An Enhanced U-Net Approach for Segmentation of Aeroengine Hollow Turbine Blade

Jia Zheng 1, Chuan Tang 1, Yuanxi Sun 2,*, Mingchi Feng 1 and Congzhe Wang 1

1 School of Advanced Manufacturing Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
2 State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing 400044, China
* Correspondence: sunyuanxi@cqu.edu.cn

Abstract: The hollow turbine blade plays an important role in the propulsion of the aeroengine. However, due to its complex hollow structure and nickel-based superalloys material property, only industrial computed tomography (ICT) could realize its nondestructive detection with sufficient intuitiveness. The ICT detection precision mainly depends on the segmentation accuracy of target ICT images. However, because the hollow turbine blade is made of special superalloys and contains many small unique structures such as film cooling holes, exhaust edges, etc., the ICT image quality of the hollow turbine blades is often deficient, with artifacts, low contrast, and inhomogeneity scattered around the blade contour, making it hard for traditional mathematical model-based methods to acquire satisfying segmentation precision. Therefore, this paper presents a deep learning-based approach, i.e., the enhanced U-net with multiscale inputs, dense blocks, focal loss function, and residual path in the skip connection to realize the high-precision segmentation of the hollow turbine blade. The experimental results show that our proposed enhanced U-net can achieve better segmentation accuracy for practical turbine blades than conventional U-net and traditional mathematical model-based methods.

Keywords: segmentation; computed tomography; U-net; hollow turbine blade

MSC: 68T07

1. Introduction

As an important part of the aeroengine, the hollow turbine blade is made of special nickel-based superalloys and has a complex internal multicavity structure with closed curved surfaces to adapt to the high temperature and pressure working conditions. Its thin-walled structure makes it hard for traditional detection methods, such as three-coordinate measurement, ultrasound, etc., to detect it qualitatively and quantitatively [1]. Computed tomography (CT) is a prospective nondestructive detection method that can “see through” the test object. Hence, CT detection has been widely applied in tasks where traditional methods cannot perform accurate detections, such as material science [2], industrial applications [3], medical imaging [4], geosciences [5], civil engineering [6], etc.

However, due to the hollow turbine blade’s complex structure and special material properties, the industrial computed tomography (ICT) images contain blurred and low-contrast small structures whose precise contours are hard to extract [7], which greatly affects their ICT detection accuracy, as shown in Figure 1. Typically, there are two ways to solve this problem: (1) improve the ICT image quality and (2) increase the precision of image segmentation methods. The first approach requires better hardware of X-ray source, flat panel detector, etc., which is expensive and laborious. Therefore, many efforts have been made to increase the segmentation accuracy of the turbine blade.
Typical ICT segmentation methods are derived from natural image processing methods, i.e., they are modified to adapt to ICT image characteristics. Popular ICT image segmentation methods include thresholding method [8], morphological method [9], edge detection [2], active contours [10], fuzzy methods [11], etc. Among them, thresholding is simple and fast. However, thresholding cannot deal with images with artifacts and noise well [12]. Morphological methods can deal with noisy images but are easily affected by artifacts [13]. Edge detection methods are fast but can only handle images without noise, inhomogeneity, and artifacts [14]. Active contours can handle noise and inhomogeneity well but is time-consuming [15]. Fuzzy methods are sensitive to noise and low contrast [16].

With the development of computer technology, deep neural networks have made great progress in the fields of machine learning and computer vision [17]. Using a deep neural network to segment an object among the background is to train the network for specific image segmentation tasks, where the network architecture and the training methods will all affect the final segmentation accuracy [18]. Common deep learning architectures include convolutional neural networks (CNNs), recurrent neural networks (RNNs), encoder–decoder models, and generative adversarial networks (GANs) [19]. Image segmentation models are basically derived from these network architectures with the help of encoders, decoders, skip-connections and dilation modifications, etc. Typical image segmentation methods are fully convolutional networks (FCNs) [20], encoder–decoder approaches [21], multiscale and pyramid network-based models [22], R-CNN approaches [23], dilated convolutional networks [24], RNN-based approaches [25], and attention- [26] and GAN [27]-based models. Among different network structures, the U-net approach is an efficient segmentation network that is very suitable for images with microscopic details, such as medical images with blood vessels and ICT images with complex microstructures [28]. Compared with conventional encoder–decoder convolution neural networks, the bypass connections in the U-net could compensate for the lost high-frequency details during pooling [29].

The most widely applied area of the U-net is medical image segmentation [30]. For example, Baltruschat et al. used the U-net to segment the bone implant, which achieved the overall best mean IoU = 0.906 [31]. Ghosh et al. combined the U-net with the VGG-16 network to segment brain tumors and achieved a pixel accuracy of 0.9975 [32]. Guo et al. used the U-net for breast ultrasound image segmentation and realized the average IOU of 82.7% ±0.02, while Khaled et al. obtained a mean dice similarity coefficient (DSC) of 0.680 (0.802 for main lesions) for dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) with the help of the U-net [33]. Li et al. also applied the U-net in isotropic quantitative differential phase contrast imaging, the result of which shows that the quantitative phase value in the ROI is recovered from 66% to 97% compared with the ground truth [34]. Lee et al. proposed the FUS-net based on the U-net for interference filtering in high-intensity-focused ultrasound images, which performs 15% better than stacked autoencoders (SAE) on evaluated test datasets [35]. Moreover, Rocha et al. compared the U-net with the conventional sliding band filter (SBF) approach, the experimental results...
of which indicate that the U-net-based models yield more identical results to the ground truth [36].

Because the conventional U-net will pass the low-resolution content through the skip-connections, the high-resolution edge information of the input image may not be sufficiently extracted, resulting in a reduced segmentation accuracy of tiny objects. Many efforts have been made to improve the performance of the conventional U-net. For example, Han et al. proposed a dual-frame U-net by adding a residual path after the pooling and before the unpooling operations for sparse-view CT reconstruction, which received a better peak-signal-to-noise ratio (PSNR) performance at the condition of x2 down-sampling factor compared with conventional U-net [29]. Seo et al. added a residual path with deconvolution and activation operations to the skip connection of the U-net to avoid duplication of low-resolution information of features, which realized a DSC of 98.51% and 89.72 % for liver and liver–tumor segmentation [37], respectively. In addition, Man et al. applied the deformable convolution in the conventional U-net for pancreas segmentation, which can capture the geometry-aware information of the head and tail end of the pancreas with a mean DSC of 86.93 ± 4.92% [38]. Hiasa et al. modified the U-net by inserting the dropout layer before each max pooling layer and after each up-convolution layer, which realized a dice coefficient (DC) of 0.891 ± 0.016 and an average symmetric surface distance (ASD) of 0.994 ± 0.230 mm for muscle segmentation [39]. Moreover, He et al. proposed a hierarchically fused U-net by incorporating contour awareness for prostate segmentation, which achieved an overall DSC accuracy of 0.878 ± 0.029 [40].

In addition to medical imaging, the U-net is also suitable for industrial applications. For example, Wang et al. used the U-net to remove artifacts produced by the projection of the static parts in CT reconstruction, which realized the recovery of rotating part details during in situ nondestructive testing of airplane engines [41]. Li et al. realized acoustic interference striation (AIS) recovery using the U-net. The test results in range-dependent waveguides with nonlinear internal waves demonstrated its effectiveness under different signal-to-noise ratios and different amplitudes and widths [42]. Xiao et al. realized precise ore mask segmentation (IoU = 92.07%) via U-net with the combination of the boundary mask fusion block [43].

In summary, the U-net architecture has good segmentation performance for grayscale images compared to conventional single-resolution neural networks. However, the ICT images of the hollow turbine blade contain severe inhomogeneity and cone beam artifacts [7], which are not conducive to accurate U-net segmentation of the turbine blade edge. At the same time, the turbine blades contain microstructures such as film cooling holes and exhaust edges that are not easy to detect via the U-net. Therefore, this paper presents an enhanced U-net with multiscale input and structural modifications to improve the segmentation performance of the conventional U-net. The contribution of this paper is that a deep-learning approach for the segmentation of the hollow turbine blade is proposed, which is a relatively novel approach in the area of aeroengine engineering. Sample expansion and modification to the U-net are conducted to improve the network’s processing effect on inhomogeneity, noise, and artifacts, providing a feasible solution for similar industrial nondestructive testing problems.

The rest of this paper is arranged as follows. Section 2 introduces our enhanced U-net; Section 3 testifies our proposed method by experiments; Section 4 concludes our paper.

2. Materials and Methods
2.1. Preprocessing of Training Data
2.1.1. Data Source

The schematic diagram of the ICT scanning system is shown in Figure 2. The workpiece to be detected is placed on the rotary table controlled by the computer numerical control (CNC) system. Then the X-ray source projects an X-ray to penetrate the workpiece and leaves a projection on the flat panel detector. By rotating the workpiece at 360 degrees,
the projection data of the workpiece in all directions can be obtained. After that, the ICT reconstruction algorithm is used to obtain the CT slice images of the workpiece.

![Schematic diagram of the ICT scanning system](image)

**Figure 2.** The working principle of the ICT system.

The segmentation objects of this paper are the high-pressure turbine blades of the CFM56-7BE and the Pratt & Whitney F100 aeroengine, as shown in Figure 3. The ICT scanning parameters are as follows: equipment model: AX-4000CT; radiation source: FineTec microfocus radiation source; detector: VAREX XRD4343N (formerly PE); scanning parameters: 400 kV voltage and 1.5 mA current.

![Turbine blades](image)

**Figure 3.** Turbine blades: (a) CFM56-7BE aeroengine and (b) Pratt & Whitney F100 aeroengine.

2.1.2. Data Augmentation

To augment the training data for segmentation of input images, this paper adopted the following augmentation approaches:

- **Translation Transformation**

  The translation transformation of an image generally moves the image as a whole along the X direction or the Y direction or both the X and Y directions at the same time. After the translation, blank pixels will be generated, which could be handled via filling with black, assuming that the entire image moves \(\Delta x\) along the X direction and \(\Delta y\) along the Y
direction. Then the transformation matrix from the original pixel \((x_0, y_0)\) to the transformed point \((x, y)\) can be expressed as Equation (1):

\[
\begin{bmatrix}
  x \\
  y \\
  1
\end{bmatrix}
= \begin{bmatrix}
  1 & 0 & \Delta x \\
  0 & 1 & \Delta y \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  x_0 \\
  y_0 \\
  1
\end{bmatrix}
\]

(1)

- **Mirror Transformation**
  
  Mirror transformation includes horizontal mirroring and vertical mirroring. The image size remains unchanged before and after the mirror transformation, and no new blank pixels are generated. Suppose the image height is \(h\), and the width is \(w\). Then the transformation matrix from the original pixel \((x_0, y_0)\) to the transformed point \((x, y)\) via horizontal mirroring and vertical mirroring can be expressed in Equations (2) and (3):

\[
\begin{bmatrix}
  x \\
  y \\
  1
\end{bmatrix}
= \begin{bmatrix}
  -1 & 0 & w \\
  0 & 1 & 0 \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  x_0 \\
  y_0 \\
  1
\end{bmatrix}
\]

(2)

\[
\begin{bmatrix}
  x \\
  y \\
  1
\end{bmatrix}
= \begin{bmatrix}
  1 & 0 & 0 \\
  0 & -1 & h \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  x_0 \\
  y_0 \\
  1
\end{bmatrix}
\]

(3)

- **Rotation Transformation**
  
  Rotation transformation typically rotates the image around its center. Rotate the original image by \(\theta\) clockwise around the center. Then the transformation matrix from the original pixel \((x_0, y_0)\) to the transformed point \((x, y)\) via rotation transformation can be expressed as Equation (4):

\[
\begin{bmatrix}
  x \\
  y \\
  1
\end{bmatrix}
= \begin{bmatrix}
  \cos \theta & \sin \theta & 0 \\
  -\sin \theta & \cos \theta & 0 \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  x_0 \\
  y_0 \\
  1
\end{bmatrix}
\]

(4)

- **Scaling Transformation**
  
  The scaling transformation in data augmentation is helpful for neural networks to learn features of different scales, which need to be cropped or filled with blank pixels (this paper filled the blank pixel with zero) to remain the same size as the original image when the scale ratio is above or below one.

2.2. Architecture of the Enhanced U-net

The traditional U-net is an end-to-end encoder–decoder-based image semantic segmentation network, which is mainly composed of the encoder, the decoder, and the jump connection. The architecture of the U-net resembles a capital “U” letter, which is a symmetrical network with an encoder on the left and the decoder on the right. The encoder is mainly used to extract the image’s features, location, semantics, and other information, which is composed of several groups of “convolution + batchnormalization + Relu” operations followed by “Maxpooling” to realize the nonlinear expression capability of the network. Unlike the encoder, the decoder is used for up-sampling, restoring, and decoding the learned abstract semantic features via up-convolution operations to recover the image’s semantic information gradually. The input of each decoder layer is fused with the features of the corresponding encoder layer through a jump connection to obtain enough image features at different scales. A 1*1 convolution operation will connect the output of the decoder and Softmax operation to map feature maps learned by the U-net network to the number of categories that need to be divided. Then each pixel in the original image is classified according to the probability value generated by the U-net network. Compared with conventional fully convolutional networks, the U-net network performs well in dealing with
image contour segmentation due to the deep multiscale fusion of high-level and low-level features through skip connection operations.

In the U-net, the deep feature maps near the end of the network can learn strong abstract semantic information. In contrast, the feature maps near the input layer can learn detailed information about the input, including target location, edge, and other features. Therefore, the noise, low contrast, and inhomogeneity of the input images will jeopardize the details of U-net feature maps and significantly affect the accuracy of the overall segmentation performance. To solve this dilemma, this paper proposed an enhanced U-net for the segmentation of the aeroengine hollow turbine blade. The following improvements are applied to the original U-net network.

1) Multiscale input: To provide the network with enough image information from different scales, this paper adopted the multiscale input for the network. The advantage of this operation is that the average pooling down-sampling of the original image will not increase the computation complexity of the parameters in the network while the network width of the decoder path is extended.

2) Dense block: Different from the common hierarchical convolutional neural network block, the dense block makes full use of the feature information at each level of the block, while the hierarchical network block only uses the feature map from the previous layer or two layers without considering the lower level. For each layer, the features of all previous layers in this block are used as input. Its own feature map is used as input for all subsequent layers, which can alleviate the gradient disappearance problem at a certain level and strengthen the full utilization of each feature map.

Let \( x_0, x_1, \ldots, x_{l-1} \) represent the output of the first \( l-1 \) layers in the current neural network block, and \( x_l \) represent the output of the \( l \)th convolutional layer; then the dense block can be expressed as follows:

\[
x_l = H_l([x_0, x_1, \ldots, x_{l-1}])
\]

where \( H_l(*) \) represents a certain nonlinear composite function of the \( l \)th convolutional layer.

3) Focal loss function: For the segmentation task of turbine blades, the proportion of blades in the input image is relatively small, i.e., the background is much larger than the area occupied by the object, which causes a serious imbalance between positive and negative samples. During neural network training, the model is more inclined to learn background features, and it is easy to judge positive samples that are difficult to learn as negative samples. To solve this imbalance issue, this paper used the Focal loss function to pay more attention to the contour pixels:

\[
L = -\left[\alpha(1 - \hat{y})^\gamma y \log \hat{y} + (1 - \alpha)\hat{y}^\gamma (1 - y) \log (1 - \hat{y})\right]
\]

In Equation (6), \( y \) and \( \hat{y} \) represent the network’s true value and prediction value; \( \alpha \) is the weighting factor, which is used to balance the imbalance of positive and negative samples; \( \gamma \) represents the modulation coefficient, which is used to control the weight of easy-to-classify and difficult-to-classify samples. For the area with extremely low contrast on the edge of the turbine blade, setting the modulation coefficient \( \gamma \) can enable the network model to strengthen the learning of difficult samples during learning, thereby improving the accuracy of image segmentation.

4) Residual path in the skip connection [37]: The input of the neural network consists of both high-resolution and low-resolution information. For the conventional U-net, the low-resolution information is passed through the next stage twice (through skip connection and pooling), while the high-resolution information is passed through only once (through the skip connection). To solve this dilemma, a residual path that contains a transposed convolution and activation is placed right after the pooling operation. The subtraction of the full-resolution information (original information passed through the skip connection) and low-resolution information (provided by the deconvolution operation of pooling, transposed convolution, and activation) can provide enough high-resolution information
for the decoder. The overall network structure of the proposed enhanced U-net is shown in Figure 4.

Figure 4. Structure of the enhanced U-net.

2.3. Training

The training platform used in this paper is based on MATLAB R2021a, where the computer is configured with Intel Core I9 11900k @3.5 Ghz, and Nvidia RTX A6000 video card. The operating system is 64-bit Windows 10.

2.3.1. Training Data Set

The original training data set contained a total of 493 slice images of two turbine blades, with the resolution cropped to 512 × 512. Another 180 and 30 slice images of these two turbine blades are used for validation and test, respectively. The ground truth of each image is manually labeled as follows:

1) manually draw the contours of the turbine blade in the CT slice image;
2) automatically fill in the area within the contour using the regional growth method.

Figure 5 shows the example of the original CT slice image and its manually labeled ground truth for both turbine blades.

Figure 5. Training data samples of two turbine blades.

2.3.2. Training Parameters

The training algorithm of the proposed network is “adam”, which could effectively decrease the learning rates of parameters with large parameter gradients and their squared values and increase the learning rates when gradients and squared values are small. The gradient moving average and decay rate of the squared gradient moving average of the training algorithm are set to 0.9 and 0.999, respectively. The initial learning rate is set to
0.001, and the epochs of the training process is 200. The minibatch size of the training is set to 8, and the validation frequency is set to every 30 iterations. The learning rate is reduced every ten epochs to improve the training effect by multiplying with a dropping factor of 0.95.

2.4. Performance Evaluation

The segmentation results are evaluated by the Jaccard similarity coefficient, as defined by the following equation:

\[ J(A, T) = \frac{|A \cap T|}{|A \cup T|} \]  \hspace{1cm} (7)

where A represents the segmentation result, T represents the ground truth, and \( \cap \) and \( \cup \) are the intersection and the union operations. || represents the number of pixels in the set. A higher Jaccard score means a better segmentation result.

Similarly, the dice similarity coefficient (DSC) is also used to evaluate the performance of the segmentation method, which can be expressed as:

\[ \text{DSC} = 2 \frac{|A \cap T|}{|A| + |T|} \]  \hspace{1cm} (8)

Another commonly used indicator is the BF score, i.e., how close the predicted boundary of an object matches the ground truth boundary. Compared with the Jaccard similarity coefficient, the BF score correlates better with human qualitative assessment. The BF score is defined as follows:

\[ \text{BF-Score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{recall} + \text{precision}} \]  \hspace{1cm} (9)

where precision is the fraction of detections that are true positives rather than false positives, while recall refers to the fraction of true positives that are detected rather than missed. Assuming TP, FP, FN, and TN represent the true positive, false positive, false negative, and true negative of the segmentation results, the precision = TP/(TP + FP), recall = TP/(TP + FN).

3. Results and Discussion

3.1. Segmentation Results

To demonstrate the effectiveness of the proposed method, it was compared with adaptively regularized kernel-based FCM (ARKFCM) [44], DRLSE [45], MAXENTROY [46], EM/MPM [47], continuous max-flow (CMF) [48], and OTSU [12]. MAXENTROPY, CMF, and OTSU methods are parameterless. Parameters of ARKFCM, DRLSE, and EM/MPM are listed in Table 1.

| Methods | Samples | 1#102 | 1#111 | 1#113 | 28#740 | 28#741 | 28#743 |
|---------|---------|-------|-------|-------|--------|--------|--------|
| ARKFCM  | winSize = 45; cNum = 2; opt = 'average' | winSize = 41; cNum = 2; opt = 'median' | winSize = 27; cNum = 2; opt = 'weighted' | winSize = 15; cNum = 2; opt = 'weighted' | winSize = 16; cNum = 2; opt = 'weighted' |
| DRLSE   | sigma = 45; iter_outer = 60; iter_inner = 10; timestep = 0.1; c0 = 1; mu = 1; | sigma = 35; iter_outer = 40; iter_inner = 10; timestep = 0.1; c0 = 1; mu = 0.1; | sigma = 21; iter_outer = 26; iter_inner = 10; timestep = 0.1; c0 = 1; mu = 1; | sigma = 21; iter_outer = 26; iter_inner = 10; timestep = 0.1; c0 = 1; mu = 1; | sigma = 21; iter_outer = 26; iter_inner = 10; timestep = 0.1; c0 = 1; mu = 1; |
| EM/MPM  | regions = 2; steps = 10; mpmSteps = 1; coolMax = 1.2; coolInc = 0.25 | regions = 2; steps = 18; mpmSteps = 1; coolMax = 1.2; coolInc = 0.025 | regions = 2; steps = 1; mpmSteps = 5; coolMax = 1.2; coolInc = 0.025 | regions = 2; steps = 10; mpmSteps = 8; coolMax = 1.2; coolInc = 0.025 | regions = 2; steps = 10; mpmSteps = 8; coolMax = 1.2; coolInc = 0.025 |
Our proposed method is compared with the conventional U-net, the dual-frame U-net [29], and the mU-net [37] approaches. The experimental results are depicted in Figure 6. In addition, a quantitative analysis of the experiments is listed in Table 2.

Figure 6. Experimental results of the proposed method, ARKFCM, DRLSE, MAXENTROPY, EM/MPM, CMF, OTSU, conventional U-net, dual-frame U-net and mU-net.
It can be known from Figure 6 that our proposed approach realized the best segmentation results compared with mathematical model-based segmentation methods and other U-net architectures. The ARKFCM method failed to segment the low contrast detail in the basin and the exhaust edge of both hollow turbine blades. In addition, small film cooling holes in the leading edges were also not segmented. For the DRLSE method, the target pixels on the edge of the hollow turbine blade are more likely to be mistakenly identified as the background, resulting in a certain dilation of the entire target object compared to the ground truth. Compared with the DRLSE method, the object dilation caused by this mis-segmentation is more obvious in the CMF method. Excessive dilation can cause severe deformation of the target, affecting the accuracy of subsequent measurements and registrations of the workpiece. Furthermore, the CMF method also failed in segmenting details in the blade basin and exhaust edge, as well as small film cooling holes scattered in the blade’s leading edge. The segmentation results of the EM/MPM and the MAXENTROPY methods were also ineffective, i.e., dilated edges, lack of details in low-contrast blade basin, and not segmented film cooling holes all appeared in both hollow turbine blades. For the Otsu method, the main issue was that the accuracy was not high, i.e., the edge details and film cooling holes were not successfully identified.

Compared with model-based segmentation methods, the comparison U-net architectures and our proposed enhanced U-net realized higher segmentation accuracies and details. The conventional U-net has a good segmentation effect in target with fewer artifacts and small details with good grayscale distribution, such as film cooling holes in the leading edge. However, it failed to segment small details with low contrast and artifacts, especially

| Evaluation Index | Methods | Experimental Results |
|------------------|---------|----------------------|
|                  | 1#102  | 1#111  | 1#113  | 2#740  | 2#741  | 2#743  |
| BF Score         |         |         |         |         |         |         |
| ARKFCM           | 0.9890  | 0.9889  | 0.9876  | 0.9801  | 0.9804  | 0.9863  |
| CMF              | 0.9403  | 0.9096  | 0.9092  | 0.8169  | 0.8127  | 0.8585  |
| DRLSE            | 0.9984  | 0.9818  | 0.9861  | 0.9713  | 0.9651  | 0.9793  |
| EM/MPM           | 0.9928  | 0.9876  | 0.9862  | 0.9769  | 0.9768  | 0.9824  |
| MAXENTROPY       | 0.9889  | 0.9872  | 0.9825  | 0.9291  | 0.9153  | 0.9177  |
| OTSU             | 0.9952  | 0.9899  | 0.9876  | 0.9799  | 0.9800  | 0.9858  |
| Conventional U-net | 1.0000  | 1.0000  | 1.0000  | 0.9960  | 0.9964  | 1.0000  |
| Proposed         |         |         |         |         |         |         |
| Dual-Frame U-net | 1.0000  | 1.0000  | 1.0000  | 0.9965  | 0.9967  | 0.9995  |
| mU-net           | 1.0000  | 1.0000  | 1.0000  | 0.9970  | 0.9967  | 0.9995  |
| Jaccard          |         |         |         |         |         |         |
| ARKFCM           | 0.8183  | 0.7622  | 0.7904  | 0.7964  | 0.7924  | 0.8121  |
| CMF              | 0.6197  | 0.5604  | 0.5865  | 0.5865  | 0.5862  | 0.6176  |
| DRLSE            | 0.7711  | 0.7013  | 0.7401  | 0.7514  | 0.7491  | 0.7815  |
| EM/MPM           | 0.8122  | 0.7425  | 0.7711  | 0.7720  | 0.7665  | 0.7916  |
| MAXENTROPY       | 0.7425  | 0.6973  | 0.7209  | 0.6853  | 0.6789  | 0.7012  |
| OTSU             | 0.8092  | 0.7388  | 0.7678  | 0.7677  | 0.7634  | 0.7879  |
| Conventional U-net | 0.9497  | 0.9311  | 0.9313  | 0.9423  | 0.9418  | 0.9472  |
| Proposed         | 0.9540  | 0.9392  | 0.9423  | 0.9547  | 0.9554  | 0.9592  |
| Dual-Frame U-net | 0.9485  | 0.9391  | 0.9404  | 0.9553  | 0.9550  | 0.9590  |
| mU-net           | 0.9474  | 0.9383  | 0.9377  | 0.9534  | 0.9518  | 0.9550  |
| DSC              |         |         |         |         |         |         |
| ARKFCM           | 0.9001  | 0.8651  | 0.8829  | 0.8867  | 0.8842  | 0.8963  |
| CMF              | 0.7652  | 0.7183  | 0.7394  | 0.7394  | 0.7391  | 0.7636  |
| DRLSE            | 0.8708  | 0.8244  | 0.8506  | 0.8581  | 0.8566  | 0.8774  |
| EM/MPM           | 0.8964  | 0.8522  | 0.8708  | 0.8713  | 0.8678  | 0.8837  |
| MAXENTROPY       | 0.8522  | 0.8217  | 0.8378  | 0.8133  | 0.8087  | 0.8244  |
| OTSU             | 0.8945  | 0.8498  | 0.8687  | 0.8686  | 0.8658  | 0.8814  |
| Conventional U-net | 0.9742  | 0.9643  | 0.9644  | 0.9703  | 0.9700  | 0.9729  |
| Proposed         | 0.9765  | 0.9686  | 0.9703  | 0.9768  | 0.9772  | 0.9792  |
| Dual-Frame U-net | 0.9736  | 0.9686  | 0.9693  | 0.9771  | 0.9770  | 0.9791  |
| mU-net           | 0.9730  | 0.9682  | 0.9678  | 0.9761  | 0.9753  | 0.9770  |
in the film cooling holes in the low-contrast leading edge and basin, as depicted in the red rectangle in Figure 6. Compared with dual-frame U-net and mU-net architectures, our proposed approach achieved better performance in segmentation accuracy and edge continuity. In fact, it is difficult for nonprofessionals to distinguish whether these pixels belong to the background or the target. The reason for this problem in conventional U-net is that its receptive field is not large enough. Simply increasing the depth of the network cannot improve the segmentation effect of tiny details with low contrast and artifacts. Therefore, this paper tried to increase the receptive field of the conventional U-net by using multiscale original images as input and adopting the dense block and residual path in the skip connection without increasing the depth of the network, thereby increasing the segmentation effect of indistinguishable details in the hollow turbine blade CT images caused by low contrast and artifacts.

To quantitatively analyze the segmentation effects of the proposed approach and the comparison methods, the BF score, the Jaccard, and the DSC indexes of each segmentation result are listed in Table 2. For the segmentation samples of 1#102, 1#111, 1#113, and 2#741, our proposed approach achieved the highest BF scores, Jaccard, and DSC indexes among all the segmentation results. Although the conventional U-net realized the highest BF score of sample 2#743, it is only 0.001 higher than our proposed approach. Moreover, the smoothness of the object contour in the segmentation results of the conventional U-net is not as good as our proposed method. More importantly, the conventional U-net cannot segment the film cooling holes in the leading edge in both samples, demonstrating our proposed approach’s superiority. The statistical T-test differences between each comparison method and our proposed approach are listed in Table 3. It can be known from Table 3 that the \( p \)-values of traditional mathematical model-based methods are all smaller than 0.01, which means that our proposed approach outperforms these traditional methods. For conventional U-net, its \( p \)-values of Jaccard and DSC indexes are smaller than 0.05, which means that our proposed approach performs better than the conventional U-net. For dual-frame U-net and mU-net, their \( p \)-values range 0.45–0.98, which means that our proposed enhanced U-net can achieve a slightly enhanced segmentation effect. It is worth noting that the \( p \)-value of dual-frame U-net is higher than that of mU-net, which indicates that the performance of dual-frame U-net is closer to our proposed approach.

Table 3. Statistical T-test difference between each comparison method and our proposed approach.

| Method          | ARKFCM | CMF  | DRILSE | EM/MMPM | MAXENTROPY | OTSU | Conventional U-net | Dual-Frame U-net | mU-net |
|-----------------|--------|------|--------|---------|------------|------|--------------------|-----------------|--------|
| \( p \)-BF      | 2.27 \times 10^{-5} | 0.000194 | 0.003279 | 0.000224 | 0.012090 | 0.000598 | 1.000000 | 0.986737 | 0.988246 |
| \( p \)-Jaccard | 6.34 \times 10^{-8} | 5.19 \times 10^{-12} | 1.05 \times 10^{-8} | 9.79 \times 10^{-9} | 3.37 \times 10^{-10} | 8.6 \times 10^{-9} | 0.009792 | 0.796774 | 0.452846 |
| \( p \)-DSC     | 1.21 \times 10^{-6} | 2.79 \times 10^{-11} | 2.9 \times 10^{-8} | 2.04 \times 10^{-8} | 9.45 \times 10^{-10} | 1.78 \times 10^{-9} | 0.050454 | 0.799691 | 0.453197 |

3.2. Analysis of the Proposed Architecture

3.2.1. Ablation Experiments

In order to quantitatively analyze the effectiveness of our proposed approach, four groups of ablation experiments were conducted in this paper to verify the improvement effect of multi-input, dense block, focal loss function, and residual path in the skip connection on turbine blade segmentation, respectively. The results are listed in Table 4. It can be known from Table 4 that the dense block and the multi-input modifications played more critical roles in the improvement of the segmentation results. The residual path and focal loss function modifications further enhanced the effect of the proposed architecture.

3.2.2. Processing Time

To study the segmentation efficiency of the proposed approach, the running time of the proposed approach and all comparison methods are listed in Table 5.
### Table 4. Ablation experiments of the proposed approach.

| Evaluation Index | Methods                              | 1#102 | 1#111 | 1#113 | 2#740 | 2#741 | 2#743 |
|------------------|--------------------------------------|-------|-------|-------|-------|-------|-------|
| BF Score         | Dense Block + U-net                  | 1.000 | 1.000 | 1.000 | 0.9961| 0.9968| 0.9991|
|                  | Multi Input + U-net                  | 1.000 | 1.000 | 1.000 | 0.9964| 0.9967| 0.9991|
|                  | Focal Loss + U-net                   | 1.000 | 1.000 | 0.9998| 0.9967| 0.9967| 0.9998|
|                  | Residual Path + U-net                | 1.000 | 1.000 | 1.000 | **0.9970**| 0.9967| 0.9995|
|                  | Conventional U-net                  | 1.000 | 1.000 | 1.000 | 0.9960| 0.9964| 1.0000|
|                  | Proposed                             | 1.000 | 1.000 | 1.000 | 0.9965| 0.9969| 0.9990|
| Jaccard          | Dense Block + U-net                  | 0.9492| 0.9397| 0.9406| 0.9535| 0.9530| 0.9578|
|                  | Multi Input + U-net                  | 0.9487| 0.9397| 0.9419| 0.9557| 0.9540| 0.9600|
|                  | Focal Loss + U-net                   | 0.9329| 0.9337| 0.9273| 0.9533| 0.9514| 0.9502|
|                  | Residual Path + U-net                | 0.9474| 0.9383| 0.9377| 0.9534| 0.9518| 0.9550|
|                  | Conventional U-net                  | 0.9497| 0.9311| 0.9313| 0.9423| 0.9418| 0.9472|
|                  | Proposed                             | 0.9540| 0.9392| 0.9423| 0.9547| 0.9554| 0.9592|
| DSC              | Dense Block + U-net                  | 0.9739| 0.9689| 0.9694| 0.9762| 0.9739| 0.9784|
|                  | Multi Input + U-net                  | 0.9737| 0.9689| 0.9701| 0.9773| 0.9765| 0.9796|
|                  | Focal Loss + U-net                   | 0.9653| 0.9657| 0.9623| 0.9761| 0.9751| 0.9745|
|                  | Residual Path + U-net                | 0.9730| 0.9682| 0.9678| 0.9761| 0.9753| 0.9770|
|                  | Conventional U-net                  | 0.9742| 0.9643| 0.9644| 0.9703| 0.9700| 0.9729|
|                  | Proposed                             | 0.9765| 0.9686| 0.9703| 0.9768| 0.9772| 0.9792|

It can be known from Table 5 that the deep learning-based methods consumed much less time compared with conventional model-based methods. The most time-saving MAX-ENTROPY method is still an order of magnitude more than the deep learning-based approaches. For the U-net architecture, the more layers the network has, the more segmentation time it consumes. However, it is worth noting that deep learning algorithms run on GPU, while conventional model-based methods run on CPU. This difference also determines that the efficiency of these methods is only horizontally comparable.

### Table 5. Processing time of the proposed approach and the comparison methods.

| Samples | ARKFCM | CMF | DRLSE | EM/MPM | MAXENTROPY | OTSU | U-net | Proposed | mU-net | Dual-Frame U-net |
|---------|--------|-----|-------|--------|------------|------|-------|----------|--------|-----------------|
| 1#102   | 6.003 s| 0.425 s | 134.397 s | 0.904 s | 0.367 s | 0.128 s | 0.0315 s | 0.0810 s | 0.0418 s | 0.0304 s |
| 1#111   | 5.927 s| 0.435 s | 59.553 s | 0.706 s | 0.360 s | 0.073 s | 0.0325 s | 0.0811 s | 0.0428 s | 0.0299 s |
| 1#113   | 6.697 s| 0.438 s | 76.094 s | 0.737 s | 0.327 s | 0.092 s | 0.0319 s | 0.0806 s | 0.0423 s | 0.0294 s |
| 2#740   | 6.117 s| 0.504 s | 25.223 s | 0.906 s | 0.320 s | 0.086 s | 0.0330 s | 0.0810 s | 0.0418 s | 0.0295 s |
| 2#741   | 5.902 s| 0.471 s | 22.218 s | 0.935 s | 0.297 s | 0.086 s | 0.0294 s | 0.0810 s | 0.0388 s | 0.0301 s |
| 2#743   | 5.899 s| 0.514 s | 22.111 s | 0.759 s | 0.424 s | 0.104 s | 0.0292 s | 0.0807 s | 0.0404 s | 0.0296 s |

### 3.2.3. Robustness of the Proposed Architecture

To study the robustness of the proposed approach, this paper evaluated it on sample 1#268 with severe inhomogeneity, noise, and low contrast. The comparison results with conventional U-net are depicted in Figure 7.

It can be known from Figure 7 that the proposed approach can realize a better segmentation effect compared with conventional U-net in terms of low-quality images. For the region with insufficient contrast, the conventional U-net can easily cause mis-segmentation. For example, in the blade tail, the severe low contrast and inhomogeneity directly affect the U-net judgment of the edge contour. Pixels that originally belonged to the object were mis-segmented as the background.

In summary, the deep learning approach outperforms conventional model-based methods in segmentation accuracy and protecting tiny details. The essential reason is that, compared with conventional mathematical model-based segmentation methods, the deep neural network-based approach has a good learning effect on the continuity between slice images. This is because the physical distance between two adjacent layers is very small.
(often around 0.1 mm), so the interlayer variation of the sliced image is relatively small. When this layer is affected by uneven contrast or artifacts, conventional mathematical model algorithms are prone to mis-segmentation. The deep neural network not only learns the association between the real image of the current layer and the corresponding ground truth but also the real image information of adjacent layers. Therefore, when an ambiguous structure only exists in a few layers, it will not have a large impact on the convergence effect of the deep neural network.

Figure 7. Robustness analysis of the proposed approach on the low-quality sample 1#268.

3.3. Limitations

Due to the particular structure and material characteristics of the aeroengine turbine blade, its CT images have very obvious artifacts and inhomogeneity at the edges and inner cavities. Therefore, the conventional model-based image segmentation methods are often unsatisfactory for the CT image of the turbine blade, where the loss of details and mis-segmentation often occur. The use of supervised deep learning methods can effectively improve the segmentation effect of turbine blade CT images. The experimental results of this paper have demonstrated that the deep learning method has completely surpassed the conventional model method, whether in terms of processing speed or segmentation accuracy. In this paper, some improvements have been made on the basis of the conventional U-net architecture. Compared with the conventional U-net, the segmentation accuracy has been improved to a certain extent even under the influence of artifacts and inhomogeneity. However, due to a large number of tiny air film cooling holes in the turbine blades and the complex exhaust edges at the tail, it is still difficult to acquire precise details, even with the help of deep learning approaches. These difficult-to-divide air cooling film holes and exhaust edges are essential for turbine blades, which play an important role in the heat dissipation of turbine blades under extremely high temperatures and high-pressure conditions. Therefore, further efforts are needed to improve the segmentation accuracy of turbine blade details.

4. Conclusions

This paper presented an enhanced U-net for the segmentation of the aeroengine hollow turbine blade. The enhanced U-net is modified based on the conventional U-net
architecture, where the multiscale input, dense block, focal loss function, and residual path in the skip connection are added to improve the receptive field of the network without increasing its longitudinal depth. Experiments were conducted based on a set of ICT slice images of two practical hollow turbine blades, namely, the CFM56-7BE and Pratt & Whitney F100 blades. The experimental results indicate that our proposed approach has the best segmentation results in terms of segmentation accuracy and protection of tiny details with low contrast and artifacts compared with typical model-based algorithms and the conventional U-net. Future work of this paper will focus on improving the segmentation accuracy of tiny objects in the blade, such as film cooling holes and exhaust edges, etc.

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

Funding: This research was funded by National Natural Science Foundation of China, grant number 62101086, 62003060, 52005070; China Postdoctoral Science Foundation, grant number 2021M693769; Scientific and Technological Research Program of Chongqing Municipal Education Commission, grant number KJQN202100648; Natural Science Foundation of Chongqing, China, grant number cstc2021jcyj-bsh0180, CSTB2022NSCQ-MSX1297, cstc2020jcyj-msxmX0886.

Data Availability Statement: Not applicable.

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

References

1. Dong, Y.W.; Li, X.L.; Zhao, Q.; Yang, J.; Dao, M. Modeling of shrinkage during investment casting of thin-walled hollow turbine blades. J. Mater. Process. Technol. 2017, 244, 190–203. [CrossRef]

2. Ohtake, Y.; Suzuki, H. Edge detection based multi-material interface extraction on industrial CT volumes. Sci. China Ser. F Inf. Sci. 2013, 56, 1–9. [CrossRef]

3. Ciliberti, G.A.; Janello, P.; Jahnke, P.; Keuthage, L. Potentials of Full-Vehicle CT Scans Within the Automotive Industry. In Proceedings of the 19th World Conference on Nondestructive Testing (WCNDT 2016), Munich, Germany, 13–17 June 2016.

4. Qian, X.; Wang, J.; Guo, S.; Li, Q. An active contour model for medical image segmentation with application to brain CT image. Med. Phys. 2013, 40, 021911. [CrossRef] [PubMed]

5. Ketcham, R.A.; Carlson, W.D. Acquisition, optimization and interpretation of X-ray computed tomographic imagery: Applications to the geosciences. Comput. Geosci. 2001, 27, 381–400. [CrossRef]

6. Lee, J.H.; Lee, J.M.; Park, J.W.; Moon, Y.S. Efficient algorithms for automatic detection of cracks on a concrete bridge. In Proceedings of the 23rd International Technical Conference on Circuits/Systems, Computers and Communications, Shimonoseki, Japan, 6–9 July 2008; pp. 1213–1216.

7. Zheng, J.; Zhang, D.; Huang, K.; Sun, Y. Cone-Beam Computed Tomography Image Pretreatment and Segmentation. In Proceedings of the International Symposium on Computational Intelligence and Design, Hangzhou, China, 8–9 December 2018; pp. 25–28.

8. Ayala, H.V.H.; Santos, F.M.d.; Mariani, V.C.; Coelho, L.d.S. Image thresholding segmentation based on a novel beta differential evolution approach. Expert Syst. Appl. 2015, 42, 2136–2142. [CrossRef]

9. Alaknanda; Anand, R.S.; Kumar, P. Flaw detection in radiographic weld images using morphological approach. Ndt E Int. 2006, 39, 29–33. [CrossRef]

10. Zhang, K.; Zhang, L.; Song, H.; Zhou, W. Active contours with selective local or global segmentation: A new formulation and level set method. Image Vis. Comput. 2010, 28, 668–676. [CrossRef]

11. Zheng, J.; Zhang, D.; Huang, K.; Sun, Y. Adaptive image segmentation method based on the fuzzy c-means with spatial information. Iet Image Process. 2017, 12, 785–792. [CrossRef]

12. Zheng, J.; Zhang, D.; Huang, K.; Sun, Y.; Tang, S. Adaptive windowed range-constrained Otsu method using local information. J. Electron. Imaging 2016, 25, 013034. [CrossRef]

13. Alaknanda; Anand, R.S.; Kumar, P. Flaw detection in radiographic weldment images using morphological watershed segmentation technique. Ndt E Int. 2009, 42, 2–8. [CrossRef]

14. Prathusha, P.; Jyothi, S. A Novel Edge Detection Algorithm for Fast and Efficient Image Segmentation; Springer: Singapore, 2018; pp. 283–291.

15. Li, Y.; Cao, G.; Yu, Q.; Li, X. Fast and Robust Active Contours Model for Image Segmentation. Neural Process. Lett. 2019, 49, 431–452. [CrossRef]
43. Wang, W.; Li, Q.; Xiao, C.; Zhang, D.; Miao, L.; Wang, L. An Improved Boundary-Aware U-Net for Ore Image Semantic Segmentation. *Sensors* **2021**, *21*, 2615. [CrossRef] [PubMed]

44. Cherfa, I.; Mokraoui, A.; Mekhmoukh, A.; Mokrani, K. Adaptively Regularized Kernel-Based Fuzzy C-Means Clustering Algorithm Using Particle Swarm Optimization for Medical Image Segmentation. In *Proceedings of the 2020 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, Poznan, Poland, 23–25 September 2020; pp. 24–29.

45. Wang, D. Extremely optimized DRLSE method and its application to image segmentation. *IEEE Access* **2019**, *7*, 119603–119619. [CrossRef]

46. Merzougui, M.; El Allaoui, A. Region growing segmentation optimized by evolutionary approach and Maximum Entropy. *Procedia Comput. Sci.* **2019**, *151*, 1046–1051. [CrossRef]

47. Masuda, Y.; Tateyama, T.; Xiong, W.; Zhou, J.; Wakamiya, M.; Kanasaki, S.; Furukawa, A.; Chen, Y.W. Liver tumor detection in CT images by adaptive contrast enhancement and the EM/MPM algorithm. In *Proceedings of the 2011 18th IEEE International Conference on Image Processing*, Brussels, Belgium, 11–14 September 2011; pp. 1421–1424.

48. Wu, M.; Fan, W.; Chen, Q.; Du, Z.; Li, X.; Yuan, S.; Park, H. Three-dimensional continuous max flow optimization-based serous retinal detachment segmentation in SD-OCT for central serous chorioretinopathy. *Biomed. Opt. Express* **2017**, *8*, 4257–4274. [CrossRef]