

Spatial characterization of prostate cancer microenvironment with DRAQ5&Eosin fluorescence imaging and quantitative feature identification

Michail Georgios Papachristos, Emiliano Spezi, Carolina Fuentes, Rachel J. Errington, Dimitris Parthimos



PII: S2153-3539(26)00208-7

DOI: <https://doi.org/10.1016/j.jpi.2026.100746>

Reference: JPI 100746

To appear in:

Received date: 3 July 2026

Revised date: 25 September 2026

Accepted date: 28 September 2026

Please cite this article as: M.G. Papachristos, E. Spezi, C. Fuentes, et al., Spatial characterization of prostate cancer microenvironment with DRAQ5&Eosin fluorescence imaging and quantitative feature identification, (2026), <https://doi.org/10.1016/j.jpi.2026.100746>

This is a PDF of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability. This version will undergo additional copyediting, typesetting and review before it is published in its final form. As such, this version is no longer the Accepted Manuscript, but it is not yet the definitive Version of Record; we are providing this early version to give early visibility of the article. Please note that Elsevier's sharing policy for the Published Journal Article applies to this version, see: <https://www.elsevier.com/about/policies-and-standards/sharing#4-published-journal-article>. Please also note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



# **Spatial Characterization of Prostate Cancer Microenvironment with DRAQ5&Eosin Fluorescence Imaging and Quantitative Feature Identification**

Michail Georgios Papachristos<sup>a,\*</sup>, Emiliano Spezi<sup>b</sup>, Carolina Fuentes<sup>c</sup>, Rachel J. Errington<sup>a</sup>,

Dimitris Parthimos<sup>a</sup>

*<sup>a</sup>School of Medicine, Division of Cancer and Genetics, Cardiff University, UHW Main Building,  
Heath Park, Cardiff, CF14 4XN, Wales, UK*

*<sup>b</sup>Department of Biomedical Engineering & Biotechnology, Khalifa University, Shakhbout Bin  
Sultan St, Hadbat Al Za'Faranah, Zone 1, Abu Dhabi, United Arab Emirates*

*<sup>c</sup>School of Computer Science and Informatics, Cardiff University, Abacws, Senghennydd Road,  
Cardiff, CF24 4AG, Wales, UK*

Acknowledgments: Biosamples were obtained from the Wales Cancer Biobank (DOI: <http://doi.org/10.5334/ojb.46>) which is funded by Health and Care Research Wales. Other investigators may have received specimens from the same subjects.

CRediT Author Roles: 1) Michail Georgios Papachristos (Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft), 2) Emiliano Spezi (Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing), 3) Carolina Fuentes (Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing), 4) Rachel J. Errington (

---

\* Corresponding author

URL: PapachristosMG@cardiff.ac.uk (Michail Georgios Papachristos)

Conceptualization, Funding acquisition, Investigation, Methodology, Resources, Supervision, Visualization, Writing – review & editing), 5) Dimitris Parthimos (Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing).

## 1. Introduction

According to the Global Cancer Observatory by the World Health Organization, in 2022 there were 1.4 million new worldwide cases of prostate cancer, ranking as the fourth cancer type in terms of incidence, causing the death of almost four hundred thousand individuals.<sup>1</sup> Prostatic adenocarcinoma incidence and progression is reflected in the tissue microstructure/microenvironment, which includes the epithelial gland and, importantly, its relationship and interaction with the surrounding stroma. This homeostatic and morphological disruption is manifested particularly during cancerous infiltration of the tissue.<sup>2</sup> In order to identify and grade prostate cancer, expert pathologists visually observe the spatial architecture of prostatic tissue stained with H&E, and assign a Gleason grade ranging from 1-5 at different epithelial regions.<sup>3</sup> This standard grading system, however, is reliant on individual expertise and suffers from inter- and intra-observer variability.<sup>4</sup> Utilizing mathematically quantifiable and standardized metrics to analyze the prostatic tissue microenvironment would lead to more robust and reproducible classification of distinct tissue patterns/architectures, and aid in more accurate clinical diagnosing. Characterizing how both the epithelial and stromal components of the tissue are affected is vital, considering that research in the role of stroma in prostate cancer is limited compared to other types of cancer.<sup>2</sup>

Over the last decade, there has been considerable advancement in the development of

digital medical imaging. This involves digital pathology such as high-resolution whole-slide microscopy imaging,<sup>5</sup> or digital radiology, focusing on radiological images such as Computed Tomography (CT), Positron Emission Tomography (PET), or Magnetic Resonance Imaging (MRI).<sup>6</sup> Digital medical imaging allows for the potential incorporation of computational analysis and artificial intelligence in both research and clinical settings.<sup>7</sup> Medical image analysis often incorporates deep learning and several other machine learning approaches.<sup>8</sup> While predictive, a major limitation of the use of deep learning is the lack of explainability and interpretability of models that are often considered black boxes. More explainable and interpretable methods of analysis include using handcrafted features for classification as opposed to the computational abstract extracted features of deep learning models. This mathematical quantification of different regions of interest is possible using feature extraction algorithms describing the intensity, morphology, and textural information, through the analytic framework known as radiomics<sup>9</sup> or pathomics<sup>10</sup> for digital radiology and digital pathology respectively.

Most deep learning and pathomics approaches usually analyze the standard H&E stained bright-field microscopy images. A more recent staining and imaging approach utilizes Fluorescence Microscopy (FM),<sup>11</sup> which is able to provide high resolution quantitative images of the nuclear and cytoplasmic aspects of the tissue using nuclear fluorescent stains.<sup>11</sup> An advantage of FM over H&E is the possibility of rapid real-time staining of the tissue,<sup>12</sup> and annotation by pathologists utilizing synthetic H&E images generated from FM data.<sup>13</sup> One method developed by Elfer et al. uses the DRAQ5 (highly-specific nuclear/DNA stain)<sup>14</sup> and Eosin fluorescent dyes to generate two different grayscale images, a DRAQ5 image containing the nuclear information of the tissue, and an Eosin image containing the cytoplasmic information of the tissue.<sup>15</sup>

In the literature, there have been approaches utilizing pathomics to classify regions of

prostatic tissue for different spatial patterns, including Gleason grading.<sup>16;17</sup> However, these studies used standard H&E staining for the pathomics analysis. To our knowledge, there is a research gap with regard to employing FM data instead, which can provide high spatial- and intensity-resolution multidimensional grayscale images, and thus can allow for specific biomarker identification of separate tissue characteristics through different stains. Identifying spatial biomarkers through pathomics can yield new insights into the biological significance of individual tissue components, as well as the interaction between the cancerous epithelium and its surrounding stromal microenvironment. Moreover, spatial profiling of tissue can potentially be used in the clinic for more precise characterization, prediction, and risk stratification of different types of prostatic adenocarcinoma.

In this work we use the pathomics analysis feature extraction methodology on DRAQ5&Eosin fluorescent images to mathematically analyze and quantify how different regions of prostatic tissue alter from healthy to cancerous, and how they differentiate across Gleason grades and histological patterns. The objective of this explanatory study is to identify important epithelial and stromal tissue spatial features using machine learning algorithms and feature selection methods through the examination of both the nuclei and cytoplasm.

We hypothesize that: (a) Fluorescence-derived engineered pathomics can distinguish healthy vs cancer tissue, and High vs Low Grade cancer, similarly to existing H&E-based pathomics models. (b) Cancer-associated structural changes are more strongly expressed in epithelial than stromal compartments when quantified from DRAQ5&Eosin fluorescence images. (c) Morphological, intensity-based, and textural pathomics feature classes capture the known disorganization associated with prostate cancer progression.

## 2. Methods and Materials

### 2.1. Overview of Datasets

The training/validation dataset used in this study consists of two Commercial Tissue Microarrays (TMAs) (“Commercial Tissue MicroArray A” and “Commercial Tissue MicroArray B”), and one additional TMA obtained from the Wales Cancer Biobank (WCB).<sup>18</sup> The external testing dataset consists of 10 additional cores also provided from WCB.<sup>18</sup> The Wales Cancer Biobank is approved as a Research Tissue Bank by Wales Research Ethics Committee (REC) 3 and is licensed by the Human Tissue Authority under the UK Human Tissue Act (2004) to store human tissue for research (research licence 12107). It should be noted that there are no overlapping patients between the training/validation dataset and the external testing dataset, and that each of the 10 cores in the external testing dataset corresponds to a different patient. Additionally, the areas of the slides/cores in the external testing dataset are significantly larger than the cores originating from the TMAs of the training/validation, allowing for tissue heterogeneity within individual crops. A detailed schematic of the datasets is provided in Figure 1.

Figure 1: Schematic of Datasets. The Commercial TMAs consisted of 49 needle biopsy prostate cores (samples from the two TMAs are the same but the slices are not sequential). The WCB TMA consisted of 103 cores. Commercial TMAs A and B along with the WCB TMA were all used as training/validation datasets (with 5-fold cross validation across the TMA cores). The WCB Tissue Slide Cores were used as the external testing dataset. Figure adapted from Papachristos et al.<sup>19</sup> (Creative Commons license: <http://creativecommons.org/licenses/by/4.0/>). Created in BioRender. <https://BioRender.com/>.

These image datasets are the same used in Papachristos et al.<sup>19</sup>, where the staining, imaging, data acquisition, and pre-processing protocols are outlined. The datasets contain for each core an unscaled 12-bit DRAQ5 and Eosin fluorescent image, a Gleason pattern mask provided by an expert pathologist (Healthy, Gleason pattern 3, Gleason pattern 4, Cribriform pattern, Glomeruloid pattern, and Gleason pattern 5), and a coarse pseudo-generated epithelial mask. For the purposes of this study, nuclear and cytoplasmic masks were generated for the DRAQ5 and Eosin images respectively. The nuclear masks of the DRAQ5 images (rescaled to 0-4095) were created using Multi Otsu thresholding<sup>20</sup> by the scikit-image package<sup>21</sup> (number of bins set to 4096, classes set to 2). The cytoplasm masks of the Eosin images (rescaled to 0-4095) were created using Multi Otsu thresholding<sup>20</sup> by the scikit-image package<sup>21</sup> (number of bins set to 4096, classes set to 3, with lower intensity threshold used to separate cytoplasm/background). Furthermore, the coarse epithelial masks were refined using a standard U-Net model trained on the training/validation dataset, and tuned on the external testing dataset, where the outputs of the U-Net model on both datasets were used as the refined epithelial masks (see Appendix A for more details regarding epithelial refinement).

The core images and masks were cropped into three different sizes: small  $(122.76 \times 122.76) \mu m^2$ , medium  $(245.52 \times 245.52) \mu m^2$ , and large  $(491.04 \times 491.04) \mu m^2$ . Before being cropped, each core image was contrast limited adaptive histogram equalized using the scikit-image package.<sup>21</sup> The medium and large crops were resized to  $256 \times 256$  pixels and the small crops were resized to  $128 \times 128$  pixels using nearest-neighbor interpolation from the opencv python package.<sup>22</sup> Each image crop also contains an associated refined epithelial mask, nuclei mask, cytoplasmic mask, Gleason pattern mask, most predominant pattern classification label, and patient of origin number.

Due to the size of the external testing dataset cores, there were some blurry crops on the edges of the tissue. The sharpness of each crop was quantified using the z-normalized variance of the Laplace filtered DRAQ5 image pixels within the crop's nuclei mask as a sharpness metric. Crops with a sharpness metric less than or equal to -1 were removed from the external testing dataset.

## **2.2. Feature Extraction**

The Python programming language<sup>23</sup> was used for all computational analysis through utilizing the Anaconda Navigator software<sup>24</sup> and the JupyterLab/ Notebook interface.<sup>25</sup> The algorithms were run on the high-performance computing cluster Falcon from the Advanced Research Computing at Cardiff (ARCCA) provided by the Wales Supercomputer infrastructure. Feature extraction was applied to the fluorescent DRAQ5&Eosin image crops. This was done using the Spaarc Pipeline for Automated Analysis and Radiomics Computing (SPAARC) (<https://www.spaarc-radiomics.io>).<sup>26;27</sup> SPAARC is a highly standardised package for the extraction of image features from both 2D and 3D medical images. It supports the extraction of 165 base features and 7 convolutional filters, all standardized following the guidelines of the Image Biomarker Standardisation Initiative (<https://theibsi.github.io>).<sup>9;27</sup>

From each DRAQ5 image crop two masks were constructed: an epithelial nuclei mask, and a stromal nuclei mask. The epithelial nuclei mask was generated by intersecting the nuclei mask with the epithelial mask corresponding to the classified pattern, while the remaining nuclei (i.e. those outside epithelial regions) were used to construct the stromal nuclei mask. From each Eosin image crop, an epithelial cytoplasm mask and a stromal cytoplasm mask were constructed

similarly. A visualization of DRAQ5&Eosin crops along with their associated epithelial nuclei, epithelial cytoplasm, stromal nuclei, and stromal cytoplasm masks for the 3 different crop sizes are presented in Figure 2. It should be noted that the stromal nuclear and cytoplasmic masks may include non-fibroblast related components, such as lymphocytes, endothelial cells, inflammatory cells, and segmentation artifacts. Thus, for the purposes of this study, stromal-based refers to regions involving the surrounding microenvironment as a whole.

Figure 2: Visualization of grayscale DRAQ5 and Eosin crops along with their associated masks for the 3 different crop sizes: small, medium and large.

Features were extracted from the DRAQ5 crop using epithelial and stromal nuclei masks, and from the Eosin crop using epithelial and stromal cytoplasm masks. The following feature categories were extracted: Morphological, Intensity-Based Statistics, Intensity Histograms, Intensity-Volume Histogram, Gray Level Co-occurrence Matrix (GLCM), Gray Level Run Length Matrix (GLRLM), Gray Level Size Zone Matrix (GLSZM), Gray Level Distance Zone Matrix (GLDZM), Neighbourhood Gray Tone Difference Matrix (NGTDM), and Neighbourhood Gray Level Dependence Matrix (NGLDM, coarseness parameter = 0) (812 total features). No interpolation or intensity range re-segmentation were used. The texture analysis type was set to 2D. Textural matrices and features (excluding GLSZM, and GLDZM) were extracted for three different distances. Short, medium, and long distances (1, 3, and 5) were used to capture texture information across fine, intermediate, and coarse spatial scales.

Because texture features can vary depending on how image intensities are discretized, we explored different intensity binning settings to identify the configuration that produced the most

stable and informative features. We used the Fixed Bin Number (FBN) and Fixed Bin Size (FBS)<sup>28</sup> methods with four bin values (8, 16, 32, 64), creating eight possible feature-extraction configurations. Combined with the three crop sizes (small, medium, large), this resulted in 12 total configurations for FBN intensity discretization models, and 12 total configurations for FBS intensity discretization models. The full machine learning pipeline described in the following section was applied to all configurations.

### ***2.3. Machine Learning Models***

Two types of models were constructed for the analysis of the DRAQ5&Eosin crops: epithelial features based, and stromal features based. For each type, we developed 3 binary classifiers: Healthy vs Low Grade Cancer (Gleason 3), Healthy vs High Grade Cancer (Gleason 4 (including cribriform and glomeruloid) or Gleason 5), and Low Grade Cancer vs High Grade Cancer. Group 5-fold cross validation was applied to the training/validation dataset using the StratifiedGroupKFold (for the Healthy vs Low Grade, and Low Grade Cancer vs High Grade classifications) and GroupKFold (for the Healthy vs High Grade classification) functions (scikit-learn package).<sup>29</sup> The group used to split the folds was the patient of origin, meaning that no two folds contained crops originating from the same tissue sample, core or patient. It should be mentioned that there was no patient overlap between the training/validation dataset, and the external testing dataset used to evaluate the developed models.

Within each training fold, the XGBoost<sup>30</sup> machine learning algorithm was used to develop the binary classification model. For the determination of the optimal hyperparameters, a grid search method was used for 6 different hyperparameters: learning rate (0.001, 0.01, 0.1), max depth (3, 5, 7), subsample (0.5, 0.75, 1), column sample by tree (0.5, 0.75, 1), number of estimators

(100, 300, 500, 700), and minimum sum of instance weight (hessian) needed in a child (3, 5, 10). The optimal hyperparameters were determined using the RandomSearchCV function with 50 random iterations (scikit-learn package)<sup>29</sup> with the cross validation strategy being selected by the GroupShuffleSplit function. Patient of origin was used as group of split, number of splits was set to 3, and train size was set to 0.8. The objective of the XGBoost was set to logistic regression for binary classification, the evaluation metric was set to Area Under the Curve (AUC), and the weights of classes were balanced to take into consideration class imbalance when training the models. Finally, the chosen models were assessed on the external testing dataset to measure their predictive power. To further examine in more detail how individual features contribute to the predictions of the models, the SHAP (SHapley Additive exPlanations)<sup>31</sup> values of the external testing dataset were utilized, and visualized using Beeswarm plots.

This machine learning pipeline was applied to develop both model categories, each of which included three binary classification models for the analysis of DRAQ5&Eosin crops. Moreover, it was repeated for each of the 8 feature extraction configurations, as well as 3 different crop sizes (small, medium, large), resulting in a total of 12 total configurations for FBN intensity discretization models, and 12 total configurations for FBS intensity discretization models. A schematic of the machine learning pipeline is presented in Figure 3.

Figure 3: Schematic of Machine Learning Pipeline. (1) The whole pipeline is run for all different possible configurations. Image adapted from Zwanenburg et al.<sup>28</sup> (Creative Commons license: <http://creativecommons.org/licenses/by/4.0/>) (2) For each combination, two categories of models were constructed (epithelial-based, and stromal-based), and for each category, 3 binary classifiers were developed. (3) The training pipeline for the development of

each model included Group 5-fold cross-validation, training using XGBoost, and hyperparameter tuning (4) Feature importance visualization involved heatmaps of importance and Beeswarm plots (Icons created in BioRender. <https://BioRender.com/>).

### 3. Results

#### 3.1. Optimal Configurations

The optimal configurations for the models were obtained based on the training/validation dataset. The metric used for selecting the models was the F1 score averaged across the five folds used for training/validation. First, the optimal intensity discretization method was determined by comparing the average F1-score across all the epithelial- and stromal-based models. Table 1 contains the average macro-F1 and AUC scores across all the epithelial- and stromal-based models. We determined that FBN discretization yielded a higher training/validation average F1 score compared to FBS discretization for both model types. We further selected the optimal bin number and crop size for each model based on its individual training/validation F1 score across all the FBN models. For each configuration, the model with the highest macro averaged F1 score across the five folds was chosen as the most predictive. In the following sections, we report results exclusively from the most predictive models. The optimal configurations across all the FBN models are presented in Table 2. The results for each individual configuration on the training/validation dataset can be found in Appendix B, on Tables B.1, B.2, B.3, and B.4.

| Average Training/Validation Model Performance |                      |               |                      |
|-----------------------------------------------|----------------------|---------------|----------------------|
| Epithelial-based                              |                      | Stromal-based |                      |
|                                               | Average F1 Scores(%) |               | Average F1 Scores(%) |

|                    | FBN          | FBS          |                    | FBN          | FBS   |
|--------------------|--------------|--------------|--------------------|--------------|-------|
| Small              | <b>69.70</b> | 69.20        | Small              | <b>64.73</b> | 64.37 |
| Medium             | 72.03        | <b>72.08</b> | Medium             | <b>65.72</b> | 64.68 |
| Large              | <b>75.22</b> | 75.20        | Large              | <b>66.18</b> | 64.94 |
| Total              | <b>72.32</b> | 72.16        | Total              | <b>65.54</b> | 64.67 |
| Average AUC Scores |              |              | Average AUC Scores |              |       |
|                    | FBN          | FBS          |                    | FBN          | FBS   |
| Small              | <b>0.785</b> | 0.781        | Small              | <b>0.734</b> | 0.730 |
| Medium             | <b>0.815</b> | 0.811        | Medium             | <b>0.745</b> | 0.732 |
| Large              | 0.851        | <b>0.857</b> | Large              | <b>0.783</b> | 0.762 |
| Total              | <b>0.817</b> | 0.816        | Total              | <b>0.754</b> | 0.741 |

Table 1: Average macro-F1 and AUC scores across all the epithelial- and stromal-based models.

| Optimal Configurations  |                  |     |               |     |
|-------------------------|------------------|-----|---------------|-----|
|                         | Epithelial-based |     | Stromal-based |     |
|                         | Crop Size        | FBN | Crop Size     | FBN |
| Low Grade VS Healthy    | Large            | 16  | Small         | 32  |
| High Grade VS Healthy   | Large            | 64  | Medium        | 8   |
| High Grade VS Low Grade | Large            | 8   | Large         | 8   |

Table 2: Optimal bin number and crop size for each model selected based on its individual training/validation F1 score across all the FBN models.

### 3.2. External Testing Data Machine Learning Results

The training/validation dataset machine learning results can be found in Appendix C, on Tables

C.1, and C.2, and Figures C.1, and C.2. To determine the predictive capability of the models developed, they were evaluated on an independent external testing dataset. Table 3 and Figure 4 summarize the accuracies, macro-averaged F1 scores, AUCs, and ROC curves for all epithelial-based and stromal-based models evaluated on the external test dataset.

For each model, performance was calculated at the core/patient level via majority voting across crops belonging to the same core. If a core contained no crops or less than 10 crops for a specific model, then that crop was not given a classification. Table 4 contains the patient-level aggregated accuracies for the 10 patients of all the epithelial-based and stromal-based models, and the actual and predicted class for each core/patient. Figure 5 contains the confusion matrices of the models on both crop-level and patient-level.

| Crop-level Results      |                      |                      |                      |                      |
|-------------------------|----------------------|----------------------|----------------------|----------------------|
|                         | Epithelial-based     |                      |                      |                      |
|                         | Accuracy(%)          | macro-F1(%)          | AUC                  | Majority Baseline(%) |
| Low Grade VS Healthy    | 62.19 [59.06, 65.21] | 55.03 [51.79, 58.18] | 0.646 [0.609, 0.683] | 78.5                 |
| High Grade VS Healthy   | 80.04 [76.41, 83.47] | 78.54 [74.69, 82.25] | 0.896 [0.867, 0.924] | 58.47                |
| High Grade VS Low Grade | 72.13 [69.35, 74.81] | 67.74 [64.65, 70.68] | 0.759 [0.724, 0.794] | 72.22                |
|                         | Stromal-based        |                      |                      |                      |
|                         | Accuracy(%)          | macro-F1(%)          | AUC                  | Majority Baseline(%) |
| Low Grade VS Healthy    | 74.05 [73.33, 74.75] | 66.05 [65.20, 66.95] | 0.777 [0.767, 0.786] | 80.37                |
| High Grade VS Healthy   | 77.25 [75.44, 79.10] | 76.87 [75, 78.64]    | 0.840 [0.823, 0.857] | 59.23                |
| High Grade VS Low Grade | 70.98 [68.20, 73.75] | 67.08 [64.01, 70.04] | 0.742 [0.703, 0.780] | 72.22                |

Table 3: Crop-level aggregated (raw) accuracies, macro-averaged F1 scores, and AUCs of all the epithelial-based and stromal-based models evaluated on the external testing dataset. 95% confidence intervals (calculated using bootstrapping with 10000 resamplings) are provided next to the metrics in square brackets.

| Patient-level Results |                              |                                |                                |
|-----------------------|------------------------------|--------------------------------|--------------------------------|
| Accuracy(%)           |                              |                                |                                |
|                       | Epithelial-based             |                                |                                |
|                       | Low Grade VS Healthy         | High Grade VS Healthy          | High Grade VS Low Grade        |
| Patient/Core 1        | NA                           | 98.31 [94.92, 100] (G45/G45)   | 96 [90, 100] (G45/G45)         |
| Patient/Core 2        | NA                           | 94.12 [86.27, 100] (G45/G45)   | 80.39 [68.63, 90.20] (G45/G45) |
| Patient/Core 3        | NA                           | 90.91 [81.82, 98.18] (G45/G45) | 23.64 [12.73, 34.55] (G45/G3)  |
| Patient/Core 4        | 78.79 [63.64, 90.91] (G3/G3) | NA                             | 87.88 [75.76, 96.97] (G3/G3)   |
| Patient/Core 5        | NA                           | 57.50 [42.50, 72.50] (G45/G45) | 0 [0, 0] (G45/G3)              |
| Patient/Core 6        | 64.91 [52.63, 77.19] (G3/G3) | NA                             | 62.07 [50, 74.14] (G3/G3)      |
| Patient/Core 7        | 81.82 [72.73, 90.91] (G3/G3) | NA                             | 62.12 [50, 74.24] (G3/G3)      |
| Patient/Core 8        | NA                           | 98.80 [96.39, 100] (G45/G45)   | 91.57 [85.54, 97.59] (G45/G45) |
| Patient/Core 9        | 48.09 [42.82, 53.37] (G3/H)  | NA                             | 79.77 [75.37, 83.87] (G3/G3)   |
| Patient/Core 10       | 68.37 [64.14, 72.61] (G3/G3) | 64.88 [58.05, 71.22] (H/H)     | 72.43 [66.91, 77.57] (G3/G3)   |
|                       | Stromal-based                |                                |                                |
|                       | Low Grade VS Healthy         | High Grade VS Healthy          | High Grade VS Low Grade        |
| Patient/Core 1        | 82.64 [77.69, 87.19] (H/H)   | 84.68 [80.24, 89.11] (G45/G45) | 88 [78, 96] (G45/G45)          |
| Patient/Core 2        | NA                           | 61.62 [54.59, 68.65] (G45/G45) | 84.31 [74.51, 94.12] (G45/G45) |
| Patient/Core 3        | NA                           | 96.12 [93.20, 98.54] (G45/G45) | 27.28 [16.36, 40] (G45/G3)     |
| Patient/Core 4        | 70.18 [66.89, 73.35] (G3/G3) | NA                             | 39.39 [24.24, 57.58] (G3/G45)  |
| Patient/Core 5        | 26.61 [19.35, 34.68] (H/G3)  | 57 [50, 64] (G45/G45)          | 22.86 [8.57, 37.14] (G45/G3)   |
| Patient/Core 6        | 79.66 [76.75, 82.43] (G3/G3) | NA                             | 79.31 [68.97, 89.66] (G3/G3)   |
| Patient/Core 7        | 85.73 [83.53, 87.93] (G3/G3) | NA                             | 87.88 [78.79, 95.45] (G3/G3)   |
| Patient/Core 8        | 70.59 [47.06, 88.24] (G3/G3) | 82.61 [78.55, 86.67] (G45/G45) | 91.57 [85.54, 97.59] (G45/G45) |
| Patient/Core 9        | 71.46 [70.22, 72.65] (G3/G3) | 55.88 [38.24, 73.53] (G45/H)   | 72.14 [67.45, 76.83] (G3/G3)   |
| Patient/Core 10       | 74.89 [73.80, 75.94] (G3/G3) | 77.89 [75.09, 80.68] (H/H)     | 70.59 [65.07, 76.10] (G3/G3)   |

Table 4: Patient-level aggregated (raw) accuracies for the 10 patients of all the epithelial-based and

stromal-based models evaluated on the external testing dataset. 95% confidence intervals (calculated using bootstrapping with 10000 resamplings) are provided next to the metrics in square brackets. Patient level class is determined through majority voting with a minimum crop threshold of 10. The actual and predicted classes for each patient are given in parenthesis in the format of (actual/predicted class). H: Healthy, G3: Low Grade, G45: High Grade.

Figure 4: ROC curves of the models along with their associated AUCs on the external testing dataset: epithelial-based models (left), and stromal-based models (right).

Figure 5: Crop- and patient-level confusion matrices of the models: low grade cancer vs healthy (left), high grade cancer vs healthy (center), and high grade cancer vs low grade cancer (right) organized by epithelial-based models (top row), and stromal-based models (bottom row).

Finally, in order to gain greater model explainability, the impact of the features on the developed XGBoost outcomes are visualized. For each of the five cross-validation folds in a model, feature importance was determined using the mean of absolute SHAP values across the training samples. The features of each fold that make up 75% of the cumulative feature importance are considered significant. Common significant features across the five folds are defined as stable. In Figures 6 and 7, the SHAP Beeswarm plots of the 20 most impactful stable features of the external testing dataset for the epithelial-based and stromal-based model outputs are presented respectively. Non-stable features are not presented in the SHAP Beeswarm plot.

Figure 6: SHAP Beeswarm plots of 20 most impactful stable features (from top to bottom) of the

external testing dataset for the epithelial-based models: low grade cancer vs healthy (left), high grade cancer vs healthy (center), and high grade cancer vs low grade cancer (right).

Figure 7: SHAP Beeswarm plots of 20 most impactful stable features (from top to bottom) of the external testing dataset for the stromal-based models: low grade cancer vs healthy (left), high grade cancer vs healthy (center), and high grade cancer vs low grade cancer (right).

Further feature visualization can be found in Appendix D, on Figures D.1, D.2, D.3, D.4, and D.5. Details about the visualized features are explained in the Image Biomarker Standardization Initiative reference manual (<https://arxiv.org/pdf/1612.07003>).<sup>28;9</sup> The textural features contain distance metrics (d1, d3, d5) where: a) d1 captures fine textures and local intensity variations, but is sensitive to noise and small-scale heterogeneity, b) d3 represents intermediate patterns, balancing local and regional texture and also reduces sensitivity to noise compared to the shorter distance, and c) d5 captures coarse textures and global heterogeneity, and can be useful for detecting large-scale structural patterns. Overall the feature codes are given here as such: (IBSI code)\_(distance metric)\_(image type) (GLSZM and GLDZM features do not contain distance metrics, so in the feature code d1 is given as a default). The computational times required for feature extraction per configuration are reported in Appendix E, on Table E.1.

## 4. Discussion

Both models based on features extracted from the epithelium, and those based on features extracted from the surrounding stroma/microenvironment were found to be predictive when it comes to differentiating between healthy and cancerous tissue patterns. In terms of predictive

power, the epithelial-based models were found to be comparable to the stromal model with the exception of the Low Grade vs Healthy classification, where the stromal-based models were more predictive. For the epithelial-based models, the most accurate was the classifier between High Grade Cancer vs Healthy tissue regions with an accuracy of 80.04% and an AUC of 0.896. Furthermore, regarding the stromal-based models, the differentiation between Healthy and Gleason 3 regions achieved an accuracy of 74.05% with an AUC of 0.777, while the cancer stratification model (High Grade vs Low Grade Cancer) achieved a lower accuracy of 70.98% with an AUC of 0.742. In general, it seems that the models developed were able to differentiate between Low Grade and High Grade cancer at a patient level, with epithelial-based models being able to classify all 5 low-grade cancer patients correctly, but misclassifying 2 out of the 5 high-grade patients as low-grade patients. The stromal-based models had the same results, with the exception of one low-grade patient that was classified as high-grade instead.

Overall, based on the SHAP values on the external testing dataset, the most influential DRAQ5 intensity features on the epithelial-based model outputs that were also positively correlated with cancer incidence or progression model prediction were: the maximum intensity, maximum gradient, coefficient of variation, variance, range, and skew. Negatively correlated DRAQ5 intensity features were: median, mean, and 10th percentile. For Eosin, intensity-based energy was positively correlated, while minimum intensity and volume fraction difference between intensity fractions 10% and 90% were negatively correlated features. In the DRAQ5 stromal regions, the maximum intensity, interquartile range, variance, robust mean absolute deviation, median absolute deviation, volume fraction difference between intensity fractions 10% and 90%, minimum histogram gradient intensity, minimum gradient, and quartile coefficient of dispersion intensity features displayed positive correlations while the maximum gradient, skew,

and 10th percentile showed the opposite pattern. Finally, Eosin stromal intensity features that had a positive correlation with SHAP values were maximum intensity, range, variance, maximum gradient, minimum gradient, and 90th percentile, while minimum intensity was negatively associated with the SHAP value.

Regarding the morphological characteristics, the morphological area density of both DRAQ5 and Eosin were found to be consistently positively associated with cancer incidence or progression prediction across all epithelial-based models, with Eosin major axis length displaying a decrease in the Low Grade vs Healthy model. In the Low Grade vs Healthy stromal-based model, morphological area density was found to be negatively associated for both DRAQ5 and Eosin, while minor axis length was positively correlated with SHAP values. For the High Grade vs Healthy stromal-based model, morphological elongation showed a positive correlation with the SHAP values, while major axis length demonstrated the opposite. This pattern is reversed in the High Grade VS Low Grade model, where morphological area density, and major axis length has a positive model impact, while elongation has a negative one .

Additionally, the most important epithelial textural characteristics positively associated with cancer incidence or progression SHAP values were: NGLDM dependence count non-uniformity, NGLDM dependence count variance, GLRLM small runs emphasis, GLDZM small distance emphasis, and GLRLM run entropy. The most negatively associated epithelial textural features were: GLDZM large distance high gray level emphasis, GLCM cluster shade, GLRLM run length variance, GLDZM zone distance entropy, NGLDM high dependence high gray level emphasis, NGLDM large distance emphasis, GLDZM zone distance entropy. Finally, the SHAP positively associated stromal features were: NGLDM low dependence emphasis, NGLDM low dependence low gray level emphasis, GLDZM zone distance non-uniformity,

GLDZM low dependence high gray level emphasis, and GLCM second measure of information theoretic correlation. The negatively correlated stromal textural features were: the GLCM correlation, GLCM inverse variance .

Related papers, such as Garcia et al.<sup>16</sup> were able to achieve an AUC of  $0.922 \pm 0.045$  for distinguishing between benign and Gleason grade 3 glands using hand-crafted features, while Brancato et al.<sup>17</sup> attained an AUC of 0.97-0.99 for differentiating between non-malignant (including stroma) and malignant regions, and an AUC of 0.72-0.73 for separating Gleason 3 vs Gleason 4 or 5 regions. In work by Papachristos et al., deep learning models were used to analyze the same DRAQ5&Eosin data as our current feature-based models for prostate cancer tissue characterization.<sup>19</sup> Comparable results were achieved, with AUCs of 0.9314, 0.8398, and 0.7715 for the High Grade vs Healthy, Low Grade vs Healthy, and High Grade vs Low Grade models, respectively.<sup>19</sup> However, it should be mentioned that these papers make use of different methodologies and datasets, while also mainly focusing on predictability rather than tissue spatial profiling. Thus, while difficult to directly compare them with our current work, our results display an encouraging pattern. Moreover, a main advantage of the present models based on feature extraction is that they provide explainability because they rely on human-interpretable features. We were thus able to generate spatial profiles for different regions of healthy and cancerous tissue which is a major benefit over the developed deep learning models by Papachristos et al.<sup>19</sup> This could potentially be clinically important, since fluorescent microscopy through nuclear markers allows for rapid real time biopsy imaging.<sup>11;32;33</sup>

Cancer-associated structural changes in the DRAQ5&Eosin fluorescence images were overall found to be equally expressed in the stromal and epithelial components. The exception was the Low Grade VS Healthy models, where the stromal-based model was more predictive. This

demonstrates the potential importance of the surrounding stromal microenvironment, especially regarding initial prostate cancer development.

The tissue disorganization associated with prostate cancer progression was captured through intensity-based features in the epithelial nuclei (maximum gradient, coefficient of variation, quartile coefficient of dispersion, robust mean absolute deviation), along with an increase in the DNA content of the nuclei (maximum intensity). Intensity variability also presented in a spatial level across the tissue, through increased textural characteristics such as NGLDM dependence count non-uniformity, NGLDM dependence count variance, NGLDM low dependence emphasis, and GLDZM small distance emphasis. This suggests that different characteristics of the spatial intensity distribution of the nuclei and cytoplasm seem to change not only from healthy to cancerous regions, but also especially between different grades of cancer, which is captured in both the epithelial tissue, but also in its surrounding stromal microenvironment. Biologically, these spatial local intensity and texture distribution alterations of the microenvironment could be linked to the tumor environment heterogeneity that is associated with prostate cancer.<sup>34</sup> Spatial intensity non-uniformity in epithelial nuclei could be connected to cancerous prostatic biological processes such as chromatin dysregulation, remodeling, epigenetic modifications, 3D chromatin structure alteration, chromatin relaxation, and transcriptional plasticity<sup>35</sup>, or epithelial-mesenchymal transition.<sup>36</sup> Furthermore, stromal nuclei and cytoplasmic heterogeneity could be linked to microstructural cancerous changes which occur in the reactive stroma, characterized by an abundance of myofibroblasts and carcinoma-associated fibroblasts, producing enzymes linked to extracellular matrix remodeling such as collagen type I, tenascin, and FAP.<sup>37;38</sup> Additionally, the identified stromal biomarkers could also be associated with non-stromal microenvironment related biological processes, including angiogenesis due to release

of vascular endothelial growth factor (VEGF),<sup>39</sup> and immune response of various different subtypes including tumor-infiltrating lymphocytes, tumor-associated macrophages, and tumor-specific antigens.<sup>40</sup> Subsequent analysis could involve a combination of DRAQ5 and Eosin staining with different types of immunohistochemical stains, to link the identified cancerous spatial profiles with more specific bio-molecular processes.

Morphological spatial biomarkers were also identified, with the most prominent one being an increase in area density of both the epithelial nuclei and cytoplasm. This could potentially be because high Gleason grades are characterized by poorer gland differentiation, potentially presenting as less compact but more dense structures. Regarding the stroma, an initial decrease in cytoplasmic area density from Healthy to Gleason 3 was pinpointed, followed by a subsequent increase from Gleason 3 to higher cancerous grades. Similarly to our study, Garcia et al.<sup>16</sup> identified 136 relevant features for classification, where 94 of them were textural. Brancato et al.<sup>17</sup> reported that the most important features for distinguishing High Grade vs Low Grade Cancer were wavelet transformation and local binary pattern features.

The main limitation of this study is the size of the datasets utilized, which affects the predictability of the models. The external test used for this study is derived from 10 tissue cores/patients, and thus the effective sample size of this study is 10 samples, which limits the statistical power. Furthermore, for the purpose of this kind of explanatory analysis and spatial profiling, it is important for the underlying masks used for feature extraction to be well defined. In our study we used pseudo-generated epithelial masks that we refined using a U-Net model trained on the training/validation dataset, and tuned on the external testing dataset. This results in a well defined separation between the epithelium/stroma which is consistent across both datasets, but also potentially reduces the predictability of the models on foreign datasets, due to the epithelial

masks on those datasets not necessarily being similarly defined, resulting in inconsistent feature extraction. In future studies, more stable mask annotations could be achieved through a combination of the DRAQ5 and Eosin dyes with a fluorescent epithelial marker.

Another limitation is related to the inter-variability of the ground-truth provided by the expert pathologists, which influences the training and validation of the proposed models, and potentially affects the robustness of the identified tissue biomarkers. From Papachristos et al,<sup>19</sup> an average quadratic weighted Cohen kappa score of 0.857 between Pathologist A and Pathologist B on a subset of 9 small cores from the WCB TMA is reported. Future studies which make use of bigger datasets and consensus-oriented annotation by multiple pathologists could allow for a higher predictive power and statistical confidence of the important features identified. Additionally, a higher number of samples would allow to investigate features for better differentiation between Gleason 4 and Gleason 5, as well as specific architectural tissue patterns such as Cribriform, Glomeruloid, Mucinous Adenocarcinoma, and Intraductal Carcinoma.<sup>41</sup>

## 5. Conclusion

This explanatory study demonstrates a computational approach for spatially profiling epithelial and stromal regions of the prostatic microenvironment using fluorescent nuclear and cytoplasmic markers. Furthermore it shows that these features can be linked to healthy, low-grade, and high-grade cancer phenotypes using feature extraction and gradient boosting algorithms. Interestingly, the stromal-based models were found to have a comparable, and in some cases higher predictive power than their corresponding epithelial-based models, which highlights how crucial the relationship between stromal and epithelial tissue is in regards to cancer development and progression. The identified spatial biomarkers could establish a biologically meaningful link

to known biological processes underlying poorly differentiated prostate cancer. Future replication and reproducibility studies using different datasets would strengthen the applicability of the identified spatial biomarkers.

## References

1. F. Bray, M. Laversanne, H. Sung, J. Ferlay, R. L. Siegel, I. Soerjomataram, A. Jemal, Global cancer statistics 2022: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA: a cancer journal for clinicians* 74 (3) (2024) 229–263.
2. L. O. Gonzalez, N. Eiro, M. Fraile, N. Beridze, A. R. Escaf, S. Escaf, J. M. Fernandez-Gomez, F. J. Vizoso, Prostate cancer tumor stroma: responsibility in tumor biology, diagnosis and treatment, *Cancers* 14 (18) (2022) 4412.
3. J. I. Epstein, W. C. Allsbrook Jr, M. B. Amin, L. L. Egevad, I. G. Committee, et al., The 2005 international society of urological pathology (isup) consensus conference on gleason grading of prostatic carcinoma, *The American journal of surgical pathology* 29 (9) (2005) 1228–1242.
4. R. N. Flach, P.-P. M. Willemse, B. B. Suelmann, I. A. Deckers, T. N. Jonges, C. van Doijeweert, P. J. van Diest, R. P. Meijer, Significant inter-and intralaboratory variation in gleason grading of prostate cancer: a nationwide study of 35,258 patients in the netherlands, *Cancers* 13 (21) (2021) 5378.
5. N. Kiran, F. Sapna, F. Kiran, D. Kumar, F. Raja, S. Shiwlani, A. Paladini, F. Sonam, A. Bendari, R. S. Perakash, et al., Digital pathology: transforming diagnosis in the digital age, *Cureus* 15 (9) (2023).
6. G. Bansal, Digital radiography. a comparison with modern conventional imaging,

- Postgraduate medical journal 82 (969) (2006) 425–428.
7. L. Pinto-Coelho, How artificial intelligence is shaping medical imaging technology: a survey of innovations and applications, *Bioengineering* 10 (12) (2023) 1435.
  8. G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, C. I. Sánchez, A survey on deep learning in medical image analysis, *Medical image analysis* 42 (2017) 60–88.
  9. A. Zwanenburg, M. Vallières, M. A. Abdalah, H. J. Aerts, V. Andrearczyk, A. Apte, S. Ashrafinia, S. Bakas, R. J. Beukinga, R. Boellaard, et al., The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping, *Radiology* 295 (2) (2020) 328–338.
  10. R. D. Bülow, D. L. Hölscher, I. G. Costa, P. Boor, Extending the landscape of omics technologies by pathomics, *npj Systems Biology and Applications* 9 (1) (2023) 38.
  11. B. Rocco, A. Cimagli, L. Sarchi, L. R. Bonetti, L. Bertoni, P. Azzoni, S. Assumma, F. Turri, G. Bozzini, A. Eissa, et al., Current and future perspectives of digital microscopy with fluorescence confocal microscope for prostate tissue interpretation: a narrative review, *Translational Andrology and Urology* 10 (3) (2021) 1569.
  12. S. Puliatti, L. Bertoni, G. M. Pirola, P. Azzoni, L. Bevilacqua, A. Eissa, A. Elsherbiny, M. C. Sighinolfi, J. Chester, S. Kaleci, et al., Ex vivo fluorescence confocal microscopy: the first application for real-time pathological examination of prostatic tissue, *BJU international* 124 (3) (2019) 469–476.
  13. A. Eissa, S. Puliatti, N. R. Peñaranda, S. Resca, S. Di Bari, J. Vella, S. Maggiorelli, L. Bertoni, P. Azzoni, L. R. Bonetti, et al., Current applications of ex-vivo fluorescent confocal microscope in urological practice: a systematic review of literature, *Chinese Clinical*

- Oncology 13 (4) (2024) 52–52.
14. V. K. Ortner, A. Sahu, M. Cordova, K. Kose, S. Aleissa, C. Alessi-Fox, M. Haedersdal, M. Rajadhyaksha, A. M. Rossi, Exploring the utility of deep red anthraquinone 5 for digital staining of ex vivo confocal micrographs of optically sectioned skin, *Journal of biophotonics* 14 (4) (2021) e202000207.
  15. K. N. Elfer, A. B. Sholl, M. Wang, D. B. Tulman, S. H. Mandava, B. R. Lee, J. Q. Brown, Draq5 and eosin ('d&e') as an analog to hematoxylin and eosin for rapid fluorescence histology of fresh tissues, *PLoS One* 11 (10) (2016) e0165530.
  16. G. García, A. Colomer, V. Naranjo, First-stage prostate cancer identification on histopathological images: Hand-driven versus automatic learning, *Entropy* 21 (4) (2019) 356.
  17. V. Brancato, M. Verdicchio, C. Cavaliere, F. Isgro, M. Salvatore, M. Aiello, Unveiling key pathomic features for automated diagnosis and gleason grade estimation in prostate cancer, *BMC Medical Imaging* 25 (1) (2025) 299.
  18. A. Parry-Jones, L. K. Spary, The wales cancer bank (wcb), *Open Journal of Bioresources* 5 (10) (2018).
  19. M. G. Papachristos, E. Spezi, C. Fuentes, I. Evangelou, D. H. Thomas, F. Akinade, M. Wiltshire, A. Wilson, R. J. Errington, D. Parthimos, Prostate cancer tissue mapping and stratification using draq5 and eosin fluorescent labels integrated with ai classification and segmentation algorithms, *PloS one* 21 (3) (2026) e0345014.
  20. N. Otsu, et al., A threshold selection method from gray-level histograms, *Automatica* 11 (285-296) (1975) 23–27.
  21. S. Van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager,

- E. Gouillart, T. Yu, scikit-image: image processing in python, *PeerJ* 2 (2014) e453.
22. G. Bradski, The OpenCV Library, *Dr. Dobbs's Journal of Software Tools* (2000).
23. G. Van Rossum, F. L. Drake, et al., *Python reference manual*, Vol. 111, Centrum voor Wiskunde en Informatica Amsterdam, 1995.
24. A. Inc., Anaconda software distribution, Computer software. Anaconda Documentation (2020).
25. T. Kluyver, B. Ragan-Kelley, F. Pérez, B. Granger, M. Bussonnier, J. Frederic, K. Kelley, J. Hamrick, J. Grout, S. Corlay, et al., Jupyter notebooks—a publishing format for reproducible computational workflows, in: *Positioning and power in academic publishing: Players, agents and agendas*, IOS press, 2016, pp. 87–90.
26. P. Whybra, C. Parkinson, K. Foley, J. Staffurth, E. Spezi, Assessing radiomic feature robustness to interpolation in 18f-fdg pet imaging, *Scientific reports* 9 (1) (2019) 9649.
27. P. Whybra, A. Zwanenburg, V. Andrearczyk, R. Schaer, A. P. Apte, A. Ayotte, B. Baheti, S. Bakas, A. Bettinelli, R. Boellaard, et al., The image biomarker standardization initiative: standardized convolutional filters for reproducible radiomics and enhanced clinical insights, *Radiology* 310 (2) (2024) e231319.
28. A. Zwanenburg, S. Leger, M. Vallières, S. Löck, Image biomarker standardisation initiative, *arXiv preprint arXiv:1612.07003* (2016).
29. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research* 12 (2011) 2825–2830.
30. T. Chen, Xgboost: A scalable tree boosting system, Cornell University (2016).

31. S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, *Advances in neural information processing systems* 30 (2017).
32. Y. Liu, R. M. Levenson, M. W. Jenkins, Slide over: advances in slide-free optical microscopy as drivers of diagnostic pathology, *The American journal of pathology* 192 (2) (2022) 180–194.
33. A. K. Glaser, N. P. Reder, Y. Chen, E. F. McCarty, C. Yin, L. Wei, Y. Wang, L. D. True, J. T. Liu, Light-sheet microscopy for slide-free non-destructive pathology of large clinical specimens, *Nature biomedical engineering* 1 (7) (2017) 0084.
34. R. Ge, Z. Wang, L. Cheng, Tumor microenvironment heterogeneity an important mediator of prostate cancer progression and therapeutic resistance, *NPJ precision oncology* 6 (1) (2022) 31.
35. K. Kukkonen, S. Taavitsainen, L. Huhtala, J. Uusi-Makela, K. J. Granberg, M. Nykter, A. Urbanucci, Chromatin and epigenetic dysregulation of prostate cancer development, progression, and therapeutic response, *Cancers* 13 (13) (2021) 3325.
36. M. Montanari, S. Rossetti, C. Cavaliere, C. D’Aniello, M. G. Malzone, D. Vanacore, R. Di Franco, E. La Mantia, G. Iovane, R. Piscitelli, et al., Epithelial-mesenchymal transition in prostate cancer: an overview, *Oncotarget* 8 (21) (2017) 35376.
37. D. A. Barron, D. R. Rowley, The reactive stroma microenvironment and prostate cancer progression, *Endocrine-related cancer* 19 (6) (2012) R187–R204.
38. J. A. Tuxhorn, G. E. Ayala, M. J. Smith, V. C. Smith, T. D. Dang, D. R. Rowley, Reactive stroma in human prostate cancer: induction of myofibroblast phenotype and extracellular matrix remodeling, *Clinical cancer research* 8 (9) (2002) 2912–2923.
39. S. Ji, W. Wu, Q. Jiang, Crosstalk between endothelial cells and tumor cells: a new era in

- prostate cancer progression, *International Journal of Molecular Sciences* 24 (23) (2023) 16893.
40. R. Novysedlak, M. Guney, M. Al Khouri, R. Bartolini, L. Koumbas Foley, I. Benesova, A. Ozaniak, V. Novak, S. Vesely, P. Pacas, et al., The immune microenvironment in prostate cancer: a comprehensive review, *Oncology* 103 (6) (2025) 521–545.
41. J. I. Epstein, L. Egevad, M. B. Amin, B. Delahunt, J. R. Srigley, P. A. Humphrey, G. Committee, et al., The 2014 international society of urological pathology (isup) consensus conference on gleason grading of prostatic carcinoma: definition of grading patterns and proposal for a new grading system, *The American journal of surgical pathology* 40 (2) (2016) 244–252.

**Declaration of interests**

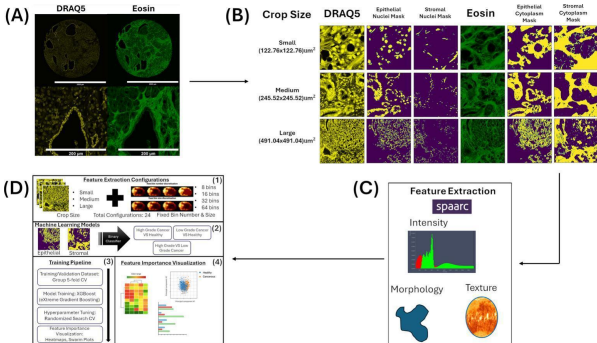
☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Emiliano Spezi reports financial support was provided by Engineering and Physical Sciences Research Council (EPSRC). Rachel J Errington reports a relationship with BioStatus Ltd that includes: equity or stocks. I, Rachel Errington, declare that I am a shareholder in BioStatus Ltd. The research presented in this paper utilizes products manufactured by BioStatus Ltd. This affiliation has been disclosed in the interest of transparency and does not influence the integrity or objectivity of the scientific findings reported. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Highlights

- Extracted hand-crafted quantitative features from fluorescent DRAQ5&Eosin images of prostatic adenocarcinoma tissue
- Spatially characterized prostate tumour microenvironment using gradient boosting algorithms
- Identified and visualized important nuclear and cytoplasmic features associated with cancer progression from both epithelial tissue and surrounding microenvironment



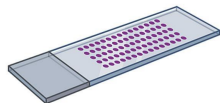
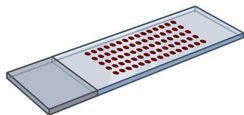
Graphics Abstract

**Training/Validation Datasets**  
5-fold Cross-Validation

**External Testing Dataset**

**Commercial  
Tissue MicroArray A**

DRAQ5&Eosin  
Synthetic H&E  
49 cores



Synthetic H&E:

Gleason Annotated by:  
Pathologist A (Ground  
Truth)

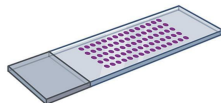
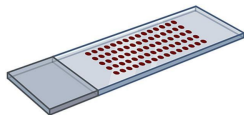
DRAQ5&Eosin:

Spatial Resolution:  
(0.69x0.69) $\mu\text{m}^2$

Core Size:  
 $\sim(2.6 \times 2.6) \text{mm}^2$

**Commercial  
Tissue MicroArray B**

DRAQ5&Eosin  
Synthetic H&E  
49 cores



Synthetic H&E:

Gleason Annotated by:  
Pathologist B (Ground  
Truth)

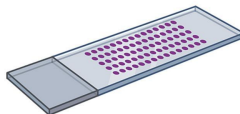
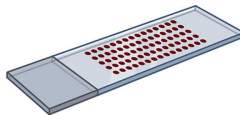
DRAQ5&Eosin:

Spatial Resolution:  
(0.52x0.52) $\mu\text{m}^2$

Core Size:  
 $\sim(2.6 \times 2.6) \text{mm}^2$

**Wales Cancer Biobank  
Tissue MicroArray**

DRAQ5&Eosin  
Sequential H&E  
103 cores



Sequential H&E:

Gleason Annotated by:  
Pathologist B (Ground  
Truth)

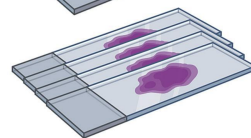
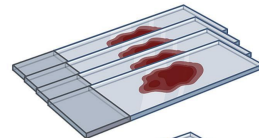
DRAQ5&Eosin:

Spatial Resolution:  
(0.52x0.52) $\mu\text{m}^2$

Core Size:  
 $\sim(1.1 \times 1.1) \text{mm}^2$

**Wales Cancer Biobank  
Tissue Slide Cores**

DRAQ5&Eosin  
Sequential H&E  
10 cores



Sequential H&E:

Gleason Annotated by:  
Pathologist B (Ground  
Truth)

DRAQ5&Eosin:

Spatial Resolution:  
(0.52x0.52) $\mu\text{m}^2$

Variable Core Size

Figure 1

**Crop Size**

**DRAQ5**

**Epithelial  
Nuclei Mask**

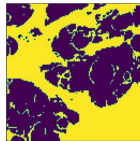
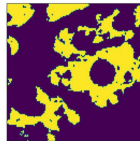
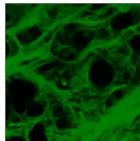
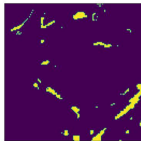
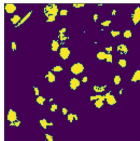
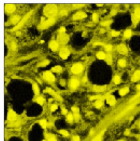
**Stromal  
Nuclei Mask**

**Eosin**

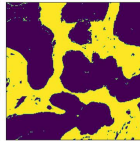
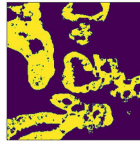
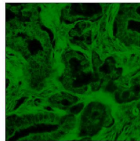
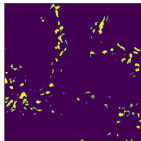
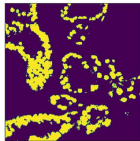
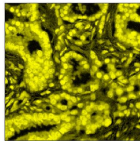
**Epithelial  
Cytoplasm  
Mask**

**Stromal  
Cytoplasm  
Mask**

**Small**  
**(122.76x122.76)um<sup>2</sup>**



**Medium**  
**(245.52x245.52)um<sup>2</sup>**



**Large**  
**(491.04x491.04)um<sup>2</sup>**

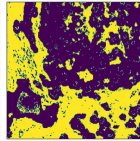
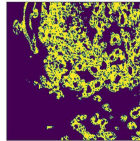
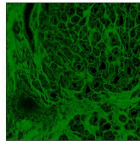
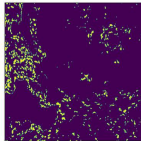
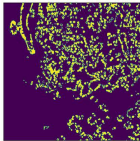
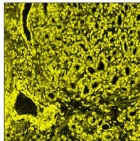
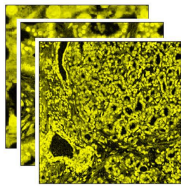


Figure 2

## Feature Extraction Configurations

(1)

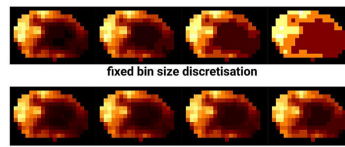


- Small
- Medium
- Large

Crop Size



Total Configurations: 24

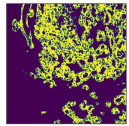


- 8 bins
- 16 bins
- 32 bins
- 64 bins

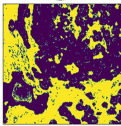
Fixed Bin Number & Size

## Machine Learning Models

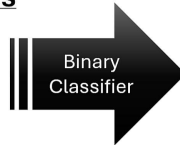
(2)



Epithelial



Stromal



High Grade Cancer  
VS Healthy

Low Grade Cancer  
VS Healthy

High Grade VS Low  
Grade Cancer

## Training Pipeline

(3)

Training/Validation Dataset:  
Group 5-fold CV

Model Training: XGBoost  
(eXtreme Gradient Boosting)

Hyperparameter Tuning:  
Randomized Search CV

Feature Importance  
Visualization:  
Heatmaps, Swarm Plots



## Feature Importance Visualization

(4)

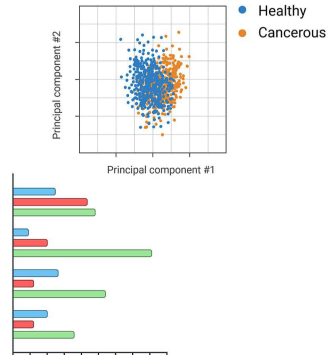
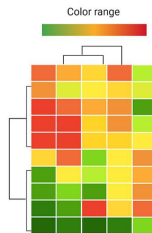


Figure 3

# ROC- External Testing Dataset

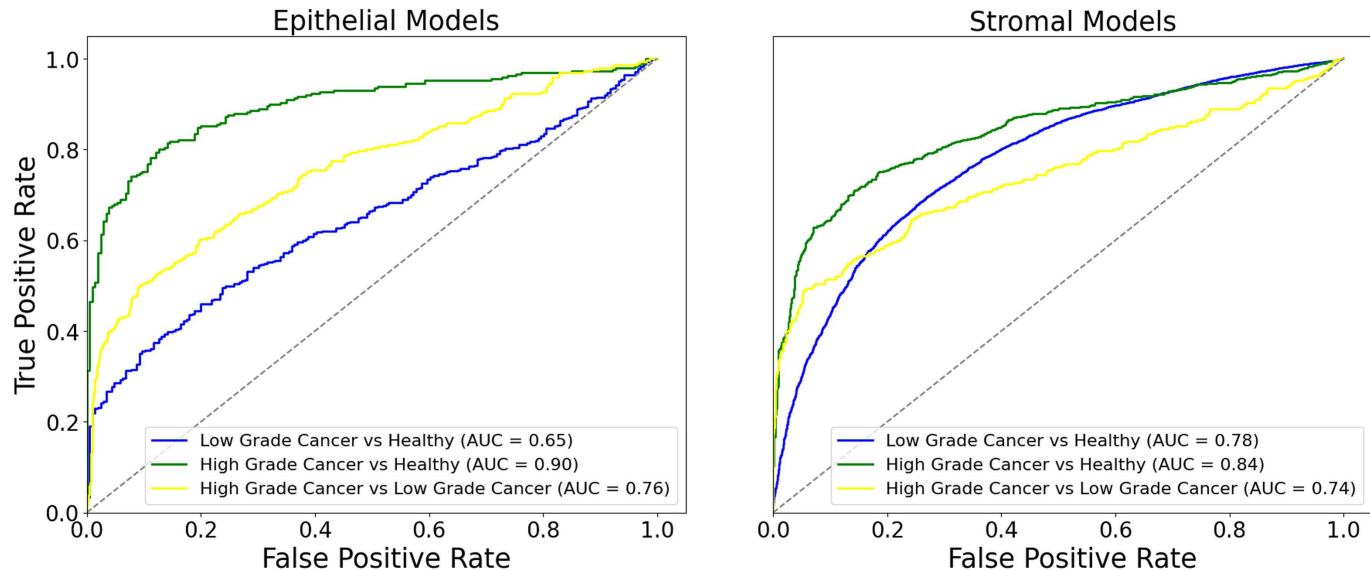
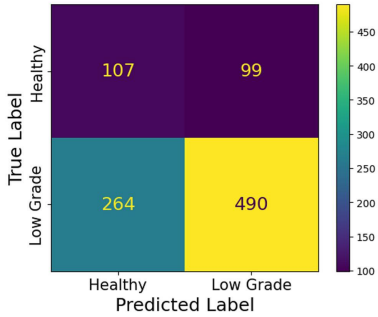


Figure 4

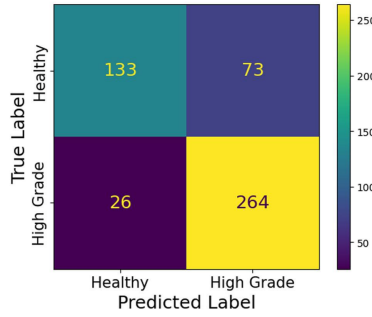
# Crop-level Confusion Matrices

Epithelial-based

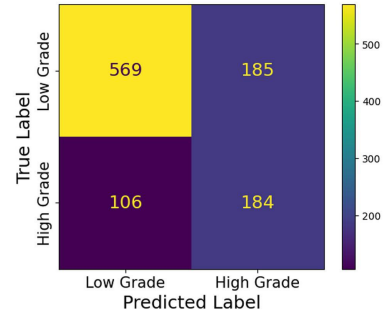
Low Grade Cancer  
VS Healthy



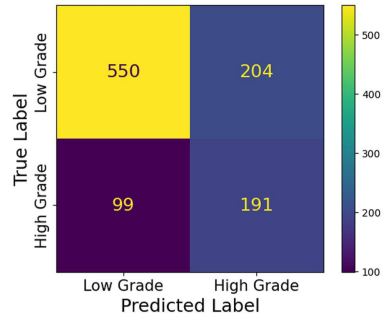
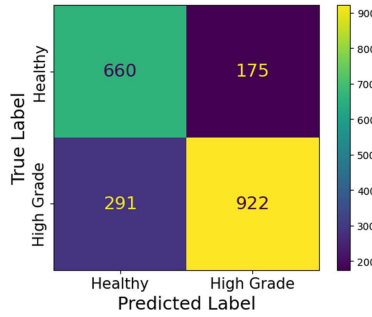
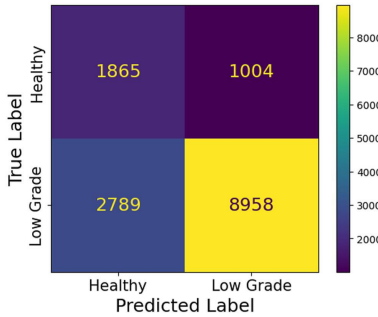
High Grade Cancer  
VS Healthy



High Grade Cancer  
VS Low Grade Cancer



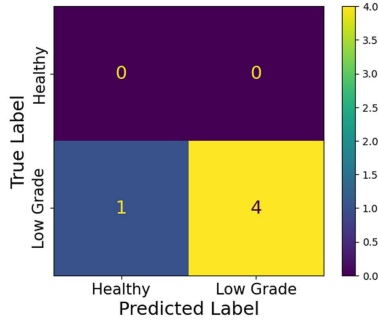
Stromal-based



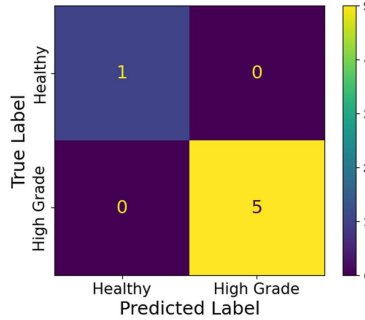
# Patient-level Confusion Matrices

Epithelial-based

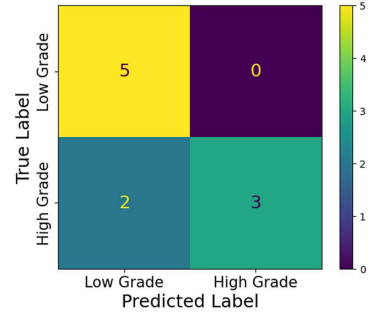
Low Grade Cancer  
VS Healthy



High Grade Cancer  
VS Healthy



High Grade Cancer  
VS Low Grade Cancer



Stromal-based

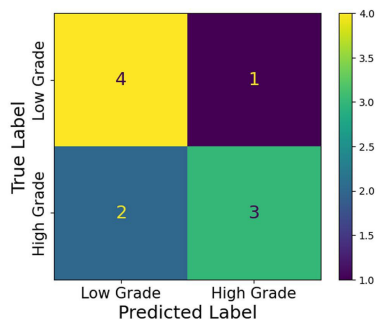
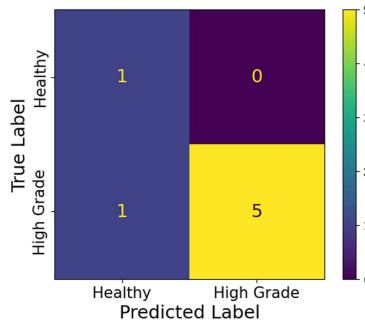
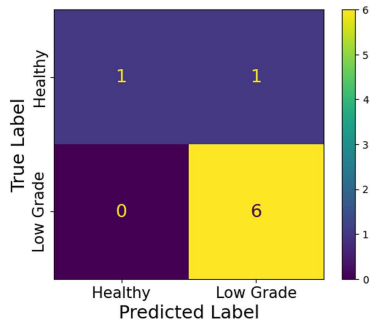
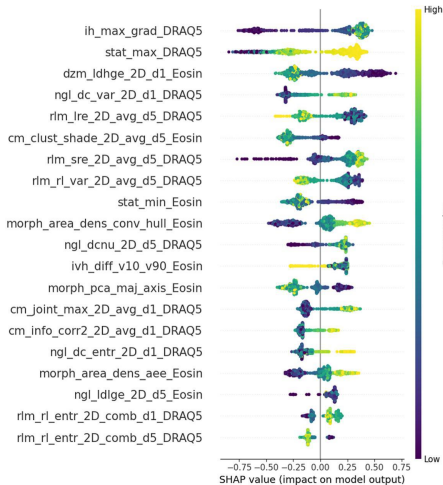


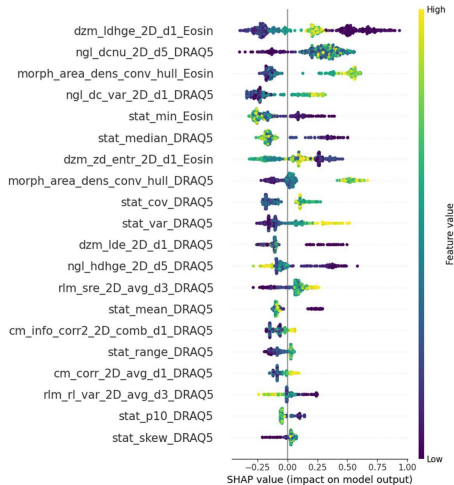
Figure 5

# Epithelial SHAP Beeswarm Plots

Low Grade VS Healthy



High Grade VS Healthy



High Grade VS Low Grade

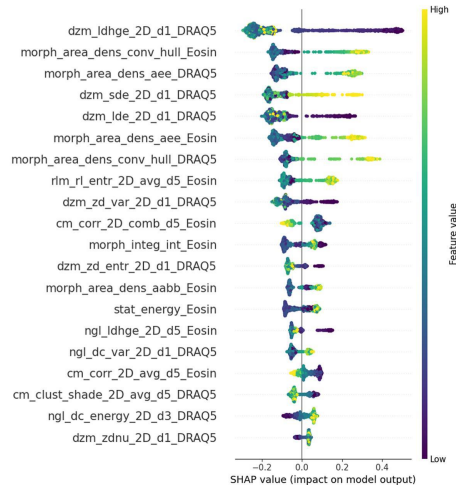


Figure 6

# Stromal SHAP Beeswarm Plots

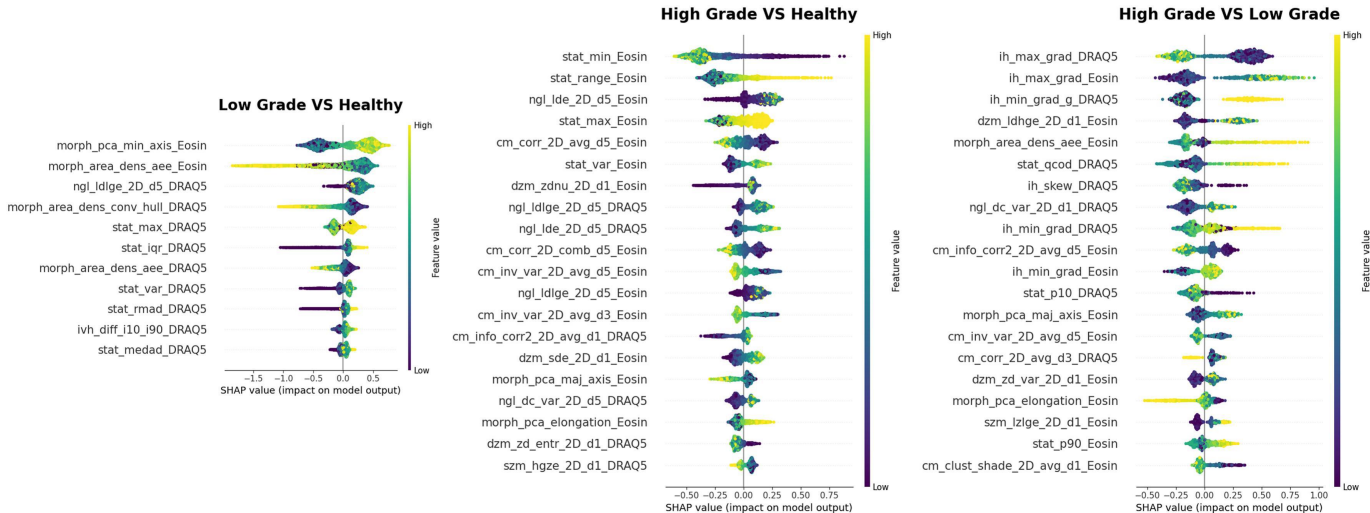


Figure 7