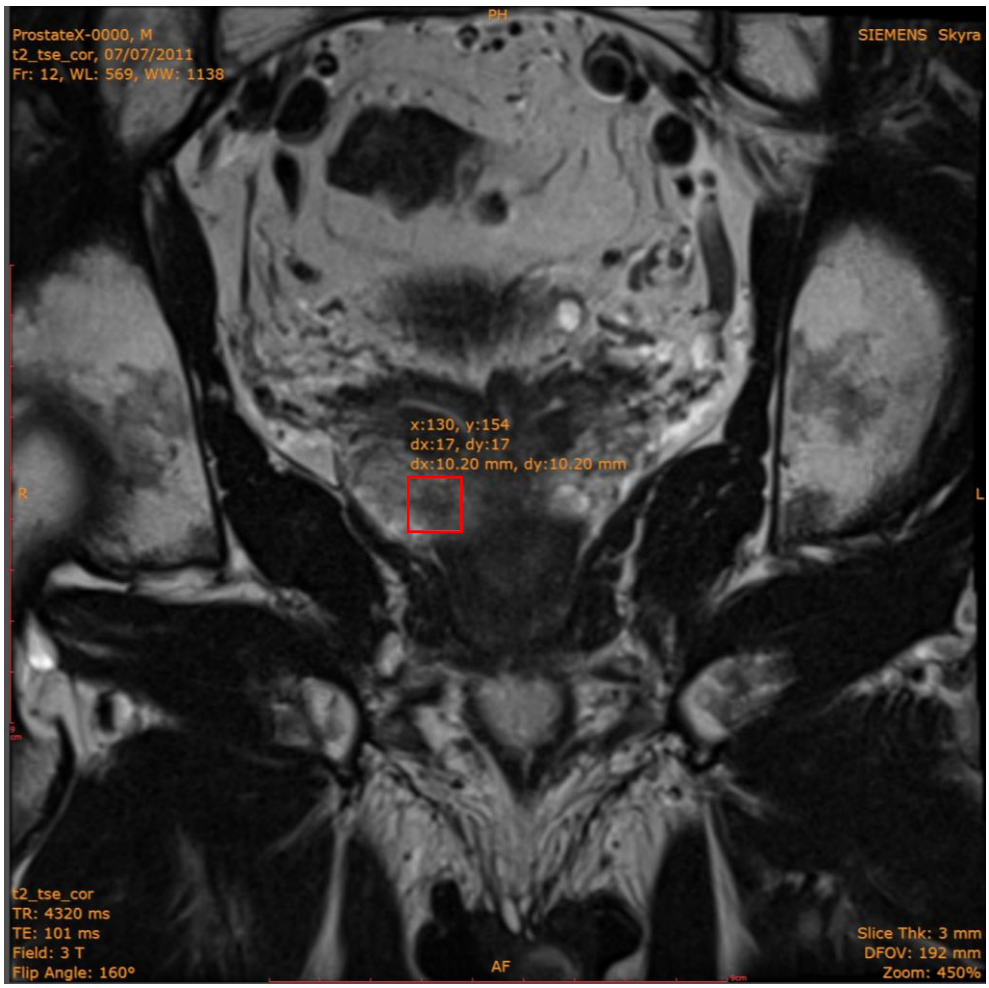

Figure 2: ijk example. This is slice 12 with the box around the position given from 'ij' of 'ijk' - in this case, around column 139 and row 166. [2] This shows how the patch is extracted

	Minimum	Maximum
Age (years)	37	78
Weight (kg)	60	120

Table 2: Distribution of ages and weights of patients in the PROSTATEx data set

Zone	TRUE	FALSE	Total
PZ: peripheral zone	36	155	191
TZ: transition zone	9	73	82
AS: anterior fibromuscular stroma	31	24	55
SV: seminal vesicles	0	2	2

Table 3: Distribution of zones of lesion findings in the PROSTATEx data set

[2] The MRI viewer used in Figure 2 is Sante DICOM Viewer. [36]

Clatterbridge Cancer Centre data

Data has been provided by the Clatterbridge Cancer Centre as part of this collaborative project. A combination of planning (diagnosis) mpMRI data and pre-treatment mpMRI data has been provided for several anonymised patients. For this project it will be shown how a located lesion within one of the patients and its relevant pixel-level data and any other required metadata could be applied to the best performing model, providing a real classification result for the given lesion – whether it is a clinically significant lesion, or not. The variety of mpMRI data held for the patients includes the data held in the PROSTATEx data set and therefore makes it applicable to use for testing purposes.

Methodology

This section describes and shows the proposed method of this machine learning process against the problem, showing the pre-processing stages and classification through SVM models, with a comparison against logistic regression models produces results to differentiate between significant and non-significant tumour tissue.

Implementation

Open source software has been used for this work, including R, oro.dicom [26], mlr [27] and kernlab [28]. All experimental tests are performed on a 64-bit Windows 10 PC with an Intel i7 processor (3.6 gigahertz) and 16 gigabyte random-access memory (RAM).

Pre-processing: Image equalisation, ROI extraction, normalisation and class weightings

Significant effort was required for the pre-processing of the PROSTATEx data set, rivalling the effort needed for implementing the machine learning models. Multiple rules were implemented on each patient to extract the most relevant and correct images from the large data set, due to various names for similar parameters of scan.

Equalisation of the images

Histogram equalisation of images flattens and enlarges the range of an images histogram through the remapping of the grey levels based on a probability density function of the image, improving its contrast. [29] An example is provided in Figure 3. The histogram of an image is a distribution of its discrete intensity levels in the range $[0, L-1]$ (where L is the number of possible intensity values in the image). The processing of histogram equalisation uses the cumulative probability function (cdf), which is a cumulative sum of all the probabilities lying in its domain and is defined by Equation 1. Through histogram equalisation, the resulting image gives a linear cdf. A linear cdf is associated to the uniform histogram that the resulting image should have. Implementing Equation 2 gives the new probability distribution function [30].

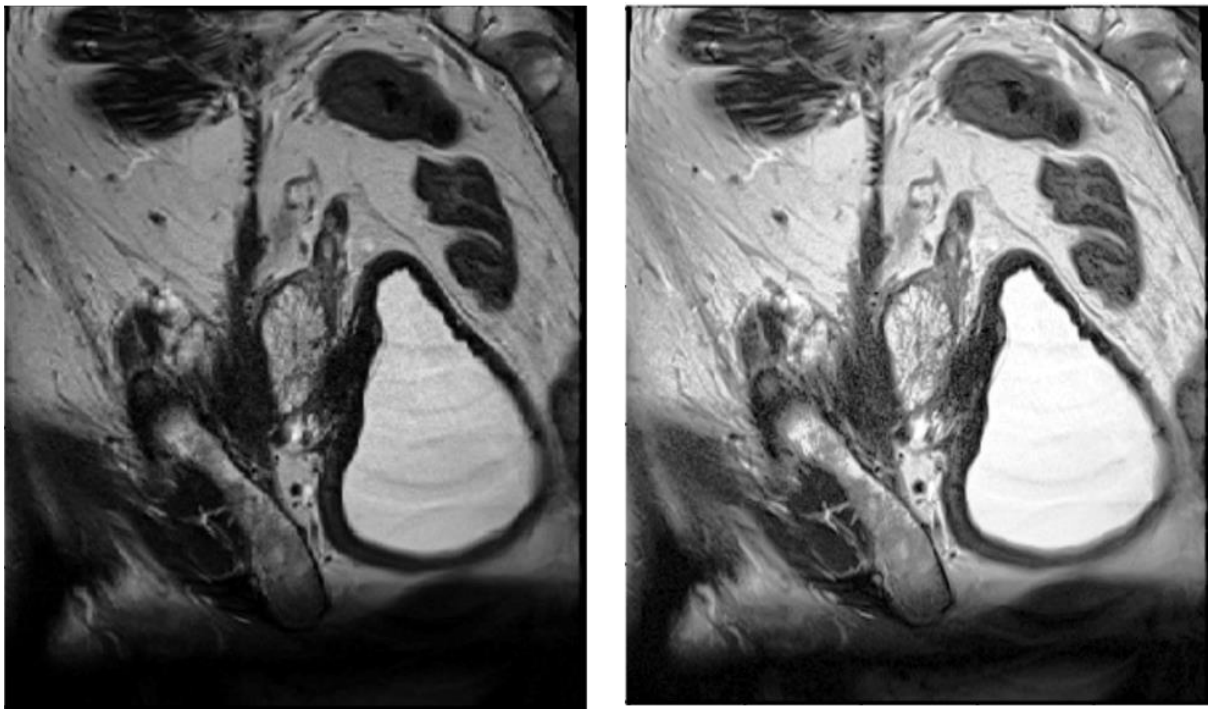

Figure 3: original image (left) and equalised image (right) through histogram equalisation

$$cdf(x) = \sum_{k=-\infty}^{x} P(k)$$

Equation 1: Cumulative distribution function (cdf) of all probabilities lying in the domain

$$S_k = (L-1)cdf(x)$$

Equation 2: Formula for new probability distribution function of equalised image

Issues with dataset/metadata

Within the metadata issues arose concerning the unique lesion finding ID's for some patients which were changed manually. This was confirmed through other sources that have used the challenge dataset [31]. The patient ID's for which the unique finding ID's were edited correctly, are 0005, 0025 and 0159 and were changed manually to be unique directly in the $.csv$ files. For some findings, the given location of the lesion was sometimes incorrect – such as the given column, row, or MRI slice number, were out of bounds. These findings were therefore ignored.

Extracting the ROI

For each findings ROI as denoted in the metadata, a centred $5\ pixel * 5\ pixel$ patch is extracted, for each parameter used in this study: T2 weighted sagittal, T2 weighted coronal, T2 weighted transverse and ADC transverse. The pixel-level data is flattened into a vector. Also combined into some of the tests, is the zone information (as a dummy variable) as well as the age and weight of the patient. The tests conducted are listed in Table 5.

Various combinations of data are evaluated and therefore multiple models were fitted and tested. Due to data cleaning, some scan parameters hold data for one finding, while another may not – this may be because of a mis-labelling in the metadata for the location of the lesion. For this reason, there may be less instances of data available for training different models. However, class weighting is amended for each test throughout and is listed in Table 5.

Normalisation

Each input dimension has the mean subtracted and is divided by the standard deviation, including the age and weight of the patient, allowing for the distribution of each dimension to be approximately normal. This was verified through plotting the histogram of the input dimensions.

Class weighting

Due to the imbalance of the class labels in the PROSTATEx data set – in the full dataset, 76 are positive (TRUE) and 254 are negative (FALSE), it is necessary to weight the classes accordingly, and this is used throughout combinations of all tests. Equation 3 shows the formula used for calculating the class weightings.

$$Class\ Weight = \frac{total\ number\ of\ examples}{(number\ of\ classes * number\ of\ examples\ in\ class)}$$

Equation 3: Class Weighting equation

For the case of the full dataset, the positive class is given a weighting of $\frac{330}{(2*76)} = 2.17$, and the negative class weighting is $\frac{330}{(2*254)} = 0.65$. This can be denoted such that it is approximately three times worse to incorrectly label a positive case than a negative case during training.

Methods: Support Vector Machines

Support Vector Machines (SVMs) are a form of supervised machine learning algorithms, which is a classifier whose function is trained based on pre-labelled data. They are used in this study for

classification of tumours based on their level of significance. First presented by Cortes & Vapnik in 1995 [4], it is a decision machine and does not provide posterior probabilities [32]. Moreover, this method optimises a margin and is known as a maximum margin classifier. Given a training data set, comprising of N input vectors $x_1, \ldots, x_N$ with corresponding target values $t_1, \ldots, t_N$ where $t_N \in \{-1, 1\}$ and new data points are classified accordingly by the sign of $y(\boldsymbol{x})$. For a linear model of the form shown in Equation 4 where $\emptyset(\boldsymbol{x})$ denotes a fixed feature-space transformation, $\boldsymbol{w}$ is the weight vector and b is the bias.

$$y(\boldsymbol{x}) = \boldsymbol{w}^T \emptyset(\boldsymbol{x}) + b$$

Equation 4: Linear model SVM.

If one assumes the training data set is linearly separable in feature space, there exists at least one choice of the parameters $\boldsymbol{w}$ and b such that the function of the form of Equation 3 satisfies $y(x_n) > 0$ for points having $t_n = 1$ and $(\boldsymbol{x}_n) < 0$ for points having $t_n = -1$, so that $t_n y(\boldsymbol{x}_n) > 0$ for all training data points.

There may be multiple solutions that separate the classes well – the support vector machine uses the concept of the margin – the smallest distance between the decision boundary and any of the samples. This is shown in Figure 4. The margin is defined as the perpendicular distance between the decision boundary and the closest of the data points (as shown on the left of Figure 4). The location of the boundary is determined by a subset of the data points, known as support vectors (as shown on the right of Figure 4) and should be as fat away from the data of both classes as possible. [32]

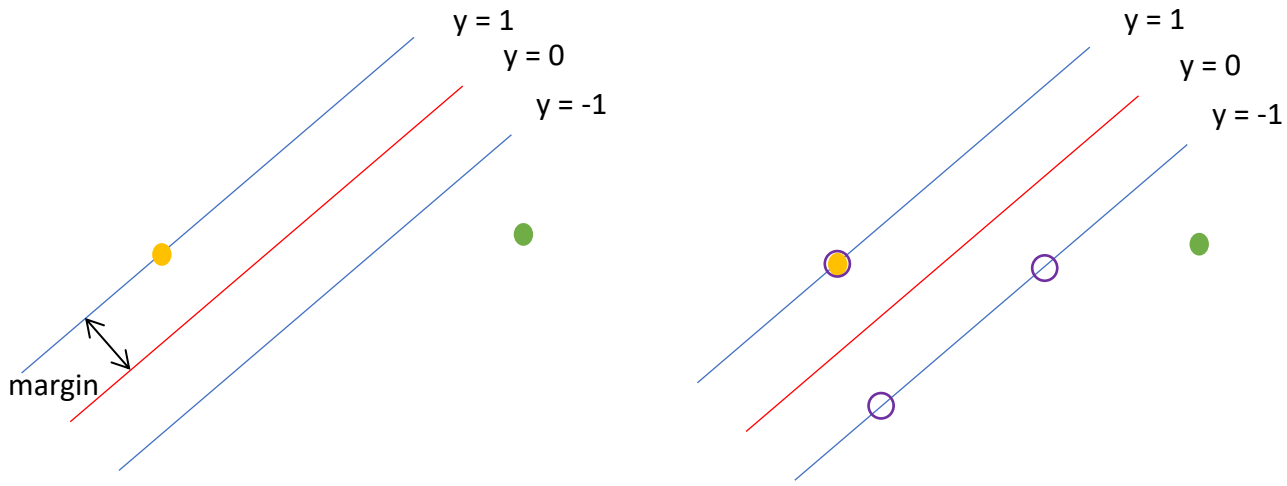

Figure 4: SVM decision boundaries

As the decision boundary should class should classify all points correctly, all data points will satisfy the constrains as in Equation 5, which is known as the canonical representation of the decision hyperplane. There will always be at least one active constraint where equality holds within the constraint, as there will always be a closest point. Once the margin has been optimised there will be at least two active constraints. This is an optimisation problem, which requires minimisation of $\|\boldsymbol{w}^2\|$, and therefore the optimisation problem in Equation 6 must be solved.

$$t_n(\boldsymbol{w}^T \emptyset(\boldsymbol{x}_n) + b) \geq 1 \quad n = 1, \ldots N$$

Equation 5: Canonical representation of the decision hyperplane.

$$\arg\min \frac{1}{2} \|\boldsymbol{w}^2\|; \; s.t. constraints \; in \; \text{Equation } 5$$

Equation 6: SVM optimisation problem: the factor of $\frac{1}{2}$ is included for convenience.

The constrained optimisation problem is solved through the introduction of Lagrange multipliers, $\alpha_n \geq 0$, with one multiplier α_n for each constraint in Equation 5, giving the Lagrangian function noted in Equation 7, where $a = (a_1, \dots, a_N)^T$. The negativity in front of the Lagrange multiplier notes that the minimisation is with respect to $\boldsymbol{w}$ and b, and maximisation is with respect to a. Setting the derivatives of $L(\boldsymbol{w}, b, \boldsymbol{a})$ with respect to $\boldsymbol{w}$ and b equal to zero, the two conditions in Equation 8 and Equation 9 are obtained.

$$L(\boldsymbol{w}, b, \boldsymbol{a}) = \frac{1}{2}\|\boldsymbol{w}^2\| - \sum_{n=1}^{N} \alpha_n\{t_n\{\boldsymbol{w}^T\emptyset(\boldsymbol{x}_n) + b) - 1\}$$

Equation 7: constrained optimisation with Lagrange multipliers for finding the decision boundary

$$\boldsymbol{w} = \sum_{n=1}^{N} a_n t_n \emptyset(\boldsymbol{x}_n)$$

Equation 8

$$0 = \sum_{n=1}^{N} a_n t_n$$

Equation 9

For classifying new data points with a trained model, the sign of $y(\boldsymbol{x})$ must be evaluated by Equation 4. This can be expressed in terms of the parameters a_n and the kernel functions by substituting for w, using Equation 8 to give Equation 10. A constrained optimisation of the form in Equation 10 satisfies a set of conditions, in this case require three properties as noted in Equation 11, Equation 12 and Equation 13, hold.

$$y(\boldsymbol{x}) = \sum_{n=1}^{N} a_n t_n k(\boldsymbol{x}, \boldsymbol{x}_n) + b$$

Equation 10: classifying new data points - evaluating the sign of $y(\boldsymbol{x})$

$$a_n \geq 0$$

Equation 11

$$t_n y(\boldsymbol{x}_n) - 1 \geq 0$$

Equation 12

$$a_n\{t_n y(\boldsymbol{x}_n) - 1\} = 0$$

Equation 13

Therefore, for every data point either $a_n = 0$ or $t_n y(\boldsymbol{x}_n) = 1$. For any data point in which $a_n = 0$ it will not appear in Equation 10 and does not affect the classification of new data points. The remaining data points are known as support vectors. As they satisfy $t_n y(x_n) = 1$, they therefore correspond to points that lie on the maximum margin hyperplanes in feature space, as shown in Figure 4. For this reason, nearly all data points can be discarded as only the support vectors are required to classify a new data point.

Kernel selection

The PROSTATEx data is not linearly separable and therefore a kernel trick is required. A kernel, $K(x)$, is based on the features, $\emptyset(\boldsymbol{x})$ and allows for a non-linear classifier. [33] Therefore, $K(x_i, x_j) =$

$\emptyset(\boldsymbol{x})^T\emptyset(\boldsymbol{x})$. Using a kernel function allows for avoidance of carrying out $\emptyset(\boldsymbol{x})$ explicitly; avoiding mapping the data to a high-dimensional feature space. This is because computing non-linear features does not scale well with the number of input features.

The kernel selected for this work is the radial basis function (RBF) and is called the RBF-SVM. This is a non-linear kernel, so increases/decreases in one dimension do not always cause a proportional change in the output. This is defined as shown in Equation 14. As the work in [8] is similar to this study, its use of an RBF kernel inspired its use here. Furthermore, a linear kernel was tested for this data and was dismissed early on, due to poor performance.

$$K(x, y) = e^{(-\frac{\|x-y\|^2}{2\sigma^2})}$$

Equation 14: Radial Basis Function kernel.

Hyperparameter selection

There are two hyperparameters that require tuning for an RBF-SVM. The first is known as the soft-margin constant, C. This is a trade-off parameter between the error and the margin. This influences the decision boundary – a smaller value of C allows for points to be ignored close to the boundary and increase the margin. However, for a larger value of C a large penalty is given to the errors/margin errors; the two points closest to the hyperplane affect its orientation giving a hyperplane that comes close to a lot of the other data points.

The kernel parameter, γ is the inverse of the standard deviation of the RBF kernel and is used as a similarity measure between two points. A smaller γ value would define a Gaussian function with a large variance, and therefore consider two points similar even if they are further away from each other. A larger γ value defines a Gaussian function with a smaller variance, so two points are considered similar just if they are closer to each other.

For each combination of hyperparameters a grid search is performed over the combinations listed in Table 4 .

$\boldsymbol{C} = \{\mathbf{0.1}, \mathbf{0.5}, \mathbf{1}, \mathbf{2}, \mathbf{5}, \mathbf{10}, \mathbf{20}, \mathbf{30}, \mathbf{50}, \mathbf{500}, \mathbf{750}, \mathbf{7500}\}$
$\boldsymbol{\gamma} = \{\mathbf{10^{-1}}, \mathbf{10^{-2}}, \mathbf{0.015}, \mathbf{0.02}, \mathbf{10^{-3}}, \mathbf{10^{-4}}, \mathbf{10^{-5}}\}$

Table 4: Hyperparameter tuning values for C and γ.

Methods: Logistic Regression

For comparison to the SVM models in this work, a logistic regression model has also been applied to each model. Binary Logistic Regression is a form of regression which allows for classification of discrete variable, when the relationship is non-linear. It is used to analyse the probability of occurrence of binary variables through fitting log odds and explanatory variables to a linear model as shown in Equation 15. The probability of occurrence is predicted by the logistic function as shown in Equation 16. For this study, the image patch pixel-level data and if required for the model fitting, the lesion zone information, patient age and patient weight are the explanatory variables and the binary variable is *ClinSig*.

$$\log\left(\frac{P(Y=1|X)}{1-P(Y=1|X)}\right) = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n$$

Equation 15: Fitting the log odds and explanatory variables to a linear model - Logistic Regression. Y = (0,1) is the binary variable, X = $(X_1, \ldots, X_n)$ are the n explanatory variables and $\beta = (\beta_0, \ldots, \beta_n)$ are the regression coefficients to be estimated based on the data.

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n)}}$$

Equation 16: Logistic function to predict probability of occurrence to assign a class label as a prediction.

Validation: 10-fold cross validation

K-fold cross validation is shown in Figure 5. K-1 of the groups are used to train a set of models, which are then evaluated on the remaining group. The procedure is then repeated for all K possible choices for the held-out group (fold), with the performance then averaged over all K runs. [32]

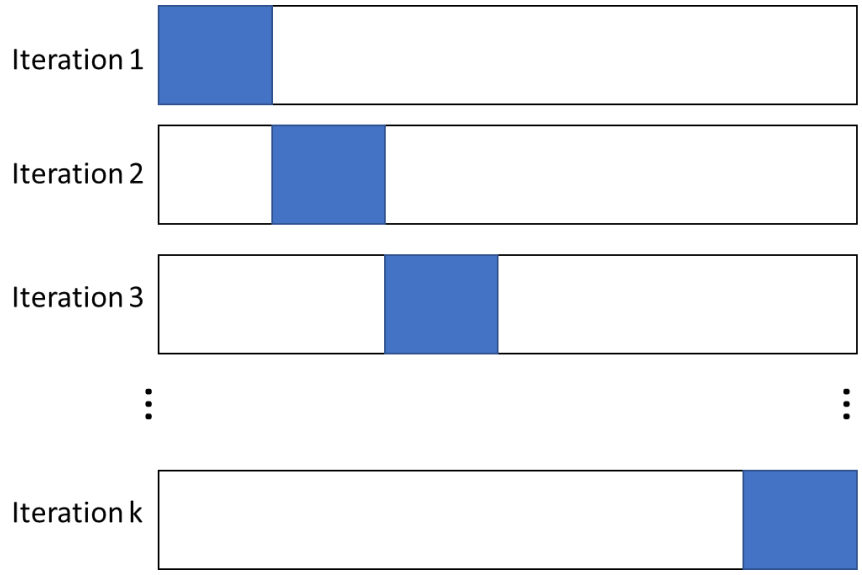

Figure 5: k-fold cross validation

Results

In this section, all experimental results have been compiled and presented. The objective of this study is to assess how well the SVM classifies various combinations of data – such as image patch data alone from various MRI scan parameters, and then combined with other metadata such as lesion zone location information, patient age and patient weight – against the level of clinical significance of the tumour.

The performance evaluation of the PROSTATEx challenge is the AUC measure and therefore the AUC measurements are reported here for continuity. Furthermore, the method for applying a model to data from the Clatterbridge Cancer Centre is presented.

Combinations of data used for models

Various combinations of data were tested for this study and are listed in Table 5. Due to data cleaning and availability of some parameters of scan, there are different amounts of data available however the class weights are adjusted accordingly and are also listed in Table 5 for completeness.

Model	Data and metadata used for the test	FALSE | TRUE	FALSE class weighting	TRUE class weighting
1	T2 transverse image patch, ADC transverse image patch	204 | 57	0.64	2.29
2	T2 transverse image patch, ADC transverse image patch, lesion zone	204 | 57	0.64	2.29
3	T2 transverse image patch, ADC transverse image patch, lesion zone, patient age, patient weight	204 | 57	0.64	2.29
4	T2 sagittal image patch, T2 coronal image patch, T2 transverse image patch	250 | 76	0.65	2.14
5	T2 sagittal image patch, T2 coronal image patch, T2 transverse image patch, lesion zone	250 | 76	0.65	2.14
6	T2 sagittal image patch, T2 coronal image patch, T2 transverse image patch, lesion zone, patient age, patient weight	250 | 76	0.65	2.14
7	T2 sagittal image patch, T2 coronal image patch, T2 transverse image patch, ADC transverse, lesion zone, patient age, patient weight	204 | 57	0.64	2.29

Table 5: Combinations of data and metadata for models fitted. Frequency of outcome class values and class weightings for SVM are also listed.

Analysis of results

A 10-fold cross validation was performed over a simple grid test of the parameters as listed in Table 4. For each run of the 10-fold cross validation, each combination of the hyperparameters were ran and the average AUC across all folds was calculated. This was completed three times for each model, with the highest AUC recorded over the grid searches along with its hyperparameters, C and γ. These are shown in Table 6 with the highest two values highlighted in red.

The 10-fold cross validation results indicate that model 2, with the combination of the T2 transverse and ADC transverse scan parameters in combination with the lesion zone information leads to the best performance which Is characterised by the highest AUC of all data combinations that were fitted of 0.796, with hyperparameters $C = 500\ and\ \gamma = 0.001$. The ROC plot is shown in Figure 6. The second highest AUC is achieved by model 3, which includes the data of the best performing model, as well as patient metadata (age and weight) at 0.746 with hyperparameters of $C = 5000\ and\ \gamma = 0.0001$. The ROC plot is shown in Figure 7.

The lowest SVM model AUC of 0.581 is achieved by the by model 4 which is a combination of the three T2 weighted image patches of lesion findings only. Including the lesion zone information in model 5 increases the AUC to 0.68, while also adding patient metadata in model 6 modestly increases the model performance to 0.705. This follows for the average AUC measures of the models also, from 0.57 to 0.68 to 0.70 respectively.

For model 1, the T2 transverse and ADC transverse image patches alone, the AUC is 0.64 for the SVM model. Including the lesion zone information increases the AUC to 0.759 in model 2 while also adding in the patient metadata also increases the AUC from just using the image patches, to 0.746 in model 3. The model performance also increases on average when extra data is included in the models. Using all information extracted for this study in model 7, decreases the model performance slightly compared to the top model to an AUC of 0.719.

For comparison, logistic regression models were fitted to the same data combinations as the SVM models. In all cases, SVM model fits performed better than logistic regression model fits. However, the higher AUC's were achieved by models 1, 2 and 3 compared to models 4, 5 and 6 respectively. This shows that there is better classification performance where transverse images are used. As a result, limiting the amount of data provided to the model has helped model performance.

	SVM		**Logistic Regression**	
Model	Best AUC	Average AUC	Best AUC	Average AUC
1	0.64	0.60	0.623	0.62
C, gamma	30, 0.1			
2	**0.796**	0.74	0.703	0.69
C, gamma	500, 0.001			
3	**0.746**	0.73	0.728	0.72
C, gamma	5000, 1e-4			
4	0.581	0.57	0.512	0.51
C, gamma	7500, 0.001			
5	0.68	0.68	0.614	0.61
C, gamma	10, 0.02			
6	0.705	0.70	0.598	0.58
C, gamma	10, 0.015			
7	0.719	0.70	0.627	0.60
C, gamma	7500, 1e-4			

Table 6: SVM and Logistic Regression has been performed on each data set combination test as referred to in Table 5. Each test was run three times, and the best AUC is listed, along with the average AUC across all three runs. The best AUC score is 0.796, SVM test 2.

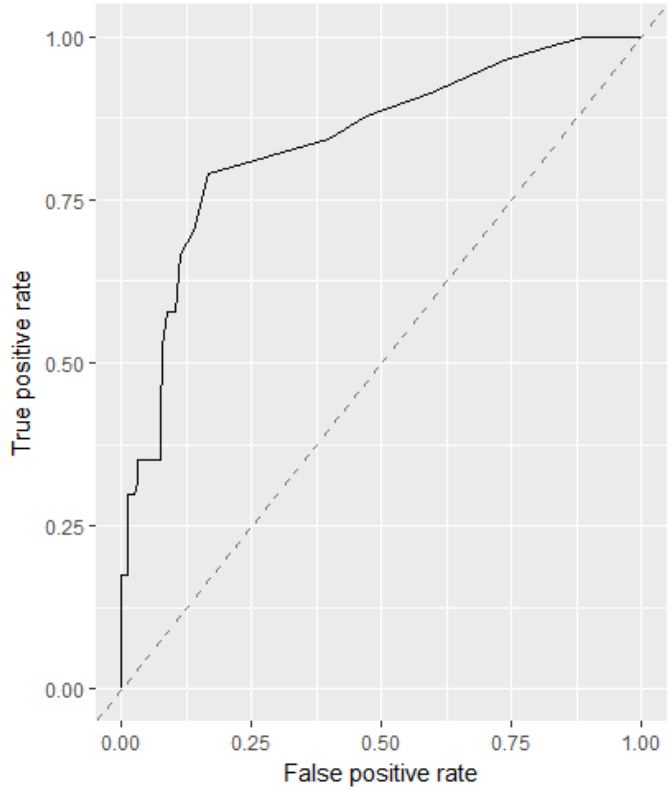

Figure 6: ROC plot for model 2: T2 transverse, ADC transverse and lesion zone. AUC 0.796.

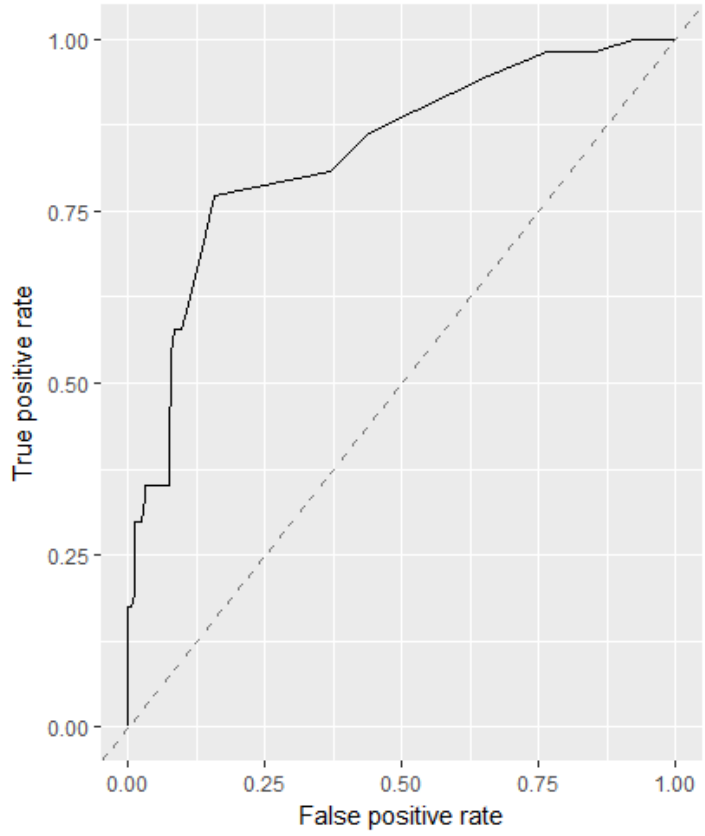

Figure 7: ROC plot for model 3: T2 transverse, ADC transverse, lesion zone and patient metadata (age and weight). AUC 0.746.

How the methodology may be applied to data from the Clatterbridge Cancer Centre

It is intended that this model will be applied onto data provided by the Clatterbridge Cancer Centre, NHS Foundation Trust. An example of how a model can be applied to the data is presented using some of the centres' anonymised patient data.

As various parameters of scan are provided including T2 transverse and ADC transverse, with professional labelling of a lesion within the prostate and the zone the lesion is present in, the model can provide information on the level of Gleason score as denoted by the PROSTATEx *ClinSig* variable. If the model predicts the *ClinSig* level of the area is "TRUE", then the area would be labelled as such and therefore that area of the lesion is predicted to have a Gleason score of 7 or higher, based on the *ClinSig* variable and vice versa.

An example is provided. [3] Figure 8 shows a box around the location of a lesion. The box shows the area that would be analysed, which would be split into sub-regions; in this case a $25 * 25\ pixel$ box was taken from each image used and split into $5 * 5\ pixel$ sub regions, all classified separately. This collection of sub-regions are classified under the classifier through the ClinSig variable, as either TRUE or FALSE. A colour-based representation is shown in the example in Figure 8. When viewed by a member of multidisciplinary healthcare staff, this could be used to aid in the diagnosis of prostate cancer.

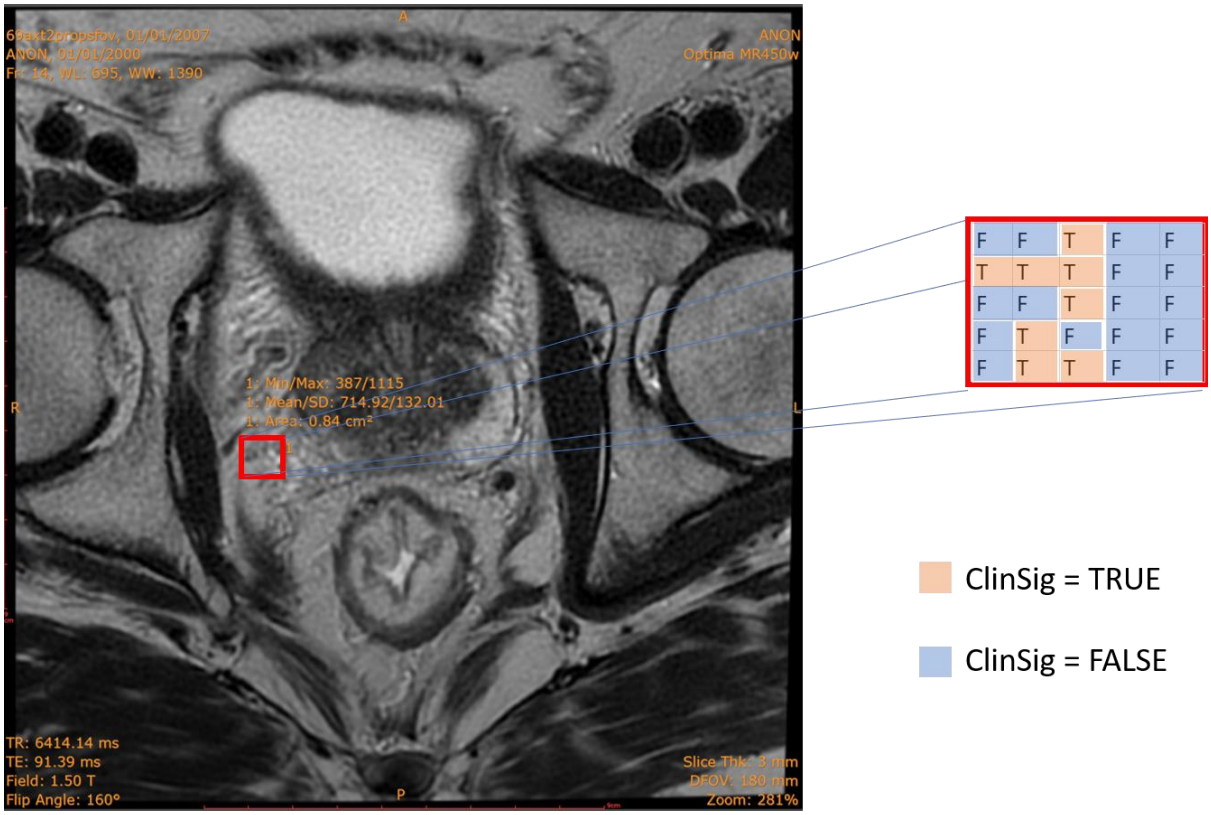

Figure 8: Application of the model onto an MRI image from the Clatterbridge Cancer Centre

[3] The image is from the Clatterbridge Cancer Centre, NHS Foundation Trust, however it is anonymised; no patient data is included.

Conclusions

The clinical diagnosis of prostate cancer is an important and delicate challenge for multidisciplinary healthcare teams in patient care. Decisions must be based on evidence gathered from non-invasive techniques. Though costly, the study of mpMRI proves effective in mitigating issues concerning other methods of diagnosis, such as biopsies which hold a low specificity.

In this study, the classification of a prostate tumour is conducted through fitting SVM models to various data combinations of the PROSTATEx challenge data set, with the classification based on the level of the Gleason grade of the lesion finding. This study provides evidence that a simple, flexible and computationally efficient model such as a trained SVM can provide strong classification results. The best performing model holds an AUC of 0.796, achieved with the T2-transversial and ADC-transversal image patch data combined with lesion location information. The SVM models have been benchmarked against logistic regression models, with the SVM consistently outperforming.

Although the AUC scores for most of the models are strong – including the highest AUC presented in this work of 0.796 as shown in Table 6, it does not score as highly as other models that entered the competition. High scoring entries of the PROSTATEx challenge obtained AUC's of 0.84 and higher. However, the highest scoring models used very complex models such as those with deep learning [12] or convolutional neural networks [34]. These methods are much more complex than the SVM models fitted in this study. Although these models fitted by SVMs may not be as accurate, it is the simplicity and flexibility of the model that provide their own benefits. This could allow for development in real-world environments, aiding multidisciplinary healthcare teams.

Further work will look at applying segmentation and delimitation of tumour borders in the prostate to further develop clinical tools for prostate cancer treatment, similar to [15].

A Self Evaluation of Project Outcome

This project was pitched as a segmentation project to start with, to produce colour maps similar to [15] by separation of sources. For this work, these sources were expected to be the significant tumour tissue, the non-significant tumour tissue, and healthy prostate tissue. However, these sources did not separate well and achieved poor performance of under 55% accuracy when attempting to re-label the sources after separation. This work will be developed further, possibly using different image enhancement techniques or different machine learning methods and has been pitched at PhD level. This will be the key objective of the study, which will likely require expert full border labelling of the tumour tissue – both significant and non-significant, similarly to the *ClinSig* variable in this study – and healthy tissue.

However, a diagnosis-assisting tool has still been produced. For the classification through SVM work presented, the result is pleasing and would be a useful add-on for a clinical diagnosis tool, especially as the model is much simpler than those submitted for the challenge. Its flexibility would allow ease of use and implementation in real world application, unlike other extensive and more sophisticated models such as in [12] and [34]. An AUC score of 0.796 is a strong score nonetheless and would be a strong aid in clinical diagnosis.

It would be interesting to test more scan parameters and assess the AUC of all combinations. This would likely be time consuming, however the abundance of MRI data available to clinicians has ramifications for being a key factor in assisting in the diagnosis of prostate cancer. Also, the PROSTATEx-2 challenge data set [35] has been made available. It would be interesting to apply SVMs similarly onto this data set to attain predictions of Gleason score groupings, on a classification system of a higher degree than just a binary classifier and if successful could provide a similar yet more powerful tool as the Gleason score could be better evaluated.

References

[1] "Prostate cancer - NHS.UK," *NHS UK*, 2015. [Online]. Available: https://www.nhs.uk/conditions/prostate-cancer/. [Accessed: 22-May-2018].

[2] BBC News, "Prostate cancer: Four in 10 cases diagnosed late, charity says - BBC News," 2018. [Online]. Available: http://www.bbc.co.uk/news/health-43669439. [Accessed: 22-May-2018].

[3] Orchid, "Orchid | Prostate Cancer," 2018. [Online]. Available: https://orchid-cancer.org.uk/prostate-cancer/. [Accessed: 17-Jun-2018].

[4] C. Cortes and V. Vapnik, "Support-Vector Networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[5] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A Training Algorithm for Optimal Margin Classiiers," in *Proceedings of the fifth annual workshop on Computational learning theory - COLT '92*, 1992, pp. 144–152.

[6] V. N. Vapnik and V. N., *The nature of statistical learning theory*. Springer, 1995.

[7] L. Bottou *et al.*, "Comparison of Classifier Methods: A Case Study in Handwritten Digit Recognition."

[8] A. Kitchen and J. Seah, "Support vector machines for prostate lesion classification," 2017, vol. 10134, p. 1013427.

[9] G. Litjens *et al.*, "Evaluation of prostate segmentation algorithms for MRI: The PROMISE12 challenge," *Med. Image Anal.*, vol. 18, no. 2, pp. 359–373, Feb. 2014.

[10] G.-F. Lin, G.-R. Chen, P.-Y. Huang, and Y.-C. Chou, "Support vector machine-based models for hourly reservoir inflow forecasting during typhoon-warning periods," *J. Hydrol.*, vol. 372, pp. 17–29, 2009.

[11] Dacheng Tao, Xiaoou Tang, Xuelong Li, and Xindong Wu, "Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 7, pp. 1088–1099, Jul. 2006.

[12] S. Liu, H. Zheng, Y. Feng, and W. Li, "Prostate Cancer Diagnosis using Deep Learning with 3D Multiparametric MRI," Mar. 2017.

[13] D. Fehr *et al.*, "Automatic classification of prostate cancer Gleason scores from multiparametric magnetic resonance images.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 46, pp. E6265-73, Nov. 2015.

[14] S. Ortega-Martorell *et al.*, "A Novel Semi-Supervised Methodology for Extracting Tumor Type-Specific MRS Sources in Human Brain Data," *PLoS One*, vol. 8, no. 12, p. e83773, Dec. 2013.

[15] S. Ortega-Martorell *et al.*, "Convex Non-Negative Matrix Factorization for Brain Tumor Delimitation from MRSI Data," *PLoS One*, vol. 7, no. 10, p. e47824, Oct. 2012.

[16] I. Reda *et al.*, "A comprehensive non-invasive framework for diagnosing prostate cancer," *Comput. Biol. Med.*, vol. 81, pp. 148–158, Feb. 2017.

[17] R. Toth, P. Tiwari, M. Rosen, A. Kalyanpur, S. Pungavkar, and A. Madabhushi, "A multi-modal prostate segmentation scheme by combining spectral clustering and active shape models," 2008, vol. 6914, p. 69144S.

[18] Xin Liu *et al.*, "Unsupervised segmentation of the prostate using MR images based on level set with a shape prior," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2009, pp. 3613–3616.

[19] X. Liu, M. A. Haider, and I. S. Yetik, “Unsupervised 3D Prostate Segmentation Based on Diffusion-Weighted Imaging MRI Using Active Contour Models with a Shape Prior,” *J. Electr. Comput. Eng.*, vol. 2011, pp. 1–11, Oct. 2011.

[20] S. D. Pathak, D. R. Haynor, and Y. Kim, “Edge-guided boundary delineation in prostate ultrasound images,” *IEEE Trans. Med. Imaging*, vol. 19, no. 12, pp. 1211–1219, 2000.

[21] “DICOM Standard.” [Online]. Available: https://www.dicomstandard.org/. [Accessed: 19-May-2018].

[22] G. Litjens, O. Debats, J. Barentsz, N. Karssemeijer, and H. Huisman, “Computer-Aided Detection of Prostate Cancer in MRI,” *IEEE Trans. Med. Imaging*, vol. 33, no. 5, pp. 1083–1092, May 2014.

[23] “SPIE-AAPM-NCI PROSTATEx Challenges - The Cancer Imaging Archive (TCIA) Public Access - Cancer Imaging Archive Wiki,” 2017. [Online]. Available: https://wiki.cancerimagingarchive.net/display/Public/SPIE-AAPM-NCI+PROSTATEx+Challenges#bff70c2399ab42098c34ed1abdf7bb77. [Accessed: 19-May-2018].

[24] D. R. Varma, “Managing DICOM images: Tips and tricks for the radiologist.,” *Indian J. Radiol. Imaging*, vol. 22, no. 1, pp. 4–13, Jan. 2012.

[25] P. A. Humphrey, “Gleason grading and prognostic factors in carcinoma of the prostate,” *Mod. Pathol.*, vol. 17, no. 3, pp. 292–306, Mar. 2004.

[26] B. Whitcher, V. J. Schmid, and A. Thornton, “Working with the DICOM and NIfTI Data Standards in R,” *J. Stat. Softw.*, vol. 44, no. 6, pp. 1–29, Oct. 2011.

[27] B. Bischl *et al.*, “mlr: Machine Learning in R,” *J. Mach. Learn. Res.*, vol. 17, no. 170, pp. 1–5, 2016.

[28] A. Karatzoglou, A. Smola, K. Hornik, and A. Zeileis, “kernlab - An S4 Package for Kernel Methods in R,” *J. Stat. Softw.*, vol. 11, no. 9, pp. 1–20, Nov. 2004.

[29] R. C. Gonzalez, R. E. Woods, L. Mcdowell, and J. Burger, *Digital Image Processing Second Edition Library of Congress Cataloging-in-Pubblication Data*, 2nd ed. 2001.

[30] A. Coste, “Image Processing.” [Online]. Available: http://www.sci.utah.edu/~acoste/uou/Image/project1/Arthur_COSTE_Project_1_report.html. [Accessed: 27-Aug-2018].

[31] Github; jspunda, “Data Set overview.” [Online]. Available: https://github.com/jspunda/prostatex/wiki/Data-set-overview. [Accessed: 01-Aug-2018].

[32] C. M. Bishop, *Pattern recognition and machine learning*, 1st ed. Cambridge: Springer, 2006.

[33] A. Ben-Hur and J. Weston, “A User’s Guide to Support Vector Machines.”

[34] A. Mehrtash *et al.*, “Classification of clinical significance of MRI prostate findings using 3D convolutional neural networks,” 2017, vol. 10134, p. 101342A.

[35] “PROSTATEx-2 Challenge.” [Online]. Available: https://www.aapm.org/GrandChallenge/PROSTATEx-2/. [Accessed: 27-Aug-2018].

[36] Santesoft, “Santesoft - Free DICOM Viewer. Advanced DICOM and PACS software.” [Online]. Available: https://www.santesoft.com/index.html. [Accessed: 27-Aug-2018].