Journal of Advanced Engineering Trends
ISSN : 2682 - 2091

Vol. 44, No. 1, January 2025

Automatic Bladder Cancer Segmentation Using Deep Learning

Lamia N. Omran[1], Kadry A. Ezzat[1,*], Hossam El Fadally[2], Emad G. Shehata[2], Gerges M. Salama[2]

[1] *Biomedical Engineering Dep., Higher Technological Institute, 10th of Ramadan City, Ash Sharqia, Egypt*

[2] *Electrical Engineering Dep., Faculty of Engineering, Minia University, Minia, Egypt*

** Corresponding author(s) E-mail: kadry_ezat@hotmail.com*

ARTICLE INFO

Article history:
Received: 26 August 2024
Accepted: 8 February 2025
Online: 26 February 2025

Keywords:
CNN
U-NET
Augmentation
Region of Interest
Diagnostic

ABSTRACT

Bladder cancer is a prevalent and potentially life-threatening disease that requires accurate diagnosis and treatment planning. Medical image segmentation plays a crucial role in the assessment of tumor location, size, and progression. In this paper, We investigated the application of a U-Net based deep learning mode for bladder cancer segmentation. To perform bladder cancer segmentation using the U-Net technique, a diverse dataset of bladder cancer images is collected, comprising various stages and types of bladder cancer. The images are preprocessed to enhance contrast and remove noise, ensuring optimal input quality for the U-Net model. Subsequently, the U-Net model is trained using a large set of annotated images, where pixel-wise annotations serve as ground truth for the segmentation task. Initial tests using five sets of the cancer imaging archive "TCIA" dataset show that the suggested algorithm achieves an average accuracy of 87.28% and an average time of 6.0 minutes without parallelized computation, clearly surpassing other current techniques for bladder segmentation.

1. Introduction

Bladder cancer ranks as the 10th most frequently identified cancer globally, with 583,319 new cases reported in 2021 [1]. Urothelial carcinoma is predominant form of bladder cancer, with a higher prevalence among men. It is one of the cancer types that have highest rate of coming back, around 50 to 70%, making it particularly difficult [2]. It necessitates a rigorous treatment and aftercare strategy, making it one of cancer types with the highest cost of treatment per patient over their lifetime [3], [4].

Pathologists utilize grading and staging in diagnostics to characterize tumor. These criteria help categorize patients based on risk level and develop an appropriate treatment and monitoring strategy. The tumor grade indicates how differentiated tumor cells appear when viewed under a microscope. Various types of cancer use various grading systems, but generally, when cancer cells closely resemble healthy cells, the grade is typically low, indicating a lower tendency for spreading [5]. Conversely, a higher grade will be given if the cells show a more abnormal appearance and lack organization. Aside from the grade, the tumor stage is also significant and is assessed based on the size of the primary tumor, extent of spreading into surrounding tissue, and the quantity of primary tumors. This paper highlights the assessment of Non-Muscle Invasive Bladder Cancer "NMIBC". Nonetheless, there is a widely recognized issue with the accurate grading of bladder cancer caused by variability within and between observers, as well as inconsistent reproducibility among pathologists[6], potentially resulting in inadequate or excessive treatment of patients. Therefore, there is a need for new tools to assist pathologists. World Health Organization "WHO" has suggested three classification systems for bladder cancer. The initial grading system, known as WHO72 [7], was implemented in 1972 and continues to be partially utilized in the present day. It is made up of three levels: grade one, grade two, and grade three, with grade 3 being the most serious level. WHO05, introduced in 2005, was later updated in 2017 as WHO17 [8], a revised version of the grading system. In these variations, instances are divided into low- and high-grade carcinoma.

Patients in grade one are categorized as low-grade patients, while grade three patients are considered high-grade patients. Patients previously classified as Grade two are now categorized as either low-grade or high-grade cases [9]. This may appear to be a small modification, but being classified as low- or high-grade as a patient could lead to varied follow-up procedures and local treatments with possible negative consequences [10].

Incorrectly diagnosing a patient as high-risk leads to unnecessary suffering, increased burden on the healthcare system, and additional expenses. The information utilized in this article was gathered and analyzed before 2017 [11], so it will concentrate on the WHO05 grading system [12]. Emergence of a new era in the field of pathology growth in medical research applications using machine technology learning and deep learning methods [13]. Majority of studies are connected to cancer conditions like breast, lung, prostate, and brain and melanoma [14]. Examining

the US Food list reveals. The Food and Drug Administration "FDA" has authorized AI-driven medical technologies in various fields, with most being approved of radiology, cardiology, and Internal Medicine/General Practice is important for improvement [15]. However, there is also a significant amount of focus directed towards images of tissues under a microscope [16].

Earlier research conducted by our team centered on bladder cancer, involving tissue segmentation as well as anticipating recurrence in NMBC patients. The results of literature studies for example, Wetteland et al. [13] tested various combinations of three different magnification levels with accuracy 73.7.

In the study by Skrede et al. [17], the whole slide image is initially divided into segments, followed by the extraction of tiles at 20x and 50x resolution with accuracy 66.2. Ten Convolutional Neural Network "CNN-"based models are utilized in an ensemble to analyze tiles from each scale. The ensembles' average score is utilized for predicting the prognosis of colorectal patients.

Certain studies have focused on urothelial carcinoma, also identified as bladder cancer. In Jansen et al. [18], two separate single-scale neural networks were employed to identify and assess 338 instances of bladder cancer gathered from 236 individuals with accuracy 77.8.

An AI model was trained to extract tissue utilizing a U-net based network, which was then utilized as input for another network trained to assign grade to tissue based on WHO05 grading system. Categorization network evaluated WHO05 classification on a slide-level by considering the majority vote of all categorized tiles. Forecasts were evaluated against assessments of three skilled pathologists. As per general agreement, classification model attained a 78% accuracy. All whole-slide images provided were obtained with a magnification of 21x (0.6 μm per pixel).

Zhang and colleagues [19] also conducted research on bladder cancer with accuracy 82.5. They utilized four separate neural networks known as s-net, d-net. The s-net model is structured similarly to U-net and categorizes every pixel as either tumor or non-tumor. d-net further defines tumor Region of Interest "ROI" then creates a clear diagnosis and compact encodings. Ultimately, a-net utilizes ROI encodings to forecast a WHO05 grading at the slide level.

Lucas et al. [20] from the same research team used the urothelium segmentation model from [18]. Sections of urothelium were inputted into a selection system that categorized tiles as either having recurrence or not having recurrence.

In their study, Hashimoto et al. [21] used a multi-scale CNN-based model to analyze WSIs of malignant lymphoma with accuracy 76.5. They evaluated the outcomes of models by analyzing tiles extracted at either 20x or 40x resolution. Nevertheless, most effective outcome was obtained by merging two scales to create a multi-scale model. Researchers in this case also verify the presence of class-specific characteristics at various levels of magnification.

CNNs were initially created and implemented in the 1980s. A CNN can achieve the maximum possible. The only thing one could do back then was identify handwritten numbers. It was primarily utilized in the industry postal areas should be identified by their corresponding zip codes or pin codes. One key aspect of any deep learning model is the need for a significant amount of data for training also necessitates a significant amount of computational resources.

This posed a significant limitation for CNNs on during that time, CNNs were only used in postal services and were unsuccessful in other sectors. explore the realm of machine learning known as convolutional neural networks (CNNs) to tackle the ImageNet challenge learning which employs complex neural networks with multiple layers ample amounts of data, referring specifically to ImageNet datasets containing millions of labeled images and researchers were able to resurrect CNNs thanks to a surplus of computing resources.

The primary accomplishments of the study include the development of an algorithm for determining the best decision threshold for classifying WSIs at the slide-level. We developed models using training sets of various sizes. Outcomes of this study offer information on how the size of the dataset impacts the models' performance, the time needed for training per epoch, and number of epochs necessary before meeting the early stopping criteria

2. Methods

The system suggested in this paper employs multi-scale models, which take tiles obtained at various magnification stages as input. To enhance readability, researchers categorize these tiles as a trio. A triplet, labelled as Qi, is a group of different tiles taken from a whole slide image at different magnification levels: 30x, 110x, and 300x. In a WSI, let Q represent a collection of triplets, where Q={Q1,Q2…Qi…Qmax}, with |Q| indicating set size.

2.1. Material Data

The dataset comprises 350 digital whole-slide images from five patients with NMBC as shown in Figure 1, obtained through transurethral resection of bladder tumor tissue. Information was gathered from TCIA in January 2024 [22]. Dataset includes all bladder cancers that are non-muscle invasive, making it a representative population dataset Units.

Figure 1: sample of bladder image dataset

The study divided 350 WSIs into training (240), validation (50), and testing (60) sets. Demographic characteristics were not

considered when dividing data into separate datasets. On other hand, WSIs were chosen randomly then divided into groups to ensure an equal representation of all diagnostic results according to WHO72 and WHO05 grading, stage, recurrence, and disease advancement, in order to accurately represent data set. The breakdown of low and high grade WSI in every dataset were used along with quantity of triplets in test set and validation.

Certain WSIs in dataset have minimal urothelium due to the small size of tissue sample or the presence of damaged or other tissue types. An augmentation technique was used for these WSI, in which a randomly chosen set of triplets were augmented. The goal of this procedure is for every WSI to make a similar or nearly equal contribution to the specified triplet number determined by N.

Rotation and vertical/horizontal mirroring were used to augment individual tiles in the triplet. Each tile in the trio was enhanced in the same way. By using both rotation and mirroring, a tile can be positioned in nine distinct orientations, which is the most ways a single tile can be enhanced. Some WSIs failed to achieve the expected triplet count, despite using 9x augmentation. The validation and test datasets remained unchanged without any augmentation.

2.2. Proposed System

A visual representation of the architecture of the -model is shown in Figure2

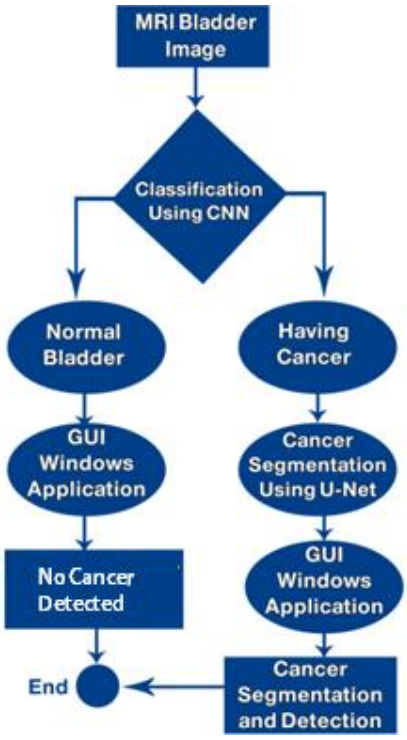

Figure 2: Flowchart of Proposed system

The pipeline illustrated includes CNN-based model for distinct purposes: classifies tissue and the grades urothelium tissue. The structure comprises of three distinct VGG16 networks, each one designed for a specific input scale.

The model structure and the TRI-terminology are derived from our earlier research on the tissue model in Wetteland [13].

The model takes in a set of three 512×512 pixel tiles (30x, 110x, and 300x) as input. The model has the capability to forecast triplets taken from any part of the whole slide image, however, a selective foreground mask is typically employed to expedite processing by eliminating the background.

The TRItissue-model produces a likelihood distribution across the six anticipated categories: muscle, stroma, urothelium, damaged tissue, blood and background. The model requires a set of three 512 × 512 pixel images taken from urothelium tissue regions as input. Model provides a probability distribution for both high-grade and low-grade carcinoma classes.

TensorFlow and Keras libraries used for the training of a Convolutional Neural Network. Convolutional Neural Network (CNN) used for categorizing images. TensorFlow and Keras are software tools utilized for creating and educating in field of machine learning. artificial intelligence mechanisms called neural networks. TensorFlow is utilized as the backend for Keras in this code.

TensorFlow offers the fundamental operations needed to carry out tasks. Performing complex calculations on extensive tensors and executing computations on graphic processing units. Keras, in contrast, offers a top-level interface for constructing and educating neural networks.

It simplifies the process of defining the neural network's structure and helps in understanding. delineate the training parameters. Keras includes integrated functionality for loading and preparing image datasets simplifies the process of handling image data.

The initial section of the code imports a set of bladder images from a directory. The pictures are divided into a training set and a validation set, with the same size. the size of the image is 512x512 pixels. The batch_size parameter determines the size of each batch.

There are numerous images to handle simultaneously while training. The CNN structure is then determined. The structure includes multiple levels, such as Conv2D levels, MaxPooling2D levels, and dense levels. The Conv2D layers use a filter on the input image to capture characteristics.

The MaxPooling2D layers reduce the size of the output from the convolutional layers. The Flatten layer transforms the output of the convolutional layers into a flat representation. The compact layers carry out classification on the compressed output. The final one A dense layer contains just one neuron that uses a sigmoid activation function.

Generates a likelihood ranging from 0 to 1. Once the model architecture is defined, the next step is to compile the model. The optimizer parameter defines the optimization process. The algorithm utilized in training employed to assess the model's performance.

The measurements parameter determines the criteria for evaluating the efficiency of the model while in the training phase. .

parameter determines how many times to loop through the entire dataset.

In the end, the model that has been trained will be utilized to forecast if a new bladder image. Whether or not someone has cancer.

The load_img method is used to load the image. The predict method is used by the model to anticipate the likelihood of cancer. In case if the probability is below 0.5, the model anticipates that the bladder is cancerous. Alternatively, it anticipates that it is typical.

The U-Net design as shown in Figure 3 is comprised of two primary parts encoder path and decoder path. The U-shaped structure of the decoder path captures contextual information from the input urinary bladder image. It usually adheres to a layout of a convolutional neural network (CNN).

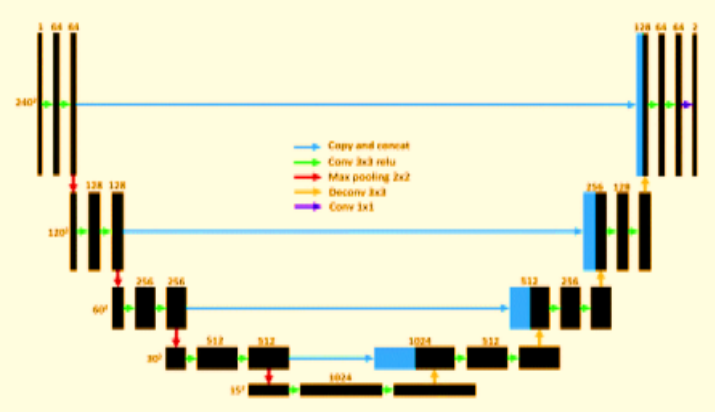

Figure 3: U-Net Technique

The image enters a sequence of convolutional layers before being processed gradually. Reduce the size of the spatial dimensions while simultaneously boosting the quantity of this reduction in feature channels, assists the network in capturing hierarchical and abstract information capturing features at various scales from the input image.

The decoder path uses skip connections and up-convolutional layers to reconstruct the segmented output. Skip connections involve concatenating feature maps from the encoder path to matching layers in the decoder path. These connections aid in restoring spatial information that may have been overlooked during the encoding stage. Up-convolutional layers, also referred to as transpose convolutions or deconvolutions, assist in this process.

By increasing the resolution of the low-resolution feature maps to match the original image size, the decoder path creates high-resolution segmentation outputs using skip connections and up-convolutions.

U-Net undergoes supervised learning with annotated biomedical image datasets during training. The aim of the training of the network is to reduce a pixel-wise loss function, like the dice loss that is calculated using either coefficient or cross-entropy, which determines how alike two things are projected result and the actual segmentation mask.

Throughout the training process, the network fine-tunes its parameters by utilizing back propagation improving segmentation accuracy by optimizing the loss function. The U-Net design is customized for the segmentation of the urinary bladder images.

The incorporation of skip connections, encoder-decoder structure, and design choices enables U-Net to function efficiently capture local and global characteristics, manage intricate formations, and restore precise details.

Due to this, U-Net has been widely embraced and attained notable success achieving success in tasks like cell segmentation and organ segmentation in biomedical image analysis partitioning, and identifying tumors.

The design of a U-Net structure for imaging of the urinary bladder is as follow:

1- The process of segmentation with the Tensor Flow-Keras framework involves importing the required layers and models such as Conv2D.
2- Normalization of batches, activation function, maximum pooling, transposed convolution.
3- The conv_block() function is stated to generate a convolutional block. Interiors of the operation, dual 3x3 convolution layers paired with batch normalization and ReLU function activation functions are used on the input tensor.
4- The purpose of the encoder_block() is processed by a convolutional block, then it undergoes max pooling. Performing (2x2) operation to reduce the size of the spatial dimensions.
5- A transpose is performed within the function. The input tensor is first up sampled by a convolutional layer before further processing combining skip connection elements through concatenation. The combined tensor stands then moved through a block of convolution.
6- The function build_unet() is created for constructing the U-Net model. It accepts the input shape as an argument. Within the function, the input tensor is established utilizing the Input() layer. The U-Net structure is created by linking encoder and decoder modules.
7- The encoder_block() function is used to create four encoder blocks. In a sequential manner, the skip connection characteristics are saved. A convolutional block is used to create a bottleneck block (b1) from the output of the previous encoder block. Four decoding blocks are generated with the decoder_block() function.
8- Each decoder block receives the bottleneck block along with the corresponding skip connection characteristics found in the encoder pathway.
9- The ultimate result is produced through the utilization of a 1x1 convolutional layer.

10- The Model() class is used to instantiate the model, including the inputs and configurations outputs are specified and linked to the variable model.
11- The input shape of the model is specified in the code's main block. The U-Net model is constructed using build_unet() with input dimensions of (512, 512, 3). A function that takes the shape as an input parameter.
12- The summary() function is invoked on the model to show a summary of the design of the model structure.

To evaluate our model, the following steps were applied:

1- The required packages are brought in: os, numpy, cv2, pandas, glob, tqdm. tensorflow, along with PIL. Certain modules and functions that are particular libraries are imported as well.
2- The function create_dir() is specified to make a directory in case it's not present. and predicted mask as images in the specified output directory. and anticipated mask combined into one image for visualization reasons. The function combines the images together and inserts lines in between as separators.
3- The outcome is stored in the designated location. The beginning of the code involves establishing the random seeds for. ensuring the ability to replicate and establishing a folder to save the findings.
4- The U-Net model that has been trained is loaded with the function tf.keras.models.load_model(). The text needs to be provided for me to be able to paraphrase it while keeping the same number of words custom metrics such as iou, dice_coef, and dice_loss.
5- CustomObjectScope is used to guarantee proper loading of the model. The paths for the input images and true masks in the test dataset are imported by loading them.
6- the glob() command. The test dataset lengths are displayed. A loop is run to assess and forecast on every test sample. In exchange for every For instance, the file's name is taken out. The image is loaded with cv2.imread() and then scaled to be within the range of [0,1] and increased in size to align with the model's input shape.
7- The actual mask is accessed using cv2.imread(), then adjusted to the [0, 1] interval. Converted to integer data type after applying threshold of 0.5 to get binary values.
8- The model forecasts the mask for the input image and deals with it through. Converting to int32 data type after setting a threshold of 0.5 for obtaining binary values
9- The save_results() function is used to store the input image, true mask, and forecasted mask as one singular image.

10- The actual mask and forecasted mask arrays are compressed. Different assessment criteria like precision, F1 measure, Jaccard index, and recall.
11- Accuracy and precision are evaluated using functions available in the sklearn.metrics library component.
12- The values for metrics are saved in a collection named SCORE.The mean of each metric is used to calculate the average metrics values across all test cases.
13- The mean values for the metrics are displayed. The metrics values are stored in a DataFrame and then exported as a CSV file.

This algorithm tests a U-Net model that has been trained on the given test dataset and then saves it projected masks alongside the corresponding input images and actual masks, and determines several assessment criteria. The outcomes are stored and ready for additional analysis examined or utilized for the purpose of reporting on the model's performance as shown in Figure 4.

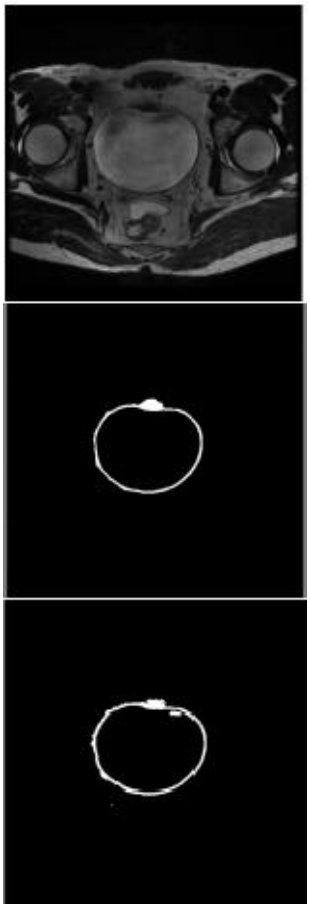

Figure 4: Segmentation evaluation of urinary bladder

3. Results

The suggested model was implemented using a software (Python). The utilization was specifically for GPUs. All the experiments were conducted using a computer equipped with an Intel Core i7 processor running at 4 GHz and 32 GB of RAM.

The findings from delineating the ROI for the bladder can be observed in Figure 5. This information served as the primary input for the remaining stages of the system

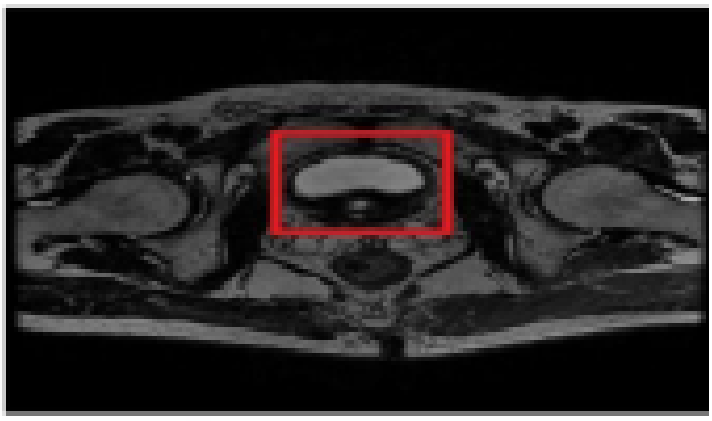

Figure 5: ROI for urinary bladder

As shown in Figure 6, the result of applying a urinary bladder segmentation.

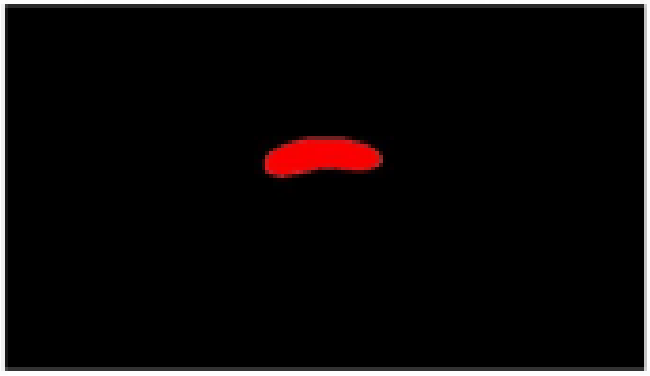

Figure 6: Segmentation of urinary bladder

This model makes predictions on the DICOM images provided using the trained model U-Net. The input image, predicted mask, and blended image are all stored. Featuring a highlighted area of interest in red square. The outcomes are shown in figure 7.

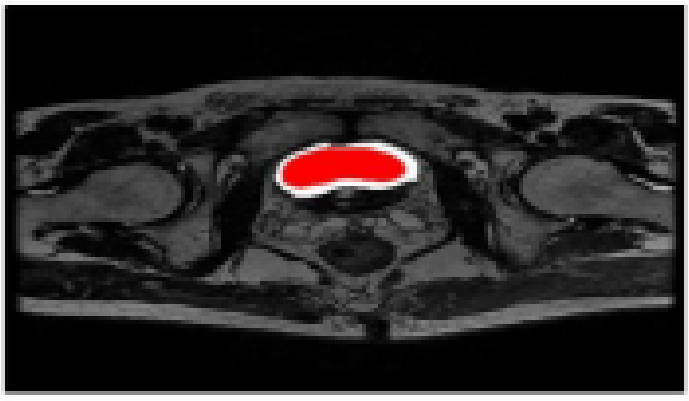

Figure 7: predicted urinary bladder

The U-Net model is trained using a large set of annotated images, where pixel-wise annotations serve as ground truth for the segmentation task.

Initial tests using five sets of TCIA dataset show that the suggested algorithm achieves an average Accuracy of 87.28% and an average time of 6.0 minutes without parallelized computation, clearly surpassing other current techniques for bladder segmentation as shown in Figure 8.

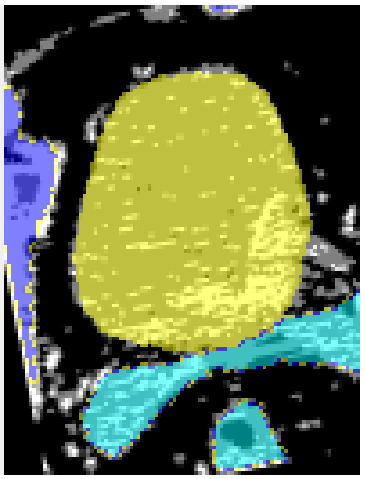

Figure 8: The resulting segmented image from the model applied to the original image

Figure 9 shows the segmented tumor from bladder

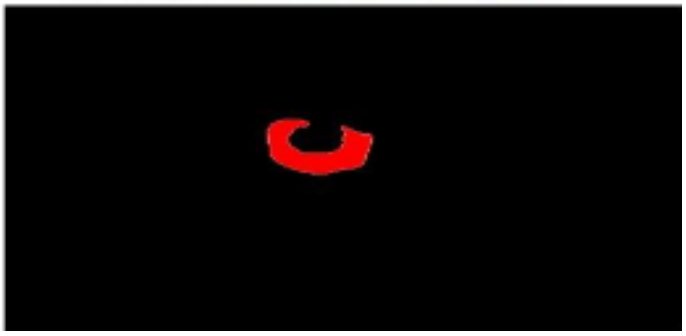

Figure 9: The resulting segmented tumor from the segmented bladder

Table 1 shows the comparison between the proposed system and other literatures

Literatures	Accuracy
Wetteland et al.[13]	73.7%
Skrede et al. [17]	66.2%
Jansen et al. [18]	77.8%
Zhang et al .[19]	82.5%
The proposed system	87.28%

4. Conclusion

Segmenting bladder cancer is an important job in the medical imaging field providing substantial advantages in the detection and management of bladder cancer. Having the arrival of advanced computational methods, such as machine learning and

innovation significant advancements have been achieved in the automation of deep learning algorithms process of dividing into segments.

Automated segmentation techniques, especially those utilizing deep learning algorithms models like convolutional neural networks (CNNs), have demonstrated outstanding performance precision in identifying bladder lesions from medical imaging.

These methods possess the ability to decrease reliance on manual engagement, ultimately saving time and effort reducing discrepancies between different observers. Accurate segmentation of bladder tumors allows clinicians to precisely determine their size study tumor properties, track disease advancement, and evaluate effectiveness of treatment.

Quantitative data, like the size and shape of tumors, assists in the planning of treatment, assessment, and individualized patient care. Moreover, the advancement of computer-assisted diagnosis systems grounded on segmenting bladder cancer improves clinicians' abilities by offering. further understanding and assistance in making decisions. These systems are able to help in the detection of issues at an early stage.

Detecting, enhancing diagnostic precision, and easing treatment enhancement. Nevertheless, difficulties still remain in the segmentation of bladder cancer, such as image quality issues differences, disturbances, and the requirement of more extensive labeled datasets.

Research to be conducted in the future efforts should be directed towards improving current algorithms and investigating multiple modes combining various imaging methods and utilizing cutting-edge techniques methods like generative adversarial networks (GANs).

References

[1] Chao-Zhe Zhu, Hua-Nong Ting, Kwan-Hoong Ng, Kein-Seong Mun, Teng-Aik Ong, "Dielectric properties of urine in relation to bladder cancer", *Physical and Engineering Sciences in Medicine*, 2023.

[2] Yupan Wu, Yingqi Meng, "Separation and characterization of cells using electrical field", *Robotics for Cell Manipulation and Characterization*, pp.355, 2023.

[3] D. Jocham, H. Stepp and R. Waidelich, "Photodynamic diagnosis in urology: State-of-the-art", *Eur. Urology*, vol. 53, no. 6, pp. 1138-1150, Jun. 2008

[4] D. M. Burke, D. C. Shackley and P. H. O'Reilly, "The community-based morbidity of flexible cystoscopy", *BJU Int.*, vol. 89, no. 4, pp. 347-349, Feb. 2002.

[5] Giorgio Ivan Russo, Nicolo Musso, Alessandra Romano, Giuseppe Caruso, Salvatore Petralia, Luca Lanzano, Giuseppe Broggi, Massimo Camarda, "The Role of Dielectrophoresis for Cancer Diagnosis and Prognosis", *Cancers*, vol.14, no.1, pp.198, 2021.

[6] Christopher Andreassen, Saul Fuster, Helga Hardardottir, Emiel A.M. Janssen, Kjersti Engan, "Deep Learning for Predicting Metastasis on Melanoma Wsis", *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, pp.1-5, 2023.

[7] Apeksha Koul, Yogesh Kumar, Anish Gupta, "A Study on Bladder Cancer Detection using AI-based Learning Techniques", *2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS)*, pp.600-604, 2022.

[8] Neel Kanwal, Fernando Pérez-Bueno, Arne Schmidt, Kjersti Engan, Rafael Molina, "The Devil is in the Details: Whole Slide Image Acquisition and Processing for Artifacts Detection, Color Variation, and Data Augmentation: A Review", *IEEE Access*, vol.10, pp.58821-58844, 2022.

[9] Julius Drachneris, Mindaugas Morkunas, Mantas Fabijonavicius, Albertas Cekauskas, Feliksas Jankevicius, Arvydas Laurinavicius, "Prediction of Non-Muscle Invasive Papillary Urothelial Carcinoma Relapse from Hematoxylin–Eosin Images Using Deep Multiple Instance Learning in Patients Treated with Bacille Calmette–Guérin Immunotherapy", *Biomedicines*, vol.12, no.2, pp.360, 2024

[10] Inanc Moran, Deniz Turgay Altilar, Muhammed Kursad Ucar, Cahit Bilgin, Mehmet Recep Bozkurt, "Deep Transfer Learning for Chronic Obstructive Pulmonary Disease Detection Utilizing Electrocardiogram Signals", *IEEE Access*, vol.11, pp.40629-40644, 2023.

[11] Jiening Zhu, Harini Veeraraghavan, Jue Jiang, Jung Hun Oh, Larry Norton, Joseph O. Deasy, Allen Tannenbaum, "Wasserstein HOG: Local Directionality Extraction via Optimal Transport", *IEEE Transactions on Medical Imaging*, vol.43, no.3, pp.916-927, 2024.

[12] Geetha C, Anusha R, Prabhu V, Abinaya Kamatchi S, Gejashree T, "Smart Tool for Earlier Prediction of Breast Cancer Using AI", *2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE)*, pp.1-7, 2023.

[13] R. Wetteland, K. Engan, T. Eftestøl, V. Kvikstad, and E. A. M. Janssen, ''A multiscale approach for whole-slide image segmentation of five tissue classes in urothelial carcinoma slides,'' Technol. Cancer Res. Treatment, vol. 19, Jan. 2020, Art. no. 153303382094678, doi: 10.1177/1533033820946787.

[14] R. Wetteland, K. Engan, T. Eftestøl, V. Kvikstad, and E. Janssen, ''Multiclass tissue classification of whole-slide histological images using convolutional neural networks,'' in Proc. 8th Int. Conf. Pattern Recognit. Appl. Methods, vol. 1, 2019, pp. 320–327, doi: 10.5220/0007253603 200327.

[15] Archana Mathur, Nikhilanand Arya, Kitsuchart Pasupa, Sriparna Saha, Sudeepa Roy Dey, Snehanshu Saha, "Breast cancer prognosis through the use of multi-modal classifiers: current state of the art and the way forward", *Briefings in Functional Genomics*, 2024.

[16] Raj Gaurang Tiwari, Himani Maheshwari, Ambuj Kumar Agarwal, Ochin Sharma, N. Bharathiraja, "Endoscopic Bladder Tissue Classification Through Fusion of Handcrafted and Deep Features", *2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, pp.708-713, 2023.

[17] O.-J. Skrede, S. De Raedt, A. Kleppe, T. S. Hveem, K. Liestøl, J. Maddison, H. A. Askautrud, M. Pradhan, J. A. Nesheim, F. Albregtsen, and I. N. Farstad, ''Deep learning for prediction of colorectal cancer outcome: A discovery and validation study,'' Lancet, vol. 395, no. 10221, pp. 350–360, 2020, doi: 10.1016/S0140-6736(19)32998-8.

[18] I. Jansen, M. Lucas, J. Bosschieter, O. J. de Boer, S. L. Meijer, T. G. van Leeuwen, H. A. Marquering, J. A. Nieuwenhuijzen, D. M. de Bruin, and C. D. Savci-Heijink, ''Automated detection and grading of non–muscle-invasive urothelial cell carcinoma of the bladder,'' Amer. J. Pathol., vol. 190, no. 7, pp. 1483–1490, Jul. 2020, doi: 10.1016/j.ajpath.2020.03.013.

[19] Z. Zhang, P. Chen, M. McGough, F. Xing, C. Wang, M. Bui, Y. Xie, M. Sapkota, L. Cui, J. Dhillon, N. Ahmad, F. K. Khalil, S. I. Dickinson, X. Shi, F. Liu, H. Su, J. Cai, and L. Yang, ''Pathologist-level interpretable whole-slide cancer diagnosis with deep learning,'' Nature Mach. Intell., vol. 1, no. 5, pp. 236–245, May 2019, doi: 10.1038/s42256-019-0052-1.

[20] M. Lucas, I. Jansen, T. G. van Leeuwen, J. R. Oddens, D. M. de Bruin, and H. A. Marquering, ''Deep learning–based recurrence prediction in patients with non–muscle-invasive bladder cancer,'' Eur. Urol. Focus, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S2405456920303102, doi: 10.1016/j.euf.2020.12.008.

[21] N. Hashimoto, D. Fukushima, R. Koga, Y. Takagi, K. Ko, K. Kohno, M. Nakaguro, S. Nakamura, H. Hontani, and I. Takeuchi, ''Multi-scale domain-adversarial multiple-instance CNN for cancer subtype classification with unannotated histopathological images,'' in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2020, pp. 3852–3861, doi: 10.1109/CVPR42600.2020.00391.

[22] https://www.cancerimageingarchive.net