**Asian Journal of Current Research in Clinical Cancer**
**ISSN: 3062-4444**
2025, Volume 5, Issue 2, Page No: 34-47

Available online at: www.galaxypub.co/page/journals

Galaxy Publication

Galaxy Publication

# Dense Multimodal Fusion AI Integrating mpMRI and Clinical Features Predicts Castration-Resistant Prostate Cancer Progression at 12 Months

**O. Shevchenko[1*], I. Kovalenko[1], Y. Melnyk[1]**

[1]Department of Oncology, Faculty of Medicine, University of Kyiv, Kyiv, Ukraine.

***E-mail** ✉ kyiv.onc.60@gmail.com

**Received:** 02 May 2025; **Revised:** 28 August 2025; **Accepted:** 02 September 2025

## ABSTRACT

This research sought to determine if individuals diagnosed with prostate cancer (PCa) would transition to castration-resistant prostate cancer (CRPC) within 12 months following hormone-based therapy. Ninety-six men with PCa who had complete baseline clinical information and underwent multiparametric magnetic resonance imaging (MRI) between September 2018 and September 2022 were retrospectively analyzed. Participants were categorized according to whether they progressed to CRPC after 12 months of hormonal treatment. A Dense Multimodal Fusion Artificial Intelligence (Dense-MFAI) framework was created, incorporating a squeeze-and-excitation module and spatial pyramid pooling into a DenseNet backbone, further combined with the eXtreme Gradient Boosting (XGBoost) learning algorithm. Model performance was assessed using accuracy, sensitivity, specificity, positive and negative predictive values, receiver operating characteristic (ROC) curves, area under the curve (AUC), and confusion matrices. The Dense-MFAI system reached an accuracy rate of 94.2% and an AUC of 0.945 when predicting PCa progression to CRPC within the 12-month treatment window. Experimental validation revealed that merging radiomic signatures with baseline clinical variables enhanced the model's predictive capability, emphasizing the advantage of multimodal integration. The proposed Dense-MFAI approach effectively forecasts whether PCa will advance to CRPC, providing clinicians with a data-driven tool for optimizing treatment plans and prognostic evaluations.

**Keywords:** Prostate cancer, Castration-resistant prostate cancer, mpMRI, Deep learning, Radiomics

**How to Cite This Article:** Shevchenko O, Kovalenko I, Melnyk Y. Dense Multimodal Fusion AI Integrating mpMRI and Clinical Features Predicts Castration-Resistant Prostate Cancer Progression at 12 Months. Asian J Curr Res Clin Cancer. 2025;5(2):34-47. https://doi.org/10.51847/fUPj9N3uIt

## Introduction

Prostate cancer (PCa) ranks as the second most frequently diagnosed malignancy among men worldwide [1]. The detection rate of PCa in China remains lower compared to many industrialized nations [2], causing a large proportion of cases to be identified at moderate or advanced phases, which contributes to poorer clinical outcomes [3]. Androgen deprivation therapy continues to be the standard management for advanced PCa [4]. Nevertheless, after roughly 12 months of such treatment, nearly all patients experience progression to castration-resistant prostate cancer (CRPC), though the duration to transition varies among individuals [5-7]. Upon reaching the CRPC phase, tumor growth accelerates rapidly, and survival outcomes are substantially diminished [8].

At present, CRPC advancement is evaluated based on three primary criteria: serum testosterone reduction below 50 ng/dL (1.7 nmol/L), continuous elevation of prostate-specific antigen (PSA), and imaging-confirmed tumor progression [9]. Yet, using single markers such as PSA or testosterone alone lacks reliability [10]. Thus, it is vital to establish an early-stage predictive framework to identify patients likely to develop CRPC after hormone therapy, enabling better treatment planning.

Even with advances in precision oncology, accurately anticipating CRPC onset within a defined period remains a challenge. Several prior works have attempted prediction using multiparametric MRI (mpMRI) combined with clinical indicators. Roest *et al*. [11] employed a U-Net model integrating both imaging and clinical data to predict

clinically significant PCa. Unlike single-architecture designs, our research utilized a staged and modular system that also embedded radiomic information. Current predictive strategies for CRPC often rely on biomarkers, gene/protein expression, or immunohistochemistry analyses [12-14]. Jun *et al*. [15] identified six genes—among them NPEPL1 and VWF—linked to CRPC through screening and regression testing across 287 genes, while Scher *et al*. [12] found a strong association between circulating tumor cell counts and CRPC progression (r = 0.84). However, these biological methods are expensive, invasive, and time-intensive.

By contrast, this study employed mpMRI-derived radiomic descriptors and baseline clinical variables to construct a multimodal predictive framework. T2-weighted imaging provides detailed structural visualization of the prostate, while dynamic contrast-enhanced MRI (DCE-MRI) assesses prostate microvasculature and helps detect recurrence [16, 17]. Radiomics quantifies tumor heterogeneity, offering deeper insight into cancer behavior and progression [18, 19]. Moreover, baseline clinical parameters have also proven valuable for forecasting disease evolution in PCa [20]. Therefore, a quantitative modeling strategy was utilized to reduce subjectivity and eliminate the labor-intensive procedures associated with manual genomic or proteomic testing.

This paper presents a multimodal fusion artificial intelligence (MFAI) technique for predicting, at approximately 12 months, the likelihood of PCa advancement to CRPC based on mpMRI data, radiomic mapping, and clinical factors. Integrating this model into clinical workflows can assist urologists in identifying high-risk patients early and customizing therapy and follow-up regimens accordingly.

## Materials and Methods

*Patients*

This study followed a retrospective design and was reviewed and approved by the Radiological Ethics Committee of the First Hospital of China Medical University, Shenyang. It complied fully with the Declaration of Helsinki's ethical guidelines. Because only preexisting clinical and imaging data were analyzed, additional patient consent was not required. The dataset consisted of men diagnosed with prostate cancer (PCa) who received treatment between September 2018 and September 2022 at our institution. Recorded baseline characteristics included age, prostate-specific antigen (PSA) level, Gleason score, and clinical TNM stage.

Patients were excluded if they:

(i) lacked pre-treatment DCE-MRI before hormone therapy or biopsy;
(ii) had incomplete apparent diffusion coefficient (ADC) sequences;
(iii) had other aggressive cancers or severe comorbidities without full clinical details;
(iv) were diagnosed with a different tumor type confirmed pathologically;
(v) were missing baseline pathological or clinical information; or
(vi) underwent any biopsy or therapy prior to MRI scanning.

A total of 96 participants met the inclusion criteria (mean age = 65 years; range: 45-88 years). Among them, 58 did not progress to CRPC within 12 months of hormone therapy, while 38 did. The recruitment and screening process is presented in **Figure 1**.

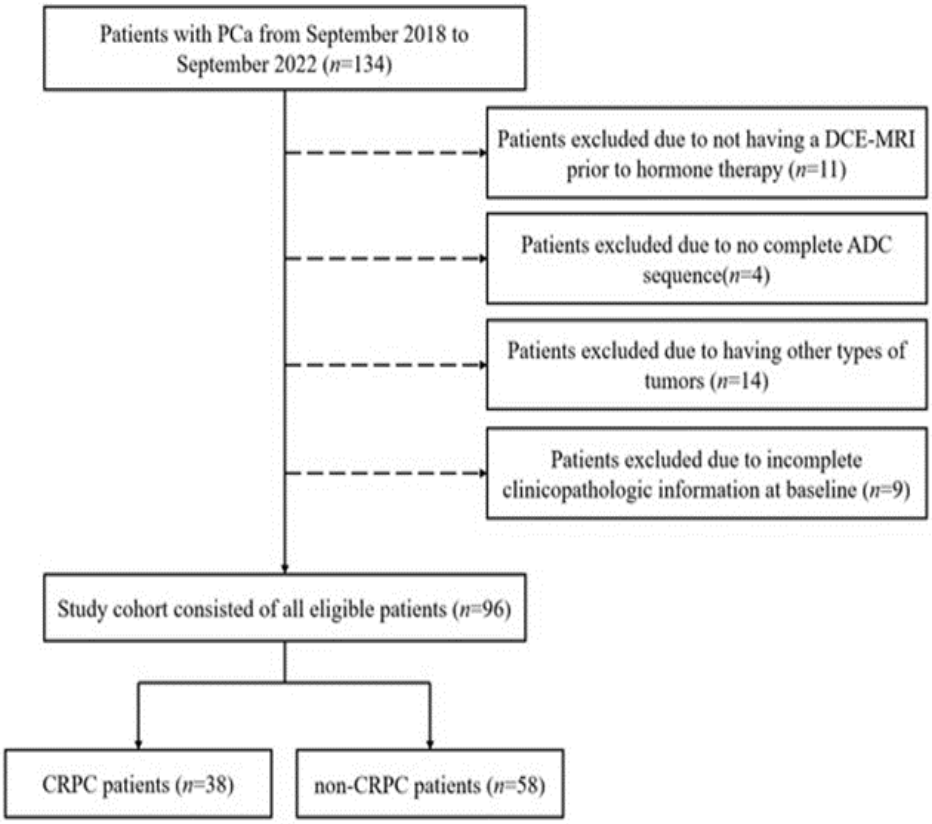



**Figure 1.** Selection workflow for patients. Abbreviations: PCa—prostate cancer; DCE-MRI—dynamic contrast-enhanced magnetic resonance imaging; ADC—apparent diffusion coefficient; CRPC—castration-resistant prostate cancer.

*Image acquisition*

MRI examinations were carried out using a Siemens VIDA 3T system (Siemens Healthineers, Erlangen, Germany). The multiparametric MRI (mpMRI) protocol included T2-weighted imaging (T2WI), diffusion-weighted imaging (DWI), and dynamic contrast-enhanced MRI (DCE-MRI). DWI sequences incorporated SPAIR (Spectral Attenuated Inversion Recovery) fat suppression and utilized b-values of 10, 50, 100, 150, 200, 400, 600, 800, 1000, and 1500 s/mm².

DCE-MRI was obtained with a temporal resolution of 7 seconds. The contrast medium, gadopentetate dimeglumine, was administered at 0.2 mmol/kg according to body weight, and the total imaging duration was 4.3 minutes. Detailed MRI parameter settings are listed in **Table 1**.

**Table 1.** MRI parameter specifications.

| Imaging Sequence | FOV (mm) | Scan Matrix Size | TE (ms) | TR (ms) | Slice Thickness (mm) | Interval (mm) | Flip Angle (°) |
|---|---|---|---|---|---|---|---|
| **T2WI** | 200 | 512 × 512 | 128 | 500 | 3 | 0.6 | 90 |
| **DWI** | 260 | 120 × 96 | 72 | 400 | 3 | 0.6 | / |
| **DCE** | 340 | 203 × 320 | 1.3 | 3.4 | 3.5 | 0 | 9 |

FOV—field of view; TE—echo time; TR—repetition time; T2WI—T2-weighted imaging; DWI—diffusion-weighted imaging; DCE—dynamic contrast-enhanced imaging.

*Data analysis*

Three quantitative features were extracted from mpMRI data: DCE-MRI signal enhancement rate (α), apparent diffusion coefficient (ADC), and T2 values. These measures have been previously shown to assist in tumor characterization [21]. Therefore, they were selected to evaluate the likelihood of progression from PCa to CRPC.

*α-mapping extraction*

The percentage signal enhancement (PSE) for each voxel on DCE-MRI was computed using Eq. 1:

$$PSE(t) = \frac{S(t) - S_0}{S_0} \cdot 100 \quad (1)$$

where S(t) is the signal intensity at time t, and $S_0$ denotes the baseline (pre-contrast) signal. The curve of PSE(t) was then modeled using the empirical function in Eq. 2:

$$PSE(t) = A(1 - e^{-\alpha t})e^{-\beta t} \quad (2)$$

Here, A represents the amplitude, α the rate of enhancement, and β the washout constant. Instead of computing a mean value across the entire region of interest (ROI) first, PSE(t) was determined for each voxel independently. Each voxel's α value was fitted to the above model, and the mean α across the ROI was used for subsequent analysis.

*T2 mapping extraction*

T2 relaxation time was derived from multi-spin echo sequences using the monoexponential decay model shown in Eq. 3:

$$S = S_0 \exp\left(-\frac{TE}{T2}\right) \quad (3)$$

where TE is echo time, S is the signal amplitude for each TE, and $S_0$ is the signal when TE = 0.

*Apparent diffusion coefficient mapping*

The ADC value was determined from DWI signals via Eq. 4:

$$S = S_0 \, exp\,(-b \cdot ADC) \quad (4)$$

where b is the diffusion-weighting factor, $S_0$ is the non-diffused reference signal, and S is the diffusion-attenuated signal. The b-values applied were 10, 50, 100, 150, 200, 400, 600, 800, 1000, and 1500 s/mm².

*Long-run high gray-level emphasis mapping*

In previous work, our group identified long-run high gray-level emphasis (LRHGLE) features—derived from the gray-level contour matrix in α maps—as key factors related to hormonal therapy resistance in PCa [22]. Hence, this texture feature was also computed.

Each pixel was treated as the center of a 5 × 5 matrix used to calculate LRHGLE values. The average of each matrix was then assigned back to the central pixel, generating the LRHGLE texture feature map utilized for radiomic analysis.

*Data preparation*

In this research, target lesions were manually and independently delineated as regions of interest (ROIs) through a semi-automated approach performed by two seasoned radiologists, using each patient's diagnostic imaging report as reference. When the Dice similarity coefficient between the two contour sets was below 0.80, the discrepancy was deemed substantial, leading to a joint review to finalize consensus contours. To maintain uniformity and dependability, once consensus was achieved, the contours drawn by the more experienced radiologist were adopted for all subsequent analyses. Both radiologists were blinded to the histopathological findings. ROIs were selected from the imaging sequence presenting the widest tumor boundary, primarily T2-weighted images (T2WI), and subsequently registered to other modalities [23]. The T2WI-based ROIs were further transferred to apparent diffusion coefficient (ADC) maps and dynamic contrast-enhanced (DCE) MRI sequences through spatial alignment. Although ROI registration across MRI sequences may exhibit minor variations, the prostate's anatomical stability and restricted mobility enhanced the overall alignment reliability [24, 25]. In addition, experienced radiologists verified ROI accuracy following registration. The procedure for lesion annotation is illustrated in **Figure 2**. After completing the feature mapping process, the extracted ROIs

were standardized and combined before being fed into the computational model. The complete data preparation framework is shown in **Figure 3a**.

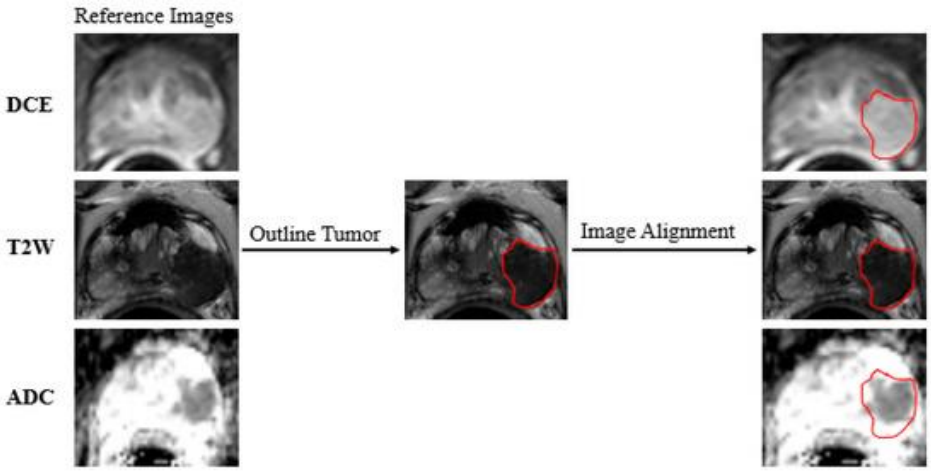



**Figure 2.** Workflow for lesion annotation. Tumor marking was conducted on T2-weighted images, which were then spatially aligned with DCE and ADC sequences. T2W: T2-weighted; ADC: apparent diffusion coefficient; DCE: dynamic contrast-enhanced.

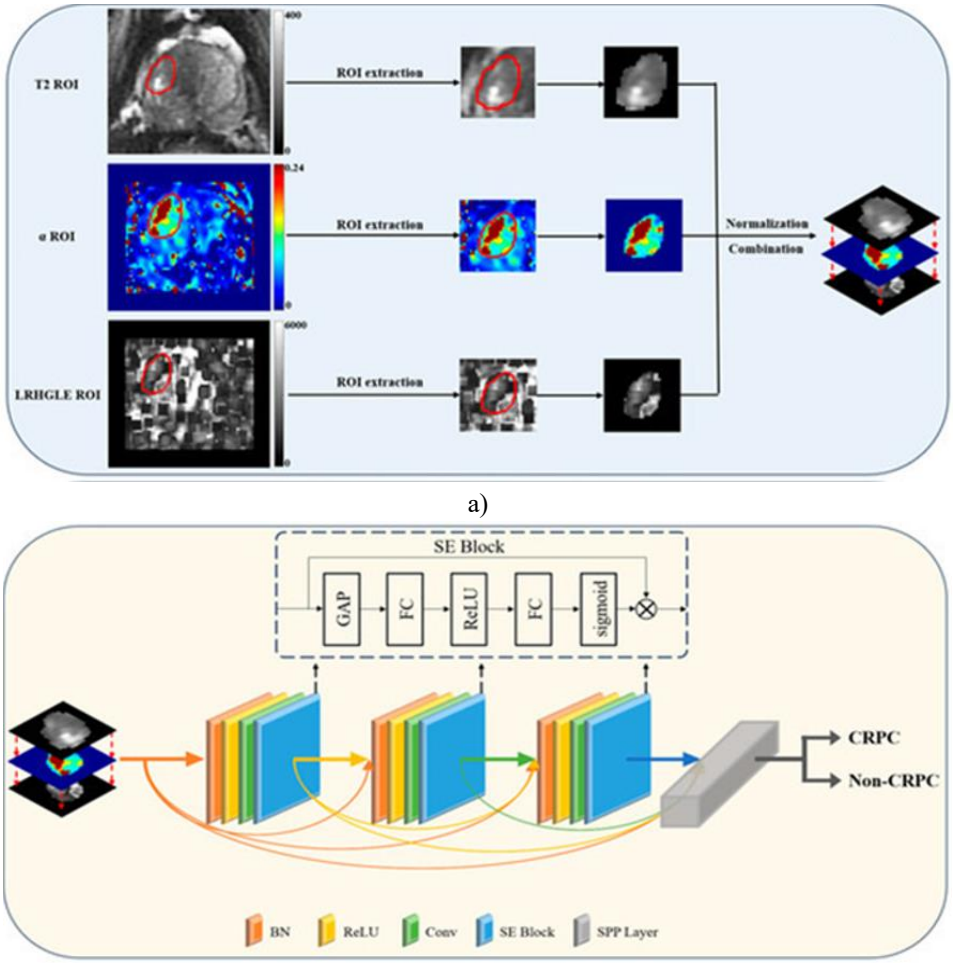



**Figure 3.** Data preparation workflow (a) and SE-SPP-DenseNet architecture (b). (a) The workflow included ROI extraction from parametric maps, normalization, and integration. (b) Preprocessed images were input into the SE-SPP-DenseNet, which adopts DenseNet as its backbone while integrating the squeeze-and-excitation (SE) module and spatial pyramid pooling (SPP) layer. ROI: region of interest; SE block: squeeze-

and-excitation block; SPP layer: spatial pyramid pooling; BN: batch normalization; ReLU: rectified linear unit; Conv: convolutional layer; GAP: global average pooling; FC: fully connected layer.

*Deep learning model construction*

All deep learning frameworks were developed using PyTorch 1.13.1 and Python 3.10 (Python Software Foundation, Fredericksburg, VA, USA). Training was performed on a workstation equipped with an Intel Core i9-13900H processor, NVIDIA RTX 4060 GPU (NVIDIA, Santa Clara, CA, USA), and 8 GB RAM.

Initially, a baseline deep learning classification model based on SE-SPP-DenseNet was established. The dataset was randomly segmented into five equal parts, maintaining a 0.8:0.1:0.1 ratio for training, validation, and testing. Prediction metrics were derived from the mean results of five training-testing cycles, ensuring consistent and unbiased outcomes. During each cycle, four subsets were allocated for training, while one was used alternately for validation and testing. This iterative cross-validation continued until the entire dataset of 96 patients had been covered, with 50 used for testing and 46 for validation. Both training and validation data underwent balancing and augmentation, while no alterations were made to the test samples.

Data augmentation employed the MedAugment algorithm, recognized for its robust automatic enhancement capabilities [26]. This framework utilized two transformation domains: six photometric (pixel-level) and eight geometric (spatial) transformations.

The adopted convolutional backbone was the DenseNet architecture **(Figure 3b)** [27], consisting of sequential dense blocks formed by multiple convolutional layers, where each block receives the output of all preceding ones to promote efficient information propagation. DenseNet121 was selected as the core due to its optimal depth for the dataset's complexity. Earlier research demonstrated that appending the squeeze-and-excitation (SE) block and spatial pyramid pooling (SPP) layer substantially boosts model accuracy [28-30]; therefore, both modules were integrated. The internal design of the SE block is illustrated in **Figure 3b**, showing that after four operations in each convolutional stage, the SE module's output serves as input to the next block.

The standard pooling component in DenseNet was substituted with an SPP layer, which operates in three stages:

(1) applying kernels of various sizes to pool input feature maps;

(2) merging outputs into a fixed-length vector;

(3) forwarding this vector into a fully connected layer for classification.

This design allows variable-scale image inputs to be effectively processed into fixed-size feature vectors. Following SPP processing, 10, 752 features were extracted per multimodal image. Input images were resized to 224 × 224 pixels using bilinear interpolation, with an average ROI dimension of 54 × 48 pixels.

For comparative analysis, the ResNet18 and Swin Transformer architectures were also trained [31-33]. The ResNet18 network captures features across different depths by employing residual connections to mitigate gradient vanishing or explosion. It was further enhanced with SE and SPP modules: SE was inserted after each convolutional block, and SPP replaced the original global average pooling layer, producing classification outputs from the final fully connected layer.

The Swin Transformer, constructed hierarchically, employs a sliding-window mechanism to integrate information across neighboring windows. Input images were progressively downsampled at 4×, 8×, and 16× through patch partition layers. This structure enables both high-resolution feature extraction and multi-scale target classification. The model converts three-channel images through a linear embedding layer, while the patch merging operation doubles feature depth and halves spatial dimensions at each stage.

*Development of the multimodal fusion AI framework*

After constructing a single deep learning classifier, this investigation advanced to designing a multimodal fusion artificial intelligence (MFAI) framework that merges neural network-based imaging analysis with machine learning-based clinical assessment. The resulting hybrid system, termed Dense-MFAI, integrates both radiological features and baseline clinical characteristics—including age, initial PSA, Gleason score, and clinical TNM staging. The workflow of this model is illustrated in **Figure 4**.

Integrating distinct modalities required the creation of two computational pathways. The first channel received imaging features derived from the trained deep network, while the second channel converted the clinical metrics into a one-dimensional numerical representation. The outcomes from both paths were concatenated into a unified

matrix, which was then passed to an eXtreme Gradient Boosting (XGBoost) classifier to estimate the likelihood of CRPC progression within approximately 12 months [34].

A dual-phase training strategy was applied: initially, training the deep network, followed by the XGBoost classifier. Data partitioning remained identical in both stages. XGBoost, a distributed gradient-boosting algorithm, was chosen for its scalability and robustness in addressing regression and classification tasks.

For the imaging channel, data originated from the fully connected layer of the neural network—the segment responsible for mapping distributed image features into label space as numerical arrays. All three deep models in this research utilized only imaging information during training. Meanwhile, clinical data were directly obtained from the patient database and applied to the machine learning stage. Alignment of image and clinical data was achieved through label-based correspondence. Full details of the model's architecture and code repository can be accessed on GitHub (https://github.com/2271371/Dense-MFAI-model.git, accessed on 26 January 2025).

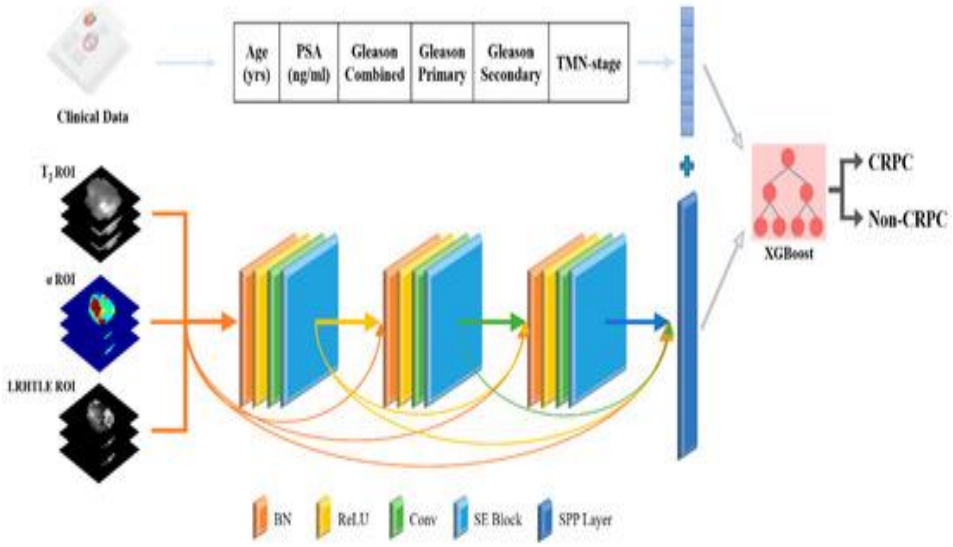



**Figure 4.** Dense-MFAI system layout. The framework integrates two primary modules: SE-SPP-DenseNet and XGBoost. Imaging features and clinical metrics are processed separately and merged within the fully connected layer. PSA: prostate-specific antigen; XGBoost: eXtreme Gradient Boosting.

*Evaluation metrics*

To evaluate the MFAI framework, Kaplan-Meier analysis was used to depict time-to-progression trends for CRPC, and intergroup differences were assessed using the log-rank test. For the independent test cohort, predictive effectiveness was quantified through accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), receiver operating characteristic (ROC) curves, and confusion matrices. Furthermore, the area under the ROC curve (AUC) served as a global index of performance.

## Results and Discussion

*Patient dataset overview*

After balancing and augmentation, imaging data included 2454 images (previously 270) from CRPC patients and 3161 images (previously 391) from non-CRPC patients. Each training cycle utilized roughly 4900 samples, composed of about 2200 CRPC and 2700 non-CRPC images. The validation dataset contained around 690 images—270 from CRPC and 420 from non-CRPC subjects—while the test dataset held approximately 60 images, with 25 CRPC and 35 non-CRPC samples. Importantly, dataset allocation was performed strictly at the patient level, ensuring no image overlap between groups.

*Model parameter configuration*

Following extensive preliminary testing, optimal hyperparameters for the DenseNet, ResNet18, and Swin Transformer models were finalized, as displayed in **Table 2**. The same parameter set was maintained across all

five training repetitions. Tuning was guided by trends in training and validation accuracy/loss, and refinement was performed through a grid search to determine the best configuration.

If validation loss failed to improve within 10 epochs, the learning rate was automatically reduced to half its prior value. Early experiments showed that all three architectures achieved convergence before 50 epochs; thus, the epoch count was fixed at that limit. To minimize overfitting, an early-stopping rule halted training if no improvement occurred for 20 consecutive epochs. The stochastic gradient descent (SGD) algorithm functioned as the optimizer, and binary cross-entropy was used as the loss function.

Image preprocessing included vertical inversion and intensity normalization (based on the mean and variance of each image set), applied solely to training data. Testing samples remained unmodified.

**Table 2.** Training settings for ResNet18, DenseNet, and Swin Transformer models.

| Parameter | Epoch | Learning Rate | Batch Size | Momentum | Weight Decay | Optimizer |
|---|---|---|---|---|---|---|
| **ResNet18** | 50 | 0.001 | 48 | 0.9 | 0.0001 | SGD |
| **DenseNet** | 50 | 0.0001 | 48 | 0.9 | 0.0001 | SGD |
| **Swin Transformer** | 50 | 0.001 | 64 | 0.9 | 0.0001 | SGD |

SGD: stochastic gradient descent.

*Predictive performance of mfai models*

This investigation introduces a hybrid artificial intelligence framework that merges deep learning with machine learning to estimate whether a patient with prostate cancer (PCa) will advance to castration-resistant prostate cancer (CRPC) within 12 months after androgen deprivation therapy. The model integrates several input streams, combining mpMRI quantitative parameters, textural radiomic characteristics, and clinical indicators into a unified predictive process.

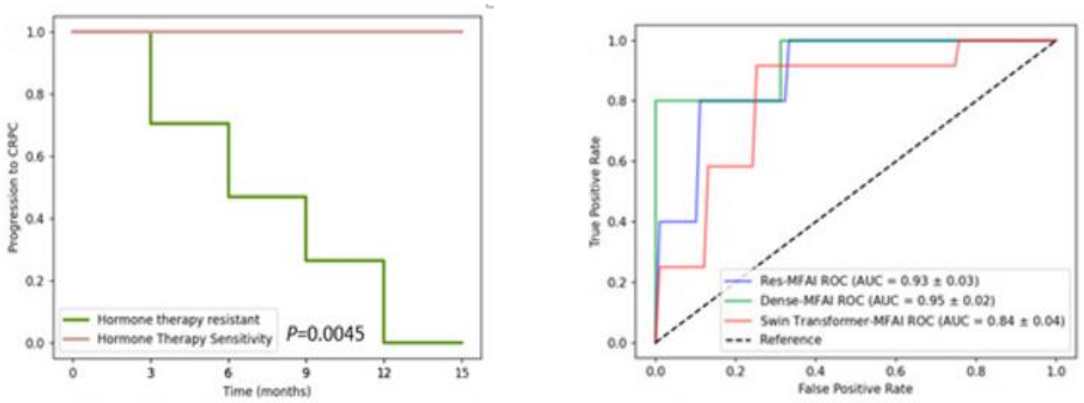
Three variants of the multimodal model were developed, each built upon DenseNet, ResNet18, or the Swin Transformer backbone. The receiver operating characteristic (ROC) plots for the three systems are depicted in **Figure 5b**.

The Dense-MFAI approach recorded an accuracy of 94.2%, sensitivity of 100.0%, specificity of 82.1%, positive predictive value (PPV) of 89.6%, and negative predictive value (NPV) of 100.0%.

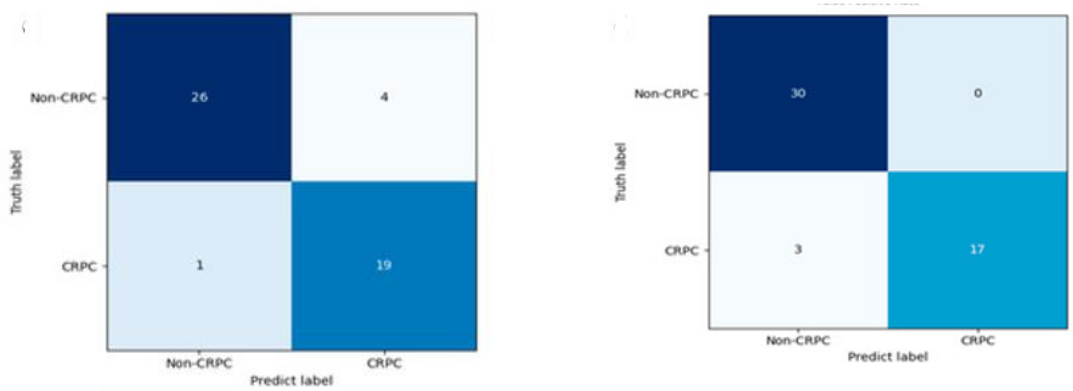
The Res-MFAI version produced an accuracy of 92.7%, sensitivity 87.5%, specificity 96.4%, PPV 97.5%, and NPV 84.1%.

The Swin Transformer-MFAI achieved an accuracy of 85.2%, sensitivity of 87.8%, specificity of 81.4%, PPV of 86.5%, and NPV of 80.6%.

The area under the ROC curve (AUC) reached 0.945, 0.925, and 0.836 for the Dense-, Res-, and Swin-based systems, respectively. The confusion matrices used to assess their CRPC prediction ability are shown in **Figure 5c-e**.



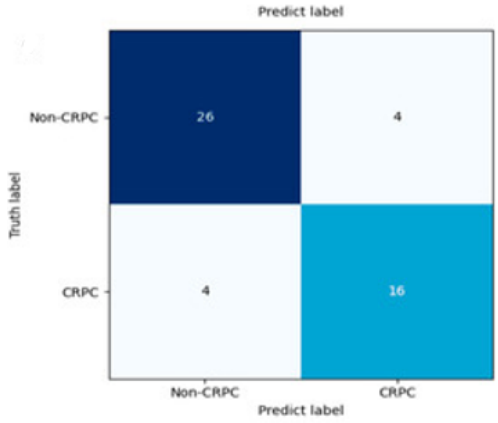

a) b)




c)




d)




e)

**Figure 5.** Assessment of MFAI predictive ability. (a) Kaplan-Meier survival distribution; (b) ROC curves of Res-, Dense-, and Swin-MFAI architectures; (c-e) confusion matrices corresponding to each model.

*Contribution of radiomics feature mapping to CRPC forecasting*

To evaluate whether radiomics feature mapping enhanced the prediction outcome, ablation analyses were carried out for all three backbone networks—DenseNet, ResNet18, and Swin Transformer. Each model was tested twice: once using only mpMRI parameter maps and again using datasets enriched with radiomic mapping information. The summarized outcomes appear in **Table 3**.

**Table 3.** Comparison between models trained solely on mpMRI mapping and those incorporating radiomics feature fusion.

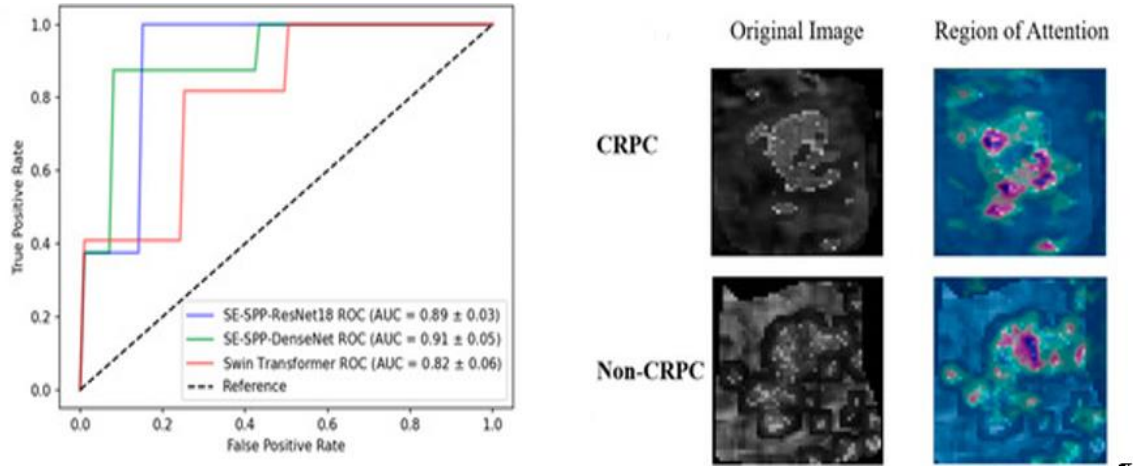
| Model | Input Combination | Accuracy | Sensitivity | Specificity | PPV | NPV |
|---|---|---|---|---|---|---|
| **SE-SPP-ResNet18** | | | | | | |
| | T2-ADC | 76.3% | 74.6% | 80.3% | 78.1% | 75.7% |
| | T2-α | 83.5% | 80.3% | 87.7% | 87.3% | 80.5% |
| | T2-ADC-α | 85.2% | 83.6% | 87.0% | 87.5% | 83.6% |
| | T2-ADC-LRHGLE | 80.4% | 80.7% | 86.4% | 86.7% | 81.8% |
| | T2-α-LRHGLE | 90.3% | 83.3% | 100.0% | 100.0% | 81.6% |
| **SE-SPP-DenseNet** | | | | | | |
| | T2-ADC | 77.6% | 80.5% | 75.1% | 77.2% | 78.5% |
| | T2-α | 85.1% | 90.4% | 75.2% | 80.4% | 91.2% |
| | T2-ADC-α | 87.7% | 89.6% | 87.8% | 88.9% | 88.7% |
| | T2-ADC-LRHGLE | 83.8% | 82.4% | 80.7% | 81.2% | 81.8% |
| | T2-α-LRHGLE | 91.2% | 91.9% | 87.3% | 91.7% | 87.8% |
| **Swin Transformer** | | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| T2-ADC | 67.9% | 75.4% | 54.7% | 62.2% | 69.3% |
| T2-α | 73.1% | 71.4% | 75.4% | 75.8% | 70.5% |
| T2-ADC-α | 75.5% | 82.4% | 68.7% | 70.3% | 76.7% |
| T2-ADC-LRHGLE | 71.5% | 79.7% | 62.7% | 69.9% | 74.0% |
| T2-α-LRHGLE | 78.9% | 98.1% | 51.7% | 73.3% | 96.8% |

*Relevance of baseline clinical variables and visualization of attention regions*

Training and evaluation were conducted using multimodal inputs composed of patient imaging and baseline clinical attributes. Previous research on CRPC forecasting primarily depended on imaging alone; therefore, this analysis compared outcomes obtained with and without clinical feature inclusion using the same three deep learning backbones.

The AUC values for SE-SPP-DenseNet, SE-SPP-ResNet18, and Swin Transformer networks were 0.906, 0.896, and 0.807, respectively. Related ROC curves and statistical summaries are provided in **Table 4** and **Figure 6a**.




a) b)

**Figure 6.** ROC curves and heat-mapped focus zones. (a) ROC curves of SE-SPP-DenseNet, SE-SPP-ResNet18, and Swin Transformer for CRPC prediction; (b) heat-based visualization from the SE-SPP-DenseNet network.

**Table 4.** Comparative prediction metrics for baseline-only, unimodal mpMRI, and multimodal fused datasets.

| Model / Feature Set | Accuracy | Sensitivity | Specificity | PPV | NPV |
|---|---|---|---|---|---|
| **Baseline Clinical Features** | 70.4% | 87.1% | 41.6% | 72.5% | 67.7% |
| **SE-SPP-ResNet18** | 90.3% | 83.3% | 100.0% | 100.0% | 81.6% |
| **SE-SPP-ResNet18 + MFAI** | 92.7% | 87.5% | 96.4% | 97.5% | 84.1% |
| **SE-SPP-DenseNet** | 91.2% | 91.9% | 87.3% | 91.7% | 87.8% |
| **SE-SPP-DenseNet + MFAI** | 94.2% | 100.0% | 82.1% | 89.6% | 100.0% |
| **Swin Transformer** | 78.9% | 98.1% | 51.7% | 73.3% | 96.8% |
| **Swin Transformer + MFAI** | 85.2% | 87.8% | 81.4% | 86.5% | 80.6% |

For the validation component, the SE-SPP-DenseNet, which achieved the best metrics, was analyzed for activation localization **(Figure 6b)**. The visualization employed a color scale: red/yellow indicated zones with maximum impact on decision output, green represented moderate influence, and blue indicated negligible contribution. Attention intensity was concentrated primarily in the tumor core, whereas the peripheral tissue exhibited minimal activation.

A dual-stream AI fusion model was developed to forecast the likelihood of CRPC progression in PCa patients after one year of hormonal therapy. According to current literature, no prior framework has jointly employed deep-learning-based image analysis and machine-learning-based clinical integration for this specific task.

Among the tested models, the Dense-MFAI displayed the greatest predictive precision (94.2%). As evident from **Figure 6a** and **Table 4**, DenseNet outperformed ResNet due to its dense connectivity, which enables continual reuse of learned features and accumulation of hierarchical information. When constructing MFAI systems, the depth and complexity of the network should align with data diversity; overly large architectures may lead to overfitting and make optimization more difficult. Therefore, ResNet18 and DenseNet121 were adopted to maintain model stability and improve convergence efficiency. Overly deep networks expand the solution search space, complicating the optimizer's path toward an optimal generalization point.

Given the limited computational capacity, model precision can be optimized by emphasizing weight distribution toward regions containing the most task-relevant information. Therefore, a squeeze-and-excitation (SE) module was incorporated to prioritize key spatial features. Additionally, the spatial pyramid pooling (SPP) component was introduced, enabling the network to manage input data of variable dimensions while reducing issues associated with shape distortion and positional inconsistencies through hierarchical pooling [30]. As a result, both SE and SPP layers were embedded into the model's final design.

This investigation integrated parametric mapping with radiomic feature extraction, merging their outputs in the fully connected layer. As shown in **Table 3**, integrating radiomic features into the model input led to a significant improvement in prediction accuracy. This confirms that multimodal imaging data deliver more comprehensive information for the algorithm [35]. Moreover, the ablation trials in **Table 3** revealed a drop in performance when α-mapping was substituted with apparent diffusion coefficient (ADC) mapping, likely due to the higher noise level and inferior spatial resolution associated with ADC-based imaging.

As summarized in **Figure 5** and **Table 4**, the Dense-MFAI approach demonstrated the highest prediction accuracy. This enhanced performance is largely attributable to the clinical variables incorporated—specifically, those routinely applied to assess whether prostate cancer (PCa) has transitioned to CRPC—alongside parameters recommended by the National Comprehensive Cancer Network (NCCN) risk classification system. While longitudinal monitoring of patient data is important in practice, the findings suggest that even baseline clinical indicators can provide sufficient predictive insight for CRPC progression.

Compared with prior predictive frameworks, the present model introduces a distinct analytical strategy with several advancements. For instance, Ali *et al*. [36] assessed radiomics-based MR image models for PCa detection and employed an advanced hybrid inferential method. Zhou *et al*. [37] extracted 2, 553 texture descriptors from mpMRI parameter maps, obtaining an AUC of 0.768 for CRPC prediction. In contrast, the Dense-MFAI developed here directly learns discriminative features from raw imaging, allowing for a fully end-to-end system, unlike conventional machine learning radiomics pipelines. Park *et al*. [6] utilized a phased LSTM network for CRPC forecasting, achieving 88.6% accuracy, whereas Jin and Zhang *et al*. [38, 39] demonstrated that integrating clinical, pathological, and mpMRI data enhanced the prediction of bone metastasis. Similarly, Zhou *et al*. [37] applied a ResNet50 combining radiomic and pathomic information, achieving an AUC of 0.86. The Dense-MFAI proposed in this work merges clinical features with imaging and outperforms these methods on the tested dataset. Additionally, Jun *et al*. [15] identified six-gene signatures for risk estimation, though their model failed to reach sufficient patient-level accuracy. Cairone *et al*. [40] also illustrated that semi-automated segmentation can support radiomics analysis, reducing manual workload. Building upon this, future research will incorporate semi-automatic delineation to improve annotation consistency and efficiency.

*study limitations*

This research faces several constraints.

First, being a retrospective investigation, it used patient data collected within a four-year timeframe, which might have introduced selection bias.

Second, because of the limited cohort size, external validation on an independent test set was not feasible; future studies should confirm the model's robustness through multi-institutional external validation.

Third, all mpMRI scans were obtained from identical imaging systems, which reduced variability but also limited generalizability. Hence, replication across diverse MRI platforms and broader patient populations is necessary.

Fourth, ROI segmentation relied on the imaging sequence demonstrating the largest visible tumor area, which could create a partial volume effect. Future work will assess multiple ROI strategies—maximum, minimum, and semi-automated segmentation—to refine both precision and reproducibility.

Lastly, extending this work toward regression-based modeling could complement classification frameworks, providing quantitative risk estimations that aid urologists in decision-making.

## Conclusion

This research presents a multimodal fusion artificial intelligence architecture (Dense-MFAI) for forecasting whether PCa patients may progress to CRPC within 12 months of androgen deprivation therapy. The model integrates the patient's mpMRI-derived parametric metrics, radiomic mappings, and clinical baseline information. The ablation assessments confirmed that combining parametric mapping with radiomic inputs significantly enhanced the classification capability of the core deep learning framework. Furthermore, incorporating clinical variables alongside imaging inputs yielded even greater predictive accuracy compared with the unimodal setup. Ultimately, the Dense-MFAI model achieved a prediction accuracy of 94.2% and an AUC of 0.945 in determining CRPC progression within the specified timeframe. This demonstrates that such a multimodal fusion approach can assist urologists in diagnosis and treatment planning by identifying individuals at heightened risk for CRPC development.

**Acknowledgments:** None

**Conflict of Interest:** None

**Financial Support:** None

**Ethics Statement:** None

## References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209-49.
2. Sharma R. The burden of prostate cancer is associated with human development index: Evidence from 87 countries, 1990-2016. EPMA J. 2019;10(2):137-52.
3. Khauli R, Ferrigno R, Guimarães G, Bulbulan M, Uson Junior PLS, Salvajoli B, et al. Treatment of localized and locally advanced, high-risk prostate cancer: Report from the First Prostate Cancer Consensus Conference for Developing Countries. JCO Glob Oncol. 2021;7:530-7.
4. Pernigoni N, Zagato E, Calcinotto A, Troiani M, Mestre RP, Calì B, et al. Commensal bacteria promote endocrine resistance in prostate cancer through androgen biosynthesis. Science. 2021;374(6566):216-24.
5. Nagappa AN, Bhatt S, Kanoujia J. Studies on structures and functions of kinases leading to prostate cancer and their inhibitors. Curr Enzym Inhib. 2020;16(1):90-105.
6. Park J, Rho MJ, Moon HW, Lee JY. Castration-resistant prostate cancer outcome prediction using phased long short-term memory with irregularly sampled serial data. Appl Sci. 2020;10(6):2000.
7. Ryan CJ, Smith MR, Fizazi K, Saad F, Mulders PF, Sternberg CN, et al. Abiraterone acetate plus prednisone versus placebo plus prednisone in chemotherapy-naive men with metastatic castration-resistant prostate cancer (COU-AA-302): Final overall survival analysis of a randomized, double-blind, placebo-controlled phase 3 study. Lancet Oncol. 2015;16(10):152-60.
8. de Jong AC, Danyi A, van Riet J, de Wit R, Sjöström M, Feng F, et al. Predicting response to enzalutamide and abiraterone in metastatic prostate cancer using whole-omics machine learning. Nat Commun. 2023;14:1968.
9. Cornford P, Bellmunt J, Bolla M, Briers E, De Santis M, Gross T, et al. EAU-ESTRO-SIOG guidelines on prostate cancer. Part II: Treatment of relapsing, metastatic, and castration-resistant prostate cancer. Eur Urol. 2017;71(4):630-42.
10. Cornford P, van den Bergh RCN, Briers E, Van den Broeck T, Cumberbatch MG, De Santis M, et al. EAU-EANM-ESTRO-ESUR-SIOG guidelines on prostate cancer. Part II—2020 update: Treatment of relapsing and metastatic prostate cancer. Eur Urol. 2021;79(2):263-82.
11. Roest C, Kwee TC, de Jong IJ, Schoots IG, van Leeuwen P, Heijmink SW, et al. Development and validation of a deep learning model based on MRI and clinical characteristics to predict risk of prostate cancer progression. Radiol Imaging Cancer. 2025;7:e240078.

12. Scher H, Armstrong A, Schonhoft J, Gill A, Zhao J, Barnett E, et al. Development and validation of circulating tumour cell enumeration (Epic Sciences) as a prognostic biomarker in metastatic castration-resistant prostate cancer. Eur J Cancer. 2021;150:83-94.
13. Jeong SH, Kyung D, Yuk HD, Jeong CW, Lee W, Yoon JK, et al. Practical utility of liquid biopsies for evaluating genomic alterations in castration-resistant prostate cancer. Cancers. 2023;15(12):2847.
14. McGrath S, Christidis D, Perera M, Hong SK, Manning T, Vela I, et al. Prostate cancer biomarkers: Are we hitting the mark? Prostate Int. 2016;4(4):130-5.
15. A J, Zhang B, Zhang Z, Hu H, Dong JT. Novel gene signatures predictive of recurrence-free survival and castration resistance in prostate cancer. Cancers. 2021;13(4):917.
16. Turkbey B, Rosenkrantz AB, Haider MA, Padhani AR, Villeirs G, Macura KJ, et al. Prostate Imaging Reporting and Data System version 2.1: 2019 update. Eur Urol. 2019;76(3):340-51.
17. Hara N, Okuizumi M, Koike H, Kawaguchi M, Bilim V. Dynamic contrast-enhanced MRI for precise detection and staging of early prostate cancer. Prostate. 2005;62(2):140-7.
18. Lambin P, Leijenaar RT, Deist TM, Peerlings J, De Jong EE, Van Timmeren J, et al. Radiomics: The bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. 2017;14(12):749-62.
19. Liberini V, Laudicella R, Balma M, Nicolotti DG, Buschiazzo A, Grimaldi S, et al. Radiomics and AI in prostate cancer: Tools for molecular hybrid imaging and theragnostics. Eur Radiol Exp. 2022;6(1):27.
20. Esteva A, Feng J, van der Wal D, Huang SC, Simko JP, DeVries S, et al. Prostate cancer therapy personalization via multimodal deep learning on randomized phase III trials. npj Digit Med. 2022;5:71.
21. Chatterjee A, He D, Fan X, Antic T, Jiang Y, Eggener S, et al. MRI-derived quantitative risk maps for prostate cancer diagnosis: A feasibility study. AJR Am J Roentgenol. 2019;213(5):W66-W75.
22. Ma Y, He D. Multiparametric MRI-based AI to predict hormone therapy effectiveness for prostate cancer: A clustering-driven ensemble learning approach. CBEC & Innovative Healthcare Summit Proceedings. 2024:2825.
23. Sun C, Chatterjee A, Yousuf A, Antic T, Eggener S, Karczmar GS, et al. Comparison of T2WI, DWI, and DCE-MRI for index lesion volume: Correlation with whole-mount pathology. AJR Am J Roentgenol. 2019;212(2):351-6.
24. Heerkens H, Hall W, Li X, Knechtges P, Dalah E, Paulson E, et al. MRI-based contouring recommendations for pancreatic cancer radiation therapy. Pract Radiat Oncol. 2017;7(2):126-36.
25. Dalah E, Moraru I, Paulson E, Erickson B, Li XA. Variability of target and normal structure delineation using multimodality imaging for pancreatic cancer radiotherapy. Int J Radiat Oncol Biol Phys. 2014;89(3):633-40.
26. Liu Z, Lv Q, Li Y, Yang Z, Shen L. MedAugment: Universal automatic data augmentation plug-in for medical image analysis. arXiv. 2023; arXiv:2306.17466.
27. Huang G, Liu Z, Van der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: CVPR 2017. p. 4700-8.
28. Zhuang H, Li B, Ma J, Monkam P, Qian W, He D. Attention-based deep learning network for predicting platinum resistance in ovarian cancer. IEEE Access. 2024;12:41000-8.
29. Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: CVPR 2018. p. 7132-41.
30. He K, Zhang X, Ren S, Sun J. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans Pattern Anal Mach Intell. 2015;37(9):1904-16.
31. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: CVPR 2016. p. 770-8.
32. Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, et al. Swin Transformer: Hierarchical vision transformer using shifted windows. In: ICCV 2021. p. 10012-22.
33. Xu X, Feng Z, Cao C, Li M, Wu J, Wu Z, et al. Improved Swin Transformer model for remote sensing object detection and instance segmentation. Remote Sens. 2021;13(23):4779.
34. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: KDD 2016. p. 785-94.
35. Zheng X, Yao Z, Huang Y, Yu Y, Wang Y, Liu Y, et al. Deep learning radiomics for predicting axillary lymph node status in early breast cancer. Nat Commun. 2020;11:1236.
36. Ali M, Benfante V, Cutaia G, Salvaggio L, Rubino S, Portoghese M, et al. Prostate cancer detection using radiomics analysis in multiparametric MRI. In: ICIAP 2023 Workshops. Springer; 2023. p. 83-92.
37. Zhou C, Zhang YF, Guo S, Huang YQ, Qiao XN, Wang R, et al. Multimodal data integration with deep learning to predict progression risk in castration-resistant prostate cancer: A multicenter study. Front Oncol. 2024;14:1287995.

38. Jin T, An J, Wu W, Zhou F. Machine learning model for bone metastasis prediction in prostate cancer based on inflammatory and nutritional indicators. Urology. 2024;190:63-70.
39. Zhang YF, Zhou C, Guo S, Wang C, Yang J, Yang ZJ, et al. Deep-learning-based multimodal MRI radiomics and pathomics to improve prediction of bone metastases in primary prostate cancer. J Cancer Res Clin Oncol. 2024;150(1):78.
40. Cairone L, Benfante V, Bignardi S, Marinozzi F, Yezzi A, Tuttolomondo A, et al. Robustness of radiomics features to varying segmentation algorithms in MRI. In: ICIAP 2022 Workshops. Springer; 2022. p. 462-72.