

ResQu-Net: Effective prostate's peripheral zone segmentation leveraging the representational power of attention-based mechanisms

Dimitrios I. Zaridis^{a,c,d}, Eugenia Mylona^a, Nikolaos Tachos^a, Charalampos N. Kalantzopoulos^a, Kostas Marias^b, Manolis Tsiknakis^b, George K. Matsopoulos^c, Dimitrios D. Koutsouris^c, Dimitrios I. Fotiadis^{a,d,*}

^a Biomedical Research Institute, FORTH, Ioannina 45110, Greece

^b Computational Biomedicine Laboratory, FORTH, Heraklion, Greece

^c Biomedical Engineering Laboratory, School of Electrical & Computer Engineering, National Technical University of Athens, 9 Iroon Polytechniou St., 15780 Athens, Greece

^d Unit of Medical Technology Intelligent Information Systems, University of Ioannina, Greece

ARTICLE INFO

Keywords:

Peripheral zone segmentation
Deep learning
MRI
Attention mechanisms

ABSTRACT

Prostate cancer is a leading cause of male cancer worldwide. With more than 70 % of prostate cancers arising in the peripheral zone of the prostate, accurate segmentation of this region is of paramount importance for the effective diagnosis and treatment of the disease. Although peripheral zone is well recognized as one of the most challenging regions to delineate within the prostate, no algorithms specifically tailored for this segmentation task are currently available. The present study introduces a new deep learning (DL) algorithm, named as ResQu-Net, which is designed to accurately segment the peripheral zone (PZ) of the prostate on T2-weighted magnetic resonance imaging (MRI). Using three publicly available datasets, the ResQu-Net outperformed the six DL segmentation models used for comparison, namely the Attention U-Net, the Dense2U-Net, the Proper-Net, the TransU-net, the U-Net, and the USE-Net, demonstrating superior performance for different anatomical regions, such as the apex, the midgland and the base. The assessment of the suggested approach was conducted not only quantitatively (Sensitivity, Balanced Accuracy, Dice Score, 95 % Hausdorff Distance, and Average Surface Distance) but also qualitatively. For the qualitative evaluation the feature maps obtained from the last layers of each model were compared with the Density Map of the Ground Truth annotations using root mean squared error. Overall, the ResQu-Net model exhibits improved performance compared to other models, of more than 5 % and 1.87 mm in terms of Dice Score and 95 % Hausdorff Distance, respectively. These advancements may contribute significantly in addressing the challenges associated with PZ segmentation, and ultimately enabling improved clinical decision-making and patient outcomes.

1. Introduction

Prostate cancer is the second most prevalent form of cancer and a leading cause of cancer-related death among men [1]. In 2020, more than 1.4 million new prostate cancer cases were diagnosed worldwide, with a crude incidence rate of 36 per 100,000 males [2]. Early detection of prostate cancer and effective treatments are of paramount importance for ensuring a successful outcome. If the disease is diagnosed at early stages, where the cancer remains local or regional and does not spread to distant body parts, the 5-year survival rate is nearly 100 % [3].

Prostate cancer develops in the prostate gland, which can be divided

into different zones according to their function, namely the central zone (CZ), the peripheral zone (PZ), and the transitional zone (TZ) [4]. The PZ, in particular, which extends posterolaterally around the gland from the apex to the base, represents the most common site in the prostate for developing prostate carcinomas due to the fact that this region contains most of the prostatic glandular tissue [5,6]. Cancers originating from the PZ account for more than 70 % of prostate cancers and are related to worse clinical outcomes than TZ cancers [7]. Therefore, accurate zonal segmentation of prostate boundaries on MRI, especially of the PZ, has a crucial role in the diagnosis and treatment of prostate cancer [8].

As part of the diagnostic pathway for prostate cancer, Magnetic

* Corresponding author.

E-mail address: fotiadis@uoi.gr (D.I. Fotiadis).

<https://doi.org/10.1016/j.bspc.2024.106187>

Received 31 July 2023; Received in revised form 19 February 2024; Accepted 1 March 2024

Available online 6 March 2024

1746-8094/© 2024 Elsevier Ltd. All rights reserved.

Resonance Imaging (MRI) has been widely adopted as the primary imaging modality for non-invasive prostate cancer detection and localization, owing to its superior soft tissue contrast. It has also been documented that prostate MRI is as effective as a routine biopsy at detecting high-grade cancers, highlighting the areas of the prostate that are most suspicious for cancer, and reducing the number of biopsies needed [9]. The Prostate Imaging, Reporting, and Diagnosis System (PIRADS) scoring system, which is being used to detect clinically significant prostate cancer based on MRI findings, recognizes that the prostate zones have different imaging properties, and this is well reflected in the scoring system [10]. Consequently, prostate cancer identification and staging on MRI rely on accurate zonal segmentation [11].

The PZ of the prostate is characterized by distinct histological characteristics. Research has emphasized the notable diversity within the PZ, which is evident not only in the range of clinical results but also in the unique genetic and cellular characteristics that contribute to tumor advancement and aggressiveness in this region [12,13]. The spatial and morphological heterogeneity of PZ, the subtle boundaries between neighboring tissues, and the large intra- and inter-individual variations render manual delineation of the PZ an extremely challenging and laborious and tedious task [8]. Several multi-reader studies have highlighted the impact and magnitude of inter-observer variability in the manually delineated prostatic regions, particularly in the extreme parts of the gland [12,13,14], such as the apex and the base, emphasizing the unmet need to develop automatic methods to streamline the segmentation process and offer robust and accurate delineation of the prostatic zones. Automated segmentation algorithms offer the potential to minimize inter-observer variability, enhance efficiency, reduce the processing time, and enable quantitative analysis for clinical decision-making.

Over the years, various algorithms and techniques have emerged to address prostate and zonal segmentation tasks. Traditionally, conventional image analysis methods, mainly based on deformable models, atlas-based approaches, and feature-based ML methods have been used for segmenting the prostatic zones [15,16,17]. While these approaches have achieved moderate success, they often struggle with accurately capturing the complex variations within the prostate. More recently, Deep Learning (DL)-based segmentation methods have seen remarkable success in medical image segmentation tasks, and prostate zonal segmentation is no exception [8].

By harnessing the capabilities of convolutional neural networks (CNNs), these algorithms can effectively learn discriminative features directly from the raw image data, thus mitigating the limitations associated with manual feature engineering. Given the continual growth in the size and complexity of prostate MRI data, the capacity of CNNs to learn hierarchical representations from such datasets has proven to be invaluable [18]. The CNNs' aptitude for capturing high-level representations, including semantic and spatial relationships, has significantly improved the accuracy and reliability of the segmentation results. However, existing studies that have employed DL approaches for segmenting the PZ either rely on algorithms originally developed for different segmentation tasks or for simultaneously segmenting the whole prostate and the prostatic zones [19,20,21,22,23,24,25]. As a result, algorithms designed for segmenting other organs and regions may not be optimized to reliably segment the PZ alone. This is also reflected in the limited segmentation performance reported in the literature when segmenting the PZ.

In this paper, we propose a novel DL algorithm, specifically tailored for segmenting the PZ of the prostate on T2-weighted MRI. We refer to this model as the Residual Spatial Attention and Squeeze and Excitation Network (ResQu-Net). To the best of our knowledge, this is the first time a DL algorithm specifically tailored for peripheral zone segmentation is being developed. The proposed DL model deploys the rich representational power of a spatial attention module, as well as squeeze and excitation blocks interconnected in a parallel manner for efficiency purposes, ultimately incorporated in a segmentation network. The

model was developed and validated using diverse population data from different clinical sites and MRI vendors, and was compared against 6 state-of-the-art DL segmentation algorithms. Models' performance was evaluated both quantitatively, using standard evaluation metrics, but also qualitatively through the extraction of feature maps and the comparison with ground truth maps.

2. Methodology

2.1. Datasets description

In this study, three publicly available datasets containing bi-parametric MRI were utilized, namely the ProstateX2 [26], the Prostate-3 T [27] and the Prostate-158 [28]. The ProstateX2 comprise 204 patients with 3200 frames, acquired on Siemens vendor and TrioTim, Skyra models with magnetic field strength 3 Tesla. The images were acquired at the Radboud University Medical Centre while the ground truth annotations were provided by the Department of Advanced Biomedical Sciences of the University of Naples "Federico II" [29]. The Prostate-3 T dataset includes 30 patients with 421 frames acquired with Siemens vendor and Skyra model with 3 Tesla magnetic field strength. This dataset was made available by Boston University. The Prostate-158 dataset contains 158 patients with 2376 frames acquired by a Siemens VIDA model with 3 Tesla magnetic field strength and it was made available by the Charité University Hospital of Berlin, Germany. The latter, being the most curated dataset among the three, was also employed to explore the inter-patient variability of the PZ in terms of shape, texture, and grayscale characteristics. All datasets combined, our analysis was conducted using data from 392 patients with approximately 6000 frames.

2.2. Proposed model architecture

The proposed model (ResQu-Net) consists of an encoder decoder network which assists on the forward passing of information, being capable of retaining the spatial information of the PZ by utilizing residual connections. The network's main components include (i) convolution layers to extract features, (ii) max pooling layers to reduce the dimensionality and ensure compactness of the network, and (iii) the novel RSA.se block which involves the parallel connection of Squeeze & Excitation (SE) [30] and Spatial Attention (SA) layers along with the input's residual connection to propagate information smoothly. The proposed architecture is depicted in Fig. 1A with the RSA.se block being detailed in Fig.1B.

2.3. Rsa.se module

To enhance the capabilities of well-known fully connected CNN architectures, such as U-Net, two visual attention mechanisms, namely the Spatial Attention and Squeeze & Excitation, were implemented and connected in a parallel manner to work complementary for the feature extraction process. Those blocks along with their unique intra-connection consist the backbone of the RSA.se block. The merits of RSA.se block are 2 fold. First, the information is computed in a parallel manner ensuring the efficiency of floating points operations (FLOPS), and therefore the fast convergence and inference time. Second, both inter-channel and intra-channel attention are deployed to retrieve information regarding the PZ features in a complementary scheme.

Specifically, for the SE layer, let there $X_{H \times W \times C}$ be the input tensor of the RSA.se block. That tensor is utilized as input for both the SE and SA layers. The SE layer could be described as an attention mechanism which applies the importance factor $f_{1 \times 1 \times C}$ to each feature map C of the input tensor X with spatial dimensions $[H, W]$ and C feature maps. The steps for the calculation of the importance factor are:

$X_{H \times W \times C}$ is global average pooled (GAP) to concentrate the relevant to the task spatial information into a single value:

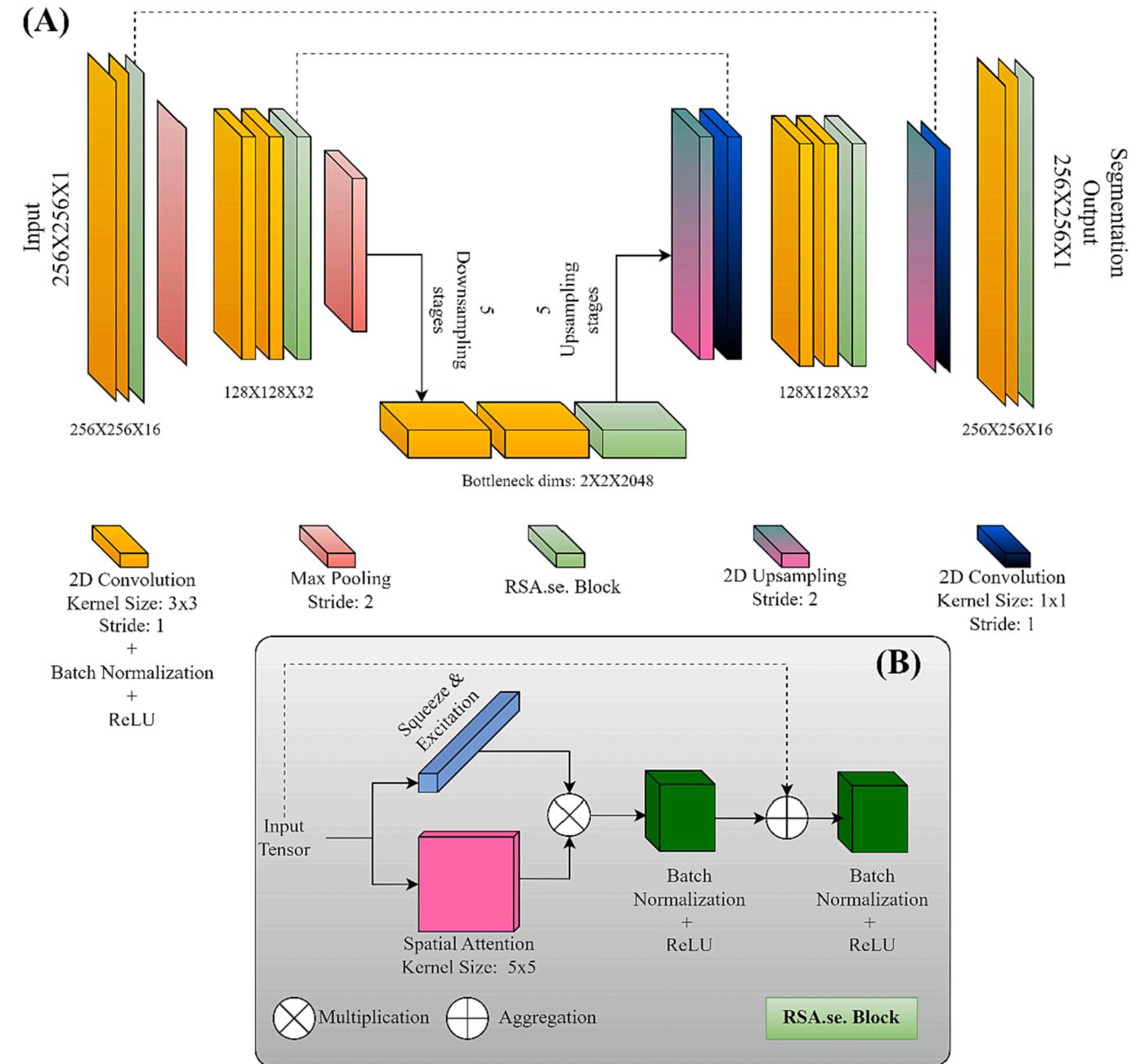


Fig. 1. Overview of: A) The proposed ResQu-Net model architecture and B) the novel RSA.se block.

$$X_{1 \times 1 \times C}^{GAP} = \frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W X_{H \times W \times C}, \quad (1)$$

$X_{1 \times 1 \times C}^{GAP}$ is then passed from 2 densely connected neural networks (nn) with 1 hidden layer with a modulation factor r which denotes the number of hidden units contained in the hidden layer. The aforementioned nn consists the cornerstone of the SE layer as it discovers the feature map C with the most informative for the task attribute. The output may be described as a set of scaling factors $f_{1 \times 1 \times C}$ which modulate the importance of each feature map C . Ultimately Eq.(2) produces the output of nn which is the scaling factor:

$$f_{1 \times 1 \times C} = \text{Sigmoid}[X_{1 \times 1 \times C}^{GAP}] \quad (2)$$

Where *Sigmoid* is the sigmoid activation function.

c) The last step to obtain the output of the SE layer is the element-wise multiplication of the input tensor $X_{H \times W \times C}$ and the importance

factor obtained from Eq.(2), $f_{1 \times 1 \times C}$. In Eq.(3) the output of the SE layer is calculated using the Eq.(2) and the input tensor to obtain the feature maps modulated by the importance factor:

$$X_{H \times W \times C}^{se} = X_{H \times W \times C} * f_{1 \times 1 \times C}, \quad (3)$$

Eq.(3) was used to selectively enhance features across channels C from input tensor $X_{H \times W \times C}$ based on the importance factor $f_{1 \times 1 \times C}$ given by Eq.(2).

SA layer is also employed in a complementary manner with the SE layer. More specifically, the effect of the SA layers is that it produces a spatial attention modulation factor to identify significant for the task spatial features (intra-channel). The steps for producing the spatial features are presented sequentially below:

a) The Input tensor $X_{H \times W \times C}$ passes from 2 pooling operations with respect to channels C . Those operations are Max pooling and Average Pooling. Eq.(4) and Eq.(5) produce the tensors after the two pooling

operations, respectively.

$$X_{H \times W \times 1}^{MAX} = \left[\sum_{c=1}^C X_{H \times W \times c} \right], \quad (4)$$

where *max* is the maximum intensity value within the spatial feature map $X_{H \times W}$ for each channel C .

$$X_{H \times W \times 1}^{AVE} = \frac{1}{C} \sum_{c=1}^C X_{H \times W \times c} \quad (5)$$

b) Tensors $X_{H \times W \times 1}^{MAX}$ and $X_{H \times W \times 1}^{AVE}$ obtained by Eq.(4) and Eq.(5), respectively, are concatenated forming the tensor:

$$X_{H \times W \times 2}^{conc} = [X_{H \times W \times 1}^{MAX}, X_{H \times W \times 1}^{AVE}] \quad (6)$$

c) The convolution operation with a kernel size of 1, $Conv_{1 \times 1 \times C}^{Sigmoid}$ and the sigmoid activation function is then applied at $X_{H \times W \times 2}^{conc}$ obtained by Eq.(6) to produce the spatial attention modulation factor. This factor is then multiplied in an element-wise manner by the input tensor $X_{H \times W \times C}$ resulting in the output tensor of the SA layer, which is presented in Eq. (7).

$$X_{H \times W \times C}^{SA} = X_{H \times W \times C} * [Conv_{1 \times 1 \times C}^{Sigmoid} * X_{H \times W \times 2}^{conc}] \quad (7)$$

Putting it all together, element wise multiplication is performed based on the Eq.(3) and Eq.(7) to distinguish features from the SE layer (inter-channel features) and SA (intra-channel features). In this way, the model achieves spatial feature awareness, as well as channel wise awareness by searching for semantic information throughout the whole feature space. Eq.(8) produces the dot production between SE and SA layers:

$$X_{H \times W \times C}^{SA.se} = X_{H \times W \times C}^{SA} * X_{H \times W \times C}^{se}, \quad (8)$$

Sequentially, the Batch Normalization and the ReLU activation function are applied to homogenize the output tensor and propagate the relevant features of Eq.(8). Ultimately, residual connection is performed between input tensor $X_{H \times W \times C}$ and $X_{H \times W \times C}^{SA.se}$ obtained by Eq.(9) to produce the output of *RSA.se* block so as to improve the representation of the output.

$$X_{H \times W \times C}^{RSA.se} = X_{H \times W \times C} + X_{H \times W \times C}^{SA.se}, \quad (9)$$

2.4. Deep learning segmentation models used for comparison

The Six deep learning segmentation models were implemented and compared against the proposed ResQu-Net. Convolutional Neural Networks (CNNs) are the backbone of the majority of the models used apart from TransU-Net. First, the baseline U-Net [31] structure was implemented, consisting of convolutional layers that are employed in the encoder route of the model to progressively reduce the spatial dimensions of the input image while extracting semantic information. These layers are in charge of gathering significant contextual information and extracting high-level characteristics. On the contrary, the decoder path is responsible for reconstructing the feature maps and along with skip connections from the encoder path to produce the final PZ segmentation. To solve the challenges of PZ segmentation, ProperNet incorporated an interconnection of CNN layers and residual connections. This model was also adopted in this work as a comparison model [32]. Furthermore, we considered the Dense2U-Net model [25], which instead of convolutional layers it is composed of densely connected CNNs with exhaustive residual mechanisms. The Attention U-Net model [33] was also implemented. Attention mechanisms have been orchestrated in segmentation models to complement the CNN layers, allowing to concentrate on informative areas and enhancing the model's capacity to recognize subtle patterns while tackling challenging segmentation tasks. The attention gating mechanisms at the end of each decoding layer make up the majority of the Attention U-Net network. Another model used for comparison was the USE-Net [34], which

leverages the attention mechanisms of Squeeze & Excitation block at the end of each encoding and decoding layer, which is a channel-wise attention mechanism capable of adaptively recalibrating feature maps, emphasizing important channels and suppressing less informative ones. Finally, the investigation of the efficiency of attention mechanisms in DL segmentation tasks brought forward the concept of Vision Transformers (ViT) encapsulation in the typical U-Net structure. Therefore, in this study, we considered the recently proposed TransU-Net [35], a novel approach that incorporates ViT into the bottleneck of U-Net. By analyzing patches within the feature maps, TransU-Net generates feature embeddings and enables adaptive feature weighting.

2.5. Study design

The Models' performance was evaluated using T2-Weighted (T2w) MR sequences in a 5-fold cross-validation pipeline. Prior to model training, all images underwent Min-Max intensity normalization, which involved scaling the pixel values to the interval [0,1]. Also, training and test sets in each fold were stratified so that the proportion of data from each dataset would remain invariant across folds. To ensure a rigorous and comprehensive study approach, the analysis was performed based on: a) the overall performance of the models tested for each metric, including the mean and standard deviation, b) the performance of the models broken down by dataset, and c) the performance of the models broken down by anatomical region, namely the apex, the midgland, and the base. The upper part of the prostate is called the base, which rests against the lower part of the bladder, while the lower, narrowed part of the prostate is called the apex. Since the apex of the prostate is defined as the inferior-most 0.5 cm portion of the gland, and the base as the superior-most 0.5 cm portion of the gland [25], for the purpose of our study, we estimated that this could be approximated by allocating the first quartile of frames of the prostate to the apex, the subsequent half to the midgland region, and the final quartile to the base.

2.6. Models' performance evaluation

The performance of the ResQu-Net and the six models in comparison, was estimated based on five performance metrics, namely the Dice Score (DS), the 95 % Hausdorff Distance (HD), the Average Surface Distance (ASD), the Sensitivity and the Balanced Accuracy (BA). The need for highlighting different aspects of models' performance led to the inclusion of various well-known segmentation measurements as the primary objective. For instance, the DS measures the overlap between Ground Truth (GT) and Model Prediction (MP), while the 95 % HD is calculated by: a) measuring the Euclidian distance between each point in MP and the nearest point in GT, b) ordering in ascending manner these distances and c) take the 95 % maximum distance. The ASD, which calculates the mean Euclidian distance for each data point between MP and GT, was employed to measure the accuracy of the surface boundary of the MP compared to GT. In addition, the Sensitivity was computed in order to provide further understanding regarding the predictive capacity of models with respect to the PZ of the prostate, while the BA was utilized to underscore the effectiveness of a given model in predicting both foreground pixels (PZ) and background pixels.

To further investigate the performance of the models, explainable AI (XAI) analysis has been performed. The general idea behind this is: a) the axial slice-wised aggregation of Saliency Maps (SM) per patient extracted by the model just before the classification layer, defined as Saliency Maps Density Image (SMDI), b) the slice-wised aggregation of GT masks per patient, defined as Ground Truth Density Image (GTDI), and c) the computation of Root Mean Squared Error (RMSE) between SMDI and GTDI in patient level. The methodology employed considers the significant characteristics of a model, rather than solely relying on a quantitative assessment [36]. As such, it serves as an indicator of the model's confidence in identifying areas that belong to the PZ, rather than simply evaluating the binary outcome (i.e., the predicted mask).

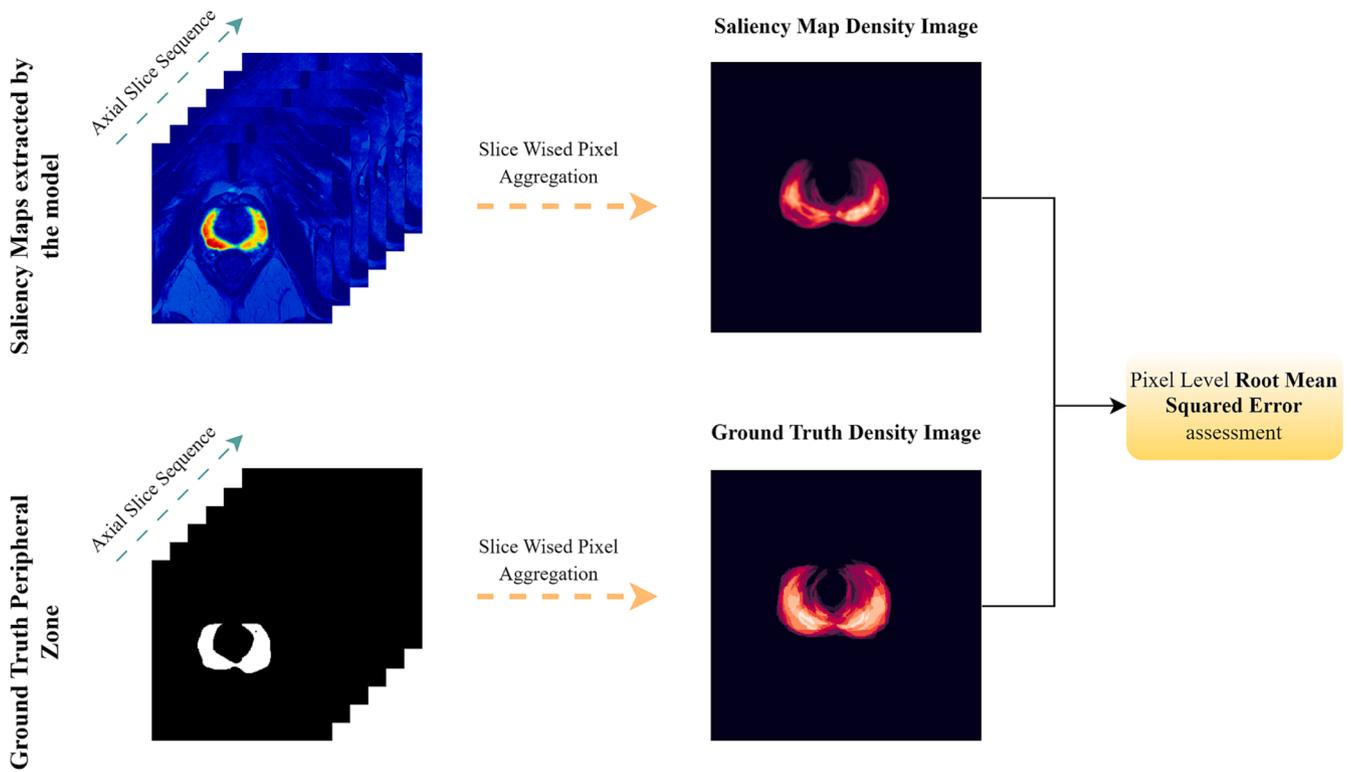


Fig. 2. Qualitative performance assessment pipeline for a given patient.

Fig. 2 depicts the XAI methodology through the computation of the aggregation of SM and GT images at the patient level, and the calculation of the Root Mean Square Error (RMSE) between them.

2.7. Experiment setup

For the experiments, different depth of convolutional filters was investigated to ensure that the comparison of ResQu-Net would be against the best available hyperparameters for each existing model and thus the best models. More specifically, batch normalization and L2 regularization approaches were used with the former having gamma and beta regularizers of 10^{-4} value. The dropout was set at 50%. The batch size was maintained at 16 frames per batch, the number of epochs was set at 120. As a loss function, the sigmoid focal cross-entropy was used, and the Adam optimizer was utilized to update the model's weights. The learning rate set to be cyclical with a periodicity of 10 epochs and values between 10^{-2} and 10^{-4} . The model training was conducted on NVIDIA Quadro P6000 GPU with driver version 441.66 while the Python packages used are Scikit-image = 0.18.3, numpy = 1.21.2, keras-unet-collection = 0.1.11, scipy = 1.7.1, tensorflow = 2.7.0, and tensorflow-addons = 0.11.2.

3. Results

Table 1 presents the average performance and the standard deviation for each metric and for each DL model across all folds. As it is shown, the ResQu-Net outperforms all the other models regardless of the performance metric. At least 5% increase was achieved in terms of DS, meaning that the overlap between the MPr and the GT masks was higher with the proposed architecture. In terms of HD and ASD, the results indicate that ResQu-Net outperformed the other models by a margin of at least 1.86 mm and 0.39 mm, respectively. Remarkable were the differences also for Sensitivity and BA, where ResQu-Net outperformed the other models by at least 4% and 3%, respectively.

Table 2 demonstrates the performance of each model in each of the

Table 1

Average scores and Standard Deviation across folds for each model and metric.

Model	Sensitivity (%)	Balanced Accuracy (%)	Dice Score (%)	Hausdorff Distance (mm)	Average Surface Distance (mm)
Attention U-Net	0.47 ± 0.05	0.74 ± 0.03	0.61 ± 0.04	6.59 ± 0.55	1.42 ± 0.14
Dense2U-Net	0.41 ± 0.08	0.7 ± 0.04	0.54 ± 0.06	7.87 ± 0.92	1.7 ± 0.2
Proper-Net	0.45 ± 0.08	0.73 ± 0.04	0.58 ± 0.06	7.3 ± 1.55	1.49 ± 0.23
TransU-Net	0.47 ± 0.06	0.73 ± 0.03	0.61 ± 0.06	6.83 ± 1.08	1.39 ± 0.15
U-Net	0.48 ± 0.06	0.74 ± 0.03	0.62 ± 0.04	7.03 ± 0.77	1.41 ± 0.17
USE-Net	0.56 ± 0.12	0.78 ± 0.06	0.64 ± 0.08	6.8 ± 1.33	1.43 ± 0.3
ResQu-Net-NET	0.62 ± 0.05	0.81 ± 0.02	0.69 ± 0.03	4.73 ± 0.35	1.12 ± 0.08

three datasets across the different performance metrics. ResQu-Net was again the best performing model across all datasets and folds. The most noteworthy difference from the second-best model was observed for the Prostate-158 dataset where DS, HD and ASD had a difference of at least 6%, 2.18 mm and 0.34 mm, respectively. On the contrary, the least notable difference arises at the Prostate-3 T dataset where ResQu-Net had an improvement of 2%, 1.84 mm and 1.35 mm for DS, HD and ASD, respectively. Similarly, to the results presented in Table 1, the USE-Net and the Attention U-Net were the second and third best performing models on each dataset. A sub-group analysis was also conducted to

Table 2
Average scores and Standard Deviation across folds for each model, metric and Dataset.

Dataset	Model	Sensitivity (%)	Balanced Accuracy (%)	Dice Score (%)	Hausdorff Distance (mm)	Average surface distance (mm)
Prostate-158	Attention U-Net	0.51 ± 0.08	0.75 ± 0.04	0.63 ± 0.06	7.95 ± 1.06	1.73 ± 0.21
	Dense2U-Net	0.46 ± 0.09	0.73 ± 0.04	0.58 ± 0.08	9.05 ± 1.44	2.0 ± 0.28
	Proper-Net	0.5 ± 0.07	0.75 ± 0.04	0.61 ± 0.06	8.51 ± 1.54	1.78 ± 0.23
	TransU-Net	0.51 ± 0.09	0.76 ± 0.05	0.64 ± 0.08	7.84 ± 1.71	1.62 ± 0.3
	U-Net	0.51 ± 0.08	0.76 ± 0.04	0.64 ± 0.06	8.06 ± 0.89	1.65 ± 0.23
	USE-Net	0.57 ± 0.12	0.78 ± 0.06	0.65 ± 0.08	8.94 ± 2.91	1.87 ± 0.58
	ReSqu-Net	0.66 ± 0.07	0.83 ± 0.03	0.71 ± 0.05	5.88 ± 1.13	1.39 ± 0.17
Prostate-3 T	Attention U-Net	0.49 ± 0.1	0.74 ± 0.05	0.62 ± 0.11	7.07 ± 2.08	1.6 ± 0.29
	Dense2U-Net	0.42 ± 0.12	0.71 ± 0.06	0.55 ± 0.14	8.59 ± 2.3	1.91 ± 0.44
	Proper-Net	0.46 ± 0.14	0.73 ± 0.07	0.6 ± 0.13	7.91 ± 2.5	1.67 ± 0.38
	TransU-Net	0.46 ± 0.11	0.73 ± 0.06	0.61 ± 0.11	6.82 ± 1.99	1.53 ± 0.3
	U-Net	0.5 ± 0.09	0.75 ± 0.04	0.63 ± 0.1	7.23 ± 2.46	1.57 ± 0.4
	USE-Net	0.57 ± 0.13	0.78 ± 0.07	0.67 ± 0.14	7.67 ± 3.48	1.61 ± 0.78
	ReSqu-NET	0.61 ± 0.08	0.8 ± 0.04	0.69 ± 0.1	5.23 ± 1.58	1.25 ± 0.34
ProstateX2	Attention U-Net	0.44 ± 0.07	0.72 ± 0.03	0.58 ± 0.06	6.09 ± 0.97	1.25 ± 0.18
	Dense2U-Net	0.37 ± 0.1	0.68 ± 0.05	0.51 ± 0.09	7.31 ± 1.32	1.51 ± 0.27
	Proper-Net	0.42 ± 0.11	0.71 ± 0.05	0.55 ± 0.08	6.88 ± 1.99	1.35 ± 0.34
	TransU-Net	0.42 ± 0.07	0.71 ± 0.04	0.57 ± 0.07	6.51 ± 1.27	1.26 ± 0.2
	U-Net	0.45 ± 0.09	0.72 ± 0.04	0.58 ± 0.08	6.52 ± 1.14	1.25 ± 0.18
	USE-Net	0.56 ± 0.12	0.78 ± 0.06	0.63 ± 0.09	5.91 ± 1.12	1.22 ± 0.26
	ReSqu-Net	0.58 ± 0.05	0.79 ± 0.03	0.67 ± 0.03	4.14 ± 0.52	0.98 ± 0.1

assess the performance of the ResQu-Net model with respect to the performance reported in the original paper of the Prostate 158 dataset using the same data splits. As it is presented in the [Supplementary Table 1](#), our model outperforms the UResNet model by 4.84 % and 4.32 % for the GT masks provided for the first and second raters, respectively. Moreover, in [Supplementary Table 2](#) and [Supplementary Fig. 1](#), we examine how RSA.se block behaves in comparison to the base ResQu-Net model and how it also compares to the Squeeze & Excitation(SE) [30], Bottleneck Attention Mechanism (BAM)[37] and Convolution Block Attention Module (CBAM) [38], both quantitatively and qualitatively.

In [Table 3](#) the average scores are presented with respect to each metric and each DL model for the three anatomical regions: the apex, the base and the midgland. The inclusion criteria related to the distribution of frames in the appropriate anatomical region were selected with respect to N. Aldo *et al.*, 2020 study [25]. It is again evident that ResQu-Net achieves a significant improvement in segmentation performance, especially for the midgland and the basal regions. More specifically, the proposed architecture achieved an increase of 5 % DS from the second

best, the USE-Net, for both the midgland and the base, and 4 % for the apex.

[Fig. 3](#) illustrates the normalized DS values for each DL model related to the anatomical region ([Fig. 3A](#)) and the dataset ([Fig. 3B](#)), averaging between 5 folds. In terms of the most effective models, the scores were standardized so that the best had a value 1 and the poorest had a value 0. Additionally, this representation offers a direct comparison between the models and presents the performance interval between them. For the different anatomical regions, notably, the ResQu-Net exhibits higher predictive capabilities in the base of PZ, with the second-best model, Use-Net, achieving only 64 % of ResQu-Net's performance. At the PZ's apex region, the performance of USE-Net is observed to be 75 % as effective as the proposed model. The ResQu-Net model was more reliable in generating PZ segmentations compared to other models tested on multi-centric datasets, including Prostate-158, Prostate-3 T, and Prostate-X2. At the Prostate-158 dataset, a significant variation is observed wherein the models of USE-Net and TransU-Net secure the second-best positions, exhibiting a performance level that is akin to 54 % of the ResQu-Net's performance. In contrast, the USE-Net model

Table 3
Average scores and Standard Deviation across folds for each model, metric and Anatomical Region.

Anatomical Region	Model	Sensitivity (%)	Balanced Accuracy (%)	Dice Score (%)	Hausdorff Distance (mm)	Average surface distance (mm)
Apex Region	Attention U-Net	0.45 ± 0.05	0.73 ± 0.03	0.59 ± 0.05	6.84 ± 0.63	1.52 ± 0.06
	Dense2U-Net	0.38 ± 0.1	0.69 ± 0.05	0.52 ± 0.09	8.25 ± 1.16	1.81 ± 0.25
	Proper-Net	0.42 ± 0.07	0.71 ± 0.04	0.57 ± 0.06	7.67 ± 1.49	1.63 ± 0.25
	TransU-Net	0.44 ± 0.07	0.72 ± 0.03	0.58 ± 0.07	7.43 ± 1.09	1.52 ± 0.23
	U-Net	0.45 ± 0.05	0.73 ± 0.03	0.59 ± 0.05	7.46 ± 0.87	1.54 ± 0.14
	USE-Net	0.55 ± 0.13	0.77 ± 0.06	0.63 ± 0.08	6.9 ± 1.33	1.45 ± 0.26
	ReSqu-Net	0.58 ± 0.04	0.79 ± 0.02	0.67 ± 0.03	4.95 ± 0.44	1.19 ± 0.08
Midgland Region	Attention U-Net	0.48 ± 0.04	0.74 ± 0.02	0.61 ± 0.04	6.32 ± 0.68	1.35 ± 0.16
	Dense2U-Net	0.41 ± 0.09	0.71 ± 0.04	0.55 ± 0.07	7.8 ± 1.23	1.66 ± 0.24
	Proper-Net	0.46 ± 0.09	0.73 ± 0.04	0.59 ± 0.06	7.03 ± 1.49	1.41 ± 0.22
	TransU-Net	0.48 ± 0.06	0.74 ± 0.03	0.62 ± 0.06	6.59 ± 1.09	1.34 ± 0.15
	U-Net	0.48 ± 0.08	0.74 ± 0.04	0.62 ± 0.06	6.83 ± 1.06	1.37 ± 0.2
	USE-Net	0.57 ± 0.11	0.78 ± 0.06	0.65 ± 0.08	6.52 ± 1.09	1.41 ± 0.3
	ReSqu-Net	0.62 ± 0.05	0.81 ± 0.03	0.7 ± 0.03	4.52 ± 0.42	1.07 ± 0.07
Basal Region	Attention U-Net	0.49 ± 0.08	0.74 ± 0.04	0.62 ± 0.07	6.86 ± 1.22	1.43 ± 0.21
	Dense2U-Net	0.43 ± 0.08	0.71 ± 0.04	0.56 ± 0.06	7.83 ± 1.14	1.66 ± 0.15
	Proper-Net	0.48 ± 0.11	0.74 ± 0.05	0.6 ± 0.08	7.51 ± 2.07	1.47 ± 0.32
	TransU-Net	0.5 ± 0.05	0.75 ± 0.03	0.64 ± 0.04	6.62 ± 1.31	1.34 ± 0.13
	U-Net	0.52 ± 0.05	0.76 ± 0.03	0.64 ± 0.03	6.85 ± 0.49	1.37 ± 0.15
	USE-Net	0.57 ± 0.12	0.79 ± 0.06	0.65 ± 0.07	7.2 ± 2.0	1.47 ± 0.38
	ReSqu-Net	0.64 ± 0.04	0.82 ± 0.02	0.7 ± 0.03	4.77 ± 0.49	1.12 ± 0.11

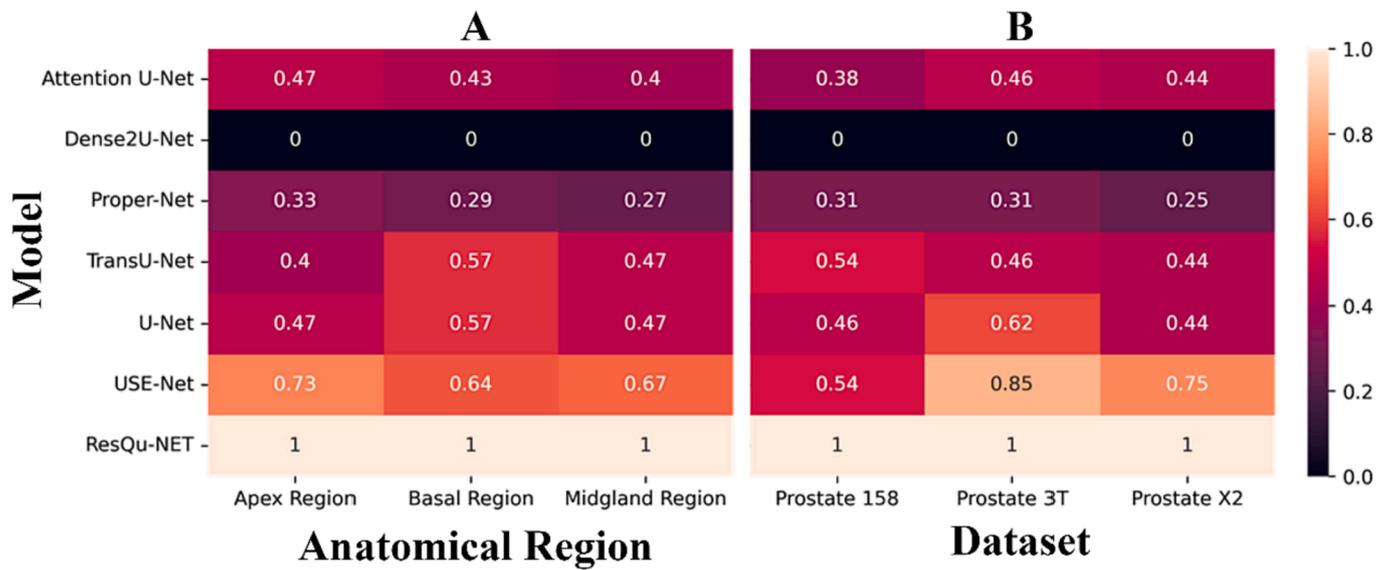


Fig. 3. Normalized Dice Score values for each Model in respect to A) Anatomical Region and B) Dataset.

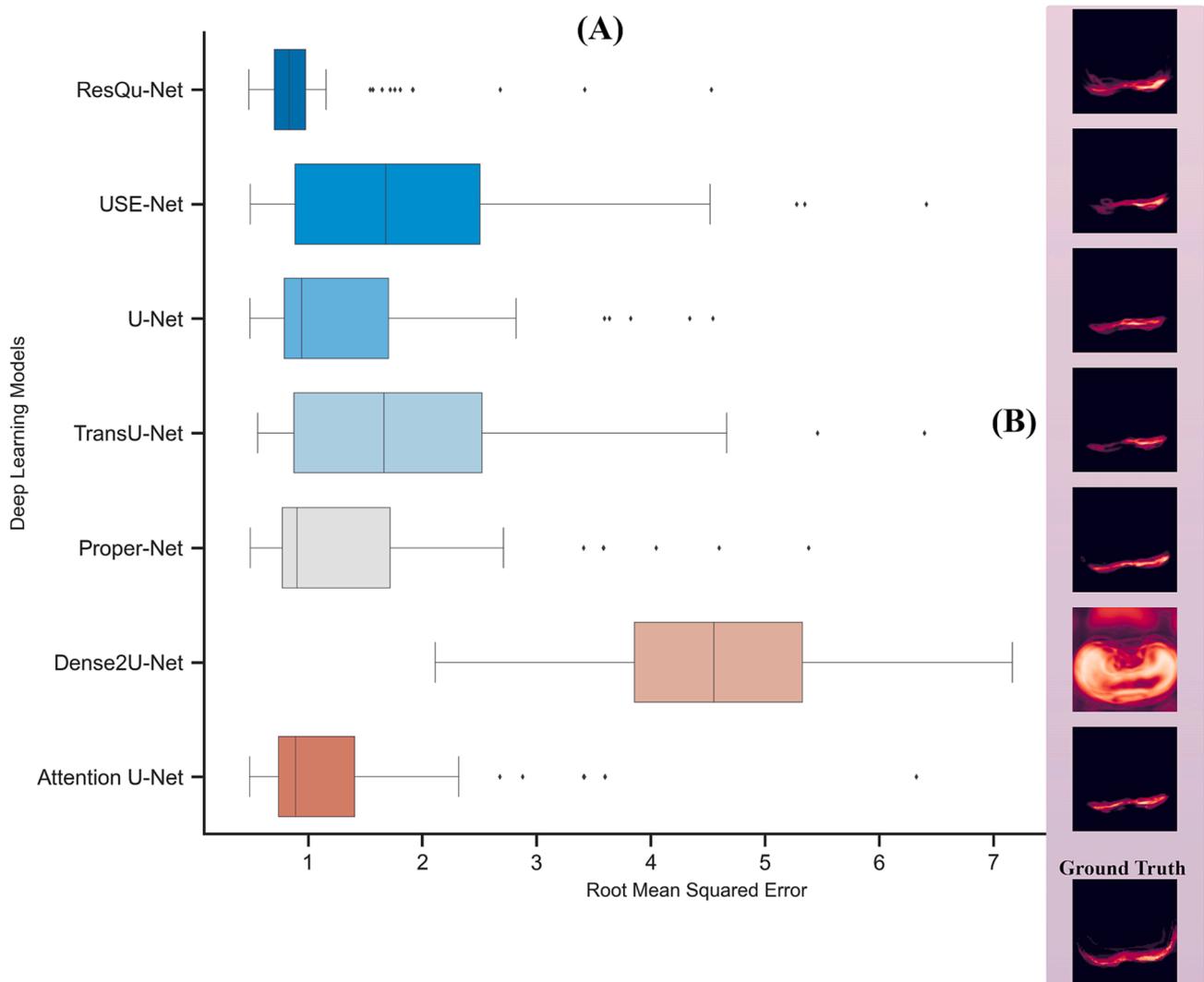


Fig. 4. Boxplots of the Root Mean Squared Error across models (A) and the corresponding SMDIs along with the GTDI for an exemplar patient (B).

demonstrates an 85 % efficacy rate in comparison to our proposed model when evaluated against the Prostate-3 T dataset. Finally, the Dense2U-Net model exhibited the lowest performance in both anatomical regions and datasets comparisons.

Fig. 4 illustrates the distribution of RMSE for each model. Each individual data point corresponds to the average RMSE between SMDI and GTDI for a given patient. The boxplot of each model depicts the distribution of all the respective data points. The lowest mean and standard deviation of RMSE was obtained using the ResQu-Net, suggesting that the model is more effective at capturing the salient features of the ground truth characteristics. This is also in line with the results presented in Tables 1–3 and Fig. 3. The most noteworthy aspect of Fig. 4 is that the USE-Net exhibits a notable deviation of its significant characteristics (SMDI) from the GTDI, which contradicts the observations provided by Table 1, Table 2, Table 3, and Fig. 3. This indicates that while the model appears to be the second-best based on quantitative assessment (mPR), the model's significant characteristics (SMDIs) are not in accordance with these obtained from the GTDIs. In the opposite direction, the Attention U-net model seems to generate higher-quality SMDIs that are similar to the GTDIs, as evidenced by the reduced RMSE. Considering all the results presented in Tables 1–3, Fig. 3 and Fig. 4, the Attention U-Net may offer higher reliability and stability compared to the USE-net. Nevertheless, the variations of observations for that model between 50 % and 75 % interquartile ranges (0.8–1.4) are much higher than these of ResQu-Net's (0.75–1.05). Finally, both quantitative and qualitative assessments highlighted that the least performant model was the Dense2U-net which was unable to identify the significant characteristics related to prostate's PZ.

4. Discussion

In this study, we proposed a novel deep learning model, named ResQu-Net, specifically designed for segmenting the PZ of the prostate on T2w MR images. Our findings demonstrate that ResQu-Net surpasses all other models in terms of segmentation performance across all metrics, datasets, and folds not only for the PZ as a whole but also for the different anatomical regions within the PZ. Additionally, the model was able to generalize across different datasets and populations, and to capture more efficiently the important features in the images, compared to other state-of-the-art DL segmentation models.

Although some previous works have partially addressed the problem of prostate zonal segmentation using CNN architectures, these models are, either more generic or are designed for simultaneously segmenting the whole prostate gland and the different zones. By solely relying on algorithms developed for other segmentation tasks, the specific characteristics and challenges of the peripheral zone may not be adequately addressed, leading to compromised segmentation accuracy and reliability. In fact, the PZ tends to have more complex boundaries, and irregular shapes than the whole gland and other prostatic zones, having a U shape at the apex and base, and a flattened U shape in the midgland region. Indicatively, different axial MRI slices are depicted in Fig. 5e, highlighting the intra- and inter-patient variability of the structure. Our analysis on the Prostate-158 dataset revealed that the distribution of elongation values (Fig. 5a) ranges from relatively spherical (elongation > 0.8) to largely elliptical (elongation < 0.8). Considering the distribution of grayscale pixel values (Fig. 5d), PZ seems to overlap to a significant degree with the grayscale values of non-relevant areas within a slice. This raises concerns regarding the ease of PZ differentiation, even by DL algorithms which are capable of capturing advanced relationships. Additionally, the distribution of PZ volume (Fig. 5b), characterized by a prominent peak, indicates a partial similarity in the size of structures in the PZ, yet some outliers exceed 40 ml while typical volumes range between 10 and 15 ml. Furthermore, the distinct and pronounced increase in texture zone entropy, shown in Fig. 5c, indicates a significant degree of textural diversity. These variations in shape, texture and grayscale values emphasize the challenges related to PZ

segmentation.

In previous studies, the reported DS performance of CNN models ranges from 0.6 up to 0.91 for PZ [8]. To date, the best performance (DS = 0.91) has been reported by the authors of the USE-NET model [34] which, in our experiments, was the second best model with DS only 0.64. One reason for this, is the implementation of different preprocessing steps by the authors, such as setting the intensities of the pixels outside of the prostate to 0, thereby forcing the algorithm to focus only in the regions within the prostate gland. Potentially the exploitation of 3D networks for segmenting the PZ, as it was recently done by Xu et al., can further enhance the performance of the models [39]. Although an increased performance was reported for the 3D U-net model (up to 0.76 in DS), the authors applied strict patient selection criteria, in both training and validation cohorts, based on the image quality and the visibility of prostate margins, which could have a significant effect on the performance of the model.

ResQu-Net's notable strength lies in its ability to segment the most challenging zone of the prostate, the PZ, demonstrating superior performance not only for the midgland region, but also for the apex and the base of the prostate, where most segmentation algorithms tend to fall short. The success of the proposed algorithm is also evidenced by the reduction in false positives and increase in true positives, illustrated in Supplementary Fig. 1. This underscores the importance of incorporating spatial attention modules and squeeze and excitation blocks in the proposed architecture, allowing ResQu-Net to leverage rich representational power and enhance the segmentation accuracy. In fact, the superiority of attention-based mechanisms in the PZ segmentation task, compared to intensive convolutions such as densely connected CNNs, is further stressed by the fact that the USE-Net and the Attention U-Net models were the second and third best-performing models, respectively, significantly surpassing the other models. However, in contrast to USE-Net and the Attention U-Net, the novelty of the ResQu-Net model lies in the integration of attention-based layers in a parallel manner, working complementarily with the convolutional layers, thereby enhancing the connectivity within the model. This parallel connectivity facilitates the exchange of information and enables the attention mechanism to effectively capture and highlight the distinctive characteristics of the PZ, leading to improved segmentation accuracy.

The experiments were conducted in a rigorous and comprehensive manner to ensure the validity and reliability of the obtained results. A 5-fold cross-validation methodology was employed, encompassing MRI data from three datasets and multiple clinical sites. By including diverse population data and accounting for variations in MRI vendors and clinical settings, enhances the reliability and validity of the findings, demonstrating that the superior performance of ResQu-Net is not limited to specific data sources or imaging protocols. In addition to evaluate the overall segmentation performance considering the PZ as a whole, a detailed analysis was conducted by breaking down the PZ into its constituent anatomical substructures, namely the base, the apex, and the midgland of the prostate. This granular assessment aimed to provide a comprehensive understanding of the models' performance across different regions of interest. By specifically examining the segmentation accuracy within each substructure, the study provided insights into the model's capability to handle variations in shape, size, and appearance within the PZ. The results showcased significant improvements for the ResQu-Net in delineating challenging regions such as the basal region, where precise segmentation is crucial for clinical decision-making as well as treatment planning.

Importantly, the comparison of ResQu-Net with state-of-the-art segmentation models was not based solely on quantitative metrics, such as the DS and the HD, but we incorporated qualitative assessments through the extraction of feature maps and comparison with the ground truth maps. This qualitative analysis enabled us to identify potential strengths and weaknesses of the algorithms, which might not be evident from quantitative metrics alone and to provide a more comprehensive understanding of models' performance. For instance, the USE-NET,

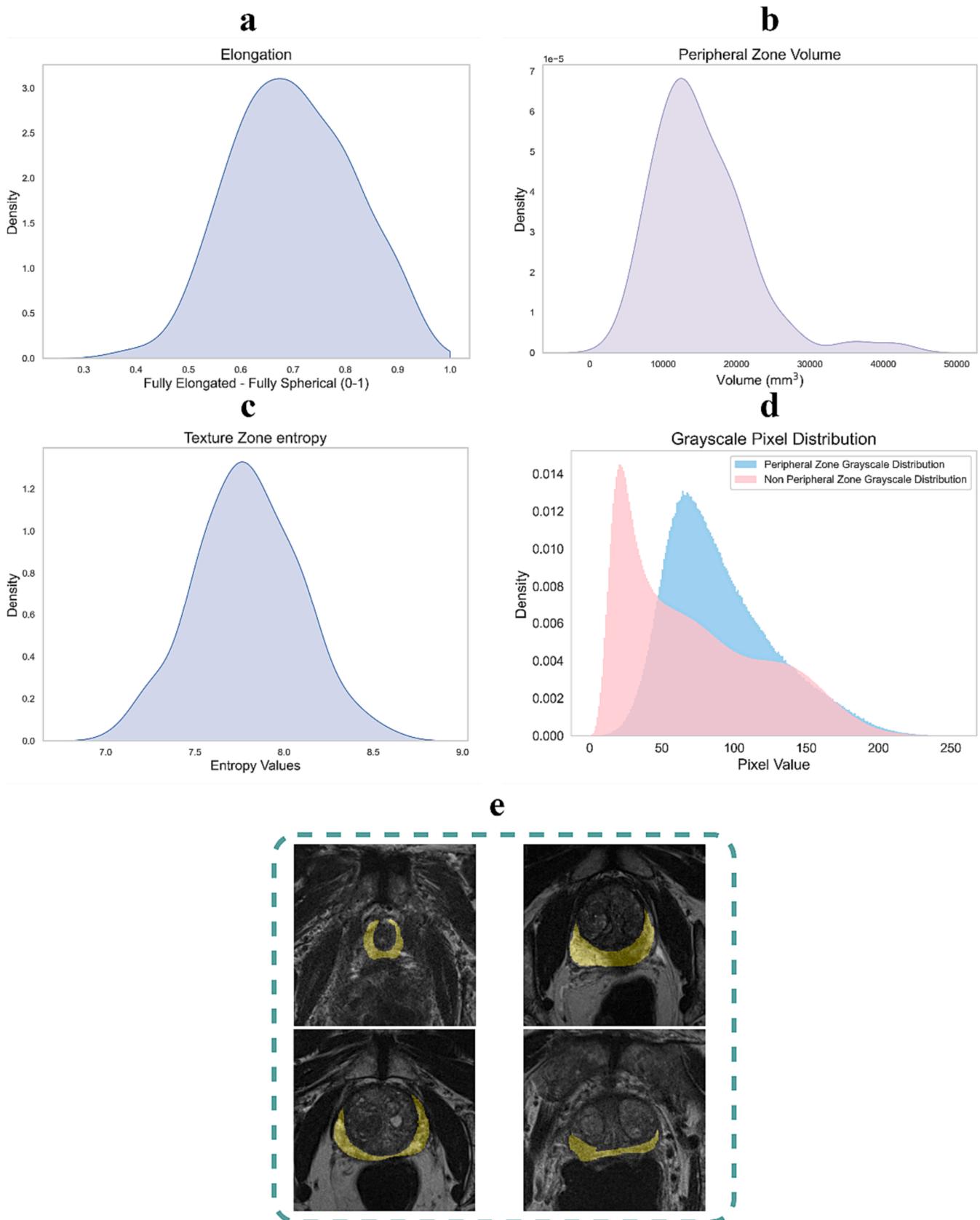


Fig. 5. Density distribution across Prostate 158 dataset for a) Elongation, b) Volume (shape features), c) gray level zone entropy (texture) features, d) grayscale pixel value distribution between peripheral zone (cyan) & Non peripheral zone (pink) pixels and e) exemplar PZ regions for different slices within the axial plane of an MRI scan.

performs relatively well in terms of DS but it tends to overlook important imaging features leading to high uncertainty in pixel wise predictions. On the other hand, the ResQu-Net processes and interprets images more efficiently, leading to a more reliable result both quantitatively and qualitatively.

5. Limitations

Despite the promising results of this study, there are some limitations that should be acknowledged and taken into consideration to guide future research and ensure the robustness and clinical utility of the proposed model. Firstly, the proposed ResQu-Net architecture was trained and evaluated using a 5-fold cross-validation approach. While cross-validation provides a reasonable estimate of model performance, it may still be susceptible to overfitting or variability in the training/validation splits. Cross-dataset and, ultimately, independent external validation on unseen datasets from different sources would provide further evidence of the model's generalizability. Also, direct comparison with the results from previous studies using datasets that were included in this analysis might not be straightforward as the data splits used herein are different. Nevertheless, a separate analysis was performed to assess the performance of the ResQu-Net model with respect to the performance reported in the original paper of the Prostate 158 datasets using the same data splits, and the results are reported in the [Supplementary material](#). Another limitation of this study is the lack of a segmentation uncertainty assessment. Quantifying and evaluating the uncertainty associated with the segmentation results is crucial for understanding the reliability and confidence of the model's predictions. Uncertainty estimation provides valuable information about areas of ambiguity or potential errors in the segmentation, enabling clinicians to make informed decisions based on the level of confidence in the model's outputs. The incorporation of uncertainty estimation techniques, such as Bayesian approaches or Monte Carlo dropout, would enhance the reliability and interpretability of the segmentation results. Future studies should aim to incorporate uncertainty assessment to provide a more comprehensive analysis of the model's performance and facilitate its clinical translation.

6. Conclusions

This work proposes an algorithmic pipeline specifically tailored for segmenting the PZ of the prostate on MRI medical imaging by capitalizing on the inherent strengths of CNNs and integrating innovative architectural designs that allowed enhancing both accuracy and efficiency. These advancements may contribute significantly in addressing the challenges associated with PZ segmentation, and ultimately enabling improved clinical decision-making and patient outcomes. To that end it is worth noting that the specific work can be part of a clinical decision support system in the domain of prostate cancer (PCa), and in particular in the diagnosis or/and characterization of cancer. Since, the performance of any software integrated in healthcare infrastructure is of paramount importance [40], the authors plan as future work to exploit various IT architectural solutions to emulate the performance of a complete pipeline in real world clinical settings.

Data Availability Statement

ProstateX2 and Prostate 3T datasets for this study are freely available throughout The Cancer Imaging Archive (TCIA) and Prostate 158 is freely available upon request from the authors through zenodo. (<https://zenodo.org/record/6481141#.ZHdfbXZBwok>).

CRediT authorship contribution statement

Dimitrios I. Zaridis: Conceptualization, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Eugenia Mylona:** Investigation, Visualization, Writing – original draft, Writing – review & editing. **Nikolaos Tachos:** Writing –

original draft, Writing – review & editing. **Charalampos N. Kalantzopoulos:** Data curation, Software. **Kostas Marias:** Writing – review & editing. **Manolis Tsiknakis:** Writing – review & editing. **George K. Matsopoulos:** Writing – review & editing. **Dimitrios D. Koutsouris:** Writing – review & editing. **Dimitrios I. Fotiadis:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have shared the link of the Data within the manuscript. The data used are openly available

Acknowledgments

This work is supported by the ProCancer-I project, funded by the European Union's Horizon 2020 research and innovation program under grant agreement No 952159 and reflects only the authors' view. The Commission is not responsible for any use that may be made of the information it contains.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bspc.2024.106187>.

References

- [1] P. Rawla, Epidemiology of prostate cancer, *World J. Oncol.*, 10(2), p. 63, 2019, 10.14740/WJON1191.
- [2] L. Wang, B. Lu, M. He, Y. Wang, Z. Wang, L. Du, Prostate Cancer Incidence and Mortality: Global Status and Temporal Trends in 89 Countries From 2000 to 2019, *Front. Public Health* 10 (Feb. 2022) 176, <https://doi.org/10.3389/FPUH.2022.811044/BIBTEX>.
- [3] R.J. Rebello et al., Prostate cancer, *Nat. Rev. Dis. Primers* 2021 7:1, 7(1), pp. 1–27, Feb. 2021, 10.1038/s41572-020-00243-0.
- [4] W.K. Oh, M. Hurwitz, A.V. D'Amico, J.P. Richie, P.W. Kantoff, *Biology of Prostate Cancer*, 2003, Accessed: May 31, 2023. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK13217/>.
- [5] D. Adler, et al., The peripheral zone of the prostate is more prone to tumor development than the transitional zone: is the ETS family the key? *Mol. Med. Rep.* 5 (2) (2012) 313–316, <https://doi.org/10.3892/MMR.2011.647>.
- [6] K.G. Holder, et al., Possible clinical implications of prostate capsule thickness and glandular epithelial cell density in benign prostate hyperplasia, *Investig. Clin. Urol* 62 (4) (Jul. 2021) 423, <https://doi.org/10.4111/ICU.20200605>.
- [7] S. Sato, T. Kimura, H. Onuma, S. Egawa, H. Takahashi, Transition zone prostate cancer is associated with better clinical outcomes than peripheral zone cancer, *BJUI Compass* 2 (3) (May 2021) 169–177, <https://doi.org/10.1002/BCO2.47>.
- [8] C. Wu, S. Montagne, D. Hamzaoui, N. Ayache, H. Delingette, R. Renard-Penna, Automatic segmentation of prostate zonal anatomy on MRI: a systematic review of the literature, *Insights Imaging* 13 (1) (Dec. 2022) 1–17, <https://doi.org/10.1186/S13244-022-01340-2/TABLES/5>.
- [9] D. Eldred-Evans, et al., Population-based prostate cancer screening with magnetic resonance imaging or ultrasonography: the IP1-PROSTAGRAM Study, *JAMA Oncol* 7 (3) (Mar. 2021) 395–402, <https://doi.org/10.1001/JAMAONCOL.2020.7456>.
- [10] B. Turkbey, et al., Prostate Imaging Reporting and Data System Version 2.1: 2019 Update of Prostate Imaging Reporting and Data System Version 2, *Eur. Urol.* 76 (3) (Sep. 2019) 340–351, <https://doi.org/10.1016/j.eururo.2019.02.033>.
- [11] Y.-H. Nai, et al., Evaluation of Multimodal Algorithms for the Segmentation of Multiparametric MRI Prostate Images, *Comput. Math. Methods. Med* 2020 (Oct. 2020) 1–12, <https://doi.org/10.1155/2020/8861035>.
- [12] S. Montagne, et al., “Challenge of prostate MRI segmentation on T2-weighted images: inter-observer variability and impact of prostate morphology”, *Insights, Imaging* 12 (1) (2021) Dec, <https://doi.org/10.1186/S13244-021-01010-9>.
- [13] A.S. Becker, et al., Variability of manual segmentation of the prostate in axial T2-weighted MRI: A multi-reader study, *Eur. J. Radiol* 121 (Dec. 2019) 108716, <https://doi.org/10.1016/J.EJRAD.2019.108716>.
- [14] K.R. Padgett, et al., Towards universal MRI atlas of the prostate and prostate zones: evaluation of performance between vendor and acquisition parameters, *Strahlenther. Onkol* 195 (2) (Feb. 2019) 121, <https://doi.org/10.1007/S00066-018-1348-5>.

- [15] O. Chilali, P. Puech, S. Lakroum, M. Diaf, S. Mordon, N. Betrouni, Gland and Zonal Segmentation of Prostate on T2W MR Images, *J. Digit. Imaging* 29 (6) (Dec. 2016) 730–736, <https://doi.org/10.1007/S10278-016-9890-0>.
- [16] N. Makni, A. Iancu, O. Colot, P. Puech, S. Mordon, N. Betrouni, Zonal segmentation of prostate using multispectral magnetic resonance images, *Med. Phys* 38 (11) (2011) 6093–6105, <https://doi.org/10.1118/1.3651610>.
- [17] G. Litjens, O. Debats, W. van de Ven, N. Karssemeijer, H. Huisman, A pattern recognition approach to zonal segmentation of the prostate on MRI, *Lecture Notes Comput. Sci. (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7511 LNCS, pp. 413–420, 2012, 10.1007/978-3-642-33418-4_51/COVER.
- [18] A. Comelli, et al., Deep Learning-Based Methods for Prostate Segmentation in Magnetic Resonance Imaging, *Appl. Sci. (basel)* 11 (2) (Jan. 2021) 1–13, <https://doi.org/10.3390/AP11020782>.
- [19] M. Baldeon-Calisto, et al., A multi-object deep neural network architecture to detect prostate anatomy in T2-weighted MRI: Performance evaluation, *Front. Nucl. Med.* 2 (Feb. 2023) 45, <https://doi.org/10.3389/FNUME.2022.1083245>.
- [20] T. Clark, J. Zhang, S. Baig, A. Wong, M.A. Haider, F. Khalvati, Fully automated segmentation of prostate whole gland and transition zone in diffusion-weighted MRI using convolutional neural networks, *J. Med. Imaging. (bellingham)* 4 (4) (Oct. 2017) 1, <https://doi.org/10.1117/1.JMI.4.4.041307>.
- [21] Y. Zhu, et al., Fully automatic segmentation on prostate MR images based on cascaded fully convolution network, *J. Magn. Reson. Imaging* 49 (4) (Apr. 2019) 1149–1156, <https://doi.org/10.1002/JMRI.26337>.
- [22] M. Bardis, et al., Segmentation of the Prostate Transition Zone and Peripheral Zone on MR Images with Deep Learning, *Radiol Imaging Cancer*, 3 (3), May 2021, 10.1148/Rycan.2021200024.
- [23] S.L. Saunders, E. Leng, B. Spilseth, N. Wasserman, G.J. Metzger, P.J. Bolan, Training convolutional networks for prostate segmentation with limited data, *IEEE Access* 9 (2021) 109214–109223, <https://doi.org/10.1109/ACCESS.2021.3100585>.
- [24] R. Cuocolo, et al., Deep learning whole-gland and zonal prostate segmentation on a public MRI dataset, *J. Magn. Reson. Imaging* 54 (2) (Aug. 2021) 452–459, <https://doi.org/10.1002/JMRI.27585>.
- [25] N. Aldoj, F. Biavati, F. Michallek, S. Stober, M. Dewey, Automatic prostate and prostate zones segmentation of magnetic resonance images using DenseNet-like U-net, *Sci. Rep* 10 (1) (Dec. 2020) 14315, <https://doi.org/10.1038/s41598-020-71080-0>.
- [26] R. Cuocolo, A. Stanzione, A. Castaldo, D.R. De Lucia, M. Imbriaco, Quality control and whole-gland, zonal and lesion annotations for the PROSTATEx challenge public dataset, *Eur. J. Radiol* 138 (May 2021) 109647, <https://doi.org/10.1016/J.EJRAD.2021.109647>.
- [27] G. Litjens, J. Futterer, H. Huisman, Data From Prostate-3T, The Cancer Imaging Archive.
- [28] L.C. Adams, et al., Prostate158 - An expert-annotated 3T MRI dataset and algorithm for prostate cancer detection, *Comput. Biol. Med* 148 (Sep. 2022) 105817, <https://doi.org/10.1016/j.compbiomed.2022.105817>.
- [29] S.G. Armato, et al., PROSTATEx Challenges for computerized classification of prostate lesions from multiparametric magnetic resonance images, *J. Med. Imaging* 5 (04) (Nov. 2018) 1, <https://doi.org/10.1117/1.JMI.5.4.044501>.
- [30] J. Hu, L. Shen, S. Albanie, G. Sun, E. Wu, Squeeze-and-Excitation Networks, *IEEE. Trans. Pattern. Anal. Mach. Intell* 42 (8) (2020) 2011–2023, <https://doi.org/10.1109/TPAMI.2019.2913372>.
- [31] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, *Lect. Notes Comput. Sci. (including Subseries Lecture Notes in Artificial Intelligence Lecture Notes Bioinformatics)* 9351 (2015) 234–241, https://doi.org/10.1007/978-3-319-24574-4_28.
- [32] E. Mylon, D. Zaridis, N. Tachos, K. Marias, M. Tsiknakis, D.I. Fotiadis, PROper-Net: A Deep-Learning Approach for Prostate’s Peripheral Zone Segmentation based on MR imaging, in: 2022 IEEE 21st Mediterranean Electrotechnical Conference (MELECON), IEEE, Jun. 2022, pp. 1124–1128. 10.1109/MELECON53508.2022.9843082.
- [33] O. Oktay et al., Attention U-Net: Learning Where to Look for the Pancreas, Apr. 2018.
- [34] L. Rundo, et al., USE-Net: Incorporating Squeeze-and-Excitation blocks into U-Net for prostate zonal segmentation of multi-institutional MRI datasets, *Neurocomputing* 365 (Nov. 2019) 31–43, <https://doi.org/10.1016/J.NEUCOM.2019.07.006>.
- [35] J. Chen et al., TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation, Feb. 2021.
- [36] D.I. Zaridis, et al., Region-adaptive magnetic resonance image enhancement for improving CNN-based segmentation of the prostate and prostatic zones, *Sci. Rep* 13 (1) (Jan. 2023) 714, <https://doi.org/10.1038/s41598-023-27671-8>.
- [37] J. Park, S. Woo, J.-Y. Lee, I.S. Kweon, BAM: Bottleneck Attention Module, Jul. 2018.
- [38] S. Woo, J. Park, J.-Y. Lee, I.S. Kweon, CBAM: Convolutional Block Attention Module, Jul. 2018.
- [39] L. Xu, et al., Development and clinical utility analysis of a prostate zonal segmentation model on T2-weighted imaging: a multicenter study, *Insights Imaging* 14 (1) (Dec. 2023) 1–11, <https://doi.org/10.1186/S13244-023-01394-W/FIGURES/4>.
- [40] V. Conti, C. Militello, L. Rundo, S. Vitabile, A novel bio-inspired approach for high-performance management in service-oriented networks, *IEEE Trans. Emerg. Top. Comput* 9 (4) (Oct. 2021) 1709–1722, <https://doi.org/10.1109/TETC.2020.3018312>.