

Review

AI in Pancreatic Image Analysis: A Review

Weixuan Liu ¹, Bairui Zhang ¹, Tao Liu ², Juntao Jiang ^{3,*} and Yong Liu ^{3,*}

¹ Sydney Smart Technology College, Northeastern University, Shenyang 110819, China

² School of Mathematics and Statistics, Northeastern University at Qinhuangdao, Qinhuangdao 066004, China

³ College of Control Science and Engineering, Zhejiang University, Hangzhou 310058, China

* Correspondence: juntaojiang@zju.edu.cn; yongliu@iipc.zju.edu.cn

Simple Summary: Pancreatic cancer has a high fatality rate and poor prognosis. Early diagnosis and accurate treatment are essential but challenging due to subtle symptoms, high misdiagnosis rates, and significant financial burdens. Artificial Intelligence (AI) can alleviate the workloads of doctors, improve decision-making, and reduce the financial burden of patients. This study mainly reviews research works on AI tasks in medical image analysis corresponding to the pancreas, known as segmentation, classification, object detection, prognosis prediction, and so on, on five modalities: CT, MRI, EUS, PET, pathological images, and integrated modalities. This study also discusses heated topics and future directions of AI-enabled automated pancreatic cancer diagnosis.

Abstract: Pancreatic cancer is a highly lethal disease with a poor prognosis. Its early diagnosis and accurate treatment mainly rely on medical imaging, so accurate medical image analysis is especially vital for pancreatic cancer patients. However, medical image analysis of pancreatic cancer is facing challenges due to ambiguous symptoms, high misdiagnosis rates, and significant financial costs. Artificial Intelligence (AI) offers a promising solution by relieving medical personnel's workload, improving clinical decision-making, and reducing patient costs. This study focuses on AI applications such as segmentation, classification, object detection, and prognosis prediction across five types of medical imaging: CT, MRI, EUS, PET, and pathological images, as well as integrating these imaging modalities to boost diagnostic accuracy and treatment efficiency. In addition, this study discusses current hot topics and future directions aimed at overcoming the challenges in AI-enabled automated pancreatic cancer diagnosis algorithms.

Keywords: pancreatic cancer; artificial intelligence; medical images; diagnosis; treatment

1. Introduction

Pancreatic cancer (PC) is a lethal tumor of the digestive system with a very poor prognosis, which is always found at an advanced stage and prone to recurrence after resection. Its mortality and morbidity are highly paralleled and pose a great threat to human health [1]. Pancreatic ductal adenocarcinoma (PDAC) is the most common type of PC, 80%–85% of which present with advanced local or distant metastatic disease, while only 15%–20% is suitable for surgical removal. In addition, the 5-year relative survival rate of PC is 12%, the lowest among all types of cancers [2]. These data suggest that early screening and diagnosis are important in improving the survival outcomes of PC patients.

Current advanced medical imaging tools are capable of visualizing tumor tissue and supporting diagnosis in a noninvasive manner, primarily magnetic resonance imaging (MRI), computed tomography (CT), endoscopic ultrasound (EUS), and positron emission tomography (PET) [3,4]. Improvements based on these imaging tools include EUS-guided fine needle aspiration (FNA) and biopsy (FNB), contrast-enhanced EUS (CE-EUS), contrast-enhanced computed tomography (CE-CT), contrast-enhanced magnetic resonance imaging (CE-MRI), and positron emission tomography (PET/CT). However, clinical diagnosis of PC is still highly dependent on biopsy after the imaging step, which is complex and

Citation: Liu, W.; Zhang, B.; Tao, L.; Jiang, J.; Liu, Y. AI in Pancreatic Image Analysis: A Review. *Sensors* **2023**, *1*, 0. <https://doi.org/>

Received:

Revised:

Accepted:

Published:

Copyright: © 2024 by the authors. Submitted to *Sensors* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

time-consuming, and may even result in patients missing critical treatment opportunities. Developing artificial intelligence (AI) processing methods in biomedical imaging can diagnose PCs quickly and accurately, with lower misdiagnosis rates and costs. This work facilitates large-scale screening and intervention for PC patients.

In recent years, AI has been applied to PC diagnosis in the experimental stage with reasonable results, marking the beginning of a shift from biopsy-dependent traditional diagnosis. AI is a computer technology that can simulate specific human behaviors, such as learning, reasoning, problem-solving, and decision-making. Deep learning is a milestone of AI and utilizes propagation algorithms, having made significant breakthroughs in automated image analysis with high Accuracy, Specificity, and Recall in diagnosing PC and differentiating it from chronic pancreatitis [5,6]. Studies have shown that in clinical trials, AI can reduce the burden of routine tasks in the medical workflow, allowing doctors to spend more time tackling other challenges [7]. Therefore, the automated analysis of PC images by AI is an efficient and convenient aid to doctors.

This paper aims to discuss the current status and future direction of AI applications in PC image processing. By summarizing the AI methods for detecting PCs, it provides an effective reference for PC early screening and diagnostic solutions and promotes potential transformations in the field of medical diagnosis.

2. Pancreatic Cancer

The pancreas consists of the head, neck, body, and tail of the pancreas, located in the abdominal cavity. PC is usually referred to as a tumor that arises within the epithelial cells of the pancreas [8]. Smoking, obesity, diabetes mellitus, alcohol, pancreatitis, allergies, the microbiome, the environment, occupation, family history of cancer, and chronic pancreatitis (CP) are risk factors of PCs [9,10]. Autoimmune pancreatitis (AIP) is a rare form of CP [10]. Pancreatic intraepithelial neoplasms (PanIN), intraductal papillary mucinous neoplasm (IPMN), and mucinous cystic neoplasm (MCN) are precursor lesions of PCs [11].

PC presents in two common types: pancreatic ductal adenocarcinoma (PDAC) and pancreatic neuroendocrine tumor (pNET) [12]. PDAC accounts for more than 90% of PCs and is the most prevalent type, while pNET is relatively rare, accounting for less than 5% [13]. Among other rare types, solid-pseudopapillary neoplasm (SPN) accounts for 0.2-2.7% of PCs, which typically affects young females [14]. Additional infrequent types of PCs include serous cystic neoplasm (SCN), pancreatic adenosquamous carcinoma (PASC), acinar cell carcinoma (ACC), solid pseudopapillary tumor (SPT), and so on [15–18]. Due to their rarity, these tumors lack large-scale clinical studies, and many issues remain unclear. Using AI to differentiate between these tumors can help doctors learn their characteristics. Figure 1 shows the relationship between these lesions.

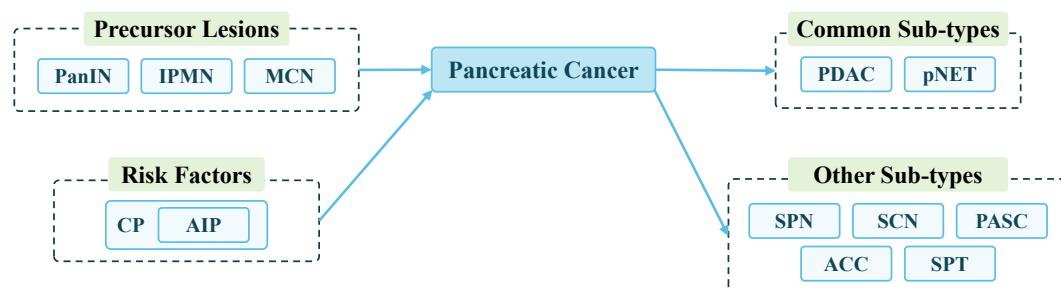


Figure 1. Precursors, risk factors, and sub-types of PC

2.1. Pancreatic Ductal Adenocarcinoma

Pancreatic Ductal Adenocarcinoma (PDAC) is a malignancy affecting the exocrine pancreas and involving acinar and duct cells with a contentious origin. Although traditionally believed to originate from duct cells, studies in rodents have suggested an alternative origin from acinar cells [19]. The development of PDAC typically commences with pancreatic intraepithelial neoplasias, marked by the accrual of genetic mutations. Part of PDAC cases

are attributed to these precursor lesions like IPMN [20]. PDAC most frequently presents in the head of the pancreas, followed by the tail, with occasional occurrences in the pancreatic body [21]. As PDAC advances, its potential for extensive spread becomes pronounced once it reaches a critical size at its primary site, underscoring the aggressive nature of this malignancy [22]. According to the gene expression profiles of malignant epithelial cells, PDAC can be divided into subtypes [23]. In the two-group classification, the main subtypes of PDAC include the classical subtype and the basal-like subtype. The basal-like subtype has been associated with a poorer prognosis and a more aggressive phenotype [24–26].

2.2. Pancreatic Neuroendocrine Tumors

Pancreatic neuroendocrine tumor (pNET) is a rare and diverse neoplasm. According to population studies, the incidence of pNET is less than 1 in 100,000 [27]. However, with increased use of computed tomographic scans, the incidence has doubled in the last few decades [28]. pNET arises from the pancreatic neuroendocrine cells. They are found in various organs and are vital for receiving signals from the nervous system and regulating numerous bodily functions. pNET can be divided into functional pNET (F-pNET), which secrete hormones leading to specific clinical syndromes, and non-functional pNET (NF-pNET), which have no symptoms [29–31]. F-pNET are relatively rare and account for about 20%. The most common F-pNET are insulinomas, which lead to hypoglycemia, and gastrinomas, which lead to excessive gastrin overproduction. Other less common types include glucagonomas, VIPomas, and somatostatinomas [32,33]. F-pNET exhibits varying degrees of malignant potential across subtypes. Specifically, the insulinomas subtype is the most benign one with a malignant potential of 5 to 15%. While other subtypes have a much higher potential ranging from 60 to 90% [34]. NF-pNET can be divided into three categories: those that do not produce hormones; those that have hormones at levels low enough to make symptoms; and those that produce hormones like pancreatic polypeptide, chromogranin A, ghrelin, calcitonin, or neurotensin that do not cause symptoms [35]. Compared with F-pNET, NF-pNET is typically discovered later and is more prone to malignancy, often leading to poorer prognoses [36]. These types of tumors generally remain asymptomatic until they reach a substantial size, at which point symptoms emerge due to the mass effects of the original tumor or its metastasis [35]. Furthermore, pNET tends to be multifocal and can metastasize to other organs, with the liver being the primary site of metastasis, significantly impacting the overall prognosis [37,38].

2.3. Clinical Challenges and AI Application

Accurately diagnosing PC poses significant challenges. While screening for early cancer precursors and the subsequent surgical removal of diseased lesions can reduce morbidity and mortality [39], the relatively low incidence of PC among diseases makes screening asymptomatic individuals unfeasible [40]. Moreover, symptoms in patients with early-stage PC are typically mild and can be mistaken for common benign diseases [41]. As a result, most of PCs are diagnosed after metastasis has occurred, with only a small number being identified at the local stage. Unfortunately, poor survival rates have not significantly improved in recent decades [42]. Compounding this issue, the lack of knowledge about PC and diagnostic pathways often results in patients reluctant to seek medical attention, causing treatment delays [43]. Additionally, existing diagnostic methods exhibit a high false positive rate and lack effectiveness [40]. The heterogeneous behavior of PC further complicates matters, as it can be challenging to determine malignant potential accurately, and overdiagnosis can potentially do more harm than good in mortality [39]. AI models are efficient and fast to detect PCs, therefore can reduce the costs of screening and the incidence of misdiagnosis.

Treating PC also presents significant challenges. Currently, the primary treatments for PC involve surgery and chemotherapy. Nevertheless, a mere 15 to 20% of patients are eligible for surgery, and post-surgical relapse is highly probable. Moreover, PC tissue exhibits low blood vessel density and a fibrotic barrier, impeding the penetration of chemotherapy

drugs and leading to drug resistance [44,45]. Although targeted therapies are available for certain PC subtypes associated with specific genetic mutations, their effectiveness is constrained by high costs, drug resistance, and the unique tissue characteristics of the pancreas [46]. AI could quickly diagnose PCs by analyzing medical images and predict the survival of patients after undergoing surgery, which could help doctors make treatment decisions.

Challenges in the diagnosis and treatment of PC emphasize the urgent need for innovative assisted diagnostic technologies. AI has the potential to reduce the burden on physicians and patients as it can automatically and accurately analyze medical images of PCs with low time and economic costs.

3. Public Data Sources

Public data sources of pancreatic medical images are essential for medical researchers, data scientists, and healthcare professionals. These datasets offer a rich source of visual information related to pancreatic tissues and cancerous cells. They facilitate research in medical imaging, machine learning, deep learning, and data science. Access to these datasets advances AI systems for analyzing pancreatic images. Detailed information on the currently publicized medical images of the pancreas used for research is provided below.

3.1. NIH (*National Institutes of Health*) [47]

This dataset comprises 82 abdominal CT scans of the pancreas. The scans have been enhanced with contrast for better visualization and have a resolution of 512×512 pixels. It includes 53 male and 27 female subjects, aged between 18 and 76 years. Among the subjects, there are 17 healthy kidney donors and additional patients without pancreatic lesions. To ensure accurate labeling, a medical student manually labeled each slice of the CT scans, under the supervision of an experienced radiologist. This dataset offers an opportunity to investigate the pancreas across different age groups, genders, and health conditions.

3.2. AbdomenCT-1K [48]

This dataset proposed by Ma et al. [48] contains more than 1,000 CT images from 12 medical centers for large-scale studies of the liver, kidney, spleen, and pancreas segmentation to enhance the generalizability of state-of-art models. They also establish benchmarks for fully supervised, semi-supervised, weakly supervised, and continuous learning segmentation and develop corresponding models for each benchmark.

3.3. BTCV (*Beyond the Cranial Vault Multi-organ Segmentation Challenge*) [49]

This dataset comprises 50 abdominal CE-CT scans (30 for training and 20 for testing) obtained from the Vanderbilt University Medical Center, with 13 organs (including the pancreas) labeled in each scan, aiming at a 13-class segmentation task. Each scan is made up by 80 to 225 slices, each with a resolution of 512×512 pixels. The images were manually labeled by the evaluator and checked for accuracy by a radiologist.

3.4. WORD (*Whole abdominal ORgan Dataset*) [50]

This dataset comprises 150 abdominal CT scans (100 for training, 20 for validation, and 30 for testing) of 30495 slices and is the first whole abdominal organ dataset. Each scan is annotated with fine pixel-level annotations for 16 organs (including the pancreas) and sparse graffiti-based annotations, a weakly supervised method that reduces labeling costs.

3.5. MSD (*Medical Segmentation Decathlon*) [51]

This dataset comprises 420 portal venous phase CT scans of PC patients who underwent resection at Memorial Sloan Kettering Cancer Center in New York. The masses include IPMNs, pNET, and PDAC. An abdominal radiologist manually segmented the pancreatic parenchyma and pancreatic mass, including cysts or tumors, on each slice using the Scout application.

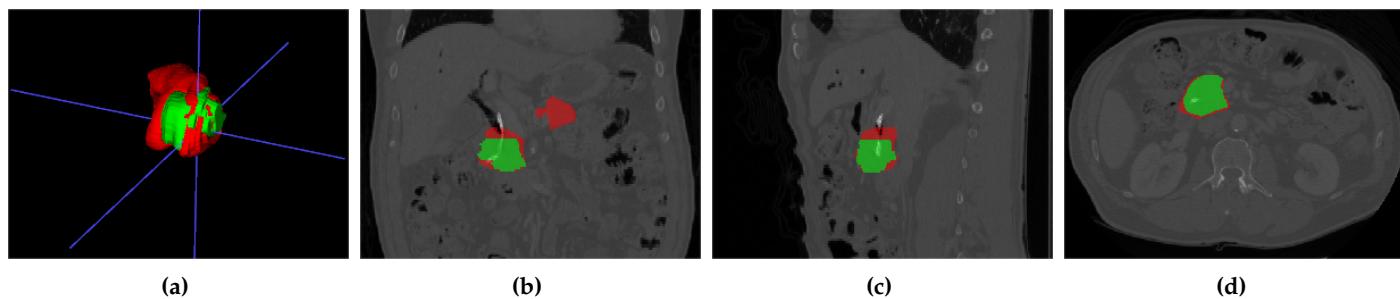


Figure 2. MSD sample data pancreas_004.nii.gz: (a) 3d visualization of pancreas and PC, (b) main view, (c) left view, (d) top view

As shown in Figure 2, ITK-SNAP [52] was used to visualize the CT images, where the red labels represent healthy pancreatic tissues and the green parts represent PC tissues. The task of this dataset is to subdivide the pancreas and the PC, the latter being relatively challenging.

3.6. Dataset of manually segmented pancreatic cystic lesions in CT images [53]

The dataset contains 221 CT images, each manually labeled with pancreatic cysts and main pancreatic ducts. In addition, it contains two nnUNet [54] models, one for segmentation of the pancreas, and one for segmentation of the cysts and main pancreatic duct.

3.7. TCGA (The Cancer Genome Atlas) [55]

The TCGA is a project funded by the public to discover the causes of cancers. It involves multiple cooperating centers that collect, process, and analyze cancer samples. TCGA provides various types of cancer-related data, including diagnostic information, tissue, samples, and radiological images. Radiological images, such as those available in the NIH [47], can be found in the Cancer Imaging Archive (TCIA). Additionally, pancreatic pathological images are accessible through the GDC portal. These resources contribute to the comprehensive collection of data in the TCGA project, aiding researchers in understanding cancer and its characteristics.

3.8. SEER (Surveillance, Epidemiology, and End Results Program) [56]

The SEER program is designed to collect cancer statistics for research to mitigate the effects of cancer. A pancreatic tissue microarray (TMA) containing tumor tissue slides from 161 cases diagnosed between 1983 and 2000 has been established. Of these cases, 154 are PDAC and 7 are pNET. The primary objective of the TMA is to explore the potential prognostic significance of PC tissue slides.

3.9. The PANORAMA Challenge (Pancreatic Cancer Diagnosis: Radiologists Meet AI) [57]

This dataset comprises CE-CT scans, including those from PDAC and non-PDAC patients. The non-PDAC group includes both individuals with healthy pancreas and those with non-PDAC pancreatic lesions. A separate test set of 400 scans has been prepared. This challenge aims to assess the clinical feasibility of modern pancreas-AI solutions for PDAC detection and diagnosis using CE-CT imaging.

3.10. LEPset [58]

The dataset is based on EUS and consists of 420 patients and 3,500 images. Its task is to classify PCs and non-PCs. Experienced physicians annotated these 3,500 images with category labels. In addition, there are 8,000 unlabeled images for pre-training. Sample images from LEPset are shown in Figure 3, with images labeled as PC or non-PC, and unlabeled data.

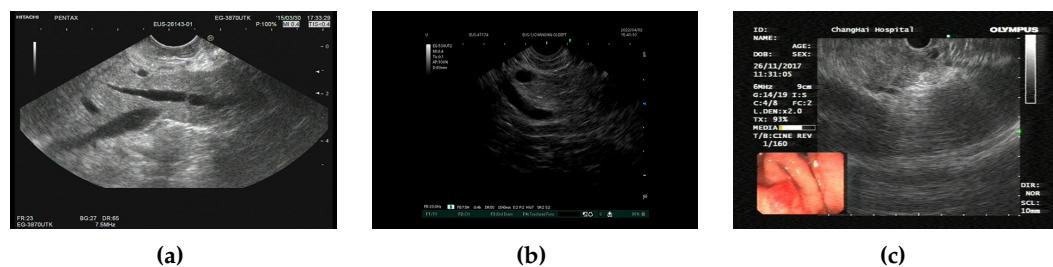


Figure 3. LEPset sample data: (a) labeled non-PC, (b) labeled PC, (c) unlabeled image

3.11. PAIP 2023 (*Tumor Cellularity Prediction in Pancreatic Cancer*) [59]

This dataset comprises 80 pancreatic pathological images (50 for training, 10 for validation, 20 for testing) for tumor cell segmentation, with a resolution of 1024×1024 . They utilize tumor cellularity (TC) as a metric between 0 to 100 to measure the remaining tumor burden in organs. The task of this dataset is to segment the tumor cell nucleus and calculate the TC. Figure 4 shows sample images from the training set, each pathological image corresponds to two masks, representing the TCs of the tumor cell nucleus and non-tumor cell nucleus, respectively.

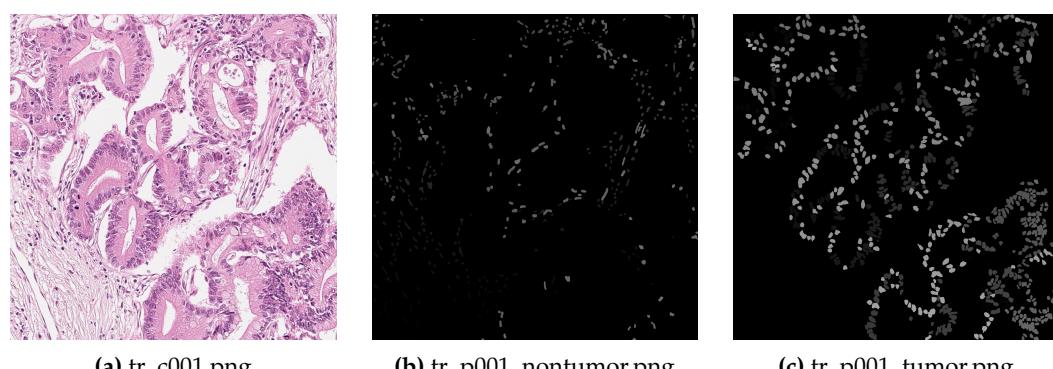


Figure 4. PAIP sample data: (a) a pathological image of PC, (b) non-tumor cell nucleus mask, (c) tumor cell nucleus mask (The masks were processed to be visible)

3.12. Dataset related to article of Grizzi et al. [60]

The dataset contains 7 patients with PDAC, 6 with chronic pancreatitis, and 5 with normal pancreas. Each category includes 10 pathological images for each case at $20\times$ objective. The objective of the dataset is to accurately quantify the amount of pancreatic collagenic extra-cellular matrix, its spatial distribution patterns, and degradation processes by computer-assisted methods.

4. AI Tasks and Evaluation Metrics

In pancreatic image analysis, researchers mainly focused on four popular AI tasks: segmentation, classification, object detection, and prognosis prediction. Figure 5 summarizes the main task applications of AI in different image modalities. To comprehensively and objectively assess the performance of an AI task, appropriate metrics must be used. This section introduces the metrics used in this review, which are widely recognized and commonly used. By utilizing these metrics, researchers can assess the effectiveness of various tasks in a standardized and comparable manner.

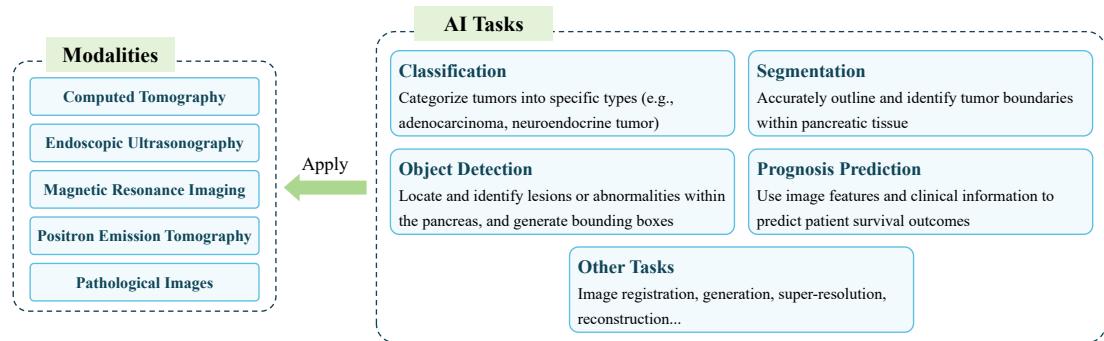


Figure 5. Summary of AI tasks on different medical imaging modalities

The workflow of AI-enabled automated PC analysis is shown in Figure 6. The medical images and clinical data (if available) are first collected and annotated, then the data will be pre-processed as the input of AI models for a certain task. Next, the AI models are trained to learn the features and generate corresponding results. Finally, AI models can be applied to support clinical workflows after they have been assessed as reliable.

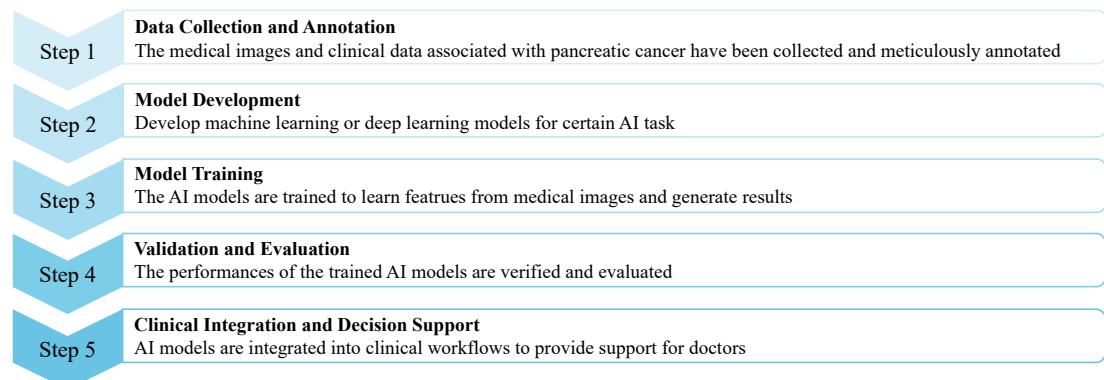


Figure 6. Flowchart of AI application in pancreatic images analysis

4.1. Classification

4.1.1. Introduction to Classification

Image classification is a well-founded task in computer vision that aims to assign labels or categories to the input image as a whole based on its content. This task forms the basis for various applications. The classification of PCs is categorizing medical images into distinct type, including PC and non-PC cases or different subtypes of PC within an image. Classification focuses on recognizing overall patterns and characteristics, rather than providing precise tumor boundaries. Specifically, some basic methods of classification using AI models are shown below:

Features extraction + machine learning Manually extracted features are often interpretable, helping to understand the physical or biological characteristics behind the data, allowing for better control over the model's input, and reducing noise and unnecessary information. In medical image classification, the process of feature extraction typically involves several common methods and follows a systematic workflow. Firstly, Regions of Interest (ROIs) are delineated within the images, focusing on areas relevant to the diagnostic task. Subsequently, various feature extraction techniques are applied to these ROIs including Shape features, which encompass parameters like height, width, perimeter, area, and others to describe geometric properties. Texture features are extracted using methods such as Gray-Level Co-occurrence Matrix (GLCM), Gray-Level Run Length Matrix (GLRLM), Gray-Level Gradient Co-occurrence Matrix (GLGCM), and Gray-Level Distribution Statistics (GLDS). Additionally, Wavelet transform can also be utilized to capture multi-scale texture information. Following feature extraction, a feature selection step is

often performed to reduce dimensionality and remove irrelevant or redundant features. Finally, normalization techniques may be applied to ensure that features are on a comparable scale.

The feature extraction workflow can be seen in Figure 7, and then the normalized data will be fed into machine learning models. Commonly used machine learning models include supervised learning and unsupervised learning. In supervised learning, common classification models include Logistic regression, Decision Trees, K-NearestNeighbor (KNN), Support Vector Machines (SVM), Random Forests (RF), Naive Bayes, and so on. In unsupervised learning, k-means clustering is usually performed on unlabeled data.

Steps of ROI extraction can be performed manually or by segmentation using deep learning algorithms (the segmentation task will be described in a later section). Segmentation can be performed as a separate feature extraction step or included in an end-to-end deep learning workflow.



Figure 7. Basic workflow of feature engineering in traditional machine learning-based image classification

Deep features + machine learning Using deep features and machine learning for image classification combines the powerful feature extraction capabilities of pre-trained CNNs with the robust performance of traditional machine learning algorithms. In this approach, high-level features are extracted by pre-trained CNNs from images, which are input to machine learning algorithms. The primary advantage is the combination of deep learning's feature extraction strength without training from scratch and traditional algorithms' efficiency and flexibility. However, a key disadvantage is that pre-trained models might not always capture the specific characteristics of the target image dataset, potentially leading to suboptimal feature representation and requiring fine-tuning or additional domain-specific training.

End-to-end deep learning Using deep-learning models like CNNs or Transformers directly for image classification can automatically learn features from raw data without time-consuming manual feature extraction. The evolution of image classification models has witnessed a shift towards increasingly sophisticated architectures and techniques. Initially, pioneering models like AlexNet [61] and VGG [62] emphasized the importance of deeper networks for capturing intricate image features. Subsequently, ResNet [63] introduced residual connections, enabling the training of even deeper networks while mitigating the vanishing gradient problem. Furthermore, attention mechanisms have emerged as a pivotal component in image classification, as evidenced by the great success of transformer-based architectures like Vision Transformer (ViT) [64] and Swin-transformer [65]. These models leverage self-attention mechanisms to capture global dependencies and contextual information, allowing for more effective feature representation and classification.

Features extraction + deep learning Some methods, while employing deep learning techniques, do not directly input images but instead undergo a feature extraction process. This approach allows for leveraging both the power of neural networks' learning ability and the explainability of manual feature extraction. However, in this case, neural networks only act as classifiers rather than feature extractors. It is unnecessary to use a very deep neural network that may involve additional computational costs. In most time, a shallow one is enough.

Deep learning + machine learning There are also works combining traditional machine learning methods with deep learning techniques. Traditional machine learning plays a role in integrating multiple deep learning models, reducing the risk of overfitting and capturing a broader range of patterns present in the data. Additionally, ensemble methods can compensate for the weaknesses of individual models, leading to improved overall performance.

4.1.2. Evaluation Metrics for Classification

Accuracy Accuracy measures the proportion of correctly classified instances out of the total instances. It quantifies the overall correctness of the classifier's predictions. The Accuracy ranges from 0 to 1, where 1 indicates the predictions are all correct, while 0 indicates the model does not make any correct predictions. Accuracy is calculated by

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \quad (1)$$

where True Negatives (TN) represent instances correctly predicted as negative, True Positives (TP) represent the number of instances correctly predicted as positive, False Positives (FP) represent the number of instances incorrectly predicted as positive when they are negative. False Negatives (FN) represent the number of instances incorrectly predicted as negative when they are positive.

Specificity Specificity is the probability of a negative test result and refers to the ability of the test to correctly reject a healthy patient without disease. A positive result on a high specificity test close to 1 can be used to diagnose a disease, as the test rarely gives a positive result in a healthy patient [66]. Specificity is calculated by

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}. \quad (2)$$

Precision Precision is the proportion of relevant instances among the retrieved instances. Higher precision means that the algorithm returns more relevant results than irrelevant ones. Precision is calculated by dividing the number of true positives by the total number of elements labeled as belonging to the positive class:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (3)$$

Recall Recall is the probability of a positive test result and refers to the ability of the test to correctly detect a sick patient. A Recall close to 1 can be used to rule out a disease because it rarely misdiagnoses someone with the disease [66]. Recall is calculated by

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (4)$$

Area Under Curve (AUC) AUC is based on receiver operating characteristic (ROC) curves, which are plots of true positive rate (TPR) versus false positive rate (FPR) for each threshold setting. The AUC calculates the area under the ROC curve and summarizes the sensitivity and specificity, but does not provide information on the precision and negative predictive value. AUC is calculated by

$$\text{AUC} = \int_0^1 \text{TPR}(f), d\text{FPR}(f), \quad (5)$$

where $\text{TPR}(f)$ represents the TPR at a specific threshold f , and $d\text{FPR}(f)$ represents the FPR at the same threshold f .

F1-score The F1 score is calculated from the Precision and Recall of the test and represents them symmetrically. The highest possible value for the F1 score is 1.0, which indicates perfect Precision and Recall; the lowest value is 0 if Precision and Recall are 0. F1-score is calculated by

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (6)$$

4.2. Segmentation

4.2.1. Introduction to Segmentation

Image segmentation refers to the precise outlining of the boundaries of the pancreatic organ or PC. With accurate pancreas and PC boundaries, doctors can get the size, shape, location, and other features of the PC for quick diagnosis and treatment planning. Instance segmentation and semantic segmentation are the two basic branches of the segmentation task. Semantic segmentation categorizes each pixel in an image into a predefined class without distinguishing different instances of the same class. Further, instance segmentation not only categorizes pixels but also distinguishes between individual object instances of the same class, assigning unique labels to each instance. Several popular segmentation topics and methods are as follows:

Super-pixel extraction Early-stage pancreas and PC segmentation before the proposal of FCN relies on super-pixel extraction. Super-pixel extraction in medical image segmentation involves grouping pixels with similar characteristics to form cohesive regions. Algorithms like Simple Linear Iterative Clustering (SLIC) or Quickshift partition the image into super-pixels based on color, intensity, or texture similarities, followed by feature extraction and segmentation to assign labels to each region.

FCN/UNet-based methods for 2D segmentation Fully Convolutional Neural Networks (FCNs) [67] revolutionized segmentation by enabling end-to-end learning, allowing pixel-wise predictions directly from input images and offering greater flexibility, efficiency, and performance in image semantic segmentation tasks compared to superpixel selection. Based on this architecture, UNet [68] stands out for its U-shaped design, efficiently capturing both high-level context and precise localization information. A lot of UNet's variants like Attention UNet [69], UNet++ [70], ResUNet++ [71], Channel UNet [72], UNet3+ [73] and so on are proposed aiming at refining segmentation accuracy and addressing specific challenges.

FCN/UNet-based methods for 3D segmentation 3D segmentation in medical imaging offers enhanced accuracy and comprehensive visualization by considering the entire volume of the image, facilitating precise treatment planning, quantitative analysis, and time efficiency in clinical practice. Unlike its 2D counterpart which processes images as two-dimensional matrices, 3D convolution considers the depth, height, and width of the input volume, using three-dimensional kernels to capture spatial dependencies along all three axes. This extension facilitates the modeling of complex volumetric structures and temporal dynamics, making 3D convolution well-suited for tasks involving volumetric medical imaging, video processing, and any application where understanding three-dimensional spatial relationships is essential for accurate analysis. Models like 3D U-Net [74], S3D-UNet [75], and V-Net [76] are advanced 3D CNN architectures designed to segment volumetric medical image.

Transformers for 2D medical image segmentation While FCN/UNet-based approaches have been highly successful, their convolutional layers suffer from flaws in accessing global and remote long-range semantic information. Thus more and more attention is paid to the success of ViT in visual tasks. Transformers can offer significant advantages in 2D medical segmentation for capturing global context information, handle variable input sizes, leverage attention mechanisms for focusing on relevant features, and utilize pre-trained models for transfer learning, ultimately improving segmentation accuracy and performance in medical imaging tasks. Transformer-based UNet variants including pure Transformer models and hybrid models of Transformer and CNNs. Typical models like TransUNet [77], Swin-UNet [78], Transformer-UNet [79], and TransAttUNet [80] show great segmentation performance on a series of medical image segmentation tasks.

Transformers for 3D medical image segmentation The volume medical image segmentation task can also be redesigned as sequence-to-sequence prediction. The transformer operates on a sequence of input embeddings of a 3D input volume $x \in \mathbb{R}^{H \times W \times D \times C}$ with resolution (H, W, D) and C input channels by dividing it into flattened uniform non-

overlapping patches $\mathbf{x}_v \in \mathbb{R}^{N \times (P^3 \cdot C)}$, where (P, P, P) denotes the resolution of each patch and $N = (H \times W \times D)/P^3$ is the length of the sequence.

Graph-based methods for medical image segmentation Graph-based methods utilize the concept of graph theory to represent the image as a graph, where pixels or voxels are nodes, and their relationships are represented by edges. Graph Neural Networks like Graph Convolution Networks (GCN) [81], Graph Attention Networks (GAT) [82] and Graph Isomorphism Networks (GIN) [83] can be used to process such graph representation.

Instance Segmentation Instance segmentation algorithms can be divided into three main branches: two-stage methods like Mask R-CNN [84] and Cascade Mask R-CNN [85] employ a two-step process involving region proposal generation followed by mask refinement; one-stage methods like SOLO [86] and SOLOv2 [87] directly predict object categories and segmentation masks without separate proposal steps, offering efficiency at the cost of some accuracy; emerging query-based approaches, such as QueryInst [88], formulate instance segmentation as a query-driven interaction problem.

4.2.2. Evaluation Metrics for Segmentation

Dice Similarity Coefficient (DSC) DSC indicates the ratio of overlapping pixels between the prediction and ground truth masks to the total number of pixels in the two masks. The coefficient ranges from 0 to 1, where 1 indicates a perfect match between the predicted and ground truth masks, and 0 indicates no overlap. DSC is calculated by

$$\text{DSC} = \frac{2 \times |A \cap B|}{|A| + |B|}, \quad (7)$$

where A represents the predicted segmentation mask or set of pixels, and B represents the ground truth segmentation mask or set of pixels.

Intersection over Union IoU, also known as the Jaccard Index, measures the overlap between a predicted region and a ground truth region. The IoU ranges from 0 to 1, where 1 indicates a perfect overlap between the masks, and 0 indicates no intersection. IoU is calculated by

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}. \quad (8)$$

Based on IoU, mean IoU (mIoU) is calculated as the average of the IoUs for each pixel class:

$$\text{mIoU} = \frac{1}{N} \sum_{i=1}^N \text{IoU}_i, \quad (9)$$

where N is the number of pixel classes. The instance segmentation tasks typically use mIoU to evaluate performance, averaging segmentation results across multiple categories.

Other Metrics In addition to DSC and IoU, there are also distance-based metrics such as Hausdorff distance (HD) [89] and Normalized Surface Distance (NSD) [90]. These metrics consider the distances between two sets of shapes that quantify the difference between the segmentation result and the true label. The larger the distance, the greater the difference between the two shapes and the worse the performance of the models.

The HD is calculated by

$$\begin{aligned} H(A, B) &= \max(h(A, B), h(B, A)), \\ h(A, B) &= \max_{a \in A} \min_{b \in B} \|a - b\|, \end{aligned} \quad (10)$$

where $H(A, B)$ is the HD between A and B, $h(A, B)$ is the directed HD, and $\|\cdot\|$ is some underlying norm on the points of A and B.

The NSD is calculated by

$$\text{NSD} = \frac{|S_i \cap B_j^{(\tau)}| + |S_j \cap B_i^{(\tau)}|}{|S_i| + |S_j|}, \quad (11)$$

where τ is the tolerance, S_i and S_j are surfaces, B_i and B_j are border regions, i is the prediction, j is the reference.

Besides, Accuracy, Specificity, Recall, Precision, AUC, and F1-score can also be used to evaluate segmentation tasks, as discussed in section 4.1. Unlike classification, these metrics evaluate the performance of the segmentation models in terms of their prediction at each pixel point.

4.3. Object Detection

4.3.1. Introduction to Object Detection

Object detection refers to the classification and localization of objects in an image, which involves assigning labels to each detected object and providing a bounding box around the object. By accurately detecting and localizing objects in an image, object detection algorithms enable machines to effectively understand and interact with visual information.

2D Object detection 2D Object detection can be categorized into two main approaches: single-stage and two-stage methods. Single-stage methods, such as SSD [91] and YOLO (You Only Look Once) series [92–99], perform object detection in a single step. They directly predict object bounding boxes and class probabilities from the entire image using a unified network architecture. While these methods are faster, they may sacrifice some accuracy compared to two-stage methods. Two-stage methods, like R-CNN [100], Fast-RCNN [101] and Faster R-CNN [84], divide the object detection process into region proposal generation stage and object classification stage. In the first stage, region proposal networks generate potential object bounding boxes, which are then refined and classified in the second stage. These methods typically achieve higher accuracy but require more computational resources. Transformer-based methods like DETR [102] can capture global context and long-range dependencies in images more effectively. This approach enables accurate object detection by attending to relevant image regions and preserving spatial information through positional encodings.

3D object detection 3D object detection is also beneficial for clinical practice. Volume of Interest (VOI) extraction is a crucial preprocessing step that enables tasks like organ segmentation and tumor classification by localizing relevant structures, reducing computational burden, and enhancing accuracy [103]. The main methods for 3D bounding box detection include generating from 2D box detection [104], generating from coarse segmentation [105], reinforcement learning [106], 3D detection models like using 3D Region Proposal Network [107] and so on.

4.3.2. Evaluation Metrics for Object Detection

IoU In addition to segmentation, IoU could also be used to assess the performance of object detection. In formula

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}, \quad (12)$$

A and B represent the predicted and ground truth bounding boxes, respectively.

mAP mAP is the mean value of Average Precision (AP). The AP involves computing the Precision and Recall values for each class and then integrating Precision to Recall. The integration is performed using the area under the Precision-Recall curve. AP is calculated by

$$\text{AP} = \frac{1}{n} \sum_{k=1}^n (P(k) \times \text{rel}(k)), \quad (13)$$

where n represents the total number of relevant items in the retrieved set, $P(k)$ represents the Precision at cut-off k , $\text{rel}(k)$ is an indicator function equaling 1 if the item at rank(k) is relevant, and 0 otherwise. mAP is calculated by

$$\text{mAP} = \frac{1}{C} \sum_{i=1}^C \text{AP}_i, \quad (14)$$

where C represents the total number of object classes, AP_i represents the AP for class i . mAP@0.5 measures the mAP when the Intersection over Union (IoU) threshold for considering a detection as a true positive is 0.5. mAP@0.5:0.95 measures the mAP averaged over different IoU thresholds ranging from 0.5 to 0.95, typically in increments of 0.05.

Other Metrics If the IoU is greater than a set threshold (e.g., 0.5), the predicted bounding box can be treated as a correct detection. The metrics mentioned by section 4.1 (accuracy, specificity, recall, precision, AUC, and F1-score) can be obtained from the number of correctly predicted bounding boxes versus the number of incorrectly predicted bounding boxes. IoU greater than the threshold is classified as TP , IoU less than the threshold as FP , and IoU of 0 as FN .

4.4. Prognosis Prediction

4.4.1. Introduction to Prognosis Prediction

The prognosis prediction for PC patients integrates medical images and clinical data to forecast the survival period of individuals. This predictive capability assists doctors in making informed prognosis decisions for their patients.

Prognosis prediction is to predict outcomes for PC patients after undergoing surgery, such as predicting whether or not the patient will survive after some time, or for overall survival (OS) time. The predicted results will usually be compared with the real results for consistency, thus evaluating the model performance. The basic methods for prognosis prediction typically include machine learning algorithms like Least Absolute Shrinkage and Selection Operator (LASSO) regression, SVM, or more advanced techniques such as random forests (RF) or gradient boosting. These methods leverage the relationship between input features, such as medical imaging characteristics and clinical variables, and the target variable, such as survival time, to learn predictive models to accurately forecast patient outcomes. Features like tumor size, shape, and texture are traditionally manually extracted from images to inform prognosis models. With the advent of deep learning, automatic feature extraction has gained traction. CNNs autonomously learn discriminative features directly from images, eliminating the need for manual feature engineering.

4.4.2. Evaluation Metrics for Prognosis Prediction

Concordance Index C-index is used to assess the predictive model accuracy in survival analysis. It measures how well a model ranks the relative order of survival times for different individuals. The C-index ranges from 0 to 1, with higher values indicating better predictive accuracy. A C-index of 1 indicates perfect concordance, where the model consistently ranks survival times correctly. A C-index of 0.5 represents random prediction, indicating that the model's predictions are not informative. C-index is calculated by

$$\text{C-index} = \frac{N_C}{N_T}, \quad (15)$$

where N_C represents the number of concordant pairs, and N_T represents the number of all evaluable pairs. Concordant pairs are pairs of individuals where the predicted survival times have the same relative order as the actual survival times. Comparable pairs are pairs where a meaningful comparison can be made, excluding tied or censored survival times.

Other Metrics Prognosis prediction also includes predicting whether a patient will die and whether PC will develop metastasis within a period after treatment. In this case, the previously mentioned metrics Accuracy, Specificity, Precision, Recall, AUC, and F1-score could evaluate the AI models on prognosis prediction task as well.

4.5. Other Tasks

In addition to the mentioned tasks, there are several other common AI tasks in medical image analysis that could potentially aid in the diagnosis and treatment of PC. These include registration, as well as various low-level visual tasks.

Image registration involves aligning multiple images from different sources or times, aiding in tracking disease progression or integrating data from different imaging modalities. Image generation encompasses techniques for creating new images based on existing ones, such as generating synthetic images to augment training data or simulating different imaging scenarios for educational purposes. Super-resolution techniques enhance image resolution, enabling the detection of finer details in PC imaging, and aiding in the identification of smaller lesions or abnormalities. Denoising methods remove noise from images, improving clarity and facilitating the identification of relevant features in PC images, especially against a noisy background. Reconstruction involves creating complete images from partial or incomplete data, such as reconstructing 3D images from 2D scans. The Medical Visual Question Answering (MedVQA) task combines computer vision and Natural Language Processing (NLP) by analyzing input medical images and related questions and outputting answers to inform medical diagnosis and treatment. The above mentioned tasks utilize a variety of AIs that can help improve the quality of imaging data and the accuracy of automated analysis, helping physicians diagnose and treat PCs more effectively.

5. Computed Tomography (CT)

5.1. Introduction to CT

CT, also known as computed axial tomography (CAT), is a non-invasive imaging technique that rapidly produces three-dimensional imaging of the inside of the body. It is the most widely used radiologic imaging method and has become a standard. Compared to conventional radiography, CT offers higher contrast. The advent of CT revolutionized the field of medical imaging, becoming an indispensable tool for diagnosis and treatment [?]. CE-CT uses iodinated contrast agents to increase the visibility of blood vessels, distinguishing them from their surroundings. This approach increases clarity and provides more detail to better analyze anatomy and potential abnormalities. However, iodinated contrast agents also have side effects, such as causing nephropathy [108].

5.2. Classification

Features extraction + machine learning Li et al. [109] used six methods for feature extraction, as shown in Table 1, and the LASSO algorithm for feature selection, and then applied the EL-SVM learner to classify normal pancreas, early-stage (stage I and stage II), stage III, and stage IV of PC. Chen et al. [110] trained an XGBoost [111] model to classify patches as cancerous or noncancerous. Patients were classified as either PDAC or non-PDAC based on the proportion of patches classified as cancerous. Mukherjee et al. [112] conducted feature extraction, normalization, and reduction, and trained four independent ML classifiers known as KNN, SVM, RF, and XGBoost to recognize PDAC at the prediagnostic stage, which achieved high accuracy.

Table 1. Features extracted by six methods in [109]

Methods	Feature name
Shape	height, width, perimeter, area, complexity, rectangularity, elongation, equivalent area radius
GLCM	mean and standard deviation of energy, entropy, moment of inertia, and correlation
GLRLM	short run emphasis, long run emphasis, gray-level nonuniformity, run percentage, run-length nonuniformity, low gray-level run emphasis, high-gray level run emphasis
GLGCM	small grads dominance, big grads dominance, gray asymmetry, grads asymmetry, energy, gray mean, grads mean, gray variance, grads variance, correlation, gray entropy, grads entropy, entropy inertia, differ moment
GLDS	mean, contrast, angular second moment, entropy
Wavelet transform	

End-to-end deep learning Liu et al. [113] used the VGG model to differentiate PC tissue from non-cancerous pancreatic tissue. Xia et al. [114] proposed a deep classification model that combined UNet with Anatomy-aware Hybrid Transformers using a single-phase non-contrast CT to facilitate more accurate, safe, and low-cost screening for distinguishing between PDAC, other abnormalities, and normal pancreas. Cao et al. [115] introduced PC detection with artificial intelligence (PANDA) method to detect and classify pancreatic lesions based on the lesion segmentation results of nnUNet. CNNs with a classification head were used to classify PDAC, pNET, SPT, IPMN, MCN, chronic pancreatitis, SCN, etc. Segmentation and classification models are included in an end-to-end scheme.

Features extraction + deep learning Vaiyapuri et al. [116] proposed an IDLDMS-PTC technique to examine the CT images for the existence of pancreatic tumors. The proposed technique comprises several sub-processes: GF-based pre-processing, EPO-MLT-based segmentation, MobileNet-based feature extraction, AE-based classification, and MLO-based parameter optimization. Huy et al. [117] used Densenet to distinguish cancerous tumors from benign tumors in CT pancreatic images.

Deep learning + machine learning To classify pancreatic SCNs and MCNs, Yang et al. [118] applied a multi-channel-multiclassifier-RF-ResNet (DNN-MMRF-ResNet). SVM, KNN, and Bayes classifiers were used after the Residual block and then the final classification was finished by an RF classifier. Bakasa et al. [119] utilized Inception V3, VGG16, and ResNet34 as weak learners in a stacking ensemble, where their first-level predictions formed the input for XGBoost that performed the final pancreas cancer classification.

Table 2. Summary of AI methods in CT images for classification task

Reference	Method	Dataset	Study population	Performance
Li et al. [109]	LASSO regression and EL-SVM learner	A private dataset	168 scans	AUC = 0.7308 (normal-early stage), 0.6587 (normal-stage III), 0.7333 (normal-stage IV)
Chen et al. [110]	XGBoost	A private dataset, MSD and NIH	27,235, 5,715, and 7,054 images	AUC = 0.97 (private test set), 0.83, and 0.89 (public test set)
Mukherjee et al. [112]	KNN, SVM, RF and XGBoost	A private dataset and NIH	596 and 82 scans	AUC = 0.95, 0.98, 0.95, and 0.96
Liu et al. [113]	VGG	A private dataset, MSD and NIH	14,780, 4,849, and 1,427 images	Accuracy = 0.986, 0.989 (private test set), and 0.832 (MSD and NIH test set)
Xia et al. [114]	UNet with Anatomy-aware Hybrid Transformers	A private dataset	1627 scans	Recall = 0.952, Specificity = 0.958
Cao et al. [115]	PANDA	Five private dataset	3,208, 786, 5,337, 18,654, and 4,815 scans	Specificity = 0.999, Recall = 0.929, AUC = 0.986-0.996
Vaiyapuri et al. [116]	IDLDMS-PTC	A private dataset	500 images	Accuracy = 0.9935, Specificity = 0.9884, Recall = 0.9935, F1-score = 0.9948
Huy et al. [117]	DenseNet	NIH and MSD	18,942 and 15,000 images	Accuracy = 0.974, Specificity = 0.966, Recall = 0.983
Yang et al. [118]	DNN-MMRF-ResNet	A private dataset	110 scans	Precision = 0.9387, Recall = 0.9136, Specificity = 0.9380, Accuracy = 0.9269
Bakasa et al. [119]	Stacking ensemble	NIH	80 scans	Accuracy = 0.988

5.3. Segmentation

Super-pixel extraction Roth et al. [120] extracted superpixels from the abdominal region are extracted via SLIC. Initial probability response maps are generated using a two-level cascade of RF classifiers, retaining superpixels with probabilities above 0.5, followed by CNN sampling bounding boxes at various scales and non-rigid deformations for refined pancreas region identification. Roth et al. [47] introduced a probabilistic bottom-up approach to segment the pancreas in abdominal CT scans, employing multi-level deep CNNs. Various ConvNets variations are evaluated for hierarchical classification on image patches and regions (superpixels), with post-processing using structured predictions.

FCN/UNet-based methods for 2D segmentation Heinrich and Oktay [121] developed BRIEFnet, which utilized binary sparse convolutions in CNNs to reduce memory cost and improve segmentation performance. Zhou et al. [122] utilized pre-trained FCN-8s incorporated with deeply-supervised nets (DSN) [123] to develop a coarse-to-fine segmentation algorithm. The model obtained a reasonable segmentation of pancreatic cysts. Lu et al. [124] proposed a Ringed Residual U-Net using the ring residual module as well as the attention mechanism. Boers et al. [125] implemented the interactive method iFCN and introduced iUNet, an interactive version of the U-net method, which is fully trained for optimal initial segmentation and additionally fine-tuned on user-generated scribbles in in-

teractive mode. Jiang et al. [126] proposed DLU-Net with deformable convolution modules to strengthen the ability to model the target edge, and the Bi-Directional Convolutional Long-Short Term Memory (BConvLSTM) was utilized to merge the features of different scales. Li et al. [127] used the skip network, residual network, and multiscale residual network strategies to efficiently address over- and under-segmentation issues through cross-domain connections and multiscale convolution operations, enhancing accuracy in pancreas shape learning. Li et al. [128] proposed a Window Attention Upsample (WAU) for upsampling, consisting of an Attention Decoder (AD) and a bilinear upsample. A window attention scheme is used to reduce computation by restricting computation in local windows instead of the global range. Juwita et al. [129] proposed M3BUNet, which fused MobileNet and UNet and was equipped with Mean-Max attention. In addition, they utilized a coarse-to-fine segmentation process to improve performance. Paithane and Kakarwal [130] introduced a 12-layer LMNS-net with 4 convolution layers, where a lightweight multiscale block dropped the unused information.

Some methods first locate the organ of interest, such as the pancreas, and then identify any abnormalities or lesions within it. This two-step process involves initially segmenting the organ from surrounding structures and then focusing on regions of interest within the organ for further analysis. Zhou et al. [131] proposed DBFE-Net with two branches. DB-Net is used to extract semantic and fine-grained features for pancreas segmentation with a coarse-to-fine strategy, and then FE-Net is used to extract fine-grained features with higher contrast for tumor segmentation in the pancreas region.

Some works focus on utilizing spiral transformation to map 3D images onto 2D planes while preserving spatial relationships, facilitating effective 3D contextual information utilization in a 2D model. Chen et al. [132] applied spiral transformation for data augmentation and incorporated a transformation-weight-corrected module based on Res-UNet [133]. This design addressed small sample size issues and ensured uniform 3D segmentation and rebuilding constraints, overcoming non-unique 3D results from uniform sampling.

FCN/UNet-based methods for 3D segmentation Roth et al. [134] investigated the 3D U-Net of two types of pancreas segmentation, one with concatenation and one with summation skip connections. Chen et al. [135] introduced a new bias-dice loss function for improved efficiency in 3D coarse segmentation, utilizes a dimension adaptation module (DAM) to incorporate 3D information into 2D networks, and proposes a fusion decision module and parallel training strategy to integrate multi-source feature cues from sub-networks for final predictions. Zhao et al. [136] proposed a two-stage framework that utilized a 3D UNet to provide candidate regions in the first stage, and another 3D UNet was trained to obtain the final results based on these candidates in the second stage. Zhang et al. [137] proposed a dynamic on-demand network (DoDNet) with a dynamic segmentation head, addressing the partially labeled issue in medical images and being applied to multiple tumors. They also proposed a large-scale partially labeled dataset MOTS for pretraining models. Zhang et al. [138] developed the scale-transferrable feature fusion module (STFFM) and prior propagation module (PPM) modules to simplify FCNs. STFFM utilized the scale-transferrable operation to learn rich fusion features, and PPM explored informative spatial priors by dynamically adapting the spatial priors to input and feature maps.

nnUNet Isensee et al. [54,139] introduced the no-new-Net (nnUNet), a robust and self-adapting framework based on 2D and 3D vanilla U-Nets without using various extension plugins (residual connections, Dense connections, and various attention mechanisms), which can automatically adapt architectures to image geometry. In addition, they defined steps for nnUNet: pre-processing, training, inference, and potential post-processing. Yao et al. [140] employed nnUNet for IPMN segmentation and achieved a better DSC than the previous studies. In recent years, nnUNet has achieved remarkable success and widespread application in medical image segmentation competitions, prompting a rethinking of the task. Effective preprocessing, post-processing, training, and inference strategies may be more important than complex network architectures.

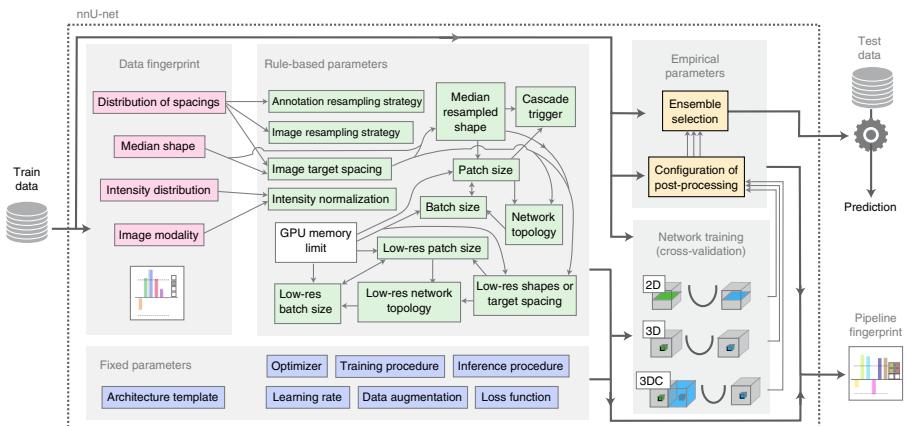


Figure 8. Automated method configuration for deep learning-based biomedical image segmentation in nnUNet [54]

Transformers for 2D medical image segmentation Sha et al. [79] proposed TransformerUnet, which combined Transformer and UNet by replacing Transformer modules in raw images with feature maps in UNet. Huang et al. [141] introduced Medical Image Segmentation tranSFormer (MISSFormer), a hierarchical encoder-decoder network. They redesigned the feed-forward network with the Enhanced Transformer Block and used the Enhanced Transformer Context Bridge to extract long-range dependencies and local context of multi-scale features. Chen et al. [77] proposed TransUNet that combined Transformers and UNet. The Transformer encoders tokenized image patches from the CNN feature map to obtain global contexts, and these encoded features were combined with high-resolution CNN feature maps for precise localization. Cao et al. [78] introduced Swin-UNet, a UNet-like pure Transformer that uses a hierarchical Swin Transformer encoder to extract context features and a symmetric decoder to restore spatial resolution. Dai et al. [142] put forward a two-stage Trans-Deformer network (TD-Net), with a 2D UNet for coarse segmentation and ViT for fine segmentation. In this framework, the multi-input module was designed to focus on high-frequency texture information, and the Scale Inter-active Fusion (SIF) module was designed to combine local and global features. Rahman et al. [143] proposed a Medical Image Segmentation Transformer (MIST) using Convolutional Attention Mixing (CAM) to capture local contexts of pixels in multimodal dimensions.

Transformers for 3D medical image segmentation Zhou et al. [144] proposed not another transFormer (nnFormer) that combined interleaved convolution and self-attention operations and utilized local and global volume-based self-attention mechanisms. Besides, they replaced the traditional concatenation or summation in skip connections with skip attention in UNet-like architecture. Hatamizadeh et al. [145] put forward UNet Transformers (UNETR) utilizing a skip-connected transformer encoder to capture global multi-scale information. Tang et al. [146] introduced Swin UNETR, a self-supervised framework, which utilized an encoder to extract features from multiple resolutions and was pre-trained on 5,050 public CT images. The model can also be applied to various proxy tasks after fine-tuning. Chen et al. [147] extended 2D TransUNet to 3D TransUNet, which tokenized image patches from a CNN feature map using a Transformer encoder and the Transformer decoder adaptively refined candidate regions by employing cross-attention between candidate proposals and U-Net features. Qu et al. [148] introduced a transformer-guided progressive fusion network (TGPFN), which supplemented long-range dependencies of convolutions by global representation captured by the transformer.

Graph-based methods for medical image segmentation Guo et al. [149] proposed a layered optimal graph image segmentation of multiple objects and surfaces (Deep LOGISMOS) method utilizing a UNet, trained on adjacent 2D patches centered at the tumor to provide contextual segmentation, refined by Gaussian Mixture Model (GMM) and morphological operations, followed by segmentation graph construction using UNet probability

maps and a max-flow algorithm for globally optimal segmentation. Soberanis et al. [150] improved UNet based on uncertainty analysis and GCNs, which trained a GCN to solve a semi-supervised graph learning problem about the uncertainty levels of a particular input volume. Hu et al. [151] proposed a distance-based saliency-aware model (DSD-ASPP-Net), a coarse-to-fine framework that trained a Dense Atrous Spatial Pyramid Pooling (DenseA-SPP) model to learn location and probability map of the pancreas for coarse stage and saliency-aware modules for fine stage. Zhao et al. [152] introduced a holistic segmentation-mesh-classification network (SMCN) that combined geometry and location information and a graph-based residual convolutional network (Graph-ResNet) with nodes fused the information of the mesh model and feature vectors of the segmentation network. Liu et al. [153] developed a graph-enhanced pancreas segmentation network (GEPS-Net), which added a graph enhancement module to UNet to extract the spatial relationship information.

Neural architecture search (NAS) NAS optimizes segmentation models by automatically finding the best network structures for improved performance. It adjusts parameters like depth and width to suit specific tasks and datasets, overcoming limitations of manual design and enhancing model accuracy and efficiency. Zhu et al. [154] employed a NAS for volumetric medical image segmentation (V-NAS), which could choose 2D, 3D, or Pseudo-3D (P3D) convolutions at each layer automatically. He et al. [155] proposed the Differentiable Network Topology Search (DiNTS) scheme, including a topology-guaranteed discretization algorithm and a discretization-aware topology loss. Besides, DiNTS could search 3D networks under different GPU memories and significantly reduce training time.

Utilizing the power of large models Large models, also known as foundation models, refer to deep learning models with extensive parameters and complex computational architectures, which can offer improved performance by capturing intricate patterns and relationships in data, enabling enhanced representation learning, flexibility across diverse domains, state-of-the-art results, and scalable handling of growing datasets and complex tasks. Liu et al. [156,157] proposed the CLIP-Driven Universal Model based on Contrastive Language-Image Pretraining (CLIP) [158]. Using transfer learning on 3,410 CT scans, they trained a universal model to capture anatomical relationships. Huang et al. [159] test the SAM on a built large medical dataset using different modes containing 18 modalities, 84 objects, 1050K 2D images, and 6033 masks. The sources of the collected dataset also included AbdomenCT-1K with images related to the pancreas. Mazurowski et al. [160]'s experimental datasets included MSD-pancreas. He et al. [161] found that SAM showed the lowest segmentation performance on the pancreas over 10 different organs (brain, chest, lung, liver, pancreas, prostate, bowel, skin, heart, and breast), and concluded that SAM is not as accurate as dataset specific deep-learning algorithms in medical images. The road to zero-shot segmentation for the pancreas and PC is still long.

Federated learning Federated learning is one of the machine learning methods and allows model training on decentralized devices or servers, preserving local data samples while protecting user privacy and data security. Models can be learned from different data sources without sharing the original data, making it particularly suitable for applications in healthcare. Knolle et al. [162] proposed a shallow and U-Net-like framework MoNet based on repeated dilated convolutions with decreasing dilation rates. This framework reduces inference time and memory compared to UNet variants and is suitable for federated learning. Wang et al. [163] put forward the conditional distillation federated learning (ConDistFL) framework, which combined federated learning with knowledge distillation. This framework was trained on images of various organs and could extract knowledge of unlabeled tumors from labeled ones. Their study also increased the stability and reduced the training time.

Reinforcement learning Reinforcement learning maximizes task rewards by training agents to observe images and take actions, and can be applied to a variety of computer vision tasks such as object detection, image segmentation, and behavior recognition. The key to this approach lies in designing appropriate state space, action space, and reward functions for effective learning strategies. Man et al. [164] introduces a deep Q network (DQN)

driven approach combined with a deformable U-Net architecture to address challenges in pancreas segmentation in medical image analysis, achieving accurate segmentation by interacting with contextual information and capturing geometry-aware features.

Instance Segmentation Dogan et al. [165] combined semantic segmentation and instance segmentation and proposed a two-phase approach. The first stage is Pancreas Localization, detecting the rough pancreas position on 2D CT slices by adopting the Mask R-CNN model. The second phase, Pancreas Segmentation, used the 3D U-Net model to refine the candidate pancreas region on 2D sub-CT slices.

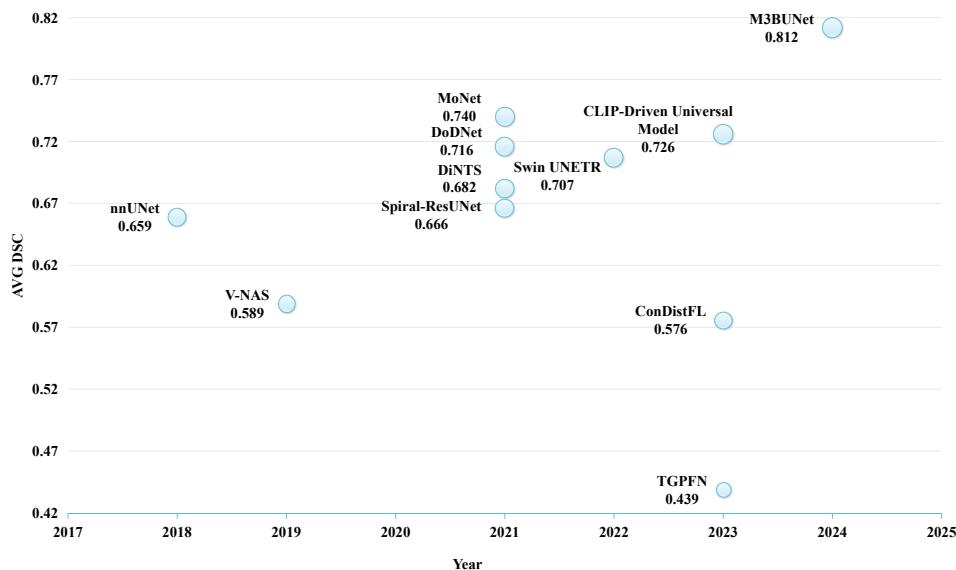


Figure 9. Summary of AI models' segmentation performance for pancreas and PCs on MSD

Figure 9 shows the average DSC of the pancreas and PCs for the AI models on the MSD dataset from 2018 to 2024. While the overall performances improve over time, they still lag behind other organs.

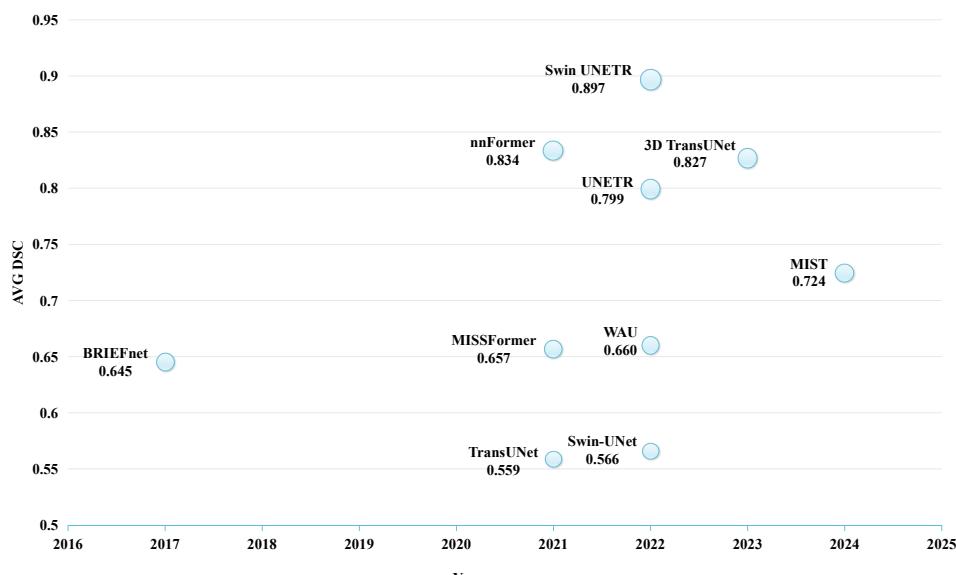


Figure 10. Summary of AI models' segmentation performance for pancreas on BTCV

Figure 10 shows the DSC of the pancreas for the AI models on the BTCV dataset from 2017 to 2024. Segmentation models for 3D images showed superior results.

Table 3. Summary of AI methods in CT images for segmentation task

Reference	Method	Dataset	Study population	Performance
Roth et al. [120]	SLIC	NIH	82 scans	DSC = 0.81
Roth et al. [47]	Probabilistic bottom-up approach	NIH	82 scans	DSC = 0.805
Heinrich and Oktay [121]	BRIEFnet	BTCV	30 scans	DSC = 0.645
Zhou et al. [122]	FCN-8s with DSN	A private dataset	131 scans	DSC = 0.6344 ± 0.2771
Lu et al. [124]	Ringed Residual U-Net	NIH	82 scans	DSC = 0.8832 ± 0.0284
Boers et al. [125]	iUNet	A combination of TCIA and BTCV, and a private dataset	90 and 1905 scans	DSC = 0.87
Jiang et al. [126]	DLU-Net	MSD and a private dataset	281 and	DSC = 0.9117 and 0.9094, Accuracy = 0.9725 and 0.9743
Li et al. [127]	Custom segmentation network	NIH	82 scans	DSC = 0.8757 ± 0.0326
Li et al. [128]	WAU	BTCV	30 scans	DSC = 0.6601
Juwita et al. [129]	M3BUNet	NIH and MSD	82 and 281 scans	DSC = 0.8952 and 0.8860, IoU = 0.8116 and 0.7990
Paithane and Kakarwal [130]	LMNS-net	NIH	82 scans	DSC = 0.8868, IoU = 0.9882, Precision = 0.6822, Recall = 0.9866
Zhou et al. [131]	DBFE-Net	Two private datasets	116 and 42 scans	Precision = 0.6573 (PCs), 0.8907 (abnormal), and 0.9147 (normal)
Chen et al. [132]	Spiral-ResUNet	MSD	281 scans	DSC = 0.6662
Roth et al. [134]	3D UNet	A private dataset	147 scans	DSC = 0.897 ± 0.038
Chen et al. [135]	CNN with Bias-dice loss function	NIH	82 scans	DSC = 0.8522
Zhao et al. [136]	3D UNet-based two-stage framework	NIH	82 scans	DSC = 0.8599
Zhang et al. [137]	DoDNet	MSD	281 scans	DSC = 0.7155, HD = 11.70
Zhang et al. [138]	CNNs with STFFM and PPM modules	NIH and MSD	82 and 281 scans	DSC = 0.8490 and 0.8556
Isensee et al. [139]	nnUNet	MSD	281 scans	DSC = 0.659
Yao et al. [140]	nnUNet	A private dataset	61 scans	DSC = 0.73
Sha et al. [79]	Transformer-UNet	NIH	82 scans	mIoU = 0.8301, DSC = 0.7966
Huang et al. [141]	MISSFormer	BTCV	30 scans (3779 slices)	DSC = 0.6567
Chen et al. [77]	TransUNet	BTCV	30 scans (3779 slices)	DSC = 0.5586
Cao et al. [78]	Swin-UNet	BTCV	30 scans (3779 slices)	DSC = 0.5658
Dai et al. [142]	TD-Net	NIH and MSD	82 and 281 scans	DSC = 0.8989 and 0.9122
Rahman et al. [143]	MIST	BTCV	30 scans (3779 slices)	DSC = 0.7243

Table 3. *Cont.*

Reference	Method	Dataset	Study population	Performance
Zhou et al. [144]	nnFormer	BTCV	30 scans	DSC = 0.8335
Hatamizadeh et al. [145]	UNETR	BTCV	30 scans	DSC = 0.799
Tang et al. [146]	Swin UNETR	BTCV and MSD	30 and 281 scans	DSC = 0.897 and 0.7071
Chen et al. [147]	3D TransUNet	BTCV	30 scans	DSC = 0.8269
Qu et al. [148]	TGPFN	Three private datasets and MSD	313, 53, 50, and 420 scans	DSC = 0.8051, 0.6717, 0.6925, and 0.4386
Guo et al. [149]	Deep LOGISMOS	A private dataset	50 scans	DSC = 0.823 ± 0.078
Soberanis et al. [150]	Improved UNet based on uncertainty analysis and GCNs	NIH	82 scans	DSC = 0.778 ± 0.063
Hu et al. [151]	DSD-ASPP-Net	NIH	82 scans	DSC = 0.8549 ± 0.0477
Zhao et al. [152]	SMCN with Graph-ResNet	A private dataset	661 scans	DSC = 0.738 (PDAC)
Liu et al. [153]	GEPS-Net	NIH	82 scans	DSC = 0.8226 ± 0.0648, IoU = 0.7036 ± 0.0887, HD = 7.88 ± 9.29
Zhu et al. [154]	V-NAS	NIH and MSD	82 and 281 scans	DSC = 0.8515 and 0.5886
He et al. [155]	DiNTS	MSD	281 scans	DSC = 0.6819, NSD = 0.8608
Liu et al. [156,157]	CLIP-Driven Universal Model	MSD	281 scans	DSC = 0.7259, NSD = 0.8976
Huang et al. [159]	SAM	AbdomenCT-1K	1,000 images	DSC = 0.7686 (box)
He et al. [161]	SAM	MSD	281 scans	DSC = 0.0547 (box)
Knolle et al. [162]	MoNet	MSD	281 scans	DSC = 0.74 ± 0.11
Wang et al. [163]	ConDistFL	MSD	281 scans	DSC = 0.5756
Man et al. [164]	DQN	NIH	82 scans	DSC = 0.8692 ± 0.0492
Dogan et al. [165]	Mask-RCNN	NIH	82 scans	DSC = 0.8615 ± 0.0445, IoU = 0.7593 ± 0.646

5.4. Object Detection

There are also several object detection works in CT images. Zhang et al. [166] proposed a pancreatic tumor detection framework that incorporated augmented feature pyramid networks, self-adaptive feature fusion, and a dependencies computation module. The framework also leveraged contextual information at multiple scales to improve detection accuracy. Baumgartner et al. [167] proposed nnDetection, a self-configuring method based on Retina U-Net [168] that could be deployed on arbitrary medical detection tasks. Juneja et al. [169] introduced a region-based CNN (RCNN)-crop method inspired by the region proposal network (RPN) and feature pyramid network (FPN). This approach extracts a cropped patch of the pancreatic region of interest (ROI) from CT images to promote accurate detection of PC. Dinesh et al. [170] proposed a novel YOLO model-based CNN (YCNN) for predicting PC in medical images. Their model utilized the YOLO architecture and CNNs to achieve efficient and accurate detection of pancreatic tumors.

762
763
764
765
766
767
768
769
770
771
772
773
774

Table 4. Summary of AI methods in CT images for object detection task

Reference	Method	Dataset	Study population	Performance
Zhang et al. [166]	Custom pancreatic tumor detection network	A private dataset	2,890 images	Recall = 0.8376, Specificity = 0.9179, Accuracy = 0.9018
Baumgartner et al. [167]	nnDetection	MSD	281 scans	mAP@0.1 = 0.766 (cross validation) and 0.791 (test set)
Juneja et al. [169]	RCNN-Crop	NIH	19,000 images	mAP@0.5 = 0.281
Dinesh et al. [170]	YCNN	A private dataset	7,245 images	AUC = 1.00, F1-score = 0.99, Accuracy = 1.00

5.5. Prognosis Prediction

Lee et al. [171] utilized ensemble learning to combine clinical data-based machine learning models (RF, GB, LR, NN, and SVM) and CT data-based deep learning models (3D ResNet-18 [172], R(2 + 1)D-18 [172], 3D ResNeXt-50 [173], and 3D DenseNet-121 [173]), leveraging preoperative data to predict postoperative survival. Zhang et al. [174] introduced a risk score-based feature fusion technique that integrated radiomics and transfer-learning features, aiming to improve the OS prediction performance for PDAC patients. Additionally, Yao et al. [140] developed a 3D contrast-enhanced convolutional long short-term memory network (CE-ConvLSTM) that leverages tumor-vascular relationships for predicting the OS of PDAC patients. Chen et al. [175] developed a dual-transformation-guided contrastive learning scheme that effectively addressed data limitations and achieved excellent performance in predicting lymph node metastasis in PC.

Table 5. Summary of AI methods in CT images for prognosis prediction task

Reference	Method	Dataset	Study population	Performance
Lee et al. [171]	Ensemble learning	A private dataset	282 scans	AUC = 0.76 (2-year OS) and 0.74 (1-year recurrence-free survival)
Zhang et al. [174]	RF	A private dataset	98 scans	AUC = 0.84
Yao et al. [140]	CE-ConvLSTM	Three private datasets, MSD and a combined dataset [176]	296, 571, 61, 281 and 90 scans	C-index = 0.651
Chen et al. [175]	Custom contrastive learning scheme	A private dataset	157 scans	Accuracy = 0.744 , AUC =0.791, Recall = 0.740, Specificity = 0.750

5.6. Other Tasks

Image reconstruction/ denosing/ super-resolution Lyu et al. [177] reviewed 47 patients with pathologically confirmed PC who underwent baseline multiphasic CE-CT scans and used deep learning method for reconstruction, which enhances spatial resolution and reduces noise texture, improving accuracy in predicting PC resectability and reducing interreader variability while optimizing the tradeoff between spatial resolution and image noise in thin-slice CT images. Noda et al. [178] reconstructed pancreatic low-dose CT using deep learning image reconstruction and compared them with those of images reconstructed using hybrid iterative reconstruction. Chi et al. [179] proposed a Low Dose CT image super-resolution network that addresses spatial resolution loss and artifacts. It featured a dual-guidance feature distillation backbone containing a dual-guidance fusion module (DGFM) and a sampling attention block (SAB) and introduced the denoising head before

and after the super-resolution head in each path to suppress residual artifacts. Takai et al. [180] found that deep-learning-based reconstruction substantially decreased background noise and enhanced both signal-to-noise ratio and contrast-to-noise ratio in pancreatic protocol CT scans at 80 kVp. Additionally, the highest quality and visibility of PDAC were achieved with the high-strength level of the deep learning reconstruction method. Shi et al. [181] proposed SR4ZCT, a self-supervised method that uses off-axis training to handle various combinations of resolution and overlap, explicitly modeling the relationship between resolutions and voxel spacings to accurately simulate training images matching the original through-plane images.

Image generation Liu et al. [182] trained a self-attention cycleGAN based on cone-beam CT (CBCT) acquired prior to the first fraction of treatment from thirty patients previously treated with pancreas SBRT to generate synthetic CTs. CT-based contours and treatment plans were then compared between first-fraction CBCTs and synthetic CTs. Similarly, Dai et al. [183] used cycleGAN to generate synthetic CT images from given CBCT images then trained the mask-scoring regional CNN (MS R-CNN) on generated images for segmentation. Shi et al. [184] introduced 3DGANet, utilizing GANs to produce realistic 3D CT images of PDAC. Its integration of a 3D U-Net architecture enhances the learning of shape and texture, improving efficiency and accuracy by preserving contextual information between slices, validated across diverse datasets, offering a promising solution to address data scarcity. Hooshangnejad et al. [185] developed a generation model named deepPERFECT that can capture minor differences and generate deformation vector fields to transform diagnostic CT into preliminary planning CT of PC, avoiding harm to patients because of separate image acquisition. Peng et al. [186] used TranscycleGAN to synthesize CECT from NECT and augment the amount of CT images. All real and synthesized CT images were used to train the modified 3D U-Net for the automatic delineation of gross tumor volume. Guan et al. [187] proposed a texture-constrained multichannel progressive GAN (TMP-GAN), using joint training of multiple channels. An adversarial learning-based texture discrimination loss is used to further improve the fidelity of the synthesized images and a progressive generation mechanism to improve the accuracy of the image synthesizer. Experiments of generating pancreatic tumor CT images were conducted.

6. Magnetic Resonance Imaging (MRI)

6.1. Introduction to MRI

Magnetic resonance imaging (MRI) is a non-invasive medical imaging technique that uses nuclear magnetic resonance (NMR) to create detailed, high-contrast, three-dimensional images of the body for diagnostic purposes [188–190]. Unlike X-rays or CT scans, MRI uses non-ionizing radiation. It produces excellent contrast images of both soft and hard tissues, by utilizing static and slowly varying magnetic fields and electromagnetic energy in the high to very high-frequency bands [188]. Compared to typical CT scans, MRI provides superior contrast images. MRI scans are generally more time-consuming, taking 20 to 90 minutes depending on the body part being imaged, but they are painless and do not cause tissue damage [191].

However, MRI is not suitable for patients with certain metallic implants due to its reliance on magnetic fields and electromagnetic energy [188]. It is also important to note that MRI tends to be relatively more expensive. Nevertheless, despite these limitations, MRI remains an important tool in clinical diagnosis, providing crucial anatomical and pathological information to assist physicians in making accurate diagnostic and therapeutic decisions. As technology continues to advance, MRI may further improve its imaging speed and expand its range of applications, offering patients more accurate and convenient diagnostic services.

6.2. Classification

Feature extraction + machine learning Cui et al. [192] applied LASSO regression to classify low and high-grade branching type IPMNs (BD-IPMNs). They determined ROIs

with radiologists and extracted features including histograms, texture parameters, RLM (run length matrix) GLCM and form factor parameters using MITK software. A linear combination of selected features with weights was used for grade prediction.

End-to-end deep learning Chen et al. [193] introduced PCN-Net for distinguishing between MCNs and SCNs in T2 and T1 weighted MRIs. The backbone of this framework utilized a pre-trained InceptionV3 [194]. The fusion of the two modalities was achieved through a fusion algorithm, followed by a voting algorithm to obtain the results. In another study, Chen et al. [195] proposed a weighted loss function and applied it to various CNNs. It is proved that this weighted loss function could improve the accuracy of most CNNs and reduce the false negatives.

Deep features + machine learning Corral et al. [196] employed a pre-trained (fast) CNN-F [197] to extract features from MRI images, resulting in formed vectors. These vectors were subsequently transformed using canonical correlation analysis (CCA) and fed into an SVM classifier. The SVM effectively classified the images into three distinct types: healthy pancreas, low-grade IPMN, and high-grade IPMN with PDAC.

Unsupervised methods Semi-supervised, weakly supervised, and unsupervised methods in machine learning and deep learning provide cost-efficient and scalable solutions by leveraging partially labeled or entirely unlabeled data. Hussein et al. [198] tried both supervised and unsupervised learning methods. 3D CNN with multi-task learning was used as a supervised methods. For unsupervised learning, they employed a proportion-SVM to classify IPMNs and normal pancreas. They initially cluster appearance features from images to estimate labels, then compute label proportions for each cluster, and finally use these initial assignments and proportions to learn tumor categorization.

6.3. Segmentation

FCN/UNet-based methods for 2D segmentation Asatryan et al. [199] used a Hausdorff-Sine loss function to address vague organ boundaries in high class-imbalanced data, optimizing boundary delineation using the modified Hausdorff metric and a sinusoidal component in medical segmentation. Chen et al. [132] proposed the Spiral-ResUNet, which incorporated a spiral transformation to enhance segmentation performance. This UNet-based framework leveraged the residual block of ResNet-34 in the encoder module, enabling effective feature extraction.

FCN/UNet-based methods for 3D segmentation Liang et al. [200] involved registering MRIs, pre-processing, patch extraction, classification with a square-window based CNN architecture, and post-processing to obtain a binary map representing tumor probability distribution, from original T1-weighted DCE MRI. Mazor et al. [201] proposed an MC3DU-Net, which utilized TSE MRI scan for pancreas ROI segmentation, transferring it to MRCP scan for cyst detection and segmentation within the ROI, employing 3D UNets trained with Hard Negative Patch Mining to address class imbalance and reduce false positives. Li et al. [202] introduced a registration-free multi-modal and multi-scale adversarial segmentation network (MMSA-Net). This innovative network eliminated the need for registration between different modalities and scales by employing a shared encoder and two separate decoders.

Graph-based methods for medical image segmentation Li et al. [203] put forward an end-to-end unsupervised domain-adaptive (UDA) segmentation method. This approach took advantage of GCN and a meta-learning strategy to address the challenges of adapting to target domains without labeled data. Cai et al. [204] conducted pancreatic detection and boundary segmentation using two CNN models: for tissue localization to differentiate pancreas and non-pancreas tissue based on spatial intensity context, and for boundary determination to delineate the semantic boundaries of the pancreas. The results from both networks are fused to initialize a conditional random field (CRF) framework, yielding the final segmentation output.

6.4. Object Detection

Chen et al. [193] developed a three-stage modified Faster-RCNN approach. Firstly, they employed a pre-trained VGG16 [205] to extract features from the input. These features were then used to identify the ROI. Subsequently, a Z-Continuity Filter (ZCF) was applied to filter the ROIs and improve the accuracy of the detection process.

6.5. Prognosis Prediction

Han et al. [206] applied logistic regression analysis and Cox proportional hazards regression to figure out the risk factors related to recurrence and disease-free survival (DFS) among pNET patients who had previously undergone surgery. They considered various MRI features such as size, location, margin, etc. The analysis revealed that certain MRI features, including portal phase iso-to hypoenhancement, dilatation of the common bile duct or main pancreatic duct, arterial invasion, and larger size, had a significant impact on poor DFS. In another study, Xu et al. [207] extracted MRI features by data-characterization algorithms in patients with PDAC. Then the LASSO algorithm was utilized to calculate risk scores based on MRI features. Then the Cox proportional hazards regression was performed to create a radiomics-based nomogram to predict survival in patients with PDAC that combined radiomics data, clinical data, and TNM information [208].

6.6. Other Tasks

Image reconstruction/ super-resolution Chaika et al. [209] used deep learning-based super-resolution gradient echo imaging to enhance MRI image quality and reduces acquisition time for pancreatic imaging, minimizing artifacts and easily integrating into post-processing workflows without protocol modifications.

Table 6. Summary of AI techniques in MRIs

Task	Reference	Method	Dataset	Study population	Performance
Classification	Cui et al. [192]	LASSO regression	A private dataset	202 scans	AUC = 0.903
Classification	Chen et al. [193]	PCN-Net	Two private datasets	52 and 68 scans	Accuracy = 0.923
Classification	Chen et al. [195]	ResNet-18	A private dataset	115 scans	Accuracy = 0.91, Precision = 0.86, Recall = 0.99, AUC = 0.90, F1-score = 0.92
Classification	Corral et al. [196]	SVM	A private dataset	139 scans	AUC = 0.78
Classification	Hussein et al. [198]	Proportion-SVM	A private dataset	171 scans	Accuracy = 0.8422, Recall = 0.972, Specificity = 0.465
Segmentation	Asatryan et al. [199]	CNN with Hausdorff-Sine loss function	Two private datasets	180 and 120 scans	DSC = 0.841 and 0.857
Segmentation	Chen et al. [132]	Spiral-ResUNet	Four private datasets	65, 69, 68 and 70 scans	DSC = 0.656, 0.640, 0.645, and 0.653
Segmentation	Liang et al. [200]	Square-window-based CNN	A private dataset	56 scans	DSC = 0.73 ± 0.09
Segmentation	Mazor et al. [201]	MC3DU-Net	A private dataset	158 scans	Precision = 0.75, Recall = 0.80, DSC = 0.80
Segmentation	Li et al. [202]	MMSA-Net	Two private datasets	67 and 67 scans	DSC = 0.6452 ± 0.1953 and 0.6560 ± 0.1532
Segmentation	Li et al. [203]	UDA	Four private datasets	67, 68, 68, and 64 scans	DSC = 0.6138, 0.6111, 0.6190, and 0.6007
Segmentation	Cai et al. [204]	CNN with CRF	A private dataset	78 scans	DSC = 0.761
Object Detection	Chen et al. [193]	Modified Faster-RCNN	Two private datasets	52 and 68 scans	Precision = 0.589 and 0.598, Recall = 0.873 and 0.889
Prognosis Prediction	Han et al. [206]	Logistic regression and Cox regression	A private dataset	99 scans	-
Prognosis Prediction	Xu et al. [207]	Cox regression	A private dataset	78 scans	C-index = 0.78

7. Endoscopic Ultrasonography (EUS)

7.1. Introduction to EUS

Endoscopic Ultrasonography (EUS) is a medical procedure that combines endoscopy and ultrasound technology to provide high-resolution imaging and detailed tissue characterization of the gastrointestinal tract and adjacent organs. It allows for the visualization of the digestive system's walls and nearby structures like the liver, gallbladder, and pancreas. EUS features the ability to perform fine needle aspiration (FNA), enabling the collection of tissue samples for analysis. This minimally invasive and well-tolerated procedure is especially effective in staging malignancies and evaluating pancreatic and biliary disorders, making it an invaluable tool for diagnosing and managing various gastrointestinal conditions.

924

925

926

927

928

929

930

931

932

933

934

EUS has demonstrated its superiority in detecting masses compared to CT scans. Studies have shown that EUS exhibits higher sensitivity in mass detection [210]. This improved sensitivity can be attributed to the close-range imaging capability of EUS, allowing for detailed examination and precise localization of abnormalities. Unlike conventional transcutaneous ultrasound examinations, EUS is not limited by pulmonary or bowel gas interference, ensuring accurate visualization and assessment of the pancreas in real-time. It provides high-resolution ultrasound images, enabling clinicians to identify and evaluate pancreatic lesions with exceptional clarity [211].

7.2. Classification

Features extraction + machine learning Ruano et al. [212] focused on identifying interest points and calculating intensity gradients, resulting in 64 features from EUS images, which were used to create a frame feature vector for analysis and classification. To distinguish between PC and non-PC cases, the authors applied SVM and AdaBoost algorithms. Notably, their results outperformed deep learning methods in noisy experiments.

End-to-end deep learning Kuwahara et al. [213] employed ResNet-50 to predict the malignant probability of IPMN, the precursor of PDAC. The accuracy of this approach was higher than the human diagnosis. Zhang et al. [214] proposed a system called BP MASTER, which utilized ResNet in EUS videos to classify pancreas stations. Udrisoiu et al. [215] combined CNN and long short-term memory (LSTM) to classify PDAC, pNET, and chronic pseudotumoral pancreatitis (CPP) in EUS images. Nguon et al. [216] used ResNet-50 in EUS images for MCN and SCN classification. Bonmati et al. [217] developed a CNN composed of two branches for voice data and image data, respectively, used to predict image labels from the spoken names of anatomical landmarks. Vilas et al. [218] applied the Xception model with pre-trained weights to classify Mucinous and Non-Mucinous pancreatic cystic lesions. Jaramillo et al. [219] used GoogleNet, ResNet-18, and ResNet-50 to distinguish PC and non-PC classes. Ren et al. [220] used ResNet-50 with a feature fusion layer to combine with clinical features to classify three types of solid pancreatic tumors in EUS images: PDAC, pNET, and SPN. Kuwahara et al. [221] applied EfficientNetV2-L [222] to categorize various types of pancreatic tumors, including PDAC, pNET, SPN, PASC, ACC, metastatic pancreatic tumor, neuroendocrine carcinoma, chronic pancreatitis, and autoimmune pancreatitis. Fleurentin et al. [223] used different CNNs and ViT models to classify pancreatic anatomical landmarks and explored the effect of LSTM modules to utilize temporal information. Li et al. [58] introduced a Dual Self-supervised Multi-Operator Transformation Network (DSMT-Net), for multi-source EUS diagnosis, which standardized region of interest extraction and employed a transformer-based dual self-supervised network for pre-training representation models using unlabeled EUS images.

Table 7. Summary of AI techniques in EUS images for classification task

Reference	Method	Dataset	Study population	Performance
Ruano et al. [212]	SVM and AdaBoost	A private dataset	55 videos	Accuracy = 0.921, Recall = 0.963, Specificity = 0.878
Kuwahara et al. [213]	ResNet-50	A private dataset	3,970 images	Accuracy = 0.940, Recall = 0.957, Specificity = 0.926
Zhang et al. [214]	ResNet	Two private datasets	21,406 and 768 images	DSC = 0.836 and 0.835
Udriștoiu et al. [215]	Combination of CNN and LSTM	A private dataset	1,350 images	Accuracy = 0.9826, AUC = 0.98
Nguon et al. [216]	ResNet-50	A private dataset	108 images	Accuracy = 0.8275, AUC = 0.88
Bonmati et al. [217]	Multi-modal CNN	A private dataset	3,575 images	Accuracy = 0.76, Precision = 0.74, Recall = 0.74, F1-score = 0.74
Vilas et al. [218]	Xception	A private dataset	5,505 images	Accuracy = 0.985, Specificity = 0.989, Recall = 0.983, AUC = 1.00
Jaramillo et al. [219]	GoogleNet, ResNet-18, and ResNet-50	A private dataset	66,249 images	Accuracy = 0.932, Specificity = 0.950, Recall = 0.877, F1-score = 0.870
Ren et al. [220]	ResNet	A private dataset	12,809 images	Accuracy = 0.9180
Kuwahara et al. [221]	EfficientNetV2-L	A private dataset	22,000 images	Accuracy = 0.91
Fleurentin et al. [223]	CNNs and ViT models	A private dataset	41 videos	Accuracy = 0.668
Li et al. [58]	DSMT-Net	LEPset	11,500 images	Accuracy = 0.877, Precision = 0.842, Recall = 0.801, F1-score = 0.822

7.3. Segmentation

FCN/UNet-based methods for 2D segmentation Zhang et al. proposed a system named BP MASTER (pancreaticobiliary master) [214] that employed a UNet++ to segment pancreatic boundaries and achieved comparable results to experts. Iwasa et al. [224] utilized UNet on contrast-enhanced EUS video images to investigate the influential factors in segmentation. They found that unclear tumor boundary (TB) negatively impacted the concordance rate, while respiratory movement (RM) had no significant effect. Oh et al. [225] employed the Attention U-Net model for automatic pancreatic cyst lesion segmentation and compared results with the Basic U-Net, Residual U-Net, and U-Net++ models. Seo et al. [226] developed DAF-Net (neural network model with deep attention features), which exhibited high accuracy and aided in effective surgical therapy for PC. Ren et al. [220] introduced an Attention UNet with a feature fusion layer for segmenting solid pancreatic tumors, assisting doctors in judging tumor scope and boundaries. Tang et al. [227] designed CH-EUS MASTER based on UNet++ with ResNet-50 as the backbone, a real-time capture and segmentation model for solid pancreatic masses using CH-EUS. The system offered equivalent tumor segmentation capabilities to trainer guidance. Studies showed that segmentation on EUS mainly relies on FCN architectures like UNet and exploration models with attention mechanisms. There have not been many attempts to use new techniques that can be further researched, like Transformer-based and other state-of-the-art methods.

Table 8. Summary of AI techniques in EUS images for segmentation task

Reference	Method	Dataset	Study population	Performance
Zhang et al. [214]	UNet++	Three private datasets	2,115 images, 768 images, and 28 videos	Accuracy = 0.942, 0.824, and 0.862
Iwasa et al. [224]	UNet	A private dataset	100 videos	IoU = 0.77
Oh et al. [225]	Attention U-Net	Two private dataset	57 and 364 scans	DSC = 0.794, IoU = 0.741, Accuracy = 0.983, Specificity = 0.991, Recall = 0.797
Seo et al. [226]	DAF-Net	A private dataset	330 images	DSC = 0.828, IoU = 0.723, AUC = 0.927, Recall = 0.890, Specificity = 0.981, Precision = 0.851
Ren et al. [220]	Attention UNet	A private dataset	1,049 images	DSC = 0.7552, mIOU = 0.6241, Precision = 0.7204, Recall = 0.8003
Tang et al. [227]	UNet++	Two private datasets	4,530 images and 270 videos	DSC = 0.763, Recall = 0.941, Precision = 0.642, Accuracy = 0.842, mIoU = 0.731

7.4. Object Detection

There have been some works of object detection in EUS images or videos. Meyer et al. [228] introduced a real-time framework named the SELSA-TROIA model. It incorporated the sequence level semantics aggregation (SELSA) [229] and the temporal ROI align (TROIA) operator [230]. The SELSA considered the sequence information and aggregated features while the TROIA extracted temporal information. This framework simplified the procedure by identifying anatomical landmarks and addressing the time-consuming nature of mastering EUS. Tian et al. [231] applied YOLOv5m to EUS images and results showed promising real-time outcomes in detecting PC and reducing misdiagnosis. Jaramillo et al. [232] proposed a method to approximate the location of tumoral masses in conventional B-mode Echoendoscopy frames combining a dedicated classifier and an object detection YOLO architecture.

Table 9. Summary of AI techniques in EUS images for object detection task

Reference	Method	Dataset	Study population	Performance
Meyer et al. [228]	SELSA-TROIA	A private dataset	50 videos	mAP@0.5 = 0.5836
Tian et al. [231]	YOLOv5m	A private dataset	1,213 images	AUC = 0.85, Recall = 0.95, Specificity = 0.75
Jaramillo et al. [232]	Combination of a classifier and YOLO	A private dataset	66,249 images	IoU = 0.42, Precision = 0.853

7.5. Other Tasks

Image Generation Grimwood et al. [233] trained a Cycle-Consistent Adversarial Network with unpaired EUS images and CT slices extracted in a manner such that they mimic plausible EUS views, to generate EUS images from the pancreas, aorta, and liver, which can be used as a data augmentation strategy when EUS data is scarce.

8. Positron Emission Tomography (PET)

8.1. Introduction to PET

Positron Emission Tomography (PET) is an advanced nuclear imaging technique that utilizes radionuclides. PET provides insights into the functioning of biological processes by using radiolabeled tracers and quantitative mapping [234]. Fluorine-18 (F-18), Carbon-11 (C-11), Nitrogen-13 (N-13), and Oxygen-15 (O-15) are key positron-emitting radioisotopes employed in PET [235]. These isotopes enable the visualization and analysis of metabolic processes and find wide applications in diagnosing and treating various malignancies. PET has an advantage over conventional imaging techniques like CT and MRI since it can detect abnormal metabolic activity even without visible structural abnormalities in organs. This makes PET a powerful tool for early detection and monitoring of cancers. It is also valuable for post-treatment evaluation in cancer patients undergoing chemotherapy or tumor resection surgeries, assisting in assessing treatment response and the possibility of recurrence [236].

However, a challenge of PET is precisely localizing functional abnormalities within anatomical structures. This limitation has been addressed by integrating PET with high-resolution anatomic imaging modalities to form new images, such as PET-CT and PET-MRI. By merging functional information from PET with detailed anatomical images, clinicians can accurately correlate metabolic activity with specific anatomical locations. This integration has significantly improved the diagnostic accuracy and clinical usefulness of PET in oncology [236–238]. The introduction of PET-CT in the early 2000s marked a significant milestone, enabling comprehensive and multimodal imaging that has greatly influenced the growth of oncology practices [237].

8.2. Classification

Features extraction + machine learning Li et al. [239] presented the hybrid feedback-support vector machine- (HFB-SVM-RF) model, which incorporated 5 different kernels (Linear, MLP, Quadratic, Polynomial, and RBF) and 3 hyperplane separation methods (QP, SMO, LS) to construct a classifier. The features used in the model were extracted through dual threshold principal component analysis (DT-PCA), which combined principal features and non-principal features. Xing et al. [240] employed the XGBoost algorithm to analyze ¹⁸F-FDG PET-CT images for preoperative classification of PDAC into grade 1 and grade 2/3. Initially, the physicians manually segmented the ROIs. Pyradiomics [241] was used to extract radiomics features from the original images and the ROIs. Following this, the XGBoost model was built using the selected features to classify PDAC into grade 1 and grade 2/3. Zhang et al. [242] extracted 251 expert-designed features from 2D and 3D PET/CT images of 111 patients and used RF, Adaboost, SVM with the Gaussian radial basis function kernel function (RBF SVM), and SVM with the linear kernel function (Linear SVM) to differentiate AIP from PDAC.

Features extraction + deep learning Zhang et al. [243] utilized a UNet encoder to extract image features and an RF algorithm to select important clinical features. Subsequently, they proposed a Trusted Multi-view Classification (TMC) algorithm to classify images as either low-grade or high-grade. Specifically, the term "low-grade" encompassed highly, moderately-highly, and moderately differentiated pathologies, whereas the "high-grade" category included undifferentiated, lowly, and moderate-lowly differentiated pathologies. Although clinical features were processed by RF, the image features were all processed in deep nets, and we regard the segmentation stage as a feature extraction process, thus we still categorized them in "features extraction + deep learning".

Deep learning + machine learning Wei et al. [244] combined deep features and radiomics features from PET and CT, which were fed into the RAD_model (the fully connected layers), the DL_model (the VGG11 network) and the MF_model (the fully connected layers), to classify PDAC and AIP.

8.3. Segmentation

Super-pixel extraction Li et al. [239] developed a method called Simple Linear Iterative Clustering (SLIC) with the Gray Interval Mapping (GIM) technique to convert CT scans into pseudo-color images. They then employed a combination of phase and frequency spectrum analysis to detect hypermetabolism areas in PET images.

FCN/UNet-based methods for 2D segmentation Zhang et al. [243] improved the UNet model by incorporating guidance from organ location and applying post-processing techniques such as erosion, expansion, and threshold segmentation (OLP).

FCN/UNet-based methods for 3D segmentation Saganuma et al. [245] used Dense-UNet for multiple organs including pancreas segmentation combining information from PET and CT images. Wang et al. [246] introduced the Multi-modal Fusion and Calibration Networks (MFCNet) for segmenting three-dimensional PET-CT images. Their framework included a Multi-modal Fusion Down-sampling Block (MFDB) with a residual structure that fused features from various modal images. Additionally, they employed a Multi-modal Mutual Calibration Block (MMCB) based on the inception structure, which combined decoding features and pathological features. Shao et al. [247] combined a cross multimodal fusion (CMF) module with a cross-attention mechanism to fuse complementary multimodal features, while a mutual information minimization (MIM) module mitigates redundant high-level modal information and computes the latent loss of PET and CT, enabling effective feature extraction and segmentation of regions of interest from PET/CT images using a semi-supervised framework.

8.4. Object Detection

Wang et al. [248] proposed the Multi-scale adaptive attention feature fusion (MAFF) network for tumor detection in PC using PET-CT imaging, which combined PET and CT strengths to improve accuracy. The network used a feature pyramid module for multi-scale feature extraction, an attention module for feature screening, and an adaptive attention feature fusion network for selecting semantic information.

8.5. Prognosis Prediction

Park et al. [249] used a semi-automatic gradient-based method to determine the volumes of interest (VOIs). They applied LASSO regression to extract clinical and radiomic features from these VOIs. Finally, a 100-layer NN was employed to predict the progression of the disease within two years for patients with pNET.

Table 10. Summary of AI techniques in PET images

Task	Reference	Method	Dataset	Study population	Performance
Classification	Li et al. [239]	HFB-SVM-RF	A private dataset	1,700 images	Accuracy = 0.965, Recall = 0.952, Specificity = 0.975
Classification	Xing et al. [240]	XGBoost	A private dataset	149 scans	AUC = 0.921
Classification	Zhang et al. [242]	RBF SVM and Linear SVM	A private dataset	111 scans	Accuracy = 0.85, Specificity = 0.84, Recall = 0.86, AUC = 0.93
Classification	Zhang et al. [243]	TMC	A private dataset	370 images	Accuracy = 0.75, Recall = 0.77, Specificity = 0.73
Classification	Wei et al. [244]	RAD_model, DL_model, and MF_model	A private dataset	159 scans	Accuracy = 0.901, Specificity = 0.930, Recall = 0.875, AUC = 0.964
Segmentation	Li et al. [239]	SLIC	A private dataset and NIH	1,700 images and 82 scans	DSC = 0.789, IoU = 0.654
Segmentation	Zhang et al. [243]	UNet with OLP	A private dataset	370 images	DSC = 0.89
Segmentation	Suganuma et al. [245]	DenseUNet	A private dataset	48,092 images	DSC = 0.751
Segmentation	Wang et al. [246]	MFCNet	A private dataset	93 images	DSC = 0.7620
Segmentation	Shao et al. [247]	CMF module and MIM strategy	A private dataset	93 scans	DSC = 0.7314, IoU = 0.6056, HD = 6.30
Object Detection	Wang et al. [248]	MAFF	A private dataset	880 images	mAP@0.5 = 0.850
Prognosis Prediction	Park et al. [249]	NN	A private dataset	58 scans	AUC = 0.830

9. Pathological images

9.1. Introduction to Pathological Images

Pathological images (or histopathological images) serve as visual representations of tissue samples observed through a microscope, playing a pivotal role in medical diagnosis, research, and treatment planning. Pathologists rely on these images to detect anomalies, characterize diseases, and provide guidance to clinicians.

The advent of computer-assisted diagnosis (CAD) in the 1990s revolutionized medical imaging and diagnostic radiology, concurrently reducing the workload of pathologists [250]. Digital pathology enables the digitalization and analysis of these images, leading to improved diagnostic accuracy [251]. At the core of digital pathology lies whole slide imaging (WSI), a technology that converts camera-captured static images into a digital format. WSI entails the scanning of slides via a scanner, followed by the analysis of resulting digital files using specialized software. In research, pathological images are invaluable for studying disease mechanisms and developing targeted therapies [252]. By integrating with other clinical data, they enable a comprehensive understanding of diseases and support personalized medicine. Ultimately, these images are indispensable in medical practice and contribute to advancing patient care.

Rapid on-site evaluation (ROSE) is a diagnostic technique that employs fine-needle aspiration (FNA). ROSE assumes critical importance in obtaining samples from deeply seated organs through non-surgical means. Furthermore, it maintains an intrinsic connection with pathology by providing images and facilitating the evaluation of masses [253,254].

9.2. Classification

Features extraction + deep learning Saillard et al. [255] proposed a deep learning-based approach named PACpAInt that accurately identifies tumor cell types and molecular phenotypes from routine histological slides, enabling comprehensive analysis of intratumor heterogeneity on a large scale and providing independent prognostic value.

End-to-end deep learning Chang et al. [256] proposed a Deep learning-based Nucleus Classification (DeepNC) method using CNN to classify cancerous and normal cells at a single-cell level. Le et al. [257] presented the Noisy Label Classification (NLC) method, also known as the NLC model, utilizing patches from WSIs to classify regions as cancerous or non-cancerous. Sehmi et al. [258] used 14 different CNN models with pre-trained models on ImageNet for PC grading in pathological images. Ono et al. [259] utilized CNN to extract features from ROSE (Rapid On-site Evaluation) images and proposed Information-Maximizing Self-Augmented Training (IMSAT) based on these features, resulting in highly accurate cluster analysis. The clustering results revealed distinct differences in features and cell density among different categories. Zhang et al. [260] introduced the Shuffle Instances-based Vision Transformer (SI-ViT) model, which effectively reduced perturbations in ROSE images, leading to significant improvements in performance. Ghoshal et al. [261] presented a Bayesian CNN for automated PC grading from MGG and H&E stained images to estimate uncertainty in model prediction. They analyzed the relationship between the accuracy and uncertainty, and leveraged uncertainty in classification error and reject tradeoff. Kou et al. [262] proposed a hybrid CNN-Transformer model incorporating deformable atrous spatial pyramids (DACTransNet), performing automated and accurate classification of histopathological images of PC.

Table 11. Summary of AI techniques in pathological images for classification task

Reference	Method	Dataset	Study population	Performance
Saillard et al. [255]	PACpAInt	Four private datasets and TCGA	424, 304, 909, 25, and 100 images	AUC = 0.86 (private test set) and 0.81 (TCGA test set)
Chang et al. [256]	DeepNC	A private dataset	60,036,000 images	Accuracy = 0.913, Specificity = 0.928, Precision = 0.926, Recall = 0.899
Le et al. [257]	NLC	TCGA and SEER	190 and 64 WSIs	AUC = 0.860 and 0.944
Sehmi et al. [258]	CNN models	A private dataset	138 images	Accuracy = 0.9561
Ono et al. [259]	CNN with IMSAT	-	-	-
Zhang et al. [260]	SI-ViT	A private dataset	5,088 images	Accuracy = 0.9400, Precision = 0.9198, Recall = 0.9068, F1-score = 91.32
Ghoshal et al. [261]	BCNN	A private dataset	3,201 patches from 138 WSIs	Accuracy = 0.7929, Precision = 0.7935, Recall = 0.7933, F1-score = 0.7915
Kou et al. [262]	DACTransNet	TCGA and three private datasets	1,336 patches from 190 WSIs, 35, 35, and 38 images	Accuracy = 0.9634 (TCGA), 0.8973 (Center A), 0.8714 (Center B), and 0.9113 (Center C)

9.3. Segmentation

FCN/UNet-based methods for 2D segmentation Gao et al. [263] put forward a selected multi-scale attention network (SMA-Net) to accomplish tumor cell segmentation, incorporating the selection unit (SU) module and the multi-scale attention (MA) module,

effectively enhancing feature filtration and information supplementation. Zhang et al. [264] developed a DCNN system based on UNet for rapid on-site cytopathology evaluation (ROSE) to improved the diagnosis efficiency. This system demonstrated exceptional robustness and generalization ability. Liu et al. [265] introduced the multi-level aggregation and global guidance network (MLAGG-Net). Gao et al. [266] devised a multi-task learning framework that adopted the EfficientNet-b0 encoding structure, featuring mobile inverted bottleneck convolution (MBConv) with squeeze-and-excitation (SE) modules to extract image features efficiently. Output utilizes a hierarchical sharing design, with three pathways designed for the main task and two auxiliary tasks, sharing more parameters as task correlation increases. Chen et al. [267] introduced a channel-spatial self-attention module, adaptable for mainstream segmentation networks, enhancing long-range dependency in feature maps and improving segmentation performance in PC pathology image segmentation. In the study of Janssen et al. [268], a single H&E-stained slide of resected PC post-NAT from 64 patients was digitized, manually segmented into the tumor, normal ducts, and remaining epithelium classes, with resulting masks and patches distributed across training, validation, and test sets. Modified U-Nets employing different encoders were trained, achieving the highest mean segmentation accuracy with a DenseNet161 encoder. Yang et al. [269] proposed a selective multi-scale attention (SMA) block for gland segmentation in the pancreas, featuring a selection unit between the encoder and decoder to amplify effective information and suppress redundant information based on a training-derived factor. Fu et al. [270] applied UNet for PDAC segmentation in WSIs.

Table 12. Summary of AI techniques in pathological images for segmentation task

Reference	Method	Dataset	Study population	Performance
Gao et al. [263]	SMANet	A private dataset	165 images	mDSC = 0.769, mIoU = 0.665
Zhang et al. [264]	UNet	A private dataset	5,345 images	F1-score = 0.929
Liu et al. [265]	MLAGG-Net	A private dataset	460 images	DSC = 0.9002, IoU = 0.8207, Accuracy = 0.9439, Recall = 0.9136
Gao et al. [266]	Multi-task learning framework	A private dataset	555,119 images	F1-score = 0.97
Chen et al. [267]	Channel-spatial self-attention module	A private dataset	329 images	DSC = 0.7393, IoU = 0.5942, Accuracy = 0.7526, Precision = 0.8030, Recall = 0.7177
Janssen et al. [268]	Modified UNet	A private dataset	16,572 patches from 64 WSIs	F1-score = 0.86
Yang et al. [269]	SMA block	A private dataset	24 WSIs	DSC = 0.8347, Precision = 0.8649, Recall = 0.8216
Fu et al. [270]	UNet	A private dataset	231 images	DSC = 0.8465

9.4. Other Tasks

Image Super-resolution Li et al. [271] reconstructed high-resolution histology images from low-resolution inputs, employing multi-scale FCN to capture hierarchical features and integrates conditional generative adversarial loss to mitigate blurriness in output images. Tissue microarray (TMA) dataset used in experiments were previously used in published PC studies.

Image Reconstruction Kugler et al. [272] proposed a fully non-rigid image registration method for 3D reconstruction of a whole PC Tumor from Pathology Images with different stains, considering the spatial continuity and smoothness of each constituent part of the

microstructures in the tissue. They further proposed a non-rigid 3D reconstruction method based on smooth and continuous internal tissue assumptions. Landmarks detected via template matching with NCC form trajectories across slices, smoothed during registration, while NCC confidence handles artifacts by rejecting unreliable landmarks [273]. Although these two works were all done on the pancreas of KPC mice, they can also be referred to in research on human beings.

10. Multiple Modalities Analysis

Several studies have employed multiple types of medical images in their AI models. Combining various modalities effectively augments the dataset. This process enables AI models to gain visual information on PC tissues from different modalities, consequently improving accuracy by compensating for individual modality limitations, and creating more robust and discriminative feature representations, just as doctors sometimes need to perform multiple imaging examinations to make a diagnosis. Especially, integrating imaging with pathological data provides a more comprehensive and in-depth understanding at both microscopic and macroscopic levels. Combining imaging with pathological data bridges the information gap between the two, facilitating more accurate and detailed diagnosis and treatment planning. Similarly, combining various MRI modalities enhances accuracy by providing complementary perspectives. Structural imaging offers anatomical details, diffusion-weighted imaging detects tissue changes, and functional MRI reveals brain activity. Analyzing these together improves diagnostic precision and physiological understanding.

10.1. Traditional Machine Learning

In traditional machine learning, using features from multiple modalities of images as input is a common practice, which enhances model performance by combining diverse information to capture richer patterns and relationships in the data, improving predictions or classifications. Koch et al. [274] employed CT and MRI images to classify malignant tissue and predict all-cause mortality. In their study, radiologists initially performed semi-automatic segmentation and feature extraction on CT images using the GrowCut algorithm. Following this, they employed the Cox proportional hazards regression to predict the survival time of patients from the time of imaging until death from any cause. Panda et al. [275] leveraged PET-MRI and CT metrics to predict OS. Principal component analysis was used to extract CT textural features while intra-class correlation, and the Cohen kappa correlation coefficient were used to extract PET-MRI features, then a Cox proportional hazards regression to predict OS using these features.

10.2. Multi-modal Fusion

Methods of multimodal fusion include feature-level fusion and decision-level fusion. Feature-level fusion combines feature vectors from different modalities into a larger feature vector, typically achieved through concatenation or concatenation. Decision-level fusion combines independent decisions or predictions from different modalities, such as through voting or weighted averaging.

Feature-level fusion involves extracting features from each modality, normalizing and aligning them if necessary, and then combining them into a single, fused representation. This fused representation captures information from multiple modalities and can be used as input for subsequent machine-learning tasks. Feature-level fusion often relies on specially designed feature fusion modules. Attention-based fusion methods are typical feature-level fusion techniques, allowing models to dynamically weight the contribution of different modalities or regions within modalities based on their relevance to the task at hand. This selective weighting helps in enhancing the discriminative power of the fused representation while suppressing noise or irrelevant information.

Hussein et al. [276] proposed a CNN-based CAD system for IPMN diagnosis and risk assessment using multi-modal MRI, employing minimum and maximum intensity

projections to mitigate annotation variations and a CNN to extract deep features from T1-weighted and T2-weighted MRI modalities. Finally, canonical correlation analysis (CCA) is utilized for feature-level fusion to derive discriminative canonical correlation features, which are then employed for classification. In the fusion part of Chen et al.'s works [195], they arranged all slices with an ROI into a volume (simply jump the blank slices when testing) with their Z-axis index, then resampled the modality with fewer slices to increase the number of slices. As a result, images of two modalities have the same number for every patient. Zhang et al. [277] developed a multi-modal fusion system Asymmetric Twinning Information Interaction Network (ATIIN) to predict the postoperative survival time of PDAC patients by utilizing both CT images and WSIs. In the ATIIN system, CT images and WSIs were processed by ResNet-101 and ResNet-50, respectively. Subsequently, feature and channel attention techniques were applied before combining the features to obtain the final results. Their study combined the advantages of radiomics and pathomics and improved the cost-benefit ratio of PCs. Chen et al. [278] introduced a model-driven multi-modal deep learning approach, using a spiral transformation algorithm to convert 3D data into 2D images, preserving spatial correlation and edge information. And they introduce prior knowledge for multi-modal fusion, enhancing performance, particularly with small sample sizes.

10.3. Cross-Modality Transfer Learning

Cross-modal transfer learning is widely applied in medical imaging to enhance image recognition, segmentation, and feature extraction tasks by transferring knowledge from one modality to another, improving diagnostic accuracy and reducing the need for labeled data. It exploits correlations and shared information between different modalities, mitigating the challenges of data scarcity and enhancing model generalization and robustness against noise and artifacts in medical images. Yao et al. [279] introduced the Transferred DenseSEMask R-CNN (TDSMask R-CNN) Network to segment pancreatic tumors, incorporating Dense and Squeeze-and-Excitation (SE) blocks to learn complementary features from both PET and MRI images. To overcome the challenge of limited labeled data in PC segmentation, they pre-trained the Dense-SENNet on PET images and then transferred its weights for MRI images.

10.4. Cross-Modality Transfer

Deep learning-based image modality conversion offers the advantage of maximizing data utilization in medical imaging and reducing annotation burdens, especially in scenarios with limited data availability. Training on one modality and converting other modalities to the trained format for inference or converting different modalities into a unified format, streamlines processing workflows can improve model performance by adapting to more suitable modalities. Li et al. [280] generated random intermediate modalities between MRIs and CT to form a larger dataset. Then they improved Res-UNet with meta-learning strategies. This framework could be easily integrated into other segmentation networks and alleviate data scarcity. Cai et al. [281] proposed a generic cross-modality synthesis approach using an end-to-end 2D/3D CNN, where mutually beneficial generators and segmentors collaborate for image synthesis and segmentation tasks. This method synthesizes realistic images without paired training data, maintains consistent anatomical structures, and improves segmentation performance by using synthetic data.

10.5. Multi Modality-tasks Models

Some models are not only designed for a single modality but also aimed to achieve good performance on multiple modal tasks. Cai et al. [282] propose a CNN-RNN model for pancreas segmentation in radiology images, aiming to improve segmentation accuracy by integrating adjacent slice information. The model combines a 2D CNN for initial segmentation with an RNN using CLSTM units for refining segmentation consistency across slices, achieving better performance on both CT and MRI images. Asaturyan et al.

[283] introduced a 2D/3D method for pancreas segmentation in multimodality radiological scans, incorporating a novel post-processing stage to enhance tissue classification through progressive contour analysis. The approach ensures detailed boundary preservation, spatial smoothness, and consistent tissue classification across slices, with potential applicability to other abdominal MRI and CT sequences and broader segmentation tasks.

Table 13. Summary of AI techniques in multiple modalities analysis

Modalities	Task	Reference	Method	Dataset	Study population	Performance
CT and MRI	Prognosis prediction	Koch et al. [274]	Cox regression	A private dataset	143 pairs	AUC = 0.995, C-index = 0.778
PET-MRI and CT	Prognosis prediction	Panda et al. [275]	Cox regression	A private dataset	44 pairs	AUC = 0.87
MRI T1w and MRI T2w	Classification	Hussein et al. [276]	CNN-based CAD system	A private dataset	139 pairs	Accuracy = 0.8280, Specificity = 0.8167, Recall = 0.8355
MRI T1w and MRI T2w	Classification	Chen et al. [195]	PCN-Net	A private dataset	52 (T1w) and 68 scans (T2w)	Accuracy = 0.800
CT and WSI	Prognosis prediction	Zhang et al. [277]	ATIIN	A private dataset	356 pairs	C-index = 0.70
MRI ADC, MRI DWI, and MRI T2w	Classification	Chen et al. [278]	Model-driven multi-modal deep learning approach	A private dataset	64 pairs	Accuracy = 0.736, Specificity = 0.680, Precision = 0.810, Recall = 0.775, AUC = 0.740, F1-score = 0.783
PET and MRI	Segmentation	Yao et al. [279]	TDSMask R-CNN	A private dataset	71 pairs	DSC = 0.7833, Recall = 0.7856, Specificity = 0.9972
CT and MRI	Segmentation	Li et al. [280]	Improved Res-UNet	A private dataset and MSD	163 and 281 scans	DSC = 0.6416 and 0.5753
CT and MRI	Segmentation	Cai et al. [281]	CNN	Two private dataset	82 (CT) and 78 (MRI) scans	DSC = 0.788 (CT) and 0.704 (MRI)
CT and MRI	Segmentation	Cai et al. [282]	CNN-RNN model	NIH and a private MRI dataset	82 (NIH) and 79 (MRI) scans	DSC = 0.833 (CT) and 0.807 (MRI), IoU = 0.718 and 0.682, Precision = 0.845 and 0.843, Recall = 0.828 and 0.783
CT and MRI	Segmentation	Asatryan et al. [283]	Custome 2D/3D method	NIH and two private MRI datasets	82 (NIH), 216, and 132 (MRI) scans	DSC = 0.793 (CT), 0.796 (MRI-A), and 0.816 (MRI-B)

11. Tools, Frameworks, and Software

For intelligent analysis of PC medical imaging, not only methods or algorithms are important, but also the support of tools, frameworks, and software, which will provide great convenience for data annotation, algorithm development, clinical usage, and the integration of medical and engineering research.

11.1. Visualization and Annotation Tools

Medical image annotation is of paramount importance for training deep learning models, particularly due to the specialized expertise required for accurate labeling. Medical images often encompass complex structures and pathologies, necessitating precise annotations crucial for training deep learning models. Moreover, annotations for medical images demand high levels of accuracy to prevent misdiagnosis or erroneous treatment plans. Moreover, 3D image annotation assumes even greater significance. Unlike 2D images, 3D images involve additional spatial dimensions, requiring more comprehensive annotation information.

Therefore, the significance of annotation and visualization platforms cannot be overstated. DicomWorks [284], free software for reading and working on medical images of DICOM format, offers several tools for analysis and annotation. 3D slicer [285] is a free open-source platform for visualization, processing, segmentation, registration, and analysis of medical, biomedical, and other 3D images and meshes, widely used by researchers, clinicians, and developers for its interactive tools and stable platform, remaining compatible with the latest hardware and software advancements. It witnessed continuous development based on a 3D slicer from the joint effort of the community. ITK-SNAP [52], a freely available, open-source software tool designed for segmenting structures within 3D and 4D biomedical images. This versatile application offers semi-automatic segmentation capabilities utilizing active contour methods, alongside manual delineation and intuitive image navigation features. RIL-contour [286] allows using fully automated deep-learning methods, semi-automated methods, and purely manual methods with voxel and/or text annotations. It uses iterative deep learning to accelerate annotation. Aiming to perform efficient semi-auto annotation on 3D medical images, EISeg-Med3D [287,288], a 3D slicer extension, is designed to help users guide a deep learning model to perform segmentation by providing positive and negative points.

User-friendly, efficient, interactive, semi-automatic medical image annotation tools will contribute to high-quality medical image datasets, research in medical deep learning algorithms, and algorithms related to PC diagnosis.

11.2. Platform, software and Packages of Radiomics

The platform, software, and packages for radiomics are essential as they enable the extraction, quantification, and analysis of radiomic features from medical images. They provide the necessary infrastructure, tools, and algorithms for efficient data processing, facilitating research and clinical applications. PyRadiomics [241] is an open-source Python package for extracting radiomics features from medical images, aiming to establish a reference standard for radiomics analysis, providing a tested and maintained platform for reproducible feature extraction. With support for both 2D and 3D analysis, it enables calculations of single values per feature for ROI or generation of feature maps. The Quantitative Image Feature Engine (QIFE) [289] is an open-source, modular system for 3D radiomics feature computation. It integrates seamlessly into existing workflows, focusing on modularity, standards, and parallelism. It offers both MATLAB code and a Docker container for easy deployment, with benchmarking showing significant time savings with parallelization. Researchers can customize components and optimize computational efficiency based on dataset characteristics.

11.3. Framework of Deep Learning Designed for Medical Image Analysis

Due to issues such as data reading formats in medical imaging, convenience is not always guaranteed. Frameworks specifically tailored for deep medical imaging, encompassing data reading, preprocessing, commonly used algorithm models, etc., not only enhance convenience but also facilitate fair comparisons of state-of-the-art methods. DLTk [290] is a toolkit based on TensorFlow developed to enable fast prototyping with a low entry threshold and ensure reproducibility in medical image analysis, containing several popular architectures of networks. However, it has not been updated for a long time. TorchIO [291] is an Open-source Python library for preprocessing, augmentation and sampling of medical images for deep learning, which supports 2D, 3D and 4D images such as X-ray, histopathology, CT, ultrasound and diffusion MRI. MONAI [292] extends PyTorch for medical data, offering specialized AI model architectures, transformations, and utilities to simplify medical AI model development and deployment, which also maintains the simplicity and compositional nature of PyTorch libraries it builds upon. MedicalSeg [287,293], an easy-to-use 3D medical image segmentation framework handling the whole segmentation process including data preprocessing, model training, and model deployment based on PaddlePaddle deep learning framework. It supports many cutting-edge models and corresponding high-precision pre-training models. Although these frameworks provide users with great convenience, the constant emergence of new methods in academia and industry poses challenges for the timely updating of these unified frameworks.

12. Special Topics and Future Directions

12.1. Efficient and Light Model Design

Given the constraints of devices' performance in hospitals, the importance of designing medical imaging models that are efficient and lightweight cannot be overstated. These models are tailored to operate seamlessly within the limitations of hospital hardware, ensuring swift and accurate processing of medical images without taxing computational resources. By prioritizing efficiency and minimizing computational overhead, such designs empower healthcare professionals to swiftly analyze medical images, facilitating timely diagnoses and enhancing patient care. Models and backbones like MobileNet [294] and ShuffleNet [295], as well as real-time detection models like tiny versions in the YOLO series, and lightweight U-Net variants like UNext [296], MALUNet [297] and EGE-UNet [298] are specifically designed with the purpose of addressing computational efficiency. However, in PC analysis, there is still relatively little focus on lightweight design and real-time performance.

12.2. Multi-Modal Tasks

Medical Visual Question Answering (MedVQA) is an AI technology designed to answer questions related to medical images. This technology combines computer vision and NLP, enabling computers to understand medical images and respond to questions about them. PMC-VQA [299], PathVQA [300] and VQA-RAD [301] are all good works as public VQA datasets. Although some public VQA datasets include questions related to the pancreas and PC, they are insufficient to cover the full spectrum of cases. There hasn't been specific Medical VQA research dedicated solely to the pancreas and PC. This is an area ripe for exploration and offers potential for utilizing multimodal large models, through which doctors and researchers can query computers about patient diagnoses, treatment plans, and more about the health of the pancreas, obtaining answers from medical images, helping improve the efficiency and accuracy of medical diagnoses.

The report generation task in clinical images is close to MedVQA, automatically generating textual descriptions or summaries based on the content of medical images, such as X-rays, MRI scans, CT scans, and histopathology slides, which can assist radiologists, pathologists, and other medical professionals in interpreting and documenting findings from medical images efficiently. Related works about the pancreas and PC in images of different modalities are also limited, which can be further explored.

Recently, Large multimodal language models (LLM) have achieved notable success in general domains but face limitations in medical scenarios due to significant differences between medical images and text. Currently, visual-language and multimodal models tailored to specific organs or diseases are also being developed, such as OphGLM [302], a newly developed ophthalmic multimodal model, demonstrating the potential for revolutionizing clinical applications in ophthalmology. Drawing inspiration of the progress in ophthalmology, the development of a dedicated Pancreatic Multimodal Language Model (PMLM) could also be expected.

12.3. Large Model Empowered Solutions

Large models' increased capacity allows them to capture and understand complex patterns and relationships within data more effectively. Moreover, large models tend to have better generalization capabilities, adapting well to new, unseen data and domains. They enable more sophisticated and nuanced representations of information, facilitating more accurate and insightful outputs.

Contrastive Language-Image Pretraining (CLIP) [158] stands as a simple yet potent pretraining paradigm. Thanks to its versatility and interpretability, it demonstrates promising results across a spectrum of tasks. It also has gained increasing attention and achieved wide application in the field of medical image analysis, serving as a pre-training paradigm for image-text alignment, or a component in different clinical tasks [303] including zero-shot classification [304], object detection [305], 2D image segmentation [306] and 3D image segmentation [156,157], as well as some cross-modality tasks [307]. As CLIP continues to evolve and adapt to the specific challenges posed by medical image analysis, its integration into clinical practice in PC diagnosis and treatment is expected to grow exponentially.

As we mentioned before, the Segment Anything Model (SAM) [308] has been trained on millions of images and over a billion masks, enabling it to produce effective segmentation masks for any input, archiving impressive zero-shot performance. And experiments show that it can also be a valuable tool in medical image segmentation if used correctly [160]. It has been witnessed the continuous effort to make SAM adapt to medical images or to train a new zero-shot medical image segmentation. Zhang et al. [309] proposed SAMed, applying the low-rank-based (LoRA) finetuning strategy to the SAM image encoder and finetunes it together with the prompt encoder and the mask decoder on labeled medical image segmentation datasets. Wu et al. [310] proposed the Medical SAM Adapter (Med-SA), incorporating domain-specific medical knowledge into the segmentation model, using Space-Depth Transpose (SD-Trans) to adapt 2D SAM to 3D medical images and Hyper-Prompting Adapter (HyP-Adpt) to achieve prompt-conditioned adaptation. Ye et al. [311] introduced SA-Med2D-20M, a large-scale segmentation dataset of 2D medical images built upon numerous public and private datasets, which consists of 4.6 million 2D medical images and 19.7 million corresponding masks, covering almost the whole body and showing significant diversity to incorporate medical knowledge into SAM. Although there have been some efforts in this regard, achieving zero-shot segmentation of the pancreas and PC remains challenging due to the limited availability of datasets containing diverse modalities of pancreatic and PC data. Larger and more diverse datasets of PCs are expected, which will facilitate the ability of large models to better address PC-related challenges.

12.4. Explanability

Deep learning models often present themselves as black boxes and medical experts have expressed their concern about such nature [312]. Particularly in the realm of medical image analysis, healthcare professionals and researchers need to comprehend the model's decision-making process and outcomes to ensure the reliability of the diagnosis and the treatment decisions. In explainability research, most current works utilize post hoc explanation methods instead of model-based explanations, providing explanations on trained neural networks rather than incorporating them during training, predominantly employing local explanations rather than global ones, particularly suited for deep learning in medi-

cal image analysis. Moving forward, the adoption of holistic approaches, integration of biological explanations, and exploration of the link between causality and Explainable AI will become increasingly important [313]. Therefore, exploring the explainability of deep learning models in the pancreas and PC research could be a promising direction for future investigation.

13. Conclusions

This study summarized applications of AI on five modalities and integrated modalities of medical images related to the pancreas and PC. AI models demonstrate reasonable results in segmentation, classification, object detection, prognosis prediction and other tasks in the experiment stage and have comparable performance to human experts in many studies, which highlights the potential of AI to assist doctors and alleviate their workload in the diagnosis and treatment of PCs. However, the overall accuracy of pancreatic analysis lags behind that of other organs, largely because of the size and variable characteristics of the pancreas. In addition, AI studies utilizing MRI, pathology images, and PET imaging for target detection and prognosis prediction are still limited. The lack of comprehensive medical image datasets and further modeling studies challenges the widespread use of AI techniques in clinical settings. Despite these challenges, lightweight model design, multi-modal tasks, large model-empowered solutions, and explainability are future directions that will enhance the efficiency and reliability of AI-based analysis. As AI evolves, its increasing importance in supporting doctors in diagnosis and treatment will ultimately benefit society.

Author Contributions: Conceptualization, W.L. and J.J.; resources, T.L. and Y.L.; writing—original draft preparation, W.L., B.Z. and J.J.; writing—review and editing, W.L., B.Z., T.L., J.J. and Y.L.; visualization, W.L. and J.J.; supervision, T.L., J.J. and Y.L.; project administration, J.J. and Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available in The Cancer Genome Atlas at <https://www.cancer.gov/ccg/research/genome-sequencing/tcga>, reference number [55]; Medical Segmentation Decathlon at <http://medicaldecathlon.com/>, reference number [314]; Surveillance, Epidemiology, and End Results Program at <https://seer.cancer.gov/>, reference number [56]; GitHub at <https://github.com>, reference number [315]; Grand Challenge at <https://grand-challenge.org/>, reference number [316]; Synapse at <https://www.synapse.org/>, reference number [317]; Zenodo at <https://zenodo.org/>, reference number [318]. These data were derived from the following resources available in the public domain: <https://academictorrents.com/details/80ecfefcabede760cdbdf63e38986501f7becd49>; <https://github.com/JunMa11/AbdomenCT-1K>; <https://www.synapse.org/Synapse:syn3193805/wiki/89480>; <https://github.com/HiLab-git/WORD>; <https://drive.google.com/drive/folders/1HqEgzS8BV2c7xYNrZdEAnrHk7osJJ--2>; <https://zenodo.org/records/4621057>; <https://panorama.grand-challenge.org/>; <https://doi.org/10.5281/zenodo.8041285>; <https://2023paip.grand-challenge.org/>; <https://zenodo.org/records/3712669>.

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

References

1. Kamisawa, T.; Wood, L.D.; Itoi, T.; Takaori, K. Pancreatic cancer. *The Lancet* **2016**, *388*, 73–85.
2. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2023. *CA: a cancer journal for clinicians* **2023**, *73*, 17–48.
3. Udare, A.; Agarwal, M.; Alabousi, M.; McInnes, M.; Rubino, J.G.; Marcaccio, M.; van der Pol, C.B. Diagnostic Accuracy of MRI for Differentiation of Benign and Malignant Pancreatic Cystic Lesions Compared to CT and Endoscopic Ultrasound: Systematic Review and Meta-analysis. *Journal of Magnetic Resonance Imaging* **2021**, *54*, 1126–1137.
4. Edward Coleman, R. Single photon emission computed tomography and positron emission tomography in cancer imaging. *Cancer* **1991**, *67*, 1261–1270.
5. Goyal, H.; Sherazi, S.A.A.; Gupta, S.; Perisetti, A.; Achebe, I.; Ali, A.; Tharian, B.; Thosani, N.; Sharma, N.R. Application of artificial intelligence in diagnosis of pancreatic malignancies by endoscopic ultrasound: a systemic review. *Therapeutic Advances in Gastroenterology* **2022**, *15*, 17562848221093873.
6. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *nature* **2015**, *521*, 436–444.

7. Huang, B.; Huang, H.; Zhang, S.; Zhang, D.; Shi, Q.; Liu, J.; Guo, J. Artificial intelligence in pancreatic cancer. *Theranostics* **2022**, *12*, 6931. [1480](#)
8. Aier, I.; Semwal, R.; Sharma, A.; Varadwaj, P.K. A systematic assessment of statistics, risk factors, and underlying features involved in pancreatic cancer. *Cancer epidemiology* **2019**, *58*, 104–110. [1481](#)
9. Klein, A.P. Pancreatic cancer epidemiology: understanding the role of lifestyle and inherited risk factors. *Nature reviews Gastroenterology & hepatology* **2021**, *18*, 493–502. [1482](#)
10. Poddighe, D. Autoimmune pancreatitis and pancreatic cancer: Epidemiological aspects and immunological considerations. *World J Gastroenterol* **2021**, *27*, 3825–3836. <https://doi.org/10.3748/wjg.v27.i25.3825>. [1483](#)
11. Distler, M.; Aust, D.; Weitz, J.; Pilarsky, C.; Grützmann, R.; et al. Precursor lesions for sporadic pancreatic cancer: PanIN, IPMN, and MCN. *BioMed research international* **2014**, *2014*. [1484](#)
12. Rawla, P.; Sunkara, T.; Gaduputi, V. Epidemiology of pancreatic cancer: global trends, etiology and risk factors. *World journal of oncology* **2019**, *10*, 10–27. [1485](#)
13. Hidalgo, M.; Cascinu, S.; Kleeff, J.; Labianca, R.; Löhr, J.M.; Neoptolemos, J.; Real, F.X.; Van Laethem, J.L.; Heinemann, V. Addressing the challenges of pancreatic cancer: future directions for improving outcomes. *Pancreatology* **2015**, *15*, 8–18. [1486](#)
14. Vassos, N.; Agaimy, A.; Klein, P.; Hohenberger, W.; Croner, R.S. Solid-pseudopapillary neoplasm (SPN) of the pancreas: case series and literature review on an enigmatic entity. *International Journal of Clinical and Experimental Pathology* **2013**, *6*, 1051. [1487](#)
15. Fang, Y.; Su, Z.; Xie, J.; Xue, R.; Ma, Q.; Li, Y.; Zhao, Y.; Song, Z.; Lu, X.; Li, H.; et al. Genomic signatures of pancreatic adenosquamous carcinoma (PASC). *The Journal of Pathology* **2017**, *243*, 155–159. [1488](#)
16. Kitagami, H.; Kondo, S.; Hirano, S.; Kawakami, H.; Egawa, S.; Tanaka, M. Acinar cell carcinoma of the pancreas: clinical analysis of 115 patients from Pancreatic Cancer Registry of Japan Pancreas Society. *Pancreas* **2007**, *35*, 42–46. [1489](#)
17. Reid, M.D.; Choi, H.; Balci, S.; Akkas, G.; Adsay, V. Serous cystic neoplasms of the pancreas: clinicopathologic and molecular characteristics. In Proceedings of the Seminars in diagnostic pathology. Elsevier, 2014, Vol. 31, pp. 475–483. [1490](#)
18. Bochis, O.; Bota, M.; Mihut, E.; Buiga, R.; Hazbei, D.; Irimie, A. Solid pseudopapillary tumor of the pancreas: clinical-pathological features and management of 13 cases. *Clujul Med* **2017**, *90*, 171–178. Epub 2017 Apr 25, <https://doi.org/10.15386/cjmed-672>. [1491](#)
19. Backx, E.; Coolens, K.; Van den Bossche, J.L.; Houbraken, I.; Espinet, E.; Rooman, I. On the origin of pancreatic cancer: molecular tumor subtypes in perspective of exocrine cell plasticity. *Cellular and molecular gastroenterology and hepatology* **2022**, *13*, 1243–1253. [1492](#)
20. Mizrahi, J.D.; Surana, R.; Valle, J.W.; Shroff, R.T. Pancreatic cancer. *The Lancet* **2020**, *395*, 2008–2020. [1493](#)
21. Artinyan, A.; Soriano, P.A.; Prendergast, C.; Low, T.; Ellenhorn, J.D.; Kim, J. The anatomic location of pancreatic cancer is a prognostic factor for survival. *Hpb* **2008**, *10*, 371–376. [1494](#)
22. Mostafa, M.E.; Erbarut-Seven, I.; Pehlivanoglu, B.; Adsay, V. Pathologic classification of "pancreatic cancers": current concepts and challenges. *Chin Clin Oncol* **2017**, *6*, 59. [1495](#)
23. Raphael, B.J.; Hruban, R.H.; Aguirre, A.J.; Moffitt, R.A.; Yeh, J.J.; Stewart, C.; Robertson, A.G.; Cherniack, A.D.; Gupta, M.; Getz, G.; et al. Integrated genomic characterization of pancreatic ductal adenocarcinoma. *Cancer cell* **2017**, *32*, 185–203. [1496](#)
24. Espinet, E.; Klein, L.; Puré, E.; Singh, S.K. Mechanisms of PDAC subtype heterogeneity and therapy response. *Trends in Cancer* **2022**. [1497](#)
25. Flowers, B.M.; Xu, H.; Mulligan, A.S.; Hanson, K.J.; Seoane, J.A.; Vogel, H.; Curtis, C.; Wood, L.D.; Attardi, L.D. Cell of origin influences pancreatic cancer subtype. *Cancer discovery* **2021**, *11*, 660–677. [1498](#)
26. Guo, W.; Zhang, Y.; Guo, S.; Mei, Z.; Liao, H.; Dong, H.; Wu, K.; Ye, H.; Zhang, Y.; Zhu, Y.; et al. Tumor microbiome contributes to an aggressive phenotype in the basal-like subtype of pancreatic cancer. *Communications Biology* **2021**, *4*, 1019. [1499](#)
27. Halfdanarson, T.R.; Rabe, K.; Rubin, J.; Petersen, G. Pancreatic neuroendocrine tumors (PNETs): incidence, prognosis and recent trend toward improved survival. *Annals of oncology* **2008**, *19*, 1727–1733. [1500](#)
28. Ellison, T.A.; Wolfgang, C.L.; Shi, C.; Cameron, J.L.; Murakami, P.; Mun, L.J.; Singhi, A.D.; Cornish, T.C.; Olino, K.; Meriden, Z.; et al. A single institution's 26-year experience with nonfunctional pancreatic neuroendocrine tumors: a validation of current staging systems and a new prognostic nomogram. *Annals of surgery* **2014**, *259*, 204. [1501](#)
29. Mpilla, G.B.; Philip, P.A.; El-Rayes, B.; Azmi, A.S. Pancreatic neuroendocrine tumors: therapeutic challenges and research limitations. *World Journal of Gastroenterology* **2020**, *26*, 4036. [1502](#)
30. Perri, G.; Prakash, L.R.; Katz, M.H. Pancreatic neuroendocrine tumors. *Current opinion in gastroenterology* **2019**, *35*, 468–477. [1503](#)
31. Pea, A.; Hruban, R.H.; Wood, L.D. Genetics of pancreatic neuroendocrine tumors: implications for the clinic. *Expert review of gastroenterology & hepatology* **2015**, *9*, 1407–1419. [1504](#)
32. Luo, S.; Wang, J.; Wu, L.; Wang, C.; Yang, J.; Li, M.; Zhang, L.; Ge, J.; Sun, C.; Li, E.; et al. Epidemiological trends for functional pancreatic neuroendocrine tumors: A study combining multiple imputation with age adjustment. *Frontiers in Endocrinology* **2023**, *14*, 1123642. [1505](#)
33. Nieven van Dijkum, E.J.; Engelsman, A.F. Diagnosis and Management of Functional Pancreatic Neuroendocrine Tumors. In *Endocrine Surgery Comprehensive Board Exam Guide*; Springer, 2022; pp. 681–693. [1506](#)
34. Tsilimigras, D.; Pawlik, T. Pancreatic neuroendocrine tumours: conservative versus surgical management. *British Journal of Surgery* **2021**, *108*, 1267–1269. [1507](#)
35. Kuo, J.H.; Lee, J.A.; Chabot, J.A. Nonfunctional pancreatic neuroendocrine tumors. *Surgical Clinics* **2014**, *94*, 689–708. [1508](#)

36. Dong, D.H.; Zhang, X.F.; Lopez-Aguiar, A.G.; Poultides, G.; Makris, E.; Rocha, F.; Kanji, Z.; Weber, S.; Fisher, A.; Fields, R.; et al. Tumor burden score predicts tumor recurrence of non-functional pancreatic neuroendocrine tumors after curative resection. *HPB* **2020**, *22*, 1149–1157. [1538](#)
37. Zerbini, A.; Falconi, M.; Rindi, G.; Delle Fave, G.; Tomassetti, P.; Pasquali, C.; Capitanio, V.; Boninsegna, L.; Di Carlo, V.; members of the AISPN-Netwok Study Group.; et al. Clinicopathological features of pancreatic endocrine tumors: a prospective multicenter study in Italy of 297 sporadic cases. *Official journal of the American College of Gastroenterology | ACG* **2010**, *105*, 1421–1429. [1542](#)
38. Nigri, G.; Petrucciani, N.; Debs, T.; Mangogna, L.M.; Crovetto, A.; Moschetta, G.; Persechino, R.; Aurello, P.; Ramacciato, G. Treatment options for PNET liver metastases: a systematic review. *World journal of surgical oncology* **2018**, *16*, 1–10. [1544](#)
39. Srivastava, S.; Koay, E.J.; Borowsky, A.D.; De Marzo, A.M.; Ghosh, S.; Wagner, P.D.; Kramer, B.S. Cancer overdiagnosis: a biological challenge and clinical dilemma. *Nature Reviews Cancer* **2019**, *19*, 349–358. [1545](#)
40. Pereira, S.P.; Oldfield, L.; Ney, A.; Hart, P.A.; Keane, M.G.; Pandol, S.J.; Li, D.; Greenhalf, W.; Jeon, C.Y.; Koay, E.J.; et al. Early detection of pancreatic cancer. *The lancet Gastroenterology & hepatology* **2020**, *5*, 698–710. [1547](#)
41. Macdonald, S.; Macleod, U.; Campbell, N.C.; Weller, D.; Mitchell, E. Systematic review of factors influencing patient and practitioner delay in diagnosis of upper gastrointestinal cancer. *British journal of cancer* **2006**, *94*, 1272–1280. [1549](#)
42. Zhang, L.; Sanagapalli, S.; Stoita, A. Challenges in diagnosis of pancreatic cancer. *World journal of gastroenterology* **2018**, *24*, 2047. [1551](#)
43. Walter, F.M.; Mills, K.; Mendonça, S.C.; Abel, G.A.; Basu, B.; Carroll, N.; Ballard, S.; Lancaster, J.; Hamilton, W.; Rubin, G.P.; et al. Symptoms and patient factors associated with diagnostic intervals for pancreatic cancer (SYMPTOM pancreatic study): a prospective cohort study. *The lancet Gastroenterology & hepatology* **2016**, *1*, 298–306. [1553](#)
44. Jiang, S.; Fagman, J.B.; Ma, Y.; Liu, J.; Vihav, C.; Engstrom, C.; Liu, B.; Chen, C. A comprehensive review of pancreatic cancer and its therapeutic challenges. *Aging (Albany NY)* **2022**, *14*, 7635. [1555](#)
45. Halbrook, C.J.; Lyssiotis, C.A.; di Magliano, M.P.; Maitra, A. Pancreatic cancer: Advances and challenges. *Cell* **2023**, *186*, 1729–1754. [1557](#)
46. Wood, L.D.; Canto, M.I.; Jaffee, E.M.; Simeone, D.M. Pancreatic cancer: pathogenesis, screening, diagnosis, and treatment. *Gastroenterology* **2022**, *163*, 386–402. [1559](#)
47. Roth, H.R.; Lu, L.; Farag, A.; Shin, H.C.; Liu, J.; Turkbey, E.B.; Summers, R.M. Deeporgan: Multi-level deep convolutional networks for automated pancreas segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part I 18. Springer, 2015, pp. 556–564. [1560](#)
48. Ma, J.; Zhang, Y.; Gu, S.; Zhu, C.; Ge, C.; Zhang, Y.; An, X.; Wang, C.; Wang, Q.; Liu, X.; et al. Abdomenct-1k: Is abdominal organ segmentation a solved problem? *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2021**, *44*, 6695–6714. [1565](#)
49. Landman, B.; Xu, Z.; Iglesias, J.; Styner, M.; Langerak, T.; Klein, A. Miccai multi-atlas labeling beyond the cranial vault—workshop and challenge. In Proceedings of the Proc. MICCAI Multi-Atlas Labeling Beyond Cranial Vault—Workshop Challenge, 2015, Vol. 5, p. 12. [1568](#)
50. Luo, X.; Liao, W.; Xiao, J.; Chen, J.; Song, T.; Zhang, X.; Li, K.; Metaxas, D.N.; Wang, G.; Zhang, S. WORD: A large scale dataset, benchmark and clinical applicable study for abdominal organ segmentation from CT image. *arXiv preprint arXiv:2111.02403* **2021**. [1569](#)
51. Simpson, A.L.; Antonelli, M.; Bakas, S.; Bilello, M.; Farahani, K.; Van Ginneken, B.; Kopp-Schneider, A.; Landman, B.A.; Litjens, G.; Menze, B.; et al. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. *arXiv preprint arXiv:1902.09063* **2019**. [1571](#)
52. Yushkevich, P.A.; Gao, Y.; Gerig, G. ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. In Proceedings of the 2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE, 2016, pp. 3342–3345. [1574](#)
53. Abel, L.; Wasserthal, J.; Weikert, T.; Sauter, A.W.; Nesic, I.; Obradovic, M.; Yang, S.; Manneck, S.; Glessgen, C.; Ospel, J.M.; et al. Automated Detection of Pancreatic Cystic Lesions on CT Using Deep Learning. *Diagnostics* **2021**, *11*. <https://doi.org/10.3390/diagnostics11050901>. [1578](#)
54. Isensee, F.; Jaeger, P.F.; Kohl, S.A.; Petersen, J.; Maier-Hein, K.H. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods* **2021**, *18*, 203–211. [1581](#)
55. Network, T.C.G.A.R. The Cancer Genome Atlas. *Nature* **2014**, *517*, 547–555. <https://doi.org/10.1038/nature13385>. [1582](#)
56. Races, A.; Males, M.W.M.B. SEER Cancer Statistics Review 1975–2017 **2020**. [1583](#)
57. PANORAMA **2024**. [1584](#)
58. Li, J.; Zhang, P.; Wang, T.; Zhu, L.; Liu, R.; Yang, X.; Wang, K.; Shen, D.; Sheng, B. DSMT-Net: Dual Self-supervised Multi-operator Transformation for Multi-source Endoscopic Ultrasound Diagnosis. *IEEE Transactions on Medical Imaging* **2023**. [1585](#)
59. PAIP2023. <https://2023paip.grand-challenge.org/> **2023**. [1587](#)
60. Grizzi, F.; Fiorino, S.; Qehajaj, D.; et al. Computer-aided assessment of the extra-cellular matrix during pancreatic carcinogenesis: a pilot study. *Journal of Translational Medicine* **2019**, *17*, 61. <https://doi.org/10.1186/s12967-019-1817-3>. [1589](#)
61. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Communications of the ACM* **2017**, *60*, 84–90. [1590](#)
62. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* **2014**. [1592](#)
63. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778. [1594](#)

64. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929* **2020**. [1597](#)
65. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the Proceedings of the IEEE/CVF international conference on computer vision, 2021, pp. 10012–10022. [1598](#)
[1600](#)
66. Altman, D.G.; Bland, J.M. Diagnostic tests. 1: Sensitivity and specificity. *BMJ* **1994**, *308*, 1552. <https://doi.org/10.1136/bmj.308.6943.1552>. [1601](#)
[1602](#)
67. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431–3440. [1603](#)
[1604](#)
68. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18. Springer, 2015, pp. 234–241. [1605](#)
[1606](#)
[1607](#)
69. Oktay, O.; Schlemper, J.; Folgoc, L.L.; Lee, M.; Heinrich, M.; Misawa, K.; Mori, K.; McDonagh, S.; Hammerla, N.Y.; Kainz, B.; et al. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999* **2018**. [1608](#)
[1609](#)
70. Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. Unet++: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE transactions on medical imaging* **2019**, *39*, 1856–1867. [1610](#)
[1611](#)
71. Jha, D.; Smedsrød, P.H.; Riegler, M.A.; Johansen, D.; De Lange, T.; Halvorsen, P.; Johansen, H.D. Resunet++: An advanced architecture for medical image segmentation. In Proceedings of the 2019 IEEE international symposium on multimedia (ISM). IEEE, 2019, pp. 225–2255. [1612](#)
[1613](#)
[1614](#)
72. Chen, Y.; Wang, K.; Liao, X.; Qian, Y.; Wang, Q.; Yuan, Z.; Heng, P.A. Channel-Unet: a spatial channel-wise convolutional neural network for liver and tumors segmentation. *Frontiers in genetics* **2019**, *10*, 1110. [1615](#)
[1616](#)
73. Huang, H.; Lin, L.; Tong, R.; Hu, H.; Zhang, Q.; Iwamoto, Y.; Han, X.; Chen, Y.W.; Wu, J. Unet 3+: A full-scale connected unet for medical image segmentation. In Proceedings of the ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, 2020, pp. 1055–1059. [1617](#)
[1618](#)
[1619](#)
74. Çiçek, Ö.; Abdulkadir, A.; Lienkamp, S.S.; Brox, T.; Ronneberger, O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016: 19th International Conference, Athens, Greece, October 17–21, 2016, Proceedings, Part II 19. Springer, 2016, pp. 424–432. [1620](#)
[1621](#)
[1622](#)
75. Chen, W.; Liu, B.; Peng, S.; Sun, J.; Qiao, X. S3D-UNet: separable 3D U-Net for brain tumor segmentation. In Proceedings of the Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part II 4. Springer, 2019, pp. 358–368. [1623](#)
[1624](#)
[1625](#)
[1626](#)
76. Abdollahi, A.; Pradhan, B.; Alamri, A. VNet: An end-to-end fully convolutional neural network for road extraction from high-resolution remote sensing data. *Ieee Access* **2020**, *8*, 179424–179436. [1627](#)
[1628](#)
77. Chen, J.; Lu, Y.; Yu, Q.; Luo, X.; Adeli, E.; Wang, Y.; Lu, L.; Yuille, A.L.; Zhou, Y. Transunet: Transformers make strong encoders for medical image segmentation. *arXiv preprint arXiv:2102.04306* **2021**. [1629](#)
[1630](#)
78. Cao, H.; Wang, Y.; Chen, J.; Jiang, D.; Zhang, X.; Tian, Q.; Wang, M. Swin-unet: Unet-like pure transformer for medical image segmentation. In Proceedings of the European conference on computer vision. Springer, 2022, pp. 205–218. [1631](#)
[1632](#)
79. Sha, Y.; Zhang, Y.; Ji, X.; Hu, L. Transformer-unet: Raw image processing with unet. *arXiv preprint arXiv:2109.08417* **2021**. [1633](#)
80. Chen, B.; Liu, Y.; Zhang, Z.; Lu, G.; Kong, A.W.K. Transattunet: Multi-level attention-guided u-net with transformer for medical image segmentation. *IEEE Transactions on Emerging Topics in Computational Intelligence* **2023**. [1634](#)
[1635](#)
81. Kipf, T.N.; Welling, M. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* **2016**. [1636](#)
82. Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y.; et al. Graph attention networks. *stat* **2017**, *1050*, 10–48550. [1637](#)
83. Xu, K.; Hu, W.; Leskovec, J.; Jegelka, S. How powerful are graph neural networks? *arXiv preprint arXiv:1810.00826* **2018**. [1638](#)
84. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems* **2015**, *28*. [1639](#)
[1640](#)
85. Cai, Z.; Vasconcelos, N. Cascade R-CNN: High Quality Object Detection and Instance Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2019**, p. 1–1. <https://doi.org/10.1109/tpami.2019.2956516>. [1641](#)
[1642](#)
86. Wang, X.; Kong, T.; Shen, C.; Jiang, Y.; Li, L. SOLO: Segmenting Objects by Locations. In Proceedings of the Proc. Eur. Conf. Computer Vision (ECCV), 2020. [1643](#)
[1644](#)
87. Wang, X.; Zhang, R.; Kong, T.; Li, L.; Shen, C. SOLOv2: Dynamic and Fast Instance Segmentation. *Proc. Advances in Neural Information Processing Systems (NeurIPS)* **2020**. [1645](#)
[1646](#)
88. Fang, Y.; Yang, S.; Wang, X.; Li, Y.; Fang, C.; Shan, Y.; Feng, B.; Liu, W. Instances As Queries. In Proceedings of the Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 6910–6919. [1647](#)
[1648](#)
89. Huttenlocher, D.P.; Klanderman, G.A.; Ruckridge, W.J. Comparing images using the Hausdorff distance. *IEEE Transactions on pattern analysis and machine intelligence* **1993**, *15*, 850–863. [1649](#)
[1650](#)
90. Nikolov, S.; Blackwell, S.; Zverovitch, A.; Mendes, R.; Livne, M.; De Fauw, J.; Patel, Y.; Meyer, C.; Askham, H.; Romera-Paredes, B.; et al. Clinically applicable segmentation of head and neck anatomy for radiotherapy: deep learning algorithm development and validation study. *Journal of medical Internet research* **2021**, *23*, e26151. [1651](#)
[1652](#)
[1653](#)

91. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer, 2016, pp. 21–37. 1654
1655
1656
92. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 779–788. 1658
1659
93. Redmon, J.; Farhadi, A. YOLO9000: better, faster, stronger. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 7263–7271. 1660
1661
94. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv preprint arXiv:1804.02767* **2018**. 1661
95. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. YOLOv4: Optimal Speed and Accuracy of Object Detection, 2020, [[arXiv:cs.CV/2004.10934](https://arxiv.org/abs/cs.CV/2004.10934)]. 1662
96. Jocher, G. Ultralytics YOLOv5, 2020. <https://doi.org/10.5281/zenodo.3908559>. 1663
97. Li, C.; Li, L.; Geng, Y.; Jiang, H.; Cheng, M.; Zhang, B.; Ke, Z.; Xu, X.; Chu, X. YOLOv6 v3.0: A Full-Scale Reloading, 2023, [[arXiv:cs.CV/2301.05586](https://arxiv.org/abs/cs.CV/2301.05586)]. 1664
1665
98. Wang, C.Y.; Bochkovskiy, A.; Liao, H.Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv preprint arXiv:2207.02696* **2022**. 1666
1667
99. Jocher, G.; Chaurasia, A.; Qiu, J. Ultralytics YOLOv8, 2023. 1668
100. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 580–587. 1669
1670
101. Girshick, R. Fast r-cnn. In Proceedings of the Proceedings of the IEEE international conference on computer vision, 2015, pp. 1440–1448. 1671
1672
102. Carion, N.; Massa, F.; Synnaeve, G.; Usunier, N.; Kirillov, A.; Zagoruyko, S. End-to-End Object Detection with Transformers. In Proceedings of the ECCV, 2020. 1673
1674
103. Kern, D.; Mastmeyer, A. 3D bounding box detection in volumetric medical image data: A systematic literature review. In Proceedings of the 2021 IEEE 8th International Conference on Industrial Engineering and Applications (ICIEA). IEEE, 2021, pp. 509–516. 1675
1676
1677
104. De Vos, B.D.; Wolterink, J.M.; De Jong, P.A.; Viergever, M.A.; Işgum, I. 2D image classification for 3D anatomy localization: employing deep convolutional neural networks. In Proceedings of the Medical imaging 2016: Image processing. SPIE, 2016, Vol. 9784, pp. 517–523. 1678
1679
1680
105. Huang, R.; Xie, W.; Noble, J.A. VP-Nets: Efficient automatic localization of key brain structures in 3D fetal neurosonography. *Medical image analysis* **2018**, *47*, 127–139. 1681
1682
106. Blair, S.I.A.S.A.; White, C.; Moses, L.D.D. Localization of lumbar and thoracic vertebrae in 3d ct datasets by combining deep reinforcement learning with imitation learning **2018**. 1683
1684
107. Xu, X.; Zhou, F.; Liu, B.; Fu, D.; Bai, X. Efficient multiple organ localization in CT image using 3D region proposal network. *IEEE transactions on medical imaging* **2019**, *38*, 1885–1898. 1685
1686
108. Hasebroock, K.M.; Serkova, N.J. Toxicity of MRI and CT contrast agents. *Expert opinion on drug metabolism & toxicology* **2009**, *5*, 403–416. 1687
1688
109. Li, M.; Nie, X.; Reheman, Y.; Huang, P.; Zhang, S.; Yuan, Y.; Chen, C.; Yan, Z.; Chen, C.; Lv, X.; et al. Computer-aided diagnosis and staging of pancreatic cancer based on CT images. *IEEE Access* **2020**, *8*, 141705–141718. 1689
1690
110. Chen, P.T.; Chang, D.; Yen, H.; Liu, K.L.; Huang, S.Y.; Roth, H.; Wu, M.S.; Liao, W.C.; Wang, W. Radiomic features at CT can distinguish pancreatic cancer from noncancerous pancreas. *Radiology: Imaging Cancer* **2021**, *3*, e210010. 1691
1692
111. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, pp. 785–794. 1693
1694
112. Mukherjee, S.; Patra, A.; Khasawneh, H.; Korfiatis, P.; Rajamohan, N.; Suman, G.; Majumder, S.; Panda, A.; Johnson, M.P.; Larson, N.B.; et al. Radiomics-based machine-learning models can detect pancreatic cancer on prediagnostic computed tomography scans at a substantial lead time before clinical diagnosis. *Gastroenterology* **2022**, *163*, 1435–1446. 1695
1696
1697
113. Liu, K.L.; Wu, T.; Chen, P.T.; Tsai, Y.M.; Roth, H.; Wu, M.S.; Liao, W.C.; Wang, W. Deep learning to distinguish pancreatic cancer tissue from non-cancerous pancreatic tissue: a retrospective study with cross-racial external validation. *The Lancet Digital Health* **2020**, *2*, e303–e313. 1698
1699
1700
114. Xia, Y.; Yao, J.; Lu, L.; Huang, L.; Xie, G.; Xiao, J.; Yuille, A.; Cao, K.; Zhang, L. Effective pancreatic cancer screening on non-contrast CT scans via anatomy-aware transformers. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part V 24. Springer, 2021, pp. 259–269. 1701
1702
1703
1704
115. Cao, K.; Xia, Y.; Yao, J.; Han, X.; Lambert, L.; Zhang, T.; Tang, W.; Jin, G.; Jiang, H.; Fang, X.; et al. Large-scale pancreatic cancer detection via non-contrast CT and deep learning. *Nature medicine* **2023**, *29*, 3033–3043. 1705
1706
116. Vaiyapuri, T.; Dutta, A.K.; Punithavathi, I.H.; Duraipandy, P.; Alotaibi, S.S.; Alsolai, H.; Mohamed, A.; Mahgoub, H. Intelligent deep-learning-enabled decision-making medical system for pancreatic tumor classification on CT images. In Proceedings of the Healthcare. MDPI, 2022, Vol. 10, p. 677. 1707
1708
1709
117. Huy, H.Q.; Dat, N.T.; Hiep, D.N.; Tram, N.N.; Vu, T.A.; Huong, P.T.V. Pancreatic Cancer Detection Based on CT Images Using Deep Learning. In Proceedings of the The International Conference on Intelligent Systems & Networks. Springer, 2023, pp. 66–72. 1710
1711

118. Yang, R.; Chen, Y.; Sa, G.; Li, K.; Hu, H.; Zhou, J.; Guan, Q.; Chen, F. CT classification model of pancreatic serous cystic neoplasms and mucinous cystic neoplasms based on a deep neural network. *Abdominal Radiology* **2022**, pp. 1–10. 1712
119. Bakasa, W.; Viriri, S. Stacked ensemble deep learning for pancreas cancer classification using extreme gradient boosting. *Frontiers in Artificial Intelligence* **2023**, 6. 1713
120. Roth, H.R.; Farag, A.; Lu, L.; Turkbey, E.B.; Summers, R.M. Deep convolutional networks for pancreas segmentation in CT imaging. In Proceedings of the Medical Imaging 2015: Image Processing. SPIE, 2015, Vol. 9413, pp. 378–385. 1714
121. Heinrich, M.P.; Oktay, O. BRIEFnet: Deep pancreas segmentation using binary sparse convolutions. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2017, pp. 329–337. 1715
122. Zhou, Y.; Xie, L.; Fishman, E.K.; Yuille, A.L. Deep supervision for pancreatic cyst segmentation in abdominal CT scans. In Proceedings of the International conference on medical image computing and computer-assisted intervention. Springer, 2017, pp. 222–230. 1716
123. Lee, C.Y.; Xie, S.; Gallagher, P.; Zhang, Z.; Tu, Z. Deeply-supervised nets. In Proceedings of the Artificial intelligence and statistics. Pmlr, 2015, pp. 562–570. 1717
124. Lu, L.; Jian, L.; Luo, J.; Xiao, B. Pancreatic segmentation via ringed residual U-Net. *IEEE Access* **2019**, 7, 172871–172878. 1718
125. Boers, T.; Hu, Y.; Gibson, E.; Barratt, D.; Bonmati, E.; Krdzalic, J.; van der Heijden, F.; Hermans, J.; Huisman, H. Interactive 3D U-net for the segmentation of the pancreas in computed tomography scans. *Physics in Medicine & Biology* **2020**, 65, 065002. 1719
126. Jiang, F.; Zhi, X.; Ding, X.; Tong, W.; Bian, Y. DLU-Net for pancreatic cancer segmentation. In Proceedings of the 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2020, pp. 1024–1028. 1720
127. Li, F.; Li, W.; Shu, Y.; Qin, S.; Xiao, B.; Zhan, Z. Multiscale receptive field based on residual network for pancreas segmentation in CT images. *Biomedical Signal Processing and Control* **2020**, 57, 101828. 1721
128. Li, Y.; Cai, W.; Gao, Y.; Li, C.; Hu, X. More than encoder: Introducing transformer decoder to upsample. In Proceedings of the 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2022, pp. 1597–1602. 1722
129. Juwita, J.; Hassan, G.; Akhtar, N.; Datta, A. M3bunet: Mobile Mean Max Unet for Pancreas Segmentation on Ct-Scans. Available at SSRN 4717898. 1723
130. Paithane, P.; Kakarwal, S. LMNS-Net: Lightweight Multiscale Novel Semantic-Net deep learning approach used for automatic pancreas image segmentation in CT scan images. *Expert Systems with Applications* **2023**, 234, 121064. 1724
131. Zhou, Z.; Bian, Y.; Pan, S.; Meng, Q.; Zhu, W.; Shi, F.; Chen, X.; Shao, C.; Xiang, D. A dual branch and fine-grained enhancement network for pancreatic tumor segmentation in contrast enhanced CT images. *Biomedical Signal Processing and Control* **2023**, 82, 104516. 1725
132. Chen, X.; Chen, Z.; Li, J.; Zhang, Y.D.; Lin, X.; Qian, X. Model-driven deep learning method for pancreatic cancer segmentation based on spiral-transformation. *IEEE Transactions on Medical Imaging* **2021**, 41, 75–87. 1726
133. Yu, L.; Yang, X.; Chen, H.; Qin, J.; Heng, P.A. Volumetric ConvNets with mixed residual connections for automated prostate segmentation from 3D MR images. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, 2017, Vol. 31. 1727
134. Roth, H.; Oda, M.; Shimizu, N.; Oda, H.; Hayashi, Y.; Kitasaka, T.; Fujiwara, M.; Misawa, K.; Mori, K. Towards dense volumetric pancreas segmentation in CT using 3D fully convolutional networks. In Proceedings of the Medical imaging 2018: image processing. SPIE, 2018, Vol. 10574, pp. 59–64. 1728
135. Chen, H.; Wang, X.; Huang, Y.; Wu, X.; Yu, Y.; Wang, L. Harnessing 2D networks and 3D features for automated pancreas segmentation from volumetric CT images. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part VI 22. Springer, 2019, pp. 339–347. 1729
136. Zhao, N.; Tong, N.; Ruan, D.; Sheng, K. Fully automated pancreas segmentation with two-stage 3D convolutional neural networks. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II 22. Springer, 2019, pp. 201–209. 1730
137. Zhang, J.; Xie, Y.; Xia, Y.; Shen, C. DoDNet: Learning to segment multi-organ and tumors from multiple partially labeled datasets. In Proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 1195–1204. 1731
138. Zhang, D.; Zhang, J.; Zhang, Q.; Han, J.; Zhang, S.; Han, J. Automatic pancreas segmentation based on lightweight DCNN modules and spatial prior propagation. *Pattern Recognition* **2021**, 114, 107762. 1732
139. Isensee, F.; Petersen, J.; Klein, A.; Zimmerer, D.; Jaeger, P.F.; Kohl, S.; Wasserthal, J.; Koehler, G.; Norajitra, T.; Wirkert, S.; et al. nnu-net: Self-adapting framework for u-net-based medical image segmentation. *arXiv preprint arXiv:1809.10486* **2018**. 1733
140. Yao, J.; Shi, Y.; Lu, L.; Xiao, J.; Zhang, L. DeepPrognosis: Preoperative Prediction of Pancreatic Cancer Survival and Surgical Margin via Contrast-Enhanced CT Imaging, 2020, [[arXiv:eeess.IV/2008.11853](https://arxiv.org/abs/2008.11853)]. 1734
141. Huang, X.; Deng, Z.; Li, D.; Yuan, X. Missformer: An effective medical image segmentation transformer. *arXiv preprint arXiv:2109.07162* **2021**. 1735
142. Dai, S.; Zhu, Y.; Jiang, X.; Yu, F.; Lin, J.; Yang, D. TD-Net: Trans-Deformer network for automatic pancreas segmentation. *Neurocomputing* **2023**, 517, 279–293. 1736
143. Rahman, M.M.; Shokouhmand, S.; Bhatt, S.; Faezipour, M. MIST: Medical Image Segmentation Transformer with Convolutional Attention Mixing (CAM) Decoder. In Proceedings of the Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2024, pp. 404–413. 1737

144. Zhou, H.Y.; Guo, J.; Zhang, Y.; Yu, L.; Wang, L.; Yu, Y. nnformer: Interleaved transformer for volumetric segmentation. *arXiv preprint arXiv:2109.03201* **2021**. 1771
1772
145. Hatamizadeh, A.; Tang, Y.; Nath, V.; Yang, D.; Myronenko, A.; Landman, B.; Roth, H.R.; Xu, D. Unetr: Transformers for 3d medical image segmentation. In Proceedings of the Proceedings of the IEEE/CVF winter conference on applications of computer vision, 2022, pp. 574–584. 1773
1774
1775
146. Tang, Y.; Yang, D.; Li, W.; Roth, H.R.; Landman, B.; Xu, D.; Nath, V.; Hatamizadeh, A. Self-supervised pre-training of swin transformers for 3d medical image analysis. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 20730–20740. 1776
1777
1778
147. Chen, J.; Mei, J.; Li, X.; Lu, Y.; Yu, Q.; Wei, Q.; Luo, X.; Xie, Y.; Adeli, E.; Wang, Y.; et al. 3d transunet: Advancing medical image segmentation through vision transformers. *arXiv preprint arXiv:2310.07781* **2023**. 1779
1780
148. Qu, T.; Li, X.; Wang, X.; Deng, W.; Mao, L.; He, M.; Li, X.; Wang, Y.; Liu, Z.; Zhang, L.; et al. Transformer guided progressive fusion network for 3D pancreas and pancreatic mass segmentation. *Medical Image Analysis* **2023**, *86*, 102801. 1781
1782
149. Guo, Z.; Zhang, L.; Lu, L.; Bagheri, M.; Summers, R.M.; Sonka, M.; Yao, J. Deep LOGISMOS: Deep learning graph-based 3D segmentation of pancreatic tumors on CT scans. In Proceedings of the 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018). IEEE, 2018, pp. 1230–1233. 1783
1784
1785
150. Soberanis-Mukul, R.D.; Navab, N.; Albarqouni, S. Uncertainty-based graph convolutional networks for organ segmentation refinement. In Proceedings of the Medical Imaging with Deep Learning. PMLR, 2020, pp. 755–769. 1786
1787
151. Hu, P.; Li, X.; Tian, Y.; Tang, T.; Zhou, T.; Bai, X.; Zhu, S.; Liang, T.; Li, J. Automatic pancreas segmentation in CT images with distance-based saliency-aware DenseASPP network. *IEEE journal of biomedical and health informatics* **2020**, *25*, 1601–1611. 1788
152. Zhao, T.; Cao, K.; Yao, J.; Nogues, I.; Lu, L.; Huang, L.; Xiao, J.; Yin, Z.; Zhang, L. 3D graph anatomy geometry-integrated network for pancreatic mass segmentation, diagnosis, and quantitative patient management. In Proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 13743–13752. 1790
1791
1792
153. Liu, S.; Liang, S.; Huang, X.; Yuan, X.; Zhong, T.; Zhang, Y. Graph-enhanced U-Net for semi-supervised segmentation of pancreas from abdomen CT scan. *Physics in Medicine & Biology* **2022**, *67*, 155017. 1793
1794
154. Zhu, Z.; Liu, C.; Yang, D.; Yuille, A.; Xu, D. V-NAS: Neural architecture search for volumetric medical image segmentation. In Proceedings of the 2019 International conference on 3d vision (3DV). IEEE, 2019, pp. 240–248. 1795
1796
155. He, Y.; Yang, D.; Roth, H.; Zhao, C.; Xu, D. Dints: Differentiable neural network topology search for 3d medical image segmentation. In Proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 5841–5850. 1797
1798
1799
156. Liu, J.; Zhang, Y.; Chen, J.N.; Xiao, J.; Lu, Y.; A Landman, B.; Yuan, Y.; Yuille, A.; Tang, Y.; Zhou, Z. Clip-driven universal model for organ segmentation and tumor detection. In Proceedings of the Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 21152–21164. 1800
1801
1802
157. Liu, J.; Zhang, Y.; Wang, K.; Yavuz, M.C.; Chen, X.; Yuan, Y.; Li, H.; Yang, Y.; Yuille, A.; Tang, Y.; et al. Universal and Extensible Language-Vision Models for Organ Segmentation and Tumor Detection from Abdominal Computed Tomography. *arXiv preprint arXiv:2405.18356* **2024**. 1803
1804
1805
158. Radford, A.; Kim, J.W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; et al. Learning transferable visual models from natural language supervision. In Proceedings of the International conference on machine learning. PMLR, 2021, pp. 8748–8763. 1806
1807
1808
159. Huang, Y.; Yang, X.; Liu, L.; Zhou, H.; Chang, A.; Zhou, X.; Chen, R.; Yu, J.; Chen, J.; Chen, C.; et al. Segment anything model for medical images? *Medical Image Analysis* **2024**, *92*, 103061. 1809
1810
160. Mazurowski, M.A.; Dong, H.; Gu, H.; Yang, J.; Konz, N.; Zhang, Y. Segment anything model for medical image analysis: an experimental study. *Medical Image Analysis* **2023**, *89*, 102918. 1811
1812
161. He, S.; Bao, R.; Li, J.; Grant, P.E.; Ou, Y. Accuracy of segment-anything model (sam) in medical image segmentation tasks. *arXiv preprint arXiv:2304.09324* **2023**. 1813
1814
162. Knolle, M.; Kaassis, G.; Jungmann, F.; Ziegelmayer, S.; Sasse, D.; Makowski, M.; Rueckert, D.; Braren, R. Efficient, high-performance semantic segmentation using multi-scale feature extraction. *Plos one* **2021**, *16*, e0255397. 1815
1816
163. Wang, P.; Shen, C.; Wang, W.; Oda, M.; Fuh, C.S.; Mori, K.; Roth, H.R. ConDistFL: Conditional Distillation for Federated Learning from Partially Annotated Data. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2023, pp. 311–321. 1817
1818
1819
164. Man, Y.; Huang, Y.; Feng, J.; Li, X.; Wu, F. Deep Q learning driven CT pancreas segmentation with geometry-aware U-Net. *IEEE transactions on medical imaging* **2019**, *38*, 1971–1980. 1820
1821
165. Dogan, R.O.; Dogan, H.; Bayrak, C.; Kayikcioglu, T. A two-phase approach using mask R-CNN and 3D U-Net for high-accuracy automatic segmentation of pancreas in CT imaging. *Computer Methods and Programs in Biomedicine* **2021**, *207*, 106141. 1822
1823
166. Zhang, Z.; Li, S.; Wang, Z.; Lu, Y. A novel and efficient tumor detection framework for pancreatic cancer via CT images. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2020, pp. 1160–1164. 1824
1825
1826
167. Baumgartner, M.; Jäger, P.F.; Isensee, F.; Maier-Hein, K.H. nnDetection: a self-configuring method for medical object detection. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part V 24. Springer, 2021, pp. 530–539. 1827
1828
1829

168. Jaeger, P.F.; Kohl, S.A.; Bickelhaupt, S.; Isensee, F.; Kuder, T.A.; Schlemmer, H.P.; Maier-Hein, K.H. Retina U-Net: Embarrassingly simple exploitation of segmentation supervision for medical object detection. In Proceedings of the Machine Learning for Health Workshop. PMLR, 2020, pp. 171–183. [1830](#)
169. Juneja, M.; Singh, G.; Chanana, C.; Verma, R.; Thakur, N.; Jindal, P. Region-based Convolutional Neural Network (R-CNN) architecture for auto-cropping of pancreatic computed tomography. *The Imaging Science Journal* **2023**, pp. 1–14. [1832](#)
170. Dinesh, M.; Bacanin, N.; Askar, S.; Abouhawwash, M. Diagnostic ability of deep learning in detection of pancreatic tumour. *Scientific Reports* **2023**, *13*, 9725. [1835](#)
171. Lee, W.; Park, H.J.; Lee, H.J.; Jun, E.; Song, K.B.; Hwang, D.W.; Lee, J.H.; Lim, K.; Kim, N.; Lee, S.S.; et al. Preoperative data-based deep learning model for predicting postoperative survival in pancreatic cancer patients. *International Journal of Surgery* **2022**, *105*, 106851. [1838](#)
172. Tran, D.; Wang, H.; Torresani, L.; Ray, J.; LeCun, Y.; Paluri, M. A closer look at spatiotemporal convolutions for action recognition. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 2018, pp. 6450–6459. [1840](#)
173. Hara, K.; Kataoka, H.; Satoh, Y. Can spatiotemporal 3d cnns retrace the history of 2d cnns and imagenet? In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 2018, pp. 6546–6555. [1842](#)
174. Zhang, Y.; Lobo-Mueller, E.M.; Karanicolas, P.; Gallinger, S.; Haider, M.A.; Khalvati, F. Improving prognostic performance in resectable pancreatic ductal adenocarcinoma using radiomics and deep learning features fusion in CT images. *Scientific reports* **2021**, *11*, 1378. [1844](#)
175. Chen, X.; Wang, W.; Jiang, Y.; Qian, X. A dual-transformation with contrastive learning framework for lymph node metastasis prediction in pancreatic cancer. *Medical Image Analysis* **2023**, *85*, 102753. [1847](#)
176. Gibson, E.; Giganti, F.; Hu, Y.; Bonmati, E.; Bandula, S.; Gurusamy, K.; Davidson, B.; Pereira, S.P.; Clarkson, M.J.; Barratt, D.C. Automatic multi-organ segmentation on abdominal CT with dense V-networks. *IEEE transactions on medical imaging* **2018**, *37*, 1822–1834. [1849](#)
177. Lyu, P.; Neely, B.; Solomon, J.; Rigioli, F.; Ding, Y.; Schwartz, F.R.; Thomsen, B.; Lowry, C.; Samei, E.; Marin, D. Effect of deep learning image reconstruction in the prediction of resectability of pancreatic cancer: Diagnostic performance and reader confidence. *European journal of radiology* **2021**, *141*, 109825. [1853](#)
178. Noda, Y.; Iritani, Y.; Kawai, N.; Miyoshi, T.; Ishihara, T.; Hyodo, F.; Matsuo, M. Deep learning image reconstruction for pancreatic low-dose computed tomography: comparison with hybrid iterative reconstruction. *Abdominal Radiology* **2021**, *46*, 4238–4244. [1855](#)
179. Chi, J.; Sun, Z.; Zhao, T.; Wang, H.; Yu, X.; Wu, C. Low-dose ct image super-resolution network with dual-guidance feature distillation and dual-path content communication. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2023, pp. 98–108. [1857](#)
180. Takai, Y.; Noda, Y.; Asano, M.; Kawai, N.; Kaga, T.; Tsuchida, Y.; Miyoshi, T.; Hyodo, F.; Kato, H.; Matsuo, M. Deep-learning image reconstruction for 80-kVp pancreatic CT protocol: Comparison of image quality and pancreatic ductal adenocarcinoma visibility with hybrid-iterative reconstruction. *European Journal of Radiology* **2023**, *165*, 110960. [1860](#)
181. Shi, J.; Pelt, D.M.; Batenburg, K.J. SR4ZCT: Self-supervised Through-Plane Resolution Enhancement for CT Images with Arbitrary Resolution and Overlap. In Proceedings of the International Workshop on Machine Learning in Medical Imaging. Springer, 2023, pp. 52–61. [1862](#)
182. Liu, Y.; Lei, Y.; Wang, T.; Fu, Y.; Tang, X.; Curran, W.J.; Liu, T.; Patel, P.; Yang, X. CBCT-based synthetic CT generation using deep-attention cycleGAN for pancreatic adaptive radiotherapy. *Medical physics* **2020**, *47*, 2472–2483. [1866](#)
183. Dai, X.; Lei, Y.; Wynne, J.; Janopaul-Naylor, J.; Wang, T.; Roper, J.; Curran, W.J.; Liu, T.; Patel, P.; Yang, X. Synthetic CT-aided multiorgan segmentation for CBCT-guided adaptive pancreatic radiotherapy. *Medical physics* **2021**, *48*, 7063–7073. [1869](#)
184. Shi, Y.; Tang, H.; Baine, M.J.; Hollingsworth, M.A.; Du, H.; Zheng, D.; Zhang, C.; Yu, H. 3DGAUnet: 3D generative adversarial networks with a 3D U-net based generator to achieve the accurate and effective synthesis of clinical tumor image data for pancreatic cancer. *Cancers* **2023**, *15*, 5496. [1871](#)
185. Hooshangnejad, H.; Chen, Q.; Feng, X.; Zhang, R.; Ding, K. deepPERFECT: Novel Deep Learning CT Synthesis Method for Expedited Pancreatic Cancer Radiotherapy. *Cancers* **2023**, *15*, 3061. [1873](#)
186. Peng, J.; Liu, Y.; Jiang, D.; Wang, X.; Peng, P.; He, S.; Zhang, W.; Zhou, F. Deep Learning and GAN-Synthesis for Auto-Segmentation of Pancreatic Cancer by Non-Enhanced CT for Adaptive Radiotherapy. *International Journal of Radiation Oncology, Biology, Physics* **2023**, *117*, e499–e500. [1875](#)
187. Guan, Q.; Chen, Y.; Wei, Z.; Heidari, A.A.; Hu, H.; Yang, X.H.; Zheng, J.; Zhou, Q.; Chen, H.; Chen, F. Medical image augmentation for lesion detection using a texture-constrained multichannel progressive GAN. *Computers in Biology and Medicine* **2022**, *145*, 105444. [1878](#)
188. Caverly, R.H. MRI fundamentals: RF aspects of magnetic resonance imaging (MRI). *IEEE Microwave Magazine* **2015**, *16*, 20–33. [1881](#)
189. Fatahi, M.; Speck, O.; et al. Magnetic resonance imaging (MRI): A review of genetic damage investigations. *Mutation Research/Reviews in Mutation Research* **2015**, *764*, 51–63. [1882](#)
190. Eshed, I.; Hermann, K.G.A. MRI in imaging of rheumatic diseases: an overview for clinicians. *Clin Exp Rheumatol* **2018**, *36*, 10–5. [1884](#)
191. Smith, N.B.; Webb, A. *Introduction to medical imaging: physics, engineering and clinical applications*; Cambridge university press, 2010. [1885](#)
192. Cui, S.; Tang, T.; Su, Q.; Wang, Y.; Shu, Z.; Yang, W.; Gong, X. Radiomic nomogram based on MRI to predict grade of branching type intraductal papillary mucinous neoplasms of the pancreas: A multicenter study. *Cancer Imaging* **2021**, *21*, 1–13. [1886](#)

193. Chen, W.; Ji, H.; Feng, J.; Liu, R.; Yu, Y.; Zhou, R.; Zhou, J. Classification of pancreatic cystic neoplasms based on multimodality images. In Proceedings of the International Workshop on Machine Learning in Medical Imaging. Springer, 2018, pp. 161–169. 1888
1889
194. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2818–2826. 1890
1891
195. Chen, X.; Chen, Y.; Ma, C.; Liu, X.; Tang, X. Classification of pancreatic tumors based on MRI images using 3D convolutional neural networks. In Proceedings of the Proceedings of the 2Nd International Symposium on Image Computing and Digital Medicine, 2018, pp. 92–96. 1892
1893
1894
196. Corral, J.E.; Hussein, S.; Kandel, P.; Bolan, C.W.; Bagci, U.; Wallace, M.B. Deep learning to classify intraductal papillary mucinous neoplasms using magnetic resonance imaging. *Pancreas* **2019**, *48*, 805–810. 1895
1896
197. Chatfield, K.; Simonyan, K.; Vedaldi, A.; Zisserman, A. Return of the devil in the details: Delving deep into convolutional nets. *arXiv preprint arXiv:1405.3531* **2014**. 1897
1898
198. Hussein, S.; Kandel, P.; Bolan, C.W.; Wallace, M.B.; Bagci, U. Lung and pancreatic tumor characterization in the deep learning era: novel supervised and unsupervised learning approaches. *IEEE transactions on medical imaging* **2019**, *38*, 1777–1787. 1899
1900
199. Asatryan, H.; Thomas, E.L.; Fitzpatrick, J.; Bell, J.D.; Villarini, B. Advancing pancreas segmentation in multi-protocol mri volumes using hausdorff-sine loss function. In Proceedings of the Machine Learning in Medical Imaging: 10th International Workshop, MLMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings 10. Springer, 2019, pp. 27–35. 1901
1902
1903
1904
200. Liang, Y.; Schott, D.; Zhang, Y.; Wang, Z.; Nasief, H.; Paulson, E.; Hall, W.; Knechtges, P.; Erickson, B.; Li, X.A. Auto-segmentation of pancreatic tumor in multi-parametric MRI using deep convolutional neural networks. *Radiotherapy and Oncology* **2020**, *145*, 193–200. 1905
1906
201. Mazor, N.; Dar, G.; Lederman, R.; Lev-Cohain, N.; Sosna, J.; Joskowicz, L. MC3DU-Net: a multisequence cascaded pipeline for the detection and segmentation of pancreatic cysts in MRI. *International Journal of Computer Assisted Radiology and Surgery* **2023**, pp. 1–10. 1908
1909
1910
202. Li, J.; Feng, C.; Shen, Q.; Lin, X.; Qian, X. Pancreatic cancer segmentation in unregistered multi-parametric MRI with adversarial learning and multi-scale supervision. *Neurocomputing* **2022**, *467*, 310–322. 1911
1912
203. Li, J.; Feng, C.; Lin, X.; Qian, X. Utilizing GCN and meta-learning strategy in unsupervised domain adaptation for pancreatic cancer segmentation. *IEEE Journal of Biomedical and Health Informatics* **2021**, *26*, 79–89. 1913
1914
204. Cai, J.; Lu, L.; Zhang, Z.; Xing, F.; Yang, L.; Yin, Q. Pancreas segmentation in MRI using graph-based decision fusion on convolutional neural networks. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2016, pp. 442–450. 1915
1916
1917
205. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. Imagenet large scale visual recognition challenge. *International journal of computer vision* **2015**, *115*, 211–252. 1918
1919
206. Han, S.; Kim, J.H.; Yoo, J.; Jang, S. Prediction of recurrence after surgery based on preoperative MRI features in patients with pancreatic neuroendocrine tumors. *European Radiology* **2021**, pp. 1–12. 1920
1921
207. Xu, X.; Qu, J.; Zhang, Y.; Qian, X.; Chen, T.; Liu, Y. Development and validation of an MRI-radiomics nomogram for the prognosis of pancreatic ductal adenocarcinoma. *Frontiers in Oncology* **2023**, *13*, 1074445. 1922
1923
208. Van Roessel, S.; Kasumova, G.G.; Verheij, J.; Najarian, R.M.; Maggino, L.; De Pastena, M.; Malleo, G.; Marchegiani, G.; Salvia, R.; Ng, S.C.; 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 surgery* **2018**, *153*, e183617–e183617. 1924
1925
1926
209. Chaika, M.; Afat, S.; Wessling, D.; Afat, C.; Nickel, D.; Kannengiesser, S.; Herrmann, J.; Almansour, H.; Männlin, S.; Othman, A.E.; et al. Deep learning-based super-resolution gradient echo imaging of the pancreas: Improvement of image quality and reduction of acquisition time. *Diagnostic and Interventional Imaging* **2023**, *104*, 53–59. 1927
1928
1929
210. Fusaroli, P.; Caletti, G. Endoscopic ultrasonography. *Endoscopy* **2003**, *35*, 127–135. 1930
211. Dimagno, E.P.; Regan, P.T.; Clain, J.E.; James, E.; Buxton, J.L. Human endoscopic ultrasonography. *Gastroenterology* **1982**, *83*, 824–829. 1931
1932
212. Ruano, J.; Jaramillo, M.; Gómez, M.; Romero, E. Robust Descriptor of Pancreatic Tissue for Automatic Detection of Pancreatic Cancer in Endoscopic Ultrasonography. *Ultrasound in medicine & biology* **2022**, *48*, 1602–1614. 1933
1934
213. Kuwahara, T.; Hara, K.; Mizuno, N.; Okuno, N.; Matsumoto, S.; Obata, M.; Kurita, Y.; Koda, H.; Toriyama, K.; Onishi, S.; et al. Usefulness of deep learning analysis for the diagnosis of malignancy in intraductal papillary mucinous neoplasms of the pancreas. *Clinical and translational gastroenterology* **2019**, *10*. 1935
1936
1937
214. Zhang, J.; Zhu, L.; Yao, L.; Ding, X.; Chen, D.; Wu, H.; Lu, Z.; Zhou, W.; Zhang, L.; An, P.; et al. Deep learning-based pancreas segmentation and station recognition system in EUS: Development and validation of a useful training tool (with video). *Gastrointestinal endoscopy* **2020**, *92*, 874–885. 1938
1939
1940
215. Udrișteiu, A.L.; Cazacu, I.M.; Gruionu, L.G.; Gruionu, G.; Iacob, A.V.; Burtea, D.E.; Ungureanu, B.S.; Costache, M.I.; Constantin, A.; Popescu, C.F.; et al. Real-time computer-aided diagnosis of focal pancreatic masses from endoscopic ultrasound imaging based on a hybrid convolutional and long short-term memory neural network model. *PLoS One* **2021**, *16*, e0251701. 1941
1942
1943
216. Nguon, L.S.; Seo, K.; Lim, J.H.; Song, T.J.; Cho, S.H.; Park, J.S.; Park, S. Deep learning-based differentiation between mucinous cystic neoplasm and serous cystic neoplasm in the pancreas using endoscopic ultrasonography. *Diagnostics* **2021**, *11*, 1052. 1944
1945

217. Bonmati, E.; Hu, Y.; Grimwood, A.; Johnson, G.J.; Goodchild, G.; Keane, M.G.; Gurusamy, K.; Davidson, B.; Clarkson, M.J.; Pereira, S.P.; et al. Voice-assisted image labeling for endoscopic ultrasound classification using neural networks. *IEEE transactions on medical imaging* **2021**, *41*, 1311–1319. 1946
1947
1948
218. Vilas-Boas, F.; Ribeiro, T.; Afonso, J.; Cardoso, H.; Lopes, S.; Moutinho-Ribeiro, P.; Ferreira, J.; Mascarenhas-Saraiva, M.; Macedo, G. Deep Learning for Automatic Differentiation of Mucinous versus Non-Mucinous Pancreatic Cystic Lesions: A Pilot Study. *Diagnostics* **2022**, *12*, 2041. 1949
1950
1951
219. Jaramillo, M.; Ruano, J.; Gómez, M.; Romero, E. Automatic detection of pancreatic tumors in endoscopic ultrasound videos using deep learning techniques. In Proceedings of the Medical Imaging 2022: Ultrasonic Imaging and Tomography. SPIE, 2022, Vol. 12038, pp. 106–115. 1952
1953
1954
220. Ren, Y.; Zou, D.; Xu, W.; Zhao, X.; Lu, W.; He, X. Bimodal segmentation and classification of endoscopic ultrasonography images for solid pancreatic tumor. *Biomedical Signal Processing and Control* **2023**, *83*, 104591. 1955
1956
221. 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*, 140–149. 1957
1958
1959
222. Tan, M.; Le, Q. Efficientnetv2: Smaller models and faster training. In Proceedings of the International conference on machine learning. PMLR, 2021, pp. 10096–10106. 1960
1961
223. Fleurentin, A.; Mazellier, J.P.; Meyer, A.; Montanelli, J.; Swanstrom, L.; Gallix, B.; Sosa Valencia, L.; Padoy, N. Automatic pancreas anatomical part detection in endoscopic ultrasound videos. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* **2023**, *11*, 1136–1142. 1962
1963
1964
224. Iwasa, Y.; Iwashita, T.; Takeuchi, Y.; Ichikawa, H.; Mita, N.; Uemura, S.; Shimizu, M.; Kuo, Y.T.; Wang, H.P.; Hara, T. Automatic segmentation of pancreatic tumors using deep learning on a video image of contrast-enhanced endoscopic ultrasound. *Journal of clinical medicine* **2021**, *10*, 3589. 1965
1966
1967
225. Oh, S.; Kim, Y.J.; Park, Y.T.; Kim, K.G. Automatic pancreatic cyst lesion segmentation on EUS images using a deep-learning approach. *Sensors* **2021**, *22*, 245. 1968
1969
226. Seo, K.; Lim, J.H.; Seo, J.; Nguon, L.S.; Yoon, H.; Park, J.S.; Park, S. Semantic Segmentation of Pancreatic Cancer in Endoscopic Ultrasound Images Using Deep Learning Approach. *Cancers* **2022**, *14*, 5111. 1970
1971
227. Tang, A.; Gong, P.; Fang, N.; Ye, M.; Hu, S.; Liu, J.; Wang, W.; Gao, K.; Wang, X.; Tian, L. Endoscopic ultrasound diagnosis system based on deep learning in images capture and segmentation training of solid pancreatic masses. *Medical Physics* **2023**. 1972
1973
228. Meyer, A.; Fleurentin, A.; Montanelli, J.; Mazellier, J.P.; Swanstrom, L.; Gallix, B.; Exarchakis, G.; Sosa Valencia, L.; Padoy, N. Spatio-Temporal Model for EUS Video Detection of Pancreatic Anatomy Structures. In Proceedings of the International Workshop on Advances in Simplifying Medical Ultrasound. Springer, 2022, pp. 13–22. 1974
1975
1976
229. Wu, H.; Chen, Y.; Wang, N.; Zhang, Z. Sequence level semantics aggregation for video object detection. In Proceedings of the Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 9217–9225. 1977
1978
230. Gong, T.; Chen, K.; Wang, X.; Chu, Q.; Zhu, F.; Lin, D.; Yu, N.; Feng, H. Temporal ROI align for video object recognition. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, 2021, Vol. 35, pp. 1442–1450. 1979
1980
231. Tian, G.; Xu, D.; He, Y.; Chai, W.; Deng, Z.; Cheng, C.; Jin, X.; Wei, G.; Zhao, Q.; Jiang, T. Deep learning for real-time auxiliary diagnosis of pancreatic cancer in endoscopic ultrasonography. *Frontiers in Oncology* **2022**, *12*, 973652. 1981
1982
232. Jaramillo, M.; Ruano, J.; Bravo, D.; Medina, S.; Gómez, M.; González, F.A.; Romero, E. Automatic Localization of Pancreatic Tumoral Regions in Whole Sequences of Echoendoscopy Procedures. In Proceedings of the 2023 19th International Symposium on Medical Information Processing and Analysis (SIPAIM). IEEE, 2023, pp. 1–5. 1983
1984
1985
233. Grimwood, A.; Ramalhinho, J.; Baum, Z.M.; Montaña-Brown, N.; Johnson, G.J.; Hu, Y.; Clarkson, M.J.; Pereira, S.P.; Barratt, D.C.; Bonmati, E. Endoscopic ultrasound image synthesis using a cycle-consistent adversarial network. In Proceedings of the Simplifying Medical Ultrasound: Second International Workshop, ASMUS 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings 2. Springer, 2021, pp. 169–178. 1986
1987
1988
1989
234. Cherry, S.R.; Dahlbom, M.; Cherry, S.R.; Dahlbom, M. *PET: physics, instrumentation, and scanners*; Springer, 2006. 1990
235. Schlyer, D.J. PET tracers and radiochemistry. *ANNALS-ACADEMY OF MEDICINE SINGAPORE* **2004**, *33*, 146–154. 1991
236. Kapoor, V.; McCook, B.M.; Torok, F.S. An introduction to PET-CT imaging. *Radiographics* **2004**, *24*, 523–543. 1992
237. Vandenberghe, S.; Moskal, P.; Karp, J.S. State of the art in total body PET. *EJNMMI physics* **2020**, *7*, 1–33. 1993
238. Townsend, D.W.; Carney, J.P.; Yap, J.T.; Hall, N.C. PET/CT today and tomorrow. *Journal of Nuclear Medicine* **2004**, *45*, 4S–14S. 1994
239. Li, S.; Jiang, H.; Wang, Z.; Zhang, G.; Yao, Y.d. An effective computer aided diagnosis model for pancreas cancer on PET/CT images. *Computer methods and programs in biomedicine* **2018**, *165*, 205–214. 1995
1996
240. Xing, H.; Hao, Z.; Zhu, W.; Sun, D.; Ding, J.; Zhang, H.; Liu, Y.; Huo, L. Preoperative prediction of pathological grade in pancreatic ductal adenocarcinoma based on 18 F-FDG PET/CT radiomics. *EJNMMI research* **2021**, *11*, 1–10. 1997
1998
241. Van Griethuysen, J.J.; Fedorov, A.; Parmar, C.; Hosny, A.; Aucoin, N.; Narayan, V.; Beets-Tan, R.G.; Fillion-Robin, J.C.; Pieper, S.; Aerts, H.J. Computational radiomics system to decode the radiographic phenotype. *Cancer research* **2017**, *77*, e104–e107. 1999
2000
242. Zhang, Y.; Cheng, C.; Liu, Z.; Wang, L.; Pan, G.; Sun, G.; Chang, Y.; Zuo, C.; Yang, X. Radiomics analysis for the differentiation of autoimmune pancreatitis and pancreatic ductal adenocarcinoma in 18F-FDG PET/CT. *Medical Physics* **2019**, *46*, 4520–4530. 2001
2002

243. Zhang, G.; Bao, C.; Liu, Y.; Wang, Z.; Du, L.; Zhang, Y.; Wang, F.; Xu, B.; Zhou, S.K.; Liu, R. 18F-FDG-PET/CT-based deep learning model for fully automated prediction of pathological grading for pancreatic ductal adenocarcinoma before surgery. *EJNMMI research* **2023**, *13*, 49. 2003
2004
2005
244. Wei, W.; Jia, G.; Wu, Z.; Wang, T.; Wang, H.; Wei, K.; Cheng, C.; Liu, Z.; Zuo, C. A multidomain fusion model of radiomics and deep learning to discriminate between PDAC and AIP based on 18F-FDG PET/CT images. *Japanese journal of radiology* **2023**, *41*, 417–427. 2006
2007
2008
245. Suganuma, Y.; Teramoto, A.; Saito, K.; Fujita, H.; Suzuki, Y.; Tomiyama, N.; Kido, S. Hybrid Multiple-Organ Segmentation Method Using Multiple U-Nets in PET/CT Images. *Applied Sciences* **2023**, *13*, 10765. 2009
2010
246. Wang, F.; Cheng, C.; Cao, W.; Wu, Z.; Wang, H.; Wei, W.; Yan, Z.; Liu, Z. MFCNet: A multi-modal fusion and calibration networks for 3D pancreas tumor segmentation on PET-CT images. *Computers in Biology and Medicine* **2023**, *155*, 106657. 2011
2012
2013
- <https://doi.org/https://doi.org/10.1016/j.combiomed.2023.106657>
247. Shao, M.; Cheng, C.; Hu, C.; Zheng, J.; Zhang, B.; Wang, T.; Jin, G.; Liu, Z.; Zuo, C. Semisupervised 3D segmentation of pancreatic tumors in positron emission tomography/computed tomography images using a mutual information minimization and cross-fusion strategy. *Quantitative Imaging in Medicine and Surgery* **2024**, *14*, 1747. 2014
2015
2016
248. Wang, H.; Wu, Z.; Wang, F.; Wei, W.; Wei, K.; Liu, Z. MAFF: Multi-Scale and Self-Adaptive Attention Feature Fusion Network for Pancreatic Lesion Detection in PET / CT Images. In Proceedings of the Proceedings of the 2022 6th International Conference on Electronic Information Technology and Computer Engineering; Association for Computing Machinery: New York, NY, USA, 2023; EITCE '22, p. 1412–1419. <https://doi.org/10.1145/3573428.3573678>. 2017
2018
2019
2020
249. Park, Y.J.; Park, Y.S.; Kim, S.T.; Hyun, S.H. A machine learning approach using [18F] FDG PET-based radiomics for prediction of tumor grade and prognosis in pancreatic neuroendocrine tumor. *Molecular Imaging and Biology* **2023**, pp. 1–14. 2021
2022
250. Mendez, A.J.; Tahoces, P.G.; Lado, M.J.; Souto, M.; Vidal, J.J. Computer-aided diagnosis: Automatic detection of malignant masses in digitized mammograms. *Medical Physics* **1998**, *25*, 957–964. 2023
2024
251. Gurcan, M.N.; Boucheron, L.E.; Can, A.; Madabhushi, A.; Rajpoot, N.M.; Yener, B. Histopathological image analysis: A review. *IEEE reviews in biomedical engineering* **2009**, *2*, 147–171. 2025
2026
252. Farahani, N.; Parwani, A.V.; Pantanowitz, L. Whole slide imaging in pathology: advantages, limitations, and emerging perspectives. *Pathology and Laboratory Medicine International* **2015**, pp. 23–33. 2027
2028
253. Michael, C.W.; Kameyama, K.; Kitagawa, W.; Azar, N. Rapid on-site evaluation (ROSE) for fine needle aspiration of thyroid: benefits, challenges and innovative solutions. *Gland Surgery* **2020**, *9*, 1708. 2029
2030
254. da Cunha Santos, G.; Ko, H.M.; Saieg, M.A.; Geddie, W.R. “The petals and thorns” of ROSE (rapid on-site evaluation). *Cancer cytopathology* **2013**, *121*, 4–8. 2031
2032
255. Saillard, C.; Delecourt, F.; Schmauch, B.; Moindrot, O.; Svrcek, M.; Bardier-Dupas, A.; Emile, J.F.; Ayadi, M.; Rebours, V.; De Mestier, L.; et al. PACpAInt: a deep learning approach to identify molecular subtypes of pancreatic adenocarcinoma on histology slides. *bioRxiv* **2022**, pp. 2022–01. 2033
2034
2035
256. Chang, Y.H.; Thibault, G.; Madin, O.; Azimi, V.; Meyers, C.; Johnson, B.; Link, J.; Margolin, A.; Gray, J.W. Deep learning based Nucleus Classification in pancreas histological images. In Proceedings of the 2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE, 2017, pp. 672–675. 2036
2037
2038
257. Le, H.; Samaras, D.; Kurc, T.; Gupta, R.; Shroyer, K.; Saltz, J. Pancreatic cancer detection in whole slide images using noisy label annotations. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I 22. Springer, 2019, pp. 541–549. 2039
2040
2041
258. Sehmi, M.N.M.; Fauzi, M.F.A.; Ahmad, W.S.H.M.W.; Chan, E.W.L. Pancreatic cancer grading in pathological images using deep learning convolutional neural networks. *F1000Research* **2021**, *10*. 2042
2043
259. Ono, N.; Iwamoto, C.; Ohuchida, K. Construction of Classifier of Tumor Cell Types of Pancreas Cancer Based on Pathological Images Using Deep Learning. *Multidisciplinary Computational Anatomy: Toward Integration of Artificial Intelligence with MCA-based Medicine* **2022**, pp. 145–148. 2044
2045
2046
260. Zhang, T.; Feng, Y.; Feng, Y.; Zhao, Y.; Lei, Y.; Ying, N.; Yan, Z.; He, Y.; Zhang, G. Shuffle Instances-based Vision Transformer for Pancreatic Cancer ROSE Image Classification. *arXiv preprint arXiv:2208.06833* **2022**. 2047
2048
261. Ghoshal, B.; Ghoshal, B.; Tucker, A. Leveraging Uncertainty in Deep Learning for Pancreatic Adenocarcinoma Grading. In Proceedings of the Annual Conference on Medical Image Understanding and Analysis. Springer, 2022, pp. 565–577. 2049
2050
262. Kou, Y.; Xia, C.; Jiao, Y.; Zhang, D.; Ge, R. DACTransNet: A Hybrid CNN-Transformer Network for Histopathological Image Classification of Pancreatic Cancer. In Proceedings of the CAAI International Conference on Artificial Intelligence. Springer, 2023, pp. 422–434. 2051
2052
2053
263. Gao, E.; Jiang, H.; Zhou, Z.; Yang, C.; Chen, M.; Zhu, W.; Shi, F.; Chen, X.; Zheng, J.; Bian, Y.; et al. Automatic multi-tissue segmentation in pancreatic pathological images with selected multi-scale attention network. *Computers in Biology and Medicine* **2022**, *151*, 106228. 2054
2055
2056
264. 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*. 2057
2058
2059

265. Liu, A.; Jiang, H.; Cao, W.; Cui, W.; Xiang, D.; Shao, C.; Liu, Z.; Bian, Y.; Zheng, J. MLAGG-Net: Multi-level aggregation and global guidance network for pancreatic lesion segmentation in histopathological images. *Biomedical Signal Processing and Control* **2023**, *86*, 105303. [\[CrossRef\]](#)
266. Gao, W.; Jiang, H.; Jiao, Y.; Wang, X.; Xu, J. Multi-tissue segmentation model of whole slide image of pancreatic cancer based on multi task and attention mechanism. *Sheng wu yi xue Gong Cheng xue za zhi= Journal of Biomedical Engineering= Shengwu Yixue Gongchengxue Zazhi* **2023**, *40*, 70–78. [\[CrossRef\]](#)
267. Chen, Z.M.; Liao, Y.; Zhou, X.; Yu, W.; Zhang, G.; Ge, Y.; Ke, T.; Shi, K. Pancreatic cancer pathology image segmentation with channel and spatial long-range dependencies. *Computers in Biology and Medicine* **2024**, *169*, 107844. [\[CrossRef\]](#)
268. Janssen, B.V.; Theijse, R.; van Roessel, S.; de Ruiter, R.; Berkel, A.; Huiskens, J.; Busch, O.R.; Wilmink, J.W.; Kazemier, G.; Valkema, P.; et al. Artificial intelligence-based segmentation of residual tumor in histopathology of pancreatic cancer after neoadjuvant treatment. *Cancers* **2021**, *13*, 5089. [\[CrossRef\]](#)
269. Yang, C.; Xiang, D.; Bian, Y.; Lu, J.; Jiang, H.; Zheng, J. Gland segmentation in pancreas histopathology images based on selective multi-scale attention. In Proceedings of the Medical Imaging 2021: Image Processing. SPIE, 2021, Vol. 11596, pp. 699–705. [\[CrossRef\]](#)
270. Fu, H.; Mi, W.; Pan, B.; Guo, Y.; Li, J.; Xu, R.; Zheng, J.; Zou, C.; Zhang, T.; Liang, Z.; et al. Automatic pancreatic ductal adenocarcinoma detection in whole slide images using deep convolutional neural networks. *Frontiers in oncology* **2021**, *11*, 665929. [\[CrossRef\]](#)
271. Li, B.; Keikhosravi, A.; Loeffler, A.G.; Eliceiri, K.W. Single image super-resolution for whole slide image using convolutional neural networks and self-supervised color normalization. *Medical Image Analysis* **2021**, *68*, 101938. [\[CrossRef\]](#)
272. Kugler, M.; Goto, Y.; Kawamura, N.; Kobayashi, H.; Yokota, T.; Iwamoto, C.; Ohuchida, K.; Hashizume, M.; Hontani, H. Accurate 3D reconstruction of a whole pancreatic cancer tumor from pathology images with different stains. In Proceedings of the Computational Pathology and Ophthalmic Medical Image Analysis: First International Workshop, COMPAY 2018, and 5th International Workshop, OMIA 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16–20, 2018, Proceedings 5. Springer, 2018, pp. 35–43. [\[CrossRef\]](#)
273. Kugler, M.; Goto, Y.; Tamura, Y.; Kawamura, N.; Kobayashi, H.; Yokota, T.; Iwamoto, C.; Ohuchida, K.; Hashizume, M.; Shimizu, A.; et al. Robust 3D image reconstruction of pancreatic cancer tumors from histopathological images with different stains and its quantitative performance evaluation. *International journal of computer assisted radiology and surgery* **2019**, *14*, 2047–2055. [\[CrossRef\]](#)
274. Koch, V.; Weitzer, N.; Dos Santos, D.P.; Gruenewald, L.D.; Mahmoudi, S.; Martin, S.S.; Eichler, K.; Bernatz, S.; Gruber-Rouh, T.; Booz, C.; et al. Multiparametric detection and outcome prediction of pancreatic cancer involving dual-energy CT, diffusion-weighted MRI, and radiomics. *Cancer Imaging* **2023**, *23*, 38. [\[CrossRef\]](#)
275. Panda, A.; Garg, I.; Truty, M.J.; Kline, T.L.; Johnson, M.P.; Ehman, E.C.; Suman, G.; Anaam, D.A.; Kemp, B.J.; Johnson, G.B.; et al. Borderline Resectable and Locally Advanced Pancreatic Cancer: FDG PET/MRI and CT Tumor Metrics for Assessment of Pathologic Response to Neoadjuvant Therapy and Prediction of Survival. *American Journal of Roentgenology* **2021**, *217*, 730–740, [<https://doi.org/10.2214/AJR.20.24567>]. PMID: 33084382, <https://doi.org/10.2214/AJR.20.24567>. [\[CrossRef\]](#)
276. Hussein, S.; Kandel, P.; Corral, J.E.; Bolan, C.W.; Wallace, M.B.; Bagci, U. Deep multi-modal classification of intraductal papillary mucinous neoplasms (IPMN) with canonical correlation analysis. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE, 2018, pp. 800–804. [\[CrossRef\]](#)
277. Zhang, Z.; Chen, E.; Zhang, X.; Yang, J.; Wang, X.; Chen, P.; Zeng, M.; Du, M.; Xu, S.; Yang, Z.; et al. Multi-Modal Fusion of Radiomics and Pathomics to Predict the Survival of Pancreatic Cancer Patients Based on Asymmetric Twinning Information Interaction Network. Available at SSRN 4260135. [\[CrossRef\]](#)
278. Chen, X.; Lin, X.; Shen, Q.; Qian, X. Combined spiral transformation and model-driven multi-modal deep learning scheme for automatic prediction of TP53 mutation in pancreatic cancer. *IEEE Transactions on Medical Imaging* **2020**, *40*, 735–747. [\[CrossRef\]](#)
279. Yao, Y.; Chen, Y.; Gou, S.; Chen, S.; Zhang, X.; Tong, N. Auto-segmentation of pancreatic tumor in multi-modal image using transferred DSMask R-CNN network. *Biomedical Signal Processing and Control* **2023**, *83*, 104583. [\[CrossRef\]](#)
280. Li, J.; Qi, L.; Chen, Q.; Zhang, Y.D.; Qian, X. A dual meta-learning framework based on idle data for enhancing segmentation of pancreatic cancer. *Medical Image Analysis* **2022**, *78*, 102342. <https://doi.org/https://doi.org/10.1016/j.media.2021.102342>. [\[CrossRef\]](#)
281. Cai, J.; Zhang, Z.; Cui, L.; Zheng, Y.; Yang, L. Towards cross-modal organ translation and segmentation: A cycle-and shape-consistent generative adversarial network. *Medical image analysis* **2019**, *52*, 174–184. [\[CrossRef\]](#)
282. Cai, J.; Lu, L.; Xing, F.; Yang, L. Pancreas segmentation in CT and MRI images via domain specific network designing and recurrent neural contextual learning. *arXiv preprint arXiv:1803.11303* **2018**. [\[CrossRef\]](#)
283. Asatryan, H.; Gligorievski, A.; Villarini, B. Morphological and multi-level geometrical descriptor analysis in CT and MRI volumes for automatic pancreas segmentation. *Computerized Medical Imaging and Graphics* **2019**, *75*, 1–13. [\[CrossRef\]](#)
284. Puech, P.A.; Boussel, L.; Belfkih, S.; Lemaitre, L.; Douek, P.; Beuscart, R. DicomWorks: software for reviewing DICOM studies and promoting low-cost teleradiology. *Journal of Digital Imaging* **2007**, *20*, 122–130. [\[CrossRef\]](#)
285. Kikinis, R.; Pieper, S.D.; Vosburgh, K.G. 3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support. In *Intraoperative imaging and image-guided therapy*; Springer, 2013; pp. 277–289. [\[CrossRef\]](#)
286. Philbrick, K.A.; Weston, A.D.; Akkus, Z.; Kline, T.L.; Korfiatis, P.; Sakinis, T.; Kostandy, P.; Boonrod, A.; Zeinoddini, A.; Takahashi, N.; et al. RIL-contour: a medical imaging dataset annotation tool for and with deep learning. *Journal of digital imaging* **2019**, *32*, 571–581. [\[CrossRef\]](#)
287. Liu, Y.; Chu, L.; Chen, G.; Wu, Z.; Chen, Z.; Lai, B.; Hao, Y. PaddleSeg: A High-Efficient Development Toolkit for Image Segmentation, 2021, [[arXiv:cs.CV/2101.06175](https://arxiv.org/abs/cs.CV/2101.06175)]. [\[CrossRef\]](#)

288. A easy-to-use, efficient, smart 3D medical image annotation platform. <https://github.com/PaddlePaddle/PaddleSeg/tree/release/2.9/EISeg/med3d>, 2022. 2119
2120
289. Echegaray, S.; Bakr, S.; Rubin, D.L.; Napel, S. Quantitative Image Feature Engine (QIFE): an open-source, modular engine for 3D quantitative feature extraction from volumetric medical images. *Journal of digital imaging* **2018**, *31*, 403–414. 2121
2122
290. Pawłowski, N.; Ktena, S.I.; Lee, M.C.; Kainz, B.; Rueckert, D.; Glocker, B.; Rajchl, M. DLTK: State of the Art Reference Implementations for Deep Learning on Medical Images. *arXiv preprint arXiv:1711.06853* **2017**. 2123
2124
291. Pérez-García, F.; Sparks, R.; Ourselin, S. TorchIO: a Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. *Computer Methods and Programs in Biomedicine* **2021**, p. 106236. 2125
2126
2127
292. Cardoso, M.J.; Li, W.; Brown, R.; Ma, N.; Kerfoot, E.; Wang, Y.; Murray, B.; Myronenko, A.; Zhao, C.; Yang, D.; et al. MONAI: An open-source framework for deep learning in healthcare, 2022, [arXiv:cs.LG/2211.02701]. 2128
2129
293. 3D Medical image Segmentation Solution. <https://github.com/PaddlePaddle/PaddleSeg/tree/release/2.9/contrib/MedicalSeg>, 2022. 2130
2131
294. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861* **2017**. 2132
2133
295. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 6848–6856. 2134
2135
296. Valanarasu, J.M.J.; Patel, V.M. Unext: Mlp-based rapid medical image segmentation network. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2022, pp. 23–33. 2136
2137
297. Ruan, J.; Xiang, S.; Xie, M.; Liu, T.; Fu, Y. MALUNet: A multi-attention and light-weight unet for skin lesion segmentation. In Proceedings of the 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2022, pp. 1150–1156. 2138
2139
298. Ruan, J.; Xie, M.; Gao, J.; Liu, T.; Fu, Y. Ege-unet: an efficient group enhanced unet for skin lesion segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2023, pp. 481–490. 2140
2141
299. Zhang, X.; Wu, C.; Zhao, Z.; Lin, W.; Zhang, Y.; Wang, Y.; Xie, W. Pmc-vqa: Visual instruction tuning for medical visual question answering. *arXiv preprint arXiv:2305.10415* **2023**. 2142
2143
300. He, X.; Zhang, Y.; Mou, L.; Xing, E.; Xie, P. Pathvqa: 30000+ questions for medical visual question answering. *arXiv preprint arXiv:2003.10286* **2020**. 2144
2145
301. Lau, J.J.; Gayen, S.; Ben Abacha, A.; Demner-Fushman, D. A dataset of clinically generated visual questions and answers about radiology images. *Scientific data* **2018**, *5*, 1–10. 2146
2147
302. Gao, W.; Deng, Z.; Niu, Z.; Rong, F.; Chen, C.; Gong, Z.; Zhang, W.; Xiao, D.; Li, F.; Cao, Z.; et al. Ophglm: Training an ophthalmology large language-and-vision assistant based on instructions and dialogue. *arXiv preprint arXiv:2306.12174* **2023**. 2148
2149
303. Zhao, Z.; Liu, Y.; Wu, H.; Li, Y.; Wang, S.; Teng, L.; Liu, D.; Li, X.; Cui, Z.; Wang, Q.; et al. Clip in medical imaging: A comprehensive survey. *arXiv preprint arXiv:2312.07353* **2023**. 2150
2151
304. Tiu, E.; Talius, E.; Patel, P.; Langlotz, C.P.; Ng, A.Y.; Rajpurkar, P. Expert-level detection of pathologies from unannotated chest X-ray images via self-supervised learning. *Nature Biomedical Engineering* **2022**, *6*, 1399–1406. 2152
2153
305. Wu, Y.; Zhou, Y.; Saiyin, J.; Wei, B.; Lai, M.; Shou, J.; Fan, Y.; Xu, Y. Zero-Shot Nuclei Detection via Visual-Language Pre-trained Models. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2023, pp. 693–703. 2154
2155
2156
306. Adhikari, R.; Dhakal, M.; Thapaliya, S.; Poudel, K.; Bhandari, P.; Khanal, B. Synthetic Boost: Leveraging Synthetic Data for Enhanced Vision-Language Segmentation in Echocardiography. In Proceedings of the International Workshop on Advances in Simplifying Medical Ultrasound. Springer, 2023, pp. 89–99. 2157
2158
2159
307. Eslami, S.; Meinel, C.; De Melo, G. Pubmedclip: How much does clip benefit visual question answering in the medical domain? In Proceedings of the Findings of the Association for Computational Linguistics: EACL 2023, 2023, pp. 1181–1193. 2160
2161
308. Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A.C.; Lo, W.Y.; et al. Segment anything. *arXiv preprint arXiv:2304.02643* **2023**. 2162
2163
309. Zhang, K.; Liu, D. Customized segment anything model for medical image segmentation. *arXiv preprint arXiv:2304.13785* **2023**. 2164
310. Wu, J.; Fu, R.; Fang, H.; Liu, Y.; Wang, Z.; Xu, Y.; Jin, Y.; Arbel, T. Medical sam adapter: Adapting segment anything model for medical image segmentation. *arXiv preprint arXiv:2304.12620* **2023**. 2165
2166
311. Ye, J.; Cheng, J.; Chen, J.; Deng, Z.; Li, T.; Wang, H.; Su, Y.; Huang, Z.; Chen, J.; Jiang, L.; et al. Sa-med2d-20m dataset: Segment anything in 2d medical imaging with 20 million masks. *arXiv preprint arXiv:2311.11969* **2023**. 2167
2168
312. Jia, X.; Ren, L.; Cai, J. Clinical implementation of AI technologies will require interpretable AI models. *Medical physics* **2020**, pp. 1–4. 2169
2170
313. Van der Velden, B.H.; Kuijff, H.J.; Gilhuijs, K.G.; Viergever, M.A. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Medical Image Analysis* **2022**, *79*, 102470. 2171
2172
314. Decathlon, M.S. Medical Segmentation Decathlon. <http://medicaldecathlon.com/>. 2173
315. GitHub. <https://github.com>. 2174
316. Challenge, G. Grand Challenge. <https://grand-challenge.org/>. 2175
317. Bionetworks, S. Synapse. <https://www.synapse.org/>. 2176
318. Zenodo. Zenodo Repository. <https://zenodo.org/>. 2177

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual ²¹⁷⁸ author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to ²¹⁷⁹ people or property resulting from any ideas, methods, instructions or products referred to in the content. ²¹⁸⁰