mathematics

MDPI

Article

A Seq2Seq Model Improved by Transcendental Learning and Imaged Sequence Samples for Porosity Prediction

Lijian Zhou [1,*], Lijun Wang [1], Zhiang Zhao [1], Yuwei Liu [2] and Xiwu Liu [2]

[1] School of Information and Control Engineering, Qingdao University of Technology, Qingdao 266525, China
[2] SinoPEC Petroleum Exploration and Production Research Institute, Beijing 100083, China
* Correspondence: zhoulijian@qut.edu.cn

Abstract: Since the accurate prediction of porosity is one of the critical factors for estimating oil and gas reservoirs, a novel porosity prediction method based on Imaged Sequence Samples (ISS) and a Sequence to Sequence (Seq2Seq) model fused by Transcendental Learning (TL) is proposed using well-logging data. Firstly, to investigate the correlation between logging features and porosity, the original logging features are normalized and selected by computing their correlation with porosity to obtain the point samples. Secondly, to better represent the depositional relations with depths, an ISS set is established by slidingly grouping sample points across depth, and the selected logging features are in a row. Therefore, spatial relations among the features are established along the vertical and horizontal directions. Thirdly, since the Seq2Seq model can better extract the spatio-temporal information of the input data than the Bidirectional Gate Recurrent Unit (BGRU), the Seq2Seq model is introduced for the first time to address the logging data and predict porosity. The experimental results show that it can achieve superior prediction results than state-of-the-art. However, the cumulative bias is likely to appear when using the Seq2Seq model. Motivated by teacher forcing, the idea of TL is proposed to be incorporated into the decoding process of Seq2Seq, named the TL-Seq2Seq model. The self-well and inter-well experimental results show that the proposed approach can significantly improve the accuracy of porosity prediction.

Keywords: porosity prediction; deep learning; transcendental learning; imaged sequence samples; logging data

MSC: 68T07

check for updates

Citation: Zhou, L.; Wang, L.; Zhao, Z.; Liu, Y.; Liu, X. A Seq2Seq Model Improved by Transcendental Learning and Imaged Sequence Samples for Porosity Prediction. *Mathematics* **2023**, *11*, 39. https://doi.org/10.3390/math11010039

Academic Editors: Hongji Xu and Zhi Liu

Received: 17 November 2022
Revised: 15 December 2022
Accepted: 19 December 2022
Published: 22 December 2022

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

As one of the critical indicators in petrophysics, porosity is a measure of the pore space in reservoir rocks, which can represent the reservoir's oil and gas content [1,2]. Therefore, predicting porosity is crucial for calculating oil reserves, organizing field development, maximizing oil recovery, etc. [3]. The traditional core porosity measurement is performed on the core extracted by logging in the laboratory [4]. Nevertheless, it is challenging to provide continuous rock porosity over a large depth range since experimental procedures are time-consuming, expensive, and only limited assessment data at discrete sample locations can be obtained [5]. The logging data are the most prevalent type of information in the oil and gas field. The data are generally continuous throughout the well, accurate in depth, high in resolution, complete in type, and easy to use [6,7]. With the development of various logging equipment and logging techniques, the amount of geological information contained in the logging data has increased exponentially. Porosity prediction using logging data has become a research hotspot [8–10].

Statistical techniques in mathematics are usually used to analyze logging data and reveal porosity trends in conventional methods. Gu et al. [11] built a porosity fitting model by employing stepwise regression and N-way analysis of variance algorithms, respectively.

A statistical approach combining N-way analysis of variance and multivariate linear fitting (MLF) [12] was further proposed to inverse porosity only from logging source data, which decreased its dependency on physical model parameters compared to the method in [11]. However, reservoirs are typically characterized by geological heterogeneity, so the obtained logging data are usually distributed in a spatially non-uniform and non-linear pattern, which leads to a complex correlation between reservoir porosity and logging features [13–15]. Common mathematical models have a limited capacity to create non-linear mappings, and as a result, they cannot accurately capture the intricate geological information inherent in logging data to some extent [16–18]. The use of ML on reservoir data has yielded promising results [19–21]. Zou et al. [22] performed porosity prediction using uncertainty quantification based on Random Forest (RF), which could achieve higher regression efficiency. Yasin et al. [23] combined Support Vector Regression machine (SVR) and a Particle Swarm Optimization algorithm (PSO) to forecast the lithology and porosity distribution using both cable logging and core data, and achieved more stable prediction results. Gamal et al. [24] developed an ANN model for building porosity prediction cubes. Chen et al. [25] designed a back propagation neural network model optimized by a genetic algorithm, which can optimize structural parameters and realize a global search for optimal solutions by self-learning and self-adaptation.

Although machine learning techniques can establish nonlinear mappings, they have limited learning ability, poor generalization ability, and are more sensitive to changes in the data environment [26]. Therefore, it is commonly effective only for data with simple relations, and cannot accurately describe the complex relations between porosity and logging features. Deep learning is a new research field in machine learning motivated by building deep neural networks that imitate the interpretive mechanisms of outer data in the human brain, such as images, sounds and text [27]. A new direction for mathematical research was provided by deep learning algorithms, and deep learning algorithms achieve continuous innovation with the development of mathematics, algorithm theory and engineering practice. There have been recent advancements in porosity prediction using deep learning algorithms like Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), etc. [28–30]. In addition, strong continuity exists in original point samples of logging data generally under the effect of stratigraphic deposition. According to this, Feng [31] proposed a reservoir porosity prediction method based on CNN in which the formation mechanism of reservoir pores is fully considered. Thus, the local intrinsic relation of rock properties in a short term can be taken into account by the filter window to obtain better predictive performance [32]. From the perspective of reservoir deposition continuity, Wang et al. [33,34] viewed the logging data as time-series data in the vertical direction and then used the RNN to learn the temporal correlations. It effectively compensated for the deficiency of traditional depth network that could not provide contextual information. On this basis, Wang et al. [35] further constructed a one-dimensional CNN to learn local correlations and hierarchical correlations of logging data, and also built a Bidirectional RNN (BRNN) to learn the tendency of features with depth and contextual information. Although the aforementioned methods have improved the prediction effect to some extent, they usually increase the computational complexity and are prone to overfitting when using the point samples directly. The Seq2Seq network model [36] has an adaptable structure, which means it can be fine-tuned to fit a wide variety of inputs and perform well in a variety of prediction and classification tasks. Therefore, the Seq2Seq network based on serial RNN is introduced first to predict the porosity in this paper.

To consider the relationship between the point samples fully, the sequence samples are constructed to map reservoir deposition relationships in this paper. Furthermore, the sequence samples of different logging features are ranged to form an ISS data. Thus, horizontal inter-feature information and vertical depositional relationships are both characterized by ISS, which can accurately represent the spatial dependency and coupling relationships between sedimentary rocks.

Although the better prediction effects are obtained using the Seq2Seq network, it has been discovered that cumulative bias can easily occur in the Seq2Seq structure [37]. That is, if an error occurs in a RNN cell, then this error is propagated to the following RNN cells. Li et al. introduced the real label values directly under the guidance of the Teacher Forcing (TF) strategy during training to alleviate the cumulative bias [38]. However, the exposure bias arises during the prediction process since real porosity values of the testing sample cannot be provided. To solve this problem, the TL module is proposed to assist the TF in the Seq2Seq network. The concept of TL is derived from the Bayesian formula in mathematics, in which the statistical results or prior experiences are taken into account. The label semantic information and encoding information are considered in the proposed TL module. Thus, a TL-Seq2Seq network, which is relatively less dependent on label semantics, is constructed. When the network prediction deviates a lot from the real value, the TL-Seq2Seq network can provide appropriate guidance to rectify the exposure bias. In addition, it can also resolve the issue of cumulative bias.

The remaining sections of this paper are laid out as follows. Section 2 introduces the original dataset and gives the experimental configuration and evaluation metrics. Section 3 introduces the methodology for preprocessing the dataset and constructing the ISS dataset. The structures of the Seq2Seq and proposed TL-Seq2Seq models are given, and their performances are analyzed in Section 4. Section 5 performs parametric and ablation experiments on the TL-Seq2Seq method and compares it with other methods. A summary of the whole paper is provided in Section 6.

2. Data Preparation and Experimental Setup

(1) Original dataset

The dataset was obtained from the real logging data of wells Luo69 and Wang76 in the exploratory zone of ShengLi Oilfield; the relative locations of the wells are shown in Figure 1. In both wells, data are selected from the bathymetry of 1878.5–2905 m. A total of 8000 samples are contained. Each logging dataset is divided into 16 segments, each of which has 500 samples, as shown in Figure 2. From the first segment, a segment is taken as the test sample every three segments (corresponding to the 1st, 4th, 7th, 10th, 13th, and 16th segments). In this way, the test set contains 6 well segments with a total of 3000 point samples, and the remaining 5000 point samples (10 segments) are used as the training set.

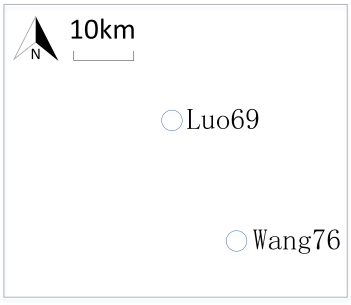

Figure 1. Relative location of logs in the study area.

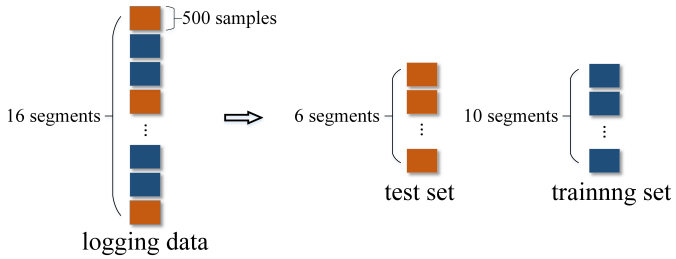

Figure 2. Dataset segmentation flowchart.

(2) Experimental environment and parameters setting

The experiments are performed on a workstation with two NVIDIA GeForce RTX 2080Ti graphics cards. The operating system is Ubuntu 18.04 LTS with torch 1.10.0 + GPU framework.

In the experiment, the parameter initialization method that obeys the uniform distribution is adopted in the model. During the training process, the batch size and number of iterations are set to 64 and 50, respectively. The standard SGD with a momentum of 0.9 and weight decay of 0.0005 is used as the optimizer. The initial learning rate is set to 0.01, and the MultiStepLR learning rate decay strategy is used. The decreasing value interval is set to 6∼20 and γ is set to 0.12. The mean square error loss function is used. The dropout rate is set to 0.5. The optimal effects have been achieved by the proposed method under these basic parameters setting. And the effects keep high levels within a certain range of each parameter value, which to some extent indicates the robustness of the network. To maintain consistency, the settings are used in all the methods mentioned in this paper except the methods quoted from others' papers.

(3) Evaluation indicators

The coefficient of determination (R^2) and the Root Mean Square Error (RMSE) are used to evaluate the performance of the model. Mean Absolute Error (MAE), Mean Squared Error(MSE), RMSE, and R^2 are commonly used to evaluate the regression prediction models. The MAE represents the average of the absolute differences between the actual and predicted values in the data set, in which the average of the residuals is measured. It is a nondifferentiable function, that is, the gradient updated by MAE in the calculation of neural networks is always the same. Therefore, MSE is selected in the loss function in this paper. For consistency, MAE is not considered here. MSE is a differentiable function that measures the variance of residuals. It is sensitive to dimensions. RMSE can eliminate the influence of dimensions, so it can describe the distribution of predicted values around the regression line. Moreover, due to the high prediction performance of the model in this paper, the calculation error is relatively small. According to this, a more intuitive numerical display can be presented by RMSE under the action of the square root. R^2 is used to measure the degree of correlation between the predicted result and the true value. It represents the proportion of dependent variables interpreted by a linear regression model. It is a unitless fraction and can eliminate the influence of the dispersion degree of the original data, so that it can objectively evaluate the fitting degree of the model.

In conclusion, both RMSE and R^2 are selected as the evaluation indicators since they can quantify the suitability of a linear regression model. When R^2 is closer to 1, the two variables are highly correlated. On the contrary, a smaller value of RMSE indicates a better prediction. These two indicators can be calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{1}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{2}$$

where n represents the number of samples, y_i means the true value of the sample, $\hat{y}_i$ denotes the predicted value of the sample, and $\bar{y}_i$ refers to the mean value of the samples.

3. Imaged Sequence Sample Dataset Construction

Due to the extremely nonlinear and geologically heterogeneous quality of the logging data, it is insufficient to accurately characterize the correlation between features and porosity only using the point sample separately in the raw data [39–41]. Therefore, the first proposal to convert the original point samples to ISS is made in this paper. Deep deposition

information is integrated into the ISS, so that the characteristics of depositional strata can be explored thoroughly to make predictions better. The construction process is divided into three parts: data normalization, logging features selection, and imaged sequence sample construction, as shown in Figure 3. The specific process is described as follows.

Figure 3. Flowchart of the ISS dataset construction.

3.1. Selection of Logging Features

Each well contains the following 10 features: spontaneous potential (SP); compensation density (DEN); natural gamma (GR); transverse wave velocity (DTS); 2.5 m bottom gradient resistivity (R25); compensated neutron log (CNL); caliper (CAL); acoustic transit time (AC); resistivity digital focused logging (RDFL); and azimuth (AZIM). The linear correlation between two variables (X,Y) can be measured by correlation coefficient (r) that is computed as Equation (3). The correlation coefficient between every feature and porosity (POR) is ranked and shown as a heat map in Figure 4. It can be seen that the RDFL and AZIM features in both wells are less correlated with porosity. Besides, considering that in some scenarios, the features may not all be available when analyzing logging data. That means fewer features are required in similar scenarios. Therefore, these two low-coefficient features are removed and the remaining eight logging features are retained as sample features. In other scenarios, the features can be determined according to the situation based on this idea.

$$r(X,Y) = \frac{\mathrm{Conv}(X,Y)}{\sqrt{\mathrm{Var}[X]\,\mathrm{Var}[Y]}} \tag{3}$$

where Conv denotes covariance, and Var denotes variance.

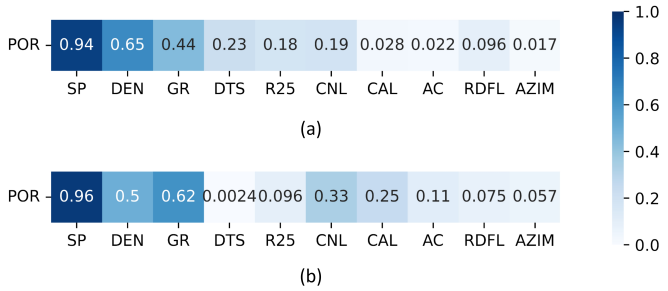

Figure 4. Heat map of correlation coefficients between 10 features and porosity. (**a**) Well Luo69. (**b**) Well Wang76.

3.2. Data Standardization

The features in the logging data have different dimensions. Take the two logging features R25 and DTS with the largest dimension difference as an example, R25 is distributed between 2–5, while DTS is distributed between 150–340, as shown in Figure 5. When they are inputted directly into the network, the role of features with higher values in the prediction will be emphasized and the role of features with lower values will be decreased, which will lead to the less accurate predictions and slower convergence [42]. Therefore, the parameter sequences need to be normalized before constructing the sequence dataset. After the standardization, the ranges of the two features are crossed, while the respective change trend is retained.

The max-min normalization method is chosen to perform a linear transformation on the original data. Assume that the data set contains M feature sequences $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_M$, each of which contains n samples. The sequences after normalization are calculated by

$$\mathbf{x}'_i = \frac{\mathbf{x}_i - \min_{1\le j\le n}\{x_{ij}\}}{\max_{1\le j\le n}\{x_{ij}\} - \min_{1\le j\le n}\{x_{ij}\}},\ 0 \le i \le M-1 \tag{4}$$

The normalized feature sequences $\mathbf{x}'_1, \mathbf{x}'_2, \ldots, \mathbf{x}'_M$ are in the range [0–1].

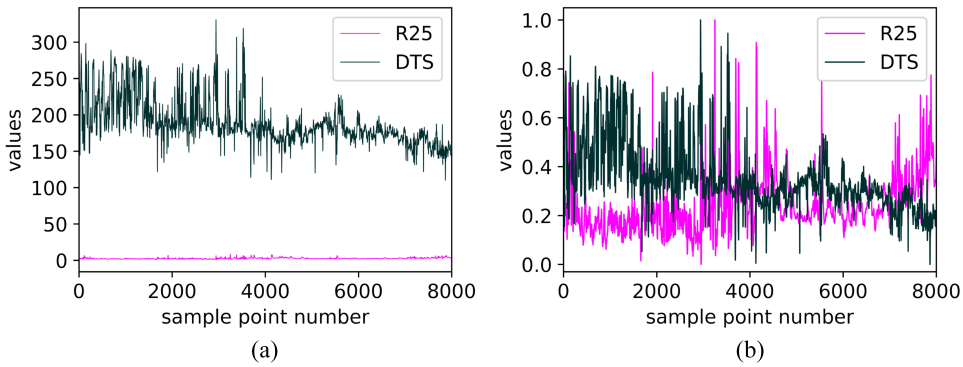

Figure 5. Distributions of R25 and DTS from the (**a**) original dataset and (**b**) standardization dataset.

3.3. ISS Dataset Construction

The variation of each feature with depth in the Wang76 well between 2065–2128 m logging depth is depicted in Figure 6. All features expect SP and DEN exhibit the opposite tendencies with porosity. Thus, it can be accounted that there is a transversal correlation among the features in the logging data. In addition, the features are affected by the layer deposition process. The ISS dataset is constructed in this paper, which can simultaneously represent the transversal correlations between the features and the depth correlation within a feature.

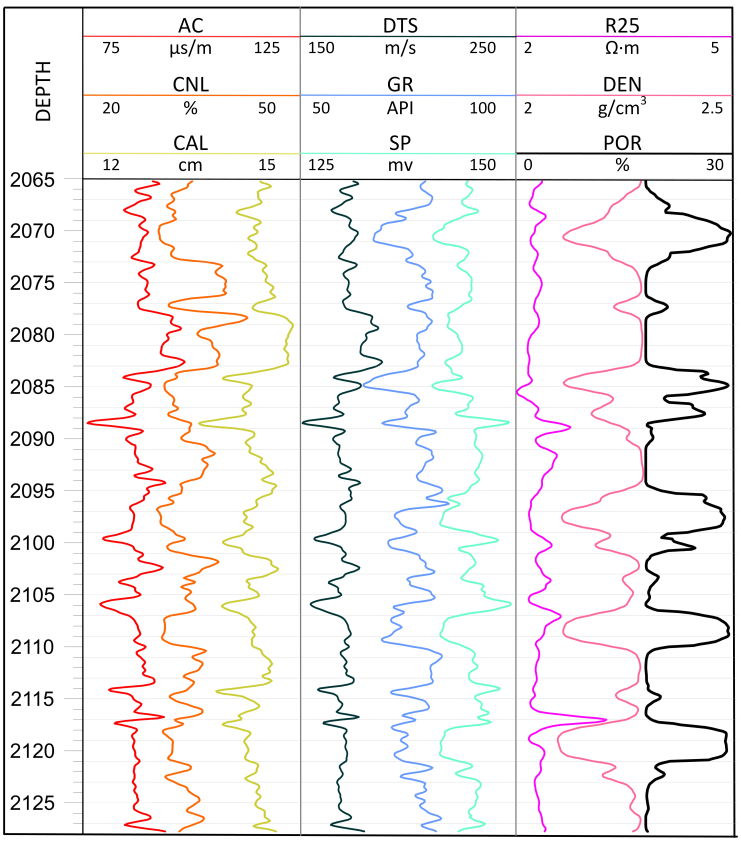

Figure 6. Well logging curve for Wang76.

The ISS construction process is depicted in Figure 7. The normalized features are laid on the left side of the dotted line, the normalized porosity is set on the right side. The sliding windows (for instance, with a window length of 4) are set to frame the data in the depth direction during construction, then the ISS can be obtained by converting the framed block data into an image.

Each column of the ISS represents a type of feature, which can be seen to vary continuously in color with depth direction and retain the correlation with the porosity. For example, DEN and SP have a negative relationship with porosity. It can be seen that DEN and SP are from dark to light in the fourth and final columns, and their porosity is from dark to light colors. The other features are positively correlated with porosity variation; the colors change consistently. The sequence length in the ISS is determined experimentally in Section 5.1.

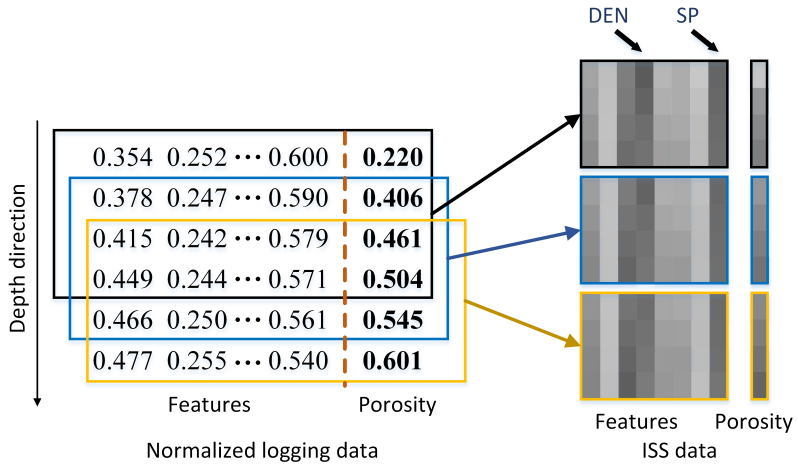

Figure 7. ISS construction process.

4. Porosity Prediction Based on Seq2Seq and TL-Seq2Seq Network Model

The Seq2Seq Network is based on a serial recurrent neural network structure and can process spatial–temporal information of the input data. It has been proven quite effective in both classification and prediction tasks [43]. Therefore, this paper first proposes that the Seq2Seq Network can be used in the porosity prediction tasks. Furthermore, the TL-Seq2Seq model that integrates the concept of transcendental learning is proposed to address the exposure bias problem that is easy to occur in Seq2Seq. The flow of this section is shown in Figure 8.

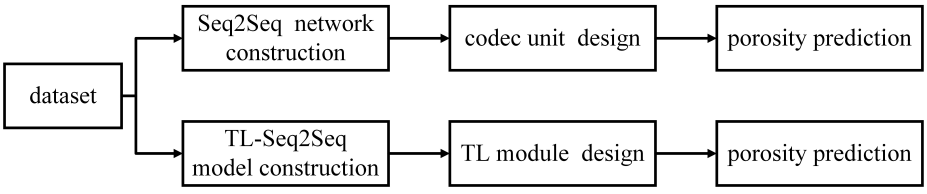

Figure 8. Flowchart of the porosity prediction based on TL-Seq2Seq model.

4.1. Porosity Prediction Based on Seq2Seq Network

4.1.1. Seq2Seq Network

A typical Seq2Seq network architecture with an encoder and a decoder is shown in Figure 9. To enhance the learning ability of the network and optimize the network parameters, the multilayer bidirectional cyclic structure can be used in the encoder part [44]. Given that the decoder function usually runs in a single step, a unidirectional cyclic structure is required. BGRU and GRU are selected as the encoder and decoder of the Seq2Seq network, respectively. Both the input and output of Seq2Seq are sequences. The input sequence is passed to the encoder according to the time step, the encoded

information C that contains the hidden parameters and the encoder output sequence is obtained. The obtained data are then sent to the decoder to decode the prediction sequence.

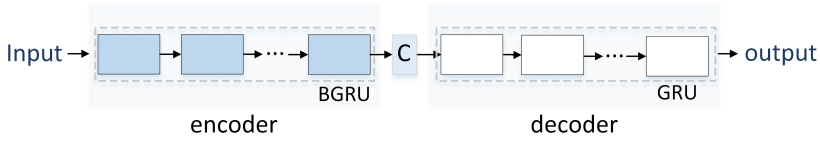

Figure 9. Seq2Seq network structure.

A start character <eos> and an end character <dos> are typically set in the input sequence to label the sequence range in Seq2Seq Network. Since the sample length of the sequence constructed in this paper is fixed, the starting and ending characters are not mandatory. Thus, the output of the final encoding stage can be sent directly into the first input of the decoder, which can maximize circular structure's encoding capabilities. Finally, the output of each time step in the decoder is integrated as the predicted sequence.

If the structures are different for the encoder and decoder, the dimensions of their hidden layers are not aligned. When the BGRU is used in the encoder and the GRU is used in the decoder, the hidden state of the first time step in the decoder can be calculated by

$$\mathbf{h}_0^d = \mathbf{W}_h \begin{bmatrix} \overrightarrow{\mathbf{h}_{t'}^e} \\ \overleftarrow{\mathbf{h}_{t'}^e} \end{bmatrix} \tag{5}$$

where $\mathbf{W}_h$ is the learnable weight of the linear layer, $\overrightarrow{\mathbf{h}_{t'}^e}$ and $\overleftarrow{\mathbf{h}_{t'}^e}$ denote the forward and backward hidden layer states, respectively, the superscripts e, d denote the encoder and decoder, respectively.

4.1.2. Codec Unit of Seq2Seq

The GRU is adopted as the basic structure of the codec in Seq2Seq, as shown in Figure 10. Similar to the ordinary RNN cell, the current state is dependent on not only the current input but also the output of the prior cell state in the GRU, which enables the processing of sequence data by associating contextual semantics [45]. Additionally, the transmission state is also controlled by the gating signals **r** and **z**, which can successfully overcome the gradient explosion and gradient disappearance phenomena that easily arise in RNN while dealing with long sequences [46]. The status transfer process is as follows.

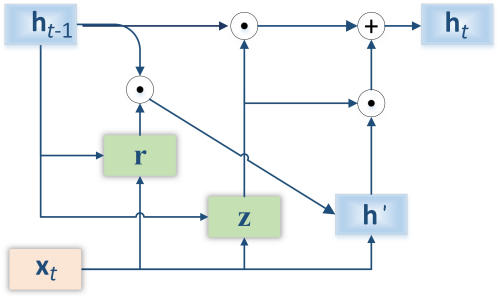

Figure 10. Flow chart of status transmission for GRU cell.

First, the reset gate signal and update gate signal **z** are calculated using the prior communicated state $\mathbf{h}_{t-1}$ and the current moment input $\mathbf{x}_t$. $\mathbf{h}'$ mainly contains the current input $\mathbf{x}_t$, and selectively receives the hidden state $\mathbf{h}_{t-1}$ at the previous moment. Finally, the current status $\mathbf{h}_t$ is obtained according to $\mathbf{h}'$ and $\mathbf{h}_{t-1}$. Thus, the processes of forgetting and remembering are happening simultaneously in the GRU. The detailed calculation is as follows:

$$\mathbf{r} = \sigma(\mathbf{U}_r \mathbf{x}_t + \mathbf{W}_r \mathbf{h}_{t-1}) \tag{6}$$

$$\mathbf{z} = \sigma(\mathbf{U}_z \mathbf{x}_t + \mathbf{W}_z \mathbf{h}_{t-1}) \tag{7}$$

$$\mathbf{h}' = \tanh(\mathbf{U}_h \mathbf{x}_t + \mathbf{W}_h(\mathbf{r} \odot \mathbf{h}_{t-1})) \tag{8}$$

$$\mathbf{h}_t = (1 - \mathbf{z}) \odot \mathbf{h}_{t-1} + \mathbf{z} \odot \mathbf{h}' \tag{9}$$

where $\sigma(\cdot)$ denotes the sigmoid activation function, $\mathbf{U}$ and $\mathbf{W}$ denote the learnable weights of the linear layer. $\odot$ symbolizes the Hadamard product. Gate signal z is between 0 and 1. The information is "remembered" more when z is closer to 1, and is "forgotten" more when it is closer to 0.

The BGRU neural network contains a forward RNN and a backward RNN, both of which connect the input and output layers [47], as shown in Figure 11. It shows how to convey every input sequence to the output layer while keeping the whole historical and prospective context.

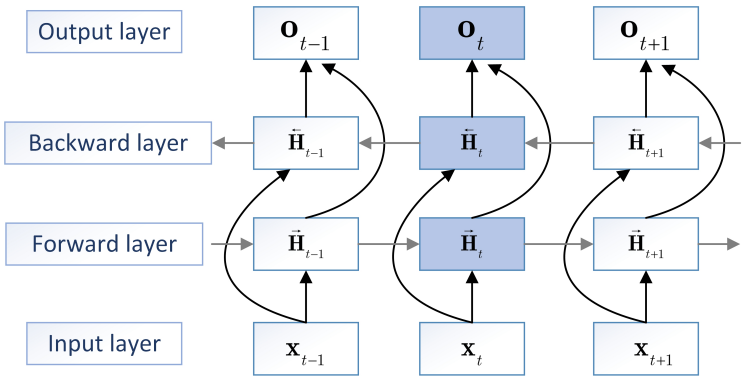

Figure 11. BGRU network.

For time step t, a batch of input data $\mathbf{x}_t \in \mathbb{R}^{n \times d}$ is given, where n is the number of samples, d is the number of inputs in each sample. The hidden layer activation function is denoted as ϕ. In the bidirectional architecture, $\overrightarrow{\mathbf{H}}_t \in \mathbb{R}^{n \times k}$ and $\overleftarrow{\mathbf{H}}_t \in \mathbb{R}^{n \times k}$ represent the forward and backward hidden layer states of the current time step, respectively, where k is the number of hidden cells. Their update formulas are as follows:

$$\overrightarrow{\mathbf{H}}_t = \phi\left(\mathbf{W}_{xh}^{(fw)} \mathbf{x}_t + \mathbf{W}_{hh}^{(fw)} \overrightarrow{\mathbf{H}}_{t-1} + \mathbf{b}_h^{(fw)}\right) \tag{10}$$

$$\overleftarrow{\mathbf{H}}_t = \phi\left(\mathbf{W}_{xh}^{(bw)} \mathbf{x}_t + \mathbf{W}_{hh}^{(bw)} \overleftarrow{\mathbf{H}}_{t+1} + \mathbf{b}_h^{(bw)}\right) \tag{11}$$

where $\mathbf{W}_{xh}^{(fw)} \in \mathbb{R}^{d \times h}$, $\mathbf{W}_{hh}^{(fw)} \in \mathbb{R}^{h \times h}$, $\mathbf{W}_{xh}^{(bw)} \in \mathbb{R}^{d \times h}$, $\mathbf{W}_{hh}^{(bw)} \in \mathbb{R}^{h \times h}$ are the linear layer learnable weights; fw and bw denote the forward and backward layers, respectively; $\mathbf{b}_h^{(fw)} \in \mathbb{R}^{1 \times h}$ and $\mathbf{b}_h^{(bw)} \in \mathbb{R}^{1 \times h}$ indicate bias. Then, the hidden layer state $\mathbf{H}_t \in \mathbb{R}^{n \times 2k}$ can be obtained by stacking the forward hidden layer state $\overrightarrow{\mathbf{H}}_t$ with the backward hidden layer state $\overleftarrow{\mathbf{H}}_t$, and is sent to the output layer. The final output $\mathbf{O}_t \in \mathbb{R}^{n \times q}$ (q is the number of output cells) is calculated by:

$$\mathbf{O}_t = \mathbf{W}_{hq} \mathbf{H}_t + \mathbf{b}_q \tag{12}$$

where weighting matrix $\mathbf{W}_{hq} \in \mathbb{R}^{2h \times q}$, and bias $\mathbf{b}_q \in \mathbb{R}^{1 \times q}$.

4.1.3. Porosity Prediction Simulation Experiments

To evaluate the proposed method, the porosity is predicted based on the point sample of wells Luo69 and Wang76 by using the BGRU (baseline method) and Seq2Seq network, respectively. In addition, the Seq2Seq network is used to predict the porosity based on the

ISS (denoted as Seq2Seq & ISS). Their network structure follows the settings in Table 1. Their performance and the computational burden: floating point operations per second (Flops), parameters (Params) and GPU memory occupation (Memory) is computed, as shown in Table 2. As can be seen from the table, although the computational burden has increased, the prediction accuracy has significantly improved by using the Seq2Seq method.

Table 1. The structural parameters of networks.

Method	Encoder Units	Decoder Units	Hidden Layers
BGRU	256	\	2
Seq2Seq	256	128	2
Seq2Seq & ISS	256	128	2

Table 2. The results obtained by the different methods based on wells Luo69 and Wang76.

Method	Luo69		Wang76		Flops/M	Params/M	Memory/G
	R^2	RMSE	R^2	RMSE			
BGRU	0.9377 ± 0.07%	0.0755 ± 0.16%	0.9463 ± 0.04%	0.0723 ± 0.06%	14.11	0.73	1.67
Seq2Seq	0.9590 ± 0.43%	0.0498 ± 0.37%	0.9590 ± 0.29%	0.0632 ± 0.13%	223.17	1.79	42.26
Seq2Seq &ISS	0.9768 ± 0.35%	0.0459 ± 0.35%	0.9679 ± 0.31%	0.0560 ± 0.09%	221.43	1.77	32.39

To observe intuitively, the predicted porosities using different methods are plotted in Figure 12 with depth ranges between 2065–2127 m and 2253–2315 m, respectively. It can be seen that the Seq2Seq method outperforms the BGRU on the porosity prediction task. The accuracy of porosity prediction using the Seq2Seq & ISS method is improved compared to only using the Seq2Seq method.

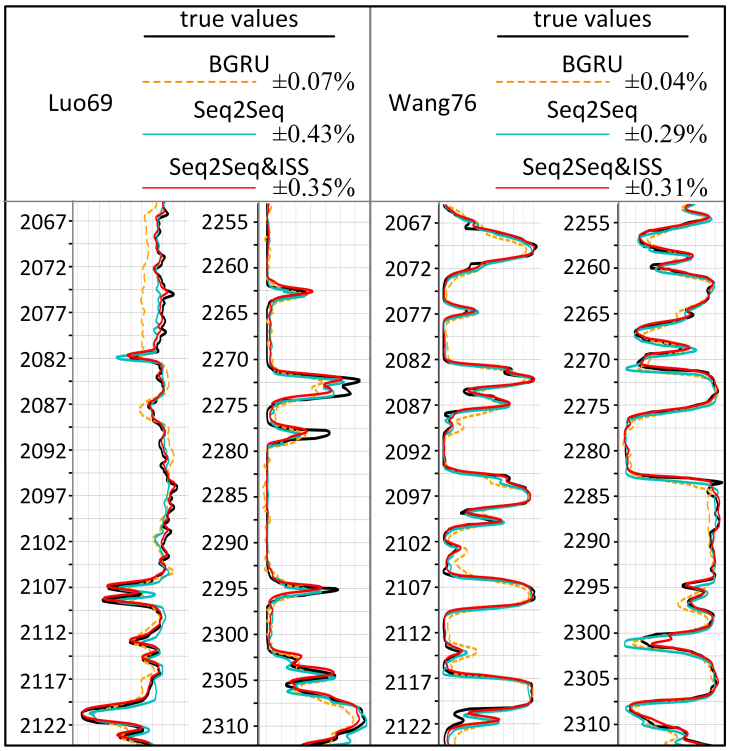

Figure 12. Prediction of Seq2Seq and BGRU for wells Luo69 and Wang76.

4.2. Porosity Prediction Based on TL-Seq2Seq Model

4.2.1. Proposed TL-Seq2Seq Model

Although the Seq2Seq network has achieved pretty good results in the porosity prediction task, the exposure bias problem is prone to occur in training the Seq2Seq network.

Motivated by the idea of TL, the truth value is introduced to optimize the Seq2Seq network and avoid the exposure bias. Thus, the TL module is incorporated into the decoding process, as shown in Figure 13. The proposed model consists of three main modules: Encoder, TL and Decoder.

First, the spatio-temporal feature vector **C** is obtained by encoding the ISS feature using the BGRU module, which contains the hidden state **H** and the output sequence (**enc_seq**). Then, **enc_seq** is sent to the TL module. In the TL module, the **enc_seq** and the ISS label (**true_seq**) are fused with the output sequence of a certain time step in the decoder (**dec_seq**), respectively. One of them is selected by the TF strategy to form the transcendental learning knowledge (**TL-seq**). Finally, the hidden layer data **H** is loaded into the decoder that contains the GRU structure and a linear layer, and the prediction result (**Pred value**) is then obtained by decoding the **TL-seq**. Details are provided below.

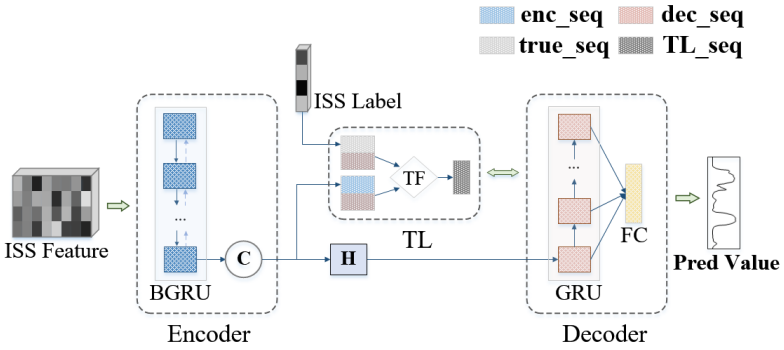

Figure 13. TL-Seq2Seq network structure.

4.2.2. TL Module

The Seq2Seq's network training can be sped up to converge by introducing real label values [48]. The generation of prediction sequences during the training process would be largely dependent on true label values that are directly fed into the decoder as context vectors. During the testing process, the entire sequence is generated alone in the decoder due to the loss of the real sequence guidance. Thus, a significant deviation can occur in the predicted sequence, called "exposure bias" [49,50], which easily leads to overcorrection.

To get rid of this exposure bias, the TL module is constructed to obtain the TL knowledge as the input of the decoder, in which both the encoded information from the encoder and the true label information are used. The hierarchical TL network module that interacts with the decoder is shown in Figure 14, which is discussed in detail below.

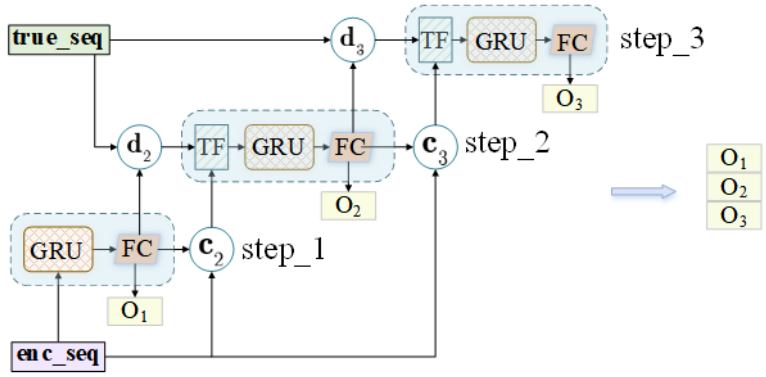

Figure 14. The hierarchical TL network module interacting with decoder.

When $t = 1$, the encoder's output sequence **enc_seq** is used as input to the decoder; when $t > 1$, the context vector $\mathbf{c_t}$ or $\mathbf{d_t}$ is used instead, which includes either the en-

coded information or the true value. Thus, the input **I** of the encoder at each moment is expressed as:

$$\mathbf{I} = \begin{cases} \mathbf{enc\_seq} & , \quad t = 1 \\ f(\mathbf{c}_t, \mathbf{d}_t) & , \quad 1 < t < N-1 \end{cases} \tag{13}$$

where the function f represents the TF strategy that can decide which context vector is assigned to **I**. The context vectors $\mathbf{c_t}$ and $\mathbf{d_t}$ are obtained by stacking the encoded sequence and the true values with the output sequence at moment t ($t > 1$) in the decoder, respectively.

$$\mathbf{c}_t = \mathbf{W_c}\begin{bmatrix} \mathbf{enc\_seq} \\ \mathbf{O}_t \end{bmatrix}, \quad \mathbf{d}_t = \mathbf{W_d}\begin{bmatrix} \mathbf{true\_seq} \\ \mathbf{O}_t \end{bmatrix} \tag{14}$$

where $\mathbf{O}_t$ is the output of the previous step in the decoder, $\mathbf{W_c}$ and $\mathbf{W_d}$ are the learnable weight of the linear layer.

The decoder's dependence on real labels can be reduced by adding the TL module, which thereby avoids the overfitting issue of the Seq2Seq and improves the generalization ability of the network. Meanwhile, the encoder's output sequence is joined at each stage of the decoder's computation, which enables the decoder to combine the encoding information that is closest to the original input sequence. Therefore, the calculation error can also be corrected, and then the accuracy of the predicted sequences will be improved.

The general TL-Seq2Seq model consists of multi-modules, and the prediction sequences (Pred_value) can be generated after the ISS data are fed into the model. The equation for the obtained TL-Seq2Seq model is expressed as follows. First, the ISS Feature is processed by BGRU to obtain center vectors (**C**) that contain the encoding information (**enc_seq**) and hidden layer data (**H**). Then, in TL, ISS Label and **enc_seq** are fused with the decoding information(**dec_seq**), respectively. The TF is used to make decisions on the two fused sequences to obtain Transcendental Learning knowledge (**TL-seq**). Next, the output ($\mathbf{O}_t$) that at every time step is obtained after the hidden layer data are loaded and the TL-seq is processed by GRU. Finally, all outputs (**O**) are inputted to a full connection layer, and the final prediction sequence is obtained.

$$\begin{cases} \mathbf{C} = (\mathbf{H}, \mathbf{enc\_seq}) = \text{BGRU(ISS Feature)}\ ) \\ \mathbf{TL\_seq} = \text{TF}((\text{ISS Label}, \mathbf{dec\_seq}), (\mathbf{enc\_seq}, \mathbf{dec\_seq})) \\ \mathbf{O}_t = \mathbf{dec\_seq} = \text{GRU}(\mathbf{TL\_seq}) \\ \mathbf{Pred\_value} = \text{FC}(\mathbf{O}) \end{cases} \tag{15}$$

4.2.3. Porosity Prediction Simulation Experiments

Based on the ISS dataset constructed in this paper, comparison experiments are conducted using the Seq2Seq network and the TL-Seq2Seq network. The TL-Seq2Seq follows the settings in Table 3. The R^2 is 0.9889 and 0.9856 for Luo69 and Wang76, respectively, when using TL-Seq2Seq, which is better than Seq2Seq (0.9768 and 0.9679, as shown in Table 2). The predicted porosities using the proposed methods are plotted in Figure 15, with depth ranges between 2090–2120 m and 2283–2317 m, respectively. The predicted results from TL-Seq2Seq are shown by the solid blue line, which fits the real curve (the solid black line) better than the results from Seq2Seq (the red dot line), especially in places where the porosity is more volatile.

Table 3. The structural parameters of the proposed method.

Method	Encoder Units	Decoder Units	Hidden Layers	TF Rate
TL-Seq2Seq & ISS	256	128	2	0.45

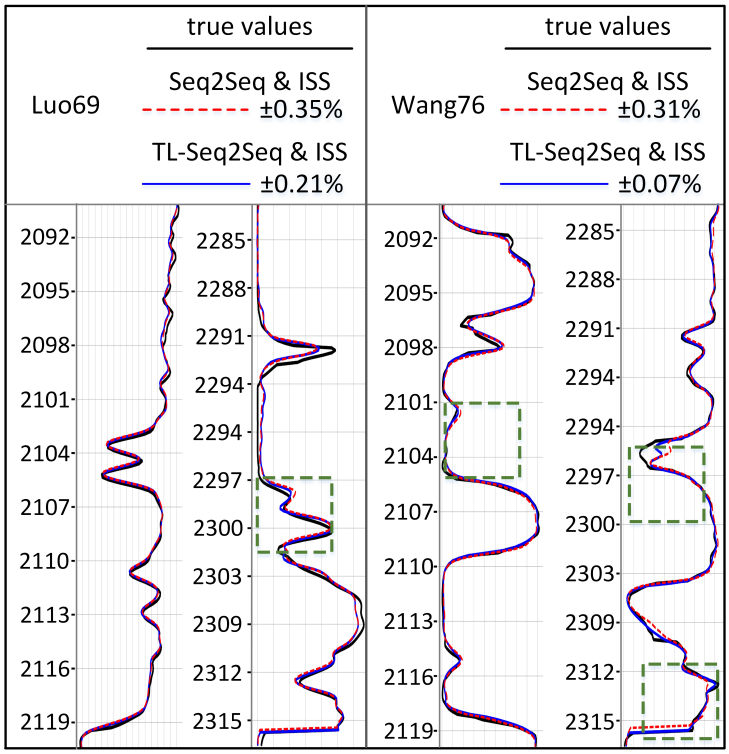

Figure 15. Comparison experimental results using Seq2Seq and TL-Seq2Seq.

5. Experimental Results and Analysis

5.1. Parameter and Structure Selection

(1) Experiments on the number and order of logging features

According to the analysis of the correlation coefficient between logging features and porosity in Section 3.2, the logging features are ranked in descending order as SP, DEN, GR, DTS, R25, CNL, CAL, and AC. To examine the contribution of the logging features to porosity prediction, comparative experiments based on the TL-Seq2Seq are performed, where the logging features are gradually increased in the aforementioned order. The experimental results are displayed in the truncation graphs as shown in Figure 16. It is clear that the final forecast accuracy is increased by using more logging features. The rising trend of prediction accuracy becomes slow at 7–10. A weak gain is generated when adding the last AC (the Feature added from 7 to 8) whose correlation with porosity is the smallest of the eight selected features. This means that AC can be retained, but not many more features should be considered. This supports our selection of features in Section 3.1.

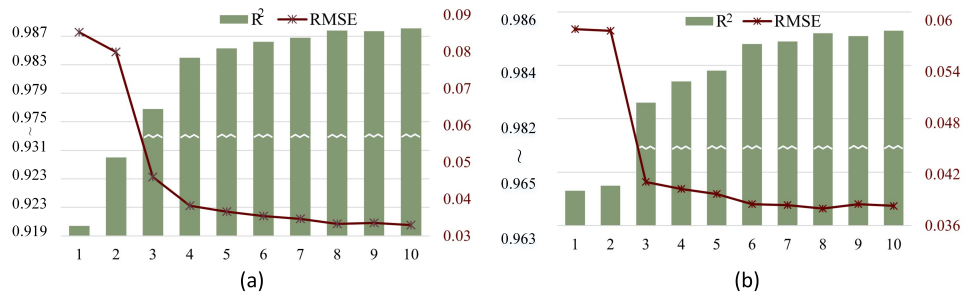

Figure 16. The experimental results using different numbers of logging features from (**a**) well Luo69 and (**b**) well Wang76.

Since the three features SP, GR and DEN have high correlation with POR, one set of comparative experiments (based on the TL-Seq2Seq) is set up: the three features are

placed at the beginning, middle and end of all the features, as well as disorder, to further investigate the influence of parameter order on the prediction results. The experimental results are shown in Table 4. It can be seen that the order of the curves has a minimal impact on the results. In this paper, the disordered features (AC, CAL, CNL, DEN, DTS, GR, R25, and SP) are adopted according to the average performance of the two wells.

Table 4. Selection experiments of logging parameter sequence.

Position	Luo69		Wang76		Average Values	
	R^2	RMSE	R^2	RMSE	R^2	RMSE
beginning	0.9878	0.0333	0.9852	0.0380	0.9865	0.0356
Middle	0.9868	0.0347	0.9863	0.0365	0.9865	0.0356
End	0.9896	0.0307	0.9846	0.0387	0.9871	0.0347
Disorder	0.9888	0.0319	0.9856	0.0374	0.9872	0.0346

(2) Experiments on the sequence length in ISS

The sequence sample may not be able to express enough deposition information when the depth range is narrow. Otherwise, it may have trouble focusing on the stratigraphic information when the depth range is broad. The sequence length is determined by the comparative experiments based on the TL-Seq2Seq. The results are shown in Figure 17. As seen from the Figure 17, the highest R^2 and the smallest RMSE are obtained when the sequence length is 4. Therefore, in this paper, the sequence length is set to 4.

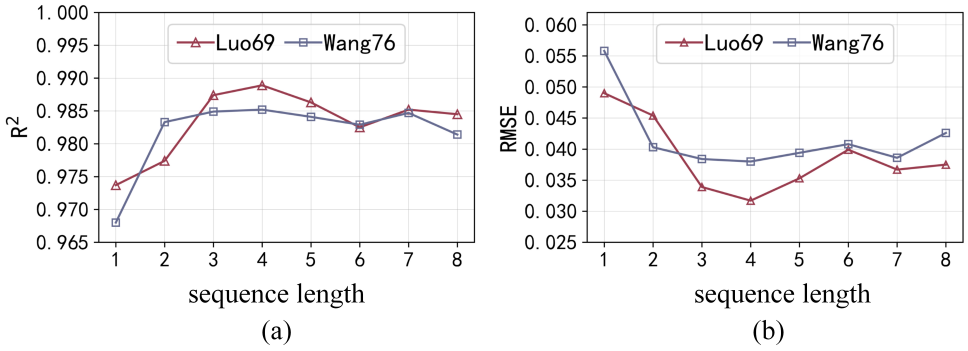

Figure 17. The sequence length selection experimental results on (**a**) R^2 and (**b**) RMSE for ISS data.

(3) Experiments on the codec structure selection in TL-Seq2Seq

A group of comparison experiments are designed to investigate the cyclic network structure in TL-Seq2Seq, which uses the RNN-RNN, BRNN-RNN, GRU-GRU and BGRU-GRU codec structures, respectively. The experimental results are shown in Figure 18. It can be seen that the GRU structure can achieve superior prediction results compared to RNN. The prediction results obtained by using the BGRU-GRU structure have the highest R^2, the smallest RMSE, so the BGRU-GRU structure is used in this paper.

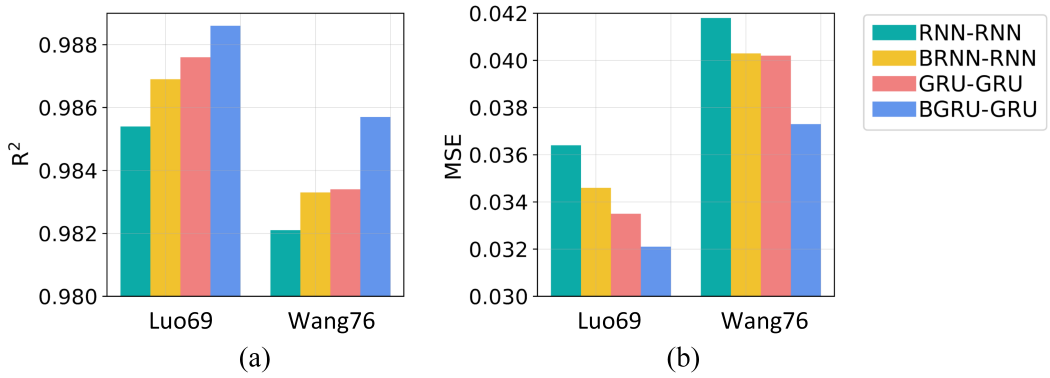

Figure 18. The Structure selection experimental results on (**a**) R^2 and (**b**) RMSE for TL-Seq2Seq.

(4) Experiments on the number of hidden layer cells in TL-Seq2Seq

The numbers of hidden layer units of the encoder and decoder in the TL-Seq2Seq are determined by experiments, whose results are shown in Figure 19. It can be seen that the results are best when the number of hidden layer units is set to 256 in the encoder and 128 in the decoder. Therefore, the numbers of hidden layer units for the encoder and decoder are set to 256 and 128 in this paper, respectively.

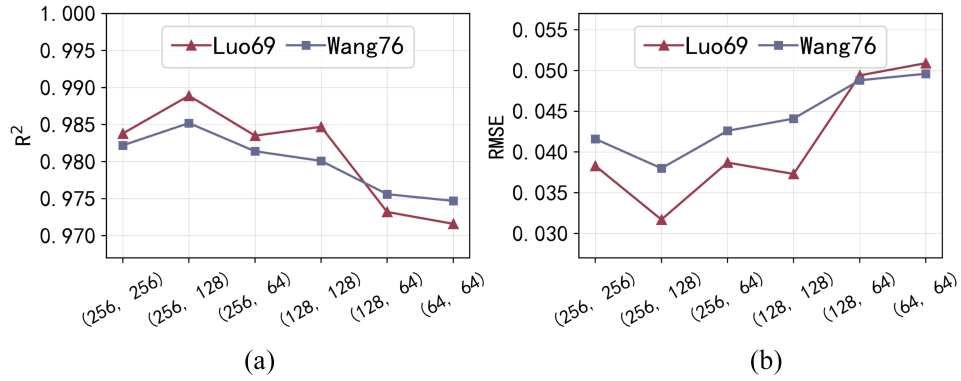

Figure 19. The number selection experimental results on (**a**) R^2 and (**b**) RMSE for hidden layer cells.

(5) Network performance analysis

The parameters and structure have been set according to the prior experiments. The trainings using the TL-Seq2Seq are performed 50 rounds. The training loss and validation loss are displayed in Figure 20. It shows that the training loss decreases quickly, and the validation loss is already at a low level after the first round of training. Eventually, the network is able to converge and maintain a steady state after 20 rounds of training, which indicates the good robustness of the proposed method.

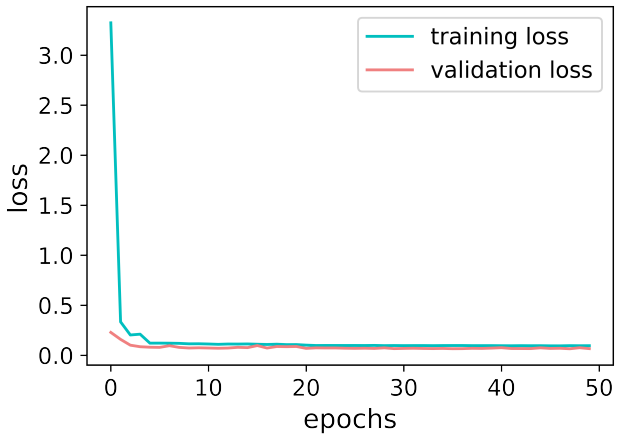

Figure 20. Loss in training using the proposed TL-Seq2Seq method (Wang76 well).

5.2. Ablation Experiments

An ablation study with several configurations is carried out to further evaluate the specific contributions of each module in the proposed approach: ISS dataset, Seq2Seq structure and TL module. The BGRU is set as the baseline method, which does not contain any of the three modules. Experiments are conducted on wells Luo69 and Wang76 to evaluate the performance, and the results are shown in Table 5.

Table 5. Ablation experimental results with different modules.

Number	Method	Luo69		Wang76		Average Value	
		R^2	RMSE	R^2	RMSE	R^2	RMSE
①	BGRU	0.9377	0.0755	0.9463	0.0723	0.9420	0.0739
②	BGRU & ISS	0.9715	0.0509	0.9598	0.0627	0.9656	0.0568
③	Seq2Seq	0.9729	0.0498	0.9590	0.0632	0.9659	0.0565
④	Seq2Seq & ISS	0.9768	0.0459	0.9679	0.0560	0.9723	0.0500
⑤	TL- Seq2Seq	0.9737	0.0490	0.9680	0.0558	0.9708	0.0524
⑥	TL- Seq2Seq & ISS	0.9889	0.0317	0.9856	0.0374	0.9872	0.0345

Firstly, by comparing ② and ①, ④ and ③, and ⑥ and ⑤, it can be concluded that the prediction effectiveness is improved to some extent after using the ISS data, which reflects the superiority of the sequence feature. Secondly, by comparing the prediction results with different network structures (③ vs. ①), the performance improvement brought by using the Seq2Seq structure can be clearly seen. The baseline network is a component of the Seq2Seq network. Seq2Seq not only inherits the prediction ability of the baseline network, but also improves the prediction performance through the serial structure. Lastly, it is proved that the performance using the hierarchical TL network module is further improved by comparing ⑤ and ③, and ⑥ and ④.

The performance using the proposed TL-Seq2Seq method surpasses all other methods' performance by learning ISS data and using hybrid decoding with label sequences. In two wells, the R^2 obtained by TL-Seq2Seq improve about 4.97% and 3.89% over the baseline method (⑥ vs. ①), which indicates that the proposed three modules can work well together to improve the prediction accuracy. The average values of these comparison methods in the two wells are plotted in Figure 21. It can be seen intuitively from the figure that the prediction performance is improving with the combination of various modules.

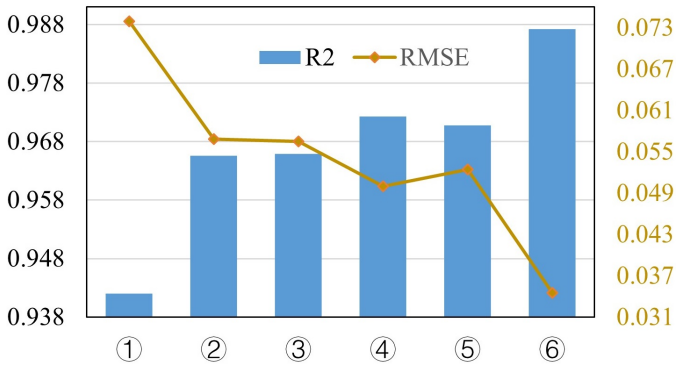

Figure 21. Average performance on two wells of the added modules.

5.3. Model Performance

The prediction models for the two wells are obtained based on the ISS dataset using the TL-Seq2Seq method, and the predicted porosities at the sample points are plotted in Figure 22. The predicted values and the real values are remarkably similar, which demonstrates how well the proposed method predicts the porosity.

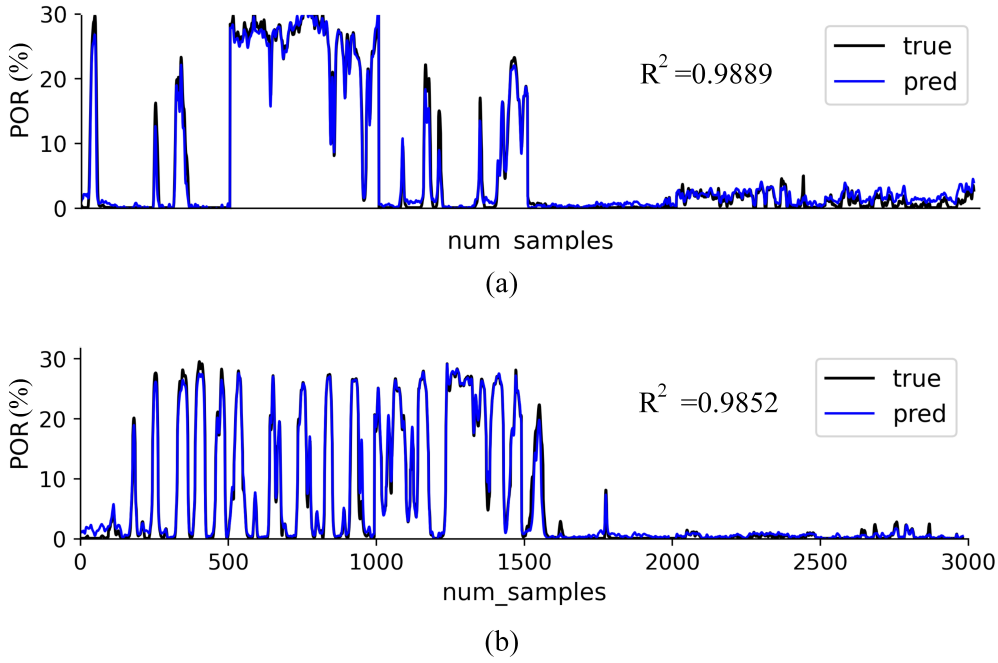

Figure 22. Fitting curve of core porosity and the prediction porosity using TL-Seq2Seq network at test samples. (**a**) Well Luo69, (**b**) Well Wang76.

5.4. Comparison of Methods

To evaluate the performance of the proposed method, this paper compares six recent advanced algorithms whose parameters setting follow the literature, and the experimental results are shown in Table 6. It shows that machine learning algorithms can produce better prediction results than conventional MLF algorithms. The CRNN is constructed by combining CNN and RNN, where the convolutional units are able to extract depth feature information. This is similar to the idea of constructing sequence samples and assigning depth information to the data in this paper, which is suitable for addressing the task of porosity prediction and achieves better prediction accuracy. In this paper, the Seq2Seq network and TL-Seq2Seq network are constructed to train the ISS dataset, respectively. It can be seen that both methods achieve better results compared with other methods, and the TL-Seq2Seq network is the best among all methods. The average values of these

comparison methods in the two wells are plotted in Figure 23. It can be seen intuitively from the figure that the prediction performance is improving with the combination of various modules.

Table 6. The comparison between the proposed methods and the advanced algorithms.

Method	Luo69		Wang76		Average Value		Flops/M	Params/M	Memory/G
	R^2	RMSE	R^2	RMSE	R^2	RMSE			
MLF [12]	0.9445	0.0713	0.9344	0.0799	0.9394	0.0756	\	\	\
RF [22]	0.9495	0.0680	0.9724	0.0518	0.9609	0.0599	\	\	\
SVR-PSO [23]	0.9477	0.0694	0.9619	0.0609	0.9548	0.0651	\	\	\
ANN [24]	0.9606	0.0600	0.9483	0.0710	0.9544	0.0655	2.52	0.14	0.01
BGRU [33]	0.9653	0.0561	0.9667	0.0568	0.9660	0.0564	14.38	0.88	1.78
CRNN [34]	0.9563	0.0632	0.9797	0.0444	0.9680	0.0538	184.17	1.63	25.99
Seq2Seq&ISS (Proposed 1)	0.9768	0.0459	0.9679	0.0560	0.9723	0.0500	221.43	1.77	32.39
TL-Seq2Seq&ISS (Proposed 2)	0.9889	0.0317	0.9852	0.0380	0.9870	0.0348	262.979	1.78	33.27

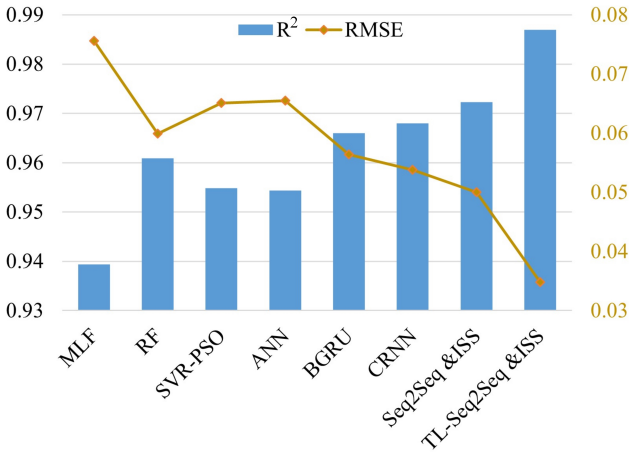

Figure 23. The performance of multiple comparison methods in the average well.

5.5. Experiments of Inter-Well

It has been demonstrated that excellent porosity prediction can be achieved within a single well using the proposed TL-Seq2Seq and ISS method. To evaluate the generalization of the proposed method, the inter-well porosity prediction experiments are performed further. The Luo69 is set as the train set. The well Wang76 is set as test set 1, and the Shi13 well is added as test set 2. The relative location of the well Shi13 is depicted in Figure 24. The Shi13 is farther from the Luo69 than the Wang76.

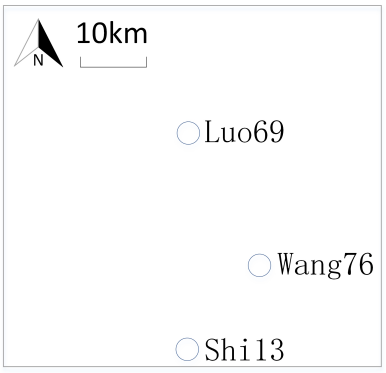

Figure 24. The relative position of the well Shi13.

5.5.1. Performance Testing of the Proposed Method on Porosity Prediction

Performance testing based on TL-Seq2Seq and baseline methods (BGRU) is conducted on two test wells. For visual observation, the predicted results of the sections (2065–2128 m) contained in two test wells are selected and plotted in Figure 25, respectively. For the Wang76 well, subtle changes in porosity with depth can be accurately predicted by the proposed method, and these changes are also able to match the lithological changes. The R^2 can reach more than 0.9730, which is much higher than the baseline method (0.8451) and is almost as good as the inner-well prediction (0.9852). In the prediction of the well Shi13, the results show considerable disagreement between true curves and predicted curves. It may be caused by the distance between the test well and the training well, as well as the different construction conditions of the two wells. Besides, the R^2 obtained by the proposed method is 0.6725, and the R^2 obtained by the baseline method is 0.6241, which indicates that the proposed method is a superior method. Although the R^2 is relatively lower than the result based on Wang76, the approximate range and change trend of predicted values is similar to that of the true values, which means the proposed method is effective.

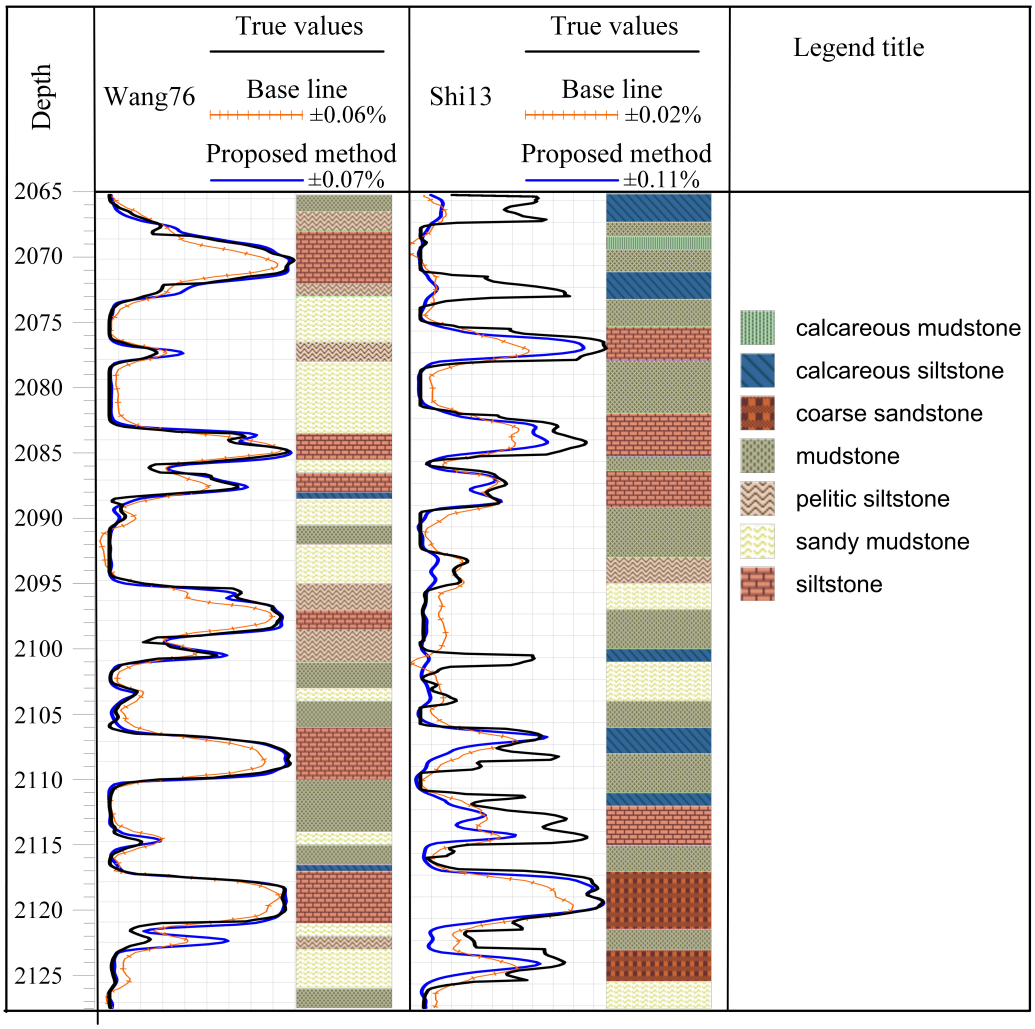

Figure 25. Partial prediction results of porosity for well Wang76 and Shi13.

5.5.2. Comparison of Methods

To better evaluate the proposed method, comparison experiments are conducted based on inter-well datasets with the same five algorithms that are already mentioned in Section 5.4. The comparison methods are implemented according to the parameters setting in the literature, and the experimental results are shown in Table 7. It can be seen that the proposed method can achieve the best results among all the compared methods. The result

using the ANN method is the worst among all methods, which are probably not suitable for the inter-well dataset prediction task due to using the original parameters setting or network structure [22]. The simplest MLF method used for the inter-well prediction task is able to obtain relatively good prediction results. CRNN performs well in the comparison experiments, which again indicates that a performance improvement in the porosity prediction can be achieved by feature processing like that described in this paper. The average values of these comparison methods in the two wells are plotted in Figure 26. It can be seen intuitively from the figure that the proposed methods can achieve better effects.

Table 7. The comparisons between the proposed methods and the other advanced algorithms based on inter-well data.

Method	Wang76		Shi13		Average Value		Flops/M	Params/M	Memory/G
	R^2	RMSE	R^2	RMSE	R^2	RMSE			
MLF [12]	0.8827	0.1068	0.6241	0.2063	0.7534	0.1565	\	\	\
RF [22]	0.6937	0.1726	0.3881	0.2632	0.5409	0.2179	\	\	\
SVR-PSO [23]	0.5890	0.1376	0.2282	0.2243	0.4086	0.1809	\	\	\
ANN [24]	−0.3932	0.3681	−0.8388	0.4564	−0.6160	0.4122	1.89	0.14	0.01
BGRU [33]	0.8451	0.1227	0.6241	0.2063	0.7346	0.1645	7.40	0.88	1.80
CRNN [34]	0.9551	0.0660	0.6397	0.2020	0.7974	0.1340	88.79	1.63	25.64
Seq2Seq &ISS (Proposed 1)	0.9624	0.0604	0.6534	0.1981	0.8079	0.1292	152.23	1.77	30.96
TL-Seq2Seq &ISS (Proposed 2)	0.9730	0.0511	0.6725	0.1926	0.8227	0.1218	152.25	1.78	30.98

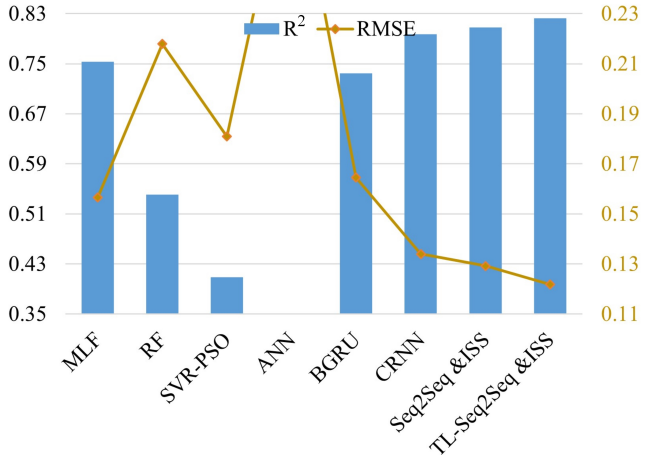

Figure 26. The performance of multiple comparison methods in average well.

6. Conclusions

The Seq2Seq network and the improved TL-Seq2Seq model with ISS data are proposed in this paper to predict porosity. The single point samples are firstly integrated into an ISS to provide small-scale spatial variability characteristics of subsurface complex reservoirs. Then, a Seq2Seq network based on BGRU-GRU is built to predict porosity, which can effectively realize the porosity prediction. Moreover, the problems of cumulative bias and exposure bias, which are prone to occur in Seq2Seq, are addressed by the proposed TL modules. In the TL-Seq2Seq network, the encoding information and label semantic information are referenced to form TL knowledge, which can make the network converge faster and better. Compared with other methods, experimental results show that the performance of the proposed TL-Seq2Seq model is the best, which has the highest R^2 (0.9872, for average well) and the smallest RMSE (0.0345, for average well). In this way, reliable basis and space constraints for further research can be provided, such as geological modeling and rock mechanics parameters modeling. Meanwhile, it also provides a new idea for the mathematical research.

Some issues still remain in the investigated porosity prediction method due to the constraints of experimental conditions and time. In the future, time-frequency analysis can be performed on the ISS to obtain multi-scale feature information, such as wavelet transform. In addition, since the logging data only provide information in the vertical direction in some wells and cannot cover the lateral information of the whole reservoir area, porosity prediction will be performed based on the seismic data with wider coverage.

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

Funding: This research was funded by the National Science and Technology Major Project "The Geophysical Identification and Prediction Method of Terrestrial Shale Oil Sweets" (2017ZX05049-002), and partially by the National Natural Science Foundation of China, grant number 62171247 and 41921781.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Katz, B.J.; Arango, I. Organic porosity: A geochemist's view of the current state of understanding. *Org. Geochem.* **2018**, *123*, 1–16. [CrossRef]
2. Khan, N.M.; Cao, K.; Emad, M.Z.; Hussain, S.; Rehman, H.; Shah, K.S.; Rehman, F.U.; Muhammad, A. Development of Predictive Models for Determination of the Extent of Damage in Granite Caused by Thermal Treatment and Cooling Conditions Using Artificial Intelligence. *Mathematics* **2022**, *10*, 2883. [CrossRef]
3. Yu, H.; Wang, Z.; Rezaee, R.; Zhang, Y.; Han, T.; Arif, M.; Johnson, L. Porosity estimation in kerogen-bearing shale gas reservoirs. *J. Nat. Gas Sci. Eng.* **2018**, *52*, 575–581. [CrossRef]
4. Mahzari, P.; AlMesmari, A.; Sohrabi, M. Co-history matching: A way forward for estimating representative saturation functions. *Transp. Porous Media* **2018**, *125*, 483–501. [CrossRef]
5. Chen, S. Automatic Identification of Lithology and Sedimentary Micro-Phase Using Logging Curves—An Example of Block 482 in Daqing Oilfield. Ph.D Thesis, Jilin University, Changchun, China, 2018.
6. Aghli, G.; Moussavi-Harami, R.; Mohammadian, R. Reservoir heterogeneity and fracture parameter determination using electrical image logs and petrophysical data (a case study, carbonate Asmari Formation, Zagros Basin, SW Iran). *Pet. Sci.* **2020**, *17*, 51–69. [CrossRef]
7. Li, X. Prediction of Reservoir Parameters Based on Logging Data. Ph.D. Thesis, Northeastern Petroleum University, Daqing, China, 2021.
8. Wang, Z.; Nie, X.; Zhang, C.; Wang, M.; Zhao, J.; Jin, L. Lithology Classification and Porosity Estimation of Tight Gas Reservoirs With Well Logs Based on an Equivalent Multi-Component Model. *Front. Earth Sci.* **2022**, *10*, 850023. [CrossRef]
9. Maurya, S.; Singh, K.; Singh, N. Qualitative and quantitative comparison of geostatistical techniques of porosity prediction from the seismic and logging data: A case study from the Blackfoot Field, Alberta, Canada. *Mar. Geophys. Res.* **2019**, *40*, 51–71. [CrossRef]
10. Liu, G.; Liu, B.; Huang, Z.; Chen, Z.; Jiang, Z.; Guo, X.; Li, T.; Chen, L. Hydrocarbon distribution pattern and logging identification in lacustrine fine-grained sedimentary rocks of the Permian Lucaogou Formation from the Santanghu basin. *Fuel* **2018**, *222*, 207–231. [CrossRef]
11. Gu, Y.; Bao, Z.; Lin, Y.; Qin, Z.; Lu, J.; Wang, H. The porosity and permeability prediction methods for carbonate reservoirs with extremely limited logging data: Stepwise regression vs. N-way analysis of variance. *J. Nat. Gas Sci. Eng.* **2017**, *42*, 99–119. [CrossRef]
12. Gu, Y.; Bao, Z.; Rui, Z.; Song, X. Efficient prediction for porosity only using logging data: A case study of lacustrine carbonate reservoirs of IARA oilfield. *Int. J. Oil Gas Coal. Technol.* **2020**, *25*, 133–160. [CrossRef]
13. Xu, J.; Zhang, B.; Xu, L. Predicting the Porosity of Natural Fractures in Tight Reservoirs. *Arab. J. Sci. Eng.* **2018**, *43*, 311–319. [CrossRef]
14. Shang, X.; Zhang, Z.; Yang, W.; Wang, J.; Zhai, C. A Thermal-Hydraulic-Gas-Mechanical Coupling Model on Permeability Enhancement in Heterogeneous Shale Volume Fracturing. *Mathematics* **2022**, *10*, 3473. [CrossRef]

15. Zhang, C.; Shan, W.; Wang, X. Quantitative evaluation of organic porosity and inorganic porosity in shale gas reservoirs using logging data. *Energy Sources Part A Recovery Util. Environ. Eff.* **2019**, *41*, 811–828. [CrossRef]
16. Khasanov, M.K.; Kildibaeva, S.R.; Stolpovsky, M.V.; Musakaev, N.G. Mathematical Model of the Process of Non-Equilibrium Hydrate Formation in a Porous Reservoir during Gas Injection. *Mathematics* **2022**, *10*, 4054. [CrossRef]
17. Ahmadi, M.A.; Chen, Z. Comparison of machine learning methods for estimating permeability and porosity of oil reservoirs via petro-physical logs. *Petroleum* **2019**, *5*, 271–284. [CrossRef]
18. Ahmadi, M.A.; Ahmadi, M.R.; Hosseini, S.M.; Ebadi, M. Connectionist model predicts the porosity and permeability of petroleum reservoirs by means of petro-physical logs: Application of artificial intelligence. *J. Pet. Sci. Eng.* **2014**, *123*, 183–200. [CrossRef]
19. Wood, D.A. Predicting porosity, permeability and water saturation applying an optimized nearest-neighbour, machine-learning and data-mining network of well-log data. *J. Pet. Sci. Eng.* **2020**, *184*, 106587. [CrossRef]
20. Zhang, Z.; Zhang, H.; Li, J.; Cai, Z. Permeability and porosity prediction using logging data in a heterogeneous dolomite reservoir: An integrated approach. *J. Nat. Gas Sci. Eng.* **2021**, *86*, 103743. [CrossRef]
21. Hou, M.; Wang, F.; Zai, Y.; Lian, P. Prediction of porosity and permeability of carbonate rocks based on machine learning and logging data. *J. Jilin Univ. Earth Sci. Ed.* **2022**, *52*, 644.
22. Zou, C.; Zhao, L.; Xu, M.; Chen, Y.; Geng, J. Porosity prediction with uncertainty quantification from multiple seismic attributes using Random Forest. *J. Geophys. Res. Solid Earth* **2021**, *126*, e2021JB021826. [CrossRef]
23. Yasin, Q.; Sohail, G.M.; Khalid, P.; Baklouti, S.; Du, Q. Application of machine learning tool to predict the porosity of clastic depositional system, Indus Basin, Pakistan. *J. Pet. Sci. Eng.* **2021**, *197*, 107975. [CrossRef]
24. Gamal, H.; Elkatatny, S. Prediction model based on an artificial neural network for rock porosity. *Arab. J. Sci. Eng.* **2022**, *47*, 11211–11221. [CrossRef]
25. Chen, L.; Lin, W.; Chen, P.; Jiang, S.; Liu, L.; Hu, H. Porosity prediction from well logs using back propagation neural network optimized by genetic algorithm in one heterogeneous oil reservoirs of Ordos Basin, China. *J. Earth Sci.* **2021**, *32*, 828–838. [CrossRef]
26. Valentín, M.B.; Bom, C.R.; Compan, A.L.M.; Correia, M.D.; de Jesus, C.M.; de Lima Souza, A.; Márcio, P.; Marcelo, P.; Faria, E.L. Estimation of permeability and effective porosity logs using deep autoencoders in borehole image logs from the brazilian pre-salt carbonate. *J. Pet. Sci. Eng.* **2018**, *170*, 315–330. [CrossRef]
27. Yasin, Q.; Du, Q.; Ismail, A.; Shaikh, A. A new integrated workflow for improving permeability estimation in a highly heterogeneous reservoir of Sawan Gas Field from well logs data. *Geomech. Geophys. Geo-Energy Geo-Resour.* **2019**, *5*, 121–142. [CrossRef]
28. Graczyk, K.M.; Matyka, M. Predicting porosity, permeability, and tortuosity of porous media from images by deep learning. *Sci. Rep.* **2020**, *10*, 21488. [CrossRef]
29. Chen, W.; Yang, L.; Zha, B.; Zhang, M.; Chen, Y. Deep learning reservoir porosity prediction based on multilayer long short-term memory network. *Geophysics* **2020**, *85*, WA213–WA225. [CrossRef]
30. Lu, H.; Li, Q.; Yue, D.; Wu, S.; Fu, Y.; Tang, R.; Zhang, Z. Study on optimal selection of porosity logging interpretation methods for Chang 73 segment of the Yanchang Formation in the southwestern Ordos Basin, China. *J. Pet. Sci. Eng.* **2021**, *198*, 108153. [CrossRef]
31. Feng, R. Estimation of reservoir porosity based on seismic inversion results using deep learning methods. *J. Nat. Gas Sci. Eng.* **2020**, *77*, 103270. [CrossRef]
32. Bedi, J.; Toshniwal, D. Features denoising-based learning for porosity classification. *Neural Comput. Appl.* **2020**, *32*, 16519–16532. [CrossRef]
33. Wang, J.; Cao, J.; Zhou, X. Reservoir porosity prediction based on deep bidirectional recurrent neural network. *Prog. Geophys.* **2022**, *37*, 267–274.
34. Wang, J.; Cao, J. Deep Learning Reservoir Porosity Prediction Using Integrated Neural Network. *Arab. J. Sci. Eng.* **2022**, *47*, 11313–11327. [CrossRef]
35. Zhang, Z.; Li, M.; Lin, X.; Wang, Y.; He, F. Multistep speed prediction on traffic networks: A deep learning approach considering spatio-temporal dependencies. *Transp. Res. Part C Emerg. Technol.* **2019**, *105*, 297–322. [CrossRef]
36. Xiang, Z.; Yan, J.; Demir, I. A Rainfall-Runoff Model With LSTM-Based Sequence-to-Sequence Learning. *Water Resour. Res.* **2020**, *56*, e2019WR025326. [CrossRef]
37. Li, M.; Miao, Z.; Xu, W. A CRNN-based attention-seq2seq model with fusion feature for automatic Labanotation generation. *Neurocomputing* **2021**, *454*, 430–440. [CrossRef]
38. Zhu, L.; Zhang, C.; Guo, C.; Jiao, Y.; Chen, L.; Zhou, X.; Zhang, C.; Zhang, Z. Calculating the total porosity of shale reservoirs by combining conventional logging and elemental logging to eliminate the effects of gas saturation. *Petrophys. SPWLA J. Form. Eval. Reserv. Descr.* **2018**, *59*, 162–184. [CrossRef]
39. Shao, R.; Xiao, L.; Liao, G.; Zhou, J.; Li, G. Research on geophysical logging reservoir parameter prediction method based on migration learning. *Chin. J. Geophys.* **2022**, *65*, 796–808.
40. Li, J.; Wu, Q.; Jin, W.; Lu, J.; Nan, Z. Logging evaluation of free-gas saturation and volume content in Wufeng-Longmaxi organic-rich shales in the Upper Yangtze Platform, China. *Mar. Pet. Geol.* **2019**, *100*, 530–539. [CrossRef]
41. Ye, Y.; Tang, S.; Xi, Z.; Jiang, D.; Duan, Y. A new method to predict brittleness index for shale gas reservoirs: Insights from well logging data. *J. Pet. Sci. Eng.* **2022**, *208*, 109431. [CrossRef]

42. Cho, K.; Van Merrienboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; pp. 1724–1734.
43. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to sequence learning with neural networks. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 1–9.
44. Song, H.; Chen, W.; Li, M.; Wang, H. A method to predict reservoir parameters based on convolutional neural network-gated recurrent unit (CNN-GRU). *Editor. Dep. Pet. Geol. Recovery Effic.* **2019**, *26*, 73–78.
45. Zhou, X.; Cao, J.; Wang, X.; Wang, J.; Liao, W. Acoustic log reconstruction based on bidirectional Gated Recurrent Unit (GRU) neural network. *Prog. Geophys.* **2022**, *37*, 357–366.
46. Schuster, M.; Paliwal, K.K. Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* **1997**, *45*, 2673–2681. [CrossRef]
47. Deng, L.; Pan, Y. Data-driven proxy model for waterflood performance prediction and optimization using Echo State Network with Teacher Forcing in mature fields. *J. Pet. Sci. Eng.* **2021**, *197*, 107981. [CrossRef]
48. Chiu, C.C.; Sainath, T.N.; Wu, Y.; Prabhavalkar, R.; Nguyen, P.; Chen, Z.; Kannan, A.; Weiss, R.J.; Rao, K.; Gonina, E.; et al. State-of-the-art speech recognition with sequence-to-sequence models. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; pp. 4774–4778.
49. Gu, J.; Lu, Z.; Li, H.; Li, V.O. Incorporating copying mechanism in sequence-to-sequence learning. *Annu. Meet. Assoc. Comput. Linguist.* **2016**, *1*, 1631–1640.
50. Lamb, A.M.; Parth, A.G.; Zhang, Y.; Zhang, S.; Courville, A.C.; Bengio, Y. Professor forcing: A new algorithm for training recurrent networks. *Adv. Neural Inf. Process. Syst.* **2016**, *29*, 2016.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.