ADDCNN: An Attention-Based Deep Dilated Convolutional Neural Network for Seismic Facies Analysis With Interpretable Spatial–Spectral Maps

Fangyu Li, Huailai Zhou, Zengyan Wang, and Xinming Wu

***Abstract*— With the dramatic growth and complexity of seismic data, manual seismic facies analysis has become a significant challenge. Machine learning and deep learning (DL) models have been widely adopted to assist geophysical interpretations in recent years. Although acceptable results can be obtained, the uninterpretable nature of DL (which also has a nickname "alchemy") does not improve the geological or geophysical understandings on the relationships between the observations and background sciences. This article proposes a noble interpretable DL model based on 3-D (spatial–spectral) attention maps of seismic facies features. Besides regular data-augmentation techniques, the high-resolution spectral analysis technique is employed to generate multispectral seismic inputs. We propose a trainable soft attention mechanism-based deep dilated convolutional neural network (ADDCNN) to improve the automatic seismic facies analysis. Furthermore, the dilated convolution operation in the ADDCNN generates accurate and high-resolution results in an efficient way. With the attention mechanism, not only the facies-segmentation accuracy is improved but also the subtle relations between the geological depositions and the seismic spectral responses are revealed by the spatial–spectral attention maps. Experiments are conducted, where all major metrics, such as classification accuracy, computational efficiency, and optimization performance, are improved while the model complexity is reduced.**

***Index Terms*— Attention map, deep learning (DL), dilated convolution, interpretability, seismic facies analysis.**

Manuscript received June 23, 2019; revised September 20, 2019 and January 30, 2020; accepted May 16, 2020. This work was supported in part by the Key Laboratory of Earth Exploration and Information Techniques of Ministry of Education in the Chengdu University of Technology and in part by the National Science and Technology Major Project Topic of China under Grant 2016ZX05026-001-005. *(Fangyu Li and Huailai Zhou are co-first authors.) (Corresponding author: Huailai Zhou.)*

Fangyu Li is with the Department of Electrical and Computer Engineering, Kennesaw State University, Marietta, GA 30067 USA (e-mail: clintfangyu@gmail.edu).

Huailai Zhou is with the State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Key Laboratory of Earth Exploration and Information Techniques of Ministry of Education, College of Geophysics, Chengdu University of Technology, Chengdu 610059, China (e-mail: zhouhuailai06@cdut.cn).

Zengyan Wang is with the Department of Computer Science, University of Georgia, Athens, GA 30602 USA (e-mail: zengyan@cs.uga.edu).

Xinming Wu is with the Laboratory of Seismology and Physics of Earth's Interior, School of Earth and Space Sciences, University of Science and Technology of China, Hefei 230052, China (e-mail: xinwucwp@gmail.com).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2020.2999365

I. Introduction

THE application of deep learning (DL) in geosciences is growing rapidly [1], including seismic event detection [2], earthquake location [3], lithology prediction [4], facies identification [5], remote-sensing image classification [6], [7], fault interpretation [8], [9], salt body characterization [10], hyperspectral image classification [11], and so on. However, interpretability is still Achilles' heel of artificial intelligence (AI) represented by the DL models [12]. The "Black box" nature of AI makes it easy to use, but resulting in low interpretability. From a recent discussion in *Science* [13], the top reason to avoid using the AI methods is that the results are hard to interpret, especially in the physical science and engineering areas. Thus, the high interpretability between the results and the inputs is necessary to enhance the AI, especially in the understanding of the complex geoscience topics.

Seismic facies analysis interprets geological deposition mechanisms, environments, and patterns from seismic data. With the dramatic growth of seismic data [14], manually interpreting a seismic volume becomes more and more challenging [15]. Previous machine learning (ML) methods, such as self-organizing maps (SOMs) [16], [17] and generative topographic mapping (GTM) [18], as well as the DL models, such as convolutional neural network (CNN) [19], and the associated autoencoder network, such as U-Net [20], show improved efficiency and accuracy in facies recognition. However, due to the low model interpretability, results obtained from the ML or DL models do not provide new insights from the geology or geophysics aspects.

Attempts to improve the DL interpretability have been made, for example, the semantic information embedded in human descriptions can be leveraged [21]. In addition, the model interpretation can be built by bridging the input features and the output regression results [22]. The emerging attention mechanism used in the medical imaging analysis is of great promise in targeting the structures with varying shapes and sizes [23]. Similarly, a complementary segmentation network was proposed to improve the brain identification by combining the labeled mask and its complementary branches [24]. Thus, if we can establish a relation between the inputs and the outputs, the interpretability can be improved.

The geological deposition process generates various structures, resulting in different seismic spectral responses [25]. The ideal interpretability in seismic facies analysis should include not only certain area belonging to facies but also how different spectral components contribute to the facies recognition, which are important for broadening the DL impacts and discovering the novel geoscience insights. Our previous work [26] explored the combination of multispectral features, which was an innovative effort in the existing DL-based work for facies analysis.

In this article, we develop an attention-based deep dilated convolutional neural network (ADDCNN) to perform an interpretable seismic facies analysis. Besides final facies-recognition results, the proposed ADDCNN also generates 3-D spatial–spectral attention maps. By visualizing the attention maps, understandings between the seismic facies and the multispectral features can be enhanced. To reduce the computation burden as well as maintain the high resolution, we propose to replace the convolution-pooling operation with the dilated convolution. Extensive quantitative analysis of a field application shows promising performances of the proposed ADDCNN in terms of accuracy, convergence rate, efficiency, and model complexity. In addition, 3-D attention maps provide insightful observations for further seismic facies analysis.

The contributions of this article are as follows.

1) We propose a novel DL architecture for automatic seismic facies analysis. The innovative design not only generates accurate results but also provides insightful interpretations for further geophysical studies.
2) We propose to adopt the dilated convolutions with the feature engineering based on multiple spectral components. Their combination enhances the facies analysis accuracy and efficiency simultaneously.
3) We propose to use the attention mechanism to map the facies-recognition results to the spatial–spectral features, which helps understand the relation between the facies and the geophysical data and how to select features in the future seismic facies analysis.

The remainder of this article is organized as follows. In Section II, we introduce related work about the ML- or DL-based seismic facies analysis. In Section III, we describe the proposed ADDCNN model in detail with the design of architecture, feature engineering, metrics, loss function, and related theoretical principles of the CNN, dilated convolution, and the attention mechanism. Using a field application in Section IV, we analyze the performances of our algorithm explicitly and quantitatively compared with some existing DL models. In addition, the benefits of the spatial–spectral attention maps are also discussed. Finally, the conclusions are drawn in Section V.

II. Related Work

The proposed ADDCNN model belongs to the automatic seismic facies analysis using the ML or DL models. In general, there are two categories of ML methods: unsupervised clustering and supervised classification. Some common unsupervised clustering methods include k-means clustering (KC) [15], fuzzy clustering [27], SOMs [16], GTM [18], and artificial neural networks (ANNs) [28]. Note that the feature extraction of unsupervised ML can also be achieved by the DL models, such as CNN-KC [29]. The unsupervised clustering methods typically require the manual feature extraction, which is time-consuming and requires domain expert knowledge [30]. Furthermore, the unsupervised clustering often does not maintain the structure information, which either results in spatially separated structures are misclassified as the same facies or the same deposition package is misclassified into multiple facies. Although the distance-preserving SOM [16] and the seismic stratigraphy-constrained SOM [17] have been proposed, the issue still remains unsolved. Supervised classification methods, such as support vector machine (SVM) [31], ANNs [32], and so on, add "supervision" information during the classification to improve the accuracy following the expert knowledge. In addition, semisupervised learning [18], which focuses on how to take advantages from the unlabeled data, also shows great potentials in the limited available data situations. In addition, unlike supervised learning, reinforcement learning [33] does not rely on the labels or "correct answers." The reinforcement agent finds the best possible action to maximize the reward in a given task. Nevertheless, there is still a long way to go before widely applying semisupervised learning and reinforcement learning in geophysics applications.

Recently, DL methods attract the attention of the geophysicists and geologists because of the automatic feature extraction and reliable and accurate seismic data analysis results [19]. The DL models have been used in the earth science society, such as microseismic detection [34], [35]. In addition, the CNN-based classification has been successfully applied in seismic applications [2], [8], [20]. In seismic facies analysis, both supervised [36] and unsupervised [37] DL applications can be found. DL is an end-to-end process to extract automatically the features from the original seismic data for various purposes, which increases the efficiency of interpretation and reduces the dependence on the handcrafted seismic attributes [38]. Note that with a careful manual selection of features, certain shallow learning methods can even be accurate than some DL methods [39]. However, DL models are usually more welcome because of the efficiency of sophisticated feature extraction without the requirement of domain knowledge. In addition, the DL models, especially CNN models, have another spatial awareness advantage over other supervised classification methods [20]. A 3-D patch-based CNN model and a 2-D whole seismic slice convolution encoder–decoder model were compared in [20]. However, the confidence of the interpretation results and the efficacy of the input data have not been discussed in the previous research. Although the efficacy of the input data has been discussed in a previous seismic lithology research [4], it has not been well established in seismic facies analysis. Furthermore, the interpretability of the DL models has not been touched by researchers. Feature engineering and model interpretability are important components in geophysical interpretation and can provide insightful understandings to help guide future research directions. Thus, our goal is to address this scarcity and promote DL in the geophysics society.

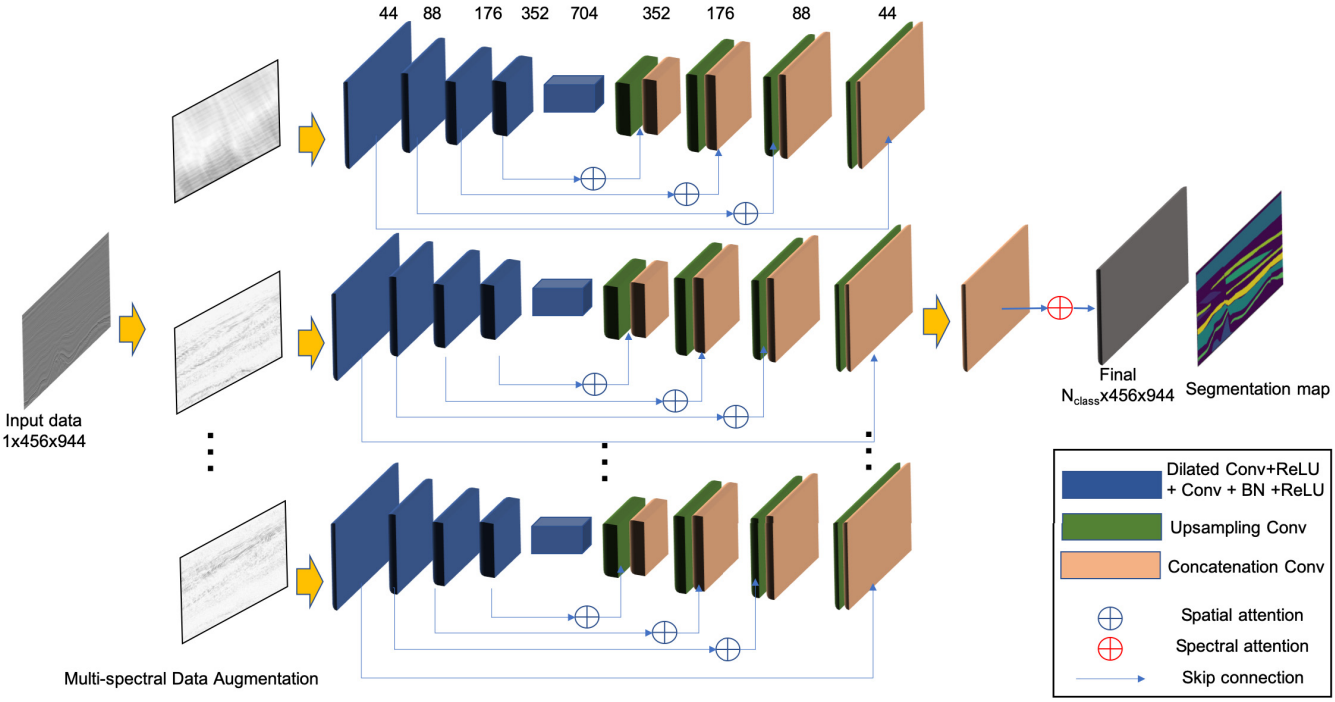


Fig. 1. Proposed ADDCNN includes three main steps: multispectral feature engineering, dilated convolutional autoencoder, and spatial–spectral attention map. The spectral decomposition-based feature engineering improves the efficiency and accuracy of the feature extraction. In addition, the dilated U-Net enhances the facies-recognition resolution and efficiency with limited model parameters. In the end, the trainable soft attention mechanism empowers the interpretability of the model to map facies-recognition results to the spatial–spectral input features. (The attention mechanism is implemented via inter- and intrastructures.) The layer numbers are annotated on the top of the architecture. Note that different block colors denote different operations, where "BN" and "Conv" are short for batch normalization and convolution, respectively.

III. ADDCNN

The proposed ADDCNN model includes three main innovative structures: feature engineering, dilated convolutions, and spatial–spectral attention mapping, as shown in Fig. 1. The ADDCNN is an end-to-end model, which processes from raw seismic data to generate final facies recognition results as well as interpretable spatial–spectral maps. Here, we introduce the ADDCNN architecture together with the principles of key steps.

A. Feature Engineering via Spectral Decomposition

Data augmentation helps the ML and DL models efficiently extract the desired features. Several well-established data-augmentation methods are applied in our model, including resize [40], rotate [41], and random cropping [42]. These regular data-augmentation operations are developed for general image-segmentation tasks rather than seismic applications. Thus, in addition to regular ones, we propose to use a spectral decomposition-based seismic feature engineering approach to enhance the hidden geological information in our model.

Seismic data contain abundant information corresponding to various responses of different structures [25]. Due to the rock properties and the seismic wave-propagation laws, geological structures with different scales generate different spectral responses [25]. Because the interpretation of individual geological features is constrained by its spectral responses [43], spectral decomposition enhances the quantitative interpretation ability [44], for example, gas hydrate accumulations can be highlighted by spectral decomposition [45]. In our previous work [26], feature engineering by spectral decomposition was adopted, which improved the image-segmentation accuracy. In addition, since the noise at different frequencies has different strengths, the combination of multispectral components can inherently improve the signal-to-noise ratio (SNR).

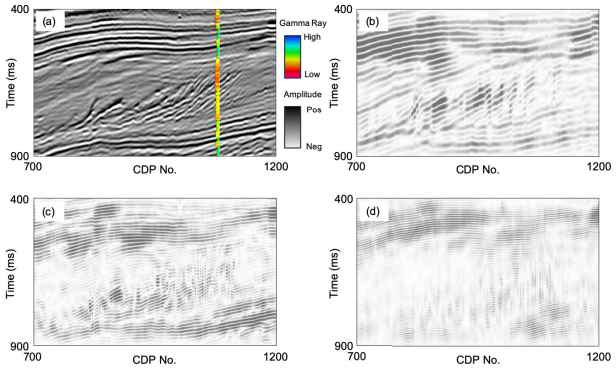


Fig. 2. Vertical slices through (a) original seismic amplitude volume and its corresponding isofrequency spectral components at (b) 10, (c) 40, and (d) 60 Hz, respectively.

In this article, we employ the wavelet transform to implement the spectral decomposition [4]. Fig. 2 shows an example of the seismic data used in this article (the data details will be introduced in Section IV-A) and its isofrequency spectral components at 10, 40, and 60 Hz, respectively. It is clear that different spectral responses highlight geological structures in different scales, which is the reason why we use the

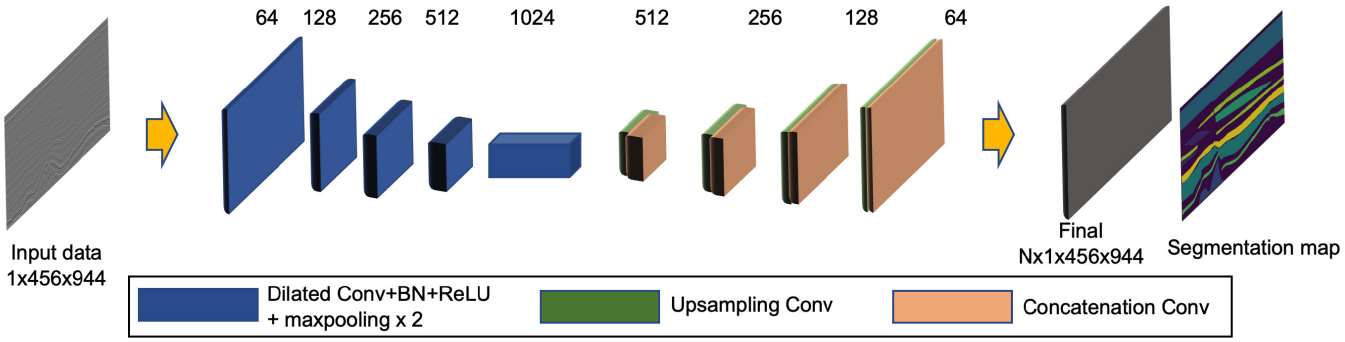


Fig. 3. Overview of the dilated U-Net architecture. Advanced features are extracted by convolution or dilated convolution layers, which have the same resolution as that of the original input.

spectral decomposition as the feature engineering to improve the robustness and effectiveness of the feature extraction.

B. CNN and Fully Convolutional Network (FCN)

The most important concept of the CNN is the convolution operation. Assuming that the output and input of layer l are $y^{(l)}$ and $z^{(l)}$, the convolution operation can be given as

$$y^{(l)} = f(z^{(l)}) \quad \text{with } z^{(l)} = w^{(l)} y^{(l-1)} + b \tag{1}$$

where $w^{(l)}$ denotes the convolutional kernel from layer $(l-1)$ to layer l, b is the bias value, and $f(\cdot)$ represents an activation function, such as, the rectified linear unit (ReLU) [$f(x) = \max(0, x)$] or sigmoid function [$f(x) = e^x/(1+e^x)$].

For computer vision problems, other than using pixel values individually, the CNN analyzes patterns among the local pixels by a convolutional kernel function w defined in (1). Unlike image classification that targets at recognizing stable objects from different images, seismic facies analysis aims at distinguishing different geological structures within the same image. Therefore, rather than a classification, seismic facies analysis is indeed an image-segmentation task to label different areas of a seismic image with different facies [20].

As an example of CNN models, the FCN is a DL framework designed for image segmentation [46]. The FCN shows the segmentation result of each corresponding area. However, because of the downsampling operations (pooling) in the FCN models, the result resolution is typically low, whereas if there is no downsampling (pooling) operations in the CNN models, the amount of parameters could be unaffordable for limited computer resources, especially the memory.

C. Dilated U-Net

To solve the low-resolution issue of the FCN, a "U"-shaped architecture U-Net [47] was proposed. The learnable upsampling convolution layers replace the pooling layers, increasing the output image resolution. Thus, the U-Net-type DL models have been used for the semantic pixelwise segmentation task in geophysics [19], [48]. The proposed ADDCNN model is built on the widely adopted U-Net, because the seismic facies-recognition problem has limited training labels but contaminated with all kinds of interferences, and U-Net has been proved an effective model for image segmentation with small training data [47].

To address further the low-resolution problem caused by the downsampling layers (such as the pooling layers) in the DL models and leverage effectively more precious spatial information, we propose to adopt the dilated convolution [49] in our model. The dilated convolution can be expressed as

$$y[i] = \sum_k x[i + d \cdot k] w[k] \tag{2}$$

where x and y are the input and output feature maps, w denotes the convolution kernel, and d corresponds to the dilated rate. The dilated convolution is equivalent to applying an upsampling operation by inserting $d-1$ zeros between two samples. Thus, a standard convolution can be viewed as a special case of dilated convolution with a dilated rate $d = 1$. In our model, optimal dilation factors are selected to expand the receptive fields efficiently and properly. Furthermore, note that, in the dilated convolution, larger regions are covered with the same amount of or even less number of parameters.

Fig. 3 shows a dilated U-Net architecture, which is used for comparison in Section IV-D. In the traditional U-Net, the spatial information is acquired by the CNN and the pooling layers [20]. However, the local information and resolution are gradually lost through layers, resulting in coarse segmentation results without subtle information [50]. Nevertheless, the dilated convolution defined in (2) expands the receptive fields and increases the kernel space without increasing the computation, which is desired for "big data" seismic processing in this article.

D. Spatial–Spectral Attention

Based on the above feature engineering and dilated U-Net model, a high-resolution seismic facies-recognition result can be obtained. However, the model is still a "black box" until now. To bridge the seismic facies analysis results and the features, we propose to use the attention mechanism. Attention-gated networks have been proposed to learn the feature-attention probabilities [23], [51], [52], where the attention mechanism was soft trainable. Similarly, we compute the attention maps to highlight the seismic image regions contributing to the facies-segmentation task. Besides the benefits of interpretability, the attention mechanism also enforces the model to suppress the interference from the untargeted information.

TABLE I
PROPOSED ADDCNN MODEL ARCHITECTURE DETAILS

Modules	Operations in Each Module	Parameters	Sizes
Contracting Block 1	Dilated Conv (1 layer), Convolution (1 layer), Batch Norm and ReLU	dilated rate = 2, kernel size = 3, padding = 1, stride = 1, output channel = 44	feature map = 452x940 receptive field = 9x9
Contracting Block 2	Dilated Conv (1 layer), Convolution (1 layer), Batch Norm and ReLU	dilated rate = 2, kernel size = 3, padding = 1, stride = 1, output channel = 88	feature map = 448 x 936 receptive field = 17 x 17
Contracting Block 3	Dilated Conv (1 layer), Convolution (1 layer), Batch Norm and ReLU	dilated rate = 2, kernel size = 3, padding = 1, stride = 1, output channel = 176	feature map = 444 x 932 receptive field = 25 x 25
Contracting Block 4	Dilated Conv (1 layer), Convolution (1 layer), Batch Norm and ReLU	dilated rate = 2, kernel size = 3, padding = 1, stride = 1, output channel = 352	feature map = 440 x 928 receptive field = 33 x 33
Bottleneck Block	Dilated Conv (5 layers), Batch Norm, and ReLU	dilated rate =4,8,16,32,64, kernel size = 3, stride = 1, padding = 4,8,16,32,64, output channel = 704	feature map = 440 x 928 receptive field = 290 x 290
Expansion Block 1	Upsample (1 layer), Convolution (1 layer), Batch Norm and ReLU	kernel size = 3, output channel = 352, upsample mode = bilinear, padding = 1, stride = 1	feature map = 440 x 928 receptive field = 290 x 290
Expansion Block 2	Upsample (1 layer), Convolution (1 layer), Batch Norm and ReLU	kernel size = 3, output channel = 176, upsample mode = bilinear, padding = 1, stride = 1	feature map = 444 x 932 receptive field = 292 x 292
Expansion Block 3	Upsample (1 layer), Convolution (1 layer), Batch Norm and ReLU	kernel size = 3, output channel = 88, upsample mode = bilinear, padding = 1, stride = 1	feature map = 448 x 936 receptive field = 294 x 294
Expansion Block 4	Upsample (1 layer), Convolution (1 layer), Batch Norm and ReLU	kernel size = 3, output channel = 44, upsample mode = bilinear, padding = 1, stride = 1	feature map = 452 x 940 receptive field = 296 x 296
Attention Block	Convolution (3 spatial layers, 1 spectral layer), Batch Norm, ReLU, and Interpolate	kernel size = 1, stride = 1, padding = 0, interpolate mode = bilinear	NA
Optimizer	Adam	start learning rate = 0.0005, weight decay=0.0001, learning rate decay = polynomial	NA

We denote the attention map as $a_i^{s,\omega}$ at the spatial location i of n total spatial locations in a given DL layer $s \in \{1, \ldots, S\}$ of the spectral component ω. For a certain spectral component $\omega*$, the "attention" $\mathcal{A}^{s,\omega*}$ is defined as the normalized compatibility scores by a softmax operation [51]

$$\mathcal{A}^{s,\omega*} = \left\{a_1^{s,\omega*}, a_2^{s,\omega*}, \ldots, a_n^{s,\omega*}\right\} \tag{3}$$

$$a_i^{s,\omega*} = \frac{\exp(c_i^{s,\omega*})}{\sum_j^n \exp(c_j^{s,\omega*})}, \quad i \in \{1, \ldots, n\} \tag{4}$$

where the set of compatibility scores $\mathcal{C}(\hat{\mathcal{L}}^s, \boldsymbol{g}) = \{c_1^s, c_2^s, \ldots, c_n^s\}$, is defined as the linear mapping from the set of feature vectors $\mathcal{L}^s = \{\boldsymbol{\ell}_1^s, \boldsymbol{\ell}_2^s, \ldots, \boldsymbol{\ell}_n^s\}$ extracted at layer s to the global feature vector $\boldsymbol{g}$. Furthermore, $\boldsymbol{g}^{s,\omega*} = \sum_{i=1}^n a_i^{s,\omega*} \cdot \boldsymbol{\ell}_i^{s,\omega*}$ can be viewed as an elementwise weighted averaging operation. At this moment, the spatial attention, indicated as blue $\bigoplus$ in Fig. 1, has been established.

Combining the spectral attention red $\bigoplus$, the spatial–spectral attention map can be obtained by a linear model

$$\mathcal{A} = \sum_{i=1}^{M} \phi_i \mathcal{A}^{\omega_i} \tag{5}$$

where M is the number of spectral features and ϕ_i are the weights (i.e., contributions). Thus, the contributions of different spectral features can be visualized easily.

E. Loss Function

We employ the cross entropy $E = \sum_{\mathbf{x}\in\Omega} \omega(\mathbf{x}) \log(p_{l_\mathbf{x}}(\mathbf{x}))$ to be the loss function, where $\ell : \Omega \to \{1, \ldots, K\}$ is the true label of each pixel and $\omega : \Omega \to \mathbb{R}$ is a weight map, at the pixel position $\mathbf{x} \in \Omega$ with $\Omega \subset \mathbb{Z}^2$.

F. ADDCNN Model Architecture

Based on Sections III-A–III-E, we propose the ADDCNN model, whose overview is shown in Fig. 1. The details of the model architecture can be found in Table I. In addition, the optimizer—Adam (adaptive moment estimation) [53], an efficient stochastic optimization method, is used. In addition, the cross entropy is employed as the loss function.

The main contribution of the proposed DL model is the model interpretability. Typically, human interpreters identify seismic facies using the morphological features and their geological understandings. We bridge the seismic facies-recognition results with the multispectral inputs using the attention mechanism. As shown in the model architecture,

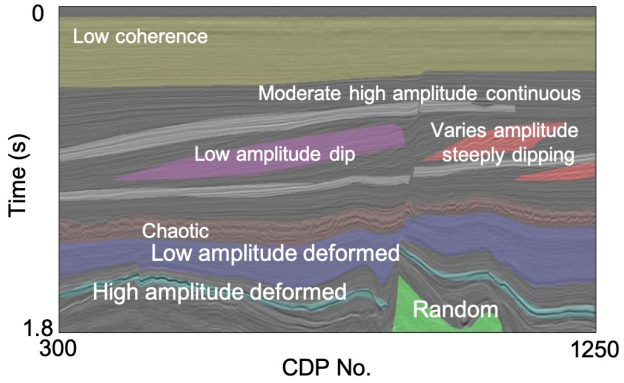


Fig. 4. Seismic facies labels corendered with seismic amplitudes. The facies names and colors are also defined in Table II.

we propose a 3-D attention mapping using two attention structures: interattention and intraattention. The intraattention helps mapping the spatial relations between the facies results and the multispectral inputs, while the interattention contributes to the spectral domain. Building on the spatial–spectral attention maps, we further implement a dilated convolution to control the parameter amount and improve the image-segmentation resolution. Our ADDCNN model is based on the U-Net architecture with the attention mechanism and the dilated convolution operation. Our model has advantages of these three DL features. All the mentioned DL architectures have been evaluated and testified using tens if not hundreds of different image data sets, including natural images, handwritten digits, and so on. Thus, in principle, the proposed model should have a good generalization ability. The detailed advantages of the proposed model will be discussed in the following sections.

IV. Experiments

A. Data Description

The 3-D seismic survey F3 [54] acquired in the North Sea is used in this article.[1] There are obvious large-scale sigmoidal beds deposited by a large fluviodeltaic system, showing clear downlap, toplap, onlap, and truncation stratigraphy structures. Based on the seismic amplitudes and the geomorphological features, we obtain the manually labeled ground truth from an independent expert. The facies and colors are defined in Fig. 4 and Table II. It is clear that the seismic facies names are based on the seismic geometric features, which means the studied seismic facies is generally defined. Because all the seismic images can be consistently normalized and the normalization does not change the seismic geometric features or the feature patterns, seismic images can be viewed as a kind of natural images.

The 3-D seismic amplitude volume is cut into 651 2-D amplitude slices. In addition, seismic data have been decomposed into 20 spectral bands from 5 to 100 Hz with a 5-Hz interval, following Section III-A. Our training data set consists of 41 slices with interval ten lines, which is less than 10% of the available data.

[1]The seismic survey is available at https://terranubis.com/datainfo/Netherlands-Offshore-F3-Block-Complete (last accessed May 6, 2019).

TABLE II
Seismic Facies and Corresponding Colors

Facies Number	Facies Name	Color
1	Varies amplitude steeply dipping	coral
2	Random	pale green
3	Low coherence	lemon
4	Low amplitude deformed	blue
5	Low amplitude dip	violet
6	High amplitude deformed	emerald
7	Moderate amplitude continuous	gray
8	Chaotic	copper
0	Everything else	transparent

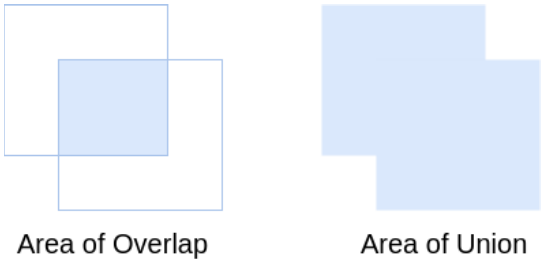


Fig. 5. Graphic definitions of IoU. The area of overlap denotes the overlap is between the prediction and the ground truth, whereas the area of union denotes the encompassed area by both the prediction and the ground truth.

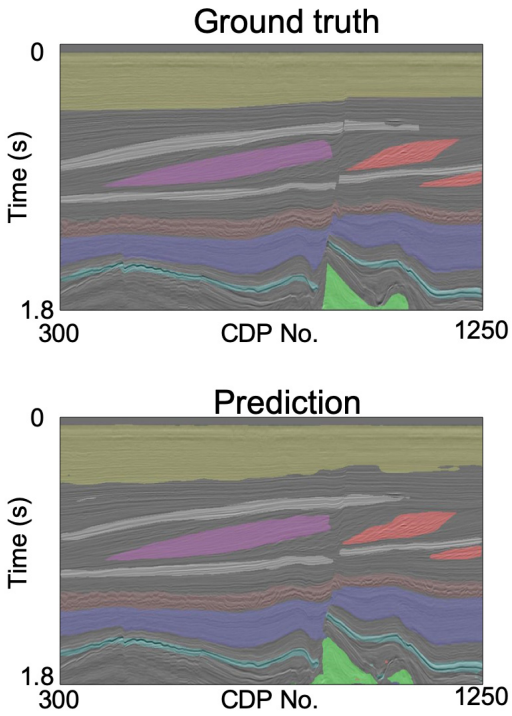


Fig. 6. Seismic facies analysis results of one seismic line using the proposed ADDCNN, compared with the ground-truth labels.

B. Experimental Setting

We implement the proposed ADDCNN model and the reference DL models (listed in Table III) using Pytorch [55]. All experiments are done on a workstation with two NVIDIA

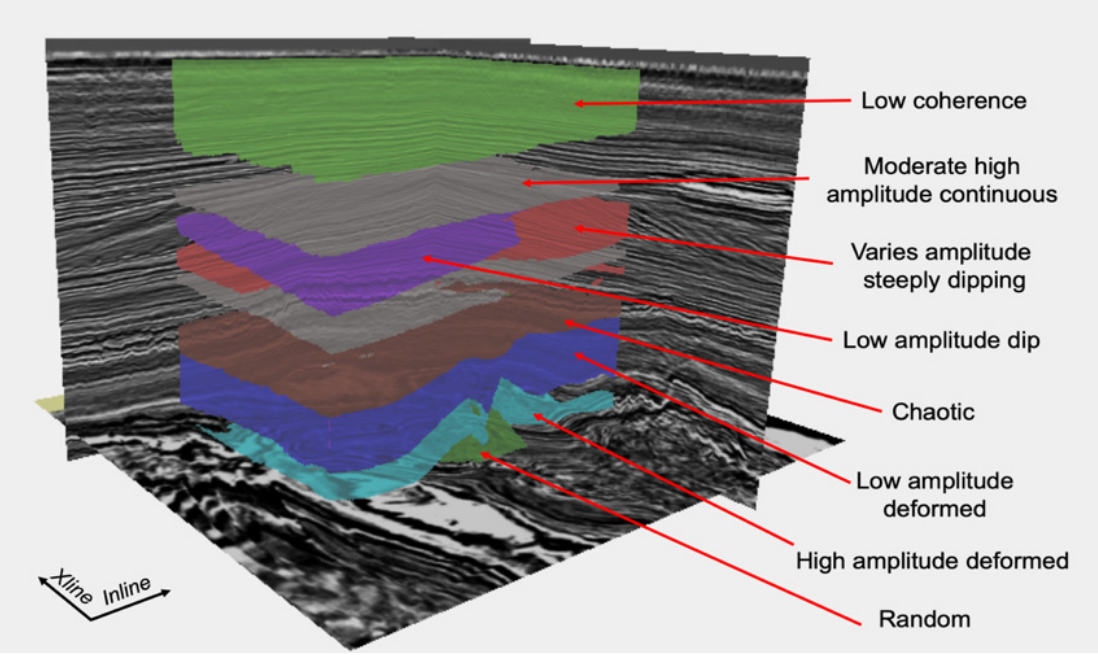


Fig. 7. 3-D seismic facies analysis result using the ADDCNN.

GeForce GTX 1080Ti GPUs. We use the tenfold random cross-validation in our experiments and divide the data set into training and test sets with a 4 : 1 proportion. A validation data set is prepared using 10% data from the training data set.

C. Metrics

Intersection over union (IoU) is adopted as the evaluation metrics [2], which is defined as the ratio between the Area of Overlap and the Area of Union, as shown in Fig. 5. In this article, instead of matching exact pixels, we want to ensure that the predicted facies match the ground truth areas as closely as possible.

D. Comparison

In this article, we compare different existing DL models: FCN 16/32 [56], SegNet [57], U-Net, U-Net with attention, Dilated U-Net, and the proposed ADDCNN. (The dilation factor is set to 2 for the dilated convolutions.) Models are compared in terms of IoU in Table III. The image-segmentation performances of different models on different seismic facies using the test data set can be compared, and the overall accuracies are also displayed. In addition, the model parameter numbers are also listed. It is obvious that the U-Net and its derivatives have less parameters than the FCN and SegNet models. Because it has the inter- and intraattention structures, the ADDCNN has more parameters than the U-Net. Since the FCN 16 has a finer resolution than the FCN 32, the FCN 16 achieves higher IoUs. SegNet [57] shares a similar architecture with the U-Net, but instead of using pooling indices, the entire feature maps are transferred from the encoder to the decoder. Therefore, the SegNet has more parameters than the U-Net. However, the boundary delineation advantage of the SegNet is not appreciated in our seismic facies task, resulting a lower overall IoU compared with the U-Net.

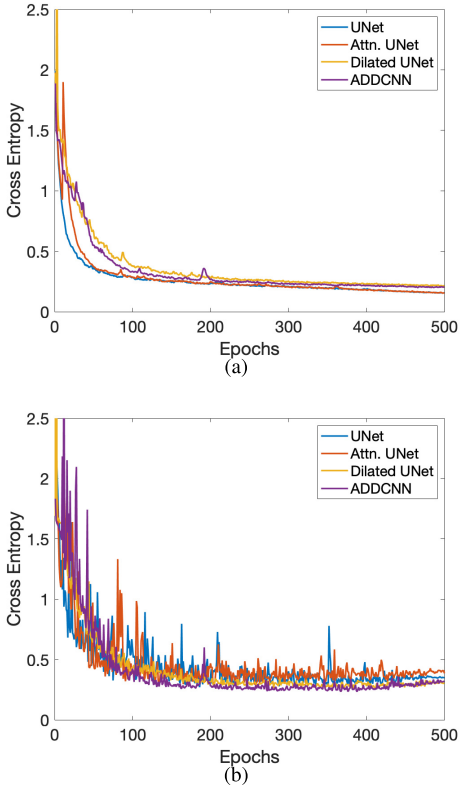


Fig. 8. Curves of validation cross entropy loss at each epoch during (a) training and (b) testing.

Fig. 6 shows the prediction results of the ADDCNN on one testing seismic line compared with the ground-truth labels. The prediction results show clean seismic facies results that do not need further processing and can be used for interpretation. The promising results indicate that the proposed ADDCNN is able to capture the seismic facies features and recognize them.

TABLE III
IOU OF EVERY FACIES OF THE TESTING DATA SET USING DIFFERENT DL MODELS

Models	Facies Number								Overall	Number of Parameter
	1	2	3	4	5	6	7	8		
FCN 32	0.0784	0.0023	0.8835	0.4844	0.4428	0.0412	0.0659	0.4157	0.302	154297417
FCN 16	0.3121	0.0072	0.9439	0.5062	0.4418	0.0998	0.6610	0.4303	0.452	153402034
SegNet	0.5079	0.0879	0.9550	0.6356	0.6810	0.1628	0.6747	0.5151	0.528	48442889
U-Net	0.7147	0.6746	0.9371	0.9243	0.7110	0.8304	0.8185	0.8942	0.813	7410313
Attention U-Net	0.7435	0.6554	0.9417	0.9031	0.7043	0.8035	0.8168	0.9104	0.811	8618252
Dilated U-Net	0.6799	0.7083	0.9642	0.9586	0.7587	0.8126	0.8035	0.8924	0.822	7864393
ADDCNN	0.7964	0.8839	0.9673	0.9633	0.8612	0.8474	0.7939	0.9041	0.877	17154731

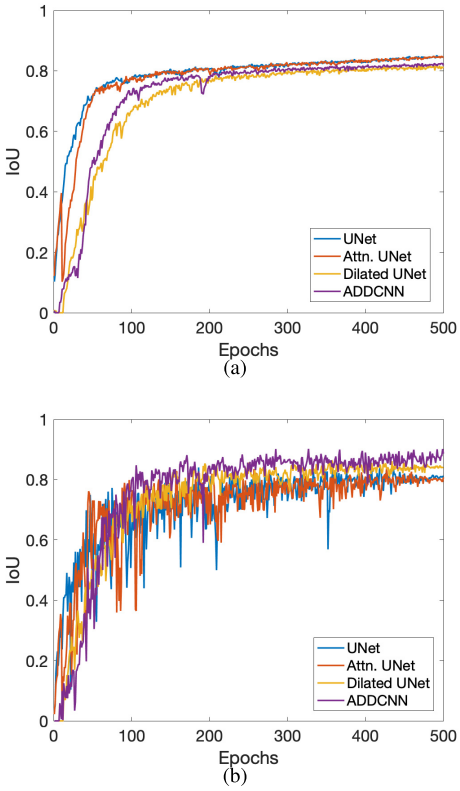


Fig. 9. Curves of validation IoU accuracy at each epoch during (a) training and (b) testing.

In addition, we also show a 3-D seismic facies analysis result in Fig. 7. Both the inline and crossline directions are available for observation, showing that our model generates acceptable lateral continuity in both directions.

To compare further the DL model performances, we demonstrate the training and validation performances of the U-Net, U-Net with attention, Dilated U-Net, and proposed ADDCNN in Figs. 8 and 9. (The other models are not as good as these four, so not compared.) The training process has hyperparameters as follows: minibatch size: 4, step decay: 1×10^{-6}, initial learning rate: 5×10^{-4}, weight decay: 4×10^{-4}, momentum factor: 0.9, epoch number: 500, and optimizer: Adam. The ADDCNN has the most promising performance, which is not prone to overfitting, is smooth, and has the best validation accuracy.

E. 3-D Attention Map

In this section, we discuss the interpretability of the proposed ADDCNN model empowered from the 3-D attention map. As discussed above, the DL models could achieve acceptable seismic facies-segmentation results, but the "black box" to recognize seismic facies does not help geoscientists extend the geological understandings of the relation between the deposition process and the seismic facies.

To unveil the correlations between the input seismic spectral components and the final facies recognition results, we build a spatial–spectral attention structure, as defined in Section III-D. Since, for a given seismic facies, other facies can be viewed as complementary information, we build inter- and intraattention structures among layers to leverage the complementary information, as shown in Fig. 1.

The studied area from the southern North Sea Basin has experienced complex stages of orogeny, rift, and subsidence that occurred during the Paleozoic and Mesozoic times, and an inversion during the Tertiary. A prograding fluviodeltaic system is constructed, because the high sediment influx filled the basin, which was from the neighboring highlands that uplifted in the late Miocene [54]. Using the low amplitude dip facies, violet color in Fig. 4, as an example, we analyze its attention maps of different spectral components to improve the seismic stratigraphic interpretation [58]. Fig. 10 shows the attention maps of 20-, 35-, and 95-Hz spectral components. The target facies area exhibits strong energies, which means the target facies has been highlighted in these spectral components. Nevertheless, we note that other untargeted areas also show strong energies, for example, the bottom on 20 and 35 Hz and the middle part on 95 Hz. Compared with the final result in Fig. 6, it is obvious that the target facies cannot be accurately recognized only using these three spectral components.

Fig. 11 demonstrates the attention maps of 10, 55, 70, and 80 Hz. It is interesting to note that the 3-D attention map for the target facies could show weak energies at the target area but strong energies at surrounding areas. Note that there are also different patterns. For example, the attention map of 10 Hz shows weak energies on a large area around the target facies, whereas the attention maps of 55 and 70 Hz highlight the surrounding areas more preciously. In addition, 70 and 80 Hz have strong energies in the top of maps, while 10 and 55 Hz have weak energies. Leveraging

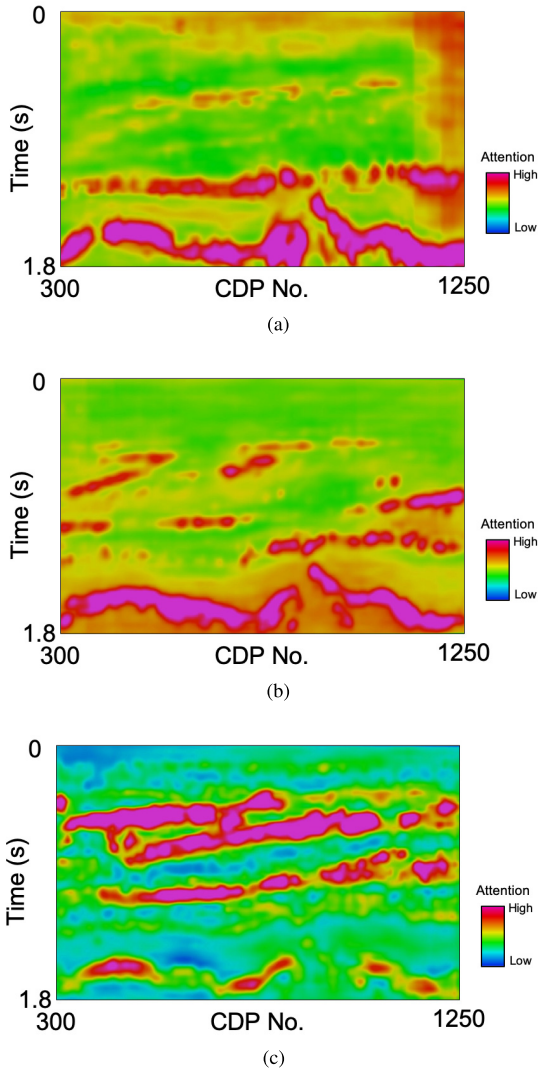


Fig. 10. Attention maps on (a) 20-, (b) 35-, and (c) 95-Hz spectral components. The target facies area and the other untargeted areas show strong energies.

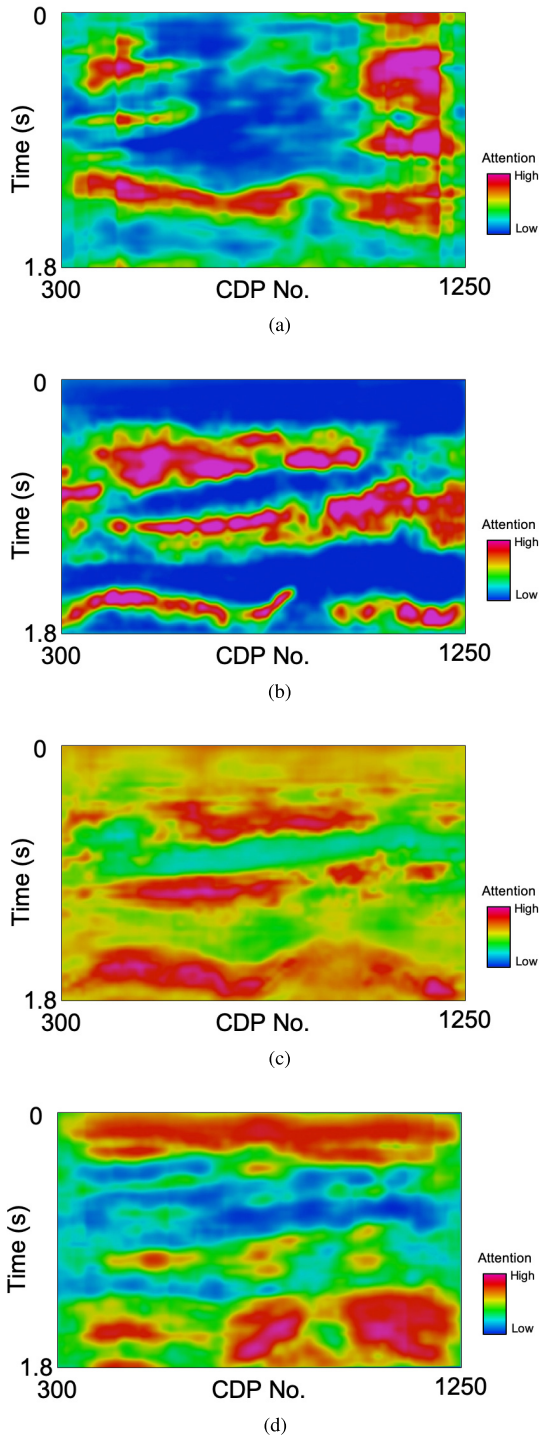


Fig. 11. Attention maps on (a) 10-, (b) 55-, (c) 70-, and (d) 80-Hz spectral components. The target facies area and other areas show strong energies.

such complementary information about the seismic facies, the robustness of the facies-segmentation method has been improved.

In general, the DL models can learn and generate what is a seismic facies and its distribution area, while they can also learn what is outside of the facies and to help the other models generate a better facies mask. Based on the feature engineering operations, the ADDCNN does not need additional inputs, as other facies can be viewed as the complementary information for the target facies. In addition, the complementary information enables the ADDCNN to handle data with visually unseen structures.

Fig. 12 shows the attention maps of 30- and 90-Hz components, where the energies are close on the whole maps. Because there is no clear difference between different areas, the attention maps with similar energies everywhere do not contribute to the facies recognition.

Given the labeled areas of one facies, an ML or a DL model focuses on extracting the features of the target areas, resulting in its difficulty of handling areas belonging to different facies but with similar features. To tackle this problem, we augment the segmentation model by adding complementary attention structures to learn the features in the untargeted region, because they provide information regarding surroundings. Direct feedback is provided to the segmentation and complementary attention structures, leading to reasonable facies predictions. In particular, the feature maps of one facies from one attention branch are converted into a probability map, which is inverted and multiplied to the feature

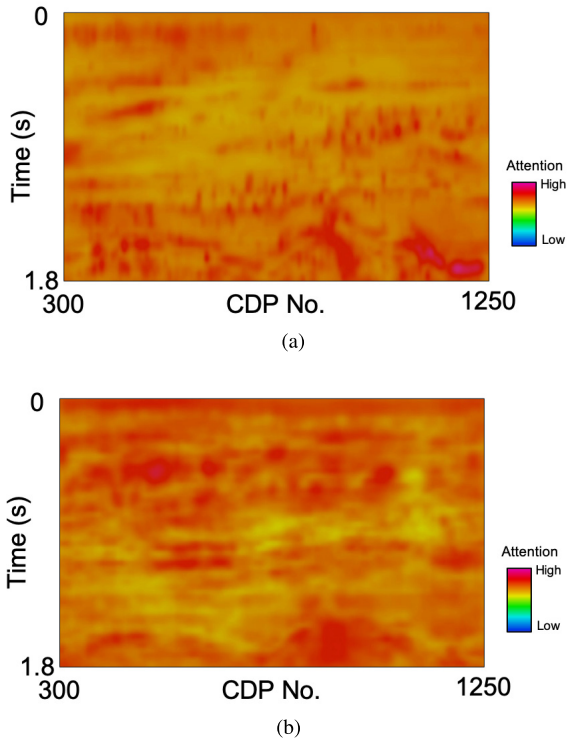


Fig. 12. Attention maps on (a) 30- and (b) 90-Hz spectral components. No facies-related information can be found.

maps of other attention branches. Essentially, one attention branch informs others to focus on the learning features of its complementary facies. According to the combination of the seismic facies-attention map and its complementary facies-attention maps, we can identify different seismic facies. This confirms that the U-Net structure works as expected [20]; more importantly, the complementary attention structure has learned features for separating the untargeted facies regions from the target facies. This enables the ADDCNN to handle the hidden geological depositional information and be insensitive to minor differences due to location differences. More specifically, the attention maps help identify the complex geological movements and seismic stratigraphy boundaries. Using different attention maps including both target and complementary facies areas, a potential relation between the geological sedimentation facies and the spectral responses can be established. Our goal is to understand why DL models can generate the facies analysis results, and the used attention mechanism provides details about the areas on which spectral components have either constructive or destructive effects on a given facies recognition. The 3-D inter- and intraattention maps present the additional "machine understanding" of the geophysical problem, leading to an improvement of the AI DL interpretability in complex geosciences topics.

V. Conclusion

We propose an ADDCNN model with a dilated U-Net structure and a 3-D soft attention mechanism to perform and improve the automatic seismic facies analysis. In terms of recognition accuracy, we have compared the IoU scores from the other six models, including the FCN 32/16, SegNet, U-Net, as well as different combinations of dilated convolution and attention mechanism. The ADDCNN generates superior segmentation results over other methods thanks to the proposed comprehensive model architecture. Furthermore, the generated attention maps can be used for explaining the geological interpretation by visualizing the spatial–spectral attention maps. We have learned that the seismic facies should be determined based on not only the target facies properties but also the surrounding properties. The complementary information provided by the untargeted facies improves the seismic facies-recognition performances.

Acknowledgment

The authors would like to thank Chuncheng Liu from the CNOOC Research Institute for his beneficial suggestions. They deeply appreciate the valuable comments from Dr. Y. Chen from Zhejiang University and another anonymous reviewer.

References

[1] K. J. Bergen, P. A. Johnson, M. V. de Hoop, and G. C. Beroza, "Machine learning for data-driven discovery in solid Earth geoscience," *Science*, vol. 363, no. 6433, Mar. 2019, Art. no. eaau0323.

[2] Y. Wu, Y. Lin, Z. Zhou, D. C. Bolton, J. Liu, and P. Johnson, "Deepdetect: A cascaded region-based densely connected network for seismic event detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 62–75, Jan. 2019.

[3] T. Perol, M. Gharbi, and M. Denolle, "Convolutional neural network for earthquake detection and location," *Sci. Adv.*, vol. 4, no. 2, Feb. 2018, Art. no. e1700578.

[4] G. Zhang, Z. Wang, and Y. Chen, "Deep learning for seismic lithology prediction," *Geophys. J. Int.*, vol. 215, no. 2, pp. 1368–1387, Aug. 2018.

[5] X. Liu, X. Chen, J. Li, X. Zhou, and Y. Chen, "Facies identification based on multikernel relevance vector machine," *IEEE Trans. Geosci. Remote Sens.*, early access, 2020, doi: 10.1109/TGRS.2020.2981687.

[6] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.

[7] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Convolutional neural networks for large-scale remote-sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 645–657, Feb. 2017.

[8] X. Wu, L. Liang, Y. Shi, and S. Fomel, "FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation," *Geophysics*, vol. 84, no. 3, pp. IM35–IM45, May 2019.

[9] X. Wu, Y. Shi, S. Fomel, L. Liang, Q. Zhang, and A. Z. Yusifov, "FaultNet3D: Predicting fault probabilities, strikes, and dips with a single convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 9138–9155, Nov. 2019.

[10] Y. Shi, X. Wu, and S. Fomel, "SaltSeg: Automatic 3D salt segmentation using a deep convolutional neural network," *Interpretation*, vol. 7, no. 3, pp. SE113–SE122, Aug. 2019.

[11] J. Feng *et al.*, "CNN-based multilayer spatial–spectral feature fusion and sample augmentation with local and nonlocal constraints for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 4, pp. 1299–1313, Apr. 2019.

[12] Q.-S. Zhang and S.-C. Zhu, "Visual interpretability for deep learning: A survey," *Frontiers Inf. Technol. Electron. Eng.*, vol. 19, no. 1, pp. 27–39, Jan. 2018.

[13] E. A. Holm, "In defense of the black box," *Science*, vol. 364, no. 6435, pp. 26–27, Apr. 2019.

[14] J. Clemente, F. Li, M. Valero, A. Chen, and W. Song, "ASIS: Autonomous seismic imaging system with *in situ* data analytics and renewable energy," *IEEE Syst. J.*, vol. 14, no. 1, pp. 1277–1284, Mar. 2019.

[15] T. Zhao, V. Jayaram, A. Roy, and K. J. Marfurt, "A comparison of classification techniques for seismic facies recognition," *Interpretation*, vol. 3, no. 4, pp. SAE29–SAE58, Nov. 2015.

[16] T. Zhao, J. Zhang, F. Li, and K. J. Marfurt, "Characterizing a turbidite system in Canterbury basin, New Zealand, using seismic attributes and distance-preserving self-organizing maps," *Interpretation*, vol. 4, no. 1, pp. SB79–SB89, Feb. 2016.

[17] T. Zhao, F. Li, and K. J. Marfurt, "Constraining self-organizing map facies analysis with stratigraphy: An approach to increase the credibility in automatic seismic facies classification," *Interpretation*, vol. 5, no. 2, pp. T163–T171, May 2017.

[18] J. Qi, T. Lin, T. Zhao, F. Li, and K. Marfurt, "Semisupervised multiattribute seismic facies analysis," *Interpretation*, vol. 4, no. 1, pp. SB91–SB106, Feb. 2016.

[19] L. Huang, X. Dong, and T. E. Clee, "A scalable deep learning platform for identifying geologic features from seismic attributes," *Lead. Edge*, vol. 36, no. 3, pp. 249–256, Mar. 2017.

[20] T. Zhao, "Seismic facies classification using different deep convolutional neural networks," in *SEG Technical Program Expanded Abstracts*. Tulsa, OK, USA: SEG, 2018, pp. 2046–2050.

[21] Y. Dong, H. Su, J. Zhu, and B. Zhang, "Improving interpretability of deep neural networks with semantic information," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4306–4314.

[22] X. Liu, J. Chen, J. Vaughan, V. Nair, and A. Sudjianto, "Model interpretation: A unified derivative-based framework for nonparametric regression and supervised machine learning," 2018, *arXiv:1808.07216*. [Online]. Available: http://arxiv.org/abs/1808.07216

[23] O. Oktay *et al.*, "Attention U-net: Learning where to look for the pancreas," 2018, *arXiv:1804.03999*. [Online]. Available: http://arxiv.org/abs/1804.03999

[24] R. Dey and Y. Hong, "CompNet: Complementary segmentation network for brain MRI extraction," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* New York, NY, USA: Springer, 2018, pp. 628–636.

[25] F. Li and W. Lu, "Coherence attribute at different spectral scales," *Interpretation*, vol. 2, no. 1, pp. SA99–SA106, Feb. 2014.

[26] Z. Wang, F. Li, T. Taha, and H. Arabnia, "2D multi-spectral convolutional encoder-decoder model for geobody segmentation," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Dec. 2018, pp. 1193–1198.

[27] A. E. Barnes and K. J. Laughlin, "Investigation of methods for unsupervised classification of seismic data," in *SEG Technical Program Expanded Abstracts*. Tulsa, OK, USA: SEG, 2002, pp. 2221–2224.

[28] T. Zhao and K. Ramachandran, "Performance evaluation of complex neural networks in reservoir characterization: Applied to Boonsville 3-D seismic data," in *SEG Technical Program Expanded Abstracts 2013*. Tulsa, OK, USA: SEG, 2013, pp. 2621–2625.

[29] Y. Chen, G. Zhang, M. Bai, S. Zu, Z. Guan, and M. Zhang, "Automatic waveform classification and arrival picking based on convolutional neural network," *Earth Space Sci.*, vol. 6, no. 7, pp. 1244–1261, Jul. 2019.

[30] T. Zhao, F. Li, and K. J. Marfurt, "Seismic attribute selection for unsupervised seismic facies analysis using user-guided data-adaptive weights," *Geophysics*, vol. 83, no. 2, pp. O31–O44, Mar. 2018.

[31] B. Zhang, T. Zhao, X. Jin, and K. J. Marfurt, "Brittleness evaluation of resource plays by integrating petrophysical and seismic data analysis," *Interpretation*, vol. 3, no. 2, pp. T81–T92, May 2015.

[32] A. Corradi, P. Ruffo, A. Corrao, and C. Visentin, "3D hydrocarbon migration by percolation technique in an alternate sand–shale environment described by a seismic facies classified volume," *Mar. Petroleum Geol.*, vol. 26, no. 4, pp. 495–503, Apr. 2009.

[33] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.

[34] Y. Chen, "Automatic microseismic event picking via unsupervised machine learning," *Geophys. J. Int.*, vol. 212, no. 1, pp. 88–102, Jan. 2018.

[35] Y. Chen, "Fast waveform detection for microseismic imaging using unsupervised machine learning," *Geophys. J. Int.*, vol. 215, no. 2, pp. 1185–1199, Nov. 2018.

[36] J. Liu, X. Dai, L. Gan, L. Liu, and W. Lu, "Supervised seismic facies analysis based on image segmentation," *Geophysics*, vol. 83, no. 2, pp. O25–O30, Mar. 2018.

[37] F. Qian, M. Yin, X.-Y. Liu, Y.-J. Wang, C. Lu, and G.-M. Hu, "Unsupervised seismic facies analysis via deep convolutional autoencoders," *Geophysics*, vol. 83, no. 3, pp. A39–A43, May 2018.

[38] H. Di, Z. Wang, and G. AlRegib, "Why using CNN for seismic interpretation? An investigation," in *SEG Technical Program Expanded Abstracts*. Tulsa, OK, USA: SEG, 2018, pp. 2216–2220.

[39] S. Qu, Z. Guan, E. Verschuur, and Y. Chen, "Automatic high-resolution microseismic event detection via supervised machine learning," *Geophys. J. Int.*, vol. 218, no. 3, pp. 2106–2121, Sep. 2019.

[40] S. van der Walt *et al.*, "Scikit-image: Image processing in Python," *PeerJ*, vol. 2, p. e453, Jun. 2014.

[41] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," 2017, *arXiv:1712.04621*. [Online]. Available: http://arxiv.org/abs/1712.04621

[42] R. Takahashi, T. Matsubara, and K. Uehara, "Data augmentation using random image cropping and patching for deep CNNs," 2018, *arXiv:1811.09030*. [Online]. Available: http://arxiv.org/abs/1811.09030

[43] S. Chopra and K. J. Marfurt, "Spectral decomposition and spectral balancing of seismic data," *Lead. Edge*, vol. 35, no. 2, pp. 176–179, Feb. 2016.

[44] F. Li, Y. Li, W. Lu, Y. Zhang, and X. Zheng, "Hydrocarbon detection for cavern carbonate reservoir using low-and-high-frequency anomalies in spectral decomposition," in *SEG Technical Program Expanded Abstracts*. Tulsa, OK, USA: SEG, 2012, pp. 1–5.

[45] Z.-L. Huang, J. Zhang, T.-H. Zhao, and Y. Sun, "Synchrosqueezing S-transform and its application in seismic spectral decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 2, pp. 817–825, Feb. 2016.

[46] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.

[47] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* New York, NY, USA: Springer, 2015, pp. 234–241.

[48] H. Wu, B. Zhang, F. Li, and N. Liu, "Semiautomatic first-arrival picking of microseismic events by using the pixel-wise convolutional image segmentation method," *Geophysics*, vol. 84, no. 3, pp. V143–V155, May 2019.

[49] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," 2017, *arXiv:1706.05587*. [Online]. Available: http://arxiv.org/abs/1706.05587

[50] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," 2015, *arXiv:1511.07122*. [Online]. Available: http://arxiv.org/abs/1511.07122

[51] S. Jetley, N. A. Lord, N. Lee, and P. H. S. Torr, "Learn to pay attention," 2018, *arXiv:1804.02391*. [Online]. Available: http://arxiv.org/abs/1804.02391

[52] J. Schlemper *et al.*, "Attention-gated networks for improving ultrasound scan plane detection," 2018.

[53] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980

[54] F. Li, B. Zhang, R. Zhai, H. Zhou, and K. J. Marfurt, "Depositional sequence characterization based on seismic variational mode decomposition," *Interpretation*, vol. 5, no. 2, pp. SE97–SE106, May 2017.

[55] A. Paszke *et al.*, "Automatic differentiation in PyTorch," in *Proc. NIPS-W*, 2017, pp. 1–4.

[56] H. Lee *et al.*, "Pixel-level deep segmentation: Artificial intelligence quantifies muscle on computed tomography for body morphometric analysis," *J. Digit. Imag.*, vol. 30, no. 4, pp. 487–498, Aug. 2017.

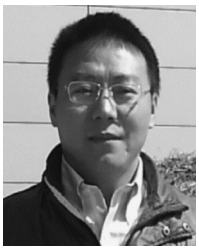
[57] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," 2015, *arXiv:1511.00561*. [Online]. Available: http://arxiv.org/abs/1511.00561

[58] Y. He, C. Kerans, H. Zeng, V. Janson, and S. Z. Scott, "Improving three-dimensional high-order seismic-stratigraphic interpretation for reservoir model construction: An example of geostatistical and seismic forward modeling of Permian San Andres shelf–Grayburg platform mixed clastic–carbonate strata," *AAPG Bull.*, vol. 103, no. 8, pp. 1839–1887, Aug. 2019.

Fangyu Li received the bachelor's degree in electrical engineering from Beihang University, Beijing, China, in 2009, the master's degree in electrical engineering from Tsinghua University, Beijing, in 2013, and the Ph.D. degree in geophysics from The University of Oklahoma, Norman, OK, USA, in 2017.

He is an Assistant Professor with the Department of Electrical and Computer Engineering, Kennesaw State University (KSU), Marietta, GA, USA. Before joining KSU, he was a Post-Doctoral Fellow with the College of Engineering, University of Georgia, Athens, GA, USA. His research interests include seismic interpretation, subsurface imaging, signal processing, seismic imaging, machine learning, deep learning, distributed computing, Internet of Things (IoT), and cyberphysical systems (CPS).

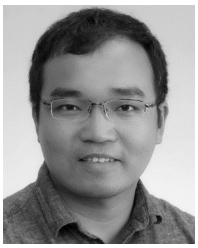
Dr. Li received the J. Clarence Karcher Award from the Society of Exploration Geophysicists (SEG) in 2020.



Huailai Zhou received the Ph.D. degree in Earth exploration and information techniques from the Chengdu University of Technology, Chengdu, China, in 2009.

He completed his post-doctoral research at the Institute of Sedimentary Geology, Chengdu University of Technology, from 2010 to 2012. He was sponsored by the China Scholarship Council to work as a Post-Doctoral Research Faculty in AASPI at The University of Oklahoma, Norman, OK, USA, from October 2013 to October 2014. He works at the Chengdu University of Technology as a Professor. His research work is mainly focused on seismic data processing methods, seismic modeling and imaging, techniques of improving seismic resolution, and seismic attribute analysis and inversion methods.

Zengyan Wang received the M.S. degree in statistics from the University of Georgia, Athens, GA, USA, where she is pursuing the Ph.D. degree with a focus on machine learning and computer vision, specifically with the development of the deep learning model for semantic segmentation in imaging spectroscopy.



Xinming Wu received the Ph.D. degree in geophysics from the Colorado School of Mines, Golden, CO, USA, in 2016.

He was a member of the Center for Wave Phenomena, Colorado School of Mines. He was a Post-Doctoral Fellow with the Bureau of Economic Geology, The University of Texas at Austin, Austin, TX, USA. He is a Professor with the School of Earth and Space Sciences, University of Science and Technology of China (USTC), Hefei, China. His research interests include image processing, machine learning, 3-D seismic interpretation, subsurface modeling, and geophysical inversion.

Dr. Wu was a recipient of the Best Paper Award in Geophysics in 2016, the Best Student Poster Paper Award at the 2017 SEG Annual Convention, and the Honorable Mention Award for Best Paper at the 2018 SEG Annual Convention. He was also honored as a recipient of the 2020 J. Clarence Karcher Award from the Society of Exploration Geophysicists (SEG).