 doi:10.20944/preprints202208.0080.v1

Article

A Weakly Supervised Tissue-level Approach for Histological Differentiation of Gastric Adenocarcinomas

Dong-Woo Hyeon [1], Dongmin Kim [1], Wi-Sun Ryu [1], Samuel Hanjoo Sung [1], Subrata Bhattacharjee [2], Cho-Hee Kim [1], Nam-Hun Cho [3], Damin Moon [1], Doo-Hee Lee [1], Hyunki Kim [3], Hyeong-Chan Shin [4], Won-Ae Lee [5], Min-Sun Cho [6], Heung-Kook Choi [1,*] and Haeyoun Kang [7,*]

[1] JLK Artificial Intelligence R&D Center, Seoul 06141, Korea; dwhyeon@jlkgroup.com (D.-W.H.); dmkim@jlkgroup.com (D.K.); wisunryu@jlkgroup.com (W.-S.R.); hjsung@jlkgroup.com (S.-H.S.); chkim@jlkgroup.com (C.-H.K.); dmmoon@jlkgroup.com (D.M.); dhlee1@jlkgroup.com (D.-H.L.);
[2] Department of Computer Engineering, u-AHRC, Inje University, Gimhae 50834, Korea; Subrata_bhattacharjee@outlook.com
[3] Department of Pathology, Yonsei University College of Medicine, Seoul 03722, Korea; CHO1988@yuhs.ac (N.-H.C.); KIMHYUNKI@yuhs.ac (H.K.)
[4] Department of Pathology, Keimyung University Dongsan Medical Center, Daegu 42601, Korea; chan@dsmc.or.kr
[5] Department of Pathology, Dankook University Hospital, Cheonan 31116, Korea; walee@dankook.ac.kr
[6] Department of Pathology, Ewha Womans University Seoul Hospital, Seoul 07804, Korea; mcho1124@ewha.ac.kr
[7] Department of Pathology, CHA University, CHA Bundang Medical Center, Seongnam 13496, Korea; hykang@cha.ac.kr
* Correspondence: hkchoi@jlkgroup.com (H.-K.C.); hykang@cha.ac.kr (H.K.); Tel.: +82-10-6733-3437 (H.-K.C.); +82-10-2273-7420 (H.K)

Simple Summary: Histological differentiation is critical given the possibility of lymph node metastasis of gastric cancer. We developed a weakly supervised learning technique for tissue-level gastric cancer adenocarcinoma histological classification (well-to-moderately or poorly differentiated). We cropped tissue regions from whole slide images (WSIs) and performed tissue-level labeling. The tissue-level areas under the receiver operating characteristic (AUROC) curves of the histological differentiation classifiers were 0.953, 0.969, and 0.943, respectively, when data from five hospitals were subjected to threefold cross-validation. Class-discriminatory regions can be visualized using Grad-CAM technique. The results of our weakly supervised model are easy to visually interpret, and indicate that tissue-level learning could aid histopathological classifications that require a wide field of view, such as grading of differentiation.

Abstract: Histologically poor differentiation is associated with lymph node metastasis. Thus, pathological evaluation of biopsy specimens is crucial when treating stomach cancers. Deep learning of WSIs is challenging because the images are enormous. Given the computing limitations, patch-level supervised learning methods have been proposed. However, valuable information is lost when dividing WSIs into smaller patches. Another drawback is the need for pixel-level annotation by a pathologist. It is acceptable to differentiate, i.e., grade, gastric cancer at the holistic tissue level (i.e., under low magnification). We developed a weakly supervised learning technique for tissue-level gastric adenocarcinoma histological differentiation (well-to-moderately or poorly differentiated) and applied global reasoning to tissue-level features. The tissue-level AUROCs of the histological differentiation classifiers were 0.953, 0.969, and 0.943, respectively when data from five hospitals were subjected to threefold cross-validation. Comparison of the Grad-CAM heatmaps of the trained classifier and the pathologists' annotations confirmed that our weakly supervised model exhibited performed well.

Keywords: gastric cancer; deep learning; digital pathology; lymph node metastasis

 doi:10.20944/preprints202208.0080.v1

1. Introduction

According to the Global Cancer Statistics report of 2020, stomach cancer ranked fifth in terms of incidence and fourth in terms of mortality [1]. In Japan, gastric cancer biopsy specimens are histologically classified as differentiated (tubular and papillary adenocarcinomas) or undifferentiated (poorly differentiated [PD] adenocarcinomas and signet ring cell carcinomas) [2]. As histologically poor differentiation is associated with lymph node metastasis, pathological examination of biopsy specimens is crucial [3].

Digital pathology employing deep learning is progressing rapidly [4]. However, deep learning of whole slide images (WSIs) remains challenging because the images are enormous. Given the computing limitations, patch-level supervised learning methods have been proposed. These methods divide WSIs into smaller patches (e.g., 224 × 224 pixels) for processing using a convolutional neural network (CNN) [5-11]. Patch-level supervised learning requires pixel-level annotation by pathologists, which is extremely laborious; moreover, pathologists often disagree [12].

Recent studies have explored weak supervision methods. Gradient-weighted class activation mapping (Grad-CAM) [13] is a visualization technique widely used by those researching weakly supervised learning. Grad-CAM highlights class-relevant discriminatory regions by calculating gradients in feature maps.

WSI-level weakly supervised methods [multiple instance learning (MIL) and whole-side training methods] are emerging [14-20]. A whole-slide training method [21] facilitates deep learning of entire WSIs using several GPU memory optimization techniques. However, such learning is highly demanding in terms of GPU memory, which slows learning [22].

We developed a weakly supervised learning technique for tissue-level gastric cancer histological differentiation (well-to-moderately differentiated [WMD] or PD). As large areas of biopsy images are white, tissue-level processing eliminates a great deal of unnecessary information. After cropping the tissue area, model training is possible without the need for GPU memory optimization. Our model is weakly supervised; it learns only tissue-level labels; there is no need for pixel-level annotation. Class-discriminatory regions can be visualized using Grad-CAM after training the deep learning classifier.

2. Materials and Methods

2.1. Subsection

A total of 680 hematoxylin-and-eosin (H&E)-stained histopathological stomach cancer specimens were collected from the surgical pathology files of five South Korean hospitals cooperating to build an AI algorithm for pathologists (the AI-ALPA consortium): CHA Bundang Medical Center (CHABMC), Keimyung University Dongsan Medical Center (KUDMC), Dankook University Hospital (DUH), Yonsei University Hospital (YUH), and Ewha Womans University Medical Center (EWUMC).

2.2. Pathological Diagnosis

Pathologists from all five institutions performed annotation using QuPath software [23]. The AI-ALPA consortium classifies gastric cancer as WMD or PD adenocarcinomas. First, the pathologists selected the gastric biopsy regions of the WSIs with the most representative high-quality adenocarcinoma areas and drew regions of interest (ROIs). Then, selected tissues were annotated at the pixel level as WMD or PD. An example is shown in Figure 1A.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 August 2022 doi:10.20944/preprints202208.0080.v1

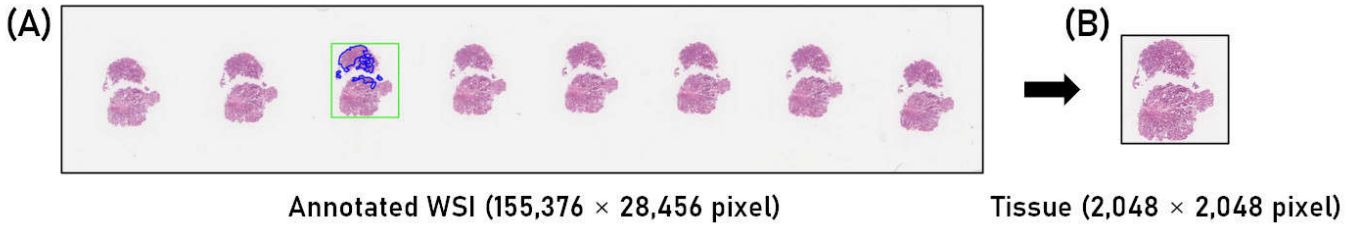

Figure 1. Extraction of tissue of interest from an annotated WSI (**A**) The annotated WSI (the green rectangle indicates the region of interest; WMD areas are indicated in blue within the rectangle). (**B**) The tissue from the annotated WSI. The rectangle was resized to 2,048 × 2,048 pixels.

2.3. The Deep Learning Model

For tissue-level labeling, we first cropped the tissue regions on all annotated WSIs. The ROIs drawn during annotation were extracted. As an example (Figure 1), the WSI is very large (155,376 × 28,456 pixels). However, the size of the green ROI rectangle containing the tissue of interest is only 11,256 × 12,384 pixels. When resizing the ROI regions, 12,384 × 12,384-pixel squares were used; each side of the square was as long as the longer side of the rectangle. Each tissue was resized to 2,048 × 2,048 pixels (Figure 1B). The OpenSlide [24] library was used to extract specific regions from WSI images.

Figure 2 illustrates the entire workflow. Figure 2A shows tissue-level labeling. As described above, 2,048 × 2,048-pixel tissue images were extracted from annotated AI-ALPA WSIs, followed by tissue-level labeling based on annotations within the ROI region. We classified the ROI type as mixed when WMD and PD annotations co-existed. Mixed-type ROIs were used for testing only, i.e., not for training. To train the classifier, we used only pure WMD and PD tissues.

The ResNet-50 [25, 26] classifier was trained to classify the 2,048 × 2,048-pixel images as WMD or PD (Figure 2B). Figure 3 shows the architecture of the ResNet-50 network; this CNN has 50 layers in five convolution blocks. The blocks differ slightly from those of conventional CNNs; identity connections between the layers deal with the vanishing gradient problem of deep neural networks. The network was implemented using the PyTorch [27] library. We used an NVIDIA RTX A6000 GPU (which supports 2,048 × 2,048-pixel images) to train the classifier and performed threefold cross-validation to test the performance of the algorithm when evaluating the multi-institutional data.

Figure 2C shows the inference step. For each WSI, the tissue regions are individually identified, and the histological classes predicted, using the ResNet-50 classifier; the Grad-CAM heatmap of the predicted class is then displayed.

The foreground extraction technique was used for automatic WSI tissue identification. Other foreground extraction algorithms have been described. Bug et al. [28] developed a filter-based algorithm, and Bandi et al. [29] a fully convolutional deep learning approach, for tissue segmentation. In contrast, we used a simpler empirical method. Note that we only require the bounding box containing the tissue region; it is not necessary to precisely segment tissue. Figure 2C shows an automatic tissue identification result. For each WSI, we converted the RGB (red, green, blue) scale to an HSV (hue, saturation, and value) scale. H, S, and V were expressed as values between 0 and 255. The S channel value indicates that 0 is achromatic (gray), while 255 is clear (pure color). We extracted only the portion wherein the S value was $\geq$ 4. Next, dilation and erosion were applied for noise reduction. Finally, we drew bounding boxes (on the original WSI) around each component of the extracted foreground. The algorithm accurately identified the foreground. To visualize the result, we subjected the trained ResNet-50 classifier to Grad-CAM analysis. We compared the Grad-CAM heatmap to the pathologists' annotations to confirm that our weakly supervised model performed well.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 August 2022 doi:10.20944/preprints202208.0080.v1

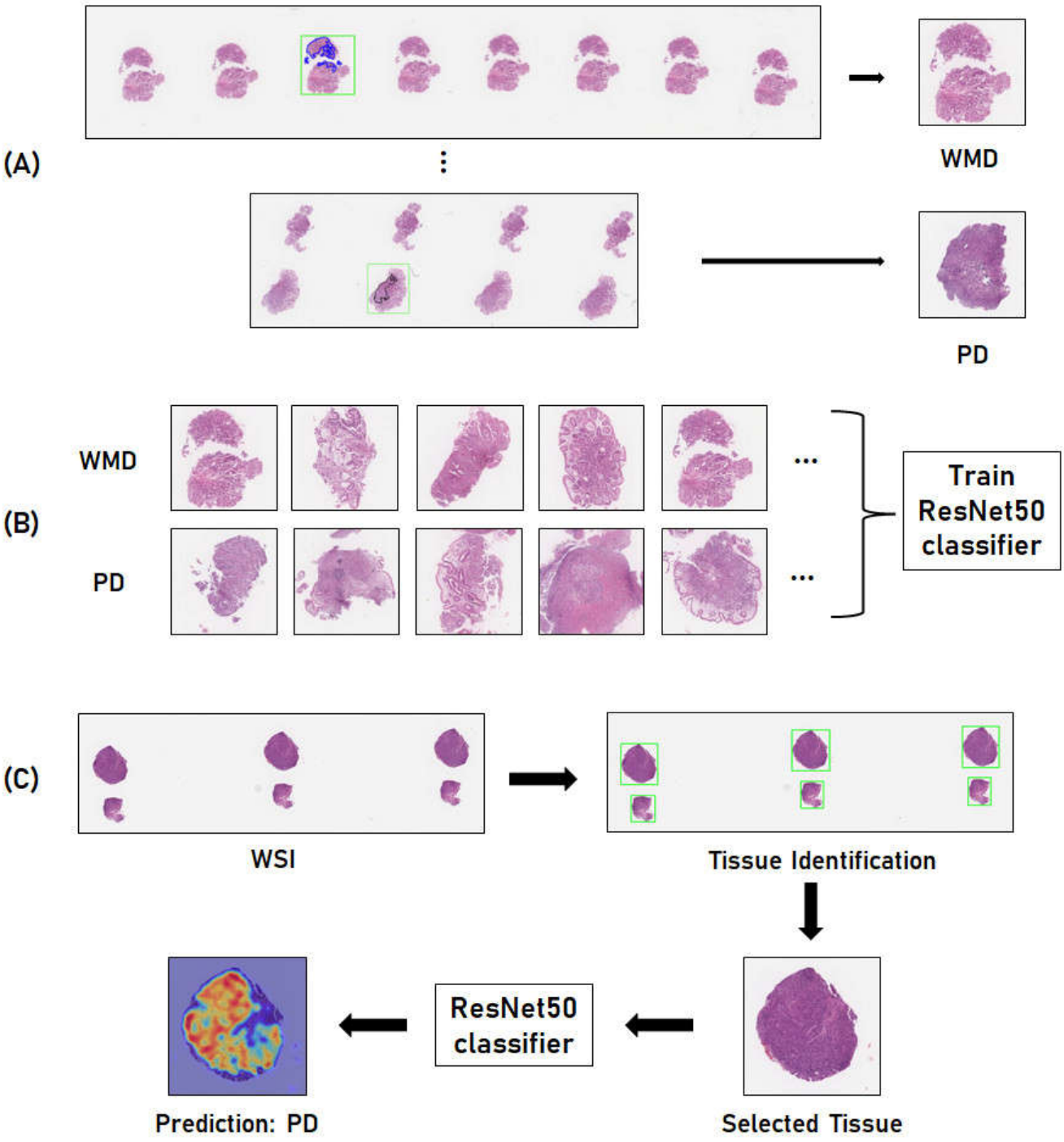

Figure 2. Workflow of the weakly supervised approach toward gastric cancer histological differentiation and adenocarcinoma classification. (**A**) Tissue extraction from the WSIs and categorization into two classes (well-to- moderately differentiated [WMD] or poorly differentiated [PD]) based on the pathologists' annotations (the ROI rectangle, WMD, and PD areas are indicated in green, blue, and black respectively). (**B**) The ResNet-50WMD/PD classifier was trained using the categorized WMD/PD 2,048 × 2,048-pixel tissues. (**C**) Inference: For each WSI, all tissue regions are identified individually using bounding boxes. The trained ResNet-50 classifier predicts the histological differentiation and the Grad-CAM heatmap interprets the predictions.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 August 2022 doi:10.20944/preprints202208.0080.v1

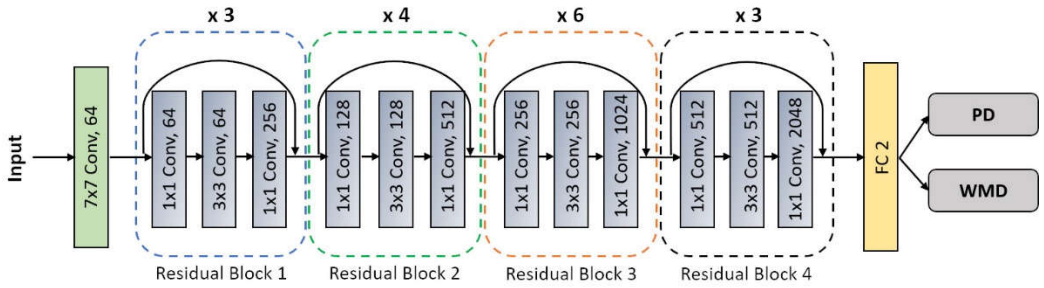

Figure 3. The ResNet-50 model architecture. The notation [n×n Conv, m] in each block denotes a filter of size n with m channels. FC 2 is the fully connected layer with two neurons. The numbers on top of the blocks are the repetitions of each unit. The connections within the residual blocks are the skip connections applied (using the summation function) to connect all predictive feature-maps.

3. Results

3.1. Tissue Extraction

A total of 714 ROIs were marked on 680 slides. Figure 4 shows the distribution of the longest ROI side lengths. A total of 673 tissues were used; we excluded areas with side lengths > 20,000 pixels. The average length (longer side) of an ROI was 8,246 pixels. When resizing to 2,048 × 2,048 pixels, the average ROI downscaling was 4.03-fold. Chen et al. [21] downscaled the WSIs fivefold (i.e., to 0.2-fold the original size) for whole-slide training. As the lung cancer type classification performance was excellent (AUROC of 0.959 for adenocarcinomas), our smaller downscaling process is acceptable.

Tissue-level labels were assigned based on annotations in the ROIs. Table 1 lists the numbers of tissues for each institution and class. CHABMC and EWUMC were grouped, as well as DUH and YUH, prior to cross-validation. We used three groups because EWUMC and YUH process fewer tissues than the other institutions (Table 2).

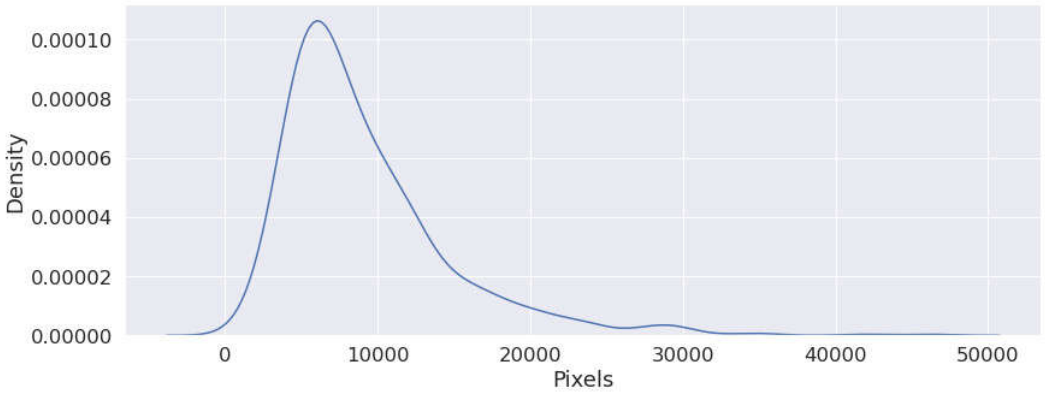

Figure 4. The ROI lengths of the AI-ALPA-annotated dataset.

Table 1. Numbers of labeled tissue of the AI-ALPA consortium.

Hospitals	WMD	PD	Mixed	Total
CHABMC	91	45	9	145
EWUMC	51	14	0	65
KUDMC	114	61	10	185
DUH	109	59	11	179
YUH	47	47	5	99
Total	412	226	35	673

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 August 2022 doi:10.20944/preprints202208.0080.v1

Table 2. The threefold cross-validation scheme involving five institutions.

	Train	Validation
Fold 1	DUH+YUH, KUDMC	CHABMC+EWUMC
Fold 2	CHABMC+EWUMC, KUDMC	DUH+YUH
Fold 3	CHABMC+EWUMC, DUH+YUH	KUDMC

3.2. Classification Performance

The WMD and PD tissue classification performances are shown in Table 3. The areas under the receiver operating characteristic (AUROC) curves and accuracies of the validation datasets for all folds are listed. The AUROC 95% confidence intervals were estimated using the Delong method [30]. The AUROC scores were > 0.943 for all folds. Figure 5 shows the confusion matrices and ROC curves for all folds.

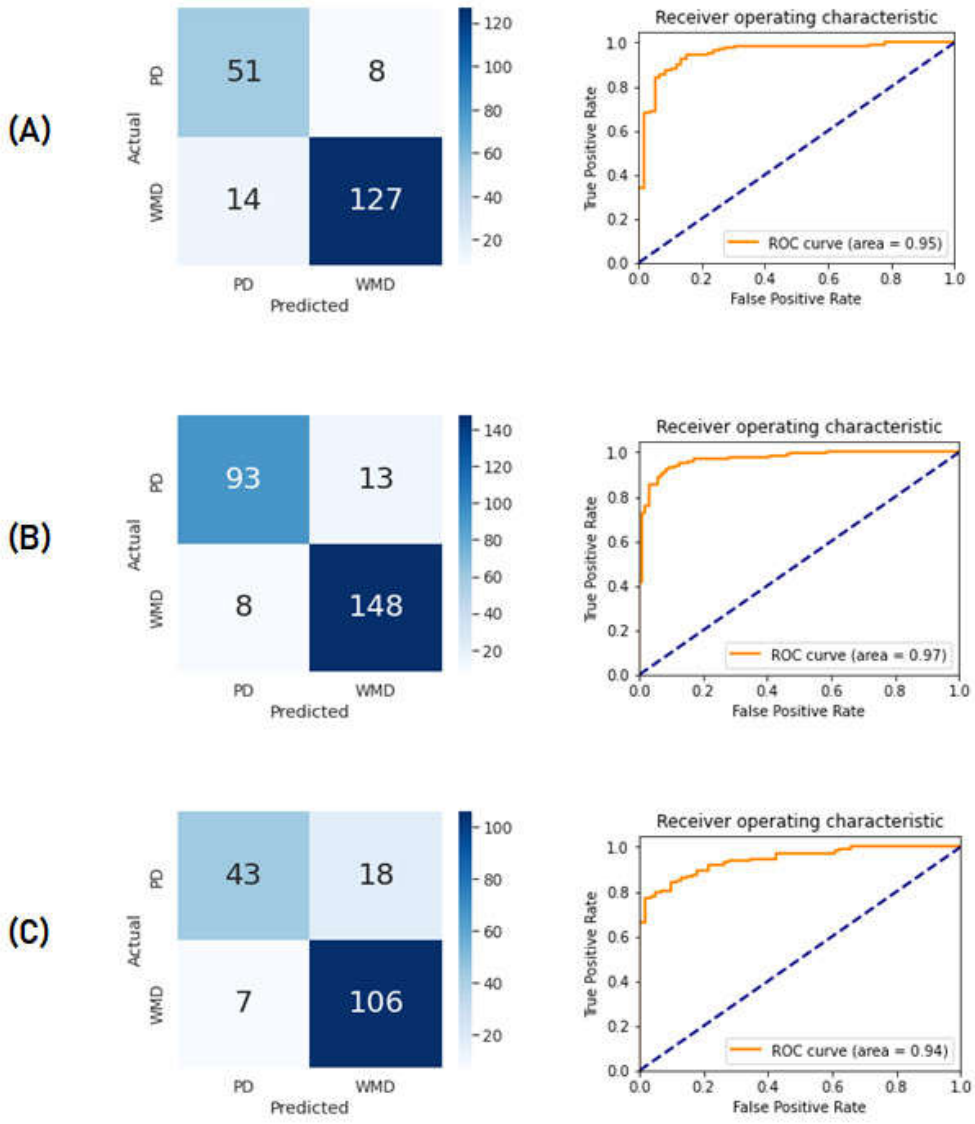

Figure 5. Three-fold validation of the tissue classifiers used for histological differentiation of gastric cancer. (**A–C**) Confusion matrices and receiver operating characteristic curves for folds 1–3, respectively.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 August 2022 doi:10.20944/preprints202208.0080.v1

Table 3. AUROC and accuracy values obtained by applying our method to the validation sets of all folds for histological classification of gastric cancer.

	AUROC	Accuracy (%)
Fold 1	0.953 [CI: 0.922-0.984]	89.0
Fold 2	0.969 [CI: 0.950-0.988]	92.0
Fold 3	0.943 [CI: 0.913-0.974]	85.6

3.3. Visual Interpretation

Our model performed impressively, but it was necessary to understand how it made decisions. Class-discriminant regions can be visualized by Grad-CAM. Figure 6 compares the Grad-CAM visualizations (right panels) and pathologist annotations (left panels) of the validation set. The numbers in parentheses are the classifier output scores for the predicted class. All output scores are between 0 and 1 because we normalized the classifier outputs using the SoftMax activation function.

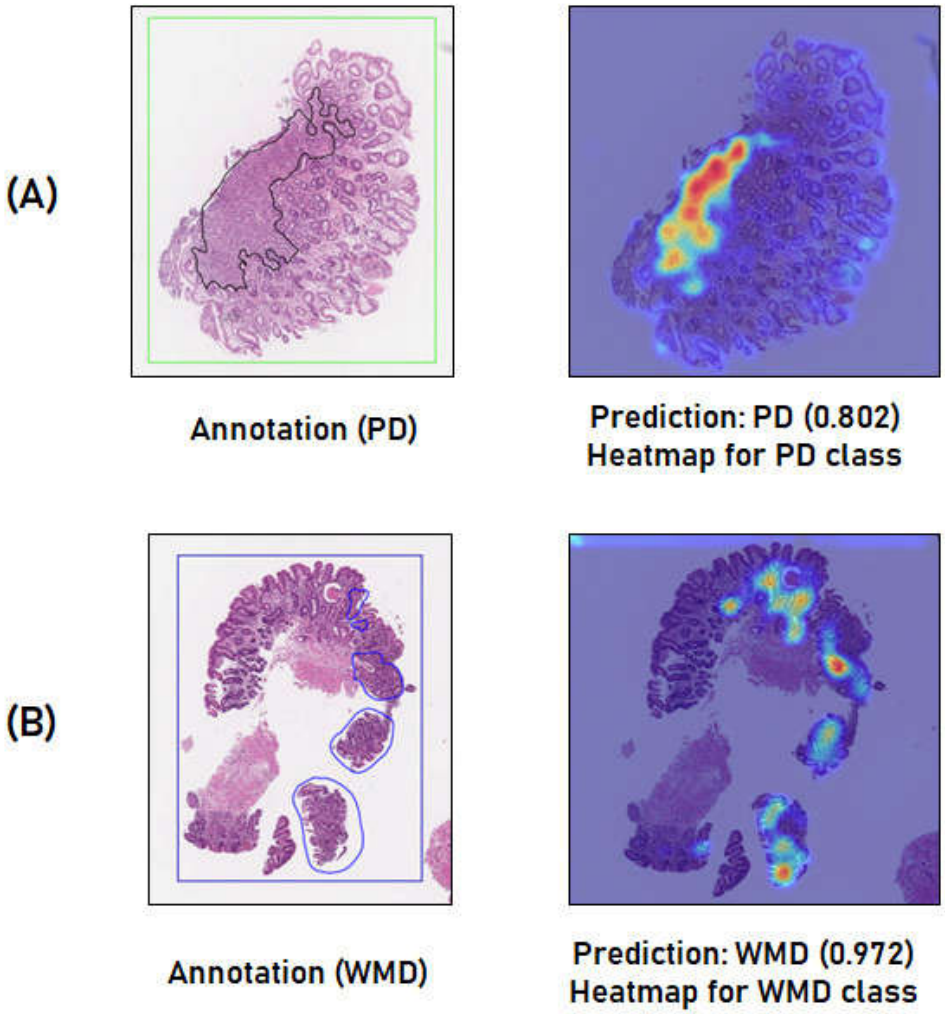

Figure 6. Comparison of pathologists' annotations and the Grad-CAM heatmaps generated by a classifier acting on the validation set. (**A**) A tissue with a PD annotation (black contour) and the heatmap for the PD class. (**B**) A tissue with a WMD annotation (blue contour) and the heatmap for the WMD class.

We next explored whether our model could identify representative information without pixel-level annotation. We used a ″pointing game″ to quantify model performance [31]. We extracted the location of the largest Grad-CAM pixel from each heatmap to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 August 2022 doi:10.20944/preprints202208.0080.v1

determine if it was included in the reference annotation (Figure 7). A ”hit” occurs when the largest Grad-CAM pixel was within the reference annotation; all other cases were “misses”. The Grad-CAM performance was given by the hit rate across the dataset (Table 4). The human experts and artificial intelligence were in good agreement.

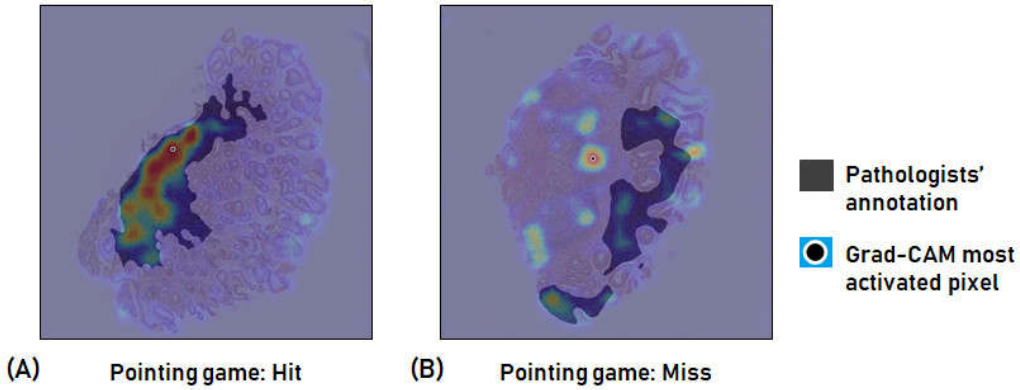

Figure 7. Evaluation of visualization performance using a pointing game. (**A**) A “hit” result; the largest Grad-CAM pixel is inside the reference annotation (**B**) A “miss” result.

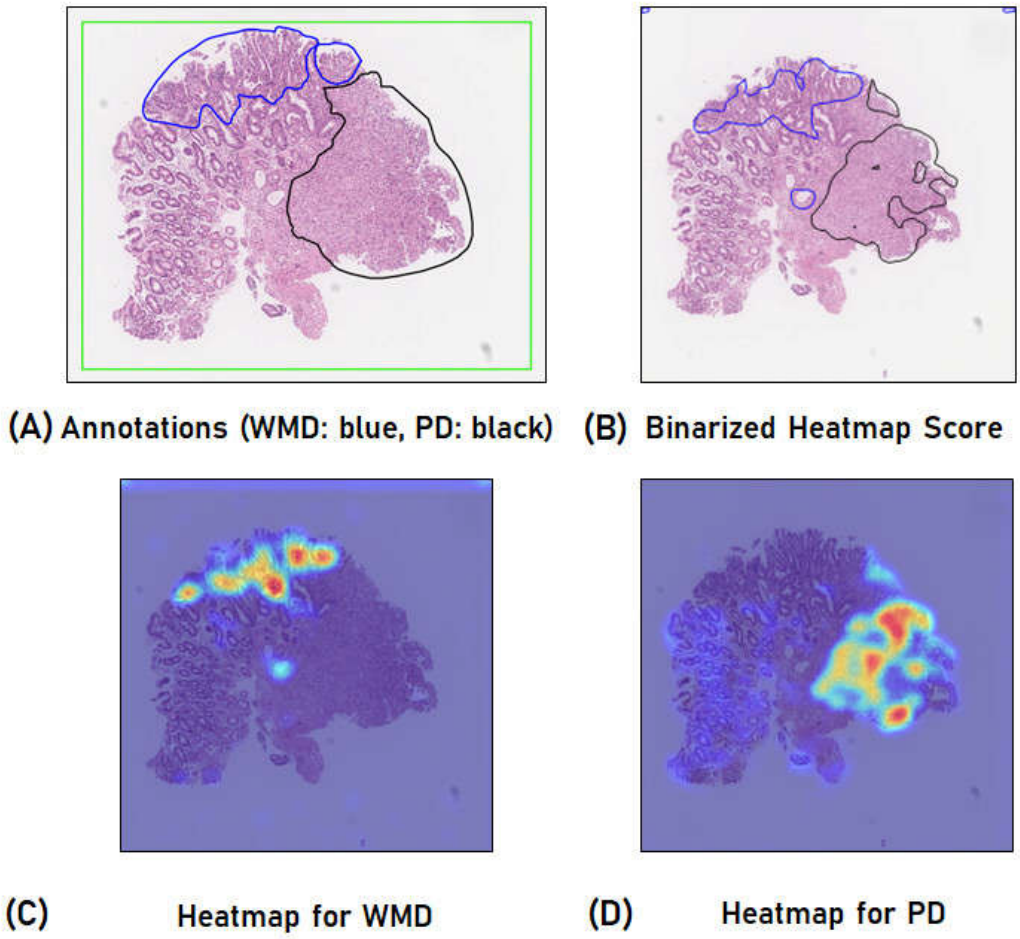

Figure 8. Drawing the inference for a mixed-type tissue. (**A**) Annotations of the mixed-type tissue (WMD and PD areas are indicated by blue and black contours, respectively). (**B**) The Grad-CAM score thresholding result for each class. (**C**) The Grad-CAM heatmap for the WMD class. (**D**) The Grad-CAM heatmap for the PD class.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 August 2022 doi:10.20944/preprints202208.0080.v1

Table 4. Grad-CAM performance, as indicated by the hit rate for each class.

	WMD	PD
Fold 1	0.859	0.780
Fold 2	0.917	0.847
Fold 3	0.886	0.770

For mixed-type tissues, (with both PD and WMD annotations), representative information is disclosed when both PD and WMD heatmaps are drawn. A pathologist's annotation of one mixed-type case is shown in Figure 8A. To compare the annotation and Grad-CAM heatmap, we used a threshold to binarize the heatmap (Figure 8B). The threshold was 20% of the maximum Grad-CAM score. Figure 8C, D show the original WMD and PD Grad-CAM heatmaps. Similarity is evident between the pathologist's annotation and binarized Grad-CAM heatmap.

4. Discussion

We developed a weakly supervised learning technique for tissue-level gastric cancer histological differentiation without the need for pixel-level annotation. The ResNet-50 binary classifier well-discriminated WMD and PD tissues. The data of five hospitals were visually interpretable and robust following threefold cross-validation.

Jang et al. [32] used deep learning to build a gastric cancer classification system. They divided WSIs into patches and used different classifiers to sequentially discriminate normal/tumor and differentiated/undifferentiated tumor patches. As each WSI was divided into small patches, a two-stage approach was necessary. The AUROC was 0.932 for differentiated/undifferentiated patch classification. However, the performance of the normal/tumor patch classifier should be considered when discussing differentiated/undifferentiated classification performance. Although the AUROC of the former classifier was high (0.993), it was not 1. Thus, when the two stages are considered together, the overall differentiated/undifferentiated patch classification performance is reduced.

For diagnosis, low magnification is used for a global overview, followed by high magnification to examine local features [33]. Our global reasoning process focuses on tissue-level features (analogous to low magnification). We expect that performance can be improved by adding other global (e.g., glandular architecture) and local (e.g., nuclear) features.

Valuable information is lost when slides are divided into small patches. In particular, geometric relationships between patches are discounted [34]. Another drawback is the need for pixel-level annotation by a pathologist. Our weakly supervised approach does not require such annotation, so the burden on experts is reduced. Unlike earlier approaches, we do not apply patching. We found that histological classification of gastric adenocarcinoma and visual interpretation were possible using only tissue-level (weak) annotations. In addition, our one-stage model is simpler than the previous model.

Chen et al. [21] used entire WSIs for training, but resized them to 0.2-fold the original size. The downscaled images were then "padded" to 21,500 × 21,500 pixels. In other words, the original WSI must be smaller than 107,500 × 107,500 pixels. We could not use this method because our WSIs were too large.

WSI-level weakly supervised learning consumes a large amount of GPU memory; patch-level learning ignores wide features and requires pathologist annotations. A tissue is larger than a patch and smaller than a WSI. Tissue-level weakly supervised learning does not consume much GPU memory, can be used to explore a wide range of histological features, and does not require annotations. Despite these advantages, few tissue-level deep learning studies have been conducted because tissue area extraction is manual [35]. However, due to developments in WSI foreground extraction, we can now automatically extract tissues from images, such that there are now no obstacles to tissue-level analysis.

Our downsizing method could be improved. We extracted 2,048 × 2,048-pixel images. Thus, each ROI was downscaled to some extent, and the microns (μm) per pixel varied.

 doi:10.20944/preprints202208.0080.v1

Also, image quality may be reduced by downscaling. Nevertheless, we obtained the global features despite these disadvantages.

The visual interpretability of the Grad-CAM heatmap is important. Grad-CAM highlights areas that strongly affect predictions, and showed excellent agreement with the pathologists. The heatmap also explains false predictions. The Grad-CAM heatmaps for some falsely predicted cases are shown in Supplementary Figure S1. Reactive atypia may have contributed to misclassification [36]. Glands that are inflamed become slightly deformed, which can lead to misclassification.

Hwang et al. [37] showed that PD gastric cancer components were associated with lymph node metastasis. Pathologist judgments of differentiation are subjective and time-consuming. Our method yields a Grad-CAM score in cases of poor differentiation. We plan to work with pathologists to understand how Grad-CAM scores vary according to the extent of differentiation. We expect that our method will accurately predict lymph node metastasis.

5. Conclusions

In general, whole WSIs are too large, while patches are too small, for pathological analysis. Tissue-level analysis is thus appropriate. We found that class-discriminative features were revealed by Grad-CAM visualization of deep learning classifications. Thus, tissue-level annotations suffice for deep learning of gastric adenocarcinoma histological differentiation. Such learning is applicable to many histopathological classification problems that require wide fields of view (such as normal/abnormal decision-making). Our work will aid the development of deep learning-based pathological diagnoses. Our weakly supervised method automatically visualizes the PD component of gastric cancer. If the algorithm can be improved to quantify PD components in WSIs, gastric cancer treatment could be optimized. Also, lymph node metastasis and the need for surgery could be predicted.

Supplementary Materials: Figure S1: The Grad-CAM heatmaps for each class of false predictions.

Author Contributions: Conceptualization, D.-W.H.; methodology, D.-W.H.; software, D.M.; validation, W.-S.R. and S.-H.S.; formal analysis, D.-W.H.; investigation, D.-W.H. and D.-H.L.; resources, H. Kim, H.-C.S., W.-A.L., and M.-S.C.; data curation, H. Kang; writing—original draft preparation, D.-W.H.; writing—review and editing, S.B., W.-S.R., H. Kang and S.-H.S.; visualization, C.-H.K. and S.B.; supervision, H.-K.C.; project administration, D.K.; funding acquisition, N.-H.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Ministry of Trade, Industry, and Energy (MOTIE), South Korea, under the ''Regional Specialized Industry Development Program (R&D, P0002072)'' supervised by the Korea Institute for Advancement of Technology (KIAT) under Grant P0002072, and grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (Grant No: HI21C0977).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and was approved by the Institutional Review Boards of CHA Bundang Medical Center (approval no. CHAMC 2021-09-021-002), Keimyung University Dongsan Medical Center (approval no. DSMC 2021-08-112), Dankook University Hospital and Yonsei University Hospital (approval no. 4-2021-0029), and Ewha Womans University Medical Center (approval no. 2021-0-013-001).

Informed Consent Statement: Not applicable.

Data Availability Statement: The data of the AI-ALPA consortium are not publicly available for privacy and ethical reasons.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. A.; 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 J. Clin.* **2021**, 71, 209–249, doi:10.3322/caac.21660.
2. Japanese Gastric Cancer Association. Japanese Gastric Cancer Treatment Guidelines 2018 (5th Edition). *Gastric Cancer* **2021**, 24, 1–21, doi:10.1007/s10120-020-01042-y.
3. Chen, J.-N.; Wang, Q.-W.; Zhang, Q.-W.; Tang, Z.-R.; Li, X.-B. Poorly Differentiated Is More Significant than Signet Ring Cell Component for Lymph Node Metastasis in Mixed-Type Early Gastric Cancer: A Retrospective Study from a Large-Volume Hospital. *Surg. Endosc.* **2021**, 35, 1558–1565, doi:10.1007/s00464-020-07532-5.
4. Wang, S.; Yang, D.; Rong, R.; Zhan, X.; Xiao, G. Pathology Image Analysis Using Segmentation Deep Learning Algorithms. *The American Journal of Pathology* **2019**, 189, 1686-1698, doi:10.1016/j.ajpath.2019.05.007
5. Kanavati, F.; Tsuneki, M. A Deep Learning Model for Gastric Diffuse-Type Adenocarcinoma Classification in Whole Slide Images. *Sci. Rep.* **2021**, 11, 20486, doi:10.1038/s41598-021-99940-3.
6. Kim, C.-H.; Bhattacharjee, S.; Prakash, D.; Kang, S.; Cho, N.-H.; Kim, H.-C.; Choi, H.-K. Artificial Intelligence Techniques for Prostate Cancer Detection through Dual-Channel Tissue Feature Engineering. *Cancers* **2021**, 13, 1524, doi:10.3390/cancers13071524.
7. Cho, B.-J.; Kim, J.-W.; Park, J.; Kwon, G.-Y.; Hong, M.; Jang, S.-H.; Bang, H.; Kim, G.; Park, S.-T. Automated Diagnosis of Cervical Intraepithelial Neoplasia in Histology Images via Deep Learning. *Diagnostics* **2022**, 12, 548, doi:10.3390/diagnostics12020548.
8. Iizuka, O.; Kanavati, F.; Kato, K.; Rambeau, M.; Arihiro, K.; Tsuneki, M. Deep Learning Models for Histopathological Classification of Gastric and Colonic Epithelial Tumours. *Sci. Rep.* **2020**, 10, 1504, doi:10.1038/s41598-020-58467-9.
9. Khened, M.; Kori, A.; Rajkumar, H.; Krishnamurthi, G.; Srinivasan, B. A Generalized Deep Learning Framework for Whole-Slide Image Segmentation and Analysis. *Sci. Rep.* **2021**, 11, 11579, doi:10.1038/s41598-021-90444-8.
10. Janowczyk, A.; Madabhushi, A. Deep Learning for Digital Pathology Image Analysis: A Comprehensive Tutorial with Selected Use Cases. *J. Pathol. Inform.* **2016**, 7, 29, doi:10.4103/2153-3539.186902.
11. Bhattacharjee, S.; Ikromjanov, K.; Carole, K.S.; Madusanka, N.; Cho, N.-H.; Hwang, Y.-B.; Sumon, R.I.; Kim, H.-C.; Choi, H.-K. Cluster Analysis of Cell Nuclei in H&E-Stained Histological Sections of Prostate Cancer and Classification Based on Traditional and Modern Artificial Intelligence Techniques. *Diagnostics* **2021**, 12, 15, doi:10.3390/diagnostics12010015.
12. Geessink, O.; Baidoshvili, A.; Klaase, J.; Ehteshami Bejnordi, B.; Litjens, G.; van Pelt, G.; Mesker, W.; Nagtegaal, I.; Ciompi, F.; van der Laak, J. Computer aided quantification of intratumoral stroma yields an independent prognosticator in rectal cancer. *Cell. Oncol.* **2019**, *42*, 331-341, doi: 10.1007/s13402-019-00429-z
13. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. IEEE International Conference on Computer Vision (ICCV) 2017, Venice, Italy, 22-29 October 2017; pp. 618-626.
14. Lu, M.Y.; Williamson, D.F.K.; Chen, T.Y.; Chen, R.J.; Barbieri, M.; Mahmood, F. Data-Efficient and Weakly Supervised Computational Pathology on Whole-Slide Images. *Nat. Biomed. Eng.* **2021**, 5, 555–570, doi:10.1038/s41551-020-00682-w.
15. Lu, M.Y.; Zhao, M.; Shady, M.; Lipkova, J.; Chen, T.Y.; Williamson, D.F.K.; Mahmood, F. Deep Learning-Based Computational Pathology Predicts Origins for Cancers of Unknown Primary. *Nature* **2021**, 594, 106–110, doi:10.1038/s41586-021-03512-4.
16. Chen, Z.; Zhang, J.; Che, S.; Huang, J.; Han, X.; Yuan, Y. Diagnose Like A Pathologist: Weakly-Supervised Pathologist-Tree Network for Slide-Level Immunohistochemical Scoring. *AAAI* **2021**, 35, 47-54.
17. Song, Z.; Zou, S.; Zhou, W.; Huang, Y.; Shao, L.; Yuan, J.; Gou, X.; Jin, W.; Wang, Z.; Chen, X.; et al. Clinically Applicable Histopathological Diagnosis System for Gastric Cancer Detection Using Deep Learning. *Nat. Commun.* **2020**, 11, 4294, doi:10.1038/s41467-020-18147-8.
18. Wang, X.; Chen, H.; Gan, C.; Lin, H.; Dou, Q.; Tsougenis, E.; Huang, Q.; Cai, M.; Heng, P.-A. Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. *IEEE Trans. Cybern.* **2020**, 50, 3950–3962, doi:10.1109/TCYB.2019.2935141.
19. Li, B.; Li, Y.; Eliceiri, K.W. 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) 2021, Virtual, 19-25 June 2021; pp. 14313-14323.
20. 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. *Sci. Rep.* **2020**, 10, 9297, doi:10.1038/s41598-020-66333-x.
21. Chen, C.-L.; Chen, C.-C.; Yu, W.-H.; Chen, S.-H.; Chang, Y.-C.; Hsu, T.-I.; Hsiao, M.; Yeh, C.-Y.; Chen, C.-Y. An Annotation-Free Whole-Slide Training Approach to Pathological Classification of Lung Cancer Types Using Deep Learning. *Nat. Commun.* **2021**, 12, 1193, doi:10.1038/s41467-021-21467-y.
22. Gadermayr, M.; Tschuchnig, M. Multiple Instance Learning for Digital Pathology: A Review on the State-of-the-Art, Limitations & Future Potential. Available online: https://arxiv.org/abs/2206.04425/ (accessed on June 25, 2022).
23. Bankhead, P.; Loughrey, M.B.; Fernández, J.A.; Dombrowski, Y.; McArt, D.G.; Dunne, P.D.; McQuaid, S.; Gray, R.T.; Murray, L.J.; Coleman, H.G.; et al. QuPath: Open Source Software for Digital Pathology Image Analysis. *Sci. Rep.* **2017**, 7, 16878, doi:10.1038/s41598-017-17204-5.
24. Goode, A.; Gilbert, B.; Harkes, J.; Jukic, D.; Satyanarayanan, M. OpenSlide: A Vendor-Neutral Software Foundation for Digital Pathology. *J. Pathol. Inform.* **2013**, 4, 27, doi:10.4103/2153-3539.119005.

25. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016, Las Vegas, USA, 26-1 June-July 2016; pp. 770-778.
26. Mahmood, A.; Ospina, A.G.; Bennamoun, M.; An, S.; Sohel, F.; Boussaid, F.; Hovey, R.; Fisher, R.B.; Kendrick, G.A. Automatic Hierarchical Classification of Kelps Using Deep Residual Features. *Sensors* **2020**, 20, 447, doi:10.3390/s20020447.
27. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada, 28-14 December 2019; pp. 1-12.
28. Bug, D.; Feuerhake, F.; Merhof, D. Foreground Extraction for Histopathological Whole Slide Imaging. In Bildverarbeitung für die Medizin 2015; Handels, H., Deserno, T.M., Meinzer, H.-P., Tolxdorff, T., Eds.; Informatik aktuell; Springer Berlin Heidelberg: Berlin, Heidelberg, 2015; pp. 419–424 ISBN 978-3-662-46223-2.
29. Bandi, P.; van de Loo, R.; Intezar, M.; Geijs, D.; Ciompi, F.; van Ginneken, B.; van der Laak, J.; Litjens, G. Comparison of different methods for tissue segmentation in histopathological whole-slide images. IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017) 2017, Melbourne, Australia, 18-21 April 2017; pp. 591-595.
30. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biom.* **1988**, 44, 837, doi:10.2307/2531595.
31. Saporta, A.; Gui, X.; Agrawal, A.; Pareek, A.; Truong, S.Q.; Nguyen, C.D.; Ngo, V.-D.; Seekins, J.; Blankenberg, F.G.; Ng, A.Y.; et al. Deep Learning Saliency Maps Do Not Accurately Highlight Diagnostically Relevant Regions for Medical Image Interpretation. *medRxiv* **2021**, 2021.02.28.21252634, doi:10.1101/2021.02.28.21252634.
32. Jang, H.-J.; Song, I.-H.; Lee, S.-H. Deep Learning for Automatic Subclassification of Gastric Carcinoma Using Whole-Slide Histopathology Images. *Cancers* **2021**, 13, 3811, doi:10.3390/cancers13153811.
33. Li, J.; Li, W.; Sisk, A.; Ye, H.; Wallace, W.D.; Speier, W.; Arnold, C.W. A Multi-Resolution Model for Histopathology Image Classification and Localization with Multiple Instance Learning. *Comput. Biol. Med.* **2021**, 131, 104253, doi:10.1016/j.compbiomed.2021.104253.
34. Chakrabarti, R.; Arandelovic, O. A whole-slide is greater than the sum of its...patches. Available online: http://hdl.handle.net/10023/25023/ (accessed on June 25, 2022).
35. Yoshida, H.; Yamashita, Y.; Shimazu, T.; Cosatto, E.; Kiyuna, T.; Taniguchi, H.; Sekine, S.; Ochiai, A. Automated Histological Classification of Whole Slide Images of Colorectal Biopsy Specimens. *Oncotarget* **2017**, 8, 90719–90729, doi:10.18632/oncotarget.21819.
36. Park, J.; Jang, B.-G.; Kim, Y.-W.; Park, H.; Kim, B.; Kim, M.-J.; Ko, H.; Gwak, J.-M.; Lee, E.-J.; Chung, Y.-R.; Kim, K.; Myung, J.-K.; Park, J.-H.; Choi, D.-Y.; Jung, C.-W.; Park, B.-H.; Jung, K.-H.; Kim, D.-I. A Prospective Validation and Observer Performance Study of a Deep Learning Algorithm for Pathologic Diagnosis of Gastric Tumors in Endoscopic Biopsies. *Clin. Cancer Res.* **2021**, 27, 719-728, doi:10.1158/1078-0432.CCR-20-3159
37. Hwang, C.-S.; Ahn, S.; Lee, B.-E.; Lee, S.-J.; Kim, A.; Choi, C.I.; Kim, D.H.; Jeon, T.-Y.; Kim, G.H.; Song, G.A.; et al. Risk of Lymph Node Metastasis in Mixed-Type Early Gastric Cancer Determined by the Extent of the Poorly Differentiated Component. *World J Gastroenterol.* **2016**, 22, 4020, doi:10.3748/wjg.v22.i15.4020.