**Automatic grain size determination in metallographic images using deep learning**

**Abstract:** Grain size is one of the most important parameters in the metallographic microstructure analysis, and in most experiments this information is usually obtained by manual processes. However, those manual processes can take a long time and increase accidental errors. At present, the segmentation of metallographic image is limited by the complexity of traditional digital image processing, which makes the application scope of one algorithm narrow and specific, without universality. On the other hand, it is difficult to identify grain boundaries in the images with low contrast and bad defined boundaries. In the paper, we apply a deep convolutional neural network segmentation model to enable new automated microstructure segmentation for complex microstructure, usually manually and subjectively evaluated. The improved U-Net network is used to realize grain segmentation. The effectiveness of this method is verified by extensive metallographic images, and a practical case is given. It has a significant advantage for low contrast images, fuzzy boundary and complex structure of metallographic image edge extraction. The grain size was calculated according to American Society for Testing Material (ASTM) standards.

**Key words:** grain size; metallographic microstructure analysis; improved U-Net network; grain segmentation (关键词可以是摘要中没有，而正文中有的词么？)

## 1.Introduction

(晶粒尺寸的重要性)The measurement of grain size plays an important role in the research of metal materials. Information related to material properties, such as yield strength, tensile strength, elongation, etc. can be obtained, which has an important impact on material properties. The traditional methods of grain measurement rely on manual operation, which is time-consuming and prone to errors. In the analysis of metal microstructure, the information of grain boundary, grain size and grain distribution can be obtained by image analysis, and these parameters can be estimated by automatic methods of image processing and mathematical morphology. (商用软件)Currently, there are some commonly commercial software for grain size rating, such as image-pro Plus, Image Tool, Image J, etc. To some extent, these tools reduce the manual work and improve the analysis efficiency. However, the operation is tedious and the scope of application is too narrow to be competent for the analysis of complex images, with fuzzy boundary or much noise. Meanwhile, commercial software has a very high cost. (现有技术) There are also many methods to contribute to extraction of grain boundary and grain size rating. Jiang et al. used the multi-scale geodesic extension algorithm to restore and reconstruct grain boundaries based on the improved extension definition. Deng et al. proposed a closed edge extraction algorithm for metallographic structure based on Canny algorithm and gray contour. Park and Yu Ding et al. used convex analysis to divide the conforming boundary into individual components, and used the missing value estimation based on FPCA (function principal component analysis) to recover the missing boundary components. Orhan Dengiz et al. used neural network and fuzzy logic algorithm to detect the boundary of the alloy. Lauch Lukasz et al. used the feedback pulse coupled neural network to optimize the bionic algorithm to achieve part of the microstructure image. These methods have made important contributions to the measurement of grain size and grain segmentation, but they are not effective in dealing with problems such as fuzzy boundary, high noise and poor picture quality.

(本文贡献)In this paper，based on the microscopic image of 690 alloy and the mixed image data, a method to automatically measure the grain size of 690 alloy is introduced. After the preliminary pretreatment of metallographic image, the grain segmentation was carried out, and the discontinuous grain boundary was reconstructed and completely segmented by the method of deep learning, so as to explore the reliability of boundary extraction.

## 2.Related Work

At present, many researchers have done a lot of theoretical and experimental research on grain boundary extraction and made remarkable progress. Now，the mainstream of the research ideas are mainly divided into two directions [12]. One is to segment the grains in the metallographic image in order to obtain a complete grain. Grain boundaries are adjacent to each other. The main methods are: watershed segmentation algorithm [13] [14], level set segmentation algorithm, clustering segmentation algorithm, etc. Secondly, some operators are used to extract grain boundaries directly from metallographic images, such as canny operator [15], Sobel operator [16] and kirsch operator. For the first research idea, the main research focus is the grain in metallographic image. Many researchers have applied some excellent digital image segmentation algorithms to the grain segmentation of metallographic image. As for the second research idea, it mainly focuses on the research of grain boundary extraction in metallographic images, and the most common method is to apply the edge extraction algorithm to grain boundary extraction in metallographic images. Many researchers continue to try on the basis of the existing algorithm, broaden their thinking, and achieve certain results by combining other algorithms.

However, according to the changes of the data set, the processing methods of these two roads vary greatly, and it is difficult to accurately segment the more complex metallographic images. In recent years, with the rapid development of computer vision, deep learning is widely used in cell segmentation, road scene segmentation, etc., and achieved excellent results, so we apply deep learning to metallographic image segmentation.

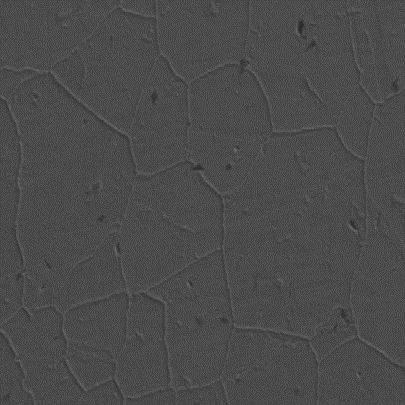
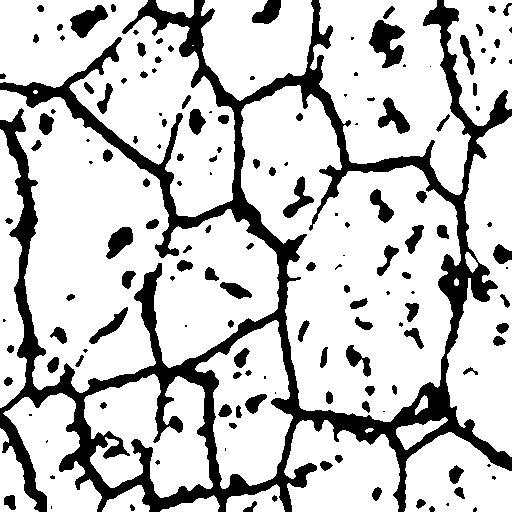
Since 2012, deep learning methods [17] have dominated many computer vision applications. This includes object recognition and detection, scene fusion, semantic segmentation and depth map prediction. The success of deep learning is often attributed to the fact that convolutional neural network (CNN) can effectively represent the visual data hierarchy, and combine the low-level image features (edges, color gradients) into the higher-level features (such as object parts) corresponding to the abstract quality of the image theme. Recently materials scientists have begun to explore the limited applications of modern computer vision techniques for flexible and universal microstructure characterization. Ref [18] and [19] explored these techniques in the context of microstructure classification. Ref [20] and [21] used the pre-trained CNN representation to study the relationship between the processing conditions and the microstructure obtained through the descending and visualization techniques. Ref [22] CNN segmentation model was used to identify the constituent phase of steel structure.

In recent years, various depth CNN architectures have been developed for intensive pixel-level tasks [23], such as semantic segmentation [24], edge detection, depth map, surface normal prediction [25], etc. Conceptually, it is the modern deep CNN that computes a highly nonlinear function by layer-by-layer combination of convolution, activation, and pooling (i.e., down-sampling) functions whose parameters are obtained from large annotated data sets through some variants of stochastic gradient descent [25,26].The classification neural network simplifies the input image into a single potential feature vector, in which the neural network designed for the pixel level task generates a potential representation for each pixel of the input image. This is usually achieved by fixed bilinear interpolation [26] or learning deconvolution operations [27]. In the latter network, the popular architectures are SegNet [24], Bayesian SegNet [28], u-net [29] with large data increment, and fully convoluted DenseNets [23]. In particular, u-net [29] is designed for medical image segmentation tasks with small data sets, and it relies on powerful data expansion to achieve good performance.

In this work, we apply the deep learning method to the image segmentation for complex microstructure data, aiming to extend the scope of quantitative analysis to metallographic microstructure systems.

## 3.Method

The preparation of metallographic samples mainly includes sampling, grinding, polishing and etching with appropriate etching agent, and then taking metallographic pictures with a microscope. However, there are many factors that affect the quality of the image, such as lighting conditions and surface irregularities. Fig.1a shows the original metallographic image, which has low contrast, high noise and many uncertain boundaries. Traditional image processing methods, such as global threshold method (bimodal method, iteration method and OTSU method) and edge detection operators such as Sobel, Prewitt, Canny operator, cannot get ideal results. Fig.1b is the result of threshold value obtained by OTSU method, indicating that the grain boundary and background cannot be separated by this method.

a b

fig1 metallographic image: a is the original image, and b is the image segmented by the threshold method

Method flow chart

**Pre-Processing**

During the preparation of metallographic samples, digital image often produces noise due to the existence of defects. Additional pretreatment should be carried out to eliminate noise. Median filtering and gaussian filtering are the most common denoising methods. Median filtering is a nonlinear filtering, which keeps edge information by removing salt and pepper noise, while gaussian filtering is a smooth linear filter, which can effectively eliminate gaussian noise. According to the requirement, we choose a specific method to eliminate noise.

**U-Net**

U-net network structure is a deformation structure of convolutional neural network, as shown in fig2. The entire u-net neural network consists of two main parts, namely the contracting path and expanding path. The contracting path is mainly used to capture the contextual semantic information in the image to extract the image feature structure. The expansion path, as opposed to it, is to precisely locate the parts of the image that need to be segmented. U-Net can take advantage of data enhancement operations to train some rare data samples. Therefore, the proposed U-Net network structure is very effective for deep learning of metallographic microscopic image with few training samples, which is mainly used for medical image segmentation.

Before applying it to the analysis of microscopic metallographic images, some super parameter of the original u-shaped network are adjusted. Inspired by [4][5], considering the deeper structure, the number of channels in each convolution layer is increased by 0.5 times to improve the ability of model fitting and generalization. In addition, set the image input dimension to 256×256 pixels.

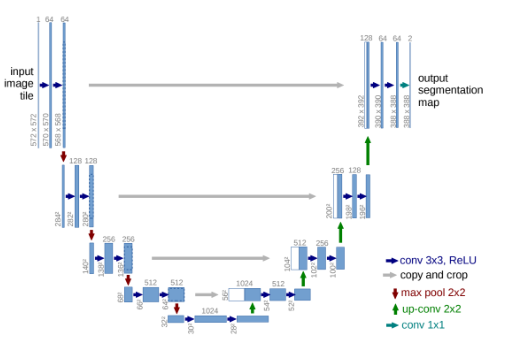


Fig2: unet network structure diagram

**Measurement of Performance**

Metallographic image segmentation is essentially a pixel-level classification task. In addition, there are far fewer boundary pixels than non-boundary pixels due to highly unbalanced instances. In order to accurately measure the performance of the proposed model, the following formula 1 scores, accuracy and recall degree are given in this paper, and then the intersection on the Union (IOU) is introduced, which is finally used in this study.



Where TP is True Positive, FP is False Positive, and FN is False Negtive, respectively. F1 considers both precision and recall, and calculates the harmonic average recall of precision and recall. F1 scores range from 0 to 1.The larger the F1 score, the better the model’s performance.

Many previous classification studies have used precision as the primary performance measure, which can be problematic. Precision may be valid if different result categories have about the same number of observations. However, metallographic data sets often have highly unbalanced distribution characteristics. In this study, another intuitive and informative metric, the IoU [18], was introduced and adopted. As shown in figure 3, an IoU is defined as the area of the intersection (the green area) divided by the area of the union (the red area).The intersection represents the ground truth value and the area covered by the prediction, and the union represents the ground truth value or the area covered by the prediction. If the prediction fits the facts perfectly, the corresponding IoU is 100%.

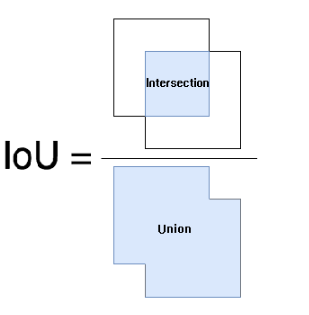
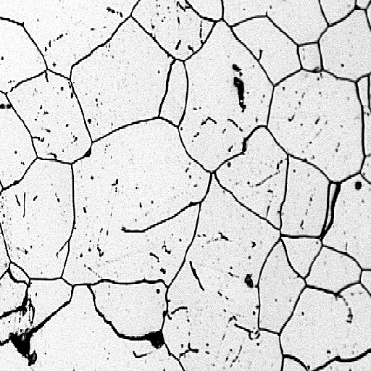


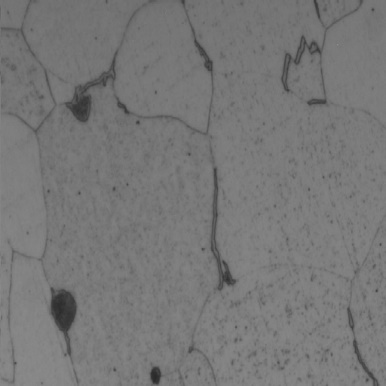
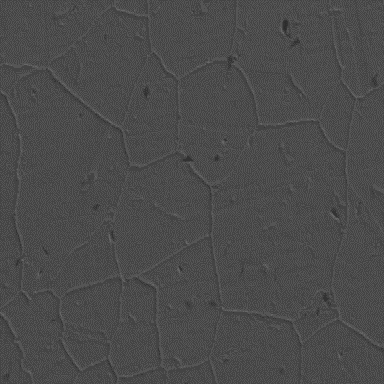
Fig3 definition of IOU

**DataSet**

In this paper, the self-made metallographic data set (690 alloy) was used to divide the selected images into blocks without overlapping. Finally, the pixel size of the metallographic image after cutting into blocks was 512\*512, with a total of 90 grayscale image samples. The metallographic image label produced by Labelme software is binary image. The prepared data set (including the original data and its corresponding label data) was divided into 60 pieces as the training set, 15 pieces as the verification set and 15 pieces as the test set. The picture of the graph of data set is light and dark, the shape is varied, have extensive.

a b

c d

fig4 original metallographic image sample

Compared with the general natural data set, the number of natural data sets is very limited. In this paper, translation, rotation and horizontal inversion are adopted to expand the data set of metallographic data that need to be trained. The data set expanded in this way makes the network model have the invariance of change and good robustness, and improves the training effect of the network model.

**Model Application**

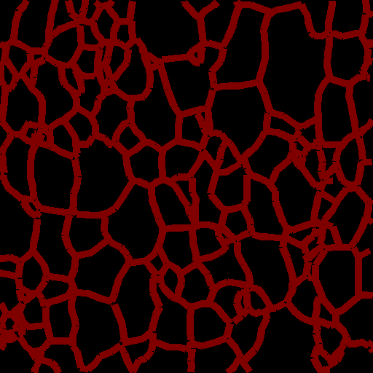
The amount of training data of u-net model is further increased by data expansion. For hyperparameter tuning, Adam optimizer is used and the learning rate is 0001. The number of training sets was set to 60 and the binary cross entropy was used as the loss function. Relu activation functions are used in all layers except the last one, which USES the sigmoid function. In addition, the batch size is set to 5.

Because deep learning methods often require a large number of annotated training images, it is important to investigate whether data from one source can be used to train models that will apply to data from different sources.(consider two training sets, one a pure 690 alloy metallographic diagram and the other a mixed single-phase picture, considering its versatility.).

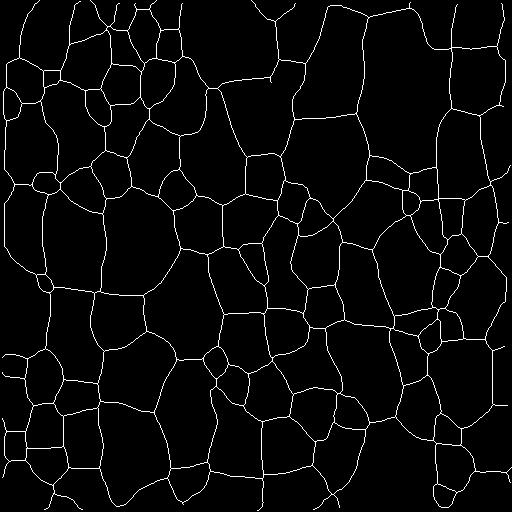
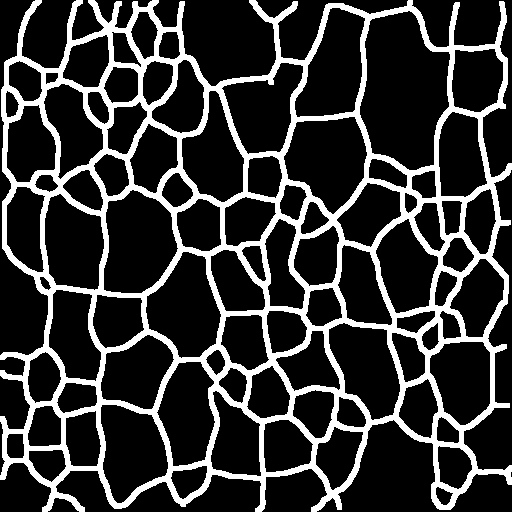
1. **Experiment and Discussion**

**Deep Learning Dataset Establishment**

Taking FIG. 4a as an example, the Labelme label was used to generate FIG. 4e. As the marked red boundary was too thick, the small crystal was almost covered. Firstly, the area of 8\*8 is used for smoothing and denoising, and then the binarization is carried out. Morphology is used to extract the skeleton, also known as binary image refinement. This algorithm can refine a connected region into the width of a pixel, which is used for feature extraction and target topology representation, such as f. however, as the boundary has only one pixel, the loss rate to 0.002 is very low, the accuracy to 98%, but the effect is very poor. This is because the proportion of the boundary is very small in the background, so it is used. Therefore, the method of corrosion in morphology is used to thicken the boundary. Finally, the label of figure a is like g.

a e

f g

fig5 Labelme is used to label the image

**Morphological method**

In order to facilitate the extraction of grain boundary and grain quantitative information, the gray scale map was converted into a binary map. After the binarization, there is a large amount of dotted salt and pepper noise, and the median filter method is used to remove it. In order to save the linear boundary and similar linear grain substructure adopts spot clean division, 8 connected neighborhood pixels model is adopted to search for "filter" image edge. the edge sealing spots pixel and spot pixel threshold set by the first comparison, proper maintenance, remove spots and reserve the substructure grain internal similar rod. The structure after noise elimination is like the "track" diagram. The "track" diagram still has many breakpoints that need to be fixed. This paper adopts the expansive corrosion technique in morphology, that is, the open operation (corrosion before expansion), to get the "opening" diagram.

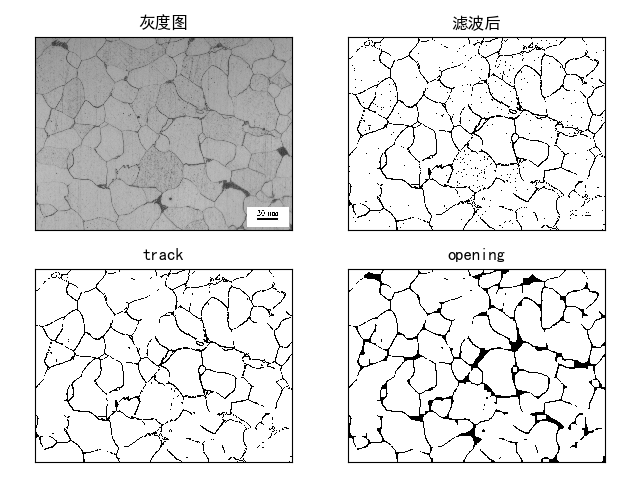


Fig 6: Results obtained by morphological methods

Although the noise is basically eliminated, the boundary becomes discontinuous and cannot achieve the effect of automatic rating.

**Watershed method**

Watershed algorithm is an important algorithm in grain segmentation, but for metallographic images with too much noise, it is easy to produce over-segmentation. The effect is shown in FIG. 7, which is very poor. It’s not going to be able to do the automatic rating.

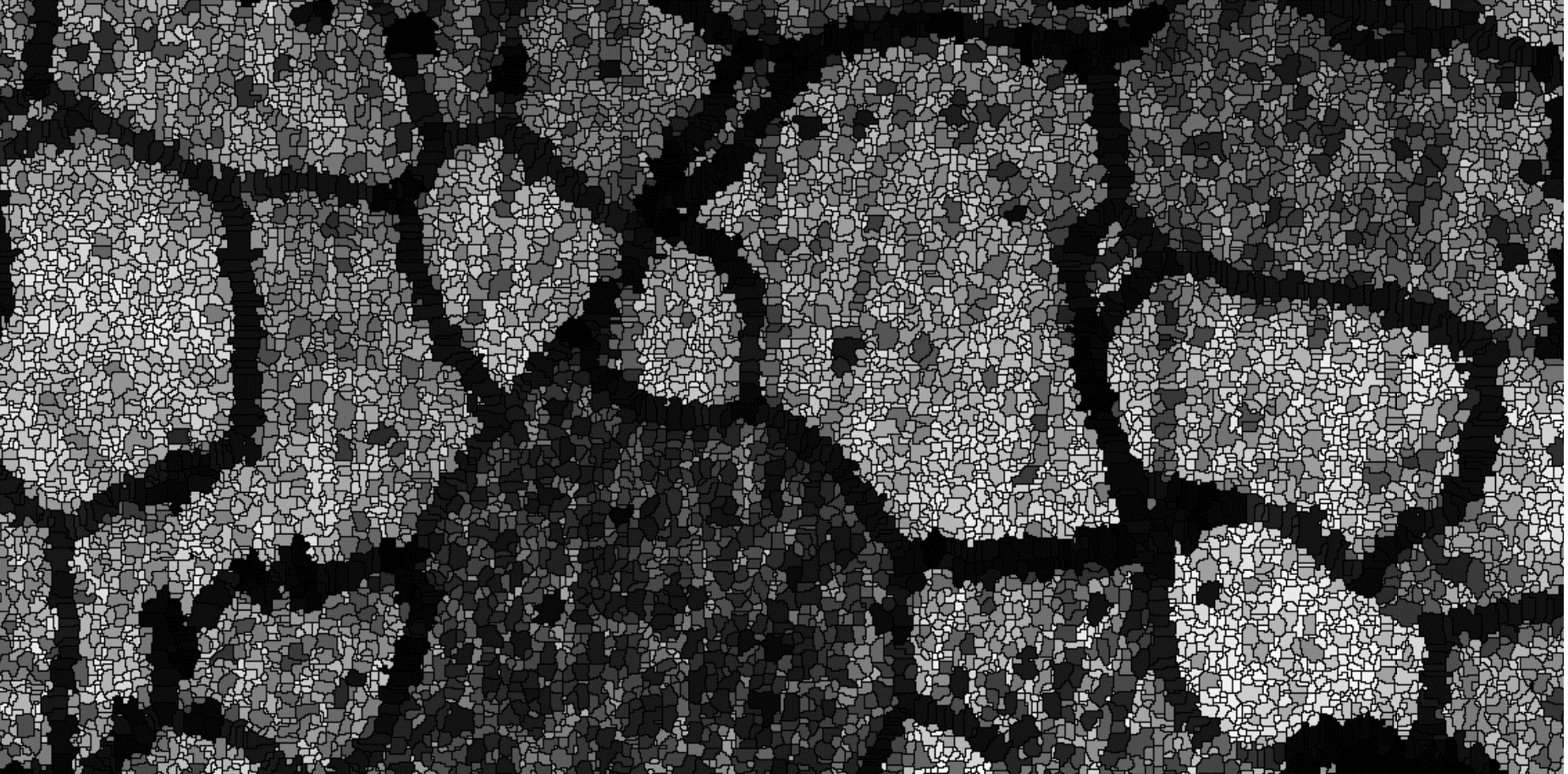


Fig.7 The results obtained through the watershed

**Directly with U-Net**

The structure of the direct u-net training network is as follows:

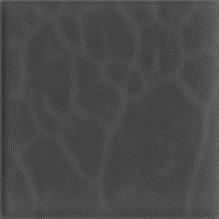
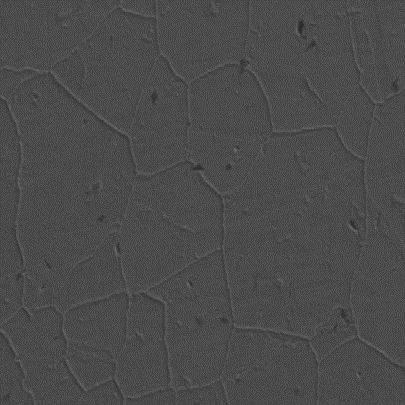


Fig 8 The results obtained directly by U-Net method

Because the contrast of the picture is very poor, if you use the u-net network for training, not very good results.

**Improved**

Firstly, the training image is processed with grayscale and median filtering, and then the 8\*8 region is used to smooth the denoising. Finally, an adaptive threshold binarization is carried out. Then it was trained with a preliminary improved u-net.

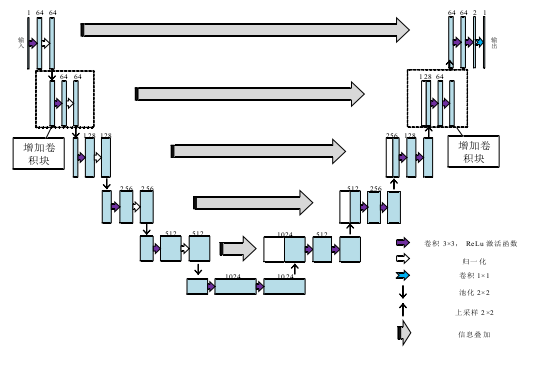


Fig9: preliminarily improved u-net

The test is as follows:

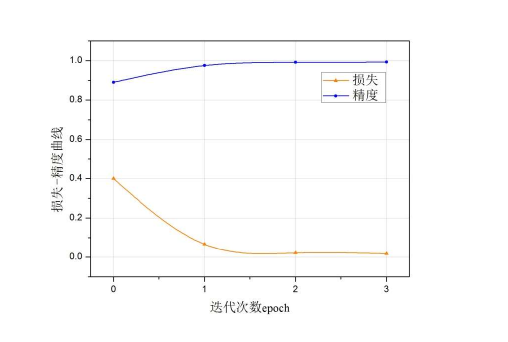
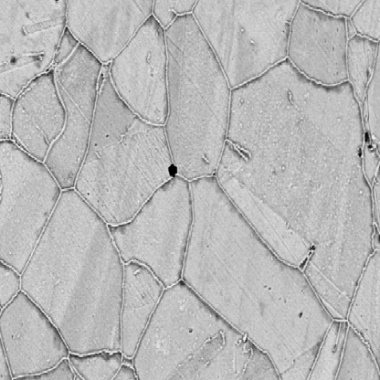
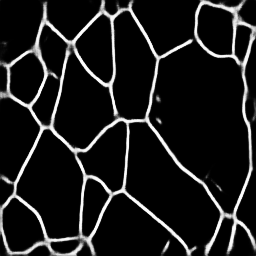
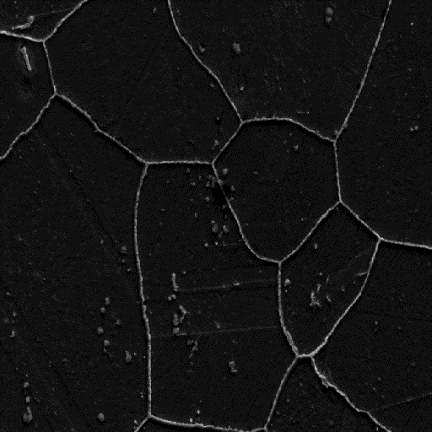
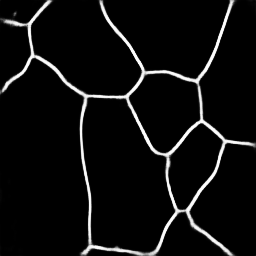
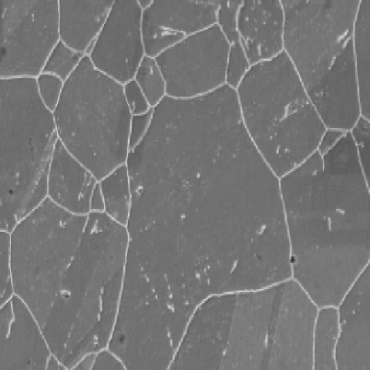
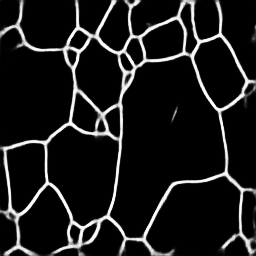


Fig 10: Number of iterations and loss relation graph

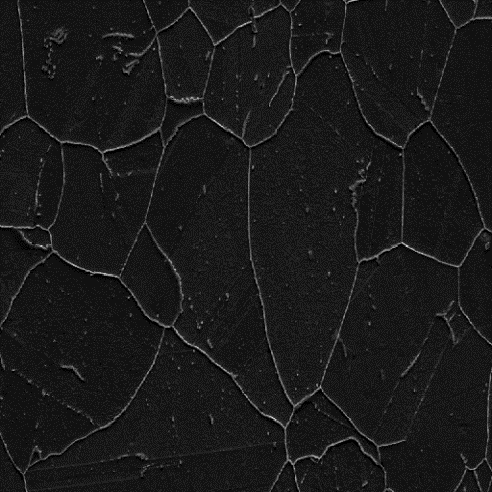
 

Fig 11: Results of our method

其他网络的对比表格（deeplab系列 segNet PixelNet）

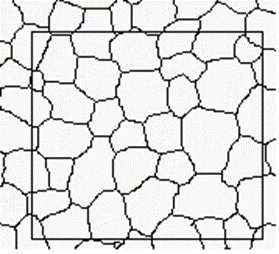
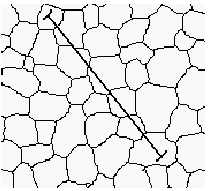
**Grain size rating**

According to ASTM standard e112-12 [25], three commonly used methods for calculating particle size are provided. These are the comparison processes，the plane measurement processes and the intercept processes. The comparison process does not require the calculation of particles, intercepts, or intersections. But it compares the images to a series of standard hierarchical images. Planar measurements involve the actual count of particles in a known area. The intercept method involves the number of grains intercepted by a test line or intersected by a test line, that is, the unit length of the test line. ASTM standard (e112-12) recommends the use of plane measurements and intercept methods to improve machining accuracy. They are repeatable and reproducible with an accuracy of 0.25 units of particle size and 0.5 units of particle size. With the same precision, the intercept method is faster than the plane method.

The intercept method and the plane method can produce the same level precision. But the intercept method is faster than the plane method. According to the advantages of the intercept method, the selected intercept method is used to determine the grain size. The grain size can be calculated as follows:



Where, l is the average intercept length in mm, and G is the grain size in the equation.

The area method Cutting line method

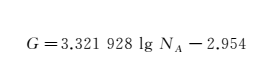
Measure the grain area step：

A. Select a rectangular area to calculate the area of the rectangular area

B. Calculate the number of grains: ( is the number of grains in the box, is the number of grains across the box boundary, minus 1 because the grains on the corner were counted twice, a total of four corners, so we need to subtract 1)

C. The grain area is obtained by dividing the rectangular area by the number of grains

D. Calculate the number of grains in unit area (mm2), and the grain size level



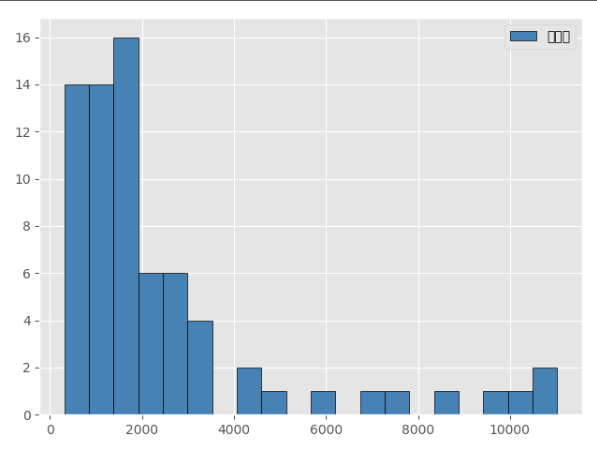
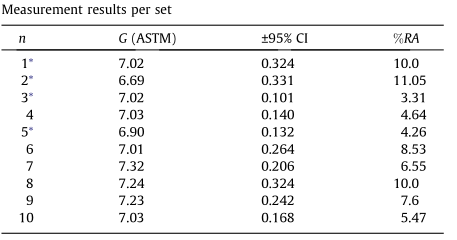


Fig 12 Area distribution statistics

Grain size level statistics：



1. **Conclusion and Future Work**

Automatic determination of grain size by deep learning technique is a new method for microstructure analysis of materials. In this study, a deep learning-based segmentation method was proposed for the defects and features in the metallographic image of 690 alloy, and the mathematical morphology and watershed methods were compared to extract and reconstruct grain boundaries. Experimental results show that this method is superior to the traditional metallographic image processing method. The grain boundary obtained by this method is complete and continuous in a short time. In order to improve the accuracy and performance of this method, further research will be carried out and an enhanced image processing algorithm for metallographic analysis will be adopted. After digital image processing, more detailed shape fact；or, fractal dimension and other parameters can be obtained to describe the microstructure characteristics of high-strength aluminum alloy, so as to establish the characterization system of high-strength aluminum alloy.