FMDB *Transactions on Sustainable Health Science Letters*

Segmentation and Classification in Prostate Cancer Using GTOA-Based Regional Attention Mechanism with Multi-Task Learning Network

B. Manjunatha[1,*], K. Aravinda[2], M. Arunadevi Thirumalraj[3], S. Gopikha[4], M. Rehena Sulthana[5]

[1]Department of Mechanical Engineering, New Horizon College of Engineering, Bengaluru, Karnataka, India.
[2]Department of Electronics and Communications Engineering, New Horizon College of Engineering, Bengaluru, Karnataka, India.
[3]Department of Computer Science and Engineering, Karunya Institute of Technology and Science, Coimbatore, Tamil Nadu, India.
[3]Department of Computer Science and Business Management, Saranathan College of Engineering, Tiruchirappalli, Tamil Nadu, India.
[4]Department of Information Technology, St. Joseph's College of Engineering, Chennai, Tamil Nadu, India.
[5]School of Information Technology and Engineering, Melbourne Institute of Technology, Melbourne, Victoria, Australia.
manjunatha.princi@gmail.com[1], aravindake@gmail.com[2], aruna.devi96@gmail.com[3], gopikha.in@gmail.com[4], rsulthana@academic.mit.edu.au[5]

Abstract: Helpful tools for pathologists include the ability to identify and segment histological areas of diagnostic interest automatically. The challenge with segmentation methods is that obtaining pixel-level annotations for Whole-Slide images (WSI) is both time-consuming and costly. As a solution, weakly supervised approaches have been developed to leverage image-level annotations. But none of these methods have been modified to handle WSIs as far as researchers are aware. This research proposes a region-aware multi-task learning framework (RMTL-Net) to detect benign or malignant prostate tumours in images and simultaneously segment their regions. This model includes a regional attention (RA) module that uses the predicted likelihood maps to automatically train the classifier to learn category-sensitive information in the tumour, peritumoral, and background regions. By seamlessly fusing these regions, it improves feature representation and, in turn, the model's segmentation and classification performance. The Group Teaching Optimisation Algorithm (GTOA) model optimally selects the hyperparameter tuning. Using three publicly available prostate datasets, the research compared the proposed RA module with three state-of-the-art methods implemented recently and conducted extensive ablation experiments.

Keywords: Tumor Classification; Whole-Slide Images; Regional Attention; Prostate Cancer; Regional-Attentive and Multi-Task Learning Framework; Group Teaching Optimization Algorithm.

Received on: 17/01/2025, **Revised on:** 24/03/2025, **Accepted on:** 27/05/2025, **Published on:** 06/12/2025

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSHSL

DOI: https://doi.org/10.69888/FTSHSL.2025.000511

Cite as: B. Manjunatha, K. Aravinda, M. A. Thirumalraj, S. Gopikha, and M. R. Sulthana, "Segmentation and Classification in Prostate Cancer Using GTOA-Based Regional Attention Mechanism with Multi-Task Learning Network," *FMDB Transactions on Sustainable Health Science Letters*, vol. 3, no. 4, pp. 218–230, 2025.

1. Introduction

[*]Corresponding author.

With 250,000 new cases reported and 35,000 fatalities in 2022, prostate cancer ranks as the second most common cancer in men in the US [1]. However, pathologists are becoming increasingly rare, despite their essential role in the diagnosis and treatment of cancer patients. In the US, the average workload increased by 42% from 2007 to 2017, while the overall workload decreased by 18%. Furthermore, there are certain difficulties inherent to uropathology [2]. Although there are established diagnostic criteria for prostate cancer grading, there is still room for disparities due to the continuum of phenotypic features across the diagnostic spectrum, which results in significant intra- and interobserver variability [3]. Automation and standardisation would be great ways to save time and effort when manually inspecting slides. Because of these factors, CAD tools for automating the diagnostic workflow have been developed [4]. Consequently, several CAD tools powered by AI have been proposed, including those for nucleus segmentation and classification, gland segmentation, and tumour detection [5]. Despite their outstanding performance, these tools often require task- and tissue-specific annotations across massive datasets, which can be painstaking, time-consuming, and even impossible to accomplish [6]. Various methods have been proposed to reduce the need for annotations, with weakly supervised approaches based on the Multiple Instance Learning (MIL) framework for automatic classification of Whole-Slide Images (WSIs) being particularly prominent [7]. Shown in Figure 1 are the images of the sample colon.

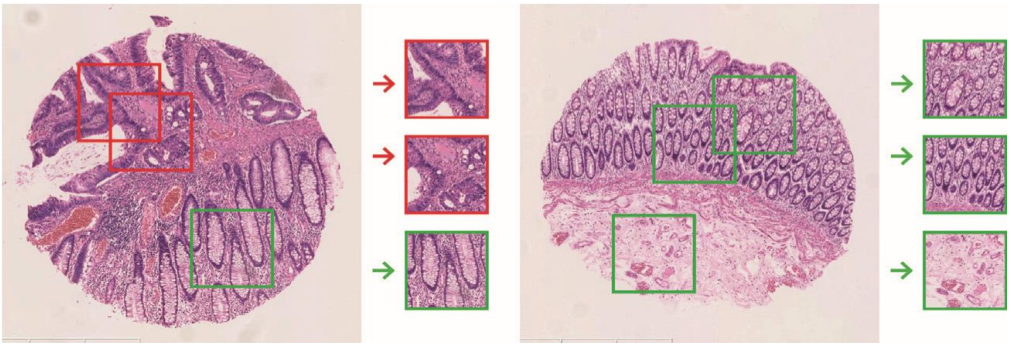

Figure 1: Sample histopathology of colon cancer, then non-cancer images

The function of classification in assisting the pathologist's focus during diagnosis is still limited, despite its usefulness. Because they enable pixel-level delineation of tissue constituents that can highlight diagnostically relevant regions, semantic segmentation methods are preferable in this context [8]; [9]. Pathologists can gain greater confidence in CAD tools with this kind of visualisation. A classifier can also use the detected areas to enhance patient diagnosis [10]. In contrast to classification tasks, semantic segmentation typically requires pixel-level labels, making it more annotation-intensive. This is why it seems the best course of action is to develop WSSS (weakly supervised semantic segmentation) methods [11]. Unfortunately, histopathology images present several obstacles for WSSS, including the following: (1) the presence of fine-grained objects with large intra-class variations; (2) the presence of ambiguous boundaries among histology components; and (3) the possibility of several gigapixels with arbitrary tissue sizes [12]. However, for different histology tasks, WSSS methods have been suggested. Methods that employ patch-level WSSS have limitations, such as the need for patch-level annotations and the inability to perform global, context-aware WSI segmentation [13]. Their computational complexity and memory requirements for working on WSIs are high, but they scale to larger tiles. Because pathologists need to annotate images beyond what is typically required in clinical practice, the methods require exact tile-level comments for model training, meaning that each lesion type in a tile must be precisely designated.

To change the subject, new WSI classification algorithms emphasise important areas by using attention mechanisms or attribute features [14]. Despite their usefulness for visual evaluation, these regions lack sufficient detail, are incomplete, and are too blurry to demarcate what is important. Producing granular saliency from WSIs is computationally expensive because it requires densely overlapping patch predictions [15]. Here, to present a RA module that can learn category-sensitive prostate images' three regions—the tumour, peritumoral, and background regions—and study how these features affect MTL. The study also shows that the proposed RA module performs well for prostate image classification and segmentation when integrated into a two-stage MTL framework. The segmentation and classification components of the suggested model are an encoder-decoder network and a lightweight network, respectively. Features retrieved from the encoder are used in both classification and segmentation. Furthermore, the RA module directs the classification network to features for improved classification using the predicted probability maps. Both tasks are improved by RMTL-Net, according to the experimental results.

2. Related Works

To segment and categorise WSIs of any size and shape simultaneously, Pati et al. [16] proposed WholeSIGHT, a weakly supervised approach. In its formal implementation, WholeSIGHT begins by building a WSI representation as a tissue graph,

with nodes representing tissue regions and edges representing their interactions. Using post-hoc feature attribution, a graph classification head sorts WSI during training and generates pseudo-labels at the node level. The next step in WSI segmentation using these pseudo labels is to train a node classification head. While testing, the two heads work in tandem to generate a WSI input and do class prediction and segmentation. Using three publicly available prostate cancer WSI datasets to assess WholeSIGHT. Across all datasets, our approach outperformed or matched the state of the art in segmentation performance. In addition, researchers measure our method's generalisability by calculating its uncertainty estimation, model calibration, segmentation and classification performance, and overall effectiveness. In a bootstrapping fashion, the segmentation and classification components of MC-DSCN can exchange information with each other, as suggested by Liu et al. [17]. With the MC-DSCN, the masks generated by the coarse segmentation constituent can be transferred to the classification component, allowing for easier classification by excluding irrelevant regions. To reduce the impact of inaccurate localisation on segmentation results, this prototype can facilitate the transfer of high-quality localisation information learnt by the classification component during the segmentation task. Two medical centres, designated centres A and B, were retrospectively surveyed for consecutive MRI exams of patients.

The results of the prostate biopsy serve as the basis for the classification, which was determined by two seasoned radiologists who segmented the prostate regions. The impact of network architectures on performance was examined and tested during the design, training, and validation of MC-DSCN with different combinations of unique MRI sequences as input. Training, validation, and internal testing used data from centre A, while external testing used data from centre B. The MC-DSCN's performance is assessed through statistical analysis. To accurately diagnose PCa using bpMRI images, Du et al. [18] proposed a two-branch Noise-Resistant Distillation Network (NRD-Net). Segmenting the labelled constrained classification response maps is the first step in reducing the impact of immaterial background classification. Next, a multi-branch online arrangement is proposed to simultaneously reduce spatial and grading noise. Private and public datasets, including the PROSTATEx-2 dataset, are subjected to extensive experiments. With an average positive predictive value (PPV) of 0.9506 and a Kappa of 0.5115, the suggested network achieves the best performance for GG forecast on the private dataset. With a state-of-the-art Kappa of 0.5058 and a PPV of 0.9473, the suggested method accomplishes impressive results on the public dataset. Radiologists can now use a computer-aided diagnostic system developed by Bouslimi et al. [19] to better segment prostate cancer (PCa) lesions from multiparametric magnetic resonance imaging (mp-MRIs) and to more accurately predict cancer versus benign lesions.

To accomplish this, the researcher proposed a method that classified clinically relevant cancers using deep learning neural network models such as ResNet and Inception. Additionally, the prostate lesion was automatically segmented from mp-MRIs using the U-Net and MultiResU-Net architectures. The Radboudumc and ProstateX datasets were utilised, both of which are publicly accessible. The PCa lesion classification model's curve exceeded 98.4%, indicating that our fully automated system performed as expected. In contrast, PCa lesion segmentation accuracy was 98.34% for the MultiResU-Net model. To illustrate the impact of AI on cancer treatment and medicine, Singla et al. [20] presented a new U-Net architecture that integrates transformers, CNNs, and medical data. The PROMISE-12 dataset, which contains all MRI images of prostate cancer, was used to test this model. Its architecture was combined with other traditional deep learning representations to assess its performance. All three architectures were evaluated for their predicted segmentation performance using Dice scores. Compared with more traditional AI architectures, our data show that U-Nets based on Transformers perform better on MRI scans of prostate cancer. This innovative U-Net architecture, based on Transformers, helps advance prostate cancer segmentation and could be used in the future for cancer diagnosis. To segment prostate cancer, Li et al. [21] proposed a network, CDA-Net, that accounts for both the localisation of cancer edges and the global localisation of cancer regions.

To be more precise, to suggest an improved localisation accuracy through feature extraction, and to create a connection mechanism between a GAN and a unit to enhance edge recognition. The results demonstrate that the suggested network segmentation methods improve segmentation performance by 1.7% in IoU, 3.8% in PA, and 1.7% in Dice. Additionally, in MRI images, the 95% Hausdorff distance is reduced by 1.82. In their groundbreaking work, Bygari et al. [22] developed a deep learning system that automatically screens prostate images for cancer and grades its severity. Visualisation using gradient-weighted class activation mapping (GradCAM) improved the classification model's explainability; this method creates a heatmap of the image that the model used to make its decision. To assess prostate cancer severity, the suggested approach employs an ensemble of deep neural networks in three steps.

The initial steps involve segmenting the histopathological image using a U-Net. Afterwards, the original image is superimposed with the segmented one to highlight the key areas that determine the cancer grade. Xception, Resnet-50, and EfficientNet-b7 form an ensemble model that uses an overlaid image to predict the final grade of the histopathology image. For training and evaluation, use the publicly available dataset from the Prostate Cancer Grade Assessment Test on Kaggle, which contains 10,000 histopathological images from the Karolinska and Radboud datasets. This technique surpasses numerous state-of-the-art approaches and achieves a classification accuracy of 92.38%.

3. Proposed System

3.1. Datasets

This research tests our approach using three datasets that include Gleason pattern segmentation and grading from whole-slide needle biopsies taken from prostate cancer patients. In grade 3 (G3), the nuclei are moderately differentiated, and the cribriform glands are poorly formed; in grade 4 (G4), the nuclei are poorly differentiated, and there are irregular masses; and in grade 5 (G5), the nuclei are less differentiated, and there are either no glands at all or only occasional ones. Benign (B) refers to normal glands and tissues that are not epithelial. A WSI's Gleason grade shows the two worst Gleason patterns, the primary (P) and the secondary (S). See below for details of the dataset:

- **Sicap Dataset:** The dataset by Silva-Rodríguez et al. [23] contains 18,783 512×512 patches from 155 WSIs across 95 patients. Each patch has full pixel-level annotation grades. Stitching the patches together enabled reconstruction of the original masks and slides. To use a Ventana iS-can Coreo scanner to scan the WSIs at 40× magnification, and then to downsample them to 10× magnification. Professional urogenital pathologists from the Hospital Cl'ınico in Valencia, Spain, annotated the slides.
- **Radboud Dataset:** The Radboud University Medical Centre in the Netherlands collected 5,759 needle biopsies from 1,243 patients [24]. The slides were scanned at 20× magnification (0.24µm/pixel resolution) using a 3D Histech Panoramic Flash II 250 scanner and then downsampled to 10×. As part of the Prostate Cancer Grade Assessment (PANDA) challenge, annotations were released, which include WSI Gleason grades and noisy pixel-level Gleason pattern masks [25]. Using common image manipulation techniques such as contextualised noise elimination, hole filling, and edge smoothing, the masks were prepared for segmentation. To utilise this dataset for method development and evaluation, given the WSIs.
- **Karolinska Dataset:** A total of 1,222 patients from different hospitals in Stockholm, Sweden, provided 5,662 core needle biopsies for the dataset [26]. The Aperio Scan Scope AT2 scanner had a pixel resolution of 0.5032µm, and the Hamamatsu C9600-12 scanner had a 20× magnification, respectively, for the slides. A highly trained uropathologist annotated the biopsies according to the Gleason scale.

3.2. Network Architecture

In this work, the proposed model is used for segmentation and classification, as shown in Figure 2.

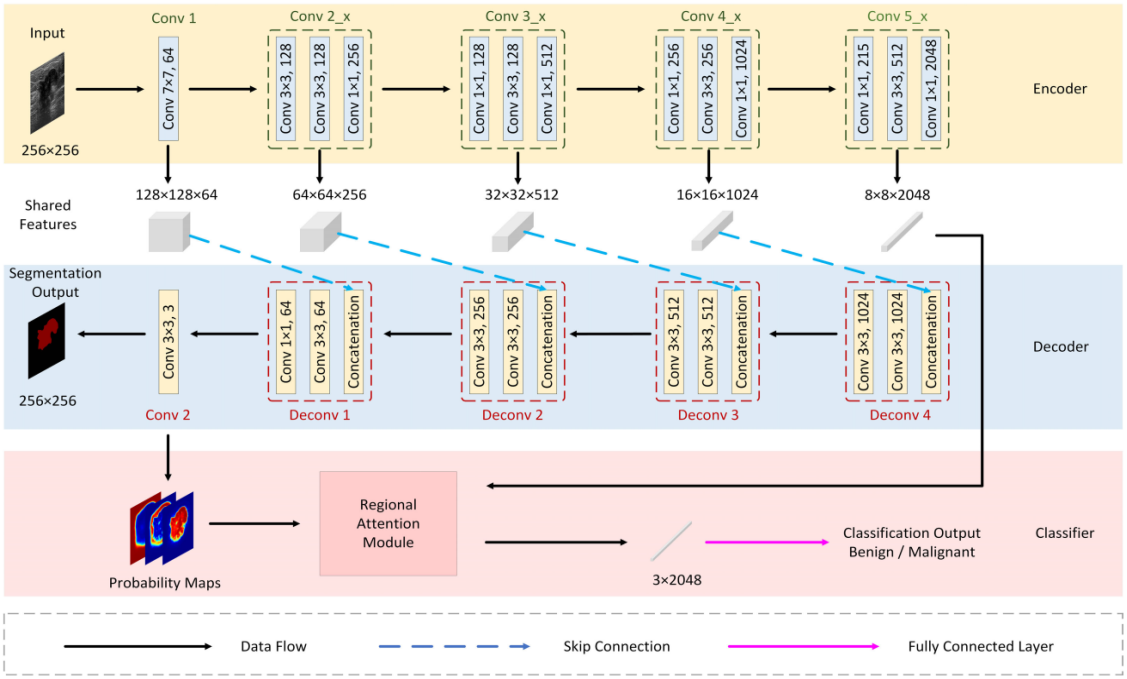

Figure 2: A comprehensive illustration of the projected RMTL-Net

Segmentation and classification are the two main components of RMTL-Net's two-stage architecture. Segmentation uses a U-shaped architecture with encoders, decoders, and skip connections to predict three separate probability maps for tumour,

peritumoral, and background regions, and to extract multi-scale features. Results from the segmentation stage's three probability maps and shared encoder features are used in the classification stage. In particular, to extract boundary features from the peritumoral region that help distinguish between benign and malignant tumours. To aid in tumour segmentation and classification, and to record tumour shape features in the tumour region. To record posterior acoustic shadowing in the background area, which is more common in cancerous lesions and less common in benign tumours, because the sonographic signal is attenuated. During training, segmentation and classification mutually support one another by sharing features. It also addresses the issue of insufficient training images for classification. In segmentation, each pixel serves as a training sample. The classification stage's overall accuracy and robustness are enhanced when features are shared with the segmentation stage, and there are enough training samples.

Because of its outstanding performance in BUS image segmentation and classification, ResNet-101 is utilised as the foundation of the segmentation stage of RMTL-Net in this study [27]. Rest assured, the architecture of ResNet-101 remains unchanged. To extract multi-scale features from input images, the encoder uses five down-sampling operations, followed by one convolutional layer (Conv1), then four blocks (Conv2_x to Conv5_x). From Conv1 to Conv5_x, multi-scale features are extracted with sizes of 128 × 128 × 64, 64 × 64 × 256, 32 × 32×512, 16×16×1024, and 8×8×2048, in that order. For upsampling, the decoder uses four deconvolutional blocks (Deconv4–Deconv1) symmetrically, followed by bilinear interpolation and softmax. To improve segmentation results and make up for the loss of sampling operations, the encoder and decoder use skip connections that combine feature maps of different sizes. This leads to the prediction of three probability maps after resizing multi-scale features to their initial input size. To achieve classification results, employ three probability maps produced by the segmentation stage of RMTL-Net in conjunction with multi-scale high-level features that are common to both the segmentation and classification phases.

3.2.1. Regional Attention Module

To further promote information sharing, to incorporate a regional attention (RA) model, distinguishing it from classical image classification networks like VGG and ResNet. A fully connected layer is used to produce more precise classification results from the weighted feature vector of size 1×2048 produced by this RA model. Benign and malignant tumours differ in their features, according to a study. To illustrate the point, benign tumours are typically round and smooth, whereas malignant tumours are invariably rough and have an aspect ratio higher than 1. The margins of benign tumours are typically thin, smooth, and regular, whereas the margins of malignant tumours are typically thick, spiculated, and irregular. Posterior acoustic shadowing in the background region is typically less pronounced in benign tumours than in malignant lesions. To aid in tasks, to suggest learning the intrinsic characteristics of tumour, peritumoral, and background regions. These characteristics include tumour features (such as component, internal echo, and aspect ratio) and tumour boundary patterns (such as smoothness, shape, and contextual texture between features). Furthermore, to promote data exchange and generate a weighted feature vector for classification, suggest integrating an RA module into the RMTL-Net's classification stage. To direct the learning of the three regions tumour, the peritumoral area, and the background, this RA module integrates multi-scale three probability maps produced at this stage. To improve classification accuracy, a fully connected layer is fed a weighted feature vector that captures category-sensitive features. The algorithmic view of the RA unit is abridged below:

Input: C5, which is the $8 \times 8 \times 2048$ feature map obtained by the encoder's Conv5_x, and P, which is the $256 \times 256 \times 3$ probability map produced by the decoder's last convolutional layer, Conv2.

Output: A weighted feature map FW of size 1×2048.

Split P into three likelihood maps P_T, P_P, and P_B of size 256 × 256, where subscripts T, P, and B represent tumour, peritumoral, and background, respectively.

Employ the nearest neighbour technique to coarse probability maps P_T^', 〖P'〗_B, and 〖P'〗_B of size 8 × 8.

Utilise a verge of 0.5 to filter coarse likelihood maps correspondingly. Specifically, values greater than 0.5 in coarse probability maps are kept intact, equal to 0.5 are oset to 0:

$$P_x'' = P_x' > 0.5 ? P_x' : 0 \quad (1)$$

Where subscript x can be replaced with T, P, or B.

Individually and element wisely multiply P_x'' channel of C_5 to generate multi channel weighted regional feature maps:

$$C_x = C_5 . P_x'' \quad (2)$$

Apply the global average pooling (GAP) on C_x to capture weights of each region in its corresponding G_x of size 1 × 2048:

$$G_x = GAP(C_x) \quad (3)$$

Concatenate G_T, G_P, and G_B to conceive a new-fangled feature vector F of size 3 × 2048:

$$F = Concatenate(G_T, G_P, G_B) \quad (4)$$

To create a weighted feature map FW with dimensions 1×2048, apply a 1×1 convolution filter to F:

$$F_w = f^{1\times1}(F) \quad (5)$$

It should be renowned non-zero pixels in P_T'', P''_B, besides P''_B the tumour's strength, peritumoral characteristics, and background features are all indicated by high likelihood values (>0.5). To divide a pixel into three categories, the researcher can choose 0.5 as the threshold. This is the product of C5 and P_T'', P''_B, besides P''_B tumour, peritumoral, and background features C_T, C_P, and C_B. The GAP operation further finds the features in each channel of C_T, C_P, and C_B to adequately portray three different areas. The three parallel channels' multi-view characteristics are weighted and summed via concatenation and 1×1 convolution. (i.e., G_T, G_P, and G_B), which can be formulated as:

$$F_W = w_1 . G_T + w_2 . G_P + w_3 . G_B \quad (6)$$

where w_1, w_2, besides w_3 suggest that the tumour, peritumoral, and background areas are significant, in that order. During training, these weights are automatically learnt. A fully connected layer receives a feature map (i.e., convoluting C_5 the vector has dimensions of 1×2048. To summarise, the suggested RA module improves segmentation and classification presentation by learning multi-view features from three input datasets while keeping radiologists' viewpoints in mind. In particular, the tumour region is useful for extracting the fundamental characteristics of prostate tumours. Tumour boundary patterns can be better captured in the peritumoral area. Gathering contextual information is made easier by the background region.

3.2.2. Loss Function

The segmentation task calculates the complete loss of RMTL-Net. L_{seg} besides loss of the task L_{cls}:

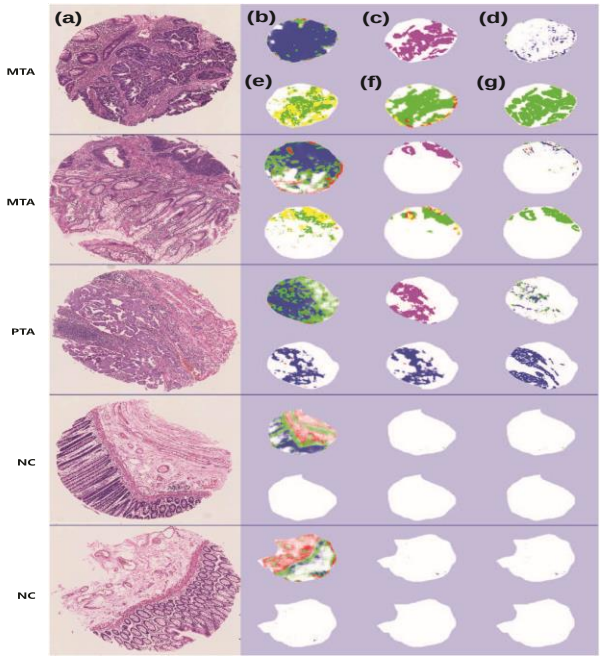

Figure 3: Sample segmented images

$$L = \lambda . L_{seg} + (1 - \lambda) . L_{cls} \quad (7)$$

Where λ and 1 − λ are also known as the weights associated with the classification task losses and segmentation task weights, respectively. The cross-entropy is used to add both. L_{seg} besides L_{cls}. For any given assignment, let K be the sum of classes, N be the sum of photos, and P be the sum of images. There are three categories in the segmentation task: tumour, peritumoral, and backdrop. K=3, to rephrase. The cross-entropy at each pixel L_{seg} of segmentation task is calculated as shadows:

$$L_{seg} = -\frac{1}{P}\sum_{p}^{P}\sum_{k}^{K} y_p k . \log \hat{y}_{p,k} \quad (8)$$

where $y_{p,k}$ and $\hat{y}_{p,k}$ that pixel p belongs to class k, and the anticipated likelihood of its belonging to that class, respectively. Since every pixel can be classified into one of three groups, the actual probability $\hat{y}_{p,k}$ can only be zero or one.The predicted likelihood $\hat{y}_{p,k}$ is in range of (0, 1). In classification, there are two classes: benign and malignant tumours. In other words, K = 2. The entropy L_{cls} of task is calculated as shadows:

$$L_{cls} = -\frac{1}{N}\sum_{n}^{N}\sum_{k}^{K} y_{n,k} . \log \hat{y}_{n,k} \quad (9)$$

where $y_{n,k}$ and $\hat{y}_{n,k}$ describe the expected and actual classification of class k image n, respectively. Both $y_{n,k}$ besides $\hat{y}_{n,k}$ are either 0 or 1. Figure 3 provides sample images of the classical segmentation.

3.2.3. Hyper-Parameter Tuning Using GTOA

A group instruction method is used to stimulate the GTOA being presented. Considering a range of student variations makes real-time execution difficult. Before making any adjustments to make this method suitable for optimisation, they take into account the student population, decision parameter, and fitness value [28]. A basic group-teaching module that retains its generalisability is subsequently developed. The technique achieves consensus using a pseudorandom witness. Adding throughput and latency to the formula makes the algorithm more scalable. It is recommended to employ a decentralised, scalable, low-latency, high-throughput consensus method. This study proposes an enhanced hybrid consensus algorithm based on the GTOA. To ensure low latency and high throughput, while ensuring fair participation by many nodes in the consensus, verifiable cryptographic sortition dynamically chooses a consensus node. The four steps that make up the given module are capability grouping, allocation, the student stage, and the teacher phase. What follows is a definition of the four steps.

3.2.4. Capability Grouping Stage

When entire classes' knowledge follows a normal distribution, there is no loss of generalisability. This is provided by:

$$f(x) = \frac{1}{\sqrt{2\pi\delta}} e^{\frac{(x-u)^2}{2\delta^2}} \quad (10)$$

x is the rate of normal delivery function, u is the mean data of all classes, and δ is the standard deviation. It is believed that a good teacher will raise their pupils' average knowledge, but an educator must devise an appropriate instructional method. Each student is divided into two smaller groups based on their aptitude for applying knowledge in GTOA. This will demonstrate the group teaching feature without sacrificing generalisability. Notably, the standard deviation of both the excellent and average groups may exceed the transfer of teaching events. To address this issue, GTOA employs a learning-cycle-based energy procedure called capability grouping.

3.2.5. Teacher Stage

What this means is that each student learns something new from their teacher that is relevant to the rule that will be determined later on. For both average and exceptional students, the teacher in the given GTOA develops a range of pedagogical approaches.

3.2.6. Teacher Stage I

The instructor places greater emphasis on expanding the exceptional group's knowledge after gaining insight into their greater capacity for knowledge adaptation. In particular, the instructor might do all in their power to raise the average level of understanding across all students. It is also important to remember that students will have different ways of applying what they have learnt. Consequently, the exceptional students could acquire their knowledge as follows:

$$x_{teacher,i}^{t+1} = x_i^t + a \times \left(T^t - F \times (b \times M^t + c \times x_i^t)\right) \quad (11)$$

$$M^t = \frac{1}{N}\sum_{i=1}^{N} x_i^t \quad (12)$$

$$b + c = 1 \quad (13)$$

Where t characterises the present sum of repetitions, N designates the sum of students, x_i^t signifies knowledge of time t, T^t designates knowledge of time t, M^t characterises the mean data of time t, F factor, $x_{teacher,i}^{t+1}$ represents the information that student i has learnt from their teacher at time t, and the values a, b, and c can take any value between 0 and 1. As manufactured, F can have a value of one or two.

3.2.7. Teacher Stage II

Given students' limited capacity for knowledge adaptation, a teacher prioritises the average group over the excellent group by adhering to guidelines that tend to improve students' knowledge based on personal perceptions. Students from the average group might thus acquire knowledge in the same way:

$$x_{teacher,i}^{t+1} = x_i^t + 2 \times d(T^t - x_i^t) \quad (14)$$

Where d might be any integer between 0 and 1; furthermore, students might not learn anything that the teacher teaches them in the beginning, when even the simplest topic is treated in the following way:

$$x_{teacher,i}^{t+1} = \begin{cases} x_{teacher,i}^{t+1}, f\left(x_{teacher,i}^{t+1}\right) < f(x_i^t) \\ x_i^t, f\left(x_{teacher,i}^{t+1}\right) \geq f(x_i^t) \end{cases} \quad (15)$$

3.2.8. Student Stage

Included are the first two steps for the third rule mentioned. During their leisure time, students have two options for expanding their knowledge: either by educating themselves or by engaging with other students, both of which are provided by:

$$x_{teacher,i}^{t+1} = \begin{cases} x_{teacher,i}^{t+1} + e \times \left(x_{teacher,1}^{t+1} - x_{teacher,j}^{t+1}\right) + g \times \left(x_{teacher,i}^{t+1} - x_i^t\right), f(x_{teacher,i}^{t+1}) \\ x_{teacher,i}^{t+1} - e \times \left(x_{teacher,i}^{t+1} - x_{teacher,j}^{t+1}\right) + g \times \left(x_{teacher,i}^{t+1} - x_i^t\right), f(x_{teacher,i}^{t+1}) \end{cases} \quad (16)$$

Where e besides g characterise two arbitrary extents of zero then one, $x_{teacher,i}^{t+1}$ designates information of student i at period t via student stage, then $x_{teacher,j}^{t+1}$ indicates knowledge of the teacher. The student $j(j \in \{1, 2, ., i-1, i+1, ., N\})$, Their selection is completely at random. The third right-hand item in Eq. (16) denotes self-teaching, while the next item suggests learning from another student. In addition, while addressing the student stage, it is possible that individual students may not acquire knowledge in the following ways:

$$x_i^{t+1} = \begin{cases} x_{teacher,i}^{t+1}, f\left(x_{teacher,i}^{t+1}\right) < f(x_{teacher,i}^{t+1}) \\ \left(x_{teacher,i}^{t+1}\right), \quad f(x_{teacher,i}^{t+1}) \geq (x_{teacher,i}^{t+1}) \end{cases} \quad (17)$$

where x_i^{t+1} characterises the data of student i at period t + 1 in the learning sequence.

3.2.9. Teacher Allocation Stage

The development of a fair method for allocating teachers is crucial to raising pupils' level of knowledge, as stated in the fourth rule. In GWO, wolf hunting is guided by the first three optimal solutions obtained. The hunting behaviour in GWO is described by the distribution provided in the technique:

$$T^t \begin{cases} x_{first}^t, & f(x_{first}^t) \leq f\left(\frac{x_{first}^t + x_{second}^t + x_{first}^t}{3}\right) \\ \frac{x_{first}^t + x_{second}^t + x_{first}^t}{3}, & f(x_{first}^t) > ff\left(\frac{x_{first}^t + x_{second}^t + x_{first}^t}{3}\right) \end{cases} \quad (18)$$

where $x_{first}^t, x_{second}^t$ and x_{first}^t show the most ideal students in the top three spots. By assigning comparable professors to the exceptional and average groups, researchers can hasten the convergence of the given GTOA.

4. Results and Discussion

The suggested approach is built on PyTorch 1.4, an open-source platform. Ubuntu 18.04 with an $\mathrm{Intel(R)Core\ (TM)\ i5-11600K}$ 3.9 processor is used for all experiments. Using the GTOA optimiser, all representations are trained and then tested on GeForce RTX 3080 Ti graphics card 12GB memory. The models have a weight decay of 0.0001, a learning rate initialised at 0.0001 that decays by 10% every 20 epochs, and an impetus β1 of 0.9. The batch size is set to 16, and the training procedure consists of 100 epochs.

4.1. Segmentation Analysis

4.1.1. Evaluation Index of Segmentation

Nuclear segmentation results are usually evaluated in medical histopathology using two indicators: (1) DICE is used to find out how near the actual sample is to the predicted model. The Aggregated Jaccard Index (AJI) is used to determine the ratio of the combined intersection to the union cardinality of the ground truth and the prediction. What follows are the DICE and AJI computations:

$$\mathrm{DICE} = \frac{2\times(X\cap Y)}{(|X|+|Y|)} \tag{19}$$

$$\mathrm{AJI} = \frac{\sum_{i=1}^{N}\left|G_i\cap P_M^i\right|}{\sum_{i=1}^{N}|G_i\cup P_M^i|+\sum_{F\in U}|P_F|} \tag{20}$$

in which X stands for the actual nucleus and Y for the expected one. A real nucleus has several nuclei denoted by N, P_M^i refers to the associated region that produces the nuclear mask, Gi mentions the connected area that P_M^i encompasses the area with the highest degree of ground-truth intersection; U is a set of entirely connected districts that do not intersect with ground truth; and PF is an element F inside U. While the first two evaluators assessed how well instance segmentation performed, they didn't verify that the predictions were based on the actual segmentation. This is accomplished by utilising the evaluation index PQ, which was initially suggested in our experiment and is calculated in the following way:

$$\mathrm{PQ} = \frac{|TP|}{|TP|+\left(\frac{1}{2}\right)|TP|+(1/2)|FN|}\times\frac{\sum_{(x,y)}\mathrm{IOU}(x,y)}{|TP|} \tag{21}$$

in which TP is the true positive, FP is the false positive, and FN is the false negative. In addition, x denotes the actual ground truth, and y denotes the anticipated segmentation area. The intersection-over-union (IoU) ratio is defined as the ratio of the intersection to the union. The detection quality is shown on the left side of the multiplier. The coldness between each positively detected instance and the ground truth it matches is shown on the right-hand side. By doing so, researchers can thoroughly and objectively assess the quality of segmentation and classification.

4.1.2. Performance Analysis of Segmentation

Table 1 shows how well four strategies perform on the Sicap dataset using measures such as DICE, AJI, DQ, SQ, and PQ. The proposed technique has the best overall ratings, beating GoogleNet, AlexNet, and U-Net.

Table 1: Segmentation on the Sicap dataset

Techniques	DICE	AJI	DQ	SQ	PQ
GoogleNet	0.794	0.505	0.544	0.725	0.397
AlexNet	0.780	0.513	0.560	0.709	0.415
U-Net	0.849	0.556	0.687	0.772	0.532
Proposed	0.856	0.558	0.690	0.777	0.546

Table 2 shows how the Radboud dataset was divided into groups based on several evaluation parameters. The Proposed technique consistently outperforms the other models, achieving higher DICE, AJI, DQ, SQ, and PQ scores.

Table 2: Radboud dataset

Techniques	DICE	AJI	DQ	SQ	PQ
GoogleNet	0.782	0.598	0.636	0.764	0.499
AlexNet	0.810	0.654	0.740	0.794	0.599
U-Net	0.818	0.651	0.757	0.783	0.609
Proposed	0.835	0.662	0.779	0.813	0.637

Table 3 compares the segmentation metrics of four approaches on the Karolinska dataset. All the strategies work about the same; however, the proposed model does a little better on most criteria.

Table 3: Karolinska dataset

Techniques	DICE	AJI	DQ	SQ	PQ
GoogleNet	0.849	0.556	0.687	0.772	0.532
AlexNet	0.849	0.558	0.685	0.774	0.531
U-Net	0.848	0.556	0.689	0.773	0.532
Proposed	0.849	0.557	0.690	0.772	0.533

4.2. Classification Analysis

After a second round of testing, the accuracy and specificity of all participants were averaged. Using these equations, researchers can evaluate how well the proposed model works:

$$\text{Recall} = \frac{\text{tp}}{\text{tp+fn}} \tag{22}$$

$$\text{Precision} = \frac{\text{tn}}{\text{tn+fp}} \tag{23}$$

$$\text{Accuracy} = \frac{\text{tp+tn}}{\text{tp+fp+tn+fn}} \tag{24}$$

$$\text{F} - \text{Measure} = \frac{\text{tp}}{\text{tp+1/2(fp+fn)}} \tag{25}$$

The correctly abnormal process is denoted by TN, the incorrectly normal by FN, the incorrectly classified by FP, and the correctly abnormal by TP (Table 4).

Table 4: Binary class results

Datasets	Precision	Recall	F-Measure	AUC	Accuracy
Sicap	0.812	0.860	0.835	0.867	0.84
Radboud	0.791	0.791	0.791	0.877	0.79
Karolinska	0.918	0.916	0.915	0.943	0.81

Table 5 shows the results of multi-class classification on three datasets. It includes metrics like accuracy, precision, recall, AUC, and F-measure. Radboud has the best overall performance across most evaluation metrics for the datasets.

Table 5: Multi-class results

Datasets	Accuracy	Precision	Recall	AUC	F-Measure
Sicap	83.87	0.8373	0.8405	0.9677	0.8379
Radboud	88.72	0.8885	0.8889	0.9775	0.8836
Karolinska	87.10	0.8675	0.8725	0.9742	0.8681

5. Conclusion

It is challenging to obtain pixel-level annotations in a real-world setting, which is necessary for accurate pattern delineation in whole-slide histopathology images. Disease diagnosis and treatment selection rely heavily on semantic segmentation of diagnostically relevant patterns. To achieve this goal, researchers present RMTL-Net, a framework for multi-task learning with regional attention that jointly performs both image tasks. With ResNet-101 as its foundation, the RMTL-Net can automatically learn weighted category-sensitive information from three images of the prostate: one each of the tumour, peritumoral, and background regions, thereby better depicting each image and improving segmentation and classification performance. Furthermore, the GTOA model is used to fine-tune hyperparameters, thereby improving segmentation and classification accuracy. Three publicly available datasets are used in the study to conduct extensive experiments. Over three datasets, RMTL-Net demonstrates superior performance compared to classification approaches and single-task segmentation methods. To throw light on the emerging field of MTL and its applications to prostate image segmentation and classification with our proposed RMTL-Net. By incorporating a classification branch, researchers can effortlessly transform any existing segmentation network architecture into its MTL equivalent, achieving both improved classification results and comparable segmentation performance at minimal cost. To improve segmentation and classification results, it is recommended to add an RA module that accounts for prior medical information on the significance of the tumour and peritumoral regions in the input images. You can modify the proposed RA module according to different prior knowledge and apply it to any existing MTL technique. While our evaluation focused on stained prostate cancer needle biopsies, the findings can be easily applied to a wide range of additional tissues, imaging techniques, and modalities.

Acknowledgement: The authors express their gratitude to New Horizon College of Engineering, Karunya Institute of Technology and Science, Saranathan College of Engineering, St. Joseph's College of Engineering, and Melbourne Institute of Technology for their academic support. Their collective guidance contributed to the successful completion of this research.

Data Availability Statement: Relevant data for this study can be obtained from the corresponding author upon reasonable request, subject to privacy and ethical considerations.

Funding Statement: This research and manuscript were carried out entirely through the authors' own efforts, with no external funding or financial assistance involved.

Conflicts of Interest Statement: The authors affirm that no competing interests influenced the development or outcomes of this research.

Ethics and Consent Statement: The study adhered to institutional ethical guidelines, with approval granted by the review board and informed consent obtained from all participants.

References

1. S. Iqbal, G. F. Siddiqui, A. Rehman, L. Hussain, T. Saba, and U. Tariq, "Prostate cancer detection using deep learning and traditional techniques," *IEEE Access*, vol. 9, no. 2, pp. 27085–27100, 2021.
2. L. M. da Silva, E. M. Pereira, P. G. Salles, R. Godrich, R. Ceballos, J. D. Kunz, A. Casson, J. Viret, S. Chandarlapaty, C. G. Ferreira, B. Ferrari, B. Rothrock, P. Raciti, V. Reuter, B. Dogdas, G. DeMuth, J. Sue, C. Kanan, L. Grady, T. J. Fuchs, and J. S. Reis-Filho, "Independent real-world application of a clinical-grade automated prostate cancer detection system," *The Journal of Pathology*, vol. 254, no. 2, pp. 147–158, 2021.
3. H. Van Poppel, M. J. Roobol, C. R. Chapple, J. W. F. Catto, J. N'Dow, J. Sønksen, A. Stenzl, and M. Wirth, "Prostate-specific antigen testing as part of a risk-adapted early detection strategy for prostate cancer: European Association of Urology position and recommendations for 2021," *Eur. Urol.*, vol. 80, no. 6, pp. 703–711, 2021.
4. F. -J. H. Drost, D. Osses, D. Nieboer, C. H. Bangma, E. W. Steyerberg, M. J. Roobol, and I. G. Schoots, "Prostate Magnetic Resonance Imaging, with or Without Magnetic Resonance Imaging-targeted Biopsy, and Systematic Biopsy for Detecting Prostate Cancer: A Cochrane Systematic Review and Meta-analysis," *Eur. Urol.*, vol. 77, no. 1, pp. 78–94, 2020.
5. S. Perincheri, A. W. Levi, R. Celli, P. Gershkovich, D. Rimm, J. S. Morrow, B. Rothrock, P. Raciti, D. Klimstra, and J. Sinard, "An independent assessment of an artificial intelligence system for prostate cancer detection shows strong diagnostic accuracy," *Modern Pathology*, vol. 34, no. 8, pp. 1588–1595, 2021.
6. C. De Vente, P. Vos, M. Hosseinzadeh, J. Pluim, and M. Veta, "Deep learning regression for prostate cancer detection and grading in bi-parametric MRI," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 2, pp. 374–383, 2020.

7. L. Duran-Lopez, J. P. Dominguez-Morales, A. F. Conde-Martin, S. Vicente-Diaz, and A. Linares-Barranco, "PROMETEO: A CNN-based computer-aided diagnosis system for WSI prostate cancer detection," *IEEE Access*, vol. 8, no. 7, pp. 128613–128628, 2020.
8. M. Hosseinzadeh, A. Saha, P. Brand, I. R. Slootweg, M. de Rooij, and H. Huisman, "Deep learning–assisted prostate cancer detection on bi-parametric MRI: Minimum training data size requirements and effect of prior knowledge," *Eur. Radiol.*, vol. 32, no. 4, pp. 2224-2234, 2022.
9. T. Tamada, A. Kido, A. Yamamoto, M. Takeuchi, Y. Miyaji, T. Moriya, and T. Sone, "Comparison of biparametric and multiparametric MRI for clinically significant prostate cancer detection with PI-RADS version 2.1," *J. Magn. Reson. Imaging*, vol. 53, no. 1, pp. 283–291, 2021.
10. A. Saha, M. Hosseinzadeh, and H. Huisman, "End-to-end prostate cancer detection in bpMRI via 3D CNNs: Effects of attention mechanisms, clinical priori and decoupled false positive reduction," *Med. Image Anal.*, vol. 73, no. 10, p. 102155, 2021.
11. I. Yusim, M. Krenawi, E. Mazor, V. Novack, and N. J. Mabjeesh, "The use of prostate specific antigen density to predict clinically significant prostate cancer," *Sci. Rep.*, vol. 10, no. 1, p. 20015, 2020.
12. M. Arif, I. G. Schoots, J. C. Tovar, C. H. Bangma, G. P. Krestin, M. J. Roobol, W. Niessen, and J. F. Veenland, "Clinically significant prostate cancer detection and segmentation using CNN on multiparametric MRI," *Eur. Radiol.*, vol. 30, no. 12, pp. 6582–6592, 2020.
13. R. Hao, K. Namdar, L. Liu, M. A. Haider, and F. Khalvati, "A comprehensive study of data augmentation strategies for prostate cancer detection in diffusion-weighted MRI using CNNs," *J. Digit. Imaging*, vol. 34, no. 4, pp. 862–876, 2021.
14. W. Lee, A. T. M. Lenferink, C. Otto, and H. L. Offerhaus, "Classifying Raman spectra of extracellular vesicles using CNNs for prostate cancer detection," *J. Raman Spectrosc.*, vol. 51, no. 2, pp. 293–300, 2020.
15. H. Van Poppel, R. Hogenhout, P. Albers, R. C. N. van den Bergh, J. O. Barentsz, and M. J. Roobol, "A European model for an organised risk-stratified early detection programme for prostate cancer," *Eur. Urol. Oncol.*, vol. 4, no. 5, pp. 731–739, 2021.
16. P. Pati, G. Jaume, Z. Ayadi, K. Thandiackal, B. Bozorgtabar, M. Gabrani, and O. Goksel, "Weakly supervised joint whole-slide segmentation and classification in prostate cancer," *Med Image Anal*, vol. 89, no. 10, p. 102915, 2023.
17. K. Liu, P. Li, M. Otikovs, X. Ning, L. Xia, X. Wang, L. Yang, F. Pan, Z. Zhang, G. Wu, H. Xie, Q. Bao, X. Zhou, and C. Liu, "Mutually communicated model based on multiparametric MRI for automated segmentation and classification of prostate cancer," *Med. Phys.*, vol. 50, no. 6, pp. 3445-3458, 2023.
18. X. Du, A. Shen, X. Wang, Z. Feng, and H. Deng, "NRD-Net: A noise-resistant distillation network for accurate diagnosis of prostate cancer with bi-parametric MRI images," *Multimedia Tools Appl.*, vol. 83, no. 9, pp. 33597–33614, 2024.
19. Y. Bouslimi, T. B. A. Gader, and A. K. Echi, "Prostate cancer detection, segmentation, and classification using deep neural networks," *in Proc. 12th Int. Conf. Pattern Recognit. Appl. Methods*, Lisbon, Portugal, 2023.
20. D. Singla, F. Cimen, and C. A. Narasimhulu, "Novel AI transformer U-NET for better identification and management of prostate cancer," *Mol. Cell. Biochem.*, vol. 478, no. 7, pp. 1439–1445, 2023.
21. Z. Li, J. Fang, R. Qiu, H. Gong, W. Zhang, L. Li, and J. Jiang, "CDA-Net: A contrastive deep adversarial model for prostate cancer segmentation in MRI images," *Biomed. Signal Process. Control*, vol. 83, no. 5, p. 104622, 2023.
22. R. Bygari, K. Rithesh, S. Ambesange, and S. G. Koolagudi, "Prostate cancer grading using multistage deep neural networks," In Machine Learning, Image Processing, Network Security and Data Sciences: Select Proceedings of 3rd International Conference on MIND 2021, *Springer Nature*, Singapore, 2023.
23. J. Silva-Rodríguez, A. Colomer, M. A. Sales, R. Molina, and V. Naranjo, "Going deeper through the Gleason scoring scale: An automatic end-to-end system for histology prostate grading and cribriform pattern detection," *Comput. Methods Programs Biomed.*, vol. 195, no. 10, p. 105637, 2020.
24. 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 deep-learning system for Gleason grading of prostate cancer using biopsies: A diagnostic study," *Lancet Oncol.*, vol. 21, no. 2, pp. 233–241, 2020.
25. W. Bulten, K. Kartasalo, P. -H. C. Chen, P. Ström, H. Pinckaers, K. Nagpal, Y. Cai, D. F. Steiner, H. van Boven, R. Vink, C. Hulsbergen-van de Kaa, J. van der Laak, M. B. Amin, A. J. Evans, T. van der Kwast, R. Allan, P. A. Humphrey, H. Grönberg, H. Samaratunga, B. Delahunt, T. Tsuzuki, T. Häkkinen, L. Egevad, M. Demkin, S. Dane, F. Tan, M. Valkonen, G. S. Corrado, L. Peng, C. H. Mermel, P. Ruusuvuori, G. Litjens, M. Eklund, and the PANDA challenge consortium, "Artificial intelligence for diagnosis and Gleason grading of prostate cancer: The PANDA challenge," *Nat. Med.*, vol. 28, no. 1, pp. 154–163, 2022.
26. P. Ström, K. Kartasalo, H. Olsson, L. Solorzano, B. Delahunt, D. M. Berney, D. G. Bostwick, A. J. Evans, D. J. Grignon, P. A. Humphrey, K. A. Iczkowski, J. G. Kench, G. Kristiansen, T. H. van der Kwast, K. R. M. Leite, J. K. McKenney, J. Oxley, C. -C. Pan, H. Samaratunga, J. R. Srigley, H. Takahashi, T. Tsuzuki, M. Varma, M. Zhou, J. Lindberg, C. Bergström, P. Ruusuvuori, C. Wählby, H. Grönberg, M. Rantalainen, L. Egevad, and M. Eklund,

"Pathologist-level grading of prostate biopsies with artificial intelligence," *arXiv*, 2019. Available: https://arxiv.org/abs/1907.01368 [Accessed by 14/11/2024].
27. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *in Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, Nevada, United States of America, 2016.
28. Y. Zhang and Z. Jin, "Group teaching optimization algorithm: A novel metaheuristic method for solving global optimization problems," *Expert Syst. Appl.*, vol. 148, no. 6, p. 113246, 2020.