



# MF-Net: Automated Muscle Fiber Segmentation From Immunofluorescence Images Using a Local-Global Feature Fusion Network

Getao Du · Peng Zhang<sup>2</sup> · Jianzhong Guo<sup>3</sup> · Xiangsheng Pang<sup>2</sup> · Guanghan Kan<sup>2</sup> · Bin Zeng<sup>2</sup> · Xiaoping Chen<sup>2</sup> · Jimin Liang<sup>4</sup> · Yonghua Zhan<sup>1</sup>

Received: 6 June 2023 / Revised: 20 July 2023 / Accepted: 24 July 2023 / Published online: 15 September 2023  
© The Author(s) under exclusive licence to Society for Imaging Informatics in Medicine 2023

## Abstract

Histological assessment of skeletal muscle slices is very important for the accurate evaluation of weightless muscle atrophy. The accurate identification and segmentation of muscle fiber boundary is an important prerequisite for the evaluation of skeletal muscle fiber atrophy. However, there are many challenges to segment muscle fiber from immunofluorescence images, including the presence of low contrast in fiber boundaries in immunofluorescence images and the influence of background noise. Due to the limitations of traditional convolutional neural network–based segmentation methods in capturing global information, they cannot achieve ideal segmentation results. In this paper, we propose a muscle fiber segmentation network (MF-Net) method for effective segmentation of macaque muscle fibers in immunofluorescence images. The network adopts a dual encoder branch composed of convolutional neural networks and transformer to effectively capture local and global feature information in the immunofluorescence image, highlight foreground features, and suppress irrelevant background noise. In addition, a low-level feature decoder module is proposed to capture more global context information by combining different image scales to supplement the missing detail pixels. In this study, a comprehensive experiment was carried out on the immunofluorescence datasets of six macaques' weightlessness models and compared with the state-of-the-art deep learning model. It is proved from five segmentation indices that the proposed automatic segmentation method can be accurately and effectively applied to muscle fiber segmentation in shank immunofluorescence images.

**Keywords** Weightless muscle atrophy · Immunofluorescence images · Deep learning · Segmentation · Transformer · Low-level feature decoder module

## Introduction

Due to the lack of gravity load, astronauts' long-term orbit flight can easily lead to weightless muscle atrophy, which greatly affects their physical activity. Therefore, it is very important to establish an accurate assessment method for muscle atrophy to ensure the flight safety of astronauts. Generally speaking, muscle atrophy is mainly manifested in the decrease of muscle volume, decrease of the cross-sectional area (CSA) of the muscle fiber, and change in muscle fiber types. Figure 1a, b show the normal and atrophied muscle fiber [1–3]. Because of this, immunofluorescence (IF) staining images are often used to evaluate this change in the field of muscle atrophy, mainly through manually identifying muscle fibers in the images and measuring muscle fiber size [4, 5]. However, this manual tracking method of a single muscle fiber is relatively subjective and time-consuming, which fundamentally makes the accuracy of segmentation

✉ Xiaoping Chen  
xpchen2009@163.com

✉ Jimin Liang  
jiminliang@gmail.com

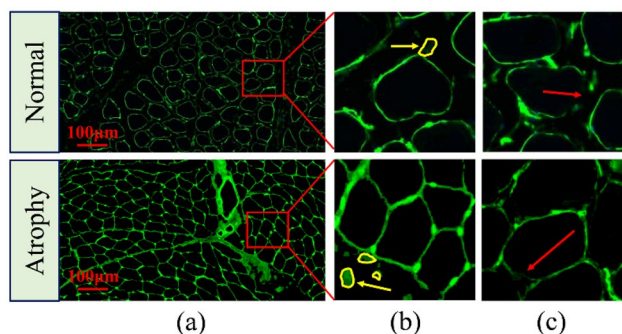
✉ Yonghua Zhan  
yhzhan@xidian.edu.cn

<sup>1</sup> School of Life Science and Technology, Xidian University & Engineering Research Center of Molecular and Neuro Imaging, Ministry of Education, Xi'an, Shaanxi 710126, China

<sup>2</sup> China Astronaut Research and Training Center, Beijing 100094, People's Republic of China

<sup>3</sup> Institute of Applied Acoustics, School of Physics and Information Technology, Shaanxi Normal University, Xi'an 710062, China

<sup>4</sup> School of Electronic Engineering, Xidian University, Xi'an, Shaanxi 710071, China



**Fig. 1** An example of IF muscle fiber image segmentation. **a** IF images of normal shank muscle tissue and atrophic shank muscle tissue. **b** An example of low contrast between muscle fibers and background noise in IF images, and the yellow arrow indicates the background noise area. **c** An example of muscle fiber boundary blurring in an IF image. The red arrow indicates the muscle fiber blurring boundary

impossible to guarantee. Therefore, it is urgent to develop an effective automatic muscle fiber segmentation algorithm to reduce manual intervention, so as to improve the accuracy of the quantitative analysis of muscle fibers. However, due to the limitations of IF imaging of muscle fiber tissue, on the one hand, the contrast difference between the muscle fiber boundary area and background noise is small, which increases the difficulty of segmentation. On the other hand, due to the low contrast and blurring of the boundary of muscle fibers, they cannot be completely segmented. In order to deal with these limitations, many studies have been carried out [6–8]. Rahmati and Rashno designed a fully automated segmentation method based on the Central Intelligence Set algorithm for analyzing the entire skeletal muscle cross-section of motion-induced regenerated muscle fibers. However, this method is cumbersome and lacks sufficient representation ability to adjust parameters [9]. Veta et al. proposed a marker controlled by watershed technology, which uses multiple marker types and a relatively simple but effective combination of segmentation generated at different scales and multiple markers [10]. Islam et al. used a locally convergent filter with a sliding band filter to deal with unique low contrast and noise problems in IF images [11]. Although these methods have improved the accuracy of IF image segmentation, the limitations of IF imaging have not been fundamentally solved.

In recent years, with the proposal and continuous development of deep learning (DL), it can learn the internal laws and representation levels of imaging data and effectively extract representative features from muscle fiber regions and background of the image [12, 13]. Ronneberger et al. proposed the U-net network based on a fully convolutional network (FCN) [14] in the International Symposium on Biomedical Imaging (ISBI) cell tracking challenge in 2015 and applied end-to-end training to medical image segmentation,

which attracted wide attention [15]. However, it is challenging to directly apply U-net to muscle fiber segmentation. The main reasons are as follows: (1) as shown by the yellow arrow mark in Fig. 1b, because of the low contrast between the muscle fiber boundary area and background noise (vessels or nerves), it is easy to increase the segmentation error and reduce segmentation accuracy in the IF image, and (2) the muscle fibers in the IF image have the problem of a fuzzy boundary (as shown in Fig. 1c). Previous studies mainly used the fusion strategy of acquiring edge information and multi-stage learning through additional edge constraints. For example, Hirohisa et al. proposed the boundary enhancement segmentation network (BESNet), in which the boundary enhancement is based on adaptive weighting of the loss of cell boundaries that are not strongly enhanced [16]. In addition, Yi et al. proposed a focused instance segmentation method to solve the problems of blurring and low structure contrast in cell images. It combines a single multiple box detector (SSD) and a U-net. In addition, an attention mechanism is used in both detection and segmentation modules to make the model pay more attention to useful features [17]. However, these methods ignore the characteristic response of background noise and the particularity of different tasks. As the transformer mechanism was formally put forward in 2017 [18], it will be combined with U-net for the first time in image segmentation in 2021. Its advantages lie in solving the inherent limitations of convolutional operations on the one hand and addressing the limitations of low-level details on localization ability on the other hand [19]. In view of this, using the DL model to carry out automatic segmentation and recognition of muscle atrophy IF images will undoubtedly provide potential opportunity for accurate evaluation of weightlessness muscle atrophy. However, there are no relevant research reports concerning this topic so far.

Inspired by the above DL segmentation model, we propose a new dual encoder branch FCN, namely, MF-Net, to solve the challenge of muscle fiber boundary segmentation in IF images. Specifically, MF-Net takes U-net as the main network, in which there are two encoder branches, low-level feature decoder (LLFD) module and a decoder branch. Unlike the previous segmentation network [20, 21], the proposed MF-Net implements a dual encoder branch that combines convolutional neural networks (CNN) and a transformer to capture both local and global feature information; its purpose is to focus on learning the features of muscle fibers during the training process, suppress the feature response of noisy areas, and thus completely segment the muscle fiber boundary areas from intermediate frequency images, improving the accuracy of segmentation. To solve the problem of the fuzzy boundary of muscle fibers in IF images, we propose an LLFD module, which up-samples the high-level features in the CNN encoder and skip connects them with the corresponding layers in the decoder branch to

mine deeper feature information and make up for lost detail pixels. In addition, the combined loss function used in this study is composed of Dice loss and cross-entropy loss. Dice loss can solve the problem of imbalance between positive and negative samples in IF images due to overlap clipping. The whole workflow is shown in Fig. 2. The main contributions of this paper are as follows:

- In this study, the macaque was used as the experimental object for the first time. The weightlessness environment

was simulated through the minus 10 degrees BR experiment. The DL model was applied to the segmentation of muscle fibers in the IF images of the macaque shank, and its feasibility was verified.

- Combined with the limitations of IF images of muscle fibers, this study proposes a dual encoder branch learning framework. Compared with a single U-net, with the help of the transformer encoder branch CNN, this method can capture both local information and global information of muscle fiber. In addition, the characteristic response

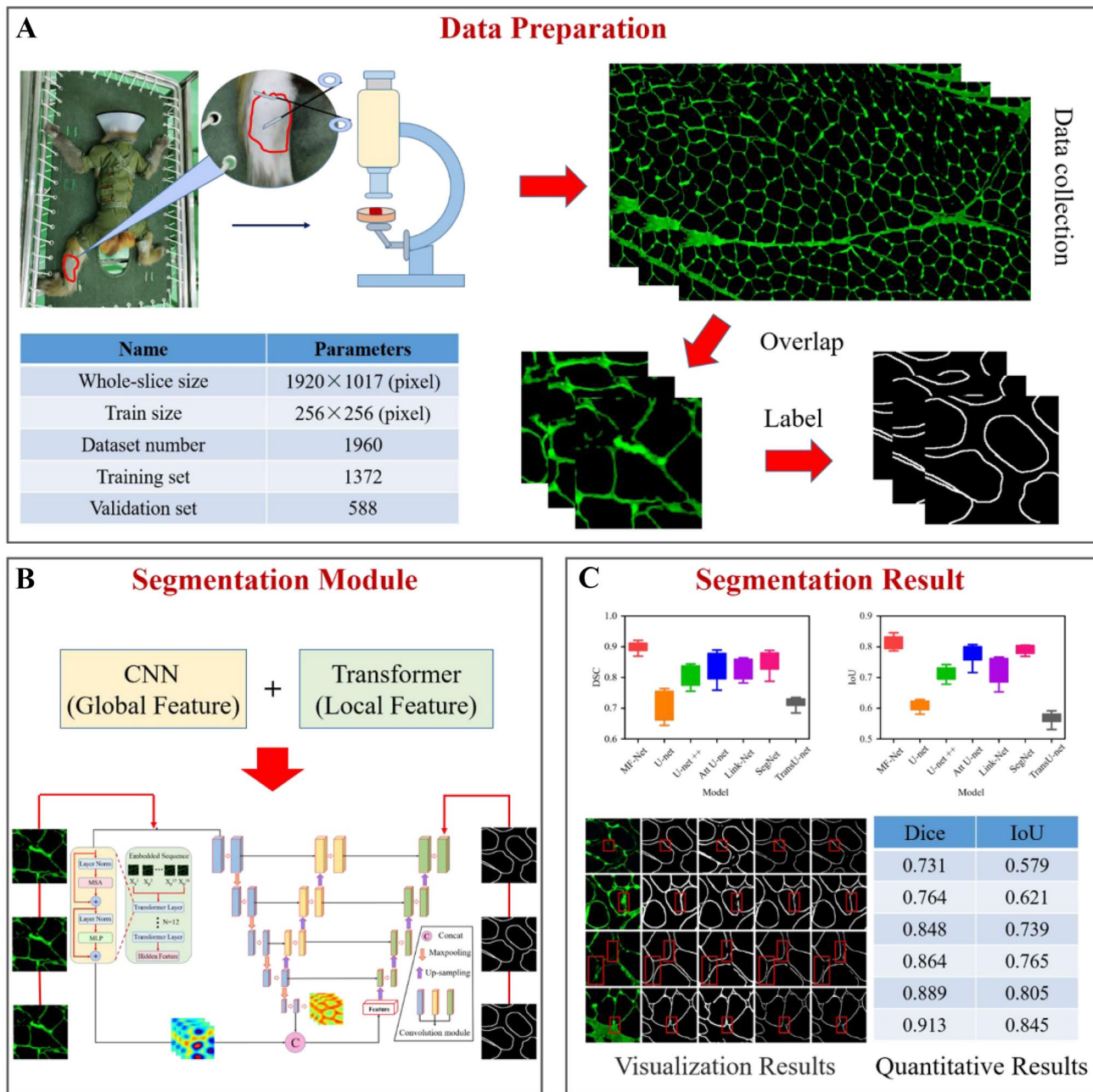


Fig. 2 Overall workflow of the muscle fiber segmentation framework

of the muscle fiber region is highlighted, and the morphological information of each muscle fiber is learned. The features of background noise are suppressed, and the complete muscle fiber boundary is segmented effectively.

- We propose an LLFD module, which uses the high-level features of CNN to conduct up-sampling to fuse the features specially extracted by the parallel dual stream encoder of the corresponding layer to obtain rich global context information and make up for lost details. The network evaluated our method on the weightlessness model of three macaques and the IF image datasets of three normal macaques. The results showed that this method was superior to the most advanced method in muscle fiber segmentation.

## Related Work

In this section, we introduce two studies related to our work, including segmentation of IF images and the application of a transformer in segmentation.

### IF Image Segmentation

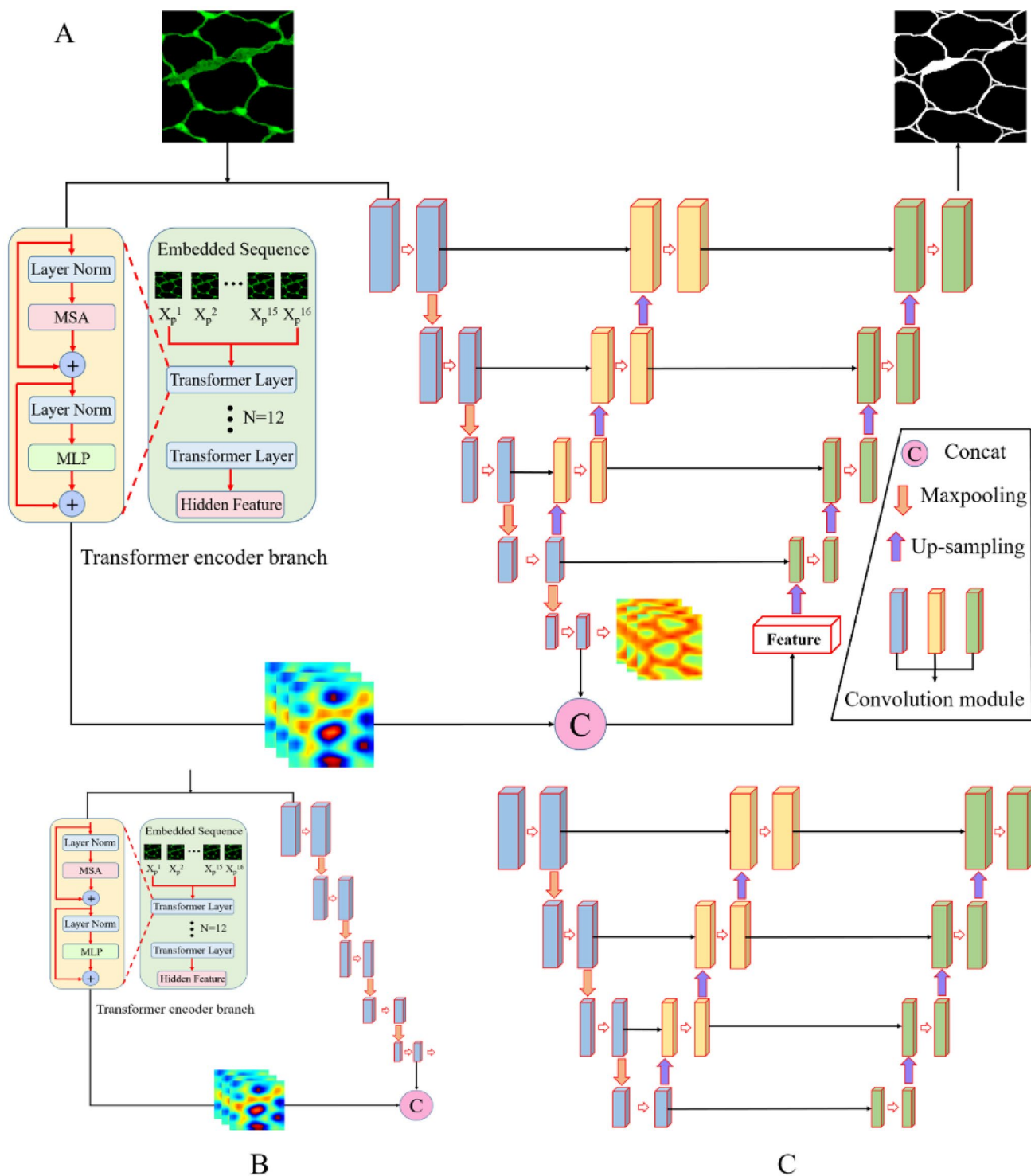
Accurately segmenting muscle fibers from IF images is still a challenge, because the contrast difference between the muscle fiber boundary area and background noise is small, the fluorescence intensity is low, and the muscle fiber boundary has low contrast and is fuzzy. Traditional methods try to use some prior knowledge or edge detection-based methods to solve these problems, such as region growth and watershed segmentation. For example, in order to solve the problem of various noises in IF images, Roy and Maji proposed a segmentation algorithm incorporating neighborhood information into a rough fuzzy clustering algorithm for the segmentation task of a human epithelial type-2 (HEp-2) cell [22]. Tonti et al. proposed a new image processing technology that automatically makes the marker selection pipeline conform to the special features of the input image to process different fluorescence intensities and coloring modes without any prior knowledge [23]. The above methods require constant adjustment of hyperparameters to obtain better segmentation results. However, because of the complex background noise and the large amount of data, the traditional segmentation method cannot show the anticipated segmentation effect.

With the great breakthrough of deep CNN in the field of image segmentation, it is widely used in cell segmentation. Through local connection and weight sharing, CNN reduces the number of weights, making the network easier to optimize, and reduces the complexity of the model [24, 25]. At first, researchers used CNN as a feature extraction

module and combined it with traditional segmentation methods to complete segmentation tasks. For example, for the complexity of histopathological images, Xing et al. first used CNN models to generate probability maps and then used region-merging methods to initialize shapes to complete kernel segmentation [26]. Jiang et al. proposed a two-stage segmentation method based on CNN to segment plant cells. First, the watershed algorithm is used to preliminarily segment plant cells, and then, the CNN-based discrimination model is used to retain segmented candidate cells with clear boundaries in the original noise cell image [27]. With the continuous improvement of CNN in image segmentation, researchers gradually paid attention to the fusion of low-level features and high-level features of IF cell images to avoid losing details in training. Kromp et al. used a variety of DL architectures to segment complex fluorescent nuclear images. Compared with traditional algorithms, the DL method achieved better results [28]. Liu et al. designed a hierarchical feature fusion attention network (HFANet). Through the hierarchical feature fusion attention (HFA) module, shallow texture features are used to supplement deep semantic features to maximize its feature extraction capability and information fusion efficiency [29]. The above network model is extremely complex and has many parameters, rarely considering the impact of the characteristic response of unrelated regions on the segmentation results. In order to segment muscle fibers accurately and improve the efficiency of the model, we propose an MF-Net segmentation network, which can extract muscle fiber features from IF images.

### Application of a Transformer in Medical Image Segmentation

Since the proposal of a transformer model to solve the sequence to sequence problem in 2017 [18], a transformer has brought rapid progress to a variety of natural language tasks. Inspired by natural language processing (NLP), the birth of vision transformer (ViT) has provided a new idea in the field of computer vision. This method is used to divide images into multiple patches for input and map them into a linear embedded sequence and encode with the encoder [30]. In recent years, a transformer has been widely used in the field of medical image segmentation and has made breakthrough progress [31–33]. Chen et al. applied a transformer and U-net to medical image segmentation for the first time. On the one hand, they solved the inherent limitations of the convolution operation, and on the other hand, they solved the problem that the positioning ability may be limited due to insufficient low-level details [19]. Wu et al. constructed a new feature adaptive transformer network (FAT-Net) to segment skin lesions, in which transformer branches can effectively capture remote dependency and



**Fig. 3** An overview of the proposed MF-Net framework. **A** Segmentation network architecture of MF-Net. **B** The modules of the backbone network. **C** The architecture of the LLFD module

global context information [34]. In order to enable a transformer to better encode the feature mapping of CNN image blocks at different stages into input attention sequences, Ma et al. used multi-scale modules to obtain richer CT image

semantic information [35]. In IF image segmentation, it is difficult to obtain samples or train the model due to the small number of samples. Valanarasu et al. proposed a medical transformer segmentation model to solve this problem.

This method shows high performance on the GLAS dataset [36]. Ji et al. proposed the multi-composite transformer (MCTrans), which integrates rich feature learning and semantic structure into a unified framework to solve the problem of a lack of learning of the cross-scale dependency of different pixels. This method has superior performance in cell segmentation [37].

### Methods

In this section, we describe the proposed MF-Net for macaque muscle fiber segmentation. First, we describe the overall structure of the proposed network. Then, the key components of the network, namely, the transformer encoder branch and LLFD module, are described in detail. By integrating CNN and a transformer in the dual encoder branch, we can not only capture rich local feature information and important global feature information but also suppress the impact of background noise on segmentation performance. In addition, we also developed an LLFD module and embedded it in the main network framework to obtain rich global context information. The proposed segmentation framework is shown in Fig. 3A.

### The Network Architecture

The proposed MF-Net is an end-to-end muscle fiber segmentation framework, as shown in Fig. 3. MF-Net consists of three key components: backbone network, transformer encoder branch, and LLFD module. Specifically, (1) backbone network is developed based on a U-net network. A dual encoder branch structure is adopted on the backbone network, which is a five-layer CNN encoder branch and a transformer encoder branch, respectively (as shown in

Fig. 3B). First,  $256 \times 256$  IF muscle fiber images are sent to the CNN branch and transformer encoder branch, respectively, and then, the image features of muscle fibers are extracted layer by layer in the CNN branch. The transformer encoder branch is used to effectively capture the global multi-scale information of muscle fibers. (2) The feature map after the fusion of CNN and transformer is sampled layer by layer and fused with the LLFD module to accurately segment the complete muscle fiber boundary. (3) The decoder reconstructs the advanced features to the same size as the input image to obtain the final segmentation result. In the following section, we will discuss the proposed transformer encoder branch and LLFD module in detail.

### Transformer Encoder Branch and LLFD Module

#### Transformer Encoder Branch

It is very important to accurately distinguish the difference between the target pixel and the background pixel in the segmentation task. For example, there is low contrast between the fibers and background noise in IF muscle fiber images, irregular shape changes of atrophied muscle fibers, and fuzzy muscle fiber boundaries. Therefore, we implemented a transformer encoder branch to capture the local characteristics and global context information of muscle fibers at the same time. Specifically, we will first input image  $Z$  with a size  $H \times W$ , which is reconstructed into one or more patches  $\{X_p^i \in R^{P^2 \times C} | i = 1, 2, 3, \dots, N\}$  for input, where the size of each patch is  $P \times P$ . The number of patches in a sequence is  $N = \frac{HW}{P^2}$ . Vectorized patch  $X_p$  is mapped to a  $D$ -dimensional embedding space by a trainable linear projection. Then, when encoding the spatial information of a patch, in order to preserve its location information, the learned specific location is embedded into the patch. The formula is as follows:

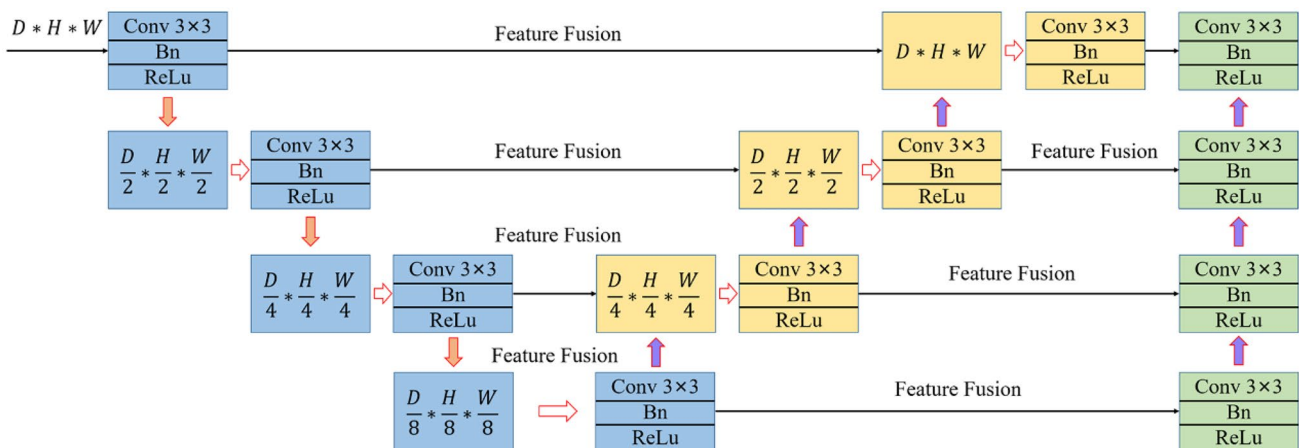


Fig. 4 The structure of the LLFD module. LLFD, low-level feature decoder

$$y = [X_p^1 E, \dots, X_p^N E] + E_{\text{pos}} \tag{1}$$

where  $E$  represents the location information of each patch and  $E_{\text{pos}}$  represents the location embedding information.

The transformer encoder is composed of a multi-head self-attention (MSA) module and a multilayer perceptron (MLP) module. MSA allows the model to learn relevant information in different representation subspaces, and MLP is composed of two linear layers. The whole process can be expressed as follows:

$$K'_l = \text{MSA}(\text{LN}(K_{l-1})) + K_{l-1} \quad l = 1, 2, 3 \dots L \tag{2}$$

$$K_l = \text{MLP}(\text{LN}(K'_l)) + K'_l \quad l = 1, 2, 3, \dots L \tag{3}$$

where LN represents the normalization algorithm and  $K_l$  represents the encoded image of each patch. In order to make the feature space conform to the input of the decoder path, the encoder feature size is changed from  $\frac{H \times W}{p^2}$  refactor to  $\frac{H}{p} \times \frac{W}{p}$ .

### LLFD Module

In traditional U-net, the original U-net is transmitted to the corresponding decoder layer through the skip connection via the multilayer feature map generated by the encoder, so that the decoder can obtain more high-resolution information when sampling and then recover the details of the original image more perfectly. However, due to the fuzzy boundary of some muscle fibers, a large amount of detailed information will still be lost in down-sampling, which affects the segmentation accuracy and integrity of the muscle fiber boundary. Therefore, we propose an LLFD module. Each layer consists of a kernel size of  $3 \times 3$ , batch normalization (BN), and the rectified linear unit (ReLU). In order to recover the detailed spatial information lost in the down-sampling phase of the encoder, we conduct the up-sampling operation on the fourth layer of the encoder, skip connect with each layer of the encoder to obtain a large amount of global context information, and then fuse each layer of LLFD with the corresponding feature layer extracted by a dual encoder to reduce the lost feature information, so as to better identify the feature information with fuzzy muscle fiber boundaries. The detailed structure of our LLFD is shown in Fig. 4.

### Implementation Details

The network is implemented on the Pytorch 1.8.0 platform. The experimental environment is GeForce GTX 3090 and Intel Core i9-10900 k. The optimizer uses Adam. The initial learning rate is set to 0.0001, and batch size and epoch are set to 25 and 100, respectively. In addition, the loss function is an

operational function used to measure the difference between the predicted value and the real value of the model. In this study, because the IF images are cropped by using the overlap method, the background pixels in some images will be larger than the positive pixels (muscle fibers). This problem of sample imbalance may make the training process difficult, resulting in over prediction of non-fiber regions. To solve this problem, we used a combined loss function composed of cross-entropy loss and Dice loss. Dice loss makes the model more inclined to mine the foreground region by finding the intersection and union between the segmentation results and ground-truth, reducing the impact of most negative pixels. The formula of the combined loss function is as follows:

$$\text{Loss} = \text{Loss}_{\text{cross}} + \text{Loss}_{\text{dice}} \tag{4}$$

$$\text{Loss}_{\text{cross}} = -y_{\text{true}} \log(y_{\text{pred}}) - (1 - y_{\text{true}}) \log(1 - y_{\text{pred}}) \tag{5}$$

$$\text{Loss}_{\text{dice}} = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \tag{6}$$

where Loss is the combined loss function,  $\text{Loss}_{\text{cross}}$  means cross-entropy loss,  $y_{\text{true}}$  is the target value, and  $y_{\text{pred}}$  is the predicted value.  $\text{Loss}_{\text{dice}}$  stands for Dice loss,  $X$  is the real image, and  $Y$  is the segmentation result.

## Experiments

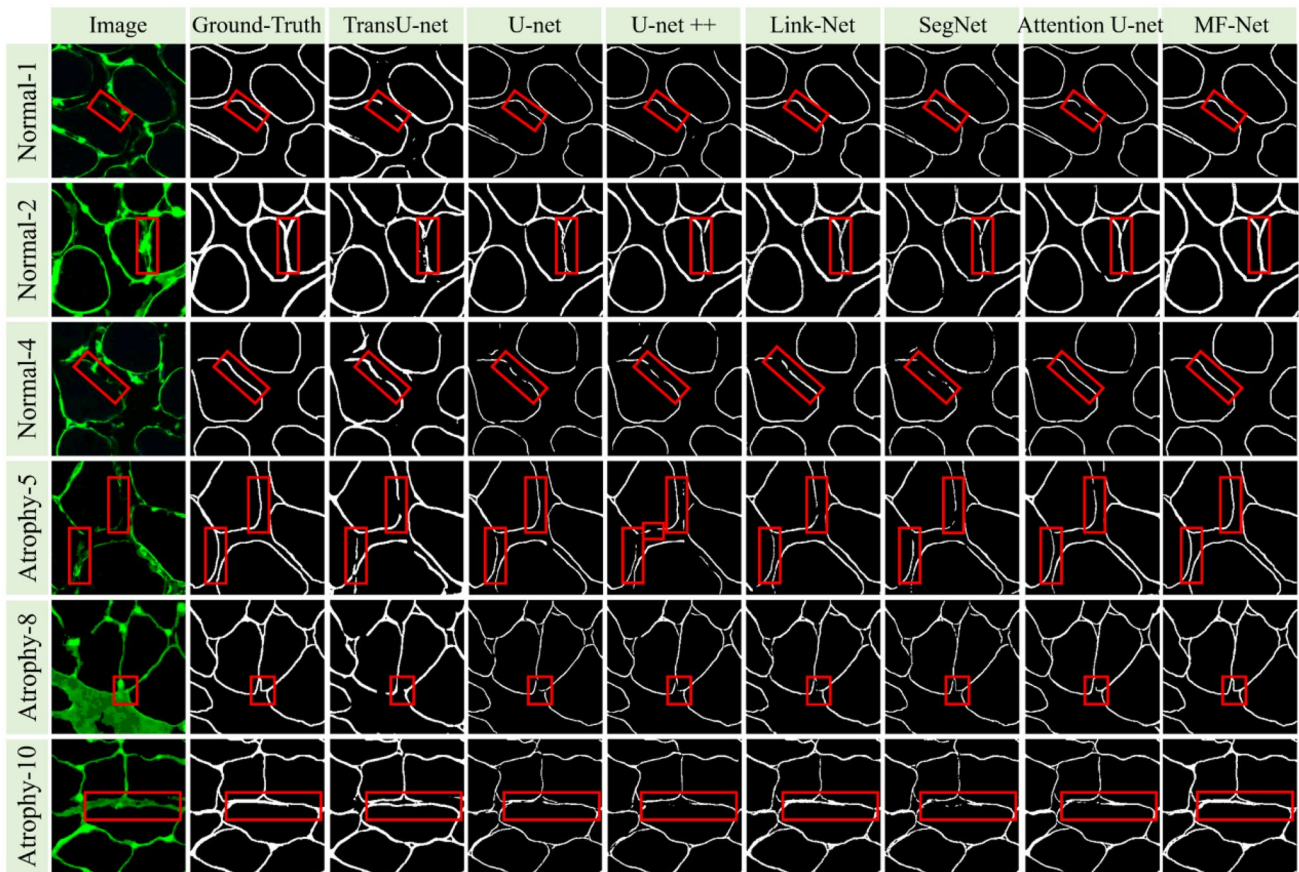
### Modeling and Dataset

#### Modeling

In this study, we cooperated with the China Astronaut Research and Training Center and took 6 normal macaques as experimental subjects. The average age was 6 years old, and the sex was male. The six macaques were divided into a normal ground-based group ( $n=3$ ) and BR group ( $n=3$ ). The BR group was kept in bed for 42 days at minus 10 degrees to simulate the weightlessness environment in space, so as to establish the model of weightlessness muscle atrophy. On the 42nd day, the shank muscle tissues of the normal group and the BR group were taken, respectively. All animal experiments were approved by the Institutional Animal Care and Use Committee of China Astronaut Research and Training Center (ACC-IACUC-2019-002).

#### Macaque Dataset

In this paper, the IF images of 6 macaques' shanks were used as the experimental data. Fix the frozen sections and add LAMININ2 (Abcam, UK) with a dilution ratio of 1:50.

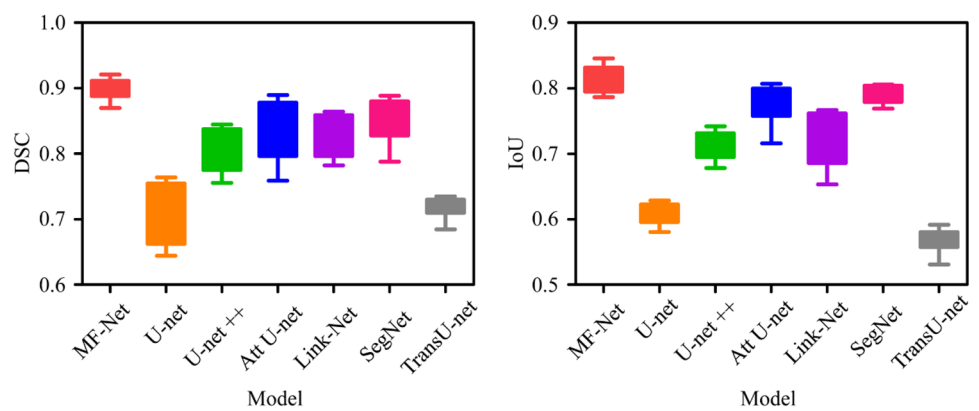


**Fig. 5** Visual comparison of segmentation results with different state-of-the-art segmentation methods. The red outline represents other methods, our method, and ground-truth, respectively

LAMININ2 antibody specifically binds to laminin to make the intercellular matrix appear green. Place the slices under a fluorescence microscope (NIKON ECLIPSE C1, Sony, Japan) to observe and collect images and obtain a total of 6 muscle tissue IF images, where the size of each IF image was  $1920 \times 1017$ . In order to adapt to the training of the network model, 6 IF images were cropped into 1960 small images with

a size of  $256 \times 256$ . The labeling of muscle fibers in all 1960 IF images was completed by two experienced pathologists (each with more than 2 years of working experience) on the MATLAB Image Labeler App. In this study, 1960 IF images were divided into 7:3 datasets, 1372 of which were used for training and 588 for validation.

**Fig. 6** DSC box diagram and IoU box diagram of MedT, TransU-net, U-net, ResU-Net, and MF-Net. DSC, Dice similarity coefficient; IoU, intersection-over-union





**Table 1** Performance comparison with the most state-of-the-art segmentation methods on the IF image muscle fiber datasets

Network	DSC	IoU	Precision	Recall	F1-score
Deeplabv3+ [43]	0.601 ± 0.0037	0.435 ± 0.0029	0.565 ± 0.0026	0.649 ± 0.0019	0.604 ± 0.0022
TransU-net [19]	0.731 ± 0.0012	0.579 ± 0.0015	0.737 ± 0.0033	0.729 ± 0.0011	0.733 ± 0.0017
U-net [15]	0.764 ± 0.0002	0.621 ± 0.0003	0.751 ± 0.0008	0.782 ± 0.0003	0.766 ± 0.0004
U-net++ [42]	0.848 ± 0.0031	0.739 ± 0.0046	0.856 ± 0.0151	0.836 ± 0.0202	0.846 ± 0.0173
Link-Net [40]	0.864 ± 0.0002	0.765 ± 0.0003	0.854 ± 0.0009	0.875 ± 0.0005	0.864 ± 0.0006
UNeXt [44]	0.883 ± 0.0011	0.794 ± 0.0017	0.896 ± 0.0025	0.869 ± 0.0021	0.882 ± 0.0023
SegNet [41]	0.889 ± 0.0026	0.805 ± 0.0016	0.888 ± 0.0031	0.891 ± 0.0025	0.889 ± 0.0028
Attention U-Net [39]	0.895 ± 0.0006	0.806 ± 0.0002	0.892 ± 0.0005	0.888 ± 0.0011	0.890 ± 0.0007
MF-Net/patches = 16	<b>0.913 ± 0.0001</b>	<b>0.845 ± 0.0002</b>	<b>0.915 ± 0.0005</b>	<b>0.909 ± 0.0004</b>	<b>0.912 ± 0.0004</b>
MF-Net/patches = 32	0.902 ± 0.0002	0.829 ± 0.0003	0.909 ± 0.0004	0.897 ± 0.0006	0.903 ± 0.0005

Values in bold indicate the best experimental results

## BBBC007 Datasets

This paper further verifies the segmentation performance of MF-Net on the BBBC007 datasets [38]. The BBBC007 dataset consists of the staining of *Drosophila melanogaster* Kc167 cells with DNA (which marks the nucleus) and actin (cytoskeletal proteins, which show the cell body), containing 32 images and ground-truth. Due to the different image sizes in the dataset, data cropping was performed and the resolution of each image was adjusted to 256 × 256. The training set and the test set are allocated according to 8:2. The training set is 410 and test set is 102. To increase the size of the dataset, the dataset was amplified (mirrored, Gaussian blur, rotated).

## Evaluation Metrics

In order to comprehensively evaluate the segmentation performance of MF-Net proposed in this paper, we use the Dice similarity coefficient (DSC), intersection-over-union (IoU), precision, recall, and F1-score to quantitatively analyze the experimental results. DSC and IoU are the most commonly used evaluation indicators to verify the segmentation algorithm. The five formulas are as follows:

$$\text{DSC} = \frac{2 * (A \cap B)}{|A| + |B|} \quad (7)$$

$$\text{IoU} = \frac{A \cap B}{A \cup B} \quad (8)$$

$$\text{Precision} = \frac{A \cap B}{A} \quad (9)$$

$$\text{Recall} = \frac{A \cap B}{B} \quad (10)$$

$$F1 - \text{score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

where  $A$  represents the ground-truth and  $B$  represents the result of the segmentation.

## Comparisons on Other State-of-the-Art Methods

### Segmentation of Muscle Fibers

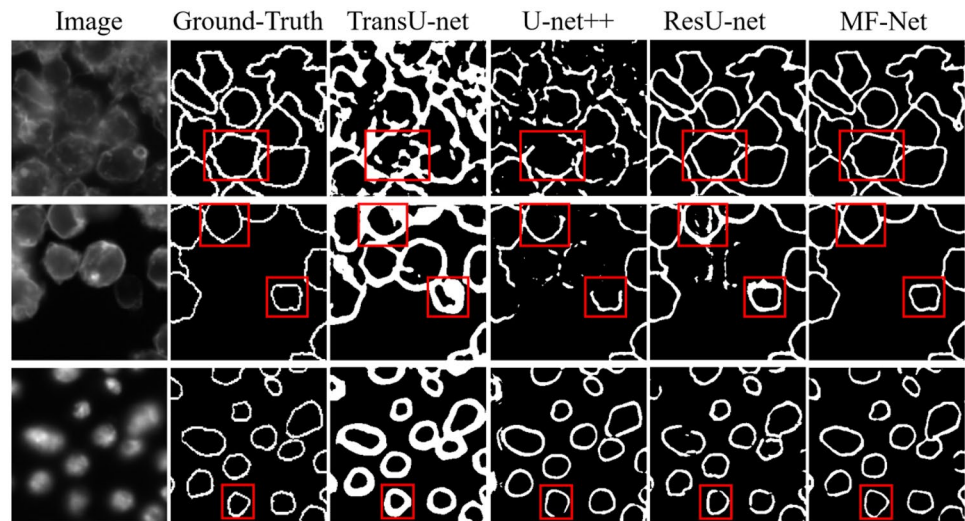
In order to compare the segmentation performance of our proposed MF-Net with other most advanced network models, we first trained nine segmentation models on the training datasets of 1372 IF images and manually adjusted all super parameters to obtain the best performance, including Deeplabv3+ [43], TransU-net [19], U-net [15], U-net++ [42], UNeXt [44], Link-Net [40], SegNet [41], Attention U-Net [39], and MF-Net. Then, we verified the validation datasets and saved the best performance model as the final segmentation model. Figure 5 shows the visualization results of muscle fiber segmentation of seven different methods. We can see that the performance of TransU-net and U-net is relatively low, which is because they cannot recognize the

**Table 2** Performance comparison with the most state-of-the-art segmentation methods on the BBBC007 datasets

Model	DSC	F1-score
TransU-net [19]	0.5373	0.5672
U-net++ [42]	0.6625	0.6667
UNeXt [44]	0.6871	0.7264
ResU-net [45]	0.7071	0.7596
MF-Net	<b>0.7316</b>	<b>0.7953</b>

Values in bold indicate the best experimental results

**Fig. 7** Visual comparison of segmentation results with different state-of-the-art segmentation methods on the BBBC007 datasets



feature information with low contrast at the boundary. The performance of U-net++ is better than that of TransU-net and U-net, but there are still some limitations in handling details. In contrast, the method proposed in this paper is closer to ground-truth, and its advantages can be seen from the red label area. It can be seen that MF-Net has identified a more complete boundary of muscle fibers. Especially for the IF image after bed rest (BR), MF-Net is more sensitive and accurate to the segmentation of irregular changes in the shape of its muscle fiber. In order to better evaluate the performance of the proposed method, a box chart of DSC and IoU index changes in the validation set is drawn as shown in Fig. 6 and quantitative evaluation results are provided as shown in Table 1. Table 1 quantifies the segmentation performance of muscle fibers in IF images by different methods. From the five indicators of DSC, IoU, precision, recall, and  $F1$ -score, we can see that our proposed MF-Net achieves the best performance. Compared with U-net, DSC and IoU are increased by 14.9% and 22.4%, respectively. Therefore, the method proposed in this paper successfully extracts the boundary of muscle fibers from IF images with good integrity.

**Table 3** Comparison of quantitative results of ablation experiments of the transformer module and LLFD module on the IF image macaque datasets

Network	DSC	IoU	Recall
MF-Net	<b><math>0.913 \pm 0.0001</math></b>	<b><math>0.845 \pm 0.0002</math></b>	<b><math>0.909 \pm 0.0004</math></b>
MF-Net/no transformer	$0.900 \pm 0.0003$	$0.823 \pm 0.0002$	$0.895 \pm 0.0012$
MF-Net/no LLFD	$0.896 \pm 0.0006$	$0.817 \pm 0.0004$	$0.896 \pm 0.0028$

Values in bold indicate the best experimental results

### Segmentation of BBBC007 Datasets

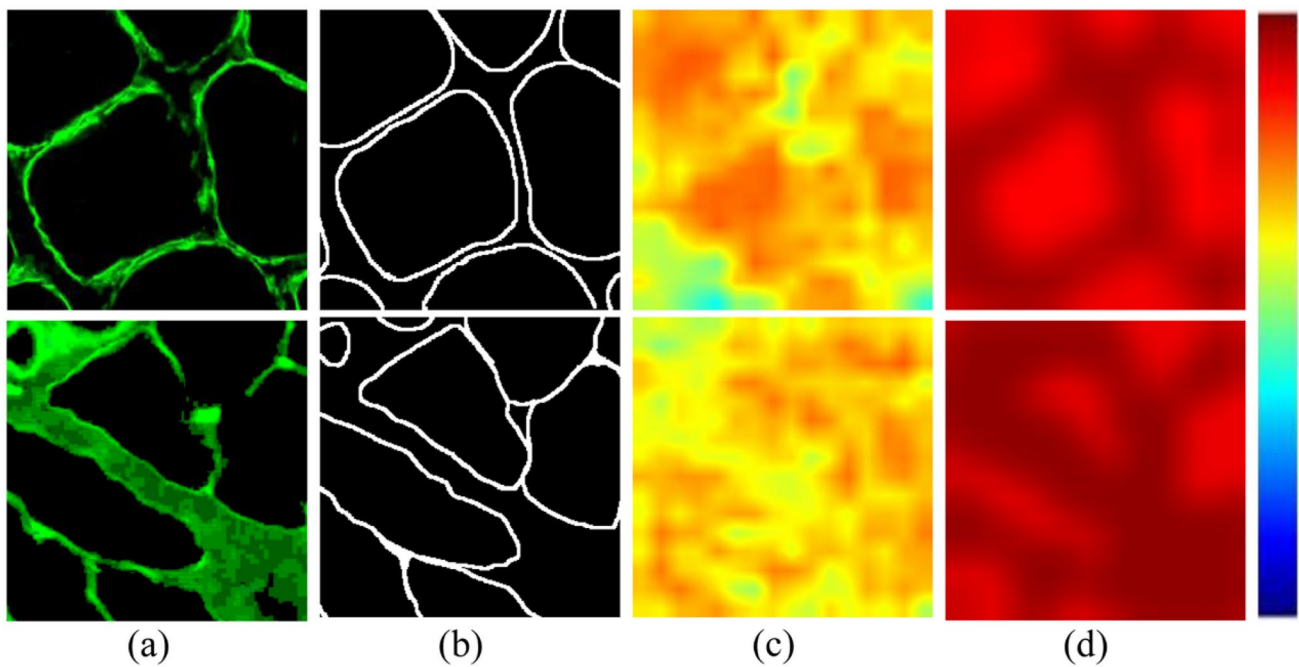
In order to further verify the segmentation ability of the proposed model, MF-Net is tested on the BBBC007 datasets. The quantitative results are shown in Table 2. We can see that MF-Net outperforms other most advanced segmentation models with 0.7316 and 0.7953 in DSC and  $F1$ -score, respectively. Because there are few BBBC007 datasets and the previous segmentation models are difficult to train on small datasets and cannot handle the details well, the proposed method can overcome this problem. In addition, we visualize the segmentation results of ResU-Net, UNeXt, U-net++, and our proposed method MF-Net in Fig. 7. As can be seen from the red box, the prediction of MF-Net captures the remote correlation well.

### Ablation Experiments

As shown in Table 3, in order to verify the performance of the transformer encoder branch and LLFD module, the ablation experiment was conducted in this study, and the effectiveness of the transformer module and LLFD module was evaluated using the U-net network as the baseline model.

#### Ablation Experiments for a Transformer

In order to discuss the performance of the transformer encoder branch, we have made an experiment to delete the transformer encoder branch. It can be seen from Tables 1 and 3 that the result of not using the transformer encoder branch is still higher than that of U-net, but lower than that of using the transformer encoder branch on DSC and IoU. In order to directly observe which feature areas the CNN



**Fig. 8** Visual comparison of different attention maps extracted by the CNN encoder and transformer encoder. **a** Image, **b** ground-truth, **c** attention map of the last layer in the CNN encoder, and **d** attention map of the last layer in the transformer encoder

encoder and transformer encoder pay more attention to, we conduct visual attention mapping from the output of the last layer of the CNN encoder and the output of the transformer encoder, as shown in Fig. 8. This shows that the transformer has certain advantages in feature extraction. It can combine global information to more accurately extract the boundary region of muscle fibers. In addition, we also discussed the impact of the size of different sequence patches on model performance. Table 1 shows that when the number of patches is 16, model performance is the best, at 1.1% higher than when the number of patches is 32.

### Ablation Experiments for LLFD

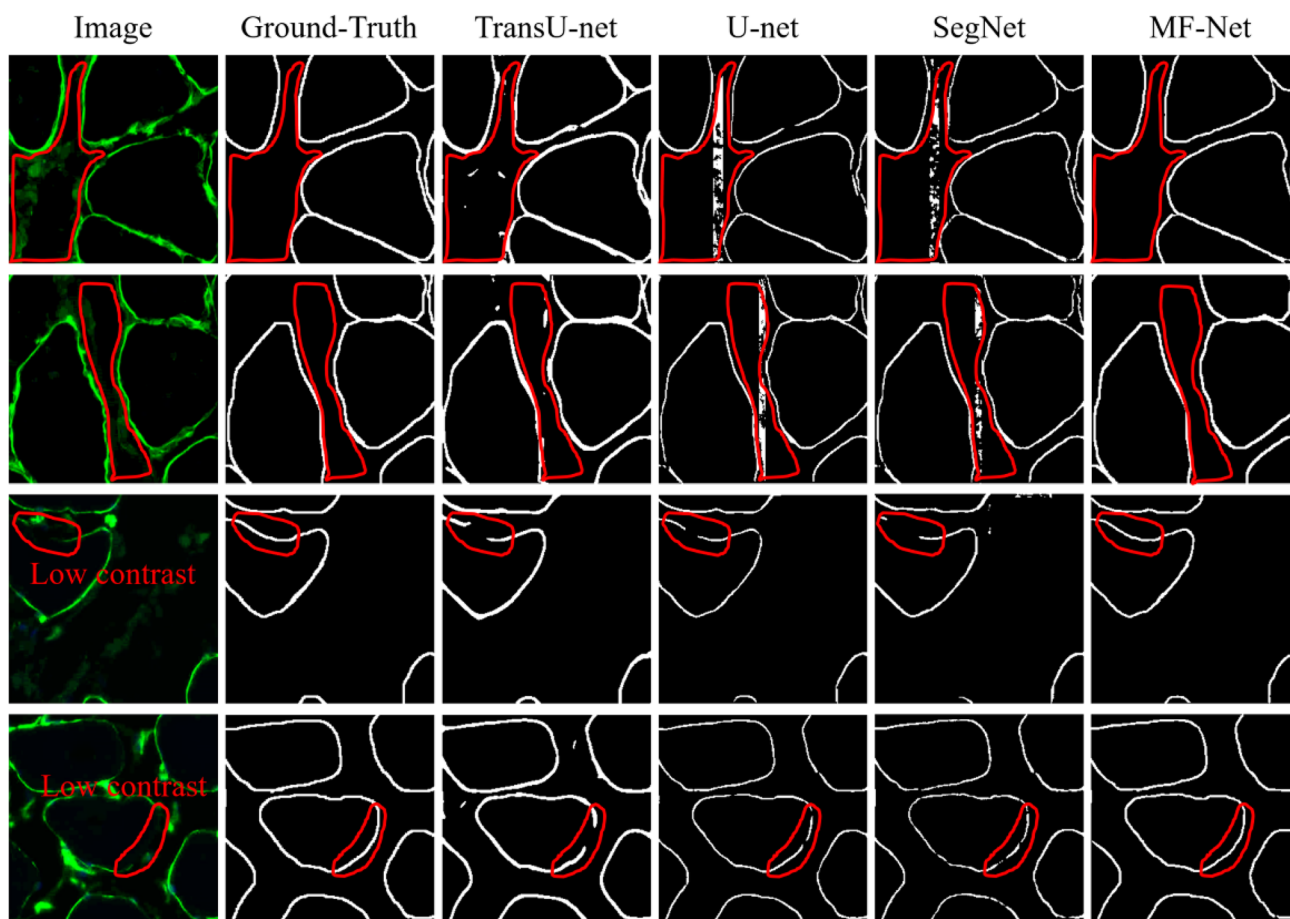
As shown in Table 3, we compared the quantitative comparison results with and without feature LLFD modules. It can be seen that the feature LLFD module does improve segmentation performance. DSC and IoU are improved from 0.896 and 0.817 to 0.913 and 0.845, respectively. This shows that the LLFD module can effectively compensate for the lost feature information in the encoder, thus improving the integrity of muscle fiber boundary segmentation. In addition, in order to verify which layer feature up-sampling process has an important impact on segmentation accuracy, we implement up-sampling operations from the fifth layer to the second layer of the encoder. The quantitative results are shown in

Table 3. DSC and IoU values obtained from up-sampling operations from the fourth layer of the encoder are the best.

Combined with the above analysis of the two modules in the MF-Net model, we carried out visualization experiments for the challenging problems in IF muscle fiber images, as shown in Fig. 9. Compared with TransU-net, U-net, and SegNet, MF-Net more effectively solved the problems of a small contrast difference between the muscle fiber boundary area and background noise and low contrast of the muscle fiber boundary.

### Effectiveness of Dice Loss

Because the pixels of positive samples in some IF images after overlap clipping are smaller than those in the background area, it is more challenging to develop accurate segmentation algorithms. To overcome this problem, we introduce Dice loss and combine it with cross-entropy loss to improve the performance of the training model. In order to verify its effectiveness, this study uses our proposed combined loss and traditional cross-entropy to carry out comparative experiments. Table 5 reports the DSC and IoU values for combined losses and cross-entropy. Compared with cross-entropy loss, our proposed combined loss can significantly improve the performance of the training model, and DSC and IoU can improve by 1.4% and 2.2%, respectively.



**Fig. 9** Visual comparison of challenges in processing IF images of four segmentation models TransU-net, U-net, SegNet, and MF-Net. The first two lines of red boxes indicate that the contrast difference

between the muscle fiber boundary area and background noise is small. The last two lines of red boxes indicate the problem of low contrast of the muscle fiber boundary

## Discussion

Weightless muscle atrophy caused by long-term microgravity environment will seriously impact the health, safety, and working ability of astronauts [46]. Therefore, accurate diagnosis of the degree of muscle atrophy is the key to timely implementation of muscle atrophy. As the gold standard for the diagnosis of weightless muscle atrophy, skeletal muscle slices play a very important role in the evaluation of muscle atrophy. For this reason, high requirements are put forward for the automatic segmentation of muscle fibers in high-resolution IF images. Although a lot of work has been done to study the problem of fiber segmentation, due to the small contrast difference between the muscle fiber boundary area and background noise, the low and blurry contrast of muscle fiber boundaries caused by boundary interruption or artifacts during muscle tissue staining or sample preparation in IF muscle fiber images, the existing research methods have some limitations. Therefore, establishing a new and effective fiber segmentation method is crucial to accurately

evaluate muscle atrophy. In this work, we propose a fully automatic muscle fiber segmentation using the MF-Net method. Compared with U-net, traditional U-net makes up for the loss of spatial information in the training process, but in terms of details, the shortcomings of low-resolution information repetition and low-resolution edge information feature extraction still exist. The key of MF-Net is to focus on the characteristic response of the boundary area of the muscle fiber by combining the global context information. The dual encoder branch composed of CNN and transformer

**Table 4** Ablation studies of LLFD in MF-Net

Network	DSC	IoU	Precision	Recall
LLFD/layer=2	0.886	0.794	0.893	0.882
LLFD/layer=3	0.906	0.837	0.913	0.897
LLFD/layer=4 (MF-Net)	<b>0.913</b>	<b>0.845</b>	<b>0.915</b>	<b>0.909</b>
LLFD/layer=5	0.898	0.814	0.902	0.896

Values in bold indicate the best experimental results

**Table 5** Segmentation results of using Dice loss on the IF image muscle fiber datasets

Loss			Evaluation index			
Cross-entropy loss	Dice loss	Focal loss	DSC	IoU	Precision	Recall
√	√		<b>0.913</b>	<b>0.845</b>	<b>0.915</b>	<b>0.909</b>
√			0.897	0.819	0.899	0.896
		√	0.892	0.811	0.906	0.881

Values in bold indicate the best experimental results

is used to effectively capture the local information and global context information and stimulate the characteristic response of the boundary area of the muscle fiber. At the same time, LLFD module is added to the network to compensate for the lost details of the muscle fiber boundary, so that the integrity and continuity of muscle fiber segmentation can be achieved. The segmentation of muscle fibers can be applied to clinical applications. We took macaques as the experimental objects and simulated the space weightlessness environment by using the method of negative 10 degrees BR. It was fully proved that the segmentation performance of the proposed model is superior to most current state-of-the-art methods on the normal and BR groups of macaque shank IF image datasets [11, 39–44]. It can be seen from Table 1 that MF-Net improves DSC and IoU of automatic muscle fiber segmentation results, which are 91.3% and 84.5%, respectively. In addition, in order to further verify the segmentation performance of the proposed model, we verified it on the BBBC007 datasets. Table 2 shows that compared with the UNeXt model proposed by MICCAI in 2021, MF-Net has improved the DSC and F1-score by 4.45% and 6.89%. In general, the experimental results on the two datasets further prove that our proposed method has a good segmentation effect and good generalization performance.

However, we found in the experiment that the problems such as the small contrast difference between the muscle fiber boundary area and background noise in the IF image and the low contrast and blurring of the muscle fiber boundary will bring bottlenecks to the training of the DL model. In order to better solve these problems, we added the transformer encoder branch and LLFD module in MF-Net. In medical image segmentation, the structure of the decoder and encoder has been widely used, but most network structures only have one encoder, so local features and global context information features cannot be captured at the same time. In the case of large-scale distribution images or high-resolution images, due to the irregular changes in the shape of muscle fibers and low contrast between muscle fibers and background noise, this has certain limitations extracting local features to global long-range dependency at the same time. However, the existing methods still cannot accurately distinguish the differences between the target and background. In view of this, we designed a dual encoder branch,

including a transformer branch and CNN branch, to capture both local features and global context information. In the IF image, some images will have the characteristic response of background noise similar to the characteristic response of the target area. If only the characteristics of the local area are considered, the target will be wrongly identified, which will affect the segmentation results. Therefore, the advantage of using dual encoders is that when recognizing the features of local areas, the self-attention mechanism in the transformer branch can extract rich global context information, so that muscle fibers and background noise areas can be accurately identified. As shown in Table 3, DSC is used to verify the influence of the transformer encoder branch on MF-Net. Without adding the transformer encoder branch, segmentation performance of MF-Net decreases by 1.3%, which indicates that the segmentation result of adding the transformer encoder branch to MF-Net is closer to the ground-truth. Due to the low contrast and blurring of the muscle fiber boundary, a large amount of boundary details will be lost in the down-sampling process, resulting in incomplete muscle fiber segmentation. Therefore, we propose an LLFD module to compensate for the loss of muscle fiber boundary details in the encoder. Unlike traditional decoders, which perform splicing operations on each layer, we up-sample the fourth layer feature map of CNN and fuse it with the corresponding encoder layer and fuse each layer feature map from LLFD with the corresponding layer feature map in the decoder. As shown in Table 3, we verify its impact on MF-Net through DSC and IoU. The experimental results show that the introduction of the LLFD module enhances the segmentation effect. It is worth noting that in the LLFD module, it is a key issue which layer of the encoder is used for up-sampling to significantly improve the segmentation accuracy. With the increase of convolution layers, the size of the image will decrease, and the feature information of the muscle fiber edge will decrease. This means that much information of the image edge position is lost, leading to the decline of segmentation accuracy. Therefore, as shown in Table 4, using the fourth layer of the encoder for up-sampling has the best segmentation effect. In addition, we noticed that in some IF images, the background pixels are larger than the foreground pixels, which will lead to a large number of background pixels dominating the loss during

the training process of the model, thus greatly reducing the accuracy of semantic segmentation [47]. The use of cross-entropy loss function for this problem has certain limitations [48], because it treats the weight of each category as the same in the calculation process, so that the model will focus on the category with more samples in the training process, and Dice loss' proposal solves the problem of data imbalance [49], so we studied the performance of Dice loss on IF unbalanced datasets. We used the joint loss of cross-entropy loss combined with Dice loss to compare with a single cross-entropy loss and a single focal loss. The experimental results are shown in Table 5. Among the three loss functions, Dice loss can help MF-Net get better segmentation results. This advantage shows that Dice loss will be more inclined to mining foreground regions during training, which alleviates the problem of sample imbalance and improves segmentation performance of MF-Net.

In conclusion, MF-Net architecture can achieve good results in muscle fiber segmentation. However, because the establishment of the experimental animal model in this study is relatively difficult, resulting in a small amount of sampling data, this has a certain impact on the training of segmentation models. Therefore, collecting more macaque IF imaging data in future work will be more conducive to improving the segmentation model's performance and accuracy of muscle atrophy evaluation.

## Conclusion

In this paper, we proposed an MF-Net segmentation architecture for muscle fibers in IF images. Unlike the traditional encoder-decoder structure, we use the dual encoder branch combined with CNN and transformer to segment images, so that our method can capture local and global information at the same time. In addition, we also adopt a more effective LLFD module to compensate for the missing details of muscle fiber boundaries, so that we can segment more complete muscle fiber boundaries. In this paper, experiments were conducted on the datasets of 3 BR macaques and 3 normal macaques to evaluate our proposed muscle fiber segmentation method. Compared with the most advanced segmentation model, the effectiveness of MF-Net is proved. The MF-Net segmentation method will help pathologists improve the efficiency of identifying muscle fibers from pathological images in the future, providing a preliminary guarantee for the subsequent evaluation of muscle atrophy.

**Author Contribution** All authors contributed to the study conception and design. Conceptualization, methodology, software, validation, formal analysis, investigation, writing—original draft preparation, and writing—review and editing were performed by Getao Du. Data acquisition and H&E slice preparation were performed by Peng

Zhang, Xiangsheng Pang, and Bin Zeng. The establishment of animal model was performed by Guanghan Kan. Resources was performed by Jianzhong Guo. Writing—review and editing was performed by Xiaoping Chen. Conceptualization, resources, and supervision were performed by Jimin Liang. Conceptualization, resources, writing—review and editing, supervision, project administration, and funding acquisition were performed by Yonghua Zhan.

**Funding** This work was supported, in part, by the State Key Laboratory Grant of Space Medicine Fundamentals and Application (No. SMFA20A02) and the National Natural Science Foundation of China (32171173, U19B2030, 61976167).

**Data Availability** All data in this study can be obtained by contacting the corresponding author upon reasonable request.

## Declarations

**Ethics Approval** The animal and the experimental protocol used in the paper were reviewed and ethically approved by the Institutional Animal Care and Use Committee of China Astronaut Research and Training Center (ACC-IACUC-2019-002).

**Competing Interests** The authors declare no competing interests.

## References

1. Vandenburg H, Chromiak J, Shansky J, Del Tatto M, Lemaire J: Space travel directly induces skeletal muscle atrophy. *FASEB JOURNAL* 13: 1031-1038, 1999.
2. Dumitru A, Radu BM, Radu M, Cretoiu SM: Muscle changes during atrophy. *Muscle Atrophy* 1088: 73-92, 2018.
3. Ventadour S, Attaix D: Mechanisms of skeletal muscle atrophy. *CURRENT OPINION IN RHEUMATOLOGY* 18: 631-635, 2006.
4. Lau YS, Xu L, Gao YD, Han RZ: Automated muscle histopathology analysis using CellProfiler. *SKELETAL MUSCLE* 8: 32, 2018.
5. Zhang P, He J, Wang F, Gong J, Wang L, Wu Q, Li WJ, Liu HJ, Wang J, Zhang KS, Li M, Huang XS, Pu CQ, Li Y, Jiang FJ, Wang FD, Min JX, Chen XP: Hemojuvelin is a novel suppressor for duchenne muscular dystrophy and age-related muscle wasting. *JOURNAL OF CACHEXIA SARCOPENIA AND MUSCLE* 10: 557-573, 2019.
6. Mayeuf-Louchart A, Hardy D, Thorel Q, Roux P, Gueniot L, Briand D, Mazeraud A, Bougle A, Shorte SL, Staels B, Chretien F, Duez H, Danckaert A: MuscleJ: a high-content analysis method to study skeletal muscle with a new Fiji tool. *SKELETAL MUSCLE* 8: 25, 2018.
7. Waisman A, Norris AM, Costa ME, Kopinke D: Automatic and unbiased segmentation and quantification of myofibers in skeletal muscle. *SCIENTIFIC REPORTS* 11: 11793, 2021.
8. Laghi V, Ricci V, De Santa F, Torcinaro A: A User-Friendly Approach for Routine Histopathological and Morphometric Analysis of Skeletal Muscle Using CellProfiler Software. *DIAGNOSTICS* 12: 561, 2022.
9. Rahmati M, Rashno A: Automated image segmentation method to analyse skeletal muscle cross section in exercise-induced regenerating myofibers. *SCIENTIFIC REPORTS* 11: 21327, 2021.
10. Veta M, van Diest PJ, Kornegoor R, Huisman A, Viergeever MA, Pluim JPW: Automatic nuclei segmentation in H&E stained breast cancer histopathology images. *PLOS ONE* 8: e70221, 2013.

11. Ul Islam I, Ullah K, Afaq M, Iqbal J, Ali A: Towards the automatic segmentation of HEP-2 cell in indirect immunofluorescence images using an efficient filtering based approach. *MULTIMEDIA TOOLS AND APPLICATIONS* 79: 34325–34337, 2020.
12. Chen L, Bentley P, Mori K, Misawa K, Fujiwara M, Rueckert D: DRINet for medical image segmentation. *IEEE TRANSACTIONS ON MEDICAL IMAGING* 37: 2453–2462, 2018.
13. Valanarasu JM, Sindagi VA, Hacihaliloglu I, Patel VM: KiU-Net: overcome complete convolutional architectures for biomedical image and volumetric segmentation. *IEEE TRANSACTIONS ON MEDICAL IMAGING* 41: 965–976, 2022.
14. Long J, Shelhamer E, Darrell T: Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39: 640–651, 2015.
15. Ronneberger O, Fischer P, Brox T: U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234–241.
16. Oda H, Roth H R, Chiba K, Sokolic J, Kitasaka T, Oda M, Hinoki A, Uchida H, Schnabel J A, Mori K: BESNet: boundary-enhanced segmentation of cell in histopathological images[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2018: 228–236.
17. Yi JR, Wu PX, Jiang ML, Huang QY, Hoepfner DJ, Metaxas DN: Attentive neural cell instance segmentation. *MEDICAL IMAGE ANALYSIS* 55: 228–240, 2019.
18. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I: Attention is all you need[C]//Proceedings of the Advances in Neural Information Processing Systems. 2017: 5998–6008.
19. Chen JN, Lu Y, Yu Q, Luo X, Adeli E, Wang Y, Lu L, Yuille AL, Zhou Y: TransUNet: transformers make strong encoders for medical image segmentation. arXiv preprint: <https://arxiv.org/abs/2102.04306>, 2021
20. Lqbal S, Qureshi AN: A Heteromorphous Deep CNN Framework for Medical Image Segmentation Using Local Binary Pattern. *IEEE ACCESS* 10: 63466–63480, 2022.
21. Tian XY, Jin Y, Tang XL: Local-Global Transformer Neural Network for temporal action segmentation. *MULTIMEDIA SYSTEMS* 29: 615–626, 2022.
22. Roy S, Maji P: Rough-fuzzy segmentation of HEP-2 cell indirect immunofluorescence images. *INTERNATIONAL JOURNAL OF DATA MINING AND BIOINFORMATICS* 17: 311–340, 2017.
23. Tonti S, Di Cataldo S, Bottino A, Ficarra E: An automated approach to the segmentation of HEP-2 cells for the indirect immunofluorescence ANA test. *COMPUTERIZED MEDICAL IMAGING AND GRAPHICS* 40: 62–69, 2015.
24. Chen W, Wei HF, Peng ST, Sun JW, Qiao X, Liu BQ: HSN: hybrid segmentation network for small cell lung cancer segmentation. *IEEE Access* 7: 75591–75603, 2019.
25. Ramesh N, Tasdizen T: Cell segmentation using a similarity interface with a multi-task convolutional neural network. *IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS* 23: 1457–1468, 2019.
26. Xing FY, Xie YP, Yang L: An automatic learning-based framework for robust nucleus segmentation. *IEEE TRANSACTIONS ON MEDICAL IMAGING* 35: 550–566, 2016.
27. Jiang WB, Wu LH, Liu SH, Liu M: CNN-based two-stage cell segmentation improves plant cell tracking. *PATTERN RECOGNITION LETTERS* 128: 311–317, 2019.
28. Kromp F, Fischer L, Bozsaky E, Ambros IM, Dorr W, Beiske K, Ambros PF, Hanbury A, Taschner-Mandl S: Evaluation of deep learning architectures for complex immunofluorescence nuclear image segmentation. *IEEE TRANSACTIONS ON MEDICAL IMAGING* 40: 1934–1949, 2021.
29. Liu HR, Zhang P, Xie YL, Li XF, Bi DJ, Zou YR, Peng L, Li GS: HFANet: hierarchical feature fusion attention network for classification of glomerular immunofluorescence images. *NEURAL COMPUTING & APPLICATIONS* 34: 22565–22581, 2022.
30. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J, Houlsby N: An image is worth 16x16 words: transformers for image recognition at scale[C]//International Conference Learning Representations. 2021.
31. Karimi D, Dou HR, Gholipour A: Medical image segmentation using transformer networks. *IEEE Access* 10: 29322–29332, 2022.
32. Huang SQ, Li JN, Xiao YZ, Shen N, Xu TF: RTNet: relation transformer network for diabetic retinopathy multi-lesion segmentation. *IEEE TRANSACTIONS ON MEDICAL IMAGING* 41: 1596–1607, 2022.
33. Dhamija T, Gupta A, Gupta S, Anjum, Katarya R, Singh G: Semantic segmentation in medical images through transfused convolution and transformer networks. *APPLIED INTELLIGENCE* 53: 1132–1148, 2022.
34. Wu HS, Chen SH, Chen GL, Wang W, Lei BY, Wen ZK: FAT-Net: feature adaptive transformers for automated skin lesion segmentation. *MEDICAL IMAGE ANALYSIS* 76: 102327, 2022.
35. Ma MJ, Xia HY, Tan YM, Li HS, Song SX: HT-Net: hierarchical context-attention transformer network for medical CT image segmentation. *APPLIED INTELLIGENCE* 52: 10692–10705, 2022.
36. Valanarasu J, Oza P, Hacihaliloglu I, Patel VM: Medical transformer: gated axial-attention for medical image segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2021: 36–46.
37. Ji YF, Zhang RM, Wang HJ, Li Z, Wu LY, Zhang ST, Luo P: Multi-compound transformer for accurate biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2021: 326–336.
38. Lehmussola A, Ruusuvaari P, Selinummi J, Huttunen H, Yli-Harja O: Computational framework for simulating fluorescence microscope images with cell populations. *IEEE TRANSACTIONS ON MEDICAL IMAGING* 26: 1010–1016, 2007.
39. Oktay O, Schlemper J, Folgoc L L, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla NY, Kainz B, Glocker B, Rueckert D: Attention U-Net: learning where to look for the pancreas. arXiv preprint: <https://arxiv.org/abs/1804.03999>, 2018.
40. Chaurasia A, Culurciello E: LinkNet: exploiting the encoder representations for efficient semantic segmentation[C]//Proceedings of the IEEE Visual Communications and Image Processing. IEEE, 2017: 10–13.
41. Badrinarayanan V, Kendall A, Cipolla R: SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE* 39: 2481–2495, 2017.
42. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J: UNet++: redesigning skip connections to exploit multiscale features in image segmentation. *IEEE TRANSACTIONS ON MEDICAL IMAGING* 39: 1856–1867, 2019.
43. Chen LC, Zhu Y, Papandreou G, Schroff F, Adam H: Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the European conference on computer vision. 2018: 801–818.
44. Valanarasu JM, Patel VM: UNeXt: MLP-Based Rapid Medical Image Segmentation Network[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2022: 23–33.
45. Diakogiannis FI, Waldner F, Caccetta P, Wu C: ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING* 1642: 94–114, 2020.

46. Demontis GC, Germani MM, Caiani EG, Barravecchia I, Passino C, Angeloni D: Human pathophysiological adaptations to the space environment. *Frontiers in Physiology* 8: 547, 2017.
47. Zheng Z, Zhong YF, Wang JJ, Ma AL: Foreground-aware relation network for geospatial object segmentation in high spatial resolution remote sensing imagery[C]//Proceedings of the 2020 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2020: 4095–4104.
48. Zhang XF, Liu XM, Zhang B, Dong J, Zhang B, Zhao SJ, Li SX: Accurate segmentation for different types of lung nodules on CT images using improved U-Net convolutional network. *Medicine* 100: e27491, 2021.
49. Li XY, Sun XF, Meng YX, Liang JJ, Wu F, Li JW: Dice loss for data-imbalanced NLP tasks[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020: 465–476.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.