An improved U-Net for cell confluence estimation*

BAI Hua[1], LU Changhao[1], MA Ming[1], YAN Shulin[2,3], ZHANG Jianzhong[2,3], and HAN Zhibo[2,3,4]****

1. Tianjin Key Laboratory of Optoelectronic Detection Technology and Systems, School of Electronic and Information Engineering, Tiangong University, Tianjin 300387, China

2. Tianjin Key Laboratory of Engineering Technologies for Cell Phamaceutical, Tianjin 300457, China

3. National Engineering Research Center of Cell Products/AmCellGene Co., Ltd., Tianjin 300457, China

4. State Key Lab of Experimental Hematology, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin 300020, China

(Received 22 August 2021; Revised 23 February 2022)

Cell confluence is an important metric to determine the growth and the best harvest time of adherent cells. At present, the evaluation of cell confluence mainly relies on experienced labor, and thus it is not conducive to the automated cell culture. In this paper, we proposed an improved U-Net algorithm (called DU-Net) for the segmentation of adherent cells. First, the general convolution was replaced by the dilated convolution to expand the receptive fields for feature extraction. Then, the convolutional layers were combined with the batch normalization layers to reduce the dependence of the network on initialization. As a result, the segmentation *accuracy* and *F1-score* of the proposed DU-Net for adherent cells with low confluence (<50%) reached 96.94% and 93.87%, respectively, and for those with high confluence (≥50%), they reached 98.63% and 98.98%, respectively. Further, the paired t-test results showed that the proposed DU-Net was statistically superior to the traditional U-Net algorithm.

Document code: A **Article ID:** 1673-1905(2022)06-0378-7

DOI https://doi.org/10.1007/s11801-022-1129-3

Cell culture in vitro for cell expansion is the basis of cell research and application. It is particularly important for stem cells which have been proposed to be a promising candidate for cell-based therapy owing to their ability to perpetuate themselves by self-renewal and to generate mature cells of a particular tissue through differentiation[1,2]. In the process of cell culture in vitro, the cell proliferation will result in the increase of cell confluence, which is defined as the ration of the area occupied by the cells to the surface area of the culture vessel. However, when the cell confluence is too high, the cell proliferation will slow down or even stop due to the cell contact inhibition. Even worse, an excessive cell confluence can cause the shortage of nutrients and the accumulation of metabolites, which will lead to a decline in cell viability. Therefore, cell confluence is an important metric to determine the growth and the best harvest time of adherent cells. Currently, the estimation of cell confluence is still done manually. This process requires experts to observe the cells frequently, and thus it is time-consuming and heavily dependent on subjective factors such as the experts' ability and experience. In particular, the manual estimation is not conducive to the large-scale and standardized cell culture. Therefore, an objective and accurate estimation of cell confluence based on machine vision is of great significance for the development of automated and standardized cell culture.

The key to automatically analyzing cell confluence is the accurate segmentation of adherent cells, which mainly faces the following difficulties. The shapes of the adherent cells in the microscopic image are not a regular circle, but irregular and diverse. In addition, changes in cell confluence will cause significant changes in cell morphology (as shown in Fig.1), which is a great challenge for cell segmentation. In the microscopic images, the difference between the adherent cells and the background is relatively small. So it is very difficult to distinguish the cells and the background (as shown in Fig.1) especially under the condition of high cell confluence, which puts forward higher requirements for the accuracy and generalizability of the algorithm.

Several traditional cell segmentation algorithms[3-6] such as thresholding, edge detection, region-based, and clustering analysis algorithms have been proposed. However, most of them are based on traditional image processing techniques, which have high requirements for the differences between the cells and the background in the images, and usually require manual adjustment of various parameters to optimize their performance. The convolutional neural network (CNN) based methods

* This work has been supported in part by the National Natural Science Foundation of China (No.61201106), and the Tianjin Special Program of Science and Technology (No.19JCTPJC47900).

** E-mails: maming@tiangong.edu.cn; zhibohan@163.com

have been widely used in object detection[7], semantic segmentation[8], image classification[9], etc. The semantic segmentation networks constantly refresh the record of deep learning in image segmentation. At present, the representative semantic segmentation networks are fully convolutional network (FCN)[10], U-Net[11], SegNet[12], pyramid scene parsing network (PSPNet)[13], DeepLabv3+[14], dual attention network (DANet)[15], etc, which can be expected to provide a more effective way to solve the problems in cell segmentation.

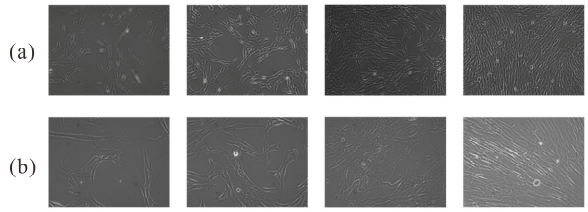


Fig.1 Microscopic images of adherent cells: (a) The cell confluences are 17%, 45%, 75%, and 96%, respectively; (b) The cell confluences are 25%, 43%, 70%, and 95%, respectively

In this paper, we propose a novel method suitable for segmentation of microscopic images of adherent cells by improving the traditional U-Net. This method can accurately segment the adherent cells with irregular shapes and automatically analyze the cell confluence under different cell densities, which is of great significance for the development of automated cell culture in vitro.

Cell segmentation is still one of the most challenging problems in medical image processing. In recent years, with the rapid development of medical image processing technology, scholars have proposed a variety of cell segmentation methods. Most of these methods are based on traditional image processing techniques. For example, SOLEIMANI et al[16] proposed a method that consists of the normalization of the uneven background, contrast adjustment and denoising with block matching 3D (BM3D) filter for cell confluence estimation. SHAO et al[17] presented an approach using the improved Gauss-Laplacian operator, the mean shift algorithm and the mathematical region to calculate cell confluence. WANG et al[18] proposed an algorithm including edge detection, entropy filtering, range filtering and a halo recognition technique to estimate cell confluence. In summary, most of the above methods used artificially formulated rules such as artificial thresholds, which limit the generalizability of the algorithms. In addition, these methods are multi-stage workflows, and thus their performance relies heavily on the previous steps[19].

Currently, semantic segmentation networks are becoming more accurate and faster. Some networks begin to dominate the tasks of the cell image segmentation. For example, BINICI et al[20] improved SegNet to realize automated segmentation of cells in phase contrast microscopy (PCM) images. TSAI et al[21] used mask region convolutional neural network (R-CNN) which has a backbone of residual neural network (ResNet)-101 to segment cells in an instance-aware manner. AYANZADEH et al[22] designed a network which applies the modified ResNet18 in the encoder and the residual blocks in the decoder for cell segmentation in PCM images. The above works show the great potential of deep learning in cell segmentation. However, these works only focus on the cases under low cell confluence, but those under high cell confluence have not been addressed.

Here, we propose a new strategy that expands the receptive fields and reduces the dependence on initialization to improve the network's adaptability and the ability to recognize adherent cells under low and high cell confluence. Considering that U-Net has a prominent advantage in dealing with a small number of samples and biomedical image processing[11], we use U-Net as the basic network architecture. Further, we introduce the dilated convolution[23] and the batch normalization[24] to build an improved U-Net network for the segmentation of adherent cells.

The pictures used in this study were the microscopic images of adherent Human Umbilical Cord Mesenchymal Stem Cells (hUC-MSCs). The umbilical cords were obtained from healthy mothers with well-developed fetuses. The donors voluntarily donated the umbilical cords and signed the informed consent for donation. Isolation and culture of hUC-MSCs were carried out in accordance with the relevant standard procedures of Tianjin AmCellGene Engineering Company Limited. A brief description of the experimental procedure[25] is as follows. First, the donors must be tested for human immunodeficiency virus (HIV), hepatitis B virus (HBV), hepatitis C virus (HCV), treponema pallidum and other pathogenic microorganisms before the separation of hUC-MSCs from their umbilical cords. Secondly, the umbilical cords were repeatedly washed with phosphate buffered saline (PBS) and then mechanically cut into pieces, and digested with digestive enzymes at 37 °C. The digested product was filtered through a 200-mesh sterile filter to obtain a cell suspension, which was centrifuged at 1 500 r/min for 10 min. Finally, after removing the supernatant, the cells were resuspended in DF-12 cell culture medium containing 10% fetal bovine serum. Cells were plated in a T75 culture flask and cultured in a humidified incubator at 37 °C and 5% CO_2. In this study, three different hUC-MSCs samples were cultured in vitro. During the cell culture process, a microscope (IX71, Olympus Corporation) was used to collect microscopic images of adherent cells every 24 h. The microscopic images of adherent cells corresponding to the three cell samples were classified into different groups (named group 1, group 2, and group 3). The labels of the adherent cells data set were manually annotated by the experts.

The shapes of adherent cells are diverse. Especially, most hUC-MSCs are spindle-shaped (as shown in Fig.1). Therefore, compared to the case of circular suspension cells, the feature extraction of the irregularly shaped adherent cells requires a larger receptive field, which is

more conducive to extracting features of large-sized or irregularly shaped objects. In this regard, the dilated convolution is an ideal choice since it can support exponentially expanding receptive fields without losing resolution or coverage[23]. Compared with the general convolution, the dilated convolution can obtain a larger receptive field without increasing the computation[26], and thus it is more suitable for the segmentation of irregularly shaped adherent cells. The general convolution can be described as

$$y(P_i, P_j) = \sum_{m=-r}^{r} \sum_{n=-r}^{r} F(P_i + m, P_j + n) K(m, n), \quad (1)$$

where the function $F(P_i, P_j)$ is the input. The function $K(m, n)$ is the filter or the convolution kernel and r is the size of it. The output $y(P_i, P_j)$ is the feature map. Based on Eq.(1), the dilated convolution operation can be defined as

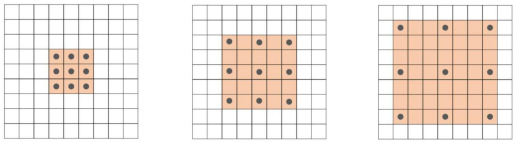
$$y(P_i, P_j) = \sum_{m=-r}^{r} \sum_{n=-r}^{r} F(P_i + lm, P_j + ln) K(m, n), \quad (2)$$

where l is the dilation factor that determines the distribution of sampling locations in the dilated convolution. In our study, the filter size is determined to be 3×3, so the distribution of sampling locations corresponding to different values of l can be illustrated in Fig.2.

Obviously, when l is equal to 1, the 1-dilated convolution is a general convolution. If l is greater than 1, there is an interval between the sampling locations of the dilated convolution, and the interval depends on the value of l. Therefore, setting the value of l reasonably can expand the receptive fields without increasing the size of the convolution kernel. The size of the receptive fields can be expressed as follows[27]

$$RF_i = RF_{i-1} + (K_i - 1) \times L_i \times \prod_{n=1}^{i-1} Stride_n, \quad (3)$$

where RF_i denotes the size of the receptive field of each element in the i-th feature map. K_i and L_i represent the filter size and the dilated factor of the i-th convolution layer, respectively. $Stride_n$ represents the stride of the n-th convolution layer. It can be observed from Eq.(3) that when the dilated convolutions are used in combinations, the receptive fields can be exponentially expanded.



(a) l=1 (b) l=2 (c) l=3

Fig.2 Illustration of the sampling locations in the dilated convolution with different values of *l*: (a) *l*=1; (b) *l*=2; (c) *l*=3 (The points represent the sampling locations, and the yellow squares represent the receptive field of each element after convolution)

The main structure of the improved U-Net algorithm (called DU-Net) is shown in Fig.3. The DU-Net is mainly composed of 3×3 general convolutions. Based on the idea of encoder-decoder, it consists of a contracting path on the left (capture context information) and an expanding path on the right (support precise localization). And the paths on the left and right are symmetrical to each other, which belongs to an end-to-end network. The input image first performs a general convolution operation, and then performs the dilated convolution operation. Two layers of 2-dilated convolution are used in combinations after the general convolution at the head of the network, which can achieve a receptive filed of 11×11. In this way, sufficient bottom features can be obtained and provide strong support for the subsequent formation of abstract features. In addition, it is necessary to perform upsampling after the contracting path ends in order to make the size of the input and output the same. At the same time, copy and crop operations are performed on the feature map, and then the feature map of the contracting path and the feature map of expanding path are performed with concatenation operations in the same stage. The concatenation operation plays a role of supplementing semantic information to ensure that the feature map incorporates more low-level features.

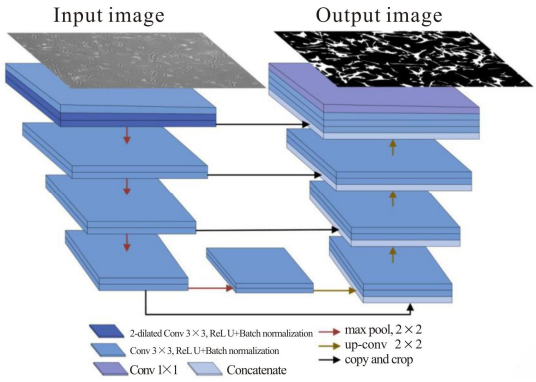


Fig.3 Network structure diagram of DU-Net (DU-Net takes the raw images with size of 512×384 as input and generates prediction maps with the same resolution)

In addition, batch normalization is used in our network. Specifically, the function of batch normalization is to regulate the input of each convolutional layer so that the mean value is 0 and the variance is 1[24]. Adding the batch normalization layer after the convolutional layer helps to enable a higher learning rate, eliminate the need for Dropout[28] and be insensitive about initialization[24].

The segmentation of adherent cells with irregular shapes is a new challenge in the field of cell segmentation. In order to explore the segmentation effect of semantic segmentation networks in this case, we employed six representative networks, including FCN, PSPNet, DeepLabv3+, DANet, UNet++ and U-Net, as well as our proposed DU-Net for adherent cell segmentation. All the networks were trained on a Nvidia Tesla V100 graphics processing

unit (GPU) and trained appropriately to achieve their best performance. The values of hyperparameter of the networks are shown in Tab.1.

For the above six networks, DeepLabv3+ is implemented on the deep learning framework Tensorflow. FCN, PSPNet, DANet and U-Net are all implemented using Keras. Keras is a machine learning library that uses TensorFlow as the backend. In addition, UNet++ is completed on the deep learning framework Pytorch. In order to obtain a reliable model evaluation, we employed three sets of microscopic images, with the numbers of 33, 34, and 33 respectively, to perform 3-fold cross validation. Specifically, during each experiment, we used two subsets to train the model and the remaining subset for testing.

Tab.1 Hyperparameters of the networks

Network	Parameter	Value
FCN	Epochs	150
	Batch size	4
	Learning rate	0.000 1
DeepLabv3+	Max iteration	5 000
	Learning rate	0.000 1
	Batch size	2
DANet/PSPNet/UNet++	Epochs	120
	Batch size	4
	Learning rate	0.000 1
U-Net	Epochs	20
	Batch size	4
	Learning rate	0.001

In order to evaluate quantitatively the pixel-level segmentation performance of the networks, two parameters including *Accuracy* and *F1-score* were used as evaluation indices. The *Accuracy* index reflects comprehensively the segmentation performance of the network on the foreground and background, and its expression is shown in Eq.(4). The *F1-score* index focuses on evaluating the accuracy of the segmentation of targets by the network, and can be calculated by Eq.(5), Eq.(6), and Eq.(7).

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}, \tag{4}$$

$$Precision = \frac{TP}{TP + FP}, \tag{5}$$

$$Recall = \frac{TP}{TP + FN}, \tag{6}$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall}, \tag{7}$$

where TP is the number of pixels correctly identified as cells. TN is the number of pixels correctly identified as the background. FP is the number of pixels incorrectly identified as cells, whereas FN is the number of pixels incorrectly identified as the background.

It should be noted that the cell confluence is an important factor affecting the segmentation of adherent cells. As shown in Fig.1, the increase in cell confluence will cause changes in cell morphology and make it more difficult to distinguish between the cells and the background. There is no doubt that it is a new challenge to the accuracy and generalization ability of the algorithms. Therefore, we evaluated the segmentation performance of the above networks under two conditions, i.e., low cell confluence (<50%) and high cell confluence (⩾50%). The evaluation results of the six semantic segmentation networks under low and high cell confluence are listed in Tab.2, and their segmentation effects are shown in Fig.4.

Tab.2 Quantitative evaluation results under different cell confluences

	Method	Low cell confluence (<50%)		High cell confluence (≥50%)	
		Accuracy	*F1-score*	*Accuracy*	*F1-score*
Group 1	FCN-8s	0.935 7±0.021	0.822 3±0.087	0.976 6±0.029	0.976 8±0.035
	PSPNet	0.927 3±0.034	0.826 7±0.074	0.975 0±0.034	0.975 2±0.041
	DeepLabv3+	0.741 5±0.082	0.663 8±0.093	0.934 4±0.117	0.955 6±0.085
	DANet	0.926 5±0.038	0.830 6±0.064	0.968 2±0.038	0.970 1±0.045
	UNet++	0.942 2±0.023	0.853 6±0.069	0.979 8±0.026	0.980 1±0.032
	U-Net	0.957 7±0.017	0.918 6±0.045	0.981 9±0.024	0.986 3±0.022
	DU-Net	0.966 3±0.016	0.937 2±0.037	0.986 3±0.018	0.989 8±0.017
Group 2	FCN-8s	0.926 5±0.006	0.804 3±0.091	0.969 9±0.025	0.974 7±0.026
	PSPNet	0.912 6±0.014	0.790 3±0.076	0.965 8±0.024	0.971 7±0.025
	DeepLabv3+	0.776 5±0.051	0.656 5±0.108	0.914 7±0.062	0.946 9±0.050
	DANet	0.913 7±0.009	0.773 1±0.105	0.955 8±0.036	0.962 4±0.043
	UNet++	0.935 0±0.007	0.851 0±0.050	0.972 4±0.024	0.976 9±0.024
	U-Net	0.926 5±0.006	0.804 3±0.091	0.972 5±0.017	0.983 2±0.015
	DU-Net	0.957 6±0.008	0.918 3±0.041	0.981 1±0.011	0.988 5±0.009

	FCN-8s	0.940 6±0.018	0.809 1±0.111	0.965 0±0.019	0.969 6±0.020
	PSPNet	0.922 0±0.035	0.793 4±0.095	0.955 5±0.025	0.963 0±0.022
	DeepLabv3+	0.500 2±0.219	0.553 6±0.136	0.841 1±0.106	0.924 5±0.047
Group 3	DANet	0.925 1±0.031	0.801 2±0.076	0.945 6±0.025	0.953 9±0.027
	UNet++	0.942 6±0.028	0.853 4±0.069	0.973 7±0.015	0.978 0±0.013
	U-Net	0.957 0±0.017	0.913 6±0.033	0.971 3±0.012	0.981 9±0.009
	DU-Net	0.969 4±0.011	0.938 7±0.021	0.977 0±0.011	0.985 5±0.008

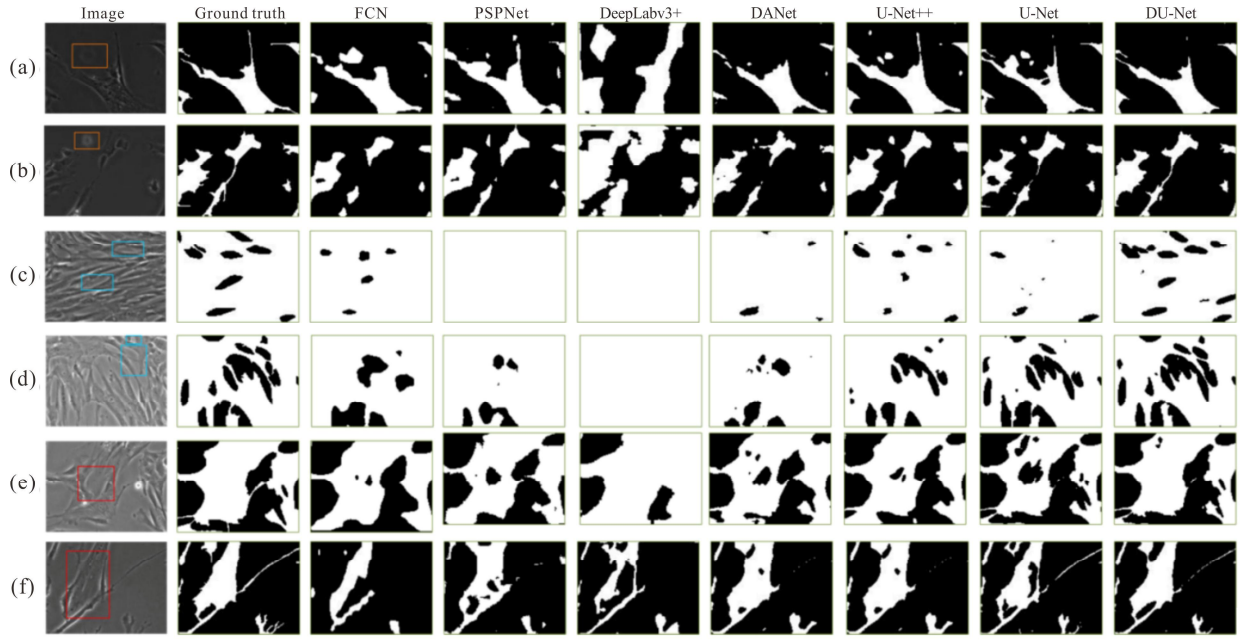


Fig.4 Intuitive comparison of segmentation effects of FCN, PSPNet, DeepLabv3+, DANet, UNet++, U-Net and DU-Net (The black parts represent the background and interference factors, the white parts represent the adherent cells, and the rectangles in the original images mark the locations that are misjudged by the U-Net): (a)—(f) Typical cell microscopic images with different brightnesses, contrasts and cell confluences

Obviously, the segmentation accuracies of DeepLabv3+, PSPNet, DANet and FCN are not high enough, especially under low cell confluence. In addition, their segmentation performance under low and high cell confluence shows large differences, indicating that they have insufficient stability and generalization ability for adherent cell segmentation. As for DeepLabv3+, the features of cells in small sizes can not be well captured by atrous spatial pyramid pooling module with large dilation rate[19]. Similar to DeepLabv3+, the pyramid pooling module of PSPNet is also not conducive to capturing the features of cells in small sizes. The cells are relatively scattered under low confluence, thus DANet can not capture global context information to establish the spatial dependencies through its position attention mechanism modules. As for FCN, when FCN completes the fusion operation of feature maps, it will directly perform upsampling by a factor of 8 to restore the size of the input image. These may be the important reasons why the segmentation results of the four networks are not accurate enough and are not sensitive to the details of the cells in the image.

Compared to FCN, PSPNet, DANet and DeepLabv3+, U-Net and UNet++ have much better performance in the segmentation of adherent cells. Between U-Net and UNet++, since there is a series of nested and dense skip pathways in UNet++, small targets are easily lost by the repeated downsampling and upsamling operations. Therefore, the overall segmentation effect of UNet++ is not as good as that of U-Net. U-Net can acquire abundant multi-scale features by concatenating the feature maps in the contracting path and the symmetric feature maps in the expanding path. Besides, the random elastic deformations of U-Net can realize data augmentation, which is conducive to solving the problem of a small number of samples. These advantages result in the good performance of U-Net in the cell segmentation. However, U-Net generally uses a 3×3 convolution kernel, making the receptive filed limited, which is not conducive to the segmentation of adherent cells with irregular shapes.

Due to the addition of the dilated convolution and the batch normalization, DU-Net has a larger receptive field and less initialization dependence. As a result, compared with those of U-Net, the *Accuracy* and *F1-score* of

DU-Net increased by 1.24% and 2.51% respectively under low cell confluence, and increased by 0.86% and 0.53% respectively under high cell confluence. Further, in order to examine whether the advantages of DU-Net are statistically significant, we used the paired t-test to compare statistically the performance of DU-Net and U-Net. The results show that the *P* values corresponding to the *Accuracy* and *F1-score* are 1.5×10^{-17} and 1.7×10^{-11} respectively, which further proves that DU-Net has a higher segmentation accuracy than U-Net. Moreover, the standard deviation of DU-Net is lower than that of U-Net, indicating that DU-Net has a better stability and a stronger generalization ability.

The intuitive comparison of segmentation effects of DU-Net and U-Net is shown in Fig.4. It can be seen that DU-Net has excellent performance in different cases. First of all, DU-Net can distinguish between the cells and interference factors such as dead cells, impurities in biological reagents, artifacts, and halos produced by optical instruments, etc. However, U-Net faces challenges in this regard, and is likely to misjudge the interference factor as the cells, as shown in the yellow rectangles in Fig.4. Secondly, under low cell confluence, the morphology of adherent cells is relatively more diverse and irregular, which makes it difficult for U-Net to correctly judge some cells, but misjudges them as back-ground, as shown in the red rectangles in Fig.4. On the contrary, DU-Net is more accurate in identifying the adherent cells. In addition, under high cell confluence, the background has very small areas and is difficult to distinguish from the cells, which makes it easy for U-Net to misjudge the background as the cells, as shown in the blue rectangles in Fig.4. In contrast, DU-Net has much better performance, and the small background areas can also be accurately segmented. These results further show that the dilated convolution combined with batch normalization can capture more features of adherent cells with irregular shapes, and thus effectively improve the performance of U-Net in the cell segmentation.

The automatic and objective estimation of cell confluence is of great significance to the in vitro culture of adherent cells. In this paper, we improved U-Net by introducing the dilated convolution and batch normalization for adherent cell segmentation. The results show that our proposed DU-Net can get a suitable receptive field to capture abundant multi-scale features of adherent cells. Whether the cell confluence is high or low, the *Accuracy* and *F1-score* of DU-Net are significantly higher than those of classic semantic segmentation networks, including FCN, DeepLabv3+, and U-Net. The results indicate that the proposed DU-Net has an excellent accuracy and generalization ability for the segmentation of adherent cells with irregular shapes, and thus can provide a new powerful tool for the automatic and objective estimation of cell confluence.

Statements and Declarations

The authors declare that there are no conflicts of interest related to this article.

References

[1] SONG N, SCHOLTEMEIJER M, SHAH K. Mesenchymal stem cell immunomodulation：mechanisms and therapeutic potential[J]. Trends in pharmacological sciences, 2020, 41(9)：653-664.

[2] BLAU H M, DALEY G Q. Stem cells in the treatment of disease[J]. The new England journal of medicine, 2019, 380(18)：1748-1760.

[3] GHANE N, VARD A, TALEBI A, et al. Segmentation of white blood cells from microscopic images using a novel combination of K-Means clustering and modified watershed algorithm[J]. Journal of medical signals & sensors, 2017, 7(2)：92-101.

[4] ARULMURUGAN R, ANANDAKUMAR H. Region-based seed point cell segmentation and detection for biomedical image analysis[J]. International journal of biomedical engineering and technology, 2018, 27(4)：273-289.

[5] JO H J, HAN J H, KIM Y S, et al. A novel method for effective cell segmentation and tracking in phase contrast microscopic images[J]. Sensor, 2021, 21(10)：3516.

[6] MOTA S M, ROGERS R E, HASKELL A W, et al. Automated mesenchymal stem cell segmentation and machine learning-based phenotype classification using morphometric and textural analysis[J]. Journal of medical imaging, 2021, 8(1)：014503.

[7] ZHAO Z Q, ZHENG P, XU S T, et al. Object detection with deep learning：a review[J]. IEEE transactions on neural networks and learning systems, 2019, 30(11)：3212-3232.

[8] TAGHANAKI S A, ABHISHEK K, COHEN J P, et al. Deep semantic segmentation of natural and medical images：a review[J]. Artificial intelligence review, 2021, 54(1)：137-178.

[9] RAWAT W, WANG Z H. Deep convolutional neural networks for image classification：a comprehensive review[J]. Neural computation, 2017, 29(9)：2352-2449.

[10] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition, June 7-12, 2015, Boston, MA, USA. New York：IEEE, 2015：3431-3440.

[11] RONNEBERGER O, FISCHER P, BROX T. U-Net：convolutional networks for biomedical image segmentation[C]//Proceedings of the Medical Image Computing and Computer-Assisted Intervention, October 5-9, 2015, Munich, Germany. Berlin Heidelberg：Springer, 2015：234-241.

[12] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet：a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(12)：

2481-2495.

[13] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York：IEEE, 2017：2881-2890.

[14] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the European Conference on Computer Vision (ECCV), September 8-14, 2018, Munich, Germany. Berlin Heidelberg：Springer, 2018：801-818.

[15] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 15-20, 2019, Long Beach, USA. New York：IEEE, 2019：3146-3154.

[16] SOLEIMANI S, MIRZAEI M, TONCU D C. A new method of SC image processing for confluence estimation[J]. Micron, 2017, 101：206-212.

[17] SHAO Y B, XU Z L, YAO T Z. Calculation of cell growth confluence based on region growth[J]. Computer science and application, 2020, 10(2)：236-244.

[18] WANG Z N, ZHAN R C, HU Y. Automated confluence measurement method for mesenchymal stem cell from brightfield microscopic images[J]. Microscopy and microanalysis, 2021, 27(5)：1093-1101.

[19] ZHANG M, LI X, XU M J, et al. Automated semantic segmentation of red blood cells for sickle cell disease[J]. IEEE journal of biomedical and health informatics, 2020, 24(11)：3095-3102.

[20] BINICI R C, AYANZADEH A, TÖREYIN B U, et al. Automated segmentation of cells in phase contrast optical microscopy time series images[C]//2019 Medical Technologies Congress, October 3-5, 2019, Izmir, Turkey. New York：IEEE, 2019：1-4.

[21] TSAI H F, GAJDA J, SLOAN T, et al. Usiigaci：instance-aware cell tracking in stain-free phase contrast microscopy enabled by machine learning[J]. SoftwareX, 2019, 9：230-237.

[22] AYANZADEH A, ÖZUYSAL Ö Y, OKVUR D P, et al. Deep learning based segmentation pipeline for label-free phase-contrast microscopy images[C]//2020 28th Signal Processing and Communications Applications Conference, October 5-7, 2020, Gaziantep, Turkey. New York：IEEE, 2021：1-4.

[23] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[C]//Proceedings of 2016 the Tenth International Conference on Learning Representations, May 2-4, 2016, Caribe Hilton, San Juan, Puerto Rico.

[24] IOFFE S, SZEGEDY C. Batch normalization：accelerating deep network training by reducing internal covariate shift[C]//Proceedings of the 32nd International Conference on Machine Learning, July 6-11, 2015, Lille, France. Pittsburg：IMLS, 2015：448-456.

[25] GONG W, HAN Z B, ZHAO H, et al. Banking human umbilical cord derived mesenchymal stromal cells for clinical use[J]. Cell transplantation, 2012, 21(1)：207-216.

[26] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation：version 3[EB/OL]. (2017-12-05) [2021-08-20]. https：//arxiv.org/abs/1706.05587.

[27] LIAO J C, CAO L B, LI W, et al. UnetDVH-Linear：linear feature segmentation by dilated convolution with vertical and horizontal kernels[J]. Sensors, 2020, 20(20)：5759.

[28] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout：a simple way to prevent neural networks from overfitting[J]. Journal of machine learning research, 2014, 15(1)：1929-1958.