ORIGINAL ARTICLE

Prostate cancer grading framework based on deep transfer learning and Aquila optimizer

Hossam Magdy Balaha[1] · Ahmed Osama Shaban[2] · Eman M. El-Gendy[1] (ORCID) · Mahmoud M. Saafan[1]

Abstract
Prostate cancer is the one of the most dominant cancer among males. It represents one of the leading cancer death causes worldwide. Due to the current evolution of artificial intelligence in medical imaging, deep learning has been successfully applied in diseases diagnosis. However, most of the recent studies in prostate cancer classification suffers from either low accuracy or lack of data. Therefore, the present work introduces a hybrid framework for early and accurate classification and segmentation of prostate cancer using deep learning. The proposed framework consists of two stages, namely classification stage and segmentation stage. In the classification stage, 8 pretrained convolutional neural networks were fine-tuned using Aquila optimizer and used to classify patients of prostate cancer from normal ones. If the patient is diagnosed with prostate cancer, segmenting the cancerous spot from the overall image using U-Net can help in accurate diagnosis, and here comes the importance of the segmentation stage. The proposed framework is trained on 3 different datasets in order to generalize the framework. The best reported classification accuracies of the proposed framework are 88.91% using MobileNet for the "ISUP Grade-wise Prostate Cancer" dataset and 100% using MobileNet and ResNet152 for the "Transverse Plane Prostate Dataset" dataset with precisions 89.22% and 100%, respectively. U-Net model gives an average segmentation accuracy and AUC of 98.46% and 0.9778, respectively, using the "PANDA: Resized Train Data (512 × 512)" dataset. The results give an indicator of the acceptable performance of the proposed framework.

Abbreviations

AdaDelta	Adaptive Delta
AdaGrad	Adaptive Gradient
Adam	An algorithm and not an acronym
AdaMax	Adaptive maximum
AI	Artificial Intelligence
AMSGrad	Adaptive Method Setup Gradient
AO	Aquila optimizer
AUC	Area Under Curve
AUROC	Area Under the Receiver Operating Characteristics
CNN	Convolutional Neural Network
CT	Computed Tomography
DA	Data Augmentation
DL	Deep Learning
DN	Data normalization
DRE	Digital Rectum Exam
DSC	Dice Similarity Coefficient
DWI	Diffusion Weighting magnetic resonance Imaging
ELU	Exponential Linear Unit
FN	False Negative
FP	False Positive
FTRL	Follow the Regularized Leader
GD	Gradient Descent
GPU	Graphics Processing Unit
IoU	Intersection over Union
ISUP	International Society of Urological Pathology
MRI	Magnetic Resonance Imaging
NAG	Nesterov Accelerated Gradient
NB	Naïve Bayes
NasNet	Neural Architecture Search Network
PANDA	Prostate cANcer graDe Assessment

✉ Eman M. El-Gendy
eman_elgendy@mans.edu.eg

1 Computers and Control Systems Engineering Department, Faculty of Engineering, Mansoura University, Mansoura, Egypt

2 Biomedical Engineering Program, Faculty of Engineering, Mansoura University, Mansoura, Egypt

PSA	Prostate Specifc Antigen
ReLU	Rectified Linear Units
ResNet	Residual Network
RF	Random Forest
RMSprop	Root Mean Square propagation
SeLU	Scaled Exponential Linear Unit
SGD	Stochastic Gradient Descent
TL	Transfer learning
TN	True Negative
TP	True Positive

1 Introduction

The prostate is a small gland inside the pelvis and surrounding the urethra with the shape of a walnut. This gland is responsible for producing the prostate fluid which creates semen when it is mixed with the sperm that is produced by testicles [1]. Prostate cancer is developed in the prostate when the cells began to grow in an unusual way and invade the surrounding organs and tissues.

According to the World Health Organization, Prostate cancer is the second most dominant and diagnosed cancer among males and represents one of the leading cancer death causes worldwide with 1,414,259 new cases (7.3% of the overall cancer new cases) and 375,304 death cases (3.8% of the total cancer death cases) globally and those number will be increased to 2,430,000 new cases and 740,000 death cases by 2040 [2]. In Egypt, prostate cancer is one of the most common cancers among Egyptian men with 4,767 new cases (3.5% of all new cases from various cancer types) and 2227 new death cases (2.5% of the overall total death cases for all cancer types) and by 2040 the morbidity and mortality cases will be doubled and reached 9,610 new cases and 4,980 new death cases [3]

These statistics for both Egypt and worldwide showed us that the mortality and morbidity of this cancer type are increased in a dramatic way that makes it the fastest malignancy of cancer among males, so the early detection and diagnosis of prostate cancer are very crucial for improving patient care and increasing the patient's survival rate [2, 3]

A lot of prostate cancer cases grow very slowly to cause any serious problems so there is no need for treatment. However in other cases, prostate cancer grows speedily and spread to the surrounding tissues and organs [4]. Adeno-carcinomas is the most common type of prostate cancer. Other types of prostate cancer include transitional cell carcinomas, neuroendocrine tumors, sarcomas, small cell carcinomas, and squamous cell carcinomas [5].

The appearance of prostate cancer symptoms depends on the stage of the disease if it was an early, advanced, or recurrent stage. The symptoms of prostate cancer may include (1) bone pain, (2) trouble urinating, (3) blood in the urine, (4) fatigue, (5) painful ejaculation, (6) blood in the semen, (7) jaundice, (8) numbness for feet or leg, (9) losing weight without trying, (10) erectile dysfunction, and (11) force decreasing in the urine stream [6, 7].

The related risk factors of prostate cancer depend on the person's lifestyle, family history, and age. The risk factors which increase the chance of having prostate cancer involves (1) obesity, (2) old age (i.e., after age 50), (3) family history, (4) ethnicity (i.e., the black men have a high probability to diagnose with prostate cancer), and (5) genes or cells DNA change [8].

There are various methods for diagnosing prostate cancer but screening tests are the most effective ways for detecting it in an early stage and it includes (1) prostate-specific antigen (PSA), (2) digital rectum examination (DRE), and (3) Biopsy [9]. PSA is a protein found on the prostate cells and made by its glands for keeping the semen in a liquid shape. Almost PSA lies in semen but also a small quantity exists in the bloodstream [10]. A high level of PSA refers to a high probability of being diagnosed with prostate cancer [11].

DRE is one of the important tools for screening prostate cancer by checking any abnormalities in the prostate area [12]. Using this test, any abnormality in the size, texture, or shape of the prostate gland can be noticed. DRE might be added beside the PSA test for screening prostate cancer [13].

Biopsy represents a medical procedure that involves taking small samples of the prostate tissues based on the gland size and examining them under a smart microscope to detect the existence of cancer cells in the prostate tissues and diagnose prostate cancer [14]. The doctor decides to make a biopsy based on the results of the DRE and PSA. Also, the doctor may require imaging tests such as magnetic resonance imaging (MRI) or Transrectal Ultrasound [15]

There are different treatment options for prostate cancer involving (1) surgery to remove the prostate cancer cells through prostatectomy procedure, (2) radiotherapy which includes brachytherapy (when the radiation pass inside the patient body) and external beam radiation therapy (when the radiation come from the body outside), (3) active surveillance which used when the cancer located only in the prostate and didn't spread to the surrounding areas and utilized for monitoring the cancer growth, (4) watchful waiting that is similar to active surveillance but it monitoring the cancer condition without treatment and focusing

on managing the symptoms, (5) focal therapy which being utilized when the cancer did not spread to another tissues and focusing on treating the prostate area affected by the cancer, otherwise systemic therapies utilized if the cancer succeed for spreading outside the prostate area, and (6) cryotherapy destroys the cancer cells through controlled freezing of the prostate gland, this type of treatment had been used for patients with a health issues to prevent them from having a radiotherapy treatment or surgery [16–18]. A graphical summary concerning the prostate cancer types, diagnosis, treatment, and symptoms is shown in Fig. 1.

In the last decades, a lot of prostate cancer detection approaches are proposed but they could not detect and diagnose cancer effectively. Recently, artificial intelligence (AI) had an important and crucial role in the diagnosis and detection process for different cancer types involving prostate cancer [19]. AI contains a subfield called Deep Learning (DL) which represents the state-of-the-art choice when we want to classify medical images from different images modalities (e.g., MRI, Ultrasound, and computed tomography (CT)) and extract features from images to determine if they contain a tumor or not [20].

For prostate cancer disease, researchers have deployed DL for detecting the disease through classifying medical images from different modalities and analyzing the biopsy images, and screening test results to detect any

abnormalities in the prostate [21]. The deployment of DL had a positive impact on the diagnosis process by decreasing time and cost and helping in the early detection of the disease that lead to improving the quality of healthcare provided for prostate cancer patients [22].

The current study introduces a hybrid framework for precise diagnosis and segmentation of prostate cancer using deep learning techniques. Diagnosis is performed using eight different pretrained CNN models, namely ResNet152, ResNet152V2, MobileNet, MobileNetV2, MobileNetV3Small, MobileNetV3Large, NASNet Mobile, and NASNet Large. Aquila optimizer is applied to tune the CNN models for better accuracy. According to the results of the diagnosis, the segmentation phase is triggered. It is crucial for physicians to determine the size and exact position of the tumor for better treatment. In the proposed framework, segmentation is performed via the U-Net model. Another merit of our proposed model is the use of three different datasets, namely (1) "PANDA: Resized Train Data (512 × 512)," (2) "ISUP Grade-wise Prostate Cancer," and (3) "Transverse Plane Prostate Dataset." The use of data from various datasets ensures the robustness of the presented framework.
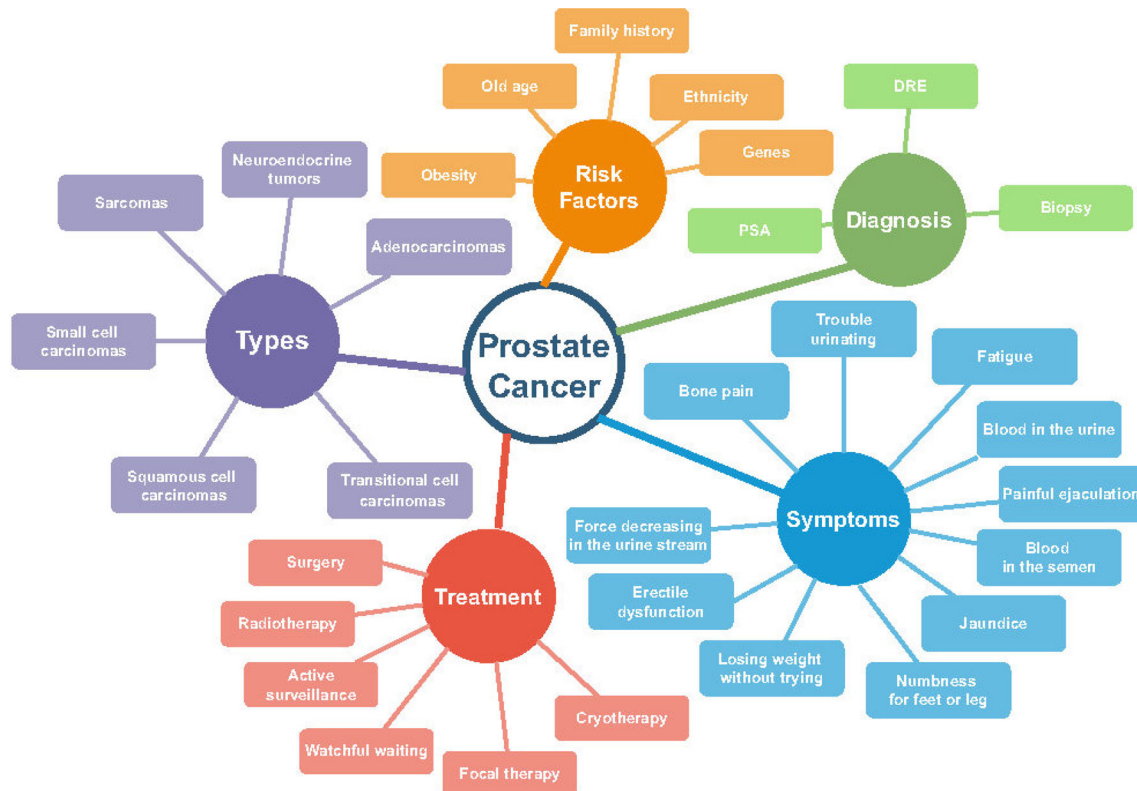
Fig. 1 A graphical summary concerning the prostate cancer types, diagnosis, treatment, and symptoms

1.1 Paper contributions

The contributions of the present study are:

- Developing a hybrid framework for precise diagnosis and segmentation of prostate cancer.
- In the diagnosis phase, eight different transfer learning CNN models have been applied in order to find the most promising model.
- Aquila optimizer is applied to tune the hyperparameters of the CNN models to find the best combinations with best performance.
- In the segmentation phase, U-Net is applied for segmenting prostate cancer to five grades.
- Comparing the performance of our framework with the state-of-the-art related studies.

1.2 Paper organization

The paper is organized in five sections. After the introduction section, the related state-of-the-art studies are presented in Sect. 2. Methodology of solution and the experimental results of the proposed framework are given in Sects. 3 and 4. The final section gives the conclusion, limitations, and the directions of future trends of the current study.

2 Related studies

Recently, there are extensive research works on prostate cancer detection and diagnosis based on artificial intelligence approaches involving machine learning and deep learning approaches [23–25]. The recent studies used different approaches, various datasets, and tools to facilitate the recognition of prostate cancer from MRI and CT images. The related studies are divided into studies that focused on (1) traditional machine learning (ML) algorithms, (2) deep learning (DL) approaches, and (3) hybrid approaches of DL and ML.

2.1 Machine learning-based studies

Zhang et al. [26] proposed a new framework for diagnosing prostate cancer and segmenting lesions in MRI by using various techniques such as Support Vector Machine, Multilayer Perceptron, and K-Nearest Neighbors. The obtained results showed that accuracy equals 80.97% and Dice value equals 0.79.

Erdem et al. [27] suggested an ML framework involving different techniques and algorithms for detecting prostate cancer such as Random Forest, Support Vector Machine,

Deep Neural Network, Multilayer Perceptron, Logistic Regression, Linear Discrimination Analysis, and Linear Regression. The result showed that the Multi-layer Perceptron had achieved the best performance with 97% accuracy, 100% recall, 0.958 AUC, 95% precision, and 97% F1-score.

Nayan et al. [28] applied different ML approaches such as fully connected Artificial Neural Network, Random Forest, Support Vector Machine, and Logistic Regression for predicting progression of the prostate cancer active surveillance to 790 patients. The best result was achieved by the Support Vector Machine classifier with F1-score equal to 0.586.

2.2 Deep learning-based studies

Gentile et al. [29] proposed a DL model framework for making optimized identification for prostate cancer by utilizing various PSA molecular forms such as free PSA, P2PSA, PSA density, and Total PSA. The proposed model was deployed with 437 patients and achieved 86% sensitivity and 89% specificity.

Shrestha et al. [30] suggested a novel approach for segmenting prostate cancer based on DL models and batch normalization on 230 MRI images. They utilized feature extraction optimization to extract multi-level features (i.e., low and high-level features) to perform prostate localization and shape recognition to improve prostate cancer segmentation and diagnosis. The proposed model achieved 95.3% accuracy for the segmentation process. Khosravi et al. [31] a DL framework for diagnosing prostate cancer by utilizing the fusion of pathology radiology and analyzing biopsies and MRI datasets for 400 patients with histological data and with prostate cancer suspicion. The proposed framework distinguishes between cancer and benign and determines the risk probability level for patients if it was in a high or a low level. Their model has Area under Curve (AUC) equal to 0.89.

Wessels et al. [32] carried out the DL approach of convolutional neural network (CNN) for predicting Lymph Node Metastasis from tumor histology of prostate cancer. The dataset consisted of stained tumor slides of 218 patients and they obtained 0.83 area under the receiver operating characteristics (AUROC) by utilizing CNN and Lymphovascular invasion. Linkon et al. [33] suggested a DL framework to diagnose prostate cancer and grade the Gleason score of histopathological images from 4 different datasets by deploying a CNN model and a post-processing technique to increase the performance of the model. They achieved 0.8921 Precision and 0.8460 F1- score.

Patel et al. [34] performed the detection of prostate cancer through DL techniques to deal with the MRI images for 158 patients. The performance of the detection process

was determined by the Ultrasound-guided and MRI-targeted biopsies. They yielded 82% accuracy, 86% precision, and 78% recall. Shao et al. [35] proposed a DL pipeline called ProsRegNet for simplifying the registration of the prostate histopathology and MRI images. They used 654 pairs of MRI and histopathology slices for 152 prostate cancer patients and obtained Dice coefficient values in the range of 0.96 to 0.98.

Amarsee et al. [36] carried out an automated tracking and detection of implanted marker seeds in prostate cancer patients by utilizing DL approaches. The marker seeds represented an identification for the prostate volume positioning in the prostate cancer treatment. CNN model analyzed 1,500 images and achieved 98% accuracy. Yang et al. [37] suggested an approach for the automatic detection of prostate cancer of a multi-parametric MRI (mpMRI) dataset for 780 patients. They used a CNN model and achieved 0.9684 AUC. Kovalev et al. [38] performed a computerized prostate cancer diagnosis that depended on DL techniques. They used 10,616 slide histology images. They deployed different CNN architectures and feature extraction techniques. The best-obtained performance was 92.77% accuracy.

John et al. [39] proposed a prostate cancer prediction approach for 330 mpMRI images from the Prostate X challenge dataset. They deployed two CNN architectures, namely DenseNet and MobileNet to analyze the prostate screening methods including DRE and PSA tests. They obtained 0.931 AUC and 0.93 F1-score. Comelli et al. [40] developed a DL framework involving three approaches for segmenting prostate cancer in MRI images for 85 patients and they yielded a 90.89% Dice coefficient. Pinckaers et al. [41] carried out prostate cancer detection in slide images of biopsies. They used a dataset that consisted of 1243 slides containing 5759 prostate biopsies. The proposed CNN model yielded 0.992 AUC.

Salvi et al. [42] proposed a hybrid DL approach called RINGS (i.e., Rapid IdentificatioN of Glandural Structures) for segmenting prostate glands of histopathological images which support a prostate cancer diagnosis process. They utilized 1500 prostate cancer biopsy images for 150 men. The suggested approach achieved a 90.16% Dice score, 91.24% precision, and 97.23% recall. Korevaar et al. 43 developed a new prostate cancer detection framework for CT scans. The dataset involved 571 CT scans (139 prostate cancer patients through MRI or Transrectal ultrasound-guided biopsy and 432 control cases for unknown prostate cancer). The results showed that the specificity is 98.8% and 0.88 AUC.

Chahal et al. [44] suggested a novel framework of CNN architecture (Xception model) that was based on the U-Net. The used dataset contained 1,689 2D gray-scaled MRI images for 50 patients. They achieved an overall Dice

value of 97.50%. Sobecki et al. [45] developed a prostate cancer diagnosis framework based on the CNN model and mpMRI dataset containing 538 images for 344 patients. They achieved 0.84 AUC. Balagopal et al. [46] suggested a DL approach for automatically segmenting Clinical Target Volumes with uncertainties of prostate cancer radiotherapy on CT images. The proposed CNN model was utilized for the localization of Organs at Risk and Clinical Target Volumes. They obtained a high Dice Similarity Coefficient value of 0.87.

Liu et al. [47] proposed a DL approach representing in Computer-Aided Diagnosis system for ultrasound image-aided prostate cancer diagnosis containing S-mask R-CNN and InceptionV3. They utilized a dataset involving 1,200 ultrasound images and yielded a Dice coefficient equal to 0.87 and a precision score for malignant and benign categories equal to 0.8 and 0.76, respectively. Abdelmaksoud et al. [44] used the transfer learning (TL) approach on the DWI dataset which consists of 470 Diffusion Weighting slices (234 malignant cases slices and 236 benign cases slices). The proposed approach achieved 91.2% accuracy, 90.1% specificity, and 91.7% sensitivity.

Ambroa et al. [48] suggested a TL approach based on the CNN model for predicting the dose volume of a histogram for prostate cancer radiotherapy. They carried out a CNN model to predict bladder and rectum Dose Volume Histogram of prostate patients. They used a database containing 2D images of CT scans for 144 patients and achieved 87.5% accuracy and 100% precision. Hao et al. [49] performed prostate cancer detection by utilizing various DA strategies on Diffusion-Weighted magnetic resonance Imaging (DWI) dataset containing 10,128 2D slices of 414 patients through utilizing different CNN models. They obtained 0.85 AUC.

Kudo et al. [50] deployed a CNN approach to diagnose prostate cancer in 32 prostate cancer biopsy images and 2594 fragments. They obtained an overall accuracy of 98.3%. Mehta et al. [51] utilized clinical features and mpMRI images on PICTURE and ProstateX datasets. Their achieved AUC for both datasets were 0.79 and 0.86, respectively. Hoar et al. [52] performed semantic segmentation for prostate cancer on mpMRI images by utilizing a combination of TL and test-time augmentation to improve the ability of the CNN model to distinguish between cancer or non-cancer. They used 154 subjects for mpMRI and yielded 0.93 AUROC. Cipollari et al. [53] suggested a CNN model to perform an automated prostate cancer classification on 316 mpMRI images for 312 men and the best performance among sequences were 100% and 96.62% accuracy.

Pellicer-Valero et al. [54] developed a DL approach for performing automated segmentation, estimation of Gleason Grade, diagnosis, and detection of prostate cancer in 490

mpMRI images for 75 patients collected from two datasets (ProstateX dataset and Valencia Oncology Institute Foundation dataset). They obtained an overall Dice Similarity Coefficient in the range between 0.894 and 0.941. Saunders et al. [55] suggested a DL framework for segmenting prostate cancer using the U-Net model. They utilized TL and DA techniques on the used 3 MRI scan datasets to increase the data and improve the model performance. The proposed model obtained a 0.9 Dice Similarity Coefficient.

Han et al. [56] developed DL models to detect bone metastasis on the body bone scan in prostate cancer patients. They utilized 9113 bone scans (i.e., 2991 positives and 6142 negatives) of 5342 prostate cancer patients and two different CNN architectures representing Global Local Unified Emphasis and Whole Body-based. The obtained results were 90% accuracy (and 0.946 AUC) and 88.9% accuracy (and 0.944 AUC), respectively.

2.3 Hybrid approaches of deep learning and machine learning-based studies

Iqbal et al. [57] proposed a prostate cancer detection based on DL approaches such as Long Short-Term Memory and ResNet, and traditional techniques for 230 MRI scans of patients with various descriptions and categories. They also used different non-DL models such as K-Nearest Neighbors Cosine, Naïve Bayes, RUS Boost tree, Support Vector Machine, and Decision Tree. The results showed that the ResNet model has obtained the best performance among the DL approaches with a 100% accuracy and 1.0 AUC.

Salama et al. [58] performed a prostate cancer detection depending on Deep Convolutional Neural Network and Support Vector Machine. They used the DWI dataset and deployed various techniques involving TL and Data Augmentation (DA). After performing the DA technique, 1,765 images existed for training the model to classify whether prostate cancer exists or not. The proposed work obtained 98.79% accuracy, 98.43% sensitivity, 0.9592 F1-score, 0.9891 AUC, and 97.99% precision.

Table 1 gives a summary of the discussed related studies in 2021.

3 Methodology

As noticed from the related studies, the common disadvantages of the presented studies are either low accuracy or lack of data. Therefore, we suggest a hybrid framework for accurate diagnosis and segmentation of prostate cancer using deep learning. To ensure the variety and robustness of the proposed model, three different datasets are applied, namely (1) "PANDA: Resized Train Data (512 × 512)," (2) "ISUP Grade-wise Prostate Cancer," and (3)

"Transverse Plane Prostate Dataset." Data preprocessing is applied to unify the different datasets to one single platform. The proposed framework is divided to two stages, namely classification stage and segmentation stage. In the first stage, eight different deep learning algorithms via transfer learning approach are used to classify patients of prostate cancer from normal ones. Aquila optimizer is used to fine-tune the hyperparameters of the various models. The second stage begins once the patient is diagnosed with prostate cancer. This phase is important to help doctors identify the infected regions and therefore, the size and shape of the tumor can be correctly recognized.

The current section presents the suggested approach for prostate cancer classification and segmentation. It is divided into different phases that are summarized in Fig. 2. In short, it starts by acquiring the datasets. The current study uses three datasets, one for segmentation and two for classification. After that, it applies different pre-processing techniques such as data augmentation, resizing, and scaling. Different transfer learning CNN models are utilized in the image classification and hyperparameters optimization phase. The Aquila optimizer is used to optimize the hyperparameters in that phase. Different performance metrics are utilized to judge the system performance. The U-Net segmentation model is utilized to determine and grade the locations of the tumors.

3.1 Data acquisition

The prostate cancer classification and segmentation in the current study are performed by utilizing 3 datasets that are obtained from Kaggle online platform. The used datasets are (1) "PANDA: Resized Train Data (512 × 512)", (2) "ISUP Grade-wise Prostate Cancer", and (3) "Transverse Plane Prostate Dataset."

3.1.1 The "PANDA: resized train data (512 × 512)" dataset

Prostate cANcer graDe Assessment (PANDA) represents as the biggest whole-slide image collection public dataset with about 11,000 slide images for digitized biopsies of H&E-stained. The dataset is provided by two centers (i.e., "Karolinska Institute" and "Radboud University Medical Center") and is utilized to perform the segmentation process. In the two centers, the prostate glands are labeled into the stroma, malignant epithelium, benign epithelium, and non-tissue [61]. The "PANDA: Resized Train Data (512 × 512)" dataset is a resized version of PANDA which is used in the current study [62].

Table 1 Related studies summarization

Reference	Category	Approach	Pros	Cons	Best performance
Zhang et al. [26]	Machine Learning	Proposed a new ML framework for diagnosing prostate cancer and segmenting lesions in MRI by using various techniques like feature extraction and ensemble learning	Utilized different algorithms and techniques. Suitable dataset size	Low accuracy	80.97% accuracy and 0.79 dice
Erdem et al. [27]	Machine Learning	Suggested ML framework involving different techniques and algorithms for predicting prostate cancer	Utilized different algorithms High accuracy	Small dataset	97% accuracy, 100% recall, 0.958 AUC, 95% precision, and 97% $F1$-score
Nayan et al. [28]	Machine Learning	Proposed ML approach for predicting progression for the prostate cancer active surveillance	Utilized different algorithms and techniques. Suitable dataset size	Low performance	0.586 $F1$-score
Gentile et al. [29]	Deep Learning	Proposed DL model framework for making optimized identification for prostate cancer by utilizing various PSA molecular forms	Suitable dataset size. Utilized different techniques	Low performance	86% sensitivity and 89% specificity
Shrestha et al. [30]	Deep Learning	Suggested a novel approach for segmenting prostate cancer based on DL models and batch normalization on MRI images	Utilized different techniques. Good accuracy	Small dataset	95.3% accuracy
Khosravi et al. [31]	Deep Learning	DL framework for diagnosing prostate cancer by utilizing the fusion of pathology-radiology and analyzing biopsies and MRI datasets	Utilized different techniques Suitable dataset size	Low performance	0.89 AUC
Wessels et al. [32]	Deep Learning	DL approach of CNN for predicting Lymph Node Metastasis from tumor histology of prostate cancer	Utilized DL approaches	Small dataset and low performance	0.83 AUROC
Linkon et al. [33]	Deep Learning	Suggested DL framework to diagnose prostate cancer and grade Gleason score of histopathological images by deploying CNN model and post processing technique	Utilized DL approaches suitable dataset size	Low performance	0.8921 precision and 0.8460 $F1$-score
Patel et al. [34]	Deep Learning	Performed the detection of prostate cancer through DL techniques to deal with MRI images	Utilized DL approaches	Small dataset and low accuracy	82% accuracy, 86% precision, and 78% recall
Shao et al. [35]	Deep Learning	Proposed new DL pipeline called prosregnet for simplifying the registration of the prostate histopathology and MRI images	Suitable dataset size. Proposed a novel approach. Good performance		(0.96 to 0.98) dice
Amarsee et al. [36]	Deep Learning	Automated tracking and detection implanted marker seeds in prostate cancer patients by utilizing DL approaches	Large dataset. Good accuracy		98% accuracy
Yang et al. [37]	Deep Learning	Automated detection of prostate cancer of mpMRI images using CNN models	Better accuracy. Suitable dataset size. Utilized different techniques for the prediction process		0.9684 AUC

Table 1 (continued)

Reference	Category	Approach	Pros	Cons	Best performance
Kovalev et al. [59]	Deep Learning	Computerized prostate cancer diagnosis depends on DL techniques for slide histology images	Large dataset. Good accuracy Utilized different techniques		92.77% accuracy
John et al. [39]	Deep Learning	Proposed a prostate cancer prediction approach based on CNN architectures for mpMRI images	Good accuracy. Utilized different techniques	Small dataset	0.931 AUC and 0.93 F1-scores
Comelli et al. [40]	Deep Learning	DL framework involving three approaches for segmenting prostate in MRI images	Good performance. Utilized different DL approaches	Small dataset	90.89% Dice Coefficient
Pinckaers et al. [41]	Deep Learning	Prostate cancer detection in slide images of biopsies using CNN model	Good performance. Large dataset		0.992 AUC
Salvi et al. [42]	Deep Learning	Proposed a new hybrid DL approach for segmenting prostate glands of histopathological images	Suitable dataset size. Good performance		90.16% dice, 91.24% precision, and 97.23% recall
Korevaar et al. [43]	Deep Learning	Developed a new DL prostate cancer detection framework for CT scans	Good performance. Utilized different DL approaches	Small dataset	98.8% specifcity and 0.88 AUC
Chahal et al. [44]	Deep Learning	Suggested a novel framework of CNN architecture (Xception model) which is based on U-Net	Good performance	Small dataset	97.50% dice
Sobecki et al. [45]	Deep Learning	Prostate cancer diagnosis framework based on the CNN model for mpMRI images	Utilized different DL approaches	Small dataset and low performance	0.84 AUC
Balagopal et al. [46]	Deep Learning	DL approach for automatically segmenting Clinical Target Volumes with uncertainties of prostate cancer radiotherapy on CT images	Suitable dataset size. Utilized different techniques	Low performance	0.87 dice
Liu et al. [47]	Deep Learning	DL approach representing in Computer-Aided Diagnosis System for ultrasound image-aided prostate cancer diagnosis	New approach. Suitable dataset size	Low performance	0.87 dice
Abdelmaksoud et al. [60]	Deep Learning	Cads for precise prostate cancer identification using TL approach for DWI dataset	Good performance. Utilized different DL approaches	Small dataset	91.2% accuracy, 90.1 specifcity, and 91.7% Sensitivity
Ambroa et al. [48]	Deep Learning	DL approach based on transfer learning and CNN model for predicting the dose volume of a histogram for prostate cancer radiotherapy	Good performance. New approach	Small dataset	87.5% accuracy and 100% precision
Hao et al. [49]	Deep Learning	Prostate cancer detection by utilizing various DA strategies on DWI MRI images	Suitable dataset size. Utilized different techniques	Low performance	0.85 AUC
Kudo et al. [50]	Deep Learning	CNN approach to diagnose prostate cancer in biopsy images and fragments	Good accuracy. Utilized different techniques	Small dataset	98.3% accuracy
Mehta et al. [51]	Deep Learning	Cads for classifying prostate cancer by utilizing clinical features and mpMRI images	Suitable dataset size	Small dataset	0.86 AUC

Table 1 (continued)

Reference	Category	Approach	Pros	Cons	Best performance
Hoar et al. [52]	Deep Learning	CNN approach of semantic segmentation for prostate cancer on mpMRI images through utilizing a combination of TL and test-time augmentation	Good performance. New approach	Small dataset	0.93 AUROC
Cipollari et al. [53]	Deep Learning	CNN approach to perform an automated prostate cancer classification on mpMRI images	Better performance	Small dataset	100% accuracy
Pellicer-Valero et al. [54]	Deep Learning	DL approach for performing automated segmentation, estimation of Gleason Grade, diagnosis and detection of prostate cancer in mpMRI images	Good performance. Utilized different DL approaches	Small dataset	0.894 to 0.941 dice
Saunders et al. [55]	Deep Learning	DL framework for segmenting prostate cancer using the U-Net model	Utilized different DL approaches. Good performance. Suitable dataset size		0.9 dice
Han et al. [56]	Deep Learning	DL approach to detect bone metastasis on the body bone scan in prostate cancer patients	Large dataset. Good accuracy. Utilized different DL techniques		90% accuracy and 0.946 AUC
Iqbal et al. [57]	Deep Learning	Prostate cancer detection based on DL approaches such as Long Short-Term Memory and Residual Net, and traditional techniques	Better performance. Utilized different DL techniques	Small dataset	100% accuracy and 1 AUC for DL approaches. 99.07% accuracy and 0.998 AUC for ML approaches
Salama et al. [58]	Deep Learning	Prostate cancer detection depends on Deep Convolutional Neural Network and Support Vector Machine on DWI dataset	Suitable dataset size. Good accuracy. Utilized different DL approaches		98.79% accuracy, 98.43% sensitivity, 0.9592 F1-score, 0.9891 AUC, and 97.99% precision

3.1.2 The "ISUP grade-wise prostate cancer" dataset

The "ISUP Grade-wise Prostate Cancer" dataset consists of 10,616 prostate images which contain a severity on a scale grading from zero to five representing the cancer condition if it was significant or non-significant [63, 64].

3.1.3 The "transverse plane prostate dataset" dataset

The "Transverse Plane Prostate Dataset" contains a total of 1,528 prostate MRI images from 64 patients in the transverse plane. The dataset is partitioned into two categories (i.e., significant and non-significant) [64, 65].

3.2 Data preprocessing

Data preprocessing is an essential process to prepare the dataset for the next phase and make the data more efficient in performing the classification and segmentation processes effectively. It involves various techniques such as data augmentation, resizing, and scaling [66].
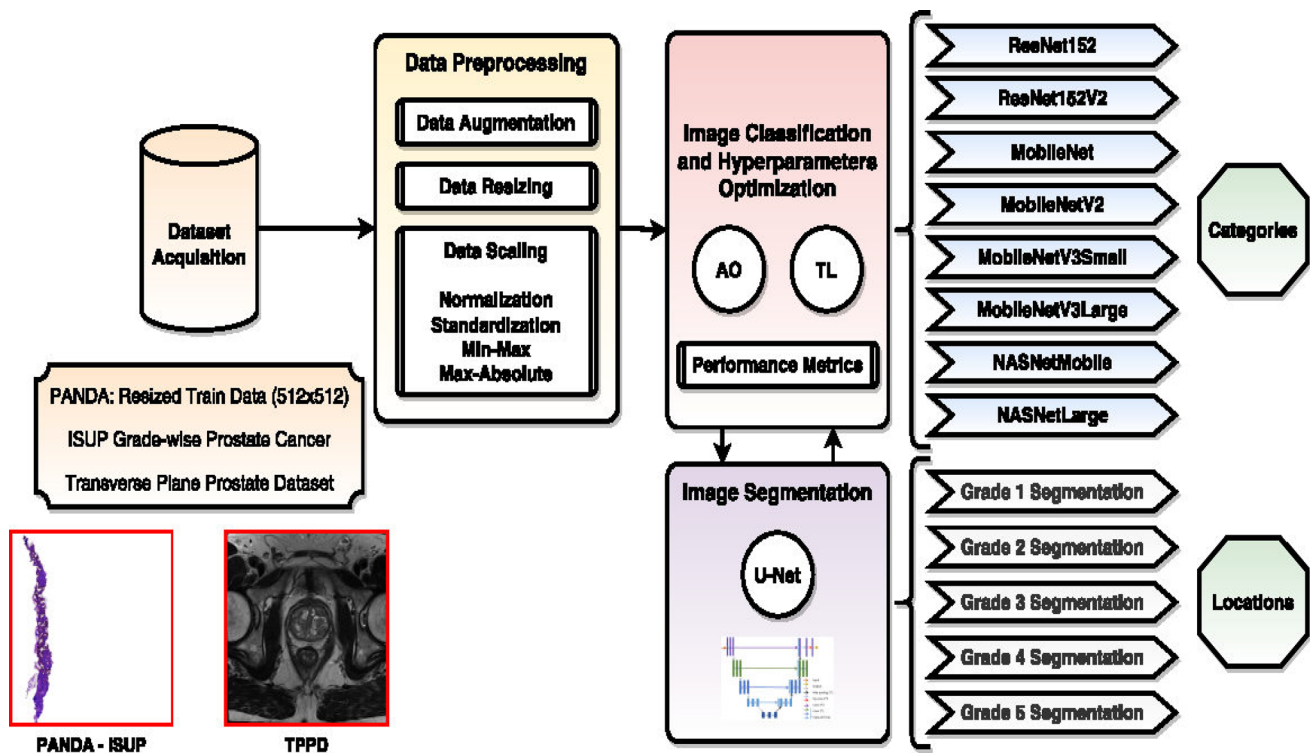
Fig. 2 The suggested approach for prostate cancer classification and segmentation

3.2.1 Data augmentation

Data Augmentation (DA) is a process involving the dataset augmenting and trying to increase it to a diverse and large one [67]. DA techniques improve the model performance. Different augmentation techniques can be applied including (1) zooming it in (or out), (2) rotation, (3) flipping horizontally (or vertically), (4) shifting vertically (or horizontally), (5) shearing vertically (or horizontally), (6) changing the brightness of the image, and (7) cropping [68–70]. Table 2 shows the used ranges in the current study.

Table 2 The used data augmentation techniques and their ranges

Technique	Range
Rotation range	0–45 with a step of 1
Width shift range	[0,0.25]%
Height shift range	[0,0.25]%
Shear range	[0,0.25]%
Zoom range	[0,0.25]%
Horizontal flip range	True and False
Vertical flip range	True and False
Brightness range	[0.5,2.0]

3.2.2 Data resizing

The images in the current study are resized to be equal in size and to make the training process quicker through smaller images. For the segmentation dataset, the used size is $(128, 128, 3)$ while in the classification datasets, the used size is $(100, 100, 3)$.

3.2.3 Data scaling

Data scaling is one of the pre-processing techniques that normalize the attribute values to be within a known scale (or range) to improve the overall performance. There are different scaling techniques but the used ones in the current study are (1) min–max scaling, (2) normalization, (3) standardization, and (4) max-absolute scaling [71, 72].

Min–Max Scaling transforms all the features into a range between 0 and 1. Also, the dataset is normalized into the range between -1 and $+1$ in case of the existence of negative values on the data [73]. Equation 1 shows how to calculate it, where out is the output, in is the input, in_{max} is the maximum value, and in_{min} is the minimum value.

$$out = \frac{in - in_{min}}{in_{max} - in_{min}} \tag{1}$$

Normalization squeezes the dataset in the range between 0 and 1. It is an efficient way in the classification process especially with datasets involving negative values [74]. Equation 2 shows how to calculate it.

$$out = \frac{in}{in_{max}} \tag{2}$$

Standardization standardizes each feature value within the distribution by removing the mean (i.e., mean equal to zero) and scaling the values into unit variance [70]. Equation 3 shows how to calculate it, where μ is the mean value and σ is the standard deviation.

$$out = \frac{in - \mu}{\sigma} \tag{3}$$

Max-Absolute Scaling looks like the min–max scaler but the values are not scaled into the range within 0 and 1. It is translated and normalized into the range between -1 and $+1$ by dividing it by the maximum absolute value [75]. Equation 4 shows how to calculate it.

$$out = \frac{in}{|in_{max}|} \tag{4}$$

3.3 Image segmentation

Image segmentation is an essential process in the computer vision field to get more detailed information. An essential need for image segmentation jobs is the usage of masks. We may get the desired result necessary for segmentation by using masking, which is essentially a binary picture consisting of zero and nonzero values. The segmentation objective is to identify a label for every pixel of the image that represents what is being represented [76]. Deep learning approaches have recently been utilized to segment medical images using a variety of architectures including U-Net, U-Net ++ , Swin-Unet, Attention U-Net, and V-Net [77].

U-Net is utilized in the current study. It is a convolutional network architecture for images segmentation that is quick and precise. It is a technique of constructing a network of sliding window architecture by treating each pixel's class label as a distinct unit and giving a local area (i.e., patch) surrounding it [78]. The goal of the U-Net is to collect both context and localization characteristics. It is built on the principle of using consecutive contracting layers, followed by upsampling operators, to get better resolution outputs on the input pictures.

With U-Net, computations may be completed in a short amount of time using a contemporary Graphics Processing Unit (GPU). The encoder and decoder are the two primary components of the U-Net design. The encoder is made up of several layers that are followed by a pooling process and it is used to extract the image's factors. To allow for localization, the decoder employs transposed convolution, and it is a fully connected layers network [79].

The U-Net model is applied for each masked layer on the "PANDA: Resized Train Data (512×512)" dataset and hence there are five segmenters one for each layer. Table 3 shows the hyperparameters used in the U-Net training process.

3.4 Learning and classification using transfer learning

Deep learning (DL) is a rapidly expanding machine learning (ML) discipline with applications in image identification, picture creation, self-driving cars, and medical sciences [80]. The architecture of DL is made up of several linked layers of weighted components that are comparable to neurons [81]. The main common characteristic of DL methods is their focus on feature learning and automatically learning representations of data [82]. Feature extraction is a technique that is utilized for capturing and extracting relevant and required attributes and information of images to use in further processing. It is deployed on rich datasets involving medical images modalities such as CT, MRI, and Ultrasound [83].

Table 3 The used U-Net hyperparameters in its training process

Technique	Range
Architecture filters	[64, 128, 256, 512, 1024]
Hidden activation function	GELU
Output activation function	Sigmoid
Backbone model	VGG19
Model pre-trained weights	ImageNet
Freeze backbone?	True
Apply batch normalization	True
Loss function	Binary Crossentropy
Optimizer	Adam
Batch size	4
Number of epochs	10
Early stopping patience	5
Apply shuffling?	True
Input image size	(128, 128, 3)
Train-to-test split	85%: 15%
Scaling technique	Min–Max scaling

3.4.1 Convolutional neural network (CNN)

When we extract DL features from the image's dataset, we utilized CNN as an automatic feature extractor. CNN is represented as one of the most powerful tools for DL tasks which take an input image, capture features from it using kernels or digital filters, classify higher dimension datasets, and transfer it to lower dimensions without missing any information [85–87].

A CNN's architecture is composed of multilayers including an Input Layer that contains the input image and holds its pixel values. The input image is passed through a convolutional layer which extracts multiple levels of information using kernels and filters with predefined widths and heights. It will determine which output neurons are linked to a particular part of the input data [88]. By sliding the filters on the input, multiple convolution operations are conducted to extract distinct features levels from the image and stack them [89].

The pooling layer downsamples the input image and decreases the image's parameters, reducing training time and overfitting without sacrificing critical data. After the pooling procedure, the fully connected layer represents a flattened feed-forward layer that facilitates classification processes. As the output of the convolutional layer, non-linear combinations of features are learned after the downsampling and feature extraction procedures [90]. The fully linked layer connects all neurons in the previous and subsequent layers. To produce predictions and classify the input data into multiple classes, it can be a nonlinear activation function [91].

The batch normalization layer is one of the main layers in the CNN architecture that improves the model's performance and speeds up the training procedure by (1) allowing a wider range of learning rates and (2) re-parametrizing optimization problems, resulting in a more stable, smoother, and faster training process [92]. The activation layer is essential to have the node's output such as Sigmoid, Hyperbolic Tangent (Tanh), Rectified Linear Unit (ReLU), Leaky ReLU, Exponential Linear Unit (ELU), Scaled Exponential Linear Unit (SeLU), and Soft-Max functions [93].

3.4.2 Transfer learning (TL)

It represents an artificial intelligence (AI) approach in which a pre-trained model that was previously used for one job is reused for another. It may be summed up as the transfer of knowledge [80]. The basic concept behind reusing pre-trained models for new processes is to get a starting point and a lot of labeled training data in a new one that doesn't have much data, rather than beginning from scratch and producing these labeled data, which is highly

expensive [94, 95]. On the ImageNet image database [96], several pre-trained CNN models were developed, but the ones utilized in this study are ResNet152, ResNet152V2, MobileNet, MobileNetV2, MobileNetV3Small, Mobile-NetV3Large, NASNet Mobile, and NASNet Large.

3.4.3 Parameters optimization

It is representing an expanded method for changing weights of the CNN model aiming at the probability of achieving accurate results and reducing losses [97]. There are many types of optimizers but the used ones in the current study are (1) Stochastic Gradient Descent (SGD), (2) Stochastic Gradient Descent-Nesterov (SGD-Nesterov), (3) Adaptive Gradient (AdaGrad), (4) Adaptive Delta (AdaDelta), (5) Adaptive Moment Estimation (Adam), (6) Adaptive Maximum (AdaMax), (7) Root Mean Square Propagation (RMSprop), (8) Nesterov-accelerated Adaptive Moment Estimation (Nadam), (9) Root Mean Square Propagation-Centered (RMSprop-Centered), (10) Follow the Regularized Leader (FTRL), and (11) Adaptive Method Setup Gradient (AMSGrad) [98–100].

3.4.4 Learning hyperparameters

The **Loss Function** is crucial in evaluating the suggested solution by referring it to the needed function and calculating model errors [101]. As a result, we can determine how good the model is and adjust the parameters to improve the model's performance and reduce the overall loss [102]. It can be used as a penalty for failing to achieve the target output. For example, if the model's predicted value differs significantly from the desired value, the function will return a large loss value and a lesser number otherwise [103]. The utilized losses in the current study are (1) Categorical Hinge [104], (2) Poisson [105], (3) Squared Hinge [106], (4) Categorical Crossentropy [107], (5) Hinge [108], and (6) KLDivergence [109].

The **Batch Size** shows the number of data records used to train the model in each iteration to ensure model generalization, parameter values, and loss function convergence. It is critical to the model's learning process since it makes it faster and more stable [99, 110]. The Dropout is a regularization strategy for training the CNN model on any or all of the architecture's hidden layers. It is critical in preventing and correcting the overfitting problem to maintain optimal performance. By setting the output of every neuron to 0 at random, it can boost generalization efficiency across all data [111]. The hyperparameters to be optimized by the Aquila optimizer are summarized in Table 4.

Table 4 The hyperparameters to be optimized by the Aquila optimizer

Hyperparameter	Ranges
Losses	Categorical Crossentropy, Squared Hinge, Hinge, Categorical Hinge, KLDivergence, and Poisson
Parameters optimizers	Adam, NAdam, Adam AMSGrad, AdaDelta, AdaGrad, AdaMax, RMSProp, SGD, SGD
Dropout	Nesterov, Ftrl, and RMSProp Centered
Batch size	$[0 \rightarrow 0.6]\%$
TL learn ratio	$4 \rightarrow 48$ (step = 4)
Scaling techniques	$1 \rightarrow 100$ (step = 1)
Apply data augmentation (DA)	Normalization, Standardization, Min–Max, and Max-Absolute
DA configurations	[Yes, No]
	Table 2

3.5 Meta-heuristic optimization using Aquila Optimizer (AO)

It is a popular choice for modeling and solving complex issues that are difficult to tackle with standard methods and optimization. The term "meta" in Meta-heuristic refers to a higher level of performance that outperforms simple heuristics [112–118]. For global exploration and local search, it employs a trade-off [119, 120]. Diversification and intensification are crucial aspects of meta-heuristic algorithms. Diversification creates a variety of alternatives for exploring the search space, whereas intensification concentrates on the search in a particular area by utilizing information to find a good answer in that area.

Aquila Optimizer (AO) is a cutting-edge meta-heuristic optimization technique. The AO algorithm's optimization process is divided into four steps (1) choosing the search space by vertical stooping (Equation 5), (2) discovering the various search spaces by contour flight by short glide attack (Equation 6), (3) swooping by grabbing prey and walking (Equation 7), and (4) exploiting through converge search space by low flight by descent attack (Equation 8), where N is the population size, $X(t + 1)$ is the solution of the next iteration, t is the iteration number, rand is a random number in the range $[0, 1]$, T is the total number of iterations, $XR(t)$ is a random solution in the current iteration t, D is the dimension space size, $X_{best}(t)$ is the best solution in the current iteration t, Levy(D) is the levy flight distribution function, U equals 0.00565, $r1$ is a value in the range $[1, 20]$, $D1$ is a value in the range $[1, D]$, α (and σ) equal to 0.1, UB is the upper bound, QF is the quality function, and LB is the lower bound. The fixed values are taken from the original AO paper.

The optimization techniques in AO begin by producing a random preset collection of candidate solutions known as the population. Equation 9 shows how to create the AO population where UB and LB are the upper and lower

bounds respectively and rand is a random vector between 0 and 1. After that, the AO search strategies examine the positions of the optimum solution using the repetition trajectory or the near-optimal one [121]. Every solution in the AO's optimization phase updates its location based on the best solution. The AO conducts a series of experiments to verify the optimizer's capacity to identify the optimum solution for various optimization tasks. The flowchart of AO is shown in Fig. 3.

$$X_{t+1} = X_{best}(t) \times \left(1 - \frac{1}{T}\right) + \left(\frac{\sum_{i=1}^{N} X(t)}{N} - X_{best}(t) \times \text{rand}\right) \tag{5}$$

$$X_{t+1} = X_{best}(t) \times \text{Levy}(D) + XR(t) + (r1 + U \times D1) \times \cos(-\omega \times D1 + 1.5 \times \pi) - \sin(-\omega \times D1 + 1.5 \times \pi)) \times \text{rand} \tag{6}$$

$$X_{t+1} = \left(X_{best}(t) - \frac{\sum_{i=1}^{N} X(t)}{N}\right) \times \alpha - \text{rand} + ((\text{UB - LB}) \times \text{rand} + LB) \times \sigma \tag{7}$$

$$X_{t+1} = QF \times X_{best}(t) - X(t) \times \text{rand} \times (2 \times \text{rand} - 1) - \text{Levy}(D) \times 2 \times \left(1 - \frac{t}{T}\right) + \text{rand} \times (2 \times \text{rand} - 1) \tag{8}$$

$$\text{population} = \text{rand} \times (\text{UB} - \text{LB}) + \text{LB} \tag{9}$$

3.6 Performance metrics

There are various ways to evaluate how the classifier work including the Accuracy, Precision, Recall, Dice Coefficient, Specificity, and Cosine Similarity. The regions in the image are being positive (or negative) depending on the type of the data where the decision resulting from the
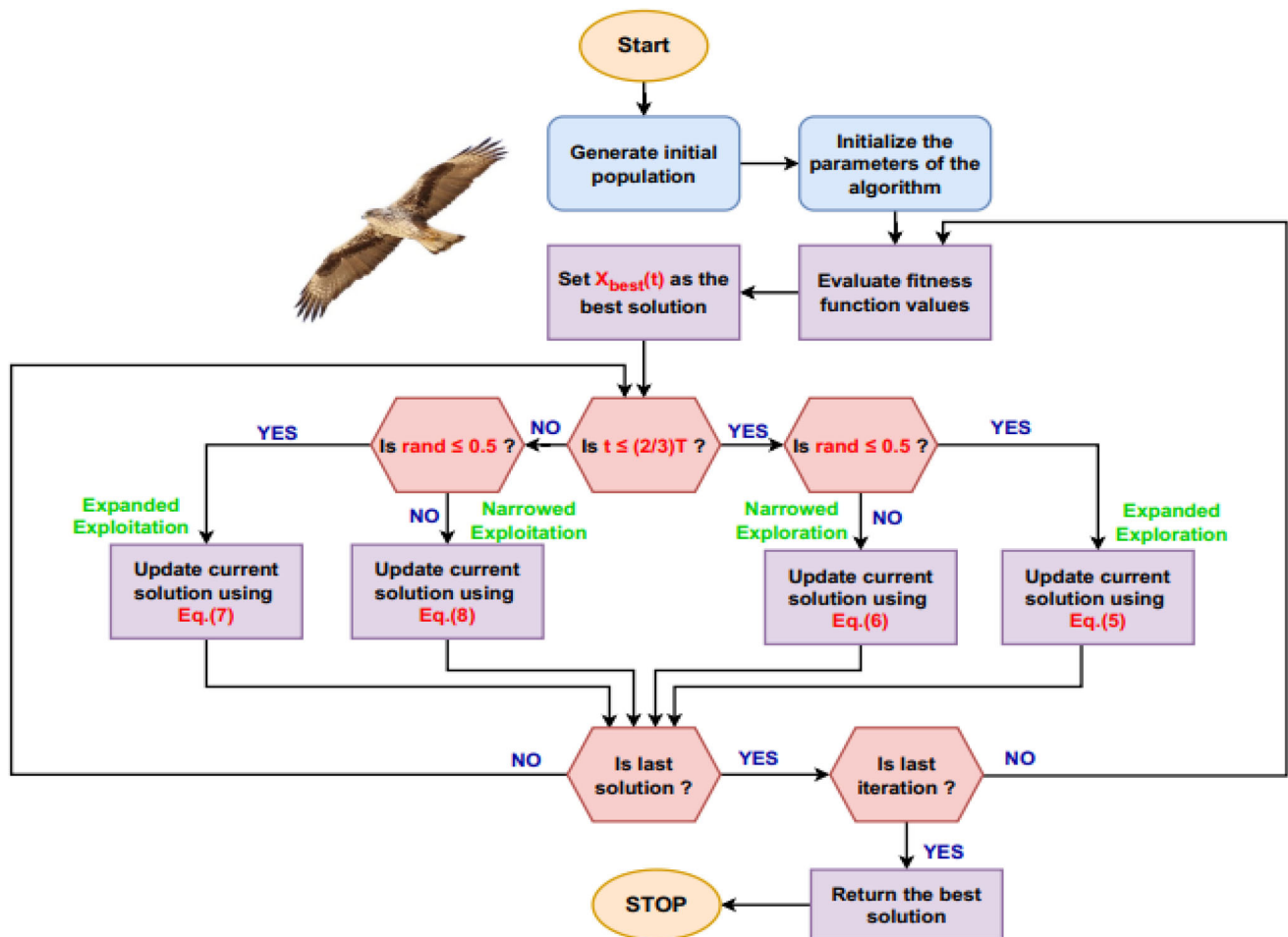
Fig. 3 The flowchart of AO

Table 5 The different utilized performance metrics

Metric	Definition	Equation
Accuracy	The model's ability for predicting correct classes effectively [122]	$\frac{T_P+T_N}{T_P+F_P+F_N+T_N}$
Area under ROC curve (AUC)	The average of each point in the ROC curve and also reflects the separability degree [123]	–
Precision	The ability of the classifier to predict correctly [124]	$\frac{T_P}{T_P+F_P}$
Recall or sensitivity	The ability to find positive examples and show how a change in an independent variable affects the dependent variable [125]	$\frac{T_P}{T_P+F_N}$
Dice coefficient or $F1$	The mean combination of recall and precision to rate the performance of the model [126]	$\frac{2*T_P}{2*T_P+F_P+F_N}$
Intersection over union (IoU)	Utilized for object segmentation by measuring the IoU for every class and make an average for all of them to know the similarity between the ground truth and the predicted region for the object inside a set of images [127]	$\frac{T_P}{T_P+F_P+F_N}$
Specifcity	The ability to detect a true negative value and identify people who did not have a specific condition or a disease [128]	$\frac{T_N}{F_P+T_N}$
Cosine similarity	Utilized to have better performance depend on the searching approach of detecting the similarities on the used dataset and its features easily [129]	$\frac{\sum_{i=1}^{n} A_i*B_i}{A*B}$

model can be true (i.e., correct) or false (i.e., incorrect) and it will be one of the possible classes which are True Positive (TP) which means that the model detected the malignant tumor (i.e., cancerous) as malignant, False Positive (FP) that refers to the benign tumor (i.e., non-cancerous) as malignant, True Negative (TN) which means that the model detected the benign tumor as benign, and False Negative (FN) which refers to the model detected the malignant tumor as benign. Table 5 shows the different utilized performance metrics in the current study.

3.7 Overall framework combination and pseudocode

The current subsection presents the framework combination and its pseudocode for grading and locating the prostate cancer. The pseudocode is presented in Algorithm 1.

Table 6 The segmentation performance metrics using the U-Net model

Grade	Loss	Accuracy	AUC	IoU	Dice
Grade 1	0.0729	96.56%	0.9888	0.9738	0.9761
Grade 2	0.0403	98.39%	0.9762	0.9840	0.9846
Grade 3	0.0196	99.36%	0.9740	0.9929	0.9931
Grade 4	0.0143	99.53%	0.9720	0.9954	0.9955
Grade 5	0.0039	99.92%	0.8978	0.9990	0.9990
Average	**0.0368**	**98.46%**	**0.9778**	**0.9865**	**0.9873**

Algorithm 1 Overall framework peseducode

```
// Step 1
1  Generate the initial population using Aquila optimizer based on Equation 9.
   // Step 2
2  Load, preprocess, and split the dataset. The split ratio is 85% for training and validation to 15% for testing.
   // Step 3
3  Loop on each solution in the population and apply the following inner steps.
   // Step 3.1
4  Map the solution cells to hyperparameters.
   // Step 3.2
5  Create the TL-CNN model with the mapped hyperparameters. The CNN model can be ResNet152, ResNet152V2, MobileNet,
   MobileNetV2, MobileNetV3Small, MobileNetV3Large, NASNet Mobile, and NASNet Large.
   // Step 3.3
6  Train, validate, and test the model.
   // Step 3.4
7  Calculate the different performance metrics and report the fitness function (i.e., accuracy).
   // Step 4
8  Sort the population according to the fitness scores in a descending order.
   // Step 5
9  Update the population using the AO optimizer using Equations 5 to 8.
   // Step 6
10 Repeat Step 3 until the completion of the iterations.
```

4 Experiments and discussions

The experiments are divided into two categories (1) experiments related to segmentation and (2) experiments related to optimization, learning, and classification. The used scripting language is Python and the used packages are Tensorflow, Keras, NumPy, OpenCV, and Matplotlib. The scripts are run on Google Colab with its GPU (i.e., Intel(R) Xeon(R) CPU @ 2.00 GHz, Tesla T4 16 GB GPU, CUDA v.11.2, and 12 GB RAM).

4.1 The "PANDA: resized train data (512 × 512)" dataset experiments

The segmentation performance metrics using the U-Net model are reported in Table 6. It shows the segmentation scores for each grade. It shows an average segmentation

Table 7 The best solutions combinations concerning each model using the "ISUP Grade-wise Prostate Cancer" dataset

Model	ResNet152	ResNet152V2	MobileNet	MobileNetV2	MobileNetV3Small	MobileNetV3Large	NASNetMobile	NASNetLarge
Loss	Categorical Crossentropy	KLDivergence	Poisson	Squared Hinge	KLDivergence	Poisson	KLDivergence	KLDivergence
Batch size	36	40	28	28	20	28	8	28
Dropout	23%	5%	34%	31%	19%	29%	45%	41%
TFlearn ratio	52%	48%	57%	89%	80%	51%	8%	64%
Optimizer	AdaGrad	AdaGrad	SGD	SGD Nesterov	SGD	AdaMax	SGD Nesterov	SGD
Scaling techniques	Standardization	Max-Absolute	Min–Max	Normalization	Standardization	Standardization	Min–Max	Max-Absolute
Apply augmentation	False	True	False	True	False	False	True	True
Rotation value	N/A	13	N/A	3	N/A	N/A	11	0
Width shift value	N/A	0.23	N/A	0.14	N/A	N/A	0.05	0.2
Height shift value	N/A	0.1	N/A	0.23	N/A	N/A	0.09	0.17
Shear value	N/A	0.2	N/A	0.05	N/A	N/A	0.11	0.23
Zoom value	N/A	0.01	N/A	0.16	N/A	N/A	0.14	0.12
Horizontal flip value	N/A	False	N/A	True	N/A	N/A	True	True
Vertical flip value	N/A	True	N/A	False	N/A	N/A	True	True
Brightness range	N/A	0.56–0.8	N/A	1.27–1.32	N/A	N/A	1.41–1.85	0.82–1.92

Table 8 The performance metrics concerning each model using the "ISUP Grade-wise Prostate Cancer" dataset

Model	ResNet152	ResNet152V2	MobileNet	MobileNetV2	MobileNetV3Small	MobileNetV3Large	NASNetMobile	NASNetLarge
Loss	0.487	0.706	0.267	0.967	1.254	0.330	1.305	1.441
TP	9,012	6,878	9,400	8,381	3,564	8,688	5,509	547
TN	51,812	52,161	51,973	50,957	51,209	51,306	49,851	52,883
FP	1,108	839	1,087	2,103	1,791	1,754	3,229	177
FN	1,572	3,722	1,212	2,231	7,036	1,924	5,107	10,065
Accuracy	86.97%	78.99%	88.91%	79.21%	52.63%	82.39%	58.55%	44.53%
F1	86.48%	73.82%	88.87%	79.39%	42.97%	82.41%	55.67%	8.99%
Precision	88.05%	87.41%	89.22%	79.84%	66.16%	83.02%	62.40%	48.81%
Recall	85.15%	64.89%	88.58%	78.98%	33.62%	81.87%	51.89%	5.15%
Sensitivity	85.15%	64.89%	88.58%	78.98%	33.62%	81.87%	51.89%	5.15%
Specificity	97.91%	98.42%	97.95%	96.04%	96.62%	96.69%	93.92%	99.67%
AUC	97.27%	95.36%	97.00%	92.01%	84.07%	94.78%	86.54%	79.76%
IoU	86.40%	69.97%	91.00%	84.48%	55.39%	86.19%	64.42%	46.62%
Dice	88.13%	74.20%	91.70%	85.31%	59.91%	87.21%	67.81%	51.03%
Precision	89.05%	89.13%	89.63%	79.94%	66.55%	83.20%	63.05%	75.55%
Cosine similarity	88.58%	80.81%	89.80%	80.58%	61.33%	84.11%	64.76%	54.14%
Hinge	89.27	96.23	87.48	90.68	103.38	89.73	99.43	107.82
Squared hinge	92.59	101.74	90.51	96.65	113.53	94.55	109.22	119.54
Logcosh error	0.015	0.026	0.013	0.026	0.047	0.022	0.045	0.054
Mean absolute error	0.059	0.129	0.042	0.073	0.200	0.064	0.161	0.245
Mean squared error	0.033	0.055	0.030	0.060	0.102	0.048	0.098	0.117
Mean squared logarithmic error	0.016	0.027	0.015	0.029	0.050	0.024	0.048	0.057
Root mean squared error	0.182	0.235	0.174	0.244	0.319	0.220	0.313	0.342
Mean IoU	0.418	0.417	0.431	0.481	0.417	0.486	0.417	0.417

Table 9 The best solutions combinations concerning each model using the "Transverse Plane Prostate Dataset" dataset

Model	ResNet152	ResNet152V2	MobileNet	MobileNetV2	MobileNetV3Small	MobileNetV3Large	NASNetMobile	NASNetLarge
Loss	KLDivergence	Squared Hinge	Poisson	Poisson	Categorical Hinge	Poisson	Squared Hinge	Squared Hinge
Batch size	24	36	28	16	12	24	8	28
Dropout	48%	7%	13%	59%	46%	29%	32%	20%
TF learn ratio	59%	21%	74%	19%	40%	47%	18%	7%
Optimizer	AdaGrad	SGD Nesterov	Adam	AdaMax	SGD Nesterov	RMSProp	SGD Nesterov	SGD Nesterov
Scaling techniques	Standardization	Standardization	Max-Absolute	Normalization	Standardization	Standardization	Max-Absolute	Standardization
Apply augmentation	False	False	True	True	True	False	True	False
Rotation value	N/A	N/A	27	38	34	N/A	6	N/A
Width shift value	N/A	N/A	0.16	0.04	0.06	N/A	0.18	N/A
Height shift value	N/A	N/A	0.06	0.04	0.03	N/A	0.16	N/A
Shear value	N/A	N/A	0.06	0.2	0.16	N/A	0.1	N/A
Zoom value	N/A	N/A	0.16	0.25	0.17	N/A	0.25	N/A
Horizontal flip value	N/A	N/A	False	False	True	N/A	False	N/A
Vertical flip value	N/A	N/A	True	True	False	N/A	False	N/A
Brightness range	N/A	N/A	1.29–1.92	0.6–0.82	0.72–0.82	N/A	0.73–1.24	N/A

Table 10 The performance metrics concerning each model using the "Transverse Plane Prostate Dataset" dataset

Model	ResNet152	ResNet152V2	MobileNet	MobileNetV2	MobileNetV3Small	MobileNetV3Large	NASNetMobile	NASNetLarge
Loss	0.005	0.504	0.500	0.501	0.294	0.603	0.554	0.514
TP	1,512	1,510	1,512	1,519	1,307	1,484	1,492	1,502
TN	1,512	1,510	1,512	1,519	1,307	1,484	1,492	1,502
FP	0	2	0	1	217	28	36	10
FN	0	2	0	1	217	28	36	10
Accuracy	100%	99.87%	100%	99.93%	85.76%	98.15%	97.64%	99.34%
F1	100%	99.87%	100%	99.93%	85.76%	98.15%	97.64%	99.34%
Precision	100%	99.87%	100%	99.93%	85.76%	98.15%	97.64%	99.34%
Recall	100%	99.87%	100%	99.93%	85.76%	98.15%	97.64%	99.34%
Sensitivity	100%	99.87%	100%	99.93%	85.76%	98.15%	97.64%	99.34%
Specificity	100%	99.87%	100%	99.93%	85.76%	98.15%	97.64%	99.34%
AUC	100%	99.93%	100%	100%	90.35%	98.79%	99.24%	99.92%
IoU	99.54%	99.74%	99.96%	99.86%	89.49%	98.74%	97.12%	99.28%
Dice	99.68%	99.79%	99.97%	99.89%	90.20%	98.75%	97.62%	99.39%
Precision	100%	99.87%	100%	99.93%	85.76%	98.15%	97.64%	99.34%
Cosine similarity	99.97%	99.87%	100%	99.92%	87.33%	98.15%	98.02%	99.42%
Hinge	50.48	50.31	50.04	50.17	64.71	51.87	53.57	50.92
Squared hinge	50.52	50.44	50.05	50.24	76.89	53.69	55.44	51.45
Logcosh error	0.000	0.001	0.000	0.000	0.054	0.008	0.009	0.002
Mean absolute error	0.005	0.003	0.000	0.002	0.147	0.019	0.036	0.009
Mean squared error	0.000	0.001	0.000	0.001	0.122	0.018	0.019	0.005
Mean squared logarithmic error	0.000	0.001	0.000	0.000	0.059	0.009	0.009	0.003
Root mean squared error	0.020	0.036	0.005	0.027	0.349	0.135	0.137	0.073
Mean IoU	0.265	0.316	0.600	0.817	0.295	0.970	0.316	0.307
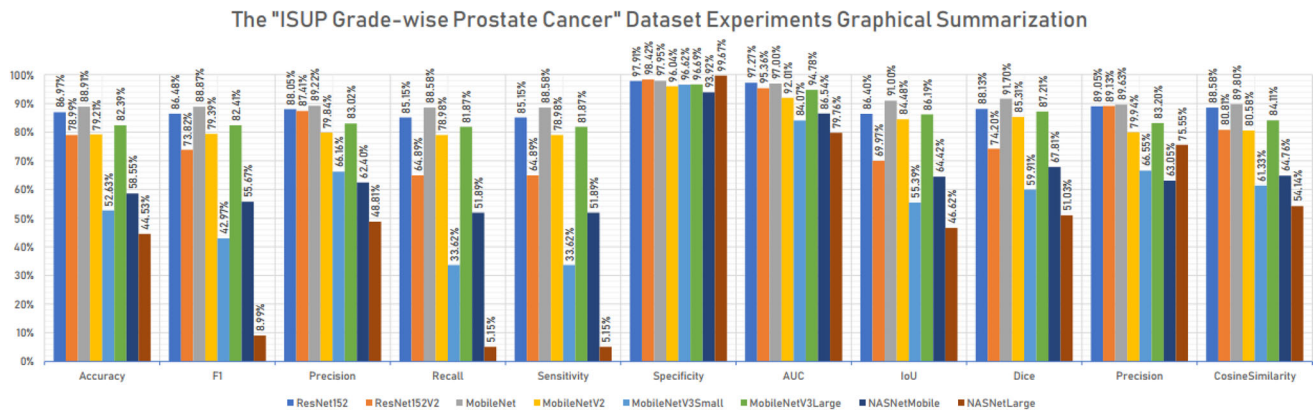
Fig. 4 Summarization of the learning an optimization experiments related to the "ISUP Grade-wise Prostate Cancer" dataset
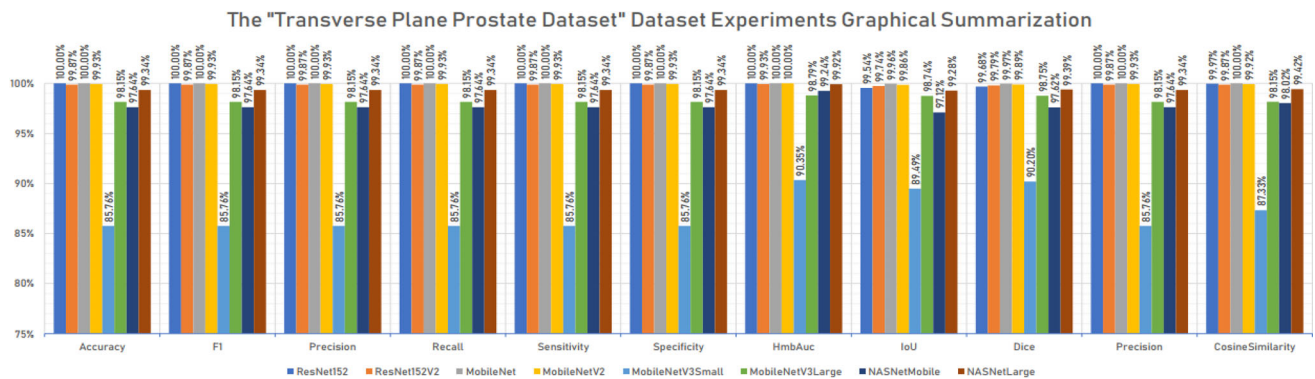

Fig. 5 Summarization of the learning and optimization experiments related to the "Transverse Plane Prostate Dataset" dataset

accuracy of 98.46% using the "PANDA: Resized Train Data (512 × 512)" dataset. The average loss is 0.0368, AUC is 0.9778, IoU is 0.9865, and Dice is 0.9873.

4.2 The "ISUP grade-wise prostate cancer" dataset experiments

Table 7 presents the finest combinations in each CNN structure using the "ISUP Grade-wise Prostate Cancer" dataset. These combinations have been optimized using AO. From this table, it can be noticed that KLDivergence loss is preferred by four models, the Standardization scaling technique and SGD parameters optimization are recommended by three models. In Table 8, the corresponding performance metrics for these finest combinations after the learning and optimization processes for each CNN model are presented. The best TP, TN, FP, and FN are 9,400, 52,883, 177, and 1,212, respectively. The best reported Accuracy, F1, Precision, Recall, Sensitivity, AUC, IoU, Dice, Precision, and Cosine Similarity are 88.91%, 88.87%, 89.22%, 88.58%, 88.58%, 99.67%, 97.27%, 91.00%, 91.70%, 89.63%, and 89.80%, respectively, using MobileNet pretrained model.

4.3 The "transverse plane prostate dataset" dataset experiments

Table 9 presents the finest combinations in each CNN structure using the "Transverse Plane Prostate Dataset" dataset. These combinations have been optimized using AO. From this table, it can be noticed that the Squared Hinge and Poisson loss are preferred by three models, the standardization scaling technique is recommended by five models, and SGD Nesterov parameters optimization is recommended by four models. In Table 10, the corresponding performance metrics for the finest combinations after optimizing the CNN models are given. The best TP, TN, FP, and FN are 1,519, 1,519, 0, and 0, respectively. The best reported Accuracy, F1, Precision, Recall, Sensitivity, Specificity, AUC, IoU, Dice, Precision, and Cosine Similarity are 100%, 100%, 100%, 100%, 100%, 100%, 100%, 99.96%, 99.97%, 100%, and 100%, respectively, using MobileNet pretrained model.

4.4 Graphical summarizations

Figure 4 is a graphical representation of the classification results obtained for the "ISUP Grade-wise Prostate

Table 11 Comparison between the suggested approach and related studies

Study	Year	Dataset	Best performance
Kiraly et al. [130]	2017	ProstateX challenge Dataset	0.9 of AUC
Liu et al. [20]	2017	ProstateX challenge Dataset	0.84 of AUC
Mehrtash et al. [131]	2017	ProstateX challenge Dataset	0.8 of AUROC
Song et al. [132]	2018	ProstateX challenge Dataset	0.944 of AUC, 87% of Sensitivity, and 90.6% of Specificity
Kwon et al. [133]	2018	ProstateX challenge Dataset	0.82 of AUC
Liu et al. [134]	2019	ProstateX challenge Dataset	86% of DSC
Nirthika et al. [135]	2020	ISUP prostate cancer Dataset	0.792 of QWK
Wang & Wang [70]	2020	ProstateX challenge Dataset	0.95 of AUROC
Aldoj et al. [136]	2020	ProstateX challenge Dataset	0.91 of AUC, 81.2% of Sensitivity, and 90.5% of Specificity
Malyshev et al. [59]	2021	ProstateX challenge Dataset	92.77% using CNN
John et al. [39]	2021	ProstateX challenge Dataset	0.931 of AUC and 0.93 of F1-Score
Pellicer-Valero et al. [54]	2021	ProstateX challenge Dataset	DSC in the range between 0.894 to 0.941
Winkel et al. [137]	2021	ProstateX challenge Dataset	0.88 of AUC
Hiremath et al. [138]	2021	ProstateX challenge Dataset	0.83 of AUC
Lai et al. [139]	2021	ProstateX challenge Dataset	90.45% of DSC
Yang and Xiao [37]	2021	PANDA challenge Dataset	0.8503 of Quadratic Weighted Kappa (QWK)
Current study	2024	"PANDA: Resized Train Data (512 × 512)", "ISUP Grade-wise Prostate Cancer", and "Transverse Plane Prostate Dataset"	88.91% for "ISUP Grade-wise Prostate Cancer", 100% for "Transverse Plane Prostate Dataset", and 98.46% for "PANDA: Resized Train Data (512 × 512)"

Cancer" dataset. It shows that the MobileNet gives the best performance according to the different metrics. Figure 5 is a graphical representation of the classification results obtained for the "Transverse Plane Prostate Dataset" dataset. It shows also that the MobileNet and ResNet152 are the best models.

4.5 Related studies comparisons

Table 11 shows a comparison between the suggested approach and related studies.

5 Conclusions, limitations, and future work

Prostate cancer is a common type of cancer worldwide. The mortality and morbidity of this type of cancer have increased dramatically in the last few years. Therefore, early and accurate diagnosis of prostate cancer is very important. In this study, we propose a hybrid framework for early and accurate classification and segmentation of prostate cancer using deep learning. This framework consists of two stages, namely classification stage and segmentation stage. In the classification stage, eight different CNN architectures via transfer learning were first fine-tuned using Aquila optimizer and then applied to diagnose patients with prostate cancer from normal ones. Once the patient is diagnosed with prostate cancer, images from the patient are passed to the segmentation phase in order to identify regions of interest. The importance of this phase is to find the shape, size, and volume of the tumor in order to help physicians in applying the correct treatment. The proposed framework is trained on three different datasets in order to generalize the proposed framework on different types of data. The proposed framework could achieve classification accuracies of 88.91% for the "ISUP Grade-wise Prostate Cancer" dataset and 100% for the "Transverse Plane Prostate Dataset" dataset. It shows an average segmentation accuracy of 98.46% using the "PANDA: Resized Train Data (512 × 512)" dataset.

Although the results of the study are promising, there are still some.

5.1 Limitations

For example, only one dataset is used for segmentation. The choice of only eight transfer learning models from all the available models is another limitation. Finally, the use of U-Net only for segmentation is among the limitations.

Despite these limitations, the results of the current study are encouraging.

As a future work, different CNN architectures can be applied in the proposed framework. Other metaheuristic optimizers can be applied such as sparrow search algorithm. Finally, the proposed framework can be applied for diagnosis of other types of tumors such as brain tumor.

Declarations

Conflict of interest No conflict of interest exists. We wish to confirm that, there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Consent to participate There is no informed consent for the current study.

Consent for publication Not Applicable.

Ethical approval The current study does not contain any studies with human participants and/or animals performed by any of the authors.

Intellectual property We confirm that, we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, concerning intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

Research ethics We further confirm, if existing, that any aspect of the work covered in this manuscript that has involved human patients has been conducted with the ethical approval of all relevant bodies and that such approvals are acknowledged within the manuscript. Written consent to publish potentially identifying information, such as details of the case and photographs, was obtained from the patient(s) or their legal guardian(s). **Authorship** We confirm that the manuscript has been read and approved by all named authors. We confirm that the order of authors listed in the manuscript has been approved by all named authors.

References

1. Rawla P (2019) Epidemiology of prostate cancer. World J Oncol 10(2):63
2. Sung H et al (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clinic 71(3):209–249
3. W (2020) I health organization international agency for research on cancer "health organization international agency for research on cancer (IARC), Globocan 2020 Estimated cancer incidence, mortality, and prevalence on egypt in 2020. https://gco.iarc.fr/today/data/factsheets/populations/818-egypt-fact-sheets.pdf. [Online; accessed 01-Oct-2021]."
4. Cha H-R, Lee JH, Ponnazhagan S (2020) Revisiting immunotherapy: a focus on prostate canceradvances and limitations of immunotherapy in prostate cancer. Can Res 80(8):1615–1623
5. Hoskin P, Neal AJ, Hoskin PJ (2009) Clinical oncology: basic principles and practice. CRC Press, UK
6. Bryant RJ, Hamdy FC (2008) Screening for prostate cancer: an update. Eur Urol 53(1):37–44
7. Sekhoacha M, Riet K, Motloung P, Gumenku L, Adegoke A, Mashele S (2022) Prostate cancer review: genetics, diagnosis, treatment options, and alternative approaches. Molecules 27(17):5730
8. Kensler KH, Rebbeck TR (2020) Cancer progress and priorities: prostate cancer. Cancer Epidemiol Biomark Prev 29(2):267–277
9. Donovan JL et al (2016) Patient-reported outcomes after monitoring, surgery, or radiotherapy for prostate cancer. N Engl J Med 375:1425–1437
10. Hamdy FC et al (2020) Active monitoring, radical prostatectomy and radical radiotherapy in PSA-detected clinically localised prostate cancer: the protect three-arm RCT. Health Technol Assess 24(37):1
11. Balk SP, Ko Y-J, Bubley GJ (2003) Biology of prostate-specific antigen. J Clin Oncol 21(2):383–391
12. Chang CM, McIntosh AG, Shapiro DD, Davis JW, Ward JF, Gregg JR (2021) Does a screening digital rectal exam provide actionable clinical utility in patients with an elevated PSA and positive MRI? BJUI Compass 2(3):188–193
13. Naji L et al (2018) Digital rectal examination for prostate cancer screening in primary care: a systematic review and meta-analysis. Ann Family Med 16(2):149–154
14. Shariat SF, Roehrborn CG (2008) Using biopsy to detect prostate cancer. Reviews in Urology 10(4):262
15. Kasivisvanathan V et al (2018) MRI-targeted or standard biopsy for prostate-cancer diagnosis. N Engl J Med 378(19):1767–1777
16. Noble SM et al (2020) The protect randomised trial cost-effectiveness analysis comparing active monitoring, surgery, or radiotherapy for prostate cancer. Br J Cancer 123(7):1063–1070
17. Sutton E et al (2021) Men's experiences of radiotherapy treatment for localized prostate cancer and its long-term treatment side effects: a longitudinal qualitative study. Cancer Causes Control 32(3):261–269

18. Swami U, McFarland TR, Nussenzveig R, Agarwal N (2020) Advanced prostate cancer: treatment advances and future directions. Trends Cancer 6(8):702–715

19. Dlamini Z, Francies FZ, Hull R, Marima R (2020) Artificial intelligence (AI) and big data in cancer and precision oncology. Comput Struct Biotechnol J 18:2300–2311

20. Liu Y and An X (2017) "A classification model for the prostate cancer based on deep learning". IEEE pp. 1–6

21. Reda I et al (2016) "Computer-aided diagnostic tool for early detection of prostate cancer". IEEE pp. 2668–2672

22. García J G, Colomer A, López-Mir F, Mossi J M, and Naranjo V (2019) "Computer aid-system to identify the first stage of prostate cancer through deep-learning techniques". IEEE pp. 1–5

23. Salman ME, Çakar GC, Azimjonov J, Kösem M, Cedimoğlu IH, (2022) Automated prostate cancer grading and diagnosis system using deep learning-based Yolo object detection algorithm. Expert Syst Appl 201:117148

24. Rabilloud N et al (2023) Deep learning methodologies applied to digital pathology in prostate cancer: a systematic review. Diagnostics 13(16):2676

25. He M et al (2023) Research progress on deep learning in magnetic resonance imaging–based diagnosis and treatment of prostate cancer: a review on the current status and perspectives. Front Oncol 13:1189370

26. Zhang L, Li L, Tang M, Huan Y, Zhang X, Zhe X (2021) A new approach to diagnosing prostate cancer through magnetic resonance imaging. Alex Eng J 60(1):897–904

27. Erdem E, Bozkurt F (2021) A comparison of various supervised machine learning techniques for prostate cancer prediction. Avrupa Bilim ve Teknoloji Dergisi 21:610–620

28. Nayan N et al. (2022) "A machine learning approach to predict progression on active surveillance for prostate cancer." vol. 40: Elsevier, 4 ed., pp. 161–e1

29. Gentile F et al (2021) Optimized identification of high-grade prostate cancer by combining different PSA molecular forms and PSA density in a deep learning model. Diagnostics 11(2):335

30. Shrestha S, Alsadoon A, Prasad PWC, Seher I, Alsadoon OH (2021) A novel solution of using deep learning for prostate cancer segmentation: enhanced batch normalization. Multimed Tools Appl 80(14):21293–21313

31. Khosravi P et al (2021) A deep learning approach to diagnostic classification of prostate cancer using pathology–radiology fusion. J Magn Reson Imaging 54(2):462–471

32. Wessels F et al (2021) Deep learning approach to predict lymph node metastasis directly from primary tumour histology in prostate cancer. BJU Int 128(3):352–360

33. Linkon AHM, Labib MM, Hasan T, Hossain M (2021) Deep learning in prostate cancer diagnosis and Gleason grading in histopathology images: an extensive study. Inform Med Unlocked 24:100582

34. Patel A, Singh SK, Khamparia A (2021) Detection of prostate cancer using deep learning framework. In: InIOP Conference Series: Materials Science and Engineering (Vol. 1022(1), p. 012073). IOP Publishing

35. Shao W et al (2021) ProsRegNet: a deep learning framework for registration of MRI and histopathology images of the prostate. Med Image Anal 68:101919

36. Amarsee K et al (2021) Automatic detection and tracking of marker seeds implanted in prostate cancer patients using a deep learning algorithm. J Med Phys 46(2):80

37. Yang H, Wu G, Shen D, and Liao S (2021) "Automatic prostate cancer detection on multi-parametric mri with hierarchical weakly supervised learning". IEEE pp. 316–319

38. Kovalev VA, Voynov DM, Malyshau VD, Lapo ED (2020) Computerized diagnosis of prostate cancer based on whole slide histology images and deep learning methods. InInformatics 17(4):48–60

39. John J, Ravikumar A, Abraham B (2021) Prostate cancer prediction from multiple pretrained computer vision model. Heal Technol 11(5):1003–1011

40. Comelli A et al (2021) Deep learning-based methods for prostate segmentation in magnetic resonance imaging. Appl Sci 11(2):782

41. Pinckaers H, Bulten W, van der Laak J, Litjens G (2021) Detection of prostate cancer in whole-slide images through end-to-end training with image-level labels. IEEE Trans Med Imaging 40(7):1817–1826

42. Salvi M et al (2021) A hybrid deep learning approach for gland segmentation in prostate histopathological images. Artif Intell Med 115:102076

43. Korevaar S et al (2021) Incidental detection of prostate cancer with computed tomography scans. Sci Rep 11(1):1–10

44. Chahal ES, Patel A, Gupta A, Purwar A (2022) Unet based xception model for prostate cancer segmentation from MRI images. Multimed Tools Appl 81(26):37333–37349

45. Sobecki P, Jóźwiak R, Sklinda K, Przelaskowski A (2021) Effect of domain knowledge encoding in CNN model architecture—a prostate cancer study using mpMRI images. PeerJ 9:e11006

46. Balagopal A et al (2021) A deep learning-based framework for segmenting invisible clinical target volumes with estimated uncertainties for post-operative prostate cancer radiotherapy. Med Image Anal 72:102101

47. Liu Z, Yang C, Huang J, Liu S, Zhuo Y, Lu X (2021) Deep learning framework based on integration of S-Mask R-CNN and Inception-v3 for ultrasound image-aided diagnosis of prostate cancer. Futur Gener Comput Syst 114:358–367

48. Ambroa EM, Pérez-Alija J, Gallego P (2021) Convolutional neural network and transfer learning for dose volume histogram prediction for prostate cancer radiotherapy. Med Dosim 46(4):335–341

49. Hao R, Namdar K, Liu L, Haider MA, Khalvati F (2021) A comprehensive study of data augmentation strategies for prostate cancer detection in diffusion-weighted MRI using convolutional neural networks. J Digit Imaging 34(4):862–876

50. Kudo MS et al (2021) The potential of convolutional neural network diagnosing prostate cancer. Res Biomed Eng 37(1):25–31

51. Mehta P, Antonelli M, Ahmed HU, Emberton M, Punwani S, Ourselin S (2021) Computer-aided diagnosis of prostate cancer using multiparametric MRI and clinical features: a patient-level classification framework. Med Image Anal 73:102153

52. Hoar D et al (2021) Combined transfer learning and test-time augmentation improves convolutional neural network-based semantic segmentation of prostate cancer from multi-parametric MR images. Comput Methods Programs Biomed 210:106375

53. Cipollari S et al (2022) Convolutional neural networks for automated classification of prostate multiparametric magnetic resonance imaging based on image quality. J Magn Reson Imaging 55(2):480–490

54. Pellicer-Valero OJ et al (2021) Deep learning for fully automatic detection, segmentation, and gleason grade estimation of prostate cancer in multiparametric magnetic resonance Images. Sci Rep 12(1):1–13

55. Saunders SL, Leng E, Spilseth B, Wasserman N, Metzger GJ, Bolan PJ (2021) Training convolutional networks for prostate segmentation with limited data. IEEE Access 9:109214–109223

56. Han S, Oh JS, Lee JJ (2022) Diagnostic performance of deep learning models for detecting bone metastasis on whole-body

bone scan in prostate cancer. Eur J Nucl Med Mol Imaging 49(2):585–595

57. Iqbal S et al (2021) Prostate cancer detection using deep learning and traditional techniques. IEEE Access 9:27085–27100

58. Salama WM, Aly MH (2021) Prostate cancer detection based on deep convolutional neural networks and support vector machines: a novel concern level analysis. Multimed Tools Appl 80(16):24995–25007

59. Malyshev V, Voynov D, & Lapo E (2021) "Computerized diagnosis of prostate cancer based on whole slide histology images and deep learning methods"

60. Abdelmaksoud IR et al (2021) Precise identification of prostate cancer from DWI using transfer learning. Sensors 21(11):3664

61. Bulten W et al. (2020) "The PANDA challenge: prostate cancer grade assessment using the gleason grading system," MICCAI challenge

62. Xhlulu, "Panda: Resized train data (512x512). https://www.kaggle.com/datasets/xhlulu/," 2020

63. Islam TN (2020) "isup_grade_wise_prostate_cancer. https://www.kaggle.com/datasets/tasnimnishatislam/isup-grade-wise-prostate-cancer

64. Van Leenders GJLH et al (2020) The 2019 international society of urological pathology (ISUP) consensus conference on grading of prostatic carcinoma. Am J Surg Pathol 44(8):e87

65. Prostata T (2021) "Transverse plane prostate dataset. https://www.kaggle.com/datasets/tgprostata/transverse-plane-prostate-dataset

66. Tang S, Yuan S, Zhu Y (2020) Data preprocessing techniques in convolutional neural network based on fault diagnosis towards rotating machinery. IEEE Access 8:149487–149496

67. Lashgari E, Liang D, Maoz U (2020) Data augmentation for deep-learning-based electroencephalography. J Neurosci Methods 346:108885

68. Mushtaq Z, Su S-F, Tran Q-V (2021) Spectral images based environmental sound classification using CNN with meaningful data augmentation. Appl Acoust 172:107581

69. Noguchi S, Nishio M, Yakami M, Nakagomi K, Togashi K (2020) Bone segmentation on whole-body CT using convolutional neural network with novel data augmentation techniques. Comput Biol Med 121:103767

70. Wang H et al (2020) Hard exudate detection based on deep model learned information and multi-feature joint representation for diabetic retinopathy screening. Comput Methods Programs Biomed 191:105398

71. Ahsan MM, Mahmud MAP, Saha PK, Gupta KD, Siddique Z (2021) Effect of data scaling methods on machine learning algorithms and model performance. Technologies 9(3):52

72. Cao XH, Stojkovic I, Obradovic Z (2016) A robust data scaling algorithm to improve classification accuracies in biomedical data. BMC Bioinformatics 17(1):1–10

73. Shaheen H, Agarwal S, Ranjan P (2020) MinMaxScaler binary PSO for feature selection. Springer, Cham, pp 705–716

74. Bhanja S and Das A (2018) "Impact of data normalization on deep neural network for time series forecasting". arXiv preprint arXiv:1812.05519

75. Ichimura S and Zhao Q (2019) "Route-based ship classification". IEEE pp. 1–6

76. Chen LC, Papandreou G, Kokkinos I, Murphy K, and Yuille AL (2014) "Semantic image segmentation with deep convolutional nets and fully connected crfs". arXiv preprint arXiv:1412.7062

77. Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL (2017) DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Trans Pattern Anal Mach Intell 40(4):834–848

78. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. Springer, Cham, pp 234–241

79. Siddique N, Paheding S, Elkin CP, Devabhaktuni V (2021) U-net and its variants for medical image segmentation: a review of theory and applications. Ieee Access 9:82031–82057

80. Balaha HM, El-Gendy EM, Saafan MM (2021) CovH2SD: A COVID-19 detection approach based on Harris Hawks optimization and stacked deep learning. Expert Syst Appl 186:115805

81. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444

82. Maier A, Syben C, Lasser T, Riess C (2019) A gentle introduction to deep learning in medical image processing. Z Med Phys 29(2):86–101

83. Bodapati JD, Veeranjaneyulu N (2019) Feature extraction and classification using deep convolutional neural networks. J Cyber Sec Mobil 2019:261–276

84. Wang S-H, Phillips P, Sui Y, Liu B, Yang M, Cheng H (2018) Classification of Alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling. J Med Syst 42(5):1–11

85. Balaha HM, Antar ER, Saafan MM, El-Gendy EM (2023) A comprehensive framework towards segmenting and classifying breast cancer patients using deep learning and Aquila optimizer. J Ambient Intell Human Comput 14(6):7897–7917. https://doi.org/10.1007/s12652-023-04600-1

86. Balaha MM, El-Kady S, Balaha HM, Salama M, Emad E, Hassan M, Saafan MM (2023) A vision-based deep learning approach for independent-users Arabic sign language interpretation. Multimed Tools Appl 82(5):6807–6826. https://doi.org/10.1007/s11042-022-13423-9

87. Talaat FM, El-Gendy EM, Saafan MM, Gamel SA (2023) Utilizing social media and machine learning for personality and emotion recognition using PERS. Neural Comput Appl 35(33):23927–23941. https://doi.org/10.1007/s00521-023-08962-7

88. Baghdadi NA, Malki A, Abdelaliem SF, Balaha HM, Badawy M, Elhosseini M (2022) An automated diagnosis and classification of COVID-19 from chest CT images using a transfer learning-based convolutional neural network. Comput Biol Med 144:105383

89. Acharya UR et al (2017) A deep convolutional neural network model to classify heartbeats. Comput Biol Med 89:389–396

90. Bahgat WM, Balaha HM, AbdulAzeem Y, Badawy MM (2021) An optimized transfer learning-based approach for automatic diagnosis of COVID-19 from chest x-ray images. PeerJ Computer Science 7:e555

91. Ma W and Lu J (2017) "An equivalence of fully connected layer and convolutional layer". arXiv preprint arXiv:1712.01252

92. Bjorck N, Gomes CP, Selman B, Weinberger KQ (2018) Understanding batch normalization. Adv Neural Inform Process Syst 31

93. Sharma S, Athaiya A (2017) Activation functions in neural networks. Towards Data Sci 6(12):310–316

94. Abdulazeem Y, Balaha HM, Bahgat WM, Badawy M (2021) Human action recognition based on transfer learning approach. IEEE Access 9:82058–82069

95. Shin H-C et al (2016) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans Med Imaging 35(5):1285–1298

96. Russakovsky O et al (2015) Imagenet large scale visual recognition challenge. Int J Comput Vision 115(3):211–252

97. Sun R (2019) "Optimization for deep learning: theory and algorithms". arXiv preprint arXiv:1912.08957

98. Balaha HM, El-Gendy EM, Saafan MM (2022) A complete framework for accurate recognition and prognosis of COVID-19 patients based on deep transfer learning and feature classification approach. Artif Intell Rev 55(6):5063–5108

99. Balaha HM, Shaban AO, El-Gendy EM, Saafan MM (2022) A multi-variate heart disease optimization and recognition framework. Neural Comput Appl 34(18):15907–15944

100. Vani S and Rao TVM (2019) "An experimental approach towards the performance assessment of various optimizers on convolutional neural network". IEEE pp. 331–336

101. Feurer M, Hutter F (2019) "Hyperparameter optimization". In automated machine learning. Springer, Cham, pp 3–33

102. Balaha HM, Hassan AE, El-Gendy EM, ZainEldin H, Saafan MM (2023) An aseptic approach towards skin lesion localization and grading using deep learning and Harris Hawks optimization. Multimed Tools Appl 28:1–29

103. Xu J, Zhang Z, Friedman T, Liang Y, and Broeck G (2018) "A semantic loss function for deep learning with symbolic knowledge". PMLR pp. 5502–5511

104. Kavalerov I, Czaja W, and Chellappa R (2021) "A multi-class hinge loss for conditional gans". pp. 1290–1299

105. Singh SK, Singh U, Kumar M (2014) Estimation for the parameter of poisson-exponential distribution under Bayesian paradigm. J Data Sci 12(1):157–173

106. Bach S, Huang B, London B, and Getoor L (2013) "Hinge-loss Markov random fields: Convex inference for structured prediction". arXiv preprint arXiv:1309.6813

107. Zhang Z and Sabuncu M (2018) "Generalized cross entropy loss for training deep neural networks with noisy labels". Adv Neural Inform Process Syst vol. 31

108. Wu Y, Liu Y (2007) Robust truncated hinge loss support vector machines. J Am Stat Assoc 102(479):974–983

109. Yu D, Yao K, Su H, Li G, and Seide F (2013) "KL-divergence regularized deep neural network adaptation for improved large vocabulary speech recognition". IEEE pp. 7893–7897

110. He F, Liu T, Tao D (2019) Control batch size and learning rate to generalize well: theoretical and empirical evidence. Adv Neural Inform Process Syst 32

111. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. J Mach Learning Res 15(1):1929–1958

112. Balaha HM, Balaha MH, Ali HA (2021) Hybrid COVID-19 segmentation and recognition framework (HMB-HCF) using deep learning and genetic algorithms. Artif Intell Med 119:102156

113. Saafan MM, El-Gendy EM (2021) IWOSSA: an improved whale optimization salp swarm algorithm for solving optimization problems. Expert Syst Appl 15(176):114901. https://doi.org/10.1016/j.eswa.2021.114901

114. Fahmy H, El-Gendy EM, Mohamed MA, Saafan MM (2023) ECH3OA: an enhanced chimp-Harris Hawks optimization algorithm for copyright protection in color images using watermarking techniques. Knowl Syst 7(269):110494. https://doi.org/10.1016/j.knosys.2023.110494

115. Saafan MM, Abdelsalam MM, Elksas MS, Saraya SF, Areed FF (2017) An adaptive neuro-fuzzy sliding mode controller for MIMO systems with disturbance. Chin J Chem Eng 25(4):463–476. https://doi.org/10.1016/j.cjche.2016.07.021

116. El-Gendy EM, Saafan MM, Elksas MS, Saraya SF, Areed FF (2019) New suggested model reference adaptive controller for the divided wall distillation column. Indus Eng Chem Res 58(17):7247–7264. https://doi.org/10.1021/acs.iecr.9b01747

117. Balaha HM, Saafan MM (2021) Automatic exam correction framework (AECF) for the MCQs, essays, and equations matching. IEEE Access 9:32368–32389. https://doi.org/10.1109/ACCESS.2021.3060940

118. Badr AA, Saafan MM, Abdelsalam MM, Haikal AY (2023) Novel variants of grasshopper optimization algorithm to solve numerical problems and demand side management in smart grids. Artif Intell Rev 56(10):10679–10732. https://doi.org/10.1007/s10462-023-10431-5

119. Gandomi AH, Yang XS, Talatahari S, Alavi AH (2013) Meta-heuristic algorithms in modeling and optimization. Meta-heuristic Appl Struct Infrastruct 1:1–24

120. El-Gendy EM, Saafan MM, Elksas MS, Saraya SF, Areed FF (2020) Applying hybrid genetic–PSO technique for tuning an adaptive PID controller used in a chemical process. Soft Comput 24:3455–3474

121. Abualigah L, Yousri D, Abd Elaziz M, Ewees AA, Al-Qaness MA, Gandomi AH (2021) Aquila optimizer: a novel meta-heuristic optimization algorithm. Comput Indus Eng 1(157):107250

122. Kim B, Im S, Yoo G (2020) Performance evaluation of CNN-based end-point detection using in-situ plasma etching data. Electronics 10(1):49

123. Moon H-C, Lee H-Y, Kim J-G (2020) Compression and performance evaluation of CNN models on embedded board. J Broadcast Eng 25(2):200–207

124. Zemčík T, Kratochvíla L, Bilík Š, Boštík O, Zemčík P, Horák K (2021) Performance evaluation of CNN based pedestrian and cyclist detectors on degraded images. Int J Image Process (IJIP) 15(1):1

125. Buckland M, Gey F (1994) The relationship between recall and precision. J Am Soc Inform Sci 45(1):12–19

126. Shamir RR, Duchin Y, Kim J, Sapiro G, and Harel H (2019) "Continuous dice coefficient: a method for evaluating probabilistic segmentations". arXiv preprint arXiv:1906.11031

127. Rahman MA, Wang Y (2016) Optimizing intersection-over-union in deep neural networks for image segmentation. Springer, Cham, pp 234–244

128. Al-Toubah T, Cives M, Valone T, Blue K, Strosberg J (2021) Sensitivity and specificity of the NETest: a validation study. Neuroendocrinology 111(6):580–585

129. Li B, Han L (2013) Distance weighted cosine similarity measure for text classification. Springer, Cham, pp 611–618

130. Kiraly AP et al (2017) Deep convolutional encoder-decoders for prostate cancer detection and classification. In: Descoteaux M, Maier-Hein L, Franz A, Jannin P, Collins D, Duchesne S (eds) Medical Image Computing and Computer Assisted Intervention − MICCAI 2017. Springer, Cham. https://doi.org/10.1007/978-3-319-66179-7_56

131. Mehrtash A, Sedghi A, Ghafoorian M, Taghipour M, Tempany CM, Wells III WM, Kapur T, Mousavi P, Abolmaesumi P, Fedorov A (2017) Classification of clinical significance of MRI prostate findings using 3D convolutional neural networks. Proc. SPIE 10134, Medical imaging 2017: computer-aided diagnosis, 101342A. https://doi.org/10.1117/12.2277123

132. Song Y, Zhang YD, Yan X, Liu H, Zhou M, Hu B, Yang G (2018) Computer-aided diagnosis of prostate cancer using a deep convolutional neural network from multiparametric MRI. J Magn Reson Imaging 48:1570–1577. https://doi.org/10.1002/jmri.26047

133. Kwon D, Reis IM, Breto AL, Tschudi Y, Gautney N, Zavala-Romero O, Lopez C, Ford JC, Punnen S, Pollack A, Stoyanova R (2018) Classification of suspicious lesions on prostate multiparametric MRI using machine learning. J Med Imaging 5(3):034502. https://doi.org/10.1117/1.JMI.5.3.034502

134. Liu Y et al (2019) Automatic Prostate Zonal Segmentation Using Fully Convolutional Network With Feature Pyramid Attention. In: IEEE Access, vol. 7, pp. 163626-163632. https://doi.org/10.1109/ACCESS.2019.2952534

135. Nirthika R, Manivannan S, Ramanan A (2020) Loss functions for optimizing Kappa as the evaluation measure for classifying diabetic retinopathy and prostate cancer images, In: IEEE 15th International Conference on Industrial and Information Systems (ICIIS), RUPNAGAR, India, pp. 144–149. https://doi.org/10.1109/ICIIS51140.2020.9342711

136. Aldoj N, Lukas S, Dewey M et al (2020) Semi-automatic classification of prostate cancer on multi-parametric MR imaging using a multi-channel 3D convolutional neural network. Eur Radiol 30:1243–1253. https://doi.org/10.1007/s00330-019-06417-z

137. Winkel DJ, Tong A, Lou B, Kamen A, Comaniciu D, Disselhorst JA, Rodríguez-Ruiz A, Huisman H, Szolar D, Shabunin I, Choi MH, Xing P, Penzkofer T, Grimm R, von Busch H, Boll DT (2021) A novel deep learning based computer-aided diagnosis system improves the accuracy and efficiency of radiologists in reading biparametric magnetic resonance images of the prostate: results of a multireader, multicase study. Invest Radiol 56(10):605–13. https://doi.org/10.1097/RLI.0000000000000780

138. Hiremath A, Shiradkar R, Fu P, Mahran A, Rastinehad AR, Tewari A, Tirumani SH, Purysko A, Ponsky L, Madabhushi A (2021) An integrated nomogram combining deep learning, Prostate Imaging–Reporting and Data System (PI-RADS) scoring, and clinical variables for identification of clinically significant prostate cancer on biparametric MRI: a retrospective multicentre study. Lancet Digital Health 3(7):e445–e454. https://doi.org/10.1016/S2589-7500(21)00082-0

139. Lai CC, Wang HK, Wang FN, Peng YC, Lin TP, Peng HH, Shen SH (2021) Autosegmentation of Prostate Zones and Cancer Regions from biparametric magnetic resonance images by using deep-learning-based neural networks. Sensors 21(8):2709. https://doi.org/10.3390/s21082709

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.