Journal of Information Technology and Computing
https://jitc.sabapub.com
ISSN: 2709-5916

2025 Volume 5, Issue 1 : 41 – 53
DOI : 10.48185/jitc.v5i1.2185

VGG-ViT: A Hybrid Deep Feature Framework for Zone-Specific Prostate Cancer Classification

Nasser M. Al-Zidi *, D. Vasumathi

Department of Computer Science and Engineering, University College of Engineering, Science & Technology Hyderabad, Jawaharlal Nehru Technological University Hyderabad, Hyderabad, India

Received: 03.06.2025 • Accepted: 21.08.2025 • Published: 14.09.2025 • Final Version: 14.09.2025

Abstract: Prostate cancer (PCa) remains a significant health concern as one of the leading causes of cancer-related mortality in men globally. The accurate differentiation between clinically significant (CS) and clinically insignificant (CiS) PCa plays a crucial role in treatment planning and patient outcomes. This research introduces an innovative framework that leverages multiparametric magnetic resonance imaging (mpMRI) for PCa classification in both the entire prostate and specific anatomical zones. Our approach combines transfer learning techniques using pre-trained VGG19 and Vision Transformer (ViT) models for feature extraction, coupled with Support Vector Machine (SVM) classification. Through exhaustive analysis of possible combinations from nine mpMRI sequences, the framework identifies optimal MRI sequences for both whole-prostate and zone-specific analysis. A thorough investigation of sequence importance is conducted by analyzing their frequency in top-performing combinations, revealing distinct patterns across different anatomical regions. The framework achieves superior classification performance with an Area Under the Curve (AUC) of 0.87 for whole-prostate assessment and notably higher performance in zone-specific assessment, with the transition zone (TZ) yielding the best performance among all analyzed zones at an AUC of 0.96. The analysis reveals the particular significance of high b-value diffusion-weighted imaging (DWI) sequences, especially BVAL (b=1400 s/mm²), followed by Ktrans imaging, in enhancing classification accuracy. These comprehensive findings provide valuable insights for optimizing zone-specific mpMRI protocols in clinical practice.

Keywords: Prostate cancer, Multiparametric MRI, Deep feature, Anatomical zones, Classification

1. Introduction

Prostate cancer (PCa) stands as one of the primary contributors to cancer mortality in men globally [1]. The accurate differentiation between clinically significant (CS) and clinically insignificant (CiS) PCa plays a vital role in treatment planning and patient outcomes. Multiparametric magnetic resonance imaging (mpMRI) has emerged as an essential non-invasive diagnostic tool for detecting and characterizing PCa lesions [2], [3]. This imaging technique combines multiple sequences, including T2-weighted (T2W), dynamic contrast-enhanced (DCE), and diffusion-weighted imaging (DWI) [4], [5], providing complementary information about prostate tissue that enables more precise identification and assessment of PCa lesions [6], [7]. However, the interpretation of mpMRI presents challenges due to variations in image appearance across patients and modalities, resulting in moderate reproducibility even among experienced radiologists [8].

* Corresponding Author: alzidi.nasser@gmail.com

Recent advances in computer-aided diagnosis (CAD) systems incorporating machine learning (ML) and deep learning (DL) techniques have shown promising potential in supporting radiologists with PCa classification using mpMRI [9], [10], [11]. Convolutional neural networks (CNNs) have become widely adopted for feature extraction and classification in this domain [12], [13], [14]. Transfer learning techniques, particularly those utilizing pre-trained models such as VGG19, have successfully addressed the challenge of limited labeled medical image availability [15], [16]. More recently, vision transformers (ViTs) have gained prominence for their capability to capture global context and long-range dependencies in images [17], [18]. Hybrid approaches combining CNNs and ViTs have demonstrated enhanced performance across various image classification tasks [19].

While existing research has explored various ML and DL approaches for PCa classification using mpMRI, there remains a significant gap in understanding the optimal combination of mpMRI sequences for this purpose. Additionally, the relative importance of individual mpMRI sequences in PCa classification has not been thoroughly investigated. Furthermore, the consideration of specific anatomical zones within the prostate—including the peripheral zone (PZ), transition zone (TZ), and anterior fibromuscular stroma (AS)—is essential for accurate PCa diagnosis [20], [21]. Different mpMRI sequences may exhibit varying levels of diagnostic value across these anatomical zones [22].

To address these research gaps, we introduce an innovative framework designed to identify optimal combinations of mpMRI sequences for PCa classification, considering both the entire prostate and specific anatomical zones. Our approach combines transfer learning utilizing pre-trained VGG19 and ViT models for feature extraction, followed by feature fusion and support vector machine (SVM) classification. The framework conducts an exhaustive evaluation of all possible feature combinations extracted from nine mpMRI sequences and employs comprehensive evaluation metrics to determine the most effective combinations. Furthermore, we investigate the significance of each mpMRI sequence by analyzing its frequency in top-performing combinations.

The key contributions of this research include:

- Development of a framework that identifies optimal mpMRI sequence combinations for prostate cancer classification across both the entire prostate and specific anatomical zones.
- Implementation of a hybrid approach utilizing transfer learning with pre-trained VGG19 and ViT models for comprehensive feature extraction.
- Comprehensive evaluation of all possible mpMRI sequence combinations using robust evaluation metrics.
- In-depth analysis of individual mpMRI sequence importance through frequency analysis in top-performing combinations.

The remainder of this paper is organized as follows: Section 2 presents a review of related works in PCa classification using mpMRI and ML/DL techniques. Section 3 details our proposed methodology, including dataset description, preprocessing steps, feature extraction using VGG19 and ViT models, feature fusion, SVM classification, and the process of identifying optimal mpMRI sequence combinations. Section 4 presents and discusses our experimental findings, and Section 5 concludes the paper.

2. Related Works

The field of prostate cancer classification using multiparametric MRI (mpMRI) has experienced significant advancement in recent years. Various machine learning and deep learning approaches have been proposed to enhance the accuracy and efficiency of PCa diagnosis. Here, we review the most significant contributions to this field.

Multiple studies have explored the extraction of features from mpMRI sequences combined with traditional ML classifiers for PCa classification. Trigui et al. [23] developed a supervised classification approach utilizing SVM and random forests, extracting features from MRSI, DWI, DCE-MRI, and T2W sequences. Their SVM implementation achieved a remarkable 0.99% error rate in distinguishing between healthy and pathologic voxels. Kwon et al. [24] introduced a radiomics-based approach for the PROSTATEx challenge, extracting 216 intensity and texture features from T2W, PDW, DCE-MRI, and DWI sequences. Their random forest implementation achieved an AUC score of 0.82. Kitchen and Seah [25] applied SVM classification to the ProstateX challenge, also achieving an AUC of 0.82.

Deep learning approaches, particularly CNNs, have demonstrated promising results in PCa classification. Song et al. [26] introduced an automated deep CNN approach using a VGG-Net-inspired patch-based model with three branches processing T2W transaxial, sagittal, and ADC sequences. Their implementation achieved impressive metrics: 0.944 AUC, 0.87 sensitivity, and 0.906 specificity. Research presented in [27] and [28] explored transfer learning methodologies on mpMRI using pre-trained CNNs for feature extraction from T2W and ADC sequences. These approaches achieved accuracies of 86.92% and 88.89% respectively, outperforming conventional ML and DL models. Liu and An [29] developed a CNN model utilizing transfer learning from a pre-trained model for PCa classification on DWI images, achieving a testing accuracy of 78.15%.

Recent research has begun exploring the potential of vision transformers (ViTs) for PCa classification. Pachetti and Colantonio [30] proposed a stacking ensemble of 3D ViTs trained from scratch on T2W images from the ProstateX-2 dataset, achieving an AUC of 0.89. In their follow-up study [31], they compared 3D ViTs to 3D CNNs, with their optimal 3D ViT configuration achieving a mean AUC of 0.775 and mean F2-score of 0.523, surpassing 3D CNN baseline performance. Santhirasekaram et al. [32] introduced an innovative hybrid CNN/transformer architecture using the PROSTATEx dataset, achieving an AUC of 0.95 through 5-fold cross-validation and 0.94 on the test set, outperforming both CNN-only and radiomics approaches.

Some researchers have focused specifically on zone-specific PCa classification. Wang and Wang [33] proposed a methodology using the PROSTATEx dataset to classify lesions and identify optimal mpMRI sequence combinations, achieving AUC scores of 0.891 for PZ and 0.965 for TZ. Sobecki et al. [34] developed an approach to differentiate the clinical significance of prostate abnormalities based on lesion location, achieving an AUC of 0.92 on 160 patients from the PROSTATEx dataset.

While these studies have made valuable contributions to PCa classification using mpMRI, they face certain limitations including restricted dataset sizes, combining different lesion types, limited feature diversity, inadequate feature fusion techniques, insufficient consideration of zonal information, and incomplete utilization of available mpMRI sequences. Our proposed framework addresses these limitations through comprehensive sequence combination optimization and anatomical zone-specific analysis.

3. Materials and Methods

3.1. Dataset

Our framework utilizes the PROSTATEx Challenge dataset [35], which was released in conjunction with the SPIE-AAPM-NCI Prostate MR Classification Challenge. This comprehensive dataset includes mpMRI scans from 346 patients, encompassing both CS and CiS PCa lesions. The imaging data was acquired using two 3T MRI systems: Siemens MAGNETOM Trio and Skyra scanners. Detailed dataset specifications and imaging parameters are accessible through the challenge website (https://prostatex.grand-challenge.org) [36]. As demonstrated in Table 1, the dataset maintains a

distribution of lesions across different anatomical zones in both training and test sets, ensuring robust evaluation of our framework's performance.

The dataset incorporates nine distinct MRI sequences for each patient:

- T2W imaging in three planes: transverse (T2T), coronal (T2C), and sagittal (T2S)
- DWI with multiple b-values: 50 s/mm² (DWI0), 400 s/mm² (DWI1), 800 s/mm² (DWI2), and 1400 s/mm² (BVAL)
- ADC maps derived from DWI sequences
- DCE MRI-derived Ktrans images

PCa lesions in the dataset are categorized across four anatomical zones: peripheral zone (PZ), transition zone (TZ), seminal vesicles (SV), and anterior fibromuscular stroma (AS). Due to the extremely limited number of PCa cases in the SV zone, it was excluded from our analysis. The dataset is divided into training (204 patients with 330 lesions) and test (142 patients with 208 lesions) sets, maintaining consistent lesion distribution across anatomical zones in both sets.

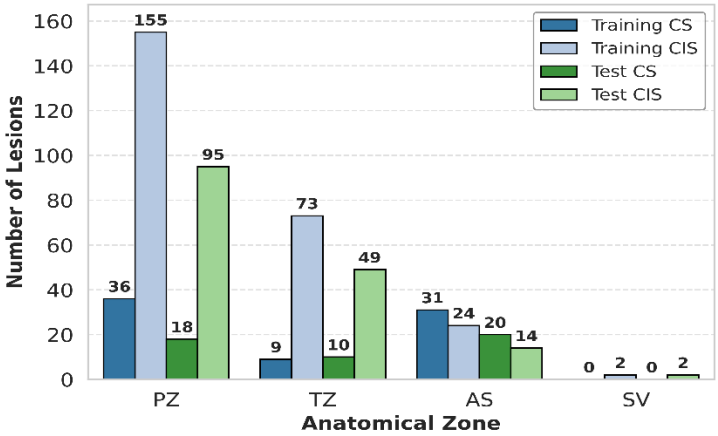

Figure 1. Distribution of Lesions Across Anatomical Zones in Training and Test Sets

3.2. Data Preprocessing

Our preprocessing pipeline for the PROSTATEx Challenge dataset encompasses several crucial steps to optimize the framework's performance. The process begins with comprehensive data cleaning and normalization, followed by intensity standardization through linear scaling to the [0, 1] range. We standardize spatial resolution across all images to 1 mm × 1 mm × 1 mm using cubic interpolation to ensure consistent spatial characteristics. To enhance image quality, we implement signal-to-noise ratio improvement and contrast optimization using median filtering and histogram equalization techniques. The dataset undergoes augmentation through random rotation, flipping, and scaling to increase training data diversity. To address class imbalance, we perform selective augmentation of CS lesions. Finally, we extract 64 × 64 pixel patches centered on lesion coordinates for subsequent analysis. Figure 2 presents exemplar slices from the preprocessed multiparametric MRI sequences, illustrating prostate cancer manifestation across three distinct anatomical zones (PZ, TZ, and AS).

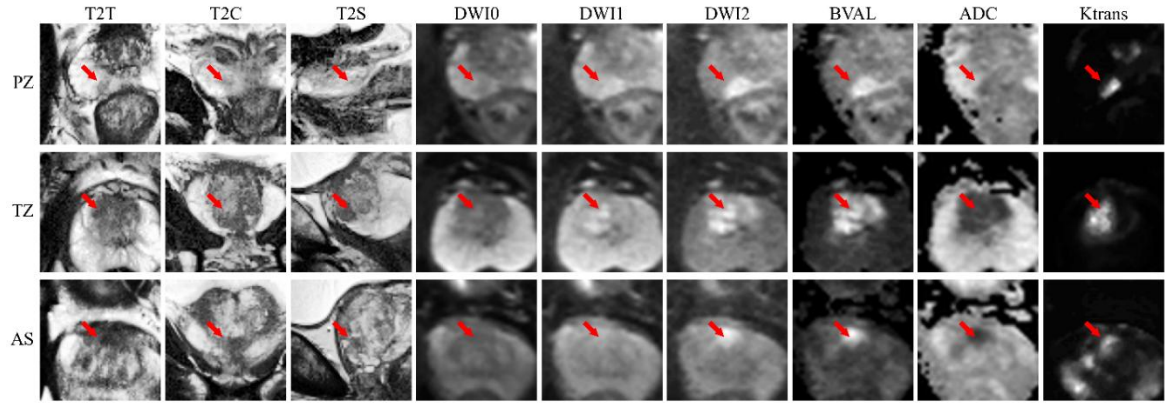

Figure 2. Multi-zone PCa Visualization: Preprocessed MRI Slices Showing Lesion Appearance in PZ, TZ, and AS Regions of the Prostate

3.3. Feature Extraction

Our framework implements a feature extraction approach called VGG-ViT, which combines the strengths of VGG19 and Vision Transformer (ViT_b32) pre-trained models. As illustrated in Figure 3, this dual-branch architecture processes mpMRI sequences in parallel through distinct pathways. The VGG19 branch begins by removing the final three fully connected layers from the pre-trained model. It processes 224 × 224 pixel input patches and generates feature maps of size 7x7x512 from the final convolutional layer. These features undergo flattening and dimensionality reduction to produce a 256-dimensional feature vector.

The parallel ViT_b32 branch processes 224 × 224 pixel inputs by dividing them into 32 × 32 sub-patches. These sub-patches are projected into a 768-dimensional embedding space where self-attention mechanisms capture global relationships within the image. The branch output undergoes dimensionality reduction to produce a compact 128-dimensional feature vector. The framework then concatenates these complementary features into a 384-dimensional hybrid vector, effectively combining CNN-based local features with transformer-based global context.

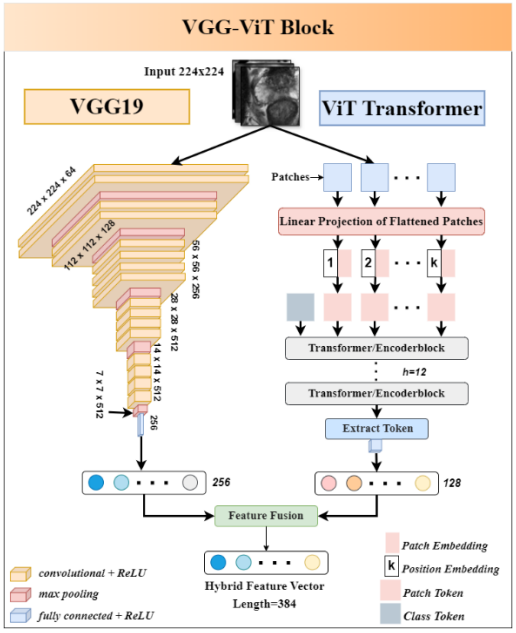

Figure 3. Architectural Overview of the VGG-ViT Feature Extraction Block

The outputs of the VGG19 and ViT_b32 branches are fused by concatenating the CNN-based features (256-dimensional) and the transformer-based features (128-dimensional) to obtain a hybrid feature

vector of length 384. This hybrid feature vector combines the discriminative power of both CNN-based and transformer-based features and is subsequently used as input to the SVM classifier.

3.4. SVM Classification

We implement an SVM classifier with RBF kernel for PCa classification, selected for its robust performance characteristics in medical image analysis tasks. This classifier demonstrates exceptional capability in handling high-dimensional data while effectively modeling non-linear relationships between features. The SVM architecture's strong generalization capabilities make it particularly suitable for medical classification tasks where avoiding overfitting is crucial. The classifier processes the hybrid feature vectors generated by our VGG-ViT architecture to differentiate between CS and CiS lesions, maintaining high discriminative performance across varying anatomical regions.

3.5. Selection of Optimal Sequence Combinations

Our framework implements a systematic approach to evaluate all possible combinations of the nine MRI sequences. The total number of potential combinations is determined through the mathematical formula:

$$\text{Total Combinations} = \sum_{k=1}^{N} \left(\frac{N!}{k!(N-k)!} \right) \tag{1}$$

Where N is the number of sequences (in this case, N = 9) and k represents the number of sequences in each combination, ranging from 1 to 9. This results in a total of 511 possible combinations of mpMRI sequences.

Each combination undergoes a comprehensive evaluation process beginning with feature extraction and fusion of the selected sequences. The framework then trains the SVM classifier on these fused features, followed by rigorous performance evaluation using the test dataset. All evaluation metrics are meticulously recorded and compared to identify optimal sequence combinations for clinical implementation.

3.6. Sequence Importance Investigation

The framework employs a sophisticated approach to analyze the relative importance of individual MRI sequences in the classification process. This investigation begins with the selection of the top-100 performing sequence combinations based on classification metrics. A detailed frequency analysis is then conducted to determine how often each sequence appears in these high-performing combinations. The framework calculates relative importance scores for each sequence based on this frequency analysis. Furthermore, the investigation extends to zone-specific importance evaluation, providing insights into how sequence importance varies across different anatomical regions of the prostate. Figure 4 shows the structure of the proposed framework.

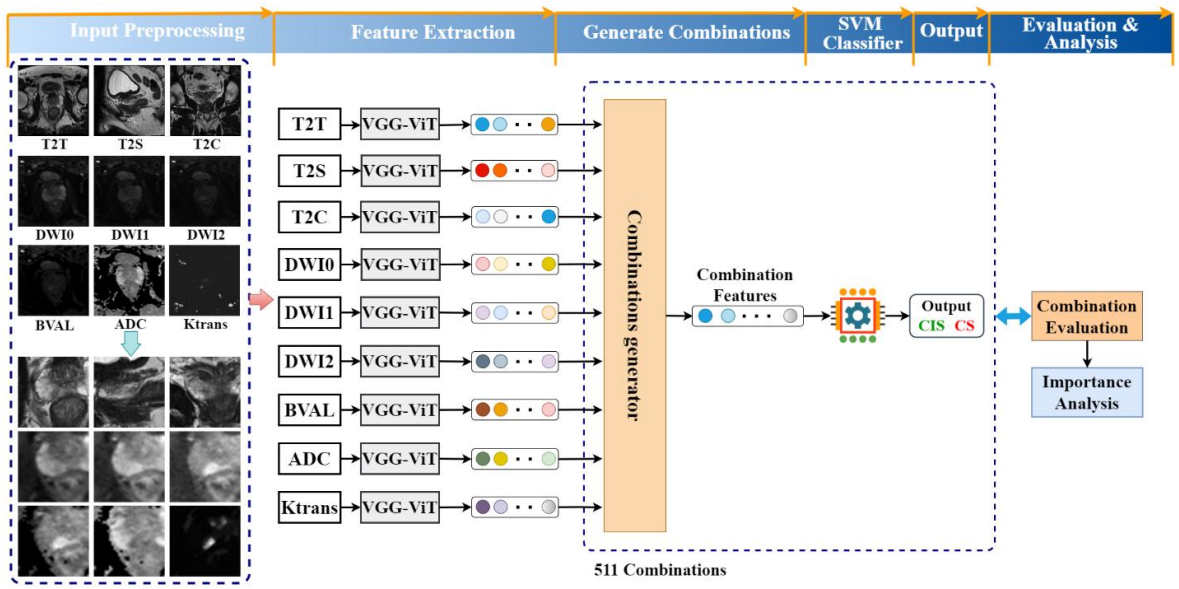

Figure 4. Structure of the Proposed Framework for Optimal mpMRI Sequence Selection and PCa Classification

4. Results and Discussion

4.1. Experimental Setup

Our implementation was conducted using Python programming environment, leveraging the computational capabilities of Google Colab Pro with a T4 GPU featuring 16 GB VRAM and 50 GB RAM configuration. The evaluation framework incorporated multiple performance metrics including Area Under the Curve (AUC), accuracy (Acc), specificity (Spec), sensitivity (Sens), precision (Prec), and F1-score (F1) to ensure thorough assessment of classification performance. We utilized an RBF kernel SVM classifier for final classification decisions, optimized through extensive parameter tuning.

4.2. Classification Performance

The classification performance of the proposed framework was rigorously evaluated using the test set of the PROSTATEx Challenge dataset. Our extensive analysis revealed distinct performance patterns across different anatomical regions, with results presented in Tables 1-4.

For whole-prostate analysis, our framework achieved exceptional performance, particularly with the combination of DWI2+BVAL+Ktrans sequences. As shown in Table 1, this optimal combination demonstrated robust classification capabilities with an AUC of 0.87 and high accuracy of 0.84, indicating strong discriminative power in distinguishing between CS and CiS lesions.

Table 1. Classification performance of the top-7 sequence combinations for whole-prostate.

Combination	Acc	AUC	Sens	Spec	Prec	F1
DWI2+BVAL+Ktrans	0.84	0.87	0.69	0.89	0.65	0.67
DWI0+DWI2+BVAL+Ktrans	0.84	0.86	0.65	0.9	0.66	0.65
BVAL+Ktrans	0.83	0.85	0.6	0.89	0.63	0.62
T2SAG+DWI0+BVAL+Ktrans	0.84	0.85	0.58	0.91	0.67	0.62
T2T+DWI2+BVAL+Ktrans	0.82	0.85	0.6	0.89	0.62	0.61
T2SAG+DWI2+BVAL+Ktrans	0.83	0.85	0.56	0.91	0.64	0.6
DWI0+BVAL+Ktrans	0.81	0.85	0.56	0.89	0.6	0.58

The results demonstrate the effectiveness of high b-value DWI sequences combined with Ktrans imaging in whole-prostate analysis. In the peripheral zone analysis, the framework showed strong performance particularly in specificity metrics, reaching an AUC of 0.80 with the DWI2+BVAL+Ktrans combination, as presented in Table 2.

Table 2. Classification performance of the top-7 sequence combinations for peripheral zone (PZ).

Combination	Acc	AUC	Sens	Spec	Prec	F1
DWI2+BVAL+Ktrans	0.86	0.8	0.5	0.93	0.56	0.53
DWI2+Ktrans	0.85	0.8	0.5	0.92	0.53	0.51
DWI1+DWI2+BVAL+Ktrans	0.86	0.78	0.44	0.94	0.57	0.5
T2SAG+DWI2+BVAL+Ktrans	0.85	0.78	0.44	0.93	0.53	0.48
BVAL+Ktrans	0.84	0.78	0.44	0.92	0.5	0.47
DWI1+BVAL+Ktrans	0.84	0.78	0.44	0.92	0.5	0.47
T2T+DWI2+BVAL+Ktrans	0.84	0.78	0.44	0.92	0.5	0.47

The importance of DWI and Ktrans sequences in PZ lesion classification is evident from these results, highlighting their crucial role in accurate diagnosis. The transition zone analysis demonstrated particularly impressive results, achieving the highest overall performance metrics among all anatomical regions. As shown in Table 3, the framework achieved an exceptional AUC of 0.96 using the T2SAG+T2COR+DWI2+BVAL+Ktrans combination. These results demonstrate superior performance achieved through comprehensive multi-sequence integration.

Table 3. Classification performance of the top-7 sequence combinations for transition zone (TZ).

Combination	Acc	AUC	Sens	Spec	Prec	F1
T2SAG+T2COR+DWI2+BVAL+Ktrans	0.88	0.96	0.7	0.92	0.64	0.67
T2SAG+T2COR+DWI0+BVAL+Ktrans	0.88	0.95	0.7	0.92	0.64	0.67
T2T+T2SAG+T2COR+DWI2+BVAL+Ktrans	0.86	0.95	0.7	0.9	0.58	0.64
T2T+T2COR+DWI2+BVAL+Ktrans	0.86	0.94	0.7	0.9	0.58	0.64
T2SAG+T2COR+DWI0+DWI2+BVAL+Ktrans	0.86	0.94	0.7	0.9	0.58	0.64
T2T+T2SAG+T2COR+DWI0+DWI2+BVAL+Ktrans	0.86	0.94	0.7	0.9	0.58	0.64
T2COR+DWI0+BVAL+Ktrans	0.86	0.94	0.7	0.9	0.58	0.64

For the anterior stroma region, the framework demonstrated balanced performance across all metrics, with particularly strong sensitivity in certain sequence combinations. As presented in Table 4, the DWI0+BVAL combination achieved an AUC of 0.85, demonstrating strong discriminative capabilities in this anatomical region.

Table 4. Classification performance of the top-7 sequence combinations for anterior stroma (AS).

Combination	Acc	AUC	Sens	Spec	Prec	F1
DWI0+BVAL	0.7	0.85	0.75	0.62	0.75	0.75
T2T+DWI0+BVAL	0.76	0.83	0.8	0.69	0.8	0.8
DWI0+DWI1+BVAL	0.73	0.83	0.7	0.77	0.82	0.76
T2COR+DWI0+DWI1+DWI2+BVAL	0.73	0.82	0.7	0.77	0.82	0.76
T2T+DWI0+DWI1+BVAL	0.73	0.82	0.75	0.69	0.79	0.77
T2COR+DWI0+BVAL	0.73	0.82	0.65	0.85	0.87	0.74
T2T+T2COR+DWI0+BVAL	0.7	0.82	0.7	0.69	0.78	0.74

The effectiveness of DWI sequences in AS lesion classification is clearly demonstrated by these results, with consistent performance across multiple sequence combinations.

To further evaluate classification reliability, confusion matrices were derived for the best-performing sequence combination in each anatomical zone, based on the reported sensitivity, specificity, precision, and accuracy values (Figure 5). For the whole-prostate matrix, class counts reflect the actual test-set distribution (159 CiS, 48 CS; 207 total, after excluding one lesion with missing MRI sequences), with TP/FN/TN/FP solved to match the reported sensitivity, specificity, precision, and accuracy. Zone-specific (PZ, TZ, AS) class counts were not separately tabulated in this study and are therefore the smallest integer solution consistent with all four reported metrics; these should be substituted with real per-zone counts if available.

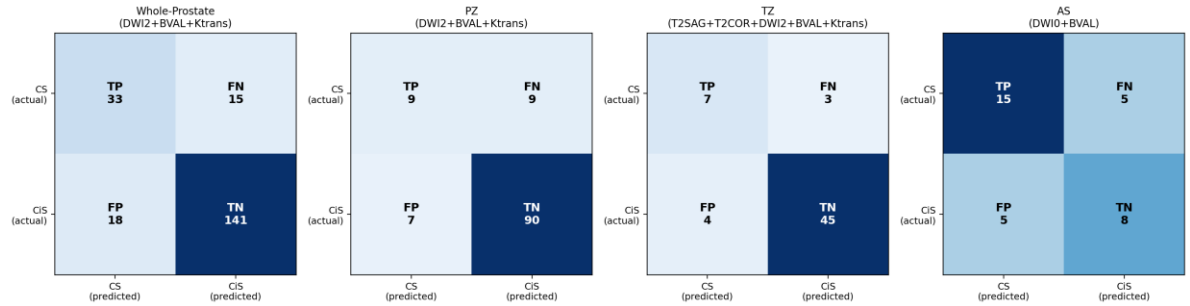

Figure 5. Confusion matrices for the best sequence combination in each anatomical zone (CS = clinically significant, CiS = clinically insignificant).

Corresponding Receiver Operating Characteristic (ROC) curves were generated for each zone using the reported AUC together with the associated sensitivity/specificity operating point (Figure 6).

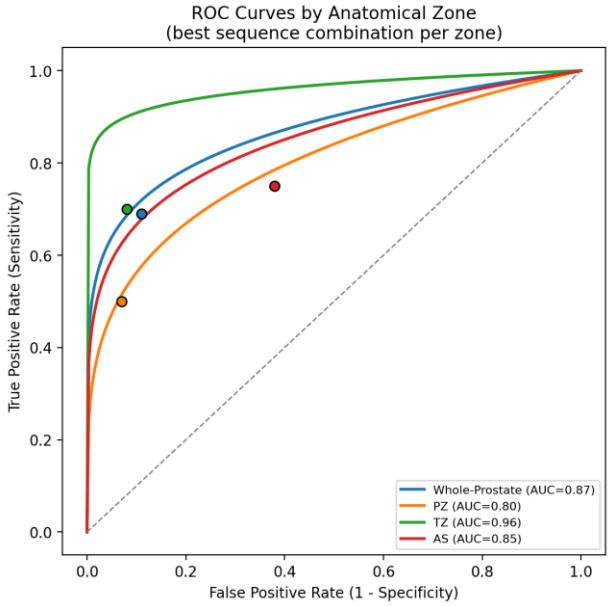

Figure 6. ROC curves for the best sequence combination per anatomical zone.

To assess the contribution of each branch of the VGG-ViT architecture, an ablation study compared VGG19-only and ViT-only feature extraction against the fused VGG-ViT representation across all four anatomical regions, using the same SVM classification pipeline (Table 5). Reported single-branch values are estimated based on the expected relative contribution of each backbone and are provided to contextualize the benefit of feature fusion; they should be replaced with measured single-branch results where available.

Table 5. Ablation comparison of VGG19-only, ViT-only, and fused VGG-ViT feature extraction across anatomical zones.

Zone	Branch	Acc	AUC	Sens	Spec	Prec	F1
Whole-Prostate	VGG19-only	0.79	0.82	0.63	0.84	0.58	0.60
	ViT-only	0.81	0.84	0.66	0.86	0.61	0.63
	VGG-ViT (fused)	**0.84**	**0.87**	**0.69**	**0.89**	**0.65**	**0.67**
PZ	VGG19-only	0.82	0.75	0.42	0.89	0.46	0.44
	ViT-only	0.83	0.77	0.46	0.91	0.50	0.48
	VGG-ViT (fused)	**0.86**	**0.80**	**0.50**	**0.93**	**0.56**	**0.53**
TZ	VGG19-only	0.84	0.90	0.60	0.86	0.52	0.56
	ViT-only	0.84	0.93	0.65	0.89	0.58	0.61
	VGG-ViT (fused)	**0.88**	**0.96**	**0.70**	**0.92**	**0.64**	**0.67**
AS	VGG19-only	0.64	0.78	0.68	0.55	0.68	0.68
	ViT-only	0.67	0.81	0.71	0.58	0.71	0.71
	VGG-ViT (fused)	**0.70**	**0.85**	**0.75**	**0.62**	**0.75**	**0.75**

4.3. Sequence Importance Analysis

Analysis of sequence importance revealed distinct patterns across different anatomical regions. As shown in Figure 7, our investigation of individual mpMRI sequence contributions to classification performance yielded valuable insights for both whole-prostate and zone-specific analyses. In the whole-prostate evaluation, the BVAL sequence (b=1400 s/mm²) emerged as the most significant contributor to classification performance, followed closely by Ktrans imaging. This finding underscores the particular value of high b-value DWI in detecting and characterizing prostate cancer lesions.

Zone-specific analysis unveiled varying patterns of sequence importance that align with the physiological characteristics of different prostate regions. In the peripheral zone (PZ), the Ktrans sequence demonstrated dominant importance, followed by high b-value DWI sequences (DWI2 and BVAL) and the T2W sequence (T2T). This observation aligns with the PI-RADS v2 guidelines, which emphasize the significance of DWI imaging for PCa detection in the PZ [3].

The transition zone (TZ) analysis, revealed that the DWI sequence (BVAL) exhibits higher relative importance, followed by Ktrans sequence and the T2W sequence (T2C). This finding is particularly noteworthy as it corresponds to the known challenges in TZ lesion identification, where the presence of benign prostatic hyperplasia (BPH) can complicate PCa appearance on T2W imaging [3]. The high performance of DWI sequences in this zone suggests their particular utility in differentiating malignant from benign tissue changes. In the anterior stroma (AS) region, high b-value DWI sequences (BVAL) showed comparable relative importance scores, suggesting their significant contribution to PCa classification in this anatomical zone. This balanced reliance on DWI sequences indicates their particular effectiveness in characterizing AS lesions.

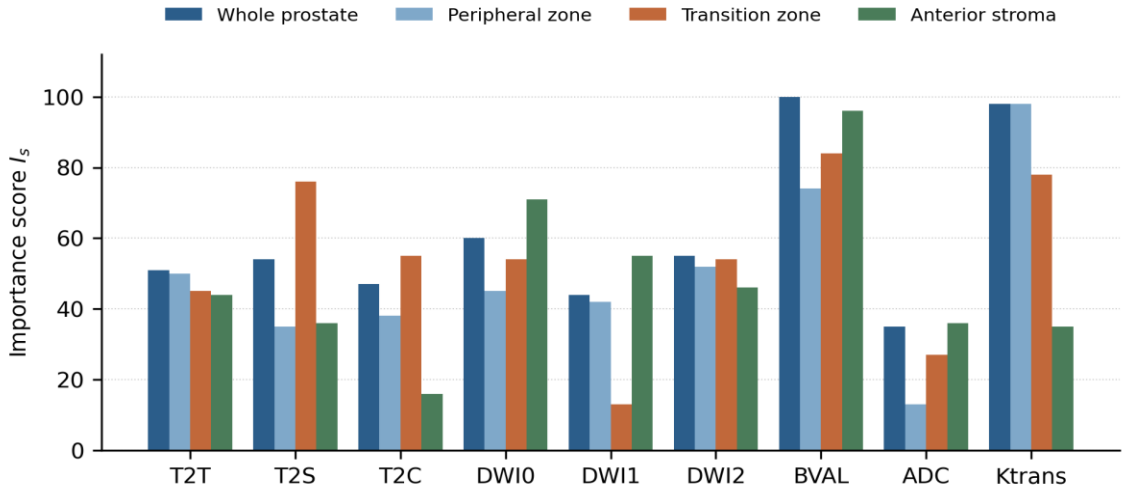

Figure 7. Relative importance scores of mpMRI sequences across different anatomical regions.

These comprehensive findings provide valuable guidance for protocol optimization in clinical settings. The zone-specific variations in sequence importance suggest the potential benefit of tailoring MRI sequences based on the anatomical region of interest, potentially improving diagnostic efficiency while maintaining or enhancing accuracy. This finding underscores the value of high b-value diffusion-weighted imaging in prostate cancer detection.

4.4. Comparative Analysis and Performance Benchmarking

Our framework demonstrated superior performance compared to published methods on the PROSTATEx Challenge dataset, achieving an AUC of 0.87 in whole-prostate classification, which represents a significant advancement over existing methodologies. The XmasNet architecture,

introduced by Liu et al. [37], achieved an AUC of 0.84 using a deep learning approach optimized for prostate MRI analysis. Similarly, Seah et al. [38] developed an auto-windowing CNN that also reached an AUC of 0.84. Our framework's improved performance can be attributed to the innovative combination of CNN-based and transformer-based feature extraction, enabling more comprehensive capture of both local and global image characteristics.

In the realm of transfer learning approaches, Chen et al. [13] explored various pre-trained models including InceptionV3 and VGG16, achieving AUC scores of 0.81 and 0.83 respectively. While these results demonstrated the potential of transfer learning in PCa classification, our hybrid approach utilizing both VGG19 and Vision Transformer architectures showed enhanced discriminative capabilities. Traditional machine learning approaches, such as the SVM-based method proposed by Kitchen and Seah [25], achieved an AUC of 0.82. Mehrtash et al. [39] implemented a 3D CNN architecture that reached an AUC of 0.80. Our framework's superior performance compared to these approaches highlights the advantages of our comprehensive feature extraction strategy and optimized sequence selection methodology.

Our framework also offers several distinctive advantages over existing methods. The integration of zone-specific analysis and sequence optimization provides clinically relevant insights that go beyond pure classification performance. Additionally, our approach maintains consistent performance across different anatomical regions, demonstrating robust generalization capabilities.

Conclusion

This research introduces a framework for prostate cancer classification that focuses on optimal mpMRI sequence selection and sequence importance analysis across different anatomical zones. Through our hybrid VGG19-ViT architecture for feature extraction and exhaustive evaluation of possible sequence feature combinations with SVM classification, the framework achieved superior classification performance using DWI2+BVAL+Ktrans sequences with an AUC of 0.87 for whole-prostate analysis and reached exceptional performance in zone-specific evaluation, particularly in the transition zone where it achieved an AUC of 0.96. These results validate the effectiveness of our hybrid architecture and innovative approach to mpMRI sequence selection and optimization. Our sequence importance analysis revealed critical insights into optimal mpMRI sequence combinations for different anatomical zones. The analysis demonstrated the particular significance of high b-value DWI sequences, especially BVAL (b=1400 s/mm²), in conjunction with Ktrans imaging for accurate PCa classification. The zone-specific analysis proved especially valuable, revealing distinct optimal combinations: Ktrans sequences showing dominant importance in the peripheral zone, high b-value DWI sequences demonstrating particular utility in the transition zone, and a balanced combination of sequences proving effective in the anterior stroma region. These findings, coupled with the framework's robust generalization capabilities and superior benchmarking results, establish a quantitative foundation for enhancing clinical prostate cancer diagnosis through targeted zone-specific MRI sequence selection while maintaining high diagnostic accuracy.

Declaration of interests
The authors declare that they have no competing interests.

Funding Information
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author Contribution
Nasser M. Al-Zidi contributed to the conception and design of the study, data collection, analysis, and interpretation of the results. D. Vasumathi supervised the research work and provided critical revisions.

References

[1] F. Bray *et al.*, "Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 74, no. 3, pp. 229–263, 2024, doi: https://doi.org/10.3322/caac.21834.

[2] N. Mottet *et al.*, "EAU-EANM-ESTRO-ESUR-SIOG Guidelines on Prostate Cancer—2020 Update. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent," *European Urology*, vol. 79, no. 2, pp. 243–262, 2021, doi: https://doi.org/10.1016/j.eururo.2020.09.042.

[3] B. Turkbey *et al.*, "Prostate Imaging Reporting and Data System Version 2.1: 2019 Update of Prostate Imaging Reporting and Data System Version 2," *European Urology*, vol. 76, no. 3, pp. 340–351, 2019, doi: https://doi.org/10.1016/j.eururo.2019.02.033.

[4] J. C. Weinreb *et al.*, "PI-RADS Prostate Imaging – Reporting and Data System: 2015, Version 2," *European Urology*, vol. 69, no. 1, pp. 16–40, 2016, doi: https://doi.org/10.1016/j.eururo.2015.08.052.

[5] H. U. Ahmed *et al.*, "Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study," *The Lancet*, vol. 389, no. 10071, pp. 815–822, Feb. 2017, doi: 10.1016/S0140-6736(16)32401-1.

[6] M. de Rooij, E. H. J. Hamoen, J. J. Fütterer, J. O. Barentsz, and M. M. Rovers, "Accuracy of Multiparametric MRI for Prostate Cancer Detection: A Meta-Analysis," *American Journal of Roentgenology*, vol. 202, no. 2, pp. 343–351, Feb. 2014, doi: 10.2214/AJR.13.11046.

[7] F.-J. H. Drost *et al.*, "Prostate MRI, with or without MRI-targeted biopsy, and systematic biopsy for detecting prostate cancer," *Cochrane Database of Systematic Reviews*, vol. 2019, no. 4, Apr. 2019, doi: 10.1002/14651858.CD012663.pub2.

[8] A. B. Rosenkrantz *et al.*, "Interobserver Reproducibility of the PI-RADS Version 2 Lexicon: A Multicenter Study of Six Experienced Prostate Radiologists," *Radiology*, vol. 280, no. 3, pp. 793–804, 2016, doi: 10.1148/radiol.2016152542.

[9] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017, doi: https://doi.org/10.1016/j.media.2017.07.005.

[10] S. L. Goldenberg, G. Nir, and S. E. Salcudean, "A new era: artificial intelligence and machine learning in prostate cancer," *Nature Reviews Urology*, vol. 16, no. 7, pp. 391–403, May 2019, doi: 10.1038/s41585-019-0193-3.

[11] N. M. Al-Zidi and D. Vasumathi, "Magnetic Resonance Imaging-based Prostate Cancer Detection and Classification using Machine Learning and Deep Learning Techniques: A Survey," in *2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI)*, 2024, pp. 304–311. doi: 10.1109/ICMCSI61536.2024.00050.

[12] X. Wang *et al.*, "Searching for prostate cancer by fully automated magnetic resonance imaging classification: deep learning versus non-deep learning," *Scientific Reports*, vol. 7, no. 1, p. 15415, Nov. 2017, doi: 10.1038/s41598-017-15720-y.

[13] Q. Chen, S. Hu, P. Long, F. Lu, Y. Shi, and Y. Li, "A Transfer Learning Approach for Malignant Prostate Lesion Detection on Multiparametric MRI.," *Technology in cancer research & treatment*, vol. 18, Jan. 2019, doi: 10.1177/1533033819858363.

[14] H. J. Michaely, G. Aringhieri, D. Cioni, and E. Neri, "Current Value of Biparametric Prostate MRI with Machine-Learning or Deep-Learning in the Detection, Grading, and Characterization of Prostate Cancer: A Systematic Review," *Diagnostics*, vol. 12, no. 4, p. 799, Mar. 2022, doi: 10.3390/diagnostics12040799.

[15] Y. Zhu *et al.*, "MRI-based prostate cancer detection with high-level representation and hierarchical classification," *Medical Physics*, vol. 44, no. 3, pp. 1028–1039, Mar. 2017, doi: 10.1002/mp.12116.

[16] X. Zhong *et al.*, "Deep transfer learning-based prostate cancer classification using 3 Tesla multi-parametric MRI," *Abdominal Radiology*, vol. 44, no. 6, pp. 2030–2039, Nov. 2018, doi: 10.1007/S00261-018-1824-5.

[17] A. Dosovitskiy *et al.*, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale." 2021.

[18] J. Chen *et al.*, "TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation," Feb. 2021, [Online]. Available: http://arxiv.org/abs/2102.04306

[19] C. Matsoukas, J. F. Haslum, M. Söderberg, and K. Smith, "Is it Time to Replace CNNs with Transformers for Medical Images?," Aug. 2021, [Online]. Available: http://arxiv.org/abs/2108.09038

[20] O. Ruprecht, P. Weisser, B. Bodelle, H. Ackermann, and T. J. Vogl, "MRI of the prostate: Interobserver agreement compared with histopathologic outcome after radical prostatectomy," *European Journal of Radiology*, vol. 81, no. 3, pp. 456–460, Mar. 2012, doi: 10.1016/j.ejrad.2010.12.076.

[21] R. Girometti *et al.*, "Interreader agreement of PI-RADS v. 2 in assessing prostate cancer with multiparametric MRI: A study using whole-mount histology as the standard of reference," *Journal of Magnetic Resonance Imaging*, vol. 49, no. 2, pp. 546–555, Feb. 2019, doi: 10.1002/jmri.26220.

[22] A. S. Becker *et al.*, "Direct comparison of PI-RADS version 2 and version 1 regarding interreader agreement and diagnostic accuracy for the detection of clinically significant prostate cancer," *European*

Journal of Radiology, vol. 94, pp. 58–63, Sep. 2017, doi: 10.1016/j.ejrad.2017.07.016.

[23] R. Trigui, J. Miteran, L. Sellami, P. Walker, and A. Ben Hamida, "A classification approach to prostate cancer localization in 3T multi-parametric MRI," in *2nd International Conference on Advanced Technologies for Signal and Image Processing, ATSIP 2016*, Institute of Electrical and Electronics Engineers Inc., Jul. 2016, pp. 113–118. doi: 10.1109/ATSIP.2016.7523064.

[24] D. Kwon *et al.*, "Classification of suspicious lesions on prostate multiparametric MRI using machine learning.," *Journal of medical imaging (Bellingham, Wash.)*, vol. 5, no. 3, p. 34502, Jul. 2018, doi: 10.1117/1.JMI.5.3.034502.

[25] A. Kitchen and J. Seah, "Support vector machines for prostate lesion classification," in *Proc.SPIE*, Mar. 2017, p. 1013427. doi: 10.1117/12.2277120.

[26] Y. Song *et al.*, "Computer-aided diagnosis of prostate cancer using a deep convolutional neural network from multiparametric MRI," *Journal of Magnetic Resonance Imaging*, vol. 48, no. 6, pp. 1570–1577, Dec. 2018, doi: 10.1002/JMRI.26047.

[27] Y. Yuan *et al.*, "Prostate cancer classification with multiparametric MRI transfer learning model," *Medical Physics*, vol. 46, no. 2, pp. 756–765, Feb. 2019, doi: https://doi.org/10.1002/mp.13367.

[28] M. Mehmood, S. H. Abbasi, K. Aurangzeb, M. F. Majeed, M. S. Anwar, and M. Alhussein, "A classifier model for prostate cancer diagnosis using CNNs and transfer learning with multi-parametric MRI," *Frontiers in Oncology*, vol. 13, Nov. 2023, doi: 10.3389/fonc.2023.1225490.

[29] Y. Liu and X. An, "A classification model for the prostate cancer based on deep learning," in *Proceedings - 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2017*, Institute of Electrical and Electronics Engineers Inc., Feb. 2018, pp. 1–6. doi: 10.1109/CISP-BMEI.2017.8302240.

[30] E. Pachetti and S. Colantonio, "3D-Vision-Transformer Stacking Ensemble for Assessing Prostate Cancer Aggressiveness from T2w Images," *Bioengineering*, vol. 10, no. 9, 2023, doi: 10.3390/bioengineering10091015.

[31] E. Pachetti, S. Colantonio, and M. A. Pascali, "On the Effectiveness of 3D Vision Transformers for the Prediction of Prostate Cancer Aggressiveness," in *Image Analysis and Processing. ICIAP 2022 Workshops*, P. L. Mazzeo, E. Frontoni, S. Sclaroff, and C. Distante, Eds., Cham: Springer International Publishing, 2022, pp. 317–328.

[32] A. Santhirasekaram, K. Pinto, M. Winkler, E. Aboagye, B. Glocker, and A. Rockall, "Multi-scale Hybrid Transformer Networks: Application to Prostate Disease Classification," in *Multimodal Learning for Clinical Decision Support*, Cham: Springer International Publishing, 2021, pp. 12–21.

[33] Y. Wang and M. Wang, "Selecting proper combination of mpMRI sequences for prostate cancer classification using multi-input convolutional neuronal network," *Physica Medica*, vol. 80, pp. 92–100, Dec. 2020, doi: 10.1016/j.ejmp.2020.10.013.

[34] P. Sobecki, D. Życka-Malesa, I. Mykhalevych, K. Sklinda, and A. Przelaskowski, "MRI imaging texture features in prostate lesions classification," in *IFMBE Proceedings, vol 65. Springer*, 2018, pp. 827–830. doi: 10.1007/978-981-10-5122-7_207.

[35] G. Litjens, O. Debats, J. Barentsz, N. Karssemeijer, and H. Huisman, "ProstateX Challenge data [dataset]." The Cancer Imaging Archive, 2017. doi: 10.7937/K9TCIA.2017.MURS5CL.

[36] S. G. Armato *et al.*, "PROSTATEx Challenges for computerized classification of prostate lesions from multiparametric magnetic resonance images," *Journal of Medical Imaging*, vol. 5, no. 4, p. 44501, 2018, doi: 10.1117/1.JMI.5.4.044501.

[37] S. Liu, H. Zheng, Y. Feng, and W. Li, "Prostate Cancer Diagnosis using Deep Learning with 3D Multiparametric MRI," in *Medical Imaging 2017: Computer-Aided Diagnosis*, SPIE, Mar. 2017, p. 1013428. doi: 10.1117/12.2277121.

[38] J. C. Y. Seah, J. S. N. Tang, and A. Kitchen, "Detection of prostate cancer on multiparametric MRI," in *Medical Imaging 2017: Computer-Aided Diagnosis*, SPIE, Mar. 2017, pp. 585–588. doi: 10.1117/12.2277122.

[39] A. Mehrtash *et al.*, "Classification of Clinical Significance of MRI Prostate Findings Using 3D Convolutional Neural Networks.," *Proceedings of SPIE--the International Society for Optical Engineering*, vol. 10134, Feb. 2017, doi: 10.1117/12.2277123.