Invasive Cancerous Area Detection in Non-Muscle Invasive Bladder Cancer Whole Slide Images

Saul Fuster*, Farbod Khoraminia†, Umay Kiraz‡§, Neel Kanwal*, Vebjørn Kvikstad‡§,
Trygve Eftestøl*, Tahlita C.M. Zuiverloon†, Emiel A.M. Janssen‡§, Kjersti Engan*

*Dept. of Electrical Engineering and Computer Science, University of Stavanger, 4021 Stavanger, Norway
†Dept. of Urology, Erasmus MC Cancer Institute, University Medical Center, 3015 GD Rotterdam, The Netherlands
‡Dept. of Pathology, Stavanger University Hospital, 4011 Stavanger, Norway
§Dept. of Chemistry, Bioscience and Environmental Engineering, University of Stavanger, 4021 Stavanger, Norway

Abstract—**Bladder cancer patients' stratification into risk groups relies on grade, stage and clinical factors. For non-muscle invasive bladder cancer, T1 tumours that invade the subepithelial tissue are high-risk lesions with a high probability to progress into an aggressive muscle-invasive disease. Detecting invasive cancerous areas is the main factor for dictating the treatment strategy for the patient. However, defining invasion is often subject to intra/interobserver variability among pathologists, thus leading to over or undertreatment. Computer-aided diagnosis systems can help pathologists reduce overheads and erratic reproducibility. We propose a multi-scale model that detects invasive cancerous areas patterns across the whole slide image. The model extracts tiles of different tissue types at multiple magnification levels and processes them to predict invasive patterns based on local and regional information for accurate T1 staging. Our proposed method yields an F1 score of 71.9, in controlled settings 74.9, and without infiltration 90.0.**

Keywords—Digital Pathology, Bladder Cancer, Staging, Pattern Detection, Machine Learning

I. INTRODUCTION

Bladder cancer is one of the most commonly diagnosed cancer worldwide, with over 573,000 new cases and 213,000 deaths estimated in 2020 [1]. Approximately 90% of newly diagnosed cases are urothelial carcinomas that arise from the urothelial lining, from these 75% are non-muscle invasive bladder cancers (NMIBC) and 25% are muscle-invasive bladder cancer (MIBC) [2]. NMIBC can be divided into non-invasive papillary tumours Ta, T1 with invasion into the lamina propria and carcinoma in situ Tis, which is a superficial lesion. The three aforementioned groups account for 70, 20, and 10% of NMIBC, respectively [3]. Ta tumours typically are non-aggressive low-risk lesions that can be treated cautiously. However, T1 tumours are considered high-risk lesions with a higher chance of progressing into muscle-invasive tumours, resulting in higher mortality rates [4]. NMIBC accurate staging of the tumour is important to categorize patients into risk groups, and guides which treatment strategy the patient will receive, and, thus, patient outcome [5].

Invasive patterns can be problematic to identify correctly, see Fig 1. Morphological features indicating lamina propria

(LP) invasion include irregular nests, refraction artifacts, stromal reaction, among others [6]. External factors such as fragmented material and tangential sectioning are often present, providing a false sense of invasion. Also, benign structures like von Brunn nests might mimic invasion [7]. Complex evaluation of multi-variate scenarios leads to high inter-and intraobserver variability, resulting in down- or upstaging [8]. Variability in diagnosing T1 tumours has significant clinical consequences; thus, the necessity for higher reproducibility in assessing staging categories arises.
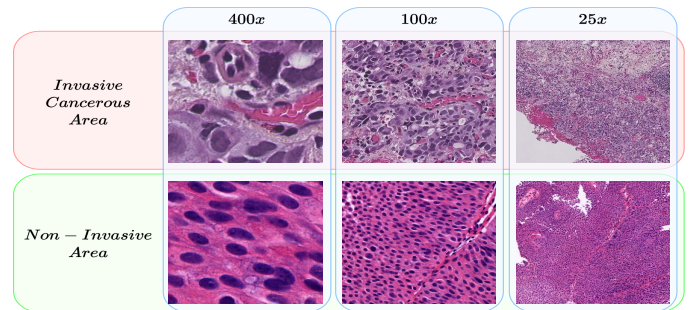

Fig. 1: Tiles of invasive (T1) and non-invasive (Ta) areas at different magnification levels. The lining contour of the basement membrane of the urothelial layer is interfered with the infiltrating tumour cells. Atypical clusters of tumour cells invade the lamina propria, resulting in irregularly shaped nests.

Computer-aided diagnosis (CAD) systems relying on machine learning methods for medical imaging analysis have proven to be an efficient manner to reduce subjectivity and accelerate the diagnostic procedure [9], [10]. Digital microscopy scanners generate high-resolution digital images from the scanned tissue sections, also named Whole Slide Images (WSI), in a fully automatic way. WSI are stored at multiple magnifications views, allowing the observer to adjust the zoom level, replicating physical microscopes. Pathologists use lower magnification levels to overview the regions of interest and analyse tissue-level morphology, while higher magnification is desirable for observing cell-level morphology. Convolutional neural networks (CNN) are the gold standard for feature extraction from histological images [11]. CNNs have been used for applications related to pattern detection of cellular features in breast, prostate, among other cancer types [12]–[15]. There are some examples in the literature where features are extracted at multiple magnification levels to incorporate
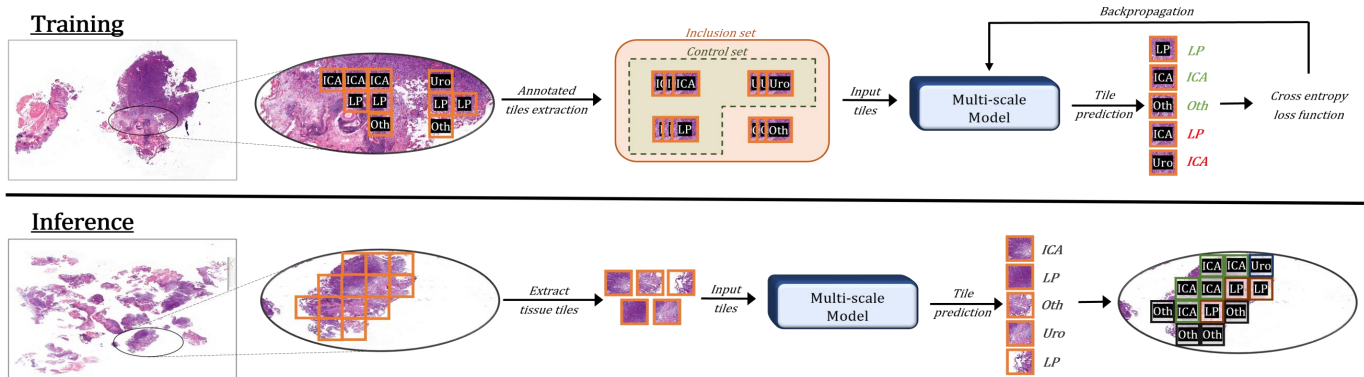
Fig. 2: During the training stage, exclusively tiles from annotated areas are extracted. Areas from all tissue types are extracted for analysis at the inference stage, resulting in a labelled colormap based on tile predictions.

contextual information of surrounding areas together with detailed information at the cellular level. Both Harmon et al. [16] and Li et al. [17] define tiles at different magnification levels independently from one another. However, Wetteland et al. [18] define one magnification to extract the tiles from and projects the centre of the pixel to other magnifications.

Bladder cancer staging should first and foremost distinguish NMIBC from MIBC. Image modalities such as CT, MRI or PET have proven to be effective, non-invasive methods to identify muscle invasion [19]–[21]. However, a detailed examination of NMIBC substaging requires histopathological analysis. Automatic applications that help to identify a possible invasion into the lamina propria include tissue segmentation to delineate the boundaries of urothelium, lamina propria and muscle tissue [18], [22]. Yin et al. [23] propose a method to classify tumour regions of H&E stained slides as stage Ta or T1 tumours using machine learning. The method was developed and evaluated in a control set containing patches of already segmented and verified Ta and T1 tumour tissue, disregarding the remaining tissue in WSIs. Analysing the entirety of the tissue in the slides is important to define the extent of invasion. Identifying focal points of invasion would provide important information to clinicians as an outcome predictor [24].

We propose an algorithm that finds invasive cancerous areas (ICA) associated with T1 stage across the WSI without the need of previous region of interest (ROI) or tumour segmentation. In ICA, the tumour has spread into the lamina propria that separates the urothelium layer from the bladder muscle beneath, whereas non-invasive areas are the regions of urothelium and lamina propria where no invasion is present. The proposed algorithm takes tissue tiles from all tissue types present in the WSI and discerns invasive from non-invasive patterns using both local and regional information extracted at different magnification levels.

II. METHODS

A. Datasets

51 WSI of NMIBC patients were collected from two independent cohorts from Erasmus MC (EMC), Rotterdam, The Netherlands and Stavanger University Hospital (SUH), Stavanger, Norway. The EMC dataset is a multi-center cohort containing 37 H&E stained WSI of high-risk NMIBC. WSI were scanned using a 3DHistech P1000 scanner at 800x magnification stored as MRXS files. The SUH dataset contains 14 WSI of NMIBC that were digitized using a Leica SCN400 scanner at 400x magnification stored in the SCN file format, 8 H&E and 6 HES stained. The combined dataset included 23 non-invasive and 28 invasive tumours. All slides were partially annotated and revised by pathologists, with the annotations serving as ground truth. Only representative regions were annotated from each WSI; thus, no slide was fully annotated. Annotated regions from EMC included mainly tissue types (urothelium, lamina propria, blood, muscle and artifacts), while SUH annotations contain primarily urothelium grading. In some cases, these tissue annotations might come with a subclass, such as grading of urothelium, presence of tumour infiltrating lymphocytes (TILs) in both urothelium and lamina propria, ICA, artifact type, among others. A dataset consisting of tiles was extracted from the annotated areas. Tiles were categorized into four main classes: invasive cancerous areas (*ICA*); ''non-invasive urothelium'' (*Uro*); ''lamina propria with no invasion'' (*LP*); and the rest of the tissue types, which we will refer as *Others*. We define non-invasive areas as the regions of *Uro* and *LP* where no invasion is present. *Others* includes blood, muscle and damaged tissue. Further details on the number of tiles extracted from our datasets can be found in Table I.

TABLE I: Number of WSI and tiles from EMC and SUH datasets.

	WSI	Tiles			
		ICA	Urothelium	Lamina Propria	Others
	Ta / T1		with / without TILs	with / without TILs	
EMC	16 / 21	14253	603 / 8859	2816 / 22154	34666
SUH	6 / 8	1333	0 / 60378	1858 / 1054	2726

From the extracted tiles, we defined two subsets; i) a *control set*, and ii) an *inclusion set*. A control set is sampled from the original dataset, containing ICA and selected non-invasive areas (*Uro*, *LP*). Extracting carefully selected tiles, we remove the surrounding tissue present in the slides while maintaining tiles related to the ROI. An inclusion set is defined including *Others* tissue types in addition to ICA, *Uro* and *LP* from the control set. The sets were split into training and test patients, ensuring no data leakage.
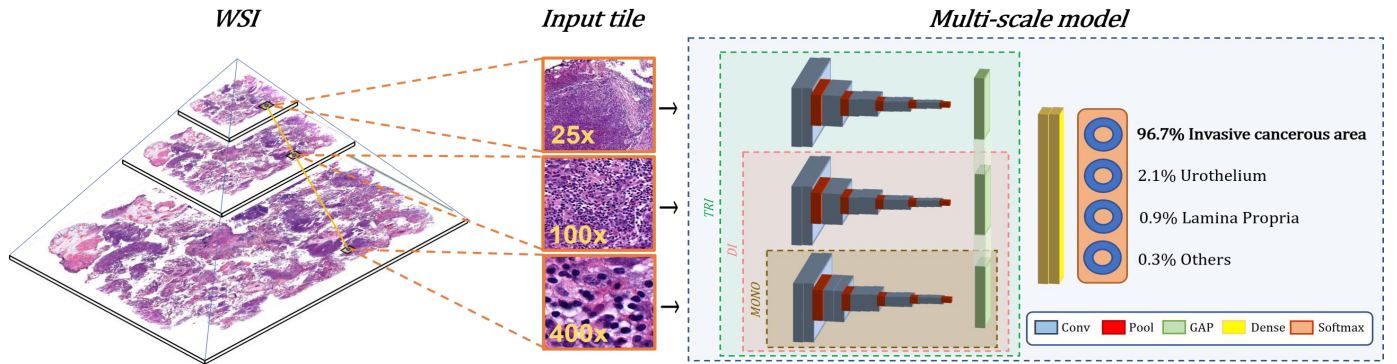
Fig. 3: Input tiles at different magnification levels are extracted from the WSI and fed into the feature extractors. Extracted features are concatenated and fed into the dense layers which will give the final prediction.

B. Multi-scale Model

We propose a CNN-based deep learning algorithm that can process all available tissue areas in a WSI and produces a segmentation map to detect ICA. Fig. 2 depicts the pipeline of the proposed method. The algorithm consists of a combination of feature extractors with input tiles at different magnification levels to obtain detailed as well as contextual information. Then, feature maps are downsampled using Global Average Pooling (GAP) to form a low-dimensional embedding. This information is later concatenated to embeddings from other magnification views and fed into the dense layers which will give the final prediction, as shown in Fig. 3. Our proposed algorithm is a TRI-scale model which analyses tiles at three magnification levels (400x, 100x, 25x), inspired by [18].

III. EXPERIMENTS

Multiple models using different combinations of magnification levels were trained and tested. The following combinations of magnification views were used: MONO (400x), DI (400x, 100x), and TRI-scale (400x, 100x, 25x). Tiles of size 256×256 were extracted from annotated regions at 400x magnification. Tiles at different magnifications were extracted so that the center of the tile remains the same for all views, as described in [25]. An example of this can be seen in Fig. 1. Extracted tiles at 400x are required to be covered by at least 70% of a tissue mask in order to be assigned such label. Background tiles were excluded using a thresholding technique. The number of annotated regions of ICA were significantly less than others. To address the data imbalance problem, we used undersampling of overrepresented classes. Tiles were converted to grayscale as datasets come from different laboratories, with different scanners and stain compositions, and the dataset was not large enough to support color variation. VGG16 was used as a feature extractor. To avoid overfitting, frozen weights were used due to the limited amount of training samples. Early stopping and learning rate decay were enabled. Cross entropy was defined as loss function. Models were trained for a maximum of 50 epochs. All models are implemented in Python 3.6 using Tensorflow machine learning library [26].

The experiments carried out are described as follows: i) for the first set of experiments, denoted E_{mag}, MONO, DI and TRI-scale models are evaluated to determine the relevance of contextual tissue morphology for discerning invasion; ii) for the second set of experiments, denoted E_{ds}, one of the original subsets was used, either control or inclusion, to ascertain the performance variation for adding non-ROI tissue; iii) for the third set of experiments, denoted E_{tils}, we excluded all tiles which included "tumour infiltrating lymphocytes" (TILs). TILs are small immune cells that can be found either in the urothelium or the lamina propria. TILs may lead to confusion for discerning them from invasion; iv) finally, for the fourth set of experiments, denoted E_{emc}, we used tiles from one of the hospitals exclusively. We chose EMC over SUH since the number of available data from EMC is far greater, especially regarding the number of ICA tiles. E_{emc} will show how impactful data balance among classes and domain shift are toward detecting invasive cancerous patterns. Also, the experiments were run in a binary and multi-class manner to assess whether grouping of non-invasive tissue improves ICA detection performance.

IV. RESULTS & DISCUSSION

Results of all experiments are collected in Table II for space efficiency. For experiment E_{mag}, TRI-scale models provide the best results for ICA detection, inferring that regional context derived from tissue morphology is beneficial in the task of discerning invasive patterns over non-invasive ones. Hence, models that incorporate lower magnification views provide better performance than those who focus on local patterns solely.

With respect to experiment E_{ds}, results on the control set demonstrate that discerning invasive from non-invasive patterns is possible, carefully selecting ROIs containing urothelium and lamina propria, with and without invasion. We observe that trained models with all sorts of tissue provide comparable results, using the inclusion set. It is, however, with models trained on the inclusion set that we can directly deploy the algorithm for WSI analysis. Models trained with the control set require a pre-processing step to sample carefully selected patches from the slides. Moreover, we further evaluate the discriminatory competence of our models for both sets representing a scatterplot of a two-component feature embedding visualization using t-SNE [27], see Fig. 4. Sample embeddings are clustered into separate locations of the feature

TABLE II: Results of all experiments are depicted with precision, recall and F1 score for the ICA class only. The experiment names explained in the text are indicated as the method (E_{mag}) and datasets (E_{ds}, E_{tils}, E_{emc}).

| Method (E_{mag}) | Control set (E_{ds}) | | | | | | Inclusion set (E_{ds}) | | | | | |
| | ICA vs (Uro+LP) | | | ICA vs Uro vs LP | | | ICA vs (Uro+LP+Others) | | | ICA vs Uro vs LP vs Others | | |
	Precision	Recall	F1 Score	Precision	Recall	F1 Score	Precision	Recall	F1 Score	Precision	Recall	F1 Score
MONO (400x)	48.5	78.4	60.0	53.8	70.1	60.8	41.0	75.6	53.2	44.3	77.5	55.3
* E_{tils}	69.3	81.2	74.8	67.6	79.9	73.2	60.9	74.3	67.0	52.9	80.3	63.8
* E_{emc}	30.4	90.4	45.5	33.9	83.6	48.3	24.9	91.1	39.1	30.5	80.8	44.3
DI (400x,100x)	65.4	76.7	70.6	64.3	81.1	71.7	64.6	74.3	69.1	59.6	79.1	68.0
* E_{tils}	91.7	79.1	84.9	87.6	87.2	**87.4**	85.9	78.2	81.9	75.3	86.0	80.3
* E_{emc}	41.5	91.3	57.0	42.7	98.2	59.6	32.3	90.6	47.6	40.3	97.5	57.0
TRI (400x,100x,25x)	66.4	81.3	73.1	70.5	79.8	74.9	65.9	79.2	71.9	63.4	80.3	70.8
* E_{tils}	93.8	86.4	**90.0**	85.5	89.1	87.3	89.0	81.2	**85.0**	77.6	88.9	**82.9**
* E_{emc}	48.1	94.1	63.6	49.2	93.3	64.4	53.1	79.5	63.7	43.1	93.2	58.9

space based on the class label; hence there is a clear separation of the classes. With regard to the number of classes, binary classification has proven to be more efficient than multi-class for ICA predictions in the inclusion set, but not for the control set. The best performing binary model of the inclusion set, TRI_{incl} binary model, was used to produce probability heatmaps over WSIs from the test set, as shown in Fig. 5. Tile predictions were overlaid on the WSI, ranging from blue to red for ICA likelihood. A pathologist could later use the generated heatmaps as a tool for guidance in detecting potential ICA.
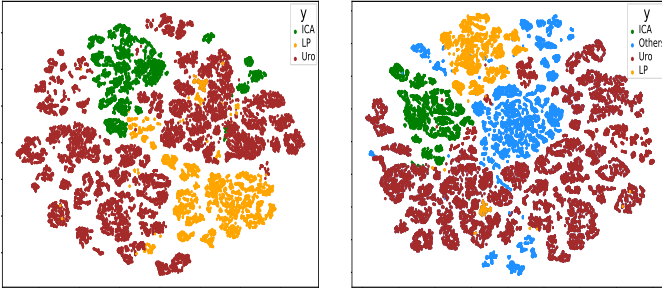

Fig. 4: tSNE scatterplot of test samples embeddings on the control and inclusion sets, respectively. Embeddings were extracted as the output of the feature extractor and reduced to two components.
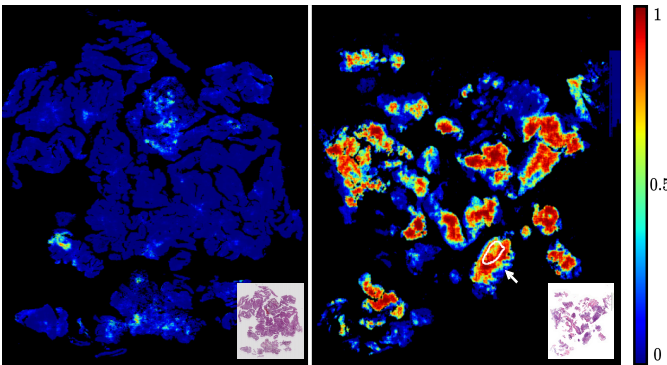

Fig. 5: Heatmaps showing the probability of a patch to belong to the ICA class, based on the TRI_{incl} binary model. Two test WSI samples with stage Ta (left) and T1 (right) are shown. An annotated ICA region is highlighted by a white border in the T1 WSI. Inclusion set models support the analysis of the entire WSI, not limited to pre-defined ROIs. A thumbnail of the original WSI is displayed at the bottom right corner.

Regarding experiment E_{tils}, we noticed that the presence of TILs dampers the predictive performance, as models struggle to differentiate such areas from ICA. Tiles where TILs are present are often missclassified for ICA. Excluding areas with abundant TILs results in a significant F1 score improvement. The positive effect of TILs exclusion is relevant for both sets and all combinations of magnification views, although it is higher when using TRI-scale models. For the TRI_{incl} binary model, the F1 score difference reaches over 13%. Based on the results obtained, a future pipeline should include a posterior classifier to discriminate ICA false positives with infiltration from genuine ICA.

Finally, for experiment E_{emc}, despite reducing scanning and staining variability, we lit upon poor performance in comparison to the other experiments. This can be due to the lack of a substantial number of urothelium tiles, which are present mostly in the excluded cohort, see Table I. Even though most of the ICA tiles remained in the dataset, discarding the majority of the *Uro* class proved unfavourable. Precision results infer a bias towards ICA. Distinguishing ICA patterns for non-invasive patterns relies mainly on telling Ta from T1 regions apart. As the tumour grows from the urothelial layer into the lamina propria, it is essential to have a significant representation of *Uro* over *LP*.

V. CONCLUSION & FUTURE WORK

A multi-scale model was developed for detecting invasive patterns in ICA across the WSI. A control set shows that CNNs distinguish invasive areas from other tumour regions, while including non-diagnostically relevant regions barely decreases the model capability of discerning invasion. Multi-scale models give best performance indicating that using contextual information combined with local patterns is highly beneficial. Predicted heatmaps for the invaded areas can be represented for user-friendly interpretation. Automatic ICA detection can also can be a first step for automated staging of T1 tumours. This can be useful for treatment planning, both alone and in combination with automated grading. The results are promising but due to the lack of large scale manual annotated regions, it is conducted over a small dataset. Future developments should be conducted over larger cohorts and with implemented color normalization schemes to alleviate scanning and staining variability. Likewise, a post-processing step for TILs exclusion should be adopted to alleviate infiltration misclassifications.

REFERENCES

[1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 71, pp. 209 – 249, 2021.

[2] M. Burger, J. W. Catto, G. Dalbagni, H. B. Grossman, H. W. Herr, P. I. Karakiewicz, W. Kassouf, L. A. Kiemeney, C. la Vecchia, S. F. Shariat, and Y. Lotan, "Epidemiology and risk factors of urothelial bladder cancer." *European urology*, vol. 63 2, pp. 234–41, 2013.

[3] Z. Kırkalı, T. Y. Chan, M. Manoharan, F. Algaba, C. Busch, L. Cheng, L. A. Kiemeney, M. Kriegmair, R. Montironi, W. M. Murphy, I. A. Sesterhenn, M. Tachibana, and J. Weider, "Bladder cancer: epidemiology, staging and grading, and diagnosis." *Urology*, vol. 66 6 Suppl 1, pp. 4–34, 2005.

[4] D. S. Kaufman, W. U. Shipley, and A. S. Feldman, "Bladder cancer," *The Lancet*, vol. 374, pp. 239–249, 2009.

[5] H. E. Schwaibold, S. Sivalingam, F. May, and R. Hartung, "The value of a second transurethral resection for t1 bladder cancer," *BJU International*, vol. 97, 2006.

[6] M. J. Magers, A. Lopez-Beltran, R. Montironi, S. R. Williamson, H. Z. Kaimakliotis, and L. Cheng, "Staging of bladder cancer," *Histopathology*, vol. 74, pp. 112 – 134, 2019.

[7] R. M. Cox and J. I. Epstein, "Large nested variant of urothelial carcinoma: 23 cases mimicking von brunn nests and inverted growth pattern of noninvasive papillary urothelial carcinoma," *The American Journal of Surgical Pathology*, vol. 35, p. 1337–1342, 2011.

[8] I. Tosoni, U. Wagner, G. Sauter, M. Egloff, H. Knönagel, G. Alund, F. Bannwart, M. J. Mihatsch, T. C. Gasser, and R. Maurer, "Clinical significance of interobserver differences in the staging and grading of superficial bladder cancer," *BJU International*, vol. 85, 2000.

[9] A. Echle, N. Rindtorff, T. J. Brinker, T. Luedde, A. T. Pearson, and J. N. Kather, "Deep learning in cancer pathology: a new generation of clinical biomarkers." *British journal of cancer*, 2020.

[10] S. Borhani, R. Borhani, and A. A. Kajdacsy-Balla, "Artificial intelligence: A promising frontier in bladder cancer diagnosis and outcome prediction." *Critical reviews in oncology/hematology*, p. 103601, 2022.

[11] M. Cui and D. Zhang, "Artificial intelligence and computational pathology," *Laboratory Investigation; a Journal of Technical Methods and Pathology*, pp. 1 – 11, 2021.

[12] J. Silva-Rodríguez, A. Colomer, M. Á. Sales, R. Molina, and V. Naranjo, "Going deeper through the gleason scoring scale: An automatic end-to-end system for histology prostate grading and cribriform pattern detection," *Computer methods and programs in biomedicine*, vol. 195, p. 105637, 2020.

[13] B. Gecer, S. Aksoy, E. Mercan, L. G. Shapiro, D. L. Weaver, and J. G. Elmore, "Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks," *Pattern recognition*, vol. 84, pp. 345–356, 2018.

[14] A.-C. Woerl, M. Eckstein, J. Geiger, D.-C. Wagner, T. Daher, P. Stenzel, A. Fernandez, A. Hartmann, M. Wand, W. Roth, and S. Foersch, "Deep learning predicts molecular subtype of muscle-invasive bladder cancer from conventional histopathological slides." *European urology*, 2020.

[15] G. García, A. Esteve, A. Colomer, D. Ramos, and V. Naranjo, "A novel self-learning framework for bladder cancer grading using histopathological images," *Computers in biology and medicine*, vol. 138, p. 104932, 2021.

[16] S. A. Harmon, T. Sanford, G. T. Brown, C. Yang, S. Mehralivand, J. M. Jacob, V. A. Valera, J. H. Shih, P. Agarwal, P. L. Choyke, and B. Turkbey, "Multiresolution application of artificial intelligence in digital pathology for prediction of positive lymph nodes from primary tumors in bladder cancer." *JCO clinical cancer informatics*, vol. 4, pp. 367–382, 2020.

[17] B. Li, Y. Li, and K. W. Eliceiri, "Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning," *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 14 313–14 323, 2021.

[18] R. Wetteland, K. Engan, T. Eftestøl, V. Kvikstad, and E. A. M. Janssen, "A multiscale approach for whole-slide image segmentation of five tissue classes in urothelial carcinoma slides," *Technology in Cancer Research & Treatment*, vol. 19, 2020.

[19] S. J. Galgano, K. K. Porter, C. M. Burgan, and S. Rais-Bahrami, "The role of imaging in bladder cancer diagnosis and staging," *Diagnostics*, vol. 10, 2020.

[20] S. S. Garapati, L. M. Hadjiiski, K. H. Cha, H.-P. Chan, E. M. Caoili, R. H. Cohan, A. Z. Weizer, A. Alva, C. Paramagul, J. Wei, and C. Zhou, "Urinary bladder cancer staging in ct urography using machine learning," *Medical Physics*, vol. 44, p. 5814–5823, 2017.

[21] X. Xu, X. Zhang, Q. Tian, H. Wang, L.-B. Cui, S. Li, X. Tang, B. Li, J. Dolz, I. B. Ayed, Z. Liang, J. Yuan, P. Du, H. Lu, and Y. Liu, "Quantitative identification of nonmuscle-invasive and muscle-invasive bladder carcinomas: A multiparametric mri radiomics analysis," *Journal of Magnetic Resonance Imaging*, vol. 49, 2019.

[22] M. K. Khan, T. E. Tavolara, V. C. Arole, A. V. Parwani, C. Lee, and M. N. Gürcan, "Automated t1 bladder risk stratification based on depth of lamina propria invasion from h and e tissue biopsies: a deep learning approach," in *Medical Imaging*, 2018.

[23] P. nien Yin, K. Kc, S. Wei, Q. Yu, R. Li, A. R. Haake, H. Miyamoto, and F. Cui, "Histopathological distinction of non-invasive and invasive bladder cancers using machine learning approaches," *BMC Medical Informatics and Decision Making*, vol. 20, 2020.

[24] C. Patriarca, R. Hurle, M. Moschini, M. Freschi, P. Colombo, M. Colecchia, L. Ferrari, G. Guazzoni, A. Conti, G. N. Conti, R. Lucianò, T. Magnani, and R. Colombo, "Usefulness of pt1 substaging in papillary urothelial bladder carcinoma," *Diagnostic Pathology*, vol. 11, 2016.

[25] R. Wetteland, K. Engan, and T. Eftesol, "Parameterized extraction of tiles in multilevel gigapixel images," *2021 12th International Symposium on Image and Signal Processing and Analysis (ISPA)*, pp. 78–83, 2021.

[26] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. A. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zhang, "Tensorflow: A system for large-scale machine learning," *ArXiv*, vol. abs/1605.08695, 2016.

[27] L. van der Maaten and G. E. Hinton, "Visualizing data using t-sne," *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605, 2008.