Research Article

Segmentation Technology of Nucleus Image Based on U-Net Network

Jie Fang ¹, QingBiao Zhou ¹, and Shuxia Wang ²

¹Zhejiang Industry Polytechnic College, Shaoxing 312000, Zhejiang, China

²School of Computing, Tonghua Normal University, Tonghua 134001, Jilin, China

Correspondence should be addressed to QingBiao Zhou; qingbiaozhou@163.com

Received 19 April 2021; Revised 28 May 2021; Accepted 2 June 2021; Published 11 June 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Jie Fang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To solve the problems of rough edge and poor segmentation accuracy of traditional neural networks in small nucleus image segmentation, a nucleus image segmentation technology based on U-Net network is proposed. First, the U-Net network is used to segment the nucleus image, which stitches the feature images in the channel dimension to achieve feature fusion, and the skip structure is used to combine the low- and high-level features. Then, the subregional average pooling is proposed to improve the global average pooling in the attention module, and an attention channel expansion module is designed to improve the accuracy of image segmentation. Finally, the improved attention module is integrated into the U-Net network to achieve accurate segmentation of the nuclear image. Based on the Python platform, the experimental results show that the proposed segmentation technology can achieve fast convergence, and the mean intersection over union (MIoU) is 85.02%, which is better than other comparison technologies and has a good application prospect.

1. Introduction

With the development of medicine, more and more medical imaging images need to be processed, and image processing technology has become more and more important [1]. Traditional medical imaging image processing and analysis only rely on the doctor's experience, which not only wastes manpower but also affects the accuracy rate because the doctor's experience and physical condition affect judgment result. Therefore, the breakthrough of automated medical image processing technology has a very critical role in improving the efficiency of medical diagnosis [2]. Medical image segmentation is an important task. Moreover, many other related tasks in medical image processing require image segmentation in advance. Medical image segmentation generally refers to the extraction of certain target regions in the entire image in some way, such as cell nuclei, an organ, or tissue [3, 4]. The results of medical image segmentation usually have no intersection in each area. Moreover, each segmented area has a certain similarity in its interior [5]. Compared with image segmentation of natural

scenes with clear outlines, medical images have great particularities. The complexity of the medical image itself causes the separation between its components to be blurred, and the boundaries between the components are not clear enough [6].

The method of medical image segmentation has developed from the initial manual segmentation to semi-automatic segmentation and then to the most recent fully automatic segmentation [7, 8]. With the gradual in-depth study of medical image segmentation by domestic and foreign researchers, many research results based on traditional algorithms have been obtained, which can be divided into three categories [9]. One is an algorithm that uses the discontinuity of boundary information to perform segmentation, such as surface fitting, parallel differential operator, deformation model, and so on [10]. The second is an algorithm for segmentation using the similarity of different regions of the image, such as region growing algorithm, threshold method, classification and clustering algorithm, and statistical-related algorithm [11]. The third is an algorithm that combines the discontinuity of boundary

information and the similarity of different regions of the image [12]. Reference [13] proposed an immune system programming (ISP) image segmentation algorithm based on a new evolutionary algorithm combined with region growth technology. The ISP algorithm with a tree data structure can segment medical images better. However, the actual area growth technology does not consider the complexity of the image boundary, and the accuracy rate needs to be optimized. Reference [14] proposed an improved multi-level threshold image segmentation method based on differential evolution. The efficiency of different parts of the allocation differential evolution algorithm is evaluated by measuring the quality of the candidate solutions, so as to generate the optimal solution of the allocation population to improve the efficiency of the algorithm when the number of thresholds increases. Reference [15] proposed a non-revisited quantum behavior particle swarm algorithm. Among them, the use of a refined search method overcomes the shortcomings of the original search method, reduces the computational cost, and has better effectiveness and robustness. Reference [16] proposed an active contour segmentation method for morphological medical images with an automatic initialization function. This method has a low computational cost, good robustness, and a high degree of automation. However, the settings before initialization are more complicated, and the adaptability to different application scenarios is low.

Traditional algorithms have obvious limitations. In recent years, deep learning methods have developed rapidly and are widely used, such as recurrent neural networks, restricted Boltzmann machines, and convolutional neural networks [17]. Reference [18] proposed a deep belief network brain tumor image segmentation method based on harmonious cuckoo search. By integrating Bayesian fuzzy clustering and the active contour model, better accuracy is obtained, but the computational efficiency is not high. Reference [19] achieved high-precision tumor segmentation in the segmentation process of the fuzzy mean clustering algorithm by extracting the features of the gray co-occurrence matrix and the grayscale run-length matrix. However, the steps in the feature extraction stage are more complicated, and the segmentation efficiency needs to be improved. Reference [20] developed a multi-graph-based tag fusion high-order feature learning framework. Fusion of the mean-covariance limited Boltzmann machine and high-level image features to segment structural brain images. Reference [21] used the evolutionary computing power potential of dense blocks and residual blocks to propose an automatic evolution model for medical image segmentation. Good results have been achieved, but there remain some difficulties for image segmentation of complex nuclei.

Aiming at the fact that the existing segmentation methods are difficult to apply to the image segmentation of the nucleus in the medical field, a U-Net-based cell nucleus image segmentation technology is proposed. Compared with traditional medical image segmentation methods, its innovations are

- (1) To solve the problems of poor segmentation of small nucleus, rough edges, and under- and over-segmentation, the U-Net network is used for image segmentation. It stitches feature maps in the channel dimension to achieve feature fusion and uses a skip structure to combine low- and high-level features to ensure the segmentation effect of the nucleus.
- (2) Due to the commonly used global mean pooling method, the extracted channel attention information is weak in interpretability, and the information is rough. Therefore, the subregional average pooling method is used in the attention module instead of the global average pooling method.

The structure of this paper is as follows: Section 1 introduces the significance and research status of medical image segmentation and summarizes the innovation points of the proposed segmentation network. Section 2 introduces the U-Net network in detail, as well as the attention mechanism and its improvement methods, thereby designing a complete nucleus segmentation network structure. Section 3 conducts experiments and evaluates the results to demonstrate that the proposed segmentation network has good feasibility and effectiveness. Finally, the full text is summarized and prospected.

2. Theory and Method

2.1. U-Net. Because the semantic segmentation effect of fully convolutional networks (FCN) is relatively rough, and the U-Net network, as a further extension of the FCN, has become the cornerstone of medical image segmentation [22, 23]. U-Net is a semantic segmentation network proposed by Olaf Ronneberger in 2015. In the upsampling process, the form of downsampling is matched to keep it consistent [24]. On this basis, a large number of feature maps in the downsampling stage are added to the upsampling to fill in the information lost in the calculation process. Its structure is shown in Figure 1.

U-Net includes a contraction path and an expansion path. The left side of Figure 1 is a contraction path, that is, downsampling, including two 3×3 convolutional layers and a 2×2 maximum pooling layer with a stride of 2. The activation function is rectified linear unit (ReLU) [25]. The classic image classification network with the fully connected layer removed is usually used. It performs a convolution kernel pooling operation on the original input picture, which can obtain contextual semantic information to solve the classification problem in image segmentation. On the right is an expansion path, that is, upsampling, which can locate segmentation tasks [26]. Up- and downsampling are symmetrical. First, a 2×2 convolutional layer is connected to reduce the number of feature channels, and then two 3×3 convolutions are used. Finally, there is a convolutional layer of 1×1 to map the required number of feature vectors, and the convolution is an unpadding structure.

FIGURE 1: U-Net network structure.

The U-Net network has a major change in structure, and it builds more characteristic channels through skip connections during upsampling. Moreover, U-Net and FCN use different methods for feature fusion. FCN adds the feature maps point by point, while U-Net splices the feature maps to make them have more channels. The model training requires fewer data sets, can converge on a small amount of data, and can quickly obtain results when performing image segmentation.

The energy function of network training uses weighted cross-entropy, which is calculated as follows:

$$E = \sum_{x \in \phi} w(x) \log(p_k(x)),$$

$$p_k(x) = \frac{\exp(\sigma_k(x))}{\sum_{k'}^K \exp(\sigma_{k'}(x))}, \quad (1)$$

where $\sigma_k(x)$ represents the activation function of the k -th feature channel at the x -th pixel. k is the number of categories. $p_k(x)$ is the approximate maximum function. ϕ is the image set. $w(x)$ is the importance of pixels in the training composition; the more the importance, the greater the weight.

U-Net has special advantages in processing medical images. To solve the problem of the lack of sample images in medical images, elastic deformation is used to complete data enhancement. Elastic deformation is a relatively common type of deformation in actual cells, so it is very suitable for medical image processing [27, 28]. The algorithm of data enhancement is adopted to make the neural network model learn the invariance of elastic deformation so that the network can have good elastic deformation adaptability when the data set is small. It can correctly complete the

segmentation when encountering the elastically deformed medical image.

2.2. Attention Module. To improve the segmentation speed of the network, a lightweight network is used as a feature extraction network, but there is a certain loss in the accuracy. In order to improve the accuracy, adding an attention module to the U-Net helps enhance the feature expression of the model [29, 30]. This module integrates different information and improves the understanding of the model, which is similar to the attention mechanism of human vision. There are two types of human visual attention mechanisms: Bottom-up data-driven attention mechanism and top-down task-driven target attention mechanism [31]. Both mechanisms can learn the parts required by the task from a large amount of data. The proposed network uses a bottom-up data-driven attention mechanism [32]. The attention module starts from the relationship between feature channels and considers the interdependence factors between feature channels. Through the self-learning of the network, the features that have little effect on the current segmentation are effectively suppressed, and the weight of beneficial features is enhanced. Its module structure is shown in Figure 2.

The attention module first performs global average pooling (GAP) on the feature map of each channel to obtain the vector of $1 \times 1 \times K$, and then performs two fully connected (FC) layer conversions. To suppress the complexity of the model, dimensionality reduction and increase are performed between the two FC layer conversions, which is similar to a “bottleneck”, and Sigmoid and ReLU activation functions are used.

FIGURE 2: The structure of attention module.

Image semantic segmentation tasks are often aimed at the processing of complex scenes. There are often multiple objects in the image; the object types and sizes are different; and the spatial scene distribution is complex. Using GAP directly means simply treating each channel as a single-category problem, and GAP has no parameters to learn from [33]. The channel information obtained by this method may be too rough to explain the meaning of the channel better [34]. To solve the problem and learn the spatial distribution of the image better, a GAP improvement method is proposed, namely, subregional average pooling (SAP). GAP directly transforms a channel feature with a dimension of $H \times W \times 1$ into a $1 \times 1 \times 1$ size feature, and the process of SAP processing is shown in Figure 3.

SAP first transforms a channel feature with a shape of $H \times W \times 1$ into a new feature map with a dimensional shape of $m \times n \times 1$ after an adaptive mean pooling operation so that the obtained channel feature retains some spatial information. Then, a convolution operation with a kernel size of $m \times n$ is used to convert the obtained new channel feature map into a channel feature with a dimension of $1 \times 1 \times 1$. The convolution operation can learn the characteristics of the image after pooling and get its spatial characteristics. It should be noted that in the experiment, in order to facilitate the calculation and not increase the amount of calculation too much, the parameter is set to $m = n = 4$, and it is better not to be too large for m and n , generally less than or equal to 7.

In order to allow low-level channel information to be transmitted to higher-level channels, low- and high-level channel information are aligned at the same time. An attention channel expansion module CE is designed in the proposed network, as shown in Figure 4.

Among them, a 1×1 convolution operation is used to increase the number of channels, and then the Sigmoid function is used [35]. The high-level channel attention weight and the low-level expanded channel attention weight are added to obtain the updated channel attention weight. Finally, the recalibrated channel weight is multiplied by the high-level features. It should be noted that the range of the updated channel weight is $(0, 2)$ instead of the usual $(0, 1)$. This will allow the weight to not only reduce the original feature value but also expand the feature value [36].

In order to facilitate the description, the input feature map is denoted as $U = [u_1, u_2, \dots, u_C]$, $u_i \in \mathbb{R}^{H \times W}$ represents the i -th channel map in U . $f \in \mathbb{R}^{m \times n \times C}$ represents the feature map of U after the pooling operation, and its dimension is $m \times n \times C$. $u_c^g \in \mathbb{R}^{h \times w}$ indicates that the pooling operation corresponds to the characteristic subregion g of

FIGURE 3: SAP process.

FIGURE 4: Improved channel attention model.

the c -th channel graph in U . The size of the subregion is $h \times w$. Then the expression of the pooling operation can be obtained as follows:

$$f_c^{(m,n)} = \frac{1}{h \times w} \sum_{i=1}^h \sum_{j=1}^w u_c^g(i, j). \quad (2)$$

Next, it is a convolution calculation, defining symbol $*$ represents the convolution, $d_c \in \mathbb{R}^{m \times n}$ represents the conjunction core corresponding to the f_c , and the result F_c is represented as follows:

$$F_c = d_c * f_c. \quad (3)$$

The activation function ReLU is used after batch normalization (BN) for $F = [F_1, F_2, \dots, F_C]$. Here, φ represents the BN operation, and σ represents the ReLU function.

$$\tilde{F} = \sigma(\varphi(F)). \quad (4)$$

Then, the obtained channel feature $\tilde{F}$ is then connected to the convolutional calculation s_r of 1×1 , and the function is activated using the Sigmoid, where $r \in [1, C]$. Remember the Sigmoid function is δ , the output δ_{SAP} of the SAP module is as follows:

$$\delta_{\text{sapm}} = [\delta_1, \delta_2, \dots, \delta_r, \dots, \delta_C],$$

$$\delta_r = \delta \left(\sum_{i=1}^C s_r * \tilde{F}_i \right). \quad (5)$$

Finally, the high-order channel attention distribution obtained by defining SAP module is $\delta'(\cdot)$. The low-order channel attention distribution of the CE module is $\delta''(\cdot)$. Then we can get the channel attention distribution δ_F for updating higher-order features.

$$\delta_F = \delta''(\cdot) + \delta'(\cdot). \quad (6)$$

After the above steps, a recalibrated feature $\hat{U}$ is $\hat{U} = U\delta_F$.

2.3. Network Structure Design. In the U-NET network, the improved attention module is integrated for nucleus image segmentation. Its network structure is shown in Figure 5.

Three layers of 3×3 convolution are used in the network; the last classification layer is removed; and the step size of the maximum pooling layer is changed from 2 to 1. At the same time, the ordinary 3×3 convolution is changed to the expansion convolution with the expansion rate of 2 so that the resolution of the output feature map is equal to 1/16 of the size of the input image. An improved atrous spatial pyramid pooling (ASPP) module is added at the top of the network, and its output channel number is 512. After 4 times of upsampling, the output of the features by ASPP is added to a low-level feature that has undergone a 3×3 convolution and has the same dimension. Then connect a 3×3 convolution operation for feature fusion. Finally, it is upsampled to restore the original image size. Among them, the SAP module is used to obtain the initial channel weight information, and the CE module is used to expand the low-level channel attention.

Finally, each layer of the network is connected by skip connections, where connections exist in all layers. And in the pooling method, ASPP is used. ASPP provides a multi-scale information model, which adds a dilated convolution with different expansion rates on the basis of spatial pyramid pooling to capture a wide range of contexts. SAP is used to combine image features to increase the global context.

3. Experimental Results and Analysis

The network is built through the TensorFlow deep learning framework released by Google. The GPU model used is RTX 2080Ti, and the card memory size is 11 GB. The main information of the experiment is shown in Table 1.

3.1. Network Parameter Setting. In the experiment, ASPP was changed to 4 parallel 3×3 expansion convolution operations, rates=(1, 6, 12, 18). At the same time, during network training, random gradient descent (SGD) is used for parameter optimization. Its momentum parameter is set to 0.98, and the weight decay rate is 1×10^{-4} . The learning rate is a reduction strategy of the initial learning rate $\gamma_0 =$

FIGURE 5: The proposed network structure.

TABLE 1: Experimental environment parameters.

Parameter	Configuration
Operating system	Ubuntu 16.04
CPU	i7-9700Ti
GPU	RTX 2080Ti
Internal storage	12G
CUDA	10.1
Python	3.6
TensorFlow	1.3.0

5×10^{-3} multiplied by $\sqrt{a^2 + b^2}$. Among them, $\text{power} = 0.99$ and $\text{iter}_{\text{current}}$ are the current number of iterations, and iter_{max} is the maximum number of iterations in the training process. The number of experimental training is 10,000, the image pixel size in the experiment is set to 512×512 , and the training batch size is $2 \times 8 = 16$. At the same time, during training, data enhancement measures such as random cropping, horizontal flipping, vertical flipping, and random sample scrambling are adopted. In addition, during the evaluation, the image will be scaled at multiple scales, with a zoom ratio of 0.55–1.55.

3.2. Experimental Data. The data used in the experiment come from the 2018 Data Science Bowl, which is manually labeled by professional doctors, and contain 670 pairs of original images with 9 resolutions and annotated segmented images of each nucleus, as shown in Figure 6.

The original image in the data set is shown in Figure 6(a). Each original image corresponds to multiple segmented images of a marked nucleus, that is, an original picture usually contains multiple nuclei, and the annotated images of multiple nuclei are merged as shown in Figure 6(b). When collecting raw data, different acquisition methods, different magnification magnifications, and different cell presentation methods are used. Moreover, the cell types collected are inconsistent. This results in the cell images in the data set with different morphology and distinct light and dark. The model needs to have a strong generalization ability and be able to adapt to a variety of different situations.

FIGURE 6: Nucleus image segmentation data: (a) original image and (b) annotated segmentation image.

3.2.1. Image Preprocessing. Due to the influence of various factors in the original data collection process, there is a large imaging difference in the cell images in the data set, which will affect the image segmentation. Therefore, it is necessary to preprocess it before segmentation. First of all, most of the pictures in the data set have a resolution of 512×512 . Therefore, the picture resolution is unified to 512×512 . Then, most of the data sets are grayscale images, and a few are color images. To improve the network processing speed, it is necessary to change the color images to gray scale.

At the same time, the contrast between the image of some nuclei in the data set and the background is small, which may make it difficult for the segmentation method to distinguish the nucleus from the background. Therefore, the data set needs to be preprocessed for histogram equalization. After the histogram equalization process, the gray value of cells and the background has a significant difference in the image, which helps the network extract more features.

In addition, in the process of collecting images, various noises are often interfered with and contaminated, resulting in a decrease in the signal-to-noise ratio of the image, and the edges of cells and the background become blurred. To improve the signal-to-noise ratio of the image, it is usually necessary to preprocess the image by filtering. In the experiment, a Gaussian smoothing filter is used to preprocess the image, and the filter $G(x, y)$ is calculated as follows:

$$G(x, y) = \frac{1}{2\pi\mu^2} e^{-\frac{(x-l-1)^2 + (y-l-1)^2}{2\mu^2}}, \quad (7)$$

where μ is the standard deviation of gray value, and l is the dimension of Gaussian convolution kernel.

3.2.2. Image Enhancement. To overcome the overfitting phenomenon of CNN, random shearing, flipping, gray perturbation, and shape perturbation are used in the experiment. Gray perturbation can transform each pixel in a small range. The gray value of the CT image is multiplied by a random number $[0.80 \sim 1.20]$, and a random number $[-0.20 \sim 0.20]$ is added. The grayscale perturbation in the training set can improve the stability of the network, thereby improving the performance of the prediction set network.

The CT and the contour images are deformed by affine transformation to form shape disturbance. The deformation

method is to obtain first the coordinates of the 3 vertices (upper left, upper right, and lower left). Then each point moves randomly, and the range of random movement is the image length. Finally, an affine transformation is performed on the entire image.

3.3. Evaluation Index. In the experiment, pixel accuracy (PA), mean pixel accuracy (MPA), and MIoU are used as indicators to evaluate the performance of the proposed network. Suppose N_{pq} represents the correct number of divisions; N_{pq} indicates that the original pixel belongs to the p category but is divided into the number of q categories; N_{qp} represents the number of pixels that originally belonged to the q category but were divided into p categories. There are $\tau + 1$ categories (including τ categories and an empty category or background category).

PA is the simplest accuracy measure for semantic segmentation, which represents the proportion of correctly marked pixels to the total pixels. The calculation is as follows:

$$PA = \frac{\sum_{p=0}^k N_{pp}}{\sum_{p=0}^k \sum_{q=0}^k N_{pq}}. \quad (8)$$

MPA calculates the proportion of pixels that are correctly segmented in each class and then finds the average of all classes. The calculation is as follows:

$$MPA = \frac{1}{\tau + 1} \sum_{p=0}^{\tau} \frac{N_{pp}}{\sum_{q=0}^{\tau} N_{pq}}. \quad (9)$$

MIoU calculates the ratio of intersection and union of two sets. The pixel intersection ratio is calculated within each pixel category, and then the average is calculated as follows:

$$MIoU = \frac{1}{\tau + 1} \sum_{p=0}^{\tau} \frac{N_{pp}}{\sum_{q=0}^{\tau} N_{pq} + \sum_{q=0}^{\tau} N_{qp} - N_{pp}}. \quad (10)$$

MIoU is highly representative, efficient, and concise and has become the current general image segmentation evaluation index. Therefore, MIoU is used as the main evaluation index of the experiment.

3.4. Training Process. When training the network, the input image undergoes local response normalization before the

FIGURE 7: Curves of MIoU, MPA, and the loss.

first layer of convolution. The objective loss function is optimized by using the Adam algorithm with an initial learning rate of 0.005 and iterated until the loss function converges. The weight decay is 0.0001, and the number of iterations is set to 10,000. In the training process, the training data set is randomly shuffled, and then the batch size is set to 20. Due to the large deviation of the number of pixels in each category in the data set, the median frequency equalization method is used to balance between classes.

In the experiment, the proposed network is iteratively trained for 100 epochs on the data set. The changes in MIoU, MPA, and the loss of the validation set during the training process are shown in Figure 7.

MIoU reached 83% when the network was trained to the 50th epoch. In the subsequent training, MIoU stabilized, reaching 90%. MPA reached 94% when the network iterated to 60 epochs. Loss is reduced to about 7% when the network is iterated to 30 epochs. After 100-epoch iterations, the model basically converged, and the loss was reduced to below 4%.

3.5. Comparison of Technical Effects. After the nucleus image segmentation network converges, it is used to segment the images on the test set, and its segmentation effect is evaluated. In order to demonstrate the segmentation performance of the proposed technology, it is compared with references [14, 21]. The segmentation result of the nucleus image is shown in Figure 8.

It can be seen from Figure 8 that reference [14] uses differential evolution improved multi-level threshold to achieve image segmentation. The segmentation of the central part of the nucleus image is very good, but the segmentation of the edge details and smaller nuclei is not very good.

Moreover, there are certain over- and undersegmentation. Reference [21] proposed an automatic evolution model to achieve image segmentation. Similarly, the central part of the nucleus image is segmented well, and the phenomenon of over- and undersegmentation is relatively reduced. However, the result of manual annotation is relatively rough, and the ability to segment a smaller nucleus is poor. The proposed technology integrates the improved attention module into the U-Net network and has a better effect than the original U-Net for edge detail and smaller nucleus segmentation. The phenomenon of over- and undersegmentation is also relatively reduced, which is closer to the result of manual labeling. This proves that the proposed technology has ideal segmentation capabilities.

To quantitatively analyze the performance of the proposed technique, experiments are carried out on the data of nucleus image segmentation. Three indexes of PA, MPA, and MIoU are used to evaluate its segmentation performance with references [14, 21]. All test set data are used in the experiment to calculate the difference between the results of each technology segmentation and the manual segmentation standard, so as to obtain the evaluation results, as shown in Table 2.

The segmentation performance of the proposed technology is the best, and its MIoU reaches 85.02%. Because the proposed technology uses the most widely used U-Net network in the medical field and, at the same time, uses an improved attention module, the segmentation accuracy is further improved. Reference [14] realizes image segmentation based on a multi-level threshold improved by differential evolution. The optimal solution of the allocation population is generated by measuring the quality evaluation of the candidate solutions. The segmentation effect is affected by the choice of the larger optimal solution. Therefore,

FIGURE 8: Comparison of nucleus image segmentation results: (a) original image, (b) image from reference [14], (c) image from reference [21], (d) proposed technology, and (e) manual marking.

TABLE 2: Performance comparison of three image segmentation networks.

%	Ref. [14]	Ref. [21]	Proposed technology
PA	82.54	85.68	89.97
MPA	84.17	87.39	91.35
MIoU	75.69	80.75	85.02

the overall performance is poor; MPA is only 84.17%. Reference [21] uses the evolutionary computing power potential of dense and residual blocks to propose an automatic evolution model, which has better image segmentation results. However, there are still some shortcomings in the image segmentation of a complex nucleus. Compared with the proposed technology, its MIoU is reduced by 4.27%.

Briefly, by comparing the experimental results, we can see that from the traditional segmentation technology in reference [14] to the deep learning algorithm in reference [21] and then to the improved U-Net network of the proposed technology, the segmentation effect and robustness of the nucleus are getting better and better. The segmentation effect of edge details and the smaller nucleus is also getting better and better. The experimental results show that the feature fusion method of skip connection and feature splicing in the U-Net network significantly improves the effect of U-Net image segmentation. Moreover, the

integration of the improved attention module has also significantly improved the accuracy of its image segmentation.

4. Conclusion

Traditional image segmentation algorithms generally need to extract some features manually in advance, such as the edges, corners, textures, and lines of the image. They have poor robustness and are easily affected by the environment. At the same time, the edges of the nucleus are more complicated, and the target is smaller. For this reason, a nucleus image segmentation technology based on the U-Net network is proposed. SAP and CE modules are used to improve the attention module, and the improved attention module is integrated into the U-Net network to segment the nucleus image, which further ensures the accuracy of segmentation. The 2018 Data Science Bowl data set is used on

the Python platform to demonstrate the proposed segmentation technology experimentally. The results show that it has better edge detail parts and smaller cell nuclei segmentation capabilities. PA, MPA, and MIoU are 89.97%, 91.35%, and 85.02%, respectively, which are better than other comparison techniques. It provides certain theoretical support for the high accuracy segmentation of the nucleus.

Because the offline method is used for image augmentation in the experiment, a larger storage space is required. In the following research, image augmentation can be added to the deep learning network to reduce the demand for storage space.

Data Availability

The data included in this paper are available without any restriction.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

The authors wish to express their appreciation to the reviewers for their helpful suggestions that greatly improved the presentation of this paper. This work was supported by the Natural Science Foundation of Zhejiang Province (no. LY18F020002).

References

- [1] M. Tofighi, T. Guo, J. K. P. Vanamala, and V. Monga, "Prior information guided regularized deep learning for cell nucleus detection," *IEEE Transactions on Medical Imaging*, vol. 38, no. 9, pp. 2047–2058, 2019.
- [2] A. Veeramuthu, S. Meenakshi, and K. A. Kumar, "A neural network based deep learning approach for efficient segmentation of brain tumor medical image data," *Journal of Intelligent and Fuzzy Systems*, vol. 36, no. 1, pp. 1–8, 2019.
- [3] G. Wang, W. Li, M. Aertsen, J. Deprest, S. Ourselin, and T. Vercauteren, "Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks," *Neurocomputing*, vol. 338, no. 4, pp. 34–45, 2019.
- [4] M. Taheri, M. Rastgarpour, and A. Koochari, "A novel method for medical image segmentation based on convolutional neural networks with SGD optimization," *Iranian Journal of Electrical and Electronic Engineering*, vol. 9, no. 7, pp. 37–46, 2021.
- [5] U. S. L. U. Fatmatülzehrä, C. Bass, and A. A. Bharath, "PERI-Net: a parameter efficient residual inception network for medical image segmentation," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 28, no. 4, pp. 2261–2277, 2020.
- [6] H. Zheng, Y. Zhang, L. Yang et al., "An annotation sparsification strategy for 3D medical image segmentation via representative selection and self-training," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 4, pp. 6925–6932, 2020.
- [7] D. A. Hasan and A. M. Abdulazeez, "A modified convolutional neural networks model for medical image segmentation," *Test Engineering and Management*, vol. 83, pp. 16798–16808, 2020.
- [8] H. Shang, S. Zhao, H. Du, J. Zhang, W. Xing, and H. Shen, "A new solution model for cardiac medical image segmentation," *Journal of Thoracic Disease*, vol. 12, no. 12, pp. 7298–7312, 2020.
- [9] T. Khin, K. Srujan Raju, G. R. Sinha, K. K. Khaing, and T. M. Kyi, "Review of optimization methods of medical image segmentation," *Proceedings of the Third International Conference on Computational Intelligence and Informatics*, vol. 1090, no. 1, pp. 213–218, 2020.
- [10] L. Fang, X. Wang, and L. Wang, "Multi-modal medical image segmentation based on vector-valued active contour models," *Information Sciences*, vol. 513, no. 4, pp. 504–518, 2020.
- [11] K. Li, L. Yu, S. Wang et al., "Towards cross-modality medical image segmentation with online mutual knowledge distillation," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 1, pp. 775–783, 2020.
- [12] J. Junlong Cheng, S. Tian, L. Yu, and H. You, "Multi-attention mechanism medical image segmentation combined with word embedding technology," *Automatic Control and Computer Sciences*, vol. 54, no. 6, pp. 560–571, 2020.
- [13] E. Mabrouk, A. Ayman, Y. Raslan, and A.-R. Hedar, "Immune system programming for medical image segmentation," *Journal of Computational Science*, vol. 31, no. FEB, pp. 111–125, 2019.
- [14] H. Tarkhaneh, "An adaptive differential evolution algorithm to optimal multi-level thresholding for MRI brain image segmentation," *Expert Systems with Applications*, vol. 138, no. 6, pp. 112820–112829, 2019.
- [15] Z. Yang and A. Wu, "A non-revisiting quantum-behaved particle swarm optimization based multilevel thresholding for image segmentation," *Neural Computing and Applications*, vol. 32, no. 16, pp. 12011–12031, 2020.
- [16] F. F. Ximenes Vasconcelos, A. G. Medeiros, S. A. Peixoto, and P. P. Rebouças Filho, "Automatic skin lesions segmentation based on a new morphological approach via geodesic active contour," *Cognitive Systems Research*, vol. 55, no. 6, pp. 44–59, 2019.
- [17] D. Karimi and S. E. Salcudean, "Reducing the Hausdorff distance in medical image segmentation with convolutional neural networks," *IEEE Transactions on Medical Imaging*, vol. 39, no. 2, pp. 499–513, 2020.
- [18] A. Ratna Raju, S. Pabboju, and R. Rajeswara Rao, "Hybrid active contour model and deep belief network based approach for brain tumor segmentation and classification," *Sensor Review*, vol. 39, no. 4, pp. 473–487, 2019.
- [19] A. Geetha and N. Gomathi, "A robust grey wolf-based deep learning for brain tumour detection in MR images," *Bio-medical Engineering/Biomedizinische Technik*, vol. 65, no. 2, pp. 191–207, 2020.
- [20] L. Sun, W. Shao, M. Wang, D. Zhang, M. Liu et al., "High-order feature learning for multi-atlas based label fusion: application to brain segmentation with MRI," *IEEE Transactions on Image Processing*, vol. 29, no. 3, pp. 2702–2713, 2019.
- [21] T. Hassanzadeh, D. Essam, and R. Sarker, "An evolutionary DenseRes deep convolutional neural network for medical image segmentation," *IEEE Access*, vol. 8, pp. 212298–212314, 2020.
- [22] N. Kadoya, "Deformable image registration and auto-segmentation for various medical imaging types," *Igaku Butsuri*:

- Nihon Igaku Butsuri Gakkai Kikanshi = Japanese Journal of Medical Physics: An Official Journal of Japan Society of Medical Physics*, vol. 39, no. 1, pp. 12–19, 2019.
- [23] K. K. Jha and H. S. Dutta, “Nucleus and cytoplasm-based segmentation and actor-critic neural network for acute lymphocytic leukaemia detection in single cell blood smear images,” *Medical & Biological Engineering & Computing*, vol. 58, no. 8, pp. 1–16, 2019.
 - [24] T. Wang, J. Huang, D. Zheng, and Y. He, “Nucleus segmentation of cervical cytology images based on depth information,” *IEEE Access*, vol. 8, no. 4, pp. 75846–75859, 2020.
 - [25] J. C. Caicedo, A. Goodman, K. W. Karhohs et al., “Nucleus segmentation across imaging experiments: the 2018 data science bowl,” *Nature Methods*, vol. 16, no. 12, pp. 1247–1253, 2019.
 - [26] T. Wan, S. Xu, C. Sang, Y. Jin, and Z. Qin, “Accurate segmentation of overlapping cells in cervical cytology with deep convolutional neural networks,” *Neurocomputing*, vol. 365, no. 6, pp. 157–170, 2019.
 - [27] H. Zhang, H. Zhu, and X. Ling, “Polar coordinate sampling-based segmentation of overlapping cervical cells using attention U-Net and random walk,” *Neurocomputing*, vol. 383, no. 3, pp. 212–223, 2020.
 - [28] J. Song, L. Xiao, M. Molaei, and Z. Lian, “Multi-layer boosting sparse convolutional model for generalized nuclear segmentation from histopathology images,” *Knowledge-Based Systems*, vol. 176, no. 15, pp. 40–53, 2019.
 - [29] N. Ramesh and T. Tasdizen, “Cell segmentation using a similarity interface with a multi-task convolutional neural network,” *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 4, pp. 1457–1468, 2019.
 - [30] A. A. Mahmoud, W. El-Shafai, T. E. Taha et al., “An efficient segmentation technique for different medical image modalities,” *Menoufia Journal of Electronic Engineering Research*, vol. 30, no. 1, pp. 22–28, 2021.
 - [31] M. Moirangthem and T. R. Singh, “Brain tumor detection through content-based medical image retrieval using ROI segmentation with Harmony search optimization,” *Journal of Green Engineering*, vol. 10, no. 10, pp. 8939–8969, 2020.
 - [32] A. F. A. Fadzil, N. E. A. Khalid, and S. Ibrahim, “Amplification of pixels in medical image data for segmentation via deep learning object-oriented approach,” *International Journal of Advanced Technology and Engineering Exploration*, vol. 8, no. 74, pp. 82–90, 2021.
 - [33] U. Reddy, P. Dhanalakshmi, and P. Reddy, “Image segmentation technique using SVM classifier for detection of medical disorders,” *Ingénierie des systèmes d’information*, vol. 24, no. 2, pp. 173–176, 2019.
 - [34] E. K. Wang, C.-M. Chen, M. M. Hassan, and A. Almogren, “A deep learning based medical image segmentation technique in Internet-of-Medical-Things domain,” *Future Generation Computer Systems*, vol. 108, no. 4, pp. 135–144, 2020.
 - [35] D. Li, N. Deng, and X. Chen, “Level set medical image segmentation aided by cooperative quantum particle optimization with Lévy flights,” *Vibroengineering PROCEDIA*, vol. 28, no. 10, pp. 93–98, 2019.
 - [36] Y. Niu and J. Cao, “Local difference-based active contour model for medical image segmentation and bias correction,” *IET Image Processing*, vol. 13, no. 10, pp. 1755–1762, 2019.