Article

Microstructure Instance Segmentation from Aluminum Alloy Metallographic Image Using Different Loss Functions

Dali Chen ^{1,2,*}, Dinghao Guo ^{1,2}, Shixin Liu ^{1,2} and Fang Liu ³

¹ State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China; dinghaoguo@foxmail.com (D.G.); liushixin@mail.neu.edu.cn (S.L.)

² College of Information Science and Engineering, Northeastern University, Shenyang 110819, China

³ College of Materials Science and Engineering, Northeastern University, Shenyang 110819, China; liufang@smm.neu.edu.cn

* Correspondence: chendali@ise.neu.edu.cn

Received: 22 February 2020; Accepted: 12 March 2020; Published: 17 April 2020

Abstract: Automatic segmentation of metallographic image is very important for the implementation of an automatic metallographic analysis system. In this paper, a novel instance segmentation framework of a metallographic image was implemented, which can assign each pixel to a physical instance of a microstructure. In this framework, we used the Mask R-CNN as the basic network to complete the learning and recognition of the latent feature of an aluminum alloy microstructure. Meanwhile, we implemented five different loss functions based on this framework and compared the influence of these loss functions on metallographic image segmentation performance. We carried out several experiments to verify the effectiveness of the proposed framework. In these experiments, we compared and analyzed six different evaluation metrics and provided constructive suggestions for the performance evaluation of metallographic image segmentation method. A large number of experimental results have shown that the proposed method can achieve the instance segmentation of an aluminum alloy metallographic image and the segmentation results are satisfactory.

Keywords: metallographic image; instance segmentation; loss function; Mask R-CNN

1. Introduction

Aluminum alloy products have been widely used in machinery manufacturing, transportation, electrical, shipbuilding, automobile, aviation, aerospace, chemical industry, construction and other fields [1–4]. Their properties highly depend on the distribution, shape and size of microstructures. Metallographic image processing is an important tool for analyzing the microstructure of metal materials [5–8]. Recently, a large number of automatic metallographic image processing methods have been proposed, which greatly improve the efficiency of metallographic analysis [9–11]. According to the use of these methods, we can classify them into four categories: grain boundary extraction, quantitative calculation, microstructural classification and segmentation.

Metallographic image segmentation is an important part of the automatic metallographic analysis system and it is also a challenging task. It aims to segment and recognize the different microstructures in the given metallographic image. At present, more and more scholars are beginning to pay attention to how to accomplish this task efficiently, and many effective methods have been introduced. To date,

these methods have been mainly divided into three groups: image processing-based methods, machine learning-based methods and deep learning-based methods.

Image processing-based methods segment the microstructure using typical image processing methods. For example, the watershed method, the typical digital image processing method, is used to segment microstructures [12]. The improved mean shift algorithm is used to segment the grains based on the characteristics of each region [13]. The directional wavelets are used for grain boundary extraction [14]. And the double-threshold binarization is applied to segment the grains [15]. In addition, the famous Markov Random Field (MRF) [16], coarse-to-fine [17] and Fuzzy C-Means (FCM) [18] are also applied for metallographic image segmentation.

Machine learning-based methods turn the metallographic image segmentation problem into a classification task. Classifiers are trained from a set of features along with labels. To date, for metallographic image segmentation, the commonly used classifiers have included multilayer perceptron [19], random forest [20], optimum-path forest [21], neural network [22] and support vector machine (SVM) [23,24]. These methods often outperform the image processing-based methods. However, the performance of these methods mainly depends on the choice of discriminable features. Without discriminable features, these methods often fail to achieve satisfactory results. Unfortunately, it is often difficult to obtain discriminable features. This greatly limits the development of machine learning-based method.

Deep Learning methods have dramatically improved the conventional machine learning methods due to their strong ability to automatically learn the discriminable features [25,26]. These methods have been successfully applied in the material science [27]. In recent years, more and more deep learning-based methods have been proposed for metallographic image segmentation. In [28], the famous FCN (Fully Convolutional Network) was used to segment some given microstructures of low carbon steel. Moreover, in reference [29], the DeepLab network was applied to segment Al-La alloy metallographic images.

These deep learning-based methods achieve satisfactory results, but they are unable to identify microstructure instances. In this paper, we were interested in a novel, more challenging problem of microstructure instance segmentation, which entails identifying, at a pixel level, where the microstructures appear as well as associating each pixel with a physical instance of a microstructure. In contrast, microstructure detection and semantic segmentation each only concern one of the two. To date, the instance segmentation methods have been widely used in medicine [30–32], transportation [33] and other fields. However, we have not found the application of this method in metallographic image segmentation. Therefore, in this paper, we applied Mask R-CNN (Mask Region-based Convolutional Neural Network) [34] to metallographic image segmentation of aluminium alloy, which is one of the most famous instance segmentation networks at present. In addition, we noticed that in the process of metallographic image processing, the impact of the loss function of neural networks has not received much attention. Therefore, we compared the performance of five loss functions and showed the importance of loss function.

To date, many performance evaluation metrics of segmentation methods have been proposed, which can effectively evaluate the different segmentation methods for natural scene images. However, metallographic images are often different from natural scene images. For example, the microstructure object or foreground in the aluminum alloy metallographic image is often far smaller than the aluminum background. Therefore many evaluation metrics cannot effectively evaluate the metallographic image instance segmentation methods. To solve this problem, we compared and analyzed the six typical performance evaluation metrics of segmentation methods, and provided constructive suggestions for the evaluation of aluminum alloy metallographic image segmentation methods.

We summarized the contributions of this paper as follows: (1) We implemented the instance segmentation framework of metallographic image, which could achieve automatic microstructure instance segmentation for given aluminum alloy metallographic images and provide a more effective tool for the quantitative analysis of metallographic images. (2) Five different loss functions were applied to the

proposed segmentation method, and their comparative analysis results are given. (3) We compared and analyzed six typical segmentation performance evaluation metrics, and obtained some conclusions which are helpful for the segmentation performance evaluation of aluminum alloy metallographic images.

This paper is organized as follows: Section 1 introduces prior works and our contributions. In Section 2, we introduced the proposed method, including parameter learning and instance segmentation. Section 3 presents the experiment results' quantitative performance comparison, qualitative performance comparison and convergence analysis. The paper is concluded in Section 4.

2. Proposed Method

2.1. Overview

The microstructure instance segmentation is an important part of the automated analysis system of aluminum alloy metallographic images. The basic framework of our proposed method is shown in Figure 1. It consists of two parts: parameter learning and instance segmentation. In the next section, we introduce the proposed method in detail.

Figure 1. The flowchart of the proposed method.

2.2. Parameter Learning

The instance segmentation method maps a given image from image space to instance space. In this process, the mapping function is very important. Because of the complexity of this process, the form of the mapping function is very complex, so we cannot obtain the exact form directly. To solve this problem, we used Mask R-CNN to model this process. Mask R-CNN contains a large number of unknown parameters. The main task of parameter learning is to estimate all unknown parameters of Mask R-CNN according to the specific sample data. The specific flow chart is shown in Figure 1.

The input of the parameter learning strategy is a given metallographic image instance segmentation training dataset. In order to improve the quality of the training dataset, we need to preprocess the given training dataset by using the image processing methods, including image resize, image flipping, Gaussian smoothing, image denoising and contrast enhancement. Then, the network parameters are learned from this enhanced dataset.

Mask R-CNN is one of the most popular instance segmentation networks at present, which has satisfactory performance. Many methods are improved on the basis of Mask R-CNN. The typical Mask R-CNN consists of five main modules: Feature Pyramid Networks (FPN), Region Proposal Network (RPN), Region of Interest (ROI) Align, Regions with CNN Features (R-CNN) and Fully Convolutional Networks (FCN). The detailed description is as follows:

(1) FPN uses feature maps from the bottom to the top, makes full use of the extracted features of each scale and can better mine multi-scale information. The input of this part is a given metallographic image and the output is a multi-resolution feature map of the metallographic image. In this module, we used a typical ResNet101 deep residual network to extract deep feature, which consists of 33 residual blocks (each residual block consists of 3 convolutional layers), 1 convolutional layer and 1 fully connected layer.

(2) RPN aims to obtain candidate regions. The input of this network is multi-resolution feature maps, and the output is a series of category probability and coordinates of the proposal regions. It consists of one 3×3 size convolutional layer, one $2k \times 1 \times 1$ size convolutional layer and one $4k \times 1 \times 1$ size convolutional layer. In this paper, the proportional parameter $k = 3$.

(3) ROI Align aims to obtain a fixed size feature map. The input of the network is the proposal region, and the output is the fixed size feature map. Compared with the typical ROI Pooling method, it eliminates the network quantization error and improves the segmentation effect of small objects.

(4) The tasks of R-CNN are classification and regression. The input of the network is the fixed size feature map, and the output is the predicted category probability and coordinate. This network consists of two convolution layers and two full connection layers.

(5) FCN aims to obtain the pixel-level segmentation results of metallographic images. Its input also is the fixed size feature map, and the output is the pixel-level segmentation image. This network consists of four 3×3 convolution layers, one deconvolution layer and one 1×1 convolution layer.

The overall loss function for training the Mask R-CNN can be denoted as follows:

$$L(\theta) = L_{RPN}(\theta) + \lambda_{rcnn} L_{RCNN}(\theta) + \lambda_{mask} L_{MASK}(\theta), \quad (1)$$

where θ is the network parameter, which is very important for generating the accurate instance segmentation results. λ_{rcnn} and λ_{mask} are the loss-balancing parameters. The RPN loss function L_{RPN} is defined by

$$L_{RPN}(\theta) = L_{cls}(c(\theta), c^*) + \lambda_{loc1} L_{loc}(r(\theta), r^*), \quad (2)$$

where c and c^* are the predicted and ground-truth labels, respectively. r and r^* are the predicted and ground-truth regression targets, respectively. λ_{loc1} is the loss-balancing parameter. $L_{cls}(\theta)$ is the softmax loss and $L_{loc}(\theta)$ is the smooth L_1 loss. The R-CNN loss function L_{RCNN} is defined by

$$L_{RCNN}(\theta) = L_{cls}(c(\theta), c^*) + \lambda_{loc2} L_{loc}(t(\theta), t^*), \quad (3)$$

where t and t^* are the predicted and ground-truth regression targets, respectively. λ_{loc2} is the loss-balancing parameter. The Mask loss function L_{MASK} is defined by

$$L_{MASK}(\theta) = L_{MASK}(m(\theta), m^*), \quad (4)$$

where m and m^* are the predicted and ground-truth mask labels, respectively. $L_{MASK}(\theta)$ is the standard binary cross-entropy loss. Using the maximum likelihood estimation, we can compute the network model parameters θ by solving the following optimization problem:

$$\theta^* = \arg \min_{\theta} L(\theta). \quad (5)$$

The stochastic gradient descent (SGD) algorithm is used to solve this optimization problem.

2.3. Instance Segmentation

The microstructure instance segmentation aims to entail associating each pixel with a physical instance of a microstructure. As shown in Figure 1, we input the given aluminum alloy metallographic images and use the Mask R-CNN to achieve the microstructure instance segmentation. Suppose that we are given a training dataset $D = \{I_n, y_n; n = 1 : N\}$. Here, I_n is the n -th metallographic image and y_n is the n -th ground-truth label. For easy description, we let f_{θ} represent the Mask R-CNN, and the parameter θ is obtained by solving Equation (5). Therefore, the instance segmentation image y' can be obtained by

$$y' = f_{\theta}(I'), \quad (6)$$

where I' represents the given aluminum alloy metallographic image. To summarize, our proposed instance segmentation method in a form of a pseudo-code is done in Algorithm 1.

Algorithm 1 Microstructure instance segmentation method.

Input: Training dataset $D = \{I_n, y_n\}$, new aluminum alloy metallographic image $\{I'\}$;

Output: The instance segmentation image y' ;

Step 1: Initializations:

- FPN: Backbone, Strides, Pyramid size.
- RPN: Anchor scales, Anchor ratios, Anchor stride, Threshold, Train anchors per image.
- R-CNN and Mask: Pool size, Mask pool size, Train ROI per image, ROI positive ratio, Mask shape, Detection max instances, Detection min confidence, Detection threshold.
- Learning rate and momentum: Learning rate, Learning momentum, Weight decay.
- Initialize parameter θ obtained from pretraining on the MSCOCO dataset .

Step 2: Optimize θ by using D :

- While not converge do.
- Compute network parameter θ^* by solving optimization problem (5) using SGD with momentum algorithm.
- Update $\theta^* \rightarrow \theta$.
- End while.

Step 3: Compute y' by using Equation (6).

2.4. Loss Functions

As opposed to common natural scene images, in metallographic images such as the ones we are processing in this work, the microstructures occupy only a very small region of the background. This often causes that the predictions are strongly biased towards background. As a result the microstructures are often missing or only partially segmented.

Many improved loss functions have been proposed to solve the problem of data imbalance, so it is necessary to analyze the impact of these loss functions on the performance of metallographic image

segmentation. However, there is still no research work in this area. To solve this problem, in this paper, we applied five different loss functions to metallographic image segmentation and compared their segmentation results.

The overall loss function for training the Mask R-CNN consists of three parts: RPN loss, R-CNN loss and Mask loss, as defined by Equation (1)–(4). RPN loss is used for classification and regression of foreground and background. R-CNN loss is used for classification and regression of specific categories. Mask loss is used for pixel level segmentation. Obviously, only Mask loss directly affects the accuracy of instance segmentation. Therefore, in this paper, we mainly focused on five different Mask loss functions, including L_{BCE} , L_{DICE} , L_{IOU} , $L_{Tversky}$ and L_{SS} . Let m_i^* and m_i be the predicted and ground truth binary labels, respectively. The detailed description is as follows:

(1) The loss function L_{BCE} is the standard binary cross-entropy loss [35], which is defined by

$$L_{BCE} = -\frac{1}{N} \sum_{i=1}^N (m_i^* \log m_i + (1 - m_i^*) \log(1 - m_i)). \quad (7)$$

(2) The Dice loss is proposed in reference [36], which is able to solve the problem that the foreground occupies only a very small region of the background. It is defined by

$$L_{DICE} = 1 - \frac{2 \sum_{i=1}^N m_i^* m_i + \zeta}{\sum_{i=1}^N m_i^* + \sum_{i=1}^N m_i + \zeta}, \quad (8)$$

where ζ is the smoothing parameter.

(3) The IOU loss is proposed in [37], which is able to solve the problem that the two classes (foreground and background) are very imbalanced. Its function L_{IOU} is defined by

$$L_{IOU} = 1 - \frac{\sum_{i=1}^N m_i^* m_i + \varepsilon}{\sum_{i=1}^N (m_i^* + m_i + m_i^* m_i) + \varepsilon}, \quad (9)$$

where ε is the smoothing parameter.

(4) The Tversky loss is proposed in reference [38], which is used to address the issue of data imbalance and achieve a much better trade-off between precision and recall. To define the Tversky loss function, we use the following formulation:

$$L_{Tversky} = 1 - \frac{\sum_{i=1}^N m_i^* m_i + \zeta}{\sum_{i=1}^N (m_i^* m_i + \alpha m_i (1 - m_i^*) + \beta m_i^* (1 - m_i)) + \zeta}. \quad (10)$$

where α and β control the magnitude of penalties for false positives and false negatives, respectively.

(5) In [39], Brosh et al. used a combination of sensitivity and specificity, which can be used together to measure classification performance even for vastly unbalanced problems. This novel loss function L_{SS} is defined by

$$L_{SS} = \gamma \frac{\sum_{i=1}^N (m_i^* - m_i)^2 m_i^* + \varepsilon}{\sum_{i=1}^N m_i^* + \varepsilon} + (1 - \gamma) \frac{\sum_{i=1}^N (m_i^* - m_i)^2 (1 - m_i^*) + \varepsilon}{\sum_{i=1}^N (1 - m_i^*) + \varepsilon}. \quad (11)$$

where γ is the sensitivity ratio used to assign different weights to the two terms.

3. Experimental Results

In order to verify the effectiveness of our proposed method, a large number of experiments were performed for the analysis of instance segmentation performance.

3.1. Experimental Setup

Datasets: In order to verify the proposed method, we built an experimental dataset, which contains 100 five-series aluminum alloy metallographic images. These aluminum alloy metallographic images were taken by a metallographic microscope and include three different types of phases, such as Mg_2Si , aluminum and Fe-containing phase. The typical aluminum alloy metallographic images are shown in Figure 2. In this figure, the two images above are two typical aluminum alloy metallographic images, and the two images below are their ground truths. In the labelled images, we set aluminum as the background and labelled it in black. At the same time, we set microstructure as the object or foreground. It should be noted that in the semantic segmentation task, we labelled different instances of the same object with the same color. As opposed to the semantic segmentation, in the instance segmentation, we labelled different instances of the same object with different colors. For example, we labelled each instance microstructure with a different color, as shown in Figure 2.

Figure 2. Aluminum alloy metallographic image and ground truth.

Implementation details: In the experiments, we used the cross validation method to ensure the accuracy of the evaluation results. We divided dataset D into five mutually exclusive subsets with the same size, $D = D_1 \cup D_2 \cup D_3 \cup D_4 \cup D_5$, and $D_i \cap D_j = \emptyset$ where $(i \neq j)$. We picked four subsets as the training set and the remaining subset as the test set, and then obtained five experimental results, as shown in the second to sixth columns in Tables 1–6. In the process of training Mask R-CNN, we used a fixed learning rate γ_1 of 0.01, a momentum ψ_1 of 0.9 in SGD with momentum algorithm which can not only improve the training stability and accelerate the training process to certain extent, but also gives a certain ability to get rid of local optimum solution. In FPN, the backbone is Resnet101, the pyramid size is 256 and the strides are 4, 8, 16, 32 and 64. In RPN, the anchor stride is 1, the non-max suppression threshold is 0.7, the number of train anchors per image is 256, anchor ratios are 0.5, 1 and 2, and anchor scales are 32, 64, 128, 256 and 512. In R-CNN and FCN, the pool size is 7, the mask pool size is 14, the train ROI per image is 200, the ROI positive ratio is 0.33, the mask shape is 28×28 , the detection max instance is 100, the detection min confidence is 0.7 and the detection threshold is 0.3. The training process reaches its convergence when the epoch number $\kappa_1 = 100$ and each epoch contains 100 steps. By using the coco pre-training model to initialize network parameters, we made the model converge faster and obtained a better result. Moreover, the L_2 regularization was used, and the regularization coefficient $\lambda = 0.0001$.

3.2. Evaluation Metrics

The aim of this section was to analyze the instance segmentation performance of the proposed method. Let $D = \{I_n, y_n\}$ represent the test dataset and f_θ represent our proposed method. We used six different evaluation metrics to evaluate this method, including Acc, Precision, Sn (Sensitivity, also named Recall), Sp (Specificity), IOU and F_1 . These metrics have been widely used in the literature. These six evaluation metrics are defined by the following formula:

(1) The Acc can be calculated by the following formula

$$Acc(f_\theta; D) = \frac{TP + TN}{TP + TN + FP + FN}, \quad (12)$$

where TN is the number of true negatives, TP is the number of true positives, FN is the number of false negatives and FP is the number of false positives.

(2) The Precision is defined by

$$Precision(f_\theta; D) = \frac{TP}{TP + FP}. \quad (13)$$

(3) The Sn is defined by

$$Sn(f_\theta; D) = \frac{TP}{TP + FN}. \quad (14)$$

(4) The Sp is defined by

$$Sp(f_\theta; D) = \frac{TN}{TN + FP}. \quad (15)$$

(5) The IOU is defined by

$$IOU(f_\theta; D) = \frac{TP}{TP + FN + FP}. \quad (16)$$

(6) The F_1 is defined by

$$F_1(f_\theta; D) = \frac{2TP}{2TP + FN + FP}. \quad (17)$$

3.3. Performance Comparison

(1) Quantitative performance comparison

In this experiment, we implemented five different mask loss functions on the basis of the proposed framework. For the easy description, we set $f_i (i = 1 : 5)$ to denote the i -th method. The first method f_1 uses loss function L_{BCE} . Similarly, f_2 uses loss function L_{DICE} , f_3 uses loss function L_{IoU} , f_4 uses loss function $L_{Tversky}$, and f_5 uses loss function L_{SS} . In addition, we show the quantitative comparison results of six different evaluation metrics, including Acc, Precision, Sn, Sp, IOU and F_1 .

In Table 1, the comparison results of Acc are shown. The first column denotes five different instance segmentation methods which were used in this experiment. The second to sixth columns show the results of five different cross experiments. The last column in the table shows the median of five experimental results. The first row denotes the experiment number. From these experimental results, we can see that the Acc of the five methods can be more than 99%. Therefore, from these experimental results, we can conclude that the five different instance segmentation methods implemented in our proposed framework can obtain a good segmentation performance. However, we can clearly observed that the distribution of this evaluation metric is very dense, so it is difficult to evaluate the performance of the metallographic

image segmentation method. The main reason for this problem is that the proportion of target and background in metallographic image is too small.

Table 1. The Acc comparison among the five different methods.

Acc	1	2	3	4	5	Median
f_1	0.999555	0.999433	0.999463	0.999551	0.999599	0.999551
f_2	0.999551	0.999364	0.999495	0.999506	0.999525	0.999506
f_3	0.999519	0.999366	0.999506	0.999551	0.999519	0.999519
f_4	0.999436	0.999277	0.999411	0.999532	0.999505	0.999436
f_5	0.999471	0.999270	0.999450	0.999535	0.999546	0.999471

Similarly, the comparison results of Sp are shown in Table 2. From these experimental results, we can see that the Sp of the five methods can reach up to more than 99%. Similarly to the Acc metric, this evaluation metric is also disturbed by background, so it cannot effectively evaluate the performance of the metallographic image segmentation method.

Table 2. The Sp comparison among the five different methods.

SP	1	2	3	4	5	Median
f_1	0.999732	0.999683	0.999656	0.999882	0.999833	0.999732
f_2	0.999723	0.999609	0.999708	0.999834	0.999760	0.999723
f_3	0.999700	0.999611	0.999714	0.999837	0.999748	0.999714
f_4	0.999601	0.999503	0.999579	0.999759	0.999706	0.999601
f_5	0.999628	0.999494	0.999633	0.999757	0.999743	0.999633

The comparison results of Precision are shown in Table 3. From these experimental results, we can see that the precision of the five methods can reach more than 60%. Therefore, from these experimental results, we can see that the proposed five different instance segmentation methods can obtain satisfactory segmentation results. This metric can avoid the interference caused by large background, so it is more suitable to evaluate the performance of the metallographic image segmentation method.

Table 3. The precision comparison among the five different methods.

Precision	1	2	3	4	5	Median
f_1	0.649482	0.653678	0.663011	0.757276	0.764655	0.663011
f_2	0.646859	0.642432	0.659344	0.740331	0.755242	0.659344
f_3	0.645198	0.648885	0.676730	0.747458	0.736909	0.676730
f_4	0.608518	0.611837	0.633345	0.725710	0.711829	0.633345
f_5	0.625438	0.590931	0.634894	0.739520	0.725018	0.634894

Similarly, the comparison results of Sn, IOU and F1 are shown in Tables 4–6, respectively. We can see that the Sn of the five methods can reach more than 62%, the IOU of the five methods can reach more than 53%, and the F1 of five methods can reach more than 61%. These three metrics can also effectively evaluate the performance of the metallographic image segmentation method. These experimental results verify the effectiveness of the proposed framework.

Table 4. The Sn comparison among the five different methods.

SN	1	2	3	4	5	Median
f_1	0.630825	0.639066	0.677610	0.669518	0.711701	0.669518
f_2	0.637324	0.633961	0.685101	0.677832	0.696087	0.677832
f_3	0.624957	0.631685	0.658808	0.689806	0.695827	0.658808
f_4	0.641941	0.661190	0.690875	0.748032	0.733668	0.690875
f_5	0.667652	0.641828	0.670456	0.746207	0.756356	0.670456

Table 5. The IOU comparison among the five different methods.

IOU	1	2	3	4	5	Median
f_1	0.562809	0.564019	0.590477	0.618716	0.648613	0.590477
f_2	0.564172	0.557246	0.588999	0.618948	0.632580	0.588999
f_3	0.556383	0.559084	0.584960	0.628808	0.627891	0.584960
f_4	0.549827	0.554891	0.579945	0.646644	0.636055	0.579945
f_5	0.567711	0.537472	0.572862	0.651192	0.653169	0.572862

Table 6. THE F1 COMPARISON AMONG THE FIVE DIFFERENT METHODS.

F1	1	2	3	4	5	Median
f_1	0.636120	0.641956	0.666217	0.705793	0.733540	0.666217
f_2	0.638624	0.633869	0.667450	0.702997	0.720067	0.667450
f_3	0.631062	0.635787	0.663737	0.712935	0.711942	0.663737
f_4	0.621341	0.631540	0.657220	0.731635	0.718488	0.657220
f_5	0.642051	0.611533	0.648007	0.737506	0.736652	0.648007

To sum up, we can obtain the following conclusions: (1) Acc and Sp evaluation metrics cannot effectively evaluate the performance of the metallographic image segmentation method. Instead, we can use Precision, Sn, IOU and F1 four evaluation metrics to effectively evaluate the performance of the metallographic image segmentation method. (2) The segmentation framework proposed in this paper can complete the task of instance segmentation and the five different methods based on the framework can achieve a satisfactory segmentation performance.

In addition, we show the box plot of the experiment results obtained by the five methods for six evaluation metrics. As shown in Figures 3 and 4, the box plot is able to reflect the characteristics of the experimental data distribution. In the figures, from top to bottom, we can clearly observe the statistical characteristics of the experimental data, including upper extreme, upper quartile, median, lower quartile and lower extreme.

In Figure 3a, we show the box plot of Acc metric. The X-axis represents accuracy (Acc), and the Y-axis represents five different methods. From this figure, we can observe the following: (1) All median values are located between 0.999436 and 0.999551 and the length of the interval is less than 0.000115. (2) The experimental results are mainly distributed between 0.999270 and 0.999599 and the length of the interval is less than 0.000163. Therefore, we can conclude that these five methods can obtain a good segmentation performance. However, the performance of the segmentation algorithm cannot be evaluated effectively because the distribution of the metric values is too centralized.

In Figure 3b, we show the box plot of the specificity metric. The X-axis represents specificity (Sp) and the Y-axis represents the five different methods. From the figure, we can observe: (1) All median values are concentrated between 0.999601 and 0.999732, and the length of the interval is less than 0.000131. (2) The

experimental results are mainly distributed between 0.999494 and 0.999882, and the length of the interval is less than 0.000388. Similarly to Acc, the distribution of metric value is too centralized, so it is not suitable to evaluate the performance of the segmentation method of the aluminum alloy metallographic image.

Figure 3. The box plot of experimental results obtained by the five methods: (a) Accuracy; (b) Specificity; (c) Precision.

In Figure 3c, we show the box plot of the precision metric. The X-axis represents precision, and the Y-axis represents the five different methods. From the figure, we can observe the following: (1) All median values are concentrated between 0.633345 and 0.676730, and the length of the interval is less than 0.043385. (2) The experimental results are mainly distributed between 0.590931 and 0.764655, and the length of the interval is less than 0.173724. Therefore, we can conclude that these five methods can get good segmentation performance and that this metric can be used to evaluate the performance of the aluminum alloy metallographic image segmentation method.

Figure 4a shows the box plot of the sensitivity metric. We can observe: (1) All median values are concentrated between 0.658808 and 0.690875, and the length of the interval is less than 0.032067. (2) The experimental results are mainly distributed between 0.624957 and 0.756356, and the length of the interval is less than 0.131399. Therefore, we can conclude that the proposed segmentation framework of the aluminum alloy metallographic image has a satisfactory robustness.

Figure 4. The box plot of experimental results obtained by the five methods: (a) Sensitivity; (b) IOU; (c) F1 Measure.

In Figure 4b, we show the box plot of the IOU metric. From the figure, we can observe: (1) All median values are concentrated between 0.572862 and 0.590477, and the length of the interval is less than 0.017615. (2) The experimental results are mainly distributed between 0.537472 and 0.653169, and the length of the interval is less than 0.115697. We can find that these five methods can obtain a good segmentation performance.

In Figure 4c, we show the box plot of F1 metric. From this figure, we can observe: (1) All median values are concentrated between 0.648007 and 0.663737, and the length of the interval is less than 0.01573. (2) The experimental results are mainly distributed between 0.611533 and 0.737506, and the length of the interval is less than 0.125973. We can find that these five methods can get good segmentation performance. Similarly, we can conclude that the proposed segmentation framework of aluminum alloy metallographic image has satisfactory robustness.

In summary, we can make the following conclusions: (1) The proposed segmentation framework of a aluminum alloy metallographic image can effectively achieve the task of instance segmentation and five different methods based on this framework can obtain a satisfactory segmentation result. (2) The distributions of Acc and specificity metric values are too centralized so they are not suitable to evaluate the performance of the aluminum alloy metallographic image segmentation method. Instead, the Precision, Sn, IOU and F1 four evaluation metrics are suitable to evaluate the performance of segmentation method of the aluminum alloy metallographic image.

(2) Qualitative performance comparison

To further evaluate the proposed instance segmentation framework of metallographic image, Figure 5 shows the qualitative performance comparison among the five different methods implemented in this framework, including the binary cross-entropy loss (f_1), Dice loss (f_2), IoU loss (f_3), Tversky loss (f_4) and SS loss (f_5). Three original metallographic images are shown in Figure 5a and their corresponding ground truth results are shown in Figure 5b. In the ground truth, the Aluminum is represented in black, and the other instances of microstructure are represented in different colors. The instance segmentation results of the five different methods are shown in Figure 5c–g. From these results, we can observe that these five methods can accomplish the task of instance segmentation of microstructure in a metallographic image.

Figure 5. Qualitative performance comparison among the five different methods: (a) original image; (b) Ground truth; (c) Binary cross-entropy loss f_1 ; (d) Dice loss f_2 ; (e) IoU loss f_3 ; (f) Tversky loss f_4 ; (g) SS loss f_5 .

For convenient observation, Figure 6 shows the local enlarged maps of instance segmentation results obtained by the five proposed methods, including the binary cross-entropy loss (f_1), Dice loss (f_2), IoU loss (f_3), Tversky loss (f_4) and SS loss (f_5). The original local enlarged map is shown in Figure 6a and the corresponding ground truth is shown in Figure 6b.

In these local enlarged maps, we can see two instances of the Fe-containing phase. As opposed to semantic segmentation, the instance segmentation method proposed in this paper can not only segment different microstructures, but also segment different instances in the microstructure. As shown in Figure 6,

we segmented different instances of the Fe-containing phase and marked them with different colors. From these experimental results, we can conclude that the instance segmentation methods proposed in this paper can achieve a satisfactory segmentation performance.

Figure 6. The local enlarged maps of instance segmentation results obtained by different methods (the “fe 1.000” inside the picture means the probability of this region being classified as Fe-containing phase is 1.000): (a) original image; (b) Ground truth; (c) Binary cross-entropy loss f_1 ; (d) Dice loss f_2 ; (e) IoU loss f_3 ; (f) Tversky loss f_4 ; (g) SS loss f_5 .

The proposed instance segmentation framework can automatically realize the representation learning of different metal alloys by using the given samples. We do not need to design features manually for different segmentation objects. Therefore, this instance segmentation framework can also be used for other types of metal alloys. In addition, for the proposed instance segmentation framework, the discriminability of the object instance has an important impact on the segmentation performance. When the phases of a given metal alloy have a stronger discriminability, this framework can obtain better instance segmentation results.

3.4. Convergence Analysis

The aim of this experiment was to analyze the convergence of the proposed framework. The loss curves with time for five different methods are shown in Figure 7. From the Figure 7, we can observe that these five methods can converge and the methods f_5 have a better convergence speed than the other methods. This verifies the convergence of the proposed instance segmentation framework.

Figure 7. The loss curves for five different methods.

4. Conclusions

A new metallographic image segmentation framework is proposed, which can achieve the task of automatic instance segmentation for aluminum alloy metallographic images. This segmentation framework provides a powerful tool for the quantitative analysis of metallographic images. Based on this proposed framework, we implemented five different instance segmentation methods by using different loss functions. A large number of experimental results verified the effectiveness of the proposed method. In order to evaluate the performance of the metallographic image segmentation method effectively, we compared and analyzed six typical segmentation performance evaluation metrics, including Accuracy, Precision, Sensitivity, Specificity, IOU and F_1 measure. From the experimental results, we found that Precision, Specificity, IOU and F_1 measure can effectively evaluate the performance of metallographic image segmentation methods. In the future, we plan to investigate the use of weakly supervised learning methods for dealing with the problem of insufficiency in high-quality hand-labeled data samples.

Author Contributions: Conceptualization, D.C. and D.G.; methodology, D.C. and D.G.; software, D.G.; validation, D.G. and F.L.; formal analysis, D.C.; investigation, D.C. and D.G.; resources, S.L.; data curation, D.G. and F.L.; writing—original draft preparation, D.C.; writing—review and editing, D.C. and D.G.; visualization, D.G.; supervision, S.L. and F.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by the National Key R&D Program of China under Grant 2017YFB0306400 and National Natural Science Foundation of China under Grant 61773104.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Heinz, A.; Haszler, A.; Keidel, C.; Moldenhauer, S.; Benedictus, R.; Miller, W. Recent development in aluminium alloys for aerospace applications. *Mater. Sci. Eng. A* **2000**, *280*, 102–107. [[CrossRef](#)]
2. Hirsch, J.; Al-Samman, T. Superior light metals by texture engineering: Optimized aluminum and magnesium alloys for automotive applications. *Acta Mater.* **2013**, *61*, 818–843. [[CrossRef](#)]
3. Martin, J.H.; Yahata, B.D.; Hundley, J.M.; Mayer, J.A.; Schaedler, T.A.; Pollock, T.M. 3D printing of high-strength aluminium alloys. *Nature* **2017**, *549*, 365. [[CrossRef](#)] [[PubMed](#)]
4. Zhang, J.; Song, B.; Wei, Q.; Bourell, D.; Shi, Y. A review of selective laser melting of aluminum alloys: Processing, microstructure, property and developing trends. *J. Mater. Sci. Technol.* **2019**, *35*, 270–284. [[CrossRef](#)]
5. Roy, N.; Samuel, A.; Samuel, F. Porosity formation in Al-9 Wt pct Si-3 Wt pct Cu alloy systems: Metallographic observations. *Metall. Mater. Trans. A* **1996**, *27*, 415–429. [[CrossRef](#)]
6. Rajasekhara, K.; Harendranath, C.; Raman, R.; Kulkarni, S. Microstructural evolution during solidification of austenitic stainless steel weld metals: A color metallographic and electron microprobe analysis study. *Mater. Charact.* **1997**, *38*, 53–65. [[CrossRef](#)]
7. Girault, E.; Jacques, P.; Harlet, P.; Mols, K.; Van Humbeeck, J.; Aernoudt, E.; Delannay, F. Metallographic methods for revealing the multiphase microstructure of TRIP-assisted steels. *Mater. Charact.* **1998**, *40*, 111–118. [[CrossRef](#)]
8. Rohatgi, A.; Vecchio, K.; Gray Iii, G. A metallographic and quantitative analysis of the influence of stacking fault energy on shock-hardening in Cu and Cu–Al alloys. *Acta Mater.* **2001**, *49*, 427–438. [[CrossRef](#)]
9. Moreira, F.; Xavier, F.; Gomes, S.; Santos, J.; Freitas, F.; Freitas, R. New analysis method application in metallographic images through the construction of mosaics via speeded up robust features and scale invariant feature transform. *Materials* **2015**, *8*, 3864–3882.
10. Povstyanoi, O.Y.; Sychuk, V.; McMillan, A.; Zabolotnyi, O. Metallographic analysis and microstructural image processing of sandblasting nozzles produced by powder metallurgy methods. *Powder Metall. Metal Ceram.* **2015**, *54*, 234–240. [[CrossRef](#)]
11. Chowdhury, A.; Kautz, E.; Yener, B.; Lewis, D. Image driven machine learning methods for microstructure recognition. *Comput. Mater. Sci.* **2016**, *123*, 176–187. [[CrossRef](#)]

12. Campbell, A.; Murray, P.; Yakushina, E.; Marshall, S.; Ion, W. New methods for automatic quantification of microstructural features using digital image processing. *Mater. Design* **2018**, *141*, 395–406. [\[CrossRef\]](#)
13. Zhenying, X.; Jiandong, Z.; Qi, Z.; Yamba, P. Algorithm based on regional separation for automatic grain boundary extraction using improved mean shift method. *Surf. Topogr. Metrol. Prop.* **2018**, *6*, 025001. [\[CrossRef\]](#)
14. Journaux, S.; Gouton, P.; Paindavoine, M.; Thauvin, G. Evaluating creep in metals by grain boundary extraction using directional wavelets and mathematical morphology. *Revue de Métall. Int. J. Metall.* **2001**, *98*, 485–499. [\[CrossRef\]](#)
15. Sun, Q.D.; Gao, S.F.; Huang, J.W.; Chen, W. Metallographical Image Segmentation and Compression. In *Applied Mechanics and Materials*; Trans Tech Publications Ltd.: Bäch SZ, Switzerland, 2012; Volume 152, pp. 276–280.
16. Simmons, J.; Przybyla, C.; Bricker, S.; Kim, D.W.; Comer, M. Physics of MRF regularization for segmentation of materials microstructure images. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 4882–4886.
17. Cheng, H.C.; Cardone, A.; Varshney, A. Interactive exploration of microstructural features in gigapixel microscopy images. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 335–339.
18. Chen, L.; Han, Y.; Cui, B.; Guan, Y.; Luo, Y. Two-dimensional fuzzy clustering algorithm (2DFCM) for metallographic image segmentation based on spatial information. In Proceedings of the 2015 2nd International Conference on Information Science and Control Engineering, Shanghai, China, 24–26 April 2015; pp. 519–521.
19. de Albuquerque, V.H.C.; de Alexandria, A.R.; Cortez, P.C.; Tavares, J.M.R. Evaluation of multilayer perceptron and self-organizing map neural network topologies applied on microstructure segmentation from metallographic images. *NDT E Int.* **2009**, *42*, 644–651. [\[CrossRef\]](#)
20. Bulgarevich, D.S.; Tsukamoto, S.; Kasuya, T.; Demura, M.; Watanabe, M. Pattern recognition with machine learning on optical microscopy images of typical metallurgical microstructures. *Sci. Rep.* **2018**, *8*, 2078. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Papa, J.P.; Nakamura, R.Y.; De Albuquerque, V.H.C.; Falcão, A.X.; Tavares, J.M.R. Computer techniques towards the automatic characterization of graphite particles in metallographic images of industrial materials. *Expert Syst. Appl.* **2013**, *40*, 590–597. [\[CrossRef\]](#)
22. de Albuquerque, V.H.C.; Silva, C.C.; Menezes, T.I.D.S.; Farias, J.P.; Tavares, J.M.R. Automatic evaluation of nickel alloy secondary phases from SEM images. *Microsc. Res. Tech.* **2011**, *74*, 36–46. [\[CrossRef\]](#)
23. DeCost, B.L.; Holm, E.A. A computer vision approach for automated analysis and classification of microstructural image data. *Comput. Mater. Sci.* **2015**, *110*, 126–133. [\[CrossRef\]](#)
24. Gola, J.; Britz, D.; Staudt, T.; Winter, M.; Schneider, A.S.; Ludovici, M.; Mücklich, F. Advanced microstructure classification by data mining methods. *Comput. Mater. Sci.* **2018**, *148*, 324–335. [\[CrossRef\]](#)
25. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
27. Sanchez-Lengeling, B.; Aspuru-Guzik, A. Inverse molecular design using machine learning: Generative models for matter engineering. *Science* **2018**, *361*, 360–365. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Azimi, S.M.; Britz, D.; Engstler, M.; Fritz, M.; Mücklich, F. Advanced steel microstructural classification by deep learning methods. *Sci. Rep.* **2018**, *8*, 2128. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Ma, B.; Ban, X.; Huang, H.; Chen, Y.; Liu, W.; Zhi, Y. Deep learning-based image segmentation for al-la alloy microscopic images. *Symmetry* **2018**, *10*, 107. [\[CrossRef\]](#)
30. Chen, H.; Qi, X.; Yu, L.; Dou, Q.; Qin, J.; Heng, P.A. DCAN: Deep contour-aware networks for object instance segmentation from histology images. *Med. Image Anal.* **2017**, *36*, 135–146. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Yi, J.; Wu, P.; Jiang, M.; Huang, Q.; Hoepfner, D.J.; Metaxas, D.N. Attentive neural cell instance segmentation. *Med. Image Anal.* **2019**, *55*, 228–240. [\[CrossRef\]](#)
32. de Bel, T.; Hermesen, M.; Litjens, G.; van der Laak, J. Structure Instance Segmentation in Renal Tissue: A Case Study on Tubular Immune Cell Detection. In *Computational Pathology and Ophthalmic Medical Image Analysis*; Springer: Berlin, Germany, 2018; pp. 112–119.

33. Guerrero-Pena, F.A.; Fernandez, P.D.M.; Ren, T.I.; Yui, M.; Rothenberg, E.; Cunha, A. Multiclass weighted loss for instance segmentation of cluttered cells. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 2451–2455.
34. Mou, L.; Zhu, X.X. Vehicle instance segmentation from aerial image and video using a multitask learning residual fully convolutional network. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 6699–6711. [[CrossRef](#)]
35. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2961–2969.
36. Milletari, F.; Navab, N.; Ahmadi, S.A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; pp. 565–571.
37. Rahman, M.A.; Wang, Y. Optimizing intersection-over-union in deep neural networks for image segmentation. In Proceedings of the International Symposium on Visual Computing, Las Vegas, NV, USA, 12–14 December 2016; pp. 234–244.
38. Salehi, S.S.M.; Erdogmus, D.; Gholipour, A. Tversky loss function for image segmentation using 3D fully convolutional deep networks. In Proceedings of the International Workshop on Machine Learning in Medical Imaging, Quebec, QC, Canada, 10 September 2017; pp. 379–387.
39. Brosch, T.; Yoo, Y.; Tang, L.Y.; Li, D.K.; Traboulsee, A.; Tam, R. Deep convolutional encoder networks for multiple sclerosis lesion segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 3–11.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).