# A hybrid model based on adaptive active contour and multi-scale attention for pearlite segmentation in metallographic images

Jingwen Cong[1,2], Lixin Tang[1] and Min Xiao[3]

1 National Frontiers Science Center for Industrial Intelligence and Systems Optimization, Northeastern University, Shenyang, 110819, China
2 Key Laboratory of Data Analytics and Optimization for Smart Industry (Northeastern University), Ministry of Education, Shenyang, 110819, China.
3 Digital Intelligence Department, Xinyu Iron & Steel Group Co., Xinyu, 338001, China.

**Abstract.** The analysis of the composition and structure of high-temperature carbon steel and cast iron, particularly the distribution and content of the pearlite, is crucial for understanding its mechanical properties and overall performance. Existing computer-vision-based methods for pearlite segmentation face challenges due to variations in metallographic images and the complex structure of metallic materials. In this study, we introduce a novel automatic segmentation method for pearlite segmentation in metallographic images, addressing the limitations of existing neural network-based methods and overcoming the obstacles presented by metallographic image variations. First, a multiscale-attention-based U-Net architecture is introduced as a backbone network, which selectively focuses on spatial and channel-wise information at different scales, capturing richer hierarchical features. Additionally, an adaptive active contour model with self-tuning evolution parameters is incorporated into the backbone network, dynamically adjusting the evolution parameters based on the input image characteristics. Furthermore, A new annotated metallographic dataset called MHCSP is constructed, and experimental results reveal that the proposed approach achieves highly competitive segmentation results on metallographic images containing diverse structural compositions and varying magnifications.

**Keywords:** Metallographic image, deep learning, pearlite segmentation, active contour model.

## 1 Introduction

The analysis of composition and structure, based on microscopic images of high-temperature carbon steel and cast iron, is a vital aspect in the realm of materials science and product performance evaluation [1]. High-temperature carbon steel and cast iron, are composed of several elements, among which pearlite serves as a significant constituent. The distribution and content of pearlite in steel materials are critical factors that govern the mechanical properties of steel, such as strength, hardness, and toughness [2]. As a

result, the identification and segmentation of pearlite within these materials is essential for understanding their quality and performance characteristics.

However, in pearlite segmentation tasks, existing computer-vision-based methods have shown limitations in handling various challenges inherent to the domain. One key challenge is the substantial variation in metallographic images, including differences in scale, texture, and illumination conditions. Due to their sensitivity to such variations, neural network-based methods may not be robust enough to cope effectively, leading to decreased segmentation accuracy. At the same time, the structure of carbon steel containing pearlite is often different, and the acquisition multiples of metallographic diagrams are often not uniform. As a result, existing methods may fail to deliver satisfactory performance, particularly when it comes to accurately capturing the intricate boundary details in pearlite segmentation tasks.

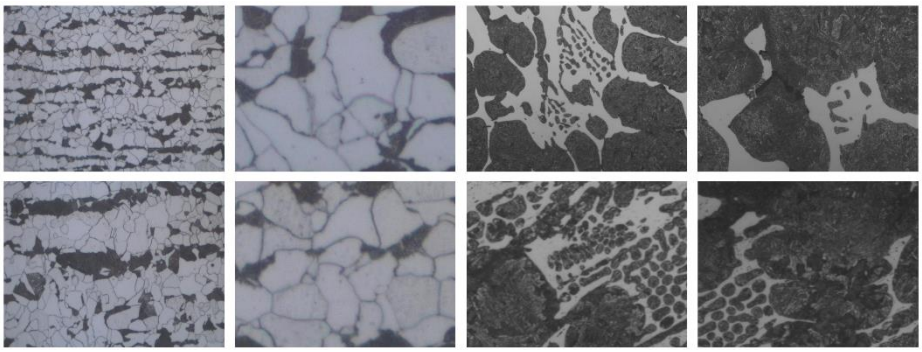

**Fig. 1.** Some examples of the self-built metallographic dataset MHCSP. The first and second columns are ferrite-pearlite images at different magnifications. The third and fourth columns are transformed ledeburite images at different magnifications.

In recent years, there have been many studies attempting to use active contour models as a complementary technique to capture more accurate boundaries in semantic segmentation tasks [3][4][5][6]. The active contour models, though a powerful tool for image segmentation, exhibits certain limitations that affect its performance in metallographic component segmentation tasks. A primary drawback is the dependence of models on fixed evolution parameters. Choosing an optimal set of parameters for diverse images with varying characteristics is challenging, which can lead to suboptimal performance when applied to different image types without fine-tuning. Furthermore, the active contour model is sensitive to initial contour placement, which can result in convergence to local minima rather than the desired global minimum, causing inaccurate segmentation.

To address these challenges, we propose a more robust and automatic segmentation framework that can effectively handle the inherent complexities of metallographic images and deliver high-quality pearlite segmentation results. In particular, a novel multiscale-attention-based U-Net architecture is first introduced as a backbone network. Unlike the common U-Net with attention module [7][8], the key innovation lies in the

integration of multiscale attention modules within both the encoder and decoder parts of the U-Net, as well as in the skip connections and bottleneck layer. The multiscale attention mechanism enables the model to capture richer hierarchical features at different scales by selectively focusing on relevant spatial and channel-wise information.

Subsequently, an adaptive active contour model with self-tuning evolution parameters is introduced and embedded in the backbone network. This adaptive model dynamically adjusts the evolution parameters based on the characteristics of the input image. By employing this module, the framework refines the initial segmentation by evolving the contour to better adhere to the boundaries of the target objects, ensuring optimal segmentation performance without manual parameter tuning and intervention. In addition, A new metallographic dataset called MHCSP was constructed, all experiments were performed on them. Experimental results reveal that the proposed approach achieves highly competitive segmentation outcomes on metallographic datasets, encompassing diverse structural compositions and varying magnifications, in comparison to existing image segmentation algorithms.

Overall, the key contributions of this paper are as follows:

1. An automatic segmentation framework is proposed, which integrates an active contour model with an improved attention module into a deep learning framework that can effectively handle the inherent complexity of metallographic images.

2. A multi-scale attention mechanism is introduced, which can selectively focus on spatial and channel information at different scales and capture richer hierarchical features.

3. An adaptive active contour model with self-tuning evolution parameters is presented, which can dynamically adjust the evolution parameters based on different images.

4. A self-built metallographic dataset has been constructed. Experimental results demonstrate that our approach can provide high-quality segmentation results.

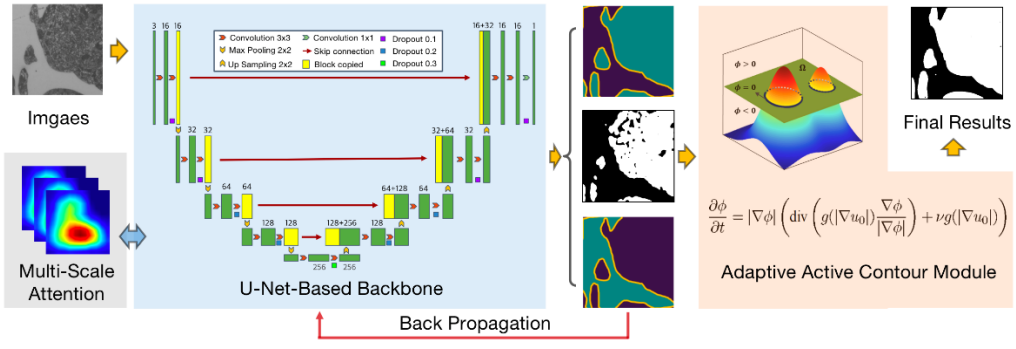


**Fig. 2.** Schematic of the proposed model.

# 2 Related Work

## 2.1 Metallographic Image Segmentation

Evolution in the field of metallographic image segmentation algorithms has seen a transition from traditional segmentation methods towards those based on neural networks. A novel watershed metallographic image segmentation algorithm was presented by Chen *et al.* [9], addressing the challenges of noise and regional texture features in metallographic images through the integration of Bayesian rules and an enhanced similar region merging algorithm. Chen *et al.* [10] proposed an improved FCM metallographic image segmentation algorithm, which improves the noise resistance and operational efficiency by integrating the two-dimensional distance measurement function and the improved objective function into the traditional FCM algorithm. Li *et al.* [11] introduced a metallographic image segmentation algorithm that leverages superpixels and transfer learning. An advanced SNL-UNet image segmentation network was suggested by Zhang *et al.* [12] that merges an enhanced non-local attention block with U-Net to address the accuracy issues caused by metallographic images captured from various angles and positions. Despite their success in certain tasks, these algorithms predominantly focus on local or pixel-level attributes of the object, often disregarding its geometric properties. Such issues induce limitations in the precision and operational efficiency of the segmentation algorithms, frequently resulting in an inability to accurately outline the boundary specifics of objects during identification.

## 2.2 Active Contour Model

The active contour algorithm is widely employed in image segmentation tasks due to its rapid convergence speed and capability to consider the geometric and contour features of objects comprehensively. Chan and Vese [13] first proposed an image segmentation algorithm independent of gradient information, constructing energy functions using regional features and paving a new direction for active contour algorithms. Zhang *et al.* [14] introduced an improved level set segmentation algorithm, combining methods such as the maximum likelihood energy function and modeling objects as Gaussian distributions to address uneven intensity in segmented images. Ma *et al.* [15] presented a new multi-scale driven active contour segmentation model based on multimodal images, merging random forest and active contour model. Liu *et al.* [16] proposed a novel active contour segmentation algorithm for SAR images, integrating region-based and edge-based active contour models to solve edge detection problems under varying initial scales. Hoogi *et al.* [17] introduced an active contour parameter adaptation method that adaptively adjusts parameters and window size based on convolutional neural network (CNN) output and the size and texture of objects, achieving satisfactory segmentation performance in low-contrast and low-resolution images. Nevertheless, factors such as the consideration of pixel-level features, contour initialization, and evolutionary parameter settings affect the segmentation performance of the active contour algorithm to some extent.

# 3 Proposed Methods

## 3.1 Multiscale-attention-based Backbone Network

In this work, a novel multiscale-attention-based U-Net architecture is first proposed for robust and efficient image segmentation. The original U-Net [18], while proven to be highly effective in various segmentation tasks, has some limitations. One key shortcoming is its inability to selectively focus on relevant features at different spatial scales and channel-wise information. The integration of attention mechanisms [19] has been proposed in several variants of U-Net to address these limitations; however, most existing attention-based U-Net models incorporate attention modules only in specific parts of the network, such as skip connections or after certain convolutional layers. Consequently, these models may not fully exploit the potential of attention mechanisms across the entire network hierarchy.

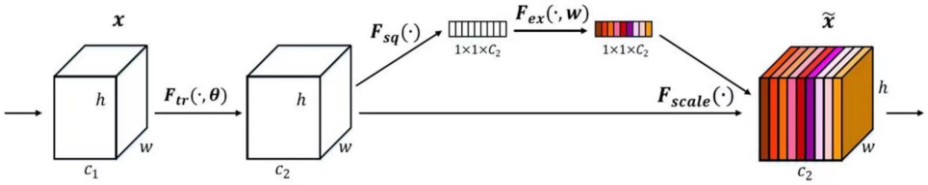


**Fig. 3.** Structure of single attention mechanism.

To overcome these limitations, our multiscale-attention-based U-Net introduces multiscale attention modules in both the encoder and decoder parts, as well as in the skip connections and bottleneck layer. These attention modules enable the model to selectively focus on the most relevant spatial and channel-wise information at various scales and levels of abstraction, leading to improved feature extraction and representation.

In the encoder part, after each convolutional layer, we add a multiscale attention module that combines self-attention and channel attention mechanisms, followed by the down-sampling operation. This allows the model to effectively capture and utilize different feature scales during the encoding process. In the decoder part, we perform the up-sampling operation and incorporate the multiscale attention module after each up-sampling layer. The output of the attention module is then concatenated or added to the corresponding feature map from the encoder through skip connections, enabling the decoder to refine the segmentation results by leveraging attention-weighted high-resolution features. Additionally, we include a multiscale attention module in the bottleneck layer, further enhancing feature extraction by capturing the most relevant contextual information across different scales. This is crucial for accurate segmentation, especially in cases where the target structures exhibit significant variations in scale, shape, or appearance.

By leveraging multiscale attention mechanisms throughout the U-Net architecture, our proposed model exhibits several key advantages compared to the original U-Net and other attention-based U-Net variants. First, it effectively captures and utilizes the most relevant features at multiple scales and levels of abstraction, leading to improved segmentation performance. Second, the multiscale attention modules provide the model with a more flexible and adaptive feature representation, allowing it to handle complex and challenging segmentation scenarios. Finally, the integration of attention mechanisms in the entire network hierarchy ensures that the model is better equipped to deal with scale variations and other challenging characteristics often encountered in real-world image segmentation tasks.

### 3.2 Adaptive Active Contour

In order to further capture the fine edges of the segmented objects, the idea of active contours is considered to be added to the proposed segmentation framework. The active contour model is a widely-used technique for image segmentation and object boundary extraction. The main idea behind the active contour model is to represent the contour of an object in an image as a parametric curve that deforms under the influence of internal and external forces. These forces are carefully designed to ensure that the contour converges to the desired object boundary. The Chan-Vese (CV) model [9], proposed by Tony F. Chan and Luminita A. Vese in 2001, is a popular variant of the active contour model. The CV model is particularly effective in segmenting images with a piecewise constant structure, where the object and background regions have relatively uniform intensities. In the CV model, the energy function is defined by the following equation:

$$\begin{aligned} F(c_1, c_2, \phi) = \mu \int_{\Omega} \delta\big(\phi(x,y)\big)\, |\nabla \phi(x,y)|\, dx\, dy \\ + \nu \int_{\Omega} H\big(\phi(x,y)\big)\, dx\, dy \\ + \lambda_1 \int_{\Omega} |u_0(x,y) - c_1|^2\, H\big(\phi(x,y)\big)\, dx\, dy \\ + \lambda_2 \int_{\Omega} |u_0(x,y) - c_2|^2 \Big(1 - H\big(\phi(x,y)\big)\Big)\, dx\, dy. \end{aligned} \quad (1)$$

In this context, the first two terms act as regular constraints on the length and area of the curve. The third term (internal force) is responsible for preserving the smoothness and regularity of the contour, ensuring it does not become overly stretched or bent. Conversely, the fourth term (external force) is derived from the image data and guides the contour towards the desired object boundary, accounting for features such as edges, gradients, and textures.

While the CV model has proven to be a powerful tool for image segmentation, it is not without its limitations. One of the primary drawbacks of the CV model is its reliance on fixed evolution parameters. These parameters, such as the weighting coefficients for the data fitting and regularizing terms, play a crucial role in the segmentation results. However, choosing a single set of optimal parameters for diverse images with varying

characteristics is a challenging task. Consequently, the performance of the CV model may be suboptimal when applied to different types of images without fine-tuning the parameters for each specific case.

To address these limitations, we extend the approaches in [20][21] and propose an adaptive CV model with self-tuning evolution parameters. This adaptive model dynamically adjusts the evolution parameters based on the characteristics of the input image, ensuring optimal segmentation performance without manual intervention. This adaptive mechanism is defined by the following equation:

$$\lambda_1(x,y) = exp\left(\frac{2 - P_{seg}(x,y)}{1 + P_{seg}(x,y)}\right),$$
$$\lambda_2(x,y) = exp\left(\frac{1 + P_{seg}(x,y)}{2 - P_{seg}(x,y)}\right), \tag{2}$$

where $\lambda_1(x,y)$ and $\lambda_2(x,y)$ represent variable parameter functions. The curve expands at coordinates (x, y) when $\lambda_2(x,y) < \lambda_1(x,y)$.$P_{seg}$ denotes the output of the backbone network. Therefore, according to (2), the energy $F$ can be rewritten as

$$\begin{aligned} F(c_1,c_2,\phi) = \mu &\int_\Omega \delta\big(\phi(x,y)\big)\,|\nabla\phi(x,y)|\,dx\,dy \\ &+\nu \int_\Omega H\big(\phi(x,y)\big)\,dx\,dy \\ &+\lambda_1(x,y)\int_\Omega |u_0(x,y)-c_1|^2\,H\big(\phi(x,y)\big)\,dx\,dy \\ &+\lambda_2(x,y)\int_\Omega |u_0(x,y)-c_2|^2\left(1-H\big(\phi(x,y)\big)\right)dx\,dy, \end{aligned} \tag{3}$$

with

$$H(z) = \begin{cases} 1, & if\ z \geq 0 \\ 0, & if\ z < 0, \end{cases} \delta_0(z) = \frac{d}{dz}H(z) \tag{4}$$

where $H$ and $\delta$ represent the Heaviside function and the Dirac metric, respectively. According to (2)(3)(4), the evolution equation can also be written as

$$\begin{cases} \dfrac{\partial\phi}{\partial t} = |\nabla\phi| div\left(\dfrac{\nabla\phi}{|\nabla\phi|}\right), & t \in (0,\infty),\ \ x \in \mathbb{R}^2 \\ \phi(0,x,y) = \phi_0(x,y), & x \in \mathbb{R}^2. \end{cases} \tag{5}$$

By incorporating an adaptive mechanism, the model becomes more robust and can handle a broader range of image types and conditions. Moreover, the adaptive CV model can potentially reduce the sensitivity to initial contour placement, improving the likelihood of converging to the correct global minimum. The enhanced flexibility and adaptability of the proposed model make it a promising solution for a variety of image segmentation tasks, paving the way for more accurate and efficient segmentation results across diverse applications.

### 3.3 Hybrid Segmentation Framework

The proposed hybrid segmentation framework is illustrated in Fig. 2. The motivation behind this framework is to leverage the strengths of both the backbone network and the parametric adaptive CV model to achieve high-quality segmentation results. Initially, the backbone network is trained using pixel-level annotations to generate the preliminary segmentation predictions. The multiscale attention mechanism within the U-Net architecture enables the network to capture contextual information and focus on relevant regions in the image, resulting in more accurate and detailed segmentation outcomes.

Following this, the predicted segmentation results are fed into the parametric adaptive CV model, where the contour of the predicted foreground region serves as the initial contour. By employing this model, the framework refines the initial segmentation by evolving the contour to better adhere to the boundaries of the target objects. The parametric adaptability of the CV model ensures optimal performance across diverse images without the need for manual parameter tuning.

The visualized evolutionary result produced by the active contour module serves as the final segmentation output, providing refined and accurate delineation of the target objects in the image. The proposed framework effectively combines the strengths of both the multiscale-attention-based U-Net and the adaptive CV model, paving the way for improved segmentation performance across a wide range of applications. This innovative approach demonstrates the potential of integrating deep learning-based segmentation techniques with classical methods, resulting in a powerful and versatile solution for challenging image segmentation tasks.

## 4 Experiments

### 4.1 Dataset

In this study, the task involves pixel-level segmentation of pearlite in high-temperature carbon steel metallographs. Due to the lack of publicly available metallographic datasets with pixel-level labels, a new dataset called MHCSP was constructed, comprising 60 self-collected metallographic images of high-temperature carbon steel containing pearlite. Relevant experts provided manual pixel-level annotations. To evaluate the performance of the proposed method on pearlite segmentation in different microstructure images, the MHCSP dataset includes 30 ferrite-pearlite microstructure images and 30 transformed ledeburite images. The transformed ledeburite is composed of pearlite and cementite at temperatures below 727 ℃. Considering the varying resolutions and magnifications of electron microscope images, the MHCSP dataset employs metallographic images with different magnifications to ensure the generalization capability the proposed method. Some examples are shown in Fig. 1. The experimental setup comprises 40 training images, 10 validation images, and 10 test images.

### 4.2 Experimental Setup

**Implementation Details.** The proposed model is implemented using the PyTorch framework. The network was trained on an NVIDIA GTX 1080 GPU with a batch size of 2, employing the Adam optimizer with a learning rate of 1e-4 and a weight decay of 1e-5. The model was trained for 100 epochs with early stopping based on the validation loss to avoid overfitting. During training, standard data augmentation techniques, including random horizontal flipping, vertical flipping, and rotation, were employed to increase the diversity of the training set and improve the generalization capabilities of models. The loss function chosen for training our model was a combination of categorical cross-entropy and the Dice loss, which encourages the model to focus on both the pixel-wise classification and the spatial coherence of the segmented regions.

**Evaluation Metrics.** To assess the performance of our proposed segmentation model, we selected the mean Intersection over Union (mIoU) as the primary evaluation metric. The mIoU is a widely used evaluation criterion in image segmentation tasks as it provides a comprehensive measure of the ability of models to accurately predict the target regions, accounting for both false positives and false negatives. Additionally, it is less sensitive to class imbalance issues, which are common in segmentation tasks.

### 4.3 Results and Comparisons

This study aims to compare the performance of five distinct models for image segmentation. The models include PSPNet [22], SegNet [23], FCN8s [24], DeepLabv3+ [25], U-Net [18], and our proposed model. Each of these models employs specialized techniques and strategies to tackle the challenges associated with image segmentation.

**Results on Ferrite-pearlite Images in MHCSP Dataset:** The experimental results are displayed in Fig. 4. It is can be seen that PSPNet [22] faces challenges in delineating target contours precisely, while SegNet [23] exhibits a tendency to overestimate the extent of certain targets. FCN8s [24] is constrained by its capacity to identify the local scope of specific targets, leading to partial segmentation. DeepLabv3+ [25], in some cases, produces noisy segmentation results, affecting the overall quality of the output. U-Net [18], on the other hand, struggles with preserving fine-grained details and maintaining accurate boundaries between adjacent objects. In stark contrast, the proposed model showcases remarkable performance in addressing these limitations, delivering accurate and reliable segmentation results. Not only does the proposed model precisely delineate target contours, but it also maintains the appropriate scale of segmented targets, ensuring accurate localization of the full extent of objects. Additionally, the proposed model effectively preserves intricate details and establishes clear boundaries between neighboring objects, contributing to the overall quality of the segmentation output.

As for quantitative metrics, Table I shows the mIoU values for PSPNet [22], SegNet [23], FCN8s [24], DeepLabv3+ [25], and U-Net [18] to be 79.62%, 90.61%, 42.19%, 80.12%, and 90.47%, respectively. The mIoU value of the proposed model is significantly higher, indicating its superior performance in the segmentation task. This advantage is further supported by qualitative visualizations of the segmentation results,

where the proposed model consistently generates more accurate and precise contours compared to the other models. Upon analyzing the experimental findings, it becomes clear that the proposed model excels over the other models in various aspects, including accuracy, stability, and effectiveness.

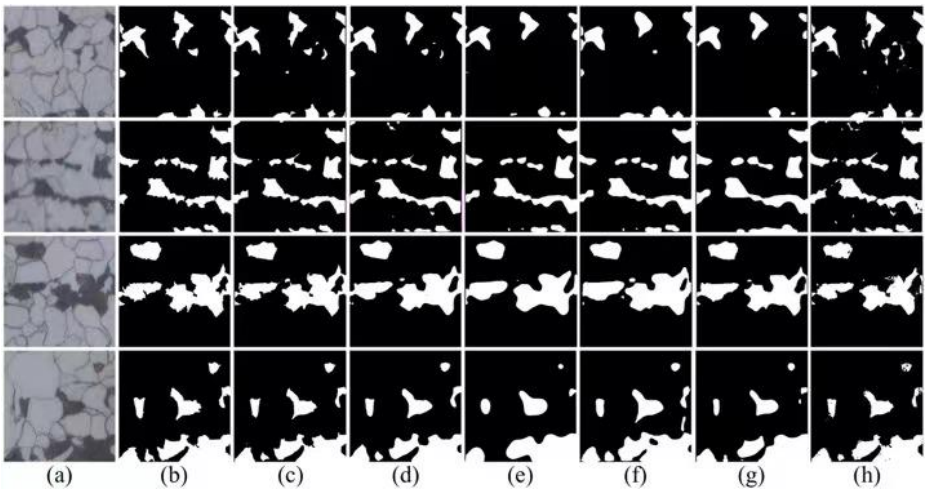


**Fig. 4.** Comparisons of segmentation results on MHCSP dataset. (a) Image. (b) Ground truth. (c) Our model. (d) PSPNet. (e) SegNet. (f) FCN8s. (g) DeepLabv3+. (h) U-Net.

**Table 1.** Quantitative results on the MHCSP dataset

| Methods | Ferrite-pearlite Images (mIoU) | Transformed ledeburite Images (mIoU) |
|---|---|---|
| PSPNet [22] | 79.62% | 84.33% |
| SegNet [23] | 90.61% | 91.80% |
| FCN8s [24] | 42.19% | 90.22% |
| DeepLabv3+ [25] | 80.12% | 89.52% |
| U-Net [18] | 90.47% | 89.43% |
| Our method | 95.75% | 96.31% |

**Results on Transformed ledeburite Images in MHCSP Dataset:** The experimental results are illustrated in Fig. 5. It is can be seen that PSPNet [5-1] encounters difficulties in accurately identifying thin structures, while SegNet [5-2] tends to merge closely spaced targets, leading to imprecise segmentation. FCN8s [5-3] exhibits reduced performance in distinguishing between similar textures, resulting in confusion between different objects. DeepLabv3+ [5-4], in certain instances, generates oversmoothed boundaries, compromising the overall sharpness of the outputs. U-Net [19], conversely, has trouble in managing occlusions and coping with variations in object appearances. In marked contrast, the proposed model demonstrates outstanding capabilities in

overcoming these challenges, yielding precise and reliable segmentation outcomes. The proposed model not only excels in detecting thin structures but also effectively separates closely spaced targets. It adeptly distinguishes between objects with similar textures and copes with variations in object appearances. Furthermore, the proposed model maintains sharp boundaries and handles occlusions gracefully, contributing to the high-quality segmentation results.

Regarding quantitative metrics, as depicted in Table I, the mIoU values for PSPNet [5-1], SegNet [5-2], FCN8s [5-3], DeepLabv3+ [5-4], and U-Net [19] are 84.33%, 91.80%, 90.22%, 89.52%, and 89.43%, respectively. In comparison, the mIoU value for the proposed model is substantially higher, reflecting its superior performance in the segmentation task. This advantage is further corroborated by the qualitative visualizations of the segmentation results, where the proposed model consistently generates more accurate and refined contours compared to the other models. Upon examining the experimental results, it is apparent that the proposed model outshines the competing models in various aspects, including precision, robustness, and adaptability. This success stems from the inventive methods and strategies implemented by the proposed model, which adeptly address the challenges inherent in pearlite segmentation tasks.

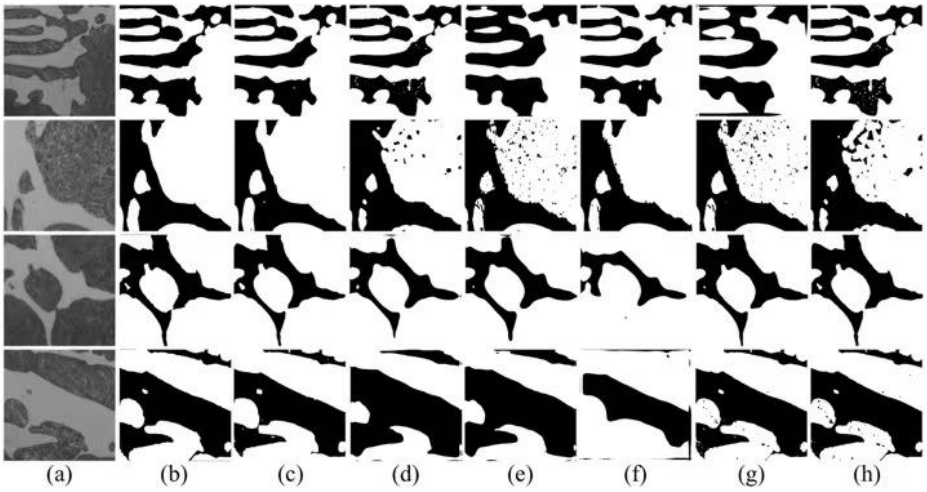


**Fig. 5.** Comparisons of segmentation results on MHCSP dataset. (a) Image. (b) Ground truth. (c) Our model. (d) PSPNet. (e) SegNet. (f) FCN8s. (g) DeepLabv3+. (h) U-Net.

### 4.4 Discussion and Analysis

To validate the sensitivity of the proposed method to the number of training images, a comprehensive series of experiments were conducted to study in more depth the performance of our method with different numbers of training samples. Specifically, five distinct experimental configurations were conducted across two diverse datasets, each varying in the number of images allocated to the training set. The experimental results are shown in Table 2, where Configuration indicates the ratio of training set, test set and validation set. Intriguingly, these results reveal that the model's performance

metrics mIoU exhibited only marginal fluctuations across the different configurations. This stability in the face of varying training set sizes lends credence to the robustness of our approach for semantic segmentation of metallography components. It suggests that the feature representations captured by our architecture are sufficiently generalizable, negating substantial overfitting or underfitting, regardless of the training sample number.

**Table 2.** Dataset Performance Metrics mIoU across the Different Configurations

| Configurations | Ferrite-pearlite Images (mIoU) | Transformed ledeburite Images (mIoU) |
|---|---|---|
| 20:5:5 | 95.75% | 96.31% |
| 16:7:7 | 93.97% | 96.02% |
| 14:8:8 | 94.32% | 96.38% |
| 12:9:9 | 93.56% | 93.83% |
| 10:10:10 | 93.14% | 96.03% |

## 5 Conclusion

In this paper, we presented a robust and automatic segmentation method for pearlite segmentation in metallographic images, overcoming the limitations of existing neural network-based methods. The proposed approach integrates a multiscale-attention-based U-Net architecture with an adaptive active contour model, allowing the framework to effectively handle the inherent complexities of metallographic images and deliver high-quality segmentation results. The experimental evaluation, conducted on a newly constructed metallographic dataset, validates the effectiveness and robustness of the proposed framework. This study contributes to the advancement of image segmentation techniques in the domain of metallographic imaging and has the potential to further enhance the analysis of metallographic structures.

**Acknowledgments.** This work was supported by the Major Program of National Natural Science Foundation of China (72192830, 72192831), the 111 Project (B16009)

## References

1. Callister W D, Rethwisch D G. Materials science and engineering: an introduction[M]. New York: John wiley & sons, 2007.
2. Deng Y, Qiao L, Zhu J, et al. Mechanical performance and microstructure prediction of hypereutectoid rail steels based on BP neural networks[J]. IEEE access, 2020, 8: 41905-41912.
3. Cheng D, Liao R, Fidler S, et al. Darnet: Deep active ray network for building segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 7431-7439.

4. Hatamizadeh A, Sengupta D, Terzopoulos D. End-to-end trainable deep active contour models for automated image segmentation: Delineating buildings in aerial imagery[C]//Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XII 16. Springer International Publishing, 2020: 730-746.
5. Zhang M, Dong B, Li Q. Deep active contour network for medical image segmentation[C]//Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part IV 23. Springer International Publishing, 2020: 321-331.
6. Qi H, Cheng L, Kong X, et al. WDLS: Deep Level Set Learning for Weakly Supervised Aeroengine Defect Segmentation[J]. IEEE Transactions on Industrial Informatics, 2023.
7. He N, Fang L, Plaza A. Hybrid first and second order attention Unet for building segmentation in remote sensing images[J]. Science China Information Sciences, 2020, 63: 1-12.
8. Qi H, Ji H, Zhang J, et al. Defect Classification Method of X-ray Images Based on Improved U-Net[C]//Proceedings of the 4th International Conference on Advanced Information Science and System. 2022: 1-7.
9. Chen Y, Chen J. A watershed segmentation algorithm based on ridge detection and rapid region merging[C]//2014 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). IEEE, 2014: 420-424.
10. Chen L, Han Y, Cui B, et al. Two-dimensional fuzzy clustering algorithm (2DFCM) for metallographic image segmentation based on spatial information[C]//2015 2nd International Conference on Information Science and Control Engineering. IEEE, 2015: 519-521.
11. Chen D, Liu S, Liu F. Metallographic Image Segmentation Method Based on Superpixels Algorithm and Transfer Learning[C]//2020 Chinese Control And Decision Conference (CCDC). IEEE, 2020: 1922-1926.
12. Zhang L, Yao H, Xu Z, et al. Segmentation and Measurement of Superalloy Microstructure Based on Improved Nonlocal Block[J]. IEEE Access, 2022, 10: 32418-32425.
13. Chan T F, Vese L A. Active contours without edges[J]. IEEE Transactions on image processing, 2001, 10(2): 266-277.
14. Zhang K, Zhang L, Lam K M, et al. A level set approach to image segmentation with intensity inhomogeneity[J]. IEEE transactions on cybernetics, 2015, 46(2): 546-557.
15. Ma C, Luo G, Wang K. Concatenated and connected random forests with multiscale patch driven active contour model for automated brain tumor segmentation of MR images[J]. IEEE transactions on medical imaging, 2018, 37(8): 1943-1954.
16. Liu C, Xiao Y, Yang J. A coastline detection method in polarimetric SAR images mixing the region-based and edge-based active contour models[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(7): 3735-3747.
17. Hoogi A, Subramaniam A, Veerapaneni R, et al. Adaptive estimation of active contour parameters using convolutional neural networks and texture analysis[J]. IEEE transactions on medical imaging, 2016, 36(3): 781-791.
18. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015: 234-241.
19. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
20. Hatamizadeh A, Hoogi A, Sengupta D, et al. Deep active lesion segmentation[C]//Machine Learning in Medical Imaging: 10th International Workshop, MLMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings 10. Springer International Publishing, 2019: 98-105.

21. Hoogi A, Subramaniam A, Veerapaneni R, et al. Adaptive estimation of active contour parameters using convolutional neural networks and texture analysis[J]. IEEE transactions on medical imaging, 2016, 36(3): 781-791.
22. Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2881-2890.
23. Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(12): 2481-2495.
24. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3431-3440.
25. Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the European conference on computer vision (ECCV). 2018: 801-818.