REVIEW PAPER

UROLOGICAL ONCOLOGY

Review of different convolutional neural networks used in segmentation of prostate during fusion biopsy

Maciej Zwolski, Andrzej Kupilas, Przemysław Cnota

Department of Urology and Urooncology, Municipal Hospital No. 4 in Gliwice, Poland

Citation: Zwolski M, Kupilas A, Cnota P. Review of different convolutional neural networks used in segmentation of prostate during fusion biopsy. Cent European J Urol. 2025; 78: 23-39.

Article history
Submitted: Mar. 11, 2024
Accepted: Oct. 16, 2024
Published online: Mar. 21, 2025

Corresponding author
Maciej Zwolski
Department of Urology and Urooncology,
Municipal Hospital No. 4 in Gliwice,
29 Kościuszki St.,
Gliwice, Poland
zwolskimd@gmail.com

Introduction The incidence of prostate cancer is increasing in Poland, particularly due to the aging population. This review explores the potential of deep learning algorithms to accelerate prostate contouring during fusion biopsies, a time-consuming but crucial process for the precise diagnosis and appropriate therapeutic decision-making in prostate cancer. Implementing convolutional neural networks (CNNs) can significantly improve segmentation accuracy in multiparametric magnetic resonance imaging (mpMRI).

Material and methods A comprehensive literature review was conducted using PubMed and IEEE Xplore, focusing on open-access studies from the past five years, and following PRISMA 2020 guidelines. The review evaluates the enhancement of prostate contouring and segmentation in MRI for fusion biopsies using CNNs.

Results The results indicate that CNNs, particularly those utilizing the U-Net architecture, are predominantly selected for advanced medical image analysis. All the reviewed algorithms achieved a Dice similarity coefficient (DSC) above 74%, indicating high precision and effectiveness in automatic prostate segmentation. However, there was significant heterogeneity in the methods used to evaluate segmentation outcomes across different studies.

Conclusions This review underscores the need for developing and optimizing segmentation algorithms tailored to the specific needs of urologists performing fusion biopsies. Future research with larger cohorts is recommended to confirm these findings and further enhance the practical application of CNN-based segmentation tools in clinical settings.

Key Words: fusion biopsy ‹› convolutional neural networks (CNNs) ‹› deep learning ‹› prostate cancer ‹› prostate segmentation

INTRODUCTION

Prostate cancer is the second most common cancer in men of all age ranges, with estimated number of new cases 1,466,718 in 2022 in the world [1]. In Poland, there is an increasing trend in prostate cancer incidence, with forecasts indicating a 55.2% increase in cases compared to the year 2019 [2]. Vital priority in coming years should be to increase the capacity of the healthcare system to diagnose prostate cancer for the early detection of clinically significant prostate cancer (CSPCa) and the implementation of appropriate therapeutic decisions. We posit that optimizing fusion biopsies holds promise as a solution to this challenge. This can be achieved through the integration of deep learning algorithms into the biopsy process. In recent years, there has been rapid development in artificial intelligence (AI) and machine learning due to the implementation of deep learning [3]. It is important to clarify that while ML and DL are related, they are not the same (Figure 1). Machine learning uses input datasets and algorithms to uncover patterns and make predictions. The two primary approaches in machine learning are supervised and unsupervised learning. The key difference lies in the type of data used to train the model. In supervised learning computer is trained on dataset that are similar to the problem at hand.

Once the model learns the relationship between the input and the output, it can classify new unknown datasets and make predictions or decisions based on them. On the other hand, unsupervised learning using unannotated data. Model learns from input data without expected values, and the available dataset does not provide answers to the given task. Thus, the main goal of supervised learning is to predict or classify new, unseen data based on the learned patterns, while unsupervised learning aims to discover hidden patterns, structures, or relationships within the data. Due to the extensive nature of classical machine learning methods and their divergence from the topic of this review, we recommend the article by Kufel et al. [4] for those interested in a more in-depth discussion of this subject.
A comprehensive review of the literature on automatic prostate segmentation in multiparametric magnetic resonance imaging (mpMRI) images reveals a significant focus on assisting radiologists in their image interpretation tasks. While these studies demonstrate the potential of segmentation algorithms to enhance diagnostic accuracy and efficiency, they largely neglect the practical application of these tools within the workflow of urologists performing fusion biopsy procedures.
Our experience indicates that a part of urologists independently delineate prostate contours on mpMRI scans, highlighting the need for segmentation algorithms that seamlessly integrate into their workflow. To fulfill this gap, future research should prioritize the development and evaluation of segmentation tools tailored to the specific needs of urologists during fusion biopsy procedures.
We undertake assess the improvement by AI in the contouring and segmentation of the prostate in MRI used for fusion biopsy and try compare the effectiveness of different algorithms build in CCNs architecture.

The principles of fusion biopsy

The distinction between clinically significant and clinically insignificant prostate cancer does not have a sharp boundary; it is more of a balance between the characteristics of the tumor and the patient himself. CSPCa can cause morbidity or death, but insignificant not. This distinction is crucial because the treatment itself carries the risk of harmful side effects for the patient. Due to the fact that transrectal ultrasound (TRUS)-guided systematic biopsy does not detect up to 20% of clinically significant cancers, a better alternative was sought [5]. Today, the most sensitive method to detect CSPCa is fusion biopsy (Figure 2) [6–8]. Currently, there are three types of Fusion biopsies: visual estimation TRUS-guided biopsy (known also as "cognitive"), in-bore MRI-guided biopsy and software-based co-registration guided biopsy with MRI to ultrasound fusion. It is based on the analysis of mpMRI in search of lesions suspicious of neoplastic features (regions of interest – ROI) and obtaining a biopsy from these lesions. It has been more than 10 years since targeted mpMRI-GB was introduced into the diagnostic pathway for detecting early prostate cancers [9] European Association of Urology (EAU) in its guidelines has approved the higher sensitivity of mpMRI compared to ultrasound-guided biopsy to detect prostate cancer for International Society of Urological Pathology (ISUP) grade ≥2 [10, 11]. Studies indicate that for the ISUP 1 group, systematic biopsy is more sensitive than mpMRI-GB; however, this often leads to overdiagnosis and possible overtreatment [12]. mpMRI consists of T2 high-resolution weighted imaging (T2w) that assesses the content of water in tissues and at least two functional MRI techniques. These two techniques include diffusion-weighted imaging (DWI; Figure 3), which evaluates the diffusion of water molecules between different tissues (from DWI, the apparent diffusion coefficient [ADC], is calculated) and dynamic contrast-enhanced imaging (DCE). DCE assesses the spread of contrast agent in the prostate gland, which, in the case of present tumor tissues, fills them more quickly (due to tumor angiogenesis) and also washes them out more rapidly [13]. T2W images are especially significant as enable the differentiation of the zonal anatomy of the prostate gland – the peripheral zone exhibits high signal intensity, the central zone shows decreased intensity, and the transition zones have heterogeneous intensity. These images also facilitate the evaluation and identification of cancer foci, which most commonly manifest as low intensity within the peripheral zone. DWI and DCE are utilized for further assessment of the

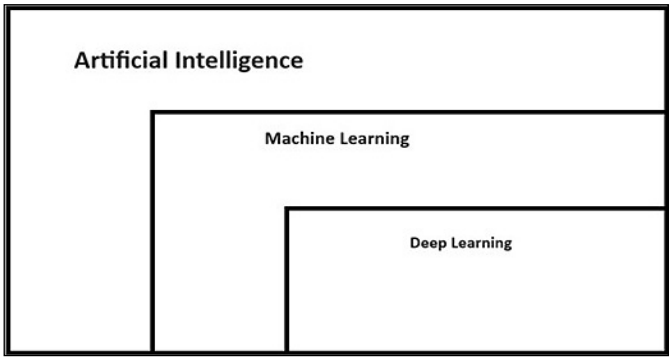

Figure 1. *Interdependence between artificial intelligence, machine learning and deep learning.*

prostate anatomy and ROI. Integral analysis of these imaging modalities significantly increases precision and sensitivity in the detection of neoplastic lesions within the prostate gland [6]. To optimize and unify the evaluation of mpMRI, the American College of Radiology, European Society of Urogenital Radiology (ESUR), and the AdMeTech Foundation developed a standardized method for assessing mpMRI images, commonly known as PI-RADS (Prostate Imaging-Reporting and Data System) [14]. PI-RADS employs a stratification system that assigns values ranging from 1 to 5. It is advised that for lesions categorized as grade 3, a biopsy should be contemplated, whereas for lesions classified with values of 4 and 5, the execution of a biopsy is emphatically recommended [15]. Over the years, the authors achieved the intended effect and enhanced the quality of prostate cancer detection due to PI-RADS [16]. PI-RADS has also been incorporated into representative clinical guidelines and has been used in clinical research as a tool for risk stratification and determining the biopsy pathway [15, 17]. Research assessing the effectiveness of PI-RADS criteria via targeted biopsy has demonstrated significant predictive power in determining the likelihood of prostate cancer, alongside a notable correlation with the Gleason score. The current version in use is PI-RADS 2v1.

The mentioned fusion biopsy techniques mainly differ in the way the needle is targeted in the ROI (Figure 4). In the cognitive method, the operator analyzes prostate gland map and, based on their own experience, collects samples from the prostate. This is a low cost method, but the learning curve is high. In the in-bore biopsy method, the biopsy is performed while the patient lies inside the MRI scanner. This method's advantage lies in its minimal image distortion detection, while its drawbacks include the high cost associated with prolonged MRI usage and the limited availability of this testing form. Prince and his research group in their study demonstrated a higher sensitivity to detect CSPCa through in-bore compared to fusion MRI-targeted biopsy [18]. The final method is software-assisted fusion biopsy, which involves manually or automatically, delineating the entire prostate on mpMRI and identifying ROIs. This phase, if manually, is not infrequently laborious, tedious,

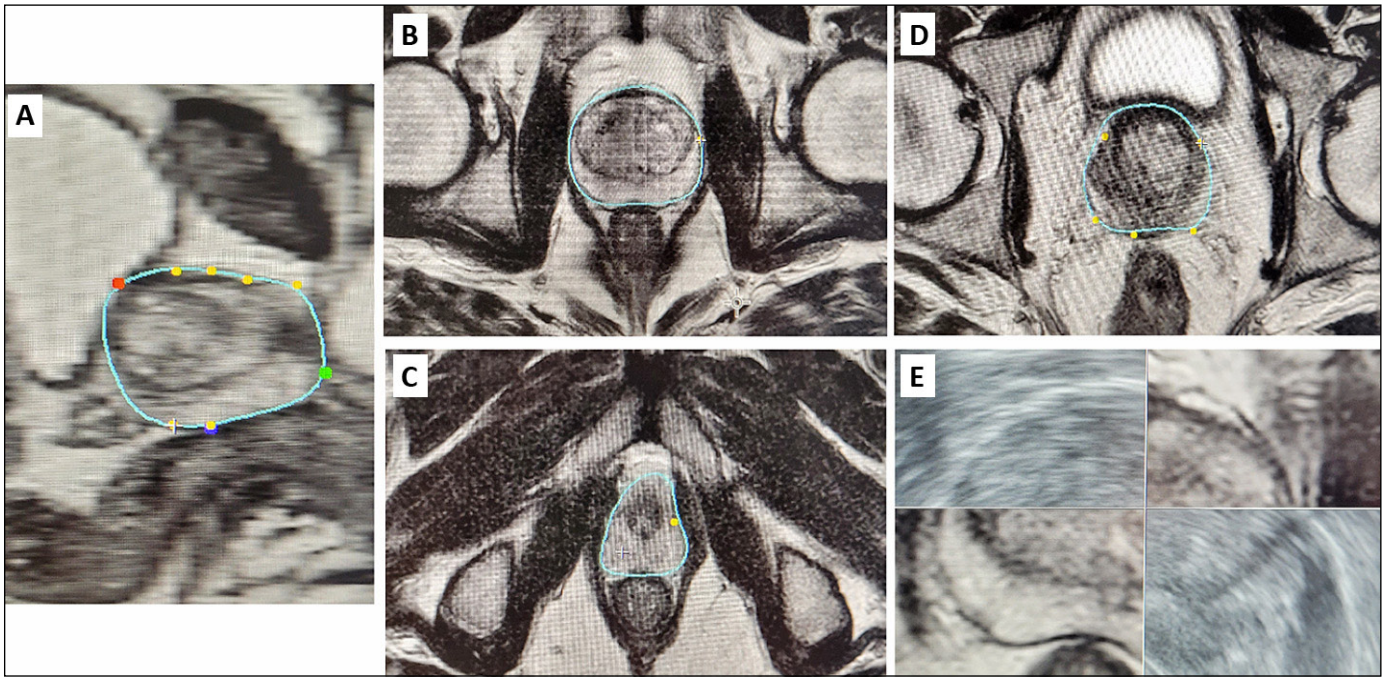

Figure 2. *Fusion performed on Urostation (Koelis). Contouring:* **A)** *The first step is to mark 3 points that generate the original outline: green – the top of the prostate gland, red – the base above the entrance to the urethra, and blue – the back wall of the gland. The software then generates the primary outline which we already contour by marking yellow points and adjusting it to the outline of the gland:* **B)** *mid,* **C)** *apex,* **D)** *base. On the finished outline we mark changes using T2 and ADC or DWI sequences, for more discrete lesions we use contrast sequences. Trinity software allows overlapping of images. The next step is to repeat these steps, but this time we work on the ultrasound image from the endorectal transducer generated live.* **E)** *After creating two contours, the apparatus performs a fusion of the images, which we check and confirm, and after confirmation we receive the superimposed lesions on the ultrasound image and can proceed to the actual part of the biopsy.*

and time-consuming, highly depending on clinician experience and skills [19].
Concurrently, TRUS is performer in real-time, and with the aid of software, the mpMRI images are merged. This co-registration is crucial step in the Fusion process, where the three-dimensionally (3D) segmented area of the prostate on MRI images is „combined" or aligned with the 3D captured area of the prostate on real-time US images. Following the completion of the co-registration, the process advances to the biopsy phase of the prostate gland. Two methods of image registration between mpMRI and real-time TRUS are available on the market: static, which directly overlays images without considering changes in the prostate visualization caused by compression from the US probe or the patient's positioning, and elastic, where the software accommodates differences in the gland's chape between MRI scans and TRUS.
The biopsy fusion stations available on the market are presented in Table 1.

The principles of deep learning

Deep learning is powered by neural networks, which mimic the information-transmitting behavior of human neurons. Similar to our brains, deep learning begins with the introduction of information, such as an image or sound. This information is then passed through the network, analyzed by successive layers until an output is generated. Each layer comprises numerous neurons that process the input data.

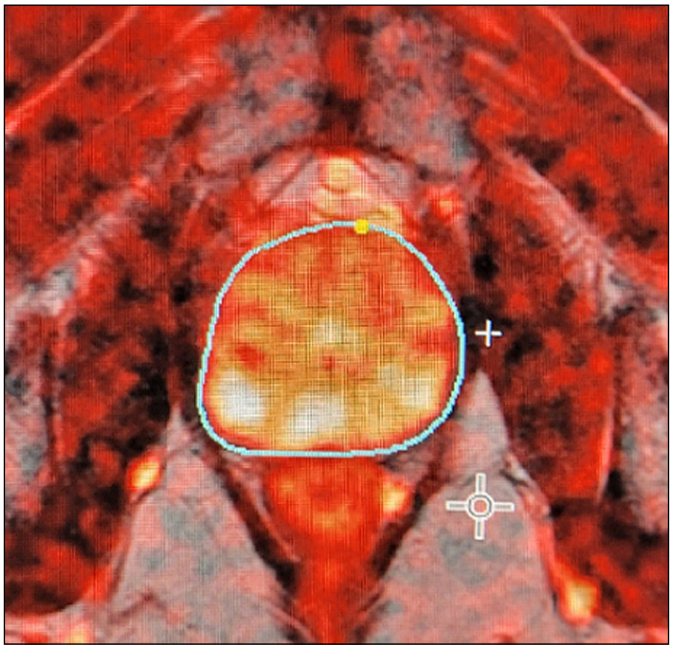

Figure 3. *Diffusion-weighted imaging.*

The number of hidden layers distinguishes between artificial neural networks (ANNs), which have one hidden layer, and deep neural networks (DNNs), which have multiple hidden layers (Figure 5).
Regardless of the training method employed, machine learning algorithms improve their performance on new data as they are exposed to more training examples. The primary distinction between ANNs and DNNs lies in their data requirements. DNNs demand significantly more input data, as they must learn to classify data independently. In contrast, ANNs rely on pre-processed and labeled data. One of the examples of DNN are CNNs [4].
Currently, the most widespread architecture in the medical industry is CNNs [21], which are primarily used for the analysis of advanced imaging, including MRI scans. CNNs are inspired by the functioning of the animal visual cortex, which attempts to emulate hierarchical, layered information processing [22]. A CNN, in contrast to traditional techniques, allows unsupervised learning and selects its own feature maps automatically while training. The basic unit of the neural network consists of: an input

Table 1. *The biopsy fusion stations available on the market*

Biopsy fusion stations	Producer
UroNav	Philips
BiopSee	MedCom
BioJet	DK Technologies
Urostation	Koelis
Semirobotic Artemis	Eigen
Virtual Navigator	Esaote

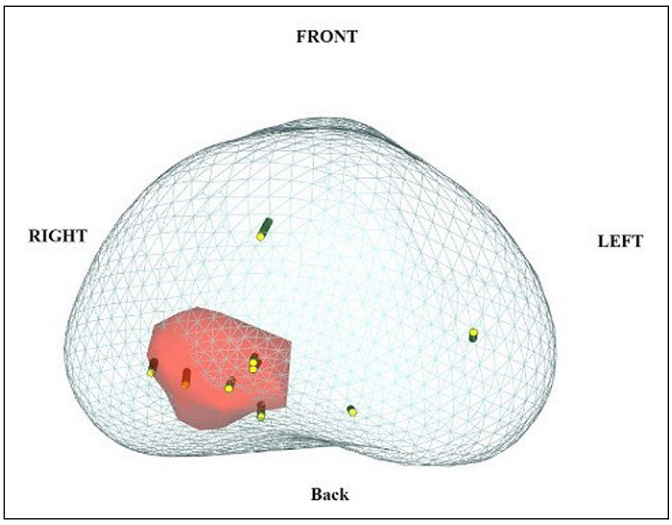

Figure 4. *Example of fusion.*

layer, connected "nodes" (equivalent of neurons) called hidden layers, and an output layer. Within the scope of hidden layers are convolutional layers (conv), pooling layers, flatten layer, and fully connected (FC) layers. Convolutional layers are made up of filters/kernels and, similar to neuronal cells, aim to extract features from the image that help in understanding the image. CNNs learn these filters, which convert the given image to extract its features like edges, colors or textures. Then, the output of the conv layer goes to an activation function layer where the ultimate goal is to map the representation in the input to a different output as per the requirements of the task. The activation function is nonlinear because without this, a neural network, no matter how many layers it has, would behave like a single layer perceptron, capable of learning only linear data. Then the image size is reduced in the pooling layer. This layer may realize the reduction using max or average pooling by converting a square (typically two by two pixels) into one pixel. In this way, overfitting is avoided, which is the phenomenon of the model being too closely fitted to the training data, critically important for maintaining the model's ability to generalize to previously unseen data. Such condensed data then go to a layer where neurons receive input from all the neurons of the previous layer. This is the FC layer, which can be likened to bridge between the features extracted by the previous layers and the final outcome/prediction, such as identifying the contours of the prostate in the mpMRI case [23]. The performance of CNNs can be influenced by various factors, with the most crucial ones being the selection of activation functions and the number of hidden layers. Networks trained on extensive datasets encompassing phenotypic and pathophysiologic variations will inherently be more robust compared to networks trained on limited datasets lacking such variety.

The ORSI Academy presents in an understandable manner the application of CNNs in creating software that assists operators during robotic surgeries as well as for training purposes [24]. If anyone is interested in more comprehensive explanation of deep learning, we highly recommend review article created by Alzubaidi et al. [21]

Analysing the experiments conducted by various researchers, we noticed that most of algorithms to a greater or lesser extent drew from the achievements of Ronneberger, Fischer, and Brox, i.e., CNNs named U-Net [25]. U-Net is a powerful tool for identifying and separating different parts or unusual areas in medical images (Figure 6). It works by analyzing images at various levels of detail. This architecture consists of an encoder-decoder structure with skip connections. The encoder part of the network gradually reduces spatial dimensions while increasing the number of feature channels through successive convolutional and pooling layers. This process captures hierarchical feature representations at multiple scales. The decoder part, on the other hand, performs upsampling and convolution operations to progressively reconstruct the segmented output. Skip connections are established between corresponding encoder and decoder layers to enable the integration of both low-level and high-level features, aiding in the preservation of fine details during the segmentation process.

One of the limitations of using these algorithms is their requirement for a vast database for training. Another limitation is the significant computational power of computer hardware that must be provided for CNNs to operate. Nowadays, researchers are trying to overcome these limitations by data augmentation, transfer learning, and creating more efficient network architectures [26–28].

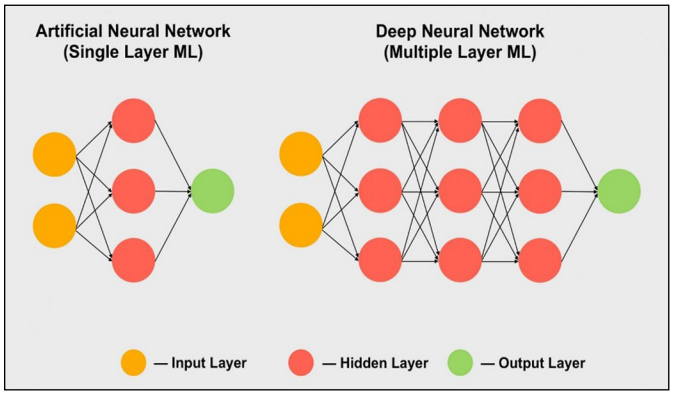

Figure 5. *A simplified example of artificial neural network and deep neural network [20].*

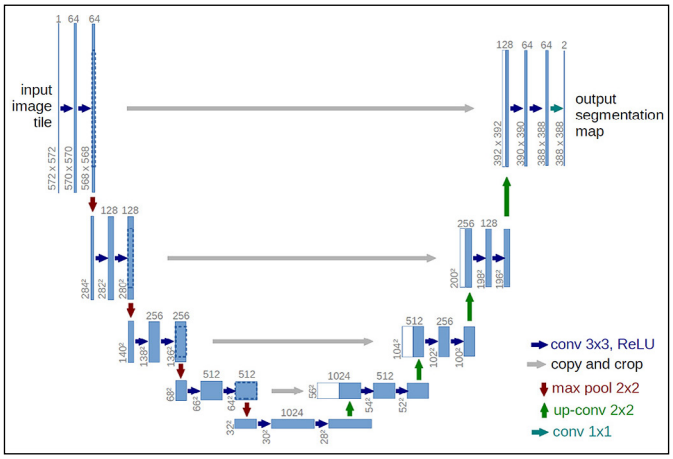

Figure 6. *U-Net architecture (example for 32 × 32 pixels in the lowest resolution) [25].*

Consumer platforms

It is important to highlight that various software platforms utilizing deep learning algorithms for the automatic contouring of the prostate gland are available on the market. However, the majority of these platforms are designed for evaluating mpMRI based on the PIRADS v 2.1 scale by radiologists, rather than as integral components of fusion biopsy workstations (Table 2).

MIM Symphony Dx™ utilizes AI to automatically segment the prostate gland and identify ROIs. Its automatic segmentation is based on multi-atlas segmentation, wherein a collection of MRI scans with manually segmented prostates by experts serves as the foundation. The atlas subjects are registered to the test case using a normalized intensity-based free-form deformable algorithm. Deformable image registration (DIR) is a technique used in medical imaging to align two or more images of the same anatomy acquired from different perspectives or at different times. The volumes of interest (VOIs) are then transformed to the test case using this deformable registration and combined using Simultaneous Truth and Performance Level Estimation (STAPLE) methods. STAPLE considers the original contours and computes a probabilistic estimate of the true representation of their combination. Atlas performance declined when tested with images of differing contrast and MRI vendors. A study published in 2018 evaluated an automatic multi-atlas-based segmentation method for generating prostate contours. The similarity coefficients between the selections of one and two experts served as benchmarks for assessing the quality of the atlas (expert radiation oncologists with 10 and 26 years of experience). The segmentation results were DSC (mean $\pm\sigma$) 0.81 $\pm$0.15 and Hausdorff distance (mm) (mean $\pm\sigma$) 2.7 $\pm$1.9. Segmentation time averaged less than 90 seconds per patient [29].

Philips DynaCAD automates segmentation for clinical use. However, its performance remains unclear due to the proprietary nature of the software and the lack of performance metrics such as the Dice score. According to the official website, UroNav does not include an algorithm for automatic segmentation but instead imports segmented images from the DynaCAD system. A significant limitation is that using this functionality requires the purchase of both products.

AI-Rad Companion also provides automated segmentation of the prostate, but no scientific reports on segmentation speed or accuracy have been found.

It is crucial to mention that the aforementioned software primarily serves to contour the prostate on desktop computers or laptops and not on fusion biopsy workstations. The literature includes review papers on AI-based software for automatic prostate segmentation, but these are predominantly intended for radiologists. We have selected software's directly associated with companies producing fusion biopsy platforms as these are directly related to the work of urologists.

ProMap Contour™ is an added software option to the Koelis Trinity® system. There are no articles assessing the time required for automatic segmentation, its accuracy, or validation. The advantage of this product lies in its integration into a platform used for performing fusion biopsies, enhancing the workflow for urologists.

ProFuse CAD Semirobotic Artemis (Eigen). This software is part of the Siemens Healthineers fusion biopsy platform and offers automated prostate gland contouring and segmentation. It utilizes deep learning algorithms to generate accurate contours on T2-weighted MRI scans. We were unable to locate any studies that validate its accuracy and speed in segmentation

In the reviewed literature, CNN-based algorithms achieved comparable accuracy in prostate segmentation tasks compared to MIM Symphony Dx™. However, CNN-based algorithms demonstrated significantly faster segmentation times in the four studies that directly assessed this metric. While MIM Symphony Dx™ achieves competitive accuracy, its segmentation speed is slower than CNN-based approaches. The segmentation speed of MIM Symphony Dx™ is reported to be under 90 seconds per patient, whereas CNN-based algorithms achieve segmentation times in the range of seconds. It is important to note that the studies referenced here have a relatively small sample size, and further research with larger cohorts is warranted to solidify these findings.

MIM Symphony Dx™ relies on a multi-atlas segmentation technique, which may be less generalizable to MRI scans with significant variations in contrast or acquired from different vendors compared to CNNs.

Table 2. *Software along with their companies*

MIM Symphony Dx™	MIM Software Inc.
Philips DynaCAD	Philips
AI-Rad Companion	Siemens Healthineers
ProMap Contour™	Koelis Trinity®
ProFuse CAD	Semirobotic Artemis (Eigen)

LITERATURE REVIEW

Search strategy

To conduct a comprehensive review, we search literature across PubMed and IEEE Xplore. These sources were selected because of their extensive publications of research in this area of study. Search included observational, randomized and non-randomized studies. It was limited to articles in English language with an abstract and published in peer-reviewed journals during the last 5 years in open access. Search was performed by title or abstract, utilizing keywords, as follows: for PubMed: (imaging-guided biopsy OR fusion biopsy) AND (mpMRI OR T2W) AND (deep learning OR CNN) AND (prostate segmentation OR prostate OR prostate contour), for IEEE Xplore ("All Metadata":fusion biopsy) AND ("All Metadata":prostate contour) OR ("All Metadata":prostate segmentation) AND ("All Metadata":convolutional neural network) Filters Applied: 2019–2024. The last search was conducted on 20.05.2024. The search was supplemented by checking published reviews, and their references. Exclusion criteria were reviews, letters, non-peer reviewed articles, conference abstracts and proceedings. Our method for identifying and evaluating data complied with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) 2020 statement and checklist.

Data extraction

The full texts of the qualified papers chosen for review were acquired, and the reviewers independently collected all study data, resolving disagreements via consensus. The references, year of publication, study setting, ML approach, improvement of segmentation time, performance measures used and accuracy attained were all extracted for every included paper, and comparative analyses were conducted on the extracted dataset. Two independent authors (MZ, PC) screened and extracted the studies. A third reviewer (AK) was consulted in the event of discordance resulting in agreement in all instances.

Risk of bias

Our study assessment aims to evaluate the methodological quality and potential sources of bias that could influence the reported findings. For instance, studies that rely solely on single-center datasets or imbalanced class distributions may introduce biases that affect the model generalizability. Additionally, the lack of clear documentation of preprocessing steps, considerations for model-fitting problems and hyperparameter tuning could hinder reproducibility.

Comparative analysis

Database searches provides 44 results. After title and abstract screening, the full texts of 31 reported studies were analyzed, but only 12 were found eligible for inclusion. The excluded articles are outlined in Supplementary Table 1. A flowchart based on PRISMA 2020 statement is shown in Figure 7.
The results section is divided into three parts: segmentation accuracy, validation, and improvement of segmentation time.

Segmentation accuracy of prostate

Results in the segmentation accuracy of prostate category are shown in Table 3.
The Dice similarity coefficient (DSC) is a diagnostic tool that assesses the degree of similarity between two sets of data. The DSC score ranges between 0 and 1, where 0 indicates no similarity, and 1 represents perfect overlap. The Hausdorff distance to evaluate the results of the algorithm developed. The Hausdorff distance is a measure used to quantify

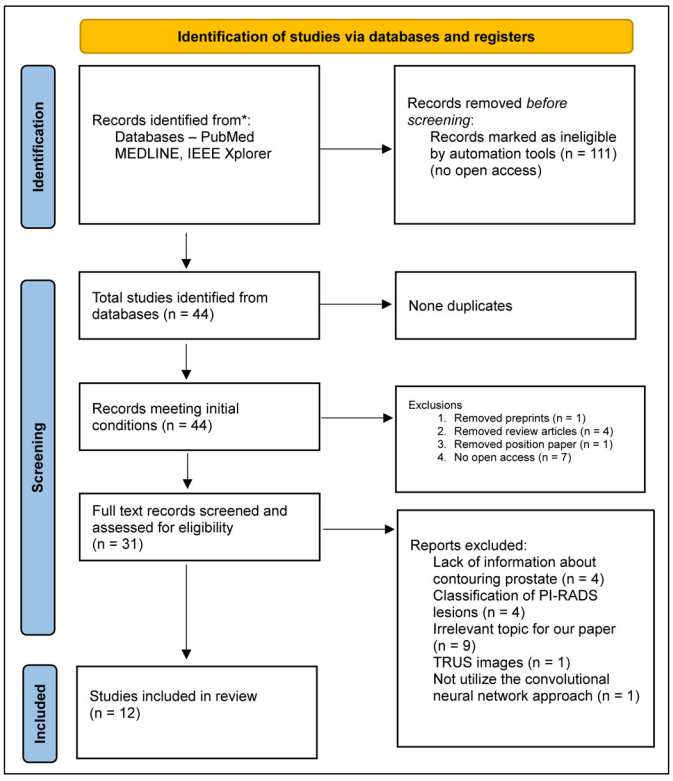

Figure 7. *PRISMA numerical flow guideline for systematic review employed in this study.*

the difference between two sets of points. The Pearson correlation coefficient values range from –1 to 1, with 1 and –1 denoting a strong linear relationship between two variables, while a value proximate to 0 indicates an absence of linear correlation between the variables.

DSC ranged from 0.74 to 0.94, indicating good to excellent agreement between CNN segmentation and expert annotations.

While DSC appears as the predominant metric for evaluating prostate segmentation algorithms in the reviewed studies, a closer examination reveals a lack of uniformity in the utilization of other validation metrics. Within the twelve studies reviewed, only two employed the Hausdorff distance, and two others utilized the Pearson correlation coefficient.

Among studies reviewed, only one compared the accuracy and time efficiency between the algorithm and human performance in segmenting the prostate [31], while another compared the consistency of humans and AI in delineating the prostate [40].

Soerensen et al. [31] created an algorithm demonstrates a narrow range in the DSC test <0.90, indicating consistency in the algorithm's accuracy in segmentation in various cases. Besides testing the algorithms on an internal database of T2-MRI scans, for further evaluation of the code's generalization capabilities, the tested its efficacy on two publicly available datasets: The results demonstrated a performance of 0.87 ±0.05 on the PROMISE12 dataset and 0.89 ±0.05 on the NCI-ISBI dataset. Researchers confirmed a high level of precision of automated segmentation obtained 2.8 mm Hausdorff distance. The culmination of the experiment was the application of ProGNet in practice, during work in the clinic. ProGNet (average DSC = 0.93 ±0.03) significantly outperformed radiology technicians (average DSC = 0.90 ±0.03, p <0.0001) in an 11-case set of prospective fusion biopsy tests.

Liu et al. [40] developed novel CNN for automatic segmentation of the prostatic transition zone (TZ) and peripheral zone. DSC was used to evaluate the segmentation performance. DSCs for peripheral

Table 3. *Segmentation accuracy of prostate*

Investigators	Mean DSC	Human experts (mean DSC)	Mean HD	Pearson correlation coefficient
Bardis et al. [30]	0.940	None	Not used	0.981 (95% CI: 0.966–0.989)
Soerensen et al. [31]	0.92 ±0.02 (100) (retrospective internal test) 0.93 ±0.03 (prospective 11 patients)	DSC = 0.89 (retrospective internal test) 0.90 ±0.03 (prospective 11 patients)	Algorithm reduced the mean HD by 2.8 mm compared to the radiology technicians	
Ushinsky et al. [32]	0.898	None	Not used	0.974
Palladino et al. [33]	0.7773 for internal dataset 0.7709 for external dataset	None	Not used	
Hassanzadeh et al. [34]	0.873	None	Not used	
Ren and Ren [35]	0.9394	none	Not used	
Huang et al. [36]	0.8782	None	10.9443	
Qian et al. [37]	0.908 (for PROMISE12) 0.892 (for ProstateX)	None	9.87 (for PROMISE12) 10.45 (for ProstateX)	
Su et al. [38]	0.9071	None	Not used	
Qian et al. [39]	0.8912 (for ProstateX)	None	Not used	
Liu et al. [40]	PZ 0.74 ±0.08 internal testing dataset TZ 0.86 ±0.07 internal testing dataset PZ external data 0.74 ±0.07 TZ external data 0.79 ±0.12	Expert 1 vs expert 2 PZ 0.71 ±0.13 p <0.05 TZ 0.75 ±0.14 p <0.051 when taking expert 1's annotations as the ground truth	Not used	
Comelli et al. [41]	ENet showed a mean DSC of 0.9089 ±0.0387, U-Net of 0.9014 ±0.0469, and ERFNet of 0.8718 ±0.0644	None	Not used	

[1]Authors used Wilcoxon Signed-Rank Test

DSC – Dice similarity coefficient; ENet – efficient neural network; ERFNet – efficient residual factorized convNet; HD – Hausdorff distance; PZ – peripheral zone; TZ – transition zone

zone (PZ) and prostatic transition zone (TZ) were 0.74 ±0.08 and 0.86 ±0.07 in the internal testing dataset (ITD) respectively. In the external testing dataset (ETD), DSCs for PZ and TZ were 0.74 ±0.07 and 0.79 ±0.12, respectively. The inter-reader consistency (Expert 2 vs Expert 1) were 0.71 ±0.13 (PZ) and 0.75 ±0.14 (TZ)

Validation

Results in the validation category are shown in Table 4.

The most commonly used technique for training and validation was k-fold cross-validation. In each iteration of this procedure, the dataset is divided into k subsets, with k-1 subsets used for training and the remaining subset used for validation. This process is repeated k times, ensuring that each subset is used exactly once for validation. This approach aims to optimize the utilization of the available dataset for robust model evaluation.

Number of epochs ranged widely, from 10 to 500.

The studies employed a diverse range of databases, encompassing both in-house image collections and publicly accessible datasets such as PROMISE12 and ProstateX:

- In-house database 6 studies,
- PROMISE12 6 studies,
- ProstateX 3 studies,
- Other datasets 2 studies.

PROMISE12 data set contained prostate MR images with well-curated prostate organ labels.

The definition of "ground truth" (gold standard) varies across studies. Certain investigations utilize radiologists, whereas others rely on urologists.

Improvement of segmentation time

Results in the improvement of segmentation time category are shown in Table 5.

Soerensen et al. [31] describe a promising result in their experiment: after a 20-hour training session, the ProGNet algorithm segmented a single case in roughly 35 seconds, resulting in the completion of a database of 100 clinical cases in about 1 hour. For comparison, radiologists took an average of 10 minutes per case. Completing the entire database took about 17 hours. ProGNet can save up to 17 times more time according to this trial. A limitation of this result may be the comparison of the algorithm with radiology technicians. In many facilities around the world, the aforementioned segmentation is performed by physicians with specialties in urology or radiology.

In another study, Ushinksy et al. modified traditional 2D-2D U-Net technique creating hybrid 3D-2D U-Net to assess its clinical utility in identifying and segmenting the entire prostate gland in mpMRI images [32]. (The algorithm is similar with developed by another researcher team Michelle Bardis et al. However, Michelle Bardis adapted their network to segment not only the entire prostate but also PZ and TZ by producing 3 separate, collaborative CNNs). Algorithm was verified by a board-certified urologist. On each axial T2WI included in the training database, two specialist radiologists contoured the prostate gland to create ground truth for assessing the segmentation quality by the algorithm. Research team from United States implemented for the quantitative evaluation of the segmentation overlap of the prostate image post manual and automatic segmentation, the authors employed the DSC. Additionally, the Pearson correlation coefficient was utilized to assess the neural network in predicting prostate volume. Findings are documented in Table 5. The authors noted that manual segmentation of the prostate requires approximately 5–10 minutes per patient. Accordingly, the neural network was able to segment the prostate in 9.4 seconds per patient.

Comelli et al. [41] presents three deep-learning approaches, namely U-Net, efficient neural network (ENet), and efficient residual factorized convNet (ERFNet), whose aim is to tackle the fully-automated, real-time, and 3D delineation process of the prostate gland on T2-weighted MRI. They found that ENet and U-Net are more accurate than ERFNet, with ENet much faster than U-Net. Specifically, ENet obtains a segmentation time of about 6 seconds using central processing unit (CPU) hardware to simulate real clinical conditions where graphics processing unit (GPU) is not always available.

Despite the absence of automatic segmentation speed assessment in the remaining eight studies, they provide significant insights into the accuracy of these algorithms, highlighting the potential for integrating CNNs into clinical practice.

DISCUSSION

Based on the studies and comparisons we collected, we have reached several conclusions, summarized below.

In comparative experiments involving ANNs and human experts, the algorithms demonstrated comparable, and in some cases even superior, performance in prostate segmentation (Soerensen et al. [31] and Liu et al. [40]). In studies where algorithms were solely compared to ground truth annotations made by radiologists or urologists, the DSC did not fall below 74%. It is important to note, however, that a significant limitation of these experiments is

Table 4. *Validation*

Investigators	Ground truth	Number of epoches	Type of used database	Evaluation	Data augmentation
Bardis et al. [30]	A board-certified subspecialty-trained abdominal radiologist with more than 10 years of experience	U-NetA, U-NetB, and U-NetC were trained for 50,000, 18,000, and 3,800 iterations U-NetA trained for approximately 7 hours, while U-NetB and U-NetC each trained for 5 hours[1]	Own database 242 T2-weighted images (6,292 axial images)	Included datasets were split into 60% training, 20% validation, and 20% test sets for model development	Not explicitly stated
Soerensen et al. [31]	Urologic oncology expert with 7 years of experience	150	Own database: 905 T2-MRI and 26 T2-MRI (PROMISE12) 30 T2-MRI (NCI-ISBI)	A deep learning model, ProGNet, was trained on 805 cases. ProGNet was retrospectively tested on 100 independent internal and 56 external cases. Algorithm was prospectively implemented as part of the fusion biopsy procedure for 11 patients	Not explicitly stated
Ushinsky et al. [32]	Two specialist radiologists images who were verified by a board-certified urologist	Not provided	Own database: 299 MRI (7,774 images)	Five-fold cross-validation	Not explicitly stated
Palladino et al. [33]	Ground truth was created using a semi-automatic procedure	Not provided (less than 2 hours)	ProstateX as training set (internal) and own dataset: MRI Local Hospital – 6 patients (external)	20% of the images were used for testing and the remaining images were split with a ratio of 0.8 in training set and 0.2 in validation set	Not explicitly stated
Hassanzadeh et al. [34]	Experienced readers at each center used a contouring tool then A second expert (C.H.) who has read more than 1,000 prostate MRIs, to make sure they were consistent	25 epochs	PROMISE12 [] 50 MRI (1377 image slices) volumes for training, 30 MRI volumes for testing (without grand truth)	Ten-fold cross-validation	Random rotation within a 10-degree range, horizontal flip, vertical flip, zoom, horizontal and vertical translation, and elastic transformation (result = 150,000 slices)
Ren and Ren [35]	Three clinicians with 5 years of experience spent 3 months annotating the segmented region. The other three specialists examined and corrected the annotated regions	400–500 epoch	Own database: 180 patients (122 healthy and 58 prostate cancer patients) scanned on the GE 3.0T 750 MR	Five cross-validation approach	Rotation, random merging, zooming in and out
Huang et al. [36]	Experienced readers at each center used a contouring tool then A second expert (C.H.) who has read more than 1,000 prostate MRIs, to make sure they were consistent. I2CVB – An experienced radiologist	200 epoch	PROMISE12 50 MRI (1,377 image slices) volumes for training, 30 MRI volumes for testing and I2CVB	Not explicitly stated	Rotation, zooming in and out, random horizontal or vertical flip, random movement along the X-axis and Y-axis
Qian [37]	Experienced readers at each center used a contouring tool then A second expert (C.H.) who has read more than 1,000 prostate MRIs, to make sure they were consistent (PROMISE12) In order to verify the generalization performance of the prostate segmentation algorithm, prostate masks were marked on the MR images of 40 randomly selected patients from the ProstateX dataset by a professional radiologist	10	PROMISE12 dataset and ProstateX dataset	Five-fold cross-validation	Not explicitly stated

Table 4. *Continued*

Investigators	Ground truth	Number of epoches	Type of used database	Evaluation	Data augmentation
Su et al. [38]	Experienced readers at each center used a contouring tool then A second expert (C.H.) who has read more than 1,000 prostate MRIs, to make sure they were consistent (PROMISE12)	Not explicitly stated	PROMISE12	Not explicitly stated	Elastic deformation, rotation with 90, 180, 270, and flip
Qian et al. [39]	Randomly selected 35 patients' MR images to label their prostate contours	60	ProstateX and PROMISE12	Five-fold cross-validation	Geometric transformations,local disturbance and combined disturbance Not explicitly stated
Liu et al. [40]	PROSTATEX – two MRI research fellows, where the contours were later cross-checked by both genitourinary (GU) radiologists (10–15 years of post-fellowship experience interpreting over 1,000 prostate mpMRI) and clinical research fellows WMHP – two clinical genitourinary (GU) radiologists research fellows, supervised by expert GU radiologists	100	PROSTATEX (internal testing) and WMHP (external testing)	Five-fold cross validation	Flipped horizontally, rotated randomly between [–5°, 5°], elastic transformations
Comelli et al. [41]	A set of trained clinical experts (FV, MP, GC, and GS authors) hand segmented the prostate region. The simultaneous ground truth estimation STAPLE tool (Warfield, Zou i Wells, 2004) was used to combine the different segmentations from the clinical experts in a consolidated reference	100	Own data set: 85 axial T2W	Five-fold cross-validation	Six different modalities – not explicitly stated

[1]An epoch refers to one complete pass through the entire training dataset. Iterations, on the other hand, can represent individual training steps within an epoch. The number of iterations per epoch can vary depending on the batch size (number of images processed at once).

Table 5. *Improvement of segmentation time*

Author, reference	Time
Bardis et al. [30]	The three U-Nets completed their tasks of bounding box creation, prostate segmentation, and prostate zone classification in 0.196 second, 0.226 second, and 0.219 second
Soerensen et al. [31]	35 seconds to segment each case
Ushinsky et al. [32]	9.4 seconds per patient
Palladino et al. [33]	Not tested
Hassanzadeh [34]	Not tested
Ren and Ren [35]	Not tested
Huang et al. [36]	Not tested
Qian [37]	Not tested
Su et al. [38]	Not tested
Qian et al. [39]	Not tested
Liu et al. [40]	Not tested
Comelli et al. [41]	ENet – 6.17 seconds ERFNet – 8.59 seconds U-Net – 42.02 seconds

ENet – efficient neural network; ERFNet – efficient residual factorized convNet

the lack of a human segmentation accuracy benchmark on the datasets used to evaluate the algorithms. Integration of CNN-based segmentation into clinical practice requires validation in larger cohorts. Of the studies available to us, four assessed the speed of prostate segmentation from T2-weighted images. In each case, segmentation of a single instance took less than one minute, with the fastest algorithms segmenting the prostate in approximately 0.3 seconds. In comparison, manual segmentation is reported by authors to take 5–10 minutes per patient [32], 10 minutes per patient [31]. From our experience, manual segmentation takes about 4–10 minutes per patient, depending on individual conditions and the quality of mpMRI. Based on this information, it can be estimated that AI-based CNNs can segment the prostate 800 to 2,000 times faster than humans. It is also important to remember that the speed of the algorithm is primarily dependent on the computational power.

Challenges and limitations

In the course of our investigation and analytical examination of deep learning algorithms applied

to prostate delineation and automated segmentation in mpMRI, we identified a notable lack of uniformity in the methodologies employed for comparing results. The considerable heterogeneity in the assessment protocols for the efficacy of neural networks significantly impedes the qualitative comparison of convolutional neural networks (CNNs). Studies utilized different datasets and evaluation metrics, making direct comparisons challenging. This diversity often leads to the inability to compare studies directly, including statistical data, thus impeding meta-analysis. As of now, no guidelines or standardization have been universally accepted. However, some studies have designed experiments specifically for network comparison [44].

We observed a significant lack of studies implementing deep learning algorithms for automatic prostate contouring in the practice of urologists performing fusion biopsies. Despite numerous studies on mpMRI contouring, these are predominantly intended as tools for radiologists, not for enhancing the workflow of urologists. In our review, we found only two studies that described the integration of deep learning algorithms into fusion biopsy stations, assessing their actual impact and accuracy in prostate contouring [31, 33].

Evaluating automated prostate delineation for fusion biopsy include significant inter-institutional variability in clinical workflows. Established protocols often entrust the delineation of prostate boundaries on T2-weighted MRI (T2W MRI) to radiologists or radiology technicians. However, at our institution, this task is handled by the urologists performing the biopsy. While the literature extensively explores deep learning algorithms for prostate delineation by radiologists, there is little research on the use of these algorithms by urologists. The final decision on the suitability of the contour for mpMRI-TRUS fusion biopsy rests with the urologist. Incorporating AI at this stage with sufficient accuracy and stability would significantly improve the urologist's workflow.

Another constraint pertains to the inherent complexity of this domain. The development and implementation of deep learning algorithms for MRI-US fusion require collaborative efforts among specialists from various scientific disciplines. This interdisciplinary team should encompass software developers, physicians specializing in urology and/or radiology, as well as professionals in computer vision, biomedical engineering, or other relevant fields such as mechatronics.

A limitation of our review is the relatively small sample size [12] which may limit generalizability to a broader population and preclude meta-analysis. To mitigate these limitations, we employed a rigorous search strategy and carefully considered the inclusion and exclusion criteria for studies. We also acknowledge that further research with larger sample sizes is needed to confirm our findings.

The last notable limitation of this review is the lack of open access to a significant number of articles. By restricting access to knowledge and expertise, it hinders comprehensive reviews, slows scientific progress, and creates an uneven playing field.

Future directions for research and clinical practice

As proven above, there is a significant gap in research on the implementation of CNNs for prostate segmentation during MRI-targeted TRUS biopsies and their integration into clinical practice. That is why the authors of this review highly recommend more experiments in this field. Benefits to clinical practice would include speeding up the biopsy process. Implementing automatic prostate segmentation to optimize the process would allow time for one more patient within the same timeframe as manual segmentation. The number may not seem staggering, but multiplying it by the number of facilities performing this examination would result in a satisfactory outcome in expanding the bottleneck. It should be noted that the rate of "machine" evaluation is predominantly determined by the computational capacity of the computer utilized. Hence, the greater the computing power dedicated to prostate segmentation, the quicker the automatic segmentation process can occur, potentially reaching to receive results even in 1 second or less as we could see in results.

Maris [45] has published a case study in "Medical Robotics" highlighting the PROST robot as a significant advancement in integrating AI into prostate biopsy procedures. The robot utilizes a prostate segmentation model known as PROST-Net, initially described by Palladino et al. [33]. In validation on ultrasound test data, PROST-Net achieved a Dice coefficient of 0.78 for prostate segmentation. Following the inclusion of additional data from a second cohort, the model's accuracy improved, exceeding a Dice coefficient of 0.80. The accuracy of prostate segmentation in MRI images was assessed using in-house MRIs, yielding a Dice coefficient greater than 0.95. Moreover, following rigid fusion, the Dice coefficient between MRI and ultrasound (US) segmentations was determined to be 0.75. Post-fusion, the distances between each lesion identified in MRI and the corresponding lesion in US were measured. The authors observed that the application of AI-driven image processing facilitates the automation of image fusion, as well as the identification and localization of prostate lesions, thereby enhancing

the procedure's efficiency and reliability. The robot is currently undergoing pre-clinical testing.
Another argument in favor would be the reduction of the risk of human errors. The human factor is prone to mistakes due to various external factors such as fatigue, low quality of images, distractions etc. [46]. Liu et al. [40] in their experiment, demonstrated that the algorithm was more consistent with the ground truth than the second expert. The directions this technological trend and its application might take include the integration of CNNs into clinical practice on a larger scale for software-co-registered MRI-targeted TRUS biopsy. Moving forward, automating the fusion biopsies could involve the development of technologies for automatic lesion delineation within the prostate glands. A significant advancement has recently been made in this field. An article published in "The Lancet Oncology", an internationally trusted source of clinical research with an Impact Factor of 41.6, investigates the performance of AI systems in detecting CSPCa on MRI compared to radiologists using the prostate imaging – reporting and Data System version 2.1 (PI-RADS 2.1). The authors demonstrate that AI was, on average, superior to radiologists using PI-RADS 2.1 in detecting CSPCa. The study compared 62 radiologists against AI in a test group comprising 400 cases [47].

CONCLUSIONS

Our review of the literature demonstrates that various CNN architectures have shown substantial potential in improving the accuracy and efficiency of prostate segmentation in mpMRI images. Tătaru et al. [48] analyzed the application of machine learning in the diagnosis of prostate cancer and highlighted that studies have reported AUC values as high as 0.91 using CNNs. This suggests that CNNs offer superior sensitivity and specificity compared to traditional radiological methods. Additionally, algorithms such as random forest classifiers and k-nearest neighbors have shown high predictive accuracy, particularly when used in conjunction with mpMRI. These techniques provide a robust set of tools for clinicians in identifying CSPCa. The integration of CNNs into the segmentation of prostate MRI images during fusion biopsy holds significant promise for improving the detection and diagnosis of CSPCa. Despite the advancements, there is a notable gap in the practical application of these tools within clinical workflows. Future research should focus on creating user-friendly, efficient segmentation tools that can be seamlessly integrated into the fusion biopsy process. Future research should focus on the development and refinement of these algorithms to ensure their robustness and applicability in clinical settings, as well as exploring methods to overcome current limitations such as the need for large training datasets and high computational power. Ultimately, the successful implementation of these advanced segmentation tools in clinical practice could lead to more accurate and efficient prostate cancer diagnoses, facilitating early detection and appropriate therapeutic decisions. The ongoing evolution of AI in medical imaging heralds a promising future for more precise and personalized prostate cancer diagnostics and treatment.

CONFLICTS OF INTEREST
The authors declare no conflict of interest.

FUNDING
This research received no external funding.

ETHICS APPROVAL STATEMENT
The ethical approval was not required.

Supplementary Table 1. *Ineligible studies for review*

van Sloun et al. [I]	TRUS images
Tollens et al. [II]	General review
Bonekamp et al. [III]	General review
Suarez-Ibarrola et al. [IV]	General review
Yilmaz et al. [V]	Based on the patient after radiotherapy
Lin et al. [VI]	No open access
Penzkofer et al. [VII]	Position paper
Lin et al. [VIII]	No open access
Hosseinzadeh et al. [IX]	Lack of information about contouring prostate
Soni et al. [X]	Segmentation of the prostate cancer lesion region. Lack of information about contouring prostate
Khosravi et al. [XI]	Lack of prostate segmentation description
Schelb et al. [XII]	Classification of PI-RADS lesions
Winkel et al. [XIII]	Classification of PI-RADS lesions
Kaneko et al. [XIV]	No open access
Jimenez-Pastor et al. [XV]	No open access
Guo et al. [XVI]	Irrelevant topic for our paper
Zhang et al. [XVII]	No open access
Labus et al. [XVIII]	No open access
Zheng et al. [XIX]	CSPCa lesion detection only
Al-Bourini et al. [XX]	No open access
Khan et al. [XXI]	General review
Saunders et al. [XXII]	Compares how the strategies of transfer learning and aggregated training
Alzate-Grisales et al. [XXIII]	Segmenting csPCa regions from MRI images. Lack of information about conturing prostate
Akhoondi and Baghshah [XXIV]	irrevelant topic for our paper – histopatological images
Elmahdy et al. [XXV]	Irrelevant topic for our paper – automatic re-contouring of follow-up scans for adaptive radiotherapy
Lee and Nishikwa [XXVI]	Mammograms segmentation/transferability of a traditionally transfer-learned CNN
Huang et al. [XXVII]	No statistical results about whole prostate contouring
Ren et al. [XXVIII]	The algorithm does not utilize the convolutional neural network approach
Shao et al. [XXIX]	Irrelevant topic for our paper – predict the Gleason Grade Group (GG-RP) of prostate cancer using bpMRI

CNN – convolutional neural networks; CSPCa – clinically significant prostate cancer; MRI – magnetic resonance imaging; PI-RADS – Prostate Imaging-Reporting and Data System; TRUS – transrectal ultrasound

I. van Sloun RJG, Wildeboer RR, Mannaerts CK, et al. Deep Learning for Real-time, Automatic, and Scanner-adapted Prostate (Zone) Segmentation of Transrectal Ultrasound, for Example, Magnetic Resonance Imaging-transrectal Ultrasound Fusion Prostate Biopsy. Eur Urol Focus. 2021; 7: 78-85.

II. Tollens F, Westhoff N, von Hardenberg J, et al. MRT-gestützte minimal-invasive Therapie des Prostatakarzinoms [MRI-guided minimally invasive treatment of prostate cancer]. Radiologe. 2021; 61: 829-838. German.

III. Bonekamp D, Schlemmer HP. Maschinelles Lernen und multiparametrische MRT in der Früherkennung des Prostatakarzinoms [Machine learning and multiparametric MRI for early diagnosis of prostate cancer]. Urologe A. 2021; 60: 576-591. German.

IV. Suarez-Ibarrola R, Sigle A, Eklund M, et al. Artificial Intelligence in Magnetic Resonance Imaging-based Prostate Cancer Diagnosis: Where Do We Stand in 2021? Eur Urol Focus. 2022; 8: 409-417.

V. Yilmaz EC, Harmon SA, Belue MJ, et al. Evaluation of a Deep Learning-based Algorithm for Post-Radiotherapy Prostate Cancer Local Recurrence Detection Using Biparametric MRI. Eur J Radiol. 2023; 168: 111095.

VI. Lin Y, Yilmaz EC, Belue MJ, et al. Evaluation of a Cascaded Deep Learning-based Algorithm for Prostate Lesion Detection at Biparametric MRI. Radiology. 2024; 311: e230750.

VII. Penzkofer T, Padhani AR, Turkbey B, et al. ESUR/ESUI position paper: developing artificial intelligence for precision diagnosis of prostate cancer using magnetic resonance imaging. Eur Radiol. 2021; 31: 9567-9578.

VIII. Lin Y, Belue MJ, Yilmaz EC, Harmon SA, et al. Deep Learning-Based T2-Weighted MR Image Quality Assessment and Its Impact on Prostate Cancer Detection Rates. J Magn Reson Imaging. 2024; 59: 2215-2223.

IX. 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: 2224-2234.

X. Soni M, Khan IR, Babu KS, Nasrullah S, Madduri A, Rahin SA. Light Weighted Healthcare CNN Model to Detect Prostate Cancer on Multiparametric MRI. Comput Intell Neurosci. 2022; 2022: 5497120.

XI. Khosravi P, Lysandrou M, Eljalby M, et al. A Deep Learning Approach to Diagnostic Classification of Prostate Cancer Using Pathology-Radiology Fusion. J Magn Reson Imaging. 2021; 54: 462-471.

XII. Schelb P, Kohl S, Radtke JP, et al. Classification of Cancer at Prostate MRI: Deep Learning versus Clinical PI-RADS Assessment. Radiology. 2019; 293: 607-617.

XIII. Winkel DJ, Wetterauer C, Matthias MO, et al. Autonomous Detection and Classification of PI-RADS Lesions in an MRI Screening Population Incorporating Multicenter-Labeled Deep Learning and Biparametric Imaging: Proof of Concept. Diagnostics (Basel). 2020; 10: 951.

XIV. Kaneko M, Fukuda N, Nagano H, et al. Artificial intelligence trained with integration of multiparametric MR-US imaging data and fusion biopsy trajectory-proven pathology data for 3D prediction of prostate cancer: A proof-of-concept study. Prostate. 2022; 82: 793-803.

XV. Jimenez-Pastor A, Lopez-Gonzalez R, Fos-Guarinos B, et al. Automated prostate multi-regional segmentation in magnetic resonance using fully convolutional neural networks. Eur Radiol. 2023; 33: 5087-5096.

XVI. Guo H, Kruger M, Xu S, Wood BJ, Yan P. Deep adaptive registration of multi-modal prostate images. Comput Med Imaging Graph. 2020; 84: 101769.

XVII. Zhang KS, Schelb P, Netzer N, et al. Pseudoprospective Paraclinical Interaction of Radiology Residents With a Deep Learning System for Prostate Cancer Detection: Experience, Performance, and Identification of the Need for Intermittent Recalibration. Invest Radiol. 2022; 57: 601-612.

XVIII. Labus S, Altmann MM, Huisman H, et al. A concurrent, deep learning-based computer-aided detection system for prostate multiparametric MRI: a performance study involving experienced and less-experienced radiologists. Eur Radiol. 2023; 33: 64-76.

XIX. Zheng H, Hung ALY, Miao Q, et al. AtPCa-Net: anatomical-aware prostate cancer detection network on multi-parametric MRI. Sci Rep. 2024; 14: 5740.

XX. Al-Bourini O, Seif Amir Hosseini A, Giganti F, et al. T1 Mapping of the Prostate Using Single-Shot T1FLASH: A Clinical Feasibility Study to Optimize Prostate Cancer Assessment. Invest Radiol. 2023; 58: 380-387.

XXI. Khan Z, Yahya N, Alsaih K, Al-Hiyali MI, Meriaudeau F, Recent Automatic Segmentation Algorithms of MRI Prostate Regions: A Review. IEEE Access. 2021; 9: 97878-97905.

XXII. Saunders SL, Leng E, Spilseth B, Wasserman N, Metzger GJ, Bolan PJ, Training Convolutional Networks for Prostate Segmentation With Limited Data. IEEE Access. 2021; 9: 109214-109223.

XXIII. Alzate-Grisales JA, Mora-Rubio A, García-García F, Tabares-Soto R, De La Iglesia-Vayá M, SAM-UNETR: Clinically Significant Prostate Cancer Segmentation Using Transfer Learning From Large Model. IEEE Access. 2023; 11: 118217-118228.

XXIV. Akhoondi PE, Baghshah MS. Semantic Segmentation With Multiple Contradictory Annotations Using a Dynamic Score Function. IEEE Access. 2023; 11: 64544-64558.

XXV. Elmahdy MS, Beljaards L, Yousefi S, et al. Joint Registration and Segmentation via Multi-Task Learning for Adaptive Radiotherapy of Prostate Cancer. IEEE Access. 2021; 9, 95551-95568.

XXVI. Lee J, Nishikawa RM. Cross-Organ, Cross-Modality Transfer Learning: Feasibility Study for Segmentation and Classification. IEEE Access. 2020; 8: 210194-210205.

XXVII. Huang H, Zhang b, Zhang X, et al. Application of U-Net Based Multiparameter Magnetic Resonance Image Fusion in the Diagnosis of Prostate Cancer. IEEE Access. 2021; 9: 33756-33768.

XXVIII Ren C, Guo Z, Ren H, Jeong D, Kim DK, Zhang S. Prostate Segmentation in MRI Using Transformer Encoder and Decoder Framework. IEEE Access. 2023; 11: 101630-101643.

XXIX. Shao L, Liu Z, Yan Y, et al. Patient-Level Prediction of Multi-Classification Task at Prostate MRI Based on End-to-End Framework Learning From Diagnostic Logic of Radiologists. IEEE Trans Biomed Eng. 2021; 68: 3690-3700.

References

1. WHO. www.gco.iarc.fr. 2022 [cited 2024 02 08. Available at: https://gco.iarc.fr/today/en/dataviz/pie?mode=cancer&group_populations=1&sexes=1&group_cancers=1&multiple_cancers=1.

2. Wojciechowska U, Barańska K, Miklewska M, Didkowska J. Cancer incidence and mortality in Poland in 2020. Nowotwory. Journal of Oncology. 2023; 73: 129-145.

3. Kaluarachchi T, Reis A, Nanayakkara S. A Review of Recent Deep Learning Approaches in Human-Centered. Sensors. 2021; 21: 2514.

4. Kufel J, Bargieł-Łączek K, Kocot S, Koźlik M, Bartnikowska W, Janik M, et al. What Is Machine Learning, Artificial Neural Networks and Deep Learning? – Examples of Practical Applications in Medicine. Diagnostics (Basel). 2023; 13: 2582.

5. Schouten M, van der Leest M, Pokorny M, et al. Why and Where do We Miss Significant Prostate Cancer with Multi-parametric Magnetic Resonance Imaging followed by Magnetic Resonance-guided and Transrectal Ultrasound-guided Biopsy in Biopsy-naïve Men? Eur Urol. 2017; 71: 896-903.

6. Bjurlin M, Mendhiratta N, Wysock J, Taneja S. Multiparametric MRI and targeted prostate biopsy: Improvements in cancer detection, localization, and risk assessment. Cent European J Urol. 2016; 69: 9-18.

7. Schoots IG, Roobol MJ, Nieboer D, Bangma C, Steyerberg E, Hunink M.

Magnetic resonance imaging-targeted biopsy may enhance the diagnostic accuracy of significant prostate cancer detection compared to standard transrectal ultrasound-guided biopsy: a systematic review and meta-analysis. Eur Urol. 2015; 68: 438-450.

8. Brock M, Bodman Cv, Palisaar J, Becker W, Martin-Seidel P, Noldus J. Detecting Prostate Cancer. Dtsch Arztebl Int. 2015; 112: 605-611.

9. Siddiqui M, Rais-Bahrami S, Turkbey B, et al. Comparison of MR/ultrasound fusion-guided biopsy with ultrasound-guided biopsy for the diagnosis of prostate cancer. JAMA. 2015; 313: 390-397.

10. Drost F, Osses D, Nieboer D, et al. Prostate MRI, with or without MRI-targeted biopsy, and systematic biopsy for detecting prostate cancer. Cochrane Database Syst Rev. 2019; 4: CD012663.

11. Exterkate L, Wegelin O, Barentsz JO, et al. Is There Still a Need for Repeated Systematic Biopsies in Patients with Previous Negative Biopsies in the Era of Magnetic Resonance Imaging-targeted Biopsies of the Prostate? Eur Urol Oncol. 2020; 3: 216-223.

12. Rouvière O, Puech P, Renard-Penna R, et al. se of prostate systematic and targeted biopsy on the basis of multiparametric MRI in biopsy-naive patients (MRI-FIRST): a prospective, multicentre, paired diagnostic study. Lancet Oncol. 2019; 20: 100-109.

13. Verma S, Turkbey B, Muradyan N, et al. Overview of dynamic contrast-enhanced MRI in prostate cancer diagnosis and management. AJR Am J Roentgenol. 2012; 198: 1277-1288.

14. Barentsz J, Richenberg J, Clements R, et al. European Society of Urogenital Radiology. ESUR prostate MR guidelines 2012. Eur Radiol. 2012; 22: 746-757.

15. Cornford P, Tilki D, Briers E, et al. EAU Guidelines on Prostate Cancer. In Edn. presented at the EAU Annual Congress; Paris 2024.

16. Woo S, Suh C, Kim S, Cho J, Kim S. Diagnostic Performance of Prostate Imaging Reporting and Data System Version 2 for Detection of Prostate Cancer: A Systematic Review and Diagnostic Meta-analysis. Eur Urol. 2017; 72: 177-188.

17. Padhani A, Barentsz J, Villeirs G, Rosenkrantz A, Margolis D, Turkbey B, et al. PI-RADS Steering Committee: The PI-RADS Multiparametric MRI and MRI-directed Biopsy Pathway. Radiology. 2019; 292: 464-474.

18. Prince M, Foster B, Kaempf A, et al. n-Bore Versus Fusion MRI-Targeted Biopsy of PI-RADS Category 4 and 5 Lesions: A Retrospective Comparative Analysis Using Propensity Score Weighting. AJR Am J Roentgenol. 2021; 217: 1123-1130.

19. Kongnyuy M, George A, Rastinehad A, Pinto P. Magnetic Resonance Imaging-Ultrasound Fusion-Guided Prostate Biopsy: Review of Technology, Techniques, and Outcomes. Curr Urol Rep. 2016; 17: 32.

20. Jeffrey C. Explainer: What is Machine Learning? TechSpot. 2020. Available at: https://www.techspot.com/article/2048-machine-learning-explained/ (accessed on 3 May 2023).

21. Alzubaidi L, Zhang J, Humaidi A, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data. 2021; 8: 53.

22. Hubel D, Wiesel T. Receptive fields of single neurones in the cat's striate cortex. J Physiol. 1959; 148: 574-591.

23. Sarvamangala D, Kulkarni R. Convolutional neural networks in medical image understanding: a survey. Evol Intell. 2022; 15: 1-22.

24. Backer D. UROwebinar: Artificial intelligence and the current applications in the field of urology 2023. Available at: https://youtu.be/0-0yD_L2j_Y?t=90.

25. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015; Springer, Cham, Munich 2015.

26. Shorten C, Khoshgoftaar TM. A survey on Image Data Augmentation for Deep Learning. J Big Data. 2019; 6: 60.

27. Wang F, Wang H, Wang H, Li G, Situ G. Learning from simulation: An end-to-end deep-learning approach for computational ghost imaging. Opt Express. 2019; 27: 25560-25572.

28. Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? Advances in Neural Information Processing Systems. 2014; 27: 3320-3328.

29. Padgett K, Swallen A, Pirozzi S, et al. Towards a universal MRI atlas of the prostate and prostate zones: Comparison of MRI vendor and image acquisition parameters. Strahlenther Onkol. 2019; 195: 121-130.

30. Bardis M, Houshyar R, Chantaduly C, et al. Segmentation of the Prostate Transition Zone and Peripheral Zone on MR Images with Deep Learning. Radiol Imaging Cancer. 2021; 3: e200024.

31. Soerensen S, Fan R, Seetharaman A, et al. Deep Learning Improves Speed and Accuracy of Prostate Gland Segmentations on Magnetic Resonance Imaging for Targeted Biopsy. J Urol. 2021; 206: 604-612.

32. Ushinsky A, Bardis M, Glavis-Bloom J, et al. A 3D-2D Hybrid U-Net Convolutional Neural Network Approach to Prostate Organ Segmentation of Multiparametric MRI. AJR Am J Roentgenol. 2021; 216: 111-116.

33. Palladino L, Maris B, Antonelli A, Fiorini P. PROST-Net: A Deep Learning Approach to Support Real-Time Fusion in Prostate Biopsy. IEEE Transactions on Medical Robotics and Bionics. 2022; 4: 323-326.

34. Hassanzadeh T, Hamey LGC, Ho-Shon K. Convolutional Neural Networks for Prostate Magnetic Resonance Image Segmentation. IEEE Access. 2019; 7: 36748-36760.

35. Ren C, Ren H. Prostate Segmentation on Magnetic Resonance Imaging. IEEE Access. 2023; 11: 145944-145953.

36. Huang X, Pang B, Zhang T, Jia G, Wang Y, Li Y. Improved Prostate Biparameter Magnetic Resonance Image Segmentation Based on Def-UNet. IEEE Access. 2023; 11: 43089-43100.

37. Qian Y. ProSegNet: A New Network of Prostate Segmentation Based on MR Images. IEEE Access. 2021; 9: 106293-106302.

38. Su C, Huang R, Liu C, Yin T, Du B. Prostate MR Image Segmentation With Self-Attention Adversarial Training Based on Wasserstein Distance. IEEE Access. 2019; 7: 184276-184284.

39. Qian Y, Zhang Z, Wang B. ProCDet: A New Method for Prostate Cancer Detection Based on MR Images. IEEE Access. 2021; 9: 143495-143505.

40. Liu Y, Yang G, Mirak SA, Hosseiny M, Azadikhah A, Zhong X. Automatic Prostate Zonal Segmentation Using Fully Convolutional Network With Feature Pyramid Attention. IEEE Access. 2019; 7: 163626-163632.

41. Comelli A, Dahiya N, Stefano A, et al. Deep Learning-Based Methods for Prostate Segmentation in Magnetic Resonance Imaging. Appl Sci (Basel). 2021; 11: 782.

42. Litjens G, Toth R, van de Ven W, Hoeks C, et al. Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge. Med Image Anal. 2014; 18: 359-373.

43. Warfield S, Zou K, Wells W. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. IEEE Trans Med Imaging. 2004; 23: 903-921.

44. Ghavami N, Hu Y, Gibson E, et al. Automatic segmentation of prostate MRI using convolutional neural networks: Investigating the impact of network architecture on the accuracy of volume measurement and MRI-ultrasound registration. Med Image Anal. 2019; 58: 101558.

45. Maris B. Advancing robotic prostate biopsy through artificial intelligence. Med Robot. 2024; 2: 1-7.

46. Ruben DM, Grace JG, Xin L, Wang G. Comparison of deep learning and human observer performance for detection and characterization of simulated lesions. J Med Imag. 2019; 6: 025503.

47. Saha A, Bosma J, Twilt J, et al. PI-CAI consortium. Artificial intelligence and radiologists in prostate cancer detection on MRI (PI-CAI): an international, paired, non-inferiority, confirmatory study. Lancet Oncol. 2024; 25: 879-887.

48. Tătaru OS, Vartolomei MD, Rassweiler JJ, et al. Artificial Intelligence and Machine Learning in Prostate Cancer Patient Management – Current Trends and Future Perspectives. Diagnostics. 2021; 11: 354.