



Enhancing automatic prediction of clinically significant prostate cancer with deep transfer learning 2.5-dimensional segmentation on bi-parametric magnetic resonance imaging (bp-MRI)

Mengjuan Li[#], Ning Ding[#], Shengnan Yin, Yan Lu, Yiding Ji, Long Jin

Department of Radiology, Suzhou Ninth People's Hospital, Suzhou Ninth Hospital Affiliated to Soochow University, Suzhou, China

Contributions: (I) Conception and design: M Li, N Ding, L Jin; (II) Administrative support: Y Ji, L Jin; (III) Provision of study materials or patients: M Li, N Ding, S Yin, Y Lu; (IV) Collection and assembly of data: M Li, S Yin, Y Lu; (V) Data analysis and interpretation: M Li, L Jin; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Long Jin, MD. Department of Radiology, Suzhou Ninth People's Hospital, Suzhou Ninth Hospital Affiliated to Soochow University, 2666 Ludang Road, Wujiang District, Suzhou, China. Email: 402588941@qq.com.

Background: The aggressiveness of prostate cancer (PCa) is crucial in determining treatment method. The purpose of this study was to establish a 2.5-dimensional (2.5D) deep transfer learning (DTL) detection model for the automatic detection of clinically significant PCa (csPCa) based on bi-parametric magnetic resonance imaging (bp-MRI).

Methods: A total of 231 patients, including 181 with csPCa and 50 with non-clinically significant PCa (non-csPCa), were enrolled. Stratified random sampling was then employed to divide all participants into a training set [185] and a test set [46]. The DTL model was obtained through image acquisition, image segmentation, and model construction. Finally, the diagnostic performance of the 2.5D and 2-dimensional (2D) models in predicting the aggressiveness of PCa was evaluated and compared using receiver operating characteristic (ROC) curves.

Results: DTL models based on 2D and 2.5D segmentation were established and validated to assess the aggressiveness of PCa. The results demonstrated that the diagnostic efficiency of the DTL model based on 2.5D was superior to that of the 2D model, regardless of whether in a single or combined sequence. Particularly, the 2.5D combined model outperformed other models in differentiating csPCa from non-csPCa. The area under the curve (AUC) values for the 2.5D combined model in the training and test sets were 0.960 and 0.949, respectively. Furthermore, the T2-weighted imaging (T2WI) model showed superiority over the apparent diffusion coefficient (ADC) model, but was not as effective as the combined model, whether based on 2.5D or 2D.

Conclusions: A DTL model based on 2.5D segmentation was developed to automatically evaluate PCa aggressiveness on bp-MRI, improving the diagnostic performance of the 2D model. The results indicated that the continuous information between adjacent layers can enhance the detection rate of lesions and reduce the misjudgment rate based on the DTL model.

Keywords: Deep learning (DL); prostate cancer (PCa); magnetic resonance imaging (MRI); classification

Submitted Mar 23, 2024. Accepted for publication May 20, 2024. Published online Jun 24, 2024.

doi: 10.21037/qims-24-587

View this article at: <https://dx.doi.org/10.21037/qims-24-587>

Introduction

Prostate cancer (PCa) is the most prevalent cancer among men globally, accounting for 27% of cancer incidence in 2022. It is also the second leading cause of cancer-related mortality worldwide, with 34,500 deaths out of 322,090 cases detected, representing nearly 1 in 5 deaths (1).

According to the updated 2020 guidelines, clinically significant PCa (csPCa) is highly aggressive and progresses rapidly, necessitating prompt intervention. In contrast, non-clinically significant PCa (non-csPCa) exhibits lower malignancy and relatively slower progression, often necessitating only observation or active monitoring (2). Accurately assessing the aggressiveness of PCa before treatment is crucial. The commonly used method for assessing invasiveness is transrectal ultrasound-guided needle biopsy (TRUS), which, although effective, is invasive, increases patient discomfort, carries a risk of complications, and can underestimate PCa invasiveness.

Magnetic resonance imaging (MRI) has become integral in PCa detection and grading. In recent years, comparative studies of bi-parametric MRI (bp-MRI), comprising T2-weighted imaging (T2WI) and apparent diffusion coefficient (ADC) mapping, and multi-parametric MRI (mp-MRI) have shown that bp-MRI not only offers superior diagnostic performance for PCa but also simplifies the imaging process, saving patient time and resources (3,4). However, current MRI image interpretation heavily relies on radiologists' subjective evaluations, leading to variability in image interpretation due to differing levels of expertise and a lack of objective assessment. With the rapid advancement of artificial intelligence (AI), deep learning (DL), a quantitative analysis method, has been increasingly applied in medical image recognition and identification. DL involves training a multi-level deep neural network model based on sample data to achieve high classification accuracy (5). Studies have demonstrated the applicability of DL in detecting, grading, and prognostic analysis of prostate diseases (6-8). However, integrating DL into medical imaging remains challenging. The need for extensive labeled data for training reliable neural networks and the laborious, specialized, and costly labeling process pose significant hurdles. Deep transfer learning (DTL), a derivative technique that utilizes pre-trained ImageNet datasets to build fine-tuned convolutional neural networks (CNNs), has successfully addressed these challenges in medical imaging (9-11).

When radiologists interpret computed tomography (CT)

or MRI images, they typically review the images layer by layer, considering the continuity between adjacent slices to extract relevant information. This study aimed to develop a 2.5-dimensional (2.5D) segmentation DTL model for predicting csPCa based on bp-MRI, where 3 consecutive sections were used as input for prediction, and to compare its performance with that of a conventional 2-dimensional (2D) model. We present this article in accordance with the TRIPOD reporting checklist (available at <https://qims.amegroups.com/article/view/10.21037/qims-24-587/rc>).

Methods

Patient selection

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This retrospective study was approved by the Institutional Review Board of the Suzhou Ninth People's Hospital (No. KYLW2024-022-01), and the requirement for written informed consent for this retrospective analysis was waived. From January 2019 to August 2023, a total of 623 patients in our hospital who underwent 3.0T MRI examination due to elevated prostate-specific antigen (PSA) or clinical symptoms with pathological results were included in our research. All patients underwent TRUS.

Data flowchart

The research workflow comprised 2 main segments: image acquisition and segmentation, as well as model construction and comparison. The entire study flowchart and specific details are depicted in *Figure 1*.

Image acquisition

All patients underwent imaging with a 3.0-T MRI scanner (GE Discovery MR750: GE Healthcare, Chicago, IL, USA). Scan sequences included sagittal and axial T2WI, T1-weighted imaging (T1WI), fat-suppressed T2WI, diffusion-weighted imaging (DWI) (b values of 50 and 1,400 sec/mm^2), and dynamic contrast-enhanced MRI (DCE-MRI). The machine automatically generated an ADC map based on the DWI sequence. This study selected the bp-MRI sequence, comprising axial T2WI and ADC images, as it aligns with PI-RADS v2.1 recommendations for accuracy and convenience (3). The parameters for T2WI were as follows: repetition time (TR) = 3,000 ms, echo time (TE)

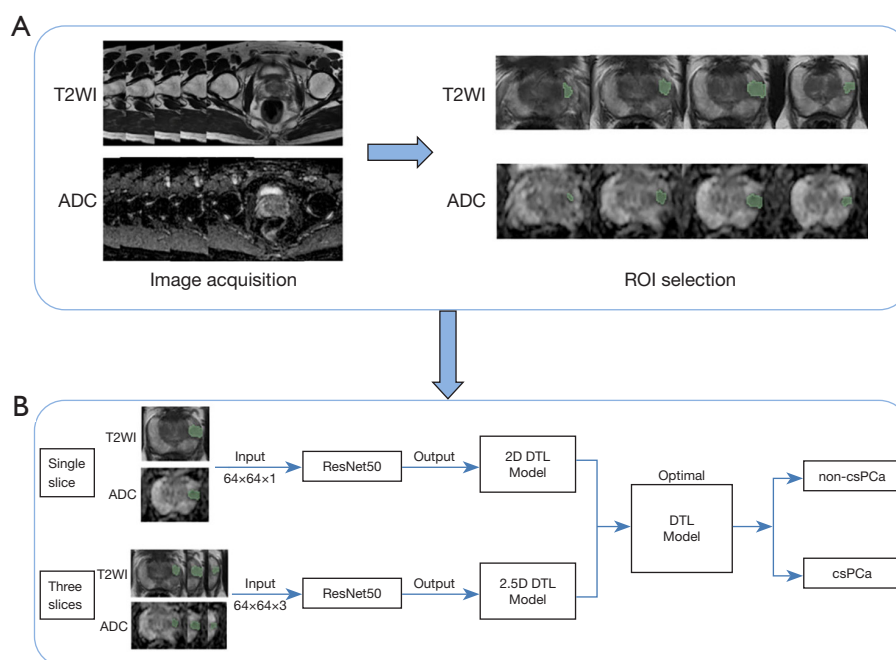


Figure 1 Workflow of this research. (A) Image acquisition and segmentation. (B) Model construction and comparison. T2WI, T2-weighted imaging; ADC, apparent diffusion coefficient; ROI, region of interest; 2D, 2-dimensional; DTL, deep transfer learning; csPCa, clinically significant prostate cancer.

=100 ms, thickness =3 mm, gap =0 mm, field of view (FOV) =220×220 mm, matrix =276×238, number of excitations (NEX) =3. The parameters for DWI were: TR =6,000 ms, TE =77 ms, thickness =3 mm, gap =0 mm, FOV =260×260 mm, matrix =104×126, NEX =2. Image preprocessing involved 3 steps: N4BiasFieldCorrection of the MRI image using the Pyradiomics package (<https://pyradiomics.readthedocs.io/en/latest/>) (12), resampling the image with a voxel size of 1×1×1 mm³, and image registration using the Elastix tool for T2WI and ADC maps. A DL network architecture known as ResNet50 was utilized to construct and validate a DL model for predicting PCa aggressiveness.

Image segmentation

The region of interest (ROI) of the lesion was manually sketched layer by layer on T2WI and ADC maps using 3-dimensional (3D) Slicer (<https://www.slicer.org/>). This step was carried out independently by a radiologist with 12 years of experience in prostate MRI diagnosis, without knowledge of clinical and pathological data.

DTL model construction

The delineated lesions were divided into a training set and a test set in an 8:2 ratio at random. A DL network architecture known as ResNet50 was utilized to construct a DL model for predicting PCa aggressiveness (13,14). The ResNet50 model was efficiently developed using transfer learning, with pre-training on 1.28 million natural images from the ImageNet dataset and subsequent fine-tuning on a training cohort of 231 PCa MRI images (15). This approach addressed the issue of poor generalization ability due to the small sample size of the original dataset and accelerated the model training speed. The input bp-MRI data into the DTL network were divided into a single-layer input (64×64×1 voxel) and a 3-layer input (64×64×3 voxel), resulting in the construction of 2D and 2.5D DL models, respectively. The single-layer input selected the largest layer of the lesion, whereas the 3-layer input selected the largest layer of the lesion and its upper and lower layers. In this study, a T2WI DTL model and an ADC DTL model were constructed based on the T2WI and ADC maps features,

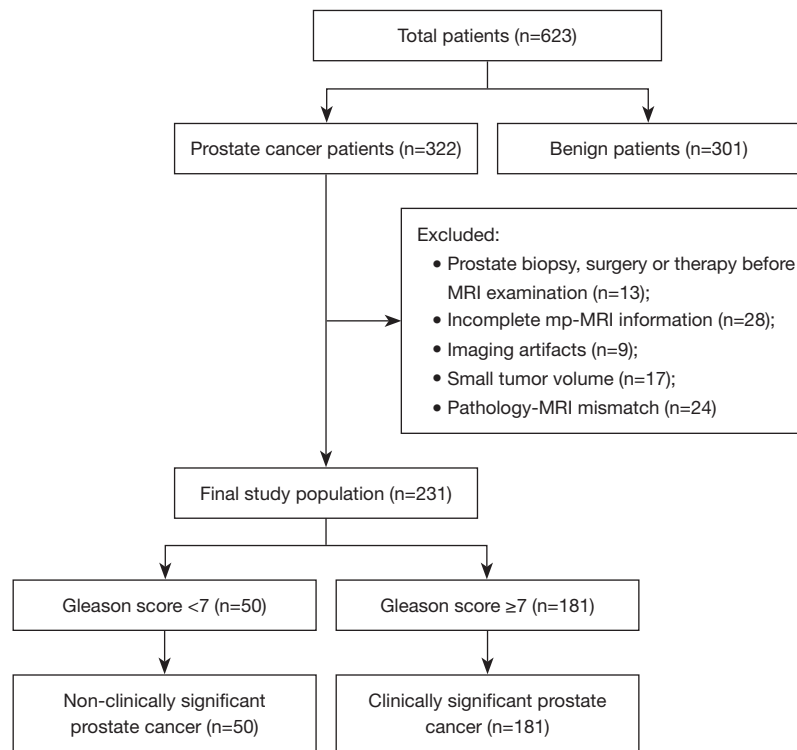


Figure 2 Flow diagram of patient selection. mp-MRI, multi-parametric magnetic resonance imaging.

respectively. Subsequently, the features of T2WI and ADC maps were combined to construct a combined DTL model.

Statistics

A DL network architecture known as ResNet50 was utilized to construct and validate a DL model for predicting PCa aggressiveness. Receiver operating characteristic (ROC) curves were used to demonstrate the ability of the DTL algorithm to evaluate the aggressiveness of PCa. The ROC curve and its related indicators, including area under the curve (AUC) value, accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and F1 score, were used to evaluate and compare the diagnostic efficiency of the constructed models.

Results

Patient characteristics

A total of 231 patients (csPCa: 181 and non-csPCa: 50) were included in this research. The details of patient selection are illustrated in *Figure 2*. All participants were then divided

into a training set [185] and a test set [46] using stratified random sampling. The characteristics of all patients are listed in *Table 1*.

Model construction and comparison

The T2WI DTL models, ADC models, and combined DTL models were constructed using logistic regression. *Figure 3* illustrates the ROC curves of the 6 models constructed, and *Table 2* provides a detailed list of their ROC-related indicators (AUC value, accuracy, sensitivity, specificity, PPV, NPV, F1 score), which comprehensively assess the diagnostic efficiency of the constructed models. In terms of T2WI, ADC map, and combined sequence, the 2.5D model outperformed the 2D model in evaluating the aggressiveness of PCa in the test set. The T2WI model demonstrated superiority over the ADC model, but was not as effective as the combined model, whether based on 2.5D or 2D.

Discussion

In this study, we established and validated DTL models based on 2D and 2.5D segmentation to assess the

Table 1 Characteristics of all patients

Characteristics (n=231)	Clinically significant PCa (n=181)	Non-clinically significant PCa (n=50)	P value
Age, years (mean ± SD)	74±8	72.8±7	0.299
Prostatic volume (mL), median (IQR)	36.7 (27, 57.9)	41.3 (32.3, 50.4)	0.363
tPSA (ng/mL), median (IQR)	37.2 (13.1, 101.2)	11.4 (8.6, 17.6)	<0.001
fPSA (ng/mL), median (IQR)	3.3 (1.5, 11.6)	1.2 (0.9, 2.2)	<0.001
PSAD (ng/mL/mL), median (IQR)	1.2 (0.4, 2.1)	0.3 (0.2, 0.6)	<0.001
f/tPSA (%), median (IQR)	0.1 (0.1, 0.2)	0.1 (0.1, 0.2)	0.979
Gleason score			
3+3=6	–	50	–
3+4/4+3=7	73	–	–
4+4=8	58	–	–
4+5/5+4=9	38	–	–
5+5=10	12	–	–

PCa, prostate cancer; SD, standard deviation; IQR, interquartile range; tPSA, total prostate-specific antigen; fPSA, free PSA; f/tPSA, ratio of free-to-total PSA; PSAD, PSA density.

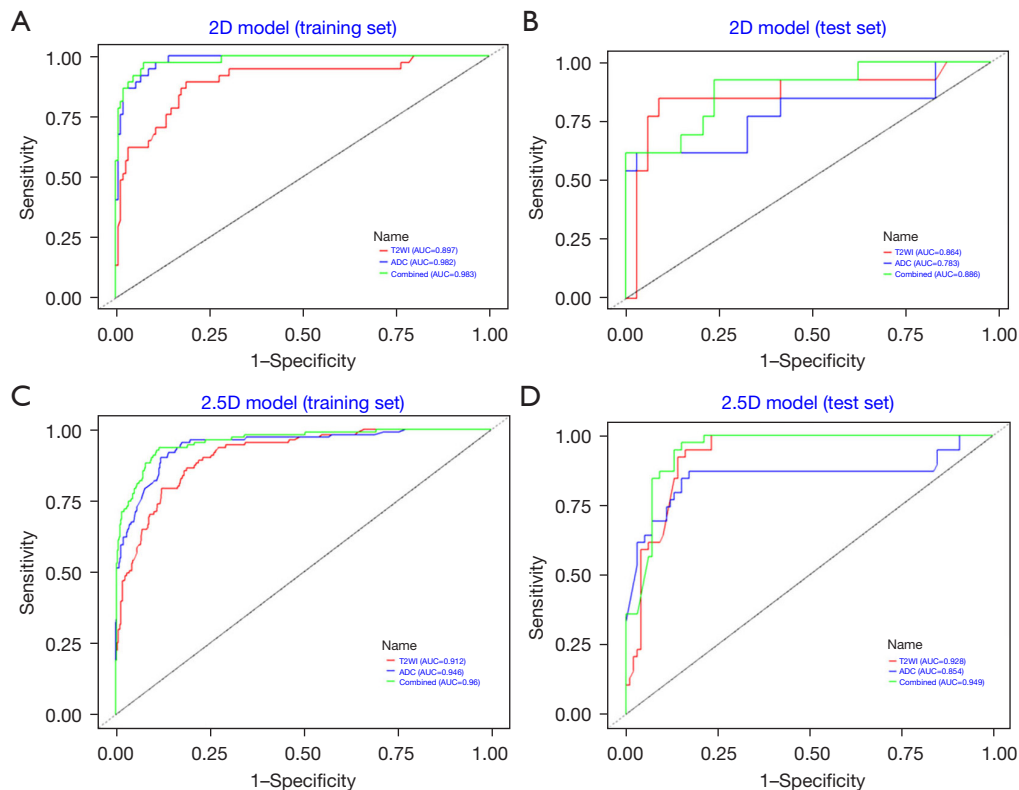


Figure 3 ROC curves of different deep transfer learning models for prediction of prostate cancer aggressiveness. (A) ROC evaluation for 2D deep transfer learning model in training set. (B) ROC evaluation for 2D deep transfer learning model in test set. (C) ROC evaluation for 2.5D deep transfer learning model in training set. (D) ROC evaluation for 2.5D deep transfer learning model in test set. 2D, 2-dimensional; T2WI, T2-weighted imaging; AUC, area under the curve; 2.5D, 2.5-dimensional; ADC, apparent diffusion coefficient; ROC, receiver operating characteristic.

Table 2 AUC results of the 2D and 2.5D deep transfer learning models for prediction of prostate cancer aggressiveness

Models	AUC (95% CI)	Accuracy	Sensitivity	Specificity	PPV	NPV	F1 score
2D T2WI model							
Training	0.897 (0.834–0.960)	0.827	0.811	0.892	0.968	0.541	0.882
Test	0.864 (0.718–1.000)	0.891	0.909	0.846	0.937	0.786	0.923
2D ADC model							
Training	0.982 (0.968–0.997)	0.908	0.892	0.973	0.992	0.692	0.940
Test	0.783 (0.603–0.964)	0.870	0.970	0.615	0.865	0.889	0.914
2D combined model							
Training	0.983 (0.966–1.000)	0.935	0.973	0.926	0.766	0.993	0.857
Test	0.886 (0.773–0.998)	0.804	0.923	0.758	0.600	0.962	0.727
2.5D T2WI model							
Training	0.912 (0.884–0.940)	0.822	0.811	0.865	0.960	0.533	0.879
Test	0.928 (0.887–0.970)	0.870	0.838	0.949	0.976	0.698	0.902
2.5D ADC model							
Training	0.946 (0.922–0.970)	0.885	0.881	0.901	0.973	0.654	0.924
Test	0.854 (0.761–0.946)	0.841	0.828	0.872	0.943	0.667	0.882
2.5D combined model							
Training	0.960 (0.941–0.980)	0.896	0.937	0.885	0.671	0.717	0.782
Test	0.949 (0.916–0.982)	0.884	0.974	0.849	0.983	0.988	0.826

AUC, area under the curve; 2D, 2-dimensional; 2.5D, 2.5-dimensional; 95% CI, 95% confidence interval; PPV, positive predictive value; NPV, negative predictive value; T2WI, T2-weighted imaging; ADC, apparent diffusion coefficient.

aggressiveness of PCa. Our results demonstrate that the diagnostic efficiency of the 2.5D DTL model surpasses that of the 2D model, regardless of whether considering a single sequence or a combined sequence. Notably, the 2.5D combined model outperformed other models in differentiating between csPCa and non-csPCa, including the 2D T2WI model, 2D ADC model, 2D combined model, 2.5D T2WI model, and 2.5D ADC model. The AUC values for the 2.5D combined model in the training and test sets were 0.960 and 0.949, respectively. Furthermore, although the T2WI model demonstrated superiority over the ADC model, it did not match the performance of the combined model, whether based on 2.5D or 2D.

AI and machine learning (ML) methods have the potential to address various everyday challenges owing to their capacity to swiftly process large volumes of data. Chutisant *et al.*'s (16) literature review on AI indicates that AI will play a significant role in training, education, patient care, and research in the future. This highlights

the strong link between AI and healthcare and scientific research, offering a more comprehensive outlook for the advancement of the field. DL, a subset of ML that utilizes deep neural networks to solve problems, has experienced rapid development in recent years due to advancements in algorithm theory, improvements in computer processing technology, and the expansion of data volume. DL has become integral to computer vision, natural language processing, healthcare, autonomous driving, and financial fields (17). Recent literature also underscores the value of DL methods in MRI segmentation, detection, and classification of prostate lesions (18–21). However, some existing literature is constrained by limited data size and slow computational speeds. In this study, we employed DTL, utilizing the pre-trained ResNet50 neural network model from the ImageNet dataset as the initial model. This approach addresses the limited generalization ability resulting from the small original dataset and accelerates model training (22,23). Feature visualization techniques

can be employed to showcase the most critical features and activation patterns within DTL models, providing insights into how the model makes predictions and identifying key features contributing to its performance. For instance, techniques such as class activation mapping (CAM) or gradient-weighted class activation mapping (Grad-CAM) can be utilized to highlight the image regions the model focuses on when making decisions. Currently, image segmentation methods primarily involve manual and automatic segmentation. With the rapid advancement of AI, automatic segmentation is poised to become the predominant method for segmentation in the future due to its ability to ensure segmentation accuracy and consistency, as well as save researchers time (24-28). This study employed manual segmentation of images, which, despite its high time cost, allows physicians to accurately identify lesions and delineate them precisely based on their extensive clinical experience.

Existing prostate detection methods have some value in assessing the aggressiveness of PCa. Prata *et al.* (29) utilized mpMRI to diagnose csPCa by comparing and combining clinical factors and radiomics methods. The results showed that the AUC values for the 2 methods were 0.698 and 0.774, respectively, with a combined AUC of 0.804. Bertelli *et al.* (30) employed MRI-based PI-RADS and radiomics methods to predict PCa aggressiveness, achieving AUC values of 0.625 and 0.682, respectively. However, their diagnostic performance was inferior to the results obtained by the DL model proposed in this study. Li *et al.* (31) showed that the DL model utilizing T2WI and ADC maps demonstrated outstanding diagnostic effectiveness in assessing the aggressiveness of PCa, achieving an AUC value of 0.940. The results of this study are similar to our findings, but currently, there is limited research on investigating the influence of input layers post-focal segmentation. Therefore, this paper aimed to compare the effectiveness of DTL models based on 2.5D and 2D segmentation for the automatic detection of csPCa.

In the process of focal segmentation, utilizing a single layer only provides cross-sectional information, leading to the loss of 3D anatomical information during the training process. This can result in unreliable training outcomes. Takao *et al.* (32) also demonstrated that incorporating continuous information between adjacent layers can mitigate false results and enhance the accuracy of DL models for lesion detection. It is widely recognized that 3D DL-based segmentation offers advantages over 2D methods by leveraging full 3D spatial information and is

already extensively used for 3D medical data. However, 3D models present several technical limitations. Firstly, substantial financial investment is required to obtain high-capacity graphics processing unit (GPU) memory to accommodate the storage and processing of data. Secondly, a large number of cases are necessary, and finally, an abundance of data and parameters may lead to overfitting of results. Furthermore, as the segmentation plane moves further from the central plane, its influence diminishes, resulting in minimal impact. Given the technical challenges associated with 3D models, 2.5D models may offer benefits for small-scale data. 2.5D models (33) encompass sagittal and coronal information that is lacking in 2D models, while being somewhat less complex than 3D models. Hence, in this study, we employed a DTL model constructed using 2.5D segmentation to assess the aggressiveness of PCa. Our approach is akin to the 2.5D network described in prior literature (32,34), which selects the largest lesion layer as the central layer and utilizes the upper and lower layers as input data. This segmentation strategy mirrors that of many other organs, such as the brain (35,36), liver (37), pancreas (38), and kidney (39). Our findings further validate that the diagnostic efficacy of the 2.5D-based DTL model surpasses that of the 2D model when evaluating the aggressiveness of PCa, whether considering a single or a combined sequence. By enhancing the accuracy of automated detection aggressiveness of PCa, this study contributes to achieving earlier diagnosis and personalized treatment plans, potentially leading to a significant improvement in patients' clinical outcomes. Furthermore, it has the potential to impact the advancement of precision medicine and personalized treatment. By more accurately identifying csPCa cases, our model aids in optimizing resource allocation, reducing unnecessary treatments, and enhancing patients' quality of life. In future research, we will aim to explore the extension of the 2.5D DTL model to larger datasets and its application to more diverse patient populations. Additionally, we plan to investigate the integration of additional imaging modalities, such as functional MRI or positron emission tomography (PET)-CT scans, to further enhance the accuracy and reliability of PCa detection. These research directions will help us gain a better understanding of the biological characteristics of PCa and provide clinicians with more dependable auxiliary diagnostic methods.

This study has several notable limitations. Firstly, it was a single-center retrospective study utilizing data from a single institution. To further assess the accuracy,

stability, generalizability, and repeatability of our model, in future research, we will include a broader population, increase sample size and geographic distribution, and conduct external validation using additional data. Secondly, the manual segmentation of lesions may introduce bias due to varying interpretations by radiologists. Future investigations will focus on automatic segmentation methods to address this limitation. Thirdly, this study only included 3 consecutive MRI images of prostate lesions, and the entire 3D lesions were compared when technical conditions permitted. Fourthly, the uneven distribution of cases in this study may have impacted the model's generalizability, particularly across different populations and medical conditions. We attempted to mitigate this issue by stratifying random sampling, but further validation of the model's robustness on more diverse datasets is still necessary. Lastly, the interpretability of the models remains a challenge, and gaining a deep understanding of how models generate predictions is crucial for clinical applications. We intend to explore model interpretation techniques in future work to enhance model transparency and credibility.

Conclusions

We developed a DTL model based on 2.5D segmentation to automatically assess PCa aggressiveness on bp-MRI, which demonstrated improved diagnostic performance compared to the 2D model. The findings indicate that incorporating continuous information between adjacent layers can enhance lesion detection rates and reduce misjudgment rates based on the DTL model.

Acknowledgments

Funding: This work was supported by the Suzhou Science and Technology Development Plan (grant No. SKYD2023240 to N.D.); Suzhou Wujiang District Health Commission "Promoting Health through Science and Education" Project (grant No. WWK202110 to N.D.); Suzhou City "Promoting Health through Science and Education" Youth Science and Technology Project (grant No. KJXW2022077 to S.Y.).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <https://qims.amegroups.com/article/view/10.21037/qims-24-587/rc>

[amegroups.com/article/view/10.21037/qims-24-587/rc](https://qims.amegroups.com/article/view/10.21037/qims-24-587/rc)

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://qims.amegroups.com/article/view/10.21037/qims-24-587/coif>). N.D. reports that this research was supported by the Suzhou Science and Technology Development Plan (grant No. SKYD2023240) and Suzhou Wujiang District Health Commission "Promoting Health through Science and Education" Project (grant No. WWK202110). S.Y. reports that this research was supported by the Suzhou City "Promoting Health through Science and Education" Youth Science and Technology Project (grant No. KJXW2022077). The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are responsible for ensuring that all aspects of the work are investigated and resolved concerning questions regarding accuracy or integrity. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Institutional Review Board of the Suzhou Ninth People's Hospital (No. KYLW2024-022-01). The requirement for individual consent for this retrospective analysis was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin* 2022;72:7-33.
2. Mottet N, 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-2020 Update. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur Urol* 2021;79:243-62.
3. Turkbey B, Rosenkrantz AB, Haider MA, Padhani AR, Villeirs G, Macura KJ, Tempany CM, Choyke PL, Cornud F, Margolis DJ, Thoeny HC, Verma S, Barentsz J, Weinreb JC. Prostate Imaging Reporting and Data System

- Version 2.1: 2019 Update of Prostate Imaging Reporting and Data System Version 2. *Eur Urol* 2019;76:340-51.
4. Chen T, Li M, Gu Y, Zhang Y, Yang S, Wei C, Wu J, Li X, Zhao W, Shen J. Prostate Cancer Differentiation and Aggressiveness: Assessment With a Radiomic-Based Model vs. PI-RADS v2. *J Magn Reson Imaging* 2019;49:875-84.
 5. McBee MP, Awan OA, Colucci AT, Ghobadi CW, Kadom N, Kansagra AP, Tridandapani S, Auffermann WF. Deep Learning in Radiology. *Acad Radiol* 2018;25:1472-80.
 6. Chaudhary K, Poirion OB, Lu L, Garmire LX. Deep Learning-Based Multi-Omics Integration Robustly Predicts Survival in Liver Cancer. *Clin Cancer Res* 2018;24:1248-59.
 7. Tran KA, Kondrashova O, Bradley A, Williams ED, Pearson JV, Waddell N. Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Med* 2021;13:152.
 8. Matsuoka Y, Ueno Y, Uehara S, Tanaka H, Kobayashi M, Tanaka H, Yoshida S, Yokoyama M, Kumazawa I, Fujii Y. Deep-learning prostate cancer detection and segmentation on biparametric versus multiparametric magnetic resonance imaging: Added value of dynamic contrast-enhanced imaging. *Int J Urol* 2023;30:1103-11.
 9. Wu B, Zhang F, Xu L, Shen S, Shao P, Sun M, Liu P, Yao P, Xu RX. Modality preserving U-Net for segmentation of multimodal medical images. *Quant Imaging Med Surg* 2023;13:5242-57.
 10. Liu X, Dong X, Li T, Zou X, Cheng C, Jiang Z, Gao Z, Duan S, Chen M, Liu T, Huang P, Li D, Lu H. A difficulty-aware and task-augmentation method based on meta-learning model for few-shot diabetic retinopathy classification. *Quant Imaging Med Surg* 2024;14:861-76.
 11. Tsuneki M, Abe M, Kanavati F. A Deep Learning Model for Prostate Adenocarcinoma Classification in Needle Biopsy Whole-Slide Images Using Transfer Learning. *Diagnostics (Basel)* 2022;12:768.
 12. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, Gee JC. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging* 2010;29:1310-20.
 13. Zhou J, Zhang Y, Chang KT, Lee KE, Wang O, Li J, Lin Y, Pan Z, Chang P, Chow D, Wang M, Su MY. Diagnosis of Benign and Malignant Breast Lesions on DCE-MRI by Using Radiomics and Deep Learning With Consideration of Peritumor Tissue. *J Magn Reson Imaging* 2020;51:798-809.
 14. Cheng N, Ren Y, Zhou J, Zhang Y, Wang D, Zhang X, Chen B, Liu F, Lv J, Cao Q, Chen S, Du H, Hui D, Weng Z, Liang Q, Su B, Tang L, Han L, Chen J, Shao C. Deep Learning-Based Classification of Hepatocellular Nodular Lesions on Whole-Slide Histopathologic Images. *Gastroenterology* 2022;162:1948-1961.e7.
 15. Abdelmaksoud IR, Shalaby A, Mahmoud A, Elmogy M, Aboelfetouh A, Abou El-Ghar M, El-Melegy M, Alghamdi NS, El-Baz A. Precise Identification of Prostate Cancer from DWI Using Transfer Learning. *Sensors (Basel)* 2021;21:3664.
 16. Kerdvibulvech C, Dong ZY. Roles of Artificial Intelligence and Extended Reality Development in the Post-COVID-19 Era. In: Stephanidis C, Kurosu M, Chen JYC, Fragomeni G, Streitz N, Konomi S, Degen H, Ntoa S. *HCI International 2021 - Late Breaking Papers: Multimodality, eXtended Reality, and Artificial Intelligence. HCII 2021. Lecture Notes in Computer Science()*, Springer, 2021;13095:445-54.
 17. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436-44.
 18. Comelli A, Dahiya N, Stefano A, Vernuccio F, Portoghese M, Cutaia G, Bruno A, Salvaggio G, Yezzi A. Deep Learning-Based Methods for Prostate Segmentation in Magnetic Resonance Imaging. *Appl Sci (Basel)* 2021;11:782.
 19. Meglič J, Sunoqrot MRS, Bathen TF, Elschot M. Label-set impact on deep learning-based prostate segmentation on MRI. *Insights Imaging* 2023;14:157.
 20. Schelb P, Kohl S, Radtke JP, Wiesenfarth M, Kickingereeder P, Bickelhaupt S, Kuder TA, Stenzinger A, Hohenfellner M, Schlemmer HP, Maier-Hein KH, Bonekamp D. Classification of Cancer at Prostate MRI: Deep Learning versus Clinical PI-RADS Assessment. *Radiology* 2019;293:607-17.
 21. Hosseinzadeh M, Saha A, Brand P, Slootweg I, de Rooij M, Huisman H. Deep learning-assisted prostate cancer detection on bi-parametric MRI: minimum training data size requirements and effect of prior knowledge. *Eur Radiol* 2022;32:2224-34.
 22. Liu Y, Li A, Zhao XM, Wang M. DeepTL-Ubi: A novel deep transfer learning method for effectively predicting ubiquitination sites of multiple species. *Methods* 2021;192:103-11.
 23. Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans Med Imaging* 2016;35:1285-98.
 24. You C, Zhao R, Liu F, Dong S, Chinchali S, Topcu U, Staib L, Duncan JS. Class-Aware Adversarial Transformers

- for Medical Image Segmentation. *Adv Neural Inf Process Syst* 2022;35:29582-96.
25. You C, Zhao R, Staib L, Duncan JS. Momentum Contrastive Voxel-wise Representation Learning for Semi-supervised Volumetric Medical Image Segmentation. *Med Image Comput Comput Assist Interv* 2022;13434:639-52.
 26. You C, Zhou Y, Zhao R, Staib L, Duncan JS. SimCVD: Simple Contrastive Voxel-Wise Representation Distillation for Semi-Supervised Medical Image Segmentation. *IEEE Trans Med Imaging* 2022;41:2228-37.
 27. You C, Xiang J, Su K, Zhang X, Dong S, Onofrey J, Staib L, Duncan JS. Incremental Learning Meets Transfer Learning: Application to Multi-site Prostate MRI Segmentation. *Distrib Collab Fed Learn Afford AI Healthc Resour Div Glob Health (2022)* 2022;13573:3-16.
 28. You C, Dai W, Min Y, Staib L, Duncan JS. Bootstrapping Semi-supervised Medical Image Segmentation with Anatomical-Aware Contrastive Distillation. *Inf Process Med Imaging* 2023;13939:641-53.
 29. Prata F, Anceschi U, Cordelli E, Faiella E, Civitella A, Tuzzolo P, Iannuzzi A, Ragusa A, Esperto F, Prata SM, Sicilia R, Muto G, Grasso RF, Scarpa RM, Soda P, Simone G, Papalia R. Radiomic Machine-Learning Analysis of Multiparametric Magnetic Resonance Imaging in the Diagnosis of Clinically Significant Prostate Cancer: New Combination of Textural and Clinical Features. *Curr Oncol* 2023;30:2021-31.
 30. Bertelli E, Mercatelli L, Marzi C, Pachetti E, Baccini M, Barucci A, Colantonio S, Gherardini L, Lattavo L, Pascali MA, Agostini S, Miele V. Machine and Deep Learning Prediction Of Prostate Cancer Aggressiveness Using Multiparametric MRI. *Front Oncol* 2021;11:802964.
 31. Li C, Deng M, Zhong X, Ren J, Chen X, Chen J, Xiao F, Xu H. Multi-view radiomics and deep learning modeling for prostate cancer detection based on multi-parametric MRI. *Front Oncol* 2023;13:1198899.
 32. Takao H, Amemiya S, Kato S, Yamashita H, Sakamoto N, Abe O. Deep-learning 2.5-dimensional single-shot detector improves the performance of automated detection of brain metastases on contrast-enhanced CT. *Neuroradiology* 2022;64:1511-8.
 33. Kim H, Lee D, Cho WS, Lee JC, Goo JM, Kim HC, Park CM. CT-based deep learning model to differentiate invasive pulmonary adenocarcinomas appearing as subsolid nodules among surgical candidates: comparison of the diagnostic performance with a size-based logistic model and radiologists. *Eur Radiol* 2020;30:3295-305.
 34. Xiong Y, Guo W, Liang Z, Wu L, Ye G, Liang YY, Wen C, Yang F, Chen S, Zeng XW, Xu F. Deep learning-based diagnosis of osteoblastic bone metastases and bone islands in computed tomograph images: a multicenter diagnostic study. *Eur Radiol* 2023;33:6359-68.
 35. Wang G, Li W, Ourselin S, Vercauteren T. Automatic Brain Tumor Segmentation Based on Cascaded Convolutional Neural Networks With Uncertainty Estimation. *Front Comput Neurosci* 2019;13:56.
 36. Xue Y, Farhat FG, Boukrina O, Barrett AM, Binder JR, Roshan UW, Graves WW. A multi-path 2.5 dimensional convolutional neural network system for segmenting stroke lesions in brain MRI images. *Neuroimage Clin* 2020;25:102118.
 37. Wardhana G, Naghibi H, Sirmacek B, Abayazid M. Toward reliable automatic liver and tumor segmentation using convolutional neural network based on 2.5D models. *Int J Comput Assist Radiol Surg* 2021;16:41-51.
 38. Zheng H, Qian L, Qin Y, Gu Y, Yang J. Improving the slice interaction of 2.5D CNN for automatic pancreas segmentation. *Med Phys* 2020;47:5543-54.
 39. Kittipongdaja P, Siriborvornratanakul T. Automatic kidney segmentation using 2.5D ResUNet and 2.5D DenseUNet for malignant potential analysis in complex renal cyst based on CT images. *EURASIP J Image Video Process* 2022;2022:5.

Cite this article as: Li M, Ding N, Yin S, Lu Y, Ji Y, Jin L. Enhancing automatic prediction of clinically significant prostate cancer with deep transfer learning 2.5-dimensional segmentation on bi-parametric magnetic resonance imaging (bp-MRI). *Quant Imaging Med Surg* 2024;14(7):4893-4902. doi: 10.21037/qims-24-587