energies

MDPI

Article

Optimization of Critical Parameters of Deep Learning for Electrical Resistivity Tomography to Identifying Hydrate

Yang Liu [1,2], Changchun Zou [1,2,*], Qiang Chen [3,4,*], Jinhuan Zhao [3,4] and Caowei Wu [1,2]

[1] School of Geophysics and Information Technology, China University of Geosciences, Beijing 100083, China; lyang@cugb.edu.cn (Y.L.); wcw@cugb.edu.cn (C.W.)
[2] National Engineering Research Center of Offshore Oil and Gas Exploration, Beijing 100028, China
[3] Key Laboratory of Gas Hydrate, Ministry of Natural Resources, Qingdao Institute of Marine Geology, Qingdao 266071, China; zhaojinhuan99@yeah.net
[4] Laboratory for Marine Mineral Resources, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao 266071, China
* Correspondence: zoucc@cugb.edu.cn (C.Z.); chenqiang_hds@126.com (Q.C.)

Abstract: As a new energy source, gas hydrates have attracted worldwide attention, but their exploration and development face enormous challenges. Thus, it has become increasingly crucial to identify hydrate distribution accurately. Electrical resistivity tomography (ERT) can be used to detect the distribution of hydrate deposits. An ERT inversion network (ERTInvNet) based on a deep neural network (DNN) is proposed, with strong learning and memory capabilities to solve the ERT nonlinear inversion problem. 160,000 samples about hydrate distribution are generated by numerical simulation, of which 10% are used for testing. The impact of different deep learning parameters (such as loss function, activation function, and optimizer) on the performance of ERT inversion is investigated to obtain a more accurate hydrate distribution. When the Logcosh loss function is enabled in ERTInvNet, the average correlation coefficient (CC) and relative error (RE) of all samples in the test sets are 0.9511 and 0.1098. The results generated by Logcosh are better than MSE, MAE, and Huber. ERTInvNet with Selu activation function can better learn the nonlinear relationship between voltage and resistivity. Its average CC and RE of all samples in the test set are 0.9449 and 0.2301, the best choices for Relu, Selu, Leaky_Relu, and Softplus. Compared with Adadelta, Adagrad, and Aadmax, Adam has the best performance in ERTInvNet with the optimizer. Its average CC and RE of all samples in the test set are 0.9449 and 0.2301, respectively. By optimizing the critical parameters of deep learning, the accuracy of ERT in identifying hydrate distribution is improved.

Keywords: deep learning; electrical resistivity tomography; hydrate distribution; numerical simulation; optimization

check for updates

Citation: Liu, Y.; Zou, C.; Chen, Q.; Zhao, J.; Wu, C. Optimization of Critical Parameters of Deep Learning for Electrical Resistivity Tomography to Identifying Hydrate. *Energies* **2022**, *15*, 4765. https://doi.org/10.3390/en15134765

Academic Editors: Yonghai Gao, Litao Chen and Xiaohui Sun

Received: 23 May 2022
Accepted: 26 June 2022
Published: 29 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The exploration and development of new energy has been very popular in recent times. As a new energy in the 21st century, the efficient and safe exploitation of hydrate is facing a tremendous challenge [1]. Submarine landslides, stratigraphic collapse, and methane leakage caused by hydrate development will cause significant harm to human safety and the environment [2–5]. Hence, how to accurately detect the dynamic distribution of hydrate needs to be solved urgently. Recently, electrical resistivity tomography (ERT) has become a promising method for the identification of hydrate because of its high efficiency, low cost, and simple operation [6–8].

ERT is used to reconstruct the resistivity distribution of the target area according to array geometry and the measured voltage data. The ERT inverse problem has complex nonlinearity and serious ill-posedness [9]. In most inversion algorithms, linearization and regularization techniques are used to solve nonlinear ill-posed problems. [10,11]. There are

two primary types of traditional ERT inversion methods: iterative and noniterative. Noniterative methods include linear back-projection method [12], singular value decomposition method, and Tikhonov regularization method [13,14], among others. Iterative methods include the conjugate gradient method [15,16], the Gauss–Newton method [17], and the Landweber method [18], among others. Using these methods can create higher-quality inversion results with the simple geological model. However, it is difficult for a complex geological model to acquire better inversion results. It is necessary to improve the accuracy of ERT inversion results to obtain a more refined underground structure.

With the rapid development of artificial intelligence, deep learning has attracted more and more attention. Based on the breakthrough success of deep learning in image classification [19], image segmentation [20], face recognition [21], and speech recognition [22], this technology has been introduced into image denoise [23], medical imaging [24], and geophysical inversion [25,26]. With powerful learning memory and nonlinear approximation, deep learning exhibits superiority in geophysical inversion [27]. Currently, many network frameworks with remarkable results in deep learning have been used for ERT inversion, including deep neural networks (DNN) and convolutional neural networks (CNN) [28–31]. When constructing a deep learning network, the choice of deep learning parameters affects the performance of the network and the accuracy of the ERT inversion results. However, the design and selection of deep learning parameters are not systematically analyzed and compared. Therefore, we must design an optimized deep learning network to realize the intelligent detection of hydrate distribution.

In this study, the DNN network framework in deep learning is applied to solve the ERT inverse problem. Furthermore, the impact of various deep learning parameters (such as activation function, loss function, number of iterations, and learning rate) on the accuracy of ERT inversion results is discussed and analyzed. The finite element method is performed to construct a deep learning sample dataset about hydrate distribution. Each sample data contains a voltage measurement vector and a resistivity vector, and the sample data set is used for network training. Different deep learning parameters are adjusted to obtain different network models. Then, the simulation data is used to compare and analyze the inversion results with different parameters. Finally, we give the optimal choice of parameters. The research results help improve the accuracy of the ERT inversion result and hydrate identification.

2. Methods

2.1. Model of ERT

Figure 1a portrays the ERT physical simulation model involved in this paper. A square study area has an edge length of 25 cm, in which different sizes and numbers of hydrate occurrence areas exist. The electrode array consists of 48 electrodes in the model. Every 24 electrodes are a group, the distance between groups is 10 cm, and the electrode spacing is 0.8 cm. The four-electrodes method conducts ERT data acquisition, containing two current injection electrodes (A, B) and two voltage measurement electrodes (M, N). Two electrodes in two groups are selected as current injection electrodes for current excitation. Two electrodes in the remaining electrodes are used as voltage measurement electrodes for voltage data collection. There are 48 current excitations. 45 voltage data for per excitations, which results in 2160 voltage data collected.

According to the ERT physical simulation model, the corresponding mesh of the forward and inverse problem is constructed. A triangular mesh is used to divide the study area into 5080 meshes, as portrayed in Figure 1b. Each mesh can be set with different resistivity. Different resistivity values are assigned to different hydrate occurrence areas to simulate the heterogeneity of hydrate in nature.

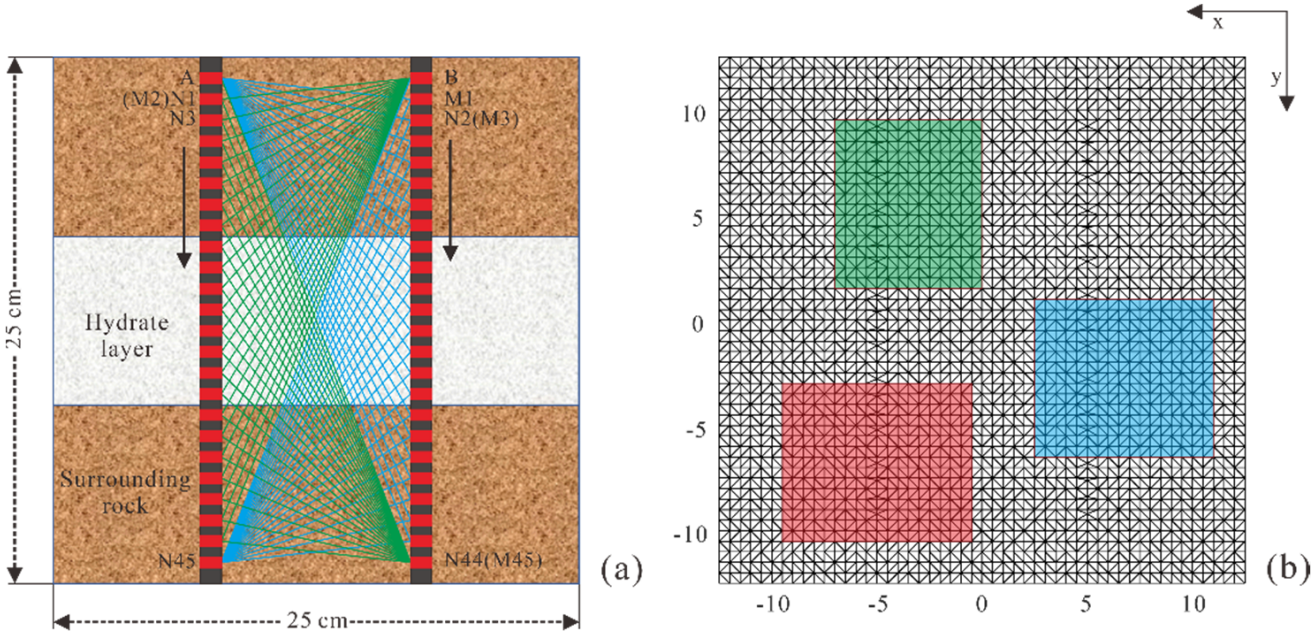


Figure 1. Model of ERT and its mesh. (**a**) The physical simulation model of ERT, where the red rectangles represent electrodes. A and B are current injection electrodes, and M and N are voltage measurement electrodes; (**b**) the forward and inverse problem mesh, where different colored grids represent different resistivity values.

2.2. Finite Element forward Method

ERT belongs to the stable current field problem. The finite element method is an effective one used to calculate the potential distribution. According to Ohm's law and the law of conservation of electric charge, the potential equation of point source is derived:

$$\frac{\partial}{\partial x}\left(\rho\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial y}\left(\rho\frac{\partial u}{\partial y}\right)=-I\delta(x_A)\delta(y_A);\ (x,y),(x_A,y_A)\in\Omega \tag{1}$$

where ρ denotes resistivity; the coordinate (x_A, y_A) is the location of the point current source; $u(x, y)$ is the potential function; I is current intensity; and Ω is the study area.

The potential function generally satisfies three types of boundary conditions: Dirichlet boundary condition (2), Neumann boundary condition (3), and Mixed boundary condition (4). The specific mathematical expression is as follows:

$$u|_{\partial\Omega}=\varphi \tag{2}$$

$$\left.\frac{\partial u}{\partial n}\right|_{\partial\Omega}=\varphi \tag{3}$$

$$\left.\left(\frac{\partial u}{\partial n}+\alpha u\right)\right|_{\partial\Omega}=\varphi \tag{4}$$

where u is the potential function of the boundary; n is the normal unit vector for the boundary $\partial\Omega$; φ is the functional value of the boundary; α is the boundary operator. The Neumann boundary condition is selected through a comprehensive analysis. The boundary value problem for the two-dimensional stable current field is obtained by combining the potential Equation (1) and the boundary condition (3).

The analytic solution can be obtained for a simple boundary value problem, while complex problems need to be solved by numerical simulation methods. Commonly, numerical simulation methods include finite element method, boundary element method, and finite volume method. The finite element method is adopted to obtain a high-precision solution.

The following linear equations are established by mesh generation, linear interpolation, element analysis, and global synthesis:

$$KU = P \tag{5}$$

where K is the stiffness matrix, which is determined by the dissection of the study area and the model parameters; U is the column vector composed of the potential of each node; P is the load vector, which is determined by the location of the point source and the current. The potential value of each node can be obtained by solving the linear equations.

2.3. Deep Learning Inversion Method

Deep learning provides a new perspective to solving the ERT inverse problem. The advantages of deep learning are manifested in two aspects: (1) deep learning is to learn the mapping law between relevant features and resistivity models from many data, which greatly reduces the interference of subjective factors; (2) deep learning is useful to establish a complex nonlinear mapping relationship between measurement data and resistivity model through a combination of multi-layer networks, which is suitable for the ERT inverse problem. A network is constructed in Tensorflow to achieve deep learning inversion. The implementation of deep learning is divided into the following three steps: datasets establishment, network structure design, and network training. In this study, the Python programming language (Python 3.8.0) was used for the deep learning inversion method.

2.3.1. Datasets Establishment

Generally, the dataset should be large enough to ensure the diversity model for deep learning inversion. Different hydrate resistivity models are established based on the constructed mesh of the forward and inverse problems. The finite element method calculates the forward response data (voltage data) according to the hydrate resistivity model and data acquisition method. The data acquisition method can be located in Section 2.1 Four types of hydrate distribution are established Figure 2. By setting the resistivity value of each mesh, the resistivity model with hydrate occurrence areas of different sizes, locations, and numbers can be obtained. The existence of hydrate leads to an increase in resistivity. The higher the hydrate saturation is, the greater the resistivity. We set the background resistivity of 1 Ω·m according to the resistivity of seabed sediments. With hydrate saturation increasing, the resistivity of hydrate occurrence area increases. The resistivity of the hydrate occurrence area is set to 5, 10, 15 Ω·m based on field data and laboratory data. 160,000 sets of data are simulated by carrying out the finite element method. Each sample data set includes the resistivity distribution vector ρ_n (n = 1, 2, 3, . . . , N, where N is the total number of samples) with 5080 elements and the voltage vector V_n (n = 1, 2, 3, . . . , N, where N is the total number of samples) with 2160 elements obtained by forward calculation Figure 2. We divide the dataset into the training set, validation set, and test set according to the ratio of 8:1:1, (Figure 2). Network training uses most of the data. Because of the dynamic change during the hydrate detecting process, more hydrate resistivity types will be added to improve the simulation data set in the future.

2.3.2. Network Structure

The primary network frameworks for deep learning include CNN, DNN, and Recurrent Neural Networks (RNN), and the rest of the network framework is derived from these. DNN with enough hidden layers can approximate any continuous function. The DNN network framework is composed of three parts: (a) the input layer, (b) two or more hidden layers, and (c) the output layer. Based on the DNN network framework in deep learning, we design a network structure for ERT inversion, as portrayed in Figure 3.

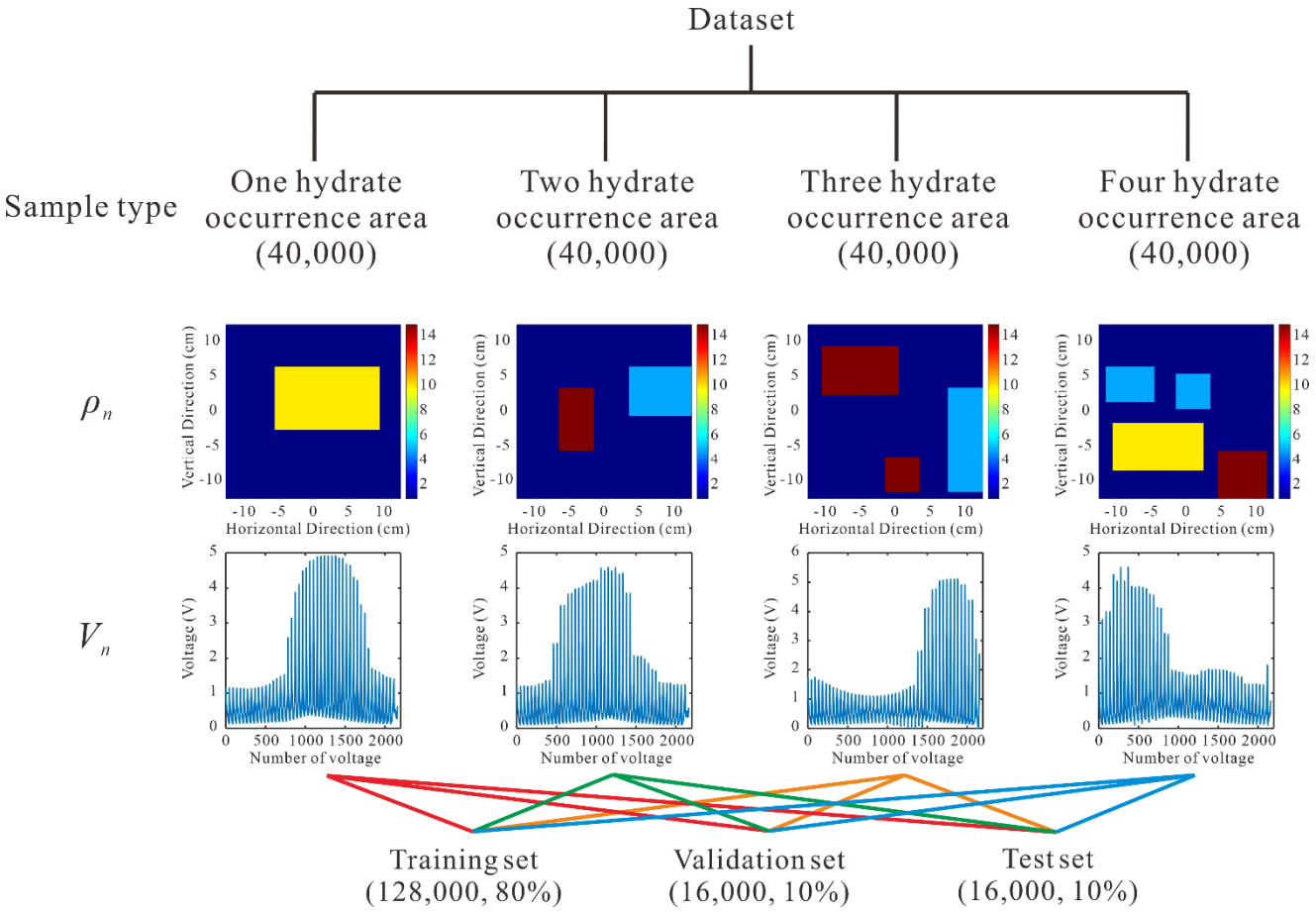


Figure 2. Information about the composition of the data set. The dataset includes four types of hydrate distribution. There are 40,000 samples of each type. Each sample data set includes the resistivity distribution vector and the voltage vector. The dataset is divided into training set, validation, and test set according to the ratio of 8:1:1.

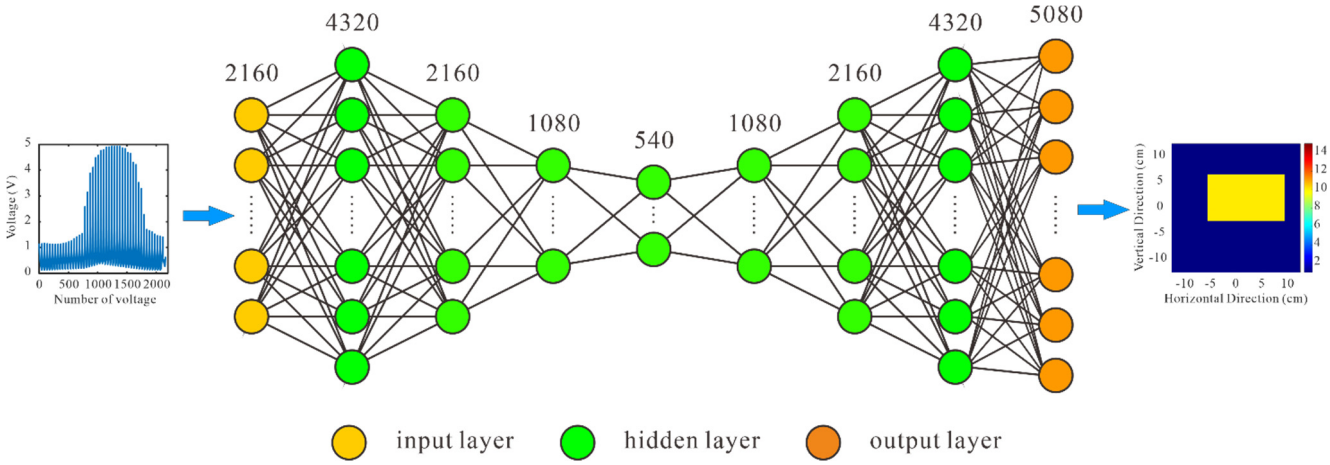


Figure 3. A network structure for ERT inversion.

The ERT inversion network (ERTInvNet) structure consists of 9 layers, which contain an input layer, seven hidden layers, and an output layer. The input layer contains 2160 neurons, which is the voltage vector (V) acquired by the finite element method according to the resistivity distribution and electrode layout. The numbers of neurons in the hidden layer are 4320, 2160, 1080, 540, 1080, 2160, and 4320 from left to right. The output layer

contains 5080 neuron elements, which is the resistivity data (ρ). Neurons between layers are connected in a fully connected form.

2.3.3. Network Training

Network training is essential for deep learning. The same network structure trained using different training parameters will generate differently. The choice of training parameters, such as active function, loss function, optimizer, minimum batch size, and epoch, will affect the accuracy, generalization, and convergence of ERTInvNet. The influence of the first three parameters on training results will be analyzed and discussed. For a detailed description of these parameters, see Section 3. During the network training process, the training set is randomly shuffled, and mini-batch sizes with 50 are input in-batch for training in the network structure, portrayed in Figure 3. Each sample traversal in the data set is called a "training epoch." One hundred epochs are performed in the training process. Meanwhile, we can observe the performance curve (loss curve and accuracy curve) of the validation and training sets. Finally, the ERTInvNet model is obtained.

3. Deep Learning Parameters

3.1. Loss Function

The loss function is one of the critical elements in deep learning. It is used to measure the difference degree between the predicted resistivity value and the real resistivity value in this paper, and to minimize the difference degree through cycling training continuously. The smaller the loss value is, the better the performance of the network model will be. For different problems, the choice of the loss function is different. Commonly used loss function includes mean square error (MSE) function, mean absolute error (MAE) function, Logcosh function, and Huber function. The expressions are as follows:

$$\text{MSE} : l = \frac{1}{n}\sum_{i=1}^{n}(\rho_i - \overline{\rho_i})^2 \tag{6}$$

$$\text{MAE} : l = \frac{1}{n}\sum_{i=1}^{n}|\rho_i - \overline{\rho_i}| \tag{7}$$

$$\text{Logcosh} : l = \frac{1}{n}\sum_{i=1}^{n} log(\cosh(\rho_i - \overline{\rho_i}))^2 \tag{8}$$

$$\text{Huber} : l = \begin{cases} \frac{1}{2n}\sum\limits_{i=1}^{n}(\rho_i - \overline{\rho_i})^2, & |\rho_i - \overline{\rho_i}| < \delta \\ \delta\frac{1}{n}\sum\limits_{i=1}^{n}|\rho_i - \overline{\rho_i}| - \frac{1}{2}\delta^2, & |\rho_i - \overline{\rho_i}| \geq \delta \end{cases} \tag{9}$$

where l is the loss value, ρ_i is the actual resistivity value, and $\overline{\rho_i}$ is the resistivity value of deep learning inversion, δ is set to 1 in this paper.

3.2. Activation Function

The selection of activation function is a significant part of building a neural network. The activation function is a step function bound by a specific threshold, which is activated once the input data is greater than the designated threshold. The activation function is equivalent to a nonlinear function, strengthening the network's learning ability and nonlinear approximation ability. Therefore, deep learning has outstanding performance in ERT inversion due to the existence of activation functions. Common activation functions include Relu, Selu, Leaky_Relu, and Softplus functions, which each have advantages and disadvantages. The mathematical form and geometric image are portrayed in Figure 4. A few common activation functions for network training were attempted to select the activation function suitable for ERTInvNet.

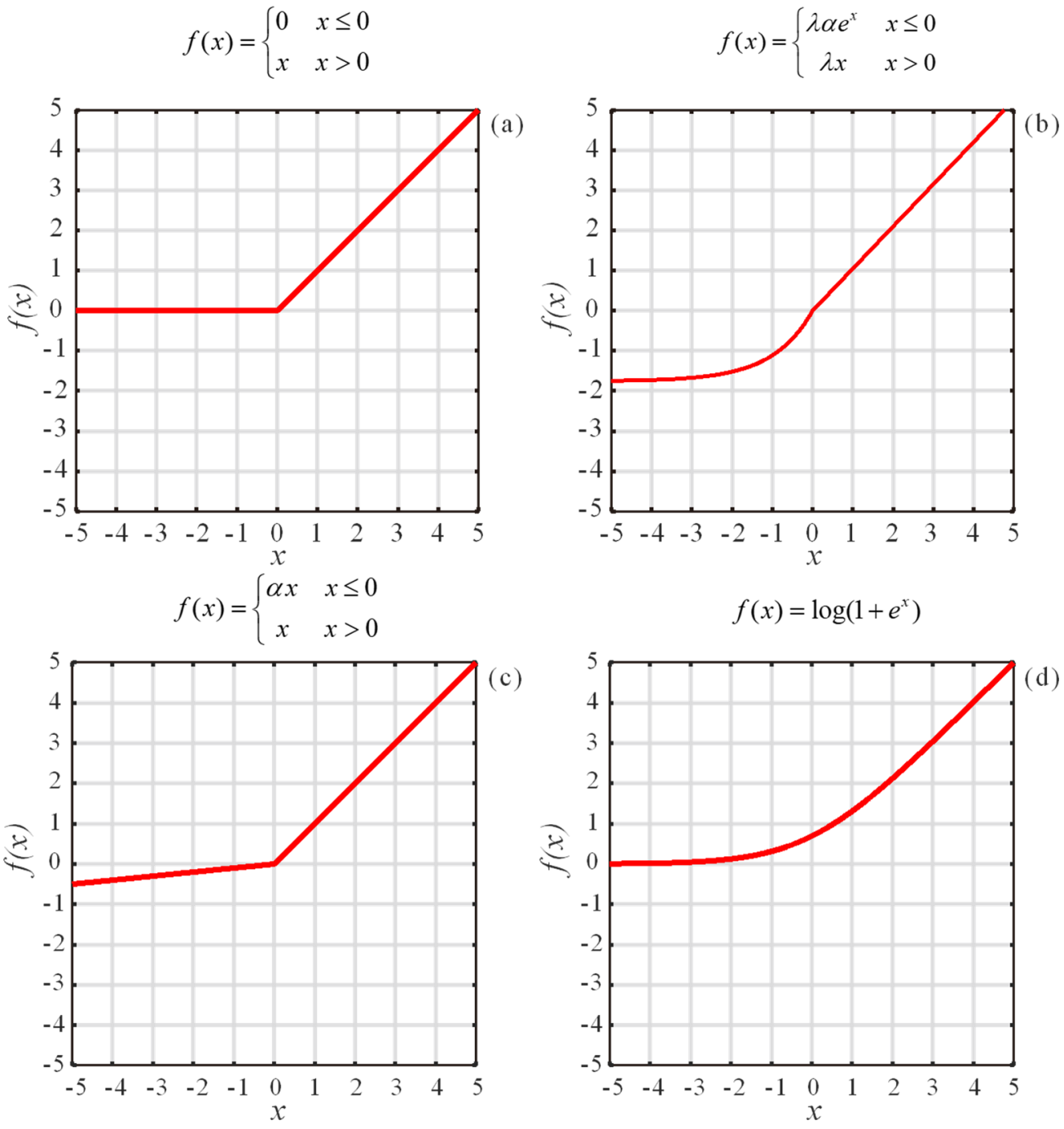


Figure 4. Geometric images of the four activation functions. (**a**) Relu function; (**b**) Selu function; (**c**) Leaky_Relu function; (**d**) Softplus function. Where λ is 1.05070098, and α is 1.67326324 and 0.1 in (**b**) and (**c**), respectively.

3.3. Optimizer

The optimization of deep learning minimizes the loss function, and the selection of its optimization algorithm directly affects the output results. Robust optimization algorithm selection is an indispensable step to improving deep learning inversion accuracy. Adaptive learning rate optimization algorithms have been developed rapidly. Since 2011, a variety of optimizers have appeared one after another. Adadelta, Adaptive Gradient (Adagrad), Adamax, and Adaptive Moment Estimation (Adam) are popular. It is necessary to select the best optimizer to accelerate convergence and correct learning. The parameters for each

optimizer are depicted in Table 1. Because Adagrad and Adadelta tend to benefit from higher initial learning rate values than other optimizers [32], An initial learning rate of 0.1 is used.

Table 1. Parameters for the different optimization algorithms.

Parameters	Adam	Adagrad	Adadelta	Adamax
Initial learning rate	1×10^{-4}	0.1	0.1	1×10^{-4}
decay	-	0.96	0.96	-
beta_1	0.9	-	-	0.9
beta_2	0.999	-	-	0.999
Epsilon	1×10^{-7}	1×10^{-7}	1×10^{-7}	1×10^{-7}
Amsgrad	False	-	-	-
Initial accumulator value	-	0.5	-	-

4. Results and Discussion

To obtain the best ERTInvNet model, different loss functions, activation functions, and optimizers are used in ERTInvNet based on DNN. A large number of simulation experiments are carried out in the test set. The relative error (RE) and correlation coefficient (CC) are used during the evaluation of the ERT inversion results.

$$\mathrm{RE} = \frac{||\hat{\rho} - \rho||_2}{||\rho||_2} \tag{10}$$

$$\mathrm{CC} = \frac{\sum_{i=1}^{N}(\hat{\rho}_i - \overline{\hat{\rho}})(\rho_i - \overline{\rho})}{\sqrt{\sum_{i=1}^{N}(\hat{\rho}_i - \overline{\hat{\rho}})^2 \sum_{i=1}^{N}(\rho_i - \overline{\rho})^2}} \tag{11}$$

where $\overline{\hat{\rho}}$ is the average of the inversion resistivity vector $\hat{\rho}$. The actual resistivity vector is the goal of the inversion network learning, and $\overline{\rho}$ is the average of ρ.

4.1. Evaluation of Inversion Network Models with Different Loss Functions

Four loss functions (MSE, MAE, Logcosh, Huber) are used in training ERTInvNet to achieve a smaller loss value. The activation function is equivalent to a nonlinear function. We need to fix the activation function of each layer and optimizer in order to study the effect of the loss function on the network performance. Selu activation function is enabled at each hidden layer and Linear activation function for the output layer. Adam optimizer is selected to optimize the loss value. The parameter settings of the Adam optimizer are displayed in Table 1.

The lower loss value and the higher accuracy value may indicate excellent performance of the network model. The performance curve on the training sets and the validation sets during the training course with different loss functions are provided in Figure 5. With the increase of epoch, the loss value decreases, and the accuracy value increases gradually. In Figure 5a, it is observable that the Logcosh loss curve during the training and validation process converges the fastest, and the MAE loss curve converges the slowest. After about 50 epochs, the Logcosh loss value in the training process falls under 0.02. Although Huber loss has also achieved a better loss value, the Logcosh loss is smaller than Huber loss. MSE loss in the validation set is about 0.04 higher than in the training set, while MAE, Logcosh, and Huber loss values in the validation set are about 0.015 greater than in the training set. Similarly, in Figure 5b, Logcosh accuracy and Huber accuracy nearly overlap in the training and validation set, reaching as high as 96.3%, 96.3%, 92.9%, and 92.8%, respectively. Through analyzing the loss curve and accuracy curve on the training set and the validation set, it can be preliminarily judged that using the Logcosh loss function can effectively improve the performance of the ERTInvNet model.

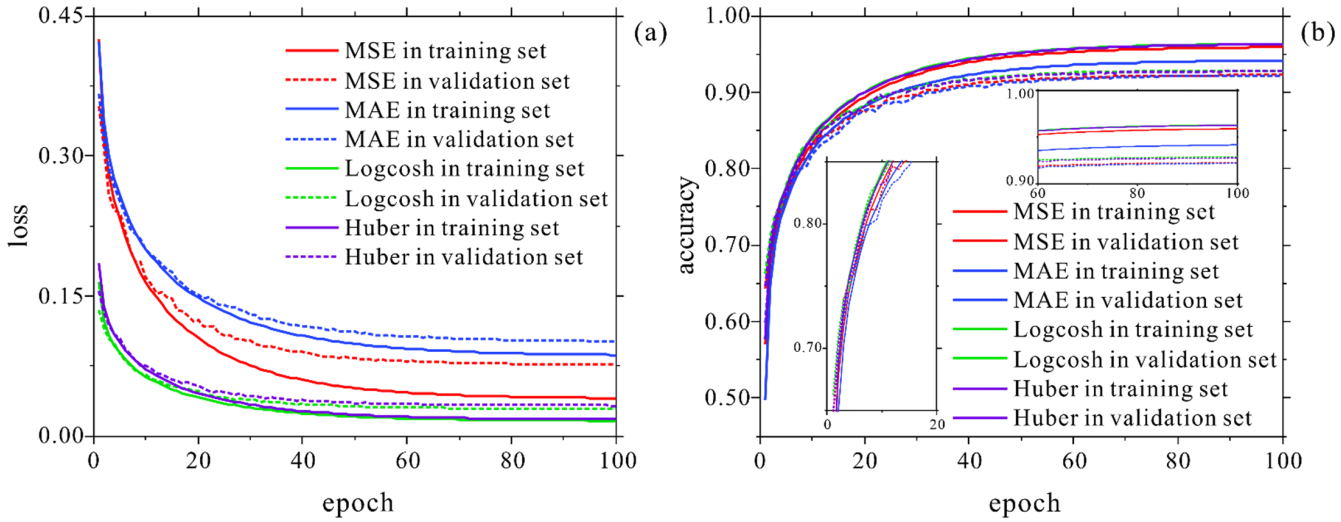


Figure 5. The model's performance curve with different loss functions. (**a**) loss curve, (**b**) accuracy curve.

To investigate the effect of the ERTInvNet model using different loss functions, some results of the four typical hydrate distributions were randomly selected. The inversion results on the test set using the ERTInvNet model with different loss functions is portrayed in Figure 6. With the increase in the number of hydrate occurrence areas, the boundary gradually blurs. The inversion results with Logcosh can still more accurately retrieve the position, shape, and resistivity value of the hydrate occurrence area, which indicates that Logcosh is the best choice.

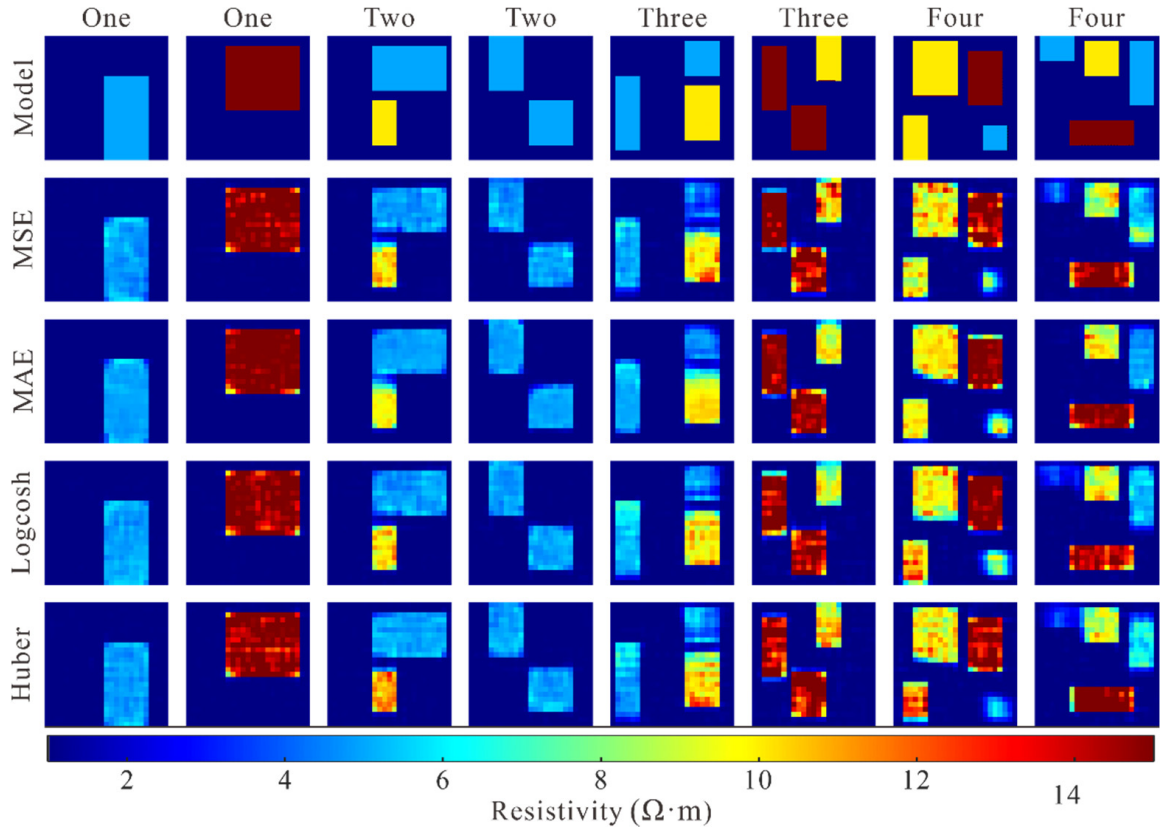


Figure 6. Model and their inversion results with using different loss functions.

The CC and RE of all samples in the test set are calculated to quantitatively analyze the quality of inversion results using different loss functions. The average RE and CC and the standard deviations of RE and CC of all samples in the test set with different loss functions are provided in Table 2. The average CC of Logcosh is 0.9511, which is higher than that of MSE, MAE, and Huber. The standard deviation of the CC of Logcosh is better than MSE and MAE, which is 0.0001 higher than Huber. The average RE of Logcosh is 0.2142, which is lower than that of MSE, MAE, and Huber. The standard deviations of RE of MSE, Logcosh, and Huber are lower than 0.11, and the standard deviation of the CC of MAE is the worst. The quantitative analysis results indicate that the ERTInvNet model with Logcosh has good performance.

Table 2. The RE and CC of all samples in the test set with different loss functions.

Loss Function		MSE	MAE	Logcosh	Huber
RE	AVE	0.2301	0.2517	0.2142	0.2155
	SD	0.1096	0.1168	0.1098	0.1086
CC	AVE	0.9449	0.9490	0.9511	0.9507
	SD	0.0502	0.0521	0.0479	0.0478

AVE = average, SD = standard deviations.

The average CC and RE of different samples in the test set using different loss functions are portrayed in Figure 7. With the gradual increase in the number of hydrate occurrence areas, the average RE and CC of four loss functions become progressively worse. In the same number of hydrate occurrence areas, the average RE and CC of Logcosh, Huber, and MAE have no apparent difference. When the number of hydrate occurrence areas is more than or equal to two, it is evident that the average RE and CC of MSE are worse than other loss functions.

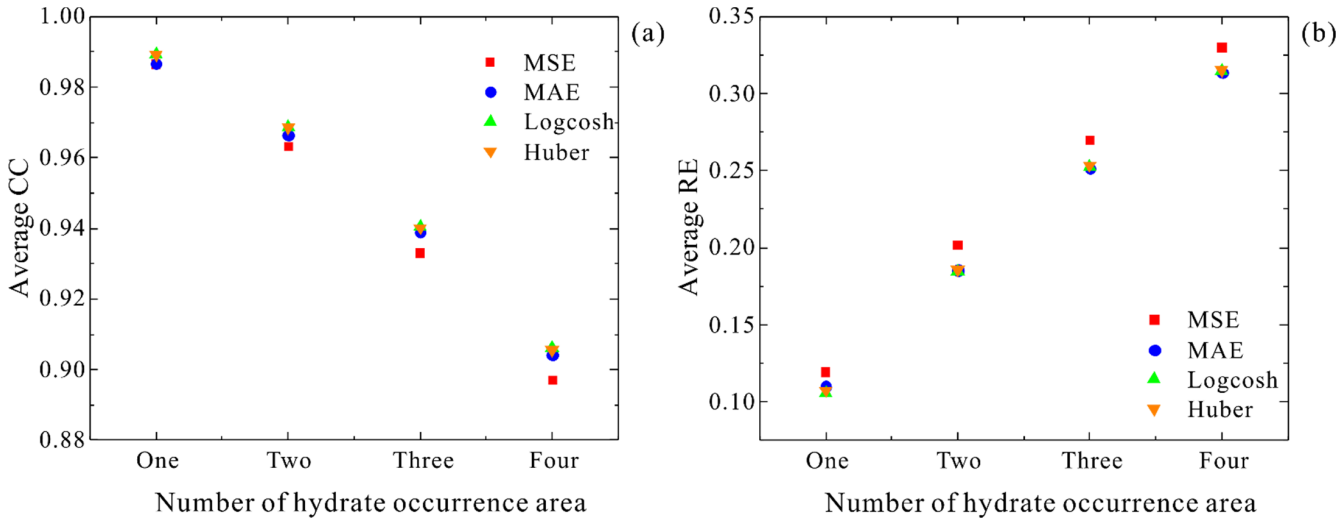


Figure 7. The average RE and CC of different samples in the test set for different loss functions. (**a**) average CC, (**b**) average RE.

4.2. Evaluation of Inversion Network Models with Different Activation Functions

Four activation functions (Selu, Relu, Leaky_Relu, Softplus) are applied to the proposed ERTInvNet in training to accurately map the nonlinear relationship between voltage and resistivity. Four activation functions are enabled at each hidden layer and the Linear activation function for the output layer. MSE is selected as the loss function, and the Adam optimizer optimizes the loss value. We set the Adam optimizer according to the parameters in Table 1.

The performance curve of the training sets and the validation sets during the training course using ERTInvNet with different activation functions are portrayed in Figure 8. With the increase of epoch, the loss value steadily decreases, and the accuracy value gradually increases without significant fluctuations. It is investigated that the loss curve with Selu during the training and validation process converges the fastest (Figure 8a). After about 50 epochs, the loss value with Selu in the training process falls under 0.05. Although the Softplus loss value slowly decreases in the training set, Softplus loss gradually increases in the validation set. This phenomenon is called overfitting. Selu, Leaky_Relu, Relu, and Softplus in validation set loss value increase sequentially when the performance curve is stable, 7.66%, 9.71%, 13.10%, and 24.90%, respectively, which demonstrates that ERTInvNet using Selu is more advantageous than other activation functions. In Figure 8b, it can be observed that Softplus accuracy is the lowest. The difference between the Selu, Leaky_Relu, and Relu accuracy curves is not apparent. Through the loss curve and accuracy curve are provided on the training set and the validation set, using the Softplus activation function cannot promote the performance of the ERTInvNet model.

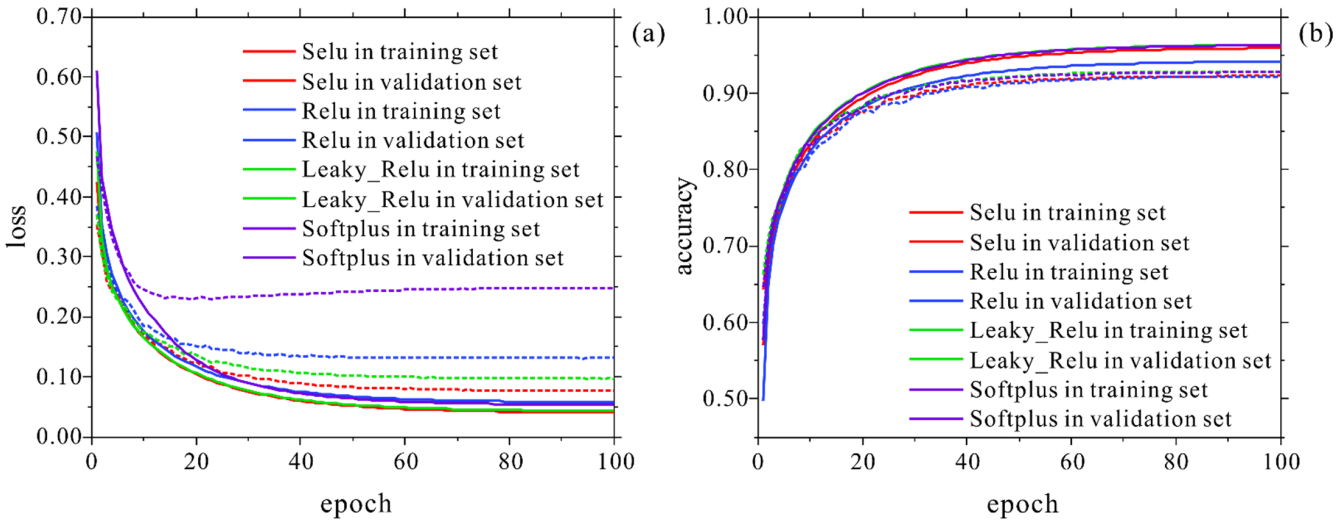


Figure 8. The model's performance curve with different activation functions. (**a**) loss curve, (**b**) accuracy curve.

Some results of the four typical hydrate distributions are randomly selected to research the effect of the network model using different activation functions. The inversion results on the test set using the network model with different activation functions is portrayed in Figure 9. With the increase in the number of hydrate occurrence areas, the boundary of the inversion results in blurs. When the number of hydrate occurrence areas is greater than 2, inversion results with Leaky_Relu, Relu, and Softplus are not ideal and cannot reflect the number of hydrate occurrence areas. The inversion results with Selu is better than with other activation functions.

After calculating the CC and RE of all samples in the test set, the quality of inversion results using different activation functions was discussed. The average RE and CC and the standard deviations of RE and CC of all samples in the test set with different activation functions are portrayed in Table 3. The average RE of Selu is 0.2301. Compared with Relu, Leaky_Relu, and Softplus, its average RE decreases by 0.0765, 0.0324, and 0.1938, respectively, indicating that Selu can improve inversion accuracy. In addition, the average RE of Selu is smaller than that of the other activation function. The average CC of Selu is 0.9449, which is much higher than those of Relu, Leaky_Relu, and Softplus. The standard deviations of RE and CC of Selu are better than the other three. The quantitative analysis results indicate that the ERTInvNet model with Selu has good performance.

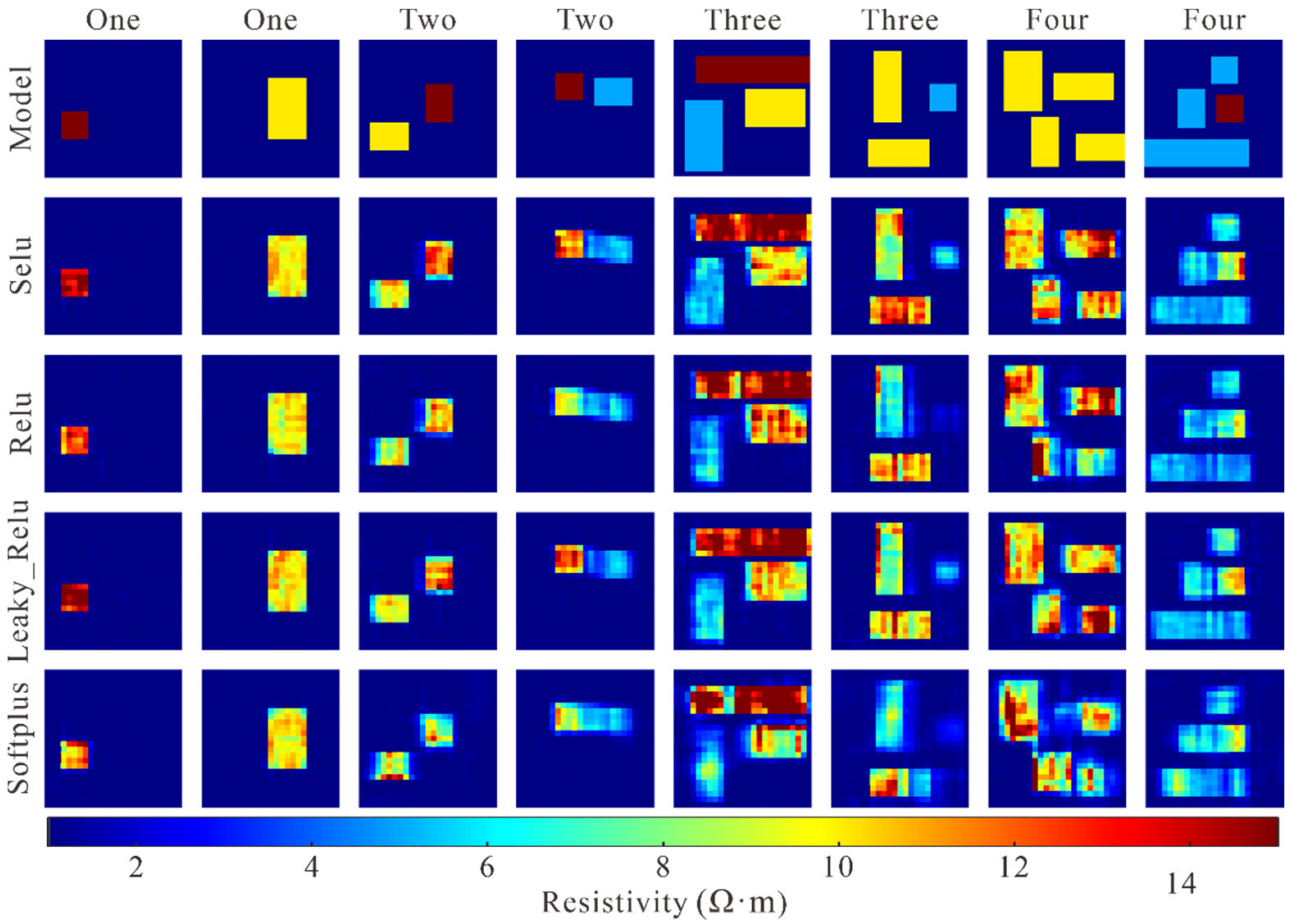


Figure 9. Models and their inversion results in different activation functions.

Table 3. The RE and CC of all samples in test set with different activation function.

Activation Function		Relu	Selu	Leak_Relu	Softplus
RE	AVE	0.3066	0.2301	0.2625	0.4239
	SD	0.1450	0.1096	0.1290	0.2555
CC	AVE	0.9044	0.9449	0.9287	0.8303
	SD	0.0809	0.0502	0.0641	0.1424

AVE = average, SD = standard deviations.

The average CC and RE of different samples in the test set using different activation functions are portrayed in Figure 10. With the gradual increase in the number of hydrate occurrence areas, the average RE is getting larger, while the average CC is getting smaller. For the same number of hydrate occurrence areas, the average CC and RE of Selu is much better than Relu, Leaky_Relu, and Softplus. When the number of hydrate occurrence areas is one, their average CC and RE are similar. Nevertheless, the greater the number of hydrate occurrence areas is, the better the average CC and RE of Selu will be. Therefore, the performance of ERTInvNet with Selu is more advantageous than other activation functions.

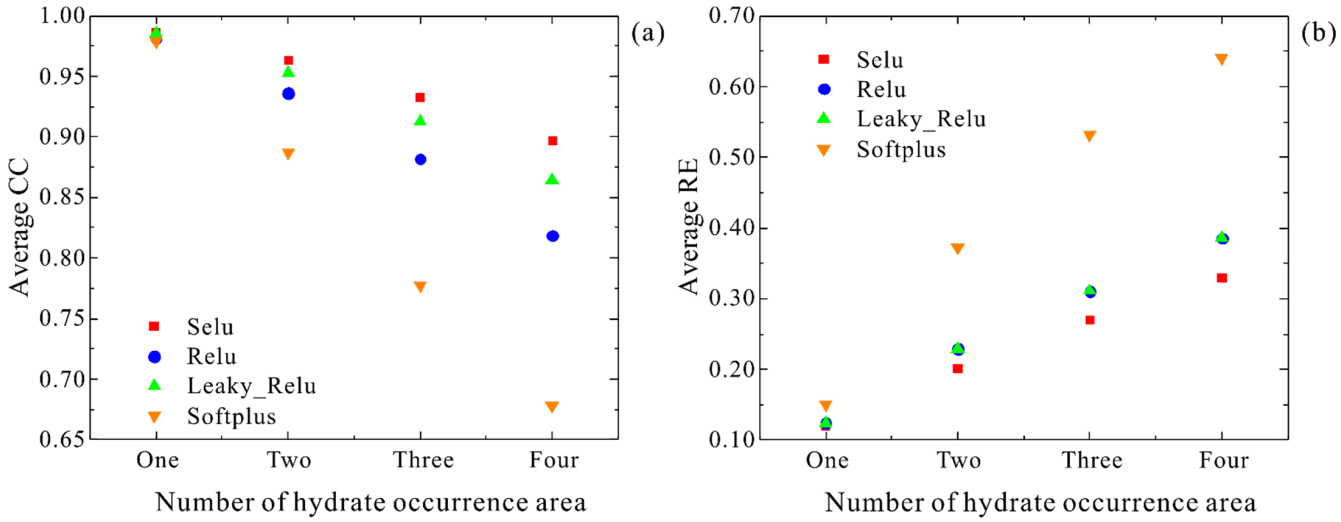


Figure 10. The average RE and CC of different samples in the test set for different activation functions. (**a**) average CC, (**b**) average RE.

4.3. Evaluation of Inversion Network Models with Different Optimizers

Four optimizers (Adadelta, Adagrad, Adamax, and Adam) are performed in training the proposed ERTInvNet to access which optimizer is most appropriate to obtain a smaller loss value. The Selu activation function is enabled in each hidden layer, and Linear activation function for the output layer. MSE loss function is chosen, and four types of optimizers minimize the loss value. The parameters used by each optimizer are portrayed in Table 1.

The performance curve of the training sets and the validation sets depicted in Figure 11 are calculated for each epoch during the training process using ERTInvNet with different optimizers. As the epoch proceeds, the loss value decreases from the initial high value, and the accuracy value increases from the initial low value. Specifically, in Figure 11a, the loss curve with Adam during the training and validation process converges slower than other optimizers. However, Adam can obtain a lower loss value of 0.05 when the curve is stable. During the training process, the loss of value with Adam is always at the lowest point, which indicates that the Adam optimizer is robust. In Figure 11b, it can be observed that the accuracy value with Adam is the highest. Four accuracy curves of Adadelta, Adagrad, Adamax, and Adam have obvious differences, and the accuracy of the first three does not exceed 0.9. Through the loss curve and accuracy curve on the training and validation sets, it can be initially inferred that using the Adam optimizer can promote the performance of the ERTInvNet model.

Some results of the four typical hydrate distributions were randomly chosen to analyze the effect of the network model using different optimizers. The inversion results on the test set using the network model with different optimizers is portrayed in Figure 12. The shape, position, and size of the inversion results with Adam are better than other optimizers. The boundary blur phenomenon in the inversion with Adam, Adamax, Adagrad, and Adadelta gradually worsens.

The quality of inversion results using different optimizers is discussed by calculating the CC and RE of all samples in the test set. The average RE and CC and the standard deviations of RE and CC of all samples in the test set with different optimizers are portrayed in Table 4. The average RE of Adam is 0.2301. Compared with Adadelta, Adagrad, and Adamax, its average RE is the smallest, which means Adam is more suitable for ERT inversion. Moreover, the average CC of Adam is 0.9449, which is significantly better than other optimizers. The standard deviation of the CC for Adam is smaller than for the other three. Although the standard deviations of RE of Adam are larger, its RE fluctuates at a low level. The results indicate that the ERTInvNet model with Adam has good performance.

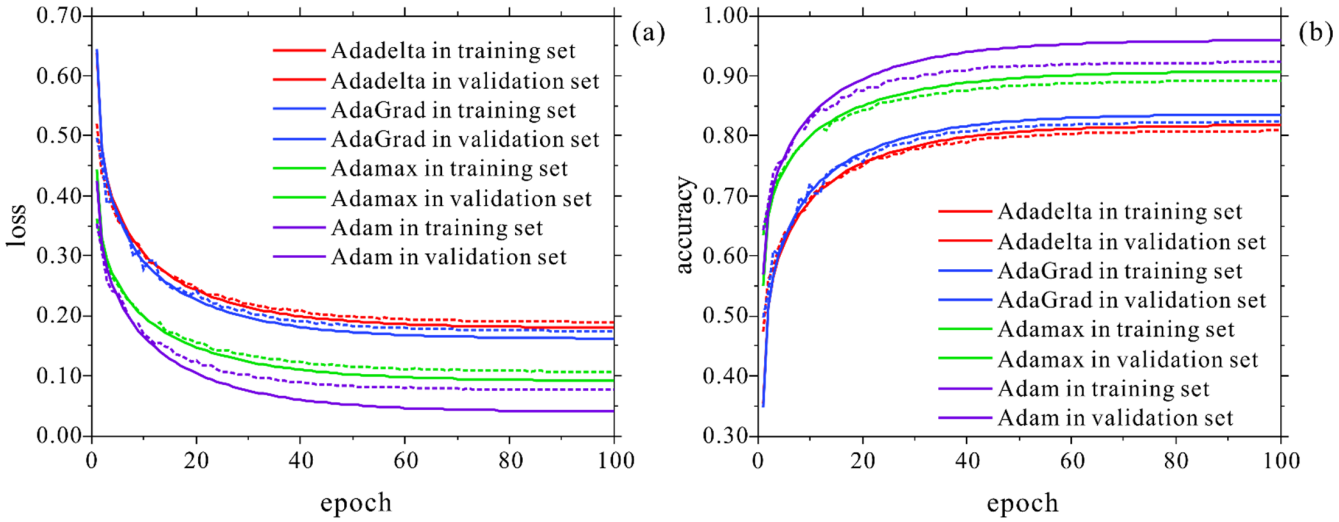


Figure 11. The model's performance curve with different optimizers. (**a**) loss curve, (**b**) accuracy curve.

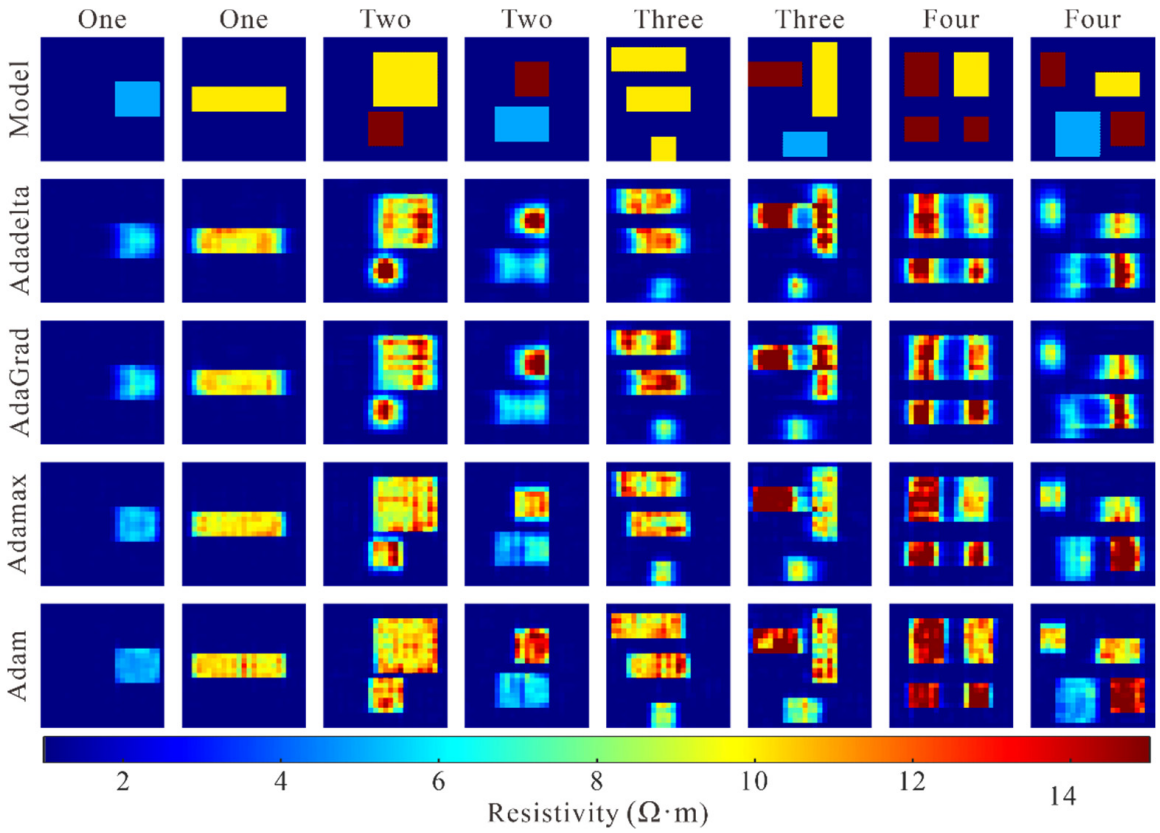


Figure 12. Models and their inversion results with different optimizers.

Table 4. The RE and CC of all samples in the test set with different optimizers.

Optimizer		Adadelta	Adagrad	Adamax	Adam
RE	AVE	0.3951	0.3790	0.3790	0.2301
	SD	0.1032	0.1041	0.1057	0.1096
CC	AVE	0.8545	0.8655	0.9180	0.9449
	SD	0.0715	0.0699	0.0565	0.0502

AVE = average, SD = standard deviations.

The average CC and RE of different samples in the test set using different optimizers are depicted in Figure 13. With the increase in the number of hydrate occurrence areas, the average RE increases linearly, and the average CC decreases linearly. For the same number of hydrate occurrence areas, Adam's average CC and RE are much better than Adadelta, Adagrad, and Adamax. The greater the number of hydrate occurrence areas, the better Adam's average CC and RE performance. Therefore, ERTInvNet with Selu is better than other optimizers in performance.

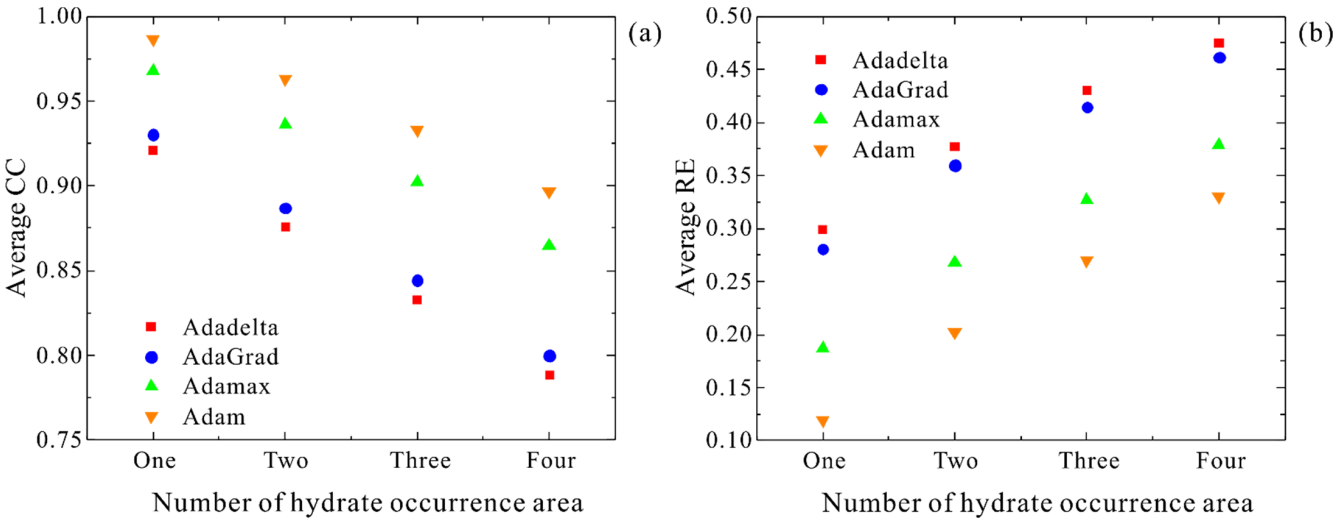


Figure 13. The average RE and CC of different samples in the test set for different optimizers. (**a**) average CC, (**b**) average RE.

5. Conclusions

Gas hydrates as new energy sources have been researched vigorously across the world. Its exploitation and utilization, however, pose new challenges for many countries. Therefore, accurate monitoring of hydrate distribution is crucial for secure and efficient hydrate exploitation.

An ERTInvNet based on the DNN network framework is proposed to solve ERT nonlinear inverse problem in this paper. The comparison is taken between various deep learning parameters to recognize better hydrate distribution, including loss function, activation function, and optimizer. Logcosh loss function should be considered preferentially in the ERTInvNet. Compared with MSE, MAE, and Huber loss function, its average CC and RE of all samples in test sets is superior. The standard deviations of CC and RE of Huber are the smallest, which indicates that the Huber loss function may also be used in the ERTInvNet. ERTInvNet with using Selu possesses obvious advantages under the condition of complex hydrate distribution. The average CC and RE of Selu are 0.9449 and 0.2301, and the standard deviations of CC and RE of Selu are 0.0502 and 0.1057, which is far superior to that of Relu, Leaky_Relu, and Softplus activation functions. The average and standard deviations of CC and RE of Adam are the best, leading to the most accurate inversion.

Deep learning has demonstrated a great potential in ERT inversion. Compared with the traditional inversion method, the deep learning inversion has the advantages of fast calculation, high resolution, and high accuracy. In the future, the combinations of deep learning parameters and some new neural network structures will be used to improve the accuracy of the ERT inversion result. We expect that ERTInvNet will be used mineral exploration and hydrocarbon mapping.

Author Contributions: Conceptualization, Y.L.; methodology, Y.L.; software, Y.L.; validation, C.W. and J.Z.; formal analysis, Y.L.; investigation, Y.L.; resources, C.Z. and Q.C.; data curation, Y.L.; writing—original draft, Y.L.; writing—review and editing, C.Z., Q.C., J.Z. and C.W.; visualization, Y.L.; supervision, C.Z. and Q.C.; project administration, C.Z. and Q.C.; funding acquisition, C.Z. and Q.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Research and Development Plan of Key Fields in Guangdong Province (No. 2020B111103003), the Natural Science Foundation of Shandong Province (No. ZR2020KE026), and Marine Geological Survey Program (No. DD20190231).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this paper are available upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sun, Y.F.; Zhong, J.R.; Chen, G.J.; Cao, B.J.; Li, R.; Chen, D.Y. A new approach to efficient and safe gas production from unsealed marine hydrate deposits. *App. Energy* **2021**, *282*, 116259. [CrossRef]
2. Sultan, N.; Cochonat, P.; Foucher, J.P.; Mienert, J. Effect of gas hydrates melting on seafloor slope instability. *Mar. Geol.* **2004**, *213*, 379–401. [CrossRef]
3. Kvalstad, T.J.; Andresen, L.; Forsberg, C.F.; Berg, K.; Bryn, P.; Wangen, M. The Storegga slide: Evaluation of triggering sources and slide mechanics. *Mar. Pet. Geol.* **2005**, *22*, 245–256. [CrossRef]
4. Hornbach, M.J.; Lavier, L.L.; Ruppel, C.D. Triggering mechanism and tsunamogenic potential of the Cape Fear Slide complex, US Atlantic margin. *Geochem. Geophys. Geosyst.* **2007**, *8*, Q12008. [CrossRef]
5. Partain, R.A.; Yiallourides, C. Hydrate occurrence in Europe: Risks, rewards, and legal frameworks. *Mar. Policy* **2020**, *121*, 104122. [CrossRef] [PubMed]
6. Priegnitz, M.; Thaler, J.; Spangenberg, E.; Ruecker, C.; Schicks, J.M. A cylindrical electrical resistivity tomography array for three-dimensional monitoring of hydrate formation and dissociation. *Rev. Sci. Instrum.* **2013**, *84*, 104502. [CrossRef]
7. Priegnitz, M.; Thaler, J.; Spangenberg, E.; Schicks, J.M.; Schroetter, J.; Abendroth, S. Characterizing electrical properties and permeability changes of hydrate bearing sediments using ERT data. *Geophys. J. Int.* **2015**, *202*, 1599–1612. [CrossRef]
8. Zhao, J.H.; Liu, C.L.; Chen, Q.; Zou, C.C.; Liu, Y.; Bu, Q.T.; Kang, J.L.; Meng, Q.G. Experimental investigation into three-dimensional spatial distribution of the fracture-filling hydrate by electrical property of hydrate-bearing sediments. *Energies* **2022**, *15*, 3537. [CrossRef]
9. Cui, Z.Q.; Wang, Q.; Xue, Q.; Fan, W.R.; Zhang, L.L.; Cao, Z.; Sun, B.Y.; Wang, H.X.; Yang, W.Q. A review on image reconstruction algorithms for electrical capacitance/resistance tomography. *Sens. Rev.* **2016**, *36*, 429–445. [CrossRef]
10. Vauhkonen, M.; Vadasz, D.; Karjalainen, P.A.; Somersalo, E.; Kaipio, J.P. Tikhonov regularization and prior information in electrical impedance tomography. *IEEE Trans. Med. Imaging* **1998**, *17*, 285–293. [CrossRef]
11. Hermans, T.; Kemna, A.; Nguyen, F. Covariance-constrained difference inversion of time-lapse electrical resistivity tomography data. *Geophysics* **2016**, *81*, E311–E322. [CrossRef]
12. Gamio, J.C.; Ortiz-Alemán, C.; Martin, R. Electrical capacitance tomography two-phase oil-gas pipe flow imaging by the linear back-projection algorithm. *Geofísica Int.* **2005**, *44*, 256–273.
13. Peng, L.H.; Merkus, H.; Scarlett, B. Using regularization methods for image reconstruction of electrical capacitance tomography. *Part. Part. Syst. Charact.* **2000**, *17*, 96–104. [CrossRef]
14. Lei, J.; Liu, S.; Li, Z.; Schlaberg, H.I.; Sun, M. An image reconstruction algorithm based on the regularized total least squares method for electrical capacitance tomography. *Flow Meas. Instrum.* **2008**, *19*, 325–330. [CrossRef]
15. Wang, M. Inverse solutions for electrical impedance tomography based on conjugate gradients methods. *Meas. Sci. Technol.* **2002**, *13*, 101–117. [CrossRef]
16. Qiang, J.K.; Han, X.; Dai, S.K. 3D DC Resistivity Inversion with Topography Based on Regularized Conjugate Gradient Method. *Geophys. J. Int.* **2013**, *2013*, 931876. [CrossRef]
17. Günther, T.; Ruecker, C.; Spitzer, K. Three-dimensional modelling and inversion of dc resistivity data incorporating topography—II. Inversion. *Geophys. J. Int.* **2006**, *166*, 506–517. [CrossRef]
18. Yan, H.; Wang, Y.F.; Zhou, Y.G. Three-dimensional electrical capacitance tomography reconstruction by the Landweber iterative algorithm with fuzzy thresholding. *IET Sci. Meas. Technol.* **2014**, *8*, 487–496. [CrossRef]
19. Bahraini, T.; Azimpour, P.; Yazdi, H.S. Modified-mean-shift-based noisy label detection for hyperspectral image classification. *Comput. Geosci.* **2021**, *155*, 104843. [CrossRef]
20. Purswani, P.; Karpyn, Z.T.; Enab, K.; Xue, Y.; Huang, X. Evaluation of image segmentation techniques for image-based rock property estimation. *J. Pet. Sci. Eng.* **2020**, *195*, 107890. [CrossRef]

21. Yu, C.; Pei, H. Face recognition framework based on effective computing and adversarial neural network and its implementation in machine vision for social robots. *Comput. Electr. Eng.* **2021**, *92*, 107128. [CrossRef]
22. Lee, K.P.; Wu, B.H.; Peng, S.L. Deep-learning-based fault detection and diagnosis of air-handling units. *Build. Environ.* **2019**, *157*, 24–33. [CrossRef]
23. Sidorenko, M.; Orlov, D.; Ebadi, M.; Koroteev, D. Deep learning in denoising of micro-computed tomography images of rock samples. *Comput. Geosci.* **2021**, *151*, 104716. [CrossRef]
24. Yu, H.; Yang, L.T.; Zhang, Q.; Armstrong, D.; Deen, M.J. Convolutional neural networks for medical image analysis: State-of-the-art, comparisons, improvement and perspectives. *Neurocomputing* **2021**, *444*, 92–110. [CrossRef]
25. Wu, X.; Liang, L.; Shi, Y.; Fomel, S. FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation. *Geophysics* **2019**, *84*, IM35–IM45. [CrossRef]
26. Liu, Z.G.; Chen, H.; Ren, Z.Y.; Tang, J.T.; Xu, Z.M.; Chen, Y.P.; Liu, X. Deep learning audio magnetotellurics inversion using residual-based deep convolution neural network. *J. Appl. Geophys.* **2021**, *188*, 104309. [CrossRef]
27. Bergen, K.J.; Johnson, P.A.; Hoop, M.V.; Beroza, G.C. Machine learning for data-driven discovery in solid Earth geoscience. *Science* **2019**, *363*, eaau0323. [CrossRef]
28. Tan, C.; Lv, S.; Dong, F.; Takei, M. Image Reconstruction Based on Convolutional Neural Network for Electrical Resistance Tomography. *IEEE Sens. J.* **2019**, *19*, 196–204. [CrossRef]
29. Li, F.; Tan, C.; Dong, F. Electrical Resistance Tomography Image Reconstruction with Densely Connected Convolutional Neural Network. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 4500811. [CrossRef]
30. Liu, B.; Guo, Q.; Wang, K.; Pang, Y.; Nie, L.; Jiang, P. Adaptive Convolution Neural Networks for Electrical Resistivity Inversion. *IEEE Sens. J.* **2021**, *21*, 2055–2066. [CrossRef]
31. Liu, B.; Guo, Q.; Li, S.C.; Liu, B.C.; Ren, Y.X.; Pang, Y.H.; Guo, X.; Liu, L.B.; Jiang, P. Deep Learning Inversion of Electrical Resistivity Data. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 5715–5728. [CrossRef]
32. Duchi, J.; Hazan, E.; Singer, Y. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *J. Mach. Learn. Res.* **2011**, *12*, 2121–2159.