

RESEARCH

Open Access



A semi-supervised convolutional neural network for diagnosis of pancreatic ductal adenocarcinoma based on EUS-FNA cytological images

Dong Fang^{1,2}, Yigeng Huang^{3,4}, Suwen Li¹, Chen Shi¹, Junjun Bao¹, Dandan Du⁵, Lanlan Xuan⁶, Leping Ye⁷, Yanping Zhang⁷, ChengLin Zhu⁸, Hailun Zheng⁹, Zhenwang Shi², Qiao Mei^{1*} and Huanqin Wang^{3,4*}

Abstract

Background The cytological diagnostic process of EUS-FNA smears is time-consuming and manpower-intensive, and the conclusion could be subjective and controversial. Moreover, the relative lack of cytopathologists has limited the widespread implementation of Rapid on-site evaluation (ROSE) presently. Therefore, this study aimed to establish an AI system for the detection of pancreatic ductal adenocarcinoma (PDAC) based on EUS-FNA cytological images.

Methods We collected 3213 unified magnification images of pancreatic cell clusters from 210 pancreatic mass patients who underwent EUS-FNA in four hospitals. A semi-supervised CNN (SSCNN) system was developed to distinguish PDAC from Non-PDAC. The internal and external verifications were adopted and the diagnostic accuracy was compared between different seniorities of cytopathologists. 33 images of "Atypical" diagnosed by expert cytopathologists were selected to analyze the consistency between the system and definitive diagnosis.

Results The segmentation indicators Mean Intersection over Union (mIoU), precision, recall and F1-score of SSCNN in internal and external testing sets were 88.3%, 93.21%, 94.24%, 93.68% and 87.75%, 93.81%, 93.14%, 93.48% successively. The PDAC classification indicators of the SSCNN model including area under the ROC curve (AUC), accuracy, sensitivity, specificity, PPV and NPV in the internal testing set were 0.97%, 0.95%, 0.94%, 0.97%, 0.98%, 0.91% respectively, and 0.99%, 0.94%, 0.94%, 0.95%, 0.99%, 0.75% correspondingly in the external testing set. The diagnostic accuracy of senior, intermediate and junior cytopathologists was 95.00%, 88.33% and 76.67% under the binary diagnostic criteria of PDAC and non-PDAC. In comparison, the accuracy of the SSCNN system was 90.00% in the dataset of man-machine competition. The accuracy of the SSCNN model was highly consistent with senior cytopathologists ($\text{Kappa} = 0.853$, $P = 0.001$). The accuracy, sensitivity and specificity of the system in the classification of "atypical" cases were 78.79%, 84.20% and 71.43% respectively.

*Correspondence:

Qiao Mei
meiqiaomq@aliyun.com
Huanqin Wang
hqwang@iim.ac.cn

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Conclusion Not merely tremendous preparatory work was drastically reduced, the semi-supervised CNN model could effectively identify PDAC cell clusters in EUS-FNA cytological smears which achieved analogically diagnostic capability compared with senior cytopathologists, and showed outstanding performance in assisting to categorize “atypical” cases where manual diagnosis is controversial.

Trial registration This study was registered on clinicaltrials.gov, and its unique Protocol ID was PJ-2018-12-17.

Keywords Pancreatic ductal adenocarcinoma, Endoscopic ultrasound-guided fine needle aspiration, Cytology, Artificial intelligence, Convolutional neural network

Introduction

Endoscopic ultrasound-guided fine needle aspiration (EUS-FNA) is currently a reliable and minimally invasive method for diagnosing pancreatic cancer. It can precisely estimate tumor location, size and peripheral vascular infiltration. Most importantly, pathological specimens can be obtained from puncture fluid and provide evidence for the formulation of treatment plans including surgery, palliative chemotherapy, and local treatment. Therefore, EUS-FNA is considered the first-line option for the diagnosis and treatment of pancreatic cancer [1].

Rapid On-Site Evaluation (ROSE) in the presence of pathologists has been confirmed beneficial in improving smear specimen quality, as well as reducing the needle puncturing times and complication rate. Hence, the demand for ROSE by endoscopists gradually increased in these decades [2, 3]. Nevertheless, the relative shortage and heavy workload of cytopathologists due to the ongoing demand for EUS-FNA operations was a common problem in most developing countries. Therefore, artificial intelligence (AI) is expected to alleviate the deficiency of cytopathologists. Additionally, the manual cytology diagnosis was often subjective and controversial, which could be influenced by the experience and fatigue of cytopathologists, leading to inconsistent conclusions among different cytopathologists when dealing with indistinguishable pathological images. AI could obviously overcome the above shortcomings, and has the advantages of convenience, high efficiency, and stable performance, so it has broad application prospects in the field of pathological diagnosis.

Few studies constructed deep learning (DL) models based on EUS-FNA cytological images have reported that the accuracy of pancreatic malignant tumors identification fluctuates between 83.4 and 94.4% [4–6]. For instance, Sohn [4] reported an AI model based on ultra-high resolution whole slide image (WSI), and Zhang [6] presented an AI segmentation system by standardized cell cluster photographs. However, to the best of our knowledge, the majority of studies published previously have built AI models under a fully supervised system, the region of interests (ROI) was extracted on the basis of tremendous manual annotations in the preliminary preparation stage. Meanwhile, the complexity of manual

annotation is aggrandized compared with another medical image because the morphology of ROI on pathological images is complicated and changeable, this makes it difficult to include large-scale samples for research; it has been regarded as one of the predominant obstacles affecting the performance of the model. In contrast, the dominant advantage of a semi-supervised deep learning model is that partial or only a small portion of images needed been labeled beforehand, the model can learn the corresponding algorithm through human-computer interaction to achieve segmentation of the target region, which is free from the dependence on the manually selected feature information to achieve higher accuracy and efficiency, and dramatically reduces the workload of the early manual annotation, so that the artificial intelligence can achieve authentically lightweight and intelligentization.

In addition, cytopathologists tend to deliberately classify suspicious lesions as “atypical” when the morphological changes in the cytoplasm and nucleus are not enough to diagnose malignant tumors according to the guidelines for pancreaticobiliary cytology from the Papanicolaou Society of Cytopathology [7]. While “atypical” is a group of heterogeneous diagnoses that include inflammatory, precancerous lesions and nontypical malignant tumors, it brings considerable confusion to endoscopy and surgeon in clinical decision-making. Artificial intelligence models can read the potential information in images, such as intensity, texture and spectral features, utilizing complex information to train DL models and fulfill the tasks of lesion identification, which may have impressive application potential in the classification of “atypical”. Unfortunately, research reported on this niche field is still rare. Hence, we designed this study to attempt to construct a semi-supervised convolutional neural network (SSCNN) system for the diagnosis of pancreatic ductal adenocarcinoma based on EUS-FNA cytological images.

Materials and methods

Patients

A total of 210 patients with pancreatic mass were admitted to The First Affiliated Hospital of Anhui Medical University, The First Affiliated Hospital of University of Science and Technology of China, Anqing Hospital

Affiliated to Anhui Medical University and The First Affiliated Hospital of Bengbu Medical University from January 2018 to April 2024 were selected (183, 12, 7 and 8 cases in order). Inclusion criteria: (1) Pancreatic mass has been detected by imageological examinations (CT, MRI or ultrasonography) which meet the indication of EUS-FNA described in the guidelines [7]. (2) Complete the EUS-FNA cytology examination with at least one qualified smear (the eligibility criteria was Papanicolaou cytology sample adequacy grade C or above [8]). (3) With a definitive final diagnosis. (4) With complete clinical and follow-up data. Exclusion criteria: (1) With contraindication of EUS-FNA. (2) The quality and the clarity of the smear could not meet the requirements of model construction. (3) The percentage of erythrocyte contamination area on the background of the smear exceeds 50% (contamination grade C [9]). (4) Diagnostic controversy existed on a single smear, with a diagnosis between clinical and cytological or cytological and histological.

This study was approved by the Ethics Committee of The First Affiliated Hospital of Anhui Medical University. It has been registered in clinicaltrials.gov, and its unique Protocol ID was PJ-2018-12-17.

Diagnostic criteria/definitions

Pancreatic masses were classified into PDAC and non-PDAC on the basis of a comprehensive analysis of serological tests, imaging findings, EUS-FNA, postoperative pathology and follow-up outcomes. The diagnostic criteria for PDAC include: (1) Confirmed by postoperative pathological results. (2) Confirmed by cytopathology, liquid-based cytology, histopathology and immunohistochemistry of EUS-FNA/FNB. (3) Imaging (enhanced CT, PET-CT, MRI), serological examination (tumor marker CA199) and other auxiliary examination evidence. (4) Death due to PDAC occurred during the 6-month follow-up period. The diagnostic criteria for non-PDAC include: (1) Non-PDAC diagnosis (such as chronic pancreatitis, pancreatic neuroendocrine tumors, autoimmune pancreatitis) were established on the basis of cytopathology, liquid-based cytology, histopathology and immunohistochemistry of EUS-FNA/FNB. (2) Survival or imaging findings of the mass were unchanged during the 6-month follow-up period [10, 11]. In this study, senior pathologists made diagnoses for each image, and the diagnoses classification was determined according to the Papanicolaou Society of Cytopathology Guidelines [12].

Cytological image collection

The EUS examination was accomplished with various types of echoendoscope (Fujinon® EG-530UT echoendoscopes, FujiFilm Medical Systems, Tokyo, Japan; GF-UCT240-AL5, or GF-UCT260-AL5, Olympus Medical Systems, Tokyo, Japan) and the puncture was performed

with two types of 22G needles (EchoTip Ultra, Cook Medical, Bloomington, IN, USA; Boston Scientific Corporation, Marlborough, MA, USA) in this study.

The standardized image acquisition process was as follows: unitive images were captured using the Olympus microscope (model BX53F2) which was connected to the digital pathology reporting system (manufactured by Nanjing Jieda Ltd, model: Path QC). The unified magnification of the microscope (eyepiece 10 times and objective 40 times) was 400 times, and the brightness of the background light source was consistent. The width and height of the target image in the captured field of view were 2592*1944 pixels. The quality of the images was reviewed by a senior cytopathologist, and images were eliminated for the inferiority of quality on account of the lack of adequate diagnostic cell cluster, stained, unevenly dyed, cells overlapping and extrusion.

ROI annotation

Previous studies have confirmed that the sensitivity and specificity of EUS-FNA cytological diagnosis could be reinforced on condition that the “Suspicious for malignancy” and “Malignant” been distinguished as positive for pancreatic cancer within the framework of Papanicolaou Society of Cytopathology Guidelines [13, 14]. Consequently, “Suspicious for malignancy” and “Malignant” cell clusters were set as PDAC with reference to previous research on the cytology AI model.

The manual annotation of ROI in binary classification was accomplished together by two experienced cytopathologists. Images with disputed ROI boundaries and diagnostic arguments should be reviewed and consulted by senior cytopathologists until a consensus was reached. The image annotation software was Labelme. Red and blue lines were selected for PDAC and non-PDAC cell cluster demarcation line annotation, respectively. Continuous image fragments were captured closely to the periphery of the cell cluster, avoiding circling background images into ROI to the greatest extent during the labeling process.

Particularly worth mentioning was a semi-supervised CNN system was built in this study, except for 100% manual annotation was required in the testing set and validation set which only included a minority proportion of images, and only 20% of images were annotated manually in the training set which contained the vast majority of images, that is, the ratio of manually annotated images to non-manual annotated images was 1:4 approximately. It only labeled a small amount of images, which significantly reduced the manual annotation workload in the preparatory stage of model construction.

Data set construction

Cytology images were divided into training set, validation set and testing set (internal testing set and external testing set) according to the requirements of SSCNN system construction and verification. The images from the First Affiliated Hospital of Anhui Medical University were randomly divided into training set, validation set and internal testing set, while the images from the First Affiliated Hospital of the University of Science and Technology of China fell into the external testing set entirely. The training set images were used for the construction of the CNN model, while the validation set was applied to the validation of the SSCNN system. Finally, the testing set was utilized to verify the segmentation performance and classification ability of the SSCNN system. Cytopathologists-System competition set constituted by 60 images from Anqing Hospital Affiliated to Anhui Medical University was devoted to the comparison of accuracy between the SSCNN system and cytopathologists with different seniority. A total of 33 images of “atypical” diagnosed by expert cytopathologists from The First Affiliated Hospital of Bengbu Medical University were selected to analyze the consistency between the system and the definitive diagnosis conclusions based on the serological test, imaging findings, EUS-FNA, postoperative pathology or follow-up outcomes.

Construction of the semi-supervised CNN system

In this study, a semi-supervised cell cluster segmentation algorithm was constructed to release the burden of manual labeling and improve the generalization performance. Semi-supervised learning was mainly divided into two steps, namely pre-training and self-training. Firstly, the teacher model was pre-trained by manually annotated labels from the training set. Afterward, pseudo-labels were generated by the teacher model by predicting the unlabeled data in each training batch of the self-training process, and pseudo-labels were mixed with manual labels for training student models. In order to achieve consistent regularization of different perturbing inputs, a weak enhancement strategy was applied to unlabeled data fed into the teacher model that generates pseudo-labels, and a strong enhancement strategy was applied to the same unlabeled data fed into the student model. Meanwhile, to improve the quality of pseudo-labels, the exponential moving average (EMA) was carried out to update the parameters of the teacher model in elevation of performance and stability. The teacher model and student models have the same network structure, both use U-Net, and the encoder backbone network is ResNet50. U-Net consists of an encoder and a decoder. The encoder backbone progressively represents the input into high-level semantic features, and the decoder will gradually decode the semantic feature representation into the

predicted probability of the mask. Moreover, the feature layer corresponding to the encoder and the decoder adopts a jump connection to supplement the rich details. The weak enhancement strategies in the training process include random clipping, angle rotation and flipping. In contrast, the strong enhancement strategies include adding random Gaussian noise and random intensity scaling based on the weak enhancement. The model construction process is shown in Fig. 1.

Evaluation indicators

The performance indicators of the SSCNN system include two aspects: (1) segmentation performance, which refers to the efficiency of segmentation of PDAC cell cluster and Non-PDAC cell cluster in the target cytological images. Indicators include: Mean Intersection over Union (mIou), precision, recall, F1-score.

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$mIou = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ij}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}}$$

TP, FP and FN represent true positive, false positive and false negative respectively, k denote the labeling results of different classes, p_{ij} expresses the number of pixels of class i predicted to belong to class j , among $i, j \in (0, 1)$. Precision and Recall are the proportion of real PDAC in the samples predicted as PDAC and the proportion of correct predictions in all PDAC, respectively. F1-score (F1) is a balanced indicator, determined by precision and recall. mIou is an accuracy metric that measures the similarity between the actual situation and the predicted result.

(2) Classification performance: refers to the capacity of the model to qualitatively classify PDAC cell cluster and non-PDAC cell cluster in the target images. Indicators include: ROC curve, accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV).

The cytopathologists were divided into senior (exceed 15 years), intermediate (10–15 years) and junior (5–9 years) groups according to the years of working (experience), with two cytopathologists in each group. The images were prepared into web links and sent to 6 cytopathologists, with PDAC and non-PDAC options under

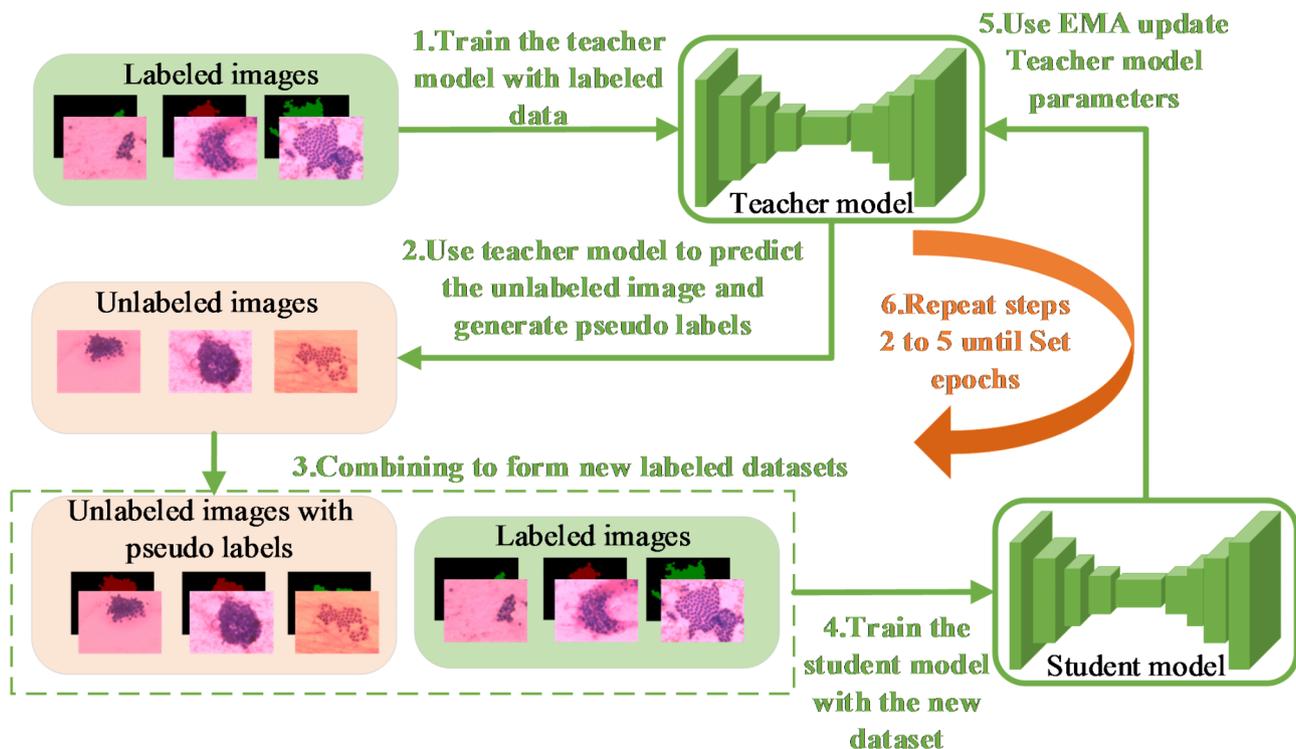


Fig. 1 The process of SSCNN system construction: The teacher model trained by manually labeled data in step 1 was the stage of model pre-training. The pseudo labels generated by the teacher model in step 2, the data re-integration in step 3, the student model generated by the new database containing manual and pseudo-labels in step 4, the teacher model update in step 5 and cycling continuous improvement in step 6 was the stage of model self-training.

each image. Each cytopathologist checked the images independently and made a judgment, then submitted them to the backend for calculating diagnosis accuracy and efficiency comparison with the SSCNN system.

“Atypical” diagnosed images by expert cytopathologists were selected and fed into the model, analyzing the consistency between the model and the definitive diagnosis conclusions. The flow chart of this study is shown in Fig. 2.

Statistics

McNemar test was adopted to compare efficacy indicators, Cohen’s Kappa test and correlation coefficient were used for consistency evaluation of diagnostic accuracy, the SPSS20.0 system was selected as statistical processing software, and the test statistic *P* value was 0.05.

Results

Patient population and basic information

A total of 183, 12 and 7 patients with pancreatic mass who underwent EUS-FNA from The First Affiliated Hospital of Anhui Medical University, The First Affiliated Hospital of the University of Science and Technology of China, Anqing Hospital Affiliated to Anhui Medical University were recruited. Qualified pathological smears were collected and 2989, 131 and 60 unified magnification images

of pancreatic cell clusters were photographed respectively. Besides, 33 qualified pathological images that were diagnosed as “atypical” from 8 patients admitted to the First Affiliated Hospital of Bengbu Medical University were collected additionally.

One hundred patients were diagnosed with PDAC and 110 patients were non-PDAC ultimately. A total of 2028 unified magnification images of PDAC cell clusters and 1152 non-PDAC cell clusters were photographed.

The basic information and pathological data of the patients were summarized in Table 1.

Data sets and annotation

A total of 1,088 images were manually annotated in the preliminary preparation stage, included 535 in the training set and the remaining 553 images in the testing and validation sets. One image may contain multiple cell clusters (ROI), ultimately, a total of 1529 ROIs were manually annotated, including 978 PDAC cell clusters and 551 non-PDAC cell clusters.

The training set consisted of 2567 images from the First Affiliated Hospital of Anhui Medical University, of which 535 images were manually annotated and the remaining 2032 images were unannotated. The Validation set, internal testing set and the external testing set consist of 134, 288 and 131 images respectively. The detailed

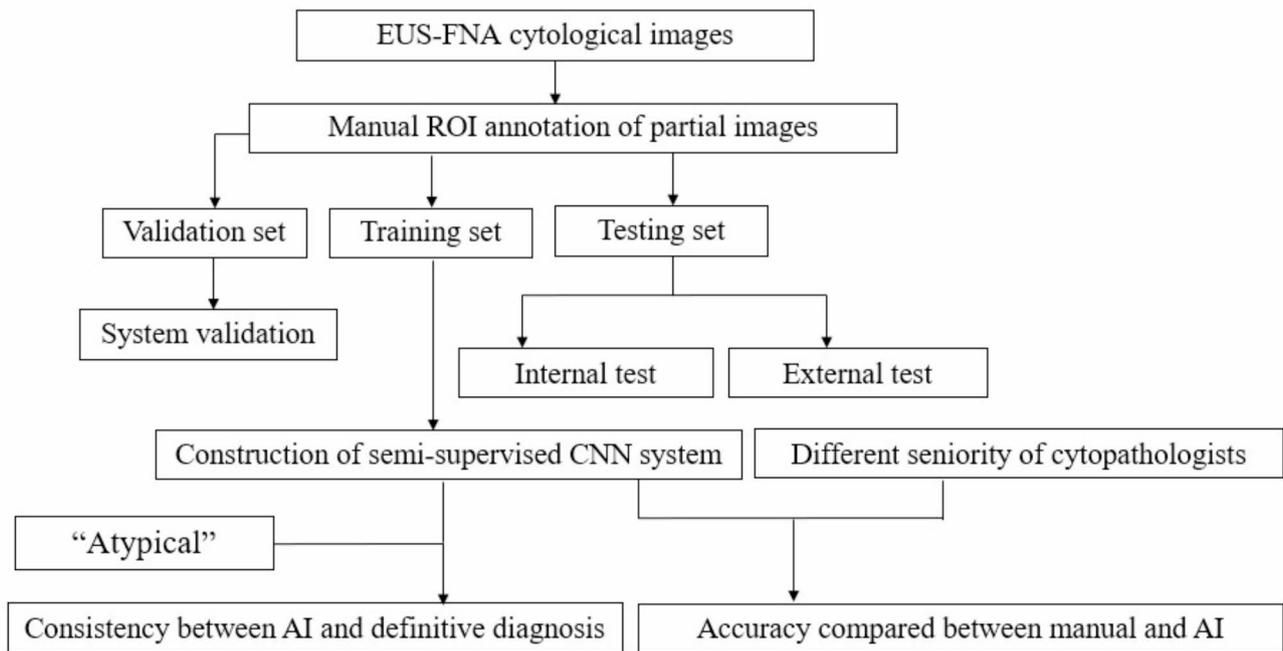


Fig. 2 Schematic flow chart

Table 1 Patient population and basic information

Clinical features		The First Affiliated Hospital of Anhui Medical University	The First Affiliated Hospital of the USTC	The Anqing Hospital Affiliated to Anhui Medical University	The First Affiliated Hospital of Bengbu Medical University
Gender (n)	male	111	8	4	5
	female	72	4	3	3
Age (years)		61.13±12.08	63.25±8.88	69.00±13.81	64.38±10.68
Puncture times ($\bar{x}\pm s$)		3.55±0.61	3.33±0.49	3.14±0.69	3.38±0.52
Smears($\bar{x}\pm s$)		6.90±2.31	6.25±1.13	8.57±1.51	7.25±1.58
Standardized images (n)		2989	131	60	33
Lesion location (n)	Uncinate process of pancreas	16	1	0	1
	Head and neck of pancreas	139	6	4	5
	Body and tail of pancreatic	28	5	3	2
Cytopathological diagnosis(n)	PDAC	83	7	4	Atypical
	Non-PDAC	100	5	3	Atypical
Definitive diagnosis(n)	PDAC	86	7	4	3
	Non-PDAC	97	5	3	5

distribution and annotation of PDAC and non-PDAC in the data sets are shown in Table 2.

Segmentation and classification performance of the SSCNN system

The segmentation efficiency indexes mIoU, precision, recall and F1-score of the SSCNN system in the internal testing set were 88.30%, 93.21%, 94.24% and 93.68% respectively. Analogously, the mIoU, precision, recall and

F1-score of the SSCNN system in the external testing set were successively 87.75%, 93.81%, 93.14% and 93.48%. The PDAC classification efficiency indexes of the SSCNN system including Area under the ROC curve (AUC), accuracy, sensitivity, specificity, PPV, NPV in internal testing set were 0.97, 95.14%, 93.82%, 97.27%, 98.23%, 90.68% respectively, and 0.99, 93.89%, 93.58%, 95.45%, 99.03% and 75.00% correspondingly in external testing set. In addition, the CNN system was trained using

Table 2 Data sets and annotations(n)

Data sets	Classification	Manual annotated	Unannotated
Training set	PDAC	351	1268
	Non-PDAC	184	764
Validation set	PDAC	93	/
	Non-PDAC	41	/
Internal testing set	PDAC	178	/
	Non-PDAC	110	/
External testing set	PDAC	109	/
	Non-PDAC	22	/
Cytopathologists-System competition set	PDAC	/	29
	Non-PDAC	/	31

Table 3 The segmentation performance of the semi-supervised CNN system (%)

Model	Data sets	mIou	precision	recall	F1-score
SSCNN	Validation set	86.20	91.82	93.22	92.44
	Internal testing set	88.30	93.21	94.24	93.68
	External testing set	87.75	93.81	93.14	93.48
CNN	Validation set	78.69	87.00	86.69	88.07
	Internal testing set	79.66	87.22	90.18	88.68
	External testing set	80.40	87.40	90.95	89.14

just 535 labeled images in a supervised learning setting to emphasize the advantages of SSCNN trained with a larger set of unlabeled images. The details are presented in Tables 3 and 4; Fig. 3.

Competition between SSCNN system and cytopathologists

The diagnostic accuracy of senior, intermediate and junior cytopathologists was 95.00%, 88.33% and 76.67%, while the diagnostic accuracy of the SSCNN system was 90.00%. The results were statistically different between the four groups ($P=0.019$). The accuracy of the SSCNN system was highly consistent with senior cytopathologists (Kappa = 0.853). The details are presented in Table 5; Fig. 3.

The performance of the system in the interpretation of “atypical”

The accuracy, sensitivity and specificity of the system in the classification of “atypical” cases were 78.79%, 84.20% and 71.43% respectively compared with the definitive

diagnosis conclusion. Additionally, representative images for model visualization are shown in Fig. 4.

Discussion

In this study, we constructed a semi-supervised CNN system for the diagnosis of PDAC based on EUS-FNA cytological images via partial annotation in advance, and the system showed outstanding performance in the segmentation and classification of PDAC cell clusters. Subsequently, we compared the differences in diagnostic performance between the system and the cytopathologist, the results indicate highly consistent between the SSCNN system and senior cytopathologists. what’s more, the system could effectively classify the images which artificially diagnosed as “atypical” with relatively precision into binary diagnosis of “PDAC” and “non-PDAC”.

The reports of artificial intelligence in the field of endoscopic ultrasonography have sprung up exuberantly in the last few years [15–18]. In the previously published researches, Sohn [4] built a DL system named “MIPCL” on the basis of a Whole slide image (WSI) with ultra-high resolution. The F1-Score and AUC of the MIPCL model in the identification of pancreatic cancer cells were 87.87% and 0.92 respectively. However, the system based on WSI always occupies an immense burden on runtime storage, which brings a great challenge to the performance of computers and quick acquisition of WSI was difficult to achieve in clinical practice. Another EUS-FNA cytological CNN system reported by Zhang Song [6] showed remarkable lesion segmentation and pathological classification performance, nevertheless, their study adopted a fully supervised learning model, and the construction of a fully supervised model depended on detailed manual annotation in the preliminary preparation stage. The target images involved in model training, verification and testing need to be completely annotated, which generates a laborious and abundant workload. The performance of AI models was ultimately depressed under the influence of insufficient image samples in partial studies. Lin Rong [5] constructed an AI model for EUS-FNA cytological image classification, and the accuracy of internal verification and external verification was 83.4% and 88.7% respectively. However, none of the above studies further explored the identification of images manually diagnosed as “atypical”, which was a

Table 4 The classification performance of the semi-supervised CNN system (95% confidence interval)

Data sets	AUC	accuracy	sensitivity	specificity	PPV	NPV
Validation set	0.98	94.78% (89.61%, 97.45%)	94.62% (88.03%, 97.68%)	95.12% (83.86%, 98.65%)	97.78% (92.26%, 99.39%)	88.64% (76.02%, 95.05%)
Internal testing set	0.97	95.14% (92.01%, 97.08%)	93.82% (89.27%, 96.51%)	97.27% (92.29%, 99.07%)	98.23% (94.94%, 99.39%)	90.68% (84.08%, 94.71%)
External testing set	0.99	93.89% (88.41%, 96.87%)	93.58% (87.33%, 96.85%)	95.45% (78.20%, 99.19%)	99.03% (94.70%, 99.83%)	75.00% (56.64%, 87.32%)

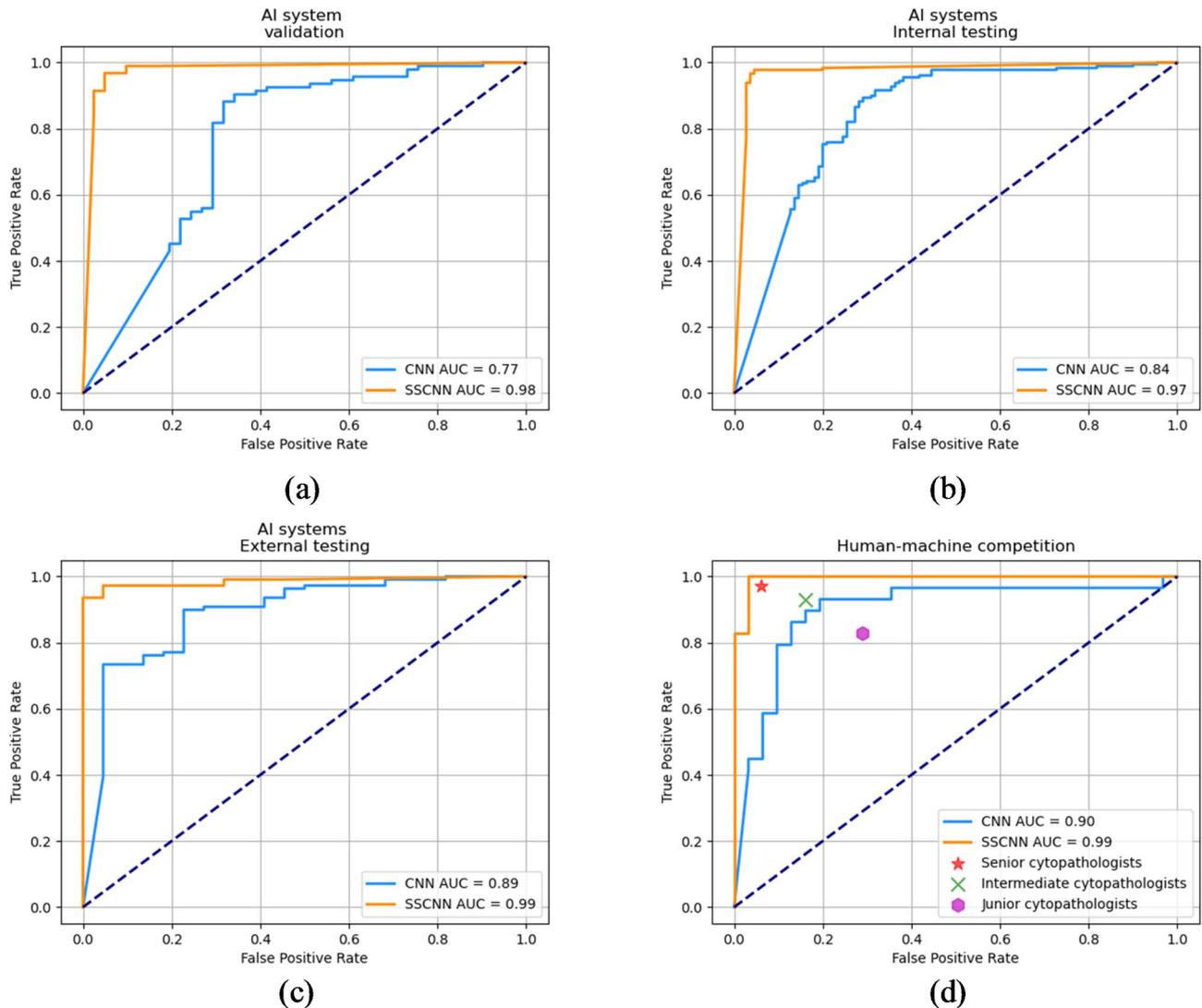


Fig. 3 The ROC of the SSCNN system in testing, validation and system-cytopathologists competition. CNN: convolutional neural networks trained only on labeled images, SSCNN: semi-supervised convolutional neural network. (a) ROC of the validation set. (b) ROC of the internal testing set. (c) ROC of external testing set. (d) ROC of competition between AI system and cytopathologists

Table 5 The performance of the SSCNN system and cytopathologists (95% confidence interval)

Participant	accuracy	sensitivity	specificity	PPV	NPV
SSCNN model	90.00% (79.85%, 95.34%)	82.76% (65.45%, 92.40%)	96.77% (83.81%, 99.43%)	96.00% (80.46%, 99.29%)	85.71% (70.62%, 93.74%)
CNN model	70.00% (57.49%, 80.10%)	96.55% (82.82%,99.39%)	45.16% (29.16%,62.23%)	62.22% (47.63%, 74.89) %	93.33% (70.18%, 98.81%)
Senior cytopathologists	95.00% (86.30%, 98.29%)	96.55% (82.82%,99.39%)	93.55% (79.28%,98.21%)	93.33% (78.68%, 98.15) %	96.67% (83.33%, 99.41%)
Intermediate cytopathologists	88.33% (77.82%, 94.23%)	93.10% (78.04%,98.09%)	83.87% (67.37%,92.91%)	84.36% (68.25%, 91.14) %	92.86% (77.35%, 98.02%)
Junior cytopathologists	76.67% (64.56%, 85.56%)	82.76% (65.45%,92.40%)	70.97% (53.41%,83.90%)	72.73% (55.78%, 84.93) %	81.48% (63.30%, 91.82%)

confusing diagnosis to clinicians: overdiagnosis of malignant tumors may bring unnecessary surgical injury to patients, on the contrary, missed diagnosis may delay the opportunity for timely treatment.

In this study, we applied a semi-supervised learning model based on EUS-FNA cytological images. Despite semi-supervised AI studies on the pathological diagnosis of prostate cancer [19], skin tumor [20], and colon cancer

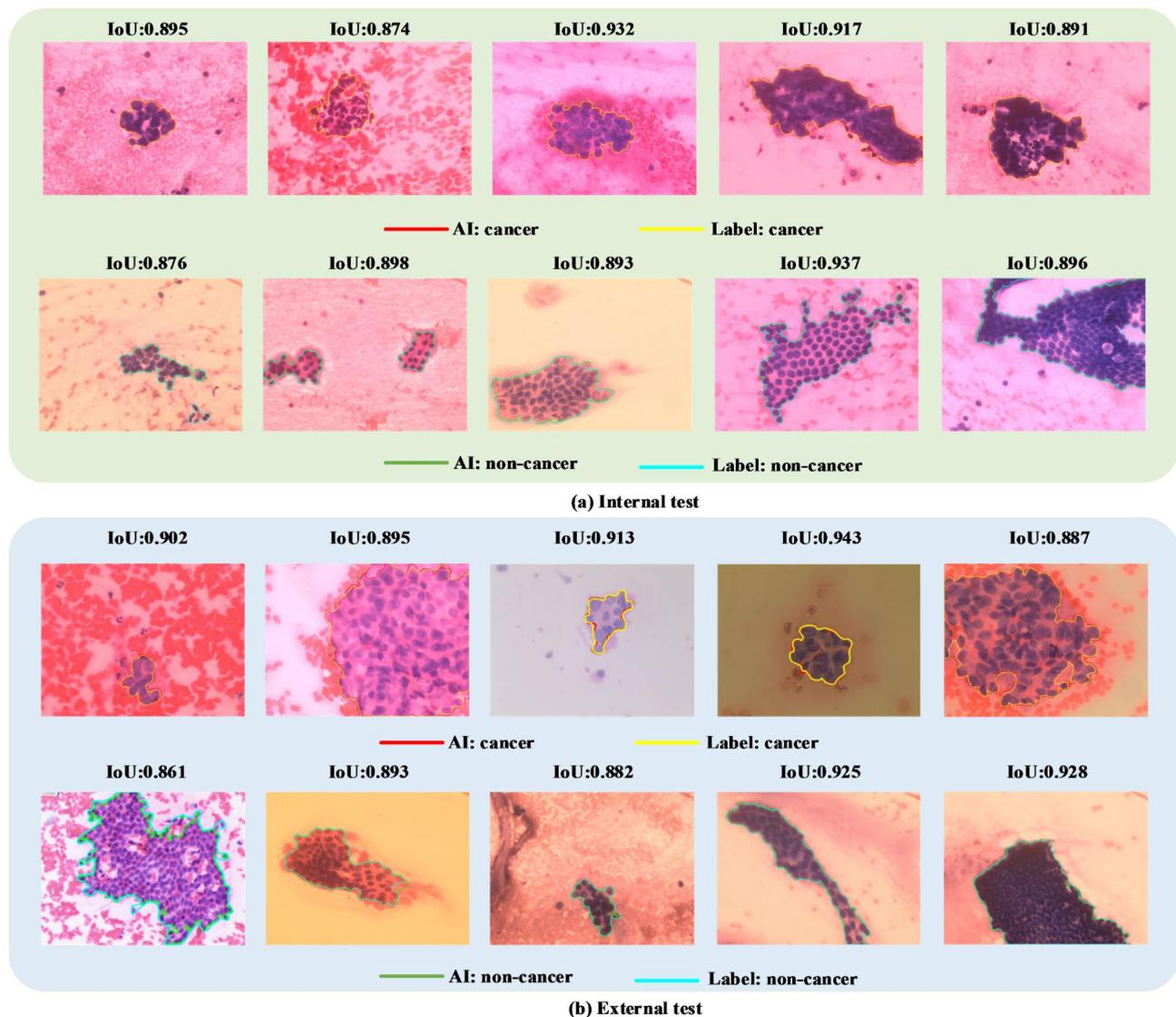


Fig. 4 Model visualization: **(a)** Representative images of manually annotated and AI system-identified PDAC and non-PDAC cell clusters in the internal test set. **(b)** Representative images of manually annotated and AI system-identified cell clusters in the external test set

[21] have already been reported, this research probably was the first report of semi-supervised AI system on EUS-FNA smears to the best of our knowledge until now. The core principle of the semi-supervised model was self-learning. The method of self-training the CNN model with pseudo-label data can achieve regularization, reduce the overfitting generated by the model training with limited labeled data, and effectively improve the generalization and prediction accuracy of the system [22]. The dominant advantage of the semi-supervised system was that only 1/3~1/5 or even fewer data need to be manually annotated in model training, thus 3~5 times or nearly 10 times of training data could be assimilated under the parallelly workload, hence, model diagnosis efficiency will be promoted with more plentiful samples as a prerequisite [23]. Previous studies have reported that

the performance of semi-supervised CNN was better than fully supervised model VGG and AlexNET, which with the most advanced algorithm reported at present [24, 25].

In this study, the accuracy, sensitivity and specificity of the SSCNN system in identifying PDAC reached 95.14%, 93.82%, 97.27% and 93.89%, 93.58%, and 95.45% in internal and external testing sets respectively. The model accomplished PDAC image information extraction on the basis of adequate sample content in the training set and showed excellent generalization ability. Besides, the remarkable resemblance of accuracy, sensitivity and specificity in internal and external testing sets indicated that the system was robust. Further analysis on the cases which was misjudged by model showed that the proportion of PDAC misjudged as non-PDAC was higher than

that of non-PDAC cases identified as PDAC, suggested that the risk of PDAC missed diagnosis was higher than the risk of upgrading pathological findings in the system. Retrospective analysis of the features of pathological degradation showed relatively low quality of these smears, some cell clusters were deformed by extrusion and lost the polarity distribution rule, which interfered with the interpretation result. It reminds us that the significance of quality improvement in smear preparation and manual review was required when the model diagnoses non-PDAC. Similarly, the model also showed excellent segmentation performance for PDAC, accurately segmenting the location of PDAC in images helps junior pathologists quickly find suspicious lesions in multitudinous cell clusters and reduces the difficulty of image reviewing. Moreover, it could be beneficial to improve lesion localization ability in endoscopists who are trained with cytopathological knowledge and create favorable conditions for ROSE implementation.

The comparative results of this study showed that there exist differences in diagnostic accuracy between the SSCNN system, senior cytopathologist, intermediate cytopathologist and junior cytopathologist. In order of accuracy from high to low, senior cytopathologist > SSCNN system > Intermediate > Junior. Kappa correlation analysis showed that it was highly consistent between the system and senior cytopathologists. Finally, we specially selected the images artificially diagnosed as “atypical” and imported them into the system for classification. The accuracy, sensitivity and specificity were 78.79%, 84.20% and 71.43% respectively compared with the definitive diagnosis, which had achieved satisfactory results and suggested that a further accurate classification by the SSCNN system should be carried out when the manual diagnosis was atypical.

There are still some shortcomings in this study. First of all, this system was set as a binary variable with PDAC and non-PDAC, whereas non-PDAC includes a variety of subtypes such as autoimmune pancreatitis, chronic pancreatitis and pancreatic neuroendocrine tumors, thus, subtypes identification was unavailable by this system. Secondly, the system has favourable performance in retrospective study, but it was still necessary to design prospective studies and include more eligible patients for further verification. Endoscopist with qualified cytological smear preparation skills can obtain the permission of SSCNN model on the condition of the system installed and debugged by computer engineers, real-time operation of the SSCNN model in an endoscopic ultrasound operating center will help achieve the clinical practice of SSCNN model assisted ROSE in the absence of cytopathologist. Therefore, large-scale, prospective, multicenter and randomized studies need to be employed to testify

the effect of ROSE accomplished by endoscopists allied with this semi-supervised model in the future.

Conclusion

This study introduced a novel AI-based method for the diagnosis of pancreatic ductal adenocarcinoma on EUS-FNA cytological images, the semi-supervised convolutional neural network demonstrated satisfactory performance in PDAC cell cluster segmentation and classification, and helpful to solve the controversial issues in manual diagnosis.

Abbreviations

AI	Artificial Intelligence
AUC	Area Under the ROC Curve
CNN	Convolutional Neural Network
DL	Deep Learning
EMA	Exponential Moving Average
EUS-FNA	Endoscopic ultrasound-guided fine needle aspiration
NPV	Negative Predictive Value
PDAC	Pancreatic Ductal Adenocarcinoma
PPV	Positive Predictive value
ROI	Region of Interests
ROSE	Rapid On-Site Evaluation
SSCNN	Semi-Supervised Convolutional Neural Network
WSI	Whole Slide Image

Acknowledgements

Not applicable.

Author contributions

DF performed the statistical analysis and wrote the paper. YH constructed the CNN system and interpreted the results. SL, CS, JB, LY, CZ and HZ collected the data. DD and LX made the ROI annotation and pathological diagnosis. YZ, ZS and QM conceived and designed this research. HW designed the Semi-supervised AI model. All authors reviewed the manuscript.

Funding

No funding was received.

Data availability

Sequence data that support the findings of this study have been deposited in the Google Cloud Drive and is available at the following URL: <https://figshare.com/s/1499f9729521ca44a743>.

Declarations

Ethics approval and consent to participate

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the ethics committee Review Board of The First Affiliated Hospital of Anhui Medical University. The need for informed consent was waived by the ethics committee Review Board because of the retrospective nature of the study.

Consent for publication

All authors expressed great pleasure to publish in this journal.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Gastroenterology, The First Affiliated Hospital of Anhui Medical University, Hefei 230022, Anhui Province, China

²Department of Gastroenterology, The Second People's Hospital of Hefei, Hefei 230011, Anhui Province, China

³Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, Anhui Province, China

⁴University of Science and Technology of China, Hefei 230026, Anhui Province, China

⁵Department of Cytopathology, The First Affiliated Hospital of Anhui Medical University, Hefei 230022, Anhui Province, China

⁶Department of Pathology, Anqing Hospital Affiliated to Anhui Medical University, Anqing 246004, Anhui Province, China

⁷Department of Gastroenterology, Anqing Hospital Affiliated to Anhui Medical University, Anqing 246004, Anhui Province, China

⁸Department of Biliary and Pancreatic Surgery, Division of Life Sciences and Medicine, The First Affiliated Hospital of USTC, University of Science and Technology of China, Hefei 230001, Anhui, China

⁹Department of Gastroenterology, The First Affiliated Hospital of Bengbu Medical University, Bengbu 233004, Anhui Province, China

Received: 17 November 2024 / Accepted: 11 March 2025

Published online: 18 March 2025

References

1. Yang X, Liu ZM, Zhou X, Yang F, Ma WZ, Sun XZ, Sun SY, Ge N. Methods to increase the diagnostic efficiency of endoscopic ultrasound-guided fine-needle aspiration for solid pancreatic lesions: an updated review. *World J Gastrointest Endosc.* 2024;16(3):117–25.
2. Matsubayashi H, Sasaki K, Ono S, Abe M, Ishiwatari H, Fukutomi A, Uesaka K, Ono H. Pathological and molecular aspects to improve endoscopic ultrasonography-guided fine-needle aspiration from solid pancreatic lesions. *Pancreas.* 2018;47(2):163–72.
3. Milluzzo SM, Olivari N, Rossi G, Bianchi D, Liserre B, Graffeo M, Lovera M, Correale L, Hassan C, Spada C. Rapid on-site evaluation improves the sensitivity of endoscopic ultrasound-guided fine needle aspiration (EUS-FNA) for solid pancreatic lesions irrespective of technique: a single-centre experience. *Cytopathology.* 2023;34(4):318–24.
4. Sohn A, Miller D, Ribeiro E, Shankar N, Ali S, Hruban R, Baras A. A deep learning model to triage and predict adenocarcinoma on pancreas cytology whole slide imaging. *Sci Rep.* 2023;13(1):16517.
5. Lin R, Sheng LP, Han CQ, Guo XW, Wei RG, Ling X, Ding Z. Application of artificial intelligence to digital-rapid on-site cytopathology evaluation during endoscopic ultrasound-guided fine needle aspiration: a proof-of-concept study. *J Gastroenterol Hepatol.* 2023;38(6):883–87.
6. Zhang S, Zhou Y, Tang D, Ni M, Zheng J, Xu G, Peng C, Shen S, Zhan Q, Wang X, et al. A deep learning-based segmentation system for rapid onsite cytologic pathology evaluation of pancreatic masses: a retrospective, multicenter, diagnostic study. *EBioMedicine.* 2022;80:104022.
7. Pitman MB, Layfield LJ. Guidelines for pancreaticobiliary cytology from the Papanicolaou society of cytopathology: a review. *Cancer Cytopathol.* 2014;122(6):399–411.
8. Dumonceau JM, Deprez PH, Jenssen C, Iglesias-Garcia J, Larghi A, Vanbiervliet G, Aithal GP, Arcidiacono PG, Bastos P, Carrara S, et al. Indications, results, and clinical impact of endoscopic ultrasound (EUS)-guided sampling in gastroenterology: European Society of Gastrointestinal Endoscopy (ESGE) clinical guideline - Updated January 2017. *Endoscopy.* 2017;49(7):695–714.
9. Pitman MB, Centeno BA, Ali SZ, Geneva M, Stelow E, Mino-Kenudson M, Fernandez-del Castillo C, Max Schmidt C, Brugge W, Layfield L, et al. Standardized terminology and nomenclature for pancreaticobiliary cytology: the Papanicolaou society of cytopathology guidelines. *Diagn Cytopathol.* 2014;42(4):338–50.
10. Marti-Bonmati L, Cerdá-Alberich L, Pérez-Girbés A, Díaz Beveridge R, Montalvá Orón E, Pérez Rojas J, Alberich-Bayarri A. Pancreatic cancer, radiomics and artificial intelligence. *Br J Radiol.* 2022;95(1137):20220072.
11. Qin X, Zhang M, Zhou C, Ran T, Pan Y, Deng Y, Xie X, Zhang Y, Gong T, Zhang B, et al. A deep learning model using hyperspectral image for EUS-FNA cytology diagnosis in pancreatic ductal adenocarcinoma. *Cancer Med.* 2023;12(16):17005–17.
12. Adler D, Max Schmidt C, Al-Haddad M, Barthel JS, Ljung BM, Merchant NB, Romagnuolo J, Shaaban AM, Simeone D, Bishop Pitman M, et al. Clinical evaluation, imaging studies, indications for cytologic study, and preprocedural requirements for duct brushing studies and pancreatic FNA: the Papanicolaou society of cytopathology recommendations for pancreatic and biliary cytology. *Diagn Cytopathol.* 2014;42(4):325–32.
13. Gonzalez-Mancera MS, Ahmadian SS, Gomez-Fernandez C, Velez-Torres J, Jorda M, García-Buitrago MT. Risk of malignancy associated with the diagnostic categories proposed by the Papanicolaou society of cytopathology for pancreaticobiliary specimens: an institutional experience. *Diagn Cytopathol.* 2022;50(2):49–56.
14. Layfield LJ, Zhang T, Esebua M. Diagnostic sensitivity and risk of malignancy for bile duct brushings categorized by the Papanicolaou society of cytopathology system for reporting pancreaticobiliary cytopathology. *Diagn Cytopathol.* 2022;50(1):24–7.
15. Dhali A, Kipkorir V, Srichawla BS, Kumar H, Rathna RB, Ongidi I, Chaudhry T, Morara G, Nurani K, Cheruto D, et al. Artificial intelligence assisted endoscopic ultrasound for detection of pancreatic space-occupying lesion: a systematic review and meta-analysis. *Int J Surg.* 2023;109(12):4298–308.
16. Kuwahara T, Hara K, Mizuno N, Haba S, Okuno N, Kuraishi Y, Fumihara D, Yanaidani T, Ishikawa S, Yasuda T, et al. Artificial intelligence using deep learning analysis of endoscopic ultrasonography images for the differential diagnosis of pancreatic masses. *Endoscopy.* 2023;55(2):140–9.
17. Yao L, Zhang C, Xu B, Yi S, Li J, Ding X, Yu H. A deep learning-based system for mediastinum station localization in linear EUS (with video). *Endosc Ultrasound.* 2023;12(5):417–23.
18. Yin M, Liu L, Gao J, Lin J, Qu S, Xu W, Liu X, Xu C, Zhu J. Deep learning for pancreatic diseases based on endoscopic ultrasound: a systematic review. *Int J Med Inf.* 2023;174:105044.
19. Marini N, Otálora S, Müller H, Atzori M. Semi-supervised training of deep convolutional neural networks with heterogeneous data and few local annotations: an experiment on prostate histopathology image classification. *Med Image Anal.* 2021;73:102165.
20. Requa J, Godard T, Mandal R, Balzer B, Whittmore D, George E, Barcelona F, Lambert C, Lee J, Lambert A, et al. High-fidelity detection, subtyping, and localization of five skin neoplasms using supervised and semi-supervised learning. *J Pathol Inf.* 2022;14:100159.
21. Neto PC, Oliveira SP, Montezuma D, Fraga J, Monteiro A, Ribeiro L, Gonçalves S, Pinto IM, Cardoso JS. iMIL4PATH: a semi-supervised interpretable approach for colorectal whole-slide images. *Cancers (Basel).* 2022;14(10):2489.
22. Chen X, Wang X, Zhang K, Fung KM, Thai TC, Moore K, Mannel RS, Liu H, Zheng B, Qiu Y. Recent advances and clinical applications of deep learning in medical image analysis. *Med Image Anal.* 2022;79:102444.
23. Tang Y, Wang S, Qu Y, Cui Z, Zhang W. Consistency and adversarial semi-supervised learning for medical image segmentation. *Comput Biol Med.* 2023;161:107018.
24. Shawi RE, Kilanava K, Sakr S. An interpretable semi-supervised framework for patch-based classification of breast cancer. *Sci Rep.* 2022;12(1):16734.
25. Li H, Zhong J, Lin L, Chen Y, Shi P. Semi-supervised nuclei segmentation based on multi-edge features fusion attention network. *PLoS ONE.* 2023;18(5):e0286161.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.