Appendix 1

Ethological Analysis

All tissue samples were reviewed by a pathologist (J.C.Z) with 11 years of experience in pancreatic pathology, blinded to other clinical and imaging results. Inflammatory lesion subtypes were distinguished by inflammatory cells, stromal reaction, and vascular response, whereas tumor lesion subtypes were determined based on tissue morphology and cellular atypia. Additional histopathological evaluations have also been completed, including features such as nuclear enlargement and increased mitotic activity in tumorous lesions, as well as tissue edema, necrosis, and cellular infiltration in inflammatory lesions, along with abnormal enhancement in both.

Lesion and Pancreas Annotation

For the annotation of multicenter datasets, eight medical professionals (H.W., K.J.W., J.F.Y., L.Y., W.X.H., L.H., M.F.S., R.D.) manually segmented the contours of normal pancreases, which were then reviewed and further annotated for simple inflammatory contours by two radiologists (Y.P.H and K.Q) with 4–5 years of experience. Two radiologists (B.Y. and Q.Y.) with 9–12 years of experience manually segmented the contours of pancreases with tumor lesions.

Abdominal radiologists employed the medical image annotation tool LabelMe (https://github.com/jiangjiawen/labelme-for-dicom) to outline the entire pancreas, including both the parenchyma and lesions, across axial slices of enhanced and non-enhanced CT scans, disregarding the patients' clinical information. If the contours were unclear on non-enhanced CT scans, the radiologists relied on lesion information from the enhanced CT images for delineation.

To eliminate the impact of intra-observer and inter-observer variability, all multicenter datasets image was segmented by different doctors during the experiment, and the Dice coefficient was calculated for each image. If the difference in the Dice coefficient exceeded 0.05, the contour was reviewed by the adjudicating doctor, L.W., who has 17 years of experience in diagnosing pancreatic tumors, to determine the true contour size.

Model Device

These models were developed using an open-source framework named Pytorch 1.4.1 on Python 3.8. The input images were loaded onto two devices for training: one was an RTX3060 (Ubuntu 14.04; 3060Ti [32GB]; Institute of Medical Imaging, Department of Radiology, Taihe Hospital, Hubei University of Medicine), and the other was an RTX4060 (Ubuntu 14.04; 4060Ti [64GB]; Institute for Network Sciences and Cyberspace, Tsinghua University, Beijing, China]. Both devices were equipped with an 11th Gen Intel i7-11700 processor (2.50 GHz) and an NVIDIA GeForce GPU with CUDA 11.6.0.

Pancreas Localization

The goal of the first stage was to segment the pancreas. In CT scans, the pancreas was usually located among the hollow organs in the center of the image, with an irregular contour and a wide variety of complex diseases. Given the capability of atrous convolution and atrous spatial pyramid pooling to handle fine features of different sizes, a DeepLabV3 model was trained in this stage to segment the entire pancreatic region (29). The goal was to obtain accurate pancreatic contours in CT scans of normal, inflammatory pancreatic lesions, and tumor-like pancreatic lesions, as shown in *Figure 2B*.

Lesion Segmentation

The purpose of the second stage was to segment lesions within the pancreas. As the contours of lesions were more irregular and diverse than those of the pancreas, and the gray values of some low-density lesions were similar to those of necrotic lesions and signs of pancreatic duct dilation, deeper networks and enhanced detail-capturing capabilities were required at this stage to improve sensitivity and specificity, thereby enabling the handling of more complex classification environments. nnUNet-MS, due to its adaptive parameter characteristics, was able to automatically adjust the depth and architecture of the

 https://dx.doi.org/10.21037/qims-2025-aw-2192

network according to different lesion contours to meet the segmentation needs of complex pancreatic lesions (30,31). Based on nnUNet-MS, the parameters of the feature extractor were fixed to optimize feature extraction at different scales, thereby achieving good lesion segmentation results, as depicted in *Figure 2C*.

Statistical Analysis

Two metrics, Dice Similarity Coefficient (Dice) and Intersection over Union (IoU), were used to evaluate segmentation. These metrics measure the difference between the ground truth and the algorithm's predictions. Dice is defined as:

$$\text{Dice} = \frac{2 \times TP}{FP + 2 \times TP + FN}$$

where TP (true positive), TN (true negative), FP (false positive), and FN (false negative) represent the correctly predicted positive target regions, correctly predicted negative target regions, incorrectly predicted positive target regions, and incorrectly predicted negative target regions, respectively.

$$\text{IoU} = \frac{TP}{FP + TP + FN}$$

Model diagnostic evaluation considered four metrics: TP, TN, FP, and FN, representing the counts of true positives, true negatives, false positives, and false negatives, respectively. Accuracy is defined as:

$$\text{Acc} = \frac{TP + TN}{TP + FP + TN + FN}$$

Results of Pancreas Segmentation

A lightweight DeepLabV3 model was used to segment the pancreatic contour in CT scans. In internal and external test sets, DeepLabV3 was demonstrated to perform well (internal test set, Dice: 0.983; IoU: 0.971; external test set, Dice: 0.981; IoU: 0.969).

References

29. Chen LC, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. arXiv: 1802.02611.
30. Isensee F, Petersen J, Klein A, Zimmerer D, Jaeger PF, Kohl S, Wasserthal J, Koehler G, Norajitra T, Wirkert S, Maier-Hein KH. nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation. arXiv:1809.10486.
31. Isensee F, Jaeger PF, Kohl SAA, Maier-Hein JPKH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods 2021;18:203-11.

https://dx.doi.org/10.21037/qims-2025-aw-2192