Article

Deep Learning-Based Screening of Urothelial Carcinoma in Whole Slide Images of Liquid-Based Cytology Urine Specimens

Masayuki Tsuneki [1],* , Makoto Abe [2] and Fahdi Kanavati [1]

[1]   Medmain Research, Medmain Inc., Fukuoka 810-0042, Japan
[2]   Department of Pathology, Tochigi Cancer Center, 4-9-13 Yohnan, Utsunomiya 320-0834, Tochigi, Japan
*   Correspondence: tsuneki@medmain.com; Tel.: +81-92-707-1977

Simple Summary: In this study, we aimed to investigate the use of deep learning for classifying whole-slide images of urine liquid-based cytology specimens into neoplastic and non-neoplastic (negative). To do so, we used a total of 786 whole-slide images to train a model using four different approaches, and we evaluated it on 750 whole-slide images. The best model achieved good classification performance, demonstrating the promising potential of use of such models for aiding the screening process for urothelial carcinoma in routine clinical practices.

Abstract: Urinary cytology is a useful, essential diagnostic method in routine urological clinical practice. Liquid-based cytology (LBC) for urothelial carcinoma screening is commonly used in the routine clinical cytodiagnosis because of its high cell collection rate. Since conventional screening processes by cytoscreeners and cytopathologists using microscopes is limited in terms of human resources, it is important to integrate new deep learning methods that can automatically and rapidly diagnose a large amount of specimens without delay. The goal of this study was to investigate the use of deep learning models for the classification of urine LBC whole-slide images (WSIs) into neoplastic and non-neoplastic (negative). We trained deep learning models using 786 WSIs by transfer learning, fully supervised, and weakly supervised learning approaches. We evaluated the trained models on two test sets (equal and clinical balance) with a combined total of 750 WSIs, achieving ROC-AUCs for WSI diagnosis in the range of 0.984-0.990 by the best model, demonstrating the promising potential use of our model for aiding urine cytodiagnostic processes.

Keywords: urothelial carcinoma; urine; liquid-based cytology; deep learning; cancer screening; whole slide image

1. Introduction

For routine clinical practices, clinicians obtain urinary tract cytology specimens for the screening of urothelial carcinoma. Urine specimens play a critical role in the clinical evaluation of patients who have clinical signs and symptoms (e.g., haematuria and painful urination) suggestive of pathological changes within the urinary tract. Urothelial carcinoma is the most common malignant neoplasm detected by urine cytology. The most common site of origin of urothelial carcinoma is bladder. According to the Global Cancer Statistics 2020 [1], bladder cancer is the tenth most commonly diagnosed cancer with 573,278 of new cases and 212,536 of new deaths worldwide in 2020. Most of the bladder cancers are urothelial in origin (approximately 90% of bladder cancers) and primary adenocarcinoma of the bladder is rare [2,3]. Bladder cancer often presents insidiously. Haematuria is the most common presentation of bladder cancer, which is typically intermittent, frank, painless and at times present throughout micturition [3]. Delayed diagnosis of urothelial carcinoma is associated with high grade muscle invasion which has the potential to progress rapidly and cancer metastasis [3]. Of course, cystoscopy with a biopsy is the gold standard for diagnosis of urothelial carcinoma in clinical practice; however, it is aggressive and relatively inconvenient as a follow-up monitoring approach [4]. It has been reported that 48.6% of

biopsy proven low-grade urothelial carcinomas had a urine cytodiagnosis of atypical or neoplastic suspicious, which could conclude that existing urine cytology screening and surveillance systems are accurate in diagnosing urothelial carcinoma [5]. Therefore, cytological urothelial carcinoma screening in urine specimens plays a key role in early stage cancer detection and treatment in routine clinical practices [6,7].

Liquid-based cytology (LBC) was developed as an alternative to conventional smear cytology in the 1990s [8]. LBC has several advantages in preparation and diagnostic process compared with conventional smear. The LBC technique preserves the cells of interest in a liquid medium and removes most of the debris, blood, and exudate either by filtering or density gradient centrifugation. LBC provides automated and standardized processing techniques that produce a uniformly distributed and cell-enriched slide. Moreover, residual specimens can be used for additional investigations (e.g., immunochtochemistry). ThinPrep (Hologic, Inc., United States) and SurePath (Becton Dickinson, Inc., United States) for LBC specimen preparation have been approved by the US Food and Drug Administration (FDA). Compared to the conventional smear cytology, LBC has lower background elements, provides better cell preservation, and has a higher satisfaction rate [9]. As for the sensitivity, it has been reported that LBC achieved at 0.58 (CI: 0.51-0.65) and conventional smear achieved at 0.38 [4,8]. It is understandable that the efficiency of diagnosis employing the LBC is high because of the cell collection rate. It was shown that the accuracy of diagnoses made employing the LBC method can be increased by understanding the characteristics of the cell morphology in suspicious cases (e.g., high-grade urothelial carcinoma and low-grade urothelial carcinoma) [10]. LBC specimens performed significantly better in urinary cytology when evaluating malignant categories especially high-grade urothelial carcinoma (HGUC), which facilitate a more accurate diagnosis than conventional preparations [11]. Moreover, from the standpoint of rationality, preparation and screening times were 2.25 and 1.33-2.00 times greater when using LBC (ThinPrep) compared with cytocentrifugation (conventional smear cytology) [12]. Therefore, computational screening (cytodiagnostic) aids for urine LBC specimens would be a great benefit for urothelial carcinoma screening as medial image analysis.

Whole-slide images (WSIs) are digitisations of the conventional glass slides obtained via specialised scanning devices (WSI scanners), and they are considered to be comparable to microscopy for primary diagnosis [13]. It has been reported that evaluation of WSI is generally equivalent to using conventional glass slides under microscopy [14]. The use of WSI has to some degree met the goals of saving pathologists working time and providing high quality pathological images with convenient access and easily navigable viewing online based software which saves resources and costs by eliminating slide glass shipping expenses [14]. The advent of WSIs led to the application of medical image analysis techniques, machine learning, and deep learning techniques for aiding pathologists in inspecting WSIs [15]. Importantly, a routine scanning of LBC slides in a single layer of WSIs would be suitable for further high throughput analysis (e.g., automated image based cytological screening and medical image analysis) [16]. Indeed, deep learning approaches and its clinical application to classify cytopathological changes (e.g., neoplastic transformation) were reported in the recent years [17–26]. In this study, we trained deep learning models based on convolutional neural networks (CNN) using a training dataset of 786 urine LBC (ThinPrep) WSIs. We evaluate the model on two test sets (equal balance and clinical balance test sets) with a combined total of 750 WSIsm achieving ROC-AUCs for WSI neoplastic classification in the range of 0.984-0.990.

2. Materials and Methods

2.1. Clinical Cases and Cytopathological Records

In this retrospective study, a total of 1,556 LBC ThinPrep Pap test (Hologic, Inc.) conventionally prepared cytopathological slide glass specimens of human urine cytology were collected from a private clinical laboratory in Japan after routine cytopathological review of those glass slides by cytoscreeners and pathologists. The LBC specimens were selected

randomly to reflect a real clinical settings as much as possible. We have also collected LBC specimens so as to compile test sets with an equal balance and a clinical balance of negative and neoplastic. The cytoscreeners and pathologists excluded inadequate LBC specimens (n=21) prior to the start of the experiments. All WSIs were scanned at a magnification of ×20 using the same Leica Aperio AT2 Digital Whole Slide Scanner (Leica Biosystems, Tokyo, Japan) and were saved in the SVS file format with JPEG2000 compression. Each WSI diagnosis was observed by at least two cytoscreeners or pathologists with the final checking and verification performed by a senior cytoscreener or pathologist. In this study, we have classified urine LBC WSIs into two classes: one is negative and the other is neoplastic (Table 1). Negative WSIs were diagnosed as Class I or Class II and neoplastic WSIs were diagnosed as Class III, Class IV, or Class V in routine clinical cytodiagnosis (Table 1). The cytoscreeners and pathologists had to agree whether the output class was negative or neoplastic on each urine LBC WSI.

2.2. Dataset

The private clinical laboratory in Japan that provided urine LBC specimen glass slides in the present study was anonymized due to the confidentiality agreement. Table 1 breaks down the distribution of the dataset into training, validation, and test (equal balance and clinical balance) sets. The split was carried out randomly taking into account the proportion of each label in the dataset. The test sets were composed of WSIs of equal balance and clinical balance urine LBC specimens. The equal balance test set consisted of 50% negative and 50% neoplastic urine LBC cases (Table 1). The clinical balance test set consisted of a ratio of 10 (negative) to 1 (neoplastic) urine LBC cases based on a real clinical setting which was reported by the Japanese Society of Clinical Cytology as the statistics on cytodiagnosis in 2016 to 2021 (https://jscc.or.jp/). The urine LBC cases in the present study were collected based on the cytodiagnoses provided by the private clinical laboratory. Two independent cytoscreeners or pathologists viewed all the WSIs with the final verification by a senior cytoscreener or pathologist and the ones they had a disagreement on were removed from the dataset. We have confirmed that cytoscreeners and pathologists were able to classify (Table 1) from the visual inspection of the LBC ThinPrep Pap test (Hologic, Inc.) stained WSIs alone.

2.3. Annotation

Cytoscreeners and pathologists who perform routine cytopathological screening and cytodiagnoses in general hospitals and clinical laboratories in Japan manually annotated 72 neoplastic WSIs from the training (62 WSIs) and validation (10 WSIs) sets (Table 1). Coarse manually drawing polygonal annotations were obtained by free-hand drawing in-house online tool developed by customising the open-source OpenSeadragon tool at https://openseadragon.github.io/ which is a web-based viewer for zoomable images. On average, the cytoscreeners and pathologists manually annotated 180 cells (or cellular clusters) per WSI. Annotated neoplastic WSIs consisted of Class III, Class IV, and Class V cytodiagnostic classes except for Class I and Class II (Table 1). We set three annotation labels for neoplastic urothelial epithelial cells: atypical cell, low-grade urothelial carcinoma (LGUC) cell, and high-grade urothelial carcinoma (HGUC) cell (Table 1 and Figure 1). For example, on the Class III (Figure 1A, B), Class IV (Figure 1C, D), and Class V (Figure 1E, F) WSIs, cytoscreeners and pathologists performed annotations around the atypical cells (Figure 1A, B), LGUC cells (Figure 1C, D), and HGUC cells (Figure 1E, F) based on the representative neoplastic urothelial epithelial cell morphology (e.g., hyperchromatism, irregular chromatin distribution, abnormalities of nuclear shape, increased nuclear/cytoplasmic ratio, irregular nuclear distribution, nuclear enlargement, abnormal cytoplasm, prominent nucleolus, cellular and nuclear polymorphism). If the WSIs were classified as Class V, for example, it would be possible to have atypical cell, LGUC cell, and HGUC cell annotations in a WSI. In contrast, the cytoscreeners and pathologists did not annotate areas where it was difficult to cytologically determine that the cells were neoplastic. The negative subset of the

training and validation sets (Table 1) was not annotated and the entire cell spreading areas within the WSIs were used. The average annotation time per WSI was about 90 minutes. Annotations performed by the cytoscreeners and pathologists were modified (if necessary) and verified by a senior cytoscreener.

2.4. Deep Learning Models

We performed training using transfer learning with fine-tuning using two different weight initialisations: ImageNet (IN) and pre-training on a uterine cervix (UC) neoplastic (x10, 1024) dataset from a previous study. We used two different approaches for training during fine-tuning: fully supervised (FS) and weakly supervised (WS) learning. We used a modified version of EfficientNetB1 (ENB1) [27] with a tile size of 1024×1024 px. This resulted in a total of four models, all trained at magnification x10 and tile size 1024x1024px: ENB1-UC-FS+WS, ENB1-UC-WS, ENB1-IN-FS+WS, and ENB1-IN-WS. We performed the fine-tuning of the models using the partial fine-tuning approach [28], which consists in only fine-tuning the affine parameters of batch-normalization layers and the final classification layer (Figure 2). starting with pre-trained weights on ImageNet. Figure 2 shows an overview of the training method and trained deep learning models. The training methodology that we used in the present study was exactly the same as reported in our previous studies [29].

We performed slide tiling by extracting square tiles from tissue regions of the WSIs. We started by detecting the tissue regions in order to eliminate most of the white background. This was conducted by performing thresholding on a grayscale version of the WSIs using Otsu's method [30]. For the CNN, we have used the EfficientNetB1 architecture [27] with a modified input size of 1024 × 1024 px to allow a larger view; this is based on cytologists' input that they usually need to view the neighbouring cells around a given cell in order to diagnose more accurately. We used the partial fine-tuning approach [28] for the tuning the CNN component.

For training and inference, we then proceeded by extracting 1024 × 1024 px tiles from the tissue regions. We performed the extraction in real-time using the OpenSlide library [31]. To perform inference on a WSI, we used a sliding window approach with a fixed-size stride of 512 × 512 px (half the tile size). This results in a grid-like output of predictions on all areas that contained cells, which then allowed us to visualise the prediction as a heatmap of probabilities that we can directly superimpose on top of the WSI. Each tile had a probability of being neoplastic; to obtain a single probability that is representative of the WSI, we computed the maximum probability from all the tiles.

During fully supervised learning, we maintained an equal balance of positively and negatively labelled tiles in the training batch. To do so, for the positive tiles, we extracted them randomly from the annotated regions (annotation label: atypical cell, LGUC cell, and HGUC cell) of neoplastic WSIs, such that within the 1024 × 1024 px, at least one annotated cell was visible anywhere inside the tile. For the negative tiles, we extracted them randomly anywhere from the tissue regions of negative WSIs (Table 1). We then interleaved the positive and negative tiles to construct an equally balanced batch that was then fed as input to the CNN. In addition, to reduce the number of false positives, given the large size of the WSIs, we performed a hard mining of tiles, whereby at the end of each epoch, we performed full sliding window inference on all the negative WSIs in order to adjust the random sampling probability such that false positively predicted tiles of negative were more likely to be sampled.

During weakly supervised learning, to maintain the balance on the WSI, we oversampled from WSIs to ensure that the model trained on tiles from all WSIs in each epoch. We then switched to hard mining tiles. To perform hard mining, we alternated between training and inference. During inference, the CNN was applied in a sliding window fashion on all the tissue regions in the WSI, and we then selected the k tiles with the highest probability for being positive. This step effectively selects tiles that are most likely to be false positives when the WSI is negative. The selected tiles were placed in a training subset, and once that

subset contained N tiles, training was initiated. We used $k = 8$, $N = 256$, and a batch size of 32.

During training, we performed real-time augmentation of the extracted tiles using variations of brightness, saturation, and contrast. We trained the model using the Adam optimisation algorithm [32], with the binary cross entropy loss, $beta_1 = 0.9$, $beta_2 = 0.999$, and a learning rate of 0.001. We applied a learning rate decay of 0.95 every 2 epochs. We used early stopping by tracking the performance of the model on a validation set, and training was stopped automatically when there was no further improvement on the validation loss for 10 epochs. The model with the lowest validation loss was chosen as the final model.

2.5. Software and Statistical Analysis

The deep learning models were implemented and trained using the open-source TensorFlow library [33]. AUCs were calculated in python using the scikit-learn package [34] and plotted using matplotlib [35]. The 95% CIs of the AUCs were estimated using the bootstrap method [36] with 1000 iterations. The ROC curve was computed by varying the probability threshold from 0.0 to 1.0 and computing both the TPR and FPR at the given threshold.

3. Results

3.1. Insufficient AUC Performance of whole slide image (WSI) neoplastic evaluation on urine LBC WSIs using existing series of LBC cytopathological model

Prior to training urine LBC neoplastic screening models, we applied existing LBC uterine cervix neoplastic screening model [20] and histopathological classification models and evaluated their AUC performances on urine LBC test sets (Table 1). This is summarised in Table 3.

3.2. High ROC-AUC performance of urine LBC WSI evaluation of neoplastic urothelial epithelial cell screening

We trained four deep learning models ([ENB1-UC-FS+WS], [ENB1-UC-WS], [ENB1-IN-FS+WS], and [ENB1-IN-WS]) using transfer learning (TL) with fine-tuning [29]with fully supervised (FS) learning [20,37], and weakly supervised (WS) learning [38] approaches as described elsewhere. These models are all based on the EfficientNetB1 convolutional neural network (CNN) architecture. To compare transfer learning models' performance ([ENB1-UC-FS+WS] and [ENB1-UC-WS]), we trained two models using EfficientNetB1 architecture at a same magnification of ×10 and tile size (1024×1024 px). To train deep learning models, we used a total of 62 neoplastic (with annotation) and 724 negative (without annotation) training set WSIs and 10 neoplastic (with annotation) and 10 negative (without annotation) validation set WSIs (Table 1). This resulted in four different models: (1) ENB1-UC-FS+WS (x10, 1024), (2) ENB1-UC-WS (x10, 1024), (3) ENB1-IN-FS+WS (x10, 1024), and (4) ENB1-IN-WS (x10, 1024). We evaluated these four different trained deep learning models on equal balance and clinical balance test sets (Table 1). For each test set (equal and clinical balance), we computed the ROC-AUC, log-loss, accuracy, sensitivity, and specificity and summarized in Table 4 and Figure 3 and 4. Overall, four different trained deep learning models achieved equivalent ROC-AUC, log-loss, accuracy, sensitivity, and specificity at whole-slide level (WSI evaluation) in both equal and clinical balance test sets (Table 4, Figure 3). However, heatmap image appearances were different among four trained deep learning models (Figure 4). The localization patterns of predicted tiles were approximately same among four trained deep learning models (Figure 4). Looking at heatmap images of the same urine LBC WSIs (WSI-1 and WSI-2) (Figure 4) that were correctly predicted (true-positive) as neoplastic WSI using four different trained models, all models could predict tiles with neoplastic urothelial epithelial cells (Figure 4A-D, M-P) satisfactorily (Figure 4E-L, Q-X). However, probabilities in each neoplastic predicted tiles were totally different among four trained models (Figure 4). Among the four trained model, ENB1-UC-FS+WS exhibited the best tile prediction overall based on inspection of the heatmap images (Figure 4E, F,

Q, R). Therefore, our results show that the ENB1-UC-FS+WS model is the best model for urine LBC neoplastic urothelial epithelial cell screening (Table 4 and Figure 4). Figures show representative WSIs of true positive (Figure 5), true negative (Figure 6), false positive (Figure 7), and false negative (Figure 8) from using the model [ENB1-UC-FS+WS].

3.3. True Positive Prediction

The model ENB1-UC-FS+WS satisfactorily predicted neoplastic urothelial epithelial cells in urine LBC WSIs (Figure 5). Cytopathologically, Figure 5A exhibited atypical urothelial epithelial cells (Figure 5B) and was diagnosed as Class III. Figure 5E showed low grade urothelial carcinoma (LGUC) cells (Figure 5F) and was diagnosed as Class IV. Figure 5I showed high grade urothelial carcinoma (HGUC) cells (Figure 5J) and was diagnosed as Class V. These three WSIs should be classified as neoplastic in this study. The heatmap images show true positive predictions of atypical utorhelial cells (Figure 5C, D), LGUC cells (Figure 5G, H), and HGUC cells (Figure 5K, L) which were confirmed by a cytoscreener and a cytopathologist by viewing original WSIs and predicted heatmap images. In contrast, in low probability tiles (light blue and blue background) (Figure 5), two independent cytoscreeners confirmed there were no neoplastic urothelial epithelial cells.

3.4. True Negative Prediction

The model ENB1-UC-FS+WS satisfactorily predicted negative cases (cytopathologically as Class I and Class II) in urine LBC WSIs (Figure 6). The heatmap images show true negative predictions of neoplastic urothelial epithelial cells (Figure 6C, F). In zero probability tiles (blue background color) (Figure 6C, F), there are no neoplastic urothelial epithelial cells in pyuria (cytodiagnosed as Class I) (Figure 6A) which consisted of infective fluid with small number of non-neoplastic epithelial cells (Figure 6B) and urothelial epithelial cells with slight nuclear enlargement (Figure 6D, E) (cytodiagnosed as Class II).

3.5. False Positive Prediction

A cytopathologically diagnosed negative (Class I) case (Figure 7A) consisted of metaplastic squamous epithelial cells and non-neoplastic urothelial epithelial cells (Figure 7B) was false positively predicted for neoplastic urothelial epithelial cells by our model [ENB1-UC-FS+WS]. The heatmap image (Figure 7C) shows false positive predictions of neoplastic urothelial epithelial cells with high probability tiles (Figure 7D). Cytopathologically, there are non-neoplastic urothelial epithelial cells with a slightly increased nuclear cytoplasmic (N/C) ratio and metaplastic squamous epithelial cells (Figure 7B), which could be a major cause of false positive.

3.6. False Negative Prediction

According to the cytodiagnosis report and additional cytoscreener and cytopathologist's review, in this urine LBC WSI (Figure 8A), there were ceuular culsters of atypical (neoplastic) urothelial epithelial cells (Figure 8B, C) with high nuclear cytoplasmic ratio, indicating this WSI (Figure 8A) should be classified as neoplastic (Class III). However, our model [ENB1-UC-FS+WS] did not predict or very low level predicted neoplastic urothelial epithelial cells (Figure 8D-F). It would be speculated that neoplastic urothelial epithelial cellular clustering could be a possible cause of false negative due to the overlapping morphology.

4. Discussion

In this study, we trained deep learning models for the classification of neoplastic urothelial epithelial cells in WSIs of urine LBC specimens. The best model (ENB1-UC-FS+WS) achieved overall a good performance with ROC-AUCs of 0.984 (CI: 0.969 - 0.995) on equal balance and 0.990 (CI: 0.982 - 0.996) on clinical balance test sets and low log-loss values of 0.180 (CI: 0.123 - 0.259) on equal balance and 0.223 (CI: 0.181 - 0.284) on clinical balance

test sets, which reflects the clinical setting based on the statistics on cytodiagnosis in 2016 to 2021 by the Japanese Society of Clinical Cytology (https://jscc.or.jp/). Both equal and clinical balance test sets (Table 1) were collected based on the cytodiagnoses, reviewed by two independent cytoscreeners or cytopathologists, then verified by a senior cytoscreener or cytopathologist. We ensured that we had consensus on the diagnoses of the test set WSIs. Our best model (ENB1-UC-FS+WS) also achieved high accuracy (0.945-0.946), sensitivity (0.940-0.960), and specificity (0.929-0.946) in WSI level. It has been reported that at urine LBC (ThinPrep) WSI level, the deep learning model predicted neoplastic (positive) WSI at 0.842 (accuracy), 0.795 (sensitivity), and 0.845 (specificity) [39]. Our latest reported uterine cervix LBC (ThinPrep) model demonstrated accuracy at 0.907, sensitivity at 0.850, and specificity at 0.911 at WSI level [20]. In this study, we have trained total four deep learning models (ENB1-UC-FS+WS, ENB1-UC-WS, ENB1-IN-FS+WS, and ENB1-IN-WS) using two different weight initialisation: ImageNet and pre-trained uterine cervix neoplastic LBC model from a previous study [20] (Figure 2). At WSI level, these four models showed almost comparable ROC-AUC, log-loss, accuracy, sensitivity, and specificity (Table 4 and Figure 3). However, there was wide variety of tile level prediction as visualized by the heatmap images between the four models (Figure 4). Based on the WSI and tile level (heatmap) evaluations, we have concluded that the model (ENB1-UC-FS+WS) which was trained using the pre-trained uterine cervix LBC model [20] weight initialisation, performed best. As for the false-negative prediction outputs in the urine LBC WSI which was cytodiagnosed as Class III (Figure 8A), the model (ENB1-UC-FS+WS) could not predict neoplastic atypical urothelial epithelial cell cluster (Figure 8B-F) in which neoplastic urothelial epithelial cells were overlapping and nuclear shapes and structures were hard to determine in the WSI (Figure 8B, C). The model (ENB1-UC-FS+WS) could predict true negative urine crystal and cell debris precisely. False negative prediction outputs were most likely due to neoplastic urothelial epithelial cell clusters that mimicked urine crystal or cell debris.

According to the annual statistics on cytodiagnosis by the Japanese Society of Clinical Cytology (https://jscc.or.jp/), in 2021, there were 2,041,547 urine cytodianosis reports in Japan. In 2021, the total number of cytodiagnosis in Japan was 7,157,413. Therefore, the population of urine cytodiagnosis was approximately 28.5%. In Japan, urine cytology was the second most common cytology in 2021, as cervical cytology was the most common (3,289,877 cases, 50.0%) (https://jscc.or.jp/). LBC of urine specimens is commonly used in cytology laboratories throughout the world and various processing methods, such as ThinPrep and SurePath, have been reported [40,41]. The LBC technique preserves the cells of interest (e.g., urothelial epithelial cells) in a liquid medium and removes most of the debris, blood, and exudate either by filtering or density gradient centrifugation. The efficiency of diagnosis employing the LBC is high because of the cell collection rate. It was demonstrated that the accuracy of diagnoses made employing the LBC method can be increased by understanding the characteristics of the urothelial epithelial cell morphology in suspicious cases [10]. Following the appropriate LBC specimen preparation steps, cell morphology (structure) is satisfactorily preserved, which allows more accurate diagnosing of LBC slides as shown by the significant concordance between cytological and histological diagnosis (92%), the significant number of LGUC (20.5%) revealed by urinary cytology and validated by histology, and the low rate (8%) of misjudgement of cytological diagnosis [42]. In addition, the leftover urine LBC material can be used for other techniques such as immunocytochemistry, molecular biology and flow cytometry [43]. Therefore, LBC has been applied with good results in urine cytology and can be regarded as an appropriate substitute for conventional smear urine cytology. LBC techniques opens new possibilities for a systemic urothelial carcinoma screening by integrating digital pathology WSI technique and deep learning model(s), resulting a standardised high-quality readout (e.g., classification).

One limitation of this study is that it primarily included urine LBC (ThinPrep) WSIs (both training and test sets) from a single private clinical laboratory in Japan. Therefore, the deep learning models could potentially be biased to such specimens. Validations on a wide variety of specimens from multiple different origins (both clinical laboratories and hospitals)

Table 1. Datasets

	training	validation	test (equal balance)	test (clinical balance)	Total
Negative	724	10	100	500	1334
Class I	360	5	50	250	665
Class II	364	5	50	250	669
Neoplastic	62	10	100	50	222
Class III	38	4	48	20	110
Class IV	11	3	23	14	51
Class V	13	3	29	16	61
Total	786	20	200	550	1556

Table 2. Annotation labels and numbers of annotation

Annotation label	Number of annotation
Atypical cell	9950
Low-grade urothelial carcinoma (LGUC) cell	1646
High-grade urothelial carcinoma (HGUC) cell	1611
Total	13207

and other LBC method(s) (e.g., SurePath) would be essential for ensuring the robustness of the models. Another potential validation study should involve the comparison of the performance of the models against cytoscreeners and cytopathologists in a clinical setting.

5. Conclusion

In the present study, we trained deep learning models for the classification of neoplastic urine LBC WSIs. We have evaluated the models on two test sets (equal and clinical balance) achieving ROC-AUCs for WSI diagnosis in the range of 0.984-0.990 by the best model (ENB1-UC-FS+WS). At WSI level, the model (ENB1-UC-FS+WS) achieved high accuracy (0.945-0.946), sensitivity (0.940-0.960), and specificity (0.929-0.946). Not only at WSI level, the model (ENB1-UC-FS+WS) satisfactorily predicted neoplastic urothelial epithelial cells (atypical, LGUC, and HGUC cells) by the heatmap images. Therefore, our model (ENB1-UC-FS+WS) can infer whether the urine LBC WSI is neoplastic (Figure 5) or negative (Figure 6) by inspecting model prediction outputs easily at WSI level as well as heatmap image, which makes it possible to use a deep learning model such as ours as a tool to aid in the urine LBC screening process in the clinical setting (workflow) for ranking cases by order of priority. Cytoscreeners and/or cytopathologists will need to perform full screening and subclassification (e.g., negative, atypical cells, suspicious for malignancy, and malignant) after the primary screening by our deep learning model, which could reduce their working time as the model would have highlighted the suspected neoplastic regions, and they would not have to perform an exhaustive search throughout the entire WSI.

Table 3. ROC-AUC and log-loss scores for existing deep learning models to classify liquid-based cytology (LBC) and histopathology whole slide images (WSIs)

Existing models	ROC-AUC	Log Loss
Liquid-based cytology (LBC)		
Uterine cervix Neoplastic (x10, 1024)	0.836 [0.775 - 0.885]	0.778 [0.620 - 0.989]

Table 4. ROC AUC, log loss, accuracy, sensitivity, and specificity results on the equal balance and clinical balance test sets.

|	Test set	
	Equal balance	Clinical balance
ENB1-UC-FS+WS (x10, 1024)		
ROC-AUC	0.984 [0.969 - 0.995]	0.990 [0.982 - 0.996]
Log-loss	0.180 [0.123 - 0.259]	0.223 [0.181 - 0.284]
Accuracy	0.945 [0.905 - 0.970]	0.946 [0.924 - 0.962]
Sensitivity	0.960 [0.920 - 0.990]	0.940 [0.861 - 1.000]
Specificity	0.929 [0.862 - 0.972]	0.946 [0.924 - 0.964]
ENB1-UC-WS (x10, 1024)		
ROC-AUC	0.990 [0.985 - 0.999]	0.990 [0.981 - 0.997]
Log-loss	0.251 [0.178 - 0.295]	0.098 [0.081 - 0.119]
Accuracy	0.955 [0.935 - 0.985]	0.940 [0.920 - 0.960]
Sensitivity	0.950 [0.911 - 0.990]	0.980 [0.933 - 1.000]
Specificity	0.960 [0.931 - 1.000]	0.936 [0.915 - 0.958]
ENB1-IN-FS+WS (x10, 1024)		
ROC-AUC	0.982 [0.957 - 0.996]	0.986 [0.963 - 0.998]
Log-loss	0.225 [0.156 - 0.321]	0.082 [0.063 - 0.106]
Accuracy	0.950 [0.910 - 0.975]	0.936 [0.918 - 0.956]
Sensitivity	0.930 [0.863 - 0.971]	0.960 [0.894 - 1.000]
Specificity	0.970 [0.930 - 1.000]	0.934 [0.914 - 0.955]
ENB1-IN-WS (x10, 1024)		
ROC-AUC	0.980 [0.960 - 0.997]	0.995 [0.990 - 0.998]
Log-loss	0.258 [0.185 - 0.289]	0.128 [0.114 - 0.144]
Accuracy	0.960 [0.940 - 0.990]	0.944 [0.924 - 0.960]
Sensitivity	0.970 [0.945 - 1.000]	1.000 [1.000 - 1.000]
Specificity	0.950 [0.914 - 0.990]	0.938 [0.915 - 0.956]
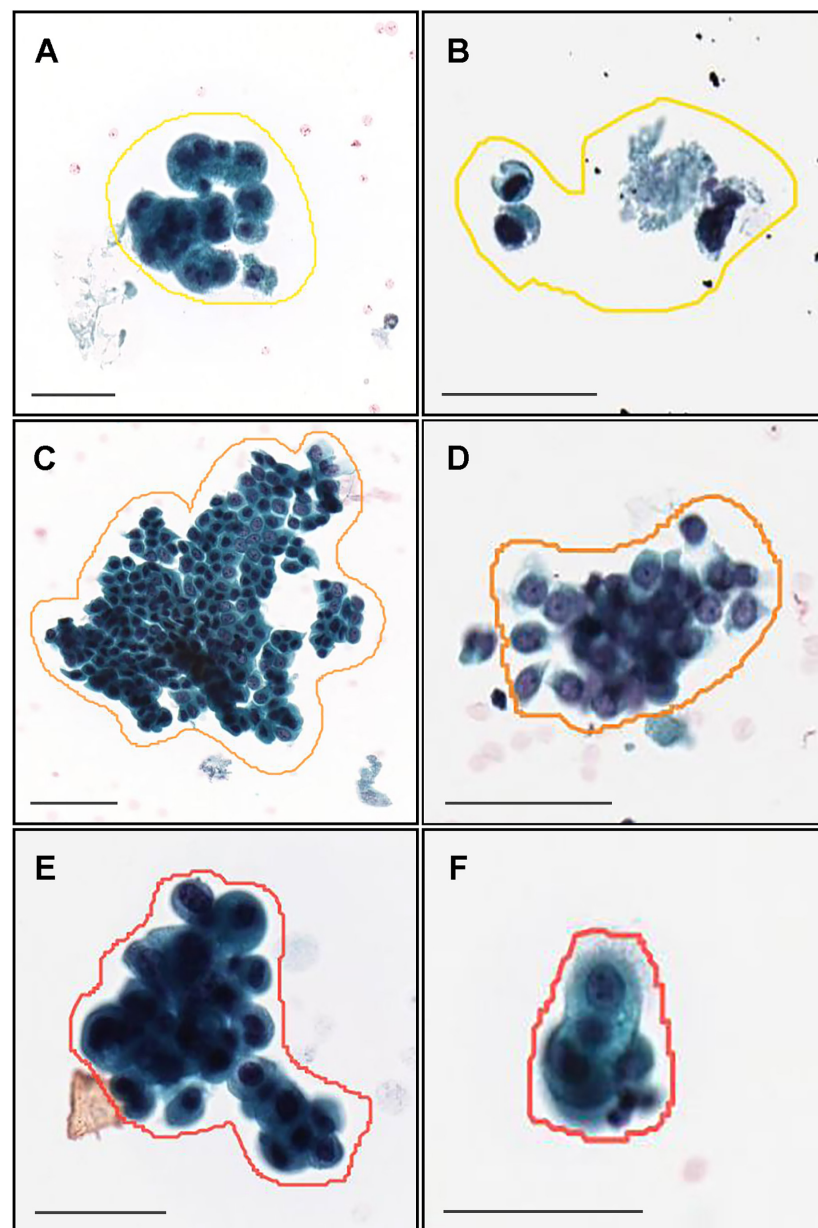
Figure 1. Representative manually drawing annotation image for neoplastic labels on urine liquid-based cytology (LBC) whole slide images (WSIs). The atypical urothelial cells (A, B) were annotated as atypical cell label. The suspected low grade urothelial carcinoma (LGUC) cells (C, D) were annotated as LGUC cell label and high grade utorhelial carcinoma (HGUC) cells (E, F) were annotated as HGUC cell label. The three labels (atypical cell, LGUC cell, and HGUC cell) were grouped as neoplastic label for fully supervised learning. Scale bars are 50 μ m.
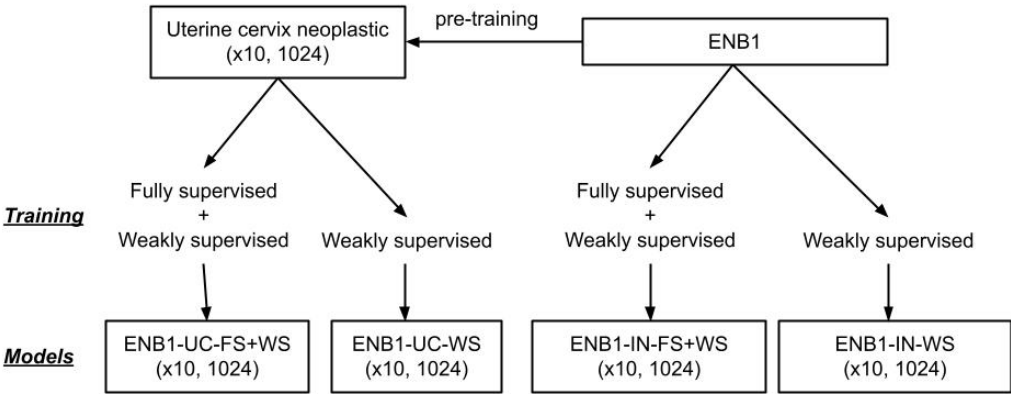
Figure 2. Training method and deep learning models overview. We performed training using two different weight initialisations: ImageNet (IN) and pre-training on a uterine cervix (UC) neoplastic (x10, 1024) dataset from a previous study. We used two different approaches for training: fully supervised (FS) and weakly supervised (WS) learning. This resulted in a total of four models, all trained at magnification x10 and tile size 1024x1024px: ENB1-UC-FS+WS, ENB1-UC-WS, ENB1-IN-FS+WS, and ENB1-IN-WS.
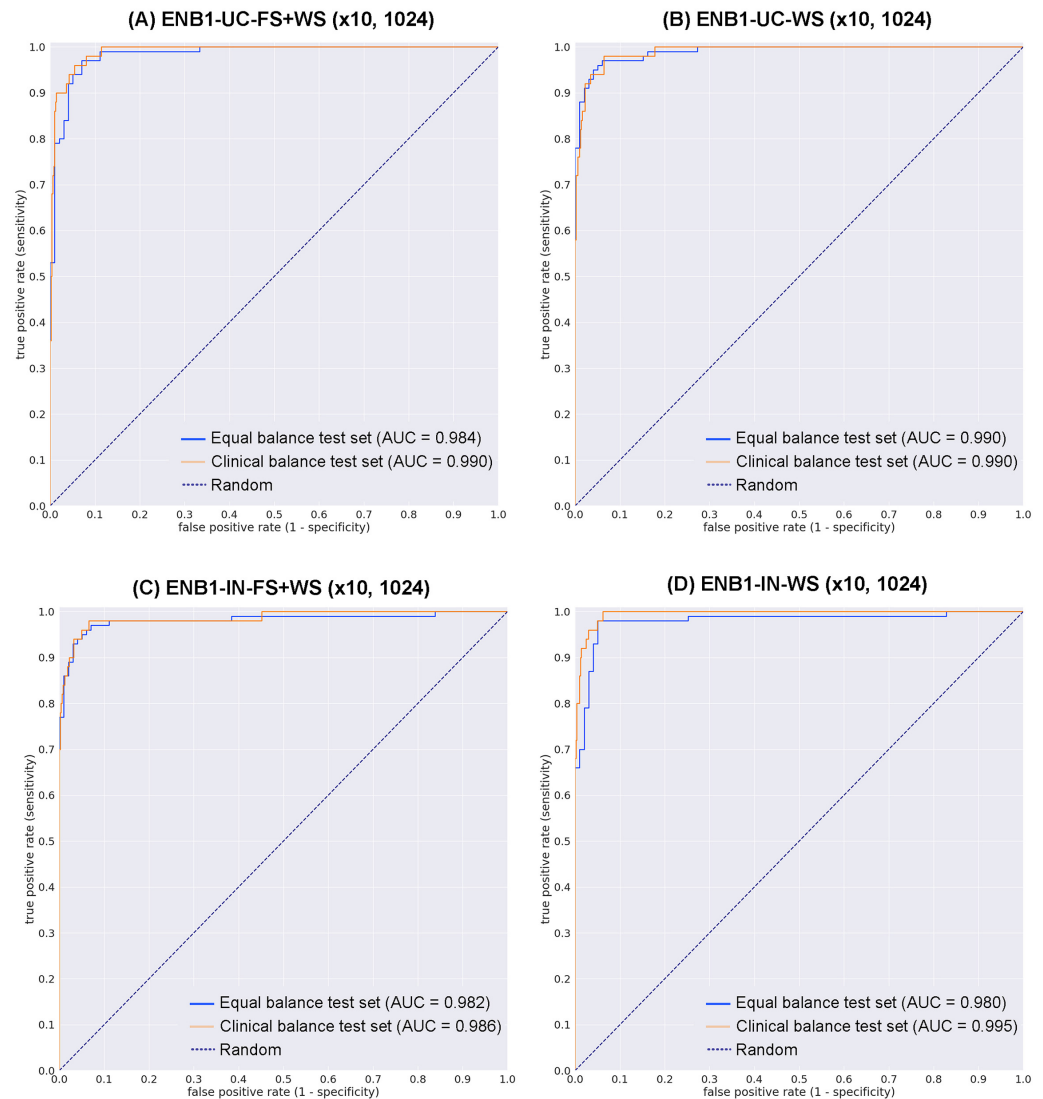
Figure 3. ROC curves for the equal balance and clinical balance test sets using four trained deep learning models. (A) transfer learning (TL) from uterine cervix liquid-based cytology (LBC) model and fully and weakly supervised learning model, magnification at x10 and tile size at 1024 x 1024 px (ENB1-UC-FS+WS (x10, 1024)); (B) TL from uterine cervix LBC model and weakly supervised learning model, magnification at x10 and tile size at 1024 x 1024 px (ENB1-UC-WS (x10, 1024)); (C) EfficientNetB1 based fully and weakly supervised learning model, magnification at x10 and tile size at 1024 x 1024 px (ENB1-IN-FS+WS (x10, 1024)); (D) EfficientNetB1 based weakly supervised learning model, magnification at x10 and tile size at 1024 x 1024 px (ENB1-IN-WS (x10, 1024)).
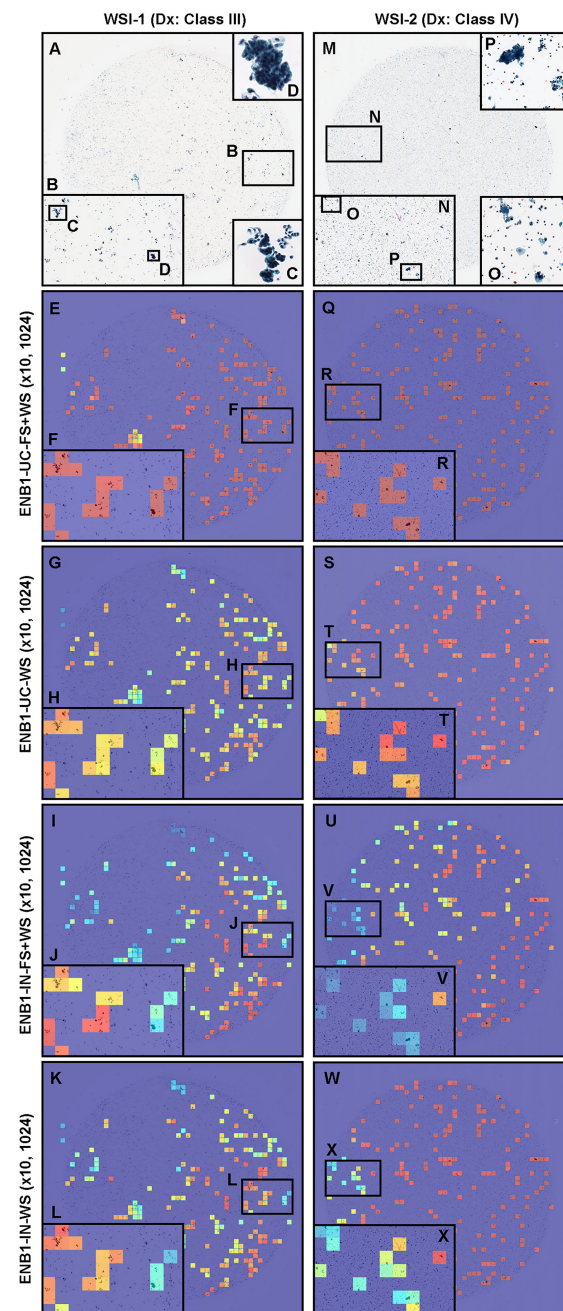
Figure 4. Comparison of neoplastic predictions in the representative two neoplastic urine liquid-based cytology (LBC) whole-slide images (WSIs) (WSI-1 and WSI-2) of four trained deep learning models (ENB1-UC-FS+WS, ENB1-UC-WS, ENB1-IN-FS+WS, and ENB1-IN-WS). According to the cytopathological diagnostic (Dx) reports, WSI-1 (A-L) was diagnosed as Class III and WSI-2 (M-X) was diagnosed as Class IV – both were classified in the neoplastic class in this study. (A-D and M-P): LBC cytopathological images for WSI-1 (A-D) and WSI-2 (M-O); heatmap prediction images for ENB1-UC-FS+WS model in WSI-1 (E-F) and WSI-2 (Q, R); heatmap prediction images for ENB1-UC-WS model in WSI-1 (G, H) and WSI-2 (S, T); heatmap prediction images for ENB1-IN-FS+WS model in WSI-1 (I, J) and WSI-2 (U, V); heatmap prediction images for ENB1-IN-WS model in WSI-1 (K, L) and WSI-2 (W, X). The localization of predicted tiles in neoplastic WSIs (WSI-1 and WSI-2) were almost same in four models (ENB1-UC-FS+WS, ENB1-UC-WS, ENB1-IN-FS+WS, and ENB1-IN-WS). However, the model pre-trained from uterine cervix LBC model with fully and weakly supervised learning (ENB1-UC-FS+WS) showed the highest neoplastic probabilities (F and R) in neoplastic tiles (B-D and N-P) as compared to other models (G-L and S-X). The heatmap uses the jet color map where blue indicates low probability and red indicates high probability.
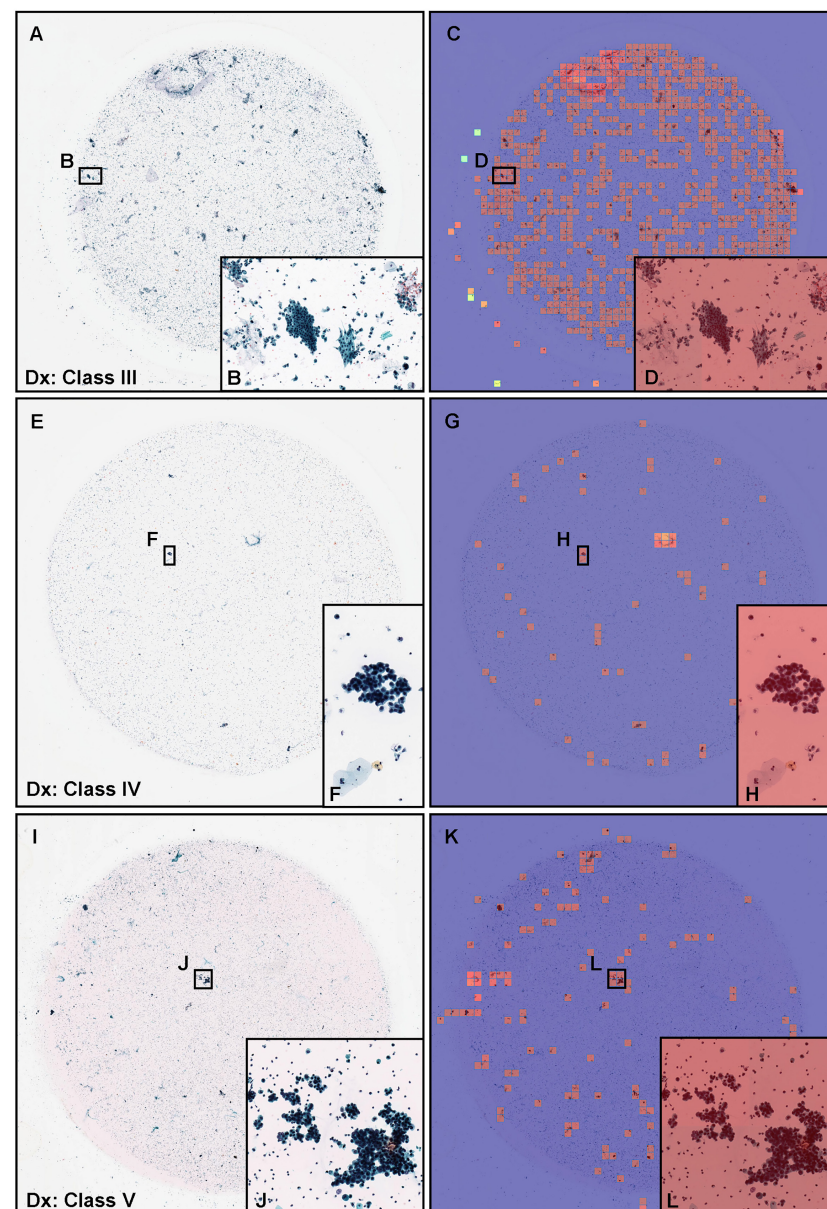
Figure 5. Representative examples of neoplastic true positive prediction outputs on urine liquid-based cytology (LBC) whole-slide images (WSIs) from test sets using the ENB1-UC-FS+WS model. According to the cytopathological diagnostic (Dx) reports, (A) was diagnosed as Class III with atypical urothelial epithelial cells (B), (E) was diagnosed as Class IV with suspected low grade urothelial carcinoma (LGUC) cells (F), and (I) was diagnosed as Class V with suspected high grade utorhelial carcinoma (HGUC) cells (J). The heatmap images (C, D, G, H, K, L) show true positive predictions of neoplastic urothelial epithelial cells (D, H, and L), which correspond respectively to atypical (B), suspected LGUC (F), and HGUC (J) cells. The heatmap uses the jet color map where blue indicates low probability and red indicates high probability.
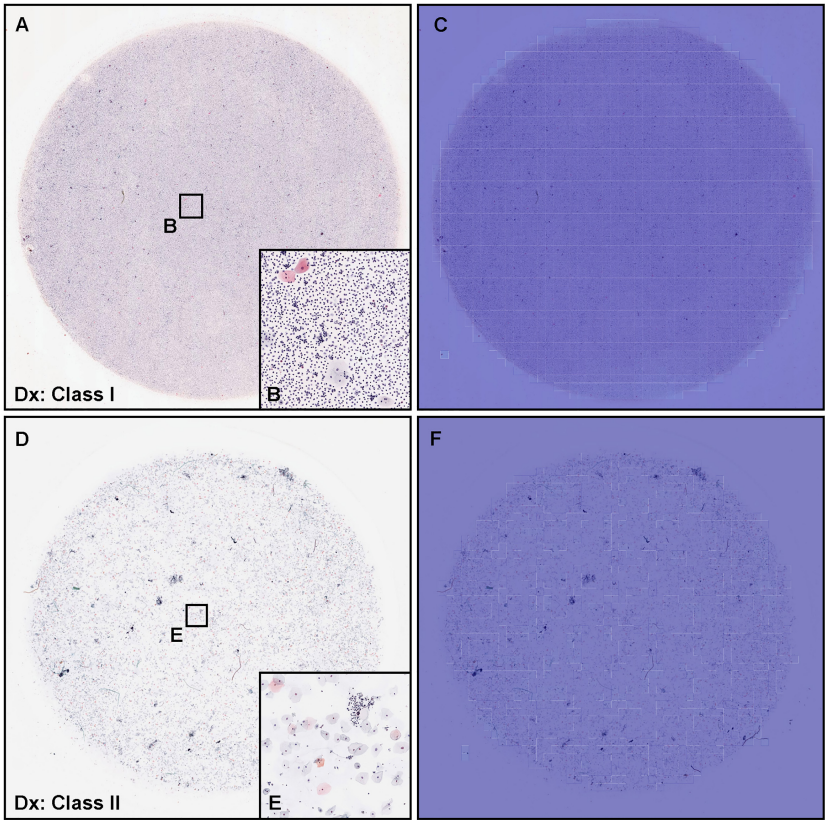
Figure 6. Two representative examples of neoplastic true negative prediction outputs on urine liquid-based cytology (LBC) whole-slide images (WSIs) from test sets using ENB1-UC-FS+WS model. According to the cytopathological diagnostic (Dx) reports, (A) was diagnosed as Class I and (B) was Class II, which were negative for urothelial neoplastic epithelial cells. Cytopathologically, (A) was pyuria which consisted of infective fluid (pus) with small number of non-atypical epithelial cells (B). (D) included urothelial epithelial cells with slight nuclear enlargement. The heatmap images (C, F) show true negative prediction of neoplastic epithelial cells. The heatmap uses the jet color map where blue indicates low probability and red indicates high probability.
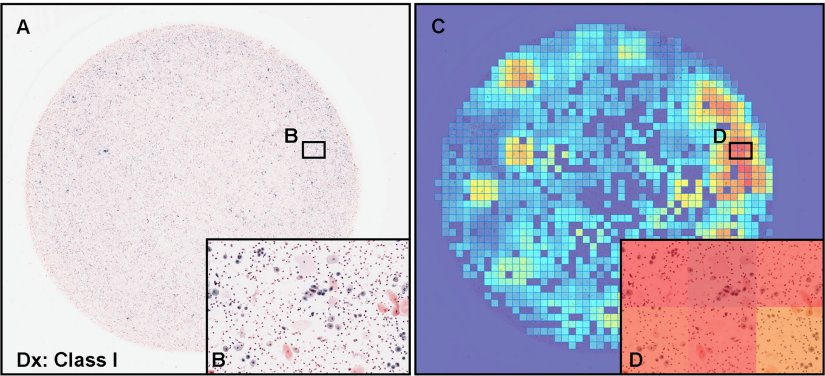

Figure 7. A representative example of neoplastic false positive prediction outputs on urine liquid-based cytology (LBC) whole-slide images (WSIs) from test sets using the ENB1-UC-FS+WS model . According to the cytopathological diagnostic (Dx) report, (A) was diagnosed as Class I and consisted of metaplastic squamous epithelial cells and non-atypical (non-neoplastic) urothelial epithelial cells with inflammatory cells (B). The heatmap images (C, D) show false positive predictions (D) which correspond respectively to (B). The heatmap uses the jet color map where blue indicates low probability and red indicates high probability.
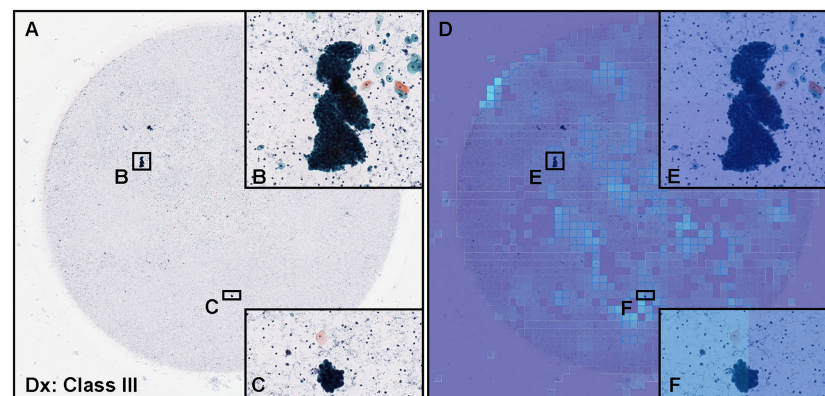
Figure 8. A representative example of neoplastic false negative prediction outputs on urine liquid-based cytology (LBC) whole-slide images (WSIs) from test sets using the ENB1-UC-FS+WS model. According to the cytopathological diagnostic (Dx) report, (A) was diagnosed as Class III and included clusters of atypical urothelial epithelial cells (B, C). The heatmap images (D-F) show false negative predictions (E, F) which correspond respectively to (B, C). The heatmap uses the jet color map where blue indicates low probability and red indicates high probability.

Institutional Review Board Statement: The experimental protocol in this study was approved by the ethical board of the private clinical laboratory in Japan. All research activities complied with all relevant ethical regulations and were performed in accordance with relevant guidelines and regulations in the private clinical laboratory. Due to the confidentiality agreement with the private clinical laboratory, the name of the private clinical laboratory cannot be disclosed.

Informed Consent Statement: Written informed consent to use cytopathological samples (liquid-based cytology glass slides) and cytopathological reports for research purposes in this study had previously been obtained from all patients and the opportunity for refusal to participate in research had been guaranteed by an opt-out manner.

Data Availability Statement: The datasets generated during and/or analysed during the current study are not publicly available due to specific institutional requirements governing privacy protection; however, they are available from the corresponding author and from the private clinical laboratory in Japan on reasonable request. The datasets that support the findings of this study are available from the private clinical laboratory (Japan), but restrictions apply to the availability of these data, which were used under a data-use agreement that was made according to the Ethical Guidelines for Medical and Health Research Involving Human Subjects as set by the Japanese Ministry of Health, Labour and Welfare (Tokyo, Japan) and, thus, are not publicly available. However, the datasets are available from the authors upon reasonable request for private viewing and with permission from the corresponding private clinical laboratory within the terms of the data use agreement and if compliant with the ethical and legal requirements as stipulated by the Japanese Ministry of Health, Labour and Welfare.

Conflicts of Interest: M.T. and F.K. are employees of Medmain Inc. All authors declare no competing interests.

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians* **2021**, *71*, 209–249.
2. Shanks, J.H.; Iczkowski, K.A. Divergent differentiation in urothelial carcinoma and other bladder cancer subtypes with selected mimics. *Histopathology* **2009**, *54*, 885–900.
3. DeSouza, K.; Chowdhury, S.; Hughes, S. Prompt diagnosis key in bladder cancer. *The Practitioner* **2014**, *258*, 23–7.
4. Luo, Y.; She, D.L.; Xiong, H.; Yang, L.; Fu, S.J. Diagnostic value of liquid-based cytology in urothelial carcinoma diagnosis: a systematic review and meta-analysis. *PloS one* **2015**, *10*, e0134940.
5. Raab, S.S.; Grzybicki, D.M.; Vrbin, C.M.; Geisinger, K.R. Urine cytology discrepancies: frequency, causes, and outcomes. *American journal of clinical pathology* **2007**, *127*, 946–953.
6. Sullivan, P.S.; Chan, J.B.; Levin, M.R.; Rao, J. Urine cytology and adjunct markers for detection and surveillance of bladder cancer. *American journal of translational research* **2010**, *2*, 412.
7. Bastacky, S.; Ibrahim, S.; Wilczynski, S.P.; Murphy, W.M. The accuracy of urinary cytology in daily practice. *Cancer Cytopathology: Interdisciplinary International Journal of the American Cancer Society* **1999**, *87*, 118–128.
8. Son, S.M.; Koo, J.H.; Choi, S.Y.; Lee, H.C.; Lee, Y.M.; Song, H.G.; Hwang, H.K.; Han, H.S.; Yun, S.J.; Kim, W.J.; et al. Evaluation of urine cytology in urothelial carcinoma patients: a comparison of CellprepPlus® liquid-based cytology and conventional smear. *Korean journal of pathology* **2012**, *46*, 68.
9. Kalantari, M.R.; Jahanshahi, M.A.; Gharib, M.; Hashemi, S.; Kalantari, S. Direct Smear Versus Liquid-Based Cytology in the Diagnosis of Bladder Lesions. *Iranian Journal of Pathology* **2022**, *17*, 56.
10. Washiya, K.; Sato, T.; Miura, T.; Tone, K.; Kojima, K.; Watanabe, J.; Kijima, H. Cytologic difference between benignity and malignancy in suspicious cases employing urine cytodiagnosis using a liquid-based method. *Analytical and Quantitative Cytology and Histology* **2011**, *33*, 169–174.
11. Laucirica, R.; Bentz, J.S.; Souers, R.J.; Wasserman, P.G.; Crothers, B.A.; Clayton, A.C.; Henry, M.R.; Chmara, B.A.; Clary, K.M.; Fraig, M.M.; et al. Do liquid-based preparations of urinary cytology perform differently than classically prepared cases? Observations from the College of American Pathologists Interlaboratory Comparison Program in Nongynecologic Cytology. *Archives of pathology & laboratory medicine* **2010**, *134*, 19–22.
12. Piaton, E.; Hutin, K.; Faynel, J.; Ranchin, M.; Cottier, M. Cost efficiency analysis of modern cytocentrifugation methods versus liquid based (Cytyc Thinprep®) processing of urinary samples. *Journal of clinical pathology* **2004**, *57*, 1208–1212.
13. Mukhopadhyay, S.; Feldman, M.D.; Abels, E.; Ashfaq, R.; Beltaifa, S.; Cacciabeve, N.G.; Cathro, H.P.; Cheng, L.; Cooper, K.; Dickey, G.E.; et al. Whole slide imaging versus microscopy for primary diagnosis in surgical pathology: a multicenter blinded randomized noninferiority study of 1992 cases (pivotal study). *The American journal of surgical pathology* **2018**, *42*, 39.
14. Malarkey, D.E.; Willson, G.A.; Willson, C.J.; Adams, E.T.; Olson, G.R.; Witt, W.M.; Elmore, S.A.; Hardisty, J.F.; Boyle, M.C.; Crabbs, T.A.; et al. Utilizing whole slide images for pathology peer review and working groups. *Toxicologic pathology* **2015**, *43*, 1149–1157.
15. Tsuneki, M. Deep learning models in medical image analysis. *Journal of Oral Biosciences* **2022**.
16. Lahrmann, B.; Valous, N.A.; Eisenmann, U.; Wentzensen, N.; Grabe, N. Semantic focusing allows fully automated single-layer slide scanning of cervical cytology slides. *PloS one* **2013**, *8*, e61441.
17. Sukegawa, S.; Tanaka, F.; Nakano, K.; Hara, T.; Yoshii, K.; Yamashita, K.; Ono, S.; Takabatake, K.; Kawai, H.; Nagatsuka, H.; et al. Effective deep learning for oral exfoliative cytology classification **2022**.
18. Ou, Y.C.; Tsao, T.Y.; Chang, M.C.; Lin, Y.S.; Yang, W.L.; Hang, J.F.; Li, C.B.; Lee, C.M.; Yeh, C.H.; Liu, T.J. Evaluation of an artificial intelligence algorithm for assisting the P aris S ystem in reporting urinary cytology: A pilot study. *Cancer Cytopathology* **2022**.
19. Tao, X.; Chu, X.; Guo, B.; Pan, Q.; Ji, S.; Lou, W.; Lv, C.; Xie, G.; Hua, K. Scrutinizing high-risk patients from ASC-US cytology via a deep learning model. *Cancer Cytopathology* **2022**.
20. Kanavati, F.; Hirose, N.; Ishii, T.; Fukuda, A.; Ichihara, S.; Tsuneki, M. A Deep Learning Model for Cervical Cancer Screening on Liquid-Based Cytology Specimens in Whole Slide Images. *Cancers* **2022**, *14*, 1159.
21. Xie, X.; Fu, C.C.; Lv, L.; Ye, Q.; Yu, Y.; Fang, Q.; Zhang, L.; Hou, L.; Wu, C. Deep convolutional neural network-based classification of cancer cells on cytological pleural effusion images. *Modern Pathology* **2022**, *35*, 609–614.
22. Lin, Y.J.; Chao, T.K.; Khalil, M.A.; Lee, Y.C.; Hong, D.Z.; Wu, J.J.; Wang, C.W. Deep Learning Fast Screening Approach on Cytological Whole Slides for Thyroid Cancer Diagnosis. *Cancers* **2021**, *13*, 3891.
23. Bhatt, A.R.; Ganatra, A.; Kotecha, K. Cervical cancer detection in pap smear whole slide images using convnet with transfer learning and progressive resizing. *PeerJ Computer Science* **2021**, *7*, e348.
24. Holmström, O.; Linder, N.; Kaingu, H.; Mbuuko, N.; Mbete, J.; Kinyua, F.; Törnquist, S.; Muinde, M.; Krogerus, L.; Lundin, M.; et al. Point-of-Care Digital Cytology With Artificial Intelligence for Cervical Cancer Screening in a Resource-Limited Setting. *JAMA network open* **2021**, *4*, e211740–e211740.
25. Lin, H.; Chen, H.; Wang, X.; Wang, Q.; Wang, L.; Heng, P.A. Dual-path network with synergistic grouping loss and evidence driven risk stratification for whole slide cervical image analysis. *Medical Image Analysis* **2021**, *69*, 101955.
26. Cheng, S.; Liu, S.; Yu, J.; Rao, G.; Xiao, Y.; Han, W.; Zhu, W.; Lv, X.; Li, N.; Cai, J.; et al. Robust whole slide image analysis for cervical cancer screening using deep learning. *Nature communications* **2021**, *12*, 1–10.
27. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. International Conference on Machine Learning. PMLR, 2019, pp. 6105–6114.

28. Kanavati, F.; Tsuneki, M. Partial transfusion: on the expressive influence of trainable batch norm parameters for transfer learning. *arXiv preprint arXiv:2102.05543* **2021**.

29. Tsuneki, M.; Abe, M.; Kanavati, F. Transfer Learning for Adenocarcinoma Classifications in the Transurethral Resection of Prostate Whole-Slide Images. *Cancers* **2022**, *14*, 4744.

30. Otsu, N. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics* **1979**, *9*, 62–66.

31. Goode, A.; Gilbert, B.; Harkes, J.; Jukic, D.; Satyanarayanan, M. OpenSlide: A vendor-neutral software foundation for digital pathology. *Journal of pathology informatics* **2013**, *4*.

32. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* **2014**.

33. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, 2015. Software available from tensorflow.org.

34. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research* **2011**, *12*, 2825–2830.

35. Hunter, J.D. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* **2007**, *9*, 90–95. doi:10.1109/MCSE.2007.55.

36. Efron, B.; Tibshirani, R.J. *An introduction to the bootstrap*; CRC press, 1994.

37. Kanavati, F.; Ichihara, S.; Rambeau, M.; Iizuka, O.; Arihiro, K.; Tsuneki, M. Deep learning models for gastric signet ring cell carcinoma classification in whole slide images. *Technology in Cancer Research & Treatment* **2021**, *20*, 15330338211027901.

38. Kanavati, F.; Toyokawa, G.; Momosaki, S.; Rambeau, M.; Kozuma, Y.; Shoji, F.; Yamazaki, K.; Takeo, S.; Iizuka, O.; Tsuneki, M. Weakly-supervised learning for lung carcinoma classification using deep learning. *Scientific reports* **2020**, *10*, 1–11.

39. Sanghvi, A.B.; Allen, E.Z.; Callenberg, K.M.; Pantanowitz, L. Performance of an artificial intelligence algorithm for reporting urine cytopathology. *Cancer cytopathology* **2019**, *127*, 658–666.

40. Ren, S.; Solomides, C.; Draganova-Tacheva, R.; Bibbo, M. Overview of nongynecological samples prepared with liquid-based cytology medium. *Acta Cytologica* **2014**, *58*, 522–532.

41. Nasuti, J.F.; Tam, D.; Gupta, P.K. Diagnostic value of liquid-based (Thinprep®) preparations in nongynecologic cases. *Diagnostic cytopathology* **2001**, *24*, 137–141.

42. Raisi, O.; Magnani, C.; Bigiani, N.; Cianciavicchia, E.; D'Amico, R.; Muscatello, U.; Ghirardini, C. The diagnostic reliability of urinary cytology: a retrospective study. *Diagnostic Cytopathology* **2012**, *40*, 608–614.

43. Rossi, E.; Mulè, A.; Russo, R.; Pierconti, F.; Fadda, G. Application of liquid-based preparation to non-gynaecologic exfoliative cytology. *Pathologica* **2008**, *100*, 461–5.