EFFECTIVE ALGORITHM FOR BIOMEDICAL IMAGE SEGMENTATION

DI ZHAO, YI TANG, ALEVTINA B. GOURINOVITCH

Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)

Abstract. Biomedical image segmentation plays an important role in quantitative analysis, clinical diagnosis, and medical manipulation. Objects in medical images have different scales, types, complex backgrounds, and similar tissue appearances, making information extraction challenging. To solve this problem, a module is proposed that takes into account the features of images, which will improve the biomedical image segmentation network FE-Net. An integral part of the FE-Net algorithm is the connection skipping mechanism, which ensures the connection and fusion of feature maps from different layers in the encoder and decoder. Features at the encoder level are combined with high-level semantic knowledge at the decoder level. The algorithm establishes connections between feature maps, which is used in medicine for image processing. The proposed method is tested on three public datasets: Kvasir-SEG, CVC-ClinicDB and 2018 Data Science Bowl. Based on the results of the study, it was found that FE-Net demonstrates better performance compared to other methods in terms of Intersection over Union and F1-score. The network under consideration copes more effectively with segmentation details and object boundaries, while maintaining high accuracy. The study was conducted jointly with the Department of Magnetic Resonance Imaging of the N. N. Alexandrov National Oncology Center. Access to the source code of the algorithm and additional technical details is available at https://github.com/tyjcbzd/FE-Net.

Keywords: biomedical image segmentation, convolution neural network, feature aware module, attention mechanism, U-shaped network.

Conflict of interests. The authors declare no conflict of interests.

ЭФФЕКТИВНЫЙ АЛГОРИТМ СЕГМЕНТАЦИИ БИОМЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ

ДИ ДЖАО, И ТАН, А. Б. ГУРИНОВИЧ

Белорусский государственный университет информатики и радиоэлектроники
(г. Минск, Республика Беларусь)

Аннотация. Сегментация биомедицинских изображений играет важную роль в количественном анализе, клинической диагностике и медицинских манипуляциях. Объекты на медицинских изображениях имеют различный масштаб, тип, сложный фон и схожий внешний вид тканей, что усложняет извлечение информации. Для решения данной проблемы предлагается модуль, учитывающий особенности изображений, что позволит усовершенствовать биомедицинскую сеть сегментации изображений FE-Net. Неотъемле-

мая часть алгоритма FE-Net – механизм пропуска соединений, обеспечивающий соединение и объединение карт признаков из различных слоев в кодере и декодере. Признаки на уровне кодера комбинируются с высокоуровневыми семантическими знаниями на уровне декодера. Алгоритм устанавливает связи между картами признаков, что используется в медицине для обработки изображений. Предлагаемый метод протестирован на трех общедоступных наборах данных: Kvasir-SEG, CVC-ClinicDB и 2018 Data Science Bowl. По результатам исследования установлено, что FE-Net демонстрирует лучшую производительность по сравнению с другими методами по показателям Intersection over Union и F1-оценки. Рассматриваемая сеть эффективнее справляется с деталями сегментации и границами объектов, сохраняя при этом высокую точность. Исследование проведено совместно с отделением магнитно-резонансной томографии Национального онкологического центра имени Н. Н. Александрова. Доступ к исходному коду алгоритма и дополнительным техническим деталям размещен по ссылке https://github.com/tyjcbzd/FE-Net.

Ключевые слова: сегментация биомедицинских изображений, свёрточная нейронная сеть, модуль распознавания признаков, механизм внимания, сеть с U-образной архитектурой.

Introduction

Medical images are images generated using various imaging technologies in the medical field and are used to diagnose, treat, and study diseases. Medical images can include image vessels, cells, and lesions. With the continuous development of biomedical imaging technology, more and more diagnostic imaging tools have emerged, such as magnetic resonance imaging, computed tomography, and positron emission tomography [1]. Normally, doctors to distinguish and make diagnoses with the naked eye do medical image segmentation. However, since the human eye is essentially an imperfect tool, inter-observer and intra-observer differences are prone to occur. At the same time, when the number of medical images increases, manual segmentation becomes tedious and expensive, and very challenging to perform. Therefore, image segmentation technology under deep learning methods can be used as an auxiliary method for detection. Image segmentation refers to the subdivision of an image into components with the same characteristics to extract regions of interest on a semi-automatic or automatic basis.

In recent years, image segmentation methods have made great progress in biomedical image segmentation and clinical decision-making. Different convolutional network models have dominated detection, classification, and segmentation problems in the biomedical field. Convolutional neural network architectures based on encoder-decoder include U-Net [2], U-Net++ [3], ResU-Net++ [4], etc. They have excellent performance in segmentation tasks. The success of these models is largely attributed to the skip connection. Through the joint high-resolution and low-resolution information, thereby improving feature expression. Skip connection algorithm solves the problem of widespread semantic errors between high- and low-resolution features. But information loss and fuzziness are ignored when extracting features at each level.

In order to address this problem and enhance the performance of image segmentation, this research proposes an improved module called the Feature-Aware Module (FAM). The proposed method utilizes an encoder-decoder architecture and incorporates attention mechanisms and residuals to allow the network to adaptively adjust the importance of each layer of features. This enables the network to focus more on key areas, suppress irrelevant noise, and prioritize the extraction of detailed and accurate features at each layer. This densely connected design enables the network to fully leverage the information exchange between low-level and high-level features, thereby enhancing multi-scale expressive capabilities and context awareness of the features.

Furthermore, this fine-grained attention regulation contributes to the enhancement of segmentation accuracy and model robustness in complex scenes. The proposed model is evaluated through experiments on three public datasets, and the results demonstrate significant improvements in image segmentation tasks. Compared to existing methods, the proposed module excels in handling segmentation details and boundaries while maintaining a high level of accuracy.

Relevant studies analysis

Research has mainly focused on the meaningful feature extraction for the medical image semantic segmentations. Efficiency is improved by using neural networks. The primary focus of research has been on extracting meaningful features for the medical images semantic segmentation. Additionally, the efficiency of segmentation has been enhanced through the utilization of neural networks.

Deep learning methods based on Convolutional Neural Networks (CNN) have demonstrated excellent performance in medical image segmentation. However, CNN only assigns categories to the entire image and does not classify individual pixels, making it unsuitable for image segmentation. In [5], a sophisticated image semantic segmentation method called "Fully Convolutional Network (FCN)" was proposed as an advancement of CNN. FCN substitutes a convolutional layer in the fully connected layer place. This configuration enables training from start to finish. This algorithm allows facilitating the per-pixel basis processing and gradual up sampling to enhance the feature dimension. Nonetheless, this approach sacrifices some detailed information from the original image, limiting further improvements in segmentation performance.

Ronneberger O. et al. [2] have developed the encoder-decoder network U-Net algorithm on the FCN network. This is a perfect tool for various types of medical image semantic segmentation tasks It fully utilizes both low-level and high-level features in the image through the use of skip connections mechanism.This is a perfect tool for various types of medical image semantic segmentation tasks.

Based on this method, a nested U-Net (U-Net++) network for medical image segmentation waspro-posed [3].The encoder and decoder subnetworks are connected through a series of nested dense skip paths in the proposed model. This method allows detailed information to be collected at different scales. This multi-scale feature fusion technique enhances the accuracy and detail retention of segmentation results. Subsequently [6] proposed an improved network Modified U-Net (mU-Net) based on the encoder-decoder network, which improves the accuracy of image segmentation by introducing an inverse convolutional layer and activation operations to skip the connected layers. However, they ignore the problem of information loss and blurring during feature extraction in each layer. To solve this problem and further improve the image segmentation performance, an improved encode-decode based network is designed in this research.

Feature aware module

While increasing the depth of a CNN can enhance the model's performance, it may also introduce challenges such as the vanishing gradient problem (gradients become very small and slow down learning) or the exploding gradient problem (gradients become very large and cause instability during training). To solve this problem, the idea of [7] is utilized to add shortcuts between layers in the learning paradigm, which is the proposed FAM, the flow chart of FAM is shown below in Fig. 1, *b*. The FAM divides the input of the network into two branches, the first of which first acquires the basic features of the image through three 3×3 convolutional and constant mappings, where each convolutional layer is followed by a batch normalization layer and a rectified linear unit (ReLU) nonlinear activation function.

Then, in order to make the module more sensitive to important features while suppressing unimportant features. Similar to the work of [8], the Squeeze and Excite (SE) module is used after the three convolutional networks, with the SE layer acting as a content-aware mechanism, re-weighting each channel accordingly to create a robust representation. First, the feature map is compressed by using global average pooling to obtain a global understanding of each channel. The squeezing operation produces a feature vector of size n, where n refers to the number of channels [9]. In the second step: excitation, this feature vector is fed through a two-layer feed-forward neural network, where the number of features is first reduced and then expanded to the original size *n*. This *n*-sized vector now represents the weights of the original feature map, which are used to scale each channel [10]. The input to the second branch is passed through 1×1 convolution and batchnorm. finally, the results of both branches of the module are subjected to cancatenate and relu operations. The computation of the proposed module can be denoted as the following formulas:

$$F_{3\times3}^{k} = \sigma(BN(W_{3\times3}^{k} \bullet I_{out}^{k})); \qquad (1)$$

$$I_{3\times3}^{k} = F_{3\times3}^{k}(I_{in}); \qquad (2)$$

$$O_{l} = SE(BN(Conv_{3\times3}(I_{out}^{3}))); \qquad (3)$$

$$O_r = Conv_{1 \times 1}(I_{in}); \tag{4}$$

$$Output = \sigma(O_l \oplus O_r), \tag{5}$$

where I_{in} is the raw input image or input feature map; $F_{3 \times 3}^k$ is the extracted features by convolutional weighting operation on the output feature map; I_{out}^k is the result of the kth convolution image, $k \in \{1, 2, 3\}$; σ is the weights of the convolution layer; $\bullet$ denotes the convolution function, σ represents the Relu activation function; $\oplus$ is the operation addition.

To address this issue, a solution proposed in [7] involves incorporating shortcuts between layers within the learning paradigm. The Residua Block, introduced in [10], adopts this approach. However, it is observed that the number of convolutional layers in each stage of the Residual Block is insufficient for effectively extracting essential feature information. To overcome this limitation, an extension called FAM module has been proposed.

Structure of FE-Net

For medical image segmentation, a novel model using the proposed FAM following the encoder-decoder architecture, which is referred to as FE-Net. The architecture of FE-Net is illustrated in Fig. 1, *a*. The classical U-Net, known for its ability to handle multi-resolution synthesis and decomposition, is employed by cascading encoders and decoders with skip connections [8]. Similarly, to address the diverse scales present in medical images, the proposed network adopts the classical encoder-decoder structure. In the proposed model, the encoder performs convolution and pooling (shrinking paths) on a given input image to create a compressed representation that can adapt to the scale of the medical image and extract local details. This compressed representation is then up-sampled in the decoder (expansion path) using deconvolution, generating a mask that highlights the region of interest. Additionally, skip connections are incorporated to concatenate these feature maps with the corresponding feature maps from the encoder layers. This merging of low-level and high-level semantic information enhances the generation of accurate masks. Specifically, the feature maps are concatenated after the skip connection, and the merged feature maps are fed into the FAM, which mitigates the problem of gradient vanishing and captures valid features. Unlike existing neural networks, the proposed network uses bilinear interpolation instead of transposed convolution in the up-sampling process of the decoder part, which can accelerate the model's computation speed and inference speed while ensuring smooth zooming of the image or feature maps, saving a lot of training memory consumption and time.
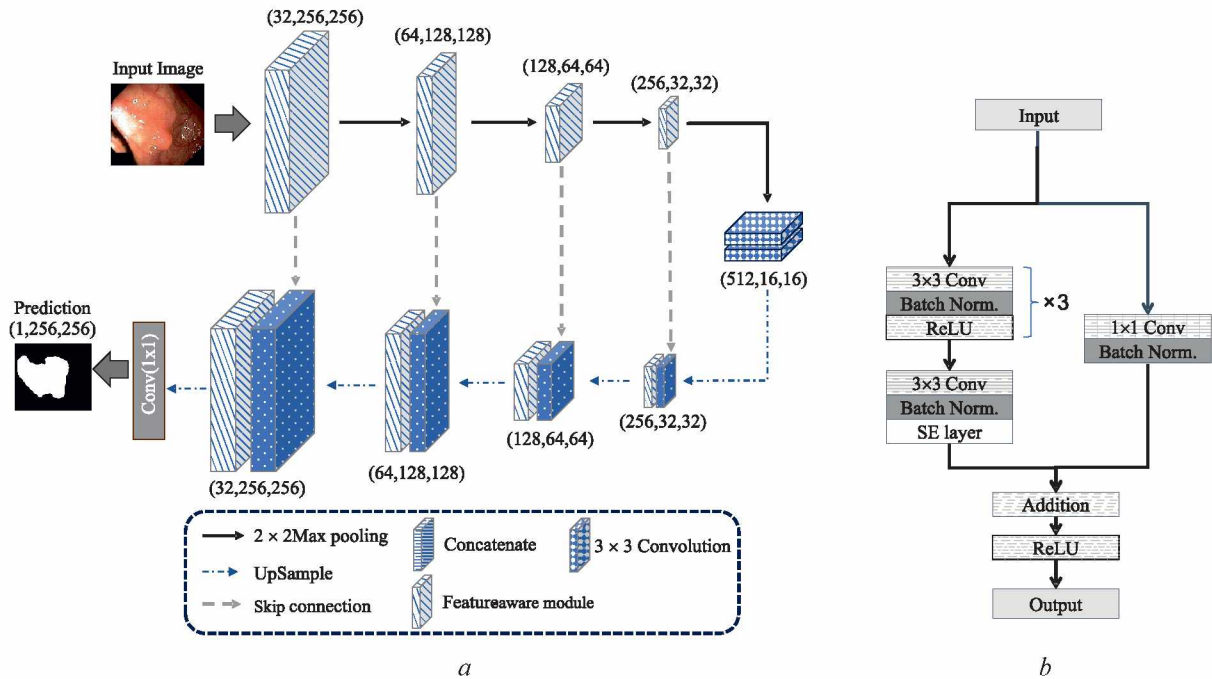

Fig. 1. The architecture of the proposed network and module:
a – architecture of FE-Net; *b* – flowchart of feature-aware module

In the proposed model, 32 filters are used in the first block and then 64, 128, 256, and 512 filters are used in each block in the network contraction path. The number of filters in each block in the network expansion path corresponds to the contraction path. Finally, a binary or multiclass segmentation mask is output using a 1×1 convolution, backed by a sigmoid layer.

Dataset description

To train and validate the proposed FE-Net models in this research, this research test it on three publicly available biomedical image datasets. The descriptions of the three datasets are as follows:

1) Kvasir-SEG: this data set contains 1.000 pictures of different resolutions. The ground truth mask corresponding to the picture consists of a large number of images annotated by professional doctors, including normal tissue, polyps, ulcers, and other lesion types;

2) CVC-ClinicDB: this dataset includes 612 RGB images, the resolution for each image is 384×288. These images include normal colon tissue, polyps, ulcers, and other structures and lesions. Different patients and different lens settings were taken into account during image acquisition to ensure diversity and realism in the data set;

3) 2018 Data Science Bowl: this data set contains 670 256×256 images. This data set provides a large amount of cell image data, including various types of cells, such as cancer cells, normal cells, etc.

For the above three datasets, they were split into training sets, verification machines and test sets through random sampling and in a ratio of 8:1:1. The details of the biomedical dataset used in the proposed experiments, denoting the total number of images, size of input images, number of training samples, number of validation samples, and number of test samples, are presented in Tab. 1.

Table 1. Details of the biomedical dataset

Dataset	Images	Size	Train	Validation	Test	Application
Kvasir-SEG	1000	Variables	800	100	100	Colonoscopy
CVC-ClinicDB	612	384×288	490	61	61	Colonoscopy
2018 Data Science Bowl	670	256×256	530	67	67	Nuclie

Evaluation metrics

In this research, various standard evaluation metrics are employed [15] to validate the effectiveness of FE-Net. These metrics include precision, dice coefficient (also known as F1 score), recall, and mean intersection over union (mIoU). All evaluation metrics can be calculated using TP and FP, where TP represents the number of true positives and FP represents the number of false positives.

Precision: an indicator to measure the proportion of true examples (True Positive) in the model prediction results. It indicates how many of the samples predicted as positive by the model are truly positive. It can be calculated in formula

$$\text{Precision} = \frac{TP}{TP+FP}. \tag{6}$$

Dice coefficient (also known as F1 score) is a metric used to evaluate the similarity between model predictions and real labels in an image segmentation task. It takes into account the degree of overlap between the predicted results and the real labels, and takes a value ranging from 0 to 1. Higher values indicate that the predicted results are more similar to the real labels. It can be calculated in formula

$$\text{Dice} = \frac{2TP}{2TP+FP+FN}. \tag{7}$$

Recall is a measure of the proportion of True Positive samples that the model successfully detects. It indicates how many of all True Positive samples are correctly detected by the model. It can be calculated in formula

$$\text{Recall} = \frac{TP}{TP+FN}. \tag{8}$$

Mean intersection over union is an average value used to evaluate the degree of overlap between the model predictions and the true labels in an image segmentation task. It calculates the ratio between the intersection and concatenation between the predicted results and the true labels and averages it overall categories. Assuming that there are a total of k classes, and its calculation formula is as below in formula

$$mIoU = \frac{1}{k} \sum_{i=0}^{k} \frac{TP}{TP + FP + FN}. \tag{9}$$

Data augmentation and experimental results

Today's deep learning researchers rely on large data sets to avoid overfitting problems and data imbalance. Usually the data sample size of biomedical images is small. Therefore, for the train dataset, the research choose to use data augmentation techniques. Random cropping and rotate the image by [90°, 180°, 270°] was utilized. In this way, new data sets are obtained, making the trained model more robust. However, for the test set, data augmentation was not used [12].

1. Results on Kvasir-SEG: Kvasir-SEG is a publicly available polyp segmentation dataset. This dataset is highly used for algorithm benchmarking. The proposed FE-Net has trained and compared with other widely used segmentation models. Tab. 2 shows the comparison results, the proposed methods outperforms other models on all most all metrics. It can be observed that the precision of FE-Net is 91.57 %, which is 8.70 and 1.52 % better than Attention U-Net [13] and recent FANet [10].

2. Results on CVC-ClinicDB: CVC-ClinicDB is a well-known dataset in the field of polyp segmentation, widely used for evaluating the performance of various algorithms. In proposed study, the proposed network was trained on the CVC-ClinicDB dataset and conducted a comprehensive comparison with other popular segmentation models. The results, as shown in Tab. 3, demonstrate the effectiveness and efficiency of the proposed approach in accurately segmenting polyps in the CVC-ClinicDB dataset. The proposed model achieved competitive performance and outperformed several state-of-the-art models in terms of segmentation accuracy and other evaluation metrics. Among them, the proposed model has a very large improvement in the ResU-Net++ [4] of SOTA, with improvements in recall and precision of 28.01 and 6.15 % respectively. Expect higher recall rates in clinical testing.

Table 2. Comparison of experimental results on Kvasir-SEG dataset

Dataset	Method	F1	mIoU	Recall	Precision
Kvasir-SEG	U-Net [2]	0.5969	0.4713	0.6171	0.6722
	ResU-Net++ [4]	0.6902	0.5721	0.7248	0.7454
	FCN [6]	0.8310	0.7365	0.8346	0.8817
	DoubleU-Net [14]	0.8129	0.7332	0.8402	0.8611
	U-Net++ [3]	0.8002	0.7000	0.8716	0.7992
	Attention U-Net [13]	0.7944	0.6959	0.8383	0.8287
	FANet [10]	0.8803	0.8153	0.9058	0.9005
	Authors	0.8896	0.8303	0.9035	0.9157

Table 3. Comparison of experimental results on CVC-ClinicDB dataset

Dataset	Method	F1	mIoU	Recall	Precision
CVC-ClinicDB	U-Net [2]	0.8230	0.7550	0.6472	0.7881
	ResU-Net++ [4]	0.7955	0.4545	0.6683	0.8877
	U-Net++ [3]	0.9377	0.8890	0.9405	0.9432
	Attention U-Net [13]	0.9325	0.8856	0.9276	0.9546
	Authors	0.9466	0.9008	0.9484	0.9492

3. Results on 2018 Data Science Bowl: the 2018 Data Science Bowl is a renowned competition that focuses on the challenging task of polyp segmentation. The proposed network was built to participate in the competition and trained proposed network using the provided dataset. The results obtained from the proposed model were compared against other top-performing models in the challenge.

As depicted in Tab. 4, the proposed approach showcased remarkable performance on the 2018 Data Science Bowl dataset, achieving high segmentation accuracy and demonstrating its capability to accurately detect and segment polyps. The proposed model's performance outperformed several strong competitors and showcased its potential for real-world application in polyp segmentation tasks [11].

Table 4. Comparison of experimental results on 2018 Data Science Bowl dataset

Dataset	Method	F1	mIoU	Recall	Precision
2018 Data Science Bowl	U-Net [2]	0.7573	0.9103	0.6059	0.9015
	DoubleU-Net [14]	0.7683	0.8407	0.6407	0.9596
	U-Net++ [3]	0.9117	0.8477	0.9203	0.9107
	Attention U-Net [13]	0.9179	0.8570	0.9183	0.9235
	Authors	0.9243	0.8642	0.9395	0.9860

Experimental details and visualization

The training and testing for FE-Net were conducted on Python version 3.9 with PyTorch 1.13.1 backend. The PyTorch library provides training and testing functions for the given datasets. The proposed method has been implemented in the code by creating a new FAM. The FAM divides the input into two branches, where the first branch extracts basic image features using three continuous convolutional layers (kernel size and padding size equal to 3 and 1 respectively), batch normalization, and ReLU activation. Additionally, the SE module is utilized after the convolutional layers to enhance sensitivity to important features through channel re-weighting. This is achieved by compressing the feature map, applying an averge pooling layer and a two-layer feed-forward network to obtain weights, and scaling each channel accordingly. Here, the compression ratio is 16. The input to the second branch undergoes 1×1 convolution (kernel size is 1 without padding) and batch normalization. Finally, the outputs of both branches are concatenated and passed through the ReLU activation function.

To achieve the best performance results, certain parameters need to be set for the neural network. The "train" function accepts the following parameters: images, labels, number of classes, epochs, and optimizer. The neural network adjusts its weights based on the results obtained after each iteration using different algorithms known as optimizers, which involve calculations on the difference between the results.In this implementation, the network utilizes stochastic gradient descent as the optimizer due to its simplicity in calculating the gradient of the network loss function. The initial learning rate value for all datasets was set to "1e-4". The learning rate decreases gradually during training. To ensure accurate timing during the training process, the model was warmed up before actual training [15]. The training batch size was set to 16, and the maximum epoch was set to 500.

Fig. 2, *b* below shows the variations in training loss on three different datasets. It can be observed that the model reaches convergence around the 300th epoch. Once the training of the model was completed, the model parameters were saved in *.pth format for future use in predicting new images. Fig. 2, *a* displays the predicted segmentation masks of the model on different datasets. It is evident that the proposed model performs well in terms of segmentation accuracy and detail retention.

Conclusion

1. The research aims are to improve the processing of segmentation details and boundaries in biomedical segmentation without compromising high accuracy. Thus, a novel encoder-decoder based model has been proposed for biomedical image segmentation. The model is comprised of the skip connection strategy and the feature aware module. The skip connections algorithm incorporating into Feature Enhancement Networks allows the reuse of features from the encoder and decoder layers. Thus the network to integrate both low-level and high-level information efficiently and improve the accuracy and precision of the generated masks. The feature aware module incorporates the idea of residual blocks by adding shortcut links between layers to facilitate feature learning. At the same time, the Squeeze and Excite module is embedded to increase the sensitivity of the module to important features and suppress unimportant features by reweighting each channel according to its importance.
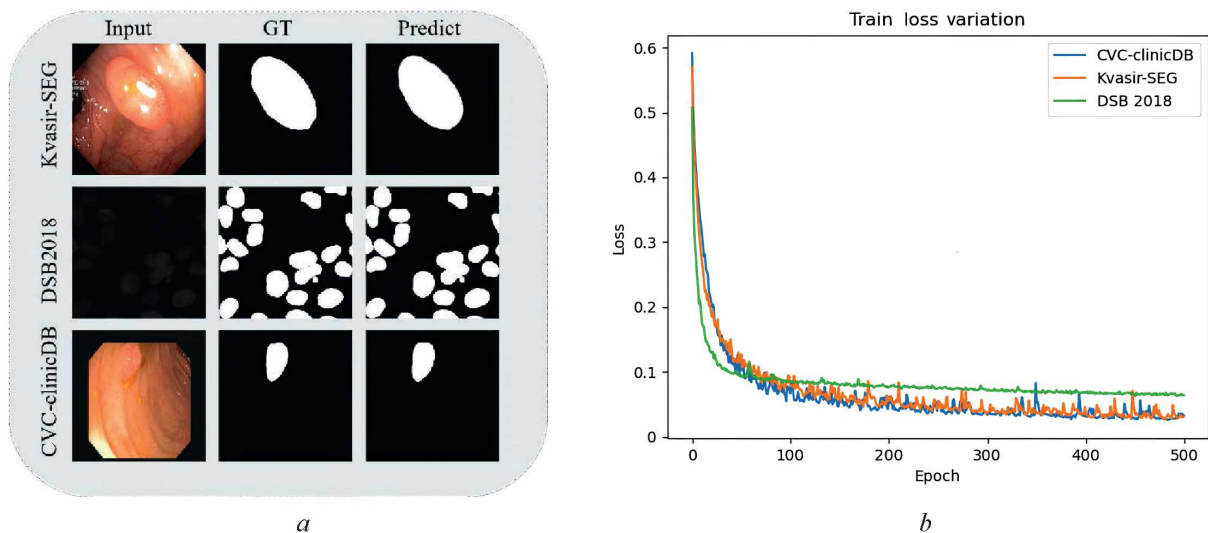
a   *b*

Fig. 2. Visualization and training loss variations:
a – comparison of the original image, ground truth mask and predicted mask
of the proposed FE-Net networks on the Kvasir-SEG, 2018 DSB, and DRIVE datasets;
b – training loss variations during 500 epochs

2. The proposed algorithm has been tested on three samples. The following parameters have been used to evaluate segmentation performance and generalization ability: Dice coefficient, recall, and Mean intersection over Union. The proposed algorithm is effective for data processing by object segmentation methods on the proposed medical image segmentation datasets.

3. The study showed the need to optimize the neural network structure. The next step is to improve performance and output speed for more medical image segmentation tasks. The research was carried out jointly with the magnetic resonance imaging department of the N. N. Alexandrov National Cancer Centre.

References

1. Litjens G., Kooi T., Bejnordi B. E., Adiyoso Setio A. A., Ciompi F., Ghafoorian M., et al. (2017) A Survey on Deep Learning in Medical Image Analysis. *Medical Image Analysis.* 60–88.
2. Ronneberger O., Fischer P., Brox T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, 18th International Conference.* 234–241.
3. Zhou Z., Rahman Siddiquee M. M., Tajbakhsh N., Liang J. (2018) Unet++: A Nested U-Net Architecture for Medical Image Segmentation. *In 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.* 3–11.
4. Jha D., Smedsrud P. H., Riegler M. A., Johansen D. (2019) Resunet++: An Advanced Architecture for Medical Image Segmentation. *In 2019 IEEE International Symposium on Multimedia.* 225–255.
5. Long J., Shelhamer E., Darrell T. (2015) Fully Convolutional Networks for Semantic Segmentation. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 3431–3440.
6. Seo H., Huang C., Bassenne M., Xiao R., Xing L. (2019) Modified U-Net (mU-Net) with Incorporation of Object-Dependent High Level Features for Improved Liver and Liver-Tumor Segmentation in CT Images. *IEEE Transactions on Medical Imaging.* 39 (5), 1316–1325.
7. He K., Zhang X., Ren S., Sun J. (2016) Deep Residual Learning for Image Recognition. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 770–778.
8. Hu J., Shen L., Sun G. (2018) Squeeze-and-Excitation Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 7132–7141.
9. Zhang Z., Liu Q., Wang Y. (2018) Road Extraction by Deep Residual U-Net. *IEEE Geoscience and Remote Sensing Letters.* 15 (5), 749–753.
10. Tomar N. K., Jha D., Riegler M. A., Johansen H. D., Johansen D., Rittscher J., et al. (2022) Fanet: A Feedback Attention Network for Improved Biomedical Image Segmentation. *IEEE Transactions on Neural Networks and Learning Systems.*

11. Yang C., Guo X., Zhu M., Ibragimov B., Yuan Y. (2021) Mutual-Prototype Adaptation for Cross-Domain Polyp Segmentation. *IEEE Journal of Biomedical and Health Informatics*. 25 (10), 3886–3897. DOI: 10.1109/JBHI.2021.3077271.
12. Chlap P., Min H., Vandenberg N., Dowling J., Holloway L., Haworth A. (2021) A Review of Medical Image Data Augmentation Techniques for Deep Learning Applications. *Journal of Medical Imaging and Radiation Oncology*. 65 (5), 545–563. DOI: 10.1111/1754-9485.13261.
13. Oktay O., Schlemper J., Folgoc L. L., Lee M., Heinrich M., Misawa K., et al. (2018) Attention U-Net: Learning Where to Look for the Pancreas. *1st Conference on Medical Imaging with Deep Learning (MIDL 2018), Amsterdam, the Netherlands.*
14. Jha D., Riegler M. A., Johansen D., Halvorsen P., Johansen H. D. (2020) Doubleu-Net: A Deep Convolutional Neural Network for Medical Image Segmentation. *In 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems*. 558–564.
15. Iqbal A., Sharif M., Khan M. A., Nisar W., Alhaisoni M. (2022) FF-UNet: A U-Shaped Deep Convolutional Neural Network for Multimodal Biomedical Image Segmentation. *Cognitive Computation*. 14 (4), 1287–1302.

Authors' contribution

The authors contributed equally to the writing of the article.

Information about the authors

Di Zhao, Postgraduate at the Department of Information Technologies in Automated Systems, Belarusian State University of Informatics and Radioelectronics (BSUIR)

Yi Tang, Master's Student at the Department of Information Technologies in Automated Systems, BSUIR

Gourinovitch A. B., Assistant Professor at the Department of Information Technologies in Automated Systems, BSUIR

Address for correspondence

220013, Republic of Belarus,
Minsk, Brovki St., 6
Belarusian State University of Informatics and Radioelectronics
Tel.: +375 29 622-69-34
E-mail: gurinovich@bsuir.by
Gourinovitch Alevtina Borisovna