Paper ID:88001

Revolutionizing Prostate Cancer Diagnosis: An Integrated Approach for Gleason Grade Classification and Explainability

Authors:

Les Sztandera*,

Dept., Computer Information Systems, Thomas Jefferson University, Philadelphia, PA 19107, USA
Anil B. Gavade,
Department of E&C, KLS Gogte Institute of Technology, Belagavi 590008, Karnataka, India

Rajendra B. Nerli,
Department of Urology, JN Medical College, KLE Academy of Higher Education and Research, Belagavi 590010, India.

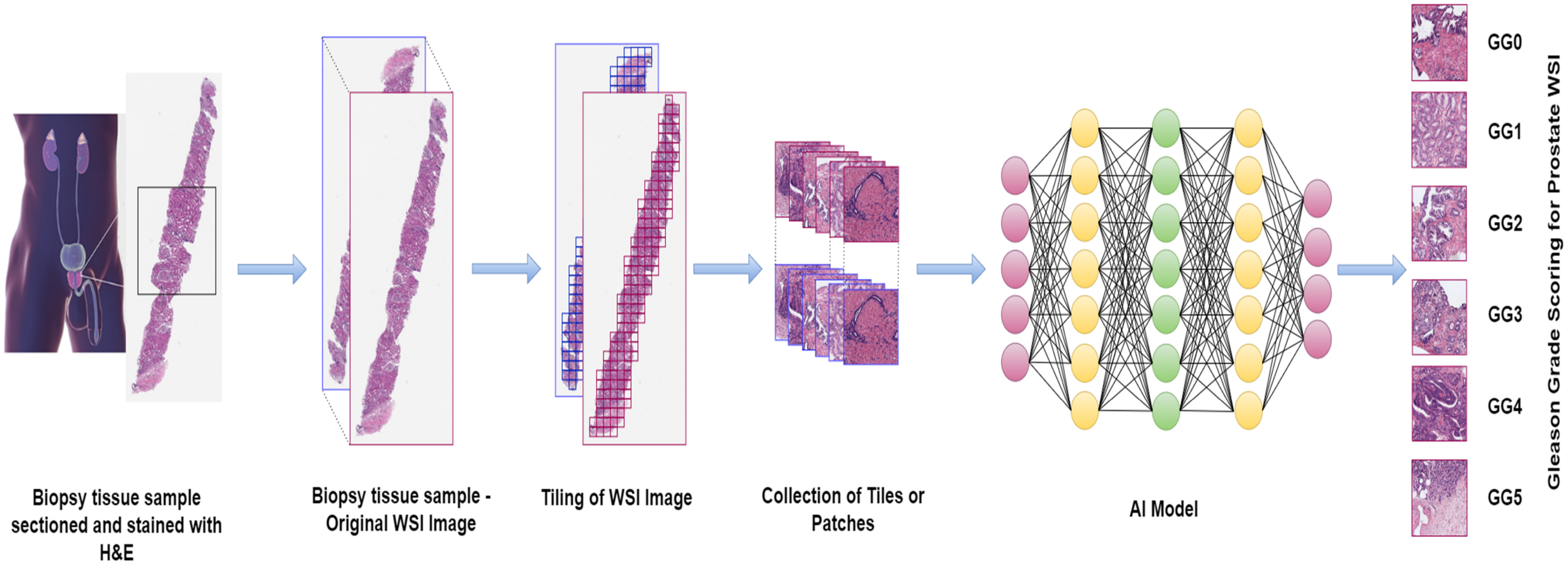
Shridhar C. Ghagane

KAHER'S Dr. Prabhakar Kore Basic Science Research Center, JNMC Campus, Belagavi-590010, India

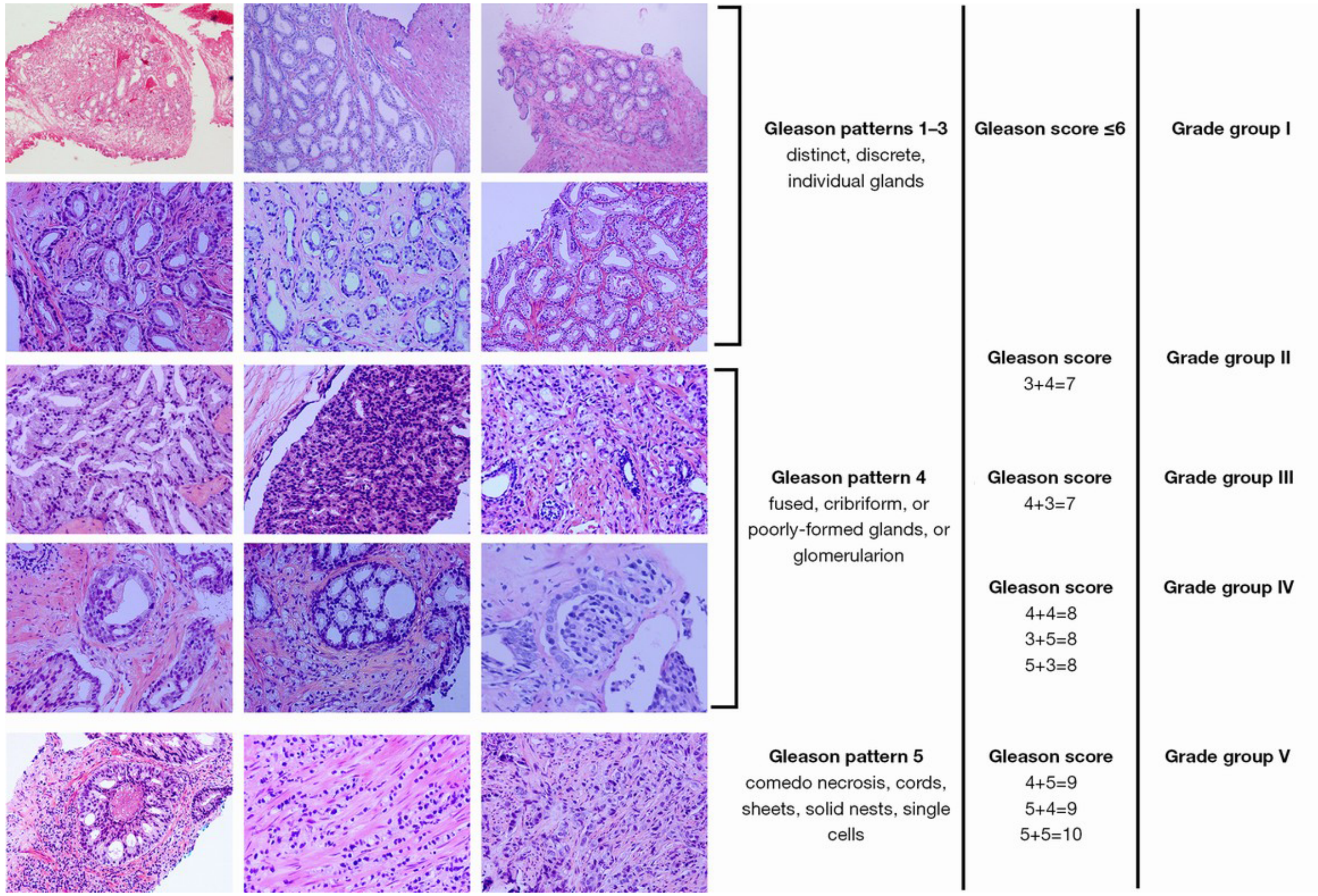
Abstract:

Accurate grading of Prostate Cancer (PCa) is vital for effective treatment planning and prognosis. This study introduces an advanced framework for Gleason Grade (GG) classification, addressing challenges in accuracy, computational efficiency, and interpretability. Utilizing the SICAPv2 dataset, which contains annotated prostate biopsy Whole Slide Images (WSIs) graded from GG0 to GG5, the framework integrates cutting-edge machine learning and deep learning techniques. Feature extraction is performed using a custom-designed Variational Autoencoder (VAE) with a VGG16 backbone, chosen for its computational efficiency, while dimensionality reduction with Principal Component Analysis (PCA) optimally selects 50 features for classification. The classification pipeline combines machine learning models, including Support Vector Machines (SVM), logistic regression, and random forests, with custom Deep Neural Networks (DNNs). SVM with an Radial Basis Function (RBF) kernel achieved an accuracy of 84% following hyperparameter tuning, while a custom five-layer dense neural network incorporating dropout and batch normalization demonstrated superior performance with an accuracy of 94.6%. Explainable AI (XAI) techniques, such as SHapley Additive exPlanations (SHAP), gradient-weighted class activation mapping (Grad-CAM), and Local Interpretable Model-Agnostic Explanations (LIME), enhance model interpretability by providing insights into feature importance and aligning predictions with clinical expertise. This framework delivers a robust, scalable, and interpretable solution for automated GG classification, bridging the gap between advanced AI techniques and clinical application.

Gleason Grade Scoring: Implement a deep learning system to automatically assign Gleason grades from WSI, enhancing precision and reproducibility.

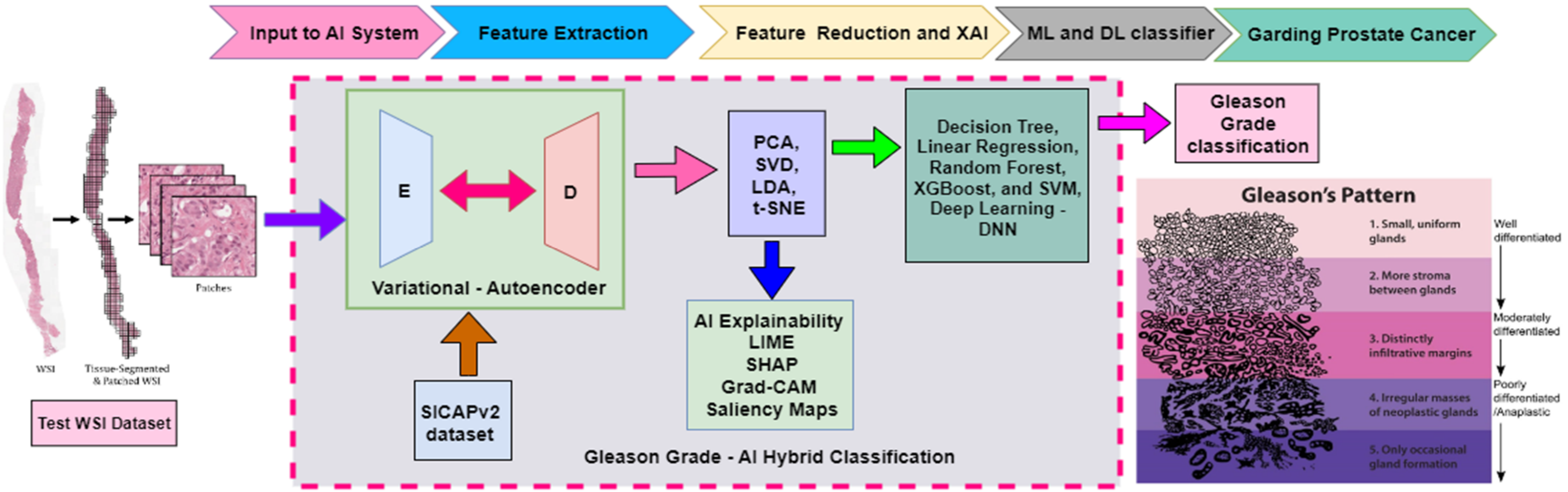


Patches of H&E Histology Samples Presenting Different Gleason Patterns from GG0 - GG5



Gleason patterns 1–3 distinct, discrete, individual glands	**Gleason score ≤6**	**Grade group I**
	Gleason score 3+4=7	**Grade group II**
Gleason pattern 4 fused, cribriform, or poorly-formed glands, or glomerularion	**Gleason score** 4+3=7	**Grade group III**
	Gleason score 4+4=8 3+5=8 5+3=8	**Grade group IV**
Gleason pattern 5 comedo necrosis, cords, sheets, solid nests, single cells	**Gleason score** 4+5=9 5+4=9 5+5=10	**Grade group V**

Implementation of an AI Pipeline



The block diagram shows a hybrid system for GG Group using WSIs, ensuring transparent and accurate PCa diagnosis and treatment with XAI techniques

SICAPv2 dataset: [Experimented]

Silva-Rodríguez, J., Colomer, A., Sales, M. A., Molina, R., & Naranjo, V. (2020). Going deeper through the Gleason scoring scale: An automatic end-to-end system for histology prostate grading and cribriform pattern detection. Computer Methods and Programs in Biomedicine, 195, 105637

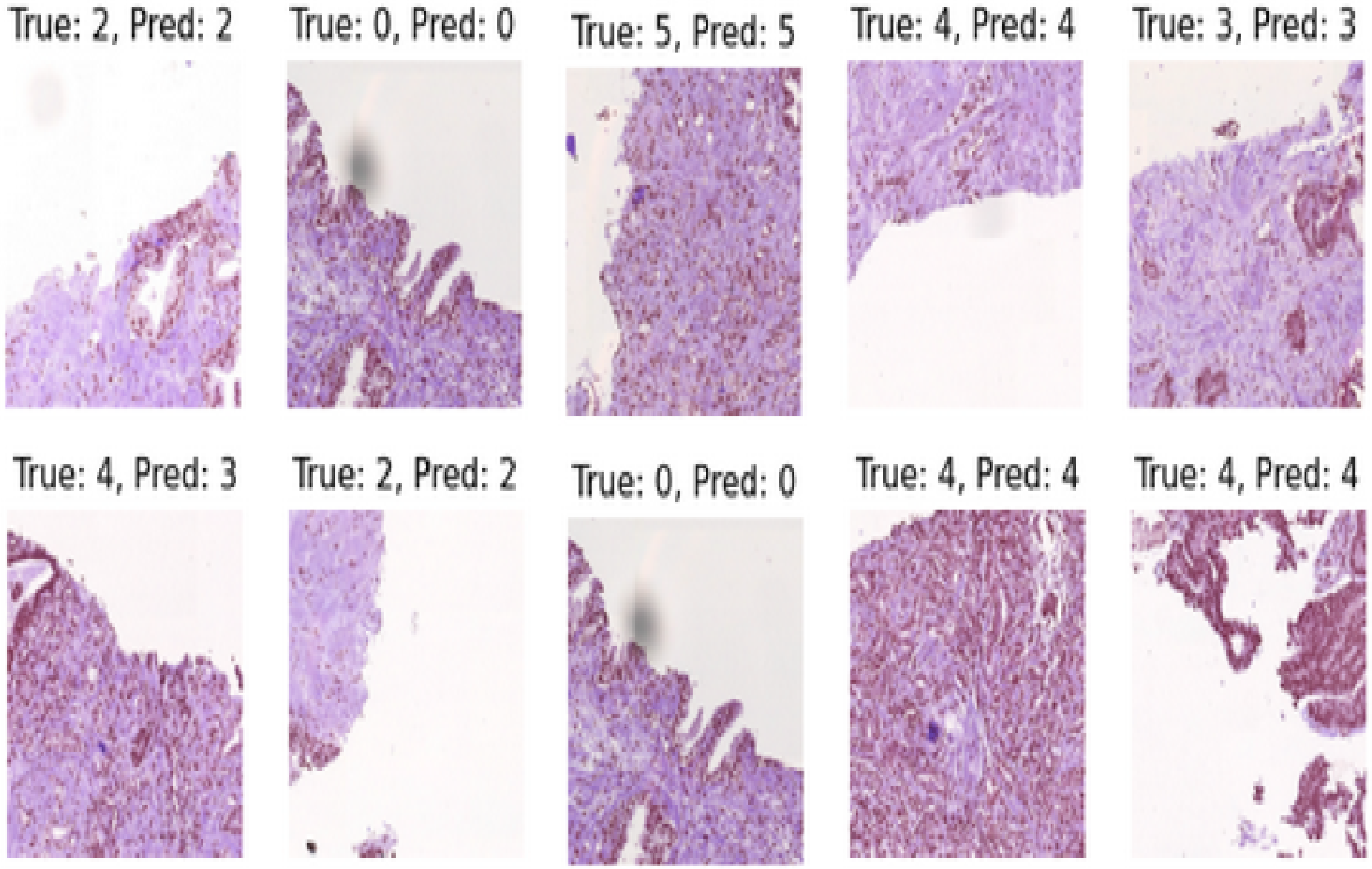
Gleason Grade Classification **Performance Metrics**

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

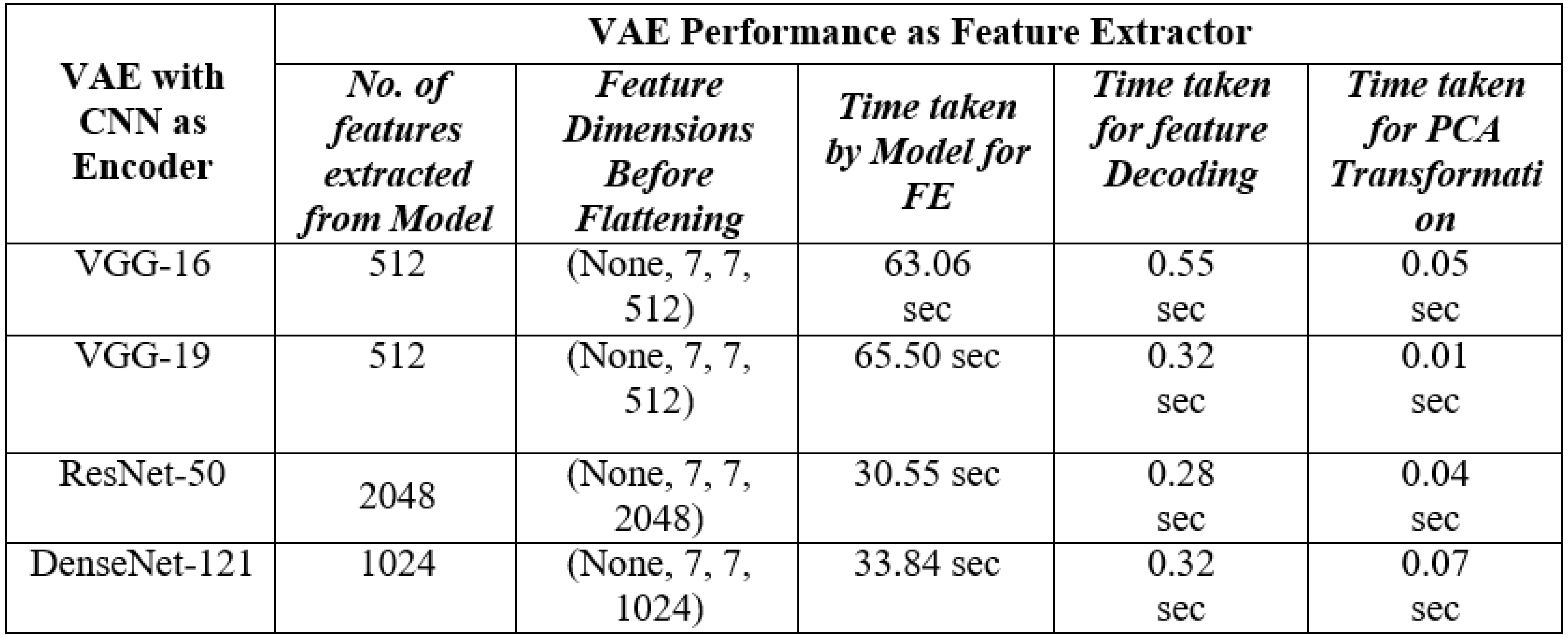
$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$



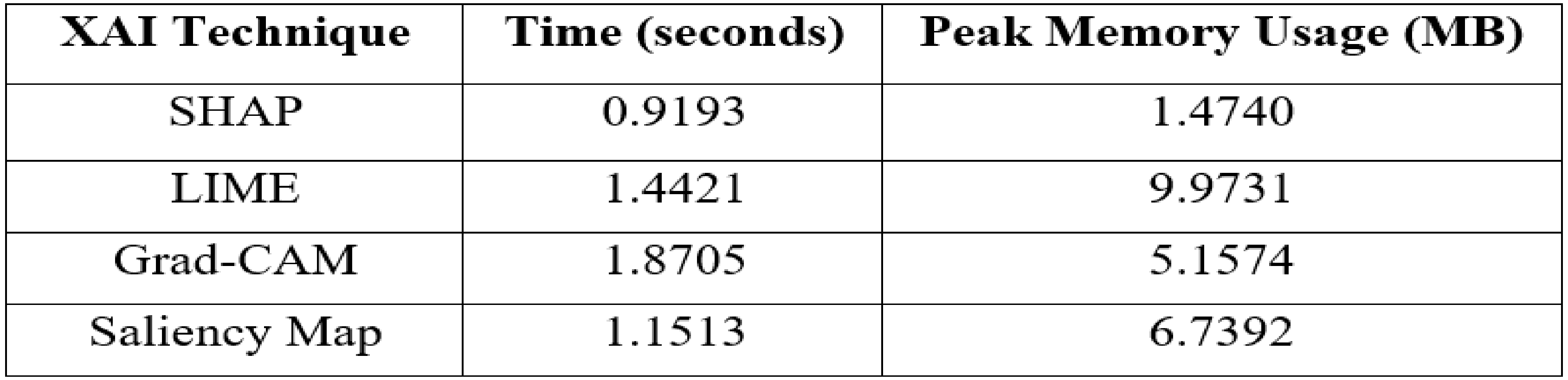
Results:

Performance feature extraction



VAE with CNN as Encoder	VAE Performance as Feature Extractor				
	No. of features extracted from Model	***Feature Dimensions Before Flattening***	***Time taken by Model for FE***	***Time taken for feature Decoding***	***Time taken for PCA Transformation***
VGG-16	512	(None, 7, 7, 512)	63.06 sec	0.55 sec	0.05 sec
VGG-19	512	(None, 7, 7, 512)	65.50 sec	0.32 sec	0.01 sec
ResNet-50	2048	(None, 7, 7, 2048)	30.55 sec	0.28 sec	0.04 sec
DenseNet-121	1024	(None, 7, 7, 1024)	33.84 sec	0.32 sec	0.07 sec

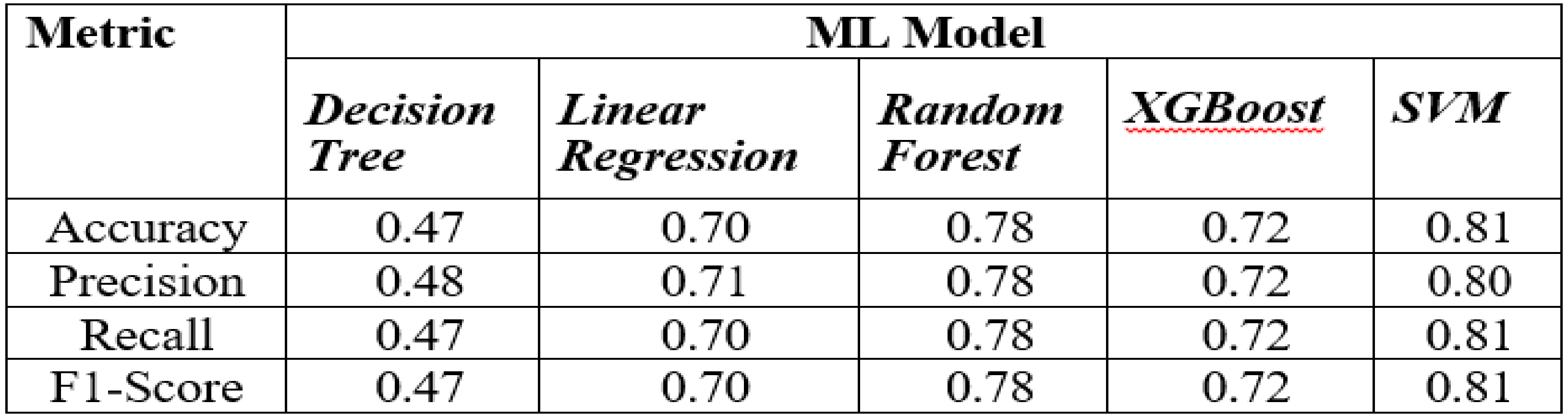
Performance explainability of feature



XAI Technique	Time (seconds)	Peak Memory Usage (MB)
SHAP	0.9193	1.4740
LIME	1.4421	9.9731
Grad-CAM	1.8705	5.1574
Saliency Map	1.1513	6.7392

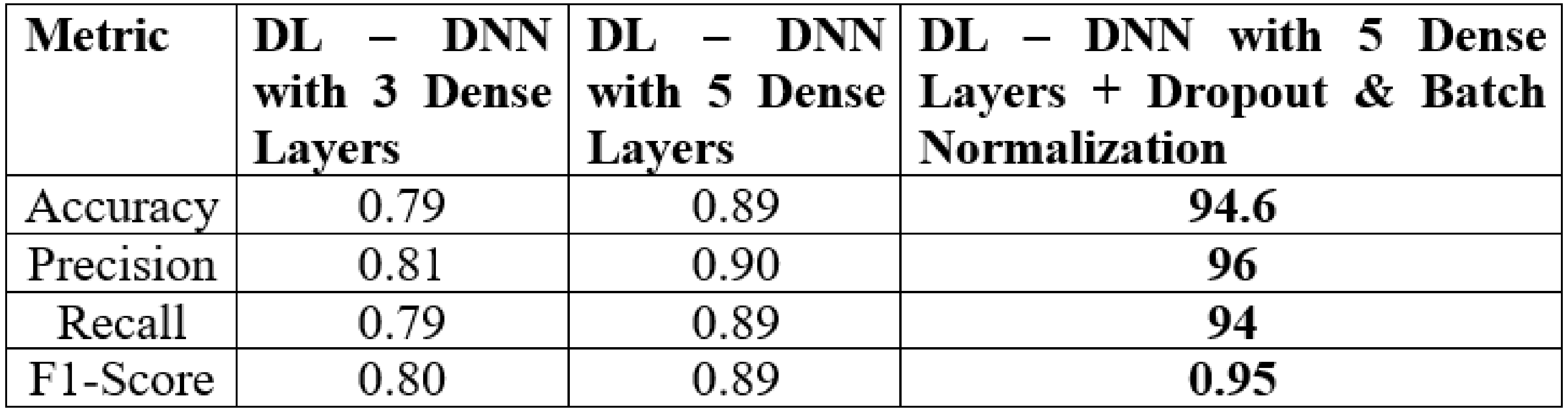
Results:

Performance Metrics for Various Classification Models



Metric	ML Model				
	Decision Tree	*Linear Regression*	*Random Forest*	*XGBoost*	*SVM*
Accuracy	0.47	0.70	0.78	0.72	0.81
Precision	0.48	0.71	0.78	0.72	0.80
Recall	0.47	0.70	0.78	0.72	0.81
F1-Score	0.47	0.70	0.78	0.72	0.81

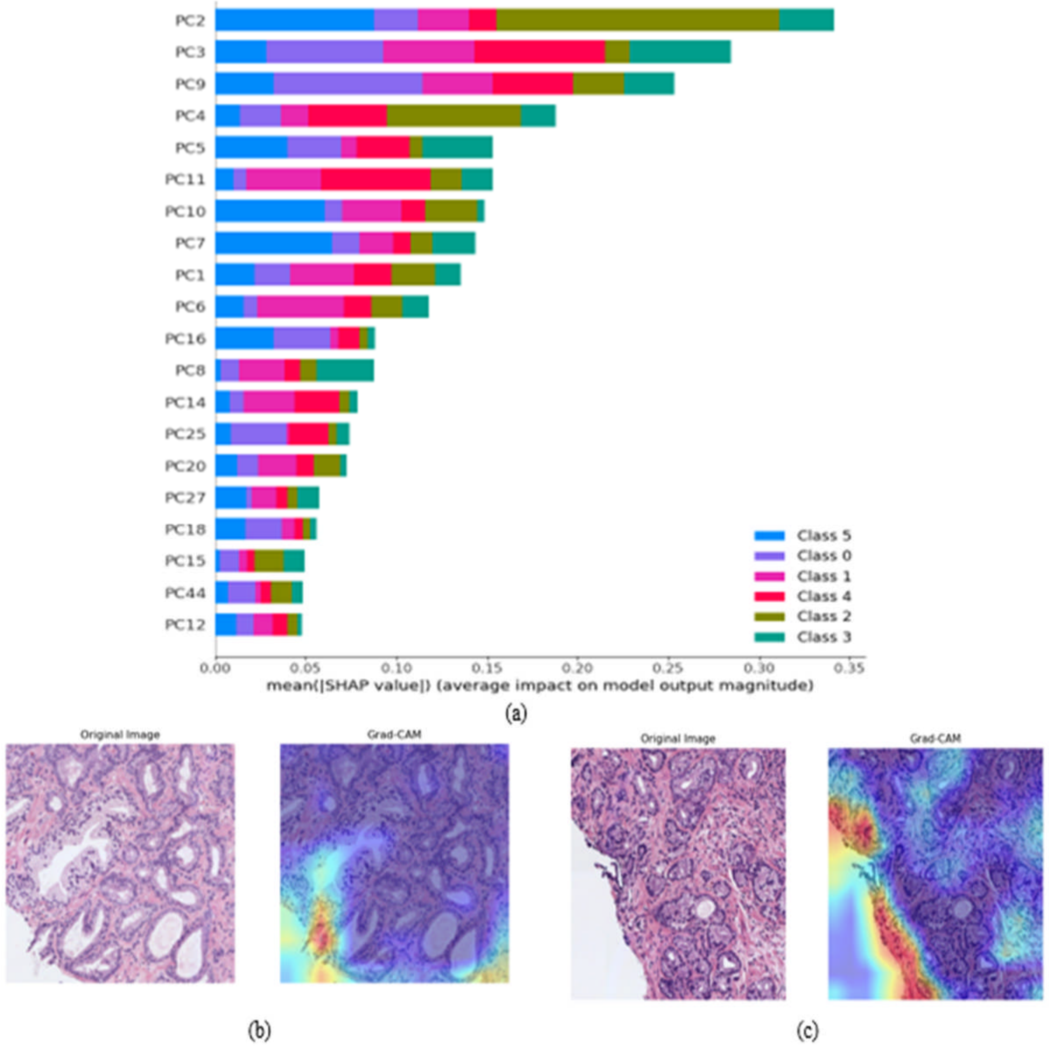
Model with Hyperparameters Tuning



Metric	DL – DNN with 3 Dense Layers	DL – DNN with 5 Dense Layers	DL – DNN with 5 Dense Layers + Dropout & Batch Normalization
Accuracy	0.79	0.89	**94.6**
Precision	0.81	0.90	**96**
Recall	0.79	0.89	**94**
F1-Score	0.80	0.89	**0.95**

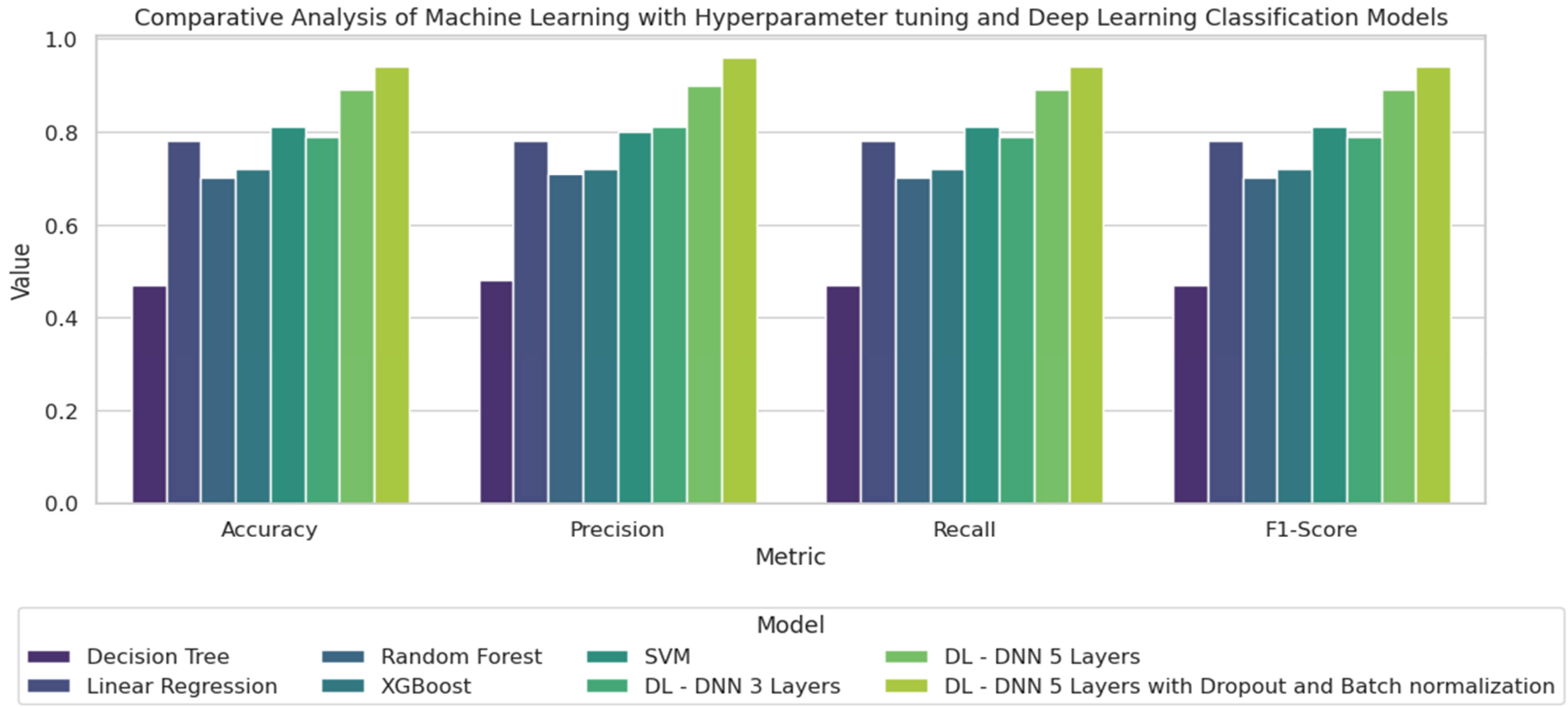
From Black Box to Clear Box:

Interpreting AI with Grad-CAM: Understanding Feature Relevance and Significance



1. Comparison of Explainability Techniques for Prostate Cancer Gleason Grade Classification (a) Significant contributions of 50 features to classification using XAI SHAP (b) Grad-CAM heatmap for GG2 and (c) Grad-CAM heat map for GG4

DL – DNN Models



Conclusion

In this study, a comprehensive framework for GG classification using WSI of PCa was developed by integrating DNN and ML models. VGG-16 was identified as the optimal feature extractor, offering a balance of feature quality and computational efficiency by extracting 512 features in 63.06 seconds. It outperformed DenseNet-121 and ResNet-50 in reducing redundancy and ensuring efficient dimensionality reduction through PCA, SVD, and t-SNE. Hyperparameter tuning enhanced ML performance, with SVM achieving the highest accuracy of 84%, while DL models incorporating dropout and batch normalization demonstrated significant improvements. A five-layer DNN achieved 94.6% accuracy, highlighting the effectiveness of regularization in preventing overfitting. A novel aspect of this research lies in the integration of XAI techniques to improve model interpretability. SHAP provided rapid, memory-efficient insights, while Grad-CAM delivered detailed visualizations, ensuring transparency in decision-making. LIME and Saliency Maps further contributed to understanding model outputs, underscoring the need for transparent AI in clinical settings. Future work will expand this framework to larger datasets and explore advanced neural architectures and XAI methods, aiming to develop scalable, interpretable, and clinically reliable AI models for PCa diagnostics. The implementation, tested on an open-access dataset, could benefit from additional testing on more benchmark and clinical datasets to enhance its clinical utility.

References :

- American Cancer Society, "Cancer Facts & Statistics," accessed Jan. 15, 2025. https://www.cancer.org/research/cancer-facts-statistics.html.
- American Cancer Society, "Cancer Statistics Center," accessed Jan. 15, 2025. https://cancerstatisticscenter.cancer.org/#/.
- PathologyOutlines.com, "Prostate WHO Classification," accessed Jan. 15, 2025. https://www.pathologyoutlines.com/topic/prostateWHO.html.
- J. G. Kench et al., "WHO Classification of Tumours fifth edition: evolving issues in the classification, diagnosis, and prognostication of prostate cancer," *Histopathology*, vol. 81, no. 4, pp. 447-458, 2022.
- A. B. Gavade et al., "Innovative Prostate Cancer Classification: Merging Auto Encoders, PCA, SHAP, and Machine Learning Techniques," presented at *Int. Conf. Adv. Robot. Control Artif. Intell. (ARCAI 2024)*, Perth, Australia, Dec. 9–12, 2024. (unpublished).
- A. B. Gavade et al., "Automated diagnosis of prostate cancer using mpMRI images: A deep learning approach for clinical decision support," *Computers*, vol. 12, no. 8, p. 152, 2023.
- K. A. Gadad et al., "Beyond Single Models: Hybrid Approaches for Multiclass Cancer Identification," in *2024 3rd Int. Conf. Adv. Technol. (ICONAT)*, pp. 1-6, IEEE, 2024.
- R. B. Nerli et al., "Artificial Intelligence and Histopathological Diagnosis of Prostate Cancer," *J. Sci. Soc.*, vol. 51, no. 2, pp. 153-156, 2024.
- A. S. Balraj et al., "PRADclass: Hybrid Gleason Grade-Informed Computational Strategy Identifies Consensus Biomarker Features Predictive of Aggressive Prostate Adenocarcinoma," *Technol. Cancer Res. Treat.*, vol. 23, p. 15330338231222389, 2024.
- A. Firjani et al., "A diffusion-weighted imaging based diagnostic system for early detection of prostate cancer," *J. Biomed. Sci. Eng.*, vol. 6, no. 3, pp. 346, 2013.
- N. Singhal et al., "A deep learning system for prostate cancer diagnosis and grading in whole slide images of core needle biopsies," *Sci. Rep.*, vol. 12, no. 1, p. 1-11, 2022.
- S. Azizi et al., "Deep recurrent neural networks for prostate cancer detection: analysis of temporal enhanced ultrasound," *IEEE Trans. Med. Imag.*, vol. 37, no. 12, pp. 2695-2703, 2018.
- W. Bulten et al., "Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study," *Lancet Oncol.*, vol. 21, no. 2, pp. 233-241, 2020.
- M. Tsuneki, M. Abe, and F. Kanavati, "A deep learning model for prostate adenocarcinoma classification in needle biopsy whole-slide images using transfer learning," *Diagnostics*, vol. 12, no. 3, p. 768, 2022.
- P. Pati et al., "Weakly supervised joint whole-slide segmentation and classification in prostate cancer," *Med. Image Anal.*, vol. 89, p. 102915, 2023.
- D. Müller et al., "DeepGleason: a System for Automated Gleason Grading of Prostate Cancer using Deep Neural Networks," *arXiv preprint arXiv:2403.16678*, 2024.
- K. Hammouda et al., "Multi-Stage Classification-Based Deep Learning for Gleason System Grading Using Histopathological Images," *Cancers*, vol. 14, no. 23, p. 5897, 2022.
- S. R. Duenweg et al., "Whole slide imaging (WSI) scanner differences influence optical and computed properties of digitized prostate cancer histology," *J. Pathol. Inform.*, vol. 14, p. 100321, 2023.
- G. Mittmann et al., "Pathologist-like explainable AI for interpretable Gleason grading in prostate cancer," *arXiv preprint arXiv:2410.15012*, 2024.
- A. Belinga, "AI-Enhanced Gleason Grading: A Comprehensive Approach," *arXiv preprint arXiv:2409.17122*, 2024.

Q&A?
Thank you