UDC 004.8

ENSEMBLE METHOD BASED ON AVERAGING SHAPES OF OBJECTS USING THE PYRAMID METHOD

Koniukhov V. D. – Postgraduate student, A. Pidhornyi Institute of Power Machines and Systems of NAS of Ukraine, Kharkiv, Ukraine.

ABSTRACT

Context. Image segmentation plays a key role in computer vision. The quality of segmentation is affected by many factors: noise, artifacts, complex shapes of objects. Classical methods cannot always guarantee good success, depending on the quality of the image and the existing noise, they cannot always achieve the desired result. The proposed method uses an ensemble of neural networks, which makes it possible to increase the accuracy and stability of segmentation.

Objective. The goal of the work is to develop a new method of combining predictions of neural network ensembles, which can improve segmentation accuracy by combining images of different image sizes.

Method. A method is proposed that averages the shapes of objects depicted on prediction masks. A pyramid of images is used to improve segmentation quality, each level of the pyramid corresponds to an increased size of the original image. This approach allows obtaining image characteristics at different levels. For a test image, a prediction is obtained from each neural network in the ensemble, after which a pyramid is built for the image. All pyramid levels are combined into the final image using SAAMC. All obtained final images for each neural network are also combined at the end using SAAMC. The use of an ensemble of neural networks combined with the pyramid method allows for reducing the impact of noise and artifacts on the segmentation results.

Results. The use of this method was compared with the usual use of individual neural networks and the ensemble averaging method. The obtained results show that the proposed method outperforms its competitors. Application of the proposed method improved the accuracy and quality of segmentation.

Conclusions. The conducted research confirmed the sense of using an ensemble of neural networks and creating a new method of combining predictions. The use of an ensemble of neural networks makes it possible to compensate for the errors and shortcomings of individual neural networks. Using the proposed method can significantly reduce the impact of noise and artifacts on segmentation. Further study and modification of this method will make it possible to further improve the quality of segmentation.

KEYWORDS: machine learning; image recognition; neural network; image segmentation, computer vision.

ABBREVIATIONS

DSC is a Dice-Sorensen coefficient;

DCNN is a Deep Convolutional Neural Network;

FCN is a Fully Convolutional Network;

FPN is a Feature Pyramid Network;

SAAMC is a Shape Averaging with Alignment to a Mean Center.

NOMENCLATURE

A is a set of pixels of ground truth mask;

$\text{Accuracy}_{\text{new}}$ is an accuracy of new method;

$\text{Accuracy}_{\text{competitor}}$ is an accuracy of competitive method;

B is a set of pixels of a predicted mask;

C_i is i-th center of mass;

C_{mean} is a mean center of mass;

d is a difference between two distance maps;

d_{all} is a difference between all distance maps;

$dt()$ is a distance transformation function;

$f(x)$ is a proposed method;

i is an index of an element;

I_i is i-th image;

I_{res} is a resulting image;

I_{resized} is a resized image;

k is a pyramid level number;

mask_i is i-th image mask;

~mask is an inverted image mask;

M is a length of one side of an image;

n is a number of elements;

N is a number of pyramid levels;

O is a length of one side of the input image;

P_i is an i-th prediction;

R_i is an image size at level i;

S is a set of images;

x_{mean} is a mean value of x coordinates;

y_{mean} is a mean value of y coordinates.

INTRODUCTION

Modern methods of image segmentation are mostly based on neural networks. The use of neural networks for segmentation tasks of various types of images has proven its effectiveness compared to classical methods [1]. Despite the achievements brought by the use of neural networks, there were still some difficulties that had to be faced: noise, artifacts, complex objects, instability of results. For this, new methods were developed, as well as new neural networks that could minimize the impact of various undesirable conditions on the segmentation result.

One of the improved approaches to solving such problems was the U-Net neural network [2], which had significant success in segmentation of medical images. But neural networks still continued to face significant challenges. For this, ensemble methods were proposed. This approach included combining the results of several neural networks and improving the quality of segmentation. Thanks to ensemble methods, it became possible to increase the accuracy and reliability of segmentation. Ensemble methods and image segmentation were not ignored, the proposed methods: averaging, weighted averaging, training one network on different data demon-

DOI 10.15588/1607-3274-2024-4-11

strated their effectiveness compared to previously created methods. The main role was played by the methods of unification. The ease of their use is one of the advantages. This article proposes to consider a new method that combines the advantage of combining masks and using different image sizes to extract image features at different levels. This method is aimed at increasing the accuracy of segmentation and smoothing out the bad influence of low-quality images.

The object of study is performance of the proposed method. The main focus is on combining the predictions of different neural networks, as well as using a different number of pyramid levels to extract different image characteristics. This study includes an analysis of the methods used and an assessment of their effectiveness on two data sets.

The subject of study is ensemble methods and the pyramid method. Existing methods have their pros and cons. The application of these methods to different data sets may have different results. The development of a new method that can improve the accuracy of segmentation and minimize the impact of existing noise is a necessity.

The purpose of the work is to develop a new ensemble method that can improve the accuracy of image segmentation and test its effectiveness.

1 PROBLEM STATEMENT

Image segmentation is a complex task in computer vision. Segmentation problems are especially relevant when using this approach in complex areas such as medicine, self-driving cars, robotics. The use of segmentation in such areas requires high reliability and accuracy of results, which traditional methods cannot cope with. One of the most difficult areas in which it is necessary to guarantee maximum reliability and accuracy of segmentation is the segmentation of medical images. The presence of noise, artifacts, incorrect exposure or the influence of poor-quality equipment creates significant difficulties in performing successful segmentation. Let us assume that a set of images of unknown quality is given:

$$S = \left\{I_i\right\}_{i=1}^{n}.$$

For a given data set, the problem of developing a new ensemble method can be represented as:

$$\text{Accuracy}_{\text{new}} = f\left(S\right) \to \max.$$

To solve the problems associated with poor segmentation quality, it is necessary that the proposed method be better than its competitors:

$$\text{Accuracy}_{\text{new}} > \text{Accuracy}_{\text{competitor}}.$$

2 REVIEW OF THE LITERATURE

Ensemble methods have proven themselves to be highly effective methods for solving complex problems. The main idea of the ensemble approach is to combine several neural networks, which makes it possible to smooth out the errors of individual networks. Thanks to the use of ensembles, it became possible to combine different architectures, which in turn makes it possible to take into account different aspects and extract the necessary characteristics in different ways. Different neural networks can produce results differently, depending on the data used. Combining them makes it possible to minimize the receipt of erroneous predictions and get a better result.

The earliest ensemble methods were: bagging [3], boosting [4] and stacking [5]. Bagging allows training several models on different subsets of data, and at the end either voting for the final result or averaging. Boosting is based on the idea of improving each subsequent model. This method is used when a neural network gets bad results. Each neural network in the ensemble learns from the errors of the previous one. In stacking, several models are trained, after which their predictions are passed to a meta-model, which forms the final prediction based on the data.

In their paper [6], the authors present an analysis of published works on ensemble learning and ensemble deep learning. A new approach for retinal vessel segmentation based on deep ensemble learning is presented in the study [7]. The obtained results showed that the proposed model outperforms existing methods on different datasets. In the paper [8], a deep ensemble model was proposed for classification of histopathological images of the breast. The proposed method was compared with others, as a result of which it was able to provide an advantage of 5-20%. Ensemble learning for brain tumor classification was proposed in the paper [9]. As a preprocessing of the data, the authors used Gabor filters to remove noise. They used three models: EfficientNet, DenseNet and MobileNet in order to achieve diversity in feature extraction. The method proposed by the authors showed good results and promising performance. In the paper [10], the authors proposed a method for using ensemble learning to detect skin cancer. The authors argue that the use of individual models is not as reliable as desired and they cannot achieve the required accuracy. In turn, thanks to the use of an ensemble of neural networks, it was possible to achieve a better result.

A three-phase approach for sclera segmentation in eye images is presented in the paper [11]. In their study, the authors use five deep learning models: Unet-DenseNet, FPN with DenseNet, Unet-ResNet50, TransUnet, and Swin-UNet. Using such models in an ensemble ensures a different approach to extracting the necessary features and improving the quality of segmentation. With this approach, the authors minimized the impact of errors and were able to achieve a good result. The paper [12] discusses a heterogeneous ensemble approach based on DCNN. In their work, the authors use weighted averag-

DOI 10.15588/1607-3274-2024-4-11

ing. They assign larger weights to the models that showed the best result during testing. This approach helped to generalize the distribution and prove the advantage of using the ensemble method.

In the paper [13], the authors propose a pyramid attention method for image restoration. This method captures features that are located far from each other from a multi-scale pyramid of features. The proposed approach can be easily integrated into different neural architectures. A multi-scale feature pyramid fusion network is proposed in the paper [14]. The authors conducted the study on three datasets and compared their method with U-Net. Ultimately, the authors were able to achieve better results than the competitive network. Spectral-Spatial Feature Pyramid Network was proposed in [15]. This network extracts multi-scale spectral information and multi-scale spatial information using the attention mechanism and feature pyramid structure. The obtained results indicate that the network can achieve good results. In the paper [16], the authors propose a context-aware network with a two-stream pyramid (CANet). They use this network to segment medical images on three datasets. Using this approach gave a positive result compared to 13 competitive methods. By using a pyramid of multi-scale features [17], the authors were able to demonstrate the effectiveness of the pyramidal multi-scale structure for segmentation of histopathological images. The authors focus on adapting to changes in image resolution, which helps to improve the quality of segmentation.

3 MATERIALS AND METHODS

At the beginning of the construction of the proposed algorithm, n number of neural networks is selected, which will comprise the ensemble.

First stage. Let there be an I_i image of size M×M, in grayscale. It is necessary to build a pyramid of levels for it. This means building an image pyramid that will accept an image of size from the data set as input, will contain n levels, where at each level the size of the original image will increase in size. In this study, the following image sizes were used in the pyramid: 64×64, 128×128, 256×256, 512×512, 1024×1024, 2048×2048 and 4096×4096. The pyramid can be designated as:

$$\text{Pyramid} = \{I_1, I_2 \ldots I_n\},$$

where at each level of the pyramid the image is designated as:

$$I_k^{R_i}.$$

The image size was determined as follows:

$$R_i = \{(O \times 2^{i-1}) \times (O \times 2^{i-1}) i \in [1, \ldots, N]\}.$$

Second stage. After the pyramid has been built, it is necessary to obtain predictions for each level of the pyramid. The size of the prediction mask is equal to the size of the image for which the prediction was made.

Third stage. Here it is necessary to bring each level of the pyramid to a single size:

$$\forall i \in \{1,2 \ldots N\} : \text{Resize}(I_i) = I_{\text{resized},i}^{M \times M}.$$

After which all the images that are in the pyramid will be presented in the same size.

Fourth stage. The method of combination of predictions must be applied to all the obtained images at the third stage, in this case SAAMC will be used. First, it is necessary to find the average value of the center for all the predictions used. Let n predictions be given $\{P_1, P_2 \ldots P_n\}$, each object on the mask has a center of mass $\{C_1, C_2 \ldots C_n\}$. The average value of the center for all objects will be $C_{\text{mean}} = (x_{\text{mean}}.\ y_{\text{mean}})$, where:

$$x_{\text{mean}} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

and

$$y_{\text{mean}} = \frac{1}{n} \sum_{i=1}^{n} y_i.$$

After the average center has been obtained, it is necessary for each prediction P_i to shift the object relative to its new center – the mean center. When all objects have been shifted relative to the mean center, it is necessary to average the shapes of the objects using the formula:

$$d_{\text{all}} = \sum_{i=1}^{n} d(\text{mask}_i),$$

where $d(mask)$ is the function given below:

$$d = dt(\text{mask}) - dt(\sim \text{mask}).$$

The distance transformation method calculates the minimum distance from the object pixel to the background pixel. The final averaging of the object shapes is as follows:

$$I_{\text{res}} = d_{\text{all}} > 0.$$

To average the shapes of objects, it is necessary to sum up the obtained results and if the sum is greater than 0, then the pixels become white, otherwise black.

Fifth stage. For each neural network in the ensemble, the final result was obtained at the fourth stage. At this stage, it is necessary to apply SAAMC, which was described in the fourth stage, to all images obtained at the fourth stage. This stage is similar to the fourth, only the input images need to be changed.

DOI 10.15588/1607-3274-2024-4-11

To summarize all of the above, we can outline the proposed method in the form of a list of tasks:

1. Create a pyramid.
2. Get predictions for all levels of the pyramid.
3. Resize all images resulting from the second step.
4. Perform averaging of objects depicted in the predictions obtained in the third stage.
5. Perform averaging of objects depicted in the predictions obtained in the fourth stage.

Visually, all stages of the proposed method can be seen in Figure 2, the stage number is indicated in the circle.

4 EXPERIMENTS

Data obtained from open sources were used to conduct the experiments [18, 19]. These were X-rays and lung masks for them. In order to conduct the study in more detail, it was decided to split this data set into two. As a result, two sets were obtained: one contains masks of the left lung, the second masks of the right lung. Each dataset contained 703 images, where 630 images were used for training and the remaining 73 were used for testing. All images were presented in 512×512 size and in grayscale. An example of the images used can be seen in Figure 1. On the left in the image is the mask of the right lung, on the right is the mask of the left lung.

The neural network used was FCN8-MobileNet. It combines the FCN architecture and the MobileNet base model. This network combines two architectures for more efficient feature extraction and improved segmentation. Its main advantage is also its lightweight architecture, which helps reduce the number of adjustable parameters and increase its speed.

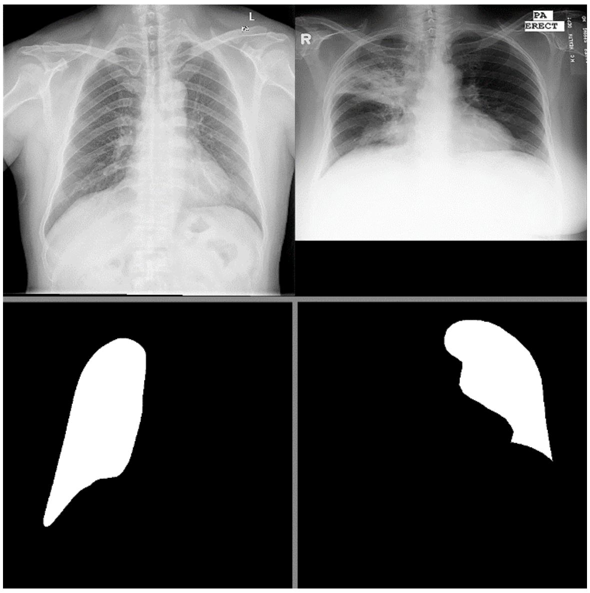

Figure 1 – The example of the images used and their masks

To check the similarity of the prediction and the ground truth mask, the Dice-Sorensen coefficient was used. It was represented by the following formula:

$$DSC = \frac{2|A \cap B|}{|A| + |B|}.$$

First of all, it was necessary to study the behavior of neural networks separately. For this purpose, 10 neural networks were trained on each data set. The next step was to obtain the DSC for different input image sizes. For this, one neural network was used, and the same operations were performed for each data set. An image resizing operation was applied to each test image. The following dimensions were used for this purpose: 64×64, 128×128, 256×256, 512×512, 1024×1024, 2048×2048 and 4096×4096. For each image with different sizes, a prediction was obtained. Then, for all images, the average DSC value was obtained and written down in a table.

The following experiment was used to justify the use of the ensemble. For this, an ensemble of 10 networks was used. Two methods were used to combine predictions in the ensemble: the well-known averaging method and SAAMC.

The use of the pyramid method for neural networks separately was also considered. For the combination of images obtained at each level of the pyramid, the same two methods were used: the averaging method and SAAMC. Table 1 shows the sizes of the images that were used to create the pyramids. The cross indicates the sizes of the images at each level of the pyramid. The images used are square, so only one side will be listed in the table.

Table 1 – Image sizes used at pyramid levels

Name	Image sizes (M×M)						
	64	128	256	512	1024	2048	4096
A	X	X	X				
B		X	X	X			
C				X	X		
D				X	X	X	
E				X	X	X	X
F	X	X	X	X	X	X	X

And at the very end, the use of an ensemble of neural networks for the pyramid method and methods of combining predictions was considered.

5 RESULTS

The DSC for all 20 trained neural networks is presented in Table 2. The obtained results indicate a good degree of segmentation. The small spread between neural networks indicates the stable operation of the model used.

Table 3 clearly shows how image size affects segmentation accuracy. The datasets used images of 512×512 pixels. Comparing this size with others, we can see that for this model, reducing the image size only worsened the situation. Increasing the image size, on the contrary, increased the accuracy of the prediction. For both sets, the best result was using images of 4096×4096.

The advantage of the ensemble is shown in Table 4. The segmentation quality of the ensemble is higher than the average quality of the neural networks separately. For

DOI 10.15588/1607-3274-2024-4-11

both data sets, the ensembles showed a better result. And SAAMC showed better accuracy than its analogue.

The studies of the use of pyramids with different levels for single networks were presented in Tables 5 and 6. They considered two different methods of combining level images. In Table 5, for the left and right lungs, the best result was using a pyramid that contained levels with the following image sizes: 512×512, 1024×1024, 2048×2048, 4096×4096. The best result for the left lung was able to exceed the average value for single networks. For the right lung, the situation is worse, the result obtained was less than the average of single networks.

In Table 6, for the left lung, to achieve the best result, it was necessary to use a pyramid with the following levels: 512×512, 1024×1024, 2048×2048, 4096×4096. For the right lung, the following sizes helped to achieve a good result: 512×512, 1024×1024, 2048×2048. In both cases, the results obtained exceeded the average for single networks.

The results obtained using the proposed method are in Table 7. For the averaging method, to achieve better results, it was necessary to use a pyramid with the following levels: 512×512, 1024×1024, 2048×2048, 4096×4096.

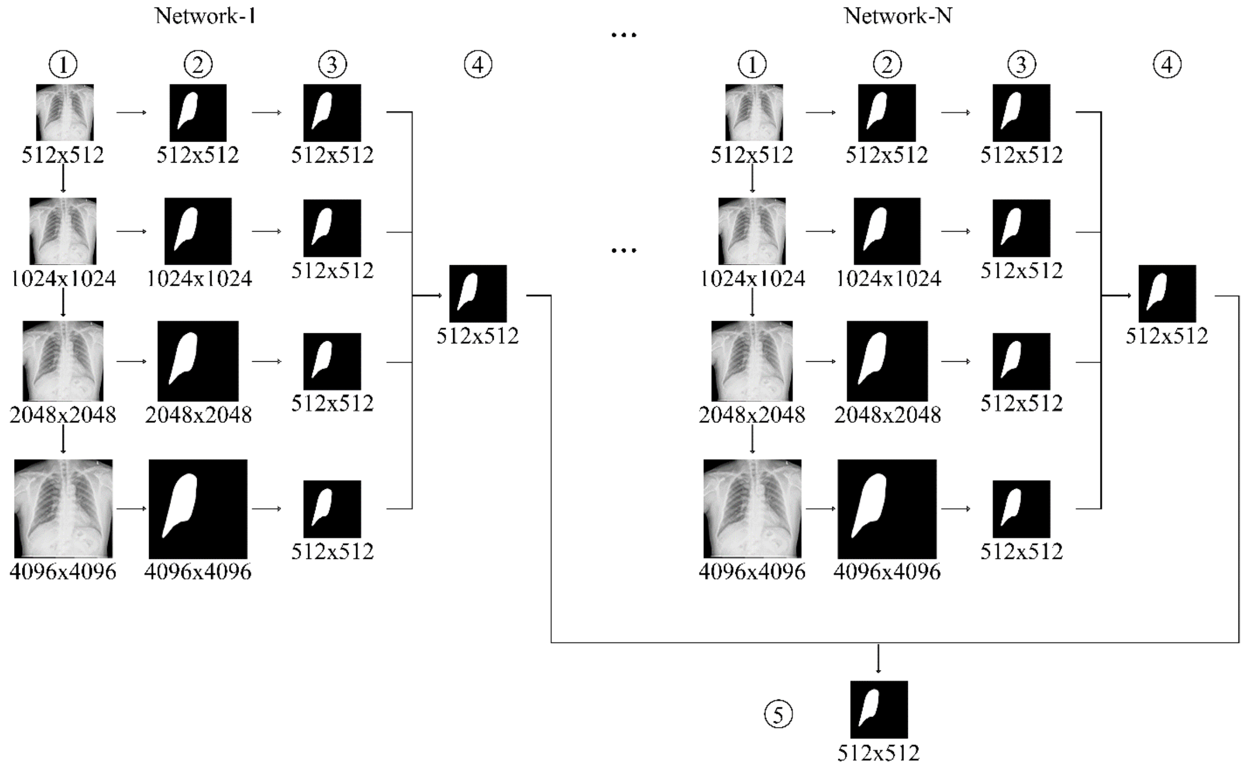

Figure 2 – Step-by-step example of the algorithm's operation.

Table 2 – Similarity measure for trained neural networks

Dataset	Neural network number										Statistics		
	1	2	3	4	5	6	7	8	9	10	Min	Max	Mean
Left lung	0.9540	0.9564	0.9559	0.9554	0.9553	0.9548	0.9575	0.9574	0.9560	0.9597	0.9540	0.9597	0.9562
Right lung	0.9485	0.9553	0.9519	0.9532	0.9488	0.9561	0.9580	0.9567	0.9568	0.9558	0.9485	0.9580	0.9541

Table 3 – Similarity measure for one neural network for different image sizes

Dataset	Image size (M×M)						
	64	128	256	512	1024	2048	4096
Left lung	0.8906	0.9356	0.9478	0.9540	0.9555	0.9563	0.9565
Right lung	0.8360	0.9219	0.9408	0.9485	0.9490	0.9499	0.9502

Table 4 – Similarity measure for an ensemble of neural networks

Dataset	Combining method	
	Mean averaging	SAAMC
Left lung	0.9599	0.9638
Right lung	0.9588	0.9610

DOI 10.15588/1607-3274-2024-4-11

Table 5 – Similarity measure using the pyramid method with mean averaging

Dataset	Approach	Neural network number										Statistics
		1	2	3	4	5	6	7	8	9	10	Mean
Left lung	A	0.9210	0.9272	0.9289	0.9241	0.9211	0.9297	0.9357	0.9319	0.9269	0.9308	0.9277
	B	0.9418	0.9457	0.9471	0.9464	0.9417	0.9455	0.9499	0.9484	0.9454	0.9498	0.9462
	C	0.9542	0.9567	0.9559	0.9555	0.9555	0.9549	0.9576	0.9575	0.9563	0.9598	0.9564
	D	0.9543	0.9563	0.9574	0.9561	0.9544	0.9560	0.9588	0.9585	0.9561	0.9602	0.9568
	E	0.9557	0.9577	0.9585	0.9570	0.9560	0.9573	0.9600	0.9598	0.9573	0.9613	0.9580
	F	0.9539	0.9559	0.9572	0.9557	0.9540	0.9557	0.9586	0.9583	0.9557	0.9599	0.9565
Right lung	A	0.9007	0.9278	0.9137	0.9172	0.9154	0.9220	0.9274	0.9291	0.9312	0.9267	0.9007
	B	0.9335	0.9477	0.9377	0.9409	0.9370	0.9457	0.9501	0.9497	0.9485	0.9459	0.9335
	C	0.9487	0.9555	0.9520	0.9533	0.9489	0.9562	0.9581	0.9568	0.9569	0.9559	0.9487
	D	0.9479	0.9556	0.9515	0.9526	0.9489	0.9551	0.9578	0.9566	0.9569	0.9556	0.9479
	E	0.9492	0.9563	0.9525	0.9538	0.9499	0.9559	0.9583	0.9571	0.9576	0.9566	0.9492
	F	0.9472	0.9554	0.9509	0.9520	0.9485	0.9548	0.9575	0.9564	0.9566	0.9552	0.9472

Table 6 – Similarity measure using the pyramid method with SAAMC

Dataset	Approach	Neural network number										Statistics
		1	2	3	4	5	6	7	8	9	10	Mean
Left lung	A	0.9424	0.9474	0.9362	0.9377	0.9498	0.9425	0.9456	0.9476	0.9447	0.9494	0.9443
	B	0.9539	0.9578	0.9529	0.9522	0.9574	0.9536	0.9563	0.9562	0.9552	0.958	0.9554
	C	0.9560	0.9571	0.9578	0.9562	0.9558	0.9562	0.9590	0.9589	0.9571	0.9605	0.9575
	D	0.9579	0.9591	0.9587	0.9575	0.9580	0.9580	0.9608	0.9606	0.9585	0.9622	0.9591
	E	0.9580	0.9590	0.9594	0.9577	0.9577	0.9584	0.9612	0.9609	0.9586	0.9624	0.9593
	F	0.9551	0.9580	0.9533	0.9518	0.9576	0.9550	0.9582	0.9580	0.9561	0.9600	0.9563
Right lung	A	0.9229	0.9342	0.9157	0.9312	0.9260	0.9415	0.9343	0.9335	0.9370	0.9345	0.9311
	B	0.9448	0.9515	0.9468	0.9501	0.9450	0.9548	0.9545	0.9528	0.9532	0.9525	0.9506
	C	0.9486	0.9561	0.9522	0.9533	0.9496	0.9562	0.9581	0.9570	0.9571	0.9561	0.9544
	D	0.9506	0.9568	0.9539	0.9550	0.9510	0.9572	0.9590	0.9577	0.9580	0.9575	0.9557
	E	0.9505	0.9568	0.9536	0.9548	0.9511	0.9569	0.9588	0.9578	0.9580	0.9575	0.9556
	F	0.9425	0.9510	0.9421	0.9492	0.9435	0.9548	0.9528	0.9515	0.9529	0.9530	0.9493

Table 7 – Similarity measure for an ensemble using the pyramid method with SAAMC

Dataset	Approach	Method	
		Mean averaging	SAAMC
Left lung	A	0.9327	0.9524
	B	0.9505	0.9629
	C	0.9607	0.9646
	D	0.9615	0.9666
	E	0.9627	0.9666
	F	0.9611	0.9634
Right lung	A	0.9268	0.9396
	B	0.9485	0.9573
	C	0.9594	0.9612
	D	0.9592	0.9623
	E	0.9602	0.9622
	F	0.9588	0.9561

When applying the proposed method for the left lung, it was necessary to use a pyramid with the following level sizes: 512×512, 1024×1024, 2048×2048. For the right lung: 512×512, 1024×1024, 2048×2048, 4096×4096. As can be seen from the tables, in all cases SAAMC was better than average averaging. For the left lung, the improvement was 1.034, for the right 0.8236 compared to using single networks.

6 DISCUSSION

Starting from studying the simple use of ensembles, one can already notice the advantage of SAAMC over averaging. The study of the use of different sizes presented in Table 3 shows that the network that was used extracted features from large-size images better. In this case, the use of the pyramidal method made it possible to improve the quality of segmentation by extracting from different levels. This suggests that using higher resolution images allows the model to detect finer details.

By combining the results of 10 networks, we were able to prove the advantage of using the ensemble method. Using the ensemble allows us to smooth out the shortcomings of individual networks and enhance the final result. In this way, we can minimize the impact of noise, artifacts, and other undesirable moments.

The most important thing was the confirmation of the advantage of SAAMC over averaging and using single networks.

One of the main advantages of the method used is its ease of use and guaranteed results. Using FPN imposes a limitation in the form of increased computational costs. The higher the image resolution, the longer it takes to process, which means the model itself will take many times longer to learn. It is worth noting that the improvement in accuracy was noticed just when the resolution increased and larger images were used. This proves the main advantage of this method. To use it, there is no need to train a special neural network, as well as configure a

DOI 10.15588/1607-3274-2024-4-11

large number of parameters. All we need is to simply apply the proposed algorithm to an ensemble of neural networks. Naturally, we can combine different architectures, which can further improve the accuracy of the prediction. And using FPN as neural networks for an ensemble can also improve accuracy, but do not forget about the speed.

CONCLUSIONS

A new ensemble method for combining predictions was proposed. The proposed method demonstrated its effectiveness in improving the quality of segmentation. Using the pyramid method as a basis for the developed method made it possible to extract features from different levels. Using the ensemble approach made it possible to smooth out the shortcomings of individual neural networks and combine their strengths. The main advantage of the method is its simplicity in implementation compared to the creation of FPN networks.

The scientific novelty of the obtained results is for the first time, a combination method was proposed that combines the advantages of the pyramid method and SAAMC. The results obtained confirmed the effectiveness of this method. The use of different resolutions at different levels of the pyramid made it possible to extract more detailed data, and the use of an ensemble of neural networks made it possible to combine the strengths of the neural networks used.

The practical significance of obtained results is that the use of the proposed method was tested on two medical data sets containing X-ray images. The use of this method can be easily implemented in automated systems for segmentation of medical images.

Prospects for further research are further developments of modifications of this method for further improvement of segmentation quality. For example, studies of using a large number of neural networks in ensembles, as well as using different architectures.

ACKNOWLEDGEMENTS

I would like to express my gratitude to A. Pidhornyi Institute of Power Machines and Systems of NAS of Ukraine for the opportunity to conduct this study.

REFERENCES

1. Voulodimos A., Doulamis N., Doulamis A., Protopapadakis E. Deep Learning for Computer Vision: A Brief Review, *Computational intelligence and neuroscience,* 2018, P. 2018:7068349. DOI: 10.1155/2018/7068349
2. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation, *ArXiv,* 2015. DOI: 10.1007/978-3-319-24574-4_28
3. Breiman L. Bagging predictors, *Machine Learning,* 1996, Vol. 24, No. 2, pp. 123–140. DOI: 10.1007/BF00058655
4. Freund Y., Schapire R. E. A decision-theoretic generalization of on-line learning and an application to boosting, *Journal of Computer and System Sciences,* 1997, Vol. 55, No. 1, pp. 119–139. DOI: 10.1006/jcss.1997.1504
5. Wolpert D. H. Stacked generalization, *Neural Networks.* – 1992, Vol. 5, No. 2, pp. 241–259. DOI: 10.1016/S0893-6080(05)80023-1
6. Yang Y., Lv H., Chen N. A survey on ensemble learning under the era of deep learning, *Artificial Intelligence Review,* 2023, Vol. 56, pp. 5545–5589. DOI: 10.1007/s10462-022-10283-5
7. Du L., Liu H., Zhang L., Lu Y., Li M., Hu Y., Zhang Y. Deep ensemble learning for accurate retinal vessel segmentation, *Computers in Biology and Medicine,* 2023, Vol. 158, P. 106829. DOI: 10.1016/j.compbiomed.2023.106829
8. Zheng Y., Li C., Zhou X., Chen H., Xu H., Li Y., Zhang H., Li X., Sun H., Huang X., Grzegorzek M. Application of transfer learning and ensemble learning in image-level classification for breast histopathology, *Intelligent Medicine,* 2023, Vol. 3, No. 2, pp. 115–128. DOI: 10.1016/j.imed.2022.05.004
9. Vaiyapuri J., Mahalingam S., Ahmad H. A. M., Abdeljaber E., Yang E., Jeong S. Y. Ensemble learning driven computer-aided diagnosis model for brain tumor classification on magnetic resonance imaging, *IEEE Access,* 2023, Vol. 11, pp. 91398–91406. DOI: 10.1109/ACCESS.2023.3306961
10. Tembhurne J. V., Hebbar N., Patil H. Y. et al. Skin cancer detection using ensemble of machine learning and deep learning techniques, *Multimedia Tools and Applications,* 2023, Vol. 82, pp. 27501–27524. DOI: 10.1007/s11042-023-14697-3
11. Al-Zebari A., Ensemble convolutional neural networks and transformer-based segmentation methods for achieving accurate sclera segmentation in eye images, *Signal, Image and Video Processing*, 2024, Vol. 18, pp. 1879–1891. DOI: 10.1007/s11760-023-02891-7
12. Iqball T., Wani M. A. Weighted ensemble model for image classification, *International Journal of Information Technology,* 2023, Vol. 15, pp. 557–564. DOI: 10.1007/s41870-022-01149-8
13. Mei Y., Fan Y., Zhang Y., Yu J., Zhou Y., Liu D., Fu Y., Huang T., Shi H. Pyramid attention network for image restoration, *International Journal of Computer Vision,* 2023, Vol. 131, pp. 3207–3225. DOI: 10.1007/s11263-023-01843-5
14. Zhang B., Wang Y., Ding C., Deng Z., Li L., Qin Z., Ding Z., Bian L., Yang C. Multi-scale feature pyramid fusion network for medical image segmentation, *International Journal of Computer Assisted Radiology and Surgery*, 2023, Vol. 18, pp. 353–365. DOI: 10.1007/s11548-022-02738-5
15. Fang L., Jiang Y., Yan Y., Yue J., Deng Y. Hyperspectral image instance segmentation using spectral-spatial feature pyramid network, *IEEE Transactions on Geoscience and Remote Sensing*, 2023, Vol. 61, pp. 1–13. DOI: 10.1109/TGRS.2023.3240481
16. Xie X., Zhang W., Pan X., Xie L., Shao F., Zhao W., An J. CANet: Context aware network with dual-stream pyramid for medical image segmentation, *Biomedical Signal Processing and Control,* 2023, Vol. 81, P. 104437. DOI: 10.1016/j.bspc.2022.104437
17. Qin P., Chen J., Zeng J., Chai R., Wang L. Large-scale tissue histopathology image segmentation based on feature pyramid, *EURASIP Journal on Image and Video Processing,* 2018. DOI: 10.1186/s13640-018-0320-8
18. Jaeger S., Karargyris A., Candemir S., Folio L., Siegelman J., Callaghan F., Xue Z., Palaniappan K., Singh R. K., Antani S., Thoma G., Wang Y. X., Lu P. X., McDonald C. J. Automatic tuberculosis screening using chest radiographs, *IEEE Transactions on Medical Imaging,* 2014, Vol. 33, No. 2, pp. 233–245. DOI: 10.1109/TMI.2013.2284099
19. Candemir S., Jaeger S., Palaniappan K., Musco J. P., Singh R. K., Xue Z., Karargyris A., Antani S., Thoma G., McDonald C. J. Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration, *IEEE Transactions on Medical Imaging,* 2014, Vol. 33, No. 2, pp. 577–590. DOI: 10.1109/TMI.2013.2290491

Received 12.08.2024.
Accepted 15.10.2024.

DOI 10.15588/1607-3274-2024-4-11

УДК 004.8

АНСАМБЛЕВИЙ МЕТОД ЗАСНОВАНИЙ НА УСЕРЕДНЕННІ ФОРМ ОБ'ЄКТІВ, ВИКОРИСТОВУЮЧИ ПІРАМІДНИЙ МЕТОД

Конюхов В. Д. – аспірант, Інститут енергетичних машин і систем ім. А. М. Підгорного НАН України, Харків, Україна.

АНОТАЦІЯ

Актуальність. Сегментація зображень відіграє ключову роль в комп'ютерному зорі. На якість сегментації впливає багато факторів: шум, артефакти, складні форми об'єктів. Класичні методи не завжди можуть гарантувати гарний успіх, в залежності від якості зображення та наявного шуму, вони не завжди можуть досягти бажаного результату. Запропонований метод використовує ансамбль нейронних мереж, що дає змогу підвищити точність та стабільність сегментації.

Мета роботи – розробити новий метод комбінування передбачень ансамблю нейронних мереж, який зможе покращити точність сегментації за рахунок комбінування зображень різного розміру зображень.

Метод. Запропоновано метод який виконує усереднення форм об'єктів зображених на масках-передбаченнях. Для досягнення покращення якості сегментації використовується піраміда зображень, кожен рівень піраміди відповідає збільшеному розміру початкового зображення. Такий підхід дозволяє отримувати характеристики зображення на різних рівнях. Для тестового зображення отримується передбачення від кожної нейронної мережі в ансамблі, після чого для зображення будується піраміда. Всі рівні піраміди комбінуються в фінальне зображення за допомогою метода усереднення форм об'єктів. Всі отримані фінальні зображення для кожної нейронної мережі в кінці також комбінуються за допомогою метода усереднення форм об'єктів. Використання ансамблю нейронних мереж та пірамідного методу дають змогу зменшити вплив шумів та артефактів на результат сегментації.

Результати. Використання даного методу було порівняно зі звичайним використанням окремих нейронних мереж та ансамблевим методом усереднення. Отримані результати показують, що запропонований метод перевершує своїх конкурентів. Застосування запропонованого методу покращило точність та якість сегментації.

Висновки. Проведене дослідження підтвердило сенс використання ансамблю нейронних мереж та створення нового методу комбінування передбачень. Використання ансамблю нейронних мереж дає можливість компенсувати помилки та недоліки окремих нейронних мереж. Використання запропонованого методу може знизити вплив шумів та артефактів на сегментацію. Подальше вивчення та модифікація цього методу дадуть змогу покращити ще більше якість сегментації.

КЛЮЧОВІ СЛОВА: машинне навчання, розпізнавання образів, нейронна мережа, сегментація зображення, комп'ютерний зір.

ЛІТЕРАТУРА

1. Deep Learning for Computer Vision: A Brief Review / [A. Voulodimos, N. Doulamis, A. Doulamis, E. Protopapadakis] // Computational intelligence and neuroscience. – 2018. – P. 2018:7068349. DOI: 10.1155/2018/7068349
2. Ronneberger O. U-Net: Convolutional Networks for Biomedical Image Segmentation / O. Ronneberger, P. Fischer, T. Brox // ArXiv. – 2015. DOI: 10.1007/978-3-319-24574-4_28
3. Breiman L. Bagging predictors / L. Breiman // Machine Learning. – 1996. – Vol. 24, No. 2. – P. 123–140. DOI: 10.1007/BF00058655
4. Freund Y. A decision-theoretic generalization of on-line learning and an application to boosting / Y. Freund, R. E. Schapire // Journal of Computer and System Sciences. – 1997. – Vol. 55, No. 1. – P. 119–139. DOI: 10.1006/jcss.1997.1504
5. Wolpert, D. H. Stacked generalization / D. H. Wolpert // Neural Networks. – 1992. – Vol. 5, No. 2. – P. 241–259. DOI: 10.1016/S0893-6080(05)80023-1
6. Yang Y. A survey on ensemble learning under the era of deep learning / Y. Yang, H. Lv, N. Chen // Artificial Intelligence Review. – 2023. – Vol. 56. – P. 5545–5589. DOI: 10.1007/s10462-022-10283-5
7. Deep ensemble learning for accurate retinal vessel segmentation / [L. Du, H. Liu, L. Zhang et al.] // Computers in Biology and Medicine. – 2023. – Vol. 158. – P. 106829. DOI: 10.1016/j.compbiomed.2023.106829
8. Application of transfer learning and ensemble learning in image-level classification for breast histopathology / [Y. Zheng, C. Li, X. Zhou et al.] // Intelligent Medicine. – 2023. – Vol. 3, No. 2. – P. 115–128. DOI: 10.1016/j.imed.2022.05.004
9. Ensemble learning driven computer-aided diagnosis model for brain tumor classification on magnetic resonance imaging / [J. Vaiyapuri, S. Mahalingam, H. A. M. Ahmad et al.] // IEEE Access. – 2023. – Vol. 11. – P. 91398–91406. DOI: 10.1109/ACCESS.2023.3306961
10. Skin cancer detection using ensemble of machine learning and deep learning techniques / [J. V. Tembhurne, N. Hebbar, H. Y. Patil, et al.] // Multimedia Tools and Applications. – 2023. – Vol. 82. – P. 27501–27524. DOI: 10.1007/s11042-023-14697-3
11. Al-Zebari A. Ensemble convolutional neural networks and transformer-based segmentation methods for achieving accurate sclera segmentation in eye images / A. Al-Zebari // Signal, Image and Video Processing. – 2024. – Vol. 18. – P. 1879–1891. DOI: 10.1007/s11760-023-02891-7
12. Iqball T. Weighted ensemble model for image classification / T. Iqball, M. A. Wani // International Journal of Information Technology. – 2023. – Vol. 15. – P. 557–564. DOI: 10.1007/s41870-022-01149-8
13. Pyramid attention network for image restoration / [Y. Mei, Y. Fan, Y. Zhang et al.] // International Journal of Computer Vision. – 2023. – Vol. 131. – P. 3207–3225. DOI: 10.1007/s11263-023-01843-5
14. Multi-scale feature pyramid fusion network for medical image segmentation / [B. Zhang, Y. Wang, C. Ding et al.] // International Journal of Computer Assisted Radiology and Surgery. – 2023. – Vol. 18. – P. 353–365. DOI: 10.1007/s11548-022-02738-5
15. Hyperspectral image instance segmentation using spectral-spatial feature pyramid network / [L. Fang, Y. Jiang, Y. Yan et al.] // IEEE Transactions on Geoscience and Remote Sensing. – 2023. – Vol. 61. – P. 1–13. DOI: 10.1109/TGRS.2023.3240481
16. CANet: Context aware network with dual-stream pyramid for medical image segmentation / [X. Xie, W. Zhang, X. Pan et al.] // Biomedical Signal Processing and Control. – 2023. – Vol. 81. – P. 104437. DOI: 10.1016/j.bspc.2022.104437
17. Large-scale tissue histopathology image segmentation based on feature pyramid / [P. Qin, J. Chen, J. Zeng et al.] // EURASIP Journal on Image and Video Processing. – 2018. DOI: 10.1186/s13640-018-0320-8
18. Automatic tuberculosis screening using chest radiographs / [S. Jaeger, A. Karargyris, S. Candemir et al.] // IEEE Transactions on Medical Imaging. – 2014. – Vol. 33, No. 2. – P. 233–245. DOI: 10.1109/TMI.2013.2284099
19. Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration / [S. Candemir, S. Jaeger, K. Palaniappan et al.] // IEEE Transactions on Medical Imaging. – 2014. – Vol. 33, No. 2. – P. 577–590. DOI: 10.1109/TMI.2013.2290491

DOI 10.15588/1607-3274-2024-4-11