Web-Based Deep Learning System for Evaluating Lu-PSMA Therapy Efficacy in Prostate Cancer Using PET/CT

Muhammed ElNajjar[1], Omar Shata[1], Abduallah Omran[1], Mohand Attya[1], Carole Bekhit[1]
[1] *Department of Systems and Biomedical Engineering, Faculty of Engineering, Cairo University, Giza, Egypt*

***Abstract*—Prostate Cancer is one of the leading causes of cancer-related death in men due to limited awareness, access to treatment, and late diagnosis. Theranostics—combining diagnostics and targeted therapy—has revolutionized metastatic prostate cancer treatment, significantly improving survival rates and quality of life. This paper introduces a Web-Based Theranostics System to support physicians in tracking the treatment and diagnosing prostate cancer patients undergoing PSMA-targeted radioligand therapy using an automated AI assistant. Our system consists of a DICOM Viewer to visualize, interact with, and compare different studies across time with the integration of an AI Module that automatically segments the prostate, classifies studies into normal and abnormal, and segments the lesions from the abnormal studies by incorporating the state-of-the-art (SOTA) models UNet and ResNet in segmentation and classification, respectively, achieving an Accuracy of 73% in the classification with 85% Precision and a Dice Score Coefficient (DSc) of 0.85 and 0.67 for prostate segmentation and lesion segmentation, respectively. By delivering accurate segmentation, efficient classifications, and precise longitudinal tracking analysis, our system provides a solid solution for the lack of technology for theranostics adoption in metastatic prostate cancer treatment. In addition, our solution emerged from a collaboration between industry and healthcare providers, ensuring a tangible impact, continuous development, and an evolving future.**

***Index Terms*—Theranostics, metastatic prostate cancer (mPC), UNet, ResNet, OHIF, Orthanic, PSMA, PET/CT, DICOM, PACS**

I. Introduction

According to the Global Cancer Observatory (GCO), Prostate cancer is the second most commonly diagnosed malignancy worldwide and the fifth leading cause of cancer-related death in men [1]. In Egypt, due to limited awareness [2], treatment accessibility, and late diagnosis it ranks as the fourth most diagnosed cancer in men and the fifth-leading cause of cancer mortality [3] [4]. Treatment for metastatic prostate cancer (mPC) has been transformed by theranostics, which combines targeted therapy and diagnostics to increase the quality of life and survival rates. Theranostics agents target to bind to molecular structures uniquely or predominantly expressed in cancer cells such as the prostate-specific membrane antigen (PSMA) in prostate cancer [5] [6]. Despite enormous improvements in theranostics, physicians still face challenges with time-consuming tasks related to analyzing large volumes of PET/CT or CT scans, as well as inaccuracies and delays resulting from manual tasks such as measurements and segmentations. Additionally, the unavailability of tools for longitudinal tracking and volume tracking makes it more difficult to evaluate the treatment over time. To overcome these challenges, we developed an AI-powered web-based system to help doctors diagnose and track prostate cancer patients receiving PSMA-targeted therapy by providing tools for longitudinal tracking, automatic classification, and segmentation of prostate and cancerous lesions all of which are integrated into an intuitive DICOM viewer. By automating these processes, our solution facilitates better and faster decision-making.

II. System Overview

A. System Architecture Diagram

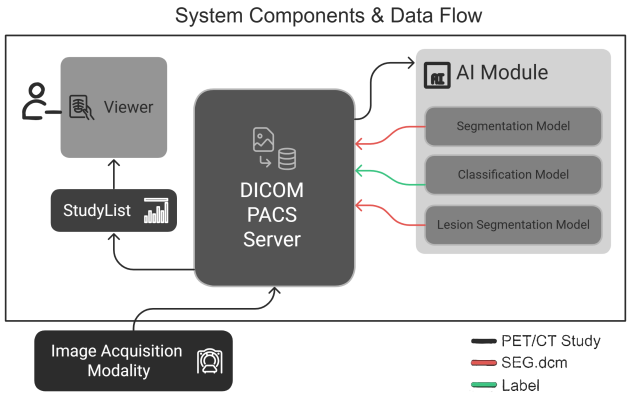

Fig. 1. The block diagram shows the system main components and the data flow.

The system is designed to streamline both the diagnostic and therapeutic processes in metastatic prostate cancer (mPC) theranostics treatment, managing within a web-based AI-driven platform. As illustrated in Figure 1, the architecture consists of four main components: the DICOM Viewer, PACS Server, Study List, and AI Module.

Once PET/CT studies are uploaded to the PACS server, they become accessible through the Study List and are visualized using the DICOM Viewer. The AI Module processes these imaging studies to automatically perform classification and segmentation. The resulting outputs are then sent to the DICOM Viewer for the physicians to visualize and perform longitudinal tracking to monitor the treatment progress.

B. System Components & Modules

1) PACS Server: The PACS (Picture Archiving and Communication System) Server manages the storage of large medical data (PET/CT imaging studies) which is essential for our system. It enables seamless integration with the DICOM Viewer and Study List, using metadata indexing to ensure fast retrieval. The system includes role-based access control and user authentication to comply with medical data privacy standards. Additionally, The PACS server contains features to enhance data management: DICOM file compression for storage and transmission time reduction, data integrity validation to ensure file consistency, automated backup routines for data protection.

2) Study List: The Study List is the centralized interface for accessing and organizing PET/CT studies, where users can search, filter, and select scans based on patient ID, date, or modality. Also, it provides a visual cue for the classification results of the study.

3) DICOM Viewer: It's a web-based viewer that provides a way of visualization for the PET/CT DICOM studies in the 3 anatomical planes; axial, coronal, and sagittal views, and their reconstructed 3D View. It includes two primary modes: **i) Default Mode:** Offers window/level adjustment, color map customization, zoom, pan, measurement tools, manual segmentation, multi-frame cine playback, and annotation capabilities. **ii) Compare Mode:** Support longitudinal study comparison with synchronized views, automated slice alignment, and trend analysis for tumor volume, tracer uptake, and lesion metrics — supporting effective treatment monitoring over time.

4) AI Module: AI module combines all the AI services that help automating mPC diagnostics and treatment tracking, and it consists of 3 state-of-the-art (SOTA) deep learning models for prostate segmentation, lesion classification, and lesion segmentation. These models will be discussed in detail in the next section.

III. DICOM VIEWER

A. Features

Multiple interactive tools and features are included in the developed DICOM viewer to improve both visualization and analysis of medical images.
The main features are:

- Measurement tools for calculating area and distance on DICOM images.
- Zoom and pan : For navigation while closely examine images
- Capture: To take photos of the current view
- Crosshair: For synchronized pointer across multiple image planes.
- Windowing for modifying contrast and brightness based on the type of tissue.
- Grid and layout support: Multiple-view grids and editable picture layouts.
- Maximum Intensity Projection (3D MIP) is used to visualize data in three dimensions.
- Undo/redo functionality
- Manual segmentation tools including brush and eraser

B. Work Flow

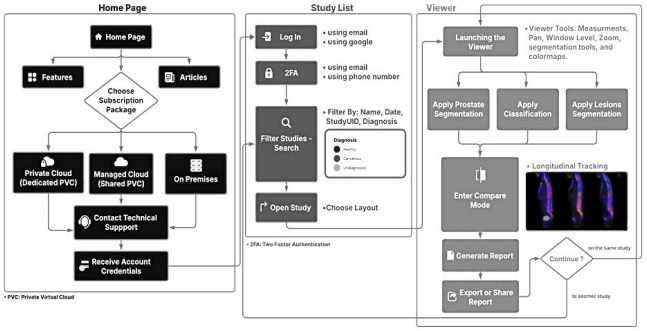

Fig. 2. User workflow

IV. AI METHODOLOGY

A. Prostate Segmentation

1) Dataset: We used a publicly available dataset from The Cancer Imaging Archive (TCIA) [7], which includes 129 CT DICOM studies. Each study contains an RTSTRUCT file that contains manual segmentations of key pelvic structures, including the prostate, bladder, rectum, and femoral heads, performed by physicians. The data set was divided into 98 (80%), 15 (10%), and 16 (10%) studies for training, validation, and testing, respectively.

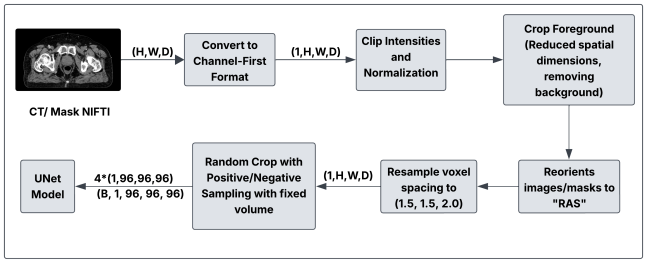

Fig. 3. Preprocessing workflow

2) Preprocessing: CT scans were converted to NIfTI , and prostate contours were extracted from RTSTRUCT files and saved as 3D binary NIfTI masks.

The preprocessing workflow as shown in Figure 3 included intensity normalization using

$$I_{\text{norm}} = \frac{\text{clip}(I - a_{\min},\, a_{\max} - a_{\min})}{a_{\max} - a_{\min}}$$

The scans were resampled to 1.5×1.5×2.0 mm, and for data augmentation, 4 cropped patches of size (96,96,96) were

extracted from each volume with a ratio of 1-1 between the mask (positive foreground) and the other parts of the image (negative background).

3) Architectures Tested: We evaluated 3 deep learning approaches: 2D UNet, Attention UNet with attention gates in the skip connections and 3D UNet [7].

4) Training and Evaluation: All models were trained using the Dice Loss function and optimized with the Adam optimizer with a learning rate of 1e-4. Each network was trained for 600 epochs. Model performance was evaluated using the Dice Similarity Coefficient (DSC) computed on the validation set. Among all tested architectures, the Attention UNet achieved the highest validation DSC and was selected for final testing as shown in Table 1.

B. Lesion Classification & Segmentation

1) Dataset: We used the publicly available FDG-PET-CT-Lesions dataset, released as part of the AutoPET 2023 Challenge[8], which includes 231 whole-body FDG PET/CT scans comprising patients with various cancer types as well as healthy individuals. Each case provides PET, CT, and segmentation NIFTI files, along with patient metadata. The dataset was split into 80%, 10%, 10% studies for training, validation, and testing, respectively.

2) Lesion Classification Methodology:

- Dataset:
  For Lesion classification, we only used Pet scans. Each scan had a binary label indicating presence or absence of cancerous lesions.
- PreProcessing:
  All PET volumes were normalized using the dataset mean and standard deviation
  $$I_{\text{norm}} = \frac{I - \mu}{\sigma}$$
  and resized to a fixed shape of (128×128×128).
- Architectures:
  First architecture was a 3D CNN built from scratch consisting of 4 sequential 3D convolutional blocks, each followed by batch normalization, ReLU activation, and max pooling. An average pooling layer was used followed by 2 fully connected layers where the last layer has a sigmoid activation function.
  A 3D ResNet50 architecture was also evaluated, the model was trained using the Binary cross-entropy loss with logits, optimized with Adam (learning rate 1×10^{-4}) and a learning rate scheduler was employed to reduce the learning rate when the validation loss stops improving.
- Evaluation:
  Model performance was assessed using accuracy , precision, recall, and F1 score, Additionally, a confusion matrix was generated to visualize the model's classification performance – as shown in Table 1 and Figure 5–

3) Lesion Segmentation Methodology:

- Preprocessing:
  For each sample, intensity normalization was applied independently to each modality: PET images were clipped to the range [0,15], while CT images were clipped to [-100,250] followed by min-max normalization to [0,1]. For data augmentation during training 4 random patches of size (128×128×32) were extracted per volume using positive-negative sampling (1:1 ratio) ensuring lesion presence in at least half of the patches.
- Architecture:
  We used a 3D UNet architecture with dual-channel input (PET and CT) and 2 output channels (foreground/background). The network consisted of five encoding-decoding levels and a downsampling stride of 2 in each layer.
- Training and Evaluation:
  The model was trained using the Dice Loss function and softmax activation on the outputs. The Adam optimizer was used with a learning rate of 1×10^{-4}, a batch size of 4, and the training was conducted for 500 epochs. Validation was performed using the Dice Similarity Coefficient (DSC), computed between predicted and ground truth lesion masks.
- Post Processing:

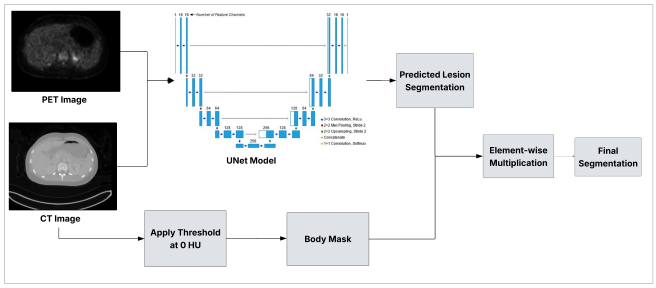

Fig. 4. Lesion segmentation pipeline

To reduce false positives outside the patient anatomy, a body mask was derived from the CT volume by applying a threshold at 0 Hounsfield Units to differentiate tissue from air. The predicted lesion mask was then element-wise multiplied with the binary body mask as shown in Figure 4 to remove any extraneous segmentations outside the anatomical region.

$$\text{cleaned\_mask} = \text{predicted\_mask} \times \text{body\_mask}$$

Performance was evaluated before and after post-processing, the cleaning step has improved the DSc with $4\% \pm 2\%$ for the testing dataset.

V. RESULTS

TABLE I
PERFORMANCE OF PROPOSED AND EXISTING METHODS

Reference	Method	Performance
Prostate Segmentation: Evaluation Metric (DSc)		
Final Approach	**Attention UNet**	**0.85**
Our Trials	3D UNet	0.84
	2D UNet	0.81
Lesion Segmentation: Evaluation Metric (DSc)		
Our Approach	**UNet**	**0.67**
Heiliger et al [9]	nnUNet + Swin UNETR*	0.72
Ahamed and Rahimm [10]	Deep Residual UNet	0.63
Scan Classification		
Custom CNN Approach	3D-Conv 4-Layers CNN	59%
Final Approach	**ResNet50**	**73%**

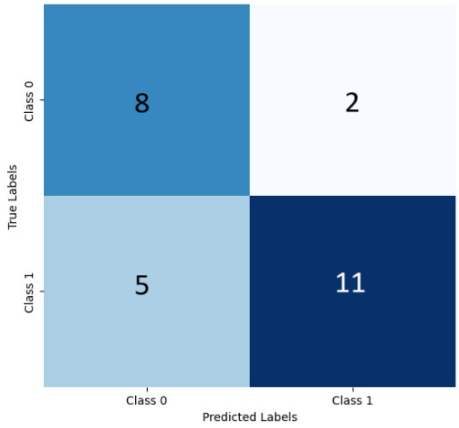

Fig. 5. Confusion Matrix

VI. DISCUSSION

A. Key Findings

Our AI-based web system has been demonstrated to possess satisfactory performance, coupled with a user-friendly and time-saving interface, making it a suitable choice for physicians working in the field of prostate cancer theranostics treatment.

In Prostate Segmentation, Attention UNet achieved a DSc of 0.86, surpassing 3D and 2D UNet variants, showing the value of the Attention Mechanism in handling complex structures.

In Classification, ResNet has reached an accuracy of 73.08%, a precision of 84.62%, and an F1 Score of 0.7586, outperforming the custom CNN architecture initially used. The confusion matrix, shown in Figure 5, illustrates the model's precision.

In Lesion Segmentation, our in-house method, which takes advantage of the 3D UNet with post-processing and preprocessing pipelines, achieved an average DSc of 0.67. While it ranks lower than Heiliger et al. [10] (DSc 0.72), it offers a lighter computationally friendly alternative for real-time web-based integration. Also, the post-processing pipeline has improved the model output by eliminating the noise and the segmentation parts outside the anatomical region of the body, increasing the DSc by 4% $\pm$ 2%.

B. Performance Variability

In Lesions-related tasks, The scores varied widely across the test dataset with DSc up to 0.89 for cases with fewer slices and larger lesions and down to 0.23 for complex scans with small or scattered lesions. After further investigations, we found that this variability results from the variation of anatomical scan range covered in the dataset itself: most scans covered the range from skull to mid-thigh, while the rest covered coverage of the entire body including the complete head and legs/feet. These anatomically different scans contained new structures that were underrepresented during model training, resulting in reduced localization and hence lower DSc.

C. Clinical Relevance

Our system is born from a collaboration between healthcare providers and the tech industry, paving the way for it to enter the healthcare market easily. In addition to supporting the clinical needs of the physicians like segmentation, classification, and longitudinal tracking, it provides a web-based solution that is compatible with PACS; making the DICOM viewer, the AI tools, and the Longitudinal tracking tools deployable on clinics and hospital systems. Since it supports DICOM images for both CT and PET modalities, it can help the physician track the effect of treatment perfectly, keep track of all the measurements, and analyze the volume of the tumors over time.

D. Limitation and future wrork

Our goal in the system is to track the efficacy of PSMA PET/CT treatment on prostate cancer, and since the data publicly available is for the FDG PET/CT we trained our model on them. This limits the models' ability to generalize and work on the PSMA PET/CT studies. Also, the classification model's results are still under development affecting its ability to be reliable for clinical decision support systems (CDSS). In our next step, we aim to train our model on a PSMA-based dataset collected from our medical partner "Misr Radiology Center". Also, we will work on improving our lesion segmentation and classification results by trying different architectures and more processing pipelines.

VII. CONCLUSION

In conclusion, we created a platform that integrates automatic PET/CT study segmentation, classification, and longitudinal tracking with visualizing and measurement tools to reduce the physicians' workload; representing an important development in the use of theranostics in the treatment of mPC. The system has shown acceptable performance across all the tasks mentioned above while being computationally efficient. Integrated with the longitudinal tracking tools gives clinicians the ability to monitor therapy response over time, enabling more data-driven decision making.

VIII. Acknowledgments

In the end, we would like to thank Dr. Ahmed Ehab and Dr. Manar Nasser for their academic supervision, advice, and their continuous support over the year. We also acknowledge the use of publicly available datasets from The Cancer Imaging Archive (TCIA) and the AutoPET Challenge. Their publicly available datasets made it possible for us to work on our solution and to reach those results.

References

[1] International Agency for Research on Cancer, "World Fact Sheet," Globocan 2020. https://gco.iarc.who.int/media/globocan/factsheets/populations/900-world-fact-sheet.pdf

[2] Aynaci, O., Tuac, Y., Mula-Hussain, L., Hammoudeh, L., Obeidat, S., Abeelh, E. A., Ibrahim, A. H., Mohammadipour, S., Alali, B., Jdaini, A., Barki, A., Mejri, N., Alhaddad, Z., Pervez, N., Hussain, H. A., Kadri, M., Elfagieh, M. A., Bounedjar, A., Junaid, M., . . . Sayan, M. (2025). Prostate cancer screening in the Middle East and North Africa: a cross-sectional study on current practices. JNCI Cancer Spectrum. https://doi.org/10.1093/jncics/pkaf019

[3] International Agency for Research on Cancer, "Egypt Fact Sheet," Globocan 2020. https://gco.iarc.who.int/media/globocan/factsheets/populations/818-egypt-fact-sheet.pdf.

[4] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021: 71: 209-249. https://doi.org/10.3322/caac.21660

[5] O'Dwyer, E., Bodei, L., and Morris, M. J. (2020). The role of theranostics in prostate cancer. Seminars in Radiation Oncology, 31(1), 71–82. https://doi.org/10.1016/j.semradonc.2020.07.004

[6] Brosch-Lenz, J., Yousefirizi, F., Zukotynski, K., Beauregard, J., Gaudet, V., Saboury, B., Rahmim, A., and Uribe, C. (2021). Role of AI in theranostics: towards routine personalized radiopharmaceutical therapies. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2107.13913

[7] The Cancer Imaging Archive (TCIA). (2023, November 27). PROSTATE-ANATOMICAL-EDGE-CASES - The Cancer Imaging Archive (TCIA). https://www.cancerimagingarchive.net/collection/prostate-anatomical-edge-cases/

[8] Ronneberger, O., Fischer, P., and Brox, T. (2015, May 18). U-NET: Convolutional Networks for Biomedical Image Segmentation. arXiv.org. https://arxiv.org/abs/1505.04597

[9] The Cancer Imaging Archive (TCIA), "FDG-PET-CT-Lesions Collection," https://www.cancerimagingarchive.net/collection/fdg-pet-ct-lesions/

[10] L. Heiliger, P. F. Christ, J. Toennies, P. Mildenberger, and B. Thomas, "AutoPET Challenge: Combining nnU-Net with Swin UNETR augmented by Maximum Intensity Projection Classifier," https://arxiv.org/abs/2209.01112

[11] S. Ahamed and A. Rahmim, "Generalized Dice Focal Loss trained 3D Residual UNet for automated lesion segmentation in whole-body FDG PET/CT images," https://arxiv.org/abs/2309.13553

[12] Wasserthal, J., Breit, H., Meyer, M. T., Pradella, M., Hinck, D., Sauter, A. W., Heye, T., Boll, D. T., Cyriac, J., Yang, S., Bach, M., and Segeroth, M. (2023). TotalSegmentator: robust segmentation of 104 anatomic structures in CT images. Radiology Artificial Intelligence, 5(5). https://doi.org/10.1148/ryai.230024