

Article

Deep Learning-Based Computational Cytopathologic Diagnosis of Metastatic Breast Carcinoma in Pleural Fluid

Hong Sik Park ^{1,†}, Yosep Chong ^{1,†} , Yujin Lee ¹ , Kwangil Yim ¹ , Kyung Jin Seo ¹ , Gisu Hwang ², Dahyeon Kim ², Gyungyub Gong ³, Nam Hoon Cho ⁴ , Chong Woo Yoo ⁵ and Hyun Joo Choi ^{1,*} 

¹ Department of Hospital Pathology, The Catholic University of Korea College of Medicine, Seoul 06591, Republic of Korea; griselbrand@gmail.com (H.S.P.); ychong@catholic.ac.kr (Y.C.); wondoocha@naver.com (Y.L.); kangse_manse@catholic.ac.kr (K.Y.); ywacko@catholic.ac.kr (K.J.S.)

² AI Team, DeepNoid Inc., Seoul 08376, Republic of Korea; kisu031@gmail.com (G.H.); anniy8920@outlook.kr (D.K.)

³ Department of Pathology, Asan Medical Center, Seoul 05505, Republic of Korea; gygong@amc.seoul.kr

⁴ Department of Pathology, Yonsei University College of Medicine, Seoul 03722, Republic of Korea; cho1988@yuhs.ac

⁵ Department of Pathology, National Cancer Center, Ilsan, Goyang-si 10408, Gyeonggi-do, Republic of Korea; cwy@ncc.re.kr

* Correspondence: chj0103@catholic.ac.kr; Tel.: +82-031-249-7634; Fax: +82-031-244-6786

† These authors contributed equally to this work.

Abstract: A Pleural effusion cytology is vital for treating metastatic breast cancer; however, concerns have arisen regarding the low accuracy and inter-observer variability in cytologic diagnosis. Although artificial intelligence-based image analysis has shown promise in cytopathology research, its application in diagnosing breast cancer in pleural fluid remains unexplored. To overcome these limitations, we evaluate the diagnostic accuracy of an artificial intelligence-based model using a large collection of cytopathological slides, to detect the malignant pleural effusion cytology associated with breast cancer. This study includes a total of 569 cytological slides of malignant pleural effusion of metastatic breast cancer from various institutions. We extracted 34,221 augmented image patches from whole-slide images and trained and validated a deep convolutional neural network model (DCNN) (Inception-ResNet-V2) with the images. Using this model, we classified 845 randomly selected patches, which were reviewed by three pathologists to compare their accuracy. The DCNN model outperforms the pathologists by demonstrating higher accuracy, sensitivity, and specificity compared to the pathologists (81.1% vs. 68.7%, 95.0% vs. 72.5%, and 98.6% vs. 88.9%, respectively). The pathologists reviewed the discordant cases of DCNN. After re-examination, the average accuracy, sensitivity, and specificity of the pathologists improved to 87.9, 80.2, and 95.7%, respectively. This study shows that DCNN can accurately diagnose malignant pleural effusion cytology in breast cancer and has the potential to support pathologists.

Keywords: breast neoplasms; cytology; pleural fluid; artificial intelligence; metastasis; deep learning



Citation: Park, H.S.; Chong, Y.; Lee, Y.; Yim, K.; Seo, K.J.; Hwang, G.; Kim, D.; Gong, G.; Cho, N.H.; Yoo, C.W.; et al. Deep Learning-Based Computational Cytopathologic Diagnosis of Metastatic Breast Carcinoma in Pleural Fluid. *Cells* **2023**, *12*, 1847. <https://doi.org/10.3390/cells12141847>

Academic Editor: Yu Xue

Received: 19 June 2023

Revised: 7 July 2023

Accepted: 11 July 2023

Published: 13 July 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

In 2020, breast cancer was the leading cancer worldwide and the fifth leading cause of cancer mortality [1]. Most of these deaths have been attributed to metastatic breast cancer. Breast cancer can spread to various organs in the body, with the most common sites being the bone, lung, liver, and brain [2,3]. In the United States, women with metastatic breast cancer have a 5-year survival rate of 30% [4]. The presence of malignant pleural effusions suggests advanced disease and is associated with poor survival rates [5]. Pleural metastasis occurs when cancer cells from a primary breast tumor spread to the pleural lining surrounding the lungs. Patients with breast cancer often experience pleural metastasis in the first few years after diagnosis, although this becomes less frequent with longer periods of disease-free survival [6]. Aspiration cytology is a widely used and efficient

method for diagnosing malignant pleural effusions. This technique is recognized for its simplicity, cost-effectiveness, and rapid results [7]. Malignant pleural effusion cytology plays a pivotal role in the diagnosis, treatment planning, and monitoring of metastatic breast cancer. Patients with malignant effusions may also experience additional symptoms, such as dyspnea. Proper management of the respiratory decline can significantly enhance the patient's quality of life. Timely information obtained from pleural effusion cytology assists in optimizing patient care, facilitating personalized treatment approaches, and improving overall patient outcomes.

However, there are concerns regarding the accuracy of cytological diagnosis, which vary depending on the category. The accuracy is higher for negative and malignant categories, at 76% and 81%, respectively [8]. On the other hand, for the intermediate suspicious malignancy cases, there is a lower agreement observed, with rates of 32% and 22%, respectively. In the previous study, the overall agreement among pathologists for cytological diagnosis was 68% [8]. A recent meta-analysis reported that overall diagnostic sensitivity for malignant pleural effusions was 58% [9,10]. Cytological diagnosis varies according to the type of malignancy, particularly in primary thoracic malignancies. The sensitivity for lung adenocarcinoma was found to be 83.6%, while for lung squamous cell carcinoma, it was 24.2% [9,10]. The report revealed that the diagnostic sensitivity of malignant pleural effusion for breast cancer was 65.3%, indicating a relatively low performance as a screening tool. Suspected malignant pleural effusion patients with negative cytology may need flow cytometry, tissue biopsy, and other tests for evaluation. Flow cytometry is useful for malignant hematological effusions, while immunochemistry and fluorescence in situ hybridization aid in identifying non-hematologic malignancies [11]. However, these tests can be costly and time-consuming. The limitations of cytological diagnosis, including its subjective nature and potential for human error, have driven researchers to explore alternative approaches that offer more reliable and accurate assessments. Consequently, researchers have actively utilized deep learning methods to diagnose cytological images and address the drawbacks and weaknesses associated with conventional cytology approaches. This study aimed to assess the diagnostic accuracy of AI in detecting malignant pleural effusion cytology, specifically that related to breast cancer. This will be accomplished by analyzing a larger number of cytopathology slides gathered from multiple institutions across the nation. Additionally, this study aimed to utilize z-stacked cytology slide images to develop an AI model.

Deep learning-based artificial intelligence (AI) image analysis technology has recently been actively researched in the medical field, and it has been showing remarkable results [12–14]. In addition, a recently published study on the classification of lung cancer cells in pleural fluid using a deep convolutional neural network showed promising results [10]. Several studies have already indicated that AI achieves diagnostic accuracy equal to or better than that of human experts in the diagnosis of cervical cytology slide images [7–9]. AI models have been increasingly used in cytology, with an initial focus on the examination of gynecological samples, showing promising results [15]. Currently, AI models are used to classify various nongynecological samples, including urinary tract cytology, thyroid fine needle aspiration cytology (FNAC), breast FNAC, and pleural fluids [16–19]. However, only a few studies have been conducted on AI classification in malignant serous fluid cytology [12,19]. In addition, no studies have diagnosed breast cancer in pleural fluid using AI to the best of our knowledge, and only two studies on breast cancer FNAC using artificial neural networks have been published [19,20]. These studies have shown perfect results in distinguishing benign from malignant tumors. However, the number of cases was relatively small; therefore, it was necessary to use a larger number of slide images for validation. Furthermore, many studies have employed image datasets to develop AI models for cytological diagnosis; however, these datasets often lack z-stacking images. For this study, we collected a larger number of cytopathology slides with z-stacked images and developed an AI model.

2. Materials and Methods

This study was approved by the Institutional Review Boards of the Catholic University of Korea, College of Medicine (UC21SNSI0064); Yonsei University College of Medicine (4-2021-0569); the National Cancer Center (NCC2021-0145); and St. Vincent Hospital, Catholic University of Korea, College of Medicine (VC22RISI0131). An outline of the method for classifying malignant and benign cells in pleural cell slide images of breast cancer patients using a deep convolutional neural network model (DCNN) is presented in Figure 1.

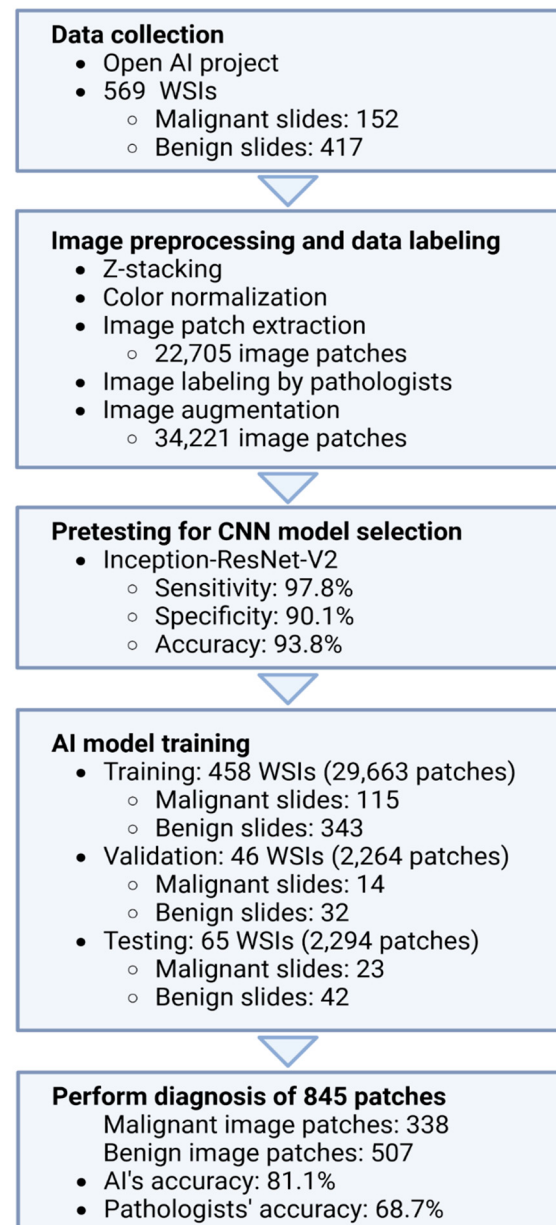


Figure 1. Overview of the study.

2.1. Data Collection

In this study, we used 569 cytologic whole-slide images (WSIs) collected from the Open AI Dataset project for cytopathology in 2021 (<https://www.aihub.or.kr> accessed on 19 June 2023). In collaboration with the Korean Society of Cytopathology, the Catholic University of Korea Uijeongbu St. Mary's Hospital, the National Cancer Center, the Korea Cancer Center Hospital, and ten other medical institutions constructed a digi-

tal cytology learning dataset. This dataset was created by refining, labeling, storing, and quality-controlling cytopathology images, consisting of 5506 WSIs for ten types of cancer and benign cases. Furthermore, 207,037 patch images were extracted from this dataset. All malignant cytopathological slides were confirmed via histological examination. The cytopathological dataset collection and preparation processes are shown in Supplementary Figures S1 and S2, respectively.

2.2. Image Preprocessing

The collected WSIs were scanned as extended depth-of-field images by merging multilayered z-stacked images through slide scanners to correct for image defocusing. The scanners used were AT2 (Leica Biosystems, Nussloch, Germany), Panoramic Flash 250 III (3DHISTECH, Budapest, Hungary), and NanoZoomer S360 (Hamamatsu, Shizuoka, Japan). After obtaining the WSIs, color normalization was performed for color constancy. Then, we split those images into non-overlapping small image patches of 1024×1024 pixels. Pathologists classified and labeled the split images as malignant or benign. After labeling, downsampling was performed to convert the images into 256×256 -pixel images. Finally, the image patch data were augmented via horizontal flip, vertical flip, and clockwise rotation to improve the sufficiency and diversity of the training data (Figure 2).

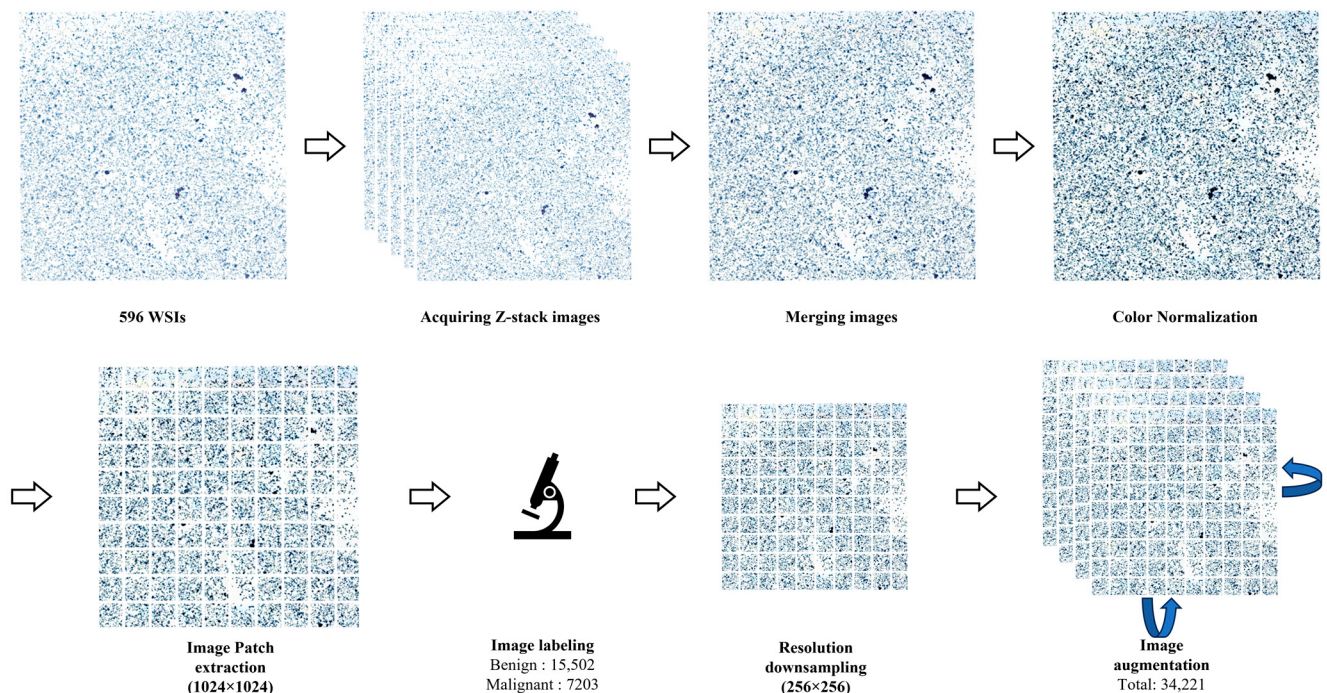


Figure 2. Overview of image preparation in the study. All cytology slides were scanned in multiple layers by z-stacking and merged to obtain an image with appropriate focus. WSIs were segmented into small image patches and labeled by pathologists as benign or malignant. Image augmentation was performed after labeling.

2.3. Pretesting for DCNN Model Selection

We pretested several DCNN models, including Inception ResNet-V2, Efficientnet-b1, ResNext50, Mobilenet v2, Densenet 121, and ResNet 50. A total of 344 WSIs, including 152 malignant and 192 benign, were used in the pre-testing. After preprocessing, 14,526 image patches were acquired (7203 malignant/7323 benign). We used 11,575 patch images (5765 malignant/5810 benign) for training, 1475 patch images (722 malignant/756 benign) for validation, and 1473 patch images (716 malignant/757 benign) for testing. After a comparison of the performances, we selected Inception-ResNet-V2 as the model for this study. To prevent overfitting, we employed data augmentation techniques. Additionally,

we implemented early stopping to obtain the best possible model. The data were made more quantitative by extracting patches from WSIs.

2.4. AI Model Training

We experimented with 165-layer Inception-ResNet-v2, as in Figure 3. The Inception-ResNet-V2 model is used in this experiment. Inception-ResNet-V2 is a convolutional neural network developed by Google that has satisfactory performance in image analysis. The Inception-ResNet approach combines the residual network with an inception architecture previously developed by Google [21]. A residual network is a neural network designed to address the issue of increasing training and test errors as the number of layers significantly increases in existing deep learning. In addition, residual networks use shortcut connections to jump directly to the classification stage when training reaches the optimal depth, thus enabling faster training than Inception V3 [21,22]. In this experiment, we substituted the softmax function in the output layer with a sigmoid function for the binary classification. The augmented patch images were randomly assigned to train, validate, and test the DCNN model (Table 1). The malignant probability of each image patch was estimated; if the probability was 50% or more, the image patch was classified as malignant. The dataset was divided into three parts, namely training, validation, and testing, following an 8:1:1 ratio as specified. The patch-wise distribution of positive and negative images is well balanced, with a ratio of 6:5. Details of the number of WSIs/image patches used for the DCNN model are shown in Table 1.

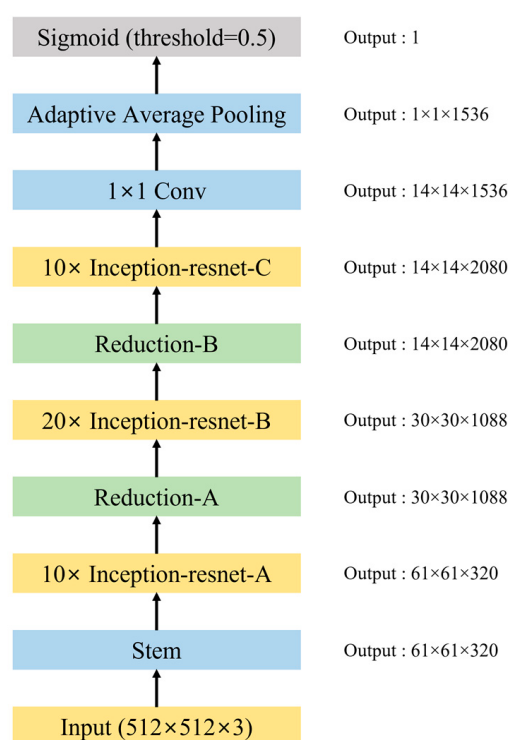


Figure 3. Schema of Inception-ResNet-V2 network used in this study for binary classification.

Table 1. Number of whole-slide images used for the deep convoluted neural network (DCNN) model.

	No. of Whole Slide Images (Image Patches)			
	Training	Validation	Testing	Total
Benign	330 (12,389)	44 (1545)	43 (1568)	417 (15,502)
Malignant	111 (17,274)	18 (719)	23 (726)	152 (18,179)
Total	441 (29,663)	62 (2264)	66 (2294)	569 (34,221)

2.5. Comparison of the Performance of the AI Model and the Pathologists

To evaluate the performance of the AI model, we randomly selected 845 Pap smear images, including 338 malignant and 507 benign, from the 34,221 image patches. Subsequently, three pathologists and the AI model diagnosed the patch images, and the results were compared. After comparison, each pathologist re-diagnosed the patch images that were inconsistent with the AI diagnosis.

3. Results

3.1. Data Characteristics

We collected Papanicolaou-stained 569 WSIs, consisting of 417 benign and 152 metastatic breast carcinoma cases, which included 564 conventional smears and 5 liquid-based prep (LBP) slides. The WSIs contained at least three z-stack layers, and 94.9% of the WSIs were scanned using a Panoramic Flash 250 III (3DHISTECH, Budapest, Hungary). Table 2 shows the characteristics of the cytological slides used in this study.

Table 2. Data characteristics of the enrolled cytologic slides.

Characteristics	Number of Cases (<i>n</i> = 569)
Age (median)	18–104 (66)
Cytologic diagnosis	
Malignant lesions	152 (26.7%)
Benign lesions	417 (73.3%)
Preparation method	
Conventional	564 (99.1%)
Liquid-based preparation	5 (0.9%)
Z-stack layers	
3-layers	490 (86.1%)
5-layers	79 (13.9%)
Scanner	
Panoramic Flash 250 III (3DHISTECH)	540 (94.9%)
AT2 (Leica)	2 (0.4%)
NanoZoomer S360 (Hamamatsu)	27 (4.7%)

3.2. Pretesting for DCNN Model Selection

The overall accuracy of the image classification ranged from 89.48% to 93.82%. Inception ResNet v2 showed the highest accuracy of 93.8%, whereas Efficientnet-b1, ResNext50, MobileNet v2, Densenet 121, and ResNet 50 showed accuracies of 92.7, 93.5, 89.5, 92.5, and 91.9%, respectively. Among the DCNN models, Inception-ResNet-v2 showed the most accurate results, with a sensitivity and specificity of 97.77 and 90.09%, respectively. The comparison of the performances of the models with the pre-testing model selection results is listed in Table 3.

Table 3. Pretesting performance of DCNN models. Inception-ResNet-V2 exhibited the highest accuracy and best performance among the pretesting models.

	Accuracy	Sensitivity	Specificity
Inception ResNet v2	0.9382	0.9777	0.9009
Efficientnet-b1	0.9267	0.8855	0.9657
ResNext50	0.9348	0.9679	0.9036
MobileNet v2	0.8948	0.9162	0.8745
Densenet 121	0.9246	0.9218	0.9273
ResNet 50	0.9192	0.8980	0.9392

3.3. AI Model Training

The Inception ResNet v2 model showed accuracy, sensitivity, and specificity values of 95.0, 93.4, and 95.8%, respectively, for an oversampled training dataset. The model showed

accuracy, sensitivity, and specificity values of 92.9, 87.6, and 96.6%, respectively, when using all training datasets and 93.8, 97.8, and 90.1%, respectively, when using a reduced training dataset.

3.4. Comparison of the Performances of the AI Model and Pathologists

The average of the three pathologists showed 72.49% sensitivity, 88.89% specificity, and 68.74% accuracy with a Fleiss kappa coefficient of 0.482. Pathologist A showed a sensitivity of 67.8%, a specificity of 94.9%, and an accuracy of 70.2%; Pathologist B showed a sensitivity of 56.8%, a specificity of 98.0%, and an accuracy of 68.1%; and Pathologist C showed a sensitivity of 92.9%, a specificity of 73.8%, and an accuracy of 68.0%. The AI showed a sensitivity of 95.0%, a specificity of 98.6%, and an accuracy of 81.1%.

The AI diagnosed 321 out of 845 true positives, 500 true negatives, 7 false positives, and 17 false negatives (Table 4). Examples of AI-diagnosed images are shown in Figure 4. In cases where both the AI and a pathologist made diagnoses, the AI alone correctly diagnosed 18 true positives and 3 true negatives. There was one false positive and two false negatives when only the AI was incorrect. Images showing the discrepancies in the diagnosis by the AI and pathologists are shown in Figure 5.

Table 4. Confusion matrix of the AI classification.

Actual Diagnosis Total (<i>n</i> = 845)	AI Diagnosis	
	Malignant (328)	Benign (517)
Malignant (338)	True positive (321)	False negative (17)
Benign (507)	False positive (7)	True negative (500)

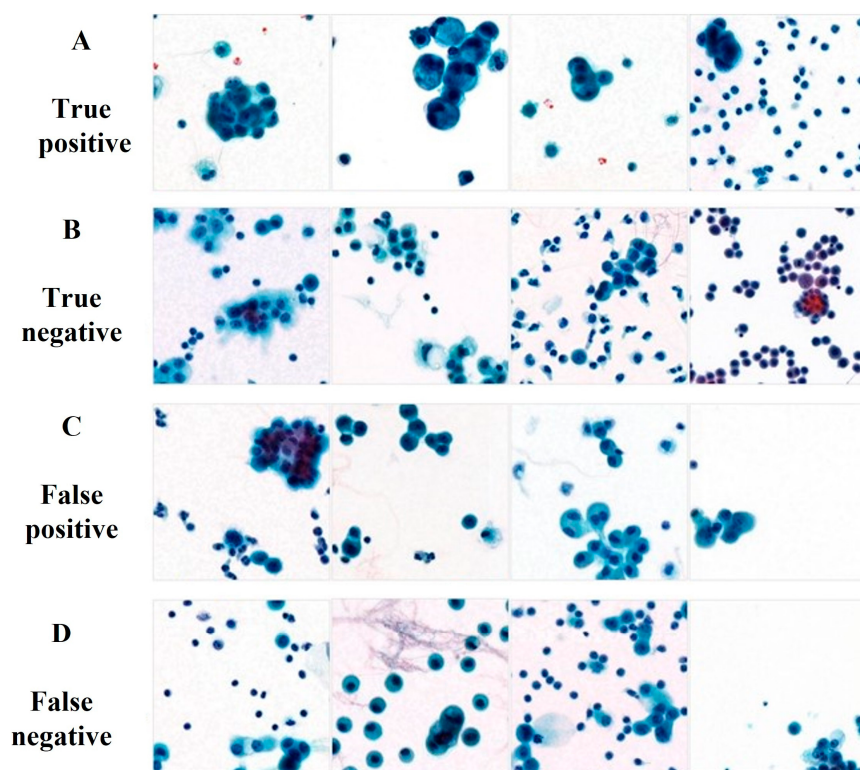


Figure 4. Images classified by AI. Patch images accurately classified patch images by the AI: (A) Malignant and (B) benign. Patch images misdiagnosed by AI (C) benign cells as malignant, and (D) malignant cells as benign.

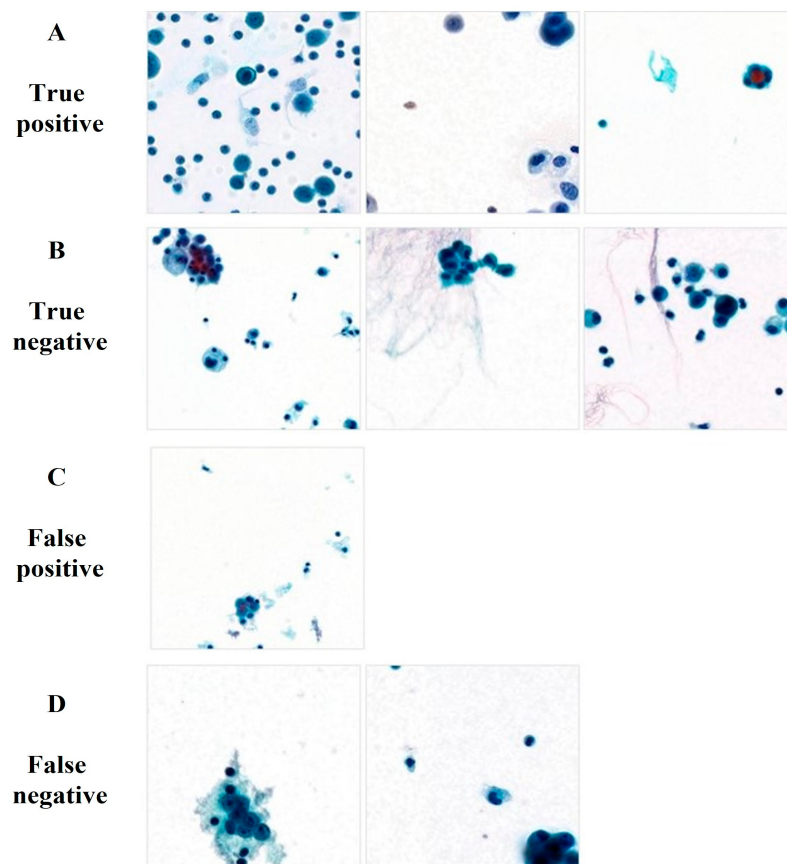


Figure 5. Patch images correctly or incorrectly diagnosed by AI. Patch images the correctly diagnosed by the AI but misdiagnosed by pathologists: (A) Malignant and (B) benign. Patch images correctly diagnosed by pathologists but misdiagnosed by the AI: (C) Benign and (D) malignant.

Cases with inconsistencies in the diagnosis by the AI and the pathologist were re-examined by all pathologists, and the average result improved to 86.19% sensitivity, 95.66% specificity, and 76.71% accuracy (Figure 6). In addition, the kappa coefficient for the pathologists improved to 0.806. The AI diagnosed 7 false-negative and 17 false-positive cases.

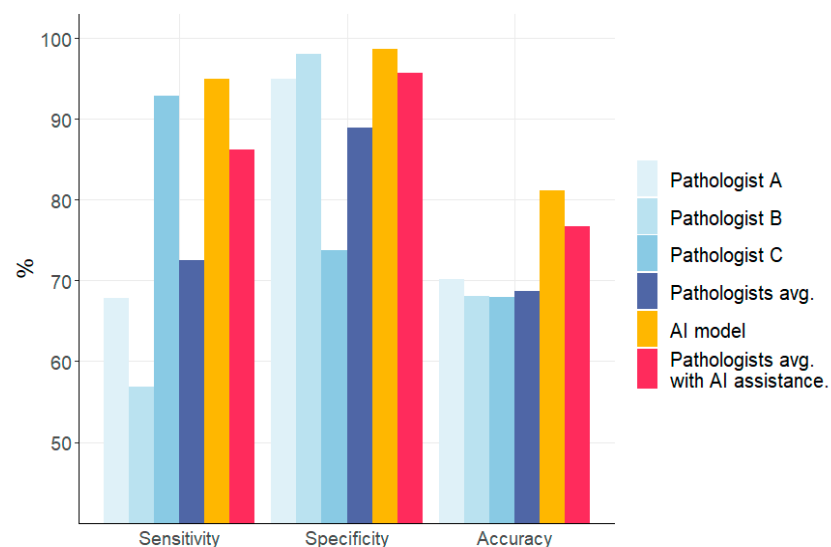


Figure 6. Comparison of the performance of the AI model and human pathologists.

4. Discussion

Diagnosing pathological WSIs in various organs using AI is currently showing considerable progress, with various patents and applications in the medical field [23,24]. A recent systematic review found that there are not many studies that have used artificial neural networks in the field of effusion cytology [25]. Consequently, additional research is imperative for exploring the potential of neural networks to aid in diagnostic cytology. In this study, we successfully demonstrated that AI exhibited remarkable accuracy in diagnosing breast cancer pleural effusion cytopathology for the first time. Our research is the first of its kind to use the largest dataset that includes z-stacking.

In this study, we used 569 qualitatively assured WSIs obtained from various medical institutions nationwide. The sample size was sufficiently large and exhibited a high level of demographic diversity and heterogeneity, surpassing the findings of previous studies. Moreover, we employed merged z-stacked images in our study, which resulted in an improved focus. The AI model developed for this experiment exhibited an accuracy of 81.13% in classifying patch images, outperforming the average accuracy of 72.49% achieved by experienced pathologists. Furthermore, when the cases in which the AI and pathologists disagreed on the diagnosis were reevaluated, the pathologists' diagnostic accuracy improved. This suggests that AI assistance can enhance cytopathologists' interpretive capabilities by striking a balance between sensitivity and specificity.

The morphological features of metastatic breast carcinomas are variable and can exhibit non-cohesive isolated cells, large cell balls, and linear arrangements [26,27]. Diagnosing metastatic breast carcinoma in pleural fluid is challenging even for experienced pathologists, especially when dealing with predominantly isolated cell patterns that are difficult to differentiate from reactive mesothelial cells and histiocytes [28]. Pleural fluid can serve as a culture medium for floating cells, leading to a decrease in the nuclear-to-cytoplasmic ratio compared to typical cancer cells; this phenomenon can mimic the appearance of mimicking normal or reactive cells [29]. Mesothelial cells can be readily activated and mimic phenotypic traits reminiscent of malignant cells, particularly when exposed to inflammatory stimuli such as pneumonia or tuberculosis. Macrophages, which are integral components of the mononuclear phagocytic system, can be activated in response to pleural effusion or inflammation. This activation can result in macrophages exhibiting characteristics that resemble malignant cells. Breast cancer cells, particularly metastatic lobular carcinomas, possess distinct characteristics such as smaller size and a more blended appearance that differentiate them from other types of carcinomas such as lung cancer [30–34]. This disparity in cellular morphology poses a challenge for accurately diagnosing these conditions via pleural fluid cytology.

In this study, AI performed well in identifying and classifying cells with distinct malignant features, such as nuclear hyperchromasia, pleomorphism, a high N:C ratio, and nuclear overlapping (Figure 4A). In addition, AI helped diagnose a small number of atypical cells that are difficult for humans to recognize (Figure 5A). However, it also tended to misdiagnose as malignant if the mesothelial cells showed severe pleomorphism, hyperchromasia, or nuclear overlap (Figures 4C and 5). In addition, there were cases in which AI could not correctly classify malignant cells because of an unclear cytoplasmic border, a low N:C ratio, or other unknown reasons (Figures 4D and 5C).

Inception-ResNet-V2 was introduced in 2016 and showed satisfactory top-1 and top-5 accuracies of 77.8 and 94.1%, respectively, in the ImageNet examination [35]. In addition, Inception-ResNet-V2 has been used in many pathological image and medical data diagnostic studies and has shown promising results. For example, a study in which Inception-ResNet-V2 classified tissue slides of skin melanoma reported high accuracy (96.5%), sensitivity (95.7%), and specificity (97.7%) [36]. Moreover, a study published in 2022, which analyzed cervical Pap smear cytology images, reported a compliance accuracy of 96.44% [37].

Currently, numerous AI-based models are emerging to aid pathologists in the diagnosis of breast cancer. Many AI-based applications have been developed in the field of breast

pathology, including primary tumor detection, breast cancer grading, the identification of histological subtypes of breast cancer, analysis of mitotic figures, and the prediction of survival outcomes [38–42]. Recently, AI algorithms have been developed to offer quantitative analysis of immunohistochemistry-stained images, specifically for evaluating the KI-67, estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor-2 (HER2) markers [41,43–45]. A recent study successfully developed an AI model for diagnosing lymph node metastasis in breast cancer, achieving an impressive accuracy of 98.8% [46]. Another study developed an AI model using pleural effusion cell blocks from breast, gastric, and lung carcinoma [47]. This model achieved an area under the curve (AUC) of 0.91, identifying benign and malignant pleural effusion, and successfully determined the primary site of metastatic carcinoma [47]. Previously, an AI model was developed to diagnose metastatic carcinoma using effusion cytology. The dataset used in this study consisted of 57 benign cases, 54 metastatic adenocarcinoma cases, 1 squamous cell carcinoma case, and 1 signet ring carcinoma case [48]. Although this study achieved an impressive 100% accuracy, the dataset used was relatively small [48]. This study did not clarify the number of patients with metastatic breast cancer pleural effusion that were included in the analysis. Our present study differs from previous studies in that this AI model specifically targeted breast cancer metastasis using a larger dataset. In another study, a computer-aided diagnostic approach was employed utilizing the nuclear structure of mesothelial cells to classify malignant mesothelioma in effusion cytology specimens [49]. This study also achieved 100% accuracy; however, it was based on a small dataset that included 16 cases of malignant mesothelioma and 18 cases of benign pleural effusion [49]. Recently, a few studies have employed AI for the FNAC cytology of breast cancer. Dey et al. successfully identified lobular carcinomas in FNAC samples [19]. Another study used an AI model to accurately identify fibroadenomas and infiltrating carcinomas of the breast in FNAC cases with 100% sensitivity and specificity [20]. However, these studies used small datasets. In this study, we utilized a larger dataset comprising 596 cytological WSIs of metastatic breast cancer, collected from various universities and hospitals throughout South Korea. Typically, malignant pleural effusion fluids contain a high number of benign background cells, leading to a low population of tumor cells. False-negative or inconclusive results may occur in samples with a low number of cancer cells during morphological analysis. The findings of the current study demonstrated superior accuracy, sensitivity, and specificity compared with pathologists in diagnosing malignant pleural effusion cytology related to breast cancer.

This study has some limitations. First, the diagnosis was limited to patch images and was not applied to the diagnosis of WSIs. Therefore, it is necessary to construct the malignancy probability of each image tile and set a cutoff value for the probability of all tiles. In addition, we divided the classification into malignant and benign two-tiers and did not designate the atypical category. Therefore, if a few atypical cells are present on WSI, this model will likely lead to an inappropriate diagnosis. Second, this AI model does not explain why cells are viewed as malignant or benign, which can confuse the pathologist and lead to an inappropriate diagnosis when the diagnoses of the pathologist and the AI do not match. Due to the black-box nature of AI systems, the mechanism by which they arrive at conclusions is not well understood. Therefore, it is necessary to introduce explainable AI (XAI) to address this shortcoming [50,51]. XAI focuses on developing AI models that can be understood by humans. XAI is used to create AI models that can not only make accurate predictions but also provide an explanation of how they arrived at those predictions. When the diagnoses of the pathologist and AI do not match, the pathologist can make a diagnosis more accurately and quickly by receiving real-time feedback from AI [52].

In future studies, the application of this AI model to WSI diagnosis and further validation with external WSIs will be necessary. If this model undergoes validation using data from a multi-ethnic population, then the performance of the model can be generalized. In addition, it is necessary to develop a more accurate AI model. Currently, various AI image classification models, showing higher classification accuracy and potential for use in

pathological image diagnosis, are being actively studied. Xie et al. reported that the noisy student training (EfficientNet-L2) method, which comprised pseudo-labeled images in the training dataset to add noise inside the dataset, improved the top-1 and top-5 accuracies of ImageNet classification to 88.4 and 98.7%, respectively [53]. In 2021, Pham et al. announced meta-pseudo labels and semi-supervised learning that improved noisy student training and enhanced the top-1 and top-5 accuracies of ImageNet classification to 90.2 and 98.8%, respectively [54].

The second aspect to be investigated is the use of fewer computer resources for model training to increase availability. In general, DCNN training requires increased computing power and data storage as accuracy increases. MobileNet is a lightweight convolutional neural network architecture for mobile and embedded vision applications. MobileNet uses depthwise separable convolutions to reduce the number of parameters and computations required for image classification. The top-1 and top-5 accuracies of this model on the ImageNet dataset are above 70 and 89%, respectively [55]. In addition, EfficientNet uses fewer parameters and calculations and is a relatively accurate model. EfficientNet-B0 showed top-1 and top-5 accuracies of 77.1 and 93.3%, respectively, in the ImageNet tests. In addition, EfficientNet-B7 showed top-1 and top-5 accuracies of 84.3 and 97.0%, respectively, but required more computer resources [56]. Other DCNN techniques can be considered when lower resources and faster training rates are required.

5. Conclusions

The AI model developed for this experiment exhibited an accuracy of 81.13% in classifying patch images, outperforming the average accuracy of 72.49% achieved by experienced pathologists. Cases with inconsistencies in the diagnosis by the AI and the pathologist were re-examined and the average accuracy, sensitivity, and specificity of the pathologists improved to 87.9, 80.2, and 95.7%, respectively. This study demonstrated that AI could provide significant accuracy in diagnosing breast cancer pleural effusion cytopathology, which could enhance the interpretative ability of cytopathologists. Future studies should focus on applying this AI model to WSI diagnosis, validating it with external WSIs. Additionally, efforts should be directed towards exploring methods to develop a more accurate AI model that requires fewer computer resources for model training, thereby increasing its availability.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cells12141847/s1>, Figure S1: 2021 NIA project for cytopathology dataset; Figure S2: Dataset preparation process.

Author Contributions: Conceptualization: G.G., C.W.Y. and Y.C.; methodology: G.G., C.W.Y. and Y.C.; software: Y.C.; validation: H.S.P., Y.L., K.J.S., K.Y., G.H., D.K., G.G., N.H.C., C.W.Y., H.J.C. and Y.C.; formal analysis: H.S.P., Y.L., K.J.S., K.Y., G.H., D.K., G.G., N.H.C., C.W.Y., H.J.C. and Y.C.; investigation: H.S.P., Y.L., K.J.S., K.Y., G.H., D.K., G.G., N.H.C., C.W.Y., H.J.C. and Y.C.; resources: Y.C.; data curation: H.S.P., Y.L., K.J.S., K.Y., G.H., D.K., G.G., N.H.C., C.W.Y., H.J.C. and Y.C.; writing—original draft preparation: H.S.P. and Y.C.; writing—review and editing: H.S.P., H.J.C. and Y.C.; visualization: H.S.P. and Y.C.; supervision: H.J.C. and Y.C.; and project administration: Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by a National Research Foundation of Korea (NRF) grant funded by the Korean Government (MSIT) (2021R1A2C2013630).

Institutional Review Board Statement: This study was approved by the Institutional Review Boards of the Catholic University of Korea, College of Medicine (UC21SNSI0064); Yonsei University College of Medicine (4-2021-0569); National Cancer Center (NCC2021-0145); and St. Vincent Hospital, Catholic University of Korea, College of Medicine (VC22RISI0131).

Informed Consent Statement: Patient consent was waived by the Institutional Review Boards of the Catholic University of Korea, College of Medicine (UC21SNSI0064); Yonsei University College of Medicine (4-2021-0569); National Cancer Center (NCC2021-0145); and Ulsan University School of Medicine (VC22RISI0131).

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Acknowledgments: We thank Muhammad Joan Ailia (The Catholic University of Korea (CUK)), Mohammad Rizwan Alam (CUK), Jamshid Abdul-Ghafar, In Park (CUK), Ah Reum Kim (CUK), Seona Shin (National Cancer Center (NCC)), Na Young Han (NCC), Joon Young Shin (Asan Medical Center), and Sook Hee Kang (Yonsei University College of Medicine) for their data collection, retrieval, management, annotation, and quality checks. We used datasets from The Open AI Dataset Project (AI-Hub, South Korea). All data information can be accessed through ‘AI-Hub (www.aihub.or.kr accessed on 19 June 2023)’.

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

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. [\[CrossRef\]](#)
2. Chen, M.-T.; Sun, H.-F.; Zhao, Y.; Fu, W.-Y.; Yang, L.-P.; Gao, S.-P.; Li, L.-D.; Jiang, H.-L.; Jin, W. Comparison of patterns and prognosis among distant metastatic breast cancer patients by age groups: A SEER population-based analysis. *Sci. Rep.* **2017**, *7*, 9254. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Wang, M.; Liu, X.; Wei, B.; Liu, N.; Li, Q.; Su, X. Mucinous breast carcinoma metastatic to thyroid gland: Report of a case diagnosed by fine-needle aspiration cytology. *Diagn. Cytopathol.* **2020**, *48*, 475–478. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Cancer.Net. Breast Cancer—Metastatic: Statistics. Available online: <https://www.cancer.net/cancer-types/breast-cancer-metastatic/statistics> (accessed on 1 June 2023).
5. Gayen, S. Malignant pleural effusion: Presentation, diagnosis, and management. *Am. J. Med.* **2022**, *135*, 1188–1192. [\[CrossRef\]](#)
6. Shinohara, T.; Yamada, H.; Fujimori, Y.; Yamagishi, K. Malignant pleural effusion in breast cancer 12 years after mastectomy that was successfully treated with endocrine therapy. *Am. J. Case Rep.* **2013**, *14*, 184.
7. Dermawan, J.K.T.; Policarpio-Nicolas, M.L. Malignancies in Pleural, Peritoneal, and Pericardial Effusions A 17-Year Single-Institution Review From 30 085 Specimens. *Arch. Pathol. Lab. Med.* **2020**, *144*, 1086–1091. [\[CrossRef\]](#)
8. Layfield, L.J.; Yang, Z.; Vazmitsel, M.; Zhang, T.; Esebua, M.; Schmidt, R. The international system for serous fluid cytopathology: Interobserver agreement. *Diagn. Cytopathol.* **2022**, *50*, 3–7. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Kassirian, S.; Hinton, S.N.; Cuninghame, S.; Chaudhary, R.; Iansavitchene, A.; Amjadi, K.; Dhaliwal, I.; Zeman-Pocrnich, C.; Mitchell, M.A. Diagnostic sensitivity of pleural fluid cytology in malignant pleural effusions: Systematic review and meta-analysis. *Thorax* **2023**, *78*, 32–40. [\[CrossRef\]](#)
10. 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. *Mod. Pathol.* **2022**, *35*, 609–614. [\[CrossRef\]](#)
11. Bode-Lesniewska, B. Flow cytometry and effusions in lymphoproliferative processes and other hematologic neoplasias. *Acta Cytol.* **2016**, *60*, 354–364. [\[CrossRef\]](#)
12. Thakur, N.; Alam, M.R.; Abdul-Ghafar, J.; Chong, Y. Recent Application of Artificial Intelligence in Non-Gynecological Cancer Cytopathology: A Systematic Review. *Cancers* **2022**, *14*, 3529. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Feng, H.; Yang, B.; Wang, J.; Liu, M.; Yin, L.; Zheng, W.; Yin, Z.; Liu, C. Identifying malignant breast ultrasound images using ViT-patch. *Appl. Sci.* **2023**, *13*, 3489. [\[CrossRef\]](#)
14. Zhuang, Y.; Chen, S.; Jiang, N.; Hu, H. An Effective WSENet-Based Similarity Retrieval Method of Large Lung CT Image Databases. *KSII Trans. Internet Inf. Syst.* **2022**, *16*, 2359–2376.
15. Wentzensen, N.; Lahrmann, B.; Clarke, M.A.; Kinney, W.; Tokugawa, D.; Poitras, N.; Locke, A.; Bartels, L.; Krauthoff, A.; Walker, J. Accuracy and efficiency of deep-learning-based automation of dual stain cytology in cervical Cancer screening. *JNCI J. Natl. Cancer Inst.* **2021**, *113*, 72–79. [\[CrossRef\]](#)
16. Muralidaran, C.; Dey, P.; Nijhawan, R.; Kakkar, N. Artificial neural network in diagnosis of urothelial cell carcinoma in urine cytology. *Diagn. Cytopathol.* **2015**, *43*, 443–449. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Sanyal, P.; Mukherjee, T.; Barui, S.; Das, A.; Gangopadhyay, P. Artificial intelligence in cytopathology: A neural network to identify papillary carcinoma on thyroid fine-needle aspiration cytology smears. *J. Pathol. Inform.* **2018**, *9*, 43. [\[CrossRef\]](#)
18. Gonzalez, D.; Dietz, R.L.; Pantanowitz, L. Feasibility of a deep learning algorithm to distinguish large cell neuroendocrine from small cell lung carcinoma in cytology specimens. *Cytopathology* **2020**, *31*, 426–431. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Dey, P.; Logasundaram, R.; Joshi, K. Artificial neural network in diagnosis of lobular carcinoma of breast in fine-needle aspiration cytology. *Diagn. Cytopathol.* **2013**, *41*, 102–106. [\[CrossRef\]](#)
20. Subbaiah, R.; Dey, P.; Nijhawan, R. Artificial neural network in breast lesions from fine-needle aspiration cytology smear. *Diagn. Cytopathol.* **2014**, *42*, 218–224. [\[CrossRef\]](#)
21. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
22. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

23. Ailia, M.J.; Thakur, N.; Abdul-Ghafar, J.; Jung, C.K.; Yim, K.; Chong, Y. Current trend of artificial intelligence patents in digital pathology: A systematic evaluation of the patent landscape. *Cancers* **2022**, *14*, 2400. [\[CrossRef\]](#)
24. Alam, M.R.; Abdul-Ghafar, J.; Yim, K.; Thakur, N.; Lee, S.H.; Jang, H.-J.; Jung, C.K.; Chong, Y. Recent applications of artificial intelligence from histopathologic image-based prediction of microsatellite instability in solid cancers: A systematic review. *Cancers* **2022**, *14*, 2590. [\[CrossRef\]](#)
25. Dey, P. Artificial neural network in diagnostic cytology. *CytoJournal* **2022**, *19*, 27. [\[CrossRef\]](#)
26. Shidham, V.B. Metastatic carcinoma in effusions. *CytoJournal* **2022**, *19*, 4. [\[CrossRef\]](#)
27. Pereira, T.C.; Saad, R.S.; Liu, Y.; Silverman, J.F. The diagnosis of malignancy in effusion cytology: A pattern recognition approach. *Adv. Anat. Pathol.* **2006**, *13*, 174–184. [\[CrossRef\]](#)
28. Chowdhuri, S.R.; Fetsch, P.; Squires, J.; Kohn, E.; Filie, A.C. Adenocarcinoma cells in effusion cytology as a diagnostic pitfall with potential impact on clinical management: A case report with brief review of immunomarkers. *Diagn. Cytopathol.* **2014**, *42*, 253–258. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Shidham, V.B. The panorama of different faces of mesothelial cells. *CytoJournal* **2021**, *18*, 31. [\[CrossRef\]](#)
30. Ashton, P.; Hollingsworth, A., Jr.; Johnston, W. The cytopathology of metastatic breast cancer. *Acta Cytol.* **1975**, *19*, 1–6.
31. Borst, M.J.; Ingold, J.A. Metastatic patterns of invasive lobular versus invasive ductal carcinoma of the breast. *Surgery* **1993**, *114*, 637–642. [\[PubMed\]](#)
32. Martinez, V.; Azzopardi, J. Invasive lobular carcinoma of the breast: Incidence and variants. *Histopathology* **1979**, *3*, 467–488. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Li, C.I.; Anderson, B.O.; Porter, P.; Holt, S.K.; Daling, J.R.; Moe, R.E. Changing incidence rate of invasive lobular breast carcinoma among older women. *Cancer Interdiscip. Int. J. Am. Cancer Soc.* **2000**, *88*, 2561–2569. [\[CrossRef\]](#)
34. Fisher, E.R.; Gregorio, R.M.; Fisher, B.; Redmond, C.; Vellios, F.; Sommers, S.C. The pathology of invasive breast cancer. A syllabus derived from findings of the National Surgical Adjuvant Breast Project (protocol no. 4). *Cancer* **1975**, *36*, 1–85. [\[CrossRef\]](#)
35. Su, D.; Zhang, H.; Chen, H.; Yi, J.; Chen, P.-Y.; Gao, Y. Is Robustness the Cost of Accuracy?—A Comprehensive Study on the Robustness of 18 Deep Image Classification Models. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 631–648.
36. De Logu, F.; Ugolini, F.; Maio, V.; Simi, S.; Cossu, A.; Massi, D.; Group, I.A.f.C.R.S.; Nassini, R.; Laurino, M. Recognition of cutaneous melanoma on digitized histopathological slides via artificial intelligence algorithm. *Front. Oncol.* **2020**, *10*, 1559. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Pramanik, R.; Biswas, M.; Sen, S.; de Souza Júnior, L.A.; Papa, J.P.; Sarkar, R. A fuzzy distance-based ensemble of deep models for cervical cancer detection. *Comput. Methods Programs Biomed.* **2022**, *219*, 106776. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Whitney, J.; Corredor, G.; Janowczyk, A.; Ganesan, S.; Doyle, S.; Tomaszewski, J.; Feldman, M.; Gilmore, H.; Madabhushi, A. Quantitative nuclear histomorphometry predicts oncotype DX risk categories for early stage ER+ breast cancer. *BMC Cancer* **2018**, *18*, 610. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Beijndorf, B.E.; Veta, M.; Van Diest, P.J.; Van Ginneken, B.; Karssemeijer, N.; Litjens, G.; Van Der Laak, J.A.; Hermsen, M.; Manson, Q.F.; Balkenhol, M. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *Jama* **2017**, *318*, 2199–2210. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Lu, C.; Xu, H.; Xu, J.; Gilmore, H.; Mandal, M.; Madabhushi, A. Multi-pass adaptive voting for nuclei detection in histopathological images. *Sci. Rep.* **2016**, *6*, 33985. [\[CrossRef\]](#)
41. Couture, H.D.; Williams, L.A.; Geradts, J.; Nyante, S.J.; Butler, E.N.; Marron, J.; Perou, C.M.; Troester, M.A.; Niethammer, M. Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype. *NPJ Breast Cancer* **2018**, *4*, 30. [\[CrossRef\]](#)
42. Al-Kofahi, Y.; Lassoued, W.; Lee, W.; Roysam, B. Improved automatic detection and segmentation of cell nuclei in histopathology images. *IEEE Trans. Biomed. Eng.* **2009**, *57*, 841–852. [\[CrossRef\]](#)
43. Basavanahally, A.N.; Ganesan, S.; Agner, S.; Monaco, J.P.; Feldman, M.D.; Tomaszewski, J.E.; Bhanot, G.; Madabhushi, A. Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology. *IEEE Trans. Biomed. Eng.* **2009**, *57*, 642–653. [\[CrossRef\]](#)
44. Senaras, C.; Sahiner, B.; Tozbikian, G.; Lozanski, G.; Gurcan, M.N. Creating synthetic digital slides using conditional generative adversarial networks: Application to Ki67 staining. In Proceedings of the Medical Imaging 2018: Digital Pathology, Houston, TX, USA, 11–12 February 2018; pp. 15–20.
45. Hossain, M.S.; Hanna, M.G.; Uraoka, N.; Nakamura, T.; Edelweiss, M.; Brogi, E.; Hameed, M.R.; Yamaguchi, M.; Ross, D.S.; Yagi, Y. Automatic quantification of HER2 gene amplification in invasive breast cancer from chromogenic in situ hybridization whole slide images. *J. Med. Imaging* **2019**, *6*, 047501. [\[CrossRef\]](#)
46. Abdollahi, J.; Davari, N.; Panahi, Y.; Gardaneh, M. Detection of Metastatic Breast Cancer from Whole-Slide Pathology Images Using an Ensemble Deep-Learning Method: Detection of Breast Cancer using Deep-Learning. *Arch. Breast Cancer* **2022**, *9*, 364–376. [\[CrossRef\]](#)
47. Ren, W.; Zhu, Y.; Wang, Q.; Jin, H.; Guo, Y.; Lin, D. Deep Learning-Based Classification and Targeted Gene Alteration Prediction from Pleural Effusion Cell Block Whole-Slide Images. *Cancers* **2023**, *15*, 752. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Barwad, A.; Dey, P.; Susheilia, S. Artificial neural network in diagnosis of metastatic carcinoma in effusion cytology. *Cytom. Part B: Clin. Cytom.* **2012**, *82*, 107–111. [\[CrossRef\]](#)

49. Tosun, A.B.; Yergiyev, O.; Kolouri, S.; Silverman, J.F.; Rohde, G.K. Detection of malignant mesothelioma using nuclear structure of mesothelial cells in effusion cytology specimens. *Cytom. Part A* **2015**, *87*, 326–333. [\[CrossRef\]](#)
50. Arrieta, A.B.; Díaz-Rodríguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; García, S.; Gil-López, S.; Molina, D.; Benjamins, R. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [\[CrossRef\]](#)
51. Mohseni, S.; Zarei, N.; Ragan, E.D. A multidisciplinary survey and framework for design and evaluation of explainable AI systems. *ACM Trans. Interact. Intell. Syst. TiiS* **2021**, *11*, 1–45. [\[CrossRef\]](#)
52. Tosun, A.B.; Pullara, F.; Becich, M.J.; Taylor, D.; Fine, J.L.; Chennubhotla, S.C. Explainable AI (xAI) for anatomic pathology. *Adv. Anat. Pathol.* **2020**, *27*, 241–250. [\[CrossRef\]](#)
53. Xie, Q.; Luong, M.-T.; Hovy, E.; Le, Q.V. Self-training with noisy student improves imagenet classification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 10687–10698.
54. Pham, H.; Dai, Z.; Xie, Q.; Le, Q.V. Meta pseudo labels. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Virtual, 19–25 June 2021; pp. 11557–11568.
55. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
56. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019; pp. 6105–6114.

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.