





CONSS: Contrastive Learning Method for Semisupervised Seismic Facies Classification

Kewen Li , Wenlong Liu , Yimin Dou , *Student Member, IEEE*, Zhifeng Xu , Hongjie Duan, and Ruilin Jing

Abstract—Recently, convolutional neural networks (CNNs) have been widely applied in the seismic facies classification. However, even state-of-the-art CNN architectures often encounter classification confusion distinguishing seismic facies at their boundaries. In addition, the annotation is a highly time-consuming task, especially when dealing with 3-D seismic data volumes. While traditional semisupervised methods reduce dependence on annotation, they are susceptible to interference from unreliable pseudolabels. To address these challenges, we propose a semisupervised seismic facies classification method called CONSS, which effectively mitigates classification confusion through contrastive learning. Our proposed method requires only 1% of labeled data, significantly reducing the demand for annotation. To minimize the influence of unreliable pseudolabels, we also introduce a confidence strategy to select positive and negative sample pairs from reliable regions for contrastive learning. Experimental results on the publicly available seismic datasets, the Netherlands F3 and SEAM AI challenge datasets, demonstrate that the proposed method outperforms classic semisupervised methods, including self-training and consistency regularization, achieving exceptional classification performance.

Index Terms—Contrastive learning, deep learning, seismic facies classification, seismic interpretation, semisupervised learning.

I. INTRODUCTION

SEISMIC facies classification is the interpretation of facies types based on seismic reflection information, which is a crucial step in petroleum exploration and reservoir characterization. However, seismic facies classification is a labor-intensive and time-consuming task. Therefore, both the industry and academia are pursuing automated or semiautomated methods for seismic facies classification.

In the past, some machine learning methods [1], [2], [3] were used for seismic facies analysis, but they required manual selection of seismic attributes as feature inputs. With the development of deep learning, convolutional neural networks (CNNs) have

been widely applied to seismic data processing [4], [5], [6] and seismic facies classification [7], [8], [9], [10]. However, there are still the following challenges.

1) *Time-Consuming Annotation*: With the rapid increase in seismic data volume, seismic interpreters face a heavy burden, core and well log data available to aid in the interpretation of seismic facies are scarce and costly.

To address the challenge, we employ a specific division strategy of dataset and propose a semisupervised method using few labeled data. The 3-D seismic data volume was uniformly divided into distinct blocks. Within each block, the first slice was assigned to the training set, whereas the remaining slices were allocated to the test set. Only 1% labels of the entire data volume was utilized for training, significantly reducing the demand for annotation.

To encompass both high-frequency regions reflecting seismic event variations and low-frequency regions reflecting overall trends, the slices spanned the entire depth direction, ensuring more comprehensive contextual and spatial information. Furthermore, we have made all the code and data openly accessible to facilitate related research and fair comparison.

2) *Difference From Natural Images*: Seismic images exhibit notable distinctions from natural images, as they lack evident semantic objects and instead possess layer-specific information associated with depth, and the simultaneous occurrence of multiple seismic facies within the same seismic profile often leads to pronounced classification confusion at the boundaries.

To cope with the challenge, we employed contrastive learning to enhance the feature representation of different seismic facies. We identified that a significant cause of classification confusion lies in the confusion of features, where features within the same class are excessively dispersed, whereas features across different classes are overly similar. The classifier of the CNN relies on the features from the last layer of the encoder–decoder for classification, thus the distribution of features directly influences the classification results.

To enhance the discriminative nature of features across different classes and promote more concentrated feature distributions within the same class, we constructed central features for each seismic facies class to capture their common characteristics. Each class's features were paired with its corresponding central feature, forming positive pairs, whereas features from different classes were paired as negative pairs. By leveraging contrastive learning, we aimed to reduce the distance between positive pairs while simultaneously increasing the distance between negative pairs. In addition, our proposed method offers the advantage of

Manuscript received 5 July 2023; revised 9 August 2023; accepted 19 August 2023. Date of publication 25 August 2023; date of current version 7 September 2023. This work was supported in part by the National Natural Science Foundation of China Major Program under Grant 51991365 and in part by the Natural Science Foundation of Shandong Province, China under Grant ZR2021MF082. (Corresponding author: Kewen Li.)

Kewen Li, Wenlong Liu, Yimin Dou, and Zhifeng Xu are with the College of Computer Science and Technology, China University of Petroleum (East China), Qingdao 266580, China (e-mail: likw@upc.edu.cn; 1707020116@s.upc.edu.cn; emindou3015@gmail.com; b21070008@s.upc.edu.cn).

Hongjie Duan and Ruilin Jing are with the Shengli Oilfield Company, SINOPEC, Dongying 257068, China (e-mail: duanhongjie.slyt@sinopec.com; jingruilin.slyt@sinopec.com).

All codes and data are available online at https://github.com/upcliuwenlong/CONSS_SEISMIC_FACIES.

Digital Object Identifier 10.1109/JSTARS.2023.3308754

end-to-end training, allowing seamless integration into existing encoder–decoder architectures.

3) *Unreliable Pseudolabels*: Although traditional semisupervised methods reduce the reliance on labeled data, they are susceptible to the influence of unreliable pseudolabels, which actually deteriorate the model’s performance.

To tackle the challenge, we devise a confidence strategy to select reliable feature samples and mitigate the interference caused by unreliable pseudolabels. The unreliable pseudolabels refer to regions that have been erroneously classified. We have observed that in these regions, the entropy of the classification probabilities tends to be relatively high. The magnitude of entropy is closely associated with the probability distribution, whereby a higher probability for a specific class corresponds to lower entropy, indicating a more reliable classification. Conversely, a more uniform probability distribution leads to higher entropy, indicating a less reliable classification outcome. Therefore, relying on the predicted probabilities, we exclude features from unreliable regions and divide the reliable regions of each class into two subregions: the weak confidence region and the strong confidence region. The strong confidence regions represent the most reliable regions and is utilized for computing the central features, and the features located within the weak confidence regions become our primary optimization target.

In summary, the contributions of this article are threefold.

- 1) We propose a contrastive learning method for semisupervised seismic facies classification. By defining positive and negative sample pairs at the feature level, we optimize the feature representation of seismic facies, significantly reducing the reliance on labeled data and mitigating classification confusion.
- 2) We introduce a confidence strategy to mitigate the influence of unreliable pseudolabels, which provides a promising solution for efficiently generating feature sample pairs in the presence of unreliable pseudolabels.
- 3) We have made our data and codes publicly available to benefit other research of the geophysical community.

II. RELATED WORK

A. Contrastive Learning

Contrastive learning [11], [12], [13], [14] is a self-supervised learning method aimed at learning feature representation by comparing the similarities and differences between data sample shown in Fig. 1. The fundamental idea of this method is to encode samples of the same class as similar representations and samples of different classes as dissimilar representations. The primary objective of contrastive learning is to bring similar samples closer together in the feature space while pushing dissimilar samples further apart. At the core of contrastive learning lies the construction of a contrastive loss function, a commonly used contrastive loss function is based on positive and negative sample pairs [15], [16], where positive sample pairs consist of samples from the same class, whereas negative sample pairs consist of samples from different classes.

Kokilepersaud et al. [17] proposed a pretraining method for seismic facies classification based on slice-level contrastive

learning, they divided seismic data into different blocks, considering slices within the same block as similar positive samples and slices from different blocks as dissimilar negative samples. This method involves two stages: unsupervised pretraining and supervised fine-tuning. In our proposed method, we employ pixel-level contrastive learning, which allows for more efficient sample sampling. Furthermore, our proposed method is an end-to-end semisupervised training method that does not involve pretraining.

B. Semi-Supervised Learning

Semisupervised learning [18] is a learning method that leverages both labeled and unlabeled data for model training. It aims to improve model performance in situations where labeled data are limited by utilizing information from unlabeled data. The most common paradigms in semisupervised learning are self-training [19], [20] and consistency regularization [21], [22]. Self-training is an iterative method where the model is initially trained on a limited amount of labeled data. The model is then applied to unlabeled data, and the model’s predictions are used as pseudolabels. These pseudolabels serve as approximate labels for the unlabeled data, which are then combined with the labeled data for the next round of model training. Consistency regularization represents another semisupervised learning paradigm, which improves model robustness and generalization by encouraging the model to produce consistent outputs for perturbed input data.

Saleem et al. [23] implemented a semisupervised method based on self-training for seismic facies classification. They trained the model on labeled data and utilized the model’s predictions of surrounding seismic facies as pseudolabels for further training. However, such pseudolabeling methods are prone to interference from unreliable pseudolabels.

C. Seismic Facies Classification

In the past, seismic facies classification was treated as a conventional image classification task. Chevotarese et al. [24] proposed a patch-based seismic facies classification model. Dramsch and L uthje [25] utilized pretraining and fine-tuning techniques to improve model performance. However, these patch-based classification methods suffer from the absence of contextual and spatial information, failing to capture high-frequency seismic variations. Considering seismic facies classification as a semantic segmentation task is more appropriate. Zhao [26] introduced an encoder–decoder CNN segmentation model for seismic facies classification and compared it with patch-based methods. Alaudah et al. [27] released the fully annotated 3-D geological model of the Netherlands F3 dataset and proposed a segmentation model based on transposed convolutions. Tolstaya and Egorov [9] incorporated positional encoding as an additional input channel to enhance spatial information and used pseudolabels for further training. Civitarese et al. [10] replaced the traditional dilated convolutions in the decoder block with transposed residual units and proposed DanetFCN. Some work in remote sensing image classification is also worthy of attention. Hong et al. [28] proposed a general multimodal

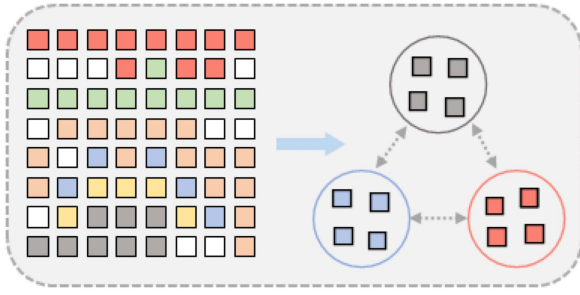


Fig. 1. Contrastive learning aims to increase the interclass distance while reducing the intraclass distance, thereby improving feature representation.

learning framework for remote sensing image classification, which achieves fine classification by data fusion of different modalities and overcomes the limitation of information diversity. In addition to CNNs, Hong et al. [29] proposed a backbone network based on the transformer architecture to mine and represent the sequence attributes of spectral signatures and built relationships between samples based on GCNs [30] to extract more diverse and discriminative feature representations for the image classification task.

Relevant to our work in [31] and [32], they computed the trace of the intraclass variance matrix and interclass variance matrix using few labeled data and minimized their ratio. They employed deep autoencoder and a greedy layer-wise pretraining strategy. In contrast, our proposed method utilizes contrastive learning with an instance discrimination task [15], it does not require pretraining and calculating variance matrix with statistical properties. More importantly, our proposed method is not limited to few labeled data and is applicable to unlabeled data as well. Furthermore, we use different dataset and division strategy.

III. METHOD

A. Overview

In our proposed CONSS as shown in Fig. 2, during each forward propagation, both labeled and unlabeled data are input simultaneously with the same batch size. Subsequently, supervised learning and contrastive learning are performed separately after a nonlinear transformation. Instead of using the direct features from the encoder–decoder output, we employ features obtained after the nonlinear transformation. This choice is based on the findings of SimCLR [14], which demonstrated that this straightforward yet effective nonlinear transformation contributes to improved representation learning, and the nonlinearity primarily stems from the ReLU activation function.

The supervised learning branch produces prediction results and undergoes learning with the aid of supervision signals. The contrastive learning branch, on the other hand, leverages the predicted probability of unlabeled data from the supervised learning branch to determine the classes of different seismic facies. To avoid interference from unreliable pseudolabels and optimize the feature representation, the contrastive learning branch employs a confidence strategy to select positive and negative samples for contrastive learning. Through the collaborative

efforts of these two branches, the CONSS effectively integrates supervised and contrastive learning, leveraging their complementary strengths. This integration leads to enhanced feature representation and improved performance in seismic facies classification. The joint optimization of the two branches involves minimizing both supervised loss and contrastive loss (1):

$$\mathcal{L} = \mathcal{L}_{\text{sup}} + \mathcal{L}_{\text{con}}. \quad (1)$$

We introduce supervised loss, confidence strategy for sample selection, definition of positive and negative sample pairs, and contrastive loss in the following sections.

B. Supervised Loss

The supervised loss is computed within the supervised learning branch, which operates on an output segmentation feature map M_{seg} of dimensions $H \times W \times C$. Each position within the feature map represents a C -dimensional vector. Following this, the vector undergoes a softmax (4) layer transformation to yield a probability distribution that sums to 1. The supervised loss is a pixel-by-pixel cross-entropy loss function with label smoothing [33] (3):

$$\mathcal{L}_{\text{sup}} = -\frac{1}{H \times W} \sum_{j=1}^{H \times W} \sum_{i=1}^C p_i^j \log(q_i^j) \quad (2)$$

$$p_i^j = \begin{cases} 1 - \epsilon, & \text{if } (i = y) \\ \frac{\epsilon}{C-1}, & \text{if } (i \neq y) \end{cases} \quad (3)$$

$$q_i^j = \frac{e_i^j}{\sum_{i=1}^C e_i^j} \quad (4)$$

where H and W represent the height and width of the model output, C is equal to the number of seismic facies classes, p_i^j is the probability that the pixel at position j belongs to class i , and ϵ is the smoothing factor responsible for dividing the small probability into other classes. Smoothing [34] hard labels to soft labels will prevent the model from overfitting due to overconfidence. The y is the true facies class, and q_i^j is the predicted probability of the model. By minimizing the cross-entropy loss function, the predicted probability distribution will be close to the true probability distribution.

C. Dual-Level Confidence Strategy

The utilization of pseudolabels strategy in semisupervised methods renders them susceptible to the influence of unreliable pseudolabels. Unreliable pseudolabels refer to regions that are prone to misclassification, the information entropy (5) of predicted probability in these regions tends to be larger. A higher information entropy value indicates a more uniform distribution of the predicted probability, resulting in increased uncertainty in the model's predictions. Conversely, a higher confidence level of the model for a specific class corresponds to a smaller entropy value, implying more reliable results shown in Fig. 3. To address this issue, we employ a simple yet effective confidence strategy, which allows us to exclude features from regions with high

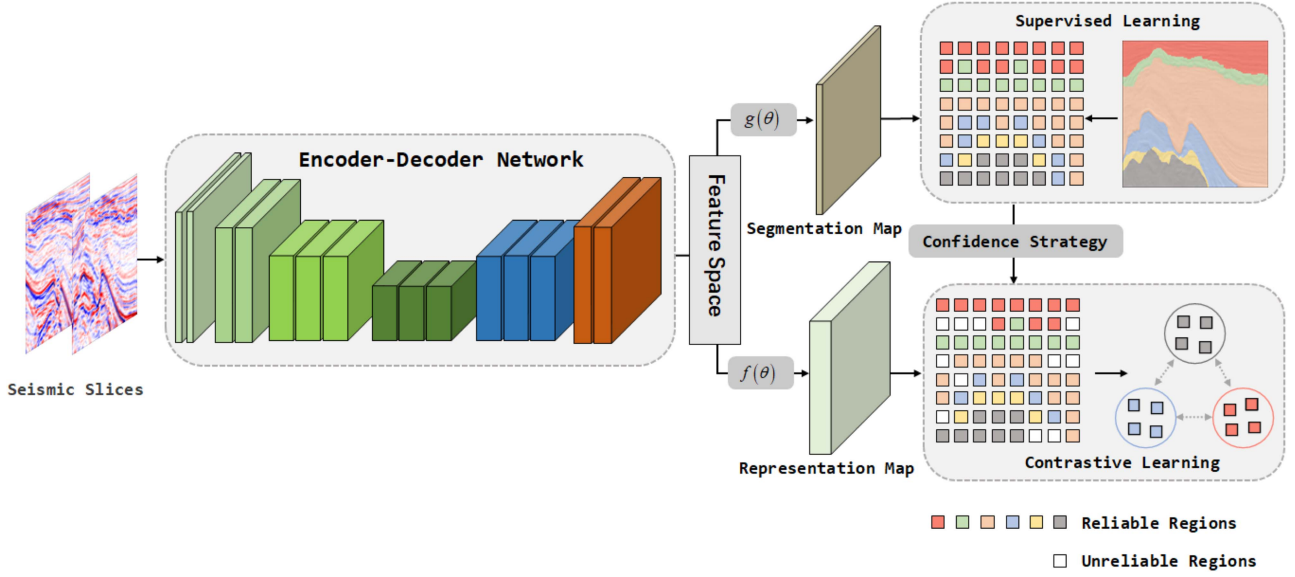


Fig. 2. CONSS pipeline consists of an encoder–decoder network that extracts features, a supervised learning branch for supervised learning, and a contrastive learning branch for representation learning. The $f(\theta)$ and $g(\theta)$ represent simple nonlinear transformations consisting of convolution, ReLU, and upsampling.

entropy values to avoid reliance on unreliable information

$$E = - \sum_{i=1}^C q_i^j \log(q_i^j). \quad (5)$$

For unlabeled seismic data, the true seismic facies class is unknown and only be inferred based on the model’s prediction results. To mitigate the impact of unreliable pseudolabels with large entropy values, we exclude regions where the maximum predicted probability is below a certain threshold t_w . In addition, we divide the reliable regions of each class i into the following two distinct subregions shown in Fig. 4.

- 1) R_i^w is the weak confidence region of class i , the model predicted probability range is $t_w < q_i^j < t_s$, and the predicted probability for class i is largest.
- 2) R_i^s is the strong confidence region of class i , the model predicted probability range is $q_i^j > t_s$, and the predicted probability for class i is largest.

It is important to highlight that the weak confidence region of unlabeled data exhibits a relatively lower predicted probability compared with the strong confidence region. However, it still falls within the realm of higher classification probability, indicating a lower entropy value for these regions. That is, weak confidence region and strong confidence region are both high-confidence regions.

For labeled data, there is no necessity to rely on unreliable pseudolabels since the true classes are known. Consequently, there are following notable distinctions in the division of confidence regions when compared with unlabeled data.

- 1) R_i^w is the weak confidence region of class i , the model predicted probability range is $q_i^j < t_s$, and the true class is i .
- 2) R_i^s is the strong confidence region of class i , the model predicted probability range is $q_i^j > t_s$, and the true class is i .

D. Positive and Negative Sample Pairs

The contrastive learning branch operates on a representation map M_{rep} with dimensions $H \times W \times D$, where each position corresponds to a D -dimensional feature vector, and D is set to 128. Based on the confidence strategy, these feature vectors are categorized into either strong confidence regions R_i^s or weak confidence regions R_i^w of different classes.

The feature vectors within the strong confidence region of class i are considered strong confidence vectors F_i^s . To capture the common characteristics of class i features, a central vector is obtained by averaging the strong confidence vectors. This averaging operation reduces the impact of unreliable pseudolabels and enhances the robustness

$$F_i^+ = \frac{\sum F_i^s}{|R_i^s|}, F_i^s \in R_i^s. \quad (6)$$

Each feature vector within the weak confidence region of class i is a weak confidence vector F_i^w . These weak confidence vectors form positive sample pairs with the central vector F_i^+ of class i

$$(F_i^w, F_i^+), F_i^w \in R_i^w \quad (7)$$

where F_i^w is also called query vector in contrastive learning.

The query vector F_i^w of class i and the strong confidence vectors F_k^s of different classes form negative sample pairs

$$(F_i^w, F_k^s), k \neq i. \quad (8)$$

The objective of contrastive learning is to bring positive sample pairs closer together while increasing the distance between negative sample pairs. The distance between two samples is evaluated using cosine similarity

$$\text{sim}(F_1, F_2) = \frac{F_1 \cdot F_2}{\|F_1\| \|F_2\|}. \quad (9)$$

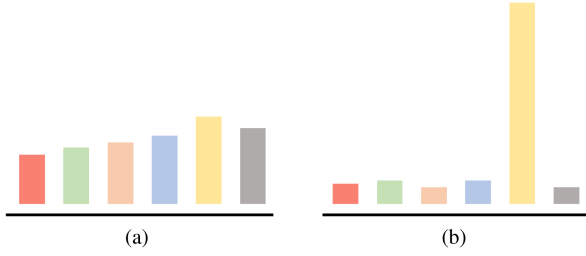


Fig. 3. Higher information entropy means the uncertainty of the prediction results of the model, and a high entropy value means that the probability distribution of the classification is relatively flat, and the prediction results will be unreliable. (a) Unreliable predicted probability. (b) Reliable predicted probability.

E. Contrastive Loss

The contrastive loss function is based on the positive and negative sample pairs defined above. In our proposed method, we employed InfoNCE [16] loss, a simplified version of NCE [35], which effectively transforms the binary classification in contrastive learning into a multiclassification problem

$$\mathcal{L}_{\text{con}} = -\frac{1}{C} \sum_{i=1}^C \log \frac{e^{\text{sim}(F_i^w, F_i^+) / \tau}}{e^{\text{sim}(F_i^w, F_i^+) / \tau} + \sum_{n=1}^N e^{\text{sim}(F_i^w, F_k^s) / \tau}} \quad (10)$$

where the C is the number of seismic facies classes, τ is the temperature coefficient used to control the discrimination of negative samples, and N is the number of negative samples.

This formulation is interpreted as an $N + 1$ classification problem, aiming to differentiate between positive and negative sample pairs within $N + 1$ pairs. Specifically, the objective is to ensure that one pair of positive samples is similar, whereas the N pairs of negative samples are dissimilar. In the context of contrastive learning, this task is referred to as instance discrimination [15]. By minimizing the contrastive loss function, the distance between positive sample pairs is minimized, whereas the distance between negative sample pairs is maximized. Consequently, this reduces the intraclass feature distance and increases the interclass feature distance of seismic facies, enabling the model to generate more distinct decision boundaries.

Furthermore, while certain contrastive learning methods [11], [12], [15], [36] rely on momentum encoders or auxiliary data structures to generate or store consistent negative samples, the nature of seismic data differs from complex natural image scenes. Seismic data exhibit continuity and similarity in its distribution, thereby facilitating the generation of more consistent samples. Consequently, our proposed method adopts an end-to-end training method that does not necessitate momentum encoders or additional data structures.

Finally, the whole process of CONSS is shown in Algorithm 1.

IV. DATASET OVERVIEW

A. Netherlands F3 Dataset

An F3 block is located in the north of the Netherlands. Alaudah et al. [27] artificially interpreted seven groups of lithostratigraphic units with reference to well logging data, and merged

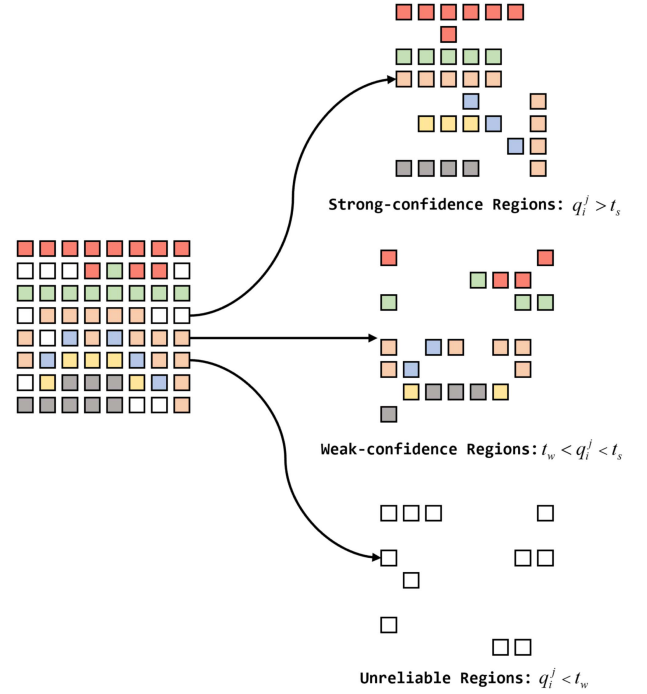


Fig. 4. Dual-level confidence strategy divides reliable regions into strong confidence regions and weak confidence regions for defining positive and negative sample pairs.

Algorithm 1: CONSS.

Input: labeled seismic data \mathcal{D}_l and unlabeled seismic data \mathcal{D}_u

Output: \mathcal{L}

$M_{\text{seg}}^l, M_{\text{rep}}^l \leftarrow \text{Model}(\mathcal{D}_l)$

$M_{\text{seg}}^u, M_{\text{rep}}^u \leftarrow \text{Model}(\mathcal{D}_u)$

Get supervised loss \mathcal{L}_{sup} by (2)

for i in classes **do**

 Get high-confidence region $\mathcal{R}_i^w, \mathcal{R}_i^s$ from $M_{\text{seg}}^l \cup M_{\text{seg}}^u$
 by confidence strategy

end for

for i in classes **do**

 Get a central vector F_i^+ from \mathcal{R}_i^s by (6)

 Get a query vector F_i^w from \mathcal{R}_i^w

 Get N vectors of other classes F_k^s from $\mathcal{R}_k^s, k \neq i$

 Form a positive sample pair by (7)

 Form N negative sample pairs by (8)

 Get contrastive loss \mathcal{L}_{con} by (10)

end for

$\mathcal{L} = \mathcal{L}_{\text{sup}} + \mathcal{L}_{\text{con}}$

Return: \mathcal{L}

Rijnland Group and Chalk Group into one, so there are following six types of seismic facies in the F3 dataset shown in Fig. 5.

- 1) Upper North Sea Group: claystones and sandstones from Miocene to Quaternary.
- 2) Lower and Middle North Sea Groups: sands, sandstones, and claystones from Paleocene to Miocene.
- 3) Rijnland Group: clay formations with sandstones of Upper Cretaceous.

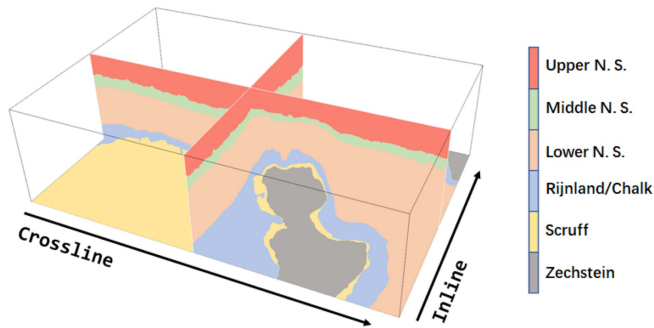


Fig. 5. Size of the Netherlands F3 dataset is $255 \times 901 \times 601$ (depth \times crossline \times inline).

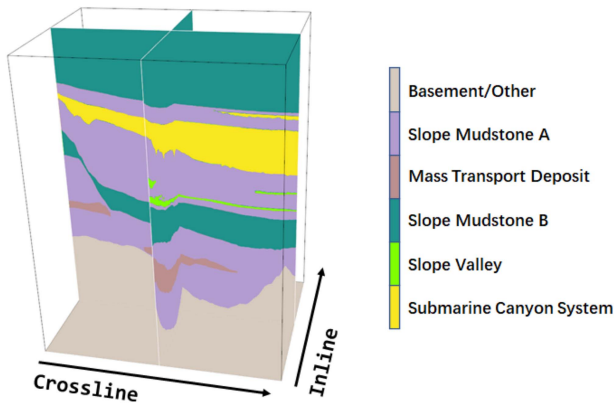


Fig. 6. Size of the SEAM AI challenge dataset is $1006 \times 782 \times 590$ (depth \times crossline \times inline).

- 4) Chalk Group: carbonates of Upper Cretaceous and Paleocene.
- 5) Scruff Group: claystones of Upper Jurassic and Lower Cretaceous.
- 6) Zechstein Group: evaporites and carbonates of Zechstein.

B. SEAM AI Challenge Dataset

The SEAM AI challenge dataset is a 3-D seismic image from a public-domain seismic survey called Parihaka, available from the New Zealand government, has been interpreted by expert geologists. The dataset consists of the following six types of seismic facies shown in Fig. 6.

- 1) *Basement/Other*: Basement—low S/N; few internal reflections; may contain volcanics in places.
- 2) *Slope Mudstone A*: Slope to basin floor Mudstones; high amplitude upper and lower boundaries; low amplitude continuous/semicontinuous internal reflectors.
- 3) *Mass Transport Deposit*: Mix of chaotic facies and low amplitude parallel reflections.
- 4) *Slope Mudstone B*: Slope to basin floor mudstones and sandstones; high amplitude parallel reflectors; low continuity scour surfaces.
- 5) *Slope Valley*: High amplitude incised channels/valleys; relatively low relief.
- 6) *Submarine Canyon System*: Erosional base is U shaped with high local relief. Internal fill is low amplitude mix

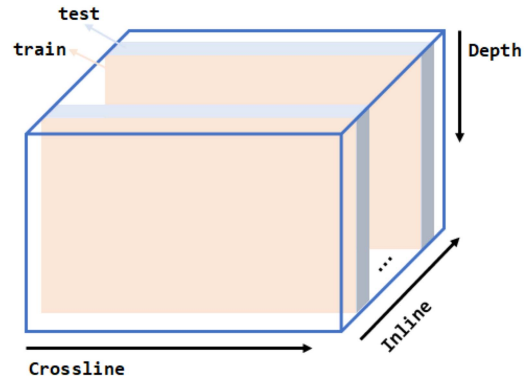


Fig. 7. 3-D seismic volume is uniformly divided into different blocks. In each block, the first slices go to the training set and the adjacent rest to the test set.

of parallel inclined surfaces and chaotic disrupted reflectors. Mostly deformed slope mudstone filled with isolated sinuous sand-filled channels near the basal surface.

C. Label Sampling

We uniformly divide 3-D seismic volume into different blocks shown in Fig. 7. In each block, the first slices go to the training set and the adjacent rest to the test set. For the Netherlands F3 dataset, we sampled a total of seven labels with size of 255×901 . The remaining labels are exclusively allocated to the test set and are not utilized during model training. The seismic data corresponding to these remaining labels are considered as unlabeled data and are used in semisupervised training. Similarly, for the SEAM dataset, we also sampled seven labels with size of 1006×782 . Subsequently, we construct the training data by utilizing sliding window in the crossline direction, with a sliding step size of 256. In the Netherlands F3 dataset, this process yields 28 labeled data slices and 2376 unlabeled data slices, and the slice size after padding is 256×256 . As for the SEAM AI challenge dataset, we obtain 28 labeled data slices, 2332 unlabeled data slices, with the slice size after padding being 1024×256 .

V. EXPERIMENT

A. Implementation Details

All experimental results in this article were obtained on NVIDIA Tesla P100 (16 G) GPU. For example, in the Netherlands F3 dataset, we use the Adam optimizer with an initial learning rate of 0.0001. The learning rate was adjusted using the StepLR strategy, where the step size was set to 4 and the gamma value was 0.1. A total of eight epochs were executed, with each epoch comprising 1188 iterations. To ensure fairness in the comparative experiments, the same encoder–decoder model, DeepLabv3+ [37], was employed, except for the semisupervised method.

B. Experimental Results and Analysis

1) *Netherlands F3 Dataset*: The supervised method exclusively utilizes labeled data for training, serving as the baseline

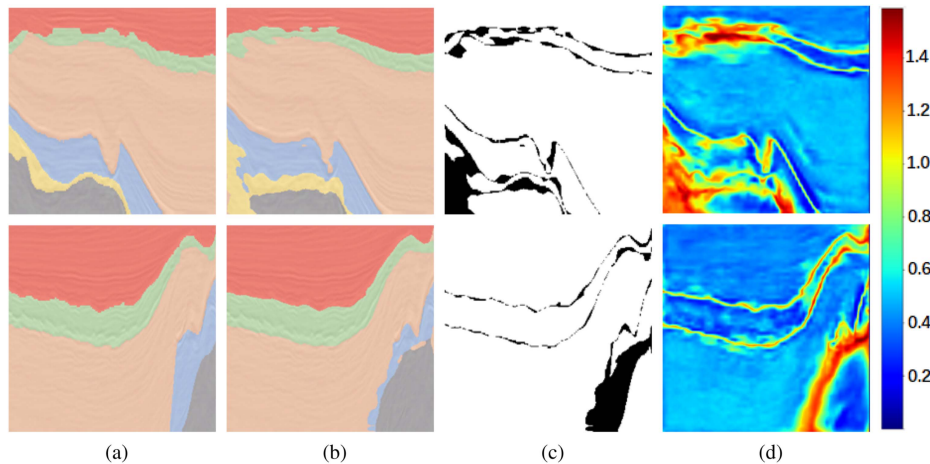


Fig. 8. Correlation between information entropy value of classification probability and unreliable pseudolabels. (a) Ground Truth. (b) Result by supervised. (c) Unreliable pseudolabels. (d) Information entropy.

TABLE I
METRICS OF NETHERLANDS F3

Method	PA	Class Accuracy						MCA	MIOU	F1
		Upper N.S.	Middle N.S.	Low N.S.	Rijnland/Chalk	Scruff	Zechstein			
Supervised	96.62	98.38	93.27	99.14	89.41	92.70	79.53	92.09	87.12	92.87
Pseudolabel	97.04	98.42	93.45	99.11	90.84	92.85	79.88	92.43	87.83(+0.71)	93.31
Mean teacher	97.38	98.80	94.57	98.76	92.82	93.26	86.92	94.19	89.06(+1.94)	94.05
Cross pseudosupervision	97.44	98.91	94.33	99.15	92.73	91.11	87.93	94.02	89.24(+2.12)	94.16
Cross-consistency training	97.48	98.78	94.26	99.05	93.71	91.62	90.04	94.58	89.78(+2.66)	94.49
CONSS w/ single-level	97.52	98.87	94.51	98.77	93.37	93.03	93.15	95.28	90.01(+2.89)	94.62
CONSS w/ dual-level	97.66	98.95	94.31	98.89	94.80	92.65	94.09	96.52	90.73(+3.61)	95.04

The bold represents the maximum value in the column.

model in this article. The pseudolabel [23] method is offline self-training, where pseudolabels are generated by the supervised method and added to the training set for retraining. The self-training achieves performance gains over the supervised baseline by +0.71 mean Intersection over Union (MIOU). The mean teacher [38] is an online self-training method that performs better than offline learning (+1.94). However, it is worth noting that the self-training method is susceptible to interference from unreliable pseudolabels shown in Fig. 8, and in some cases, its classification performance may deteriorate.

The cross pseudosupervision [22] method based on consistency regularization has demonstrated superior performance compared with self-training (+2.12). In contrast to the self-training method, this method utilizes soft pseudolabels represented as probability distributions. However, the reliability of the soft pseudolabels is affected when the entropy value is high. The cross-consistency training [21] is also a consistency regularization method designed to promote consistency under different perturbations (+2.66).

We observed a strong correlation between unreliable pseudolabels and high entropy values shown in Fig. 8, so the proposed CONSS employed a confidence strategy to mitigate the influence of unreliable pseudolabels. Furthermore, it utilizes fine-grained pixel-level contrastive learning at the feature level instead of directly using pseudolabels as supervisory signals, resulting in enhanced robustness and interpretability. The single level is

a traditional single confidence threshold strategy, whereas the dual level utilizes the dual-level confidence strategy proposed in this article. Our proposed method demonstrates state-of-the-art performance (+3.61) given in Table I, as evidenced by both quantitative metrics and visual effects shown in Fig. 9.

2) *SEAM AI Challenge Dataset*: Compared with the Netherlands F3 dataset, the seismic facies change of the SEAM AI challenge dataset shown in Fig. 10 is more complex, even so, our propose method still achieves state-of-the-art performance (+4.59) given in Table II.

C. Confusion Matrix and Visualization of the Feature Space

By analyzing the confusion matrix shown in Fig. 11 of classification on the Netherlands F3 dataset, we observed that the supervised model frequently misclassifies the Zechstein as Scruff, as well as encounters similar misclassifications with the Rijnland/Chalk and its neighboring facies. Through the utilization of t-SNE [39] for dimensionality reduction and visualization of the features, we obtained feature distributions for both the supervised model and the proposed CONSS, as shown in Figs. 13 and 14. It is obvious that the feature distribution of the supervised model exhibits significant scattering within the same class, particularly for the Rijnland/Chalk, Scruff, and Zechstein. This scattered distribution of the same class leads to substantial confusion between the distinct class features, thereby hindering

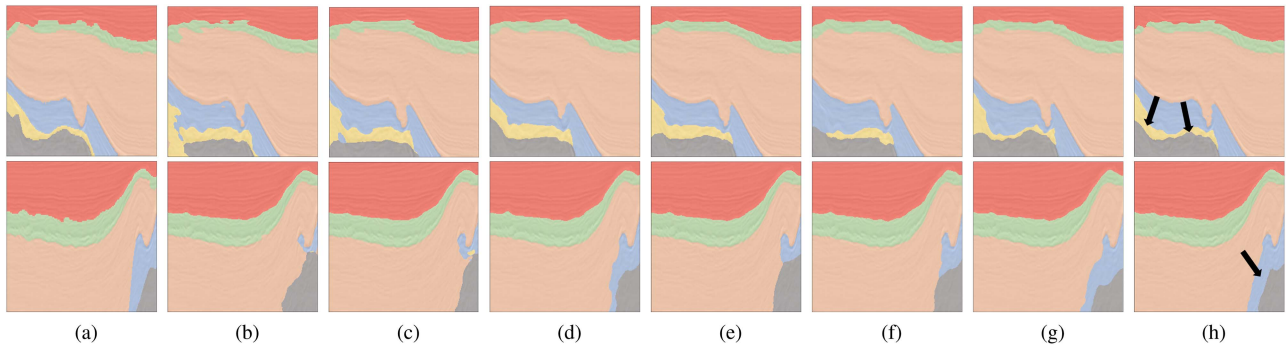


Fig. 9. Visualized results on the Netherlands F3 dataset, and the black arrows indicate the improvement. (a) Ground Truth. (b) Result by supervised. (c) Result by pseudolabel. (d) Result by mean teacher. (e) Result by cross pseudosupervision. (f) Result by cross-consistency training. (g) Result by CONSS w/ single level. (h) Result by CONSS w/ dual level.

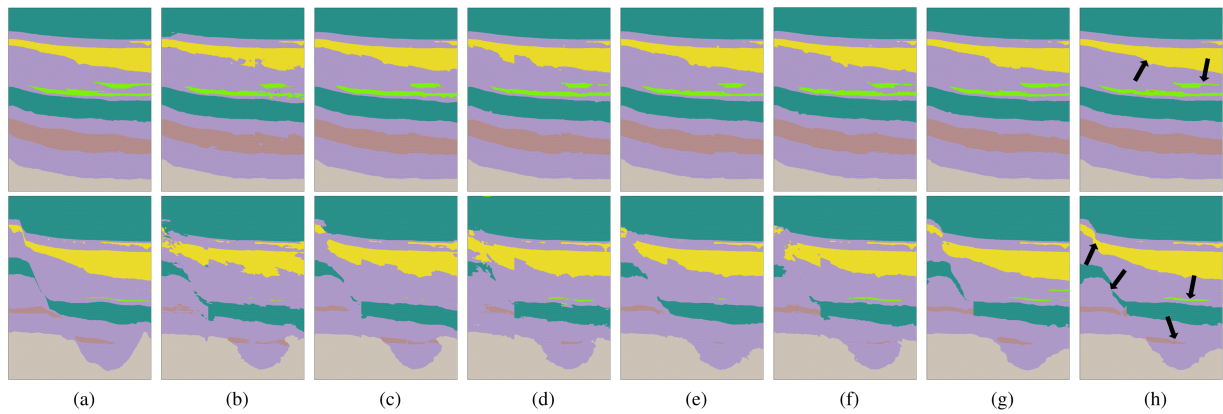


Fig. 10. Visualized results on the SEAM AI challenge dataset, and the black arrows indicate the improvement. (a) Ground Truth. (b) Result by supervised. (c) Result by pseudolabel. (d) Result by mean teacher. (e) Result by cross pseudosupervision. (f) Result by cross-consistency training. (g) Result by CONSS w/ single level. (h) Result by CONSS w/ dual level.

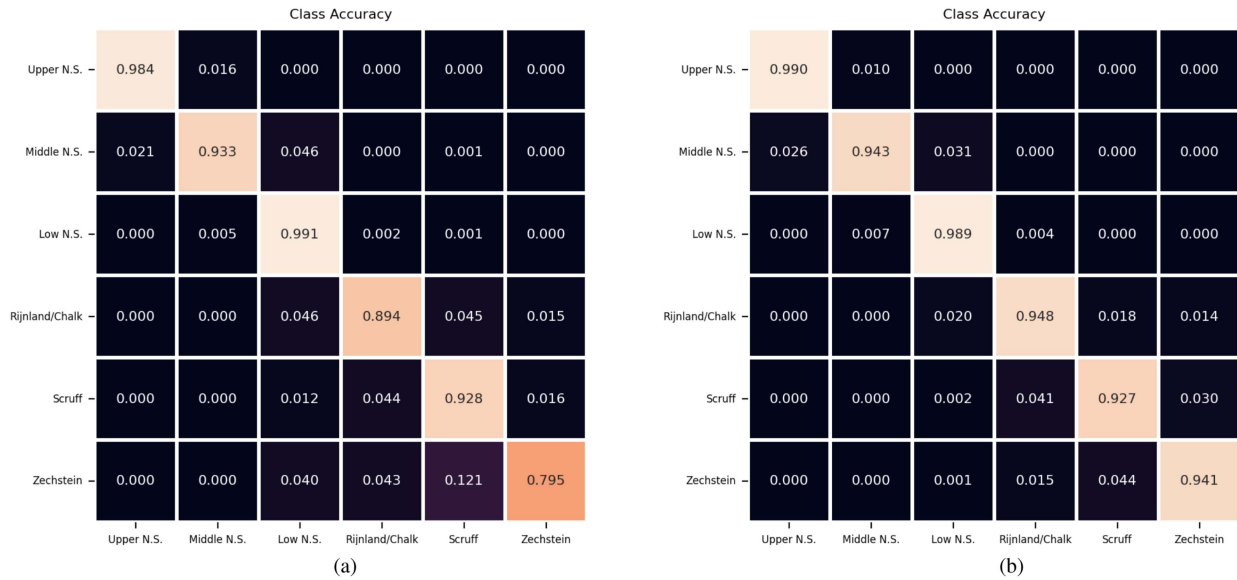


Fig. 11. Confusion matrix of classification on the Netherlands F3 dataset. Classification confusion mainly occurs in Zechstein and Scruff facies, Rijnland/Chalk, and its adjacent facies. (a) Supervised. (b) CONSS.

TABLE II
METRICS OF SEAM AI CHALLENGE

Method	Class Accuracy									
	PA	B/O	SMA	MTD	SMB	SV	SCS	MCA	MIOU	F1
Supervised	95.04	98.43	93.99	85.15	98.36	71.47	90.76	89.69	83.40	90.38
Pseudolabel	95.57	98.71	94.35	87.01	98.36	75.94	93.24	91.27	85.01(+1.61)	91.44
Mean teacher	95.74	98.69	94.23	89.03	98.37	74.33	94.99	91.61	85.38(+1.98)	91.67
Cross pseudosupervision	96.55	98.25	96.71	87.19	98.60	69.34	93.63	90.62	86.54(+3.14)	92.28
Cross-consistency training	96.18	98.54	95.06	91.12	98.47	78.61	94.38	92.70	86.56(+3.16)	92.41
CONSS w/ single-level	96.57	98.34	95.83	90.23	98.90	75.85	94.87	92.34	87.09(+3.69)	92.69
CONSS w/ dual-level	96.67	98.74	95.68	91.98	98.90	78.32	94.88	93.08	87.99(+4.59)	93.30

The bold represents the maximum value in the column.

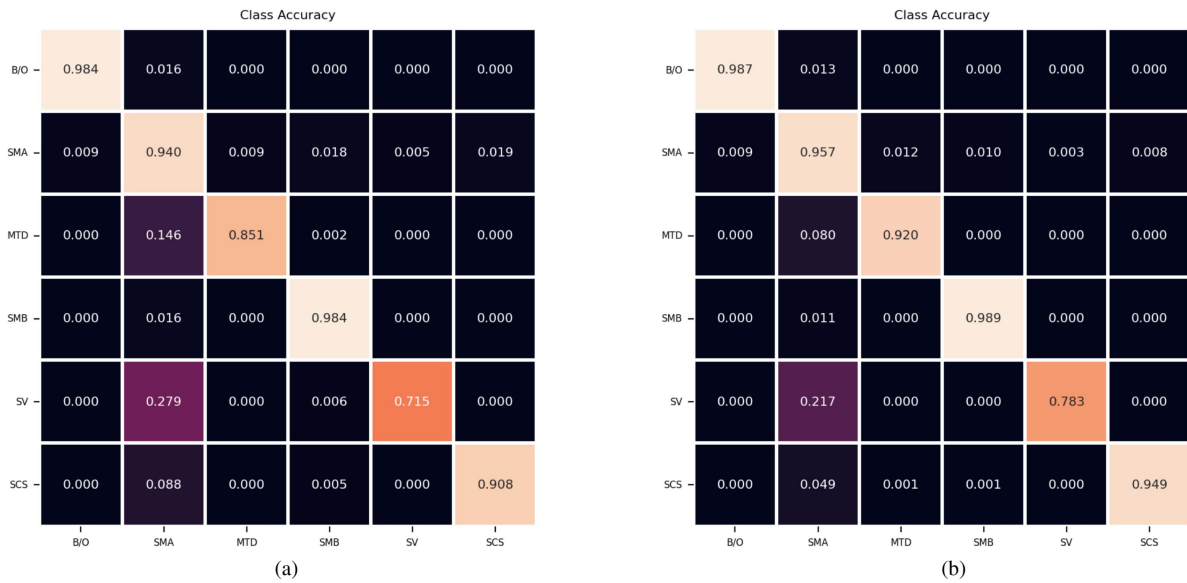


Fig. 12. Confusion matrix of classification on the SEAM AI challenge dataset. Classification confusion mainly occurs in Mass Transport Deposit, Slope Valley, and Slope Mudstone A. (a) Supervised. (b) CONSS.

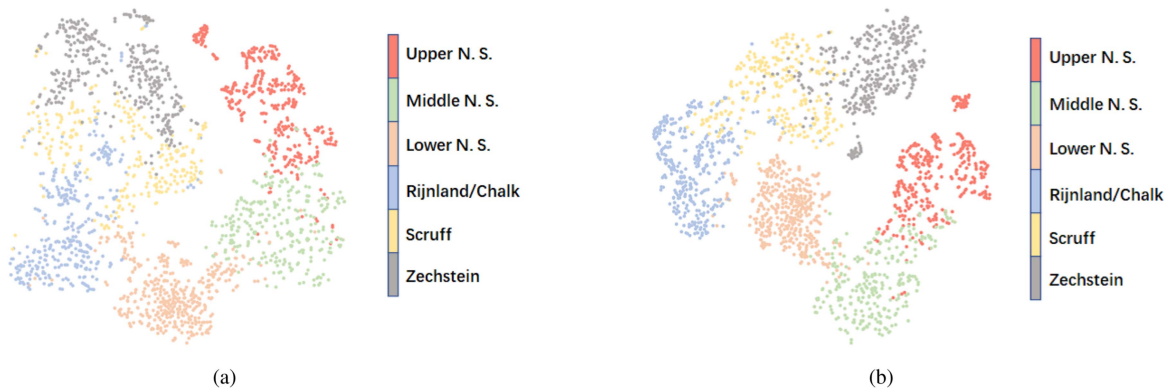


Fig. 13. Visualization of the feature space learned on the Netherlands F3 dataset, using t-SNE [39]. (a) Supervised. (b) CONSS.

accurate classification. In contrast, the feature distribution in same class of CONSS is more concentrated.

In the SEAM AI challenge dataset, we observe a similar occurrence of misclassifications in the Mass Transport Deposit, Slope Valley, and Slope Mudstone A shown in Fig. 12. In CONSS,

through the contrastive learning, the boundaries between these classes become more discernible.

In aggregate, these feature visualization outcomes offer a compelling and intuitive explanation of the pivotal role played by contrastive learning.

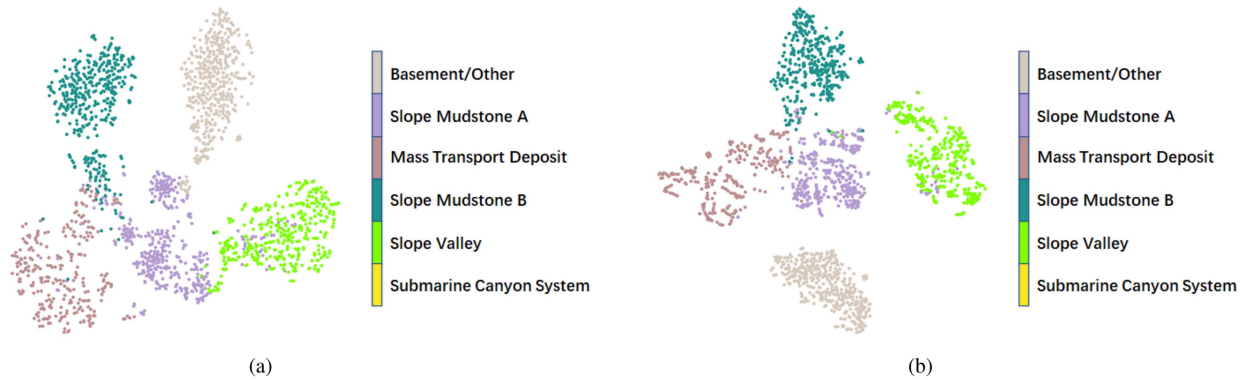


Fig. 14. Visualization of the feature space learned on the SEAM AI challenge dataset, using t-SNE [39]. (a) Supervised. (b) CONSS.

TABLE III
SMOOTHING FACTOR

ϵ	0	0.05	0.1	0.2
Netherlands F3	89.56	90.31	90.73	90.71
SEAM AI challenge	87.51	87.41	87.99	82.87

The bold represents the maximum value in the column.

TABLE IV
CONFIDENCE THRESHOLD

t_w	0.6	0.6	0.7	0.7
t_s	0.8	0.85	0.9	0.95
Netherlands F3	90.23	89.57	90.73	90.30
SEAM AI challenge	87.58	87.34	87.32	87.99

The bold represents the maximum value in the column.

TABLE V
FLOATING POINT OPERATIONS

Method	FLOPs
Supervised	16.38GFLOPs
Pseudolabel	16.38GFLOPs
Mean teacher	16.38GFLOPs
Cross pseudosupervision	16.38GFLOPs
Cross-consistency training	26.08GFLOPs
CONSS	18.93GFLOPs

D. Ablation Experiments and Discussion

In order to evaluate the impact of label smoothing in the supervised loss, we performed ablation experiments on the two datasets. Specifically, we varied the parameter ϵ for label smoothing, where $\epsilon = 0$ denotes the absence of label smoothing. The results indicate that appropriate label smoothing can significantly enhance the model's performance on both datasets given in Table III.

We empirically preset different confidence intervals, and selected the optimal t_w and t_s values for the two datasets based on the experimental results given in Table IV.

We computed the floating point operations (FLOPs) for each method given in Table V. The proposed CONSS method extends the contrastive learning branch, resulting in higher FLOPs compared with other semisupervised methods without branch expansion. Among the methods evaluated, cross-consistency

TABLE VI
NETHERLANDS F3

labeled data	unlabeled data	\mathcal{L}_{sup}	\mathcal{L}_{con}	MIOU
✓		✓		87.12
✓		✓	✓	87.82
✓	✓	✓	✓	90.73

TABLE VII
SEAM AI CHALLENGE

labeled data	unlabeled data	\mathcal{L}_{sup}	\mathcal{L}_{con}	MIOU
✓		✓		83.40
✓		✓	✓	83.78
✓	✓	✓	✓	87.99

training exhibits the highest FLOPs due to its reliance on multiple perturbation branches.

Furthermore, contrastive loss can serve as a regularization technique for supervised training, allowing training solely on labeled data given in Tables VI and VII. However, the current method is only applicable to situations where multiple seismic facies exist within a single seismic data slice. In the extreme case where a single slice contains only one class of seismic facies, contrastive learning cannot be employed due to the lack of negative samples from other seismic facies. Therefore, the data slice must be large enough to encompass multiple seismic facies. In the future, we will explore the method for offline storage of negative samples to address this issue.

VI. CONCLUSION

In this article, we point out a series of challenges faced in seismic facies classification and propose a simple yet effective semisupervised method, CONSS, which significantly reduces the reliance on labeled data. To address classification confusion, we employ pixel-level contrastive learning to enhance feature representation. Furthermore, to mitigate the interference caused by unreliable pseudolabels, we introduce a confidence strategy for selecting positive and negative samples. Experimental results demonstrate the superiority of our proposed method over supervised baselines, semisupervised self-training, and consistency regularization methods. The visualization of features confirms

that the proposed method significantly improves feature representation. In addition, our proposed method offers end-to-end training and seamless integration into existing encoder–decoder models.

APPENDIX

A. Evaluation Metrics

- 1) Pixel accuracy (PA) is the percentage of pixels over all classes that are correctly classified

$$PA = \frac{\sum_i |P_i \cap G_i|}{\sum_i G_i} \quad (11)$$

where the set of pixels that belong to class i is denoted as G_i , and the set of pixels classified as class i is denoted as P_i .

- 2) Class accuracy for class i (CA_i) is the percentage of pixels that are correctly classified in a class i

$$CA_i = \frac{|P_i \cap G_i|}{G_i}. \quad (12)$$

- 3) Mean class accuracy (MCA) is the average of CA over all classes

$$MCA = \frac{1}{C} \sum_i CA_i. \quad (13)$$

- 4) MIOU is defined as averaging the number of elements of the intersection of G_i and P_i over the number of elements of their union set

$$MIOU = \frac{1}{C} \sum_i \frac{|P_i \cap G_i|}{|P_i \cup G_i|}. \quad (14)$$

- 5) F1 score is the harmonic mean of precision and recall

$$F1 = \frac{1}{C} \sum_i \frac{2 \times \text{Recall}_i \times \text{Precision}_i}{\text{Recall}_i + \text{Precision}_i}. \quad (15)$$

B. Other Hyperparameter Ablation Experiments

TABLE VIII
LEARNING RATE

lr	10^{-2}	10^{-3}	10^{-4}	10^{-5}
Netherlands F3	89.27	90.22	90.73	81.04
SEAM AI challenge	86.29	87.99	87.01	78.73

TABLE IX
TEMPERATURE COEFFICIENT

τ	0.05	0.5	1	10
Netherlands F3	89.83	90.06	90.73	90.19
SEAM AI challenge	87.54	87.99	87.70	86.61

REFERENCES

- [1] T. Wrona, I. Pan, R. L. Gawthorpe, and H. Fossen, "Seismic facies analysis using machine learning," *Geophysics*, vol. 83, no. 5, pp. O83–O95, 2018.
- [2] A. Amendola, G. Gabriellini, P. Dell'Aversana, and A. Marini, "Seismic facies analysis through musical attributes," *Geophysical Prospecting*, vol. 65, no. S1, pp. 49–58, 2017.
- [3] M. M. Saggaf, M. N. Toksöz, and M. I. Marhoon, "Seismic facies classification and identification by competitive neural networks," *Geophysics*, vol. 68, no. 6, pp. 1984–1999, 2003.
- [4] Y. Dou, K. Li, H. Duan, T. Li, L. Dong, and Z. Huang, "MDA GAN: Adversarial-learning-based 3-D seismic data interpolation and reconstruction for complex missing," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, Feb. 2023, Art. no. 5905014, doi: [10.1109/TGRS.2023.3249476](https://doi.org/10.1109/TGRS.2023.3249476).
- [5] Y. Dou, K. Li, J. Zhu, T. Li, S. Tan, and Z. Huang, "MD loss: Efficient training of 3-D seismic fault segmentation network under sparse labels by weakening anomaly annotation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Aug. 2022, Art. no. 5919014, doi: [10.1109/TGRS.2022.3196810](https://doi.org/10.1109/TGRS.2022.3196810).
- [6] S. Li et al., "Deep-learning inversion of seismic data," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 3, pp. 2135–2149, Mar. 2020, doi: [10.1109/TGRS.2019.2953473](https://doi.org/10.1109/TGRS.2019.2953473).
- [7] F. Li, H. Zhou, Z. Wang, and X. Wu, "ADDCNN: An attention-based deep dilated convolutional neural network for seismic facies analysis with interpretable spatial–spectral maps," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1733–1744, Feb. 2021.
- [8] X. Chai et al., "An open-source package for deep-learning-based seismic facies classification: Benchmarking experiments on the SEG 2020 open data," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Jan. 2022, Art. no. 4507719, doi: [10.1109/TGRS.2022.3144666](https://doi.org/10.1109/TGRS.2022.3144666).
- [9] E. Tolstaya and A. Egorov, "Deep learning for automated seismic facies classification," *Interpretation*, vol. 10, no. 2, pp. SC31–SC40, 2022.
- [10] D. Civitarese, D. Szwarcman, E. V. Brazil, and B. Zadrozny, "Semantic segmentation of seismic images," 2019, *arXiv:1905.04307*.
- [11] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 9729–9738.
- [12] X. Chen, H. Fan, R. Girshick, and K. He, "Improved baselines with momentum contrastive learning," 2020, *arXiv:2003.04297*.
- [13] T. Chen, S. Kornblith, K. Swersky, M. Norouzi, and G. E. Hinton, "Big self-supervised models are strong semi-supervised learners," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, pp. 22243–22255.
- [14] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *Proc. Int. Conf. Mach. Learn.*, 2020, pp. 1597–1607.
- [15] Z. Wu, Y. Xiong, S. X. Yu, and D. Lin, "Unsupervised feature learning via non-parametric instance discrimination," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 3733–3742.
- [16] A. v. d. Oord, Y. Li, and O. Vinyals, "Representation learning with contrastive predictive coding," 2018, *arXiv:1807.03748*.
- [17] K. Kokilepersaud, M. Prabhushankar, and G. AlRegib, "Volumetric supervised contrastive learning for seismic semantic segmentation," in *Proc. 2nd Int. Meeting Appl. Geosci. Energy Soc. Exploration Geophys. Amer. Assoc. Petroleum*, 2022, pp. 1699–1703.
- [18] J. E. Van Engelen and H. H. Hoos, "A survey on semi-supervised learning," *Mach. Learn.*, vol. 109, no. 2, pp. 373–440, 2020.
- [19] D.-H. Lee et al., "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks," in *Proc. Workshop Challenges Representation Learn.*, 2013, Art. no. 896.
- [20] L. Yang, W. Zhuo, L. Qi, Y. Shi, and Y. Gao, "ST: Make self-training work better for semi-supervised semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 4268–4277.
- [21] Y. Ouali, C. Hudelot, and M. Tami, "Semi-supervised semantic segmentation with cross-consistency training," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 12674–12684.
- [22] X. Chen, Y. Yuan, G. Zeng, and J. Wang, "Semi-supervised semantic segmentation with cross pseudo supervision," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 2613–2622.
- [23] A. Saleem, J. Choi, D. Yoon, and J. Byun, "Facies classification using semi-supervised deep learning with pseudo-labeling strategy," in *Proc. SEG Int. Expo. Annu. Meeting*, 2019, Art. no. D023S018R003.
- [24] D. S. Chevitarese, D. Szwarcman, E. V. Brazil, and B. Zadrozny, "Efficient classification of seismic textures," in *Proc. Int. Joint Conf. Neural Netw.*, 2018, pp. 1–8.
- [25] J. S. Dramsch and M. Luthje, "Deep-learning seismic facies on state-of-the-art CNN architectures," in *Proc. SEG Int. Expo. Annu. Meeting*, 2018, Art. no. SEG-2018.
- [26] T. Zhao, "Seismic facies classification using different deep convolutional neural networks," in *Proc. SEG Int. Expo. Annu. Meeting*, 2018, Art. no. SEG-2018.

- [27] Y. Alaudah, P. Michałowicz, M. Alfarraj, and G. AlRegib, "A machine-learning benchmark for facies classification," *Interpretation*, vol. 7, no. 3, pp. SE175–SE187, 2019.
- [28] D. Hong et al., "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4340–4354, May 2021.
- [29] D. Hong et al., "SpectralFormer: Rethinking hyperspectral image classification with transformers," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Nov. 2021, Art. no. 5518615, doi: [10.1109/TGRS.2021.3130716](https://doi.org/10.1109/TGRS.2021.3130716).
- [30] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5966–5978, Jul. 2021.
- [31] X. Liu et al., "Deep classified autoencoder for lithofacies identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Dec. 2021, Art. no. 5909914, doi: [10.1109/TGRS.2021.3139931](https://doi.org/10.1109/TGRS.2021.3139931).
- [32] X. Liu, B. Li, J. Li, X. Chen, Q. Li, and Y. Chen, "Semi-supervised deep autoencoder for seismic facies classification," *Geophys. Prospecting*, vol. 69, no. 6, pp. 1295–1315, 2021.
- [33] R. Müller, S. Kornblith, and G. E. Hinton, "When does label smoothing help?," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 4694–4703.
- [34] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2818–2826.
- [35] M. Gutmann and A. Hyvärinen, "Noise-contrastive estimation: A new estimation principle for unnormalized statistical models," in *Proc. 13th Int. Conf. Artif. Intell. Statist. Workshop*, 2010, pp. 297–304.
- [36] S. Liu et al., "Bootstrapping semantic segmentation with regional contrast," in *Proc. Int. Conf. Learn. Representations*, Oct. 2021.
- [37] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 801–818.
- [38] Y. Liu, Y. Tian, Y. Chen, F. Liu, V. Belagiannis, and G. Carneiro, "Perturbed and strict mean teachers for semi-supervised semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 4258–4267.
- [39] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 2579–2605, 2008.



Kewen Li received the Ph.D. degree in information management and information system from Tianjin University, Tianjin, China, in 2012.

He is currently a Professor with the College of Computer Science and Technology, China University of Petroleum (East China), Qingdao, China. His research interests include artificial intelligence, machine learning, and data mining in oil and gas.



Wenlong Liu received the B.S. degree in software engineering in 2021 from the China University of Petroleum (East China), Qingdao, China, where he is currently working toward the M.S. degree in computer science and technology.

His research interests include seismic interpretation and deep learning.



Yimin Dou (Student Member, IEEE) received the B.S. degree in environmental engineering and the M.S. degree in computer technology from the University of Jinan, Jinan, China, in 2017 and 2020, respectively. He is currently working toward the Ph.D. degree in computer technology and resource information engineering with the China University of Petroleum (East China), Qingdao, China.

His research focuses on interdisciplinary research in geosciences and machine learning.



Zhifeng Xu received the B.S. degree in computer science and technology and the M.S. degree in software engineering from the Shandong University of Science and Technology, Qingdao, China, in 2018 and 2021, respectively. He is currently working toward the Ph.D. degree in advanced science and engineering computing with the China University of Petroleum (East China), Qingdao.

His research focuses on interdisciplinary research in geosciences and machine learning.

Hongjie Duan received the Ph.D. degree in mining engineering from the China University of Mining and Technology, Beijing, China, in 2001.

He is currently a Professor with the Shengli Oilfield Company, SINOPEC, Dongying, China. His research focuses on oilfield information technology and application.

Ruilin Jing received the M.S. degree in computer technology from Shanghai Jiao Tong University, Shanghai, China, in 2003.

He is currently a Senior Engineer with the Shengli Oilfield Company, SINOPEC, Dongying, China. His research focuses on oilfield data analysis and application.