



## OPEN ACCESS

## EDITED BY

Sharon R. Pine,  
University of Colorado Anschutz Medical  
Campus, United States

## REVIEWED BY

Antonella Argentiero,  
National Cancer Institute Foundation (IRCCS),  
Italy  
Stefano Marletta,  
University of Verona, Italy

## \*CORRESPONDENCE

Sungmoon Jeong  
✉ jeongsm00@gmail.com

†These authors have contributed equally  
to this work and share first authorship

RECEIVED 29 November 2023

ACCEPTED 31 May 2024

PUBLISHED 05 July 2024

## CITATION

Kim J, Bae S, Yoon S-M and Jeong S (2024)  
Roadmap for providing and leveraging  
annotated data by cytologists in the PDAC  
domain as open data: support for AI-based  
pathology image analysis development and  
data utilization strategies.  
*Front. Oncol.* 14:1346237.  
doi: 10.3389/fonc.2024.1346237

## COPYRIGHT

© 2024 Kim, Bae, Yoon and Jeong. This is an  
open-access article distributed under the terms  
of the [Creative Commons Attribution License  
\(CC BY\)](#). The use, distribution or reproduction  
in other forums is permitted, provided the  
original author(s) and the copyright owner(s)  
are credited and that the original publication  
in this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Roadmap for providing and leveraging annotated data by cytologists in the PDAC domain as open data: support for AI-based pathology image analysis development and data utilization strategies

Jongkwang Kim<sup>1†</sup>, Sumok Bae<sup>1†</sup>, Seong-Mi Yoon<sup>2</sup>  
and Sungmoon Jeong<sup>1\*</sup>

<sup>1</sup>Department of Medical Informatics, School of Medicine, Kyungpook National University, Daegu, Republic of Korea, <sup>2</sup>Department of Pathology at the Medical Center, Medical Research Institute, ORTHOTECH, Daegu, Republic of Korea

Pancreatic cancer is one of the most lethal cancers worldwide, with a 5-year survival rate of less than 5%, the lowest of all cancer types. Pancreatic ductal adenocarcinoma (PDAC) is the most common and aggressive pancreatic cancer and has been classified as a health emergency in the past few decades. The histopathological diagnosis and prognosis evaluation of PDAC is time-consuming, laborious, and challenging in current clinical practice conditions. Pathological artificial intelligence (AI) research has been actively conducted lately. However, accessing medical data is challenging; the amount of open pathology data is small, and the absence of open-annotation data drawn by medical staff makes it difficult to conduct pathology AI research. Here, we provide easily accessible high-quality annotation data to address the abovementioned obstacles. Data evaluation is performed by supervised learning using a deep convolutional neural network structure to segment 11 annotated PDAC histopathological whole slide images (WSIs) drawn by medical staff directly from an open WSI dataset. We visualized the segmentation results of the histopathological images with a Dice score of 73% on the WSIs, including PDAC areas, thus identifying areas important for PDAC diagnosis and demonstrating high data quality. Additionally, pathologists assisted by AI can significantly increase their work efficiency. The pathological AI guidelines we propose are effective in developing histopathological AI for PDAC and are significant in the clinical field.

## KEYWORDS

pancreatic ductal adenocarcinoma, deep convolutional neural network, whole slide image, histopathology, supervised learning, dice score, high quality

## 1 Introduction

Pancreatic cancer is one of the most lethal malignancies, with a five-year survival rate of approximately 5%–9%, which has remained virtually unchanged since the 1960s (1, 2). More than 85% of pancreatic cancers are adenocarcinomas (PDACs), which arise from the pancreatic duct epithelium in the head, body, and tail of the pancreas (2). The head of the pancreas is the most common site of PDAC. PDAC is not effectively preventable or screened for and is associated with 98% of expected lifetime loss and 30% of disability-adjusted life years (3, 4). In addition, recent studies have suggested that a molecular subgroup of PDAC characterized by bone metastases may have an unfavorable outcome, suggesting that this subgroup of patients may have distinctive prognostic features and may be potential candidates for specific targeted therapies (5). Further molecular-level research is needed to explore this, which could contribute to better PDAC treatments and AI development. Nevertheless, research funding for PDAC remains markedly lower than for other cancer types; the European Commission and the United States Congress designated it as a neglected cancer (3). The rapid progression and high frequency of pancreatic cancer distant metastases pose a challenge in pathology, where the misdiagnosis consequences can be severe (6–8). Multidetector computed tomography, magnetic resonance imaging, and endoscopic ultrasound are recommended initial imaging modalities for timely PDAC diagnosis (9). The gold standard for clinical diagnosis is the histopathologic imaging assessment by a pathologist (10); however, during the diagnostic process, pathologists must repeatedly zoom in and out of the field of view, determine areas critical for diagnosis, and classify them according to features because of the large slide sizes. Thus, the manual analysis of pathological slides is extremely time-consuming and labor-intensive and may miss important diagnostic information (11). In modern medicine, artificial intelligence (AI) is emerging as a revolutionary technology that can help make faster and more accurate decisions in the medical field. This has led to its application in a wide range of medical fields, including radiology, pathology, pharmacology, infectious diseases, and personalized decision-making, and it has shown the potential to improve current standards of care (12).

Digital pathology has become a rapid and convenient standard of practice in pathology, as it allows for the management and analysis of data from digitized specimen slides using high-resolution digital imaging (13). With the significant advances in artificial intelligence (AI) algorithms and data management capabilities, combining digital pathology and AI has emerged as a front-runner in modern clinical practice (13, 14). The number of publications on AI for clinical decision-making in oncology has increased exponentially in recent years (15).

In surgical pathology, AI can be used to evaluate lymph nodes (LNs) for the presence of metastatic disease by automatically identifying metastatic cancer cells in whole slide images (WSI), which can help in the staging of cancer patients and the prediction of prognosis (16). Other examples include the use of pathology AI for microbial identification to supplement manual microscopy, which is a time-consuming process for the efficient identification of many microbes (17). Digital pathology has shown promising

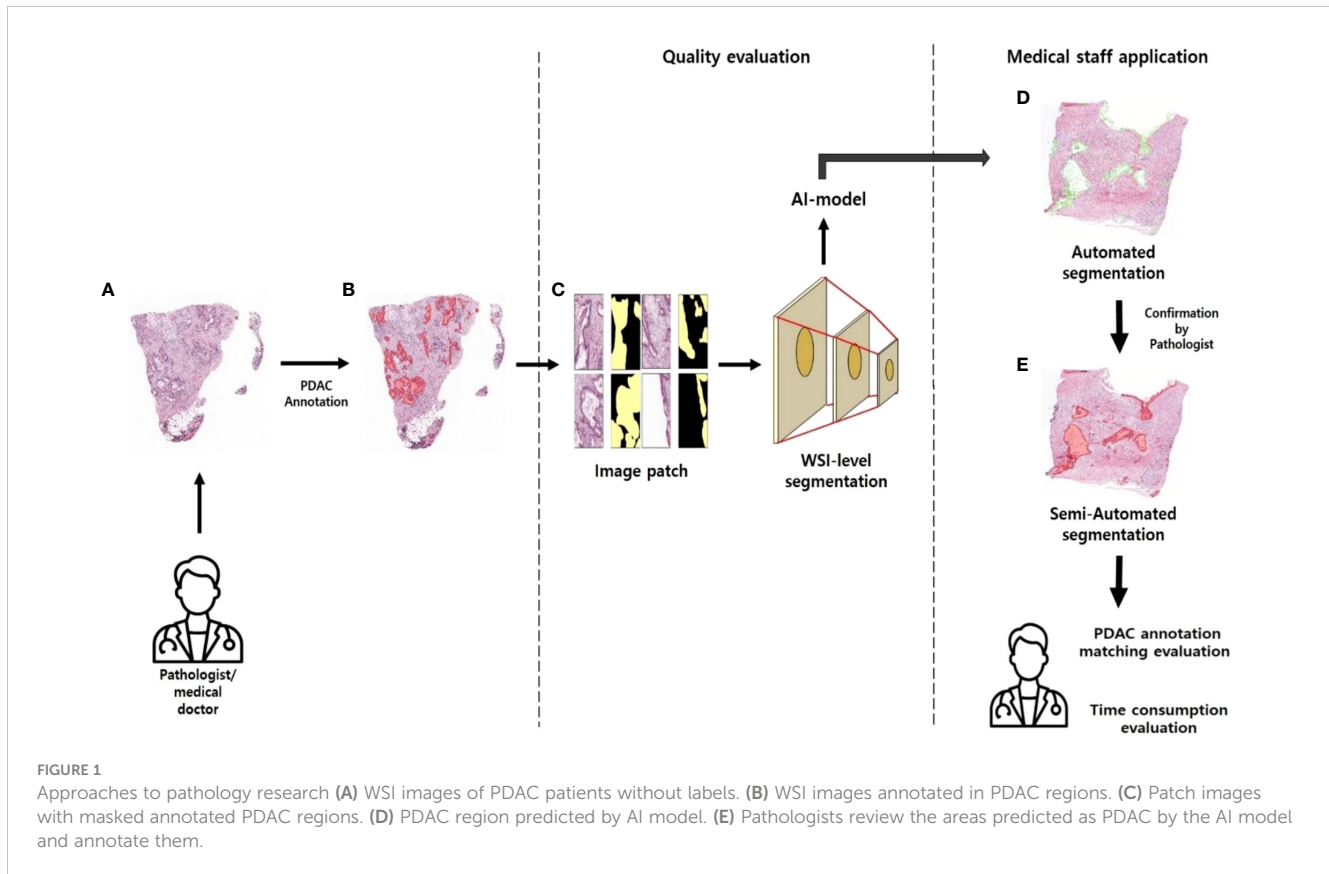
results with regard to the digital evaluation of cytological samples, with the development of portable mobile devices such as smartphones that allow pathologists to examine both surgical and cytological samples (18). The development of digital pathology and AI in various pathology fields has the potential to improve the quality of healthcare in resource-limited settings, where there is a shortage of specialized healthcare professionals. Digital pathology systems can enable remote patient samples to be easily sent to experts, and AI-based automated analysis can be used. Whole slide imaging (WSI) is a major innovation in pathology, which digitizes glass slides to improve pathology workflow, reproducibility, availability of educational materials, outreach to underserved populations, and inter-institutional collaboration (19). However, due to the limited computing resources available currently, performing image analysis using whole slide images (WSIs) as input to convolutional neural network (CNN) classification models (20), which are currently widely used in image-based AI, remains challenging. Here, we adopted a novel scheme to realize whole slide analysis while preserving the high resolution and accuracy of pathological slide analysis. Deep learning approaches to WSI analysis have major limitations: labeled data for histopathology images are particularly scarce; WSIs are large; experienced pathologists must invest significant time and cost to annotate them using specialized labeling tools; and pathological images have rich background regions (e.g., vessels or lymphocytes) that can affect the analysis (21). Here, we provide high-quality data hand-drawn by Hepatobiliary-pancreatic pathologists in an open-access manner—so that anyone can easily use it—to address the abovementioned issues. We applied basic supervised learning (SL), already open to the public, as a data quality assessment and application method. SL algorithms rely on a training dataset that depends on ground truth labels provided by human annotations for input variables (i.e., features) to predict the corresponding output, allowing SL models to mimic expert annotators in predicting features of unknown inputs (22). This study suggests an effective application method for the quality assessment of open-annotation data provided by Hepatobiliary-pancreatic pathologists and the development of pathology AI (Figure 1).

## 2 Materials and methods

### 2.1 Data collection

#### 2.1.1 Dataset

The primary data set comprises pathology images of Clinical Proteomic Tumor Analysis Consortium (CPTAC) patients collected and publicly released by The Cancer Imaging Archive to enable researchers to investigate cancer phenotypes that may be correlated with the corresponding proteomic, genomic, and clinical data. Pathology images are collected as part of the CPTAC qualification workflow (23, 24). The data collection includes hospitals from three institutions (Beaumont Health System, Royal Oak, MI; Boston Medical Center, Boston, MA; St. Joseph's Hospital and Medical Center, Phoenix, AZ) and medical research institutes from three institutions (International Institute



for Molecular Oncology, Poznań, Poland; University of Calgary, Alberta, Canada; Cureline, Inc. team and clinical network, Brisbane, CA) and includes subjects from the National Cancer Institute's CPTAC Pancreatic Adenocarcinoma (CPTAC-PDA) cohort. All CPTAC cohorts are released as single-cohort data sets or, where appropriate, are split into discovery and validation. For this study, we selected 11 high-resolution WSIs of cancerous pancreatic tissue samples as the dataset. Each sample was collected via surgical resection and stained with hematoxylin and eosin (H&E) and stored as high-resolution WSIs (Figure 2A). The inclusion criteria for patient samples were as follows: organ: pancreas; tumor site: head, body, tail; disease: PDAC; patient age: 40–80 years old; and staining type: H&E. WSIs are typically about 100 MB, with a resolution of about  $10,000 \times 10,000$  pixels, but the size can vary between WSIs. The data utilization method we present leverages our provided labeled data to generate tiles to train segmentation models.

### 2.1.2 Data annotation

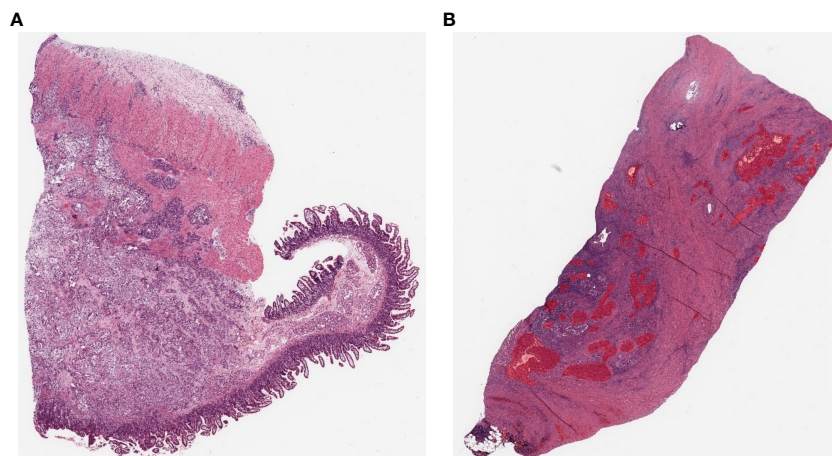
In this study, model training is conducted using an SL approach. Training, test, and validation sets were prepared to train and validate the PDAC detection algorithm on labeled WSIs. All annotations for the annotation dataset were validated by a common golden standard of at least two double board-certified cytopathologists & Hepatobiliary-pancreatic pathologists who agreed on the annotation placement (Figure 3). The WSI information for all annotation datasets is listed in Table 1. To generate ground truth SL labels, human encoders hand-drew

annotations using the open-source pathology and bioimage analysis software QuPath (v0.1.3.5). For each WSI, the PDAC regions were annotated by Hepatobiliary-pancreatic pathologists with a red line (Figure 2B).

## 2.2 Data preprocessing

### 2.2.1 WSIs to patch images

In pathology diagnosis, high-resolution images are necessary for accurate diagnosis. Most WSIs are  $10,000 \times 10,000$  pixels or larger and are typically stored as SVS files. However, directly using such high-resolution images in a deep learning model is not feasible due to the GPU memory limitations, which prevents implementing WSI convolutional operations. Therefore, it is necessary to reduce the images' size. However, directly downsizing high-resolution images to low-resolution ones can result in losing important features. To address this problem, a patch-based approach was adopted, allowing us to maintain the original resolution while dividing the image into smaller patches. The patch dataset comprised three types. First, each image was divided into partially overlapping patches for the training dataset to enhance the model's learning capability. Second, the validation and test datasets used in the model's quantitative evaluation did not require overlapping patch images; each image was divided into non-overlapping patches. These patch data types required the original and mask images to be divided into patches in the same manner, which was achieved using the scripting function provided by QuPath also used for



**FIGURE 2** Example pathology slide images of PDAC patients (A) Open H&E-stained WSI image. (B) WSI image after hand-drawn annotation.

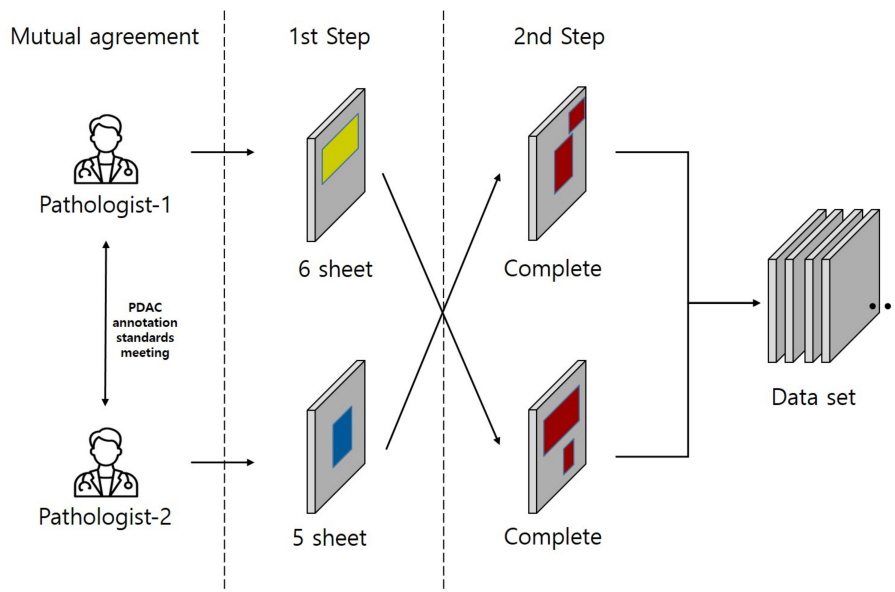
annotation. Last, the test dataset used in clinical evaluation required merging the patch images back into a single large image during the postprocessing stage when the WSI was divided into patch images. The PyHIST library was utilized, which outputs the x and y coordinates of each patch image during the patch division process (25), allowing tracking of each patch’s spatial information and reconstructing the original image by aligning the patches based on their respective coordinates. These patches were saved as PNG files of 512 × 512 pixels and maintained the highest resolution of the WSI, which is 20X [0.5 microns per pixel (MPP)], to prevent

resolution degradation. The MPP value was calculated as shown in Equation 1.

$$MPP = \frac{1 \mu m}{pixel} \tag{1}$$

### 2.2.2 Augmentation

Data augmentation techniques are essential in data preprocessing to prevent overfitting and improve the AI models’ performance during training. Various data augmentation methods



**FIGURE 3** This study was conducted on a total of 11 Whole Slide Images (WSIs) where two pathologists agreed on a common annotation range for Pancreatic Ductal Adenocarcinoma (PDAC), also referred to as the gold standard, from each of 6 and 5 WSIs respectively. The annotation for PDAC was carried out in two steps. In the first step, the two pathologists individually annotated the images. In the second step, the pathologist who annotated in the first step had their work reviewed by the other pathologist. This second pathologist added any missed PDAC areas to the annotation, thus completing the annotation process. Yellow region: Annotation by Pathologist 1, Blue region: Annotation by Pathologist 2, Red region: Final completed annotation.

TABLE 1 Clinical information and characteristics of patients with pathology slides.

Slide ID	Tumor	Topographic Site	Tumor Site	THT	PTN	PTC	PN	Gender	Age
C3L-00017-22	PDA	Pancreas	Head	Ductal adenocarcinoma	20	100	0	Male	60-70
C3L-00017-23	PDA	Pancreas	Head	Ductal adenocarcinoma	20	100	0	Male	60-70
C3L-00102-21	PDA	Pancreas	Head	Ductal adenocarcinoma	20	98	0	Male	40-50
C3L-00102-22	PDA	Pancreas	Head	Ductal adenocarcinoma	10	100	0	Male	40-50
C3L-01637-21	PDA	Pancreas	Body	Ductal adenocarcinoma	20	90	5	Female	70-80
C3L-00277-21	PDA	Pancreas	Tail	Ductal adenocarcinoma	70	90	10	Male	60-70
C3L-00277-22	PDA	Pancreas	Tail	Ductal adenocarcinoma	70	90	0	Male	60-70
C3L-00277-23	PDA	Pancreas	Tail	Ductal adenocarcinoma	70	90	0	Male	60-70
C3L-01453-22	PDA	Pancreas	Tail	Ductal adenocarcinoma	60	90	0	Male	60-70
C3L-01662-23	PDA	Pancreas	Head	Ductal adenocarcinoma	70	100	0	Female	60-70
C3L-00819-22	PDA	Pancreas	Head	Ductal adenocarcinoma	60	100	0	Male	70-80

THT, Tumor Histological Type; PTN, Percent Tumor Nuclei; PTC, Percent Total Cellularity; PN, Percent Necrosis.

are available, and using appropriate augmentation techniques for each task is essential. For our task, which involved segmentation, applying the same augmentation techniques to the original and mask images was crucial as they were matched. Therefore, we implemented effective image transformations using the Albumentations library, which provides most of the commonly used augmentation techniques in deep learning while simultaneously transforming the original and mask images (26). We normalized the images for image transformations and then added noise through ColorJitter. Additionally, we applied various data augmentation techniques by randomly choosing one of three methods: HorizontalFlip, RandomRotate, and VerticalFlip. This approach allowed us to augment the data diversely. By implementing these image transformations, we created a training environment for the AI model to effectively learn the features of the target region, even in extreme conditions.

### 2.3 AI model architecture

Accurate segmentation of histopathological images is increasingly recognized as a key challenge in diagnosis and treatment. An appropriate deep learning model is essential for accurately segmenting histopathological features with various sizes and characteristics. Therefore, we adopted the DeepLabV3+ model and used ResNet18 as its backbone. Additionally, we employed transfer learning by applying pretrained weights from ImageNet to ResNet18,

enabling the model to learn the general features of the images. Subsequently, we trained the model using histopathological images relevant to the main task and performed fine-tuning for the histopathological features. ResNet18 is a well-known model for image feature extraction and effectively overcomes the gradient vanishing problem when training deep neural networks through residual connections (27). This characteristic contributes to effectively extracting histopathological features with various sizes and complexities. Moreover, in DeepLabV3+, the features extracted from ResNet18 are utilized using the Atrous Spatial Pyramid Pooling (ASPP) method. ASPP employs parallel convolution layers with different dilate rates to capture receptive fields of various sizes (28), allowing accurate target classification at different scales without losing spatial information. In particular, for model training using histopathological images where features of various sizes are important, ASPP can comprehensively recognize features of various sizes, enabling more accurate training of the model. Therefore, we adopted DeepLabV3+ with ASPP as the base model and upsampled the features through the decoder part of DeepLabV3+. This process involved restoring the low-resolution feature maps to their original input size, thus obtaining the segmentation results as the final step of the model.

### 2.4 Data postprocessing

Unlike typical deep learning segmentation tasks, deep learning on WSIs requires a data preprocessing step to convert WSIs into

patch images. Additionally, during model training, unnecessary background images need to be removed. As a result, the mask images predicted from the model are output as patch images without including background images, similar to the input images. In typical quantitative AI model evaluation processes, the generated patch images and the corresponding label patch images can be compared using evaluation metrics to evaluate the model's performance. However, in our study, we conducted quantitative and qualitative evaluations to assess the effectiveness of AI assistance in histopathological diagnosis scenarios. Therefore, visualizing the mask patch images generated by the model to assist pathologists is essential. It requires a postprocessing step that comprises two main processes. First, the binary mask patch images obtained from the model's predictions are overlaid onto the original image patches. Second, the overlaid patch images are combined into a single large-sized image. The 1-channel grayscale mask images are converted into 3-channel RGB images while using distinctive colors to make them visually stand out. Then, utilizing the x and y coordinates, which represent the location information of each patch obtained during the image segmentation, the mask patch images are accurately overlaid onto the corresponding positions of the original WSIs. By merging the patch images into an image of the same size as the original image, we prevent a decrease in resolution. The images obtained through the postprocessing step are used for clinical evaluation.

## 3 Experiments and results

### 3.1 Dataset description

The patch dataset comprised three types. For the training dataset used in model training, each patch image had partial overlap and was generated by dividing 23,239 images from 8 WSIs. The mask patch images, corresponding to the patch images, were also created, resulting in 23,239 mask patch images. The validation and test datasets used in the model's quantitative evaluation were generated using the same method as the training dataset but without overlapping patch images. Therefore, they were composed of fewer patch images. The validation dataset comprised 630 patch images (with corresponding mask patch images) generated from 1 WSI, and the test dataset included 1,202 patch images (with corresponding mask patch images) generated from 2 WSIs. In total, the validation and test datasets were composed of 25,071 patch images (with corresponding mask patch images) generated from 11 WSIs. The detailed distribution of this dataset is listed in [Table 2](#).

Additionally, the test dataset used in the model's qualitative evaluation was generated from the same WSIs as the test dataset used in the quantitative evaluation. It comprised 1,214 patch images, and a data table containing coordinate information for each patch image was also created for postprocessing purposes. All

TABLE 2 Dataset used for model training and quantitative evaluation.

	WSIs	Image patches	Mask patches
Train	8	23,239	23,239
Validation	1	630	630
test	2	1,202	1,202

patch datasets comprised  $512 \times 512$ -pixel images with the background removed.

### 3.2 Training and evaluation metrics

In this study, we conducted experiments using two GPUs, namely NVIDIA QUADRO RTX 6000 and NVIDIA TITAN RTX, in parallel, with CUDA 11.6 and cuDNN 8. A total of 48GB of GPU memory, with each GPU having 24GB, was utilized for the experiments. The deep learning framework used was PyTorch 1.13.1. During the AI model training process, small batch training iterations were used with a batch size set to 128, and the total number of training epochs was set to 50. The training was configured to terminate early if the validation Dice score did not improve for 30 consecutive epochs. The training time took about 3 hours. We utilized the Adam optimization algorithm with a learning rate set to  $1e^{-4}$  and used the Dice score as an evaluation metric. Dice score is one of the most common methods for evaluating image segmentation performance in medical imaging (29); it measures the similarity between the predicted mask by the model and the ground truth label mask. The Dice score value was calculated as shown in [Equation 2](#). Dice loss was employed, commonly used as a loss function in image segmentation tasks, was used for the loss function, aiming to train the model to maximize the Dice score. The Dice loss value was calculated as shown in [Equation 3](#).

$$Dice = \frac{2 \times Area\ of\ Overlap}{Total\ Sum\ of\ Pixels\ in\ All\ Areas} \quad (2)$$

$$Dice\ Loss = 1 - \frac{2 \times Area\ of\ Overlap}{Total\ Sum\ of\ Pixels\ in\ All\ Areas} \quad (3)$$

### 3.3 AI model results

In a quantitative evaluation of the AI model based on DeepLabV3+, we achieved a specificity of 96.37%, an accuracy of 93.77%, and a Dice score of 73% ([Table 3](#)). For qualitative evaluation, we visualized the predicted segmentation for each patch image in the test dataset ([Figure 4](#)). The AI model has achieved high performance and demonstrated the ability to predict and segment the lesion areas ([Figure 5A](#)). Compared to

**TABLE 3** Metrics for the DeepLabV3+ model's various scores for PDAC on WSIs.

Metrics	Score (%)
DICE	73.20
Specificity	96.37
Accuracy	93.77
Sensitivity	75.22
Precision	73.04

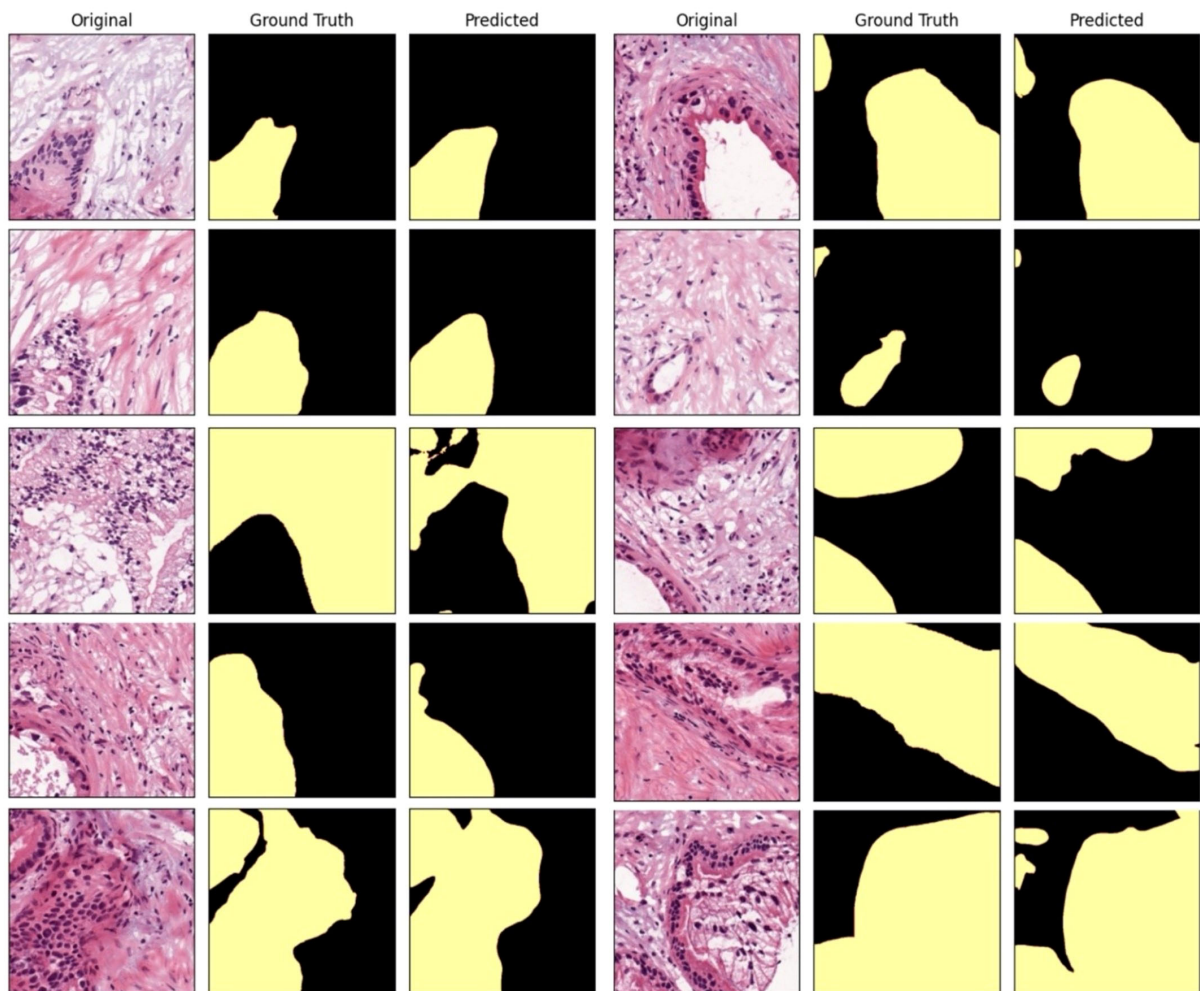
the ground truth, it excelled in representing PDAC regions of various shapes, especially in the main pancreatic and interlobular ducts. However, the accuracy was lower due to the false positive rate, as the predicted region recognized an area larger than the actual PDAC annotation or recognized some non-PDAC areas. Visualizing the whole image through postprocessing, converting patch images to WSIs, confirmed the consistency with the Hepatobiliary-pancreatic pathologist's annotation (Figure 5B) level. In addition, our test results were confirmed at low and high

magnifications (Figure 5). The AI model trained with the annotated WSIs data we provided displayed high sensitivity to PDAC, the cancerous area of the pancreas.

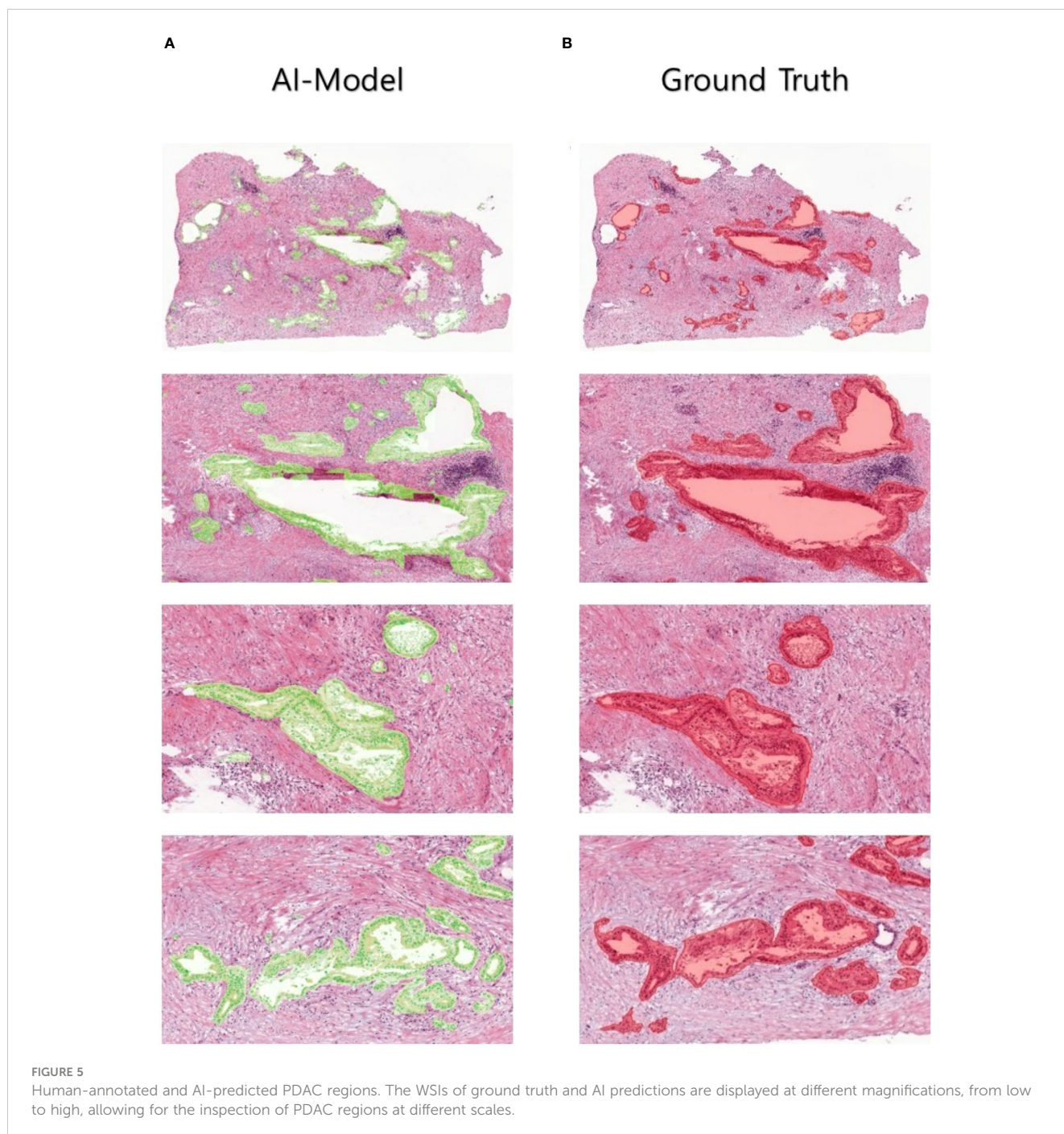
### 3.4 Technical application

#### 3.4.1 Progress assessment

To compare the annotation rates of PDAC regions in WSIs, a pathologist hand-annotated PDAC regions in two different WSIs under two experimental conditions. We performed repeated experiments in which the pathologist annotated two WSI images in four consecutive cycles, one for each image, in the absence and presence of AI model assistance. Each cycle lasted 15 min, with a 3-min break, timed using the iPhone 13 stopwatch. We used an evaluation metric called the sensitivity to evaluate the area annotated by the pathologist within a limited time compared to the ground truth area in each cycle. The sensitivity value is calculated as shown in Equation 4). When the pathologist annotated the PDAC regions in WSIs without AI assistance, the rate of the overall annotation achieved a relatively low sensitivity



**FIGURE 4** Comparison between the AI-predicted segmented patch images in the test dataset and the ground truth.



average of 44.64% in the final four cycles (Table 4). In contrast, when the pathologist confirmed and annotated the PDAC regions identified by the AI model using WSI-level segmentation, the annotation rate was overwhelmingly higher than without AI model assistance from the first cycle, and the overall PDAC annotation rate also achieved a significantly high sensitivity average of 85.54% (Table 4). As a result, AI assistance helped achieve a significantly higher annotation rate than the human without AI assistance. We can also expect the annotation accuracy to be significantly higher when the human is assisted by the AI model. We also visualized the images to increase the understanding of these clinical trial results (Figure 6).

$$\text{Sensitivity} = \frac{2 \times \text{Area of Overlap}}{\text{Total Sum of Pixels in Ground Truth Area}} \quad (4)$$

## 4 Discussion

In this study, we demonstrated the AI potential to aid the diagnosis and prognostic assessment of PDAC, a deadly cancer classified as a public health emergency. Although the majority of PDACs occur in the head of the pancreas, the WSI dataset used in this study contains WSIs with patterns of various PDAC regions that



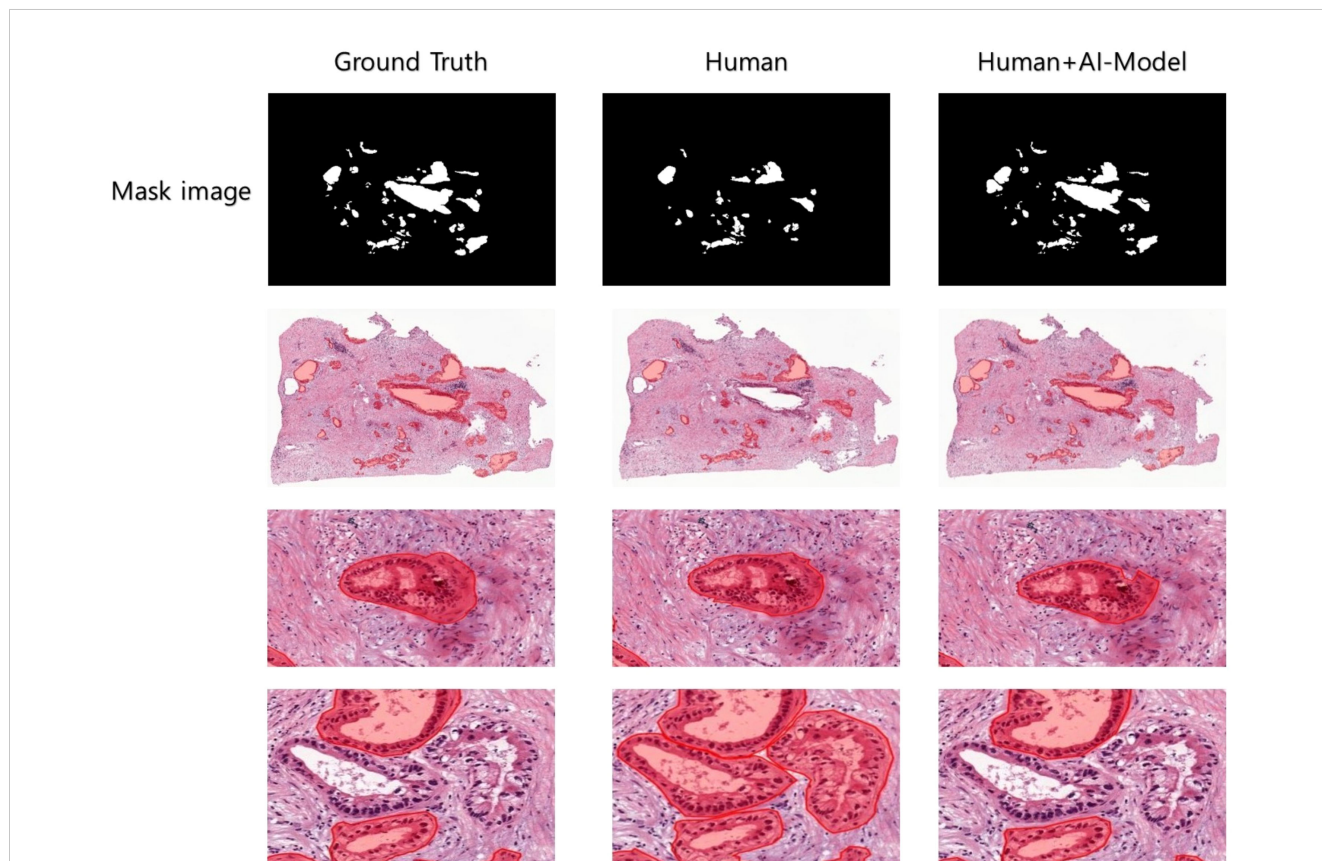
**TABLE 4** Clinical trial: Sensitivity for PDAC annotation rate in WSIs of humans using humans and AI.

Sensitivity (%)	Test1		Test2	
	AI (O)	AI (X)	AI (O)	AI (X)
1 Cycle	<b>67.26</b>	10.54	<b>80.29</b>	8.16
2 Cycle	<b>81.31</b>	30.97	<b>83.25</b>	23.32
3 Cycle	<b>84.60</b>	35.49	<b>86.07</b>	24.12
4 Cycle	<b>84.65</b>	43.70	<b>86.43</b>	45.59

The bold values are highlighted to emphasize that the Sensitivity performance was significantly higher with AI assistance than without it as cycles progressed.

occur in the body and tail, which are less common than PDACs in the head (30), to explore PDAC in depth. The results represent a significant step forward in AI application to the tissue pathological diagnosis and prognostic assessment of PDAC. The research findings suggest that AI, especially CNN deep learning models, can be effectively used to segment and analyze PDAC tissue pathological WSIs, thereby simplifying and improving the accuracy of PDAC diagnosis. One key aspect highlighted in the study is the challenge posed by the limited access to medical data, especially public pathology data (31). This issue has been a persistent obstacle in pathology AI research. For this study, two pathologists collaboratively

annotated PDAC regions in WSIs (Figure 3), and the WSI data used in the study is publicly available for anyone to use, including high-quality annotations. This approach increases the amount of high-quality data available for training AI models and ensures that these models are trained with reliable and accurate information. SL using a deep CNN architecture to segment 11 annotated PDAC WSIs presented promising results. It displayed high Dice scores on the whole tissue image, including PDAC regions, indicating accurate segmentation, and identified areas important for PDAC diagnosis through image visualization. It also showed high specificity and accuracy with a specificity of 96.37% and an accuracy of 93.77% through a precise analysis. These observations demonstrate our high-quality dataset and suggest that AI can play an essential role as an auxiliary tool to improve the efficiency and accuracy of histopathological analysis. In addition, when the whole image was visualized and patch images were converted to WSIs through post-processing, the performance was not significantly different from the pathologist’s annotations, but some parts of the small pancreatic ducts, intercalated ducts, and intralobular ducts showed false positives. This is an impressive achievement considering the complexity of pancreatic cancer in interpreting tissue pathological images, but it is expected that increasing the number of pathologists and adding training data will minimize false positives while improving the reliability of the data. Visualization techniques, such



**FIGURE 6** Clinical experiments were conducted to visualize the PDAC areas predicted by Human and Human+AI Model in WSI after cycle 4. Human and Human+AI Model drew similar shapes to the correct values, but the Human part made errors in recognizing TIS (carcinoma *in situ*) as PDAC, probably due to decreased concentration and increased fatigue.

as postprocessing techniques that convert patch images back to WSIs, were crucial in validating model performance against expert annotations. In some cases, the AI models achieved high performance, but when visualized and compared with the pathologist's annotations, the AI model recognized areas other than the annotated lesion area. This observation reinforces how essential visualization tools are in evaluating the interpretability of AI models in medical imaging tasks. Our study results indicated that SL deep learning models trained on hand-drawn annotated WSIs displayed high sensitivity for malignant pancreatic areas (i.e., PDAC areas). One important aspect of this study was to confirm the significant improvement in the efficiency of annotation work by pathologists assisted by AI, as AI provides a user-friendly, intuitive interface that minimizes complex technical content and allows pathologists to focus on pathological findings. When pathologists were assisted by the SL model in annotating PDAC in WSIs, the annotation accuracy of pathologists increased while the area of PDAC regions did not differ significantly from the ground truth, and the average annotation progress rate increased by about 2 times compared to the same time spent, which indicates that the annotation time was significantly reduced. Therefore, AI-assisted pathology interpretation of PDAC can diagnose a large number of clinical specimens quickly and accurately, and a cohort study on the prognosis of patients after diagnosis is needed to consider the survival of patients. In addition, if pathological image data for Acinar Cell Carcinoma and Pancreatic Neuroendocrine Tumors (PNETs), which are very rare pancreatic cancers in addition to PDAC, are collected together and used for pathological AI research, the performance of the model can be evaluated in a more comprehensive range for pancreatic cancer, and the applicability of pathological research is expected to increase significantly. Moreover, previous pathological image AI studies mainly used classification models, but due to the reduced image resolution, it is difficult for pathologists to accurately identify the lesion area predicted by AI, so there are limitations in using AI as an auxiliary tool for diagnosis in the clinical pathology field. However, there are few studies that can compensate for this using segmentation, and in the case of PDAC, which has fewer patient cases than other diseases, the application of segmentation is limited to patch-level segmentation rather than whole-slide images, which limits its use (32). To address these issues, this study provides a clear analysis result that identifies PDAC regions with high resolution at low and high magnifications through segmentation in the whole pathological slide images of PDAC patients, and proves that pathologists in the actual pathological clinical field are assisted by AI models. It has significant value in annotation and diagnosis. In addition, it can contribute to the development of pathology AI for pancreatic cancer by providing high-quality pathology annotation data for free. We used open tissue pathology data from six hospitals and medical research institutions to ensure data diversity. As well as, by continuously uploading public data with PDAC annotations to the [https://github.com/moksu27/PDAC\\_pathological\\_image\\_segmentation](https://github.com/moksu27/PDAC_pathological_image_segmentation), we can resolve the data imbalance for data with a small number of cases. Also, with the increase in data, data diversity can be achieved through external validation using data from various hospitals, preventing overfitting of AI models and reducing bias to improve the generalization performance of AI models and give objectivity. This will increase

the reliability of AI performance for pathologists who will receive direct assistance in the clinical setting, and AI will play the role of an auxiliary tool, or co-pilot, in the pathologist's diagnostic process. Direct diagnosis will still be made after review by a pathologist, so patients will be free from anxiety and prejudice about AI. This is expected to contribute significantly to cost-effectiveness and improved patient outcomes. If our annotated data and AI model manual are used in pathology AI research, AI will be able to assist in the diagnosis of the WHO classification screening reading and 8th-edition AJCC pTNM staging (33) defined by the American Joint Committee on Cancer (AJCC) for PDAC patient slides in clinical practice, and pathologists will be able to quickly and accurately diagnose many clinical specimens through digital pathology. However, several obstacles must be overcome before the results of this study can be applied to actual clinical practice. First, the need for data standardization between hospitals. It is difficult to ensure the compatibility of AI tools because the data format or structure used by each hospital is different, making it difficult to apply AI tools to the clinical field. It is necessary to ensure technical compatibility through standardization of data between hospitals, and systematic integration between medical institutions is required for this. Second, there is the problem of increasing the understanding of medical personnel about AI technology. For medical personnel with a low understanding of AI technology, the use of AI tools may be difficult. To solve this, it is important to support additional promotion and education to enable medical personnel to effectively use AI tools. This will encourage the use of AI tools in multiple institutions and provide a safer and more standardized medical environment. Finally, I would like to emphasize that in order to effectively use AI in pathology interpretation, not only technical development but also institutional structure and education system that support it must develop together. This will be the future research direction of this study, and will play an important role in further expanding the use of AI in the field of pathology.

## 5 Conclusion

This study provides essential insights to develop effective AI solutions for the specific diagnosis of PDAC and significantly contributes to the pathological AI guidelines, which may have broader implications, even within oncology. Making high-quality annotated datasets publicly accessible and applying advanced machine learning techniques, such as SL, can revolutionize our approach to annotating and diagnosing complex diseases, like pancreatic cancer. We also reiterate the importance of public access to high-quality datasets for AI research while encouraging active research in pathology AI to develop more sophisticated models with improved diagnostic capabilities.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

## Ethics statement

Ethical approval was not required for the studies involving humans because using open data. The dataset is collected and publicly released by The Cancer Imaging Archive. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

## Author contributions

JK: Writing – review & editing, Writing – original draft, Project administration, Methodology, Data curation. SB: Validation, Software, Writing – review & editing, Writing – original draft, Data curation. SY: Writing – review & editing, Investigation, Data curation. SJ: Writing – review & editing, Project administration, Investigation, Funding acquisition, Conceptualization.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2023-2020-0-01808) supervised by the IITP (Institute of

Information & Communications Technology Planning & Evaluation) And This research was supported by a grant of 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 number: HR22C1832).

## Acknowledgments

The authors would like to express their sincere gratitude to the National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC) for providing the data necessary for this study.

## Conflict of interest

Author SY was employed by company ORTHOTECH.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Hidalgo M. Pancreatic cancer. *N Engl J Med.* (2010) 362:1605–17. doi: 10.1056/NEJMra0901557
- Rawla P, Sunkara T, Gaduputi V. Epidemiology of pancreatic cancer: global trends, etiol-ogy and risk factors. *World J Oncol.* (2019) 10:10. doi: 10.14740/wjon1166
- Michl P, Löhr M, Neoptolemos JP, Capurso G, Rebours V, Malats N, et al. UEG position paper on pancreatic cancer. Bringing pancre-atic cancer to the 21st century: pre-vent, detect, and treat the disease earlier and better. *United Eur Gastroenterol J.* (2021) 9:860–71. doi: 10.1002/ueg2.12123
- Löhr JM. Pancreatic cancer should be treated as a medical emer-gency. *BMJ.* (2014) 349:g5261. doi: 10.1136/bmj.g5261
- Argentiero A, Calabrese A, Solimando AG, Notaristefano A, Panarelli MG, Brunetti O. Bone metastasis as primary presentation of pancreatic ductal adenocarcinoma: A case report and literature review. *Clin Case Rep.* (2019) 7:1972–6. doi: 10.1002/ccr3.2412
- Dal Molin M, Zhang M, De Wilde RF, Ottenhof N, Rezaee L, Wolfgang C, et al. Very long-term survival following resection for pancreatic cancer is not explained by commonly mutated genes: results of whole-exome sequencing analysis. *Clin Cancer Res.* (2015) 21:1944–50. doi: 10.1158/1078-0432
- Strobel O, Neoptolemos J, Jager D, Buchler MW. Optimizing the outcomes of pancreatic cancer surgery. *Nat Rev Clin Oncol.* (2019) 16:11–26. doi: 10.1038/s41571-018-0112-1
- Yamamoto T, Yagi S, Kinoshita H, Sakamoto Y, Okada K, Uryuhara K, et al. Long-term survival after resection of pancreatic cancer: a single-center retrospective analysis. *World J Gastroenterol.* (2015) 21:262. doi: 10.3748/wjg.v21.i1.262
- McGuigan A, Kelly P, Turkington RC, Jones C, Coleman HG, McCain RS. Pancreatic cancer: A review of clinical diagnosis, epidemiology, treatment and outcomes. *World J Gastroenterol.* (2018) 24:4846–61. doi: 10.3748/wjg.v24.i43.4846
- Xu Y, Jia Z, Wang L-B, Ai Y, Zhang F, Lai M, et al. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC Bioinf.* (2017) 18:1–17. doi: 10.1186/s12859-017-1685-x
- De Matos J, Ataky STM, de Souza Britto A, Soares de Oliveira LE, Lameiras Koerich A. Machine learning methods for histopathological image analysis: A review. *Electronics.* (2021) 10:562. doi: 10.3390/electronics10050562
- Koteluk O, Wartecki A, Mazurek S, Kołodziejczak I, Mackiewicz A. How do machines learn? Artificial intelligence as a new era in medicine. *J Personalized Med.* (2021) 11:32. doi: 10.3390/jpm11010032
- Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. *Lancet Oncol.* (2019) 20:e253–61. doi: 10.1016/S1470-2045(19)30154-8
- Tizhoosh HR, Pantanowitz L. Artificial intelligence and digital pathology: challenges and opportunities. *J Pathol Inform.* (2018) 9:38. doi: 10.4103/jpi.jpi\_53\_18
- Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM.* (2017) 60:84–90. doi: 10.1145/3065386
- Caldonazzi Nicolò, Rizzo PC, Eccher A, Girolami I, Fanelli GNicolò, Naccarato AG, et al. Value of artificial intelligence in evaluating lymph node metastases. *Cancers.* (2023) 15:2491. doi: 10.3390/cancers15092491
- Marletta S, L'Imperio V, Eccher A, Antonini P, Santonicco N, Girolami I, et al. Artificial intelligence-based tools applied to pathological diagnosis of microbiological diseases. *Pathol - Res Pract.* (2023) 243:154362. doi: 10.1016/j.prp.2023.154362
- Santonicco N, Marletta S, Pantanowitz L, Fadda G, Troncone G, Brunelli M, et al. Impact of mobile devices on cancer diagnosis in cytology. *Diagn cytopathology.* (2022) 50:34–45. doi: 10.1002/dc.24890
- Zarella MD, Bowman D, Aeffner F, Farahani N, Xthona A, Absar SF, et al. A practical guide to whole slide imaging: A white paper from the digital pathology association. *Arch Pathol Lab Med.* (2019) 143:222–34. doi: 10.5858/arpa.2018-0343-RA
- Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. ICLR (2015) Conference Track Proceedings, May 7-9, 2015. San Diego, CA, USA.
- Jiang Y, Sui X, Ding Y, Xiao W, Zheng Y, Zhang Y. A semi-supervised learning approach with consistency regularization for tumor histopathological images analysis. *Front Oncol.* (2023) 12:1044026. doi: 10.3389/fonc.2022.1044026

22. Rashidi HH, Tran NK, Betts EV, Howell LP, Green R. Artificial intelligence and machine learning in pathology: the present landscape of supervised methods. *Acad Pathol.* (2019) 6:2374289519873088. doi: 10.1177/2374289519873088
23. National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC). The clinical proteomic tumor analysis consortium pancreatic ductal adenocarcinoma collection (CPTAC-PDA) (Version 14) [Data set]. *Cancer Imaging Arch.* (2018). doi: 10.7937/K9/TCIA.2018.SC20FO18
24. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The cancer imaging archive (TCIA): maintaining and operating a public information repository. *In J Digital Imaging.* (2013) 26:1045–1057. doi: 10.1007/s10278-013-9622-7
25. Muñoz-Aguirre M, Ntasis VF, Rojas S, Guigó R. PyHIST: a histological image segmentation tool. *PloS Comput Biol.* (2020) 16:e1008349. doi: 10.1371/journal.pcbi.1008349
26. Buslaev A, Iglovikov VI, Khvedchenya E, Parinov A, Druzhinin M, Kalinin AA. Albumentations: fast and flexible image augmentations. *Information.* (2020) 11:125. doi: 10.3390/info11020125
27. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *IEEE Xplore (2016) Conference Track Proceedings*, June 27–30, 2016. Las Vegas, NV, USA.
28. Liu M, Fu B, Xie S, He H, Lan F, Li Y, et al. ...Comparison of multi-source satellite images for classifying marsh vegetation using DeepLabV3 Plus deep learning algorithm. *Ecol Indic.* (2021) 125:107562. doi: 10.1016/j.ecolind.2021.107562
29. Crum WR, Camara O, Hill DL. Generalized overlap measures for evaluation and validation in medical image analysis. *IEEE Trans Med Imaging.* (2006) 25:1451–61. doi: 10.1109/TMI.2006.880587
30. Tomasello G, Ghidini M, Costanzo A, Ghidini A, Russo A, Barni S, et al. Outcome of head compared to body and tail pancreatic cancer: a systematic review and meta-analysis of 93 studies. *J Gastrointest Oncol.* (2019) 10:259–69. doi: 10.21037/jgo.2018.12.08
31. Schuurmans M, Alves N, Vendittelli P, Huisman H, Hermans JPANCAIM Consortium. Artificial intelligence in pancreatic ductal adenocarcinoma imaging: A commentary on potential future applications. *Gastroenterology.* (2023) 165:309–16. doi: 10.1053/j.gastro.2023.04.003
32. Janssen BV, Theijse R, Roessel SV, Ruitter RD, Berkel A, Huiskens J, et al. Artificial intelligence-based segmentation of residual tumor in histopathology of pancreatic cancer after neoadjuvant treatment. *Cancers.* (2021) 13:5089. doi: 10.3390/cancers13205089
33. Roessel S, Gyulnara G, Verheij J, Najarian R, Maggino L, Pastena M, et al. International validation of the eighth edition of the american joint committee on cancer (AJCC) TNM staging system in patients with resected pancreatic cancer. *JAMA Surg.* (2018) 153:e183617. doi: 10.1001/jamasurg.2018.3617