

Article

Deep Learning for the Pathologic Diagnosis of Hepatocellular Carcinoma, Cholangiocarcinoma, and Metastatic Colorectal Cancer

Hyun-Jong Jang ¹, Jai-Hyang Go ², Younghoon Kim ³ and Sung Hak Lee ^{3,*}

¹ Department of Physiology, CMC Institute for Basic Medical Science, College of Medicine, The Catholic University of Korea, Seoul 06591, Republic of Korea; hjjang@catholic.ac.kr

² Department of Pathology, Dankook University College of Medicine, Cheonan 31116, Republic of Korea; cyjy555@daum.net

³ Department of Hospital Pathology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul 06591, Republic of Korea; glabella85@gmail.com

* Correspondence: hakjang@catholic.ac.kr; Tel.: +82-2-2258-1617

Simple Summary: The pathologic diagnosis of primary and secondary liver cancers can often be difficult. Artificial intelligence (AI) presents potential solutions to these difficulties by aiding in the histopathological diagnosis of tumors using digital whole slide images (WSIs). We developed an AI diagnostic assistant using a deep learning model for distinguishing hepatocellular carcinoma, cholangiocarcinoma, and metastatic colorectal cancer using WSIs. Overall, the classifiers were highly accurate, showing significant potential for improving liver cancer diagnosis and advancing precision medicine. However, additional research is required to further refine and validate these promising tools.

Abstract: Diagnosing primary liver cancers, particularly hepatocellular carcinoma (HCC) and cholangiocarcinoma (CC), is a challenging and labor-intensive process, even for experts, and secondary liver cancers further complicate the diagnosis. Artificial intelligence (AI) offers promising solutions to these diagnostic challenges by facilitating the histopathological classification of tumors using digital whole slide images (WSIs). This study aimed to develop a deep learning model for distinguishing HCC, CC, and metastatic colorectal cancer (mCRC) using histopathological images and to discuss its clinical implications. The WSIs from HCC, CC, and mCRC were used to train the classifiers. For normal/tumor classification, the areas under the curve (AUCs) were 0.989, 0.988, and 0.991 for HCC, CC, and mCRC, respectively. Using proper tumor tissues, the HCC/other cancer type classifier was trained to effectively distinguish HCC from CC and mCRC, with a concatenated AUC of 0.998. Subsequently, the CC/mCRC classifier differentiated CC from mCRC with a concatenated AUC of 0.995. However, testing on an external dataset revealed that the HCC/other cancer type classifier underperformed with an AUC of 0.745. After combining the original training datasets with external datasets and retraining, the classification drastically improved, all achieving AUCs of 1.000. Although these results are promising and offer crucial insights into liver cancer, further research is required for model refinement and validation.

Keywords: deep learning; pathology; hepatocellular carcinoma; cholangiocarcinoma; colorectal cancer



Citation: Jang, H.-J.; Go, J.-H.; Kim, Y.; Lee, S.H. Deep Learning for the Pathologic Diagnosis of Hepatocellular Carcinoma, Cholangiocarcinoma, and Metastatic Colorectal Cancer. *Cancers* **2023**, *15*, 5389. <https://doi.org/10.3390/cancers15225389>

Received: 4 October 2023

Revised: 1 November 2023

Accepted: 9 November 2023

Published: 13 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Primary liver cancer is the sixth-most common malignancy and the third leading cause of cancer-related deaths worldwide [1]. Hepatocellular carcinoma (HCC) and cholangiocarcinoma (CC) are the two primary types of liver malignancy. HCC is a malignant epithelial tumor that displays features of hepatocellular differentiation [2]. HCC has become one of the most common cancers worldwide, with a growing incidence rate over the past three decades [3]. While great achievements have been made in the histopathological classification of HCC, the diagnostic process is labor-intensive and requires pathologists

to meticulously evaluate pathological images, with potential inter- and intra-observer variations [4,5]. In addition, significant intratumoral heterogeneity in HCC tissues may present a challenge when attempting to analyze morphological features solely through visual examination [6]. Along with HCC, CC is the second-most common primary liver cancer, comprising 15% of primary hepatic malignancies [7]. CC is a malignant tumor that exhibits biliary epithelial features [8]. Most CCs are diagnosed histopathologically as adenocarcinomas.

The diagnostic distinction between HCC and CC is of significant importance in terms of therapeutic implications. For instance, although orthotopic liver transplantation is widely accepted as a therapeutic approach for patients with HCC, it is contraindicated in patients with CC [9]. However, diagnostic differentiation between the two can occasionally be challenging, even for highly specialized pathologists [10,11].

Secondary liver cancer originates from a malignant tumor outside the liver that subsequently metastasizes. Secondary liver tumors predominantly consist of carcinomas, followed by melanoma, sarcoma, and lymphoma, in that order of frequency. The primary cancers that most frequently metastasize to the liver include colorectal, breast, lung, and gastric carcinomas [12]. About 20–30% of patients diagnosed with colorectal cancer (CRC) already have metastatic manifestations [13,14]. The liver is the most common site of CRC metastasis. When identifying liver metastases from CRC, the main differential diagnosis is CC, as most CRCs are predominantly adenocarcinomas [15].

Recently, advances in artificial intelligence (AI) have resulted in the development of highly efficient algorithms for various purposes in the medical field including pathology, with a subset of devices making their way into commercialization for clinics [16–18]. A representative task is the histopathological classification of tumors, which is essential for predicting outcomes and determining treatment [19,20]. More and more pathology laboratories are transitioning to the regular use of digital slides in the form of whole slide images (WSIs) for their daily diagnostic workflows [21–23]. The transformation from traditional microscopy to WSIs has facilitated the incorporation of AI support systems into pathology, thereby enhancing the efficiency, accuracy, and consistency. Consequently, this has enabled the emergence of cutting-edge techniques via deep learning (DL) [24–27]. Recent studies have confirmed the efficacy of AI in pathology for identifying tumors in several organs, including the lungs, stomach, and breasts [19,20,28–30]. However, the application of DL-based histopathological analysis for the diagnosis of HCC and CC has rarely been documented.

In this study, we aimed to develop a deep learning-based, fully automated model for the differential diagnosis of HCC, CC, and metastatic CRC (mCRC) using histopathological images. The potential utility of our results in the clinical setting is also discussed.

2. Materials and Methods

2.1. Datasets

Formalin-fixed paraffin-embedded (FFPE) slides of HCC and CC samples were obtained from The Cancer Genome Atlas (TCGA). After a basic quality review, 366 and 39 WSIs were selected for HCC (TCGA-LIHC) and CC (TCGA-CHOL), respectively. As the number of WSIs was too different between TCGA-LIHC and TCGA-CHOL to train a DL-based classifier, we collected 156 CC WSIs from Dankook University Hospital (DKUH-CHOL). Furthermore, 179 mCRC WSIs were collected from Dankook University Hospital (DKUH-META). For the external validation of the trained classifier, 31 and 29 WSIs of HCC and CC, respectively, were collected from Seoul St. Mary's Hospital (SSMH dataset). All tissues included in this study were stained with hematoxylin and eosin (H&E). Patient characteristics and clinicopathological features of all cohorts are summarized in Table S1.

2.2. Deep Learning Models

Because the WSI was too large to be classified by a DL model as a whole, we trained the classifiers on small tissue image patches of 360×360 pixels at $20\times$ magnification.

To classify different cancer types, comparisons should only be made for cancer tissues. Therefore, cancer tissue image patches should be collected before classifying the cancer types. We sequentially applied two different DL-based tissue classifiers to collect cancer tissue image patches (Figure 1a).

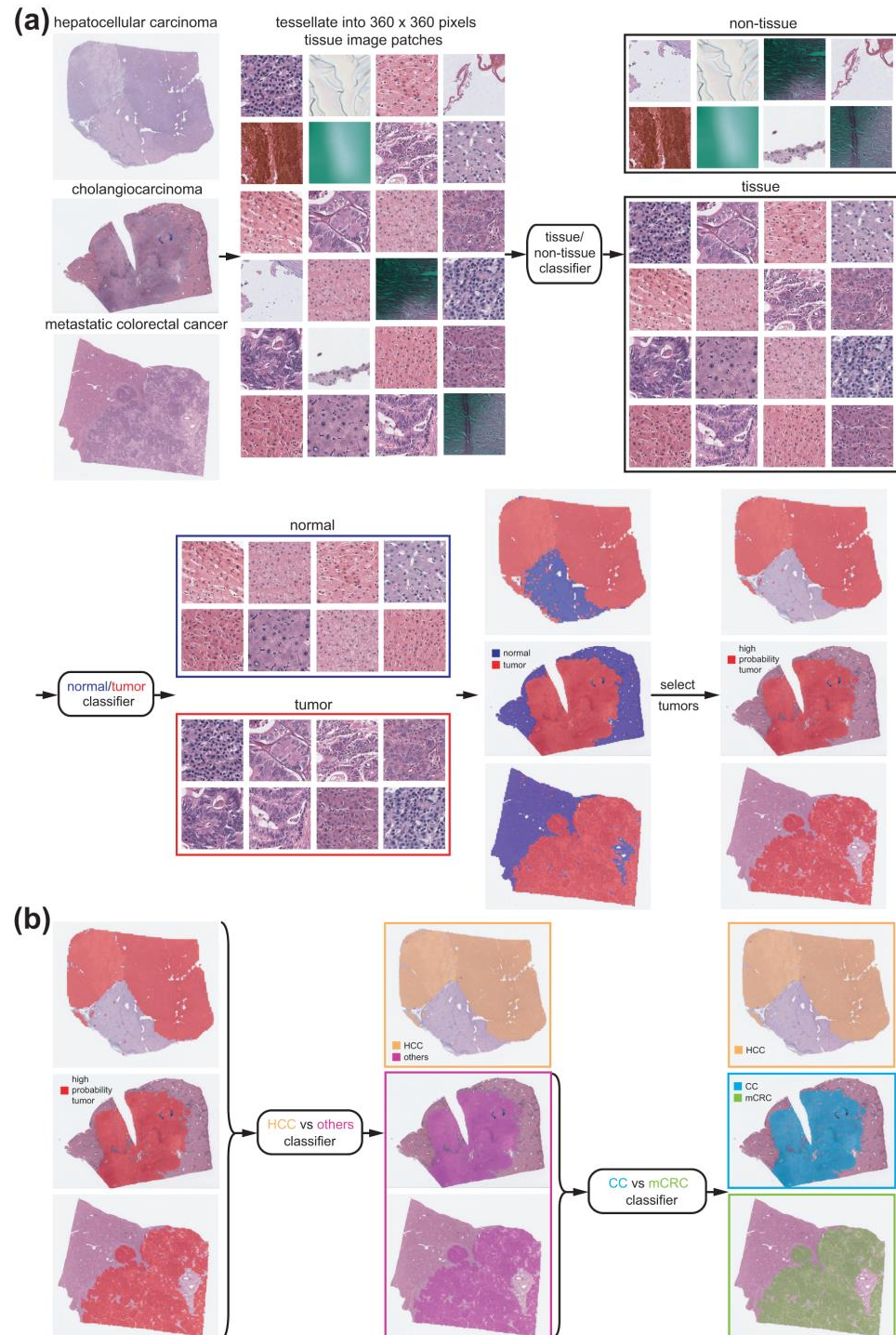


Figure 1. Procedure for the discrimination of different cancer types. **(a)** Tissue/non-tissue and normal/tumor classifiers are sequentially applied to separate proper tumor tissues from the whole slide images. **(b)** The first classifier discriminates hepatocellular carcinoma (HCC) from other cancer types. Then, the second classifier discriminates cholangiocarcinoma (CC) from metastatic colorectal cancer (mCRC).

First, proper tissue images in WSIs should be discriminated from multiple artifacts, including air bubbles, compression artifacts, out-of-focus blurring, pen markings, tissue folding, and white backgrounds. We reused a tissue/non-tissue classifier from our previous study to eliminate these artifacts [31]. A normal/tumor classifier was then trained to discriminate cancerous tissues from normal tissues. Three pathologists annotated the normal/tumor tissue regions in the WSIs of HCC, CC, and mCRC (Figure 2, left panels).

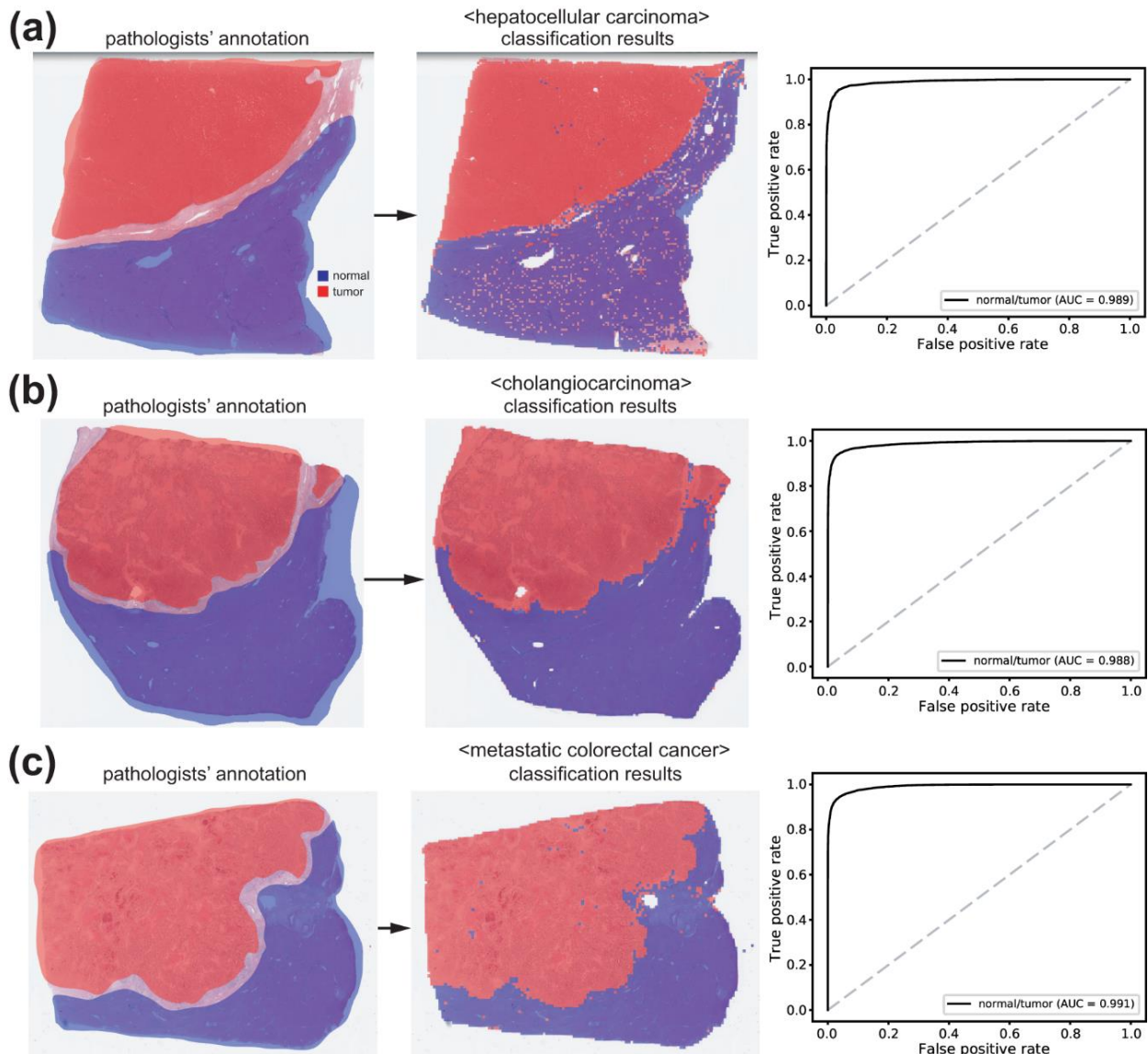


Figure 2. Results for normal/tumor classification. The patch-level classification results for hepatocellular carcinoma (a), cholangiocarcinoma (b), and metastatic colorectal cancer (c) were demonstrated with representative tissue images. Left panels: pathologists' annotation. Middle panels: classification results by the normal/tumor classifier. Right panels: the receiver operating characteristic curves for the patch-level classification results. AUC: area under the curve.

Normal and tumor tissue image patches were collected from all three cancer types to train a classifier that could discriminate normal/tumor tissues from the three cancer types simultaneously. We randomly selected 90% of normal and tumor tissue image patches to train the classifier and evaluated the performance of the classifier on the remaining 10%. We adopted three popular convolutional neural network (CNN) models to train the normal/tumor classifier: AlexNet, ResNet-50, and Inception-v3. The TensorFlow DL library (version 1.15) was used to train each DL model (<http://tensorflow.org>, accessed

on 23 February 2023). Before training and testing the classifiers, the tissue images were color-normalized. Data augmentation techniques, including random horizontal/vertical flipping and random rotation by 90° , were applied to the tissue image patches during the training of the classifiers.

For the classification of cancer types, we collected tumor patches with a tumor probability higher than 0.9 to include patches with prominent tumor features. Since we trained the models based on slide-level diagnoses, all tumor tissue image patches from a given slide inherit the same label from the slide-level diagnosis. We implemented slide-level 5-fold cross-validation for cancer type classification. In this scheme, five different sets of sides were used for the training, validation, and testing with proportions of 60%, 20%, and 20%, respectively. The validation datasets were used to evaluate the performance of the classifiers during training. The classifiers with the highest performances for the validation datasets were then used for the testing datasets. We implemented a two-step approach to fully discriminate between HCC, CC, and mCRC. First, a classifier was trained to discriminate HCC from other cancer types (CC and mCRC) (Figure 1b, left). Another classifier was trained to discriminate between CC and mCRC (Figure 1b, right).

Six computer systems equipped with Intel Core i9-12900K (Intel Corporation, Santa Clara, CA, USA) processors and dual NVIDIA RTX 3090 GPUs (NVIDIA corporation, Santa Clara, CA, USA) were used to train and test the DL models.

2.3. Visualization and Statistics

The classification results were overlapped on the WSIs with color-coded heatmaps to visualize the distribution of the different tissue types. The averages of the patch classification results were used to obtain the slide-level classification results. To demonstrate the performance of each classifier, receiver operating characteristic (ROC) curves and areas under the curves (AUCs) for the ROC curves are presented. In the case of cancer-type classifiers, ROC curves for the folds with the lowest and highest AUCs and for the concatenated results of all five folds are provided for a more precise evaluation of the classification performance on the 5-fold cross-validated datasets. For the concatenated results of all five folds, 95% confidence intervals (CIs) are also presented. Accuracy, sensitivity, specificity, and F1 score were calculated with cutoff values yielding a maximal Youden index (sensitivity + specificity − 1).

To compare the ROC curves, the Venkatraman's permutation test with 1000 iterations was applied [32]. Statistical significance was set at $p < 0.05$.

3. Results

3.1. Normal/Tumor Classification

Normal/tumor tissue image patches were collected based on annotations from pathologists. Tissue images from HCC, CC, and mCRC were mixed to train a single classifier for normal/tumor discrimination of all tissue types. AlexNet, ResNet-50, and Inception-v3 models were used to train the normal/tumor classifiers. Overall, Inception-v3 exhibited the best classification performance (Supplementary Table S2). Therefore, we adopted the Inception-v3 model for further analysis in the present study. The normal/tumor classification results obtained using the Inception-v3 model are presented in Figure 2. When the classifier was applied to the WSIs for HCC, CC, and mCRC in the test sets, the AUCs were 0.989, 0.988, and 0.991, respectively. Although the classification performance was better for mCRC than for HCC or CC (both $p < 0.05$, Venkatraman's permutation test), we concluded that the classification performance of the classifier was sufficient for collecting tumor tissues from all cancer types based on a qualitative review of the classification results (Figure 2, middle panels).

3.2. Classification of HCC/Other Cancer Types

Using a normal/tumor classifier, high-probability tumor tissues were collected to build a classifier to discriminate HCC from other cancer types (CC and mCRC). We adopted

five-fold cross-validation for cancer type classification. Training was performed at least four times for each fold and the classifiers with the best AUC for the testing datasets were adopted to present the results in Figure 3. The average number of tissue image patches used for training the classifiers is summarized in Supplementary Table S3. As shown in the upper panels of Figure 3, most WSIs of HCC were clearly discriminated from those of CC and mCRC. The AUCs were 0.997 and 0.999 for fold-changes with the lowest and highest AUCs, respectively. The AUC for the concatenated results was 0.998 (95% CI, 0.997–0.999). The accuracy, sensitivity, specificity, and F1 score of the classifier are listed in Supplementary Table S4.

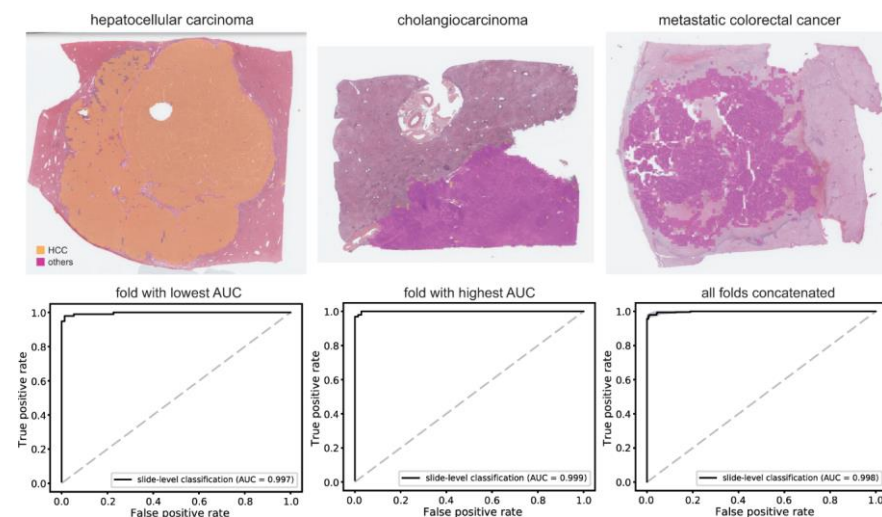


Figure 3. Classification results between hepatocellular carcinoma and other cancer types (cholangiocarcinoma and metastatic colorectal cancer). Upper panels: representative tissue images of hepatocellular carcinoma, cholangiocarcinoma, and metastatic colorectal cancers that were correctly classified by the classifier. Lower panels: the receiver operating characteristic curves of slide-level classification results for folds with the lowest and highest area under the curve (AUC) and concatenated results of all 5-folds. HCC: hepatocellular carcinoma.

3.3. CC/mCRC Classification

Subsequently, a classifier was developed to discriminate between CC and mCRC. The classification results are shown in Figure 4. Representative WSIs of clear CC, clear mCRC, and confusing cases with mixed classification results are shown in the upper panel. The AUCs were 0.992 and 0.998 for fold-changes with the lowest and highest AUCs, respectively. The AUC for the concatenated results was 0.995 (95% CI, 0.992–0.998). The accuracy, sensitivity, specificity, and F1 score of the classifier are listed in Supplementary Table S4.

3.4. Performance on an External Dataset

To test whether the trained DL model performed well on an external dataset, we tested hepatocellular carcinoma/other cancer type classifiers on the SSMH dataset. The performance was poor with an AUC of 0.745, suggesting poor generalizability of the classifier to external datasets (Supplementary Figure S1). We split the SSMH dataset into five-folds and mixed the data with TCGA and DKUH data to retrain the classifier. When we retrained the classifier with mixed data from TCGA, DKUH, and SSMH tissue images, the classification results for the SSMH dataset significantly improved (Figure 5). The AUCs were all 1.000 for every fold, suggesting perfect classification. The classification performance of the new classifier on the original TCGA+DKUH datasets did not significantly improve despite the enlarged training datasets ($p = 0.329$ through Venkatraman's permutation test, Supplementary Figure S2).

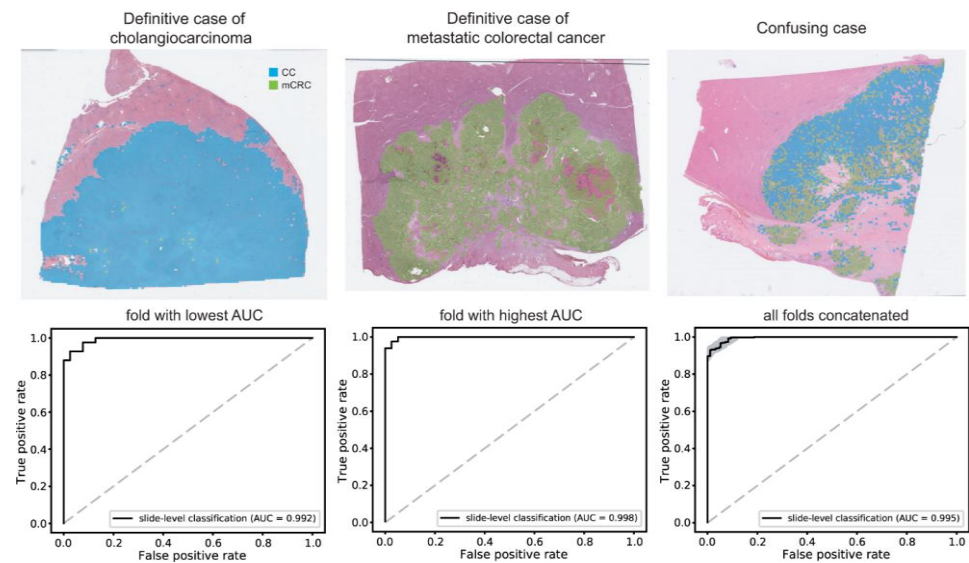


Figure 4. Classification results between cholangiocarcinoma and metastatic colorectal cancer. Upper panels: the representative whole slide images of clear cholangiocarcinoma, clear metastatic colorectal cancer, and confusing case with mixed classification results. Lower panels: the receiver operating characteristic curves of slide-level classification results for folds with the lowest and highest area under the curve (AUC) and concatenated results of all 5-folds. CC: cholangiocarcinoma, mCRC: metastatic colorectal cancer.

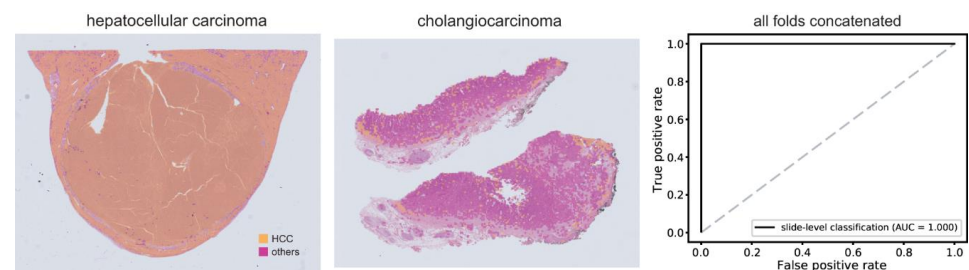


Figure 5. Classification results for the Seoul St. Mary's Hospital (SSMH) dataset by a classifier trained with mixed datasets. Representative whole slide images of hepatocellular carcinoma (left) and cholangiocarcinoma (middle) from the SSMH dataset. The receiver operating characteristic curve of slide-level classification results for the concatenated results of all 5-folds is presented in the right panel. AUC: area under the curve, HCC: hepatocellular carcinoma.

4. Discussion

In the present study, we adopted a two-step approach to discriminate between HCC, CC, and mCRC. Although DL-based classifiers can perform three classes of classification tasks directly, we split the task into two steps based on the dataset size. As summarized in Supplementary Table S3, more tissue imaging data are available for HCC. In general, DL-based classifiers do not perform well when the amount of training data is severely unbalanced [20]. As the data size can be balanced when CC and mCRC are mixed into a class, we first trained a classifier to discriminate between HCC and other cancer types. Furthermore, the second classifier for discriminating between CC and mCRC can also be trained on balanced datasets because the numbers of tissue images were similar for CC and mCRC. The overall AUCs for the two classifiers were 0.998 and 0.995, respectively. The higher AUC for the first classifier suggests that discriminating between HCC and other cancer types is a more obvious task than discriminating between CC and mCRC. Alternatively, it is a mere reflection of the larger training dataset sizes for the first classifier. We plan to collect more data for CC and mCRC to test whether the classification performance can be improved with more data.

While the discrimination performance showed promise, there were instances of misclassified cases and confusion due to mixed classification results, as illustrated in Figure 4. Enhancing the performance for challenging cases can be achieved through additional training methodologies, such as hard negative mining [33]. Additionally, a recent study has shown that a transformer network can enhance classification performance in comparison to a CNN [34]. These techniques hold the potential to enhance the performance of the HCC-CC-mCRC classifier for clinical applications.

The generalizability of DL-based classifiers is an important issue for the wide application of DL models. Unfortunately, our previous studies have shown that classifiers for the discrimination of various molecular traits from tissue images based on TCGA datasets did not perform well on Korean datasets [35–38]. In the present study, although we attempted to enhance the generalizability using color normalization and data augmentation techniques, the performance of the first classifier was poor on the SSMH dataset. Because the tissue image data of HCC came exclusively from the TCGA dataset for training the classifier, the results also indicate that the cancer-type classifiers trained on TCGA datasets show poor generalizability for Korean data. Overfitting on the training dataset can account for this lack of generalizability. Nevertheless, we made efforts to mitigate overfitting by implementing early stopping during training and utilizing validation datasets. Thus, the poor generalizability might not solely stem from overfitting. We expect that the poor generalizability originates mainly from differences in tissue characteristics. The quality of H&E-stained tissue slides can vary depending on tissue preparation and staining procedures, including fixation methods, cutting methods, dye concentration, and staining time [39]. Furthermore, differences in slide scanners and scanning settings can affect the quality of the scanned tissue slides. Finally, ethnic differences in the datasets may have resulted in poor generalizability. When we mixed the SSMH data into the original training data to train a new classifier, the SSMH dataset achieved a perfect AUC of 1.000. To enhance the generalizability of DL-based tissue classifiers, it is important that the classifier is exposed to a variety of data. Therefore, it is necessary to collect larger datasets from multiple institutes [40]. We are currently conducting a multi-center research project that collects and annotates the representative pathology WSIs for major cancers including liver and colorectal carcinomas (2021–2025, <https://www.codipai.org/>, accessed on 23 February 2023), and plan to implement a large-scale validation study using these cohorts in the near future.

The pathological evaluation of tissue samples is a key step in the differential diagnosis of HCC, CC, and mCRC. Nevertheless, even with histological analysis, distinguishing between HCC, CCA, and mCRC can be challenging. In most cases, HCC is suspected or directly identified on H&E-stained sections. HCC is characterized by the presence of sheets of polygonal cells with cytological atypia and architectural abnormalities such as thickened hepatic plates. In contrast, CC consists of cells that exhibit diverse structural arrangements, such as glandular and solid growth patterns. CCs induce a marked desmoplastic fibrotic reaction [8]. However, CCs do not always show identifiable tubular structures; rather, tumor cells often form solid sheets that mimic HCC. When metastasizing to the liver, mCRC also has morphological features similar to those of primary liver cancer and may display characteristics of desmoplasia or fibrotic changes [12].

For time-sensitive tasks such as histopathologic tumor classification, an initial diagnostic impression needs to be promptly established from the examination of routine H&E-stained slides. Several ancillary tests, including immunohistochemistry, fluorescence in situ hybridization, and molecular assays, depend on this preliminary determination. Therefore, our deep learning algorithm has the potential to be used as a screening tool for the diagnosis of HCC, CC, and mCRC. Moreover, apart from simply diagnosing CRC from pathology WSIs, the results of molecular testing such as *K-RAS*, *TP53*, and *BRAF* mutations, which are crucial for determining the treatment strategy for CRC patients, can be predicted directly from the H&E-stained tissue slides [41]. In addition, machine learning and AI algorithms can incorporate genomic data including gene expression profiling with WSI datasets to develop predictive models for cancer classification, potentially enhancing optimized treatment decisions [42,43].

Recently, the rapid growth of the internet of things (IoT) and cloud computing has tremendous prospects in the field of surgical pathology [44]. With the integration of DL technologies with IoT and cloud computing, it is possible to develop a portable AI-assisted pathology diagnosis platform and increase the accuracy of diagnosis including primary liver cancers by using it as an initial diagnostic tool.

Since the approval of WSIs as primary diagnostic materials [45], many pathology laboratories have adopted slide scanners. The era of digital pathology has enabled the accumulation of digital tissue image archives for training various DL models. The discrimination of cancer tissue subtypes is the most basic task in cancer diagnosis. Recently, DL has been widely used for tissue subtyping in various cancers [31,46–49]. In the present study, we successfully trained classifiers to discriminate between HCC, CC, and mCRC. Because inter- and intra-observer variability in pathologists' diagnoses is an important issue that limits the reliability of pathologic reports [50], DL-based classifiers can be adopted to reduce errors in pathologists' diagnostic decisions. However, low generalizability is an important hurdle in the adoption of DL-based assistant systems. We expect that the accumulation of digitized tissue data will eventually help to develop more generalized DL models and will help to improve diagnostic accuracy in the near future.

5. Conclusions

Our pipeline may provide significant information for patients with liver cancer and thus contribute to precision medicine. The findings of this study suggest that an AI-powered platform may be able to detect and properly classify liver cancer with high accuracy and efficiency, facilitating the application of screening tools for histopathologic diagnosis. Further large-scale studies are required to refine the models developed and validate our results.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/cancers15225389/s1>, Figure S1: The receiver operating characteristic curve of slide-level classification results for the external datasets (Seoul St. Mary's Hospital datasets); Figure S2: Classification results between hepatocellular carcinoma and other cancer types (cholangiocarcinoma and metastatic colorectal cancer), classified by a classifier trained with mixed datasets; Table S1. Baseline characteristics of all cohorts; Table S2: Performance of AlexNet, ResNet-50, and Inception-v3 models for the normal/tumor classification represented as area under the curves for the receiver operating characteristic curves; Table S3: The average number of patches in each training fold for the 5-fold cross-validation scheme; Table S4: Accuracy, sensitivity, specificity, and F1 score of the classification results.

Author Contributions: Conceptualization, H.-J.J. and S.H.L.; methodology, H.-J.J. and S.H.L.; software, H.-J.J.; validation, J.-H.G., Y.K. and S.H.L.; formal analysis, J.-H.G., Y.K., H.-J.J. and S.H.L.; data curation, H.-J.J. and S.H.L.; writing—original draft preparation, H.-J.J. and S.H.L.; writing—review and editing, H.-J.J. and S.H.L.; visualization, H.-J.J.; supervision, S.H.L.; funding acquisition, H.-J.J. and S.H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a grant from the National Research Foundation of Korea (NRF-2022R1A2C2010644) and a grant from the National Research Foundation of Korea (NRF-2021R1A4A5028966).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of the College of Medicine at The Catholic University of Korea (KC19SESI0787, approved on 21 November 2019).

Informed Consent Statement: Not applicable.

Data Availability Statement: The TCGA data presented in this study are openly available in the GDC data portal (<https://portal.gdc.cancer.gov/>, accessed on 10 July 2023). Further information is available from the corresponding authors upon request.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

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 Cancer J. Clin.* **2021**, *71*, 209–249. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Torbenson, M.; Ng, I.; Park, Y.; Roncalli, M.; Sakamoto, M. *WHO Classification of Tumours. Digestive System Tumours*; International Agency for Research on Cancer: Lyon, France, 2019; pp. 229–239.
3. Mak, L.Y.; Cruz-Ramon, V.; Chinchilla-Lopez, P.; Torres, H.A.; LoConte, N.K.; Rice, J.P.; Foxhall, L.E.; Sturgis, E.M.; Merrill, J.K.; Bailey, H.H.; et al. Global Epidemiology, Prevention, and Management of Hepatocellular Carcinoma. *Am. Soc. Clin. Oncol. Educ. Book* **2018**, *38*, 262–279. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Cooper, L.A.; Kong, J.; Gutman, D.A.; Dunn, W.D.; Nalisnik, M.; Brat, D.J. Novel genotype-phenotype associations in human cancers enabled by advanced molecular platforms and computational analysis of whole slide images. *Lab. Investig.* **2015**, *95*, 366–376. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Chen, C.; Chen, C.; Ma, M.; Ma, X.; Lv, X.; Dong, X.; Yan, Z.; Zhu, M.; Chen, J. Classification of multi-differentiated liver cancer pathological images based on deep learning attention mechanism. *BMC Med. Inform. Decis. Mak.* **2022**, *22*, 176. [\[CrossRef\]](#)
6. Friemel, J.; Rechsteiner, M.; Frick, L.; Bohm, F.; Struckmann, K.; Egger, M.; Moch, H.; Heikenwalder, M.; Weber, A. Intratumor heterogeneity in hepatocellular carcinoma. *Clin. Cancer Res.* **2015**, *21*, 1951–1961. [\[CrossRef\]](#)
7. Massarweh, N.N.; El-Serag, H.B. Epidemiology of Hepatocellular Carcinoma and Intrahepatic Cholangiocarcinoma. *Cancer Control* **2017**, *24*, 1073274817729245. [\[CrossRef\]](#)
8. Chung, T.; Park, Y.N. Up-to-Date Pathologic Classification and Molecular Characteristics of Intrahepatic Cholangiocarcinoma. *Front Med.* **2022**, *9*, 857140. [\[CrossRef\]](#)
9. Amin, M.B.; Edge, S.B.; Greene, F.L.; Byrd, D.R.; Brookland, R.K.; Washington, M.K.; Gershenwald, J.E.; Compton, C.C.; Hess, K.R.; Sullivan, D.C. *AJCC Cancer Staging Manual*; Springer: New York, NY, USA, 2017; Volume 1024.
10. Altekruse, S.F.; Devesa, S.S.; Dickie, L.A.; McGlynn, K.A.; Kleiner, D.E. Histological classification of liver and intrahepatic bile duct cancers in SEER registries. *J. Registry Manag.* **2011**, *38*, 201–205.
11. Lei, J.Y.; Bourne, P.A.; diSant’Agnese, P.A.; Huang, J. Cytoplasmic staining of TTF-1 in the differential diagnosis of hepatocellular carcinoma vs. cholangiocarcinoma and metastatic carcinoma of the liver. *Am. J. Clin. Pathol.* **2006**, *125*, 519–525. [\[CrossRef\]](#)
12. Park, J.H.; Kim, J.H. Pathologic differential diagnosis of metastatic carcinoma in the liver. *Clin. Mol. Hepatol.* **2019**, *25*, 12–20. [\[CrossRef\]](#)
13. van der Geest, L.G.; Lam-Boer, J.; Koopman, M.; Verhoef, C.; Elferink, M.A.; de Wilt, J.H. Nationwide trends in incidence, treatment and survival of colorectal cancer patients with synchronous metastases. *Clin. Exp. Metastasis* **2015**, *32*, 457–465. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Riihimäki, M.; Hemminki, A.; Sundquist, J.; Hemminki, K. Patterns of metastasis in colon and rectal cancer. *Sci. Rep.* **2016**, *6*, 29765. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Fleming, M.; Ravula, S.; Tatishchev, S.F.; Wang, H.L. Colorectal carcinoma: Pathologic aspects. *J. Gastrointest. Oncol.* **2012**, *3*, 153–173. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Rajpurkar, P.; Irvin, J.; Ball, R.L.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.P.; et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med.* **2018**, *15*, e1002686. [\[CrossRef\]](#)
17. Hannun, A.Y.; Rajpurkar, P.; Haghpanahi, M.; Tison, G.H.; Bourn, C.; Turakhia, M.P.; Ng, A.Y. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat. Med.* **2019**, *25*, 65–69. [\[CrossRef\]](#)
18. Joshi, G.; Jain, A.; Adhikari, S.; Garg, H.; Bhandari, M. FDA approved Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices: An updated 2022 landscape. *medRxiv* **2022**. [\[CrossRef\]](#)
19. Campanella, G.; Hanna, M.G.; Geneslaw, L.; Miralflor, A.; Werneck Krauss Silva, V.; Busam, K.J.; Brogi, E.; Reuter, V.E.; Klimstra, D.S.; Fuchs, T.J. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat. Med.* **2019**, *25*, 1301–1309. [\[CrossRef\]](#)
20. Ehteshami Bejnordi, B.; Veta, M.; Johannes van Diest, P.; van Ginneken, B.; Karssemeijer, N.; Litjens, G.; van der Laak, J.; the, C.C.; Hermesen, M.; Manson, Q.F.; et al. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA* **2017**, *318*, 2199–2210. [\[CrossRef\]](#)
21. 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. [\[CrossRef\]](#)
22. 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). *Am. J. Surg. Pathol.* **2018**, *42*, 39–52. [\[CrossRef\]](#)
23. Retamero, J.A.; Aneiros-Fernandez, J.; Del Moral, R.G. Complete Digital Pathology for Routine Histopathology Diagnosis in a Multicenter Hospital Network. *Arch. Pathol. Lab. Med.* **2020**, *144*, 221–228. [\[CrossRef\]](#) [\[PubMed\]](#)
24. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Vintch, B.; Zaharia, A.D.; Movshon, J.A.; Simoncelli, E.P. Efficient and direct estimation of a neural subunit model for sensory coding. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 3113–3121. [\[PubMed\]](#)

26. Kermamy, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.S.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell* **2018**, *172*, 1122–1131.e1129. [\[CrossRef\]](#)
27. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.; van Ginneken, B.; Sanchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [\[CrossRef\]](#)
28. Coudray, N.; Ocampo, P.S.; Sakellaropoulos, T.; Narula, N.; Snuderl, M.; Fenyo, D.; Moreira, A.L.; Razavian, N.; Tsirigos, A. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat. Med.* **2018**, *24*, 1559–1567. [\[CrossRef\]](#)
29. Yoshida, H.; Shimazu, T.; Kiyuna, T.; Marugame, A.; Yamashita, Y.; Cosatto, E.; Taniguchi, H.; Sekine, S.; Ochiai, A. Automated histological classification of whole-slide images of gastric biopsy specimens. *Gastric Cancer* **2018**, *21*, 249–257. [\[CrossRef\]](#)
30. Bandi, P.; Geessink, O.; Manson, Q.; Van Dijk, M.; Balkenhol, M.; Hermesen, M.; Ehteshami Bejnordi, B.; Lee, B.; Paeng, K.; Zhong, A.; et al. From Detection of Individual Metastases to Classification of Lymph Node Status at the Patient Level: The CAMELYON17 Challenge. *IEEE Trans. Med. Imaging* **2019**, *38*, 550–560. [\[CrossRef\]](#)
31. Song, J.; Im, S.; Lee, S.H.; Jang, H.J. Deep Learning-Based Classification of Uterine Cervical and Endometrial Cancer Subtypes from Whole-Slide Histopathology Images. *Diagnostics* **2022**, *12*, 2623. [\[CrossRef\]](#)
32. Venkatraman, E.S. A permutation test to compare receiver operating characteristic curves. *Biometrics* **2000**, *56*, 1134–1138. [\[CrossRef\]](#)
33. Dimitriou, N.; Arandjelovic, O.; Caie, P.D. Deep Learning for Whole Slide Image Analysis: An Overview. *Front. Med.* **2019**, *6*, 264. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Wagner, S.J.; Reisenbuchler, D.; West, N.P.; Niehues, J.M.; Zhu, J.; Foersch, S.; Veldhuizen, G.P.; Quirke, P.; Grabsch, H.I.; van den Brandt, P.A.; et al. Transformer-based biomarker prediction from colorectal cancer histology: A large-scale multicentric study. *Cancer Cell* **2023**, *41*, 1650–1661.e1654. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Jang, H.J.; Lee, A.; Kang, J.; Song, I.H.; Lee, S.H. Prediction of clinically actionable genetic alterations from colorectal cancer histopathology images using deep learning. *World J. Gastroenterol.* **2020**, *26*, 6207–6223. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Lee, S.H.; Song, I.H.; Jang, H.J. Feasibility of deep learning-based fully automated classification of microsatellite instability in tissue slides of colorectal cancer. *Int. J. Cancer* **2021**, *149*, 728–740. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Jang, H.J.; Lee, A.; Kang, J.; Song, I.H.; Lee, S.H. Prediction of genetic alterations from gastric cancer histopathology images using a fully automated deep learning approach. *World J. Gastroenterol.* **2021**, *27*, 7687–7704. [\[CrossRef\]](#)
38. Lee, S.H.; Lee, Y.; Jang, H.J. Deep learning captures selective features for discrimination of microsatellite instability from pathologic tissue slides of gastric cancer. *Int. J. Cancer* **2023**, *152*, 298–307. [\[CrossRef\]](#)
39. Nam, S.; Chong, Y.; Jung, C.K.; Kwak, T.Y.; Lee, J.Y.; Park, J.; Rho, M.J.; Go, H. Introduction to digital pathology and computer-aided pathology. *J. Pathol. Transl. Med.* **2020**, *54*, 125–134. [\[CrossRef\]](#)
40. Serag, A.; Ion-Margineanu, A.; Qureshi, H.; McMillan, R.; Saint Martin, M.J.; Diamond, J.; O'Reilly, P.; Hamilton, P. Translational AI and Deep Learning in Diagnostic Pathology. *Front. Med.* **2019**, *6*, 185. [\[CrossRef\]](#)
41. Bousis, D.; Verras, G.-I.; Bouchagier, K.; Antzoulas, A.; Panagiotopoulos, I.; Katinioti, A.; Kehagias, D.; Kaplanis, C.; Kotis, K.; Anagnostopoulos, C.-N. The role of deep learning in diagnosing colorectal cancer. *Gastroenterol. Rev./Przegląd Gastroenterol.* **2023**, *18*, 266–273. [\[CrossRef\]](#)
42. Yaqoob, A.; Musheer Aziz, R.; verma, N.K. Applications and Techniques of Machine Learning in Cancer Classification: A Systematic Review. *Hum. Centric Intell. Syst.* **2023**, 1–28. [\[CrossRef\]](#)
43. Afreen, S.; Bhurjee, A.K.; Aziz, R.M. Gene selection with Game Shapley Harris hawks optimizer for cancer classification. *Chemom. Intell. Lab. Syst.* **2023**, *242*, 104989. [\[CrossRef\]](#)
44. Mulita, F.; Verras, G.I.; Anagnostopoulos, C.N.; Kotis, K. A Smarter Health through the Internet of Surgical Things. *Sensors* **2022**, *22*, 4577. [\[CrossRef\]](#)
45. Evans, A.J.; Bauer, T.W.; Bui, M.M.; Cornish, T.C.; Duncan, H.; Glassy, E.F.; Hipp, J.; McGee, R.S.; Murphy, D.; Myers, C.; et al. US Food and Drug Administration Approval of Whole Slide Imaging for Primary Diagnosis: A Key Milestone Is Reached and New Questions Are Raised. *Arch. Pathol. Lab. Med.* **2018**, *142*, 1383–1387. [\[CrossRef\]](#) [\[PubMed\]](#)
46. 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. *Technol. Cancer Res. Treat.* **2021**, *20*, 15330338211027901. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Im, S.; Hyeon, J.; Rha, E.; Lee, J.; Choi, H.J.; Jung, Y.; Kim, T.J. Classification of Diffuse Glioma Subtype from Clinical-Grade Pathological Images Using Deep Transfer Learning. *Sensors* **2021**, *21*, 3500. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Yu, K.H.; Wang, F.; Berry, G.J.; Re, C.; Altman, R.B.; Snyder, M.; Kohane, I.S. Classifying non-small cell lung cancer types and transcriptomic subtypes using convolutional neural networks. *JAMA* **2020**, *27*, 757–769. [\[CrossRef\]](#) [\[PubMed\]](#)
49. 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. [\[CrossRef\]](#)
50. Niazi, M.K.K.; Parwani, A.V.; Gurcan, M.N. Digital pathology and artificial intelligence. *Lancet Oncol.* **2019**, *20*, e253–e261. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.