Segmentation and Measurement of Superalloy Microstructure Based on Improved Nonlocal Block

LIXIN ZHANG, HAOTIAN YAO, ZHENGGUANG XU, AND HAN SUN

School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

Corresponding author: Haotian Yao (yxyht183@163.com)

ABSTRACT The microstructure of superalloy materials has a decisive impact on its service performance. When preparing the material and photographing the microstructure, different depths of metallography perpendicular to the cut plane appear in the microstructure image. These metallographic features are major factors causing inaccurate segmentation. Aiming at the problems of traditional image processing methods, such as large noise and poor robustness of the edge extraction, the deep learning method is introduced. The receptive field of the traditional convolutional neural network method is too local to obtain remote dependencies, and dense feature extraction also brings problems such as excessive noise. To address the problem of non-tangential metallographic information in microstructure images that is not easily distinguishable. We consider how to combine and utilize the feature maps of the intermediate process as effectively as possible, and find that reusing affinity matrices from different angles and stacking them can improve the overall effect, and can solve the problem of inaccurate segmentation of metallographic images at different depths. We propose to improve and optimize the non-local attention module and further combine the module with the UNet network to form a new improved SNL-Unet image segmentation structure, which significantly raises the accuracy and efficiency of image segmentation, Additionally, we measure the characteristic parameters such as volume fraction, average thickness and the degree of rafting. The code for this paper will be available at github.com/ustbjdl1021/improved-snl-unet.

INDEX TERMS Characteristic parameter measuremen, deep learning, image segmentation, self-attention, superalloy.

I. INTRODUCTION

The microstructure is the key to determining the properties of superalloy materials. Microstructure analysis is mainly based on the shape, distribution, and other characteristics of different tissues in microscopic images. Manual analysis methods have problems such as the influence of subjective factors, excessive time consumption, high cost, limited local statistics, and low accuracy. Therefore, seeking more advanced microstructure image analysis methods for superalloys and improving the calculation speed and accuracy of microstructure features have become the key to the metallographic analysis.

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Jiang.

When measuring and calculating characterization parameters such as volume fractions, more accurate information on the metallic phases in the same cutting plane is required. However, when preparing materials and taking pictures of microstructures, due to the limitations of on-site process conditions and errors in manual slicing operations, different depths of metallography perpendicular to the tangent plane will be present in the microstructural image. These image information that is not in the same cutting plane are not in the scope of characterization parametric measurement, so they need to be accurately rejected during image segmentation.

Therefore, in order to employ the information in microstructure images more efficiently, it is necessary to improve the network's ability to utilize image features and cope with the problem of identifying depth information.

We propose an improved nonlocal bolck, which can better segment deep and shallow information in effect. Further, we combine the improved non-local attention module and UNet network to form an improved SNL-UNet network structure, which significantly improves the accuracy and efficiency of image segmentation.

Experimentally, the segmentation result of this method is better than that of other network structures, and the ability to recognize the depth information of the image is also significantly improved. Additionally, based on more accurate segmentation results, the microstructure characterization parameters such as volume fraction, rafting degree of γ' phase and thickness of γ' phase are measured.

A. RELATED WORK

Some studies [1]–[3] show that the distribution of the γ' phase have an important influence on the mechanical properties and hardness of superalloys. Recent years, more and more attention has been paid to the study of γ-phase [4], [5]. Hou *et al.* [6] find that the yield strength and deformation mechanism are temperature-dependent. Sun *et al.* [7] find that morphology and size of secondary γ' are sensitive to temperature and strain rate.

There are some feature representation methods that combine imaging techniques and algorithms. Payton *et al.* [8] study that backscatter electron imaging of specimens in which the γ' phase has been selectively etched yields images that can be more readily segmented with image processing algorithms than other imaging techniques. Tiley *et al.* [9] measure γ' precipitates in a nickel-based superalloy using energy-filtered transmission electron microscopy coupled with automated segmenting techniques.

Traditional image processing methods mainly focus on image texture, gray difference, and other single features for segmentation, such as Otsu [10] method threshold segmentation, Canny [11] edge detection operator, Sobel [12] edge operator, etc. These methods have problems in edge extraction of material microstructure, such as blurred edges, low precision, and many noises.

There are methods based on traditional algorithms that perform well in segmenting microstructure images. Cao *et al.* [13] develop a new Multichannel Edge-Weighted Centroidal Voronoi Tessellation (MCEWCVT) algorithm to automatically segment all the 3D grains from microscopic images of a super-alloy sample. Chuang *et al.* [14] study the combination of a region merging segmentation method called the stabilized inverse diffusion equation (SIDE), and a stochastic segmentation method, the expectation-maximization/maximization of the posterior marginals (EM/MPM) algorithm. Ewees *et al.* [15] presents a hybrid meta-heuristic approach for multi-level thresholding image segmentation by integrating both the artificial bee colony (ABC) algorithm and the sine-cosine algorithm (SCA). Alwerfali *et al.* [16] develop an alternative MTI segmentation method by using a modified version of the salp swarm algorithm (SSA).

In recent years, the continuous improvement of the computing power of computer hardware equipment has provided the possibility for the employment of deep learning in many fields. Deep learning [17]–[19] methods can learn from intensive material data to quickly find new materials with target property.

Convolutional neural network (CNN), as a classic application structure of deep learning, can automatically extract features from massive data and has good generalization capabilities. Long *et al.* [20] proposed a fully convolutional neural network (FCN) based on semantic segmentation, replacing the fully connected layer in CNN with a convolutional layer. Ronneberger *et al.* [21] proposed a symmetrical U-Net network based on the idea of encoder-decoder structure based on FCN and flexibly used the jump connection between deep and shallow networks, thus overcoming the fact that FCN cannot retain part of the pixel spatial position information. Moreover, context information leads to the shortcomings of loss of local features and global features. Zhou *et al.* [22] proposed the U-Net++ network, rethinking the down-sampling and up-sampling times of the UNet network, redesigning the multi-scale connection nodes in the original network, adding in-depth supervision to pruning different segmentation tasks to the appropriate network, Achieving a good segmentation effect. Mehta and Sivaswamy [23] proposed an M-Net network based on U-Net for brain magnetic resonance imaging image segmentation, three times faster than random forest and 2D CNN in volume segmentation. At present, the U-Net network and a series of derivative network structures [24] have become one of the most popular image semantic segmentation methods.

Although the UNet network and its series of derivative network structures perform well in image segmentation [25], [26], they are limited by a large number of intensive use of convolution operators, obtaining the interdependence between remote information is inefficient, and the receptive field of image feature extraction is excessively localized. Stacking more layers of convolution operators does not always increase the effective receptive field, so convolution operators still lack a mechanism for modeling remote dependency information.

There are many works on fusing multi-scale features. Hu *et al.* [27] propose a joint feature pyramid (JFP) module, and built a spatial detail extraction (SDE) module, design a bilateral feature fusion (BFF) module, making full use of the correspondence between high-level features and low-level features. Benvcevic *et al.* [28] propose training a neural network on polar transformations of the original dataset, such that the polar origin for the transformation is the center point of the object. Mohamed *et al.* [29] propose a hybrid SI based approach that combines the features of two SI methods, marine predators algorithm (MPA) and moth flame optimization (MFO). Liu *et al.* [30] The author proposes an improved Itti model and an improved GrabCut image segmentation algorithm for PET images that

are low-resolution grayscale images to solve the problems existing in the original algorithm in grayscale images.

The Nonlocal block [31] can capture remote dependencies more robustly and flexibly to help the deep network better integrate Nonlocal information. Some Nonlocal network modules currently proposed are NL [31], A2 [32], NS [33], CC [34], CGNL [35], SNL [36], etc. Chen *et al.* [32] propose the Double Attention block, which first collects the features in the entire space and then assigns them back to each location. Yue *et al.* [35] proposed a compact generalized Nonlocal block to capture cross-channel cues, which inevitably increases the noise of the attention map. Huang *et al.* [34] proposed a lightweight Nonlocal block called an interleaved attention block, which decomposes the positional attention of NL into conterminously column-wise and row-wise attention. In order to improve the stability of the NL block, Tao *et al.* [33] proposed using the Laplacian of the incidence matrix as an attention map, and the Nonlocal stage (NS) module can follow the diffusion characteristics. Zhu *et al.* [36] proposed The SNL (Spectral Nonlocal Block) module, which symmetrically processes the attention feature block. From the perspective of a new Chebyshev approximation and graph filtering, it uniformly explains the above five famous Nonlocal modules and provides a theoretical explanation.

The Nonlocal operator in the SNL block is equivalent to filtering the signal with a set of graph filters. This article takes this feature as a starting point, considering that the use of the SNL module for the incidence matrix is equivalent to filtering in the form of row transformation and lack of another corresponding form of column transformation filtering, proposed an improved SNL block and combined with the excellent performance of the image segmentation field UNet network structure. Experimental results show that the performance is state-of-the-art, improving mIoU significantly, and better than other segmentation networks such as Unet.

In a nutshell, our contributions are threefold:
- This paper proposes an improved nonlocal bolck, which can better segment deep and shallow information in effect.
- This paper combines an improved non-local attention module and UNet network, which significantly improves the accuracy and efficiency of image segmentation.
- Based on more accurate segmentation results, the microstructure characterization parameters such as volume fraction, rafting degree of γ phase and thickness of γ phase are measured.

II. METHOD

This paper proposes an improved SNL block, combined with the UNet network, to form an improved SNL-UNet network structure. The network inserts an improved SNL module after multiple downsampling to perform attention feature transformation, which has significantly improved the depth

information recognition ability of the image, and compared with other network structure methods, it can more accurately distinguish the target area γ' phase and background (non-γ' phase area) of the superalloy microstructure. Briefly mention that, The γ' phase is the ideal area for the white part of the image in Figure 3, and correspondingly, the γ' phase is the ideal area for the black part.

A. NETWORK STRUCTURE

The method network in this paper is an encoder-decoder structure as a whole, as shown in Fig. 1. The network consists of encoders, decoders, jump connectors, and improved SNL block.

The input image enters the encoder for image feature extraction and reduces network parameters through multiple maxpool downsampling to obtain a larger receptive field. The encoder part specifically includes a structure combining four repetitions of convolution and down-sampling. Each structure module includes convolution, ReLU activation function, BN layer, and maxpool down-sampling.

The image features collected through the encoder structure are sent to the improved SNL block. After passing through the improved SNL block, the output image feature block is up-sampled in the decoder part through linear interpolation and then concatenates with the features of the corresponding layer of the encoder on the channel. Then proceed to the traditional convolution module. After the above steps are repeated many times, the feature size is the same as the original image, and then convolution is used to perform image pixel-level classification and output feature map channels equal to the number of pixel classification categories for subsequent loss function calculation and network parameter training. In order to achieve the purpose of segmenting the microstructure image of the superalloy.

B. NONLOCAL BLOCK

The Nonlocal block [31] calculates a sum of a relationship between each position's characteristics. The nonlocal operator is defined:

$$F\left(\mathbf{X}_{i,:}\right) = \frac{\sum_{j}\left[f\left(\mathbf{X}_{i,:}, \mathbf{X}_{j,:}\right) g\left(\mathbf{X}_{j,:}\right)\right]}{\sum_{j} f\left(\mathbf{X}_{i,:}, \mathbf{X}_{j,:}\right)} \tag{1}$$

where $\mathbf{X} \in \mathbf{R}^{N \times C_{In}}$ is the input feature map, i, j are the position indexes in the feature map, $f(\cdot)$ is the affinity kernel with a finite Frobenius norm, $g(\cdot)$ is a linear embedding that is defined as $g\left(\mathbf{X}_{j,:}\right) = \mathbf{X}_{j,:}\mathbf{W}_Z$ with $\mathbf{W}_Z \in \mathbf{R}^{C_{In} \times C_T}$. Here N is the total positions of each feature. C_{In} and C_T are the numbers of channels for the input features and the transferred features implied by intermediate processes.

The NL block then adds a linear transformation with weight matrix $\mathbf{W} \in \mathbf{R}^{C_T \times C_{In}}$ and a residual connection to accommodate various network models:

$$\mathbf{Y}_{i,:} = \mathbf{X}_{i,:} + F\left(\mathbf{X}_{i,:}\right) \mathbf{W} \tag{2}$$
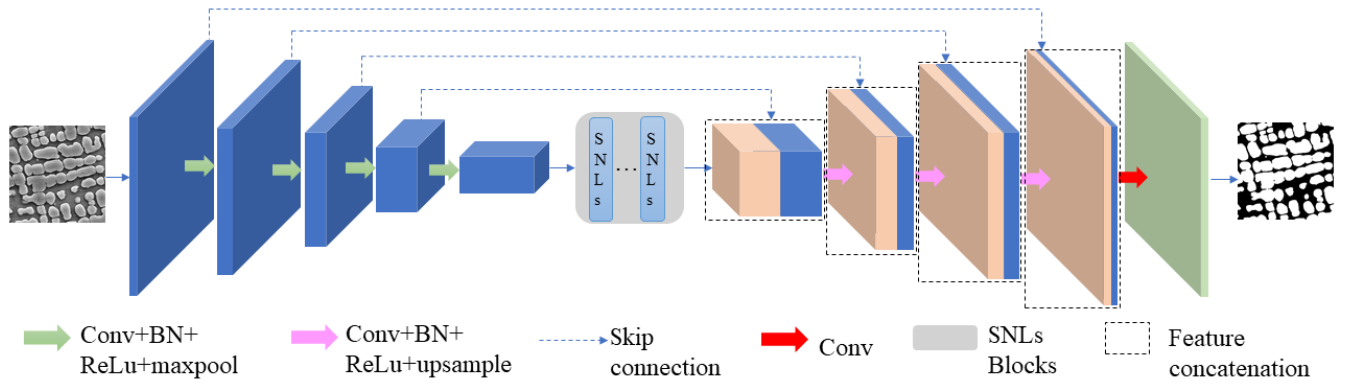
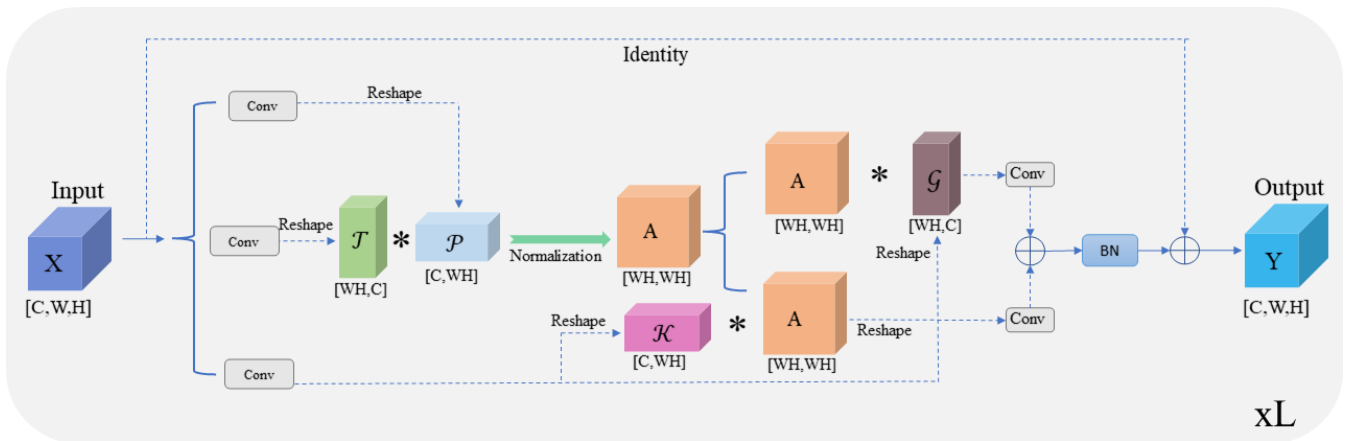**FIGURE 1.** Improve SNL-UNet network structure.

FIGURE 2. Improved SNL Block.

C. IMPROVED SNL BLOCK

As shown in Fig. 2, this paper proposes an improved SNL block based on the SNL block. This block is essentially a non-local attention block. In view of the fact that the receptive field of the convolutional network is too local, this block can capture remote dependencies more robustly and flexibly to help the deep network better integrate non-local information.

The input image features with the shape [C,W,H] are reshaped into a feature map with the shape [WH,C] after the 1×1 Conv on the two branches, Then Reshape into a feature map $\mathcal{T}$ with the shape [WH,C] and a feature map $\mathcal{P}$ with the shape transposed [C,WH], multiply the two feature maps by matrix. After symmetric normalization, the affinity matrix A is obtained, and the formula is as follows:

$$A = D_{\hat{M}}^{-\frac{1}{2}} \hat{M} D_{\hat{M}}^{-\frac{1}{2}}, \quad \hat{M} = \left(M + M^{\top} \right)/2 \quad (3)$$

In the above formula $M = TP$, $D_{\hat{M}}$ is a diagonal matrix containing the degree of each vertex of $\hat{M}$. Symmetrically processed $\hat{M}$ is composed of pairwise similarities between pixels,the affinity matrix A normalized from $\hat{M}$ in Figure 2 represented by a feature block att with the shape [WH,WH]. The affinity matrix A is the attention

block of image features, including the pairwise similarity relationship between image pixels and non-local attention information.

The input image features with the shape [C,W,H] are reshaped into a feature map with the shape [WH,C] after the 1×1 Conv on the other branch, Then Reshape into a feature map $\mathcal{K}$ with the shape [WH,C] and a feature map $\mathcal{G}$ with the shape transposed [C,WH], Feature map $\mathcal{K}$ and feature map $\mathcal{G}$ perform matrix left multiplication and matrix right multiplication with matrix A, respectively. Feature element-wise added after passing through the matrices W_1 and W_2 in the form of convolution, and then through the BN layer.On this basis, the output feature block is added with the identity element of the input feature block to form a residual jump connection. The formula is as follows:

$$Y = X + f_{BN}(\mathcal{K}AW_1 + A\mathcal{G}W_2) \quad (4)$$

The input feature map X passes through the entire module, and the output is the feature map Y. Such a module is called a stage, which can be selectively repeated several times according to the processed data set and task requirements to achieve the best feature attention extraction effect.

III. RESULTS AND ANALYSIS OF EXPERIMENTS

A. DATESET

The image dataset for this task is the microstructure image of the DZ125 directionally solidified superalloy blade material obtained by the ZEISS SUPRA 55 field emission scanning electron microscope. In DZ125's microstructures, such as carbides, grains, grain boundaries, dendritic core(DC) and interdendritic(ID) region. A grain boundary is the interface between two grains, or crystallites, in a polycrystalline material. A dendrite in metallurgy is a characteristic tree-like structure of crystals growing as molten metal solidifies, interdendritic region is the alloy region corresponding to the dendrite core. the γ' phase of dendrites is an important strengthening phase of superalloys, and the volume of γ' phase particles directly affects the strengthening ability of superalloys. It has the properties of uniform size and shape that can sensitively reflect the difference in service temperature and stress of blade materials. Therefore, this article focuses on the study of the γ' phase of the DZ125 microstructure.

Due to the difficulty of obtaining electron microscope images of this material, the long time to take microstructure images, and the high cost, the data sets are all small samples. The training set has a total of 200 images, and the image size is 512×512 pixels. During the training process, 10% of the images are randomly sampled as the verification set through cross-validation. We selected about ten images with different metallic morphology as the test set. Classification of the data set based on the different morphological characteristics exhibited by metallography and the different directional cut planes used for image acquisition. The dataset is composed of images of dendritic core(DC) and interdendritic(ID) region mixed together to form a training set. On this basis, the online method is adopted in the data set loading process, the CPU is used to generate transformed images, horizontal and vertical flipping, random mirroring and other data enhancement methods are carried out to expand the amount of data to reduce the risk of overfitting. Nevertheless, compared with other image segmentation tasks, this dataset can be described as a fairly small scale. The method in this paper still solves the problem of semantic segmentation implemented on small datasets of superalloy microstructure images. In the field of deep learning, there are many difficulties in the learning of small datasets, and most tasks require massive data for learning. Therefore, the method in this paper is of great significance for the realization of small dataset image segmentation.

B. EXPERIMENT

Since the network structure in this article is used to segment the microstructure image of superalloys, the purpose of the task is not only to classify the category of a single-pixel accurately but also to separate the phase structure of the image as a whole, so the combination of cross-entropy and Dice loss is adopted.

TABLE 1. Training hyperparameters.

hyperparameters	value
Initial learning rate	0.0001
Learning decay rate	1×10^{-8}
Batch size	20
Transfer matrix channels	16
Epoch	100

The experiment is based on the PyTorch1.9.0 open-source deep learning framework, and the training environment is Anaconda2020.11, CUDA10.2, cuDNN7.6.5. Training is performed on a server with GPU model NVIDIA Tesla V100-SXM2. The training parameters are as Table 1

C. SEGMENTATION RESULTS AND ANALYSIS

In order to verify the performance of the network structure proposed in this paper in the image segmentation task, it is compared with other semantic segmentation network structures on the same superalloy microstructure image data set. We conducted a comparison experiment with the same hardware conditions, the same hyperparameter configuration and the same number of image datasets. A variety of evaluation indicators were compared. The pixel-based evaluation methods mainly include pixel accuracy and mean accuracy, and the evaluation methods based on intra-class coincidence mainly include (Mean intersection over union, MIoU), (Frequency weighted intersection over union, FWIoU) and Dice score [37]. Cluster-based evaluation methods mainly include Rand index (RI) [38], [39] and Adjusted rand index (ARI) [40]. Table 2 shows the comparison of evaluation indexes corresponding to each segmentation method and the time consumed per image prediction in seconds.

The segmentation results show that for microscopic images of superalloys, the grain boundaries are not obvious or even difficult to distinguish with the naked eye due to the unique properties of metallography. The traditional threshold segmentation methods are not ideal. Although the segmentation effects of deep learning segmentation methods such as UNet, ResUnet [41] and the recently popular ViT [42] have excellent performance, there are apparent segmentation errors and unnecessary noise. As shown in Fig. 3, the area in the red box is a non-target area. Non-target areas refer to metallic phases that are not in the tangent plane of the image acquisition and they are perpendicular to the tangent plane. The segmentation effect of this method is better than that of other network structures, and the ability to recognize the depth information of the image is also significantly improved.The segmentation effect of this method is significantly better than other network structures, and the depth information recognition ability of the image has been significantly improved. Compared with other network structure methods, it can be more accurate. The γ' phase and the background (non-γ' phase area) of the target area of
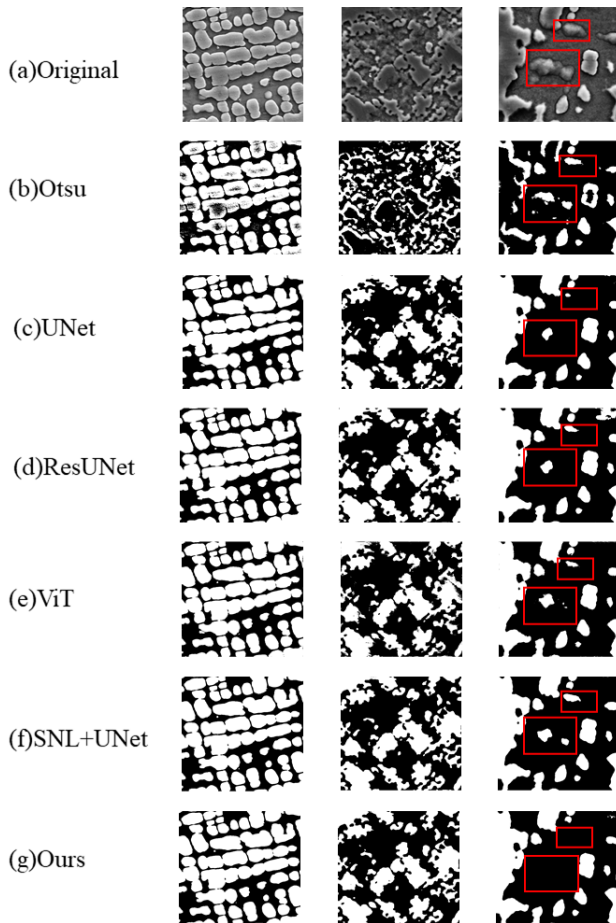
FIGURE 3. Comparison of image segmentation effects of different methods.

TABLE 2. Comparison of segmentation results of different methods.

Method	Pixel acc	Mean acc	mIoU	FWIoU
UNet	0.9595	0.9590	0.9209	0.9232
ResUNet	0.9594	0.9589	0.9203	0.9228
ViT	0.9285	0.9326	0.8671	0.8705
SNL+UNet	0.9602	0.9601	0.9220	0.9244
Ours	0.9713	0.9423	0.9320	0.9342

Method	Dice	RI	ARI	Time
UNet	0.9573	0.9302	0.8406	0.1381
ResUNet	0.9574	0.9318	0.8407	0.1820
ViT	0.9278	0.8763	0.7210	0.1403
SNL+UNet	0.9577	0.9337	0.8469	0.1851
Ours	0.9636	0.9375	0.8559	0.1296

the superalloy microstructure can be divided into the ideal level division and more accurate grain boundaries, and the evaluation indicators of the mIoU and the pixel accuracy are also significantly improved. Table 2 shows the comparison of evaluation indexes corresponding to each segmentation method.

Based on more accurate segmentation results, the characteristic parameters such as volume fraction, γ' phase thickness, and rafting degree are measured and calculated.

TABLE 3. Characterization parameters under different service time.

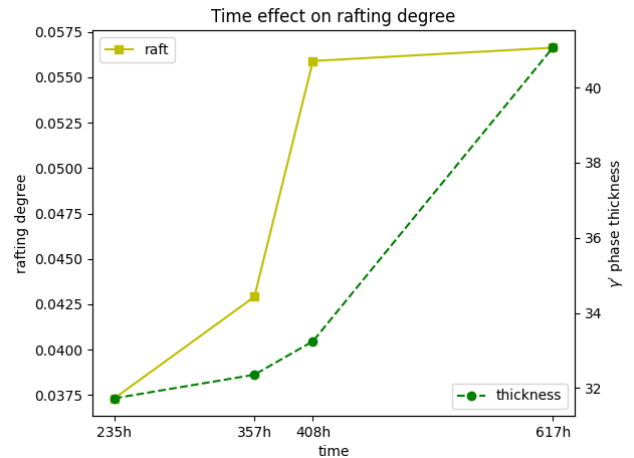| time	[Inter]dendritic	V_f	D	Ω
235h	dendritic	0.5815	28.37	0.0373
	interdendritic	0.5798	34.47	0.0373
357h	dendritic	0.5115	30.20	0.0421
	interdendritic	0.5301	34.89	0.0436
408h	dendritic	0.5488	32.05	0.0561
	interdendritic	0.5614	34.49	0.0556
617h	dendritic	0.5643	34.93	0.0536
	interdendritic	0.5935	47.05	0.0595

FIGURE 4. Schematic diagram of rafting degree and thickness change with time.

The phase volume fraction is defined as follows:

$$V_f = P/P_T \qquad (5)$$

where, P and P_T are the phase volume and total volume respectively. The γ' phase rafting degree Ω, which measures the degree of deformation of a material under the influence of temperature, stress, etc, is defined as follows:

$$\Omega = \frac{P_L^\perp - P_L''}{P_L^\perp + P_L''} \qquad (6)$$

where P_L represents the number of crossings and interruptions of the γ' raft-shaped tissue within a unit length in a specified direction. Ω is a value between 0-1. When Ω is 0, isometric organization. When Ω is 1, ideal raft organization. The larger the Ω value, the more perfect the raft shape. The γ' phase thickness D is the average pixel thickness of the γ' phase under 20k magnification. $P_L^\perp$ and P_L'' are two mutually perpendicular P_L, one of which is in the same direction as the metallurgical growth.

As shown in Figure 4, it is easy to know from the measurement results that as the service time increases, the degree of rafting gradually increases. The statistical results of the γ phase volume fraction distribution within the same service time are shown in Figure 5. The measured data and statistical results show that the distribution phenomena such as the increase of the raft with time and the increase
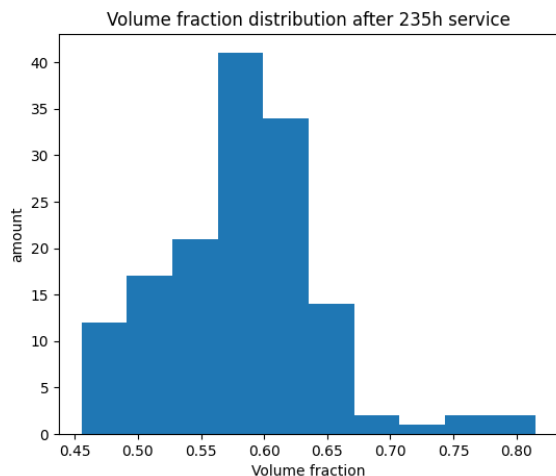
FIGURE 5. Volume fraction distribution after 235h service.

of thickness with time are consistent with the objective laws.

IV. CONCLUSION

In this paper, an improved module is proposed to solve the problem of metallographic features of different tangent planes that are difficult to accurately segment in microstructure images. And further combine the module with the UNet network to form a new improved SNL-Unet image segmentation structure, based on the improved SNL-Unet and conducts experimental verification with DZ125 alloy material microstructure images. The comparison with other methods shows that this method can more accurately distinguish the non-tangential metallographic phase that is difficult to segment, not only can meet the needs of complex image segmentation of superalloys, but also can further improve the segmentation accuracy and robustness of microstructure images of superalloys. In particular, the extraction of depth information is more accurate, ensuring the accuracy of the same plane tissue feature extraction and calculation, and it is significantly improved under various evaluation indicators. Based on more accurate segmentation results, the microstructure characterization parameters such as volume fraction, rafting degree of γ' phase and thickness of γ' phase are measured. The measured data and statistical results are consistent with the distribution of material properties with time and under stress, i.e., phenomena such as the increase in rafting with time and the increase in thickness with time. However, there are still many problems that need to be solved by follow-up research. For example, the data set is more specific and unique, and the restrictions are large, so the method is not general enough. The follow-up work will continue to increase the image analysis research work of different types of superalloy microstructure image organization, and carry out more research on the relationship between organization and performance in combination with material properties.

REFERENCES

[1] Y.-K. Kim, D. Kim, H.-K. Kim, E.-Y. Yoon, Y. Lee, C.-S. Oh, and B.-J. Lee, "A numerical model to predict mechanical properties of Ni-base disk superalloys," *Int. J. Plasticity*, vol. 110, pp. 123–144, Nov. 2018.

[2] E. Galindo-Nava, L. Connor, and C. Rae, "On the prediction of the yield stress of unimodal and multimodal γ' Nickel-base superalloys," *Acta Mater.*, vol. 98, pp. 377–390, Oct. 2015.

[3] S. Li, B. Wang, D. Shi, X. Yang, and H. Qi, "A physically based model for correlating the microstructural degradation and residual creep lifetime of a polycrystalline Ni-based superalloy," *J. Alloys Compounds*, vol. 783, pp. 565–573, Apr. 2019.

[4] Y. Li, Y. Liu, S. Luo, Z. Wang, K. Wang, Z. Huang, H. Zhao, and L. Jiang, "Neural network model for correlating microstructural features and hardness properties of nickel-based superalloys," *J. Mater. Res. Technol.*, vol. 9, no. 6, pp. 14467–14477, 2020.

[5] L. Rakoczy, O. Milković, B. Rutkowski, R. Cygan, M. Grudzień-Rakoczy, F. Kromka, and A. Zielińska-Lipiec, "Characterization of γ' precipitates in cast Ni-based superalloy and their behaviour at high-homologous temperatures studied by TEM and *in Situ* XRD," *Materials*, vol. 13, no. 10, p. 2397, 2020.

[6] K. Hou, M. Wang, P. Zhao, M. Ou, H. Li, Y. Ma, and K. Liu, "Temperature-dependent yield strength and deformation mechanism of a casting Ni-based superalloy containing low volume-fraction γ' phase," *J. Alloys Compounds*, vol. 905, Jun. 2022, Art. no. 164187.

[7] B. Sun, T. Zhang, and L. Song, "Flow behavior and dynamic recrystallization of a power metallurgy nickel-based superalloy during hot compression in ($\gamma + \gamma'$)-phase region," *J. Alloys Compounds*, vol. 891, Jan. 2022, Art. no. 161944.

[8] E. Payton, P. Phillips, and M. Mills, "Semi-automated characterization of the γ' phase in Ni-based superalloys via high-resolution backscatter imaging," *Mater. Sci. Eng., A*, vol. 527, nos. 10–11, pp. 2684–2692, 2010.

[9] J. Tiley, G. Viswanathan, A. Shively, M. Tschopp, R. Srinivasan, R. Banerjee, and H. Fraser, "Measurement of γ' precipitates in a nickel-based superalloy using energy-filtered transmission electron microscopy coupled with automated segmenting techniques," *Micron*, vol. 41, no. 6, pp. 641–647, 2010.

[10] X. Xu, S. Xu, L. Jin, and E. Song, "Characteristic analysis of Otsu threshold and its applications," *Pattern Recognit. Lett.*, vol. 32, no. 7, pp. 956–961, May 2011.

[11] L. Ding and A. Goshtasby, "On the Canny edge detector," *Pattern Recognit.*, vol. 34, no. 3, pp. 721–725, Mar. 2001.

[12] W. Gao, X. Zhang, L. Yang, and H. Liu, "An improved Sobel edge detection," in *Proc. 3rd IEEE Int. Conf. Comput. Sci. Inf. Technol. (ICCSIT)*, vol. 5, Jul. 2010, pp. 67–71.

[13] Y. Cao, L. Ju, Q. Zou, C. Qu, and S. Wang, "A multichannel edge-weighted centroidal Voronoi tessellation algorithm for 3D super-alloy image segmentation," in *Proc. CVPR*, Jun. 2011, pp. 17–24.

[14] H.-C. Chuang, L. M. Huffman, M. L. Comer, J. P. Simmons, and I. Pollak, "An automated segmentation for nickel-based superalloy," in *Proc. 15th IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 2280–2283.

[15] A. A. Ewees, M. Abd Elaziz, M. A. A. Al-Qaness, H. A. Khalil, and S. Kim, "Improved artificial bee colony using sine-cosine algorithm for multi-level thresholding image segmentation," *IEEE Access*, vol. 8, pp. 26304–26315, 2020.

[16] H. S. N. Alwerfali, M. Abd Elaziz, M. A. A. Al-Qaness, A. A. Abbasi, S. Lu, F. Liu, and L. Li, "A multilevel image thresholding based on hybrid salp swarm algorithm and fuzzy entropy," *IEEE Access*, vol. 7, pp. 181405–181422, 2019.

[17] B. Kailkhura, B. Gallagher, S. Kim, A. Hiszpanski, and T. Y.-J. Han, "Reliable and explainable machine-learning methods for accelerated material discovery," *NPJ Comput. Mater.*, vol. 5, no. 1, pp. 1–9, Dec. 2019.

[18] C. Wang, H. Fu, L. Jiang, D. Xue, and J. Xie, "A property-oriented design strategy for high performance copper alloys via machine learning," *NPJ Comput. Mater.*, vol. 5, no. 1, pp. 1–8, Dec. 2019.

[19] W. Huang, P. Martin, and H. L. Zhuang, "Machine-learning phase prediction of high-entropy alloys," *Acta Mater.*, vol. 169, pp. 225–236, May 2019.

[20] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.

[21] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham, Switzerland: Springer, 2015, pp. 234–241.

[22] Z. Zhou, S. Rahman, M. Mahfuzur, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, D. Stoyanov, Z. Taylor, G. Carneiro, T. Syeda-Mahmood, A. Martel, L. Maier-Hein, J. M. R. S. Tavares, A. Bradley, J. P. Papa, V. Belagiannis, J. C. Nascimento, Z. Lu, S. Conjeti, M. Moradi, H. Greenspan, and A. Madabhushi, Eds. Cham, Switzerland: Springer, 2018, pp. 3–11.

[23] R. Mehta and J. Sivaswamy, "M-Net: A convolutional neural network for deep brain structure segmentation," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 437–440.

[24] P. Ahmad, H. Jin, R. Alroobaea, S. Qamar, R. Zheng, F. Alnajjar, and F. Aboudi, "MH UNet: A multi-scale hierarchical based architecture for medical image segmentation," *IEEE Access*, vol. 9, pp. 148384–148408, 2021.

[25] A. Chowdhury, E. Kautz, B. Yener, and D. Lewis, "Image driven machine learning methods for microstructure recognition," *Comput. Mater. Sci.*, vol. 123, pp. 176–187, Oct. 2016.

[26] B. L. DeCost, B. Lei, T. Francis, and E. A. Holm, "High throughput quantitative metallography for complex microstructures using deep learning: A case study in ultrahigh carbon steel," *Microsc. Microanal.*, vol. 25, no. 1, pp. 21–29, Feb. 2019.

[27] X. Hu, L. Jing, and U. Sehar, "Joint pyramid attention network for real-time semantic segmentation of urban scenes," *Int. J. Speech Technol.*, vol. 52, no. 1, pp. 580–594, Jan. 2022.

[28] M. Benčević, I. Galić, M. Habijan, and D. Babin, "Training on polar image transformations improves biomedical image segmentation," *IEEE Access*, vol. 9, pp. 133365–133375, 2021.

[29] A. Mohamed, A. Ahmed, Y. Dalia, S. Husein, A. Qamar, and L. Songfeng, "An improved Marine Predators algorithm with fuzzy entropy for multi-level thresholding: Real world example of COVID-19 CT image segmentation," *IEEE Access*, vol. 8, pp. 125306–125330, 2020.

[30] H. Liu, W. Chu, and H. Wang, "Automatic segmentation algorithm of ultrasound heart image based on convolutional neural network and image saliency," *IEEE Access*, vol. 8, pp. 104445–104457, 2020.

[31] X. Wang, R. Girshick, A. Gupta, and K. He, "Non-local neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7794–7803.

[32] Y. Chen, Y. Kalantidis, J. Li, S. Yan, and J. Feng, "A^2-Nets: Double attention networks," 2018, *arXiv:1810.11579*.

[33] Y. Tao, Q. Sun, Q. Du, and W. Liu, "Nonlocal neural networks, nonlocal diffusion and nonlocal modeling," 2018, *arXiv:1806.00681*.

[34] Z. Huang, X. Wang, L. Huang, C. Huang, Y. Wei, and W. Liu, "CCNet: Criss-cross attention for semantic segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 603–612.

[35] K. Yue, M. Sun, Y. Yuan, F. Zhou, E. Ding, and F. Xu, "Compact generalized non-local network," 2018, *arXiv:1810.13125*.

[36] L. Zhu, Q. She, D. Li, Y. Lu, X. Kang, J. Hu, and C. Wang, "Unifying nonlocal blocks for neural networks," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 12292–12301.

[37] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[38] K. Y. Yeung and W. L. Ruzzo, "Principal component analysis for clustering gene expression data," *Bioinformatics*, vol. 17, no. 9, pp. 763–774, 2001.

[39] W. M. Rand, "Objective criteria for the evaluation of clustering methods," *J. Amer. Statist. Assoc.*, vol. 66, no. 336, pp. 846–850, 1971.

[40] L. Hubert and P. Arabie, "Comparing partitions," *J. Classification*, vol. 2, no. 1, pp. 193–218, 1985.

[41] F. I. Diakogiannis, F. Waldner, P. Caccetta, and C. Wu, "ResUNet–A: A deep learning framework for semantic segmentation of remotely sensed data," *ISPRS J. Photogramm. Remote Sens.*, vol. 162, pp. 94–114, Apr. 2020.

[42] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16×16 words: Transformers for image recognition at scale," 2020, *arXiv:2010.11929*.

LIXIN ZHANG was born in Qiqihar, Heilongjiang, in 1977. She received the B.S. and M.S. degrees in detection technology and automatic equipment from the University of Science and Technology Beijing, in 2000 and 2003, respectively, where she is currently pursuing the Ph.D. degree in control theory and control engineering. From 2007 to 2018, she was the Director of Research and Development and a Senior Engineer at the National Center of Materials Service Safety. Since 2018, she has been a Senior Engineer with the Experimental Center, School of Automation, University of Science and Technology Beijing. Her research interests include the study of applied methods of pattern recognition in the field of materials engineering, the study of technical methods of material microstructure image recognition, and the study of applied techniques of artificial intelligence in the industrial field.

HAOTIAN YAO was born in Chengdu, Sichuan, in 1997. He received the B.S. degree from the University of Science and Technology Beijing, in 2019, where he is currently pursuing the M.S. degree in control engineering. His research interests include image processing and machine vision-related.

ZHENGGUANG XU received the B.S. degree from the Shandong Institute of Building Materials, in 1982, and the M.S. degree in automatic control from Chongqing University, in 1988, and the Ph.D. degree from the University of Science and Technology Beijing, in 2001. He has been a Postdoctoral Supervisor with the Department of Automation, University of Science and Technology Beijing, since 2001. His research interests include the study of pattern recognition methods in complex production processes, motion pattern control theory and methods, biometric recognition, image processing, and pattern recognition.

HAN SUN was born in Baoding, Hebei, in 1998. He received the B.S. degree from the Xi'an University of Posts and Telecommunications, in 2020. He is currently pursuing the M.S. degree with the University of Science and Technology Beijing. His research interests include image processing and machine vision-related.

• • •