Asian Journal of Engineering and Applied Technology
ISSN: 2249-068X (P)
Vol.13 No.2, 2024, pp.44-51

www.crijournals.org
DOI: https://doi.org/10.70112/ajeat-2024.13.2.4261

Enhancing Prostate Cancer Prognosis through Digital Pathology and Machine Learning: A Systematic Review and Meta-Analysis

Chidiebere T. Ogbonna*, F. Y. Ayankoya and S. O. Kuyoro
Department of Computer Science, School of Computing, Babcock University, Ilishan-Remo, Nigeria
E-mail: ayankoyaf@babcock.edu.ng, kuyoros@babcock.edu.ng
*Corresponding Author: ogbonnac@student.babcock.edu.ng
(Received 17 September 2024; Revised 19 October 2024, Accepted 10 November 2024; Available online 15 November 2024)

***Abstract* - Analyzing digital pathology scans with machine learning can significantly improve prostate cancer prognosis. In recent years, machine learning (ML) algorithms have shown impressive capabilities in automating Gleason grading and prostate cancer prognostication, addressing challenges such as inter-observer variability among pathologists. This study investigates the development and validation of machine learning models specifically designed for prostate cancer prognostication. The study was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The review includes studies conducted between 2010 and 2024 that applied machine learning techniques to analyze digital pathology scans for prostate cancer prognosis. Inclusion criteria were studies focused on machine learning, prostate cancer, and digital pathology or whole-slide scans in the context of prostate cancer prognosis. Exclusion criteria included studies not involving machine learning, digital pathology, or prostate cancer, as well as studies not published in English or outside the specified time frame. Convolutional Neural Networks (CNNs) were the most commonly used machine learning approach in the reviewed studies, with brief mentions of other techniques. Meta-analysis was conducted using GraphPad Prism to create graphical representations of machine learning techniques employed in prostate cancer prognosis. The findings underscore the transformative potential of combining machine learning with digital pathology in revolutionizing prostate cancer prognosis. The integration of deep learning algorithms with digital pathology scans offers more accurate and efficient prognostication, significantly improving patient outcomes in the fight against prostate cancer.**
Keywords: **Prostate Cancer Prognosis, Machine Learning, Digital Pathology, Convolutional Neural Networks (CNNs), Meta-Analysis**

I. INTRODUCTION

In cancer-related causes of death, prostate cancer is the second leading cause of mortality in males. Despite recent advancements, prostate cancer continues to claim a significant number of lives due to its high prevalence. Bone metastases, which are common in prostate cancer patients, greatly affect patients' quality of life [1].

Consequently, improving disease outcome forecasts remains a critical clinical need. To enable precision care delivery, a robust process for assessing a patient's disease grade, stage, and trajectory must be established following diagnosis. Accurate pathological stage prognosis is essential for selecting the optimal treatment strategy in prostate cancer management [2]. Although the Gleason grade is a powerful predictor of outcomes, its utility is limited by significant inter observer variability. Incorporating standardized assessment techniques for prognostic biomarkers into conventional clinical protocols offers potential for improving the precision of identifying aggressive prostate tumors in the future [3].

Prostate-specific antigen (PSA) levels are one factor influencing prostate cancer staging. However, the Gleason score (GS) is the primary determinant of prostate cancer aggressiveness, based on the Gleason patterns observed during examination. GS typically ranges from 6 to 10, with lower values indicating low-grade cancer characterized by moderate growth, and higher values representing high-grade cancer associated with rapid spread [4]. Gleason scores are derived from Gleason grades and are used to establish prostate cancer prognosis.

Models built using various biomarkers, such as TMPRSS2: ERG fusion status, prostate-specific antigen (PSA), TP53 mutation, and collagen IV, demonstrate the potential of machine learning techniques in predicting the progression and prognosis of prostate cancer [5]-[7]. Examples of machine learning techniques used include support vector machines (SVM) [6], k-nearest neighbors (kNN), convolutional neural networks (CNN) [4], pyramid semantic parsing network (PSPNet) [8], and Gaussian classifiers.

This systematic review focuses on using machine learning techniques to analyze digital pathology scans and estimate a patient's prognosis or survival likelihood for prostate cancer. Machine learning is increasingly being integrated into the analysis of digital pathology images. Digital pathology image analysis algorithms, powered by machine learning and computer vision, have significantly improved the ability to recognize, segment, label, and classify various histological features linked to the molecular and spatial characteristics of prostate cancer [6]. The Gleason score, or the International Society of Urological Pathology (ISUP) grading system, obtained from prostate biopsies, remains the industry standard for assessing cancer stage and progression [9]. Therefore, integrating this grading system into machine learning methodologies is crucial.

II. RATIONALE

Prostate cancer is a leading cause of cancer-related deaths in men, underscoring the importance of accurate prognosis for effective treatment. However, traditional methods, such as Gleason grading, often encounter challenges, particularly due to inconsistencies arising from inter-observer variability. Machine learning, particularly deep learning approaches like Convolutional Neural Networks (CNNs), has shown significant potential in automating the analysis of digital pathology scans, reducing variability, and improving diagnostic accuracy. This study investigates the role of machine learning in enhancing prostate cancer prognosis, addressing existing gaps, and highlighting its potential to improve patient care.

III. OBJECTIVES OF THE STUDY

The primary objectives of this review are as follows.

1. To conduct a systematic review on the analysis of digital pathology scans in prostate cancer prognosis using machine learning (ML) techniques.
2. To evaluate the achievements of ML techniques in prostate cancer prognosis through the analysis of digital pathology scans.
3. To identify the benefits and limitations of ML techniques in prostate cancer prognosis using digital pathology scans.
4. To synthesize and analyze data extracted from selected articles to provide quantitative insights into the performance of ML techniques in prostate cancer prognosis.

IV. METHODOLOGY

A. Criteria for Eligibility

The PRISMA guidelines were followed to perform an exhaustive search on both Scopus and PubMed. The search included articles published between 2010 and 2024 in English. Keywords used in the search were “Prostate cancer,” “Digital pathology,” “Artificial intelligence,” “Deep learning,” “Machine learning,” and “Prognosis.”

The search strategy incorporated both keywords and Medical Subject Headings (MeSH) terms related to prognosis, digital pathology, machine learning, and prostate cancer. The search was conducted on PubMed and Scopus, targeting English-language articles published between 2010 and 2024.

This systematic review focused on articles utilizing any machine learning technique to analyze digital pathology scans in the context of prostate cancer prognosis.

B. Defining Criteria for Inclusion

1. Research focusing on prostate cancer prognosis using digital pathology scans analyzed with machine learning.
2. Studies with prostate cancer as the primary cohort.
3. Studies addressing the prognosis of prostate cancer.
4. Studies published in English between 2010 and 2024.

C. Defining Criteria for Exclusion

1. Research that does not focus on digital pathology scan analysis with machine learning for prostate cancer prognosis.
2. Studies not involving prostate cancer as the primary cohort.
3. Studies that do not include prostate cancer prognosis or digital pathology.
4. Studies not published in English.
5. Review articles or resources that are inaccessible.
6. Studies published before 2010 or after 2024.

V. SEARCH APPROACH USED

The search approach used and advanced key on Scopus and PubMed extensive supply of materials to identify studies that were of use to this systematic review.

The search queries used on Scopus gave 64 articles:
TITLE-ABS-KEY ((“Digital” OR “Automated”) AND (“Pathology” OR “Study”) AND (“Scans” OR “Analysis”) AND (“Machine Learning” OR “Deep Learning” OR “Artificial Intelligence” OR “AI” OR “Big Data”) AND (“Prostate Cancer” OR “Prostate Carcinoma”) AND (“Prognosis” OR “Prediction” OR “Outcome”)) AND PUBYEAR > 2009 AND PUBYEAR < 2025 AND PUBYEAR > 2012 AND PUBYEAR < 2025 AND (LIMIT-TO (SUBJAREA , “MEDI”) OR LIMIT-TO (SUBJAREA , “BIOC”) OR LIMIT-TO (SUBJAREA , “COMP”) OR LIMIT-TO (SUBJAREA , “HEAL”)) AND (LIMIT-TO (LANGUAGE , “English”))

The search query used on PubMed gave 56 articles:
((“digit pathol 2019”[Journal] OR (“digital”[All Fields] AND “pathology”[All Fields]) OR “digital pathology”[All Fields] OR ((“whole”[All Fields] OR “wholeness”[All Fields] OR “wholes”[All Fields]) AND (“slide”[All Fields] OR “slides”[All Fields] OR “sliding”[All Fields] OR “slidings”[All Fields]) AND (“image”[All Fields] OR “image s”[All Fields] OR “imaged”[All Fields] OR “imager”[All Fields] OR “imager s”[All Fields] OR “imagers”[All Fields] OR “images”[All Fields] OR “imaging”[All Fields] OR “imaging s”[All Fields] OR “imagings”[All Fields]))) AND (“radionuclide imaging”[MeSH Terms] OR (“radionuclide”[All Fields] AND “imaging”[All Fields]) OR “radionuclide imaging”[All Fields] OR “scanning”[All Fields] OR “scan s”[All Fields] OR “scanned”[All Fields] OR “scannings”[All Fields] OR “scans”[All Fields] OR (“analysis”[MeSH Subheading] OR “analysis”[All Fields])) AND (“machine learning”[MeSH Terms] OR (“machine”[All Fields] AND “learning”[All Fields]) OR “machine learning”[All Fields] OR (“deep learning”[MeSH Terms] OR (“deep”[All Fields] AND “learning”[All Fields]) OR “deep learning”[All Fields])) AND (((“prostatic

neoplasms"[MeSH Terms] OR ("prostatic"[All Fields] AND "neoplasms"[All Fields]) OR "prostatic neoplasms"[All Fields] OR ("prostate"[All Fields] AND "cancer"[All Fields]) OR "prostate cancer"[All Fields]) AND ("prognosis"[MeSH Terms] OR "prognosis"[All Fields] OR "prognoses"[All Fields])) OR (("prostat"[All Fields] OR "prostate"[MeSH Terms] OR "prostate"[All Fields] OR "prostates"[All Fields] OR "prostatic"[All Fields] OR "prostatism"[MeSH Terms] OR "prostatism"[All Fields] OR "prostatitis"[MeSH Terms] OR "prostatitis"[All Fields]) AND ("carcinoma"[MeSH Terms] OR "carcinoma"[All Fields] OR "carcinomas"[All Fields] OR "carcinoma s"[All Fields]) AND ("predict"[All Fields] OR "predictabilities"[All Fields] OR "predictability"[All Fields] OR "predictable"[All Fields] OR "predictably"[All Fields] OR "predicted"[All Fields] OR "predicting"[All Fields] OR "prediction"[All Fields] OR "predictions"[All Fields] OR "predictive"[All Fields] OR "predictively"[All Fields] OR "predictiveness"[All Fields] OR "predictives"[All Fields] OR "predictivities"[All Fields] OR "predictivity"[All Fields] OR "predicts"[All Fields])) OR (("prostatic neoplasms"[MeSH Terms] OR ("prostatic"[All Fields] AND "neoplasms"[All Fields]) OR "prostatic neoplasms"[All Fields] OR ("prostate"[All Fields] AND "cancer"[All Fields]) OR "prostate cancer"[All Fields]) AND ("predict"[All Fields] OR "predictabilities"[All Fields] OR "predictability"[All Fields] OR "predictable"[All Fields] OR "predictably"[All Fields] OR "predicted"[All Fields] OR "predicting"[All Fields] OR "prediction"[All Fields] OR "predictions"[All Fields] OR "predictive"[All Fields] OR "predictively"[All Fields] OR "predictiveness"[All Fields] OR "predictives"[All Fields] OR "predictivities"[All Fields] OR "predictivity"[All Fields] OR "predicts"[All Fields])) OR (("prostat"[All Fields] OR "prostate"[MeSH Terms] OR "prostate"[All Fields] OR "prostates"[All Fields] OR "prostatic"[All Fields] OR "prostatism"[MeSH Terms] OR "prostatism"[All Fields] OR "prostatitis"[MeSH Terms] OR "prostatitis"[All Fields]) AND ("carcinoma"[MeSH Terms] OR "carcinoma"[All Fields] OR "carcinomas"[All Fields] OR "carcinoma s"[All Fields]) AND ("prognosis"[MeSH Terms] OR "prognosis"[All Fields] OR "prognoses"[All Fields]))))) AND (2010:2024[pdat])

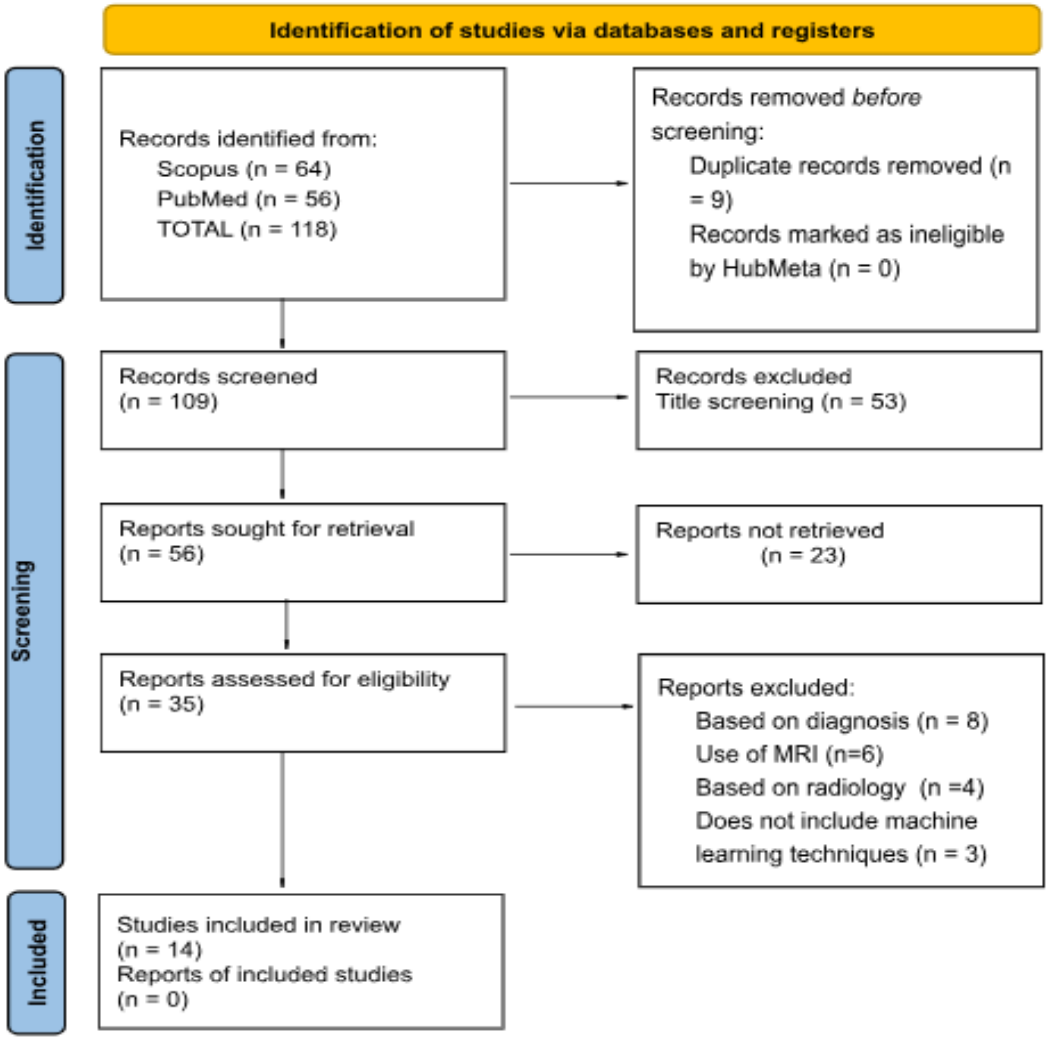

Fig. 1 A flowchart diagram showing the process of screening articles

A. Selection Process

This review employed the PICOS framework to define the inclusion and exclusion criteria.

1. *Participants:* Males diagnosed with prostate cancer.
2. *Interventions:* Studies utilizing machine learning techniques to analyze digital pathology scans or whole-slide imaging for prostate cancer prognosis.
3. *Comparison:* Various machine learning techniques used to analyze digital pathology scans or whole-slide imaging.
4. *Outcomes:* Analysis of prognostic accuracy metrics for machine learning models.
5. *Study Design:* Retrospective and prospective studies.

VI. DATA COLLECTION PROCESS

A. Data Extraction, Sorting, and Selection

A systematic approach was employed to extract and organize data from selected studies. Discrepancies during the process were resolved using predefined eligibility criteria, AI-supported tools, and expert judgment. The extracted data included essential study details such as participant demographics, interventions, outcomes, and key findings.

The screening process consisted of three distinct stages:

1. Search results were imported into Mendeley and transferred to HubMeta, where duplicates were removed, and articles were organized for review.
2. Titles and abstracts were screened using an AI assistant in conjunction with a human reviewer, following the PICOS framework.
3. Full-text articles meeting the inclusion criteria were thoroughly reviewed to confirm their relevance.

B. Data Items

The extracted data included the following: study titles, authors, publication years, journals, methodologies, sample sizes, inclusion/exclusion criteria, and data sources.

C. Assessment of Quality

Each study was evaluated based on four core criteria:

1. *Selection Bias:* Assessment of participant selection methods.
2. *Instrument Reliability:* Evaluation of the dependability of measurement tools.
3. *Handling of Missing Data:* Review of methods for managing incomplete data.
4. *Accuracy of Results Reporting:* Verification of transparent and accurate results reporting.

To assess the utility of machine learning in analyzing digital pathology for prostate cancer prognosis, metrics such as accuracy, mean, and standard deviation were employed. Additional performance indicators included AUC-ROC, Gleason score, and the Dice Similarity Coefficient (DSC).

From an initial pool of 118 studies sourced from Scopus and PubMed, 14 were deemed eligible after applying the PICOS framework and inclusion/exclusion criteria. The selection process was visually represented using a PRISMA flowchart (Figure 1). After duplicates were removed and abstracts screened, the final selection, spanning 2010 to 2024, was summarized in Table I.

D. Measures of Summary

The performance of predictive models was summarized using metrics such as odds ratios (OR) and hazard ratios (HR), presented with 95% confidence intervals (CI).

E. Strategy for Data Integration and Synthesis

A PRISMA flowchart (Figure 1) illustrated the article selection process, highlighting the number of studies retained at each stage. Key findings were summarized in a narrative synthesis, with a table detailing datasets, machine learning models, accuracy, and outcomes. When sufficient data was available, a meta-analysis was conducted using a random-effects model, with heterogeneity measured by the I^2 statistic.

F. Bias Assessment

Potential biases - including selection, performance, detection, attrition, and reporting biases - were critically examined. Disagreements were resolved through reassessment, and studies with significant bias were excluded from the final review. Although machine learning models generally demonstrated strong performance in predicting prostate cancer outcomes, challenges persisted in managing large datasets. Recent algorithmic advances could reduce the need for dataset segmentation, enabling real-time patient data integration to enhance predictive accuracy and clinical applications.

G. Study Design

This systematic review focused on predictive models for prostate cancer prognosis using machine learning and digital pathology scans. The study adhered to PRISMA guidelines, ensuring rigor in the selection process and quality assessment.

VII. DATA EXTRACTION, SORTING, AND SELECTION

Study data were collected using a standardized extraction form, and inconsistencies were resolved through a combination of AI-assisted screening and expert judgment. Extracted data included details on study characteristics, participant demographics, interventions, outcomes, and results.

The selection process involved three steps:

1. Search results were imported into Mendeley and HubMeta, where duplicate entries were removed.
2. Titles and abstracts were evaluated by an AI tool and a reviewer based on the PICOS criteria.
3. Full-text articles were reviewed to ensure they met the eligibility requirements.

TABLE I DATA EXTRACTED FROM 14 SYSTEMATIC REVIEWED PUBLICATIONS

Sl. No.	Author(s), Year	ML Models Used	Dataset Source	Accuracy	AUC-ROC	Outcome
1	Hammouda *et al.,* (2021) [4]	CNN (pyramidal deep learning)	608 Whole Slide Images (WSIs)	Patch: High	Not stated	Improved automatic Gleason grade classification with precision ~80%, recall 60-80%, F1-score.
2	Qiu *et al.,* (2022) [8]	Pyramid Semantic Parsing Net	Vancouver Prostate Centre, MICCAI 2019	Not stated	Not stated	Effective segmentation for Gleason grades, distinguishing low-risk and high-risk cancers.
3	Pizurica *et al.,* (2023) [7]	TiDo deep learning model	Whole Slide Images (WSIs)	Not stated	Not stated	Predicted TP53 mutations linked to aggressive disease phenotypes.
4	Karageorgose *et al.,* (2024) [6]	Deep learning-based pipeline	CD31, CD34, Collagen IV images	Precision: ~93-95%	Not stated	Detected blood vessels; correlations with Gleason grades and 5-year recurrence rates.
5	Melo PAS *et al.,* (2021) [1]	Inception v3, Mask R-CNN	Radical prostatectomy WSIs	91.2%-94.1%	Not stated	Identified Gleason patterns but lower concordance with pathologists in test set (44%).
6	Kott *et al.,* (2021) [10]	Residual CNN	85 Prostate Core Biopsies	85.4%-91.5%	Not stated	Effective coarse/fine patch classification but struggled with adjacent Gleason patterns.
7	Jake *et al.,* (2022) [11]	3D CNN (nnU-Net)	PET/CT scans ([68Ga]Ga-PSMA-11)	>90%	PPV: 97.2%	Biomarker extraction correlated strongly with patient survival outcomes.
8	Nishio *et al.,* (2023) [12]	EfficientNet with LDL	10,616 Whole Slide Images	40.7%	0.364	Improved diagnostic performance using label distribution learning (LDL).
9	Paulson *et al.,* (2022) [2]	VGG-16, Extreme Gradient Boosting	Hematoxylin and eosin-stained biopsies	Not stated	0.72 (avg)	Predicted adverse pathology in GG 2 and 3 prostate biopsies.
10	Sang *et al.,* (2011) [13]	SVM, ANN	Transrectal ultrasound-guided biopsies	Not stated	SVM: 0.805	SVM outperformed ANN in predicting advanced prostate cancer (>pT3a).
11	Ikromjanov *et al.,* (2023) [9]	EfficientNetB2 U-Net	Whole Slide Images (WSIs)	Not stated	Not stated	Segmented benign, cancerous, and stroma tissue for histological analysis.
12	Marini *et al.,* (2021) [14]	CNN (semi-supervised learning)	Multiple heterogeneous datasets	Not stated	Not stated	Enhanced Gleason grading performance across datasets with sparse annotations.
13	Omar *et al.,* (2024) [5]	Attention-based deep learning	TCGA prostate adenocarcinoma WSIs	Not stated	0.73	Predicted TMPRSS2 fusion, identifying morphologic features linked to survival outcomes.
14	Blessin *et al.,* (2023) [3]	AI with multiplex fluorescence	Tissue Microarrays (TMA)	Not stated	Not stated	Automated Ki-67 labeling index correlated with Gleason score and provided robust prognosis.

VIII. REPORT ON META-ANALYSIS

The analysis indicates that the mean AUC-ROC is 0.655 (95% CI: 0.487-0.822), suggesting that the predictive models exhibit moderate to good ability in distinguishing outcomes for prostate cancer prognosis. Among these, Support Vector Machines (SVMs) demonstrate superior performance.

A. Performance Distribution

The AUC-ROC values across the studies ranged from 0.364 to 0.805, highlighting variability in model effectiveness. A breakdown of the models is as follows.

B. Relative Model Ranking

1. *Lowest-Performing Model:* Nishio *et al.,* (AUC-ROC: 0.364) exhibited poor accuracy, potentially due to dataset challenges or methodological limitations.
2. *Highest-Performing Model:* Sang *et al.,* (AUC-ROC: 0.805) demonstrated strong predictive capability, likely reflecting better algorithm optimization or dataset alignment.

TABLE II SHOWING THE AUC-ROC VALUES

Studies	Model	Auc-Roc
1. Sang *et al.,*	SVM	0.805
2. Omar *et al.,*	DL attention-based	0.73
3. Paulson *et al.,*	CNN, XGBoost	0.72
4. Nishio *et al.,*	EfficientNet + LDL	0.364

This ranking, presented in Table II, suggests that SVM and attention-based deep learning (DL) models outperform others, indicating a potential advantage for models utilizing support vector-based optimization and attention mechanisms.

C. Consistency and Variability

1. *Standard Deviation:* The standard deviation of AUC-ROC values (0.179) indicates moderate variability among the models.
2. *Implications:* This variability could stem from differences in dataset sizes, preprocessing techniques, feature selection, or the complexity of the algorithms used.

D. Confidence Intervals

The 95% CI (0.487-0.822) highlights the range of expected model performances:

1. *Lower Bound (0.487):* Indicates suboptimal performance, nearing random guessing.

E. Result

AUC-ROC values above 0.7 (observed in three out of four models) are generally considered clinically meaningful, suggesting that these models provide reliable decision support. The outlier model (Nishio *et al.*) may benefit from methodological improvements or additional training on more diverse datasets.

F. Visual Analysis

A deeper visual analysis can provide further insights. A box plot and performance comparison chart will be generated to better illustrate the distribution and variability.

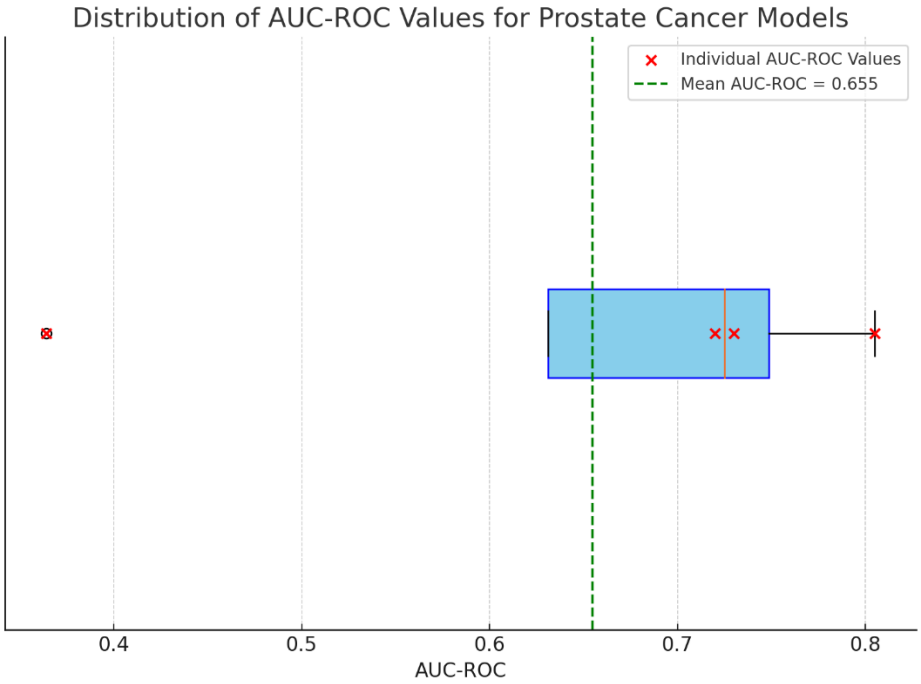

Fig. 2 Bar graph visually representing the distribution of AUC-ROC values for Prostate cancer models

The AUC-ROC box plot for prostate cancer models (Fig. 2) illustrates the distribution of model performance across studies, showcasing the range, median, and spread of values. Individual red dots represent the AUC-ROC values for each study, while the box highlights the interquartile range (middle 50% of the data).

The green dashed line in Fig. 2 marks the mean AUC-ROC (0.655), indicating moderate overall model performance. The variability in values, ranging from near-random guessing (e.g., Nishio *et al.,*) to strong predictive accuracy (e.g., Sang *et al.*), reflects differences in dataset quality, preprocessing techniques, and algorithmic methodologies among the studies. This visualization emphasizes the potential for improving underperforming models through enhanced training and optimized feature selection.

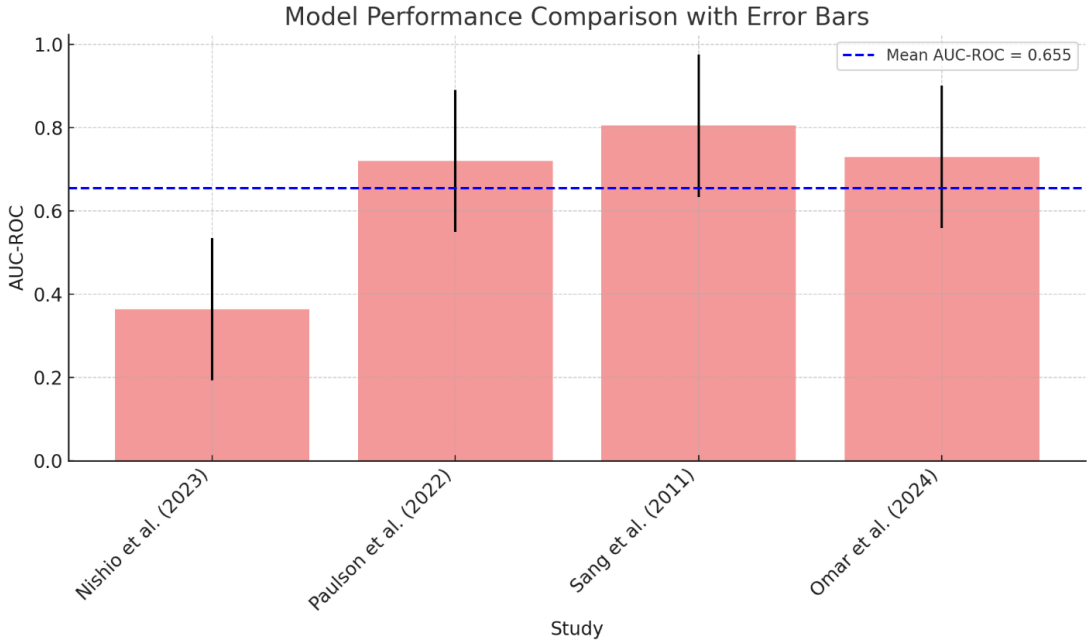

Fig. 3 Bar graph visually representing the AUC-ROC values for the included studies, with the mean, error bars and confidence interval highlighted

The bar graph in Fig. 3 visually represents the AUC-ROC values for the included studies, with the mean, error bars, and confidence intervals highlighted.

The visualizations provide additional insights:

1. *Box Plot (Fig. 2):* Highlights the range and distribution of AUC-ROC values, with a noticeable spread between the highest and lowest values. The mean (green dashed line) shows a central tendency around 0.655.
2. *Bar Chart with Error Bars (Fig. 3):* Compares AUC-ROC values across studies while accounting for variability (error bars). Models with higher AUC-ROC values, such as Sang *et al.*, demonstrate superior performance, while Nishio *et al.,* stands out as an outlier with significant room for improvement.

TABLE III SHOWING THE DESCRIPTIVE STATISTICS

Mean (Average AUC-ROC)	**0.655**
Standard Deviation	0.198 (indicating moderate variability)
Minimum Value	0.364 (Nishio *et al.,*)
Maximum Value	0.805 (Sang *et al.,*)
Range	0.441 (difference between highest and lowest values)

IX. DISCUSSION

This research emphasizes the growing impact of machine learning in analyzing digital pathology scans to improve prostate cancer prognosis. The reported AUC-ROC values indicate a range of performance levels, with some models, such as SVM and deep learning approaches like CNNs and attention-based architectures, showing strong predictive abilities. However, the variability observed, with AUC-ROC scores ranging from 0.364 to 0.805, highlights important factors such as the quality of the dataset, preprocessing techniques, and the complexity of the models. Advanced approaches, including label distribution learning (LDL) and semi-supervised techniques, demonstrate the potential to enhance accuracy even when dealing with limited annotations or heterogeneous datasets. In contrast, studies with lower performance suggest opportunities to refine methods, especially in handling complex data and improving feature selection.

Combining machine learning with digital pathology offers great promise in identifying aggressive prostate tumors and supporting personalized treatment decisions. However, achieving consistent reliability remains a challenge. Expanding the datasets, standardizing preprocessing methods, and integrating clinical data, such as PSA levels, Gleason scores, and biomarkers, could significantly improve model performance and usability in clinical settings.

A. Prognosis

The analysis reveals an average AUC-ROC of 0.655, indicating moderate predictive performance among the models. The lowest-performing study (Nishio *et al.*, 0.364) suggests variability in effectiveness, while the highest (Sang *et al.*, 0.805) demonstrates strong predictive capability. Moderate variability (standard deviation: 0.198) highlights differences in methodology and dataset quality. While most models achieve clinically relevant AUC-ROC (>0.7), improvements are needed to bridge the performance gap and ensure consistent reliability for prostate cancer prognosis.

X. CONCLUSION

This study shows that machine learning has significant potential to improve prostate cancer prognosis through digital

pathology analysis. The mean AUC-ROC of 0.655 reflects moderate overall performance, with some models demonstrating strong potential for clinical use. However, the variation in outcomes points to the need for further refinement of algorithms and evaluation techniques. Future efforts should focus on improving model design, leveraging diverse datasets, and incorporating standardized methodologies to ensure better reliability and accuracy. By addressing these areas, machine learning could become an indispensable tool in prostate cancer care, enabling earlier diagnoses and more precise treatment.

REFERENCES

[1] P. A. de S. Melo, C. L. N. Estivallet, M. Srougi, W. C. Nahas, and K. R. M. Leite, "Detecting and grading prostate cancer in radical prostatectomy specimens through deep learning techniques," *Clinics (Sao Paulo)*, vol. 76, p. e3198, 2021, doi: 10.6061/clinics/2021/e3198.

[2] N. Paulson *et al.,* "Prediction of adverse pathology at radical prostatectomy in grade group 2 and 3 prostate biopsies using machine learning," *JCO Clin Cancer Inform.*, vol. 6, p. e2200016, 2022, doi: 10.1200/CCI.22.00016.

[3] N. C. Blessin *et al.,* "Automated Ki-67 labeling index assessment in prostate cancer using artificial intelligence and multiplex fluorescence immunohistochemistry," *Journal of Pathology*, vol. 260, no. 1, pp. 5-16, 2023, doi: 10.1002/path.6057.

[4] K. Hammouda *et al.,* "A deep learning pipeline for grade groups classification using digitized prostate biopsy specimens," *Sensors (Basel)*, vol. 21, no. 20, 2021, doi: 10.3390/s21206708.

[5] M. Omar *et al.,* "Semi-supervised, attention-based deep learning for predicting TMPRSS2: ERG fusion status in prostate cancer using whole slide images," *Mol. Cancer Res.*, 2024, doi: 10.1158/1541-7786.MCR-23-0639.

[6] G. M. Karageorgos *et al.,* "Deep learning-based automated pipeline for blood vessel detection and distribution analysis in multiplexed prostate cancer images," *Frontiers in Bioinformatics*, vol. 3, 2023, doi: 10.3389/fbinf.2023.1296667.

[7] M. Pizurica *et al.,* "Whole slide imaging-based prediction of TP53 mutations identifies an aggressive disease phenotype in prostate cancer," *Cancer Res.*, vol. 83, no. 17, pp. 2970-2984, 2023, doi: 10.1158/0008-5472.CAN-22-3113.

[8] Y. Qiu *et al.,* "Automatic prostate Gleason grading using pyramid semantic parsing network in digital histopathology," *Front. Oncol.*, vol. 12, 2022, doi: 10.3389/fonc.2022.772403.

[9] K. Ikromjanov *et al.,* "Region segmentation of whole-slide images for analyzing histological differentiation of prostate adenocarcinoma using ensemble EfficientNetB2 U-Net with transfer learning mechanism," *Cancers (Basel)*, vol. 15, no. 3, 2023, doi: 10.3390/ cancers15030762.

[10] M. Pizurica *et al.,* "Whole slide imaging-based prediction of TP53 mutations identifies an aggressive disease phenotype in prostate cancer," *Cancer Res.*, vol. 83, no. 17, pp. 2970-2984, 2023, doi: 10.1158/0008-5472.CAN-22-3113.

[11] O. Kott *et al.,* "Development of a deep learning algorithm for the histopathologic diagnosis and Gleason grading of prostate cancer biopsies: A pilot study," *Eur. Urol. Focus*, vol. 7, no. 2, pp. 347-351, 2021, doi: 10.1016/j.euf.2019.11.003.

[12] J. Kendrick, R. J. Francis, G. M. Hassan, P. Rowshanfarzad, J. S. L. Ong, and M. A. Ebert, "Fully automatic prognostic biomarker extraction from metastatic prostate lesion segmentations in whole-body [68Ga]Ga-PSMA-11 PET/CT images," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 50, no. 1, pp. 67-79, 2022, doi: 10.1007/s00259-022-05927-1.

[13] M. Nishio, H. Matsuo, Y. Kurata, O. Sugiyama, and K. Fujimoto, "Label distribution learning for automatic cancer grading of histopathological images of prostate cancer," *Cancers (Basel)*, vol. 15, no. 5, 2023, doi: 10.3390/cancers15051535.

[14] S. Y. Kim *et al.,* "Pre-operative prediction of advanced prostatic cancer using clinical decision support systems: accuracy comparison between support vector machine and artificial neural network," *Korean J. Radiol.*, vol. 12, no. 5, pp. 588-594, 2011, doi: 10.3348/kjr.2011.12. 5.588.

[15] N. Marini, S. Otálora, H. Müller, and M. Atzori, "Semi-supervised training of deep convolutional neural networks with heterogeneous data and few local annotations: An experiment on prostate histopathology image classification," *Med. Image Anal.*, vol. 73, p. 102165, 2021, doi: 10.1016/j.media.2021.102165.

[16] E. E. Onuiri, O. Akande, O. B. Kalesanwo, T. Adigun, K. Rosanwo, and K. C. Umeaka, "A systematic review of machine learning prediction models for colorectal cancer patient survival using clinical data and gene expression profiles," 2023, doi: 10.18280/ria.370520.

[17] E. E. Onuiri and O. J. Adeniyi, "Evaluating machine learning models for predicting prostate cancer progression using lifestyle factors: A systematic review and meta-analysis," *Asian Journal of Engineering and Applied Technology*, vol. 13, no. 1, pp. 44-56, Apr. 2024, doi: 10.70112/ajeat-2024.13.1.4241.

[18] V. Pushpalatha, "A novel framework for detection of cervical cancer," *Asian Journal of Engineering and Applied Technology*, vol. 7, no. 2, pp. 26-30, Nov. 2018, doi: 10.51983/ajeat-2018.7.2.1016.

[19] A. Arokiyamary Delphina, M. Kamarasan, and S. Sathiamoorthy, "Self-organization map based segmentation of breast cancer," *Asian Journal of Engineering and Applied Technology*, vol. 7, no. 2, pp. 31-36, Nov. 2018, doi: 10.51983/ajeat-2018.7.2.1015.

[20] A. O. Olosunde and E. E. Onuiri, "Deep learning models for early detection of ovarian cancer: A systematic review of ultrasound-based diagnostic tools," *Asian Journal of Engineering and Applied Technology*, vol. 13, no. 2, pp. 15-21, Oct. 2024, doi: 10.70112/ajeat-2024.13.2.4246.

[21] S. T. Hsu, Y. J. Su, C. H. Hung, M. J. Chen, C. H. Lu, and C. E. Kuo, "Automatic ovarian tumors recognition system based on ensemble convolutional neural network with ultrasound imaging," *BMC Med. Inform. Decis. Mak.*, vol. 22, no. 1, Dec. 2022, doi: 10.1186/S12911-022-02047-6.