Is it possible to detect cribriform adverse pathology in prostate cancer with magnetic resonance imaging machine learning-based radiomics?

Hüseyin Bıçakçıoğlu1, Sedat Soyupro1, Onur Ertunç2, Şehnaz Evirmler3, Tekin Ahmet Serel1,*

1 Department of Urology, School Medicine, Süleyman Demirel University, 32260 Isparta, Turkey
2 Department of Pathology, School Medicine, Süleyman Demirel University, 32260 Isparta, Turkey
3 Department of Radiology, School Medicine, Süleyman Demirel University, 32260 Isparta, Turkey
* Corresponding author: Serel Tekin Ahmet, tekinserel@sdu.edu.tr

Abstract: Rationale and objectives: Cribriform patterns are accepted as aggressive variants of prostate cancer. These adverse pathologies are closely associated with early biochemical recurrence, metastasis, castration resistance, and poor disease-related survival. A few publications exist to diagnose these two adverse pathologies with multiparametric magnetic resonance imaging (mpMRI). Most of these publications are retrospective and are not studies that have made a difference in diagnosing adverse pathology. It is also known that fusion biopsies taken from lesions detected in mpMRI are insufficient to detect these adverse pathologies. Our study aims to diagnose this adverse pathology using machine learning-based radiomics data from MR images. Materials and methods: A total of 88 patients who had pathology results indicating the presence of cribriform pattern and prostate adenocarcinoma underwent preoperative MRI examinations and radical prostatectomy. Manual slice-by-slice 3D volumetric segmentation was performed on all axial images. Data processing and machine learning analysis were conducted using Python 3.9.12 (Jupyter Notebook, Pycaret Library). Results: Two radiologists, SE and MAG, with 7 and 8 years of postgraduate experience, respectively, evaluated the images using the 3D-Slicer software without knowledge of the histopathological findings. One hundred seventeen radiomic tissue features were extracted from T1 weighted (T1W) and apparent diffusion coefficient (ADC) sequences for each patient. The interobserver agreement for these features was analyzed using the intraclass correlation coefficient (ICC). Features with excellent interobserver agreement (ICC > 0.90) were further analyzed for collinearity between predictors using Pearson’s correlation. Variables showing a very high correlation (r ≥ ±0.80) were disregarded. The selected features for T1W and ADC images were First-order maximum, First-order skewness, First-order 10th percentile for ADC, and Gray level size zone matrix, Large area low gray level emphasis for T1W. As a result of the classification of PyCaret, the three best models were found. A single model was obtained by blending these three models. AUC, accuracy, recall, precision, and F1 scores were 0.79, 0.77, 0.85, 0.82, and 0.83, respectively. Conclusion: ML-based MRI radiomics of prostate cancer can predict the cribriform pattern. This prognostic factor cannot be determined through qualitative radiological evaluation and may be overlooked in preoperative histopathological specimens.

Keywords: prostate cancer; cribriform pathology; MRI; radiomics; machine learning

1. Introduction

Cribriform pattern, among the adverse pathological features of prostate cancer, leads to a higher incidence of extraprostatic extension, lymph node involvement, metastasis, and biochemical recurrence compared to prostate adenocarcinoma. Due to these characteristics, the mortality rates associated with prostate cancer are higher.
in patients exhibiting these patterns. These lesions are a diverse group of prostate pathologies that can be challenging to diagnose. These lesions encompass a spectrum of conditions, including normal anatomical variations, benign proliferative lesions, premalignant conditions, suspicious findings, clearly malignant tumors, and aggressive entities. In recent years, our understanding of cribriform prostate adenocarcinoma (CrP4) and intraductal carcinoma of the prostate (IDC-P) has significantly advanced. There is now a growing body of evidence suggesting that the presence of these morphologies plays an important role in clinical decision-making for managing prostate cancer. Therefore, it is crucial to accurately recognize and report the architectural features of CrP4 and IDC-P [1–8].

Prostate multiparametric magnetic resonance imaging (mpMRI) is a valuable technique for detecting tumor lesions in the prostate. However, despite its usefulness, there are still cases where clinically significant lesions are undetected [9]. It has been concluded that mpMRI may accurately identify prostate cancer (PCa) tumors located in the peripheral zone (PZ). Additionally, the mean apparent diffusion coefficient (ADC) value and ADC ratio can serve as predictors for the presence of the cribriform pattern in PCa [10].

In recent years, there have been significant advancements in imaging technology and analysis methods, leading to the emergence of a framework called radiomics, which is a method of texture analysis that extracts imaging features from digital images by converting them into mineable, quantifiable data, revealing the heterogeneity of tumors. This framework involves extracting comprehensive high-dimensional features from imaging data and utilizing data mining techniques to develop analytical models that enhance decision support. Radiomics encompasses various features, including texture and shape, thereby providing rich information for precision medicine [11–13].

Numerous studies focusing on prostate radiomics have demonstrated promising outcomes in evaluating pathological characteristics, predicting treatment response, and stratifying risk groups. Radiomics features show promise as markers for assessing the aggressiveness of prostate cancer at both the histopathological and genomic levels. By correlating radiomics features with histopathological findings and genomic data, researchers are gaining insights into the aggressiveness and potential outcomes of prostate cancer. This has the potential to improve risk stratification, treatment selection, and personalized management for patients with prostate cancer [14,15].

The role of artificial intelligence in prostate cancer has been investigated because artificial intelligence (AI) may have the potential to revolutionize pathologic diagnosis and cancer patient management by serving as a predictive and prognostic biomarker. AI-based systems can analyze digitally scanned histopathology slides, enabling the differentiation between benign and malignant cells and low-grade and high-grade tumors. Deep learning models can process patient data from various sources, including individual or multimodal combinations, to identify patterns that can predict responses to different therapeutic options, assess the risk of recurrence or progression, and determine the prognosis for newly diagnosed patients. By leveraging AI-based models, treatment planning for prostate cancer patients can be significantly improved, and the efficiency and cost-effectiveness of pathology
laboratories can be enhanced [16–20]. However, to the best of our knowledge, there have been no studies in the English literature evaluating the predictive performance of ML-based MRI radiomics and the cribriform pattern for prostate cancer using multi-classifier models. Our study aims to assess the predictive performance of machine learning (ML)-based prostate cancer radiomics using Magnetic Resonance Imaging (MRI) to detect the cribriform pattern.

1.1. Patients

This retrospective study was approved by the institutional review board on 13 September 2022, with approval number 20/280. The data were collected from patients with prostatic carcinoma who underwent radical prostatectomy between January 2018 and November 2022. Only patients who underwent radical prostatectomy were included because the cribriform pattern can be present focally and may be missed in prostate biopsy specimens. A total of 88 patients who underwent preoperative MRI examination were identified. All patients had PIRADS 3 or above lesions. The cohort was divided into a training set (n = 70 subjects) and a test set (n = 18 subjects).

1.2. Imaging technique and radiomics analysis

All patients underwent standardized multiparametric MRI, which included multiplanar T1-weighted imaging (T1WI), axial diffusion-weighted imaging (DWI), and dynamic contrast-enhanced MRI (DCE-MRI) sequences at 3 Tesla using a phased array body coil. An endorectal coil was not used, and the imaging protocol adhered to the PI-RADS v2.1 standards. The MRI scans were performed on a MAGNETOM Avanto system from Siemens Medical Solutions in Erlangen, Germany. Apparent diffusion coefficient (ADC) maps were generated based on the DW images with all acquired b-values, and a calculated \( b = 1600 \text{ s/mm}^2 \) image was produced for PI-RADS classification. DCE-MRI was performed using intravenous gadobutrol (Gadavist, Bayer) contrast agent at a dose of 0.1 mL/kg body weight (0.1 mmol/kg), infused at a rate of 2 mL/s, with imaging initiated simultaneously with the administration of contrast.

Two radiologists, SE and MAG, with 7 and 8 years of post-graduate experience, evaluated the images using the 3D-Slicer software (version 4.11) [21]. They were blinded to the histopathological findings. DICOM images of the axial T1-weighted acquisition were obtained from the PACS system. The texture features, including first-order and second-order features, were extracted using Slicer-Radiomics (PyRadiomics v3.0.1) software (Figure 1). Before radiomics processing, the T1-weighted signals within the lesion volumes of interest (VOIs) were normalized to a range between the mean value plus three times the standard deviation of the signal intensity within the VOI, as previously described [22,23]. The spatial resampling voxel size was set to \( 1 \times 1 \times 1 \text{ mm} \), and gray-level discretization was performed using a fixed bin width of 64 gray levels.
The pure type of cribriform pattern of prostate cancer in H&E section (200X magnification)

**Figure 1.** MRI-based radiomics workflow and the Hematoxylin Eosin (H&E) section are examples of cribriform patterns in prostate cancer.

GLCM, Gray-level co-occurrence matrix; GLDM, Gray-level difference method; GLRLM, Gray-level run-length matrix; GLSZM, Graylevel size zone; NGTDM, Neighbouring Gray Tone Difference Matrix.

### 1.3. Statistical evaluation

Statistical analysis was conducted using IBM SPSS Statistics software (version 24.0, IBM Corp., Armonk, NY). Descriptive statistics were reported as frequencies for categorical variables and mean ± standard deviation or median for numerical variables. The normality of continuous variables was assessed using the Kolmogorov-Smirnov test. A p-value less than 0.05 was considered statistically significant.

### 1.4. Dimension reduction

The interobserver agreement for these features was assessed using the intraclass correlation coefficient (ICC). Features that demonstrated excellent interobserver
agreement (ICC > 0.90) were further examined for collinearity between predictors using Pearson’s correlation. Variables that exhibited a very high correlation ($r \geq \pm 0.80$) were excluded from the analysis. The selected features for T1-weighted (T1W) and apparent diffusion coefficient (ADC) images were first-order Maximum, first-order Skewness, and first-order 10th Percentile for ADC, as well as Gray level zone matrix, large area low gray level emphasis for T1W.

1.5. Data handling and machine learning analysis

Data handling and machine learning analysis were conducted using Python 2.3 in Jupyter Notebook with the Pycaret library. [24]. The “classic method” random forest classifier was employed to select and reduce features [25]. The interobserver agreement for these features was assessed using the intraclass correlation coefficient (ICC). Features demonstrating excellent interobserver agreement (ICC > 0.90) were further examined for collinearity between predictors using Pearson’s correlation. The tune model function in Pycaret classification was utilized to identify the optimal hyperparameters. This function tunes the model’s hyperparameters and generates a scoring grid with cross-validated scores by fold. The best model is selected based on the defined metric in the optimized parameter. The get_metrics function allows access to the metrics evaluated during cross-validation, and custom metrics can be added or removed using the add_metric and remove_metric functions.

The data was split into a training set (70%) and a test set (30%). Z-score scaling was applied for normalization, and 5-fold cross-validation was performed to avoid overfitting due to limited dataset size. The synthetic minority oversampling technique was used to address the issue of imbalanced datasets.

Fourteen machine learning algorithms were employed to determine significant models. The predictive performance of these algorithms was compared using metrics such as AUC, accuracy, recall, precision, and F1 scores. The top three models based on accuracy and AUC were evaluated on the test set. To optimize results, these three models were combined into a single ensemble model, leveraging the strengths of each model to improve overall performance and predictive accuracy. A receiver operating characteristic (ROC) and learning curves were plotted. AUC, accuracy, recall, precision, and F1 scores were provided along with the confusion matrix.

2. Results

The patients had a mean age of 66.05. Among them, there were 49 cases (55%) of pure prostatic adenocarcinoma and 39 cases (45%) of cribriform pattern prostatic adenocarcinoma, all of which were tumors. A total of 117 texture features were extracted. The 14 machine learning algorithms demonstrated varying AUC and accuracy values on the train set, ranging from 0.50% to 0.76% and 31% to 72%, respectively. The top three models identified were the AdaBoost classifier, with an AUC of 0.76 and an accuracy of 72%; the random forest classifier, with an AUC of 0.80 and an accuracy of 71%; and the extra trees classifier, with an AUC of 0.78 and an accuracy of 71% (Table 1). The voting classifier, which combines these models, yielded the following scores: AUC of 0.77, accuracy of 0.79, recall of 0.85, precision of 0.82, and F1 score of 0.83 (Table 2). The confusion matrix and classification
report, along with the receiver operating curve (ROC), illustrating the predictive performance of the blended model, are presented in Figure 2. Figure 3 shows the decision boundary and hyperparameters of the blended model.

Table 1. The predictive performances of machine learning based MRI radiomics models for cribriform pattern of prostate cancer.

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
<th>AUC</th>
<th>Recall</th>
<th>Prec.</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Forest Classifier</td>
<td>0.7143</td>
<td>0.8061</td>
<td>0.8155</td>
<td>0.7906</td>
<td>0.7978</td>
</tr>
<tr>
<td>Extra Trees Classifier</td>
<td>0.7145</td>
<td>0.7864</td>
<td>0.8156</td>
<td>0.7916</td>
<td>0.7964</td>
</tr>
<tr>
<td>Gradient Boosting Classifier</td>
<td>0.6429</td>
<td>0.7609</td>
<td>0.6889</td>
<td>0.7912</td>
<td>0.7106</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>0.6714</td>
<td>0.7639</td>
<td>0.7111</td>
<td>0.8076</td>
<td>0.7591</td>
</tr>
<tr>
<td>Ada Boost Classifier</td>
<td>0.7286</td>
<td>0.7622</td>
<td>0.7033</td>
<td>0.8178</td>
<td>0.7699</td>
</tr>
<tr>
<td>Light Gradient Boosting Machine</td>
<td>0.7143</td>
<td>0.7417</td>
<td>0.7489</td>
<td>0.8256</td>
<td>0.7837</td>
</tr>
<tr>
<td>K Neighbors Classifier</td>
<td>0.6714</td>
<td>0.8075</td>
<td>0.6887</td>
<td>0.8363</td>
<td>0.7291</td>
</tr>
<tr>
<td>Linear Discriminant Analysis</td>
<td>0.7000</td>
<td>0.6522</td>
<td>0.7733</td>
<td>0.8061</td>
<td>0.7382</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>0.6514</td>
<td>0.6428</td>
<td>0.6467</td>
<td>0.8181</td>
<td>0.6143</td>
</tr>
<tr>
<td>Decision Tree Classifier</td>
<td>0.6714</td>
<td>0.6006</td>
<td>0.7911</td>
<td>0.7386</td>
<td>0.7615</td>
</tr>
<tr>
<td>Quadratic Discriminant Analysis</td>
<td>0.5857</td>
<td>0.5872</td>
<td>0.6057</td>
<td>0.7086</td>
<td>0.6547</td>
</tr>
<tr>
<td>Dummy Classifier</td>
<td>0.5413</td>
<td>0.5000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>SVM-Linear Kernel</td>
<td>0.5000</td>
<td>0.0000</td>
<td>0.6179</td>
<td>0.4945</td>
<td>0.5258</td>
</tr>
<tr>
<td>Ridge Classifier</td>
<td>0.7000</td>
<td>0.0000</td>
<td>0.7733</td>
<td>0.8051</td>
<td>0.7832</td>
</tr>
</tbody>
</table>

Table 2. The predictive performances of blended machine learning models (random forest, extra trees, Ada boost) for cribriform pattern of prostate cancer.

<table>
<thead>
<tr>
<th>Fold</th>
<th>Accuracy</th>
<th>AUC</th>
<th>Recall</th>
<th>Prec.</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.7857</td>
<td>0.7760</td>
<td>0.9000</td>
<td>0.8182</td>
<td>0.8571</td>
</tr>
<tr>
<td>1</td>
<td>0.8571</td>
<td>0.9000</td>
<td>0.8000</td>
<td>1.0000</td>
<td>0.8889</td>
</tr>
<tr>
<td>2</td>
<td>0.7143</td>
<td>0.7200</td>
<td>0.8000</td>
<td>0.8000</td>
<td>0.8000</td>
</tr>
<tr>
<td>3</td>
<td>0.6429</td>
<td>0.6887</td>
<td>0.7778</td>
<td>0.0000</td>
<td>0.7368</td>
</tr>
<tr>
<td>4</td>
<td>0.8571</td>
<td>0.9556</td>
<td>1.0000</td>
<td>0.8182</td>
<td>0.9000</td>
</tr>
<tr>
<td>Mean</td>
<td>0.7714</td>
<td>0.7994</td>
<td>0.8556</td>
<td>0.6273</td>
<td>0.8056</td>
</tr>
<tr>
<td>Std</td>
<td>0.0833</td>
<td>0.1119</td>
<td>0.0837</td>
<td>0.5089</td>
<td>0.0608</td>
</tr>
</tbody>
</table>
Figure 2. Plot graphics of blended three machine learning classifiers: (a) confusion matrix; (b) learning curve; (c) classification report; (d) receiver operating characteristic curve; (e) Kolmogorov Smirnov statistic plot.
Figure 3. Plot graphic and hyperparameters of blended three machine learning classifiers: (a) decision boundary; (b) hyperparameters of blended models with Pycaret classification.

3. Discussion

This study assessed the predictive performance of machine learning-based MR radiomics for prostatic adenocarcinoma, focusing on the cribriform pattern. The machine learning algorithms exhibited high rates of AUC and accuracy. The voting classifier model, in particular, demonstrated excellent accuracy in predicting the cribriform status.

It is worth noting that the cribriform pattern has been associated with a poorer prognosis, characterized by lower survival rates and a higher prevalence of advanced stage, positive surgical margin, lymph node metastasis, lymphovascular invasion, and recurrence. Therefore, it has been suggested that the cribriform pattern could be utilized for risk stratification and guiding therapy management [4,26–29].

Early diagnosis in patients with prostate cancer having cribriform pathology is of critical importance in selecting the most appropriate patient management. Detecting the cribriform pattern in prostate biopsies beforehand is challenging. Thus, risk stratification to better identify patients at risk is necessary to identify low-grade prostate cancer that needs active surveillance. This difficulty has been attempted to be overcome through imaging techniques.

In a study using PSMA PET/CT, the incidence of the cribriform pattern in patients with prostate cancer was reported to be 39% [30]. Truong et al. published their study, which compares the results of fusion biopsies using MRI combined ultrasound fusion with the specimens obtained after radical prostatectomy; they demonstrated that only 17.3% of the cribriform pattern could be detected using MRI [31]. These studies indicate that even with targeted biopsies using MRI, the sensitivity for detecting patients with cribriform architecture is not high. Therefore, there is a need for advanced imaging techniques that can detect the lesion with higher accuracy. To our knowledge, this is the first report that has investigated the potential role and the predictive performance of machine learning-based MRI radiomics for detecting the cribriform pattern in prostate cancer in the existing English literature.

Our study exposed that the best three models for predicting the cribriform
pattern in prostate cancer were the AdaBoost, random forest, and extra trees classifiers. These models demonstrated high AUC and accuracy rates, indicating their predictive solid performance. It is important to note that increasing the number of training samples likely improved the generalization of these models. Therefore, we recommend conducting further studies with a larger patient population to enhance and validate our results.

This study has several limitations that should be acknowledged. Firstly, it is essential to note that the study design was retrospective, which inherently carries the risk of data loss and potential biases. Secondly, the patient population included in the study was relatively small. We specifically focused on the histopathological findings of radical prostatectomy specimens to ensure the presence of the cribriform pattern, as tumor heterogeneity could lead to its absence in prostate biopsy specimens. However, a larger patient population would be beneficial for further validation and generalization of the results. To address these limitations, we recommend conducting a prospective, multi-center study that evaluates the predictive value of machine learning models for the cribriform pattern and its impact on prognosis. Such a study would provide more robust and reliable evidence for this purpose.

4. Conclusion

Qualitative radiological examinations alone are insufficient for accurately determining the presence of the cribriform pattern in prostate cancer lesions. Additionally, this pattern can be missed on preoperative histopathological specimens due to the inherent heterogeneity of the lesion. Therefore, there is a clear need for an objective method to identify and predict the cribriform pattern reliably. In our study, we propose that machine learning-based MRI radiomics of prostate cancer can effectively predict the presence of the cribriform pattern. By utilizing advanced algorithms and analyzing a wide range of imaging features, this approach has the potential to provide valuable guidance to clinicians in making informed decisions regarding therapy management. It offers an objective and quantifiable method to assist in accurately identifying and characterizing the cribriform pattern, ultimately improving patient care and outcomes.

Author contributions: Conceptualization, TAS and SS; methodology, ŞE; software, TAS; validation, TAS, HB and ŞE; formal analysis, HB; investigation, SS; resources, TAS and OE; data curation, TAS; writing—original draft preparation, OE; writing—review and editing, TAS; visualization, TAS; supervision, OE; project administration, HB. All authors have read and agreed to the published version of the manuscript.

Conflict of interest: The authors declare no conflict of interest.

References


