ISSN 2522-3488 Vol. 7, Issue 1: 53-59, DOI: 10.25236/IJNDES.2023.070109

Image segmentation based on improved UNET++

Yiming Zhang

Artificial Intelligence School, Jianghan University, Wuhan, Hubei, 430056, China

Abstract: *In order to help researchers to perform cell segmentation, computer vision techniques are used to perform cell segmentation on neural cells. The addition of the Attention mechanism based on U-net++ suppresses the activity of irrelevant regions and improves the segmentation efficiency and segmentation accuracy. According to the experimental data, the segmentation accuracy of the method reaches 0.78 and the loss rate is 0.3, which can achieve a better segmentation effect compared with the traditional algorithm.*

Keywords: *Unet; Unet++; image segmentation*

1. Introduction

In the medical field, the analysis of medical images plays a vital role in the analysis of diseases and scientific research. In image analysis of cells, it is generally necessary to segment and count the morphology of the cells, which helps in quantitative and qualitative analysis. Segmenting and labelling cells is a very difficult task, especially for nerve cells with complex shapes, which takes a lot of time if they are manually labelled through traditional manual work, so computer vision is becoming increasingly popular in the field of medical research.

Traditional cell segmentation algorithms are based on watershed [1] segmentation and contour [2] segmentation, which are relatively inefficient. With the development of deep neural networks, it has been found that using machine learning for cell segmentation can substantially improve the accuracy and efficiency, and the convolutional neural network [3] is a very efficient neural network that can be used without complex pre-processing of the input image. Among convolutional neural networks, U-net [4] is one of the models that are widely used in the field of medical cell segmentation, but it is easy to misidentify cell edges that are close together, imaging blurred locations, and has a high value of semantic feature loss in image processing.

To address the above problems, we used the U-net++ [5] model, which adds more connected paths to the network, connecting convolutional blocks of different depths together. We added the Attention [6] mechanism to the U-net++ network, paying more attention to cellular regions rather than irrelevant regions when training the model, reducing the weight of background regions and improving the model training efficiency.

2. Model section

The model used in this paper is shown above and is an improvement on the U-net++ model, which adds Deep Supervision to the model, adding network branches for auxiliary classification in the middle hidden layer, which supervise the backbone network, thus speeding up convergence during neural network training and avoiding the problem of gradient disappearance. In the classical U-network model, the corresponding image layers are superimposed during upsampling and downsampling in order to avoid losing the precise details of the samples, but this operation will also superimpose a lot of redundant information, thus affecting the feature values and reducing the efficiency. This allows the training to automatically focus on the regions with higher weights and reduce the computation of irrelevant regions, thus reducing the computational effort and avoiding the introduction of a large amount of irrelevant parameter information, as shown in Figure 1.

ISSN 2522-3488 Vol. 7, Issue 1: 53-59, DOI: 10.25236/IJNDES.2023.070109

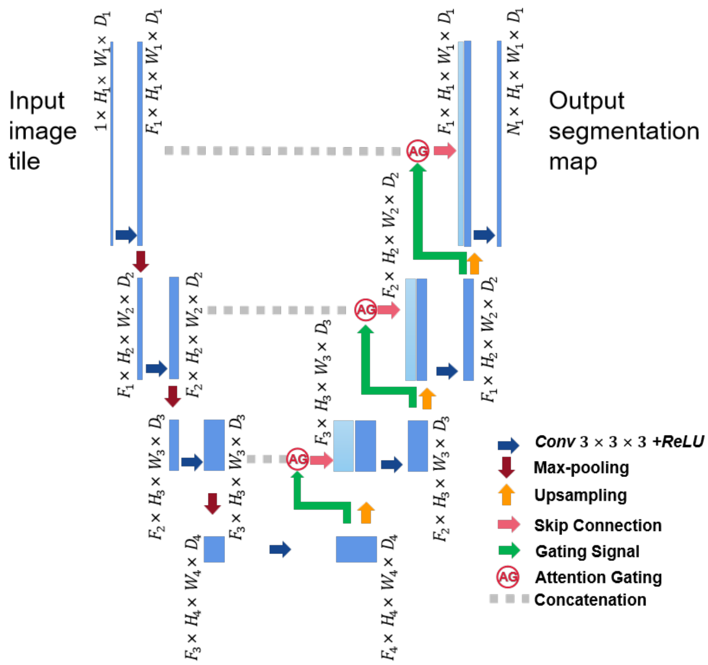

Figure 1: Model schematic diagram

The left half of the model is made up of a convolution operation and a downsampling operation. After inputting the sample information, the model first performs the convolution and ReLU operations to extract the different features of the input sample, and then performs the Max-pooling operation to reduce the pixels of the sample. In the operation, padding is set to 0 and striding is set to 1. Because there is no padding, the height and width of the feature map become smaller after each convolution, as shown in Figure 2.

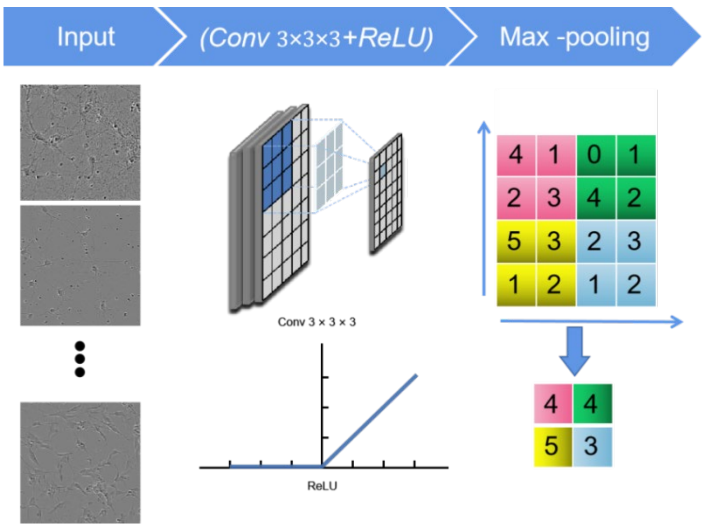

Figure 2: Convolution flow chart

After the up-sampling operation, in order to recover the original resolution of the feature map, in addition to the convolution operation, the up-sampling and Skip connection are also performed. The Skip connection is to superimpose the four previously acquired feature layers with the up-sampled feature layer according to the number of feature map channels, and the Attention mechanism is added in this process, and the two The feature layers are input into the following integral function, and the two maps are superimposed to distinguish the active region from the irrelevant region, and combined with the formula we can see that x_i^l and g_i After the convolution operation, the two feature maps of the same dimension are superimposed, in effect strengthening the signal of the same region of interest (the red part), and the respective different regions (the orange region) are also in them as auxiliary or complementary, see Figure 3 for a schematic diagram.

ISSN 2522-3488 Vol. 7, Issue 1: 53-59, DOI: 10.25236/IJNDES.2023.070109

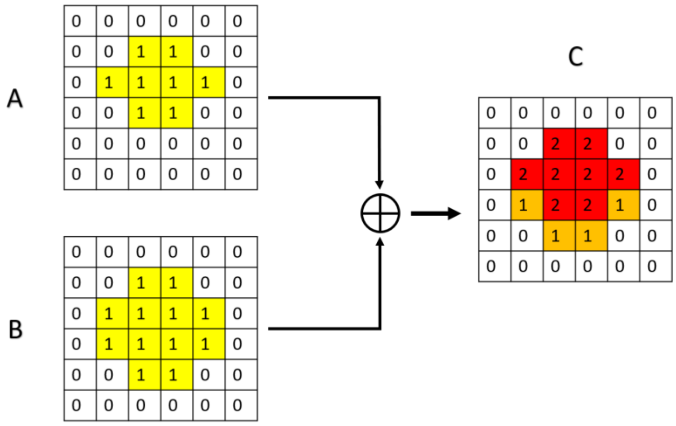

Figure 3: Attention schematic diagram

The data after processing in the previous step is then linearly and non-linearly transformed to obtain the weight values of Attention forx_i^l The process is shown in Figure 4. The above operation is repeated to output the prediction network and then perform pixel point classification.

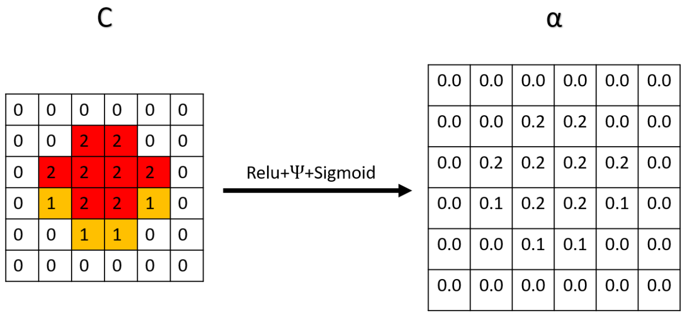

Figure 4: Attention weight map

$$q_{att}^{l} = \psi^{T}(\sigma_1(W_x^T x_i^l + W_g^T g_i + b_g)) + b_\psi \quad (1)$$

$$\alpha_i^l = \sigma_2(q_{att}^l(x_i^l, g_i; \Theta_{\text{att}})) \quad (2)$$

$$\frac{\partial(\hat{x}_i^l)}{\partial(\phi^{l-1})} = \frac{\partial(\alpha_i^l f(x_i^{l-1};\phi^{l-1}))}{\partial(\phi^{l-1})} = \alpha_i^l \frac{\partial(f(x_i^{l-1};\phi^{l-1}))}{\partial(\phi^{l-1})} + \frac{\partial(\alpha_i^l)}{\partial(\phi^{l-1})} x_i^l \quad (3)$$

3. Experimentation and analysis

In this section, the experimental environment and server hardware configuration, the detailed parameters of the training samples, the experimental procedure and analysis are presented. Because of the long training time of this experiment, one server is used for training. The following validation experiments were conducted to prove the effectiveness of the model proposed in this paper.

3.1 Experimental environment and hardware configuration

Table 1: Experimental environment

Operating System/Version	Python Version	Jupyter Notebook Version	Tensorflow Version
LINUX/CENTOS 7.6	3.9.6	6.4.5	1.5
CUDA	CPU	GPU	Memory
10.1	Intel E5-2680 v4 (2.4 GHz)	Tesla P40*6	50GB

The experimental environment and hardware configuration are shown in Table 1, and this

ISSN 2522-3488 Vol. 7, Issue 1: 53-59, DOI: 10.25236/IJNDES.2023.070109

experiment was implemented in the Jupyter Notebook environment.

3.2 Experimental sample set

In this experiment, shsy5y, cort, and astro cells were selected as segmentation objects, and the training set image samples used are shown in Figure 5. Pre-labelling the sample data and splitting the statistics by cell type helps with training and also allows for cross-validation at a later stage.

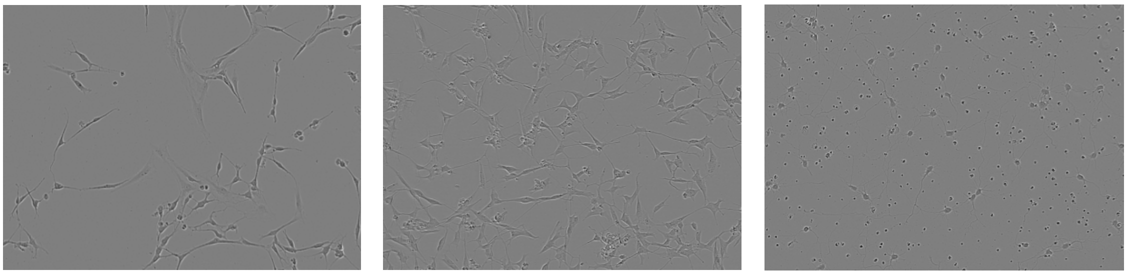

Figure 5: Import original image

3.3 Experimental procedure

We first pre-processed the images, writing the tag value information into the images to facilitate our use in the later training process, next we classified the types of images according to the pre-given dataset, first counting the number of various cells and then counting the statistics of different types of cells see Figure 6 Figure 7, we can see that the number of shsy5y is higher and the number of cort and astro are lower.

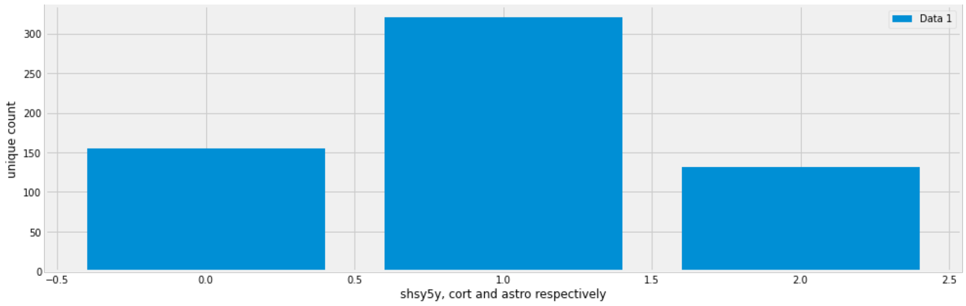

Figure 6: Cell type bar chart

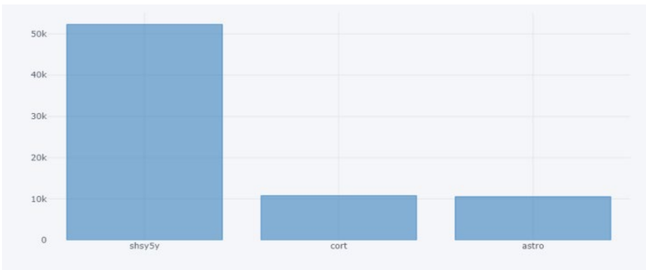

Figure 7: Cell classification

Next we generated masks of various types of each cell for comparison, and then for various types of shsy5y cells we sorted and collected the images for edge detection, thresholding, and obtaining the number of image channels to find suitable image processing methods for later image segmentation tasks, as shown in Figure 8 and Figure 9.

ISSN 2522-3488 Vol. 7, Issue 1: 53-59, DOI: 10.25236/IJNDES.2023.070109

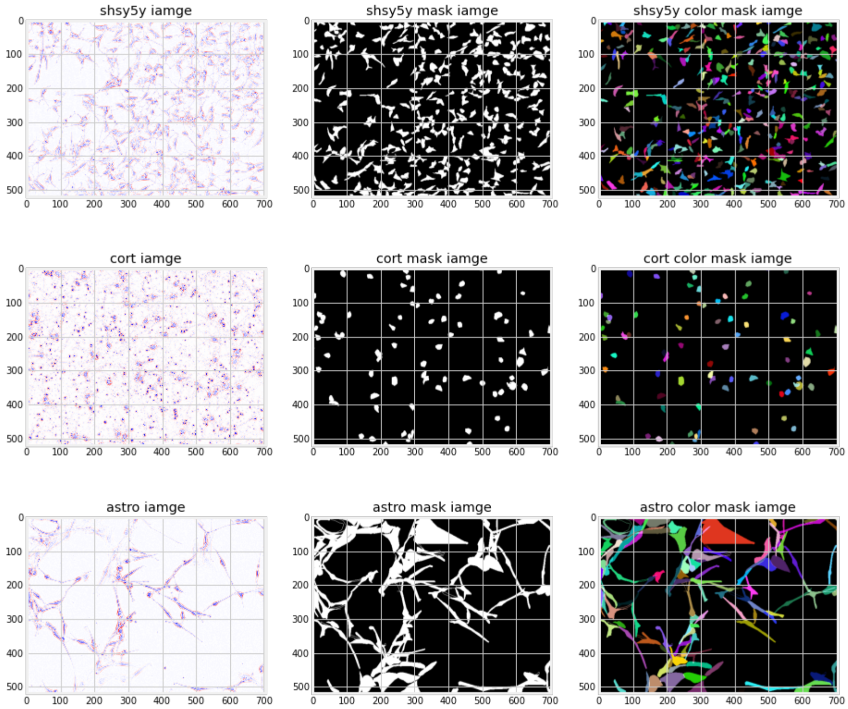

Figure 8: Subplots of shsy5y,cort and astro

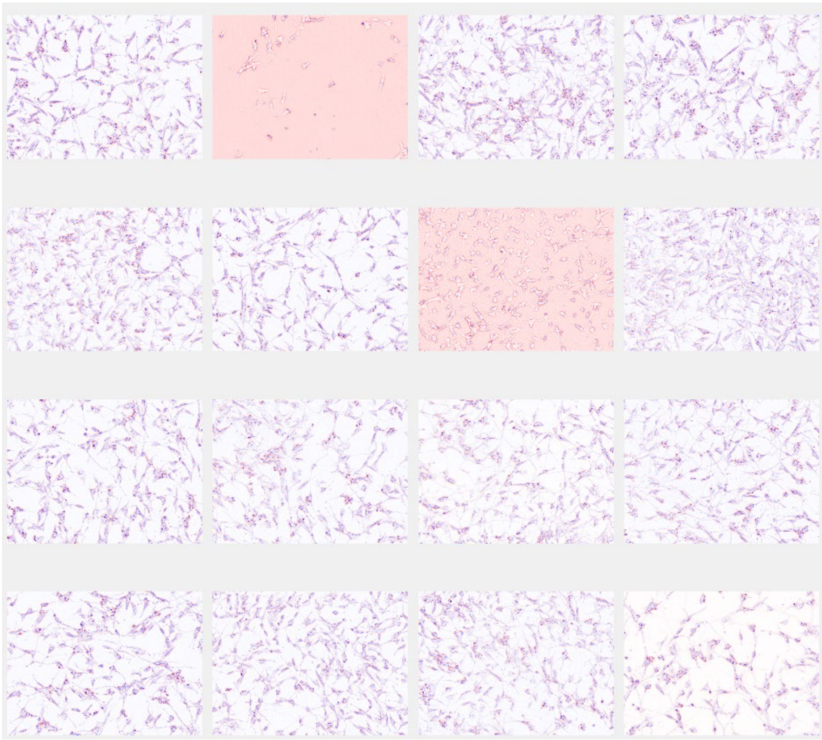

Figure 9: All shsy5y type cells

3.4 Experimental analysis

After analysing the morphology of the images and the choice of processing methods, we improved the u-net++ model and then trained our sample set. Due to the small number of images in our sample set, we chose to perform multiple iterations to improve our segmentation accuracy and reduce the loss rate. The number of iterations for this experiment was set to 60, and the loss value, IoU and accuracy during the experiment are shown in Figure 10. We can see that the loss value of the model decreases faster in the first 30 iterations, and gradually stabilises in the next 30 iterations.

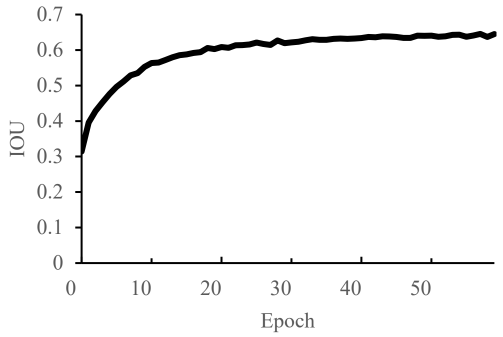

Figure 10 -1

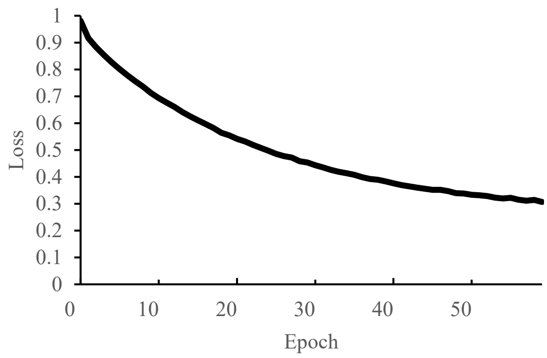

Figure 10-2

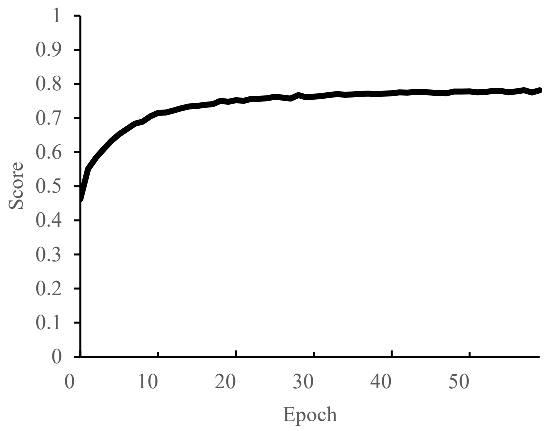

Figure 10-3

Figure 10: Training sample loss value, IoU, accuracy

4. Conclusion

This paper introduces a method for cell segmentation based on deep neural networks, adding the Attention mechanism to the U-net++ model to effectively suppress irrelevant region feature activation and improve the training efficiency of the model. The post-experimental results show that, compared with the traditional algorithm and the classical u-net model's, both achieve better cell segmentation results after segmenting cells such as shsy5y, effectively improving the training accuracy and reducing the loss rate, and the accuracy of cell segmentation is significantly improved from 30%~40% of the traditional segmentation method to about 80%, and for closer cell edges, imaging blurred This demonstrates that U-net++ with the Attention mechanism has better cell segmentation results. The experiments also further demonstrate that the deep neural network-based image processing method is more useful for medical-related research.

ISSN 2522-3488 Vol. 7, Issue 1: 53-59, DOI: 10.25236/IJNDES.2023.070109

References

[1] KOWAL M, ŻEJMO M, SKOBEL M, et al. Cell nuclei segmentation in cytological images using convolutional neural network and seeded watershed algorithm [J]. Journal of Digital Imaging, 2020, 33(1): 231 - 242.
[2] Chen H, Yu X. S., Wu C. D., et al. Fast image segmentation algorithm for parametric level set active contour model. Journal of Northeastern University (Natural Sciences Edition), 2019, 40(1): 6-10.
[3] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition [J]. Neural Computation, 1989, 1(4): 541 - 551.
[4] Byra M, Jarosik P, Szubert A, Galperin M, Ojeda-Fournier H, Olson L, O'Boyle M, Comstock C and Andre M. 2020. Breast mass segmentation in ultrasound with selective kernel U-Net convolutional neural network. Biomedical Signal Processing and Control, 61: #102027 [DOI: 10. 1016 / j. bspc. 2020. 102027]
[5] ZHOU, ZONGWEI, SIDDIQUEE, MD MAHFUZUR RAHMAN, TAJBAKHSH, NIMA, et al. UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation [J]. 2020, 39(6):1856-1867. doi:10.1109/TMI.2019.2959609.
[6] Zhang W X, Zhu Z C, Zhang Y H, Wang X Y and Ding G P. 2020e. Cell image segmentation based on residual block and attention mechanism. Acta Optica Sinica, 40(17): 76-83: 76-83