Polish Journal *of* Radiology

Original paper

Reproducibility of MRI-derived radiomic features in prostate cancer detection: a methodological approach

Javad Zarei[1,A,C,D,E], Asma Soleimani[1,A,B,C,E], Marziyeh Tahmasbi[2,C,D,F], Mohsen Sarkarian[3,B,C],
Seyed Masoud Rezaeijo[4,A,B,C,D,E,F,G]

[1]Department of Health Information Technology, School of Allied Medical Sciences, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran
[2]Department of Radiologic Technology, School of Allied Medical Sciences, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran
[3]Department of Urology, School of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran
[4]Department of Medical Physics, Faculty of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

Abstract

Purpose: We aim to evaluate the reproducibility of these features and apply machine learning algorithms to predict cancer diagnosis.

Material and methods: We analyzed magnetic resonance (MR) images from a cohort of 82 individuals, split between 41 prostate cancer patients and 41 healthy controls. A total of 215 radiomic features were extracted from T2-weighted and ADC images using the Software Environment for Radiomic Analysis (SERA). Intraclass correlation coefficient (ICC) analysis was used to assess the reproducibility of features, and Pearson's correlation was applied to remove redundant features. After feature selection, seven dimensionality reduction techniques, including principal component analysis (PCA), kernel PCA, linear discriminant analysis, and locally linear embedding, were applied to preprocess the radiomic features. Ten machine learning algorithms, including support vector machines (SVM), random forests, neural networks, logistic regression, and ensemble methods such as CatBoost and AdaBoost, were utilized to classify cancerous versus non-cancerous tissues. Model performance was evaluated using accuracy and AUC-ROC metrics.

Results: The results showed that features with high reproducibility (ICC > 0.75) contributed significantly to the performance of machine learning models. SVM, neural networks, and logistic regression achieved the highest accuracy (0.88-0.9) and AUC (up to 0.93) when using features from the good and excellent reproducibility categories. PCA emerged as the most effective dimensionality reduction method, preserving the discriminative power of reproducible features across all models.

Conclusion: The results indicate that radiomic feature extraction from MR images, combined with dimensionality reduction and machine learning algorithms, provides a robust approach for prostate cancer diagnosis.

Key words: dimensionality reduction, ICC, machine learning, MRI, prostate cancer, radiomics.

Introduction

Prostate cancer (PCa) stands as a leading health issue for men globally, ranking as the most common cancer among males and the second leading cause of cancer-related deaths [1-4]. While traditional screening methods like prostate-specific antigen (PSA) testing, digital rectal examinations, and transrectal ultrasound-guided biopsies are widely used, they come with significant drawbacks, such as being invasive or having limited accuracy [5]. To address these limitations, multiparametric magnetic resonance imaging (mp-MRI) has emerged as a powerful diagnostic

Correspondence address:
Seyed Masoud Rezaeijo, Department of Medical Physics, Faculty of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran,
e-mail: rezaei-sm@ajums.ac.ir

Authors' contribution:
A Study design · **B** Data collection · **C** Statistical analysis · **D** Data interpretation · **E** Manuscript preparation · **F** Literature search · **G** Funds collection

tool, providing a non-invasive means to detect and evaluate PCa with a high degree of accuracy [6]. However, interpreting mp-MRI to determine the aggressiveness of PCa can be challenging and often relies heavily on the expertise of skilled radiologists.

Recent innovations in artificial intelligence (AI) and machine learning have begun transforming PCa diagnostics by enhancing the capabilities of mp-MRI. These technologies aim to reduce human error, expedite the diagnostic process, and improve overall accuracy by using advanced algorithms to analyze MRI images. Computer-aided diagnostic (CAD) systems [7-9], particularly those utilizing texture analysis and convolutional neural networks (CNNs), have shown promise in distinguishing between clinically significant and insignificant PCas [10]. This technology not only aids radiologists in making more precise diagnoses but also helps in standardizing interpretations, which could lead to better patient outcomes and more tailored treatment plans.

The growing field of radiomics – extracting vast amounts of quantitative data from medical images – represents a frontier in personalized medicine. By analyzing these data, radiomics can provide detailed insights into tumor characteristics that go beyond what is visible to the naked eye, potentially predicting disease progression and treatment response more accurately [11-13]. However, the adoption of radiomics in clinical practice faces several hurdles, including variations in imaging techniques and inconsistencies in how data are processed, which can lead to unreliable results [14-16]. Addressing these technical challenges is crucial for integrating radiomics into routine clinical workflows.
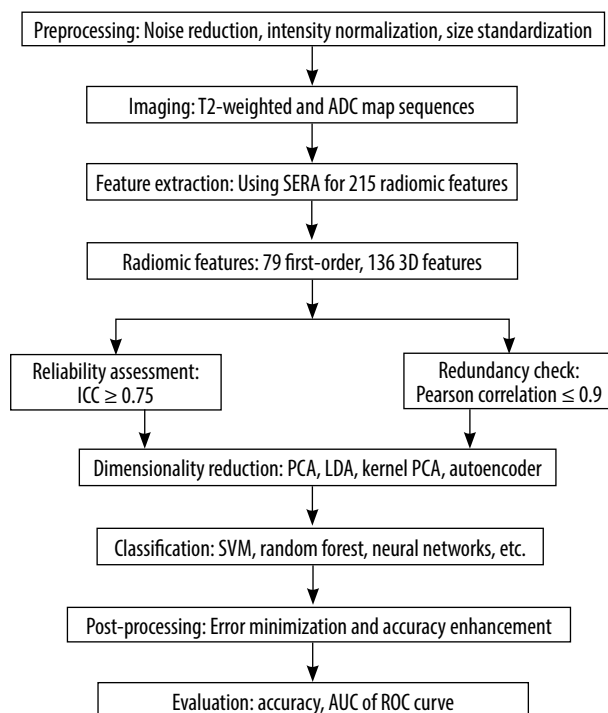
Ensuring the reliability of radiomics is where statistical tools like the intraclass correlation coefficient (ICC) come into play [17-19]. The ICC is used to measure the consistency and agreement of radiomic features across different imaging datasets, accounting for variations in equipment, patient demographics, and imaging protocols. Its application is essential for validating the robustness of radiomic features and ensuring that they can be reliably used in clinical setting [19-21].

This study aims to enhance the reliability and clinical utility of radiomics in PCa diagnosis by identifying robust radiomic features using the ICC. By focusing on features with high ICC values, we can ensure more consistent and accurate extraction of data from MRI images, which can then be applied to advanced AI classifiers. This approach not only promises to improve the accuracy of PCa diagnosis but also supports the broader goal of developing more personalized and effective treatment strategies for patients.

Material and methods

Research methodology steps

The methodology for developing a Prostate Tumor Classification System involves several key stages. The process begins with pre-processing, where prostate imaging data is prepared through noise reduction, intensity normalization, and standardization of MRI image sizes to ensure consistency across all images. Next, feature extraction is performed to generate feature vectors that represent unique aspects of the prostate tissue. Due to the high dimensionality of these extracted features, dimensionality reduction techniques are applied to decrease computational complexity while preserving essential information. The reduced feature vectors are then used to train a classifier to distinguish between tumor types, with regions identified as tumorous or benign based on imaging data. Following classification, post-processing steps are implemented to minimize errors and enhance accuracy. Finally, the system's performance is evaluated using metrics such as accuracy, sensitivity, and specificity, ensuring its reliability and effectiveness for clinical application. In Figure 1, the study flowchart is shown.

Study cohort and imaging techniques

This study analyzed prostate MRI images from a total of 82 individuals, consisting of 41 patients diagnosed with PCa and 41 healthy controls. The images were randomly selected and acquired using two distinct imaging sequences: T2-weighted and apparent diffusion coefficient (ADC) map sequences. The T2-weighted images, known for their high contrast, were utilized to enhance the visualization of soft tissue structures, while the ADC maps were employed to provide insights into the diffusion characteristics of water molecules within the tissue. All imaging was performed

Figure 1. Graphical abstract of the study

using an MRI system, which is well-regarded for its high precision and resolution, thereby ensuring high-quality medical images suitable for detailed analysis. The MRI images used in this study were sourced from a publicly available dataset [https://prostate158.grand-challenge.org].

Radiomics feature extraction methodology

The radiomics feature extraction process yielded a total of 215 features distributed across various categories, including first-order statistics, 3D texture features, and morphological characteristics, among others. Table 1 provides a detailed summary of the extracted radiomics features, highlighting their number and functional relevance in analyzing MRI data. These features were generated using the Software Environment for Radiomic Analysis (SERA) to capture comprehensive intensity, shape, and texture information.

The MRI images were acquired with a 1.5 Tesla system using the following parameters: a field of view (FOV) of 240 mm, a slice thickness of 3 mm, and an in-plane resolution of 0.5×0.5 mm. Pre-processing steps included noise reduction using Gaussian filtering, intensity normalization to standardize the dynamic range across images, and resizing all images to a consistent matrix size of 256×256 pixels for uniform analysis. These steps ensured consistency and minimized variability due to acquisition artifacts or patient-specific differences.

Assessment of reproducibility

The reproducibility of the RFs was assessed using the ICC with specific parameters, including two-way random effects, absolute agreement, and multiple raters or measurements. The ICC is a commonly used metric for evaluating the reliability and agreement between continuous variables in reproducibility studies, providing a ratio between 0 and 1. Based on the ICC value, reliability is classified as poor (ICC < 0.5), moderate ($0.5 \leq$ ICC < 0.75), good ($0.75 \leq$ ICC < 0.9), or excellent (ICC $\geq$ 0.9). Features with an ICC of 0.75 or higher were deemed reliable and included in further analyses. The ICC was calculated using custom Python code developed in-house. Additionally, Pearson's correlation coefficient was applied to detect redundant features, with those showing a correlation coefficient greater than 0.9 considered redundant. Consequently, features exhibiting both high reliability (ICC $\geq$ 0.75) and low redundancy (correlation coefficient $\leq$ 0.9) were averaged and selected for subsequent analysis.

Machine learning algorithms

The machine learning algorithms utilized in this study include support vector machine (SVM), random forest, neural networks, logistic regression, LightGBM, k-nearest neighbors (k-NN), decision trees, CatBoost, and Ada-Boost. These models were selected to represent a broad range of classification techniques, encompassing both linear and nonlinear approaches, as well as ensemble-based methods. SVM and logistic regression were chosen for their effectiveness in linear decision-making, while ensemble methods such as random forest, CatBoost, and AdaBoost were included to evaluate their ability to capture complex patterns. Neural networks were employed to explore the performance of deep learning approaches on feature-extracted data, while k-NN and decision trees were selected for their simplicity and interpretability. LightGBM, as a gradient-boosting algorithm, was chosen for its speed and adaptability to complex datasets.

Table 1. Overview of radiomics features extracted from magnetic resonance images

Feature category	Number of features	Description
First-order features	79	Basic statistical features derived from the intensity values of the MRI images.
3D features	136	Features extracted in three dimensions from the MRI images.
Morphology features (Morph)	29	Features describing the shape and structure of the regions of interest.
Local intensity features (LOC)	2	Features capturing local intensity variations.
Statistical features (STAT)	18	Statistical measures computed from the image data.
Intensity histogram features (IH)	24	Features derived from the intensity histogram of the image.
Intensity volume histogram features (IVH)	7	Features based on the intensity volume histogram.
Co-occurrence matrix features (CM)	50	Features extracted from the co-occurrence matrix (3D, averaged, and merged).
Run length matrix features (RLM)	32	Features derived from the run length matrix (3D, averaged, and merged).
Size zone matrix features (SZM)	16	Features based on the size zone matrix (3D).
Distance zone matrix features (DZM)	16	Features derived from the distance zone matrix (3D).
Neighborhood grey tone difference matrix features (NGT)	5	Features capturing local texture variations using the NGT matrix (3D).
Neighboring grey level dependence matrix features (NGL)	16	Features based on the neighboring grey level dependence matrix (3D).

For dimensionality reduction, the study employed autoencoder, isomap, kernel principal component analysis (kernel PCA), laplacian eigenmaps, linear discriminant analysis (LDA), locally linear embedding (LLE), and principal component analysis (PCA). These techniques were selected to explore both linear and nonlinear transformations of the data. Linear methods, such as PCA and LDA, were used for their ability to preserve global variance and maximize class separability, respectively. Nonlinear methods, including kernel PCA, isomap, laplacian eigenmaps, and LLE, were included to assess their capability in capturing complex, manifold-based structures in high-dimensional data. Autoencoder, as a neural network-based approach, was employed to learn compact and representative feature spaces. This combination of models and dimensionality reduction techniques provides a comprehensive framework for evaluating the interplay between algorithms and feature transformations.

Each machine learning algorithm was fine-tuned using the following hyperparameters: SVM with a radial basis function kernel (C = 1, gamma = 0.1), random forest with 100 estimators and a maximum depth of 10, neural networks with two hidden layers (64 and 32 neurons, respectively) and a learning rate of 0.001, logistic regression with L2 regularization (C = 1), and LightGBM with 200 estimators and a learning rate of 0.05. The CatBoost and AdaBoost ensemble methods were configured with default parameters, with 200 estimators each. For validation, we employed 10-fold cross-validation to ensure the robustness of the models. This approach splits the dataset into ten parts, using nine for training and one for testing, iteratively rotating the test set. Model performance was averaged across folds to mitigate biases due to data partitioning.

Evaluation of models applied

To assess the performance of the classification models applied to prostate data, accuracy and the area under the curve (AUC) of the receiver operating characteristic (ROC) curve were used as primary evaluation metrics. Accuracy measures the proportion of correctly classified instances, providing an overall indication of model performance. AUC, on the other hand, evaluates the discriminatory ability of the models by quantifying the trade-off between the true positive rate (sensitivity) and the false positive rate across different classification thresholds. A higher AUC value indicates better performance in distinguishing between positive and negative classes. These metrics were selected to offer both a straightforward evaluation of overall correctness (accuracy) and a more detailed understanding of the models' ability to differentiate between classes (AUC). This dual approach ensures a robust evaluation of model effectiveness across various dimensionality reduction techniques and classification algorithms. The formula for accuracy is:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

where *TP*, *TN*, *FP*, and *FN* represent true positives, true negatives, false positives, and false negatives, respectively.

Results

This study focused on the automatic diagnosis and classification of PCa using tissue characteristics extracted from MRI images, leveraging machine learning algorithms and feature reduction techniques. A cohort of 82 patients was analyzed, divided into two groups: those with prostate tumors (Class I) and those without (Class II). MRI scans, specifically T2-weighted images (T2W) and ADC maps, were used to extract quantitative features. The ViSERA software facilitated the extraction of 215 quantitative radiomic features from each image sequence. These features were processed to differentiate between healthy and cancerous tissues using machine learning algorithms. To ensure the reliability of the extracted features, ICCs were computed to evaluate their reproducibility across different MRI sequences. Features were classified based on their ICC values into four categories: poor, moderate, good, and excellent. Features with ICC values greater than 0.75 were deemed good or excellent and were used in subsequent analyses.

Reproducibility of T2W and ADC map features

The extracted features were grouped into 11 categories: NGL, NGT, DZM, SZM, RLM, CM, IVH, IH, STAT, LOC, and MORPH. As shown in Figure 2, the distribution of selected features across these groups varied between the T2W and ADC map sequences. For the T2W sequence, the number of selected features in each group was 17, 0, 8, 10, 6, 35, 16, 7, 8, 3, and 7, respectively. In comparison, for the ADC map sequence, the numbers were 15, 0, 5, 8, 3, 25, 8, 6, 6, 2, and 3, respectively. Figures 3 and 4 illustrate the reproducibility of these features, with the horizontal axis representing the radiomic groups and the vertical axis showing the ICC values and their respective quartiles.

Selection of features based on ICC values for T2W sequences and ADC maps

Based on ICC values, four categories of feature reproducibility were established: poor, moderate, good, and excellent. Out of the 215 extracted features for both T2W sequences and ADC maps, those categorized as good and excellent were selected for further analysis. As depicted in Figures 4 and 5, for T2W sequences, 66, 34, 63, and 52 features were identified in the categories of poor, moderate, good, and excellent, respectively. Features in the good and excellent categories (groups 63 and 52) were prioritized for the next phase of analysis (Figure 4). Similarly, for ADC maps, 67, 67, 55, and 26 features were
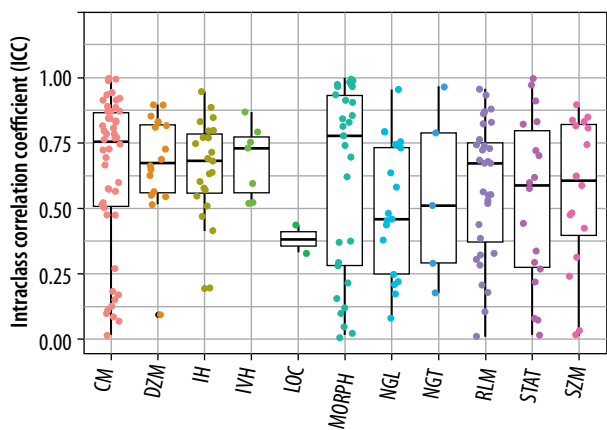
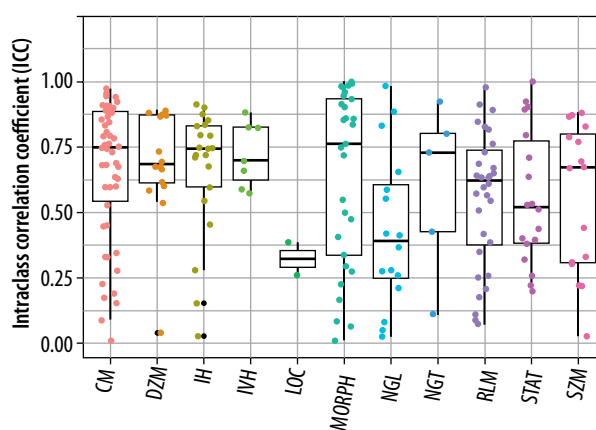**Figure 2.** Reproducibility of features in T2W-weighted images

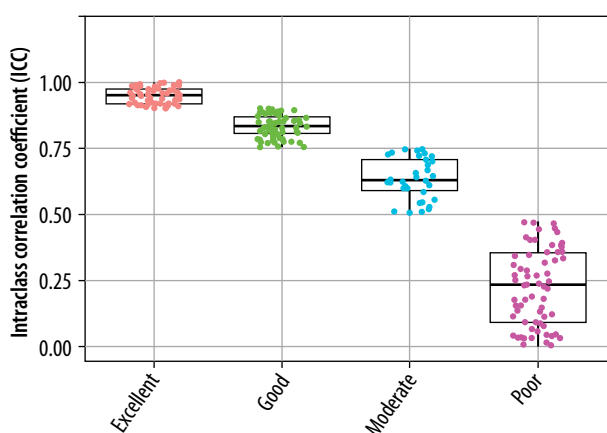**Figure 3.** Reproducibility of features in ADC maps

Figure 4. Classification of reliability based on ICC values across four categories: poor, moderate, good, and excellent for the T2W sequence
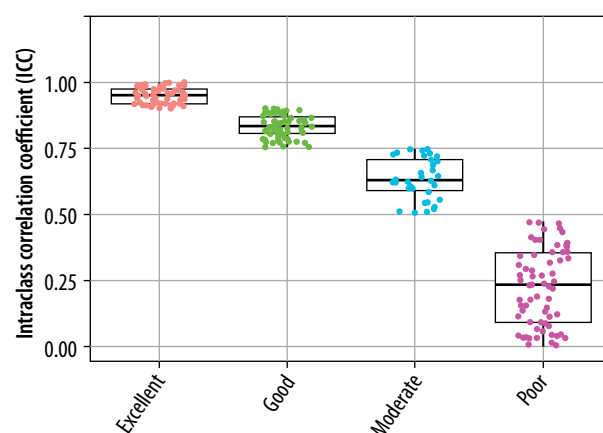

Figure 5. Classification of reliability based on ICC values across four categories: poor, moderate, good, and excellent for the ADC sequence

identified in the respective categories, with groups 55 and 26 being optimal for subsequent analysis (Figure 5). In these figures, the horizontal axis represents the four reproducibility categories, while the vertical axis corresponds to the number of features.

Combining selected radiomic features and machine learning classification

The analysis of accuracy values highlights the robustness of certain algorithms across various dimensionality reduction techniques. SVM, Neural Networks, and Logistic Regression consistently achieve high accuracy (0.88-0.9), demonstrating their adaptability and reliability regardless of feature transformation (Figure 6). PCA emerges as the most effective dimensionality reduction technique, providing high accuracy across all algorithms, while Kernel PCA and Laplacian Eigenmaps also perform well in maintaining accuracy. In contrast, methods like Isomap and LLE introduce significant variability, particularly for algorithms such as LightGBM and AdaBoost, which exhibit lower accuracy (as low as 0.64 and 0.76, respectively). k-NN shows unusual behavior with extremely low accuracy (0.10) for Autoencoder, suggesting

sensitivity to specific feature transformations. Overall, the results underscore the importance of aligning algorithm selection with dimensionality reduction methods to optimize accuracy.

The AUC analysis further reveals important patterns regarding the algorithms' performance in distinguishing between classes (Figure 7). SVM demonstrates exceptional consistency with AUC values of 0.91 across most methods, while Logistic Regression and CatBoost also deliver strong performance (up to 0.91 and 0.93, respectively). Neural Networks achieve the highest AUC (0.93) with PCA and Kernel PCA but show significant drops (e.g., LLE: 0.5), indicating potential instability under certain transformations. k-NN and decision trees exhibit comparable AUC performance (up to 0.93 and 0.91, respectively) but show variability with LLE and isomap, likely due to challenges in adapting to transformed feature spaces. LightGBM performs poorly in AUC for several methods (as low as 0.50), suggesting sensitivity to the quality of feature representation. These findings reinforce PCA as a reliable preprocessing technique for maintaining strong AUC, while highlighting the need for caution with methods like isomap and LLE, which may compromise the discriminatory power of certain algorithms.
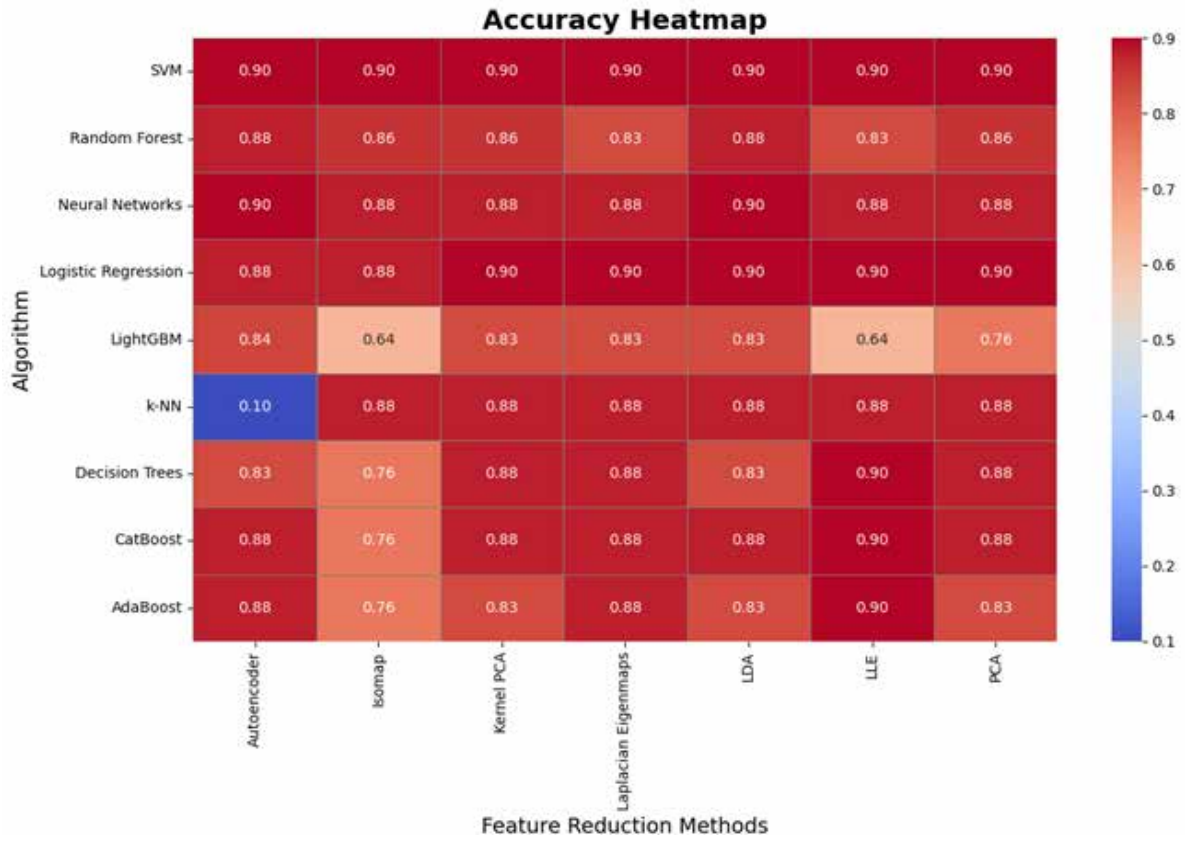
Figure 6. Comparison of accuracy across machine learning algorithms and dimensionality reduction techniques
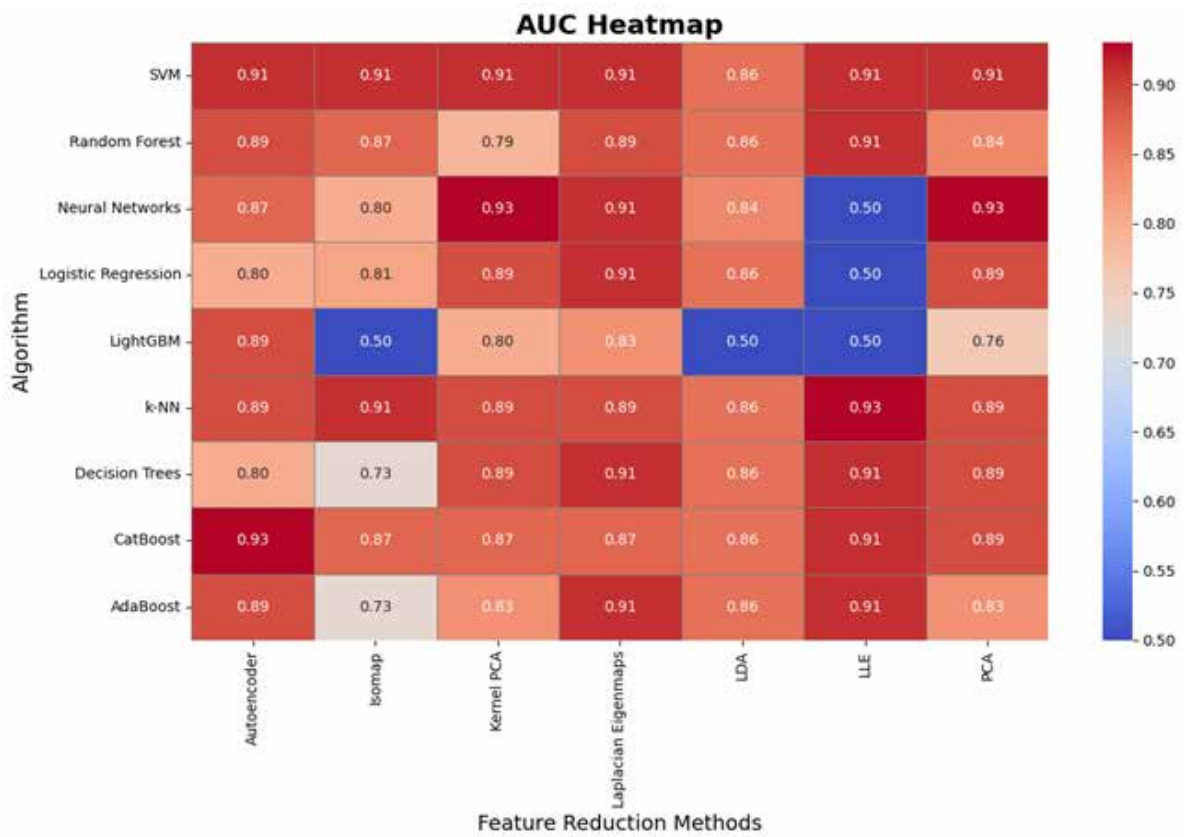

Figure 7. AUC performance of algorithms with various feature transformation methods

Discussion

PCa is the most common cancer among men and the second leading cause of cancer-related deaths. Traditional diagnostic methods like PSA testing and biopsies have limitations in accuracy and invasiveness. mp-MRI offers a non-invasive alternative with high accuracy, but its interpretation requires significant expertise. To improve diagnostic precision, this study explores the use of AI and machine learning techniques, such as CNNs and texture analysis, to enhance mp-MRI capabilities. The methodology includes preprocessing MRI images, extracting radiomic features, and applying dimensionality reduction techniques like PCA to manage data complexity. Machine learning algorithms, including SVM and random forest etc., are used to classify tumor types. To ensure the reliability of radiomic features, the study employs the ICC to select robust features. This approach aims to improve the accuracy and consistency of PCa diagnostics, enabling more personalized treatment strategies.

Our study and the work by Li *et al.* [22] both emphasize the potential of radiomics in improving PCa diagnosis. While our study focuses on a balanced cohort of 82 individuals with 215 radiomic features extracted from T2-weighted and ADC map sequences, Li *et al.* [22] analyzed a larger cohort of 381 patients and used biparametric MRI. Both studies employed machine learning for prediction, but we utilized a broader range of algorithms and dimensionality reduction techniques. In contrast, Li *et al.* [22] focused on logistic regression models and combined clinical data with radiomics for better prediction accuracy. Both works highlight the diagnostic value of radiomics, but our study further explores feature reproducibility and dimensionality reduction methods. Zhang *et al.* [23] developed and validated a multiparametric MRI-based radiomics signature to distinguish between indolent and aggressive PCa, extracting 1576 features from T2WI and diffusion-weighted imaging. In comparison, our study focused on a smaller dataset of 82 individuals and extracted 215 features from T2-weighted and ADC map sequences. Both studies utilized machine learning techniques; however, Zhang *et al.* [23] employed Select K Best and LASSO regression to build their predictive model, while we implemented a broader set of algorithms and dimensionality reduction methods. While Zhang *et al.* [23] concentrated on classifying PCa aggressiveness, our study aimed at overall diagnostic classification, with a greater focus on feature reproducibility.

Rodrigues *et al.* [24] explored the development of robust classifiers for predicting PCa aggressiveness by combining handcrafted and deep radiomic features. They tested several training strategies, including removing unstable features based on ICC, averaging features between radiologists, and using heterogeneous datasets from radiologist-segmented masks. In contrast, our study focused on feature reproducibility and dimensionality reduction in a balanced dataset of 82 patients. While

Rodrigues *et al.* [24] integrated deep learning features alongside handcrafted ones, we exclusively used radiomic features extracted from MRI sequences. Both studies highlight the challenges of overfitting and variability, but our approach emphasizes radiomic feature stability, whereas Rodrigues *et al.* [24] found that heterogeneous data training improved classifier robustness.

Khanfari *et al.* [25] introduced a novel multi-flavored feature extraction approach, combining radiomics and deep features from four mpMRI images using eight fusion techniques, creating 52 datasets per patient. In contrast, our study focused on a smaller dataset of 82 individuals, with radiomic features extracted from T2-weighted and ADC sequences. While Khanfari *et al.* [25] leveraged a wide range of fusion techniques and feature sets, including deep learning, to grade PCa, our study emphasized feature reproducibility and dimensionality reduction in diagnostic classification. Both studies highlight the strength of integrating multiple feature extraction techniques, but Khanfari *et al.* [25]. observed stronger performance from deep features, while our study remained focused on handcrafted radiomic features.

Limitations, challenges, and areas for further investigation in future research

While our study successfully demonstrated the potential of radiomic features and machine learning algorithms for the classification and diagnosis of PCa, several limitations should be acknowledged. First, the relatively small sample size of 82 patients may limit the generalizability of the results to larger and more diverse populations. Additionally, our analysis was confined to two MRI sequences (T2-weighted and ADC maps), and incorporating additional imaging modalities such as DWI or contrast-enhanced MRI could further enhance diagnostic accuracy. Another limitation is the exclusion of clinical data such as patient age, PSA levels, or prostate volume, which could provide valuable complementary information when combined with radiomic features. Future work should focus on expanding the dataset, integrating multi-modal imaging, and incorporating clinical data to improve model robustness. Additionally, exploring more advanced deep learning methods alongside traditional radiomic features could provide further insights into cancer classification and prognosis. Finally, longitudinal studies could assess the effectiveness of these models in monitoring disease progression and treatment response over time.

Conclusions

Our study demonstrates the significant potential of radiomic features extracted from prostate MRI scans combined with machine learning algorithms in enhancing the non-invasive diagnosis of PCa. By focusing on feature reproducibility and employing a diverse range of machine learning models

and dimensionality reduction techniques, we have provided a robust framework for distinguishing between healthy and cancerous tissues. While challenges such as small sample size and limited imaging modalities remain, the results underscore the promise of radiomics as a valuable tool in precision medicine. The ability to leverage quantitative imaging biomarkers offers a pathway toward more accurate, personalized treatment strategies for PCa patients. As advancements in AI and imaging technologies continue, integrating radiomic features with clinical data and deep learning models holds great promise for future improvements in cancer diagnosis, prognosis, and treatment planning.

Disclosures

1. Institutional review board statement: The study was approved by the Ethics Committee of Ahvaz Jundishapur University of Medical Sciences (approval number: IR.AJUMS.REC.1401.533).
2. Assistance with the article: None.
3. Financial support and sponsorship: None.
4. Conflicts of interest: None.

References

1. Chen FZ, Zhao XK. Prostate cancer: current treatment and prevention strategies. Iran Red Cres Med J 2013; 15: 279-284.
2. Rawla P. Epidemiology of prostate cancer. World J Oncol 2019; 10: 63-89.
3. Ohnuma H. 4. Pathology of prostate cancer. Nihon Hoshasen Gijutsu Gakkai Zasshi 2017; 73: 571-577 [Article in Japanese].
4. Tefekli A, Tunc M, Tugcu V, Esen T. Prostate cancer: all aspects. ScientificWorldJournal 2013; 2013: 265234. DOI: 10.1155/2013/265234.
5. Gandaglia G, Leni R, Bray F, Fleshner N, Freedland SJ, Kibel A, et al. Epidemiology and prevention of prostate cancer. Eur Urol Oncol 2021; 4: 877-892.
6. Kumar V, Bora GS, Kumar R, Jagannathan NR. Multiparametric (mp) MRI of prostate cancer. Prog Nucl Magn Reson Spectrosc 2018; 105: 23-40.
7. Reda I, Shalaby A, Elmogy M, Aboulfotouh A, Khalifa F, El-Ghar MA, et al. Image-based computer-aided diagnostic system for early diagnosis of prostate cancer. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2016, pp. 610-618.
8. Xing X, Zhao X, Wei H, Li Y. Diagnostic accuracy of different computer-aided diagnostic systems for prostate cancer based on magnetic resonance imaging: a systematic review with diagnostic meta-analysis. Medicine (Baltimore) 2021; 100: e23817. DOI: 10.1097/MD.0000000000023817.
9. Litjens GJS, Barentsz JO, Karssemeijer N, Huisman HJ. Clinical evaluation of a computer-aided diagnosis system for determining cancer aggressiveness in prostate MRI. Eur Radiol 2015; 25: 3187-3199.
10. Yoo S, Gujrathi I, Haider MA, Khalvati F. Prostate cancer detection using deep convolutional neural networks. Sci Rep 2019; 9: 19518. DOI: 10.1038/s41598-019-55972-4.
11. Bijari S, Sayfollahi S, Mardokh-Rouhani S, Bijari S, Moradian S, Zahiri Z, et al. Radiomics and deep features: robust classification of brain hemorrhages and reproducibility analysis using a 3D autoencoder neural network. Bioengineering 2024; 11: 643. DOI: 10.3390/bioengineering11070643.
12. Mahboubisarighieh A, Shahverdi H, Jafarpoor Nesheli S, Alipoor Kermani M, Niknam M, Torkashvand M, et al. Assessing the efficacy of 3D Dual-CycleGAN model for multi-contrast MRI synthesis. Egypt J Radiol Nucl Med 2024; 55: 1-12.

13. Rezaeijo SM, Harimi A, Salmanpour MR. Fusion-based automated segmentation in head and neck cancer via advance deep learning techniques. In: 3D Head and Neck Tumor Segmentation in PET/CT Challenge. Springer; 2022, pp. 70-76.
14. Whybra P, Zwanenburg A, Andrearczyk V, Schaer R, Apte AP, Ayotte A, et al. The image biomarker standardization initiative: standardized convolutional filters for reproducible radiomics and enhanced clinical insights. Radiology 2024; 310: e231319. DOI: 10.1148/radiol.231319.
15. Salmanpour MR, Hosseinzadeh M, Akbari A, Borazjani K, Mojallal K, Askari D, et al. Prediction of TNM stage in head and neck cancer using hybrid machine learning systems and radiomics features. In: Medical Imaging 2022: Computer-Aided Diagnosis. SPIE; 2022, pp. 648-653.
16. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, et al. Introduction to radiomics. J Nucl Med 2020; 61: 488-495.
17. Bijari S, Rezaeijo SM, Sayfollahi S, Rahimnezhad A, Heydarheydari S. Development and validation of a robust MRI-based nomogram incorporating radiomics and deep features for preoperative glioma grading: a multi-center study. Quant Imaging Med Surg 2025; 15: 1121138-1125138.
18. Fatan M, Hosseinzadeh M, Askari D, Sheikhi H, Rezaeijo SM, Salmanpour MR. Fusion-based head and neck tumor segmentation and survival prediction using robust deep learning techniques and advanced hybrid machine learning systems. In: 3D Head and Neck Tumor Segmentation in PET/CT Challenge. Springer; 2021, pp. 211-223.
19. AkbarnezhadSany E, EntezariZarch H, AlipoorKermani M, Shahin B, Cheki M, Karami A, et al. YOLOv8 outperforms traditional CNN models in mammography classification: insights from a multi institutional dataset. Int J Imaging Syst Technol 2024; 35: e70008. DOI: 10.1002/ima.70008.
20. Paeenafrakati MS, Hajianfar G, Rezaeijo SM, Ghaemi M, Rahmim A. Advanced automatic segmentation of tumors and survival prediction in head and neck cancer. Lect Notes Comput Sci Challenges; 2022.
21. Rezaeijo SM, Chegeni N, Baghaei Naeini F, Makris D, Bakas S. Within-modality synthesis and novel radiomic evaluation of brain MRI scans. Cancers (Basel) 2023; 15: 3565. DOI: 10.3390/cancers-15143565.

22. Li M, Chen T, Zhao W, Wei C, Li X, Duan S, et al. Radiomics prediction model for the improved diagnosis of clinically significant prostate cancer on biparametric MRI. Quant Imaging Med Surg 2020; 10: 368-379.

23. Zhang Y, Chen W, Yue X, Shen J, Gao C, Pang P, et al. Development of a novel, multi-parametric, MRI-based radiomic nomogram for differentiating between clinically significant and insignificant prostate cancer. Front Oncol 2020; 10: 888. DOI: 10.3389/fonc.2020.00888.

24. Rodrigues A, Rodrigues N, Santinha J, Lisitskaya MV, Uysal A, Matos C, et al. Value of handcrafted and deep radiomic features towards training robust machine learning classifiers for prediction of prostate cancer disease aggressiveness. Sci Rep 2023; 13: 6206. DOI: 10.1038/s41598-023-33339-0.

25. Khanfari H, Mehranfar S, Cheki M, Mohammadi Sadr M, Moniri S, Heydarheydari S, et al. Exploring the efficacy of multi-flavored feature extraction with radiomics and deep features for prostate cancer grading on mpMRI. BMC Med Imaging 2023; 23: 195. DOI: 10.1186/s12880-023-01140-0.