Computer Aided Diagnosis System for Prediction of Gleason Score in Prostate Cancer

Bharath Vishal G, Anusha S, Sam DJ, Jino Hans W, Kishore L
Department of Electronics and Communication Engineering, Sri Sivasubramaniya Nadar College of Engineering, Chennai, India

***Abstract*—Prostate Cancer (PCa) is the most common type of cancer in males and the third most diagnosed cancer overall in the world. Early Detection of PCa increases the chances of successfully treating it. Present detection methods like transrectal ultrasound (TRUS) biopsy are invasive while Prostate-Specific Antigen (PSA) tests have high false positive rates. Magnetic Resonance Imaging (MRI) has revolutionized the way prostate cancer is detected with the help of Computer-Aided Diagnosis (CAD). This research work developed a fully automated end-to-end detection and grading system with mpMRI from Prostate X-2 dataset which was used to detect the cancerous lesions and grade cancer aggressiveness according to Gleason Grade Groups (GGG) scores. This study developed an end-to-end system with a VGG19 feature extraction model to extract features, which were fed into a Random Forest classifier for the classification of Gleason Grade Groups. This system can assist healthcare professionals in the early detection of PCa lesions and classifying the aggressiveness using the GGG scores. Additional research in this field has the potential to advance both healthcare technology and improve patient care.**

Keywords—Computer-aided detection; mp-MRI; cancer lesions; Gleason Grade Group; Convolutional Neural Networks

1. INTRODUCTION

The prostate is a gland present in men that is part of the reproductive system. Prostate cancer develops when a normal prostate cell begins to grow uncontrollably. Prostate cancer has a higher chance of being cured if it is detected early. The prostate contains four different zones, which include the peripheral zone, central zone, transitional zone, and anterior fibromuscular stroma. The peripheral zone is the most common location for prostate cancer to develop, accounting for approximately 70% of all prostate cancers. Prostate cancer is less common in the central and transitional zones, and these tumors may differ from those in the peripheral zone in terms of their characteristics. The anterior zone only makes up a small portion of the prostate gland, and the majority of prostate cancer is located in the peripheral zone.

Prostate cancer (PCa) is the third most common type of cancer, and it affects men across the world. According to the World Health Organization (WHO), prostate cancer has accounted for 1,414,259 cases (7.3% of the total cancer cases), which preceded only lung and colorectal cancer [1]. There are different methods to detect prostate cancer, including the Prostate Specific Antigen (PSA) test. The PSA test is not effective, leading to false-positive results, which in turn lead to overdiagnosis and unnecessary biopsies [2]. Transrectal Ultrasound (TRUS) biopsy is another common method for assessing the aggressiveness of prostate cancer. During a TRUS biopsy, a small needle is inserted into the prostate and through the rectum to remove a sample of tissue from the prostate. The drawbacks of TRUS biopsy are that the results cannot be generalized for the whole organ as the biopsy only samples a small area of the gland, hence, the results cannot be generalized for the whole organ [3].

Multiparametric MRI is a more detailed version of the standard MRI, which, when compared to the traditional TRUS and other techniques, has been shown to increase sensitivity, potentially reduce unnecessary biopsies, and decrease overdiagnosis. However, prostate cancer analysis using mp-MRI is a time-consuming process that requires the expertise of an experienced radiologist. The time-consuming nature of this process has prevented the majority of the healthcare sector from implementing MRI-guided decision-making, which could be solved by automated image analysis [4]. With multiparametric MRI, the automated prostate cancer detection system can help the radiologist receive more information about the aggressiveness of the cancer lesions. This research focuses on solving the important challenge faced in the detection of prostate cancer, which is figuring out how to spot and distinguish clinically insignificant or harmless PCa lesions from clinically significant PCa lesions that could cause harm to the patient.

Accurately classifying Gleason Grade Groups (GGGs) is vital for determining cancer aggressiveness and guiding treatment. However, traditional biopsy analysis relies on subjective assessments, leading to potential variability and misdiagnosis among pathologists. This highlights the need for reliable tools to aid in GGG classification. This research explores using deep learning models for automatic feature extraction from prostate biopsy images. The goal is to capture subtle morphological patterns and intratumoral heterogeneity. Incorporating these features into robust machine learning classifiers aims to achieve objective and accurate GGG classification. The study addresses the critical need for objective GGG classification in prostate cancer. Leveraging deep learning and machine learning aims to improve diagnostic accuracy, enable personalized treatment approaches, and ultimately enhance patient outcomes.

2. RELATED WORK

A two-dimensional U-Net that takes input from MRI slices and output from a lesion segmentation map that encodes the Gleason Grade Groups (GGG) scale of cancer aggressivity was proposed by Coen de Vente et al. This study compares various ensemble approaches and previous information-based

prostate region segmentation algorithms. The ordered classes are taken advantage of in a method for encoding GGG into model measures. Additionally, an ensemble technique was applied and prostate region segmentation was taken into account as prior information. Additionally, the suggested model measure performs better than multi-label ordinal regression and conventional multi-class classification [5]. Mirasol et al proposed a repeatable framework for prostate cancer lesion binary semantic segmentation using convolutional neural networks. The publicly available QIN-PROSTATE-Repeatability dataset containing annotated MRI images from 15 patients was utilized and the core architecture employed was U-Net with batch normalization, tested with different models like VGG16, ResNet34, and DenseNet121. The best-performing combination achieved a Dice score of 0.47, comparable to previous studies [6]. In a study by Bejoy Abraham et al, a unique approach that differs from most of the research work that is performed by researchers in Prostate Cancer Detection uses texture features and stacked sparse autoencoder to classify prostate cancer grade groups from MRI images. Initially, First-order features like mean, skewness, and variance are calculated. This is followed by a matrix-based feature extraction method which is further used to extract useful information in the form of handcrafted features. These features include Energy, Contrast, Homogeneity, and Correlation which are used to find the similarity between the feature vectors obtained. A stacked sparse autoencoder is used to extract high-level features where the output of every layer is given as input to the next stacked layer and thus gives information-rich feature vectors when compared to a normal autoencoder. The softmax classifier is used after this to map the output vector to the label of the input data [7].

Ruiming Cao et al. proposed the use of FocalNet, which has a foveal attention architecture when compared to the general self-attention networks. There are T2 weighted and ADC images in the dataset and this FocalNet helps in taking both into the imaging channel to predict the label of the classes of Gleason Scores. It helps in identifying aggressiveness by identifying the ground-truth mask. Ordinal encoding of labels is performed before this step to enhance the classification performance because, in one hot encoding, the CNN assumes that the different labels are unrelated but in reality, there exists some correlation between the Gleason score labels for the input data. FocalNet is trained based on two parameters- mutual finding loss for features in either of the imaging components because different components will contain different information and only a part of it is similar; focal loss with relation to both T2w and ADC which combines both lesion and non-lesion pixels. Pre-processing techniques involve the Registration of ADC images to the T2-weighted ones as well as the normalization of the intensity of the pixels in the images[8]. Carina Jensen et al explored a KNN classifier to identify the lesions. It involves extracting the ROI which is highly essential to perform accurate classification as it involves zone-specific features. This is extracted using the T2 weighted and DWI images alone. Texture features are also extracted which along with KNN helps in classification. Preprocessing involves re-sampling of T2 weighted and DWI images and then feature selection is done through a semi-exhaustive feature search. Validation is done through K-fold cross-validation which was done separately to the different zones of the prostate MRI images [9]. Audrey Duran et al. suggested a novel architecture that has an attention mechanism to perform multi-class segmentation of lesions. The proposed network has two branches where the first branch does segments of the input images and the second branch uses the information obtained about the zones as prior information for the attention gate and grades the input into the corresponding Gleason Grade Groups. Preprocessing involves spatial resampling and no registration of T2W images to ADC images is required, unlike the other methods. This architecture was focused only on the peripheral zone as it generally has more lesion and zonal information when compared to other zones [10].

In another paper, Bejoy Abraham et al. created a method for the detection and grading of PCa using VGG-16. Pre-processing involves ROI extraction for the T2W, ADC, and BIVAL images. The Feature extraction is done using a pre-trained CNN model, VGG-16 which allows training without annotated images which is highly advantageous. It recognizes all the patterns from coarse to fine which gives a lot of generalization and better training accuracy for the model. Principal Component Analysis is done after that which is followed by classification done through an ordinal class classifier because a normal classifier will not identify the correlation between the Gleason Grades [11]. Ruiming Cao et al. also implemented a novel method of using a selective dense conditional random field after the CNN to improve the results from the CNN into lesion segmentation. Preprocessing is done through the normalization of T2W images after which registration of ADC images to T2W images is done. The post-processing step of the conditional random field is done to map the corresponding output from the CNN and checks which of T2W and ADC gives better lesion detection results and then compares the L2 distance of that with the area having the cancerous tissues in the prostate gland. Validation and evaluation are done through ROC characteristics that help in identifying the false positives after prediction [12].

Furthermore, a conditional generative adversarial network approach was used by S. Tahri et al. The Pix2Pix model was used to create synthetic CT scans in addition to normal T2-weighted pictures. Even before pre-processing, the generator and loss function were evaluated, allowing the model to produce high-quality synthetic CT images from MRI data. The Pix2Pix model takes MRI pictures as input and outputs a probability of 0 or 1 that is transferred to a true or false CT image. This approach is compared to others such as patch-based and deep learning methods [13]. Tsehay et al proposed a deep-learning architecture based on convolutional neural networks where the initial Pre-processing involves converting the input MRI images to three-channel RGB images. This is followed by histogram equalization and the three channels of images were combined to give a single RGB image. The CNN takes the input images and gives the output as a probability map image. Cross-validation is done by evaluating the model over multiple intermediate steps, and ROC analysis is done on the obtained probability maps [14]. Nader Aldoj et al. 's research work focused on a deep learning

method that performs semi-automatic prostate cancer classification. This is done through a 3D-CNN and unlike other methods, it does not require any zonal information as prior. It just requires information about the lesion center. Pre-processing involves resampling and registration as well as image augmentation. The CNN takes in the input images of every modality into a separate channel and classifies them into GGG [15].

The proposed model in this paper addresses the additional challenge of detecting and grading Prostate cancer into the corresponding Gleason Grade Groups, aiding in the determination of appropriate treatment strategies for patients. Drawing insights from a comprehensive literature review, this research work evaluates multiple models to establish a fully automated end-to-end system for detection and grading using mp-MRI from the prostate X-2 dataset.

3. MATERIALS AND METHODS

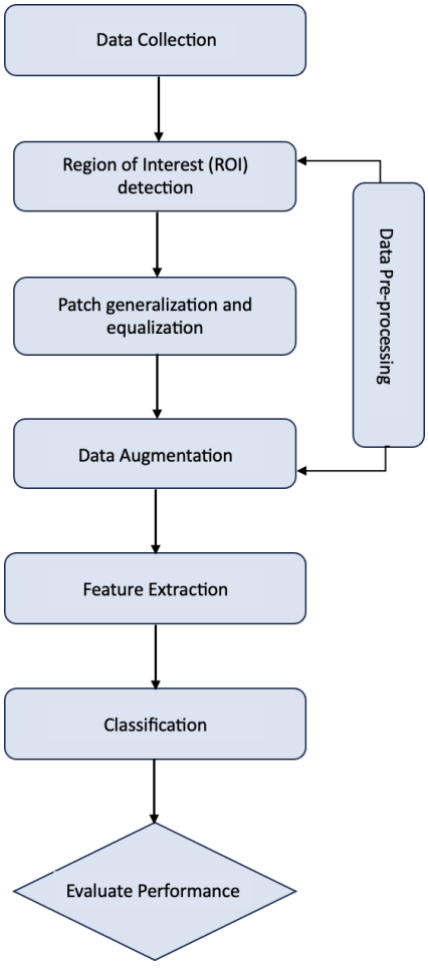

Fig 1. Flowchart of the proposed methodology

Figure 1 shows the proposed methodology flow which involves Data pre-processing techniques like Region of Interest detection, patch generation, equalization, and data augmentation techniques. Dicom files available in the dataset for individual subjects were imported using a slicer python interactor and were converted to NRRD files.Data augmentation was performed to make the model more generalized on the data by generating five images for every image through rotation and shearing. This paper experimented with three feature extraction models which include the VGG16, VGG19, and DenseNet 121 models, and also combined them with each of the following three machine learning classifier models which include Random Forest Classifier, Decision Tree Classifier, and XGBoost Classifier. Performance evaluation was done using Gleason Grade Groups and performance metrics like the AUC-ROC curve.

3.1 DATA COLLECTION

This research work utilized the dataset from the ProstateX2 challenge conducted by the American Association of Physicists in Medicine (AAPM), along with the SPIE (the International Society for Optics and Photonics) and the National Cancer Institute (NCI).

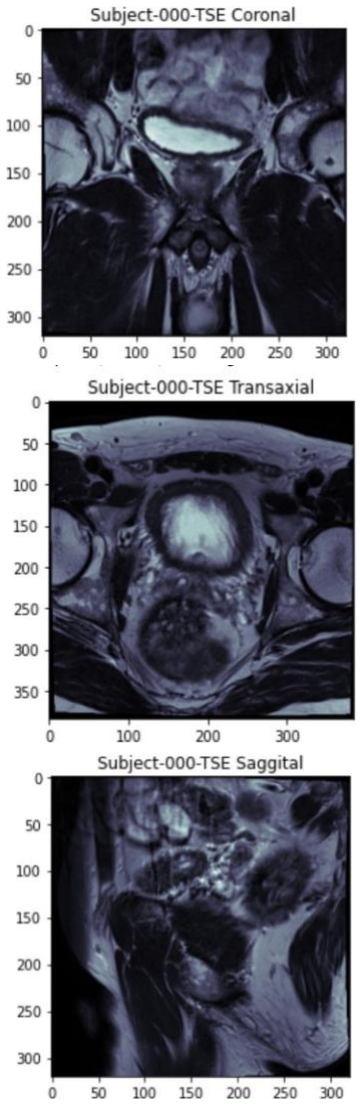

Fig.2.Sample of Prostate MRI image in 3 different orientations from the prostateX2 dataset (Top to bottom: Coronal, Transaxial, and Sagittal plane)

The dataset obtained contained a total of 182 lesions in these cases out of which 112 of them are present in the training data

and 70 of them are present in the testing data. Each lesion will include its precise location, reference thumbnail image, and Gleason Grade Group (GGG). Figure 2 shows a sample of the prostate MRI with three different orientations namely the Coronal plane, Transaxial plane, and Sagittal plane.

This dataset from the ProstateX2 challenge is open source and publicly available, it can be downloaded from here https://www.aapm.org/GrandChallenge/PROSTATEx-2/ [16].

3.2 Data pre-processing

3.2.1 *Region Of Interest (ROI) Detection*

One of the important steps in the pipeline of preprocessing prostate MRIs is the generation of the region of interest (ROIs). Generally, early detection of diseases is necessary to prevent any fatalities, similarly, Prostate cancer must be detected early to improve the rate of survival of the patients. However, due to the low contrast of the MRI images, it is difficult to differentiate between the cancerous cancer cells and the clinically insignificant regions which shows that generating ROI is a difficult process. ROI-based feature extraction methods for prostate cancer detection help in reducing false negatives. It eliminates redundant information, which helps in defining the boundaries of a tumor on an image or in a volume, to measure its size.

3.2.2 *Patch Generation And Equalization*

Patch generation was performed where a patch of size 64*64*3 was generated for T2 weighted images and the rest 16*16*3 size patch was generated. Then the location of the findings was calculated and equalization was performed. Rescaling was done followed by which the patches were extracted and stored in a NumPy array.

3.2.3 *Data Augmentation*

Augmentation was performed to generalize the model for better performance on unseen data. For one image, five new images were generated using augmentation by applying the transformation. These augmented images are added to the original set of images to increase the available data size. The following transformations are applied: Rotation through an angle is distinctly chosen between 80,170,-80,-170, and shearing of the image by a factor of 0.2 or -0.2 to get five augmented images

3.3 Feature Extraction

In the subsequent stage of the pipeline, Feature Extraction was conducted to identify and select instrumental features for classifying lesions into distinct Gleason grade groups. This was performed mainly to capture the aggressiveness of prostate cancer using information extracted from T2, ADC, and Bval images. Drawing from previous literature and research endeavors, automated feature extraction techniques were employed to represent the content within the images. The research utilized Convolutional Neural Network (CNN) architectures, specifically VGG16[18], VGG19[18], and DenseNet121[19], for feature extraction. Subsequently, machine learning classifiers were employed to evaluate and make predictions based on the learned features. This approach aims to enhance the accuracy and efficiency of Gleason grade group classification in prostate cancer diagnosis.

VGG16 and VGG19 were mainly utilized for identifying patterns and structures within the different sequences of mpMRI. Hierarchical feature extraction starting from low-level features such as edges and textures to more complex features representing structures and patterns was performed due to the stackable nature of the models which helps in extracting subtle features which is helpful for further classification into Gleason Grade Groups. DenseNet121 has a dense network that captures the features across different planes and modalities of the prostate MRI images. The inter-layer connectivity helps preserve the information from multiple modalities, improving the model's sensitivity for feature extraction. The T2-weighted images provide insights into tissue characteristics like the total amount of lesion area in the prostate gland and ADC images indicate the tissue microstructure and cellularity within the prostate gland. Areas with reduced ADC values indicate the presence of cancerous tissue in the gland. The selected architectures were efficient in handling the hierarchical features and provided adaptability to diverse imaging modalities and planes.

3.4 Classification

After the extraction of features, Decision Tree[20], Random Forest[21], and XGBoost[22] models were employed as classifiers in the proposed prostate cancer detection model. Decision trees help understand the complex relationships within the data but may be vulnerable to overfitting and this limitation was addressed by hyperparameter tuning during grid search. To overcome this limitation, Random Forest was chosen to improve generalization and accuracy by aggregating the predictions of multiple trees. In this study, Random Forest excels at capturing complex relationships within the multiparametric features while providing increased stability and reduced sensitivity to individual data fluctuations concerning the modalities and planes. Grid search is applied to the Decision Tree, and randomized search to Random Forest and XGBoost's hyperparameters are optimized to ensure optimal performance.

4. Results And Discussion

The results of this research work were assessed for the five Gleason Grade Groups which are thresholds (GGG1, GGG2, GGG3, GGG4, GGG5) at the lesion level to evaluate the lesion significance. The lesions for which the predicted GGG score was equal to or above the chosen threshold were chosen for further evaluation. Metrics such as the area under the receiver operating characteristic (AUC-ROC) were employed. The AUC provides a comprehensive overview of sensitivity and specificity across various threshold values. It is determined by calculating the area under the curve, which is constructed by graphing sensitivity against specificity for all conceivable threshold values.

In this comprehensive experimentation involving diverse feature extraction models, such as VGG16, VGG19, and DenseNet121, in conjunction with various machine learning classifiers, including Decision Tree, Random Forest, and XGBoost, the investigation has yielded various insights as shown in Table I. The results suggest that ensemble methods, such as Random Forest, may be particularly effective for feature extraction in the context of the evaluated dataset. The thorough analysis demonstrates that the combination of the

VGG19 feature extraction model and the Random Forest classifier consistently outperforms other configurations, providing the most promising results in accuracy and robustness. This optimized model pairing emerges as the preferred choice for accurate Gleason Grade Groups classification in prostate cancer. This combination of VGG19 and Random Forest feature extraction performs well due to VGG19's deep architectural design, which captures complex and intricate patterns in MRI images that are pivotal for discerning between different Gleason Grade Groups.

TABLE I. COMPARISON BETWEEN DIFFERENT DEEP LEARNING FEATURE EXTRACTION AND MACHINE LEARNING CLASSIFIER MODELS

Feature Extraction Model	**Overall Accuracy**		
	Decision Tree	***Random Forest***	***XGBoost***
VGG16	40.60	44.83	42.80
VGG19	36.25	46.25	41.22
DenseNet121	40.83	44.17	41.22

The proposed model was applied in three different imaging modalities which are T2 Transverse, T2 Sagittal, and (ADC) Apparent Diffusion Coefficient images. To conclude, the proposed model performed better in ADC images when compared to T2 Sagittal and T2 Transverse images. Figure 3 shows the AUC-ROC values of each class mapping to GGG scores from ADC Images and similarly, Figure 4 shows the AUC-ROC values of each class mapping to GGG scores from T2-Sagittal Images and Figure 5 AUC-ROC of each class mapping to GGG scores from T2-Transversal Images. The AUC scores from Table II highlight distinct performance patterns among imaging modalities across Gleason Grade Groups (GGG). For GGG $\geq$ 5, ADC Images exhibit the highest AUC score at 0.7588, surpassing T2 Sagittal (0.6681) and T2 Transverse (0.5725). Conversely, in GGG $\geq$ 2, T2 Sagittal outperforms other modalities with an AUC score of 0.5224, while GGG $\geq$ 1 sees the highest AUC score for T2 Sagittal at 0.6321. These findings underscore the varying efficacy of imaging modalities in differentiating Gleason Grade Groups.

TABLE II. COMPARISON BETWEEN AUC SCORES FOR DIFFERENT IMAGING MODALITY AND EACH GGG SCORE

Gleason Grade Group	**AUC scores**		
	ADC Images	***T2 Sagittal***	***T2 Transverse***
GGG $\geq$ 1	0.6044	0.6321	0.5447
GGG $\geq$ 2	0.6108	0.5224	0.5599
GGG $\geq$ 3	0.5715	0.5601	0.6643
GGG $\geq$ 4	0.6154	0.5501	0.6263
GGG $\geq$ 5	0.7588	0.6681	0.5725

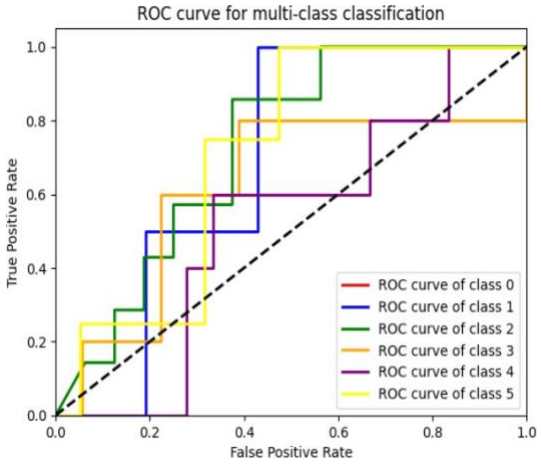

Fig 3 AUC-ROC of each class mapping to GGG scores from ADC Images

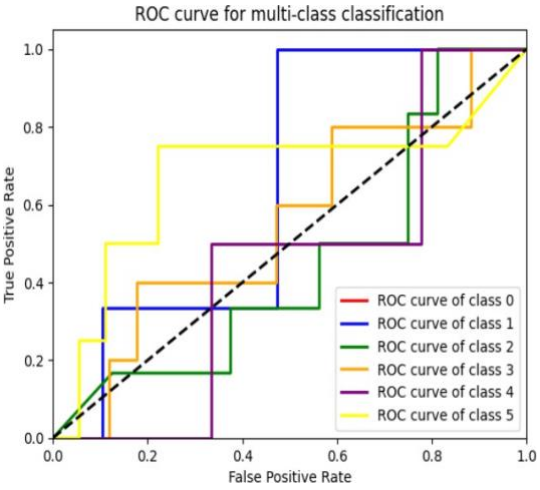

Fig 4 AUC-ROC of each class mapping to GGG scores from T2-Sagittal

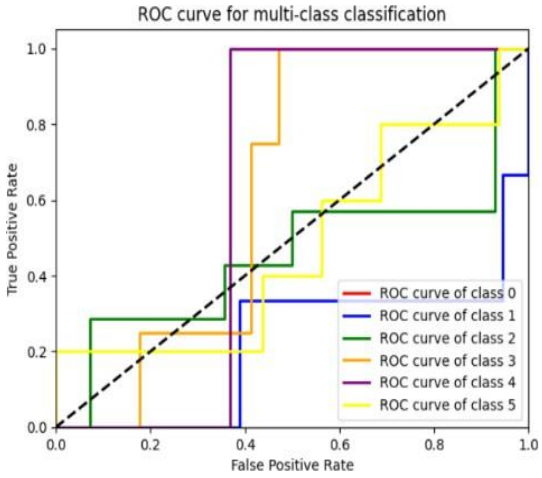

Fig 5 AUC-ROC of each class mapping to GGG scores from T2-Transversal

The difference in performance of the model across the different imaging modalities, specifically the results obtained from Apparent Diffusion Coefficient (ADC) images, highlights the importance of selecting optimal imaging modality for grading of prostate cancer. ADC images give better results because of their sensitivity to the diffusion of water molecules, giving important insights into tumor aggressiveness and cellular density, which are crucial for the accurate classification into different Gleason grades. This observation suggests that adding ADC images into the diagnostic workflow could significantly improve the predictive accuracy of CAD systems, providing doctors with a more detailed understanding of tumor features.

5. CONCLUSION AND FUTURE WORK

In conclusion, leveraging transfer learning with models such as VGG16, VGG19, and DenseNet121, coupled with effective classifiers like Decision Tree, Random Forest, and XGboost, the proposed architecture demonstrates exceptional performance. This contributes to accurate cancer detection but also holds promise in guiding treatment decisions for patients based on Gleason Grade Groups. Future work in this area of research could employ ensemble learning techniques to enhance the accuracy of the detection and grading. State-of-the-art image registration methods could be explored to combine the different imaging modalities and generate better results, hence extending the research work to benefit healthcare professionals.

REFERENCES

[1] World Health Organization. (2020). "Globocan report: Estimated cancer incidence and mortality numbers worldwide." International Agency for Research on Cancer. [Online]. Available: https://gco.iarc.fr/globocan/2020/estimates/summary.html

[2] Duffy, M. J. (2014). "PSA in screening for prostate cancer: more good than harm or more harm than good?" Adv Clin Chem, 66, 1-23. PMID: 25344984.

[3] Moe, A., & Hayne, D. (2020). "Transrectal ultrasound biopsy of the prostate: does it still have a role in prostate cancer diagnosis?" Translational Andrology and Urology, 9(6), 1123-1132. doi:10.21037/tau.2019.09.37.

[4] American Urological Association. (2015). "Standard Operating Procedure for Multiparametric Magnetic Resonance Imaging in the Diagnosis, Staging and Management of Prostate Cancer." Translational Andrology and Urology, 9(6), 1123-1132. doi:10.21037/tau.2019.09.37.

[5] Vente, C., et al. (2021). "Deep Learning Regression for Prostate Cancer Detection and Grading in Bi-Parametric MRI." IEEE Trans Biomed Eng, 68(2), 374-383. doi: 10.1109/TBME.2020.2993528.

[6] Mirasol, Ian & Abu, Patricia Angela & Reyes, Rosula (2022). Construction of a Repeatable Framework for Prostate Cancer Lesion Binary Semantic Segmentation using Convolutional Neural Networks. International Journal of Advanced Computer Science and Applications. 13. 10.14569/IJACSA.2022.0130697.

[7] Abraham, B., & Nair, M. S. (2018). "Computer-aided classification of prostate cancer grade groups from MRI images using texture features and stacked sparse autoencoder." Computerized Medical Imaging and Graphics. doi: 10.1016/j.compmedimag.2018.08.006.

[8] Cao, R., et al. (2019). "Joint Prostate Cancer Detection and Gleason Score Prediction in mp-MRI via FocalNet." IEEE Transactions on Medical Imaging, 38(11), 2496-2506. doi: 10.1109/TMI.2019.2901928.

[9] Jensen, C., et al. (2019). "Assessment of prostate cancer prognostic Gleason grade group using zonal-specific features extracted from biparametric MRI using a KNN classifier." J Appl Clin Med Physics. doi: 10.1002/acm2.12542.

[10] Duran, A., et al. (2022). "ProstAttention-Net: A deep attention model for prostate cancer segmentation by aggressiveness in MRI scans." Medical Image Analysis, 77, 102347. ISSN 1361-8415.

[11] Abraham, B., & Nair, M. S. (2019). "Automated grading of prostate cancer using convolutional neural network and ordinal class classifier." Informatics in Medicine Unlocked, 17, 100256. doi:10.1016/j.imu.2019.100256.

[12] Cao, R., et al. (2019). "Prostate Cancer Detection and Segmentation in Multi-parametric MRI via CNN and Conditional Random Field." 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), pp. 1900-1904. doi: 10.1109/ISBI.2019.8759584.

[13] Tahri, S., et al. (2022). "A high-performance method of deep learning for prostate MR-only radiotherapy planning using an optimized Pix2Pix architecture." Physica Medica, 103. doi:10.1016/j.ejmp.2022.10.003.

[14] Tsehay, Y., et al. (2017). "Convolutional neural network based deep-learning architecture for prostate cancer detection on multiparametric magnetic resonance images." doi: 10.1117/12.2254423.

[15] Aldoj, N., 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(2), 1243-1253. doi: 10.1007/s00330-019-06417-z.

[16] Litjens, G., et al. (2017). "SPIE-AAPM PROSTATEx Challenge Data (Version 2) [dataset]." The Cancer Imaging Archive.

[17] LeCun, Y., et al. (1998). "Gradient-based learning applied to document recognition." Proceedings of the IEEE, 86(11), 2278-2324. doi: 10.1109/5.726791.

[18] Simonyan, K., & Zisserman, A. (2015). "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1518.

[19] Huang, G., et al. (2017). "Densely connected convolutional networks." IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[20] Quinlan, J. R. (1994). "Induction of Decision Trees." Machine Learning, 1(1), 81-106.

[21] Breiman, L. (2001). "Random forests." Machine Learning, 45(1), 5-32.

[22] Chen, T., & Guestrin, C. (2016). XGboost: A scalable tree boosting framework. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 785-794.

[23] Rundo, L., Tangherloni, A., Galimberti, S., Militello, C., Sala, E., & Mauri, G. (2019). "Automated Prostate Cancer Diagnosis and Gleason Grading of Biopsy Images: A Deep Learning-Based Approach." Entropy, 21(4), 375.

[24] Song, Z., Zou, S., Zhou, W., Huang, Y., Shao, L., Yuan, J., ... & Gao, J. (2020). "Deep Learning Enables Accurate Diagnosis and Gleason Grading of Prostate Cancer in Biopsies and MRI." Nature Communications, 11(1), 2336.

[25] Lucas, M., Jansen, I., Savci-Heijink, C. D., Meijer, S. L., & de Boer, O. J. (2021). "Deep Learning for Automatic Gleason Pattern Classification for Grade Group Determination of Prostate Biopsies." Virchows Archiv, 477(3), 395-402.

[26] Nagpal, K., Foote, D., Tan, F., Liu, Y., Chen, P. C., Steiner, D. F., ... & Stumpe, M. C. (2019). "Development and Validation of a Deep Learning Algorithm for Improving Gleason Scoring of Prostate Cancer." npj Digital Medicine, 2, 48.