Migration Letters

Volume: 21, No: S5 (2024), pp. 2138-2154

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)

www.migrationletters.com

Gleason Grading in Prostate Cancer Images: An Effective Segmentation Method and an Optimized Dense Convolutional Neural Network

Majed Aborokbah *

Abstract

To qualitatively characterize the various tumor histology observed in the prostate, recognized as the most prevalent and the second most lethal type of cancer in men globally, pathologists employ a range of screening procedures. The GG (Gleason-grade) classification of PCs (prostate cancers), which is based on photographs of the illness acquired via transrectal ultrasound imaging, is a significant instrument that is utilized in risk assessment as well as in the process of planning for patients. Subsequently, cancer-affected areas are discerned using Compactness Fuzzy C-means (CFCM). This method incorporates an adaptive processing approach rooted in the Least Mean Square (LMS) technique to determine the clip limit for CLAHE (Contrast Limited Adaptive Histogram Equalization) during the denoising and segmentation processes. Classification of PCs based on GG utilizing histological images is crucial for risk assessment and therapy planning, and an optimal deep model finally clas-sified the segmented images. The Dense Convolutional Deep Neural Network (DCDNN) architecture is utilized for multi-task prediction that uses the Modified Dunnock Search algorithm (MDSA) for optimal hyperparameter tuning of the CDCNN model, improving classification performance. This model has achieved the highest possible accuracy on both epithelial cell recognition and Gleason grading at the same time.

Keywords: *image processing, Prostate cancer, segmentation, preprocessing, classification, CLAHE, FCM, Deep neural network, dunnock search*

INTRODUCTION

PC is a widespread malignancy that impacts males on a global scale, holding the second position in terms of incidence and ranking sixth in terms of mortality [1]. Timely identification of this disease is of ut[1]most importance as it significantly impacts the efficacy of treatment and enhances the overall prognosis for affected individuals. The utilization of histological analysis on prostate tissue samples is a generally acknowledged approach to identifying and categorizing PC. In recent times, there have been notable improvements in the field of digital pathology and image processing techniques. These developments have brought about a significant transformation in the manner in which histological images are interpreted.

Consequently, there has been an improvement in the precision and efficiency of PC detection and catego-rization. The primary focus of this study is to explore the probable of integrating histological image analysis and machine learning (ML) techniques to improve the accuracy and effectiveness of PC diagnosis. Histological im-ages, which are acquired from biopsies or surgical specimens, offer comprehensive insights into cellular shape, tissue architecture, and other essential characteristics that might assist in the detection and classification of PC.

Numerous methods exist to evaluate the condition of the prostate, including digital rectal

Faculty of Computers and Information Technology University of Tabuk, Saudi Arabia.

examination and the application of the PSA (prostate-specific antigen) blood test. Various medical imaging techniques are available for prostate cancer identification and grading, with multiparametric MRI (magnetic resonance imaging) standing out as the predominant approach [2]. Furthermore, alternative imaging modalities have been employed [3]. Nevertheless, it is imperative to validate and assess the risk factors associated with malignancy through histological scoring of biopsy tissue by a skilled pathologist. In this procedure, the excised tissue is thinly sec-tioned and subsequently subjected to staining with H&E (hematoxylin and eosin) dyes. The pathologist conducts a microscopic examination of stained tissue sections to analyze the cellular and morphological patterns and to assess the aggressiveness of PC using the GG method.

Since its inception in 1966, the GG system has been revised in light of new information [4, 5, 6]. Notably, the GGs 1 and 2 secreters exhibit distinct and independent structures without any infiltration into surrounding tissues. GGs 3 is characterized by the presence of well-defined and discernible glands, exhibiting variability in size although appearing smaller and more densely arranged compared to benign prostate tissue. Gleason pattern four is categorized by the attendance of inadequately developed glands that exhibit fusion with neighboring glands, resulting in the absence of stromal separation. Additionally, this pattern may include formations known as cribriform structures. Gleason pattern 5 encompasses glands that exhibit a significant lack of differentiation. The Gleason score, as presented, represents the combined value of the two most prevalent patterns observed in the tissue sample. It has been shown that there is a strong link between the Gleason evaluation and clinical results, including metastasis and survival [7], and the lowest Gleason score presently assigned is 6. As a result, it is frequently used in the process of making decisions when deciding between continuing surveillance for low-risk diseases and various treatment selections for more severe forms of sickness.

The process of histopathology image grading by a pathologist, known as the GG system, is characterized by its time-intensive nature, significant cost implications, and susceptibility to substantial interobserver variability. In recent times, numerous endeavors have been undertaken to establish automated techniques for GG histo-pathological images. These strategies have the potential to enhance the efficiency and replicability of the out-comes. Generally, these methodologies operate by calculating a collection of characteristics from a patch or re-gion of interest inside the image. Typical instances of a zone of interest encompass various entities such as individual cells, glands, or nuclei. Certain proposed methodologies rely on an independent apparatus to effectively partition these regions of interest to calculate the necessary attributes. Several features have been identified as valuable in the categorization of prostate histopathological images into distinct GGs [8,9].

The use of CAD systems, which are designed to help doctors make more informed decisions in the clinic, has increased dramatically in recent years. Evaluation and forecasting of PCs rely heavily on the early identification of rapid recognition. The success of cancer diagnosis and treatment relies heavily on medical imaging. The speedy and successful detection of abnormalities in tissue findings constitutes a considerable challenge for physicians. The process of manual processing is characterized by a significant expenditure of time, a lack of cost-effectiveness, and the potential for treatment delays. ML algorithms are frequently employed in CAD applications that utilize medical imagery for cancer detection. Over the past span, there has been substantial advancement in the field of ML and DL (Deep learning) technology.

Moreover, this enhancement also has a positive impact on CAD applications. DL can acquire and comprehend complex visual patterns and attributes from images. The utilization of DL techniques has the potential to attain a significant level of detection accuracy without the need for manually designed features, as the process of feature extraction may be included in the training phase. Furthermore, the use of extensive parallel computing in recent times has led to the widespread adoption of DL methodologies in the field of PC detection and GG. In addition to introducing the DCDNN method, this work

aims to provide a thorough analysis of PC-detecting applications and GG. The following is a list of the study's primary objectives:

• To automate the identification of PC in histology images, this research aims to create DCDNN algorithms. The task at hand entails the detection and localization of malignant areas within the tissue specimens, a task that presents difficulties owing to the nuanced discrepancies in tissue morphology.
• The second objective of this proposed research is to classify PC based on the GG system, which is widely recognized as the standard method for evaluating the level of aggressiveness in PC after identifying malignant spots. The categorization of patients is crucial to ascertain the most suitable course of therapy and forecast po-tential outcomes for the individuals involved.
• Feature extraction involves the identification and extraction of pertinent features from histological images, including textural patterns, glandular structures, and nuclear properties. These extracted features serve as input for the DCDNN model. The inclusion of these features is crucial to facilitate the training of precise classification models.
• The focus of this study is on the development and training of DCDNN models to analyze histology images. The models above will acquire the ability to identify patterns and structures that are indicative of PC and its corresponding grade.
• Validation and clinical application involve the assessment of the created models using a varied dataset of histological images to ascertain their dependability and applicability in real-world scenarios. Moreover, there is a need to investigate the potential therapeutic utility of these tools as a means of aiding pathologists in their diagnostic processes.

The succeeding sections of the article are structured as follows: The "Literature Review" section provides a comprehensive analysis of the existing literature about the classification of medical images. The "Proposed Methodology" section delineates the experimental design. The section titled "Experimental Results" provides an account of the outcomes obtained from the conducted experiments. In contrast, the subsequent section titled "Discussion" engages in an analysis and interpretation of these data. Ultimately, the present study culminates in the formulation of a conclusive statement, accompanied by an exposition of potential avenues for future research. Next comes a massive list of sources used throughout the paper.

2. RELEATED WORK

Exploration into the limitations of current methods for detecting and classifying prostate cancers has been undertaken. Li et al. [10] pioneered the development of a DL model aimed at enhancing prostate cancer diagnostic capabilities by utilizing multiparametric MRI (mpMRI) and comprehensive whole-mount histopathology data. The DL model, derived from whole-mount histology, integrated distinct segmentation and classification networks to delineate the thyroid gland and identify areas affected by prostate cancer for diagnostic purposes. The performance of the prostate classification networks was evaluated using the AUC. The prostate identification network achieved an AUC of 0.871 on the validation dataset and 0.797 on the test dataset when employing the DL model. When applied to the validation set, the DL model achieved a sensitivity is 0.710, a specificity is 0.690, a precision of 0.696, and an accuracy of 0.700 for diagnosing PCa. The foundation of the system is made up of three preprocessing modules that were trained separately and still needed pixel-wise annotations.

Hassan et al. [11] offer a unique automated classification system for detecting PCs from US and MRI images by merging several DL approaches. Furthermore, the suggested system provides justifications for each decision made in light of the supplied US or MRI image. Multiple DL models, each with its own set of custom-built layers, are applied to the datasets after being pre-trained. The optimal model achieves a best-case accuracy of 97% on the test set's US images and 80% on the set's MRI images. Tissue borders are a common site for incorrect detection.
The RMANet (Robust Multi-modal Feature Autoencoder Attention net) was proposed by

Li et al. [12] and is a unique and actual multi-modal CNN for identifying the clinical severity grade of PC. Two branches of the model are used to learn from the T2- and Diffusion-weighted MRI data: one employs a ten-layer CNN (convolutional neural network) with two input weights that are shared to acquire the global features of the two different modalities, and the other utilizes an auto-encoder framework with a classical U-net as its backbone to acquire the distinctive features of each modality and reimburse for the absence of data for the most serious cases.

Two branches of the model are used to learn from the T2- and Diffusion-weighted MRI data: one employs a ten-layer CNN. In the PANDA challenge, the largest histopathological challenge to date, Kartasalo et al. [13] used 10,616 digitalized prostate samples to speed up the creation of repeatable AI algorithms for GG. While the pathologists remained in the dark, the submitted procedures were tested on separate cross-continental cohorts. The algorithms demonstrated a concordance of 0.862 (quadratically balanced, 95% CI, 0.840-0.884) and 0.868 (95% CI, 0.835-0.900) with human uropathologists on externally validated sets from the United States and Europe, respectively.

Ayyad et al. [14] provided examples of PC histology slides. However, large-scale studies analyzing PC histopathology images are lacking. This work provides a thorough analysis of the research on histopathological images for diagnosing PC. The difficulties of preparing histopathology images are introduced at the outset of the survey. Also, briefly go over some of the common computational techniques used in image processing, catego-rization, selection of features, and labelling that can aid in the identification of PC in histopathology images. But at each iteration, the model was practiced on a fresh set of mini-patches.

The model utilizes two branches to learn from T2- and Diffusion-weighted MRI data. In a study by Han et al. [15], one branch employs a ten-layer CNN. In this study, tissue component maps (TCMs) were generated from the images, with each pixel labeled as "lumina," "nuclei," or "other." Whole-mount RP tissue slices were analyzed using seven different ML approaches for cancer detection and classification. Three non-DL classifiers employed features collected from TCMs, whereas the transfer learning technique was used by four DL algorithms on TCMs, luminance maps, nucleus maps, and raw pictures. They needed to incorporate more features into their feature extraction procedure to improve classification accuracy.

After testing their DL-based method for GG of prostatic adenocarcinomas in multiple well-characterized validation cohorts, Tolkach et al. [16] concluded that it achieved human-level performance in the prognostic classification of patients. In addition, the optimum minimal tumour size for reliable GG of the total tumour focus was identified (actual size of roughly 560 560 m). The strategy is

implemented in the unified digital pathology pipeline, which provides all the important tumour metrics for a pathology report. The model was tested on pho-tographs that may have varied appearances after training it on a small set of images.

For gland-oriented segmentation and classification, Gurav et al. [17] used the CS (Color Space) transfor-mation and the SSA-RideNN (Salp Swarm Optimization Algorithm-based Rider Neural Network) technique. Maximum significant regions are extracted as features from the gland region utilizing MK-SIFT (multiple-kernel scale-invariant feature transform) for cancer detection. Here, SRA (Salp-Rider Algorithm), an integration of SSA and ROA (Rider Optimization Algorithm), was proposed to train the RideNN classifier to its fullest potential. Experimental results employing histopathology images show that the presented strategy for detecting PC achieved the highest possible levels of accuracy (0.8966), sensitivity (0.8919), and specificity (0.8596). Although it takes less time, its sensitivity, accuracy, and specificity are all capped at 90% at most.
To do this, Karimi et al. [18] merged the results of three CNNs trained on varying patch sizes. Because of this, the approach was able to make use of both the increased quantity of

smaller patches present in the labelled training data and the greater contextual information present in bigger patches. A logistic regression model, trained inde-pendently after the CNN training, is then used to aggregate the predictions from the three CNNs. Novel data augmentation strategies were devised to enhance the training process, and their effects on classification precision were experimentally explored. When comparing malignant and benign patches, the suggested technique has a 92% success rate, and when separating low-grade (i.e., GG 3) from high-grade (i.e., GGs 4 and 5), it has an 86% success rate. Its effectiveness was measured solely by how it affected the functioning of the ensemble framework.

Two branches of the model are used to learn from the T2- and Diffusion-weighted MRI data: one employs a ten-layer CNN. Li et al. [19] devised a unique Region-Based Convolutional Neural Network (R-CNN) design for multi-task prediction, amalgamating an Epithelial Network Head with a Grading Network Head. This multi-task strategy holds the potential to surpass single-task models by offering more comprehensive contextual information. Simultaneously, it outperformed state-of-the-art systems in detecting epithelial cells and assigning Gleason Grades (GGs). The model demonstrated an average Area Under the Curve (AUC) of 0.998 and a detection accuracy of 99.07% in epithelial cell testing through five-fold cross-validation. For GG, the model achieved an overall pixel accuracy of 89.40% and a mean intersection over union of 79.56%. Due to a lack of patient-level data, a more thorough patient-level stratification was not feasible. This study introduces a deep Convolutional Neural Network (CNN)-based model for the most precise detection and classification of prostate cancers, effectively addressing the aforementioned constraints.

3. **MATERIAL AND METHODS**

Using a DCDNN trained with an MDSA model, this section covered PCa detection and classification. Fig 1 de-picts the overall architecture diagram. Two branches of the model are used to learn from the T2- and Diffu-sion-weighted MRI data: one employs a ten-layer CNN. At first, the input image is pre-processed using adaptive CLAHE. Here, LMS chooses an adaptive clip limit for CLAHE, which is the only

cause of the noise reduction throughout image enhancement, and uses it until the most suitable (CDF) depiction of the image very nearly converges to a straight line. Standard metrics are used to determine the optimal window size for this strategy. After achieving high accuracy in CFCM segmentation of the Prostate region, a DCDNN model is used to determine if the resulting image is malignant or normal. Time spent on calculations is cut down by using the MDSA to determine the hyperparameter of DCDNN. At last, the outcomes of the performance tests are analysed and compared to current PCa categorization methods.

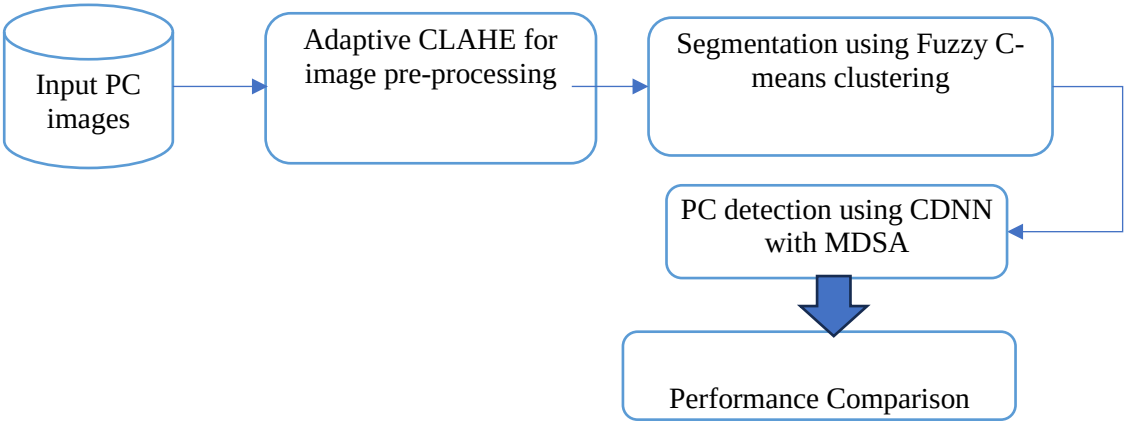

Fig 1: Architecture diagram of DCDNN with MDSA

3.1 Pre-processing Using Adaptive CLAHE

The proposed procedure has two stages. Selecting the ideal window size based on the image's AMBE (absolute bright mean square error) [20] and PSNR (peak signal to noise ratio) [21] characteristics requires first estimating the ideal clip limit adaptively. The next step is to apply CLAHE based on the calculated clip boundary. The LMS method is used in Adaptive CLAHE to estimate the clip limit, although it has many other applications besides just picking adaptive variables. Inconsistencies in the image capture process, as well as variations in contrast distribution and dynamic range, do not affect the LMS algorithm. The procedure for estimating the clip limit is depicted in a diagrammatical form in Fig. 2.

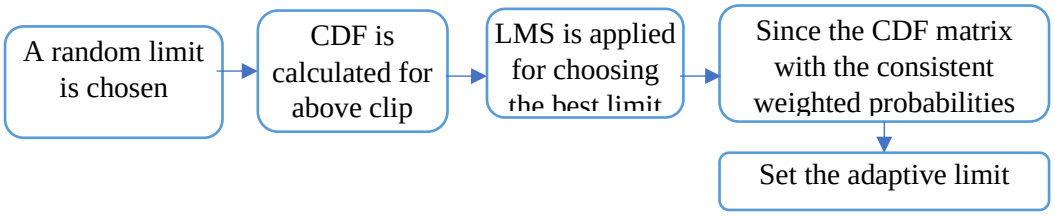

Fig 2: Adaptive CLAHE processing diagram

The histology image's Histogram Distribution (HD) is first determined using the formula $HD(a) = \{fr_1, fr_2, fr_3, \ldots\ldots\ldots fr_{IL-1}\}$. For an image of IL intensity, HD(a) is a list of all the frequencies present in the image from fr_1 to fr_{IL-1}. Probability of occurrence, or $Pr_i(k)$, for each passion level, k is then determined as follows:

$$Pr_i(k) = r.fr_k/N(0 < k < IL) \quad (1)$$

Where N is the entire number of images' pixels, with the pixel distribution computed using traditional CLAHE, the CDF is then calculated using a random clip limit (r).

$$CDF_i(k) = (IL - 1)\sum_{i=0}^{k} pr_i(k) \quad (2)$$

In addition, the distribution matrix from 0 to IL-1, denoted by the symbol A_k, can be constructed using the CDF obtained from the equation above. Currently, the LMS method is employed to determine the optimal weight Wt^* for a given set of inputs (DO_k) and error (er_ke). The linear fluctuation of the content's matrix, DO_k, now uniformly ranges from 0 to IL-1. Where A_k is the input matrix used to get the best-fit value for Wt^*. Since the nonlinear content of the CDF must be removed in this methodology, the er_k is considered to be a continuous variable throughout. The Weiner weight, denoted by Wt_o, is defined as $R^{-1}P$. In this case, $P = DO_kA_k$ and $R = A^kA_k^T$. The weight Wt_{k+1} of the k+1 iteration can be found using the steepest descent [22] technique, which is as follows:

$$Wt_{k+1} = Wt_k - \mu\nabla_k \quad (3)$$

Where Wt_k is the weight for iteration k, μ is the gain constant set to 0.01, and ∇_k is the image gradient estimate. Gradient estimation calculation is unnecessary here since

$$\nabla_k = 2\mu er_kA_k \ . \quad (4)$$

$$Wt_{k+1} = Wt_k + 2\mu er_kA_k \quad (5)$$

The variable's expected value is represented by the operator Ex

$$Ex[\nabla_k] = -2Ex[er_kA_k] = 2Ex[DO_kA_k - A_kA_kA_k^TWt = (2\,RW - P) = \nabla \quad (6)$$

$$Ex[Wt_{k+1}] = Ex[Wt_k] + 2\mu Ex[er_kA_k] = Ex[Wt_k] + \mu(Ex[DO_kA_k - Ex[A_kA_K^TW_{tk} =$$

$$Ex[Wt_k + 2\mu(P - REx[Wt_k] = (I\ -\ 2\mu R)Ex[Wt_k] + 2\mu R20.Wt^* \quad (7)$$

$$W^* = 5.0\mu^{-1}R^{-1}(Ex[W_{k+1}] - (I - 2\mu R)Ex[Wt_k] \quad (8)$$

The optimal body mass is calculated using equation (9). The new CDF is then the best-fit line, which is found by calculating AA_kW^*. Currently, the clip limit is drawn using A_kW^*on the previous CDF plot, and the weighted probability of each pixel intensity is traced back to the new CDF. The following formula can be used to derive the revised probability of recurrence of the ith pixel intensity value from this information ($Pr_i^l(k)$):

$$Pr_i^I(k) = pr(xa) - Pr_{i-1}^I(k) \qquad 1 < i < IL - 1 \quad (9)$$

$$Pr_0^I(k) = Pr(0) \quad (10)$$

Two branches of the model are used to learn from the T2- and Diffusion-weighted MRI data: one employs a ten-layer CNN Maximum for the probability of getting an intensity degree from (10), and (11) is used to estimate the new clip boundary for a particular window size. The next part discusses CFCM-based segmentation, which is performed after pre-processing.

3.2 Segmentation using CFCM

In this subsection, a novel algorithm called CFCM is proposed for prostate segmentation. There are two problems with the FCM algorithm which include: The first is that it is extremely sensitive to the choice of the number of clusters (c) and the starting membership matrix (M0). Because of these two drawbacks, FCM has a difficult time determining the optimal number of clusters, and its output is highly volatile. Assuming that the initials are closer to that of actual "cluster centres" is a general requirement for a successful initial cluster centre in partition-based clustering. The goal of CFCM is to produce more stable and accurate clustering results in a variety of data clustering applications by leveraging compactness-based initialization and outlier robustness methods.
It is common practice in density-based clustering to first calculate the density of each sample (here, a histologically pre-processed image designated by a_i, whose value is ρ_i).

The proposed model then ranks the samples based on their decreasing densities. $As = \{a_i\}, i = 1, 2, \cdots, n$ represents the sorted samples, and the related densities are given by $\{\widetilde{\rho_{\iota_i}}\}, i = 1, 2, \cdots, n$. A density threshold dr_c defined as

$$r_c = (\sum_{i=1}^{n} \rho_i) * rd, \quad (11)$$

where density rate rd is a quantity between zero and one. Here, the technique picks the possible cluster hubs with densities higher than the mean. Specifically, CC_pDefines the set of possible cluster centres, and the formula is:

$$CC_p = \{\tilde{a}_i | \sum_{i=1}^{p} \tilde{\rho}_i < dr_c, \sum_{i=1}^{p} \tilde{\rho}_i \geq dr_c, i = 1, 2, \cdots, p\}. \quad (12)$$

It can be seen from (12) and (13) that the values of dr_c, which determines the number of possible cluster centres, are affected by the parameter rd. The parameter rd determines how many candidates for cluster centres are ultimately chosen. The final step is to determine the total number of clusters (c) and then pick the best candidates for initial cluster centres. There needs to be some space between the centres of each cluster at the outset to prevent two samples from the same cluster from being picked using equation. (14)

$$th = \frac{1}{p}(\sum_{i,j=1}^{p} \widetilde{ds}_{i\,j}) * rd, \quad (13)$$

where $\widetilde{ds}_{i\,j}$ is the distance between $\tilde{a}_i$and $\tilde{a}_j$. In this case, use $\tilde{a}_1$ as the starting point for the cluster and $\tilde{a}_2$ as a reference point to determine the distance. If $\widetilde{ds}_{12} > \delta$, then $\tilde{a}_2$ is the initial class centre; otherwise, $\tilde{a}_2$ is not. The sample $\tilde{a}_3$ serves as the first cluster centre if the distances computed are all greater than the threshold value (th). The p-th sample is

compared to the first p-1 samples to establish whether or not it is the primary cluster's nucleus. Assuming c samples meeting the criteria mentioned above are located in the centres of the p possible clusters, this approach can achieve the goal of estimating the value of c. This method uses the obtained initial cluster centres and the density of each sample to build the initial membership matrix, which is a reasonable starting point without the need for complex math. M. restricts samples to the most centrally dense cluster. The corresponding $\tilde{u}_{ij}$ is defined as follows: where a_i is a sample from the dataset and A_s is the dataset whose samples are in descending order through local densities.

$$\widetilde{M}_{i\,j} = \begin{cases} 1, & if\tilde{a}_i \in V \; and \; j = k \\ 0, & if\tilde{a}_i \in V \; and \; j \neq k \\ \frac{\tilde{\rho}_i/\rho_{v_1}}{k-1}, & if\tilde{a}_i \notin V, \rho_{CC_k} \leq \tilde{\rho}_i \leq \rho_{CC_{k-1}} and \; j \leq k-1 \\ \frac{1-\tilde{\rho}_i/\rho_{v_1}}{c-(k-1)}, & if j > k-1 \; and \; \tilde{a}_i \notin V, \rho_{CC_k} \leq \tilde{\rho}_i \leq \rho_{CC_{k-1}} \end{cases} \quad (14)$$

where $i = 1, 2, \cdots, n, k = 1, 2, \cdots, c$ and ρ_{CCk} is the density of the cluster center CC_k and suppose $\rho_{CCk-1} \geq \rho_{CCk}, k = 1, 2, \ldots, c$. In this study, the density ρ_i [3] of a sample a_i defined as

$$\rho_i = \sum_j \chi\,(ds_{i\,j} - ds_c), \quad (15)$$

where $\chi(a) = \begin{cases} 1 & if\ a < 0 \\ 0 & otherwise \end{cases}$, ds_c is a cutoff distance and $ds_{ij} = ds(a_i, a)$ is the Euclidean distance between sample a_i and image sample a_j. To put it simply, ρ_i is equal to the number of samples that are within a distance of ds_c of image sample a_i. The CFCM algorithm 1 can be described as follows

Algorithm 1: CFCM-based prostate segmentation Algorithm
Input: Dataset A, density rate rd, distance rate rd, and fuzziness index m
Output: Final membership matrix $M^{(t+1)}$ and cluster centres $CC^{(t+1)}$.

1. Put t = 0 in the iteration counter.
2. Calculate the neighbourhood density. ρ_i for each sample a_i using the formula (16).
3. Arrange the density samples and their corresponding densities from highest to lowest or from $\tilde{\rho}_i$ to $\{\tilde{a}_i\}$.
4. Determine rd_c and CC_p using equations (4) and (5).
5. Let $\tilde{a}_1$ be the starting point of the cluster, and compute the using equation (14).
6. For i = 2: p,
7. Find $\tilde{d}s_{1i}, \widetilde{ds}_{2i}, \cdots, \tilde{d}s_{(i-1)i}$.
8. If $min\{d\widetilde{ds}_{1i}, \widetilde{ds}_{2i}, \cdots, \widetilde{ds}_{(i-1)i} > th$,
9. An initial cluster center is denoted by $\tilde{a}_i$, and additional centres are obtained in the form of $CC_1, CC_2, \cdots, CC_c$.
10. Replace $M^{(0)}$ with Equation.(15) in the membership matrix M.
11. Using Equation.(3), the updated cluster centres $CC^{(t+1)}$determined.
12. Calculate the new membership matrix $M^{(t+1)}$.
13. If $CC^{(t+1)} = CC^{(t)}$ (or $M^{(t+1)}=M^{(t)}$) then stop the process.
14. Else, proceed to step 16 and set t = t + 1.
15. Return the result of $M^{(t+1)}$ and $CC^{(t+1)}$
16. End if
17. End for

3.3 PCa Prediction and Classification Using DCDNN with MDSA

At this stage, the suggested classifier is used to categorize PCa cases into either of two histological subtypes. The appropriate hyperparameters of the DDCNN approach are

chosen with the use of an MDSA-based optimization algorithm to raise the classification accuracy. Extensive explanations of the DDCNN and MDSA algorithms are provided below.

Dense Convolutional Deep Neural Network for Classification

As a form of FFNN (Feed Forward Neural Network) with a unified design between the neurons, DCDNN stands out even among the DL techniques. It detects PCa through neurons that share districts with neurons that are otherwise distinct. The MDSA technique is also used to optimize the DCDNN framework. This model is used to identify PCa in histology images, and its design features a sophisticated architecture of stacked layers. When it comes to extracting a feature from an image, DCDNN is both sensitive and sturdy. The DCDNN model comprises standard components, including an input and output layer, a compressed layer, two convolutional layers, and a fully connected layer. The initial layer is recognized as the input layer, followed by the application of Rectified Linear Units (ReLu) and Batch Normalization (BN) layers within the convolutional layer. To extract features, filters in the convolution layer "travel" along the time axis. Features from the time dimension can be gleaned using the third layer and the horizontal axis [23], and the fourth layer, called the compressed layer, is represented by the C3 vector. Following the convolutional layers, the output classes are classified using a fully connected layer along with a Softmax layer. This combination is employed to differentiate between the classes associated with prostate cancer in the image. In DDCNN, a neuron is revealed as

$$n(CL, NF, PF) \tag{16}$$

Multiple position feature maps (PF), a large number of feature maps (NF), and a large number of NN layers (CL). To define the neuron's input ($A_{NF}^{CL}(Pf)$) and output ($B_{NF}^{CL}(PF)$), the following equation is used:

$$B_{NF}^{CL}(PF) = F(A) = FB_{NF}^{CL}(PF) \tag{17}$$

The AF (activation function), can be expressed as

$$AF(A) = ln(1 + ea) \tag{18}$$

The source image is labelled as,

$$A_{IJ}^{1}(PF)(N_C@N_{FR} * N_T) \tag{19}$$

This equation is the output of the second convolution layer.

$$B_K^2 = F(A) = F\left(\sum_{I=1}^{N} A_{IJ}^{1} * wf_k^2 + bi_k^2(PF)\right) k = 1, 2, 3 \ldots . N \tag{20}$$

In this context, B_K^2 represents the output layer of C3. The term bi_k^2 refers to the biases associated with this layer, while wf_k^2represents the filter that is horizontally combined with the FV (feature vector) to get the desired vectors. The regularization of these vectors can be achieved by incorporating a BN layer before the transfer of the input to the activation layer. Following this, the 4th input layer is merged with a vector, and the 5th output layer is calculated subsequently.

$$B^5 = F\left(\sum_{I=1}^{N} B^4 wt_I^5 + bi^5(PF)\right) \tag{21}$$

In this context, the notation wt_I^5represents the weights associated with the fifth layer, while bi^5represents the biases of the 5th layer. The Softmax layer, which constitutes the sixth layer of the neural network, is composed of two neurons. The output of the sixth layer is determined mathematically in the following manner.

$$B^6 = F(\sum_{I=1}^{N} B^5(I) wt^6(I) + bi^6(PF)) \tag{22}$$

The forward propagation calculation flow of the DCDNN network is persistent based on the equations provided above. The neural network is trained using a labelled training dataset, as well as the discrepancy between the location value and the forecast value. The weight and bias parameters of the DCDNN network undergo updates through the gradient descent optimization algorithm, as illustrated in the following manner:

$$wt^K = wt^K - \frac{\partial E}{\partial wt^K} \tag{23}$$

$$bi^K = bi^K - \eta \frac{\partial E}{\partial bi^K} \tag{24}$$

During the training stage of DCDNN, the minimum error rate is calculated and mitigated, and a notable absence of a pooling layer is observed in the network design. The exclusion of a pooling layer leads to a streamlined architecture that avoids incorporating suboptimal features. FVs are transmitted to the input layer of the ReLU, allowing each FV to demonstrate consistency. In the fourth layer, the time domain FV and frequency FV are amalgamated before being input into the fully connected layer. The Softmax layer, commonly used for classification tasks, is then employed. The utilization of the MDSA algorithm facilitates the optimal selection of design parameters for DDCNNs. In the subsequent part, the fundamental behaviour and characteristics of MDSA will be delineated before delving into its explanation.

Modified Dunnock Search Algorithm

Individuals can be categorized as either discoverers, members, or alertness in the MDSA algorithm. Food (i.e., optimal hyperparameter) and population search (i.e., weight values) are both the responsibility of the discoverer. The letter is aware of environmental dangers and informs the dunnock population to relocate to a safer region, and the followers follow the discoverer to obtain food. If the finder has a higher fitness than the participants, then the dunnock population as a whole will be more fit. If the finder has lower fitness than the participants, then dunnocks with lower energy levels will fly elsewhere to increase

their energy levels. When a warner in a group of dunnocks detects a threat to their surroundings, they raise the alarm, and if the value of the alarm is high enough, the finder leaves the area and leads the group to safety.

Assume that N is the total number of dunnocks, and let $P_i^{ti} = \left(P_{i,1}^{ti}, P_{i,2}^{ti}, P_{i,3}^{ti}, \ldots\ldots, P_{i,d}^{ti}, \ldots\ldots, P_{i,Di}^{ti}\right), Di$ be the position of the ith bird at time ti and where P is the population. In a D-dimensional search space, the entire dunnock population A can be written as

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1d} & \cdots & p_{1Di} \\ p_{21} & p_{22} & \cdots & p_{2d} & \cdots & p_{2Di} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{Nd} & \cdots & p_{NDi} \end{bmatrix} \tag{25}$$

the discoverer failed to identify the danger, and they were accountable for instructing the community to forage and search thoroughly. When certain members of the population detect danger ($\sigma < \delta$) and sound the alarm ($> \delta$), the rest of the population is led to the secure zone. The location update is described as follows:

$$P_{i,j}^{ti+1} = \begin{cases} P_{i,j}^{ti} \cdot exp\left(\frac{-i}{\alpha \cdot M}\right), & \sigma < \delta, \\ P_{i,j}^{ti} + Qc.\mathrm{Lm}, & \sigma > \delta, \end{cases} \tag{26}$$

where α is a random number between 0 and 1, $\sigma \in [0,1]$ is the early warning value, and is the current-environment security threshold. Q regulates the step size, which is a normally distributed random number. Lm is a one-dimensional matrix of $1 \times d$ elements, each of which is 1. Participants are divided into two groups: those who accompany and monitor the

discoverer as they gather food (where i is less than or equal to n/2), and those who forage for food independently (where i is greater than n/2). Consequently, the location update description for the participants is delineated as follows:

$$P_{i,j}^{ti+1} = \begin{cases} Qc.\exp\left(\frac{P_{wt}^{ti}}{i^2}\right), & i > \frac{n}{2} \\ P_{Pos}^{ti+1} + \left|P_{i,j}^{ti} - P_{pos}^{ti+1}\right|.P^{+} > Lm, \quad otherwise, \end{cases} \tag{27}$$

where A_{wt} is the worst possible position in the current population and A_p is the best possible position that the discoverer currently holds. Since A controls the direction of the $1 \times d$ matrix, $P^{+} = P^{T}(PP^{T})^{-1}$ (1 or -1) has only two possible values, 1 and -1. When the dunnock population realizes the threat (when the parameter value is ineffective), the birds take measures to protect themselves from predators. The dunnock has to move closer to the population centre To reduce the risk of predation when $fe_i \neq fe_g$. It would seem from this that the dunnock is now at the periphery of the population and is aware of the threat. When $fe_i = fe_g$, the dunnock in the intermediate of the flock is aware of the threat and must leave its current location. The location update explanation of the alertness is as follows:

$$P_{i,j}^{ti+1} = \begin{cases} P_{bt}^{ti} + \beta.\left|P_{i,j}^{ti} - P_{bt}^{ti}\right| & fe_i \neq fe_g, \\ P_{i,j}^{ti} + k.\left(\frac{\left|P_{i,j}^{ti} - P_{wt}\right|}{(fe_i - fe_{wt}) + \varepsilon}\right), & fe_i = fe_g \end{cases} \tag{28}$$

The optimal population density is denoted by A_{bt} and the β value determines the step size, which is a normally distributed random variable. Both the direction and pace of the dunnock's flight are determined by K. The number was chosen at random. As the fitness value of the ith person, fe_g, fe_i , and fe_w stands for the greatest, current, and worst population fitness values, respectively. There needs to be at least one positive real number ε to prevent having a null denominator.

Late in the iteration process, a local optimum will inevitably be reached, leading to inaccurate convergence [24]. A refined algorithm was formulated by incorporating a chaotic adaptive inertia weight and an improved boundary restriction, resulting in a substantial enhancement of the algorithm's performance. The capability of the algorithm to identify the best solution depends critically on the quality of the original solution. Ergodicity, underlying regularity, and long-term unpredictability are hallmarks of chaos, a type of random phenomenon [25]. Because they can move through all possible stages in an optimization algorithm's population without ever repeating themselves, chaotic sequences are frequently utilized in the population initialization process. Its function is expressed as

$$\begin{cases} \mu(ti+1) = 3.5\mu(ti)^2 + 3.3\mu(ti)^2 - 0.265 \\ \bar{P}(ti+1) = \frac{1}{\pi}\arcsin\left(-\frac{7}{4}\mu(ti+1) - \frac{33}{40}\right) \end{cases} \tag{29}$$

Here, μ ∈ [-73/70, 1/10] represents the initialization arrangement, and $\overline{P}$ is entirely chaotic within the range [-1/2, 1/2]. The procedure for transforming the chaotic sequence $\overline{P}$ into the resultant space is expressed as follows:

$$P = \frac{Upb + Lwb}{2} + \bar{A}(Upb - Lwb) \tag{30}$$

In this context, Upb and Lwb denote the upper and lower boundary values of the optimized variables, respectively. The pseudocode of DCDNN is predicted as follows algorithm 2. Thew overall flow chart is shown in fig 3.

Algorithm 2: hyperparameter optimal value selection in DCDNN model
Input: N, objective function dimension Di; maximum number of iterations ti; upper and lower bounds Upb, Lwb; and safety threshold δ.

Output: optimal weight value selection (i.e., optimal fitness)

1. Set the population in MDSA.
2. Find each person's fitness level (denoted by fe_i) and rank them accordingly; then, label the best fitness level (denoted by fe_g) and its corresponding position (A_{best}); similarly, label the worst fitness level (denoted by fe_{wt}) and its corresponding position (A_{wt}).
3. By informing the discoverer's position using Eq.(27) and recording the optimal position A_P currently held, the individual selected with the pre-discoverer proportion PN*N of fitness value as creator.
4. Chaos is determined by using equation (30).
5. Select the long-term members as agreements, and revise the order of the new members following equation (28).
6. The location of the alerters was modernized according to formula (4), and the individuals with alerted proportion SN*N were chosen at random.
7. Update the dunnock's location and the fitness values (fe_g, A_{best}, fe_{wt}, and A_{worst}) using the updated data.
8. If the criterion for output is met, the cycle finishes and the result is outputted; otherwise, steps 2 through 8 are repeated.
9. Selecting for maximum fitness.

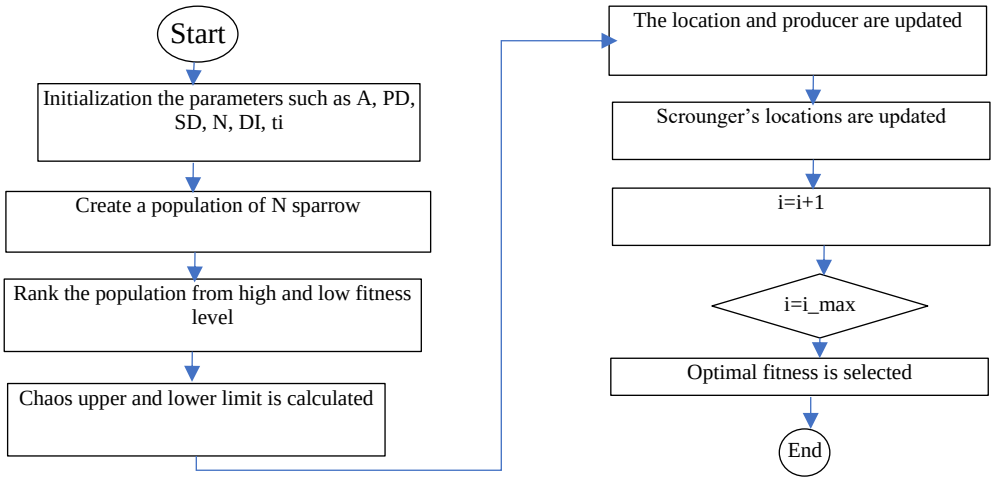

Fig 3: Flowchart of MDSA algorithm for optimal parameter selection

4. RESULTS AND DISCUSSION

Here, the effectiveness of the proposed DCDNN model PCa detection and classification compared to that of R-CNN [19], CNN [18], and DL [10]. Precision, sensitivity, F1-score, and accuracy are some of the metrics used to assess how well a test or procedure is performed.

Dataset Description

The study population comprised 102 patients who underwent radical prostatectomy at Radboudumc between 2006 and 2011 (IRB number 2016-2275). Adjuvant-treated patients were not eligible for surgery. Based on the pathologist's initial Gleason report, one formalin-fixed, paraffin-embedded tissue block was chosen from each prostatectomy. The GG group for each block was calculated from the reported grades. Since a single tissue block may contain specimens of varying grades, it is necessary to record each instance of each grade separately. In 24% of tissue blocks, grade 2 was present; in 69%, grade 3 was; in 63%, grade 4 was; and in 33%, grade 5 was present. High-grade tumours (grades 4 and 5) are overrepresented in this study because of the study's use of a more selective selection strategy. The ability to investigate DCDNN-based cancer segmentation algorithms'

performance in the challenging scenario of high-grade PCa was made possible by the generous size of the sample. The data collected for this investigation may be found at https://doi.org/10.5281/zenodo.1485967 in the Zenodo repository. Fig 4 shows some example photos, while Fig 5 contrasts normal and tissue images.

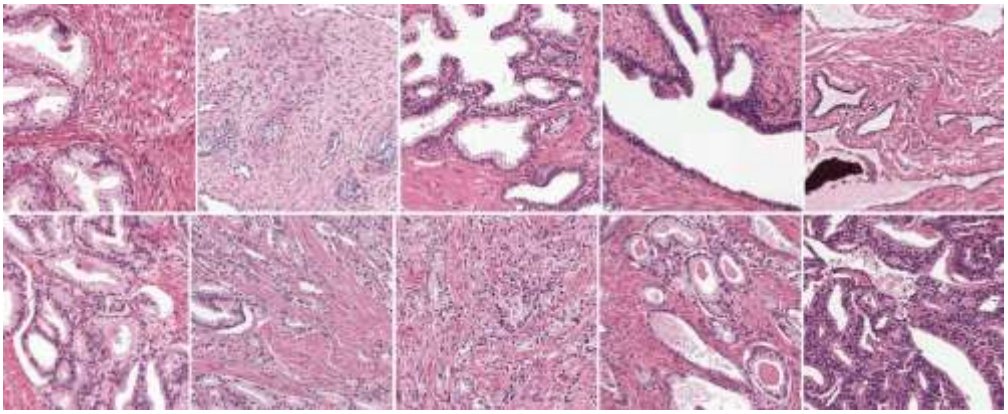

Fig. 4. A spectrum of GGs, from (A) benign to (H) badly formed glands and solitary cells, as well as (B) highly-formed glands to (C) badly formed glands to (D) cribriform glands**.**

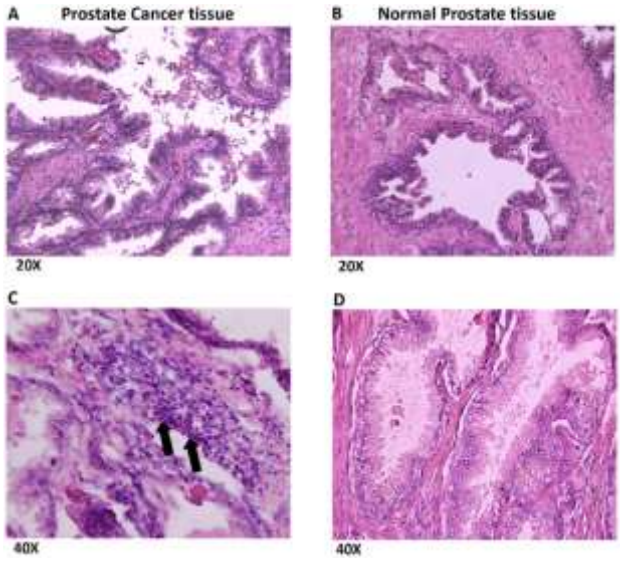

Fig. 5. PC tissue (left) vs Normal Prostate tissue (right)

As outlined in the dataset description, the quantity of images and, consequently, the number of patches available for training were similar for both high and low Gleason Grade (GG) classes. In the absence of data augmentation, the training set for Gleason 6-7 (low Gleason) class comprised approximately 150,000 patches, while the training set for Gleason 8-10 (high Gleason) class consisted of roughly 150,000 patches. Throughout several iterations, the network was fed 64-patch batches, and its optimal parameters were calculated via stochastic gradient descent optimization. The final layer employed Softmax as a loss function, and the training was terminated if the total model accuracy did not show a discernible improvement after 10–50 epochs. Once optimal accuracy has been reached, the learning weights are frozen, and the resulting model is used to classify image patches that were never seen before.

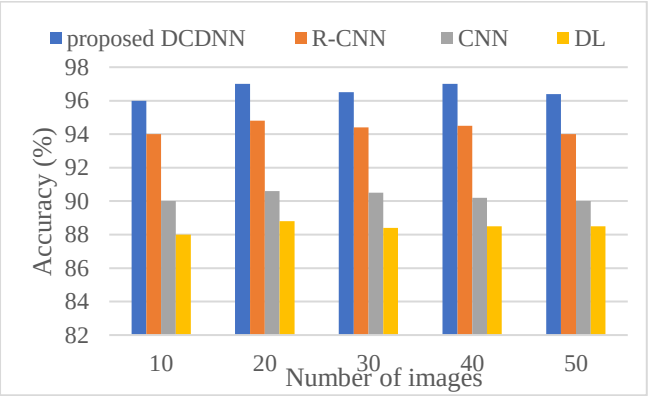

Fig 6: Accuracy comparison among DCDNN and others

Using the same number of histological images, Fig. 6 compares the accuracy performance of the proposed DCDNN to that of existing classification systems like R-CNN, CNN, and DL. As the number of epochs in the DCDNN model grew, so did its precision. The computation time is decreased while simultaneously improving accuracy. By optimizing local minima with MDSA, the proposed DCDNN system can achieve high accuracy—97.5%, to be exact—when compared to existing algorithms. The efficient ACLAHE preprocessing boosts DCDNN's accuracy and precision while cutting down on computation time.

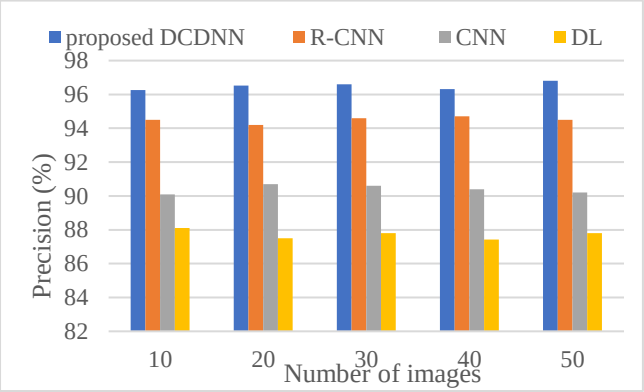

Fig 7: Precision comparison among DCDNN and others

For a given number of histological images in a given database, Fig. 7 compares the precision performance of the proposed DCDNN to that of existing classification systems like R-CNN, CNN, and DL. As the number of epochs in the DCDNN model grew, so did its accuracy. When compared to other algorithms, DCDNN's accuracy is significantly higher at 97.3%. The suggested CFCM achieved superior accuracy over alternative schemes under its effective illness segmentation and clustering. To see how the proposed DCDNN stacks up against other classification methods, such as R-CNN, CNN, and DL, for a given set of histology images, check out Fig. 8. The F1-score of the DCDNN model rose in tandem with the number of epochs. The DCDNN achieves a high F1-score of 97.5% compared to other algorithms by using processes that greatly enhance the quality of the hyperparameters of the hybrid CFCM scheme while reducing the computational cost.

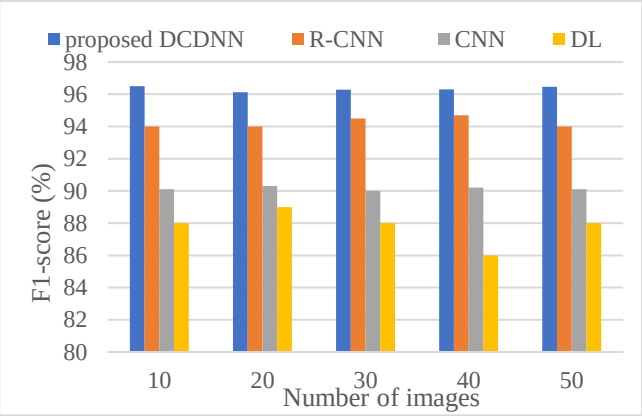

Fig 8: F1-score comparison among DCDNN and others

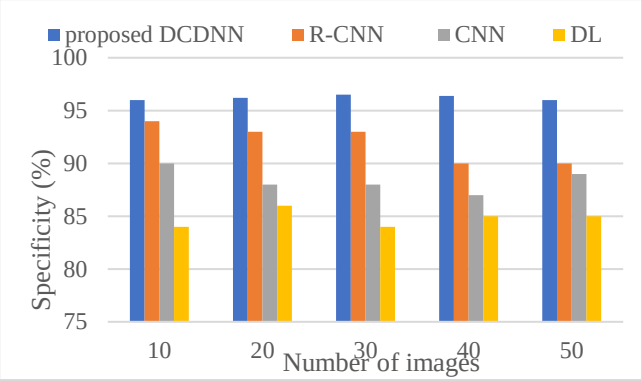

Fig 9: Specificity comparison among DCDNN and others

For the number of histology images in a particular database, Fig. 9 compares the specificity performance of DCDNN to that of existing classification techniques like R-CNN, CNN, and DL. When the number of epochs used to train the DCDNN model is increased, the model becomes more particular. While the computation time for DCDNN is decreased, its specificity is improved. When compared to other algorithms, the hybrid FCM-ACSO achieves an impressively high level of accuracy (96.34 %). The suggested CFCM achieved good specificity compared to existing schemes thanks to effective illness segmentation and effective clustering. Existing approaches are underfitting because they are based on simplistic models that perform unwell when applied to high-dimensional datasets.

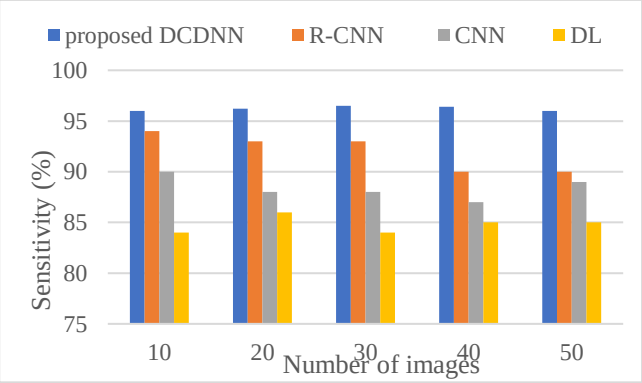

Fig 10: Sensitivity comparison among DCDNN and others

For a given number of histology images in a database, Fig. 10 compares the proposed DCDNN to existing classification systems like R-CNN, CNN, and DL in terms of sensitivity performance. The model's specificity improved as the number of epochs grew. The sensitivity of the DCDNN is improved while processing time is decreased. When

compared to other algorithms, the sensitivity achieved by the hybrid FCM-ACSO is significantly higher: 96.8%. The proposed CFCM and ACLAHE achieved excellent sensitivity in comparison to other schemes since they segmented diseases well and clustered them. Given that CFCM exhibits variable sensitivity-boosting potential, the suggested DCDNN's ability to optimize independently for each layer is an important feature for maximizing efficiency.

5. CONCLUSIONS

PC is one type of lung cancer as an important cause of cancer death; during his lifetime, nearly one out of every seven men will be identified with PC. The benefits of CFCM and ACLAHE are leveraged in this research to present a DCDNN with an MDSA algorithm for PC detection and classification in histological images. At first, ACLAHE is applied to images during preprocessing to lessen noise and heighten the precision of detection. For disease segmentation, FCM uses compactness through candidate minimization (CFCM) to avoid the local minimal en-trapment problem. The DCDNN model is then introduced for grading diseases as mild, moderate, or severe. Here, the MDSA is proposed to improve the hyperparameter to cut down on computation time and boost classification accuracy. The experiment was run on a dataset, and the results were compared to one another in terms of accuracy (97.5%), precision (97.3%), sensitivity (96.8%), f1-score (97.5%), and specificity (96.34%) for several classification tasks. The proposed DCDNN outperforms the state-of-the-art R-CNN, CNN, and DL algorithms on the benchmark dataset. The DCDNN can boost performance results by up to 33.5% across the board. In the future, a different real-time dataset will be used in conjunction with alternative categorization methods for assessment purposes.

REFERENCES

Jemal, Ahmedin, et al., “Global cancer statistics,” CA: a cancer journal for clinicians, vol. 61, no. 2, pp. 69–90, 2011.

Lemaître, Guillaume, et al., “Computer-aided detection and diagnosis for prostate cancer based on mono and multiparametric MRI: A review,” Computers in Biology and Medicine, vol. 60, pp. 8 – 31, 2015.

Sahebjavaher, R. S., Nir, G., Honarvar, M., Gagnon, L. O., Ischia, J., Jones, E. C., ... & Salcudean, S. E. (2015). MR elastography of prostate cancer: quantitative comparison with histopathology and repeatability of methods. NMR in Biomedicine, 28(1), 124-139.

Inamura, K. (2018). Prostatic cancers: understanding their molecular pathology and the 2016 WHO classifi-cation. Oncotarget, 9(18), 14723.

Epstein, Jonathan I., et al. (2016). The 2014 international society of urological pathology (isup) consensus conference on gleason grading of prostatic carcinoma, The American journal of surgical pathology, vol. 40, no. 2, pp. 244–252, 2016.

Epstein, J. I., Egevad, L., Amin, M. B., Delahunt, B., Srigley, J. R., Humphrey, P. A., & 2014 International Society of Urological Pathology (ISUP). (2015). Consensus conference on Gleason grading of prostatic car-cinoma: definition of grading patterns and proposal for a new grading system. Am J Surg Pathol..

Gordetsky, J., & Epstein, J. (2016). Grading of prostatic adenocarcinoma: current state and prognostic impli-cations. Diagnostic pathology, 11, 1-8.

Gleason Donald F. Histologic grading of prostate cancer: a perspective. Hum Pathol 1992;23(3):273–9

Bulten W, Pinckaers H, van Boven H, et al. Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study. Lancet Oncol 2020;21(2):233–41.

Li, D., Han, X., Gao, J., Zhang, Q., Yang, H., Liao, S., ... & Zhang, B. (2022). Deep learning in prostate cancer diagnosis using multiparametric magnetic resonance imaging with whole-mount histopathology referenced delineations. Frontiers in medicine, 8, p.810995.

Hassan, M. R., Islam, M. F., Uddin, M. Z., Ghoshal, G., Hassan, M. M., Huda, S., & Fortino, G. (2022). Prostate cancer classification from ultrasound and MRI images using deep learning based Explainable Artificial Intelligence. Future Generation Computer Systems, 127, pp.462-472.

Li, Bochong, Ryo Oka, Ping Xuan, Yuichiro Yoshimura, and Toshiya Nakaguchi. "Robust multi-modal prostate cancer classification via feature autoencoder and dual attention." Informatics in Medicine Unlocked 30 (2022): 100923.

Kartasalo, K., Ström, P., Eklund, M., Bulten, W., Pinckaers, H., Litjens, G., ... & Ruusuvuori, P.

(2022, April). Gleason grading of prostate cancer using artificial intelligence: lessons learned from the PANDA challenge. In Medical Imaging with Deep Learning.

Ayyad, Sarah M., Mohamed Shehata, Ahmed Shalaby, Mohamed Abou El-Ghar, Mohammed Ghazal, Moumen El-Melegy, Nahla B. Abdel-Hamid, Labib M. Labib, H. Arafat Ali, and Ayman El-Baz. "Role of AI and histopathological images in detecting prostate cancer: a survey." Sensors 21, no. 8 (2021): 2586.

Han, W., Johnson, C., Gaed, M., Gómez, J. A., Moussa, M., Chin, J. L., ... & Ward, A. D. (2020). Histologic tissue components provide major cues for machine learning-based prostate cancer detection and grading on prostatectomy specimens. Scientific reports, 10(1), p.9911.

Tolkach, Y., Dohmgörgen, T., Toma, M., & Kristiansen, G. (2020)High-accuracy prostate cancer pathology using deep learning. Nature Machine Intelligence, 2(7), pp.411-418.

Gurav, S. B., Kulhalli, K. V., & Desai, V. V. (2019). Prostate cancer detection using histopathology images and classification using improved RideNN. Biomedical Engineering: Applications, Basis and Communications, 31(06), 1950042.

Karimi, D., Nir, G., Fazli, L., Black, P. C., Goldenberg, L., & Salcudean, S. E. (2019). Deep learning-based gleason grading of prostate cancer from histopathology images—role of multiscale decision aggregation and data augmentation. IEEE journal of biomedical and health informatics, 24(5), pp.1413-1426.

Li, W., Li, J., Sarma, K. V., Ho, K. C., Shen, S., Knudsen, B. S., ... & Arnold, C. W. (2018). Path R-CNN for prostate cancer diagnosis and gleason grading of histological images. IEEE transactions on medical imaging, 38(4), 945-954.

Chen, S. D., & Ramli, A. R. (2004). Preserving brightness in histogram equalization-based contrast en-hancement techniques. Digital Signal Processing, 14(5), 413-428.

Liang, L., Wang, S., Chen, J., Ma, S., Zhao, D., & Gao, W. (2010). No-reference perceptual image quality metric using gradient profiles for JPEG2000. Signal Processing: Image Communication, 25(7), 502-516.

Barzilai, Jonathan, and Jonathan M. Borwein. "Two-point step size gradient methods." IMA Journal of Nu-merical Analysis 8.1 (1988): 141-148.

M. Heidari, S. Mirniaharikandehei, A. Z. Khuzani, G. Danala, Y. Qiu et al., “Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms,” In-ternational Journal of Medical Informatics, vol. 144, pp. 104284, 2020.

C. Zhang and S. Ding, “A stochastic configuration network based on chaotic dunnock search algorithm,” Knowledge-Based Systems, vol. 220, article 106924, 2021.

S. Li and Y. Li, Intelligent Optimization Algorithm Theory and Applications, Publishing House of Harbin Institute of Technology, Harbin, 2012.