A Novel Attention-Based Model for Semantic Segmentation of Prostate Glands Using Histopathological Images

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 5 August 2023, accepted 21 September 2023, date of publication 2 October 2023, date of current version 10 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3321273

RESEARCH ARTICLE

A Novel Attention-Based Model for Semantic Segmentation of Prostate Glands Using Histopathological Images

MAHESH ANIL INAMDAR[1], U. RAGHAVENDRA[2], ANJAN GUDIGAR[2], SARVESH BHANDARY[1], MASSIMO SALVI[3], RAVINESH C. DEO[4], (Senior Member, IEEE), PRABAL DATTA BARUA[5,6,7], EDWARD J. CIACCIO[8], FILIPPO MOLINARI[3], (Senior Member, IEEE), AND U. RAJENDRA ACHARYA[4,9], (Senior Member, IEEE)

[1]Department of Mechatronics, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal 576104, India
[2]Department of Instrumentation and Control Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal 576104, India
[3]Biolab, PolitoBIOMed Laboratory, Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Turin, Italy
[4]School of Mathematics, Physics, and Computing, University of Southern Queensland, Springfield, QLD 4300, Australia
[5]Cogninet Brain Team, Cogninet Australia, Sydney, NSW 2010, Australia
[6]School of Business (Information Systems), Faculty of Business, Education, Law & Arts, University of Southern Queensland, Toowoomba, QLD 4350, Australia
[7]Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW 2007, Australia
[8]Department of Medicine, Columbia University, New York, NY 10032, USA
[9]International Research Organization for Advanced Science and Technology (IROAST), Kumamoto University, Kumamoto 860-0862, Japan

Corresponding author: U. Raghavendra (raghavendra.u@manipal.edu)

ABSTRACT One of the foremost causes of death in males worldwide is prostate cancer. The identification, detection and diagnosis of the same is very crucial in saving lives. In this paper, we present an efficient gland segmentation model using digital histopathology and deep learning. These methods have the potential to revolutionize medicine by identifying hidden patterns within the image. The recent improvements in data acquisition, processing and analysis of Deep Learning Models has made Artificial Intelligence driven healthcare a very lucrative area, in terms of data inference and delivering meaningful insights. This study presents an automated method for segmenting histopathological images of human prostate glands. The main focus is developing new methods for segmenting histopathological images of prostate gland using a multi-channel algorithm with an attention mechanism to detect important areas. We compare our results with a host of contemporary techniques and show that our method performs better at the segmentation task for histopathological imagery. Our method is able to delineate gland and background parts with an average Dice-coefficient of 0.9168. In this attention-based model we propose for semantic segmentation of prostate glands the potential to provide accurate segmentation versus tumor features, which has significant implications for medical screening applications.

INDEX TERMS Prostate cancer, image processing, histopathology images, digital image analysis, computational pathology, artificial intelligence.

I. INTRODUCTION

Prostate cancer is the one of the most common cancers in males [1]. In 2020, there were close to 1.5 million prostate cancer cases diagnosed worldwide. Besides being the fifth leading cause of cancer death among men, the burden increases owing to lifestyle, stress, and an aging population [2]. With a 10% fatality ratio, no serious clinical symptoms, and the geriatric patient population being particularly vulnerable, early detection of prostate cancer is important [3], [4], [5], [6].

As a result of the severity of the condition, diagnosis becomes a challenging task. The Gleason Score framework has been developed by researchers in order to identify tumors.

The associate editor coordinating the review of this manuscript and approving it for publication was Carmelo Militello.

It has become the most used method for tumor identification. An assessment of this score is based on the microscopic examination of the cancer cell. It has been observed that malignant tumors have a propensity to spread, unlike benign tumors which are generally dormant. The pathologist score is done on a scale of 1-5 by assessing the shape and cell arrangement. Cancer cells largely correlated to the characteristics of healthy cells are considered innocuous and hence receive a low score [7].

The architectural configurations of prostatic carcinoma serve as the sole basis for the Gleason grading system, which is a key factor in treatment decision-making. Grades are ranged between G1 to G5, with G1 denoting tissue with the most severe degree of resemblance to healthy tissue and a good prognosis, on the other hand G5 denoting tissue with the lowest severe degree of differentiation and the poorest prediction [8]. Artificial intelligence has the potential to enhance Gleason grading accuracy. Numerous automated Gleason grading techniques were suggested, which improved consistency.

Although the Gleason Score Technique is widely applicable, it has certain limitations [9], [10], [11], [12], [13], [14]. Due to its non-exact scoring and the need for manual input, it is time-consuming. To overcome the shortcomings of this method, deep learning models have been used to analyze, identify, and detect cancerous tissue within prostate tissue samples. To add to new knowledge in this area, the purpose of this study is to develop a novel attention-based model for semantic segmentation of prostate glands utilizing histopathological images in order to fill research gaps and identify related challenges in existing methods of detecting cancerous tissues. The study has therefore implemented a UNet-based model with a capability to segment the contours of prostate glands using a U-Net architecture with a ResNet50 backbone, as described on the methodology section.

II. BACKGROUND: HISTOPATHOLOGY AS AN IMAGING MODALITY

Histopathology is a vital domain of biology involving study and research of the microscopic structure of cells and tissues within living things. It is a valuable tool for assessing the health of organs and for detecting the presence of disease [15], [16]. In the case of the prostate, its biological structure typically consists of glands and stroma, with the stroma serving as supportive tissue around the gland [17]. Each gland unit has a lumen (cavity within a gland) that is surrounded by rows of epithelial layers. Typically, tissues are dehydrated and embedded in molten paraffin wax after being collected (often as biopsies) and preserved. The resulting block is mounted on a microtome and separated into thin slices [18]. Following their attachment to microscope slides, the tissue slices are rehydrated and made ready for staining once the wax has been removed using a solvent [19]. One of the most prominently used staining techniques is Hematoxylin and Eosin (H&E). This procedure involves applying hematoxylin combined with a metallic salt, or mordant, followed by bluing in moderately alkaline water. This procedure is frequently followed by a rinse in a weak acid solution to eliminate extra staining (differentiation) [20], [21]. Eosin (most frequently eosin Y) is used to counterstain the tissue after hematoxylin has been applied [18], [22], [23]. Conventionally, the results are quite consistent, with cell nuclei stained blue and the cytoplasm and extracellular matrix colored pink [18]. A detailed procedure for staining is mentioned here [24].

III. RELATED RESEARCH WORK

In recent years, segmentation techniques using deep learning have provided better results as compared with many traditional implementations, contributing to further progress in the research and development of this topic. A plethora of related work has recently been published, particularly for the improvement of accuracy in developing techniques for gland segmentation from histopathological imagery [25]. Here in this study, we primarily have focused on digital histopathology images, but Magnetic Resonance Imaging (MRI) also has been shown to be an effective modality for the segmentation of prostate glands [26]. Here we also provide a review of work performed using histopathological images for prostate glands/ cancer cells segmentation and in the general overall domain (breast, lung). MRI is one of the most powerful imaging modality that can be used to capture detailed anatomical information, such as structure of the prostate gland. It can also be used to detect cancer cells in the prostate, which makes it a valuable tool for segmenting prostate glands and cancer cells. In this study, we used histopathology images to segment prostate glands and cancer cells, but the technique can be applied to other domains such as breast and lung.

Many of the studies have used CNN/U-Net as the basic segmentation architecture, due to its profound results in object segmentation and region identification in medical imaging. Here we describe the following studies pertaining to this architecture. Li et al. proposed a Multi-scale U-Net (with tile size 400,200 and 100) for the segmentation of histopathological tiles into benign, stroma, G3 & G4 classes. They compared the U-Net and Multi-Scale U-Nets Pixel-wise CNN as a base model. They showed the superiority of having the model trained over variable tile sizes due to their ability to acquire global information within the region [27]. The same authors experimented with a semi-supervised approach with Expectation Maximization to generate pixel-wise probability maps using variable-sized tiles [28]. Ing et al. experimented with variants of Convolutional Nets for semantic segmentation of high- and low-grade tumors [29]. Kalapahar et al. showed the feasibility and superiority of Residual nets as the backbone for the U-Nets, owing to their smooth gradients and preventive actions for vanishing gradients [30].

The study of Singh et al. have used local and spatial information to combine the findings of pixel-level classification

into object-level segmentation for recognizing these regions using an array of pixel and object-level classifiers [31]. Ali et al. performed multistage segmentation using wavelet packet features and entropy values [32]. Ren et al. performed semantic segmentation using an encoder-decoder network. The sliding window approach was utilized for larger images [33].

In their studies, Xu et al. focused on channel-based segmentation with CNN (VGG Model). The three channels were used to discriminate or segment background from foreground pixels, to detect gland boundaries, to detect, discriminate, or segment background from foreground pixels, and to detect gland boundaries and individual glands. The network has five structures for pooling with $512 \times 512 \times 3$ as the input shape. This model enhanced performance but suffered in segmenting glands where the lumen is not visible, and it could not detect glands in severe cases [34]. Silva-Rodríguez et al. proposed a weakly supervised technique (WeGleNet) for correctly segmenting cancerous regions in the gland cells with global spatial information level Gleason score from the clinical records [35].

Here we present some of the most prominent works in the domain that can be used in prostate cancer diagnosis. They are mainly focused on works that have published new architectures in the deep learning domain. For example, Oda et al. presented Boundary-Enhanced Segmentation (BESNet) of Cells, which comprises UUNet-type structures but with two decoders (Boundary, Region). This novel net trains on the boundaries of the cell, and adaptively pays more attention to difficult parts [36]. Naylor et al. devise an approach to solving the problem of segmentation of overlapped nuclei, by transforming the RGB space to distance maps (from the pixels to the nearest background). A Neural net is then trained with regression loss to minimize the predicted and actual values [37].

The study of Graham et al. provided a unique approach to deal with the heterogeneity in the appearance of nuclei by building a multi-scale stain-aware network (SAM-Net) with a unique loss function. The net exploits the residual connections and transfers the deep feature across the layers. The loss function is pixel-wise cross entropy weighted by a pre-defined weight map [38]. Xu et al., proposed a Semi-supervised multiple-instance learning technique CAMEL. This process further refines the labels using Max-Max and Max-Min criteria [39]. In their study, Graham et al., attempted to tackle the challenge of indistinguishable histological structures due to variability in appearance by building a net with Minimizing Information Loss Dilation (MILD Net). The net re-introduces the original image multiple times to counter the loss due to max pooling and maintains the resolution through spatial pyramid pooling with varying dilations [40]. Mahmood et al., shows the Conditional Generative Adversarial Net (cGAN) enforces consistency in higher order spatial space which is missing in traditional CNNs. This approach performs better for organ segmentation in peculiar cases of overlapping nuclei [41].

The study of Chan et al., proposed a method for identifying a region of interest in WSI. Song et al., proposed Multi-layer boosting Sparse Convolution (ML-BSC) for bettering the nuclei segmentation, by using a boosting technique with discriminative probabilistic Binary Decision Trees as weak learners [42]. Typically, CNNs consider instances of images being rotated as different images and hence attempts to learn weights specifically to each orientation. Graham et al., incorporated G-Convolution to make them equivariant to rotations, this is advantageous to segment regions where there is more variation in different orientations. But this study increased the computational loads, as it required more kernels [43]. To produce fine and crisp segmentation maps, a framework must be able to ingest additional information. HookNet proposed by Rijthoven et al., attempts to combine (hook) context and high-resolution information into the network across many branches. The authors claim that the typical confusion between class instances is reduce due to the contextual information [44].

Atrous spatial pyramid pooling has been widely assessed in many investigations, as it enables the capture of a detailed view without loss of resolution. Wan et al., combined the UNet model with the contour concave point detection algorithm for accurate nuclei segmentation and dealt with cases like nucleus occlusions [45]. Qu et al. entered into the semi/weakly and self-supervised learning domain to detect the nuclei region and to refine the segmented maps using conditional random field loss [46]. Lal et al., proposed a Nuclei Segmentation Net (NucleiSeg) having the residual connection and attention module to address the variable size and overlapping nuclei regions. The attention module (part of decoder) specializes in object detection and reduces false positives [47]. Furthermore, the novel approach of NucleiSeg yielded superior results compared to the conventional approaches and demonstrated that attention-based models can be advantageous for biomedical segmentation tasks. It is therefore evident that by incorporating semi-supervised learning, weakly supervised learning, and self-supervised learning into the model, the model can learn from both labeled and unlabeled data, allowing it to better detect the nuclei regions. Additionally, the residual connection and attention module helps to address the issue of variable size and overlapping nuclei regions, as well as reduce false positives.

In Sections IV (and subsections), we describe the database description and methodology of the entire process, including the preprocessing, training, and morphological post-processing operations. As a final step, in Section V and VI we demonstrate Results and the Ablation Study conducted. Lastly in Section VII and VIII we discuss the findings; present the Conclusion and Future Works.

IV. MATERIALS AND METHODS

A. DATA

The datasets used in this work were acquired by the Division of Pathology of the A.O.U. Città della Salute e della Scienza in Italy and were previously used by Salvi et al. [48].

Following a number of quality control checks, we created two subsets of the designated database: a TRAINING subset (with a total of 1000 histopathological images) and a TESTING subset (with a total of 500 histopathological images) in order to develop and fully evaluate the proposed attention-based model for semantic segmentation of prostate glands.

B. METHODOLOGY - DATA PRE-PROCESSING

The first step in preprocessing for analysis of histopathological images was the normalization of the respective image. Here we examine the theoretical underpinnings of the color deconvolution approach and determined how hematoxylin, eosin stains, and the stroma mask could be obtained. After this, the stroma mask was fused with the predicted output of the model with region and boundary channels.

The goal of color deconvolution was to distinguish between the effects of the various dyes used during the staining process. It is noteworthy that the hematoxylin binds itself to the cell nuclei with a deep blue-purple color, while eosin stains stroma with a pinkish colour. As a result, there were two stains total since the histology slides stained with H&E. In accordance with the Beer-Lambert law, the amount of stain absorbed tends to determine the colour intensity that a particular cell component is likely to exhibit. Therefore, the logarithm of the incident ratio (I_0) to transmitted (I) light intensity is used to define a medium's absorbance, and it may alternatively be represented as the product of the light beam's path length l, the molar coefficient (extinction) ε, and the concentration c of the absorbing quantum.

$$\boldsymbol{A = -log_{10}(\frac{I}{I_0}) = \varepsilon.c.l} \tag{1}$$

The input (converted into a vector) is then mapped into optical density (OD) space using Equation 1. It is important to note that the OD space can be utilized to separate the contributions of each stain because the corresponding OD value of each pixel per channel is correspondent with the concentration of absorbing quantum. The OD of an image can be expressed as the product of the stain appearance color matrix (W) and stain density map product (H). This allows for the separation of an input (RGB) image into two distinct channels, each correlated to the concentration of the specific stain.

Once the stains are separated and estimated, they are processed via range normalization. We have followed the paradigm for stain normalization developed by Salvi et al. [52]. This procedure involves normalizing the image as, during the staining process, the histological specimen goes through color variations after the interaction of dye and chemicals [53], [54]. Color normalization is the most prominent step to standardize the stain appearance of the image (source). This operation reduces the stain variability and improves the robustness of computer-aided diagnostics and algorithms for image quantification [49], [50], [51], [52].

To obtain the stroma mask, we first detected the white regions of the images using fixed thresholding technique. The hematoxylin and eosin stains were separated using the technique mentioned in [58]. To localize nuclei, we solved an energy function Equation (2) (object-based thresholding) to obtain the probability values of pixel values equal or lower T(p_0) and higher than T(p_1) where in p_0, p_1 are the background and nuclei (pixel) distribution and T being the threshold point. This technique (Multi adaptive nuclei analysis) [59] helps us find the optimal threshold point for demarcation. On similar lines, for stroma regions, two clusters are identified in the Eosin Stain using k-mean clustering, and stroma pixels are identified as the cluster with highest intensity (mean).

$$\mathbf{E(T)} = \boldsymbol{p_0^2 * var_0 * \log(var_0) + p_1^2 * var_1 * \log(var_1)} \tag{2}$$

where var represents the variance of the pixel distributions.

To determine the Dice coefficient, which is a statistic used to gauge the similarity of any two data samples, this image was compared with the appropriate ground truth image. The average and best values of all Dice coefficients were determined once all the pictures in a given dataset had passed the full procedure. Figure 3 shows the entire block diagram of post-processing operations.

C. METHODOLOGY - NETWORK ARCHITECTURE

The primary purpose of this study was to develop and implement a UNet-based model to segment the contours of prostate glands using a U-Net architecture with a ResNet50 backbone. It should be noted that the study of Chen et al. [53] has adopted a CNN trained purely on glandular regions in order to increase the accuracy in the detection of cancer by approximately 26.9%. It should also be noted that the network follows an encoder-decoder architecture. The encoder network, based on the ResNet50, down-samples the image features to a lower-dimensional representation.

Immediately following this is a decoding algorithm that is an exact copy of the encoder, except that it utilizes a transposed convolution where the objective is to transfer all the distinguishable features learned from the encoder onto the higher dimensional space. This process allows for the extraction of high-level features that are then used by the decoder network to generate the final output. The model output assigns a probability value to each pixel based on how well it fits into a specific class. Consequently, this model takes in 480 × 480 dimensional RGB images and produces a probability map for three classes. There are three classes of targets (pixels), namely glands, boundaries of gland, and background. Consequently, we have adopted a method that follows previous studies showing that such methods outperform traditional approaches [58].

In this study, is important to note that the encoder network is pre-trained with ImageNet weights and is not trained [55]. During training, the weights of the lower decoder layer are updated in order to fine-tune the network to the current objective. This strategy enables the transferring of knowledge about identifying basic features of images learned from the

ImageNet weights. This approach is helpful to deal with smaller datasets and optimize the time taken for training [56]. Additionally, data augmentation techniques were applied to make the model more robust to variations in the images and to prevent overfitting.

Finally, the network was trained for 100 epochs, with early stoppage being triggered after 10 epochs of no improvement. For each image, 9 patches were used during training, with a learning rate of 10^{-3} and a scheduler. The categorical cross-entropy and Dice-Coefficient loss function and the Adam optimizers were used. The proposed model was trained on Google Colab Pro with NVIDIA-SMI 460.32.03 Tesla T4 GPU computer system.

D. METHODOLOGY - INCORPORATING THE ATTENTION MECHANISM

In the field of deep learning, the attention mechanism has played a significant role. This method allows a selective focus on the necessary important objects (and relevant patterns) to simplify calculation and improve accuracy. While doing so, it pays relatively less attention to others (in this example, concentrating on the gland region and ignoring the stromal areas) [57].

An attention mechanism has been incorporated into the layers to significantly improve the performance. A relationship between an encoder and a decoder is what is meant by an attention mechanism. Using this framework, it is feasible to selectively focus on the necessary or more significant portions of the supplied image. Additionally, it makes it easier for the model to handle large input picture sizes.

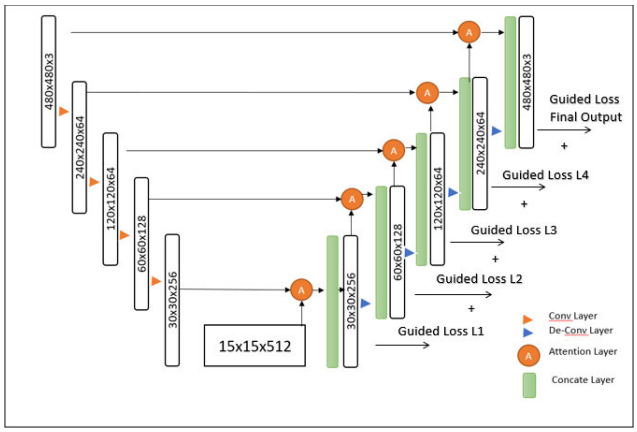

FIGURE 1. The proposed attention-based U-Net model for semantic segmentation of prostate glands using histopathological images.

In addition to reducing false positives, attention gates also promote the updating of model parameters for relevant spatial regions. The attention gate is built to accept two inputs – one containing the spatial and contextual information, the second being the gating signal from the decoder layer underneath it as shown in Figure 1. The output from the attention gate is concatenated and further processed. A pictorial representation of the attention gate is shown in Figure 2.

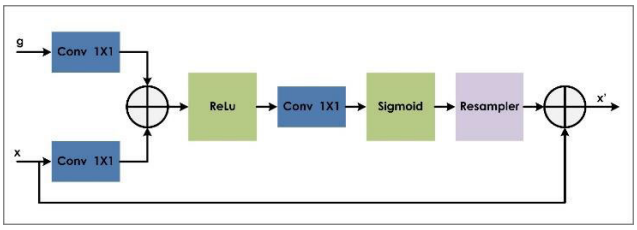

FIGURE 2. Attention gate used to focus on important areas.

E. METHODOLOGY - MODEL DEVELOPMENT

Prior to model development, in total, we applied a 10-fold data augmentation to each of the 1000 training images, resulting in a dataset of 10,000 images. Moreover, we divided the image into nine equal-sized patches, giving the model a total of 90,000 images. The segmented output result is a hybrid instance level process which fuses the region, edge and stroma mask to obtain the final (before post processing) image. Figure 3 shows the entire process of the fusion of images.

Next, we trained the model using the Dice Loss and Cross Entropy Loss metrics across each encoder layer, which is referred to as the guided loss, or GL. As there are four decode layers, the summation was weighted, with the highest 50% to the last layer and equal weight assignment to the rest of 12.5%. Ultimately, the output was acquired only from the last layer following the mathematical expression for the loss function:

$$LExp = wDiceLDice + wCrossLCross \tag{3}$$

where, wDice and wCross are the respective weights of the Dice loss (LDice) and the weighted cross-entropy (LCross).

The images were divided into 9 patches of dimension 480×480 each as a requirement since training on a large image (of dimension 1500×1500) was computationally expensive. By computing more images with a smaller size, the model is able to train faster and more efficiently. Figure 4 displays flowchart of the modelling process.

The modification taken over each process is shown in Figure 5.

F. METHODOLOGY - POST-PROCESSING OPERATIONS

Although CNN's semantic segmentation can accurately identify most glandular areas, it may not recognize all of them. As a result, it is possible for the CNN output to contain incorrect or incomplete gland outlines. The proposed strategy uses a hybrid segmentation approach based on identifying all the stromal areas, which in practice means locating anything that is not a gland, as direct segmentation of the prostate glands is a difficult task. During the stroma segmentation, the background of glands can be identified, allowing the glands to be recognized more precisely. Stroma recognition is easier than gland detection due to the great variety of glandular

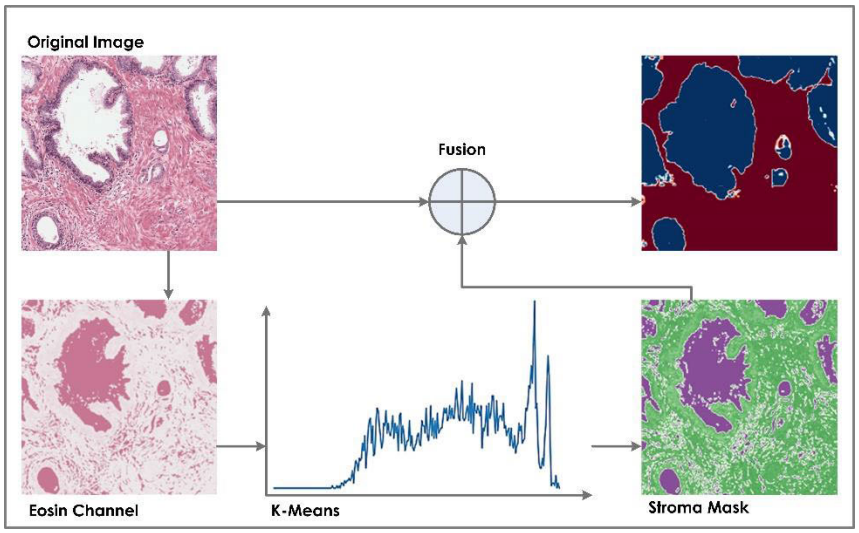

FIGURE 3. Process to obtain the fused images for constructing and evaluating the proposed attention-based U-Net model for semantic segmentation of prostate glands.

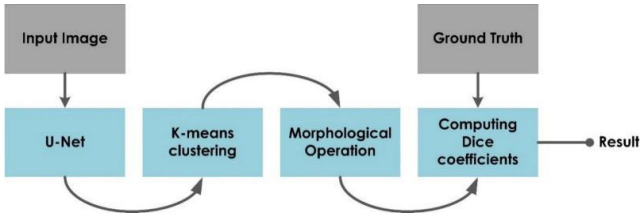

FIGURE 4. Flowchart of the methodology to showcase the schematic representation of the algorithm.

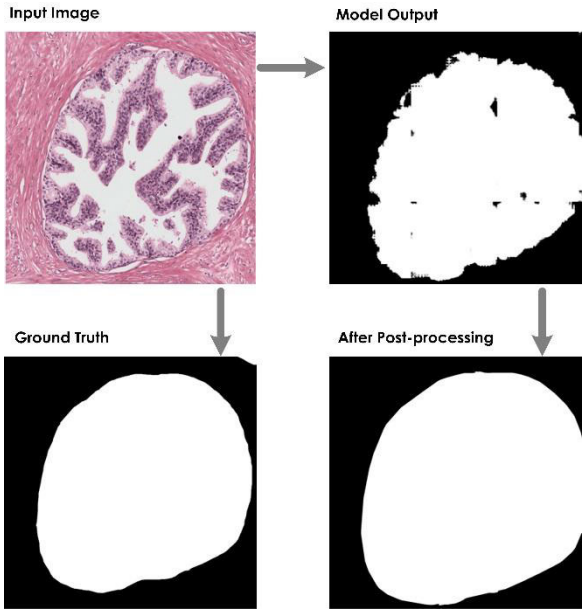

FIGURE 5. The results showing the image after the processing operations.

patterns, especially in diseased conditions. To identify all stromal regions, a BGR image combining structure-based detection and CNN output is produced. To identify every stromal area, the suggested technique applies a SoftMax-driven active contour model to the fusion (final) image.

One popular robust and adaptable technique called the Chan-Vese model for active contours [60] has the ability to segment a wide variety of pictures. These include an image that could be very challenging to segment especially when using ''classical'' segmentation techniques (e.g., thresholding or gradient-based techniques). This approach is based on the Mumford-Shah functional [61] for segmentation built on energy minimization issues.

To describe this model, let Ω be an open bounded set of R^2, with dΩ its boundary. Let ko: $\Omega \rightarrow$R be a given image and C(s) is a piecewise parameterized representation of C_1 [0, 1] a curve. The region inside C is denoted by ω, and the region outside C as $\Omega \setminus \omega$. Moreover, C_1 denotes the average pixel intensity inside C, and C_2 will denote the average intensity outside C (i.e., $C_1 = C_1(C)$, $C_2 = C_2(C)$). The region for the objective function is shown in Figure 6.

The objective of Chan-Vese algorithm is to minimize the energy functional F(C_1, C_2, C) defined by:

$$\begin{aligned} \boldsymbol{F}(C_1, C_2, \boldsymbol{C}) = \boldsymbol{\mu}.\boldsymbol{Length}(\boldsymbol{C}) + \boldsymbol{v}.\boldsymbol{Area}(\boldsymbol{inside}(\boldsymbol{C})) \\ + \lambda_1 \oint |u_0(x,y) - C_1|^2 dxdy \\ + \lambda_2 \oint |u_0(x,y) - C_2|^2 dxdy \end{aligned} \tag{4}$$

Here the first and second integral are for inside and outside C region where in μ, λ_1, $\lambda_2 > 0$ and are hyperparameters.

Hence, we are in search of the best values of C_1, C_2, and C to minimize the above-mentioned function. In the level set method, C $\subseteq \Omega$ is represented by the zero-level set of some Lipschitz function α: $\Omega \rightarrow$ R

Which can also be written as

$$C = \boldsymbol{d\omega} = \{(\boldsymbol{x},\boldsymbol{y})\Omega : \alpha(\boldsymbol{x},\boldsymbol{y}) = 0\} \tag{5}$$

FIGURE 6. The region for the objective function.

$$\text{Inside (C)} = \boldsymbol{\omega} = \{(\boldsymbol{x}, \boldsymbol{y})\ \boldsymbol{\Omega}\ \boldsymbol{\mathcal{V}}\ \boldsymbol{\alpha}\ (\boldsymbol{x}, \boldsymbol{y}) > \mathbf{0}\}\} \tag{6}$$

$$\text{Outside (C)} = \boldsymbol{\Omega} \backslash \boldsymbol{\omega} = \{(\boldsymbol{x}, \boldsymbol{y})\ \boldsymbol{\Omega}\ \boldsymbol{\mathcal{V}}\ \boldsymbol{\alpha}\ (\boldsymbol{x}, \boldsymbol{y}) < \mathbf{0}\} \tag{7}$$

Hereby, now framing the question which involves, evolving α(x,y), when the evolved contour C in each time t is the zero-level set α(x,y),

Hence, the final function in terms of α(x,y) will be

$$\begin{aligned} F\left(C_1, C_2, C\right) = \mu . \oint \delta(\alpha\ (\mathrm{x}, \mathrm{y})\ |\partial \alpha\ (\mathrm{x}, \mathrm{y})|\ dxdy \\ + v \oint H(\alpha\ (\mathrm{x}, \mathrm{y}))dxdy \\ + \lambda 1 \oint |u_0\ (x, y) - C_1|^2 H\ (\alpha\ (\mathrm{x}, \mathrm{y}))\ dxdy \\ + \lambda 2 \oint |u_0\ (x, y) - C_2|^2\ H\ (1 - \alpha\ (\mathrm{x}, \mathrm{y}))\ dxdy \end{aligned} \tag{8}$$

where, H is a pointwise function. Figure 7 shows the obtained output using the Chan-Vese Algorithm.

As a result of the CV algorithm, the output image was subjected to morphological operations including erosion, dilation, contour identification, convex hull [62] and filling. It is important to note that we have used bounding contours to minimize dilation and improve results. Our final step was to limit the presence of objects with a minimum area to remove unnecessary regions.

V. RESULTS

In order to test the proposed attention-based U-Net model for semantic segmentation of prostate glands and interpret our findings, we have used the Dice coefficient, which measures how similar any two binary masks are by comparing them directly. In the Dice analysis approach, a coefficient ranges between 0 and 1, with a value near to 1 suggesting a greater degree of resemblance. To assess model performance, we compute the average Dice coefficient of the training or testing dataset. Figure 8 shows the Dice scores and Intersection over Union (IoU) values of 5 images randomly sampled from our test dataset, along with the respective images:

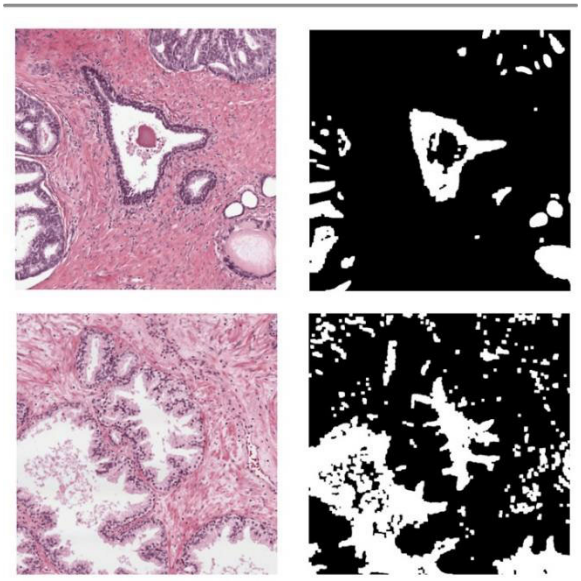

FIGURE 7. Original image and output from Chan-Vese algorithm.

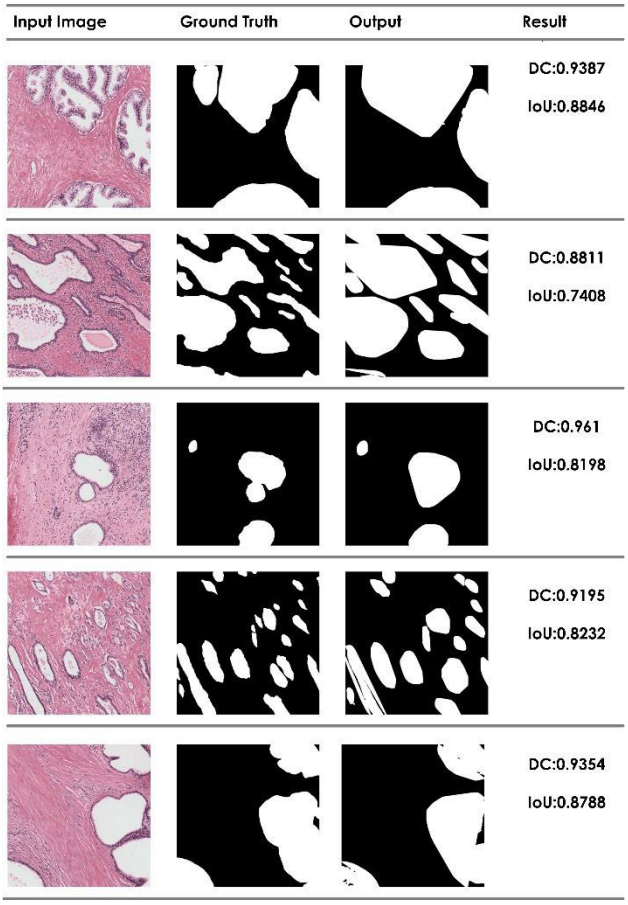

FIGURE 8. Samples outputs from the tested dataset utilizing the attention-based U-Net model for semantic segmentation of prostate glands.

The proposed model appears to have performed much better in the task. Table 1 compares and summarizes the results.

TABLE 1. A summary of the results obtained from previous investigations.

Authors	Dataset	Results
Lt j et al.; [27]	Pvt Dataset	10-fold cross validation 65.8% JACC across stroma, Gleason 3, Gleason 4 and benign glands; and 75.5% across (stroma, benign glands, prostate cancer
Lt j et al.; [28]	Pvt Dataset	Semi-supervised approach 49.5% on an independent set
Ing et al. [29]	Pvt Dataset	FCN-8s: mIOU of 0.759 and an ACC of 0.87, two SegNet variants, and multi-scale U-Net: mIOU of 0.738 and accuracy of 0.885.
Kalapahar et al. [30]	Pvt Dataset	Pixel-level Cohen's quadratic Kappa of 0.52 using ResNets
Singh et al. [31]	Pvt Dataset	DC 0.5203 ± 0.2517 (Train) DC 0.4931 ± 0.2557 (Test)
Ali et al. [32]	Pvt Dataset	Bin Acc: 90%
Ren et al. [33]	Pvt Dataset	DC 0.8394 ± 0.1382 (Train) DC 0.8308 ± 0.1495 (Test)
Xu et al. [34]	Pvt Dataset	DC 0.8106 ± 0.1257 (Train) DC 0.8079 ± 0.1264 (Test)
Rodriguez et al. [35]	Pvt Dataset	pixel-level κappa of 0.61 and a macro-averaged F1-score of 0.58,
Salvi et al. [48]	Pvt Dataset	DC 0.9073 ± 0.0989 (Train) DC 0.9016 ± 0.1087 (Test)
Our proposed method	Pvt Dataset	**Best** DC 0.9978 (Gland) (Train) DC 0.9787 (Tumor) (Train) DC 0.9682 (Gland) (Test) DC 0.9611 (Tumor) (Test) **Average** DC 0.9328 (Gland) (Train) DC 0.9217 (Tumor) (Train) DC 0.9111 (Gland) (Test) DC 0.9043 (Tumor) (Test)

JACC: Jaccard Index, ACC: Accuracy, mIOU: mean Intersection over Union

In this study, K-fold cross-validation (with 10 splits) was conducted to validate our results. Initially, the dataset was divided into 10 equal parts, nine of which were used for training, and the remaining part for testing. In this manner, all combinations of partitioned data parts were incorporated into training and testing. This process yielded an average Dice coefficient of approximately 0.92±0.0011 (for the Glands) and 0.91±0.0041 (for the Tumors). Thus, it can be said that post-processing techniques, especially morphological operations, improve Dice scores.

When evaluating the model with performance metrics, the output masks were assessed based on Precision, recall, IoU and Dice scores. The precision is meant to assess false positives; the recall is meant to assess the model's accuracy in selecting the outputs in designated pixels; and the Dice score measures the intersection of two masks. Table 2 shows the magnitude of the metrics during training the model, and Figure 9 shows how recall and precision vary over epochs. It is evident from these results that our model is generalizable in the sense that it can adapt to new data images and still produce credible results.

VI. ABLATION STUDY

In this research, in addition to presenting the efficacy of the attention-based U-Net model for semantic segmentation of prostate glands, we have performed ablation studies to

TABLE 2. The results with more metrics of the proposed attention-based U-Net model for semantic segmentation of prostate glands.

METRIC TYPE	METRIC VALUE
SENSITIVITY	0.8818
SPECIFICITY	0.9987
PRECISION	0.9822

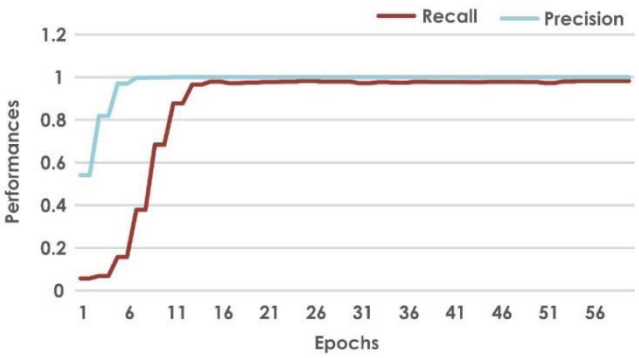

FIGURE 9. The performance metric represented as the Recall vs. Epochs, and Precision vs. Epochs.

provide an in-depth look at each proposed module's contribution within the proposed overall architecture, to generate the results. With the State-Of-The-Art (SOTA) models, we conducted ablation studies on the proposed module and corroborated the claims. It is noteworthy that the skip links in U-Net can serve as communication channels between the corresponding encoding and the decoding layers.

In reality, U-Net serves as the framework for the proposed model, but it does lack distinct network highlights, resulting in a segmented image that is significantly less realistic. Additionally, the fundamental idea of residual blocks is related to the addition of an ''identity connection'' that can cross several layers. With this link, the current layer's result is combined with the input from the preceding layer, and the input for the next layer is the sum of these outputs.

During our ablation studies, the residual connections have also been added to the UNet model's design to address the degradation and saturation problems, as well as to extract additional features at each layer. In principle, the Res-UNet model provides better segmentation quality than the UNet model when residual connections are combined with it. By restricting activations to relevant regions, the attention gate improves spatial precision at each layer. As a result, segmentation correctness was improved in terms of the Dice score and the mean value of the Intersection Over Union (IoU) with attention gates over the backbone.

In Figure 10, we show the comparison between learnings of mid-layers of the net with- and without the Attention Mechanism in order to demonstrate the importance of this module on the overall efficacy of the proposed U-Net model. Evidently, the learning curve i.e., the curvatures with the proposed attention mechanism appears to be similar to the ground truth value. Nevertheless, the guided decoding stage is the network's primary architectural feature, which enhances the learning process by incorporating it at the individual layers. As a result, the segmentation outcomes have been further enhanced by the better development of features and their use in the overall loss function.

Here we wished to connect the interpretability of the model (through Figure 10) with the Explainable Artificial Intelligence (XAI) [65] which recently has been gaining tremendous popularity in the AI domain. While the terminology is new, the concept of making the system's working components interpretable and explainable has been already established. This includes the introduction of ethics, security and fairness to the system; and removal of 'favoring' and 'ignorance' related to a specific class. In this study, we attempted to understand the working of the model specifically regarding the proposed attention module, hence exploring it through the 'Model specific approach'. In the Figure 10, as explained in that section, we show the outputs from the decoder part of U-Net. It could be seen that the model makes an informed decision (of predicting a pixel values) by cumulative loss gathered by its previous loss (except the first). This could also be corroborated by the gradual decrease in the Guided Loss across the decoder part.

Table 3 shows the quantitative contrasts demonstrated between the objective model on predominantly employed architectures. It is evident that the highest dice score and IOU were generated by the proposed ResNet U-Net with Attention, followed by the InceptionResNetV2 U-Net and the ResNet U-Net model. Therefore, the presented results demonstrate that the proposed model has achieved a superior segmentation accuracy relative to several SOTA models.

As an additional task, we have trained our objective model using two main types of loss functions, namely the Dice loss and the weighted cross-entropy loss function. In order to find the best combination among each, we experimented with all these combinations and exploited the best technique. Table 4 shows the results. According to our results, Cross-Entropy with Dice Loss is the best loss function for this application.

VII. DISCUSSION

The domain of prostate cancer detection has been of particular interest to many research efforts. This interest stems from the fact that it is challenging both in terms of pattern recognition and classification, due to the overlap in the featuring and similarity in identifiable patterns [63], [64]. Several open segmentation and classification tasks in leading competitions have demonstrated remarkable results using deep learning techniques. These are based on the Gleason score, the most trusted method, but they still suffer from internal variability. Our study presents an improved segmentation algorithm for

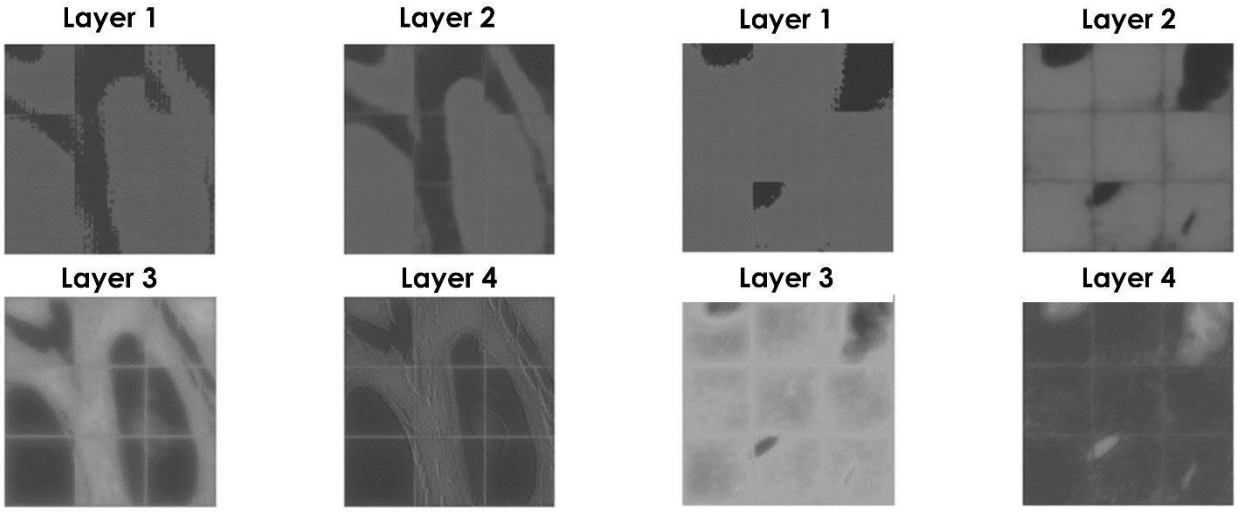

FIGURE 10. Comparative analysis of intermediate layers with, and without the attention mechanism with left begin with attention layer and the right without the attention layer.

TABLE 3. Comparison of the proposed attention-based U-Net model for semantic segmentation of prostate glands relative to the other model architectures. The best model is boldfaced.

METHODS	DICE SCORE	IOU
INCEPTIONRESNETV 2 U-NET	0.8921 (Mean) 0.9413 (Best)	0.875 (Mean) 0.9231 (Best)
RESNET U-NET	0.8892 (Mean) 0.9112 (Best)	0.8325 (Mean) 0.8929 (Best)
RESNET U-NET WITH ATTENTION (OBJECTIVE MODEL)	0.9278 (Mean) 0.9788 (Best)	0.91 (Mean) 0.9574 (Best)

TABLE 4. Comparison of the proposed attention-based U-Net model using a range of loss functions. The best model is boldfaced.

LOSS VARIANTS	DICE SCORE	IOU
DICE LOSS	0.9064 (Mean)	0.8853(Mean)
WEIGHTED CROSS ENTROPY	0.8973 (Mean)	0.8762 (Mean)
DICE LOSS WITH WEIGHTED CROSS ENTROPY (OBJECTIVE MODEL)	0.9278 (Mean)	0.910079 (Mean)

healthy versus cancerous glands. With the addition of the Attention mechanism to the current method (U-Net), we have been able to streamline the focus areas and better delineate the glands.

Although our method focuses solely on CNN-based segmentation, the use of appropriate post-processing techniques and attention mechanisms contributed significantly to the improved results. Moreover, we believe that enlarging the dataset to include more variation in terms of region and conditions enhance the model robustness and improve its performance. Additionally, we are confident that the proposed model, when combined with hybrid techniques, has the potential to yield even better results in future implementations.

To instill further confidence in the proposed model and to prove its superiority we conducted the ablation study summarized in Figure 10, which clearly shows the effect of the attention mechanism. The two subplots depict the mid layer outputs of the model. The pictures have been produced by up-sampling the outputs from the respective layers. These mid-layers are the ones where the guided loss is being added and evaluated. It is evident that the outputs delineation (without the post-processing) is improved in the model with the attention module. This is mainly due to a special focus on the specific regions given by the attention module. Experimentation with different combinations of loss function is conducted

and shown in Table 3, which shows that the employed combination works best especially in such scenarios. Overall, the presented study demonstrates an approach for segmentation of gland and cancerous cell regions in histopathological images with a novel attention mechanism and a combination of loss function for achieving better results.

In terms of output, the attention module with dual combination of loss function works much better, as shown in Table 3 and 4. The dual loss combination, i.e. Dice coefficient and cross entropy, focuses on each pixel level correction with accurate border delineation. Notwithstanding, in terms of computational complexity the model is bulky, and consists of many learnable parameters, enabling the model to extract and learn the hidden patterns, but at a cost of being resource intensive. Knowledge distillation [66] could help in reducing the resources and to trim the model with the same output results.

VIII. CONCLUSIONS, LIMITATIONS AND FUTURE WORKS

Based on histopathology imagery, a novel Attention Res-UNet model has been developed for the segmentation of prostate glands. The results demonstrated that the model learns from Guided loss at each decode layer, resulting in refined feature maps. In the future, improved combinations of loss functions may be used in conjunction with weight optimization procedures. With respect to the limitations of the present study, we note that a larger network was used, consequently, requiring a substantial computation time. Therefore, future studies could investigate and optimize the network size versus performance efficiency of the proposed Attention Res-UNet model based on ablation-based studies. Furthermore, the current training procedure could be further improved by considering more sophisticated optimization strategies, such as adaptive learning rate and momentum, which have been successfully proven in the field of deep learning. Additionally, a deeper investigation of the model performance improvement under varying network sizes should be conducted in future studies.

Inculcating XAI techniques into our study was done to debunk the 'black-box' notation prominently given to DL models due to their predictions made without valid decision-making logic. The presented combination of Guided Loss function enhances the learnability of model through the layers, and the attention module enables it to provide special focus to specific regions for better precision and accuracy. Our paradigm has the potential to greatly improve healthcare services for quick and better identification of major diseases, while providing interpretable insights. As a future direction to the work, we think that XAI techniques will greatly add to the value of this domain, particularly in terms of visualization and the working of the model, and to understand the back-end logics prior to making decisions, such as discrimination within class region.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENT

The study was conducted in Advanced Healthcare Laboratory, Department of Instrumentation and Control Engineering. The authors would like to thank the Manipal Academy of Higher Education (MAHE) for providing the required facility to carry out this research.

REFERENCES

[1] *Key Statistics for Prostate Cancer*. Accessed: Aug. 2, 2023. [Online]. Available: https://www.cancer.org/cancer/prostate-cancer

[2] L. Wang, B. Lu, M. He, Y. Wang, Z. Wang, and L. Du, "Prostate cancer incidence and mortality: Global status and temporal trends in 89 countries from 2000 to 2019," *Frontiers Public Health*, vol. 10, Feb. 2022, Art. no. 811044, doi: 10.3389/fpubh.2022.811044.

[3] V. Narayan, S. Jiang, and C. A. Warlick, "Early stage cancer in older adults: Prostate—Avoiding overtreatment and undertreatment," *Cancer J.*, vol. 23, no. 4, pp. 238–241, Jul. 2017.

[4] V. J. Gnanapragasam, O. Bratt, K. Muir, L. S. Lee, H. H. Huang, P. Stattin, and A. Lophatananon, "The Cambridge prognostic groups for improved prediction of disease mortality at diagnosis in primary non-metastatic prostate cancer: A validation study," *BMC Med.*, vol. 16, no. 1, pp. 1–10, Dec. 2018.

[5] T. J. Daskivich, K.-H. Fan, T. Koyama, P. C. Albertsen, M. Goodman, A. S. Hamilton, R. M. Hoffman, J. L. Stanford, A. M. Stroup, M. S. Litwin, and D. F. Penson, "Prediction of long-term other-cause mortality in men with early-stage prostate cancer: Results from the prostate cancer outcomes study," *Urology*, vol. 85, no. 1, pp. 92–100, Jan. 2015.

[6] A. M. Hoogland, C. F. Kweldam, and G. J. V. Leenders, "Prognostic histopathological and molecular markers on prostate cancer needle-biopsies: A review," *BioMed Res. Int.*, vol. 2014, pp. 1–12, Jan. 2014.

[7] Accessed: Aug. 2, 2023. [Online]. Available: https://www.cancer.net/cancer-types/prostate-cancer/stages-and-grades

[8] W. Bulten, H. Pinckaers, H. van Boven, R. Vink, T. de Bel, B. van Ginneken, J. van der Laak, C. Hulsbergen-van de Kaa, and G. Litjens, "Automated Gleason grading of prostate biopsies using deep learning," 2019, *arXiv:1907.07980*.

[9] T. A. Ozkan, A. T. Eruyar, O. O. Cebeci, O. Memik, L. Ozcan, and I. Kuskonmaz, "Interobserver variability in Gleason histological grading of prostate cancer," *Scandin. J. Urol.*, vol. 50, no. 6, pp. 420–424, Nov. 2016.

[10] P. A. Rodriguez-Urrego, A. M. Cronin, H. A. Al-Ahmadie, A. Gopalan, S. K. Tickoo, V. E. Reuter, and S. W. Fine, "Interobserver and intraobserver reproducibility in digital and routine microscopic assessment of prostate needle biopsies," *Hum. Pathol.*, vol. 42, no. 1, pp. 68–74, Jan. 2011.

[11] W. C. Allsbrook, K. A. Mangold, M. H. Johnson, R. B. Lane, C. G. Lane, and J. I. Epstein, "Interobserver reproducibility of Gleason grading of prostatic carcinoma: General pathologist," *Human Pathol.*, vol. 32, no. 1, pp. 81–88, Jan. 2001.

[12] E. T. Sadimin, F. Khani, M. Diolombi, A. Meliti, and J. I. Epstein, "Interobserver reproducibility of percent Gleason pattern 4 in prostatic adenocarcinoma on prostate biopsies," *Amer. J. Surgical Pathol.*, vol. 40, no. 12, pp. 1686–1692, Dec. 2016.

[13] M. Zhou, J. Li, L. Cheng, L. Egevad, F.-M. Deng, L. P. Kunju, C. Magi-Galluzzi, J. Melamed, R. Mehra, S. Mendrinos, A. O. Osunkoya, G. Paner, S. S. Shen, T. Tsuzuki, K. Trpkov, W. Tian, X. Yang, and R. B. Shah, "Diagnosis of 'poorly formed glands' Gleason pattern 4 prostatic adenocarcinoma on needle biopsy," *Amer. J. Surgical Pathol.*, vol. 39, no. 10, pp. 1331–1339, Oct. 2015.

[14] A. Meliti, E. Sadimin, M. Diolombi, F. Khani, and J. I. Epstein, "Accuracy of grading Gleason score 7 prostatic adenocarcinoma on needle biopsy: Influence of percent pattern 4 and other histological factors," *Prostate*, vol. 77, no. 6, pp. 681–685, May 2017.

[15] J. Arevalo, A. Cruz-Roa, and F. A. Gonzalez, "Histopathology image representation for automatic analysis: A state-of-the-art review," *Rev. Med.*, vol. 22, no. 2, pp. 79–91, 2014.

[16] C. Li, H. Chen, X. Li, N. Xu, Z. Hu, D. Xue, S. Qi, H. Ma, L. Zhang, and H. Sun, "A review for cervical histopathology image analysis using machine vision approaches," *Artif. Intell. Rev.*, vol. 53, no. 7, pp. 4821–4862, Oct. 2020.

[17] K. Sirinukunwattana, J. P. W. Pluim, H. Chen, X. Qi, P.-A. Heng, Y. B. Guo, L. Y. Wang, B. J. Matuszewski, E. Bruni, U. Sanchez, A. Böhm, O. Ronneberger, B. B. Cheikh, D. Racoceanu, P. Kainz, M. Pfeiffer, M. Urschler, D. R. J. Snead, and N. M. Rajpoot, "Gland segmentation in colon histology images: The GlaS challenge contest," *Med. Image Anal.*, vol. 35, pp. 489–502, Jan. 2017.

[18] A. Stevens, "The haematoxylins," in *The Theory and Practice of Histological Techniques*, B. John and S. Alan, Eds., 2nd ed. London, U.K.: Longman, 1982, p. 109.

[19] K. Larson, H. H. Ho, P. L. Anumolu, and M. T. Chen, "Hematoxylin and eosin tissue stain in mohs micrographic surgery: A review," *Dermatologic Surg.*, vol. 37, no. 8, pp. 1089–1099, Aug. 2011, doi: 10.1111/j.1524-4725.2011.02051.x.

[20] C. Ortiz-Hidalgo and S. Pina-Oviedo, "Hematoxylin: Mesoamerica's gift to histopathology. Palo de campeche (logwood tree), pirates' most desired treasure, and irreplaceable tissue stain," *Int. J. Surgical Pathol.*, vol. 27, no. 1, pp. 4–14, Feb. 2019, doi: 10.1177/1066896918787652.

[21] J. Kiernan, "Does progressive nuclear staining with hemalum (alum hematoxylin) involve DNA, and what is the nature of the dye-chromatin complex?" *Biotechnic Histochemistry*, vol. 93, no. 2, pp. 133–148, Feb. 2018, doi: 10.1080/10520295.2017.1399466.

[22] D. Wittekind, "Traditional staining for routine diagnostic pathology including the role of tannic acid. 1. Value and limitations of the hematoxylin-eosin stain," *Biotechnic Histochemistry*, vol. 78, no. 5, pp. 261–270, Oct. 2003, doi: 10.1080/10520290310001633725.

[23] M. Titford, "Progress in the development of microscopical techniques for diagnostic pathology," *J. Histotechnol.*, vol. 32, no. 1, pp. 9–19, Mar. 2009, doi: 10.1179/his.2009.32.1.9.

[24] S. M. Ayyad, M. Shehata, A. Shalaby, M. Abou El-Ghar, M. Ghazal, M. El-Melegy, N. B. Abdel-Hamid, L. M. Labib, H. A. Ali, and A. El-Baz, "Role of AI and histopathological images in detecting prostate cancer: A survey," *Sensors*, vol. 21, no. 8, p. 2586, Apr. 2021, doi: 10.3390/s21082586.

[25] H. Qu, Z. Yan, G. M. Riedlinger, S. De, and D. N. Metaxas, "Improving nuclei/gland instance segmentation in histopathology images by full resolution neural network and spatial constrained loss," in *Proc. Int. Conf. Med. Image Comput. Comput. Interv.* Cham, Switzerland: Springer, 2019, pp. 378–386.

[26] Z. Khan, N. Yahya, K. Alsaih, M. I. Al-Hiyali, and F. Meriaudeau, "Recent automatic segmentation algorithms of MRI prostate regions: A review," *IEEE Access*, vol. 9, pp. 97878–97905, 2021, doi: 10.1109/ACCESS.2021.3090825.

[27] J. Li, K. V. Sarma, K. C. Ho, A. Gertych, B. S. Knudsen, and C. W. Arnold, "A multi-scale U-Net for semantic segmentation of histological images from radical prostatectomies," in *Proc. AMIA Annu. Symp.*, 2017, p. 1140.

[28] J. Li, W. Speier, K. C. Ho, K. V. Sarma, A. Gertych, B. S. Knudsen, and C. W. Arnold, "An EM-based semi-supervised deep learning approach for semantic segmentation of histopathological images from radical prostatectomies," *Computerized Med. Imag. Graph.*, vol. 69, pp. 125–133, Nov. 2018.

[29] N. Ing, Z. Ma, J. Li, H. Salemi, C. Arnold, B. S. Knudsen, and A. Gertych, "Semantic segmentation for prostate cancer grading by convolutional neural networks," *Proc. SPIE*, vol. 10581, pp. 343–355, Mar. 2018.

[30] A. Kalapahar, J. Silva-Rodríguez, A. Colomer, F. López-Mir, and V. Naranjo, "Gleason grading of histology prostate images through semantic segmentation via residual U-Net," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Abu Dhabi, United Arab Emirates, Oct. 2020, pp. 2501–2505, doi: 10.1109/ICIP40778.2020.9191250.

[31] M. Singh, E. M. Kalaw, D. M. Giron, K.-T. Chong, C. L. Tan, and H. K. Lee, "Gland segmentation in prostate histopathological images," *J. Med. Imag.*, vol. 4, no. 2, Jun. 2017, Art. no. 027501.

[32] T. Ali, K. Masood, M. Irfan, U. Draz, A. A. Nagra, M. Asif, B. M. Alshehri, A. Glowacz, R. Tadeusiewicz, M. H. Mahnashi, and S. Yasin, "Multistage segmentation of prostate cancer tissues using sample entropy texture analysis," *Entropy*, vol. 22, no. 12, p. 1370, Dec. 2020, doi: 10.3390/e22121370.

[33] J. Ren, E. Sadimin, D. J. Foran, and X. Qi, "Computer aided analysis of prostate histopathology images to support a refined Gleason grading system," *Proc. SPIE*, vol. 10133, Feb. 2017, Art. no. 101331V.

[34] Y. Xu, Y. Li, Y. Wang, M. Liu, Y. Fan, M. Lai, and E. I.-C. Chang, "Gland instance segmentation using deep multichannel neural networks," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 12, pp. 2901–2912, Dec. 2017.

[35] J. Silva-Rodríguez, A. Colomer, and V. Naranjo, "WeGleNet: A weakly-supervised convolutional neural network for the semantic segmentation of Gleason grades in prostate histology images," *Computerized Med. Imag. Graph.*, vol. 88, Mar. 2021, Art. no. 101846.

[36] H. Oda, H. R. Roth, K. Chiba, J. Sokolic, T. Kitasaka, M. Oda, A. Hinoki, H. Uchida, and J. A. Schnabel, "BESNet: Boundary-enhanced segmentation of cells in histopathological images," in *Medical Image Computing and Computer Assisted Intervention—MICCAI*. Granada, Spain: Springer, 2018, pp. 228–236.

[37] P. Naylor, M. Laé, F. Reyal, and T. Walter, "Segmentation of nuclei in histopathology images by deep regression of the distance map," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 448–459, Feb. 2019.

[38] S. Graham and N. M. Rajpoot, "SAMS-NET: Stain-aware multi-scale network for instance-based nuclei segmentation in histology images," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 590–594.

[39] G. Xu, Z. Song, Z. Sun, C. Ku, Z. Yang, C. Liu, S. Wang, J. Ma, and W. Xu, "CAMEL: A weakly supervised learning framework for histopathology image segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 10682–10691.

[40] S. Graham, H. Chen, J. Gamper, Q. Dou, P.-A. Heng, D. Snead, Y. W. Tsang, and N. Rajpoot, "MILD-net: Minimal information loss dilated network for gland instance segmentation in colon histology images," *Med. Image Anal.*, vol. 52, pp. 199–211, Feb. 2019, doi: 10.1016/j.media.2018.12.001.

[41] F. Mahmood, D. Borders, R. J. Chen, G. N. McKay, K. J. Salimian, A. Baras, and N. J. Durr, "Deep adversarial training for multi-organ nuclei segmentation in histopathology images," *IEEE Trans. Med. Imag.*, vol. 39, no. 11, pp. 3257–3267, Nov. 2020, doi: 10.1109/tmi.2019.2927182.

[42] J. Song, L. Xiao, M. Molaei, and Z. Lian, "Multi-layer boosting sparse convolutional model for generalized nuclear segmentation from histopathology images," *Knowl.-Based Syst.*, vol. 176, pp. 40–53, Jul. 2019, doi: 10.1016/j.knosys.2019.03.031.

[43] S. Graham, D. Epstein, and N. Rajpoot, "Rota-Net: Rotation equivariant network for simultaneous gland and lumen segmentation in colon histology images," in *Proc. Eur. Congr. Digit. Pathol.* (Lecture Notes in Computer Science), 2019, pp. 109–116, doi: 10.1007/978-3-030-23937-4_13.

[44] M. van Rijthoven, M. Balkenhol, K. Silina, J. van der Laak, and F. Ciompi, "HookNet: Multi-resolution convolutional neural networks for semantic segmentation in histopathology whole-slide images," *Med. Image Anal.*, vol. 68, Feb. 2021, Art. no. 101890.

[45] T. Wan, L. Zhao, H. Feng, D. Li, C. Tong, and Z. Qin, "Robust nuclei segmentation in histopathology using ASPPU-Net and boundary refinement," *Neurocomputing*, vol. 408, pp. 144–156, Sep. 2020, doi: 10.1016/j.neucom.2019.08.103.

[46] H. Qu, P. Wu, Q. Huang, J. Yi, Z. Yan, K. Li, G. M. Riedlinger, S. De, S. Zhang, and D. N. Metaxas, "Weakly supervised deep nuclei segmentation using partial points annotation in histopathology images," *IEEE Trans. Med. Imag.*, vol. 39, no. 11, pp. 3655–3666, Nov. 2020.

[47] S. Lal, D. Das, K. Alabhya, A. Kanfade, A. Kumar, and J. Kini, "NucleiSegNet: Robust deep learning architecture for the nuclei segmentation of liver cancer histopathology images," *Comput. Biol. Med.*, vol. 128, Jan. 2021, Art. no. 104075.

[48] M. Salvi, M. Bosco, L. Molinaro, A. Gambella, M. Papotti, U. R. Acharya, and F. Molinari, "A hybrid deep learning approach for gland segmentation in prostate histopathological images," *Artif. Intell. Med.*, vol. 115, May 2021, Art. no. 102076, doi: 10.1016/j.artmed.2021.102076.

[49] T. A. A. Tosta, P. R. de Faria, L. A. Neves, and M. Z. do Nascimento, "Computational normalization of H&E-stained histological images: Progress, challenges and future potential," *Artif. Intell. Med.*, vol. 95, pp. 118–132, Apr. 2019.

[50] M. Niethammer, D. Borland, J. S. Marron, J. Woosley, and N. E. Thomas, "Appearance normalization of histology slides," in *Proc. Int. Workshop Mach. Learn. Med. Imag.* Cham, Switzerland: Springer, 2010, pp. 58–66.

[51] A. Anghel, M. Stanisavljevic, S. Andani, N. Papandreou, J. H. Rüschoff, P. Wild, M. Gabrani, and H. Pozidis, "A high-performance system for robust stain normalization of whole-slide images in histopathology," *Frontiers Med.*, vol. 6, p. 193, Sep. 2019, doi: 10.3389/fmed.2019.00193.

[52] M. Salvi, F. Molinari, U. R. Acharya, L. Molinaro, and K. M. Meiburger, "Impact of stain normalization and patch selection on the performance of convolutional neural networks in histological breast and prostate cancer classification," *Comput. Methods Programs Biomed. Update*, vol. 1, 2021, Art. no. 100004, doi: 10.1016/j.cmpbup.2021.100004.

[53] C. Chen, Y. Huang, P. Fang, C. Liang, and R. Chang, "A computer-aided diagnosis system for differentiation and delineation of malignant regions on whole-slide prostate histopathology image using spatial statistics and multidimensional DenseNet," *Med. Phys.*, vol. 47, no. 3, pp. 1021–1033, Mar. 2020.
[54] M. Salvi, U. R. Acharya, F. Molinari, and K. M. Meiburger, "The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis," *Comput. Biol. Med.*, vol. 128, Jan. 2021, Art. no. 104129.
[55] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, p. 1097.
[56] D. Sarkar, R. Bali, and T. Ghosh, *Hands-on Transfer Learning With Python: Implement Advanced Deep Learning and Neural Network Models Using TensorFlow and Keras*. Birmingham, U.K.: Packt, 2018.
[57] D. Maji, P. Sigedar, and M. Singh, "Attention res-UNet with guided decoder for semantic segmentation of brain tumors," *Biomed. Signal Process. Control*, vol. 71, Jan. 2022, Art. no. 103077, doi: 10.1016/j.bspc.2021.103077.
[58] M. Salvi, N. Michielli, and F. Molinari, "Stain color adaptive normalization (SCAN) algorithm: Separation and standardization of histological stains in digital pathology," *Comput. Methods Programs Biomed.*, vol. 193, Sep. 2020, Art. no. 105506.
[59] M. Salvi and F. Molinari, "Multi-tissue and multi-scale approach for nuclei segmentation in H&E stained images," *Biomed. Eng. OnLine*, vol. 17, no. 1, pp. 1–13, Dec. 2018, doi: 10.1186/s12938-018-0518-0.
[60] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, 2001.
[61] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math.*, vol. 42, no. 5, pp. 577–685, Jul. 1989.
[62] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa, "The quickhull algorithm for convex hulls," *ACM Trans. Math. Softw. (TOMS)*, vol. 22, no. 4, pp. 469–483, 1996.
[63] S. Sahran, D. Albashish, A. Abdullah, N. A. Shukor, and S. H. M. Pauzi, "Absolute cosine-based SVM-RFE feature selection method for prostate histopathological grading," *Artif. Intell. Med.*, vol. 87, pp. 78–90, May 2018.
[64] M. Talo, "Automated classification of histopathology images using transfer learning," *Artif. Intell. Med.*, vol. 101, Nov. 2019, Art. no. 101743.
[65] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52138–52160, 2018.
[66] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," 2015, *arXiv:1503.02531*.