Image Anal Stereol 2025;44:1-9 doi: 10.5566/ias.3279
*Original Research Paper*

# APPLICATION OF SEMI-SUPERVISED MEAN TEACHER TO ROCK IMAGE SEGMENTATION

JIASHAN LI[1] AND YUXUE WANG[✉1]

[1]Northeast Petroleum University, Daqing 163318, China
e-mail: lijiashan1012@163.com; wangyx0629@nepu.edu.cn
*(Received May 27, 2024; revised November 14, 2024; accepted November 22, 2024)*

## ABSTRACT

Accurate segmentation of rock images is crucial for studying the internal structure and properties of rocks. To address the issue of requiring a large number of labeled images for model training in traditional image segmentation methods, this paper proposes an improved semi-supervised Mean Teacher algorithm based on ResNet34-UNet. This method achieves relatively accurate rock image segmentation using only a small amount of labeled data. Initially, we use ResNet34-UNet as the base model to create Student Model and Teacher Model with identical structures. Then, we introduce self-attention mechanism into the semi-supervised Mean Teacher algorithm to further enhance its performance in rock image segmentation. Finally, by comparing the performance of supervised and semi-supervised Mean Teacher algorithms on image segmentation tasks, we validate the effectiveness of semi-supervised learning in rock image segmentation.

Keywords: Image segmentation; ResNet; Rock image; Self-attention; Semi-supervised learning; Unet.

## INTRODUCTION

Rock images are a vital source of geological data that are essential for understanding the composition and characteristics of rocks. In contrast to natural photographs, accurate rock image segmentation typically depends on experts annotating the images based on their own experience, which makes obtaining large-scale labeled datasets a difficult and time-consuming operation. The topic of image segmentation has seen a lot of activity and applications recently (Cai *et al.* (2024); Huang *et al.* (2024); Lou *et al.* (2024); Liu *et al.* (2024); Mazher *et al.* (2024); Liu *et al.* (2023); Zhao *et al.* (2024)). The work of rock segmentation can be completed more precisely by combining image processing and rock image recognition technologies. This has significant practical implications for geological research and resource optimization.

The application of picture segmentation in petrology has grown with the swift advancement of deep learning and computer vision technology. While segmentation can be achieved to some extent using traditional pixel-based techniques like threshold segmentation, region-growing, and edge detection, handling complicated rock textures and structures can frequently be challenging.In the field of image segmentation, deep learning-based techniques like U-Net (Ronneberger *et al.* (2015)), SegNet (Badrinarayanan *et al.* (2017)), and Mask R-CNN (Murray *et al.* (2021)) have made impressive strides. These techniques are well-suited for challenging feature extraction and segmentation tasks in rock images and can better capture the semantic information of the image.

Enhancing the accuracy and efficiency of segmentation can be achieved by processing the complex features of rock images more efficiently through the use of deep learning methods, particularly convolutional neural networks (CNN) (Cao *et al.* (2022);Chen *et al.* (2024);Niu *et al.* (2020);Roslin *et al.* (2023)). Researchers are able to do automatic rock images segmentation by building suitable neural network topologies and optimization algorithms. This lessens the need for manual annotation by experts and expedites data processing and analysis.

RockSeg network combining CNN and Transformer methods (Fan *et al.* (2023)) is proposed to effectively segment rock images in deep space environments; an automatic segmentation method combining texture analysis and deep learning is proposed (Manzoor *et al.* (2023)) in the task of segmenting the porosity and mineral composition of digital rock images.An algorithm combining local adaptive thresholding and a improved watershed transform (Dong and Jiang (2013)) was used in the rock image segmentation task. This algorithm solves the incorrect segmentation problem caused by sticking, stacking, and edge blurring in the blasting rock images more quickly and accurately. Xia Lin et al. also proposed an algorithm for classifying microscopic images of rock sheets based on the deep residual contraction network and the attention mechanism (Zhang *et al.* (2023)), which is one of several techniques to improve the accuracy of segmentation of rock images.

However, in order to train rock images segmentation models, these approaches typically require a big amount of labeled data, which will take a lot of time and labor and may restrict the algorithm use in practical situations. Semi-supervised learning methods prove to be a useful substitute in resolving this issue. This research is motivated by the fact that semi-supervised learning algorithms may learn features from a significant amount of unlabeled image data, reducing the demand for labeled data and improving model performance and generalization ability.

Semi-supervised image segmentation is one of the important methods for segmenting images, which can effectively use limited labeled data to improve the performance and generalization ability of the model by combining labeled and unlabeled data. In complex scenarios such as rock images, semi-supervised learning methods provide new ideas and solutions to improve model performance and reduce the demand for labeled data, such as a method combining simple linear iterative clustering algorithm (SLIC) and semi-supervised self-training has been proposed (Liu and Lv (2023)) for automated segmentation and component recognition of rock images, and an improved semi-supervised SVM-FCM based on chaotic algorithm (CSVM-CMF) was proposed (Liang and zou (2020))for segmentation of pores and rocks. Semi-supervised learning techniques offer fresh ideas and solutions to improve the model performance and reduce the demand for labeled data in complex scenarios like rock images. In addition, semi-supervised learning has also shown potential and importance in medical image analysis (Weng *et al.* (2024)) disease diagnosis (Han *et al.* (2024);Sinha *et al.* (2023)), etc. such as in the segmentation of liver images in clinical diagnosis (Huang *et al.* (2024)) and the segmentation of skin lesions (Zhang *et al.* (2022)), which further validate the value and advantages of semi-supervised image segmentation in a variety of complex scenarios, and provide useful references and lessons for future image segmentation research and applications.

Therefore, this paper proposes an improved Mean Teacher algorithm based on Resnet34-UNet for rock segmentation. By using a significant amount of unlabeled data for feature learning, the approach lessens the reliance on labeled data and enhances the robustness and accuracy of segmenting rock images. Furthermore, this research introduces a self-attention mechanism that helps the model better focus on the significance of various parts in the image, Compared to traditional semi-supervised learning methods, our algorithm demonstrates higher robustness and accuracy in dealing with complex textures of rocks. In particular, with limited labeled samples, by introducing a self-attention mechanism, the model better focuses on the importance of different regions in the image and improves the ability to capture detailed information such as rock structure and texture. We anticipate that our work will advance the development and use of rock images segmentation, lead to new discoveries, and advance related fields' practice and research.

# RELATED WORKS

## RESNET34-UNET

A neural network model called ResNet34-Unet combines the architectures of Unet and ResNet34 to perform image segmentation tasks. In this model, Unet serves as the decoder component, producing the segmentation results of the picture, and ResNet34 serves as the encoder component, extracting the characteristics of the image. ResNet34-Unet is able to provide precise picture segmentation and feature extraction by fusing the potent feature extraction capabilities of ResNet34 with the segmentation capabilities of Unet.

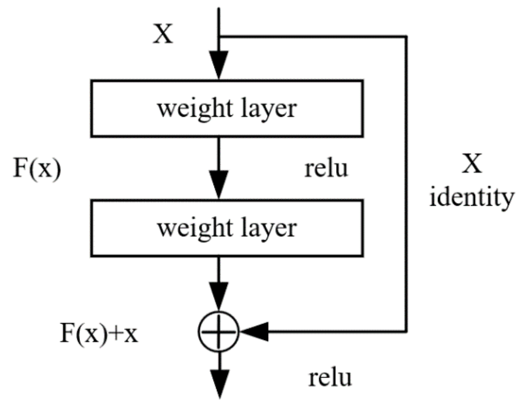


Fig. 1. *Residual structure*

Among the ResNet family, ResNet34 is a traditional deep residual network model with a stronger feature extraction capacity and a deeper network structure. The model utilizes residual structure, as shown in Fig. 1, which through a shortcut connection, the input signal is directly added to the output signal, so as to realize the identity mapping, solve the problem of gradient disappearance and gradient explosion in the training process, and improve the training efficiency of the deep network. ResNet34 extends the network depth while preserving model performance in contrast to the shallower network structure, which allows it to more effectively capture the image's semantic content and abstract properties.

The Unet structure has jump connections and an upsampling layer in addition to acting as a decoder. By using the jump connection to retain more detailed information, the decoder improves the accuracy of the segmentation process by reducing the feature maps extracted by the encoder to a segmentation result of the same size as the input image through upsampling.

In summary, ResNet34-Unet integrates the benefits of both ResNet34 and Unet, demonstrates strong performance in image segmentation assignments, and can successfully achieve precise segmentation and feature extraction of rock images.

## MEAN TEACHER ALGORITHM

In order to enhance the model performance and capacity for generalization, semi-supervised learning combines a large number of unlabeled examples with a limited number of labeled samples. Semi-supervised learning has a larger range of application scenarios than supervised learning because it is more difficult to gather huge amounts of labeled data in real-world scenarios. Conversely, semi-supervised learning can train a model with a similar level of accuracy to a supervised model by combining a big amount of unlabeled data with a little amount of labeled data. Semi-supervised learning methods are mainly classified into two categories: consistent regularization method and proxy labeling.

In this paper, the Mean Teacher algorithm based on consistent regularization (Tarvainen and Valpola (2017)) is used to segment rock images with the flow chart shown in Fig. 2, the whole algorithm consists of two networks, teacher network and student network,in which the student network updates the parameters through gradient descent, and the teacher network updates the parameters through the student network parameters. The algorithm mainly includes the following five steps:
(1) Set up the Teacher networks $G_t$ and Student networks $G_s$, which often share the same network topology.
(2) Load the labels training datasets $D_l = (x_i^l, y_i^l)_{i=1}^N$ and the unlabeled training dataset $D_l = (x_i^u)_{i=1}^M$, where $x_i^l$ is the labeled image, $y_i^l$ is the corresponding label, $x_i^u$ is the unlabeled image, $N$ is the number of labeled samples, and $M$ is the number of unlabeled samples.
(3) Describe the loss function $L_{total} = L_{labeled} + \omega^* L_{unlabeled}$, where $\omega$ is a hyperparameter that weighs the two losses.
(4) Training loop, each time through the training process:
a. Draw a batch of data from the labeled training dataset.
b. Draw a batch of data from the unlabeled training dataset.
c. Use the Student network $G_s$ to forward propagate the labeled data and compute the loss $L_{labeled}$.
d. Use the Teacher network $G_t$ to forward propagate the unlabeled data and compute the loss $L_{unlabeled}$.
e. Calculate the total loss $L_{total} = L_{labeled} + \omega^* L_{unlabeled}$, use backpropagation to update the network's parameters, and optimize the total loss.
f. Use the Exponential Moving Average (EMA) to update the teacher network parameters $G_t$, the exponential moving average update formula is $\theta_t = \alpha\theta_{t-1} + (1-\alpha)\theta_t$, where $\theta_t$ is the parameter value following the $t$th update, $\theta_{t-1}$ is the parameter value following the last update, and $\alpha$ is the attenuation coefficient of the exponential moving average. Typically, $\alpha$ takes a value between 0 and 1, which is used to control the amount of influence that the historical parameter values have on the current parameter update, which is chosen as $\alpha = 0.999$ in this paper.
(5) Validate the model performance using the validation set and evaluate the generalization ability of the model.
(6) Adjust hyperparameters

such as learning rate, etc. to optimize model performance. In order to guarantee that the model has good initial performance and stability, we combine the Mean Teacher algorithm on the base model for model training in this paper. We also use the pre-training parameters of the base model to investigate in detail the potential role of semi-supervised learning in rock images segmentation and to assess the effects of various training strategies on the model performance. We are dedicated to enhancing the precision and efficacy of rock images segmentation through our research technique, offering helpful references and perspectives for scholarly studies and real-world applications in this domain.

## SELF-ATTENTION MECHANISM

The Self-Attention mechanism (Vaswani *et al.* (2017)) is an attention mechanism used to compute the relationship between elements in a sequence. The algorithm generates attention weights by calculating

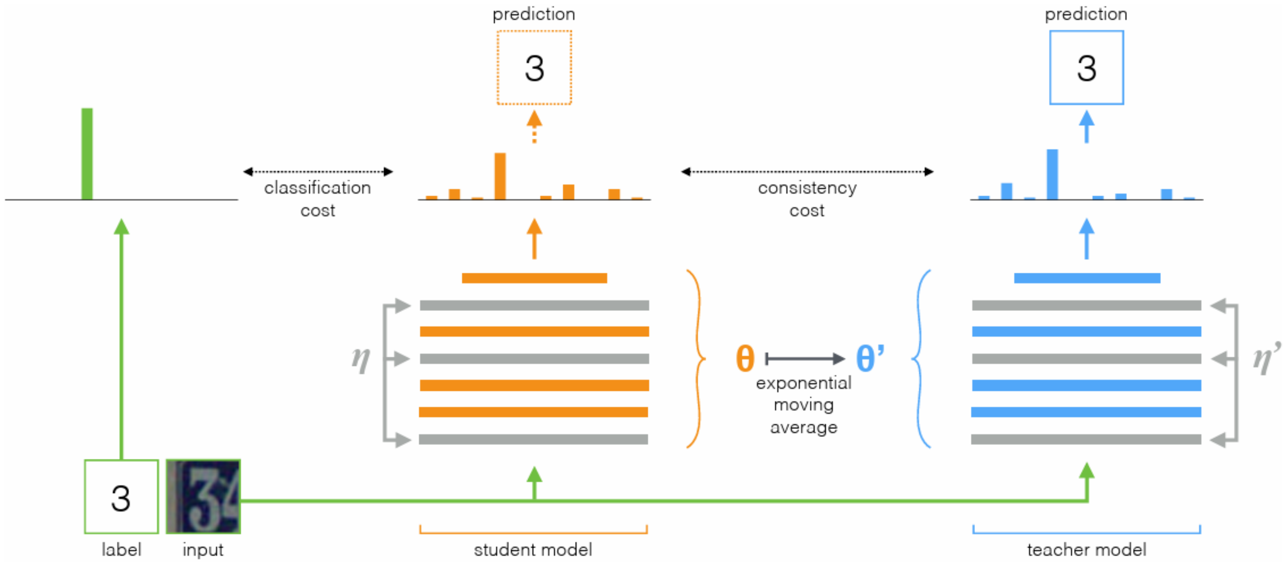


Fig. 2. *Mean Teacher training method diagram*

the similarity between the input features, thus improving the model attention to features at different spatial locations. In particular, a self-attention function $Attention(Q,K,V)$ is constructed, in which $Q$ stands for the query , $K$ for the key, and $V$ for the value.

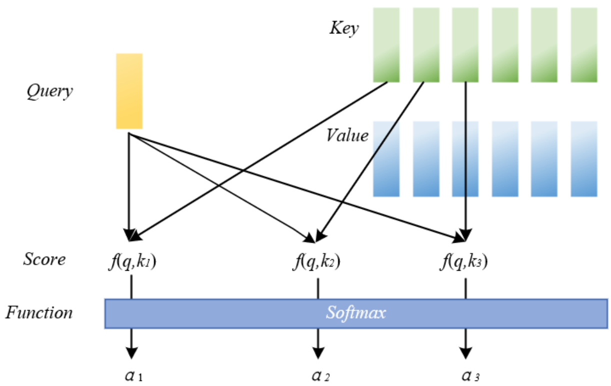


Fig. 3. *Self-attention mechanism schematic diagram*

The schematic of the self-attention mechanism is shown in Fig. 3, this mechanism obtains the attention weight $\alpha$ by calculating the similarity between the query and the key. Afterwards, the weighted feature representation is obtained by applying the attention weight to the value $V$. The following is the calculation formula

$$Attention(Q,K,V) = softmax(\frac{QK^T}{\sqrt{d_k}})V, \qquad (1)$$

where $d$ is the vector dimension.

This procedure can enhance the model depiction of spatial location attributes and efficiently capture the connections among various places.

# THE METHOD OF IMPROVED SEMI-SUPERVISED MEAN TEACHER

## SELF-ATTENTION ENHANCED MEAN TEACHER MODULE

The Mean Teacher algorithm has been widely used in semi-supervised learning tasks, and its core idea is to improve model performance by integrating labeled and unlabeled data. To further enhance the model ability to understand the global context, we propose the method of integrating self-attention in the Mean Teacher algorithm (MTSA). By introducing the self-attention module, we enable the model to adaptively focus on important regions in the input data, thus improving its performance in image classification tasks. The network structure is illustrated in Fig. 4.

Specifically, we build on the Mean Teacher algorithm by embedding the self-attention module after the third and fourth layers of the downsampling phase of the ResNet34-Unet network. This module allows the network to automatically learn and pay attention to the dependencies between different regions of the input data while extracting features, allowing the model to better capture global contextual information during feature extraction and dynamically adjust its

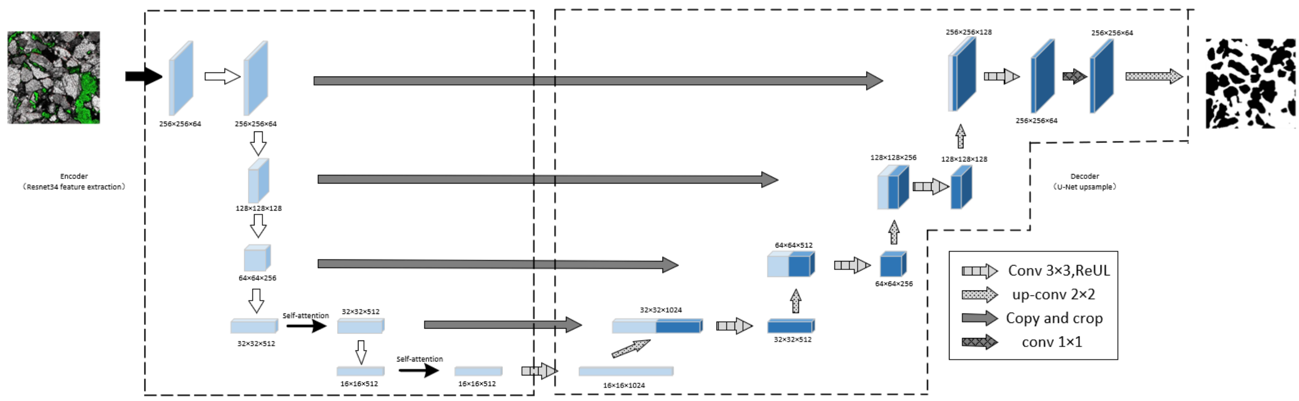


Fig. 4. *MTSA network structure*

attention during training, thus improving the accuracy of image segmentation.

## EVALUATING INDICATOR

In machine learning tasks, evaluation metrics are one of the most important criteria for assessing the performance of a model. By measuring a model performance on a particular job, evaluation metrics allow us to compare the performance of various models. Confusion matrix is a commonly used tool for evaluating the performance of a classification model, it compares the predicted results of the model with the true labels and shows how the model classifies on different categories.In the rock image segmentation task, we will use the confusion matrix Table 1 to help understand the classification results of the model and compute the evaluation metrics to evaluate the model performance.

Table 1. *Confusion matrix.*

| | Predicted negative | Predicted positive |
|---|---|---|
| Actual negative | TN | FP |
| Actual positive | FN | TP |

In this paper, we will use the mIoU (Mean Intersection over Union) value as the only evaluation metric to evaluate the performance of the model in the rock image segmentation task.mIoU value is a commonly used evaluation metric for semantic segmentation tasks, which calculates the ratio of the intersection and concatenation between the predicted segmentation results and the real segmentation results, which reflects the model segmentation accuracy for different categories. accuracy, the mIoU value is calculated as follows:

$$Miou = \frac{TP}{TP+FP+FN} \tag{2}$$

# THE RESULTS OF EXPERIMENT

## DATASET AND HYPERPARAMETERS

The rock images dataset used in this paper consists of 430 unlabeled and 50 labeled rock images. To ensure the diversity and representativeness of the experiment, these rock images feature a variety of rock sample kinds and shapes.

In order to guarantee the fairness and practicability of the studies, the models in the experiments all use the same parameters and experimental environment configurations, which include the method of terminating the training iterations, the optimizer, the learning rate, the size of the batch size, and the selection of the loss function, etc, and the specific parameters are listed in the Table 2:

Table 2. *Training parameter settings.*

| Parameter name | Parameter setting |
|---|---|
| Accuracy degradation limits | 5 |
| Categorical Loss Function | Cross Entropy Loss |
| Consistency Loss Function | Mean Square Error Loss |
| Optimizer | SGD |
| Learning rate | 0.01 |
| Batch_size | 4 |

In the semi-supervised control experiments, the overall size of batch_size is set to 4, but it should be noted that each batch actually contains 22 image samples. Specifically, each batch contains 2 labeled samples and the corresponding 20 unlabeled samples to maintain a ratio of labeled to unlabeled data of 1:10.This setting helps to efficiently utilize the unlabeled data in the semi-supervised learning task and improve the model performance.

## THE RESULTS OF RESNET34-UNET MODEL

When training the ResNet34-Unet model using a fully supervised approach, the best mIoU value we obtained on the data validation set was 0.7687. As can be observed from the Fig. 5, the variation of mIoU values during fully supervised training shows a relatively smooth upward trend in the training set, whereas the mIoU in the test set varies more drastically, which suggests that the model performance is limited when trained with only a small amount of This indicates that the performance of the model is limited when it is trained with only a small amount of labeled data, and it cannot fully capture the complex features of the data to realize its potential.


Fig. 5. *Variation in mIoU during Supervision training*

## THE RESULTS OF MEAN TEACHER

When the model is trained by the Mean Teacher method, the change of mIoU value is not stable on the test set, which may indicate that the Mean Teacher method may have encountered training difficulties or challenges in the model generalization ability at some stages, and this phenomenon may originate from the training mechanism of the Mean Teacher method, which introduces supervised signals from unlabeled data, resulting in the the model to suffer from more data interference during the learning process, which in turn leads to fluctuations in performance performance. In this paper, the weights of the unsupervised loss function are adjusted by combining the grid search method to further optimize the model performance and reduce the impact of performance fluctuations, and the optimal parameters are finally determined as $\omega = 0.6$.

As of right now, the model best mIoU value on the data validation set is 0.8237. Compared to the supervised, the Mean Teacher method shows some performance enhancement throughout the training phase, according to the findings of the observed experiments in Fig. 6. All things considered, the Mean Teacher approach has to contend with training instability in addition to performance gains, which offers helpful information for additional study and method optimization. It enhances the model performance and capacity for generalization and fosters the growth of the semi-supervised learning field by thoroughly examining the Mean Teacher method's optimization techniques and performance features in many contexts.


Fig. 6. *Variation in mIoU during Semi-supervision training*

The advantage of Mean Teacher may lie in its ability to utilize unlabeled data for training, thus improving the model ability to learn data features and hence performance performance. This semi-supervised learning approach helps to address the problem of costly data labeling while improving model performance and generalization.

## EXPERIMENTAL RESULTS OF MEAN TEACHER INCORPORATING SELF-ATTENTION MECHANISMS

Although Mean Teacher has made some progress in improving the model performance, there is still room for improvement in dealing with data labeling noise and model instability, and the ability to capture complex features still needs to be improved, in order to further improve the model performance, we introduce a self-attention mechanism based on Mean Teacher algorithm, which enhances the model's ability to capture global dependencies. In order to further improve the performance of the model, we introduce the self-attention mechanism based on Mean Teacher algorithm to enhance the model ability to capture global dependencies. After experimental validation, we observe that after adding the self-attention mechanism, the optimal mIoU value of the model is further improved to 0.8792, which is higher than that of the Mean Teacher algorithm only. This indicates that the introduction of the self-attention mechanism effectively enhances the model attention to spatial location features, which further improves the model performance in the image segmentation task.


Fig. 7. *Variation in mIoU during Mean Teacher Incorporating Self-Attention Mechanisms training*

According to Fig. 7 analyzing the experimental results comprehensively, we can find that although Mean Teacher has achieved some success in semi-supervised learning, it still has deficiencies in dealing with data labeling noise and model instability. The introduction of the self-attention mechanism can effectively improve the performance of the ResNet34-Unet model in the image segmentation task, so that it can better capture the complex features of the data and thus achieve better segmentation results. This experimental result provides a useful reference and inspiration for further research and application of the self-attention mechanism in the field of image segmentation.

## COMPARATIVE ANALYSIS

In the comparative analysis section, we compare the performance of three different models in the image segmentation task, namely ResNet34, Mean Teacher and Meanteacher+attention models. The Table 3 shows their mIoU values on the training and test sets:

Table 3. *Performance of different models.*

| Model | Miou_train | Miou_test |
|---|---|---|
| ResNet34-UNet | 0.7906 | 0.7687 |
| Mean Teacher | 0.8609 | 0.8237 |
| Meanteacher+attention | 0.9408 | 0.8792 |

From the Table 3, it can be seen that the mIoU value of Mean Teacher model is significantly improved on both the training and test sets compared to the model ResNet34-UNet.On the training set, the mIoU value of the Mean Teacher model reaches 0.8609, which is 0.0703 higher than that of ResNet34-UNet, and on the test set, the mIoU value reaches 0.8237, which is 0.055 higher than that of ResNet34-UNet, which suggests that the semi-supervised learning method Mean Teacher achieves better results in the image segmentation task has achieved better results and can effectively improve the performance of the model.

Further, the performance of the Mean Teacher model is further improved on the training and test sets by introducing the self-attention mechanism based on the Mean Teacher model. In the training set, the mIoU value reaches 0.9408, which is 0.0799 higher than that of the Mean Teacher model, and in the test set, the mIoU value reaches 0.8792, which is 0.0555 higher than that of the Mean Teacher model, indicating that the introduction of the self-attention mechanism effectively enhances the model focus on spatial location features, which further enhances the model performance in the image segmentation task. performance in image segmentation tasks. A comparison of the results of the rock image samples segmented under different methods is shown in Fig. 8.

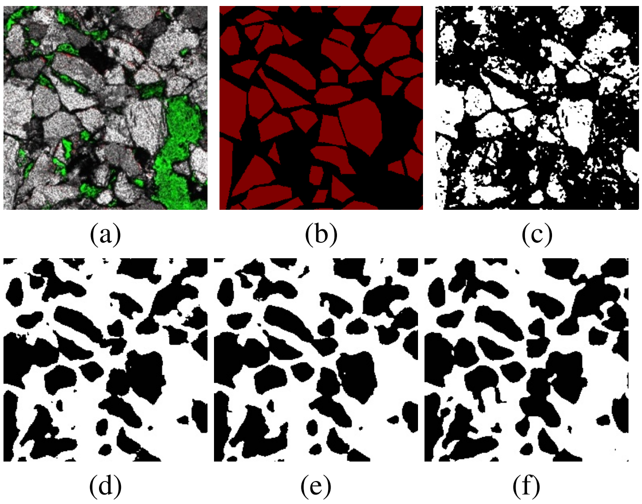


Fig. 8. *Test samples and model outputs for the rock dataset, (a) test sample images, (b) test sample labels, (c) graph of threshold segmentation results, (d) graph of segmentation results for the supervised Rsnet34-Unet model, (e) graph of segmentation results for Mean teacher, and (f) graph of segmentation results for Meanteacher+Self attention.*

In summary, the Mean Teacher Incorporating Self-Attention Mechanisms model achieves the best performance in this study, with a mIoU value of 0.8792 on the test set, which is 0.1105 higher than the benchmark model, ResNet34. This further validates the effectiveness of the self-attention This further validates the effectiveness of the self-attention mechanism in the image segmentation task and provides strong support for the model performance improvement.

## CONCLUSION

The advantage of the semi-supervised Mean Teacher approach is its ability to utilize unlabeled data for training, which improves the model ability to learn the features of the data, and thus improves the performance performance.Although this method helps to solve the problem of high cost of data labeling, but it is still unavoidable that the Mean Teacher approach also faces the challenge of unstable training on the validation set. The Mean Teacher model with the introduction of the self-attention mechanism further improves the performance of the image segmentation task. The self-attention mechanism can effectively capture the complex features of the data and enhance the model attention to the spatial location features, which makes the model more accurate and stable in handling the image segmentation task. However, the method still faces the challenge of unstable training, which leads to fluctuations in model performance or difficulty in convergence.

To further improve the performance and stability of the model, future research can focus on the following aspects: first, exploring more stable training strategies, such as introducing more effective loss function design, optimizer selection, or learning rate adjustment methods, in order to improve the stability of the model performance on the validation set. Then, to address the problem that the pseudo-labels generated in the Mean Teacher algorithm may have certain noise and inaccuracy, we explore the improvement of the pseudo-label generation strategy and the introduction of a more accurate and reliable pseudo-label generation method. By improving the training strategy of the algorithm and introducing different pseudo-label generation methods to improve its accuracy, it is expected to achieve more significant performance enhancement and stability improvement in image segmentation tasks in the future.

## FUNDING

The research presented in this paper was funded by the National Natural Science Foundation of China (421721163).

## REFERENCES

Badrinarayanan V, Kendall A, Cipolla R (2017).Segnet: A deep convolutional encoder-decoder architecture for image segmentation. Ieee T Pattern Anal 39(12):2481–95.

Cao D,Ji S,Cui R,Liu Q (2022). Multi-task learning for digital rock segmentation and characteristic parameters computation. J Petrol Sci Eng 208:109202.

Chen J, Zhao Z, Zhang J (2024). Predicting peak shear strength of rock fractures using tree-based models and convolutional neural network. COMPUT GEOTECH 166:105965.

Cai F, Wen J, He F, Xu W (2024). SC-Unext: A Lightweight Image Segmentation Model with Cellular Mechanism for Breast Ultrasound Tumor Diagnosis. J Imaging Inform Med 37:1505–15.

Dong K and Jiang D (2013).Ore image segmentation algorithm based on improved watershed transform. Computer Engineering and Design 34(3):899–903.

Fan L, Yuan J, Niu X, Zha K, Ma W (2023). RockSeg: A Novel Semantic Segmentation Network Based on a Hybrid Framework Combining a Convolutional Neural Network and Transformer for Deep Space Rock Images. Remote Sens-Basel 15(16):3935.

Han K, Sheng V, Song Y, Liu Y, Qiu C, Ma S, Liu Z (2024). Deep semi-supervised learning for medical image segmentation: A review. Expert Syst Appl 245:123052.

Huang B, Huang T, Xu J, Min J, Hu C, Zhang Z (2024). RCNU-Net: Reparameterized convolutional network with convolutional block attention module for improved polyp image segmentation. Biomed Signal Proces 93:106138.

Huang S, Luo J, Ou Y, Pang Y, Nie X, Zhang G (2024). Sd-net: a semi-supervised double-cooperative network for liver segmentation from computed tomography (CT) images. J Cancer Res Clin 150(2):79.

Lou L, Xin Y, Qian J, Dong Y (2024). A detail-oriented super-2D network for pulmonary artery segmentation. Biomed Signal Proces 93:106183.

Liu Y, Wang X, Zhang Z, Deng F (2023). Deep learning in image segmentation for mineral production: A review. Comput Geosci-Uk 180:105455.

Liang H and Zou J (2020). Rock image segmentation of improved semi-supervised SVM¨CFCM algorithm based on chaos. Circ Syst Signal Pr 39(2):571–85.

Liu Y, Shen Y, Song H, Yan F, Su Y (2024). Solar radio spectrogram segmentation algorithm based on improved Fuzzy C-Means clustering and adaptive cross filtering. Phys Scripta 99(4):5005.

Murray G, Bourlai T, Spolaor M (2021). Mask R-CNN: Detection performance on SPEED spacecraft with image degradation. In 2021 IEEE International Conference on Big Data (Big Data) ,Orlando, FL, USA.IEEE.

Manzoor S, Qasim T, Bhatti N, Zia M (2023). Segmentation of digital rock images using texture analysis and deep network. Arab J Geosci 16(7):436.

Mazher M, Razzak I, Qayyum A, Tanveer M, Beier S, Khan T, et al. (2024). Self-supervised spatialtemporal transformer fusion based federated framework for 4D cardiovascular image segmentation. Inform Fusion 106:102256.

Niu Y, Mostaghimi P, Shabaninejad M, Swietojanski P, Armstrong R (2020). Digital rock segmentation for petrophysical analysis with reduced user bias using convolutional neural networks. Water Resour Res 56(2):e2019WR026597.

Roslin A, Lebedev M, Mitchell TR, Onederra IA, Leonardi CR (2023). Processing of micro-C Timages of granodiorite rock samples using convolutional neural networks (CNN). Part III: Enhancement of Scanco micro-CT images of granodiorite rocks using a 3D convolutional neural network superresolution algorithm. Miner Eng 195:108028.

Ronneberger O, Fischer P, Brox T (2015). U-net: Convolutional networks for biomedical image segmentation. Medical image computing and computer-assisted intervention-MICCAI 2015,Cham.Springer International Publishing.

Sinha A, Aljrees T, Pandey SK, Kumar A, Banerjee P, Kumar B, et al. (2023). Semi-Supervised Clustering-Based DANA Algorithm for Data Gathering and Disease Detection in Healthcare Wireless Sensor Networks (WSN). Sensors-Basel 24(1):18.

Tarvainen A and Valpola H (2017). Weight-averaged consistency targets improve semi-supervised deep learning results.ArXiv abs: 1703.01780.

Vaswani A, Shazeer NM, Parmar N, Uszkoreit J, Jones L, Gomez AN,et al. (2017). Attention is All you Need.ArXiv abs: 1706.03762.

Weng Y, Zhang Y, Wang W, Dening T (2024). Semi-supervised information fusion for medical image analysis: Recent progress and future perspectives. Information Fusion 106:102263.

Liu Y and Lv J (2023). Semi-supervised Rock Image Segmentation and Recognition Based on Superpix. Adv Model Simul Eng 55(2):171-83.

Zhang Z, Tian C, Gao X, Wang C, Feng X, Bai H, et al. (2022). Dynamic prototypical feature representation learning framework for semi-supervised skin lesion segmentation. Neurocomputing 507:369–82.

Zhao F, Tang Z, Xiao Z, Liu H, Fan J, Li L (2024). Ensemble CART surrogate-assisted automatic multi-objective rough fuzzy clustering algorithm for unsupervised image segmentation. Eng Appl Artif Intel 133: 108104.

Zhang X, Wang Z, Liu D, Sun Q, Wang J (2023). Rock thin section image classification based on depth residuals shrinkage network and attention mechanism. Earth Sci Inform 16(2):1449–57.