



ResDAC-Net: a novel pancreas segmentation model utilizing residual double asymmetric spatial kernels

Zhanlin Ji¹ · Jianuo Liu¹ · Juncheng Mu¹ · Haiyang Zhang² · Chenxu Dai¹ · Na Yuan³ · Ivan Ganchev^{4,5,6} 

Received: 8 October 2023 / Accepted: 13 February 2024 / Published online: 8 March 2024
© The Author(s) 2024

Abstract

The pancreas not only is situated in a complex abdominal background but is also surrounded by other abdominal organs and adipose tissue, resulting in blurred organ boundaries. Accurate segmentation of pancreatic tissue is crucial for computer-aided diagnosis systems, as it can be used for surgical planning, navigation, and assessment of organs. In the light of this, the current paper proposes a novel Residual Double Asymmetric Convolution Network (ResDAC-Net) model. Firstly, newly designed ResDAC blocks are used to highlight pancreatic features. Secondly, the feature fusion between adjacent encoding layers fully utilizes the low-level and deep-level features extracted by the ResDAC blocks. Finally, parallel dilated convolutions are employed to increase the receptive field to capture multiscale spatial information. ResDAC-Net is highly compatible to the existing state-of-the-art models, according to three (out of four) evaluation metrics, including the two main ones used for segmentation performance evaluation (i.e., DSC and Jaccard index).

Keywords ResDAC-Net · Pancreatic segmentation · Image segmentation · Medical image processing

1 Introduction

Pancreatic cancer is a highly malignant disease that is difficult to diagnose early and progresses rapidly [1–3]. Therefore, accurate pancreatic segmentation can effectively improve the diagnostic accuracy of pancreatic cancer and help doctors assess treatment progress. Doctors need to meticulously delineate

patients' medical images, converting 2D slices into 3D volumetric images in their minds [4]. This process allows them to understand the spatial relationships between pathological tissues and surrounding organs. It is essential for doctors to possess substantial experience in image interpretation. However, the presence of imaging artifacts, such as pseudo-shadows and gas artifacts, introduces noise interference that can potentially impact doctors' assessments [5–7]. Additionally, different doctors may yield subjective and inconsistent diagnostic results, prone to misjudgment due to fatigue or insufficient experience. On the other hand, automatic pancreatic segmentation is a challenging task due to several reasons [8]. Firstly, the pancreas has a relatively small volume in the abdomen, requiring a considerable amount of non-target image information for organ segmentation [9]. Secondly, the shape and boundaries of the pancreas in computed tomography (CT) images are highly irregular and variable, even within adjacent 3D slices, exhibiting significant diversity [10]. Lastly, the pancreas exhibits similar intensity to surrounding organs, making it difficult to distinguish it using image enhancement techniques such as contrast, leading to unclear slice boundaries [11]. The application of deep learning techniques has shown promising results in addressing these pancreatic segmentation problems by offering a potential solution for quick and accurate extraction of pancreas regions in medical images [12, 13].

✉ Na Yuan
yuanna@tsc.edu.cn

✉ Ivan Ganchev
ivan.ganchev@ul.ie

¹ Department of Artificial Intelligence, North China University of Science and Technology, Tangshan 063009, China

² Department of Computing, Xi'an Jiaotong-Liverpool University, Suzhou, China

³ Intelligence and Information Engineering College, Tangshan University, Tangshan 063000, China

⁴ Telecommunications Research Centre (TRC), University of Limerick, Limerick V94 T9PX, Ireland

⁵ Department of Computer Systems, University of Plovdiv "Paisii Hilendarski", Plovdiv 4000, Bulgaria

⁶ Institute of Mathematics and Informatics—Bulgarian Academy of Sciences, Sofia 1040, Bulgaria

In recent years, 2D deep learning has been widely applied to pancreatic segmentation. These methods split the 3D volume data of CT images into 2D slices and perform segmentation on individual slices through pixel-wise classification. The U-Net model, proposed by Ronneberger et al. [14], is widely used for pixel-level classification tasks in medical images. However, the labeling of medical image datasets requires annotations from medical experts, resulting in a scarcity of clinically meaningful labeled datasets [15]. The U-Net model reduces the number of parameters by dynamically adjusting the number of channels and model depth for each layer [16]. Therefore, information from each scale of the image is meaningful [17]. The skip connections in the U-Net model address the issue of high-resolution feature loss during the up-sampling process, thereby enhancing the expression capability of local and global features [18, 19]. Expanding on this groundwork, researchers have put forth several enhancement strategies that integrate pancreatic features, aiming to further augment the effectiveness of segmentation [20].

Okta et al. [21] integrated attention gates into the U-Net model, proposing the Attention U-Net architecture. The attention gates can suppress irrelevant parts of the model while enhancing the learning of task-relevant features [22]. This approach abandons the coarse-to-fine segmentation framework, which performed well in pancreatic segmentation, and eliminates the need for cascaded networks to achieve coarse segmentation and localization of organs. As a result, this approach performs well on various datasets. Cai et al. [23] utilized Convolutional Long Short-Term Memory (CLSTM) modules to address the problem of learning temporal features in pancreatic CT slice data. The CLSTM module achieves spatially consistent segmentation of individual slices for label predictions and further optimizes Convolutional Neural Network (CNN)-based segmentation results by considering the context features of the pancreas [24]. Heinrich et al. [25] proposed a high-performance ternary network architecture aiming to reduce reliance on computer memory [22].

In the current paper, we propose a pancreatic segmentation model, called ResDAC-Net, which utilizes a combination of residual networks and asymmetric convolutions to construct its backbone. The specific contributions of our work are as follows:

1. Based on the U-Net architecture, we replace two standard 3×3 convolutional blocks with newly designed Residual Double Asymmetric Convolution (ResDAC) blocks, which contain residual networks and asymmetric convolutions. This substitution enhances feature fitting, preserves more feature information in the target region, and alleviates issues like vanishing and exploding gradients.
2. We strengthen feature fusion between adjacent encoding layers through additional skip connections, thus fully utilizing the low-level and deep-level features extracted by the ResDAC blocks.

3. We employ parallel dilated convolutions with different dilation rates in a novel feature fusion PDC module, to increase the receptive field between encoding and decoding layers.
4. Inspired by ExFuse [26], we propose a new Adjacent Layer Feature fusion (ALF) block that performs feature fusion on additional skip connections between adjacent layers.

2 Related work

2.1 ResNet

CNNs are a popular topic in the field of image processing, particularly widely used in image segmentation. Among them, ResNet (Residual Network) is a special type of CNN architecture [27]. Compared to conventional CNN architectures, ResNet introduces residual units with identity mappings [28]. In traditional deep neural networks, as the number of layers increases, issues like vanishing or exploding gradients can occur, making it difficult for the network to converge [29, 30]. To address this problem, ResNet proposes the concept of residual units [31]. These units allow the network to directly learn the shallow representation of the data and then focus on learning the differences in deeper layers.

Specifically, ResNet achieves this by incorporating shortcut connections, also known as skip connections, which add the shallow input features to the deep output features. This enables the deep layers to directly reference information from the shallow layers, a connection referred to as “identity mapping” [32]. As a result, ResNet possesses enhanced learning capabilities and has demonstrated outstanding performance in various types of image segmentation tasks. Its essence lies in the fusion of two types of features from different receptive fields, using the fused features as the learning target, rather than simply relying on the output features from the previous layer [28].

2.2 U-Net

U-Net is a deep learning network architecture used for image segmentation, particularly well-suited for biomedical image segmentation tasks. The design of U-Net draws inspiration from its U-shaped network structure, consisting of a down-sampling path (encoder) and an up-sampling path (decoder) [33]. The down-sampling path is similar to common CNNs. It gradually reduces the spatial dimensions of the image through a series of operations like convolution and pooling, while simultaneously extracting abstract features [34].

The up-sampling path in U-Net is different from traditional CNNs. In this path, U-Net employs operations like up-sampling to gradually restore low-resolution feature maps to prediction maps with the same size as the input

image. Additionally, to fuse feature information from different scales, U-Net connects the features from the down-sampling path to the up-sampling path. This cross-layer connection helps the model better localize segmentation targets and preserve detailed information. The network architecture of U-Net allows it to perform effective image segmentation even with limited data, leading to remarkable success in the field of biomedical image segmentation.

2.3 ACNet

ACNet is an improved CNN that enhances image recognition capabilities by introducing asymmetric convolutional kernels [35]. In traditional convolution operations, the kernels are symmetric. For example, a 3×3 convolutional kernel has the same weights in the horizontal and vertical directions. However, in real-world images, features in different directions may have varying importance. To better capture these features, the asymmetric convolutional network introduces asymmetric convolutional kernels, designed to have different weights in the horizontal and vertical directions. With this design, the network can better distinguish features in different directions within the image and more effectively capture directional information. Such asymmetric design enables the network to achieve better performance in image recognition tasks and provides new solutions for image segmentation tasks [36–38].

3 Proposed model

3.1 Residual double asymmetric convolution (ResDAC) block

In the usage of CNNs, when images are flipped or rotated, the image features extracted by standard square convolutions may change, leading to biases in the model's segmentation results for the same object and reducing the model's generalization ability. To address the issues associated with standard square convolution models, such as decreased accuracy due to parameter reduction, underfitting with separable convolutions, and model ineffectiveness at small input sizes, we propose a novel convolutional block, called ResDAC, based on spatial kernels. ResDAC aims to overcome the challenges posed by standard square convolutions by introducing a double asymmetric convolution approach. This new approach effectively addresses the decrease in accuracy caused by parameter reduction, mitigates underfitting issues with separable convolutions, and ensures the model's effectiveness even at small input sizes.

ResDAC overcomes the problem of decreased accuracy when using standard square convolutions by combining residual networks with asymmetric convolutions. In this

design, we connect two asymmetric convolutions using a residual network approach and incorporate the Squeeze-and-Excitation (SE) attention mechanism [39]. This allows us to capture channel dependencies after feature fusion, highlight important feature dimensions in the pancreas, and achieve superior accuracy while reducing the number of model's parameters. This innovation liberates the model from the limitations of accuracy loss when dealing with standard square convolutions and effectively avoids overfitting and gradient explosion issues that may arise from deep network stacking. By introducing the ResDAC block, we aim to provide a more efficient and accurate solution, fully leveraging the advantages of spatial kernel-based convolution models and optimizing their performance across various tasks.

The elaborated ResDAC block takes input in the format of particular height (H) $\times$ width (W) $\times$ input channels (C_{in}) and outputs in the format of height (H) $\times$ width (W) $\times$ output channels (C_{out}). The block consists of one standard square convolutional layer (with a 3×3 kernel) and two asymmetric convolutional layers (with 3×3 kernels), uniquely forming a residual network convolutional block for discriminative feature extraction. In addition, the last layer also incorporates an SE block, as shown in Fig. 1.

Initially, the input feature map of size $H \times W \times C_{in}$ is fed into a 3×3 2D convolution. Let x_i denote the input sample, where i is the index of the convolutional layer. Then, the output after the convolution of the first layer is calculated similarly to [36], as follows:

$$P_{i=1} = [BNReLU(x_{i=1} * \omega_{i=1}^{k \times k} + b)]_{k=3} \quad (1)$$

where $\omega_{i=1}^{k \times k}$ denotes the kernel weights, k denotes the size of the convolutional kernel, b denotes the bias term, BNReLU refers to the combined operation of Batch Normalization (BN) and Rectified Linear Unit (ReLU) activation, symbol $*$ denotes the convolution operation, and $P_{i=1}$ denotes the output feature map. In contrast to the initial layer, the second and third convolutional layers are composed of asymmetric convolutions with a kernel size of $k = 3$. The convolutional output for these layers is calculated as follows:

$$P_{i=2} = [BNReLU(x_{i=2} * \omega_{i=2}^{k \times k} + x_{i=2} * \omega_{i=2}^{1 \times k} + x_{i=2} * \omega_{i=2}^{k \times 1} + b)]_{k=3} \quad (2)$$

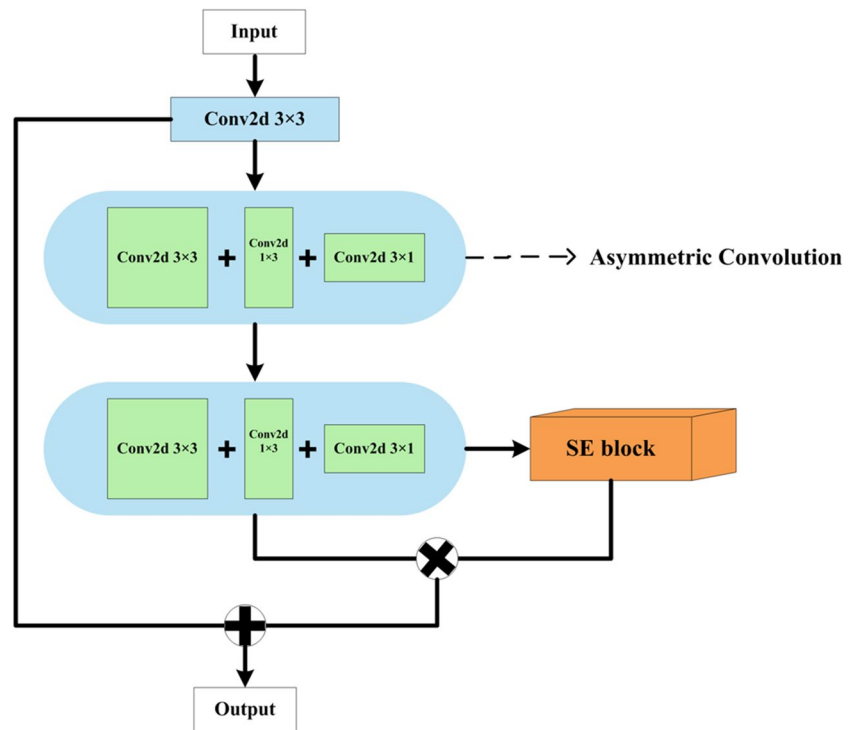
$$P_{i=3} = [BNReLU(x_{i=3} * \omega_{i=3}^{k \times k} + x_{i=3} * \omega_{i=3}^{1 \times k} + x_{i=3} * \omega_{i=3}^{k \times 1} + b)]_{k=3} \quad (3)$$

The SE block performs the following operations:

$$A_{sq} = \frac{1}{H \times W} x_{i=4} \quad (4)$$

$$A_{ex} = \sigma[W_2 * (W_1 * A_{sq})] \quad (5)$$

Fig. 1 The elaborated ResDAC block



$$P_{i=4} = A_{ex} \times x_{i=4} \quad (6)$$

where A_{sq} denotes global average pooling on the input sample $x_{i=4}$ for the generation of a $1 \times 1 \times C$ feature vector; A_{ex} is obtained through two fully connected layers (W_1 and W_2), followed by a Sigmoid operation σ ; and $P_{i=4}$ performs element-wise multiplication (denoted by $\times$) of the generated feature vector A_{ex} and the weights from each of the $H \times W$ values in the input sample $x_{i=4}$ for each channel.

Then, the outputs of the last convolutional layer and the SE block are element-wise multiplied, and the result is element-wise added to the output of the first convolutional layer to strengthen feature extraction and obtain discriminative representations. So, the ResDAC block's output P_{out} is obtained as follows:

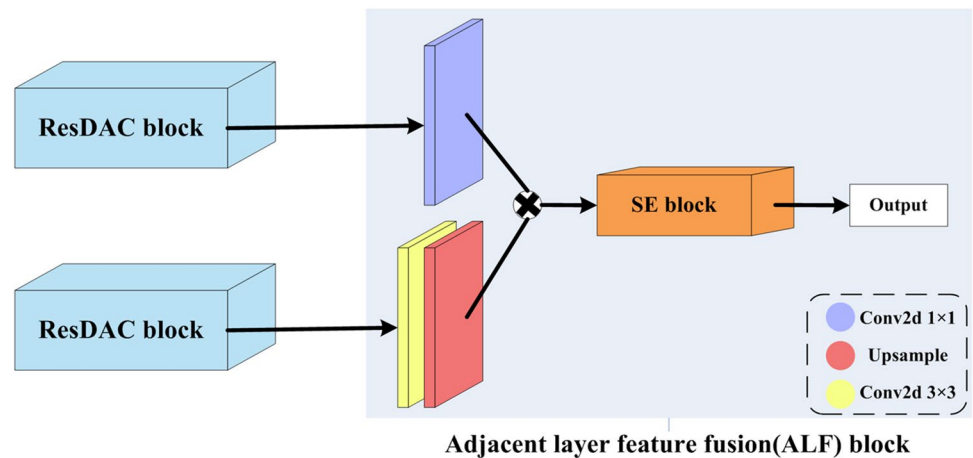
$$P_{out} = P_{i=3} \times P_{i=4} + P_{i=1} \quad (7)$$

By combining the outputs of multiple different convolutional layers through element-wise multiplication and addition, further feature fusion is achieved in generating the ResDAC output. The effectiveness of this asymmetric residual block lies in its ability to circumvent issues such as overfitting and gradient explosion, which may arise from the stacking of deep networks. Simultaneously, it attains superior performance.

3.2 Adjacent Layer Feature fusion (ALF) block

We have also designed a novel feature fusion block, called ALF, which fuses the outputs of adjacent convolutional blocks during the encoding stage to refine and enrich the feature maps. This block utilizes a 3×3 2D convolution to transform the high-level feature maps into feature maps at the same level as the previous layer, facilitating the extraction of detailed information from the high-level features and improving the capturing of pancreatic boundary features. For the low-level feature maps, a 1×1 2D convolution is applied to introduce non-linearity and expand the model's options. This is why a “deep” network is preferable to a “wide” network [36], as the non-linear layer expands the model's capacity. The two processed feature maps are then fused through element-wise multiplication, and an SE block is employed for implementing an attention mechanism as to explicitly model the interdependence between channels and adaptively recalibrate the channel-wise characteristics, thus enhancing the representation capability of the convolutional block for fused information. This improves boundary misalignment and forms an effective attention-based feature fusion execution framework. The structure of the elaborated ALF block is shown in Fig. 2.

Fig. 2 The elaborated ALF block



3.3 Parallel dilated convolutions (PDC) block

Compared to standard convolutions, dilated convolutions, also known as atrous convolutions, can enlarge the receptive field of the convolutional kernel by changing the dilation rate, thereby capturing multi-scale information from the feature maps. When the dilation rate is set to 1, the receptive field size of the dilated convolution is the same as that of the standard convolution. However, when the dilation rate is greater than 1, dilated convolutions can obtain a larger receptive field size and capture richer image information than standard convolutions [40]. In traditional convolution operations, a single convolutional kernel can only capture limited ranges of input features, thus restricting the network's ability to understand large-scale information. Therefore, parallel dilated convolutions, through clever design, allow multiple convolutional kernels to parallelly span different receptive field sizes, providing a broader perception range. Parallel dilated convolutions introduce multiple parallel branches within the convolutional kernel, each with a different dilation rate, which allows to control the sampling span of the kernel when taking input. By using parallel convolutional kernels with different dilation rates, the network can capture multi-scale input features, including both local and global ones. Furthermore, compared to traditional dilated convolutions, parallel dilations can reduce the number of parameters and computational complexity of a model to some extent. This is because they share some convolutional layer parameters, and the computations of multiple branches can be executed in parallel. The calculation of the receptive field size for dilated convolutions is done according to the following formula:

$$M_i = m_i + (k_i - 1) \prod_{j=1}^i s_j \quad (8)$$

where M_i denotes the receptive field after dilation for the i -th layer, m_i denotes the actual receptive field for the i -th layer,

k_i denotes the kernel size used in the dilated convolution for the i -th layer, s_i denotes the offset of a convolutional kernel expanding outward, and $\prod_{j=1}^i s_j$ denotes the cumulative offset at which each convolutional kernel expands outward. The structure of the designed PDC block is illustrated in Fig. 3.

3.4 ResDAC-Net

The proposed ResDAC-Net model consists of three core types of blocks, described in the previous subsections, as detailed in Fig. 4. The U-shaped encoder-decoder form of U-Net is adopted as a basic backbone. The encoder performs down-sampling on the input image and extracts essential features from the image. The encoder path is composed of a series of ResDAC blocks, followed by max-pooling. The output of a ResDAC block is combined with a 2×2 max-pooling to form an encoding unit (down-sampling unit) as follows:

$$E_j = \text{Maxpool}(P_{out}, 2) \quad (9)$$

where E_j denotes the output of the j -th encoding layer, Maxpool denotes the max-pooling operation, and P_{out} denotes the output of the j -th ResDAC block.

Similarly to the encoder path, the decoder path is also composed of a series of ResDAC blocks for feature extraction. Additionally, before the skip connection, an up-sampling operation is applied to further refine the output features. The up-sampling unit with a scale of 2 is used to up-sample the feature map received from the lower layers of the network, as shown below:

$$up_j = \text{Upsample}(x_{j-1}, 2) \quad (10)$$

The up-sampled output, denoted as up_j , is further feature fused with the output of the ALF block at the

Fig. 3 The elaborated PDC block

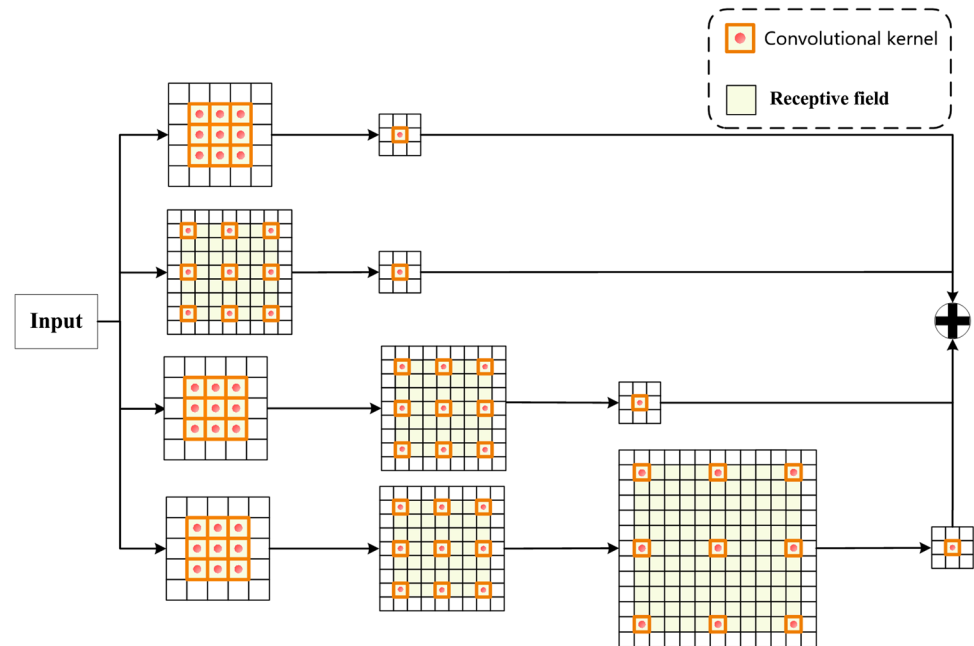
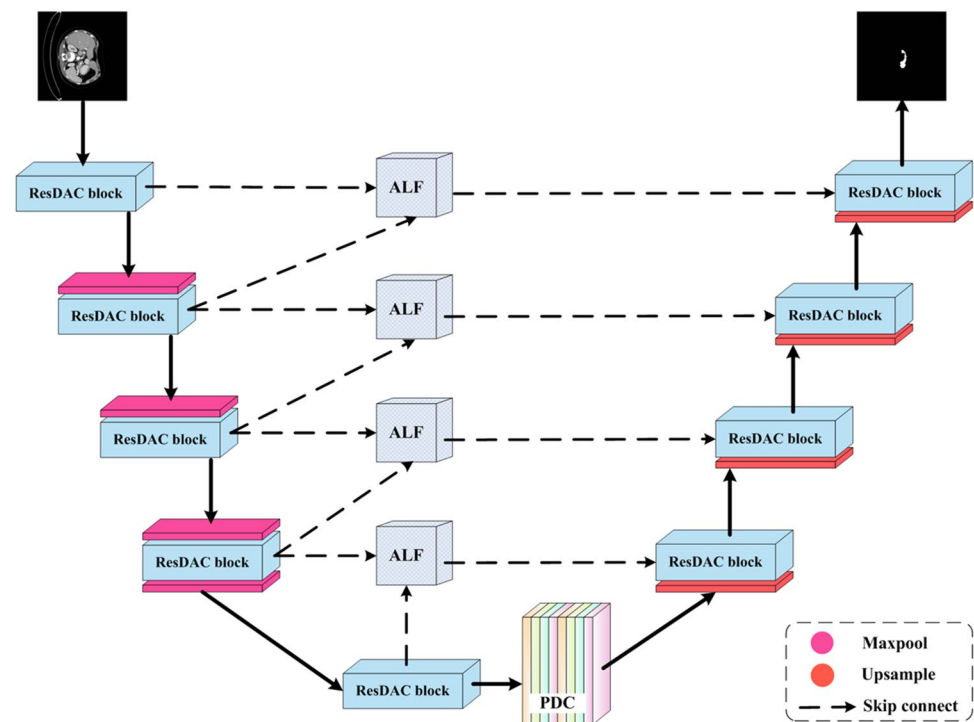


Fig. 4 The proposed ResDAC-Net model



corresponding encoder layer through a skip connection. By copying the feature maps from the encoder, the model can capture multi-scale information. The skip connections allow utilization of both high-resolution and low-resolution information from the feature maps during the up-sampling process [36], as follows:

$$D_j = P_{out} \{ \text{concat} [up_j, ALF(E_j, E_{j-1})] \} \quad (11)$$

where D_j denotes the output of the j -th decoder layer, concat denotes the skip connection, $ALF(E_j, E_{j-1})$ denotes the output of the feature fusion through the ALF block between the j -th

and $(j - 1)$ -th encoder layers, and P_{out} denotes the output after the convolution with the ResDAC block.

3.5 Training and optimization methods

The proposed ResDAC-Net model utilizes a composite BCE-Dice loss and determines appropriate weight factors based on experimental results. The BCE-Dice loss [41] combines the binary cross-entropy (BCE) loss [42] and Dice loss [43], which are commonly used for binary segmentation tasks. The BCE loss is defined in [44], as follows:

$$BCE(p, q) = -\frac{1}{N} \sum_{i=1}^N (y_i \log(p_i) + (1 - y_i) \log(1 - p_i)) \quad (12)$$

where N denotes the number of pixels, y_i denotes the actual label (0 or 1) of the i -th pixel, and p_i denotes the predicted probability of the i -th pixel belonging to a class [45].

The Dice loss is calculated as follows:

$$Dice(p, q) = \frac{2 \sum_{i=1}^N p_i \cdot q_i}{\sum_{i=1}^N p_i^2 + \sum_{i=1}^N q_i^2} \quad (13)$$

where q_i denotes the target label of the i -th pixel, which is the binarized ground-truth label.

The composite BCE-Dice loss is calculated as follows:

$$BCEDice(p, q) = \alpha \cdot BCE(p, q) + (1 - \alpha) \cdot Dice(p, q) \quad (14)$$

where α denotes the weight factor.

As an optimizer, the Adam optimizer [46], which adapts the learning rate based on the first and second moment estimates of the gradients, was used, as it converges quickly

during training and avoids getting stuck in local minima. In each iteration step t ($t = \{1, 2, \dots, T\}$), the following computations are performed:

$$m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t \quad (15)$$

$$v_t = \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2 \quad (16)$$

where g_t denotes the gradient at the current iteration step t , m_t denotes the exponentially decaying average of the gradients, v_t denotes the exponentially decaying average of the squared gradients (uncentered variance), and β_1 and β_2 denote the decay rates. Then, the bias-corrected estimates are calculated as follows:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (17)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (18)$$

The parameters are updated as follows:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} \cdot \hat{m}_t \quad (19)$$

where η denotes the learning rate, ϵ denotes a small constant used to prevent division by 0, and θ_{t+1} denotes the updated parameter.

The hardware configuration used in the conducted experiments, described in the next section, includes an Intel Core i5-12490 processor with a clock speed of 3.0 GHz, and a single NVIDIA RTX3060 graphics card with 12-GB memory. The neural model's hyperparameters for training were

Table 1 Ablation study results on NIH dataset

Models	DSC (%)	Precision (%)	Recall (%)	Jaccard index (%)
U-Net (<i>baseline</i>)	78.69	77.88	80.88	65.66
U-Net + ResDAC	82.42	81.77	84.04	70.75
U-Net + ResDAC + ALF	83.89	83.58	85.10	72.74
U-Net + ResDAC + ALF + PDC (i.e., ResDAC-Net)	84.56	81.69	88.48	73.70

The best results are shown in bold

Table 2 Ablation study results on MSD dataset

Models	DSC (%)	Precision (%)	Recall (%)	Jaccard index (%)
U-Net (<i>baseline</i>)	64.36	69.26	66.78	53.40
U-Net + ResDAC	67.07	71.68	69.05	56.64
U-Net + ResDAC + ALF	68.21	69.41	71.66	57.77
U-Net + ResDAC + ALF + PDC (i.e., ResDAC-Net)	69.15	71.10	72.31	58.76

The best results are shown in bold

Table 3 Pancreas segmentation performance comparison results obtained on NIH dataset

Models	DSC (%)	Precision (%)	Recall (%)	Jaccard index (%)
U-Net	78.69	77.88	80.88	65.66
Attention U-Net	79.77	80.64	80.60	67.25
U-Net++	79.41	73.49	88.07	66.62
ResUNet	77.09	81.37	74.87	63.79
ResNet	78.52	78.15	80.61	65.80
ResDAC-Net (ours)	84.56	81.69	88.48	73.70

The best results are shown in bold

set as follows: Batch_Size=4, Epochs=100 (validation was performed on each epoch, and the model was trained using the Adam optimizer), Initial_Learning_Rate= 1×10^{-4} , momentum=0.9, Minimum_Learning_Rate= 1×10^{-5} . The network structure is implemented by using PyTorch.

4 Experiments and results

4.1 Datasets and evaluation metrics

CT image slices from the pancreas CT dataset of the National Institutes of Health (NIH) Clinical Center Cancer Imaging Archive [47–49] and the Medical Segmentation Decathlon (MSD) dataset [50] were used in the conducted pancreatic segmentation experiments, since these are mainstream datasets for pancreatic organ segmentation.

The NIH dataset contains 82 abdominal contrast-enhanced CT scans. The ground-truth annotations for the pancreatic region in each CT scan were provided by experienced radiologists. The CT scan images are stored in the Digital Imaging and Communications in Medicine (DICOM) medical image format. Each CT scan has an original size of 512×512 pixels, and the number of slices varies between 181 and 466 for different patients. The slice thickness ranges from 1.5 to 2.5 mm. The scans were performed using Philips and Siemens multidetector computed tomography (MDCT) machines. For the experiments, only the CT scans, containing a pancreas image, were selected. To ensure a fair comparison, this dataset was divided into 60 individuals' slices for the training set, whereas the validation and test sets contained slices from 11 individuals each.

The MSD dataset contains 281 abdominal contrast-enhanced CT scans. Each CT scan has an original size of 512×512 pixels. The slice thickness ranges from 0.7 to 7.5 mm. For the experiments, only the CT scans, containing a pancreas image, were selected. To ensure a fair comparison, this dataset was split in a 3:1:1 ratio for creating the training, testing, and validation sets.

In order to objectively evaluate the performance, training was conducted on the same dataset while keeping certain parameters constant. Common metrics such as Dice

similarity coefficient (DSC), precision, recall, and Jaccard index were used to evaluate the results. These metrics are defined as follows:

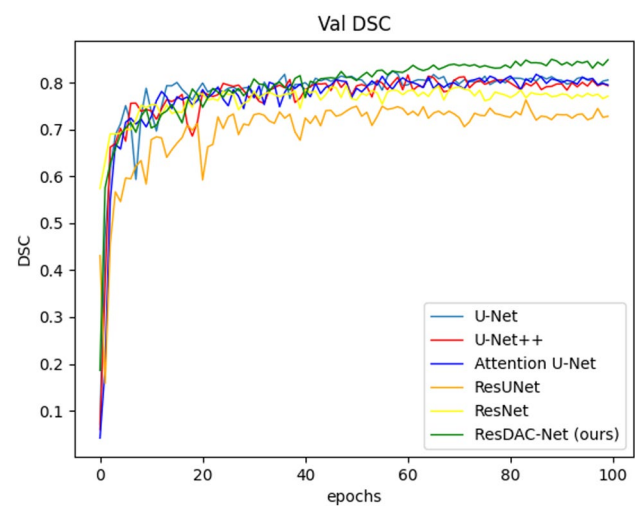


Fig. 5 The DSC validation curves of the compared models

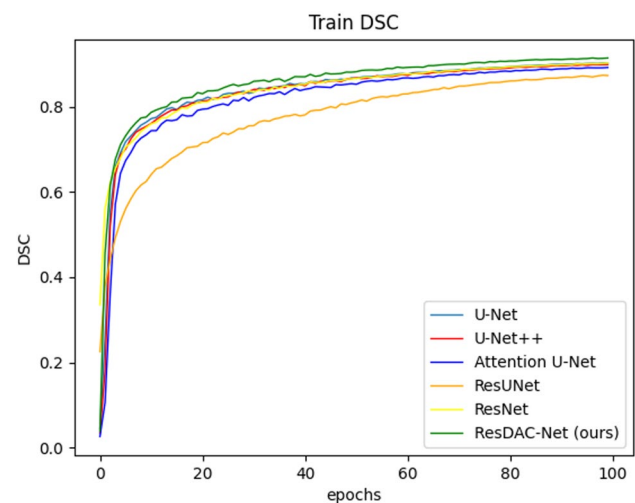


Fig. 6 The DSC training curves of the compared models

$$\text{DSC} = \frac{2TP}{2TP + FP + FN} \quad (20)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (21)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (22)$$

$$\text{Jaccard index} = \frac{TP}{TP + FN + FP} \quad (23)$$

where TP denotes the true positive counts, FP denotes the false positive counts, and FN denotes the false negative counts. These selected metrics provide a comprehensive evaluation of the segmentation results, enabling a fair comparison of different models on the same dataset.

4.2 Ablation study experiments

To evaluate the effectiveness of different blocks incorporated into the baseline (U-Net), ablation study experiments were conducted on the NIH dataset first and then repeated on the MSD dataset. In these experiments, the elaborated ResDAC blocks, ALF blocks, and PDC block were subsequently added one by one to U-Net, by keeping the blocks added in the previous steps. The obtained results, presented in Tables 1 and 2, show gradual improvements in segmentation performance in each step, compared to the previous step, based on three (out of four) metrics, including the two main ones used for segmentation performance evaluation (i.e., DSC and Jaccard index). The

Table 4 Pancreas segmentation performance comparison results obtained on MSD dataset

Models	DSC (%)	Precision (%)	Recall (%)	Jaccard index (%)
U-Net	64.36	69.26	66.78	53.40
Attention-U-Net	64.80	80.64	80.60	67.25
UNet++	64.66	73.49	88.07	66.62
ResUNet	62.23	68.21	63.32	50.30
ResNet	64.09	67.93	67.34	52.75
ResDAC-Net (ours)	69.15	71.10	72.31	58.76

The best results are shown in bold

only exception was precision, where the second step performed on the NIH dataset, resp. first step on MSD dataset, achieved a better result than the final step.

More specifically, in the first step performed on the NIH (resp. MSD) dataset, adding ResDAC blocks to the baseline model allowed to increase DSC by 3.73 (resp. 2.71) percentage points, precision by 3.89 (resp. 2.42) percentage points, recall by 3.16 (resp. 2.27) percentage points, and Jaccard index by 5.09 (resp. 3.24) percentage points. This indicates that the ResDAC blocks enable effective utilization of channel information from different receptive fields and adaptive refinement of encoded features to retain multi-scale information. Additionally, these blocks allowed the model to better distinguish features in different directions and capture specific directional information more effectively. In the second step, the addition of ALF blocks allowed to further improve the

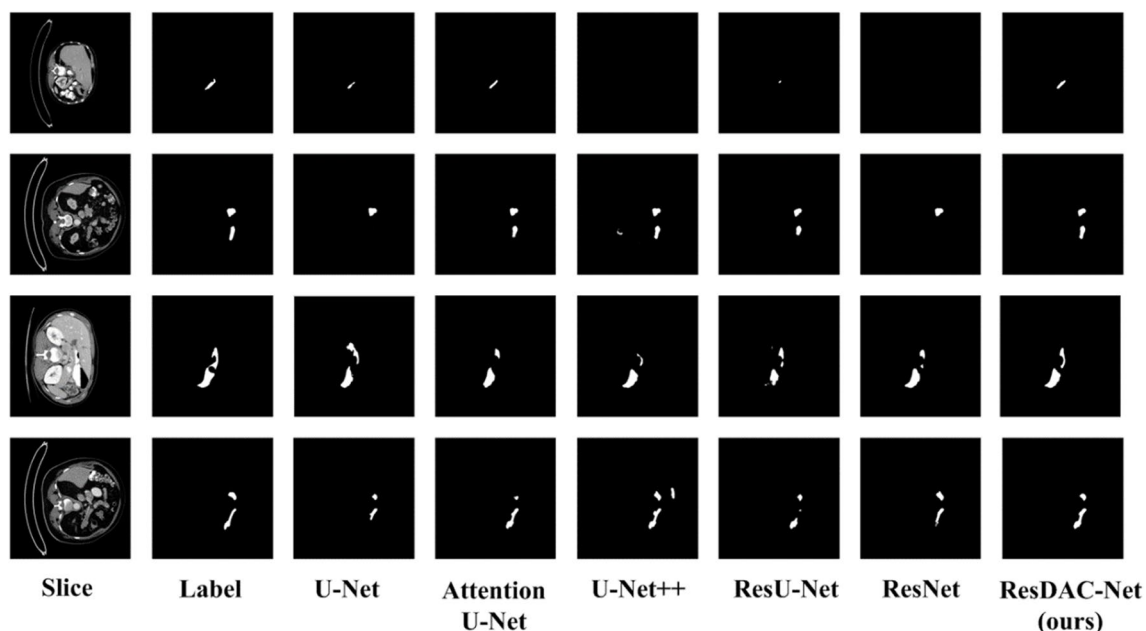


Fig. 7 Sample visualizations of pancreas segmentation results of different models on NIH dataset

Table 5 Pancreas segmentation performance comparison results according to literature

Models	DSC (%)	Precision (%)	Recall (%)	Jaccard index (%)
DAN [55]	83.31	84.09	83.30	71.76
ADR-U-Net [54]	83.03	-	84.60	-
MBU-Net [56]	82.87	89.29	77.37	70.97
ADAU-Net [53]	83.76	-	-	72.38
SE-PResUnet [57]	83.13	-	-	-
MCF [58]	75.02	-	-	61.27
SGU-Net [59]	80.72	-	-	-
MDAG-Net [22]	83.04	81.71	84.42	-
ResDAC-Net (ours)	84.56	81.69	88.48	73.70

The best results are shown in bold

metrics values (except for precision when experimenting on the MSD dataset) by 1.47 (resp. 1.14) percentage points for DSC, 1.81 points for precision, 1.06 (resp. 2.61) percentage points for recall, and 1.99 (resp. 1.13) percentage points for Jaccard index. This highlights the explicit modeling of channel dependencies using the ALF blocks, enabling adaptive recalibration of channel characteristics, which enhances the convolutional block's ability to represent fused information, improves mismatched boundaries, and provides powerful feature extraction capabilities for complex edges. Finally, in the last step, the addition of the PDC block allowed to further improve the metrics values (except for precision when experimenting on the NIH dataset) by 0.67 (resp. 0.94) percentage points for DSC, 1.69 percentage points for precision, 3.38 (resp. 0.65) percentage points for recall, and 0.96 (resp. 0.99) percentage points for Jaccard index. This indicates that using parallel dilated convolutions of different scales before the decoding stage allows to fuse feature maps with different scales and semantics, resulting in different scales of perception fields and effective extraction of multi-scale features.

Overall, the ablation study experiments demonstrated the effectiveness of the resultant ResDAC-Net structure in improving segmentation accuracy and effectively capturing multi-scale features in complex image segmentation tasks. Obviously, the combination of ResDAC, ALF, and PDC blocks plays a crucial role in achieving state-of-the-art segmentation performance.

4.3 Performance comparison with U-Net-based models and ResNet models on NIH dataset

In the second set of experiments, the NIH dataset was used for pancreas segmentation performance comparison of the proposed ResDAC-Net model with U-Net-based models and ResNet models, namely U-Net [14], Attention U-Net [21], U-Net++ [51], ResUNet [52], and ResNet [27]. The obtained results are shown in Table 3.

As can be seen from Table 3, ResDAC-Net outperforms all other models, taking part in this comparison, based on all evaluation metrics. More specifically, the second-best performing model was left behind by 4.79 percentage points (w.r.t. Attention-UNet) according to DSC, 0.32 percentage points (w.r.t. ResU-Net) based on precision, 0.41 percentage points (w.r.t. UNet++) according to recall, and 6.45 percentage points (w.r.t. Attention-UNet) based on Jaccard index. These results demonstrate the effectiveness of ResDAC-Net in performing pancreas segmentation tasks, and its strong capabilities in extracting features from small objects and learning meaningful representations. This is due to the fact that ResDAC-Net takes into consideration the channel information, which allows it to extract the intra-slice relationships between pancreas and background pixels. Additionally, the inclusion of attention modules enables the effective utilization of multi-scale information among feature maps, facilitating communication and integration of features from different levels. As a result, ResDAC-Net can capture multi-scale targets, reduce the impact of inter-class imbalance, and adequately learn the relationship between the overall structure and fine details in pancreas images, leading to improved segmentation results.

The DSC validation and training curves of the compared models are shown in Figs. 5 and 6, respectively. From these figures, it is evident that the ResDAC-Net validation and training (represented by green color) converge faster than that of the other models.

Figure 7 contains sample visualizations of pancreas segmentation results achieved by the compared models on the NIH dataset. The proposed ResDAC-Net model achieves more accurate pancreas segmentation compared to other models. It outperforms them in extracting pancreas edge pixels and effectively suppresses similar background information and noise surrounding the pancreas. Overall, ResDAC-Net is more effective in capturing detailed features of the pancreas, leading to improved segmentation performance.

4.4 Performance comparison with U-Net–based models and ResNet models on MSD dataset

In order to validate the generalization ability of the proposed ResDAC-Net model, pancreas segmentation experiments were conducted with it, along with U-Net–based models and ResNet models, using a different dataset (i.e., MSD) from that on which the model was trained (i.e., NIH). The obtained results are shown in Table 4. As can be seen, ResDAC-Net achieved the highest value (among all compared models) for one of the main metrics used for segmentation performance evaluation, by scoring 69.15% for DSC. This result demonstrates robust generalization and strong pancreas pixel recognition capabilities of the proposed model across different datasets. On the other evaluation metrics, the proposed model took third place respectively on each one. While there were variations in the ResDAC-Net performance here, it is noteworthy that the baseline model (U-Net) generally performed lower on this dataset compared to the other dataset which it was trained on (i.e., NIH).

4.5 Performance comparison with state-of-the-art models

Finally, we compared the segmentation performance of the proposed ResDAC model with that of state-of-the-art models as reported in the corresponding literature sources. The results are shown in Table 5. As can be seen, ResDAC outperforms all other models according to three (out of four) evaluation metrics, including the two main ones used for segmentation performance evaluation (i.e., DSC and Jaccard index). More specifically, the second-best performing model was left behind by 0.80 percentage points (w.r.t. ADAU-Net [53]) according to DSC, 3.88 percentage points (w.r.t. ADR-U-Net [54]) based on recall, and 1.32 percentage points (w.r.t. ADAU-Net [53]) according to Jaccard index.

5 Conclusions and future directions

This paper has proposed a novel pancreas segmentation model, called ResDAC-Net, utilizing residual networks and asymmetric convolutions in its backbone, allowing it to significantly enhance the feature fitting and capture the pancreas' position and shape edge features, thereby further improving its segmentation performance. The obtained experimental results demonstrate that ResDAC-Net achieves significant improvements in metrics, such as Dice similarity coefficient (DSC), precision, recall, and Jaccard index, compared to other existing models.

Although ResDAC-Net has achieved success in pancreas segmentation, there are still some unresolved issues. For

instance, more effective preprocessing methods need to be used, such as targeted approaches to remove artifacts and noise, performing image registration, etc., which will contribute to the enhancement of the segmentation performance of the proposed model. Addressing distribution differences among different datasets is also a challenging problem worthy of in-depth research. By tackling these issues, more reliable organ segmentation can be achieved, which will have a positive impact in the field of medical imaging and clinical applications. In the future, we will strive to improve the generalization ability of the proposed model, design models for segmentation of other organs, and apply them to clinical experiments combined with hardware devices.

Funding Open Access funding provided by the IReL Consortium. This paper has emanated from research conducted with the financial support of the National Key Research and Development Program of China under the Grant No. 2017YFE0135700, the Tsinghua Precision Medicine Foundation under the Grant No. 2022TS003, the Bulgarian National Science Fund (BNSF) under the Grant No. КП-06-ИП-КИТАЙ/1 (KP-06-IP-CHINA/1), and the Telecommunications Research Centre (TRC) of University of Limerick, Ireland.

Declarations

Ethics approval This paper does not contain any experiments with human participants or animals performed by any of the authors.

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Xie L, Cai W, Gao Y (2022) DMCGNet: a novel network for medical image segmentation with dense self-mimic and channel grouping mechanism. *IEEE J Biomed Health Inform* 26:5013–5024. <https://doi.org/10.1109/JBHI.2022.3192277>
2. Yue G, Zhuo G, Li S, Zhou T, Du J, Yan W, Hou J, Liu W, Wang T (2023) Benchmarking polyp segmentation methods in narrow-band imaging colonoscopy images. *IEEE J Biomed Health Inform* 27:3360–3371. <https://doi.org/10.1109/JBHI.2023.3270724>
3. Zeng J, Li Y (2022) Knee joint magnetic resonance images recognition based on iterative threshold segmentation. In *Proceedings of the 2022 2nd International Conference on Computer Graphics, Image and Virtualization (ICCGIV)*, 23–25 Sept. 2022, pp 139–144

4. Zhang X, Liu Z, Wang Q, Chen B (2022) Pneumothorax segmentation of chest X-rays using improved UNet++. In Proceedings of the 2022 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), 5–7 Aug. 2022, pp 24–28
5. Ru K, Li X (2023) A dermoscopic image segmentation algorithm based on U-shaped architecture. In Proceedings of the 2023 2nd International Conference on Big Data, Information and Computer Network (BDICN), 6–8 Jan. 2023, pp 1–5
6. Shao W, Wang Z, Zhang Z, Zhou Q, Wang R, Tan W (2022) A segmentation method of airway from chest CT image based on VGG-Unet neural network. In Proceedings of the 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 6–8 Dec. 2022, pp 1702–1705
7. Shen T, Huang F, Xu H (2023) HardNet and dual-code attention mechanism based model for medical images segmentation. IEEE Access 11:47827–47835. <https://doi.org/10.1109/ACCESS.2023.3275966>
8. Poorani B, Khilar R (2023) Identification of polycystic ovary syndrome in ultrasound images of ovaries using distinct threshold based image segmentation. In Proceedings of the 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT), 5–6 May 2023, pp 570–575
9. Patel S (2023) An overview and application of deep convolutional neural networks for medical image segmentation. In Proceedings of the 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS), 2–4 Feb. 2023, pp 722–728
10. Ning Z, Zhong S, Feng Q, Chen W, Zhang Y (2022) SMU-Net: saliency-guided morphology-aware U-Net for breast lesion segmentation in ultrasound image. IEEE Trans Med Imaging 41:476–490. <https://doi.org/10.1109/TMI.2021.3116087>
11. Monkam P, Jin S, Tang B, Zhou X, Lu W (2022) A disentanglement and fusion data augmentation approach for echocardiography segmentation. In Proceedings of the 2022 IEEE International Ultrasonics Symposium (IUS), 10–13 Oct. 2022, pp 1–4
12. Li C, Liu T, Chen Z, Pang S, Zhong L, Feng Q, Yang W (2022) SPA-RESUNET: strip pooling attention resunet for multi-class segmentation of vertebrae and intervertebral discs. In Proceedings of the 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), 28–31 March 2022, pp 1–5
13. Liu Y, Zhou J, Liu L, Zhan Z, Hu Y, Fu Y, Duan H (2022) FCP-Net: a feature-compression-pyramid network guided by game-theoretic interactions for medical image segmentation. IEEE Trans Med Imaging 41:1482–1496. <https://doi.org/10.1109/TMI.2021.3140120>
14. Ronneberger O, Fischer P, Brox T (2015) 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, pp 234–241
15. Jung K, Nguyen TD, Le DT, Bum J, Woo SS, Choo H (2023) Hand bone X-rays segmentation and congregation for age assessment using deep learning. In Proceedings of the 2023 International Conference on Information Networking (ICOIN), 11–14 Jan. 2023, pp 565–568
16. Du J, Guan K, Liu P, Li Y, Wang T (2023) Boundary-sensitive loss function with location constraint for hard region segmentation. IEEE J Biomed Health Inform 27:992–1003. <https://doi.org/10.1109/JBHI.2022.3222390>
17. Gao Q (2023) Medical image segmentation algorithm based on deep learning and convolutional neural network. In Proceedings of the 2023 IEEE International Conference on Integrated Circuits and Communication Systems (ICICACS), 24–25 Feb. 2023, pp 01–05
18. Gut D, Tabor Z, Szymkowski M, Rozynek M, Kucybała I, Wojciechowski W (2022) Benchmarking of deep architectures for segmentation of medical images. IEEE Trans Med Imaging 41:3231–3241. <https://doi.org/10.1109/TMI.2022.3180435>
19. Huang X, Deng Z, Li D, Yuan X, Fu Y (2023) MISSFormer: an effective transformer for 2D medical image segmentation. IEEE Trans Med Imaging 42:1484–1494. <https://doi.org/10.1109/TMI.2022.3230943>
20. Ify MAH, Shajid MSS (2023) Implementation of liver segmentation from computed tomography (CT) images using deep learning. In Proceedings of the 2023 International Conference on Electrical, Computer and Communication Engineering (ECCE), 23–25 Feb. 2023, pp 1–6
21. Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla NY, Kainz B, Glocker B, Rueckert D (2018) Attention U-Net: learning where to look for the pancreas. ArXiv, abs/1804.03999
22. Cao L, Li J, Chen S (2023) Multi-target segmentation of pancreas and pancreatic tumor based on fusion of attention mechanism. Biomed Signal Process Control 79:104170. <https://doi.org/10.1016/j.bspc.2022.104170>
23. Cai J, Lu L, Xie Y, Xing F, Yang L (2017) Improving deep pancreas segmentation in CT and MRI images via recurrent neural contextual learning and direct loss function. arXiv, abs/1707.04912
24. Song E, Zhan B, Liu H (2024) Combining external-latent attention for medical image segmentation. Neural Netw 170:468–477. <https://doi.org/10.1016/j.neunet.2023.10.046>
25. Heinrich MP, Blendowski M, Oktay O (2018) TernaryNet: faster deep model inference without GPUs for medical 3D segmentation using sparse and binary convolutions. Int J Comput Assist Radiol Surg 13:1311–1320
26. Zhang Z, Zhang X, Peng C, Xue X, Sun J (2018) ExFuse: enhancing feature fusion for semantic segmentation. In Proceedings of the European Conference on Computer Vision, pp 273–288
27. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
28. Miao J, Xu S, Zou B, Qiao Y (2022) ResNet based on feature-inspired gating strategy. Multimedia Tools Appl 81:19283–19300. <https://doi.org/10.1007/s11042-021-10802-6>
29. Celik G (2023) Detection of COVID-19 and other pneumonia cases from CT and X-ray chest images using deep learning based on feature reuse residual block and depthwise dilated convolutions neural network. Appl Soft Comput 133:109906. <https://doi.org/10.1016/j.asoc.2022.109906>
30. Guo S, Sheng S, Lai Z, Chen S (2022) Trans-U: transformer enhanced U-Net for medical image segmentation. In Proceedings of the 2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA), 20–22 May 2022, pp 628–631
31. Wang J, Tang Y, Xiao Y, Zhou JT, Fang Z, Yang F (2023) GREnet: gradually recurrent network with curriculum learning for 2-D medical image segmentation. IEEE Trans Neural Netw

- Learning Syst: 1–15. <https://doi.org/10.1109/TNNLS.2023.3238381>
32. Xiao X, Lian S, Luo Z, Li S (2018) Weighted res-unet for high-quality retina vessel segmentation. In Proceedings of the 2018 9th international conference on information technology in medicine and education (ITME), pp 327–331
 33. Kim Y, Kang D, Mok Y, Kwon S, Paik J (2023) A review on few-shot learning for medical image segmentation. In Proceedings of the 2023 International Conference on Electronics, Information, and Communication (ICEIC), 5–8 Feb. 2023, pp 1–3
 34. Qian L, Wen C, Li Y, Hu Z, Zhou X, Xia X, Kim S-H (2024) Multi-scale context UNet-like network with redesigned skip connections for medical image segmentation. *Comput Methods Programs Biomed* 243:107885. <https://doi.org/10.1016/j.cmpb.2023.107885>
 35. Ding X, Guo Y, Ding G, Han J (2019) Acnet: strengthening the kernel skeletons for powerful cnn via asymmetric convolution blocks. In Proceedings of the Proceedings of the IEEE/CVF international conference on computer vision, pp 1911–1920
 36. Dayananda C, Yamanakkanavar N, Nguyen T, Lee B (2023) AMCC-Net: an asymmetric multi-cross convolution for skin lesion segmentation on dermoscopic images. *Eng Appl Artif Intell* 122:106154. <https://doi.org/10.1016/j.engappai.2023.106154>
 37. Jena PK, Khuntia B, Palai C, Nayak M, Mishra TK, Mohanty SN (2023) A novel approach for diabetic retinopathy screening using asymmetric deep learning features. *Big Data Cogn Comput* 7(1):25. <https://doi.org/10.3390/bdcc7010025>
 38. Ma Y, Wu L, Gao Y, Gao F, Zhang J, Luo Z (2023) ULFAC-Net: ultra-lightweight fully asymmetric convolutional network for skin lesion segmentation. *IEEE J Biomed Health Inform* 27(6):2886–2897
 39. Hu J, Shen L, Sun G (2018) Squeeze-and-excitation networks. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, pp 7132–7141
 40. Li X, Zhai M, Sun J (2021) DDCNNC: dilated and depthwise separable convolutional neural network for diagnosis COVID-19 via chest X-ray images. *Int J Cogn Comput Eng* 2:71–82. <https://doi.org/10.1016/j.ijcce.2021.04.001>
 41. Montazerolghaem M, Sun Y, Sasso G, Haworth A (2023) U-Net architecture for prostate segmentation: the impact of loss function on system performance. *Bioengineering* 10(4):412. <https://doi.org/10.3390/bioengineering10040412>
 42. Ruby U, Yendapalli V (2020) Binary cross entropy with deep learning technique for image classification. *Int J Adv Trends Comput Sci Eng* 9(4):5393–5397. <https://doi.org/10.30534/ijatcse/2020/175942020>
 43. Milletari F, Navab N, Ahmadi SA (2016) V-Net: fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), 25–28 Oct. 2016, pp 565–571
 44. Tu WC, Liu MY, Jampani V, Sun D, Chien SY, Yang MH, Kautz J (2018) Learning superpixels with segmentation-aware affinity loss. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 18–23 June 2018, pp 568–576
 45. Clough JR, Byrne N, Oksuz I, Zimmer VA, Schnabel JA, King AP (2022) A topological loss function for deep-learning based image segmentation using persistent homology. *IEEE Trans Pattern Anal Mach Intell* 44:8766–8778. <https://doi.org/10.1109/TPAMI.2020.3013679>
 46. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. *CoRR*, abs/1412.6980
 47. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M (2013) The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging* 26:1045–1057
 48. Roth HR, Farag A, Turkbey E, Lu L, Liu J, Summers RM (2016) Data from pancreas-CT. The Cancer Imaging Archive, vol. 32
 49. Roth HR, Lu L, Farag A, Shin H-C, Liu J, Turkbey EB, Summers RM (2015) DeepOrgan: multi-level deep convolutional networks for automated pancreas segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Cham, 2015//, pp 556–564
 50. Antonelli M, Reinke A, Bakas S, Farahani K, Kopp-Schneider A, Landman BA, Litjens G, Menze B, Ronneberger O, Summers RM (2022) The medical segmentation decathlon. *Nat Commun* 13:4128
 51. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J (2018) Unet++: a nested U-Net architecture for medical image segmentation. In Proceedings of the Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4, pp 3–11
 52. Diakogiannis FI, Waldner F, Caccetta P, Wu C (2020) ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data. *ISPRS J Photogramm Remote Sens* 162:94–114. <https://doi.org/10.1016/j.isprsjprs.2020.01.013>
 53. Li M, Lian F, Li Y, Guo S (2022) Attention-guided duplex adversarial U-Net for pancreatic segmentation from computed tomography images. *J Appl Clin Med Phys* 23:e13537. <https://doi.org/10.1002/acm2.13537>
 54. Li M, Lian F, Wang C, Guo S (2021) Accurate pancreas segmentation using multi-level pyramidal pooling residual U-Net with adversarial mechanism. *BMC Med Imaging* 21:168. <https://doi.org/10.1186/s12880-021-00694-1>
 55. Li M, Lian F, Guo S (2021) Automatic pancreas segmentation using double adversarial networks with pyramidal pooling module. *IEEE Access* 9:140965–140974. <https://doi.org/10.1109/ACCESS.2021.3118718>
 56. Huang M-L, Wu Y-Z (2022) Semantic segmentation of pancreatic medical images by using convolutional neural network. *Biomed Signal Process Control* 73:103458. <https://doi.org/10.1016/j.bspc.2021.103458>
 57. Zhu Y, Hu P, Li X, Tian Y, Bai X, Liang T, Li J (2022) Multiscale unsupervised domain adaptation for automatic pancreas segmentation in CT volumes using adversarial learning. *Med Phys* 49:5799–5818. <https://doi.org/10.1002/mp.15827>
 58. Wang Y, Xiao B, Bi X, Li W, Gao X (2023) MCF: mutual correction framework for semi-supervised medical image segmentation. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp 15651–15660
 59. Lei T, Sun R, Du X, Fu H, Zhang C, Nandi AK (2023) SGU-Net: shape-guided ultralight network for abdominal image segmentation. *IEEE J Biomed Health Inform* 27:1431–1442. <https://doi.org/10.1109/JBHI.2023.3238183>



Zhanlin Ji (Member, IEEE) received his MEng and PhD degrees from Dublin City University and University of Limerick in 2006 and 2010, respectively. He is a Professor at the North China University of Science and Technology (China) and an Associated Researcher at the Telecommunications Research Centre (TRC), University of Limerick (Ireland). Prof. Ji's research interests include Ubiquitous Consumer Wireless World (UCWW), Internet of Things (IoT), cloud computing, big data management, and

data mining. He has authored/co-authored 100+ research papers in refereed journals and conferences.



Jianuo Liu was born in 2000. She received the B.S. degree from the North China University of Science and Technology, in 2022. She is currently pursuing the Master's degree with the North China University of Science and Technology. Her research interests include machine vision and graphic image processing.



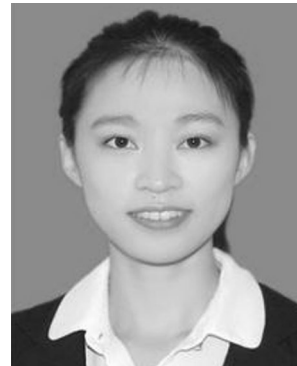
Juncheng Mu was born in 2000. He received his B.S. degree from North China University of Science and Technology in 2022. Currently, he is pursuing a Master's degree at North China University of Science and Technology. His research interests include machine vision and graphic image processing.



Haiyang Zhang is currently a lecturer at Xi'an Jiaotong-Liverpool University, China. She received her PhD degree in the Department of Electronic & Computer Engineering, University of Limerick, Ireland, in 2018, and her B.S. degree from the School of Software Engineering, Jilin University, China, in 2013. Dr. Zhang's current research interests include recommender systems, data mining, collaborative filtering, and natural language processing.



Chenxu Dai was born in 2000, he received the B.S. degree from the North China University of Science and Technology, in 2022. She is currently pursuing the Master's degree with the North China University of Science and Technology. Her research interests include machine vision and graphic image processing.



Na Yuan obtained a Bachelor's degree from Heilongjiang University in July 2011 and a Master's degree from Hebei University of Technology in January 2014. She is currently a lecturer at Tangshan University. Her research interests include intelligent control, machine vision, and graphic image processing.



Ivan Ganchev is an IEEE Senior Member, an ITU-T Invited Expert, and an IET Invited Lecturer, currently affiliated with University of Limerick, Ireland, University of Plovdiv "Paisii Hilendarski," and Institute of Mathematics and Informatics—Bulgarian Academy of Sciences, Bulgaria. He received his doctoral and engineering (*summa cum laude*) degrees from Saint-Petersburg University of Telecommunications in 1995 and 1989, respectively. Prof. Ganchev was involved in more than 40 international and

national research projects. He has served on the TPC of more than 400 prestigious international conferences/symposia/workshops, and has authored/co-authored 1 monographic book, 3 textbooks, 4 edited books, and more than 300 research papers in refereed international journals, books, and conference proceedings. Prof. Ganchev is on the editorial board of and has served as a Guest Editor for multiple prestigious international journals.