

Article

Optimizable Image Segmentation Method with Superpixels and Feature Migration for Aerospace Structures

Chengwei Fei ^{1,2} , Jiongran Wen ¹ , Lei Han ^{1,3,*}, Bo Huang ² and Cheng Yan ⁴¹ Department of Aeronautics and Astronautics, Fudan University, Shanghai 200433, China² Shanghai Key Laboratory of Aircraft Engine Digital Twin, Shanghai 200241, China³ Department of Materials Science and Engineering, Pohang University of Science and Technology, Pohang 37673, Korea⁴ School of Aerospace Engineering, Xiamen University, Xiamen 361102, China

* Correspondence: han_lei@fudan.edu.cn

Abstract: The lack of high-quality, highly specialized labeled images, and the expensive annotation cost are always critical issues in the image segmentation field. However, most of the present methods, such as deep learning, generally require plenty of train cost and high-quality datasets. Therefore, an optimizable image segmentation method (OISM) based on the simple linear iterative cluster (SLIC), feature migration model, and random forest (RF) classifier, is proposed for solving the small sample image segmentation problem. In the approach, the SLIC is used for extracting the image boundary by clustering, the Unet feature migration model is used to obtain multidimensional superpixels features, and the RF classifier is used for predicting and updating the image segmentation results. It is demonstrated that the proposed OISM has acceptable accuracy, and it retains better target boundary than improved Unet model. Furthermore, the OISM shows the potential for dealing with the fatigue image identification of turbine blades, which can also be a promising method for the effective image segmentation to reveal the microscopic damages and crack propagations of high-performance structures for aeroengine components.



Citation: Fei, C.; Wen, J.; Han, L.; Huang, B.; Yan, C. Optimizable Image Segmentation Method with Superpixels and Feature Migration for Aerospace Structures. *Aerospace* **2022**, *9*, 465. <https://doi.org/10.3390/aerospace9080465>

Academic Editor: Andrea Da-Ronch

Received: 9 July 2022

Accepted: 17 August 2022

Published: 21 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: semantic segmentation; small samples; SLIC; Unet; transfer learning; turbine blade images

1. Introduction

With the rapid development of artificial intelligence and computer vision technologies, Image Recognition (IR) has been applied in many frontier fields, such as the liver tumor segmentation and cell recognition in biomedicine, as well as the defect detection, microstructural analysis of key components in the aerospace field [1–6]. Therefore, it is vital to conduct high-precision target recognition for improving the measurement and statistics of important characteristics. Meanwhile, the accurate and reliable target segmentation is actually one of the major problems in the image recognition.

The current popular image segmentation methods can be roughly categorized into two groups [7,8]: handcrafted features, and Deep Learning (DL) method. Generally, the typical handcrafted feature methods mainly use the handcrafted low-level and middle-level features for segmentation [9]. (1) The low-level features include the color [10], Gray-Level Co-occurrence Matrix (GLCM) [11], Local Binary Patterns (LBP) [12], Gabor descriptors [13], and Scale-Invariant Feature Transform (SIFT) [14], etc. Although these features are easy to be comprehended and conducted, it is usually unable to accurately capture the high-level image semantic information. Thence, the accuracy of the low-level features for image segmentation is relatively low. (2) The middle-level features, which include the Spatial Pyramid Matching (SPM) [15], Spatial Co-occurrence Kernel (SCK) [16], Probabilistic Latent Semantic Analysis (PLSA) [17], and Fisher Kernel [18], etc., can be obtained by encoding the low-level features. Meanwhile, compared with the low-level features, the

middle-level features can have a higher accuracy. However, the methods using the middle-level features generally require clever design features or specific constraints to increase feature discrimination. Thus, many factors must be considered in feature design and their generalization is poor, especially in constructing middle-level features.

The deep learning technology has been widely applied in the field of image semantic segmentation [19]. For example, Shelhamer et al. [20] adopted the full convolutional network (FCN) technology to solve the pixel-level segmentation task, but this method is not ideal in the fine-grained segmentation, which is due to that the end-to-end single pixel classification training made much information in model pooling operation lost. Ronneberger et al. [21] proposed the Unet model through combining the deconvolution with feature layer by skip connections, and avoiding the image feature information lost in the multi-layer convolution. The Unet model is widely applied due to the advantages of simple frame, few parameters and strong optimizable ability. However, it is found that Unet models usually have the problem of weak edge segmentation break. Zhu et al. [22] combined the handcrafted features with high-level semantic features which are extracted by the deep CNNs to improve the performance of scene classification, while this method cannot be effectively trained and implemented in the end-to-end method. Contrast to the traditional methods, training an existing deep learning method can achieve the outstanding performance. Nevertheless, the deep learning image segmentation methods [23–25] generally require large training datasets, which makes them difficult to have the universal applicability in small samples.

Recently, some image segmentation algorithms began to regard pixels as the basic unit of image analysis, but most of them have ignored the spatial relationship between pixels, resulting in poor image boundary segmentation [26–28]. Meanwhile, the superpixels algorithm can fit the boundary of generated hyperparameter region to the edge of the object or background in the image, i.e., the segmentation effect of the target and the background turns to be good. Among them, the simple linear iterative clustering (SLIC) superpixels algorithm based on K-means has the advantages of fast convergence, stable segmentation, and wide application range [29]. In addition, the fusing superpixel and deep learning method also has been revealed to exhibit good image segmentation effects [30–32]. For instance, Kanezaki et al. [33] proposed an image segmentation method based on unsupervised algorithm, which is not enough robust and difficult to be applied directly. Lv et al. [34] put forward a method based on superpixels and the stacked contractive autoencoder in synthetic aperture radar images. Xiong et al. [35] raised a rice panicle segmentation algorithm based on the SLIC, CNN, and entropy rate superpixel optimization methods. Farhat et al. [36] proposed the hierarchical framework based on two-dimensional superpixels and deep learning algorithm. In general, the methods combined superpixels and deep learning have obtained a good performance, but they are not suitable for the small samples' image segmentation, and not improvable according to the prior knowledge.

The Random Forest (RF) method, as one of the primary machine learning methods, can overcome the over-fitting problem to obtain better prediction accuracy and generalization ability owing to the high efficiency and explainability for datasets [37], which makes it fairly promising in solving small sample classification problems [38–40]. Accordingly, to address the problem of image segmentation with few training samples and strong specialization, an optimizable image segmentation method (OISM) is proposed through uniting the Unet feature migration technology, SLIC superpixel algorithm and RF model together. In the method, the SLIC is used to deal with the spatial organization relationship of pixels, and to make the boundary of superpixel region well fit to the edge of image object or background. Furthermore, the improved Unet model is applied to extract the multi-dimensional feature information of input images and expand the super-pixel features. Therefore, the method can solve the problem with insufficient feature data in the superpixels algorithm. Then, the RF method is applied for modeling, predicting and updating the superpixels feature data, to improve the accuracy of image segmentation. In final, the proposed OISM is verified to be feasible and effective by applying in solving the engineering problems. Under

preserving the semantic information and boundaries, a single image can be segmented quickly and accurately by using the OISM.

To sum up, the paper is structured as follows. Section 2 introduces the theory and methods of OISM, which involves the image filtering process, superpixels partition, Unet and SLIC superpixels feature extraction algorithm. In Section 3, the cell Kaggle dataset is applied for verifying the tunable image segmentation method proposed in the paper. Section 4 focuses on the application of the proposed method in the fatigue damage image analysis of turbine blades. Furthermore, major conclusions are summarized at last.

2. Materials and Methods

The framework of the proposed OISM is displayed in Figure 1, which including (I) image filter processing, (II) extracting superpixels features based on the Unet, and (III) image recognition and updating according to the OISM.

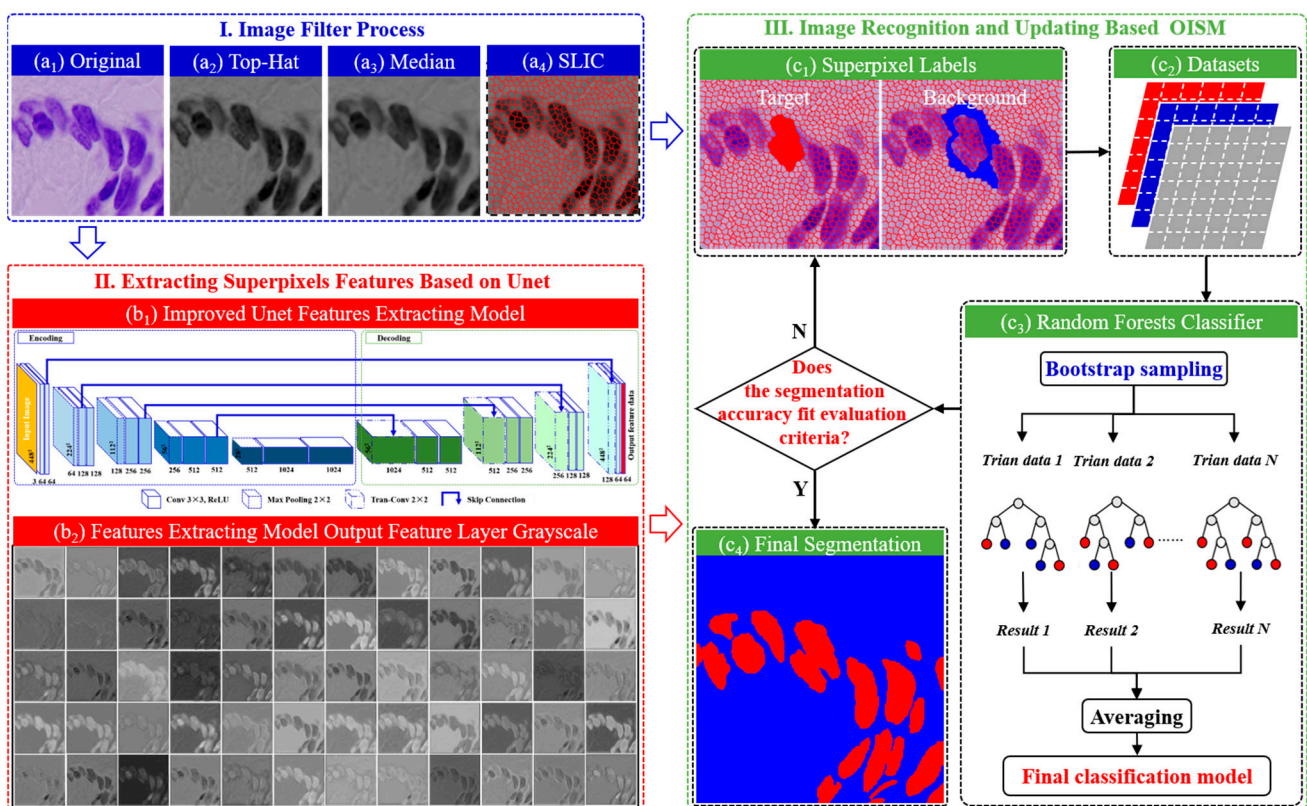


Figure 1. Framework of the proposed OISM with Unet feature migration model and SLIC superpixels. (I) the filter processes and SLIC superpixels results of the original image; (II) the improved Unet migration model and the multidimensional feature layers of input image; (III) the image segmentation and iteration process of the OISM based on RF.

(I) In the image filter processes obtain the superpixels' mask from the original image using the Top-Hat, Median filter, and SLIC superpixels algorithms;

(II) In the extracting process of superpixels features, obtain multidimensional feature layers of input images using the improved Unet model, then calculate the superpixels feature data by combining the superpixels mask in (I);

(III) In the image recognition and updating process, the RF classifier is trained by adding superpixels labels and feature datasets, then if the results are not good enough, the segmentation accuracy will be improved by new superpixels labels based on OISM.

2.1. Image Filter Processing and Superpixels Partition

As shown in Figure 1(a₂,a₃), the Top-Hat and Median filter are employed to preprocess the original images for preparing the superpixels division [41,42]. The Top-Hat filter is used to reduce the uneven illumination of the input image by the morphological opening operation. The opening operation is one of the morphology enhancement algorithms for gray image and often used to de-noise or smooth the image edges. It is defined as

$$FOB = F \ominus B \oplus B \tag{1}$$

where O is the opening operator; \ominus and \oplus are dilation and erosion operations; $F(x, y)$ is the input gray image; $B(x, y)$ is the structural element.

Figure 2c,d shows the results of dilation and erosion operations of the image (a) using the circular structuring element (b). The Top-Hat operation has some characteristics of high-pass filtering, which can subtract the dark background from original images to reduce uneven lighting. It is defined as:

$$H = F - (FOB) \tag{2}$$

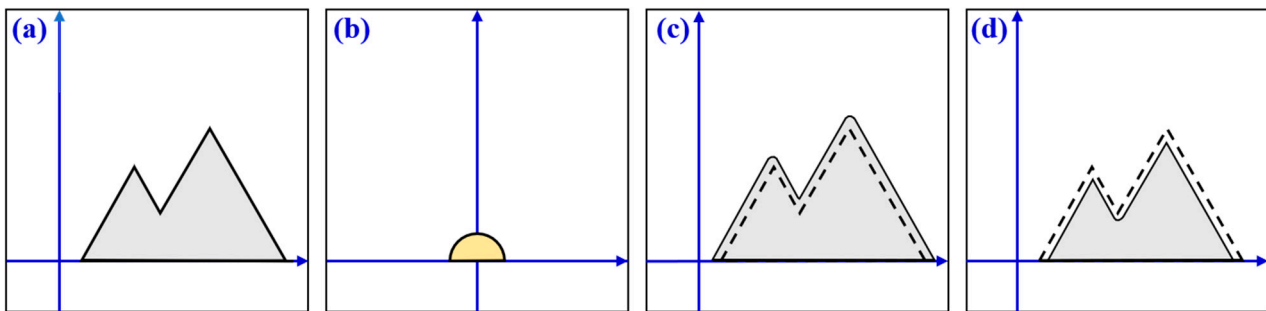


Figure 2. The description of dilation and erosion operations. (a) original image; (b) the structuring element; (c) the processing result of dilation operation; (d) the processing result of erosion operation.

The Median filter is one of Sequential statistical filters, which can replace the target pixel value with the median gray value of its neighborhood. It is applied to reduce the salt and pepper noise of images and explained by Figure 3.

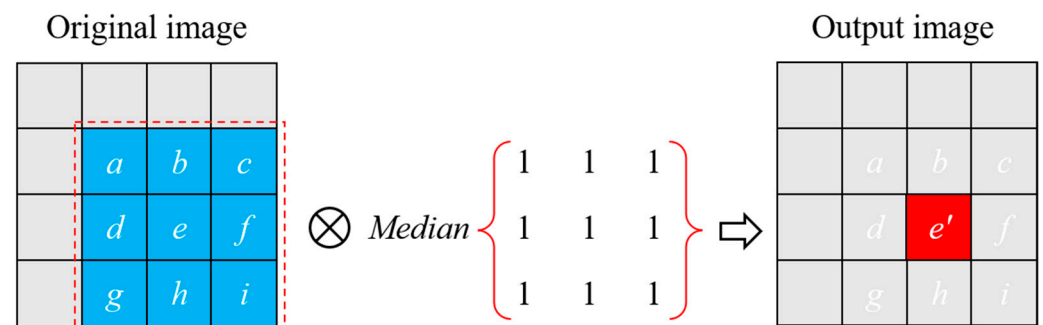


Figure 3. The principle of the Median filter.

\otimes is defined by

$$e' = median\{a, b, c, d, e, f, g, h, i\} \tag{3}$$

where a, b, \dots, i are the values of the pixel of the gray image; e' is the Median filter result of shape 3×3 ;

In addition, the SLIC algorithm is used to obtain the superpixels mask according to the color and spatial similarity of each pixel in the processed image. The distance, D , between pixels and clustering center in SLIC algorithm is defined by

$$D = \sqrt{\left(\frac{d_c}{N_c}\right)^2 + \left(\frac{d_s}{N_s}\right)^2}$$

$$\text{s.t.} \begin{cases} d_c = \sqrt{(l_i - l_j)^2 + (m_i - m_j)^2 + (n_i - n_j)^2} \\ d_s = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \end{cases} \quad (4)$$

in which d_c is color distance; d_s is spatial distance; N_c is the maximum color distance within [1, 40]; N_s is the maximum spatial distance; i is the i th pixel; j is the j th clustering center; $l, m,$ and n are the LAB color space values of pixels. After each iteration, the mean value of all pixel values in the cluster is used to update the cluster center. Figure 1(a₄) visually shows the boundary and shape of superpixels.

2.2. Superpixels Feature Extraction Algorithm with Improved Unet and SLIC

After the image filtering and SLIC superpixels algorithm processing, the Unet feature migration model is used to extract the superpixels feature dataset. As shown in Figure 4, the improved Unet model is applied to extract the multi-dimensional feature information of input images and enhance the super-pixel features. In the improved Unet model, we change the stride and the kernel size of padding of Conv2d to make the input image and output results to be the same in the Pytorch frame. The size of the input image is 448^2 and the number of the channels is 3. To obtain the multi-scale feature datasets, the same-level layers' features of encoding and decoding are combined in the skip connection. In addition, we clip the final layer of the improved Unet model to retain the 64-dimensional pixel feature matrix.

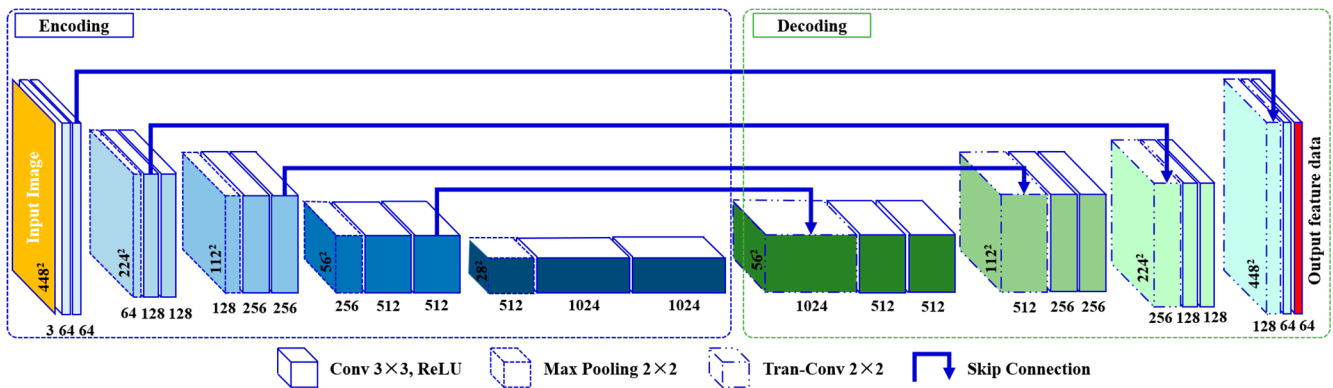


Figure 4. The improved Unet feature migration model.

Figure 5 shows the way of calculating the superpixels feature datasets in OISM. The 64-dimensional output feature data obtained by the improved Unet model is converted into 64 gray images, which have the same size as the original image. Through combining the superpixels mask and 64 gray images, each superpixel has 64 gray regions. The superpixels' features are defined by,

$$S_{r,q} = \frac{\sum_{t=1}^n P_{q,t}}{n} \quad (5)$$

where $S_{r,q}$ is the q -dimension feature of r th superpixel; n is the number of pixels contained in r th superpixel; $P_{q,t}$ is the t -dimension feature of q th pixel. Then, combining the superpixels' feature data to obtain 64-dimensional superpixels feature datasets.

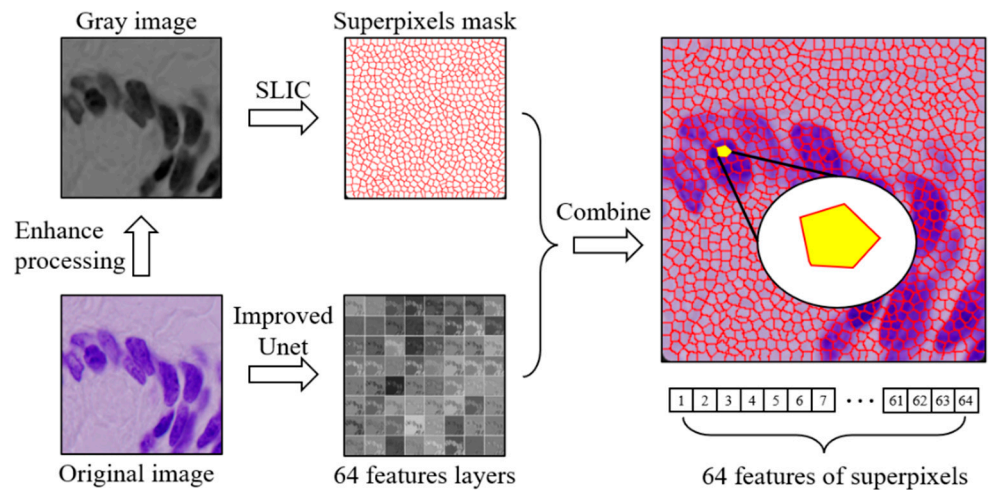


Figure 5. Superpixel feature dataset with the improved Unet model and SLIC algorithm.

2.3. Optimizable Image Segmentation Method

After acquiring the superpixels feature datasets, the OISM based on RF is proposed in Figure 6. In the method, the image segmentation is completed by

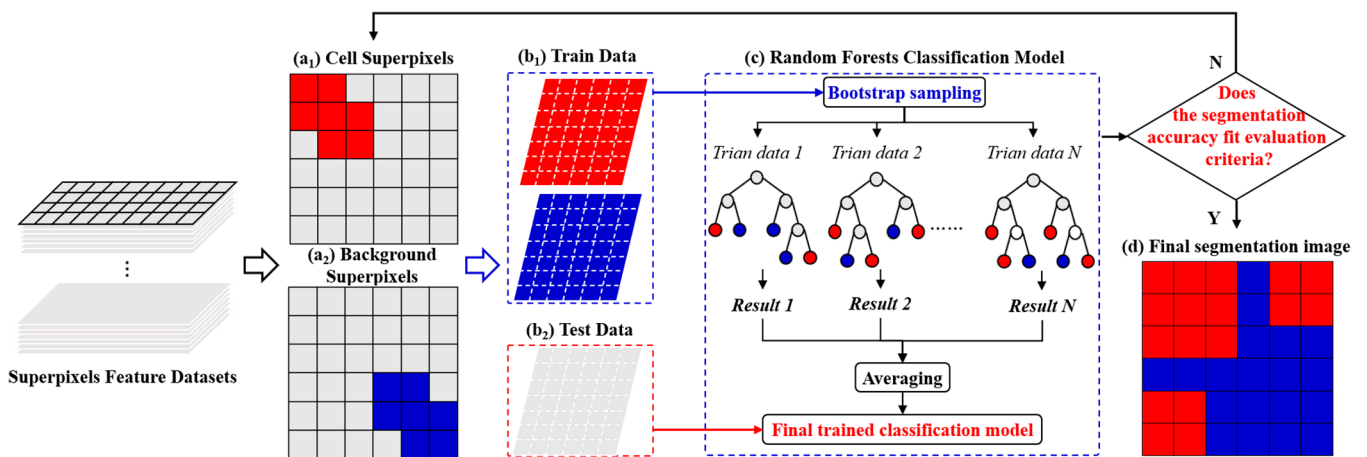


Figure 6. Flowchart of optimizable image semantic segmentation with RF classifier. (a₁) and (a₂) are the labeled cell and the background class superpixels of the image; (b₁) and (b₂) are the train data of labeled superpixels and test data of non-labeled superpixels of the image; (c) is the superpixels classifier based on the RF; (d) is the final segmentation image.

- (1) Label the interest superpixels to add the category, such as cell superpixels and background superpixels, according to human’s prior knowledge;
- (2) The superpixels feature data and labeled data are divided into the train datasets (labeled superpixels) and the test datasets (non-labeled superpixels);
- (3) Train the RF classification model using the train datasets, and predict the category of the test datasets in the final trained model;
- (4) If the evaluation indicators of the image segmentation are good, output the final segmentation map; Otherwise, return to Step (1) to add new superpixels labels.

To describe the number of labeled superpixels clearly, the proportion of labeled superpixels in the image, K is defined by

$$K = \frac{\sum_{r=1}^n A_r}{A} \tag{6}$$

where A_i is r th superpixel area; A is the total area of single image; n is the total number of superpixels.

To verify the effectiveness of the proposed OISM, the confusion matrix [43] is regarded as the index to evaluate the relationship between the predicted figure and the class figure. The confusion matrix is used to distinguish true positive (TP), true negative (TN), false positive (FP) and false negative (FN) value (Among them, TP is correctly predicted target pixels, TN is wrong predicted target pixels, FP is correctly predicted background pixels, and FN is wrong predicted background pixels), which are specifically defined in Table 1 and Figure 7.

Table 1. The definition of TP , TN , FP , and FN in the confusion matrix.

Ground Truth	Prediction	
	Positive	Negative
Positive	Ture Positive	False Negative
Negative	False Positive	Ture Negative

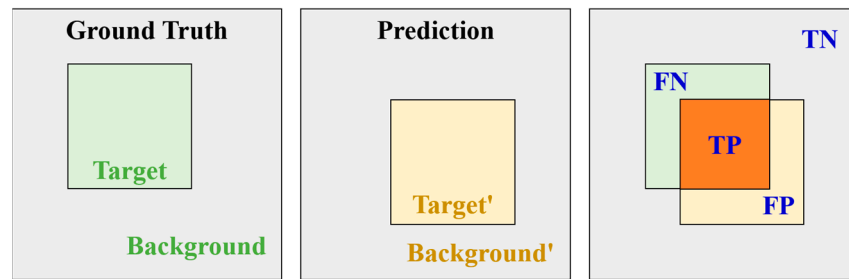


Figure 7. The schematic diagram of TN , TP , FN , and FP in the analysis of image segmentation.

In the comparative test, four evaluation indexes [44–47], i.e., pixel accuracy (PA), mean pixel accuracy (MPA), mean intersection union ($MIoU$), and frequency weighted intersection over union ($FWIoU$), are selected to evaluate the performance of the proposed method. The PA is the proportion of pixels with correct classification to the total pixels, indicating the proportion of the correct prediction values to the total predicted values. The PA in the confusion matrix is gained by,

$$PA = \frac{TP + TN}{TP + FP + FN + TN} \tag{7}$$

The MPA is the mean proportion of the number of correctly-classified pixels for each category to the number of predicted pixels for the same category. When $n+1$ is the total number of categories, the MPA is,

$$MPA = \frac{1}{n + 1} \frac{TP}{TP + FP} \tag{8}$$

The $MIoU$ is the standard measure of the accuracy of image segmentation algorithms. The larger the $MIoU$ value is, the higher the segmentation accuracy is. The $MIoU$ is computed by the average value of the ratios of the intersection between predicted region and real region to the union of predicted region and real region, i.e.,

$$MIoU = \frac{1}{n + 1} \sum_{i=1}^n \frac{TP}{TP + FP + FN} \tag{9}$$

The *FWIoU* is the improvement of the *MIoU*, in which the weight each category is selected by the frequency of its occurrence. The *FWIoU* is calculated by,

$$FMIOU = \frac{TP + FN}{TP + FP + TN + FN} \frac{TP}{TP + FP + FN} \quad (10)$$

3. Results and Discussion

To verify the feasibility and effectiveness of the proposed OISM, a comparative experiment is conducted with respect to the 2018 Kaggle cell dataset [48]. Meanwhile, the microscopic image identification of a fractured aeroengine turbine blade [49,50] is regarded as an example to validate the applicability of the proposed method. The Kaggle dataset is divided into two datasets, which contains 735 training and testing cell images, as shown in Figure 8. Wherein, one dataset is Unet model training dataset (Figure 8a), which is used to train and improve the Unet model. The other is the superpixels multi-dimensional feature dataset which is obtained from the developed OISM (Figure 8b). In the validation, the Torch API with Pytorch as the backend is adopted for the Unet training/testing. Furthermore, the platform is configured with the CUDA 11.2 and cuDNN 8.1.0, which is equipped with the Nvidia RTX3070 graphics card in Windows 11(Python = 3.8) and one i7-12700KF CPU/32 GB RAM as well.

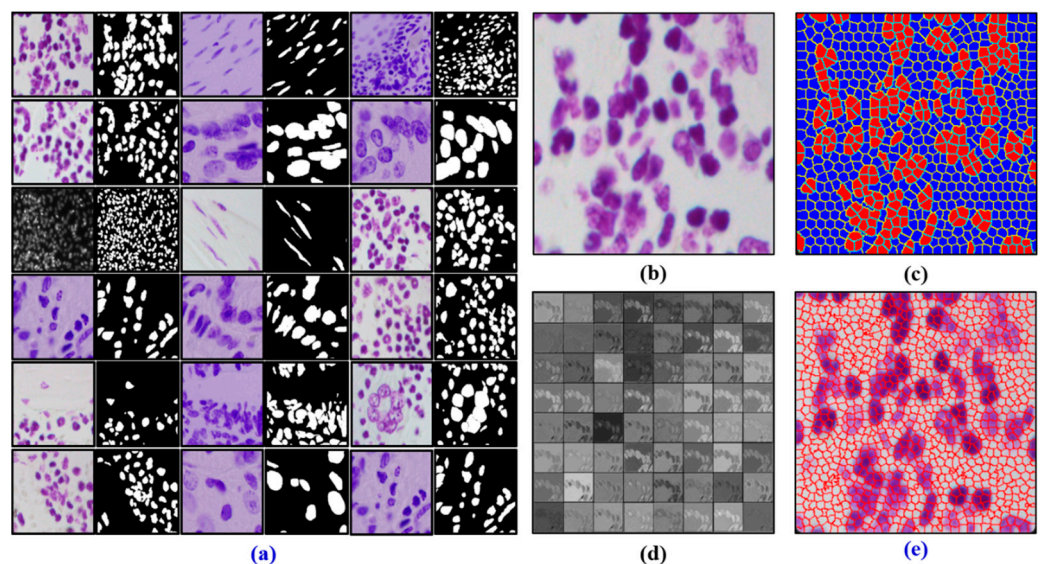


Figure 8. Image segmentation datasets of the Unet model and proposed OISM: (a) is the Unet model training dataset; (b) is the original image; (c) is labeled superpixels image; (d) is the improved Unet feature layers; (e) is the superpixels features dataset from the OISM.

3.1. Method Validation

The proposed OISM is verified based on the Kaggle dataset in the paper, specific steps of which are shown as follows:

(1) The Kaggle dataset is uniformly filtered, then it is used to train the improved Unet to obtain the feature extraction model. The average evaluation indexes *PA*, *MPA*, *MIoU*, and *FWIoU* of the model in the test dataset are, respectively, 0.853324, 0.823063, 0.782891, and 0.834420.

(2) The prediction image is preprocessed by the same filtering and SLIC superpixels algorithm, in order to acquire the 64-dimension superpixels feature dataset through combining the improved Unet model and Equation (5). Then, add the category labels to the training dataset and testing dataset with superpixels forms. Next, the training dataset is applied to train the RF classifier (model). Furthermore, the built RF model is utilized for predicting the unlabeled superpixels classification (testing dataset), and fusing the same class superpixels to obtain the segmentation images. Lastly, the misclassified superpixels is

corrected by adjusting the K value based on the Equation (6), then the accurate RF classifier is retrained with the corrected superpixels.

(3) To compare the segmentation effects of the Unet and the proposed OISM methods, the indexes in Equations (7)–(10) are used to evaluate images with different cell sizes, morphologies, and aggregation states. The corresponding results are listed in Table 2, in which the bold data indicate the best methods among different indicators. The segmentation effects of the Unet and the OISM methods for 5 typical test images and their intersection with the actual calibration data are illustrated in Figure 9.

Table 2. Comparison of the evaluation indexes between the Unet and the OISM methods with typical samples.

NO.	Method	PA	MPA	MIoU	FWIoU
a	Unet	0.834259	0.851848	0.679433	0.733431
	OISM ($K = 0.8\%$)	0.887360	0.902457	0.763478	0.808452
	OISM ($K = 5.0\%$)	0.957993	0.971772	0.899906	0.922028
b	Unet	0.838364	0.871663	0.673284	0.746117
	OISM ($K = 0.8\%$)	0.913635	0.943578	0.799415	0.852301
	OISM ($K = 5.0\%$)	0.951767	0.969375	0.876222	0.912522
c	Unet	0.762253	0.762783	0.576126	0.640748
	OISM ($K = 0.8\%$)	0.686020	0.776503	0.510159	0.548241
	OISM ($K = 5.0\%$)	0.941086	0.959623	0.865410	0.893312
d	Unet	0.834915	0.895797	0.617159	0.764119
	OISM ($K = 0.8\%$)	0.946335	0.964344	0.816673	0.908712
	OISM ($K = 5.0\%$)	0.979523	0.988267	0.917164	0.961958
e	Unet	0.806793	0.811859	0.632902	0.697995
	OISM ($K = 0.8\%$)	0.937119	0.958537	0.855316	0.887218
	OISM ($K = 5.0\%$)	0.957596	0.972039	0.897415	0.921520

In Table 2, most of the evaluation indexes with the OISM ($K = 0.8\%$) are better than those with the improved Unet, and the segmentation accuracy of the OISM is rising with the increasing K .

The dashed black boxes in Figure 9c show the segmentation accuracy inside different cells. For instance, compared with (6) in Figure 9c, there are some blue areas in (5), which means that the cells' features extracted from the improved Unet have the relatively poor effect. However, the OISM using the SLIC has exhibited a much better performance.

In addition, the dashed red boxes in Figure 9d show the segmentation accuracy outside the cells. Compared with (6) in Figure 9d, there are some pink areas in (5), which means the relatively poor ability of the target edges of the improved Unet. Although, the figure with OISM ($K = 0.8\%$) in (6) appears the same problem, the accuracy of segmentation is gradually improved as the increasing of K . It is because that the OISM can add some new labeled superpixels to improve the segmentation results. Thence, the OISM method developed in this paper can retain the segmentation boundary well and exhibits little misclassification rate, which can largely improve the segmentation accuracy compared with the Unet method.

Moreover, to analyze the proposed OISM more comprehensively, 20 images from the test datasets of the cells are randomly selected and calculated with the PA , MPA , $MIoU$, and $FWIoU$ compared with the improved Unet. A box plot is presented with the statistical results in Figure 10, which consists of the minimum (0th percentile, Q_0), the maximum (100th percentile, Q_4), the median (50th percentile, Q_2 , ★), first quartile (25th percentile, Q_1) and lower quartile (75th percentile, Q_3). In addition, the interquartile range (IQR), which is the distance between the upper and lower quartiles is defined by:

$$IQR = Q_3 - Q_1 \quad (11)$$

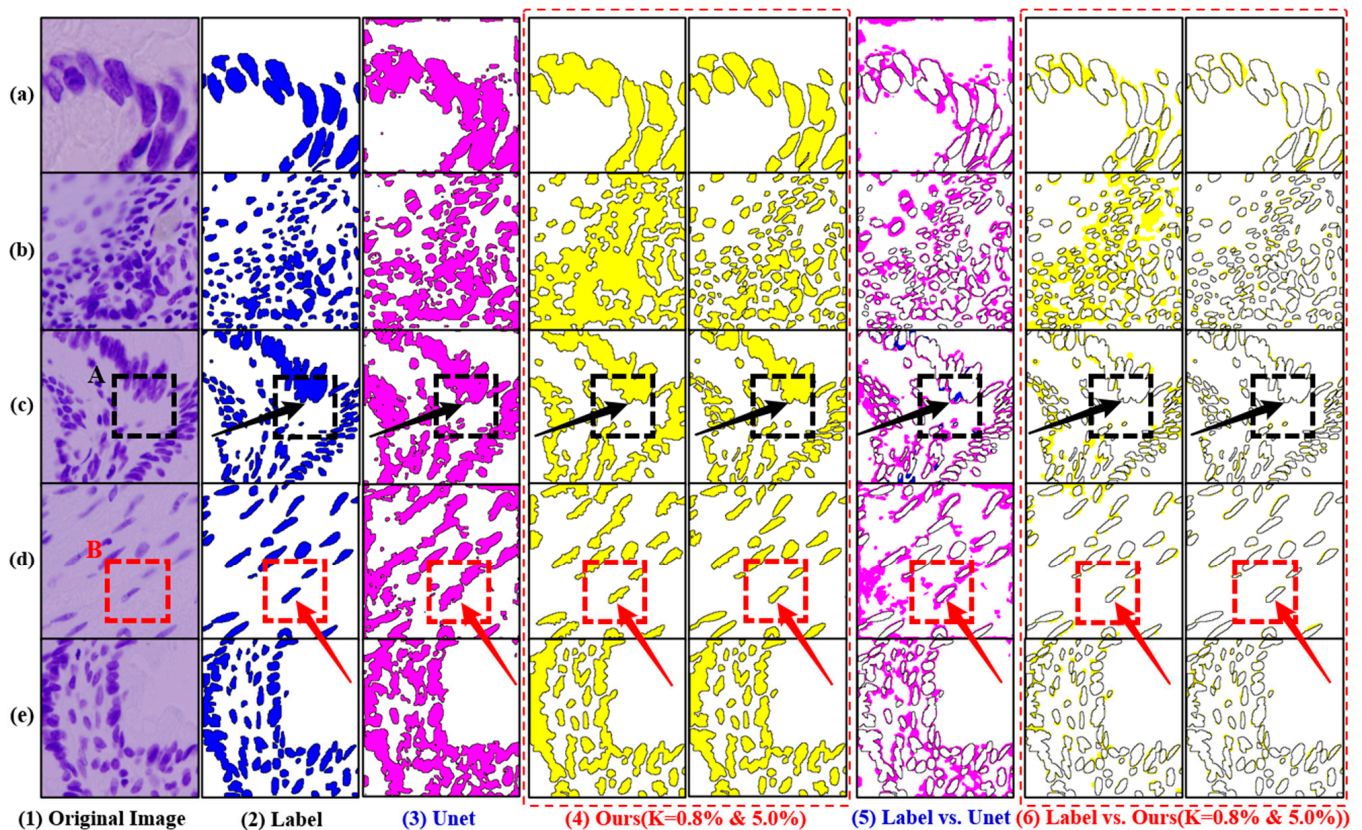


Figure 9. The results from the Unet and the OISM methods in typical samples: (1) (a–e) are five cell images selected for semantic segmentation; (2) denotes five true labeled images; (3) indicates the segmentation results of Unet; (4) is the segmentation results with the OISM at $K = 0.8\%$ and $K = 5.0\%$; (5) and (6) are the relationships among the true labeled images, Unet segmentation images, and the OISM images; A is the area of dashed red box where compare the results of inside cells segmentation; B is the area of dashed black box where compare the results of outside cells segmentation.

(1) Compared with the *IQR* values of *PA*, *MPA*, *MIoU*, and *FWIoU*, the data of the improved Unet is smaller than that of the OISM ($K = 0.8\%$), while it is bigger than that of the OISM ($K = 5.0\%$). Thus, it shows that the accuracy of the OISM is more concentrated with the increasing of K value.

(2) Compared with the median values (50th percentile) of four segmentation accuracy indicators of the improved Unet, the data of the OISM is relatively better except the *MIoU*. Owing to the different areas of cells and background, the accuracy weight of the background pixels is higher than the cells. Therefore, the *FWIoU* value calculated by the frequency weight of the occurrence is more reasonable than that of the *MIoU*.

(3) The *PA* of the OISM is concentrated between the median and the lower quartile, while the *PA* of the improved Unet is concentrated around the median. The *MPA* of the OISM is concentrated around the median, while the *MPA* of the improved Unet is concentrated between the median and the lower quartile. The larger weight of the background class impacts more on the *MPA* of the OISM. It shows that the background class prediction accuracy of the OISM is poor than the cell class, which makes the *MPA* non-concentrated on the lower quartile.

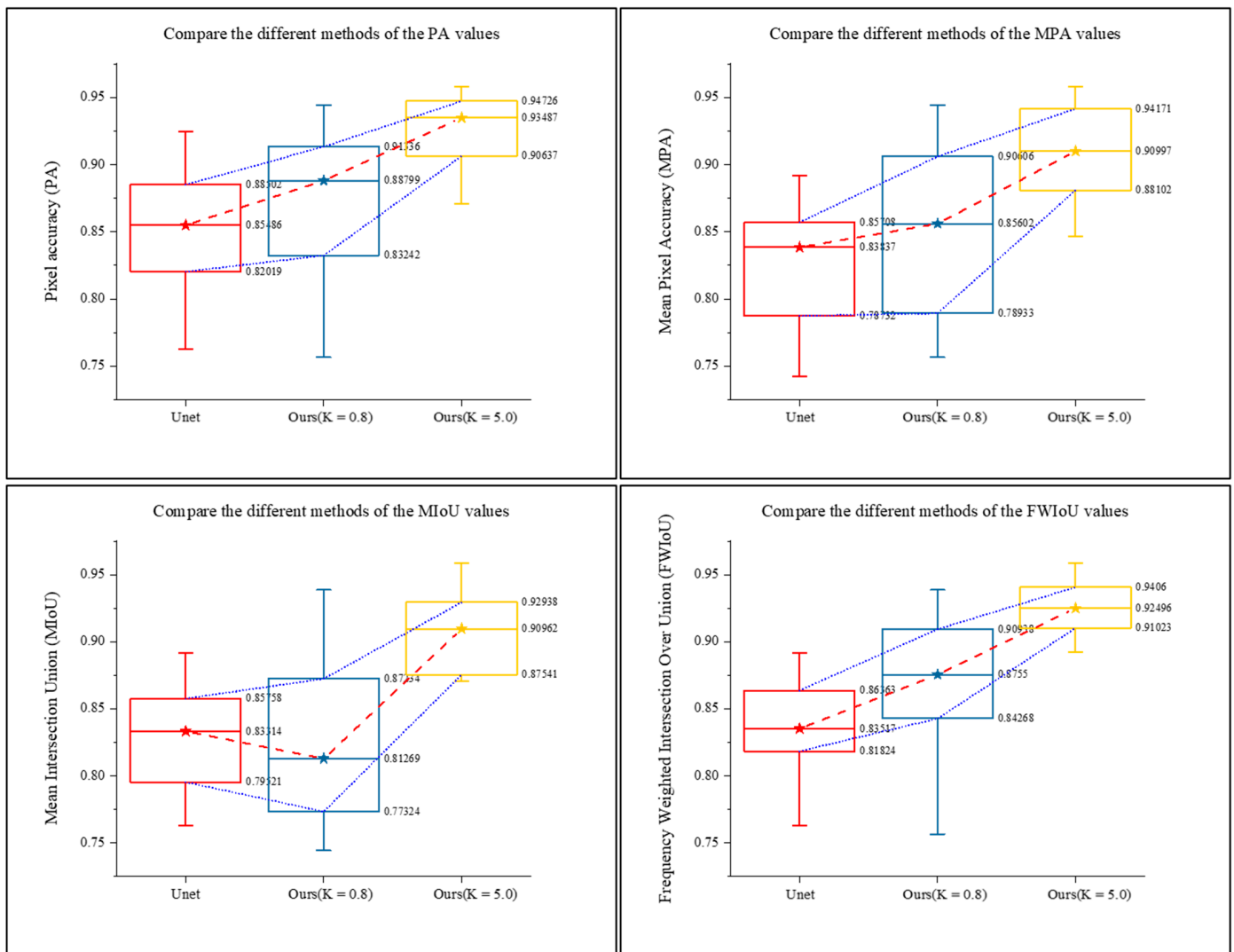


Figure 10. Image segmentation datasets for the verification of proposed OISM and Unet models (the star is the median value of the box plot).

3.2. Engineering Application in Turbine Blade Image Analysis

In Figure 11(1), the microscopic damage image analysis of the aeroengine turbine blades is conducted as well to verify the applicability and application potential of the proposed OISM. In the research, the cracks, voids, and microstructures of different turbine blades are recognized and tested, respectively. The corresponding results are shown in Figure 11 and Table 3. It is demonstrated that the proposed OISM can obtain the excellent segmentation results at a small K value for single image, and the segmentation accuracy increases with the increasing K value.

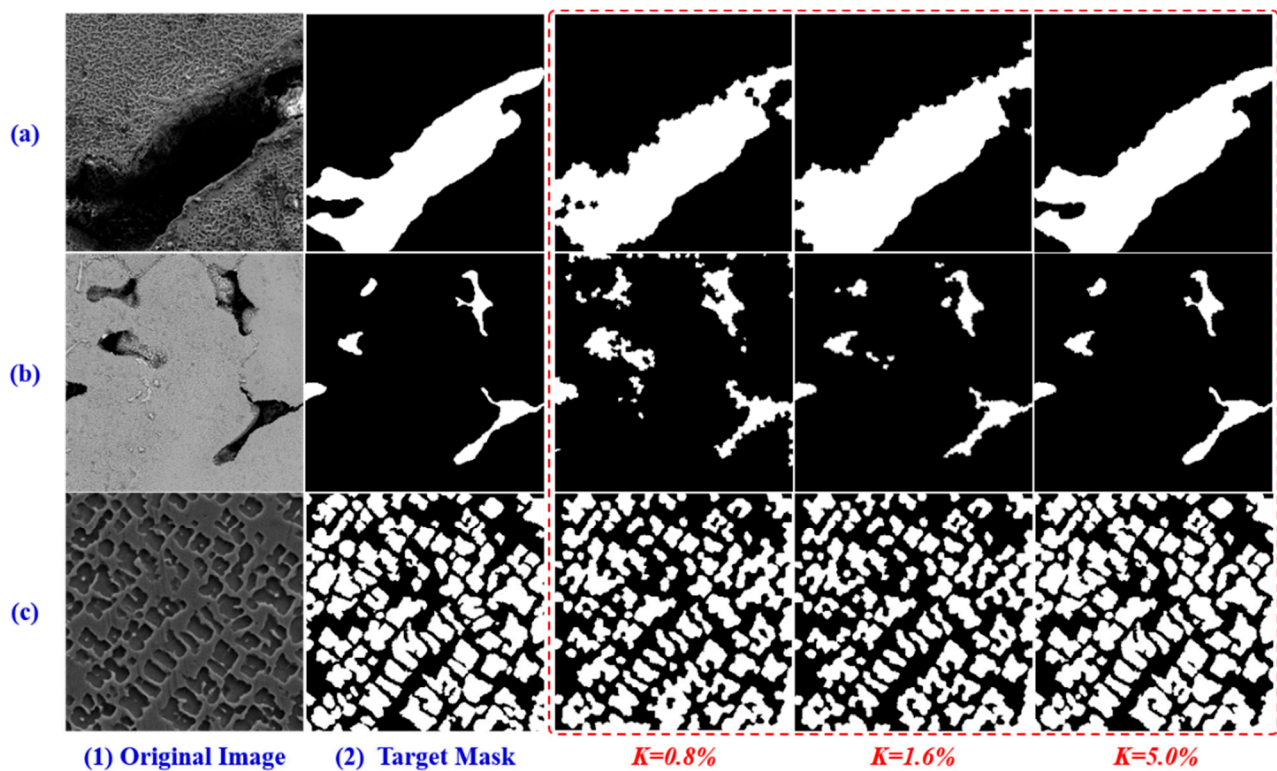


Figure 11. The identification of fatigue cracks, voids, and carbides in the turbine blades with the OISM. (a–c) are the images of crack, voids, and microstructure; (1) and (2) are the original images and the segmentation target mask of the fatigue test.

Table 3. Evaluation indexes of the image segmentation in different turbine blades with the OISM.

Type	K Value (%)	PA	MPA	MIoU	FWIoU
Crack	$K = 0.8$	0.874330	0.865002	0.762658	0.778716
	$K = 1.6$	0.875681	0.856561	0.761054	0.779182
	$K = 5.0$	0.947255	0.954521	0.894341	0.901013
Void	$K = 0.8$	0.942570	0.945785	0.862369	0.895216
	$K = 1.6$	0.932300	0.951606	0.844659	0.879264
	$K = 5.0$	0.981622	0.985598	0.952157	0.964483
Microstructure	$K = 0.8$	0.944552	0.899092	0.646619	0.922001
	$K = 1.6$	0.979505	0.893149	0.778808	0.964833
	$K = 5.0$	0.992388	0.971998	0.903307	0.985876

4. Conclusions

The paper is aimed to propose an optimizable image segmentation method (OISM) to improve the accuracy of image segmentation in structural health monitoring with small dataset and strong specialization, by integrating simple linear iterative cluster (SLIC) superpixels algorithm with K-means, improved Unet feature migration technology and the random forest (RF) classification method. The SLIC is to well fit the boundary of the superpixel region to the image edge of the object or background, by handling the spatial organization relationship of pixels. The improved Unet model is used for extracting the multidimensional features of superpixels by feature migration technology. The RF method is employed to build the model based on the superpixels feature data and update the model for image recognition, to improve the accuracy of image segmentation. The effectiveness and practicability of the proposed OISM are verified by two cases. Main conclusions are listed as follows:

(1) Through introducing the SLIC superpixels algorithm, the proposed OISM can preserve edge information and can address the problems of unclear boundary and misclassification of Unet segmentations;

(2) Compared with the Unet model, the proposed method has relatively higher segmentation accuracy as $K = 0.8\%$, and the accuracy gradually improves with the increase of K value;

(3) In light of the fatigue test image analysis of turbine blades, the developed OISM shows a good segmentation effect when $K = 5.0\%$, which further demonstrates the applicability and engineering application potential of the proposed method.

The efforts of this study provide a promising method for accurate image segmentation with small dataset and strong specialization, and develops a useful technology for the damage and crack propagation identification of aeroengine components in structural health monitoring. In the future work, the OISM will be applied in the fatigue and fracture performance research of aviation structural parts, which is quantitatively analyzed from multiple-scales by image segmentation technologies. Furthermore, unsupervised algorithms will be used to improve the efficiency of the OISM.

Author Contributions: Conceptualization, C.F., J.W. and L.H.; methodology, J.W. and C.F.; software, J.W.; validation, B.H. and J.W.; formal analysis, L.H., C.F. and J.W.; investigation, J.W. and C.F.; resources, L.H. and J.W.; data curation, J.W.; writing—original draft preparation, J.W., L.H. and C.F.; writing—review and editing, C.F., L.H. and J.W.; visualization, J.W., L.H. and C.Y.; supervision, C.Y., L.H. and C.F.; project administration, L.H. and C.F.; funding acquisition, C.F., L.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was co-funded by the National Natural Science Foundation of China (No. 51975124), China Postdoctoral Science Foundation (No. 2021M700783), China Postdoctoral Science Foundation (No. 2020M682584), and State Key Laboratory for Strength and Vibration of Mechanical Structures (No. SV2021-KF-05).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used to support the findings of this study are included within the article.

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

References

1. Li, W. Automatic segmentation of liver tumor in CT images with deep convolutional neural networks. *J. Comput. Commun.* **2015**, *3*, 146. [\[CrossRef\]](#)
2. Zhou, Z.-H.; Jiang, Y.; Yang, Y.-B.; Chen, S.-F. Lung cancer cell identification based on artificial neural network ensembles. *Artif. Intell. Med.* **2002**, *24*, 25–36. [\[CrossRef\]](#)
3. Adeniji, D.; Oligee, K.; Schoop, J. A Novel Approach for Real-Time Quality Monitoring in Machining of Aerospace Alloy through Acoustic Emission Signal Transformation for DNN. *J. Manuf. Mater. Process.* **2022**, *6*, 18. [\[CrossRef\]](#)
4. Fei, C.W.; Li, H.; Lu, C.; Han, L.; Keshtegar, B.; Taylan, O. Vectorial surrogate modeling method for multi-objective reliability design. *Appl. Math. Model.* **2022**, *109*, 1–20. [\[CrossRef\]](#)
5. Li, H.; Bu, S.; Wen, J.-R.; Fei, C.-W. Synthetical Modal Parameters Identification Method of Damped Oscillation Signals in Power System. *Appl. Sci.* **2022**, *12*, 4668. [\[CrossRef\]](#)
6. Fei, C.W.; Liu, H.T.; Rhea PLiem Choy, Y.S.; Han, L. Hierarchical model updating strategy of complex assembled structures with uncorrelated dynamic modes. *Chin. J. Aeronaut.* **2022**, *35*, 281–296. [\[CrossRef\]](#)
7. Cheng, G.; Han, J.; Lu, X. Remote sensing image scene classification: Benchmark and state of the art. *Proc. IEEE* **2017**, *105*, 1865–1883. [\[CrossRef\]](#)
8. Yang, Y.; Newsam, S. Bag-of-visual-words and spatial extensions for land-use classification. In Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, San Jose, CA, USA, 2–5 November 2010; pp. 270–279.
9. Yu, D.; Xu, Q.; Guo, H.; Zhao, C.; Lin, Y.; Li, D. An Efficient and Lightweight Convolutional Neural Network for Remote Sensing Image Scene Classification. *Sensors* **2020**, *20*, 1999. [\[CrossRef\]](#)
10. Swain, M.J.; Ballard, D.H. Color indexing. *Int. J. Comput. Vis.* **1991**, *7*, 11–32. [\[CrossRef\]](#)

11. Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *6*, 610–621. [[CrossRef](#)]
12. Bhagavathy, S.; Manjunath, B.S. Modeling and detection of geospatial objects using texture motifs. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 3706–3715. [[CrossRef](#)]
13. Risojević, V.; Momić, S.; Babić, Z. Gabor descriptors for aerial image classification. In Proceedings of the International Conference on Adaptive and Natural Computing Algorithms, Ljubljana, Slovenia, 14–16 April 2011; Springer: Berlin/Heidelberg, Germany, 2011; pp. 51–60.
14. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [[CrossRef](#)]
15. Lazebnik, S.; Schmid, C.; Ponce, J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 17–22 June 2006; Volume 2, pp. 2169–2178.
16. Yang, Y.; Newsam, S. Spatial pyramid co-occurrence for image classification. In Proceedings of the 2011 International Conference on Computer Vision, Washington, DC, USA, 6–13 November 2011; pp. 1465–1472.
17. Zhong, Y.; Cui, M.; Zhu, Q.; Zhang, L. Scene classification based on multifeature probabilistic latent semantic analysis for high spatial resolution remote sensing images. *J. Appl. Remote Sens.* **2015**, *9*, 095064. [[CrossRef](#)]
18. Zhao, B.; Zhong, Y.; Zhang, L.; Huang, B. The Fisher Kernel Coding Framework for High Spatial Resolution Scene Classification. *Remote Sens.* **2016**, *8*, 157. [[CrossRef](#)]
19. Minaee, S.; Boykov, Y.Y.; Porikli, F.; Plaza, A.; Kehtarnavaz, N.; Terzopoulos, D. Image segmentation using deep learning: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 3523–3542. [[CrossRef](#)]
20. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
21. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Cham, Switzerland, 2015; pp. 234–241.
22. Zhu, Q.; Zhong, Y.; Liu, Y.; Zhang, L.; Li, D. A Deep-Local-Global Feature Fusion Framework for High Spatial Resolution Imagery Scene Classification. *Remote Sens.* **2018**, *10*, 568. [[CrossRef](#)]
23. Lu, Y.; Chen, Y.; Zhao, D.; Chen, J. Graph-FCN for image semantic segmentation. In *International Symposium on Neural Networks*; Springer: Cham, Switzerland, 2019.
24. Yan, W.; Wang, Y.; Gu, S.; Huang, L.; Yan, F.; Xia, L.; Tao, Q. The domain shift problem of medical image segmentation and vendor-adaptation by Unet-GAN. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Shenzhen, China, 13–17 October 2019; Springer: Cham, Switzerland, 2019.
25. Cao, H.; Wang, Y.; Chen, J.; Jiang, D.; Zhang, X.; Tian, Q.; Wang, M. Swin-unet: Unet-like pure transformer for medical image segmentation. *arXiv* **2021**, arXiv:2105.05537.
26. Ahmad, P.; Jin, H.; Alroobaea, R.; Qamar, S.; Zheng, R.; Alnajjar, F.; Aboudi, F. MH UNet: A multi-scale hierarchical based architecture for medical image segmentation. *IEEE Access* **2021**, *9*, 148384–148408. [[CrossRef](#)]
27. Shuvo, M.B.; Ahommed, R.; Reza, S.; Hashem, M.M.A. CNL-UNet: A novel lightweight deep learning architecture for multimodal biomedical image segmentation with false output suppression. *Biomed. Signal Process. Control* **2021**, *70*, 102959. [[CrossRef](#)]
28. Kaymak, Ç.; Uçar, A. Semantic image segmentation for autonomous driving using fully convolutional networks. In Proceedings of the 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), Malatya, Turkey, 21–22 September 2019; pp. 1–8.
29. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2274–2282. [[CrossRef](#)]
30. Sornapudi, S.; Stanley, R.J.; Stoecker, W.V.; Almubarak, H.; Long, R.; Antani, S.; Thoma, G.; Zuna, R.; Frazier, S.R. Deep learning nuclei detection in digitized histology images by superpixels. *J. Pathol. Inform.* **2018**, *9*, 5. [[CrossRef](#)]
31. Yang, F.; Ma, Z.; Xie, M. Image classification with superpixels and feature fusion method. *J. Electron. Sci. Technol.* **2021**, *19*, 100096. [[CrossRef](#)]
32. Cai, L.; Xu, X.; Liew, J.H.; Foo, C.S. Revisiting superpixels for active learning in semantic segmentation with realistic annotation costs. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 25 June 2021; pp. 10988–10997.
33. Kanezaki, A. Unsupervised image segmentation by backpropagation. In Proceedings of the 2018 IEEE international conference on acoustics, speech and signal processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; pp. 1543–1547.
34. Lv, N.; Chen, C.; Qiu, T.; Sangaiyah, A.K. Deep learning and superpixel feature extraction based on contractive autoencoder for change detection in SAR images. *IEEE Trans. Ind. Inform.* **2018**, *14*, 5530–5538. [[CrossRef](#)]
35. Xiong, X.; Duan, L.; Liu, L.; Tu, H.; Yang, P.; Wu, D.; Chen, G.; Xiong, L.; Yang, W.; Liu, Q. Panicle-SEG: A robust image segmentation method for rice panicles in the field based on deep learning and superpixel optimization. *Plant Methods* **2017**, *13*, 1–15. [[CrossRef](#)] [[PubMed](#)]
36. Afza, F.; Sharif, M.; Mittal, M.; Khan, M.A.; Hemanth, J. A hierarchical three-step superpixels and deep learning framework for skin lesion classification. *Methods* **2022**, *202*, 88–102. [[CrossRef](#)] [[PubMed](#)]
37. Ali, J.; Khan, R.; Ahmad, N.; Maqsood, I. Random forests and decision trees. *Int. J. Comput. Sci. Issues (IJCSI)* **2012**, *9*, 272.

38. Liu, B.; Guo, W.; Chen, X.; Gao, K.; Zuo, X.; Wang, R.; Yu, A. Morphological attribute profile cube and deep random forest for small sample classification of hyperspectral image. *IEEE Access* **2020**, *8*, 117096–117108. [[CrossRef](#)]
39. Kong, Y.; Yu, T. A deep neural network model using random forest to extract feature representation for gene expression data classification. *Sci. Rep.* **2018**, *8*, 1–9. [[CrossRef](#)] [[PubMed](#)]
40. Luan, J.; Zhang, C.; Xu, B.; Xue, Y.; Ren, Y. The predictive performances of random forest models with limited sample size and different species traits. *Fish. Res.* **2020**, *227*, 105534. [[CrossRef](#)]
41. Zeng, M.; Li, J.; Peng, Z. The design of top-hat morphological filter and application to infrared target detection. *Infrared Phys. Technol.* **2006**, *48*, 67–76. [[CrossRef](#)]
42. Brownrigg, D.R.K. The weighted median filter. *Commun. ACM* **1984**, *27*, 807–818. [[CrossRef](#)]
43. Townsend, J.T. Theoretical analysis of an alphabetic confusion matrix. *Percept. Psychophys.* **1971**, *9*, 40–50. [[CrossRef](#)]
44. Li, S.; Zhao, X.; Zhou, G. Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network. *Comput.-Aided Civ. Infrastruct. Eng.* **2019**, *34*, 616–634. [[CrossRef](#)]
45. Pan, Y.; Zhang, L. Dual attention deep learning network for automatic steel surface defect segmentation. *Comput.-Aided Civ. Infrastruct. Eng.* **2022**, *37*, 1468–1487. [[CrossRef](#)]
46. Shi, J.; Dang, J.; Cui, M.; Zuo, R.; Shimizu, K.; Tsunoda, A.; Suzuki, Y. Improvement of Damage Segmentation Based on Pixel-Level Data Balance Using VGG-Unet. *Appl. Sci.* **2021**, *11*, 518. [[CrossRef](#)]
47. Ye, S.; Wu, K.; Zhou, M.; Yang, Y.; Tan, S.; Xu, K.; Song, J.; Bao, C.; Ma, K. Light-weight calibrator: A separable component for unsupervised domain adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 13736–13745.
48. Caicedo, J.C.; Goodman, A.; Karhohs, K.W.; Cimini, B.A.; Ackerman, J.; Haghghi, M.; Heng, C.K.; Becker, T.; Doan, M.; McQuin, C.; et al. Nucleus segmentation across imaging experiments: The 2018 Data Science Bowl. *Nat. Methods* **2019**, *16*, 1247–1253. [[CrossRef](#)]
49. Han, L.; Li, P.Y.; Yu, S.J.; Chen, C.; Fei, C.W.; Lu, C. Creep/fatigue accelerated failure of Ni-based superalloy turbine blade: Microscopic characteristics and void migration mechanism. *Int. J. Fatigue* **2022**, *154*, 106558. [[CrossRef](#)]
50. Li, X.Q.; Song, L.K.; Bai, G.C. Deep learning regression-based stratified probabilistic combined cycle fatigue damage evaluation for turbine bladed disks. *Int. J. Fatigue* **2022**, *159*, 106812. [[CrossRef](#)]