Automatic pancreas segmentation in abdominal CT image with contrast enhancement block

Pan, Shengxue, Xiang, Dehui, Bian, Yun, Lu, Jianping, Jiang, Hui, et al.

Shengxue Pan, Dehui Xiang, Yun Bian, Jianping Lu, Hui Jiang, Jianming Zheng, "Automatic pancreas segmentation in abdominal CT image with contrast enhancement block," Proc. SPIE 11596, Medical Imaging 2021: Image Processing, 115961B (15 February 2021); doi: 10.1117/12.2581040

SPIE.

Event: SPIE Medical Imaging, 2021, Online Only

Automatic Pancreas Segmentation in Abdominal CT Image with Contrast Enhancement Block

Shengxue Pan[1], Dehui Xiang[1*], Yun Bian[2], Jianping Lu[2], Hui Jiang[3], Jianming Zheng[3]

[1]School of Electronic and Information Engineering, Soochow University, Suzhou, Jiangsu Province, 215006, China
[2]Department of Radiology, Changhai Hospital, The Navy Military Medical University, Shanghai, China
[3]Department of Pathology, Changhai Hospital, The Navy Military Medical University, Shanghai, China

ABSTRACT

Automatic pancreas segmentation in abdominal CT images plays an important role in clinical applications. It can provide doctors with quantitative and qualitative information. Due to the small size, unclear edges, and the high anatomical differences between patients, it is a challenging task to accurately segment the pancreas with diseases. In this paper, we propose a new method to automatically segment the pancreas in abdominal CT images. First, we propose a contrast enhancement block. The block generates edge information and uses gating mechanism to enhance edge details of the pancreas. Second, we leverage a reverse attention block. This block utilizes the decoder feature map to guide the network to mine complementary discriminative regions. The proposed method is trained on 63 3D CT images, validated on 15 3D CT images, and tested on 28 3D CT images. Compared with manual segmentation, the mean Dice similarity coefficient can reach 86.11±8.02%. Experimental results show that our method can obtain more accurate segmentation results compared with existing segmentation methods.

KEYWORDS: diseased pancreas segmentation, contrast enhancement block, reverse attention block

1. INTRODUCTION

Pancreas is the second largest exocrine organ and an important endocrine gland of the human body. It can be divided into pancreatic head, neck, body, and tail. Pancreatic cancer is one of the common malignant tumors of the digestive system, which is characterized by late clinical discovery, high degree of malignancy and poor prognosis. The 5-year relative survival rate of pancreatic cancer is less than 10%, which is the worst among common malignant tumors and shows a trend of deterioration year by year [1]. Early detection of pancreatic cancer is the key to radical resection and survival. Computed tomography (CT) is the most important imaging method for pancreatic disease, which has the advantages of high density resolution and repeatability. Automatic pancreas segmentation in abdominal CT images can provide more accurate information for the diagnosis and treatment of pancreatic cancer.

As shown in Fig.1, it is still a challenging task to segment pancreas. First, the pancreas shows very high anatomical variation between patients in its size, shape, position, and texture. Second, the edges of the pancreas are not clear and the contrast between the surrounding tissues and organs is low in CT images. Especially at the head of the pancreas, the boundary with the duodenum is indistinguishable. Third, diseases such as pancreatitis and pancreatic cancer can greatly change the appearance of the pancreas, resulting in diffuse enlargement, uneven density, blurred boundaries and other problems of the pancreas [2].

In recent years, many methods have been used to segment pancreas in abdominal CT images. Roth et al. [3] proposed a random forest algorithm and used the semantic clues of internal and boundary maps of organs obtained by deep learning to generate pixel-level labeling of pancreas segmentation. Ma et al. [4] proposed a new Bayesian model

*Corresponding author: Dehui Xiang, E-mail: xiangdehui@suda.edu.cn

Medical Imaging 2021: Image Processing, edited by Ivana Išgum, Bennett A. Landman, Proc. of SPIE Vol. 11596, 115961B · © 2021 SPIE · CCC code: 1605-7422/21/$21 · doi: 10.1117/12.2581040

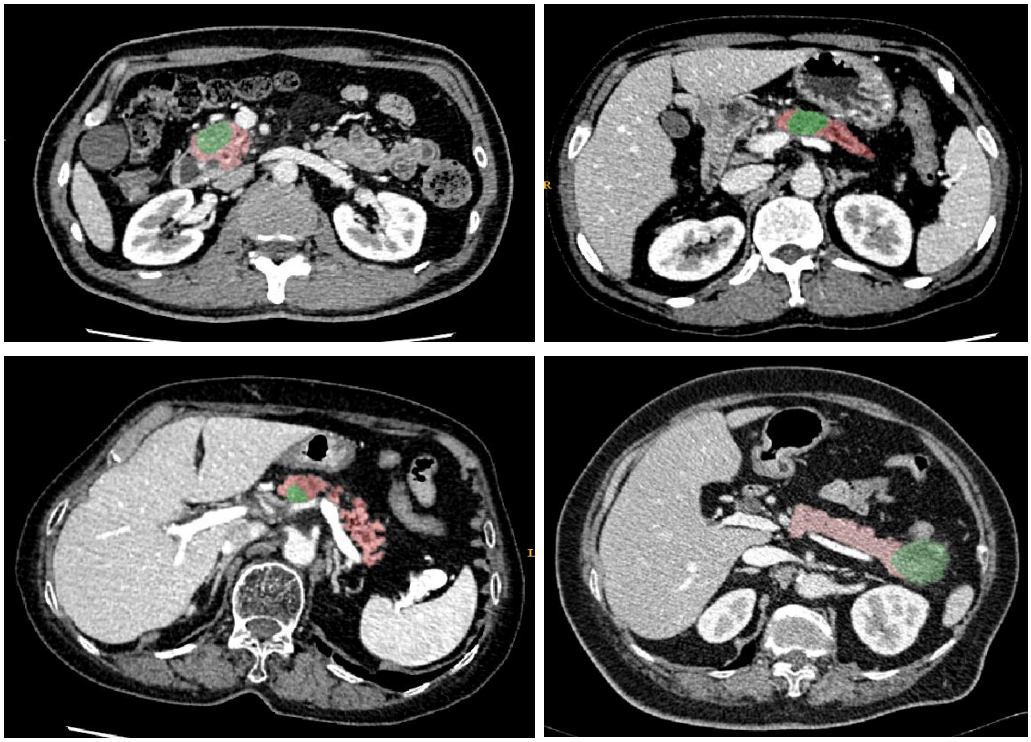


Fig.1. Illustration of the challenges in pancreas segmentation. The green area is the pancreatic tumor and the red area is the rest of the pancreatic organ.

combining deep neural network and statistical shape model to refine the segmentation results of pancreas. Both of these two methods combine traditional methods and deep learning methods to optimize the segmentation of the pancreas. Zhou et al. [5] used a two-stage coarse-to-fine segmentation method to segment the pancreas. They trained two FCN-8 segmentation networks to process the entire input area and the area cropped according to the bounding box. Then, they sent the coarse segmentation results of the first network to the second network in the test phase and used an iterative process to optimize the fixed-point model to achieve more accurate segmentation. Zhu et al. [6] proposed a new three-dimensional coarse-to-fine convolutional neural network framework to use rich spatial information to finely segment the pancreas. However, as the edges of the pancreas are often blurry, many existing methods do not perform well near the region boundaries.

To deal with this problem, we propose a convolutional neural network with contrast enhancement block and reverse attention block, which can achieve automatic end-to-end segmentation of pancreases from abdominal CT images. Experiments show the method helps to extract features and emphasize the edges, and the proposed network outperforms the baseline and some existing deep networks.

2. METHODS

We introduce the proposed method in the following four parts: the structure of the proposed network, the contrast enhancement block (CEB), reverse attention block (RAB) and the loss function.

2.1 Network Structure

As we know, U-Net [10] achieves good results in medical image segmentation. In this paper, we also use the idea of encoding and decoding network. Without changing the U-Net depth, the number of feature channels is reduced by half at each stage to obtain a lightweight benchmark network. The overview of our proposed method is shown in Fig.2. In order to further improve the image segmentation performance, we first add a contrast enhancement block after the feature map

of the encoder to guide the network to pay more attention to the edge details of the pancreas. Second, the reverse attention block is leveraged, using decoder information to guide the mining of complementary discriminative regions.

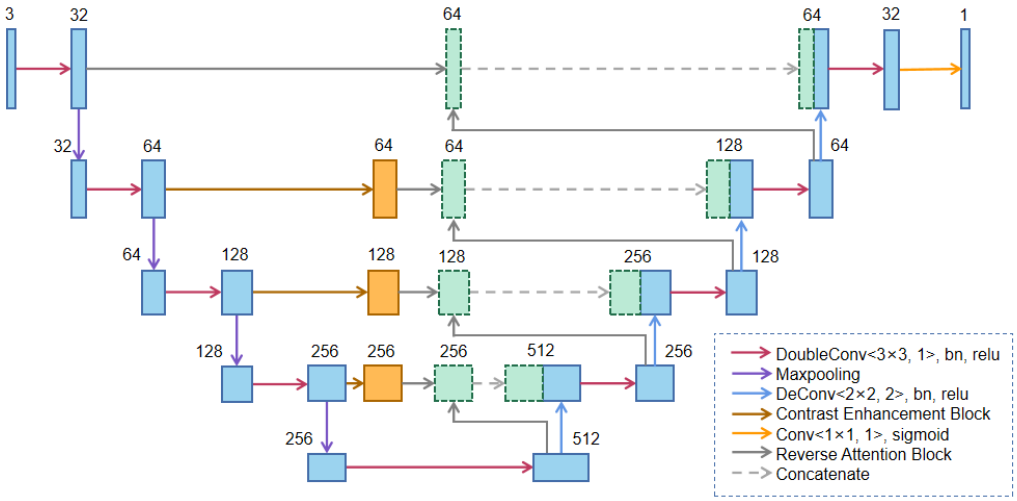


Fig.2. The architecture of proposed method for diseased pancreas segmentation.

2.2 Contrast Enhancement Block

The U-Net merges the feature map of the encoder layer and the feature map of the decoder layer through a simple concatenate operation in the skip-connection part. In order to highlight the edge details of the pancreas in the feature map of the encoder layer, we propose a contrast enhancement block, as shown in Fig.3. The feature map of the encoder layer can be denoted as F_{en}, and the output of the contrast enhancement block O_{ceb} can be formulated as

$$O_{ceb} = (1 + G_1) \times F_{en} + (1 - G_1) \times (G_2 \times D_{eem}) \tag{1}$$

where D_{eem} is the edge extraction module, which can be formulated as

$$D_{eem} = F_{en} - Average\_pooling\ (F_{en}) \tag{2}$$

Average pooling can extract local average information. Then, the edge details can be obtained by subtracting the local average information from F_{en}. G_1 and G_2 are the gate maps. They are computed by 1×1 convolution and sigmoid activation functions respectively for F_{en} and D_{eem}, which represent the deterministic feature map of each position of F_{en} and D_{eem}. The uncertain position of F_{en} can be obtained through (1-G_1). The deterministic edge information with the F_{en}'s uncertain position can supplement and enhance the information of the encoder feature map in detail.

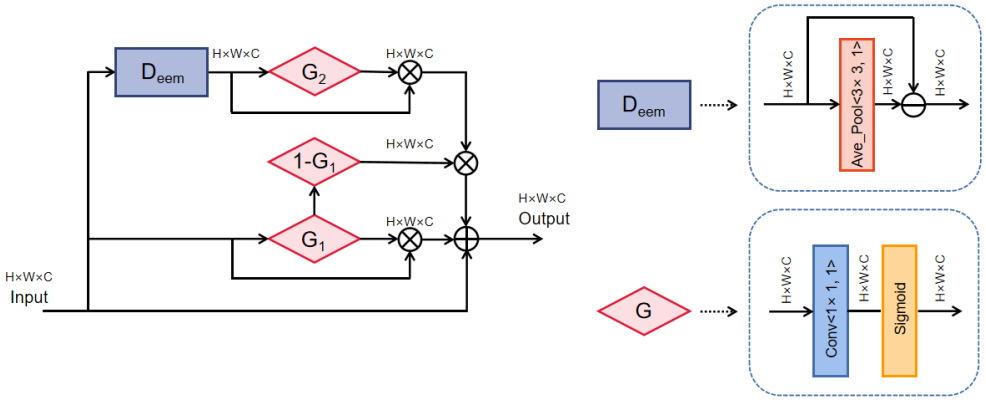


Fig.3. The structure of contrast enhancement block (CEB).

2.3 Reverse Attention Block

In general, the U-shaped convolutional neural network generates the segmentation results of pancreas through up-sampling. It can only capture a relatively rough location without detail structure information. To solve this problem, we leverage a reverse attention block, as shown in Fig.4. The block utilizes the output feature of each layer of decoder to mine the complementary regions and details of the previous layer. Specifically, given the high-level skip-connection inputs I_s and the reverse attention weight W_r, the attention feature map F_a can be denoted as

$$F_a = \sigma_1(f_1(I_s)) \times W_r + I_s \tag{3}$$

The reverse attention weight W_r can be formulated as

$$W_r = \sigma_2(-f_2(D_{de})) \tag{4}$$

where $f_1(\cdot)$ denotes the convolution operation, $\sigma_1(\cdot)$ denotes the relu activation function, $f_2(\cdot)$ denotes the deconvolution operation, $\sigma_2(\cdot)$ denotes the sigmoid activation function, and D_{de} denotes the feature map of the current decoder layer. Inverting the rough feature map obtained by deconvolution and multiplying it with the high-level encoded feature map can refine the inaccurate and rough estimation into an accurate and complete prediction map.

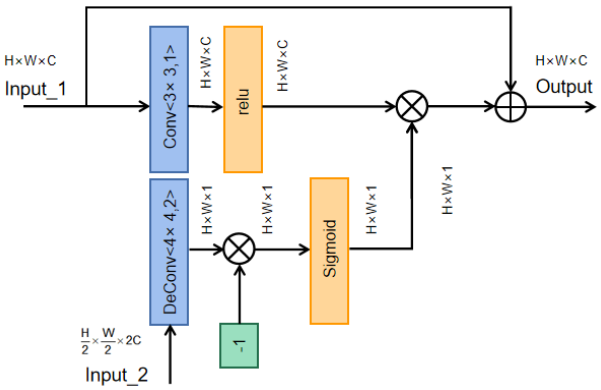


Fig.4. The structure of reverse attention block (RAB).

2.4 Loss function

As this is a two-class segmentation task with class imbalance, we use Dice similarity coefficient (DSC) rather than cross entropy to measure the similarity. Given prediction map Y' and ground truth Y, Dice loss is formulated as

$$L(Y, Y') = 1 - \frac{2 \times |Y \cap Y'|}{|Y| + |Y'|} \tag{5}$$

3. RESULTS

3.1 Datasets

The dataset contains 106 contrast-enhanced 3D abdominal CT images from different patients with pancreatic cancer. The size of each 3D CT image is 512×512×L, where L∈[174, 376] is the number of slices along the long axis of the body. The pancreases were manually labeled as ground truth by clinical experts.

3.2 Data Preprocessing

In our experiments, we divided them into 63 training images, 15 validation images and 28 test images. For the data pre-processing, we simply truncated the raw intensity values to be in [-100, 240] to improve the display of the details of the pancreas. Then, we normalized each raw CT case to be zero mean and unit variance to decrease the data variance caused

by the physical processes of medical images. Our model is trained on those slices which the pancreas occupies at least 50 pixels to prevent the model from being heavily impacted by the background. In order to obtain the spatial information in the 3D image without increasing the amount of calculation, three consecutive slices of these 3D images are concatenated as the input.

3.3 Implementation Details

Our proposed network was trained on a workstation equipped with one NVIDIA Tesla K40m GPU with 12G memory for 20 epochs. We used Adam algorithm with an initial learning rate of 10^{-4} to optimize the weight of the network in the training process. In our experiment, the batch size was set to 3 and our model was trained for 20 epochs. The model was implemented based on the Tensorflow framework. Since the pancreas is a whole organ, we take the largest connected region of the segmentation results to further improve the segmentation accuracy.

3.4 Experimental results

In order to quantitatively evaluate the performance of our proposed method, we compared the segmentation results with ground truth based on the following four measures: dice similarity coefficient (DSC), intersection over union (IoU), precision and recall.

Both DSC and IoU are usually used to measure the similarity between the network segmentation results and the ground truth. They are defined as

$$DSC = \frac{2TP}{FP + 2TP + FN} \tag{6}$$

$$IoU = \frac{TP}{FP + TP + FN} \tag{7}$$

where TP represents the number of true positives, FP represents the number of false positives and FN represents the number of false negatives. Precision and recall are defined as

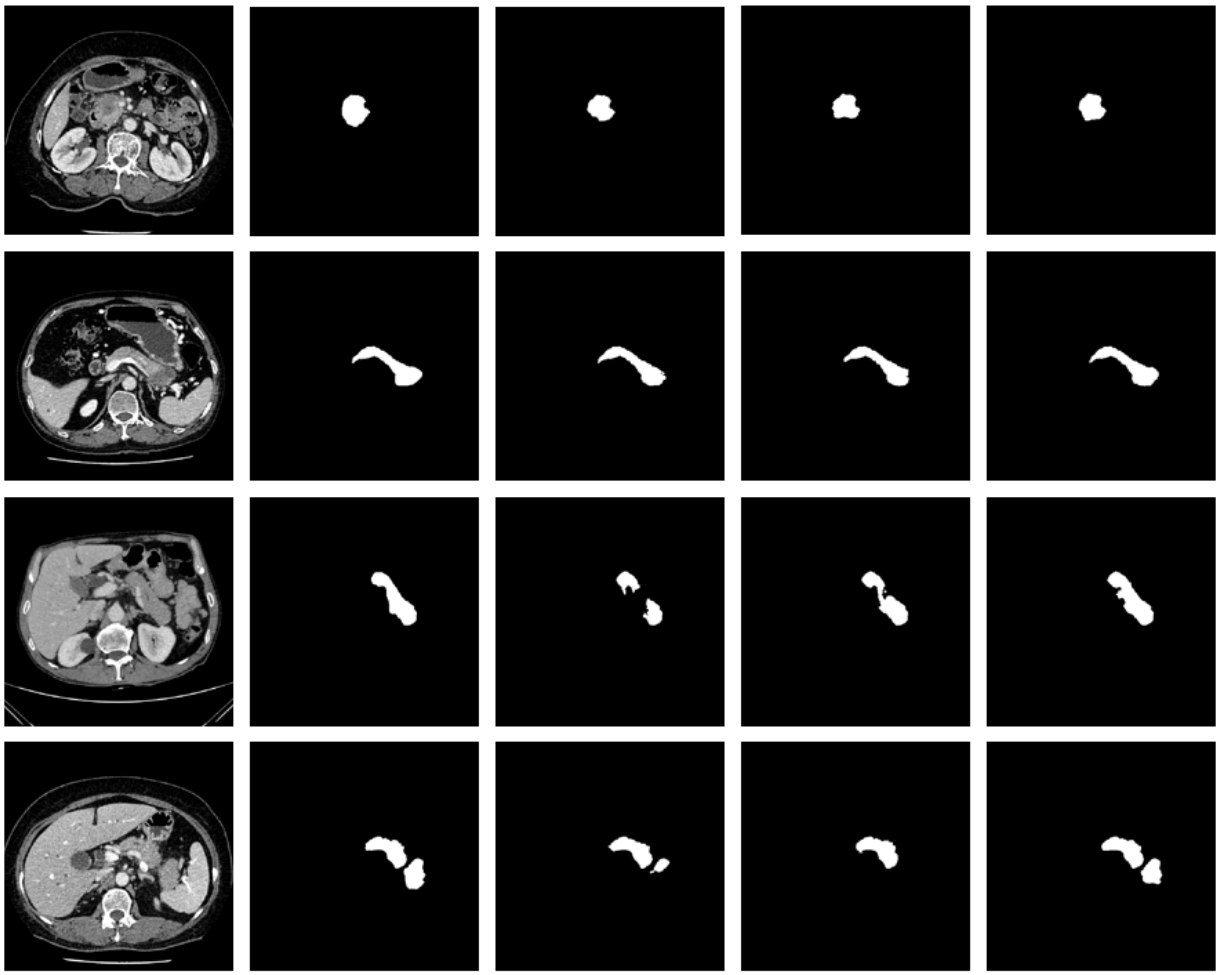
$$precision = \frac{TP}{TP + FP} \tag{8}$$

$$recall = \frac{TP}{TP + FN} \tag{9}$$

In order to demonstrate the effectiveness of our approach, we compared our network with some other existing networks, including FCN-8s, SegNet, U-Net and Attention U-Net. We also did some experiments to prove the effectiveness of our blocks. As can be seen from Table 1, compared with these existing networks, our method achieved remarkable improvements in DSC, IoU and recall. Our method also improves DSC by more than 2% compared to U-Net.

Table 1. Comparison of quantitative segmentation results for different methods (mean±standard deviation).

Methods	DSC(%)	IoU(%)	Precision(%)	Recall(%)
FCN [7]	75.47±14.67	62.35±14.95	84.40±7.14	71.60±17.72
SegNet [8]	75.85±16.85	63.30±16.55	89.41±6.68	69.05±18.55
Attention U-Net [9]	82.58±13.32	71.99±14.86	89.93±6.72	78.87±16.49
U-Net [10]	83.54±14.51	73.68±15.89	**90.95±5.94**	79.66±17.18
U-Net+CEB	84.83±13.17	75.32±14.81	90.69±5.65	82.25±16.31
U-Net+CEB+RAB	**86.11±8.02**	**76.39±11.04**	89.51±6.24	**84.23±12.06**



(a) (b) (c) (d) (e)

Fig.5. Segmentation results of different methods. (a) Original image. (b) Ground truth. (c) U-Net. (d) Attention U-Net. (e) Our method.

Fig.5 shows some CT image slices from different patients and the corresponding segmentation results. We selected different parts of the pancreas to show the results, including the head, body, and tail. Although the pancreases in different images have different sizes and shapes, our method still can achieve good results. When compared to other methods, the results are more accurate, indicating that the proposed blocks make the network perform better.

4. CONCLUSIONS

In this paper, a new approach for automatic pancreas segmentation is introduced. First, a contrast enhancement block is proposed to guide the network to pay more attention to the edge details of the pancreas. Second, a reverse attention block is used to utilize decoder information to guide the mining of complementary discriminative regions. Our method achieves more accurate segmentation results compared to state-of-the-art approaches.

5. ACKNOWLEDGEMENTS

This work has been supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61971298, and in part by the National Key R&D Program of China under Grant 2018YFA0701700.

6. REFERENCE

[1] Chu LC, Park S, Kawamoto S, et al. Utility of CT Radiomics Features in Differentiation of Pancreatic Ductal Adenocarcinoma From Normal Pancreatic Tissue[J]. AJR Am J Roentgenol, 2019, 213(2): 349-357.

[2] Kaur, Sukwinder, et al. "Early diagnosis of pancreatic cancer: challenges and new developments." Biomarkers in Medicine 6.5 (2012): 597-612.2.

[3] Roth H R, Lu L, Farag A, et al. Spatial aggregation of holistically-nested networks for automated pancreas segmentation[C]//International conference on medical image computing and computer-assisted intervention. Springer, Cham, 2016: 451-459.

[4] Ma J, Lin F, Wesarg S, et al. A novel Bayesian model incorporating deep neural network and statistical shape model for pancreas segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2018: 480-487.

[5] Zhou Y, Xie L, Shen W, et al. A fixed-point model for pancreas segmentation in abdominal CT scans[C]//International conference on medical image computing and computer-assisted intervention. Springer, Cham, 2017: 693-701.

[6] Zhu Z, Xia Y, Shen W, et al. A 3D coarse-to-fine framework for volumetric medical image segmentation[C]//2018 International Conference on 3D Vision (3DV). IEEE, 2018: 682-690.

[7] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." computer vision and pattern recognition (2015): 3431-3440.

[8] Badrinarayanan, Vijay, Alex Kendall, and Roberto Cipolla. "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation." IEEE Transactions on Pattern Analysis and Machine Intelligence 39.12 (2017): 2481-2495.

[9] Oktay, Ozan, et al. "Attention U-Net: Learning Where to Look for the Pancreas." arXiv: Computer Vision and Pattern Recognition (2018).

[10] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation." medical image computing and computer assisted intervention (2015): 234-241.