Research Article

Research on Carbon Foam Image Segmentation Based on Deep Learning

Peng Gu, Ping Zhang, Mingfei Jiang, Xiaofeng Qiu, Xinyan Wang, Chun Du, and Tianyou Peng

1Tibet University, China
2University of Science and Technology Liaoning, China

Correspondence should be addressed to Ping Zhang; 6093198@qq.com

Received 27 April 2022; Revised 25 May 2022; Accepted 24 June 2022; Published 14 July 2022

Academic Editor: Kuruva Lakshman

Copyright © 2022 Peng Gu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As a substitute for metal, carbon foam is vital in the electromagnetic shielding industry. Nevertheless, the diameter and density of carbon foam cells are still mostly assessed manually. This research offers a Deep-Res-MixAttention segmentation method to effectively minimize manual labor and increase measurement efficiency. Moreover, the method consists of two modules: the MixAttention module is intended to improve feature extraction skills, and we use the multiscale deep residual module to collect edge information. In addition to enhancing the segmentation capability of incomplete carbon foam, the loss function is adjusted to address the dataset imbalance issue. Additionally, we propose the bidirectional selection rotation calipers algorithm to intelligently determine the density and diameter. The results reveal that the optimized network’s IoU and acc carbon reach 91.05% and 88.31%. Finally, the calculation errors of the average diameter and density are under control at 1.79% and 7.09%, respectively. The approach has a high application value for assessing the electromagnetic shielding effectiveness of carbon foam.

1. Introduction

With the widespread use of current electronic items, electromagnetic waves serve as the primary route of information transmission in people’s daily environments. Simultaneously, both the electronic components used in electronic devices and different signal towers will create electromagnetic pollution. This pollution will decrease the life of electronic equipment and diminish its sensitivity. The most effective way to prevent electromagnetic pollution from damaging human health and the surrounding environment is to apply electromagnetic shielding to different electrical items. Due to their low density, great electrical conductivity, ease of modification, and high chemical stability, carbon materials have increasingly replaced metal materials in the area of electromagnetic shielding during the last several decades. Among carbon materials, carbon foam has a tunable three-dimensional network cell structure: a material with low weight, a broad absorption frequency, and specific structural and functional properties [1–4]; therefore, it has been highly desired by researchers. The majority of research has shown that differences in the cell structure of carbon foam [5] have a large influence on its electromagnetic shielding performance. However, typical statistical approaches for cell structure are mostly manual statistics performed one by one, which is inefficient and leads to error rates.

The threshold-based image segmentation approach generates one or more grayscale thresholds using the grayscale characteristics, compares the grayscale value of each pixel in the picture to the threshold, and then classifies the pixels into the relevant categories. This approach is susceptible to noise and focuses only on pixel attributes, ignoring spatial aspects.

Two primary types of region-based picture segmentation algorithms exist: one involves beginning with a single-pixel and progressively joining suitable nearby pixels to generate the requisite segmentation area, while the other involves separating and merging regions. Target extraction is accomplished by beginning with the full picture, constantly subdividing to produce each subregion and then combining
the foreground areas to create the to-be-segmented targets. A segmentation approach based on edge detection including distinct areas is to tackle the segmentation problem [6–8]. The Fourier transform moves the picture from the spatial domain to the frequency domain, and the edge corresponds to the high-frequency component by radically altering the gray values of pixels on distinct borders. Alternatively, we may discover edge locations by using the discontinuous characteristic of pixel values in nearby regions. Wavelet analysis-based image segmentation algorithms offer good local characteristics in both the time and frequency domains and may examine signals by combining the time and frequency domains. The core contour model-based edge detection approach features a consistent and accessible description format. In a given picture, the target is detected using curve evolution, and the edge contour may be refined.

In recent years, with the advancement of deep learning [9, 10] in the fields of big data [11–13] and IoT [14], more and more individuals are focusing on the practical applications of deep learning in many sectors. Image segmentation based on deep learning is basically divided into supervised learning and unsupervised learning. The most prevalent approach for picture segmentation is supervised learning since supervised learning segmentation is superior, the accuracy is higher, and the number of application situations is bigger. FastFCN [15], Gated-SCNN [16], DeepLab [17, 18], Mask R-CNN [19], etc., are among its primary backbone network concepts. Unsupervised learning relies on unknown category (unlabeled) training samples to address numerous problems in pattern recognition; its primary studies include Double DIP [20] and SSL-ALPNet [21], but its segmentation impact is inferior to that of supervised learning, and its data scale requirements are enormous. If datasets cannot be expanded to a specific size, it is impossible to guarantee the accuracy of picture segmentation in the field of unsupervised learning.

The convex hull is the first geometric model that has been extensively researched in computational geometry, and its unique qualities have led to a wide range of applications. The convex hull is also significant in estimating the average radius and density of carbon foam. Images of segmented carbon foam are sometimes convex sets. To address this problem, we extract the carbon foam contours using the technique of contour detection. Then, based on the real circumstances, delete nonconvex data using Andrew’s Monotone Chain technique. Utilizing the enhanced Rotating Caliper technique, we calculated the remaining data, obtaining the average radius and density of carbon foam, beginning with the convex hull’s attributes. This technique introduces a novel concept to this subject.

2. Related Work

2.1. Carbon Foam. During the operation of energized equipment, the internal magnetic field will generate electromagnetic waves. Electromagnetic pollution can cause mutual interference between devices and threaten people’s health. Figure 1 shows a schematic diagram: the transmission line method of electromagnetic wave incident inside the carbon foam. The porous structure like carbon foam has a rich microstructure, which increases the reflection and scattering of electromagnetic waves at the porous structure’s surface and carbon matrix’s interface. Each time the electromagnetic wave is in contact with the conductive carbon matrix, one more loss attenuation will be performed on the electromagnetic wave. The more micropores in the carbon foam, the more interfaces are generated in the carbon matrix, so the loss attenuation of the electromagnetic wave and the conductive carbon matrix can be effectively increased, enhancing the electromagnetic shielding effect inside the carbon foam shown in Figure 1. In the preparation and research of electromagnetic shielding materials, the characterization and statistical processing of the internal structure are incredibly significant.

2.2. U-Net. The U-Net network [22] structure is similar to the FCN [23] structure and consists of an upsampling stage and a downsampling stage. There are only convolutional layers and pooling layers in the network structure, and no fully connected layers are included. The high-resolution layer with fewer channels in the network is to solve the pixel localization problem [24], and the layer with more channels is to solve the pixel classification problem and finally realize the image segmentation task. The U-Net network includes a contraction path for finding different convolution layer features and a symmetrical expansion path for precise positioning. This method uses tiny data to complete end-to-end training and works well. The upsampling and downsampling stages use the same convolution operations and utilize the skip connection structure to connect the two layers, so the downsampling layer features can be directly transferred to the upsampling layer. This method ensures that the network can segment pixels more accurately.

2.3. Attention Mechanism. In recent years, attention-based methods have been recognized by both academia and industry for their interpretability and effectiveness. SEM first performs AvgPool on the spatial dimension, then learns the relationship between channels through two fully connected networks, uses sigmoid to obtain the Channel Attention Map, and finally multiplies the Channel Attention Map with the original features to obtain the weighted feature results. The Convolutional Block Attention Module is similar in
structure to the SEM. It performs AvgPool and MaxPool on the original features in the spatial dimension, respectively, and then uses the structure of SE to extract channel attention. After adding the two features, the attention matrix is obtained by normalization. ECA-Attention replaces the two fully connected layers in SEM with convolutions. One is that the two fully connected layers introduce too many parameters and calculations, and then, it is not necessary to calculate the attention between all channels.

2.4. Two-Dimensional Convex Hull and Its Geometric Solution. Solving the two-dimensional convex hull problem [25, 26] generally needs to be divided into two stages: constructing and solving the convex hull. When constructing a convex hull, Andrew’s algorithm is an improvement of Graham’s algorithm. For a point set in a two-dimensional space, some points can always form a convex polygon. A convex hull point set is finally formed by constructing the upper and lower convex hulls. It continuously adds new points to the convex hull and removes points that affect convexity by setting up a stack of candidate points. In addition to directly calculation, we could calculate the convex hull’s geometric characteristics by using the Rotating Calipers Algorithm (its principle is like rotating a pair of cards around a polygon).

### 3. Method

#### 3.1. Model. The segmentation criteria for supervised learning largely depend on obtaining label image priors. The more sufficient the carbon foam feature extraction, the higher the segmentation accuracy and the better the model generalization. However, the microscopic images of carbon foams are characterized by complex and diverse morphological structures and variable scale information. Therefore, U-Net is used as the basic framework in this paper. U-Net reuses the context feature information in the encoding structure to learn the contour information and can also recover the detailed information. By combining depth and shallow information, the ability to determine the target location is improved, and a multiscale module [27] is introduced to enhance the adaptability to extract scale features of carbon foam. In addition, we have listed those central and confusing symbols in Table 1.

The connection of each network layer in the Deep-Res-MixAttention model is shown in Figure 2. It mainly includes an encoding network for extracting image depth features, a decoding network for upsampling and fusing multiscale features, a classifier for classifying pixels into center regions and backgrounds, and a regressor for predicting pixel-to-center distances.

The module details are shown below:

1. **Encoding network structure**

   To solve the problem of losing spatial information in deep feature mapping, this paper uses a Multiscale Deep Residual Full Convolution Module. This module enables the network to perform self-learning and optimization according to the label image characteristics, capture more carbon foam morphological changes and scale information, improve the network’s recognition and positioning of contour features, and prevent the incomplete boundary information problem.

2. **Middle structure**

   Although the feature extraction adopts a multiscale residual module and a dense multiscale convolution module, hence it can fully extract more scales of carbon foam boundary features and internal shape features. However, noise interference such as incomplete images and the loss of detailed information during the downsampling process brings some noncarbon foam features in the upsampling as well as restoring process. The attention mechanism is introduced in the decoding part not only to make the network focus on the feature extraction and eliminate the noise interference but also to improve the recovery ability and further improve the model robustness. The middle structure is shown in Figure 3, the channel attention mechanism determines each feature channel’s weight, and the spatial attention mechanism is used to model the spatial relationship between any two pixels in the feature.

3. **Decoding network structure**

   Contrary to the downsampling effect of the encoding network, the decoding network first stacks adjacent features

---

### 1: Symbol description.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_{ij} )</td>
<td>Pixel true category label</td>
</tr>
<tr>
<td>( \mu_{ij} )</td>
<td>Pixel predicted label</td>
</tr>
<tr>
<td>( W )</td>
<td>Width of the image</td>
</tr>
<tr>
<td>( H )</td>
<td>Height of the image</td>
</tr>
<tr>
<td>( L_{\text{BCE}} )</td>
<td>The Binary Cross-Entropy Loss</td>
</tr>
<tr>
<td>( p )</td>
<td>Normalization function</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>Weight of positive sample loss</td>
</tr>
<tr>
<td>( \beta )</td>
<td>Weight of the negative sample</td>
</tr>
<tr>
<td>( L_{\text{Dce}} )</td>
<td>Sample similarity</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>Weight of Binary Cross-Entropy Loss</td>
</tr>
<tr>
<td>( \text{num}_{\text{carbon true}} )</td>
<td>Number of correct prediction</td>
</tr>
<tr>
<td>( \text{num}_{\text{carbon false}} )</td>
<td>Number of wrong predictions</td>
</tr>
<tr>
<td>( N_{\text{cell}} )</td>
<td>Cells density in the picture</td>
</tr>
<tr>
<td>( M )</td>
<td>Magnification of the scanning electron microscope</td>
</tr>
<tr>
<td>( A )</td>
<td>Area of the test area and ( n ) is the number of cells</td>
</tr>
<tr>
<td>( T )</td>
<td>True value</td>
</tr>
<tr>
<td>( P )</td>
<td>Prediction value according to ours methods</td>
</tr>
<tr>
<td>( E )</td>
<td>Error rate</td>
</tr>
</tbody>
</table>
through the Concat Layer and then uses the multiscale depth convolution layer to reduce the fused feature channel number and data redundancy. Using bilinear interpolation upsampling doubles the resolution of the feature map, so that it can be fused with the previous layer feature map and finally form a feature pyramid with a dimension of $224 \times 224$.

3.2. MSDRN: Multiscale Deep Residual Network. With a fixed $3 \times 3$ convolution kernel, the convolutional layer of the classical residual module [28] is not conducive to obtaining the carbon foam contour features with various scale transformations. To increase the network scale diversity, we improved the classical residual module. The structure of the improved residual module is shown in Figure 4.

Consider using convolutional layers with varied sizes to extract features. The multiscale residual module uses a combination of $3 \times 3$ convolutional layers and two consecutive $3 \times 3$ convolutional layers. Among them, the receptive field of two consecutive $3 \times 3$ convolutional layers are the same as that of $5 \times 5$ convolutional layers, but with less computation and stronger network nonlinear transformation ability. Adding the ReLU activation function after each convolutional layer and adding Batch Normalization can simplify the gradient disappearance problem to some extent. Use the concatenate operation to perform feature fusion on the two branches of different receptive fields, so that the latter layer can share the two branches’ information and improve the network’s multiscale adaptability. To avoid parameter overfitting caused by adding convolutional layers and reduce the difficulty of network training, we add a random Dropout between two stacked units. Dropout can make some neurons stop working and avoid overreliance on a local feature of the input image, thereby improving the generalization ability. The multiscale residual module also uses a $1 \times 1$ convolution layer to reduce network parameters by first reducing the network channel number and then restoring the channel number. This method ensures the input and output channels’ number remains unchanged, reduces the channel number
in the middle layer, and thus reduces the overall parameter amount of the network.

3.3. Loss Function. The segmentation problem can be attributed to judging whether a pixel is a background, so the most commonly used image segmentation convolutional network is Binary Cross-Entropy Loss (BCE) segmentation [29], as shown in the formulas (1) and (2). Binary Cross-Entropy Loss uses the interclass competition mechanism to make the neural network focus on samples with a high proportion and cannot solve the data class imbalance problem.

\[
L_{\text{BCE}}(P, T) = - \sum_{i=1}^{W} \sum_{j=1}^{H} \left[ t_{ij} \log(p(\mu_{ij})) + (1 - t_{ij}) \log(1 - p(\mu_{ij})) \right].
\]

\[
p(\mu_{ij}) = \text{sigmoid}\left(\pi(\mu_{ij})\right).
\]

Among them, \(t_{ij}\) is the pixel true category label, \(\mu_{ij}\) is the pixel predicted label, and \(W\) and \(H\) are the width and height of the image. For segmentation problems, if the negative sample number is greater than the positive sample number, only using Binary Cross-Entropy Loss will cause the network to fall into local minima. The Dice loss function is often used to calculate the similarity of samples. It focuses on whether the foreground image is correctly classified and does not pay attention to the background pixels. This function can effectively alleviate the imbalance between foreground and background samples, as shown in formula (3).

\[
L_{\text{Dice}}(P, T) = 1 - \frac{2 \times \sum_{i=1}^{W} \sum_{j=1}^{H} \left| t_{ij} \cap \mu_{ij} \right|}{\sum_{i=1}^{W} \sum_{j=1}^{H} \left( \left| t_{ij} \right| + \left| \mu_{ij} \right| \right)}.
\]

The total loss function represents as formula (4):

\[
L_{\text{RD}} = \gamma \times (\alpha L_{\text{bce}}(P_{\text{pos}}, T_{\text{pos}}) + \beta L_{\text{bce}}(P_{\text{neg}}, T_{\text{neg}})) + (1 - \gamma) \times (\alpha L_{\text{Dice}}(P_{\text{pos}}, T_{\text{pos}}) + \beta L_{\text{Dice}}(P_{\text{neg}}, T_{\text{neg}})).
\]

\(\alpha\) is the weight of positive sample loss, \(\beta\) is the weight of negative sample, and \(\gamma\) is the weight of Binary Cross-Entropy Loss.

3.4. Calculation of Pore Cell Evaluation Index. Further, based on the two-dimensional convex hull construction algorithm and computational geometry method, we propose an improved evaluation method (convex hull construction and bidirectional selection rotation calipers algorithm) for carbon foam diameter calculation.

We add thresholds to deal with some outliers and use the improved rotation calipers algorithm to solve the convex hull’s constructing and solving problem. Experiments show the real-time performance and robustness are improved. Pseudocode flow is as follows:

4. Experiment and Result

4.1. Experiment Platform. To make an objective and fair evaluation, all experiments are conducted in the same environment. The experimental environment is Linux, and GPU is NVIDIA RTX 3090; the operating system is Ubuntu16.04.7; the program implementation software is Python3.6, and third-party libraries such as deep learning framework tensorflow2.8 and Opencv are installed.

4.2. Datasets. The dataset used in this paper is provided by the Coal-based New Materials and Coal Coke Measurement and Control Technology Research Institute, and the carbon foam’s cell topography characterized by JSM-6480LV and Zeiss EVOMA10 scanning electron microscope (SEM) is mainly divided into carbon foam pictures of the spherical pore-like structure. Carbon foams are mainly classified into the following states: complete, incomplete at the edges, densely overlapped, and extensively damaged. As shown in Figure 5. We used data augmentation to expand the original dataset to 1040 images, and another 260 images were used as the test dataset, and the image pixel size was 224 × 224. The incomplete classification of the training set is shown in Table 2.

4.3. Result. This paper trains the U-Net [22], Res-Net, and the deep residual network with improved attention mechanisms, followed by four comparative experiments for these models. The relative proportion of training sets and verification sets is 8:2, the initial learning rate is 10e-4, we select the Adam optimization algorithm, and the batch size is 16. Different training cycles are to ensure network convergence, and that their classification and segmentation capabilities are optimal. This article uses the IoU [30, 31] coefficient to evaluate model segmentation capabilities, we generally believe its value is greater than 0.7 indicating a high repeating degree, and the segmentation effect is promising. We use
the incomplete identification accuracy: $acc_{\text{carbon}}$ as the evaluation model index and the calculation formula (5) is as follows.

$$acc_{\text{carbon}} = \frac{\text{num}_{\text{carbon true}}}{\text{num}_{\text{carbon true}} + \text{num}_{\text{carbon false}}}.$$  

num_{\text{carbon true}} means the number of correct prediction; num_{\text{carbon false}} represents the number of wrong prediction.

4.4. Network Comparative Analysis. The training hyperparameters of each model and the experimental results on the test set are shown in Table 3.

To confirm the improved network is indeed effective, comparative experiments were conducted using AlexNet, DenseNet, Xception, and Mask RCNN networks. The original Unet network has slow convergence speed, seriously missed detection, and false detection and often fails to complete the boundary prediction, and the two classification networks AlexNet and DenseNet are far less accurate than the improved Res-Unet. After analysis, it is also found the
Xception network judgment on whether the picture contains defects is inaccurate, and the location information prediction is not clear enough. The Mask-RCNN network has incomplete and discontinuous segmentation.

We also found only adding Res18-Unet, the network segmentation ability is slightly improved, and the detection accuracy of the incomplete part is improved to some extent, but the convergence speed becomes slower, the network training time increases, and the incomplete edge detection problem still occurs. By adding the spatial attention mechanism module and channel attention mechanism module, respectively, IoU and acc_carbon can be improved by up to 6.14% and 4.33%. Compared with Res50-Unet+SEBlock and Res50-Unet+CBAMBlock, the Deep-Res-MixAttention network framework proposed in this paper has a maximum improvement of 5.43% and 1.80% in IoU and acc_carbon. And after visualizing the prediction results, we found it is clearer for small boundary segmentation and more accurate for the incomplete carbon foams identification.

Comparing each network segmentation results with the labels, Figure 6 shows a 224 × 224 size image of carbon foam at 30x magnification, and the order is the real label, each network’s (C1–C9, C11) segmentation results. The U-Net segmentation is discontinuous, and the edge segmentation is not clear; Res50-Unet+CBAMBlock and Res18-Unet+ECA-Block segmentation are ideal, but they are easily disturbed by noise and lead to unclear edge segmentation. The training set data augmentation can improve the network robustness, so the network can better focus on the incomplete carbon foam. In the meantime, the MixAttention module focuses the network on the carbon foam’s edge contours. The multiscale deep residual network improves the small defect recognition and also increases the accuracy of delineating edges. In the segmentation network comparison experiment, Xception did not detect the incomplete carbon foam, and Mask RCNN had some missed detections and incomplete segmentation. Both segmentations are not as superior as the improved Deep-Res-MixAttention.

4.5. Prediction of Carbon Foam Performance. To verify the reliability of this method in practical measurement, the average diameter, maximum diameter, number, and density of...
cells were selected as evaluation indicators. Formula (6) is to calculate the cell density.

$$N_{\text{cell}} = \left(\frac{nM^2}{A}\right)^{3/2}.$$  \hspace{1cm} (6)

$N_{\text{cell}}$ is the cell density in the picture, $M$ is the magnification of the scanning electron microscope, $A$ is the area of the test area, and $n$ is the numbers of cells.

In Table 4, the $T$ represents the true value, the $P$ is the prediction value according to our method, and the $E$ is the error rate.

Compared with the true data, the error rate of the average cell diameter and the maximum diameter of cells is stably controlled within 2%. The recognition rate of the cell number is as high as 96.23%, and the error rate of cell density is 7.06%. According to the experimental data, the method proposed in this paper can effectively solve the bottleneck of carbon foam microscopic image measurement, improve the calculation efficiency, and make a progress in promoting the research and development of carbon foam in various fields.

5. Conclusion

This research presents a Deep-Res-MixAttention approach to handle the picture segmentation problem of carbon foam. This approach integrates multiscale, deep residual networks to produce various receptive fields. Comparing various attention mechanisms, we demonstrate the MixAttention method has a beneficial effect on carbon foam edge detection, and the optimized loss function enhances the fraction of positive samples, which successfully easing the dataset’s intraclass imbalance constraint. Moreover, IoU and acc carbon achieve a high level in the enhanced technique. Using the revised Rotating Calipers method, we computed the average diameter and density of carbon foam intelligently, with errors reduced to within 1.1% and 1.2% when compared to human measurement. Based on the above, we have confirmed this approach has a high application value when analyzing the electromagnetic shielding performance of carbon foam.
In the future, we intend to add a transformer mechanism to the network and increase the amount and diversity of dataset to confirm their precision and dependability have been enhanced. Explore cellpose [38] more to investigate and fix these issues. Additionally, the network topology is lightweight, making it simple to implement IoT devices.

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

**Conflicts of Interest**
The authors declare that they have no competing interests.

**Acknowledgments**
This research was funded in part by the National Science Foundation of China (No. 51868068).

**References**


