

## Research Article

# Prediction of Shear Wave Velocity Based on a Hybrid Network of Two-Dimensional Convolutional Neural Network and Gated Recurrent Unit

Tengfei Chen <sup>1</sup>, Gang Gao <sup>1</sup>, Peng Wang <sup>1</sup>, Bin Zhao <sup>1</sup>, Yonggen Li <sup>2</sup>,  
and Zhixian Gui <sup>1</sup>

<sup>1</sup>Key Laboratory of Exploration Technologies for Oil and Gas Resources (Yangtze University), Ministry of Education, Wuhan 430100, China

<sup>2</sup>Research Institute of Petroleum Exploration and Development, CNPC, Beijing 100083, China

Correspondence should be addressed to Gang Gao; [dragon\\_china316@163.com](mailto:dragon_china316@163.com)

Received 9 March 2022; Accepted 7 May 2022; Published 2 June 2022

Academic Editor: Zhenzhen Wang

Copyright © 2022 Tengfei Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Compressional and shear wave velocities ( $V_p$  and  $V_s$ , respectively) are important elastic parameters to predict reservoir parameters, such as lithology and hydrocarbons. Due to acquisition technologies and economy, the shear wave velocity is generally lacking. Over the last few years, some researchers proposed deep learning algorithms to predict the shear wave velocity using conventional logging data. However, these algorithms focus either on spatial feature extraction for different physical properties of rocks or on sequential feature extraction in the depth direction of rocks. Only focusing on feature extraction in a direction of rocks might lead to a decrease in prediction accuracy. Therefore, we propose a hybrid network of a two-dimensional convolutional neural network and the gated recurrent unit (2DCNN-GRU), which can establish more complex nonlinear relationships between the input and output data based on the spatial features extracted by 2DCNN and the sequential features extracted by GRU. In this study, two cases are used to validate the reliability and prediction accuracy of the proposed network. Comparing the prediction results of 2DCNN, GRU, and the proposed network, the proposed network shows better performance. Meanwhile, for improving the prediction accuracy of the proposed network, the relationship is analyzed between the prediction accuracy of the proposed network and the length of the input sample.

## 1. Introduction

Compressional and shear wave velocity ( $V_p$  and  $V_s$ , respectively) are very important parameters in hydrocarbon fields for characterizing and evaluating reservoir, identification of the pore types, and estimation of the dynamic properties of rocks [1–4]. Due to various reasons, shear wave velocity is generally lacking. Therefore, it is necessary to study a shear wave prediction method with high prediction accuracy and strong generalization ability to improve the reservoir prediction accuracy.

Currently, empirical regression methods, rock physics methods, and machine learning methods are the main methods for shear wave velocity prediction. Since empirical methods are the fastest and easiest to apply, linear or nonlinear empirical relationships between compression and shear

wave velocities have been proposed by various researchers [5–12]; however, they are constrained by site-specific and the rock type.

A variety of methods for predicting the shear wave velocity on the basis of rock physics have been proposed. These rock physics models focused on the modeling of the modulus of the rock matrix, dry rock, and saturated rock of the equivalent medium. In particular, for the modeling of the dry rock modulus of an equivalent medium, complex pore shapes were the research focus. Jørstad et al. [13] used both DEM and self-consistent approximation (SCA) for the shear wave velocity prediction in sandstones and concluded that the effective-medium theories were more accurate by comparing the results with those predicted from empirical regression methods. Xu and White [14] proposed a hybrid

approach to predict the shear wave velocity based on a shaly sandstone formation using a combination of the Kuster and Toksöz (KT) model [15] and the differential effective medium (DEM) model [16]. Based on the widely used Xu-White model and Gassmann's equations [17], Bai et al. [18] analyzed the influence of errors of input parameters of rock matrix, fluid inclusions, porosity, and aspect ratio (AR) on the prediction accuracy of shear wave velocity. Bai et al. [19] illustrated that a variable aspect ratio method in the Xu-White model was significantly improved. Liu et al. [20] proposed a differential Kuster-Toksöz (DKT) model to predict shear wave velocity and focused on the process in which the porosity with certain geometric shapes is gradually increased from zero to its final value to overcome a diluted concentrated pore of the KT model. Yang et al. [21] developed a revised Xu-White model and improved the estimated shear wave velocity for a calciferous sandy shale formation by considering the effect of the volume fraction of limestone. In the past ten years, with the development of unconventional oil and gas, rock physics models of complex reservoirs have been developed rapidly. Xu and Payne [22] extended the Xu-White model, originally designed for clastic rocks, to carbonate rocks and proposed a carbonate rock physics model with complex pore types. Zhang [23] established an anisotropic rock physics model to predict shear wave velocity, which was suitable for rocks with high-angle fractures. Based on the dual pore theory, an anisotropic rock physics model of tight oil sandstone was proposed, and the influence of clay content and type and pore connectivity and type on it was systematically studied by Huang et al. [24]. Assuming that shale is a laterally isotropic medium, Gui et al. [25] proposed a shear wave velocity prediction method that considered the microscopic characteristics of the rock. Liu et al. [26] proposed a method for predicting shear wave velocity suitable for organic-rich rocks. The accuracy of these methods for predicting shear wave velocity depended on the accurate calculation of reservoir geophysical parameters such as porosity, pore type, pore shape, mineral composition, and water saturation. However, these high-precision parameters are difficult to obtain, which increases complexity and indeterminacy of rock physics.

With the rapid development of software and hardware technology, some researchers use machine learning algorithms to predict shear wave velocities using logging data [27–33]. Deep learning developed from artificial neural network algorithms is a research hotspot in academic and industrial circles. Comparing with traditional shallow learning, deep learning improves the accuracy of prediction or classification by constructing many hidden layer machine models with complex function approximation and layer-by-layer feature transformation. The convolutional neural network (CNN) with spatial feature capture has achieved good results in different geophysical fields including interpretation of reservoir parameters from logging data [34, 35], seismic interpretation [36–38], and seismic inversion [39–42]. Based on the characteristics of logging data based on long-term dependencies, a long-short-term memory (LSTM) network was proposed to predict the shear wave velocity and its application in the identification

of geophysical parameters of complex reservoirs [43–46]. Comparing with LSTM which required a long training time, the gated recurrent unit (GRU) has the characteristics of faster speed and basically unchanged accuracy by simplifying the internal structure of LSTM [47]. Sun and Liu [48] proposed a GRU-based shear wave velocity prediction method. The above applications show that deep learning models have been successfully applied in the field of geophysics and are in rapid development.

Predicting shear wave velocity is essentially a typical regression problem in data processing. Compared with the empirical and rock physics methods, deep learning is better at handling regression problems by building many hidden-layer machine models with complex function approximations and layer-by-layer feature transformations. In order to fully mine the sequential features in the depth direction of rocks and spatial features of different physical properties of rocks, a hybrid network of two-dimensional convolutional neural network and gated recurrent unit (2DCNN-GRU) was constructed to predict the shear wave velocity using conventional logging data. This network takes full advantage of the powerful spatial features extracted by 2DCNN and the sequential features extracted by GRU. The process of predicting  $V_s$  using the 2DCNN-GRU hybrid network included data normalization, generating sample datasets, and constructing the 2DCNN-GRU hybrid network and its training and prediction. The  $V_s$  prediction of the two cases confirmed that the 2DCNN-GRU hybrid network was an accurate and reliable method of  $V_s$  prediction.

## 2. Methodology

*2.1. Convolutional Neural Network (CNN).* CNN, which is a feed-forward artificial neural network, is widely used in the field of vision and image. With the rapid development of deep learning, it has been proven to successfully solve various geological problems, such as fault recognition, reservoir prediction, lithofacies classification, and geological parameter inversion [39, 40]. CNN typically consists of the convolutional layers, the pooling layers, and the fully connected layers (Figure 1). In the convolutional layers, the data of the input layer is convolved with the convolutional kernels of the convolutional layers, which can mine the local features between the data. Its weight sharing feature greatly reduces the complexity of the network. The nonlinear relationship of the data is added through the activation function, usually the rectified linear function (“ReLU”) to avoid overfitting. When the data is passed into the convolution layer, the output features can be expressed as

$$y_j^l = \sigma \left( \sum \left( a_i^{l-1} w_{ij}^l \right) + b_j^l \right), \quad (1)$$

where  $y_j^l$  represents the  $j$ -th feature map of the  $l$ -th layer,  $a_i^{l-1}$  represents the  $i$ -th feature map of the previous layer,  $w_{ij}^l$  denotes the weight matrix of the  $l$ -th layer,  $b_j^l$  represents the corresponding bias term, and  $\sigma$  represents the activation function.

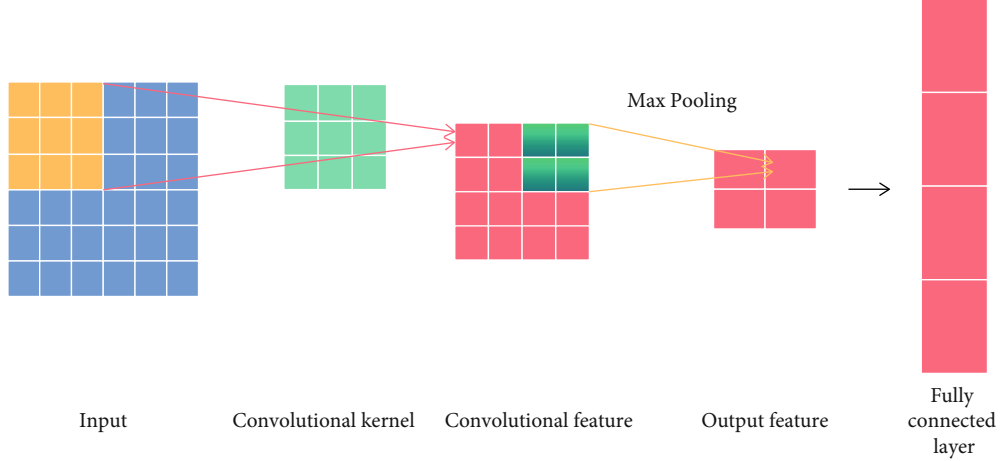


FIGURE 1: The structure of CNN and its unrolled network.

Considering the intrinsic relationship between shear wave velocity and various logging data, inspired by the extremely high feature extraction ability of a 2D convolutional neural network, the 2D convolution is used to extract more high-dimensional information and preserve topology as well as type and depth of log data.

**2.2. Gated Recurrent Unit (GRU).** A recurrent neural network (RNN) [49] is very effective for mining data with sequence characteristics. The hidden unit of the RNN with long-term sequence storage contains a loop that can combine the output at the current moment with the input at the next moment as the input at the next moment. Therefore, the RNN is particularly suitable for processing logging data that varies with sedimentary facies in the depth direction. However, due to the relatively simple structure of RNN, problems such as gradient disappearance or explosion are prone to occur in practical applications [50], and it can only hold memory functions for short-term data. In response to the above problems, the RNN variants LSTM [51] and GRU [52] were proposed. LSTM sets three-unit gates (forget gate, input gate, and output gate) to update the input data and obtain the ability of long-term memory data. However, the hidden unit of LSTM has many parameters, a complex structure, and a long training time. Compared with the LSTM network, the reset gate and update gate of GRU can reduce the network training parameters, shorten the training time, and improve the generalization ability of the network under the premise of ensuring the prediction accuracy [53] (Figure 2).

The structure of GRU combined the reset gate ( $r_t$ ), the update gate ( $z_t$ ), the output of the hidden state at  $t-1$  ( $h_{t-1}$ ), the output of the hidden state at  $t$  ( $h_t$ ), and the input at  $t$  ( $x_t$ ); they can be expressed as

$$r_t = \sigma(W_r[h_{t-1}, x_t] + b_r), \quad (2)$$

$$z_t = \sigma(W_z[h_{t-1}, x_t] + b_z), \quad (3)$$

$$\tilde{h}_t = \tanh(W_{\tilde{h}}[h_{t-1} \circ r_t, x_t] + b_{\tilde{h}}), \quad (4)$$

$$h_t = (1 - z_t) \circ h_{t-1} + z_t \circ \tilde{h}_t, \quad (5)$$

where  $W_r$ ,  $W_z$ ,  $W_{\tilde{h}}$  and  $b_r$ ,  $b_z$ ,  $b_{\tilde{h}}$  are the weights and biases, respectively, which are learned, “ $\sigma$ ” is the logistic function sigmoid,  $\tilde{h}_t$  is the new hidden state at  $t$ , “ $\circ$ ” represents the dot product, and “[ ]” represents that two vectors are connected. The reset gate controls how much information from the previous state is retained. On the other hand, the update gate is contrary to the function of the reset gate [52].

**2.3. Building a Hybrid Network of 2DCNN-GRU.** The shear wave velocity changing with time has a certain periodicity and has a nonlinear relationship with various factors such as density, porosity, Vp, and resistivity. Therefore, a 2DCNN-GRU hybrid network was proposed in this study to solve the problem of lack of shear wave velocity. The structure of 2DCNN uses the convolution kernels to fully excavate the high-dimensional features of different logging data, while the series data of time and depth cannot be accurately excavated. The structure of GRU has a strong ability to capture features in sequence data, while it is easy to introduce noise and lose some features during the calculation process [54], which is difficult to express the spatial features of the data and ultimately leads to deviations in the prediction results. To make up for the shortcomings of a separate network, the 2DCNN and the GRU are combined to make full use of the spatial convolution characteristics of 2DCNN and the sensitivity of GRU to sequence data to establish a nonlinear relationship between input and output. The structure of the 2DCNN-GRU hybrid network (Figure 3) and the flow chart of the shear wave velocity prediction (Figure 4) are as follows.

It can be seen clearly from Figure 3 that the first part of the 2DCNN-GRU is the CNN, which convolves with the input logging data through the convolution kernels to obtain the spatial characteristics and uses padding to fill it which can keep the size of the input sample unchanged after convolution. The second part of 2DCNN-GRU is the

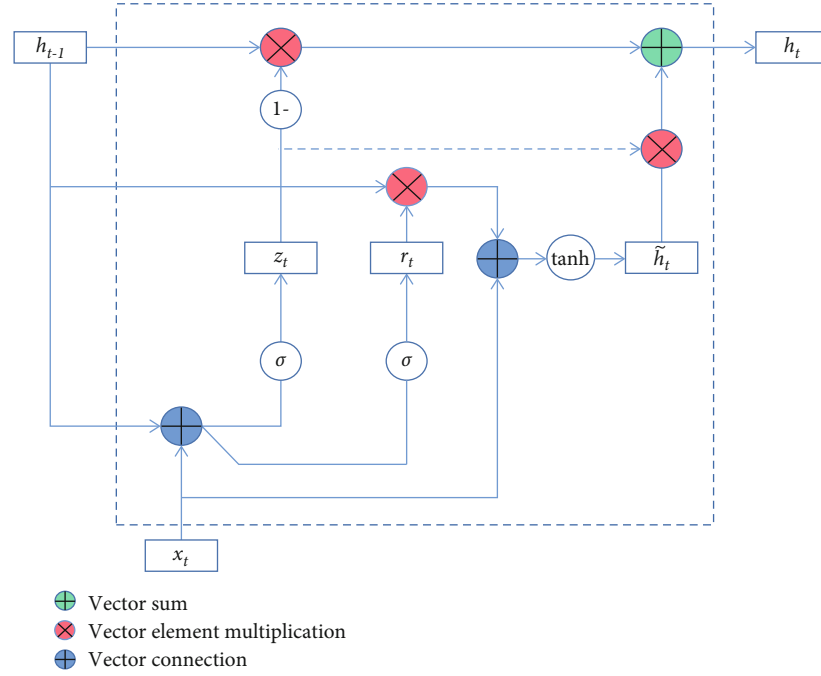


FIGURE 2: The internal structure of GRU.

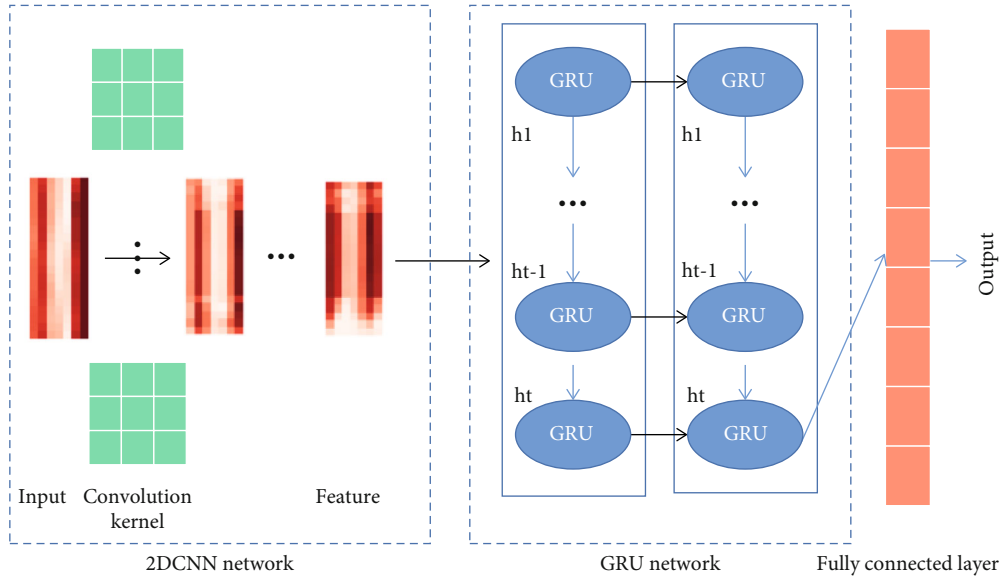


FIGURE 3: The structure of the 2DCNN-GRU hybrid network. A 2DCNN-GRU network is mainly composed of an input layer, a CNN Network, a GRU Network, and a fully connected layer.

GRU, which uses the spatial features extracted by the first part as the input of this layer. In particular, the first layer of GRU adopts the method of returning intermediate values. Both of the networks use activation functions to increase the nonlinearity of the network and use dropout to prevent overfitting and increase the generalization ability of the network. Finally, these features are taken into the fully connected layer to obtain the prediction of the shear wave velocity.

### 3. Prediction Framework Based on 2DCNN-GRU

Figure 4 shows the shear wave velocity prediction framework based on the 2DCNN-GRU hybrid network, and the specific process includes the following 4 parts.

*3.1. Feature Selection.* Deep learning networks are often used to deal with classification and regression problems.

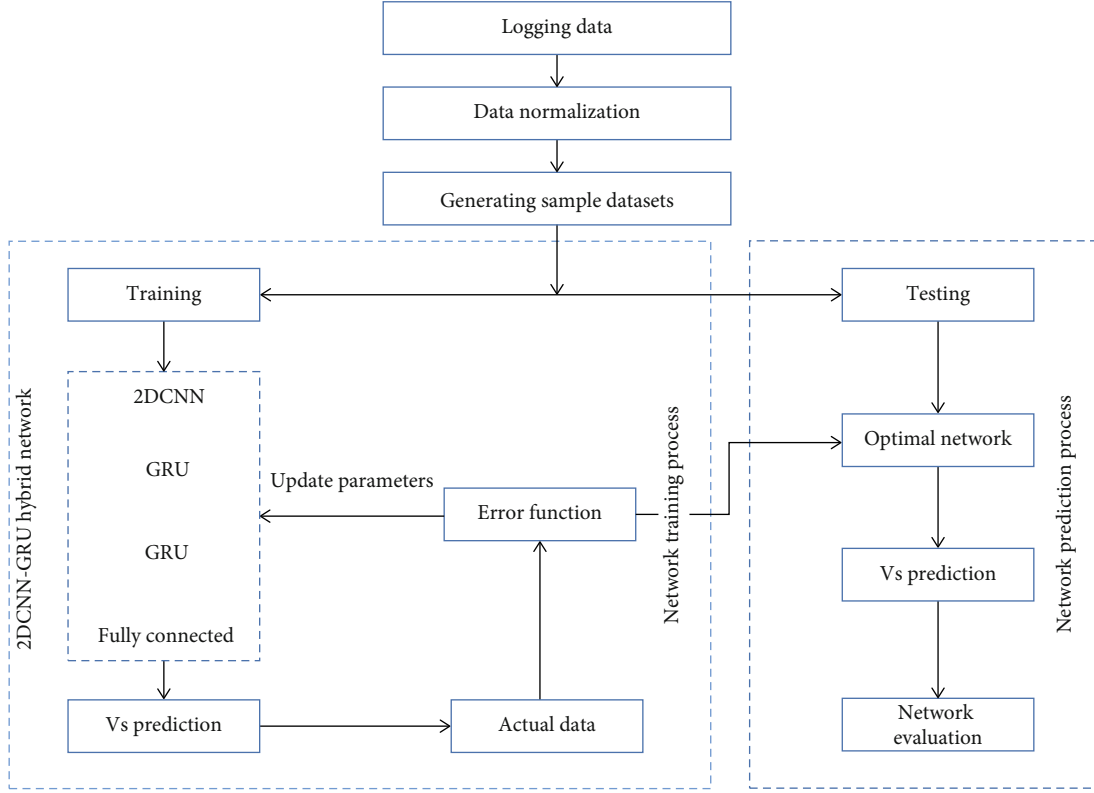


FIGURE 4: The VS prediction framework of the 2DCNN-GRU hybrid network.

Predicting shear wave velocity using conventional logging data is a typical regression problem. The assumptions that deal with regression problems often require correlations between input and output data. The correlation coefficient between logging data and shear wave velocity is shown in the cross-plot (Figure 5). The correlation coefficients between shear wave velocity ( $V_s$ ) and compression wave velocity ( $V_p$ ), neutron porosity (CNL), gamma ( $Gr$ ), shale volume ( $Sh$ ), density (RHOB), and water saturation ( $Sw$ ) are, respectively, 0.791, 0.576, 0.324, 0.300, 0.004, and 0.003. In these selected logs, the correlation between density,  $Sw$ , and shear wave velocity is small. In addition, the correlations between other logging data and shear wave velocity are all above 0.3, which satisfies the assumption that deep learning deals with regression problems.

**3.2. Data Normalization.** Since there are different degrees of differences between different logging data, it is necessary to normalize the logging data to speed up the training process, which can reduce the impact on the network accuracy [55]. The logging data have mapped the range of [0, 1] with the MinMaxScaler normalization method. The normalization formula can be expressed as

$$Y = \frac{X - X_{\min}}{X_{\max} - X_{\min}}, \quad (6)$$

where  $X_{\min}$  and  $X_{\max}$  are the minimum and maximum of a sequence  $X$ , respectively, and  $Y$  represents the result of normalization.

**3.3. Generating Sample Datasets.** The recurrent neural network has various network structures in dealing with time series problems, such as one-to-one, one-to-many, many-to-one, and many-to-many. Due to the depositional law of the subsurface in the depth direction, a many-to-one structure is adopted in the process of the prediction framework based on 2DCNN-GRU (Figure 6).

**3.4. Network Training and Evaluation.** To speed up the network training, the loss function mean square error (MSE) was used to calculate the gap between the predicted values and the true values in this study; at the same time, the Adaptive Moment Estimation [56] was used to back-propagate to update the weight parameters. The prediction performance of the network was evaluated by mean absolute error (MAE) and correlation coefficient ( $R^2$ ) in this study, which can be expressed as

$$MAE = \frac{1}{m} \sum_{i=1}^m |(y_i - \tilde{y}_i)|, \quad (7)$$

$$R^2 = \frac{\sum_{i=1}^m (\tilde{y}_i - \bar{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y}_i)^2}, \quad (8)$$

where  $m$  represents the number of samples,  $y_i$  represents the real value,  $\tilde{y}_i$  represents the predicted value, and  $\bar{y}_i$  represents the mean of samples.

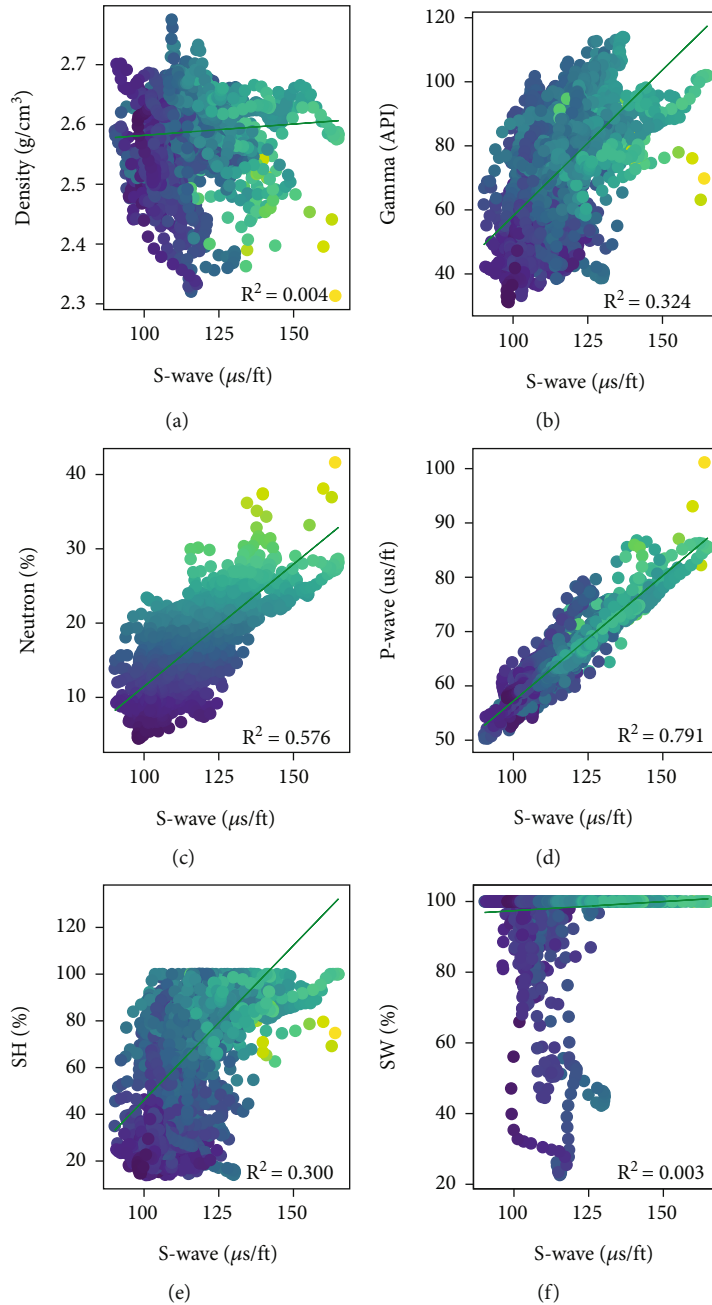


FIGURE 5: The cross-plot of conventional logging data and shear wave velocity. (a) RHOB-Vs; (b) Gr-Vs; (c) CNL-Vs; (d) Vp-Vs; (e) Sh-Vs; (f) Sw-Vs.  $R$  is the correlation coefficient between the two variables.

#### 4. Testing and Analysis

The logging data used in our study were derived from the Tarim Basin (Figure 7). The target layer is buried at a depth of about 5500 m, mainly composed of medium-fine sandstone, and the reservoir porosity is less than 10%, which are typical characteristics of deep tight sandstone. In order to verify the prediction accuracy of the 2DCNN-GRU hybrid network proposed and optimize its parameters, the network is trained with the logging data from 8 wells in a certain area, and tested with another 2 wells to verify its accuracy and generalization, and two cases are adopted using the optimi-

zation algorithm Adaptive Moment Estimation (Adam), the loss function Mean Squared Error (MSE), and Dropout to avoid overfitting of the network. In case I, the results of the 2DCNN, GRU, and 2DCNN-GRU hybrid network were analyzed and compared to verify the prediction accuracy of the 2DCNN-GRU hybrid network. In case II, the influence of sample length on the prediction accuracy of the 2DCNN-GRU hybrid network was analyzed.

*4.1. Case I.* Predicting shear wave velocity based on deep learning is essentially a sequence prediction problem. Fully considering the spatial and sequential features of the logging



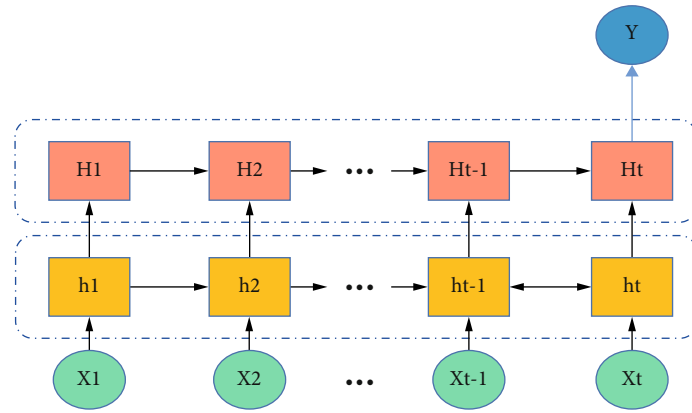


FIGURE 6: The structure of recurrent neural network (many-to-one).

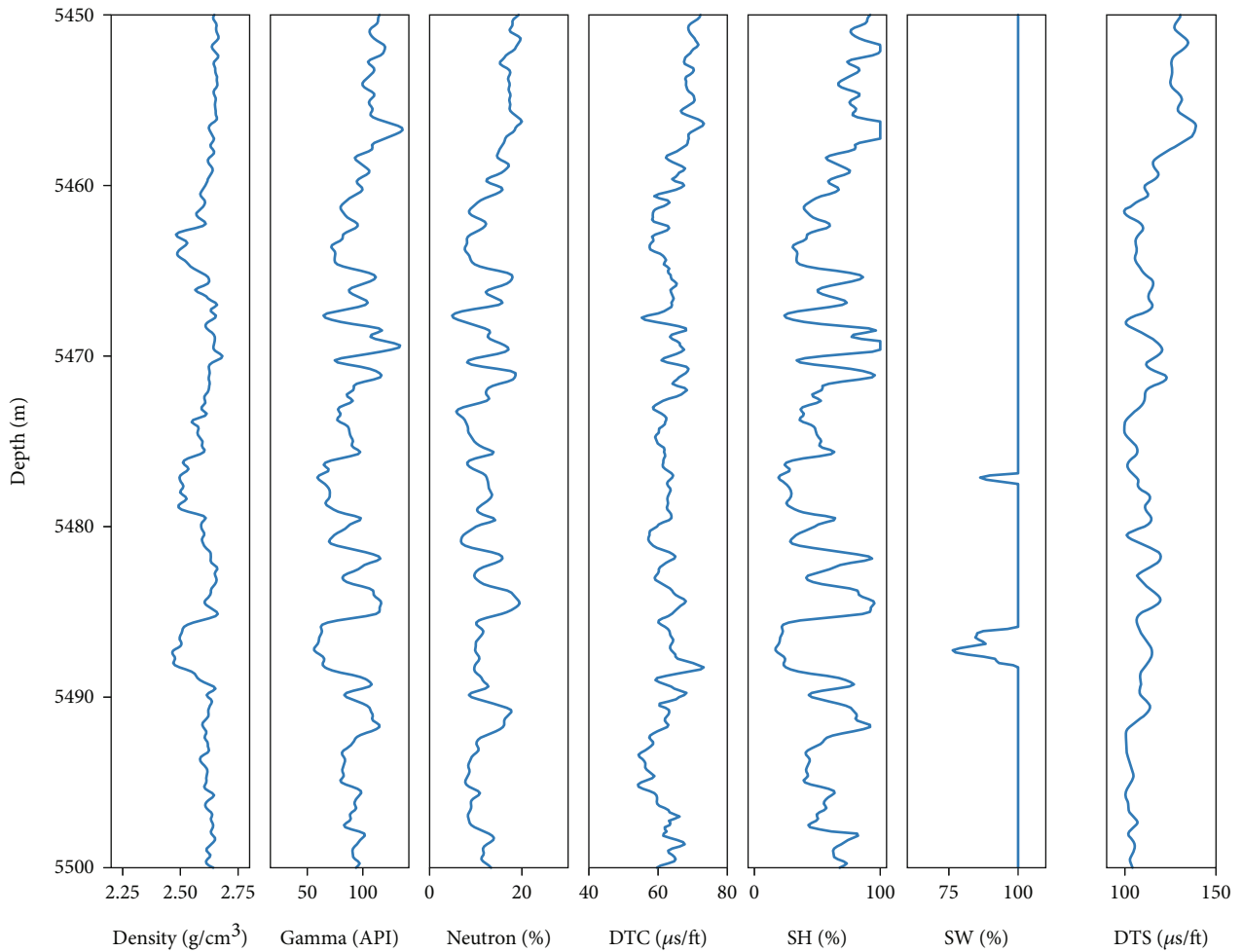


FIGURE 7: The visualization of logging data.

data, the 2DCNN-GRU hybrid network was established to predict the shear wave velocity, and its results were compared with those of the separate 2DCNN and GRU. The structures of the 2DCNN-GRU, 2CNN, and GRU networks are listed in Table 1.

Figure 8 shows the loss errors of 2DCNN-GRU, 2DCNN, and GRU networks. After a period of training, the loss values of all networks reach the minimum value and remain the same. It can be seen that the 2DCNN-GRU hybrid network has the lowest loss error in that the

TABLE 1: The structures of the 2DCNN-GRU, GRU, and 2DCNN.

Network name	Architecture	Kernel size of 2DConv layers	Number of conv filters and GRU units	Trainable parameters	Other parameters
2DCNN-GRU	Conv+GRU+GRU+dense	(15 * 3)	16 + 32 + 64	32097	Padding = "same" Return_sequences = true
2DCNN	Conv+conv+dense	(15 * 3) + (15 * 3)	16 + 32	30529	Padding = "same"
GRU	GRU+GRU+dense	0	32 + 64	22721	Return_sequences = true

Conv: 2D convolution+ReLU activation.

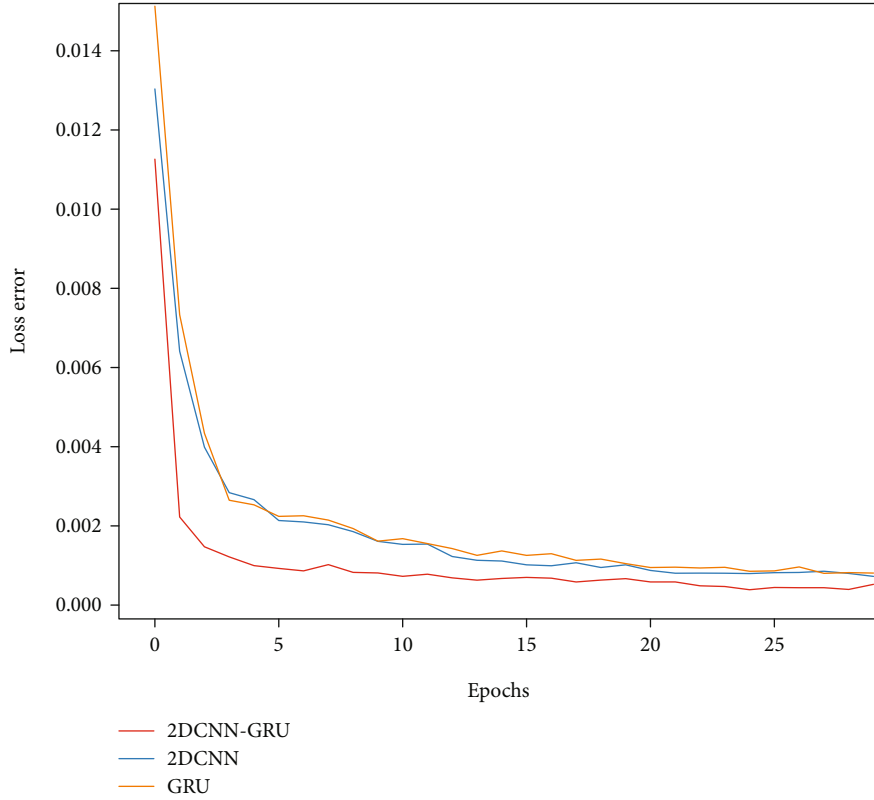


FIGURE 8: Loss error curves of 2DCNN-GRU, 2DCNN, and GRU networks.

shear wave velocity prediction values are closer to the true values than the other two networks. The logging data are convolved with the convolution kernels to extract the high-dimensional spatial features of the logging data. However, the logging data has time-series features in the depth direction, so the extracted spatial features are inputted into the GRU for time-series feature extraction, which can combine the spatial and time-series features of the logging data to predict shear wave velocity.

To compare the difference between the prediction performances of the three networks, the 2DCNN-GRU, 2DCNN, and GRU perform shear wave velocity prediction on the same test set and the experimental results are shown in Figure 9. Although the predicted values of the three

networks are generally similar in trend to the true values, the predicted values of the 2DCNN-GRU hybrid network are closer to the true values than those of the other two networks at 5570-5590 m.

As can be seen from Figure 10, the prediction of a single 2DCNN or GRU at this stage is always slightly higher than the true values, but the prediction effect of the 2DCNN-GRU that integrates spatiotemporal features has been greatly improved. That is to say, combining with the spatiotemporal features of the logging data can better predict shear wave velocity.

In order to analyze the prediction results of the three networks more precisely, the mean absolute error (MAE) and correlation coefficient ( $R^2$ ) were used to quantitatively evaluate the prediction accuracy of the three networks



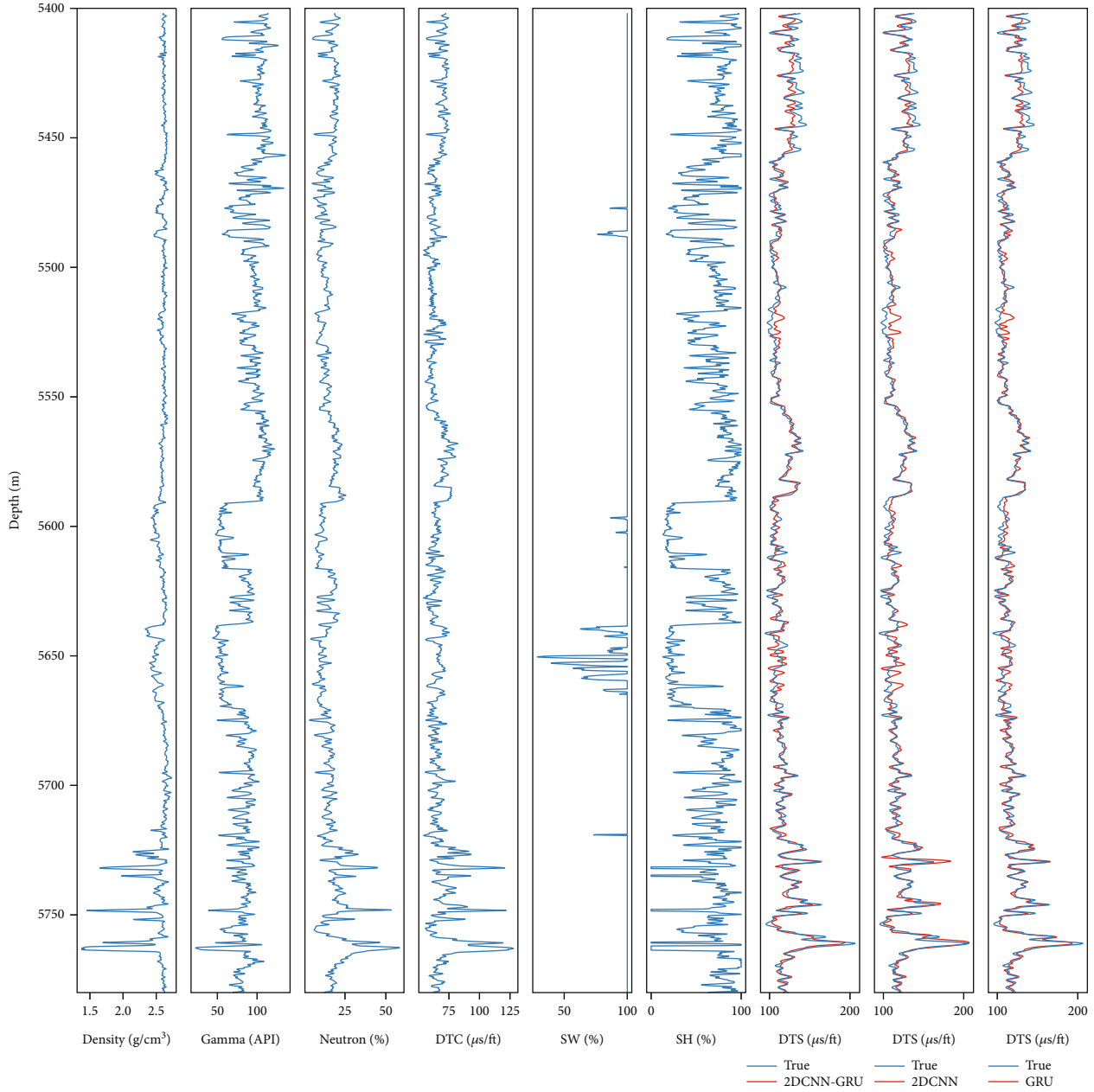


FIGURE 9: Prediction results of three networks in case I. The last three subgraphs indicate the final prediction results. The blue line represents the true values and the red line represents different prediction results based on three networks. The left one is the prediction result from 2DCNN-GRU, the middle one is that from 2DCNN, and the right one is that from GRU. To analyze the prediction performance of the three networks more clearly, the differences between the prediction of the three networks were compared in the form of local amplification (Figure 10).

(Figure 11). The correlation coefficient between the data predicted by the 2DCNN-GRU hybrid network and the true values was higher than that by the other two networks and was as high as 0.866. Moreover, the MAE of 2DCNN-GRU was lower than that of 2DCNN and GRU and was as low as 0.0165. Compared with the 2DCNN-GRU hybrid network, the reason for the low prediction accuracy of 2DCNN and GRU is that both of them only predict shear waves from single spatial or temporal features of logging data. Therefore,

the 2DCNN-GRU hybrid network that comprehensively considers spatiotemporal features improves the accuracy of shear wave velocity prediction.

4.2. Case II. Due to the depositional law of the subsurface in the vertical direction, there is a certain correlation between the sequence sampling points, which indicates that the length of the input sample plays an important role in the prediction of shear wave velocity by deep learning. In order to analyze

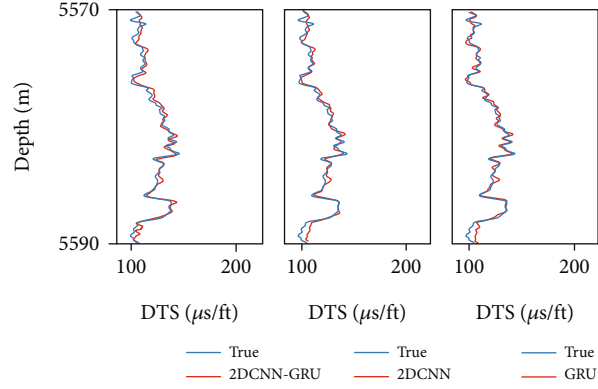


FIGURE 10: The local amplification comparison of 2DCNN-GRU, 2DCNN, and GRU.

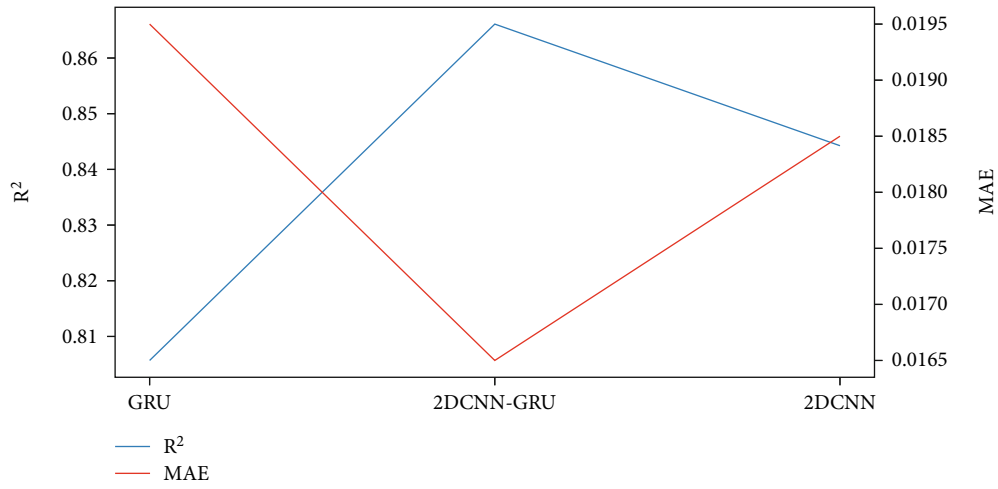


FIGURE 11: The evaluation metrics of 2DCNN-GRU, 2DCNN, and GRU networks.

TABLE 2: The structures of the 2DCNN-GRU hybrid network in 9 experiments.

Network name	Architecture	The length of the convolution kernel and sample	Number of conv filters and GRU units	Trainable parameters	Other parameters
2DCNN-GRU	Conv+GRU+GRU	3	32 + 32 + 64	40897	Padding = "same" Return_sequences = true
		10	32 + 32 + 64	41569	
		15	32 + 32 + 64	42049	
		20	32 + 32 + 64	42529	
		25	32 + 32 + 64	43009	
		30	32 + 32 + 64	43489	
		35	32 + 32 + 64	43969	
		40	32 + 32 + 64	44449	
		65	32 + 32 + 64	46849	

Conv: 2D convolution+ReLU activation+dropout.

the optimal sample length, the vertical length of the convolution kernel in the CNN was consistent with the input sample length, and the horizontal length was set to 3. A total of 9 experiments were performed with sample lengths set to 3,

10, 15, 20, 25, 30, 35, 40, and 65. The structure of the 2DCNN-GRU hybrid network is shown in Table 2.

Prediction results of the 2DCNN-GRU hybrid network are shown in Figure 12. It can be seen that under different

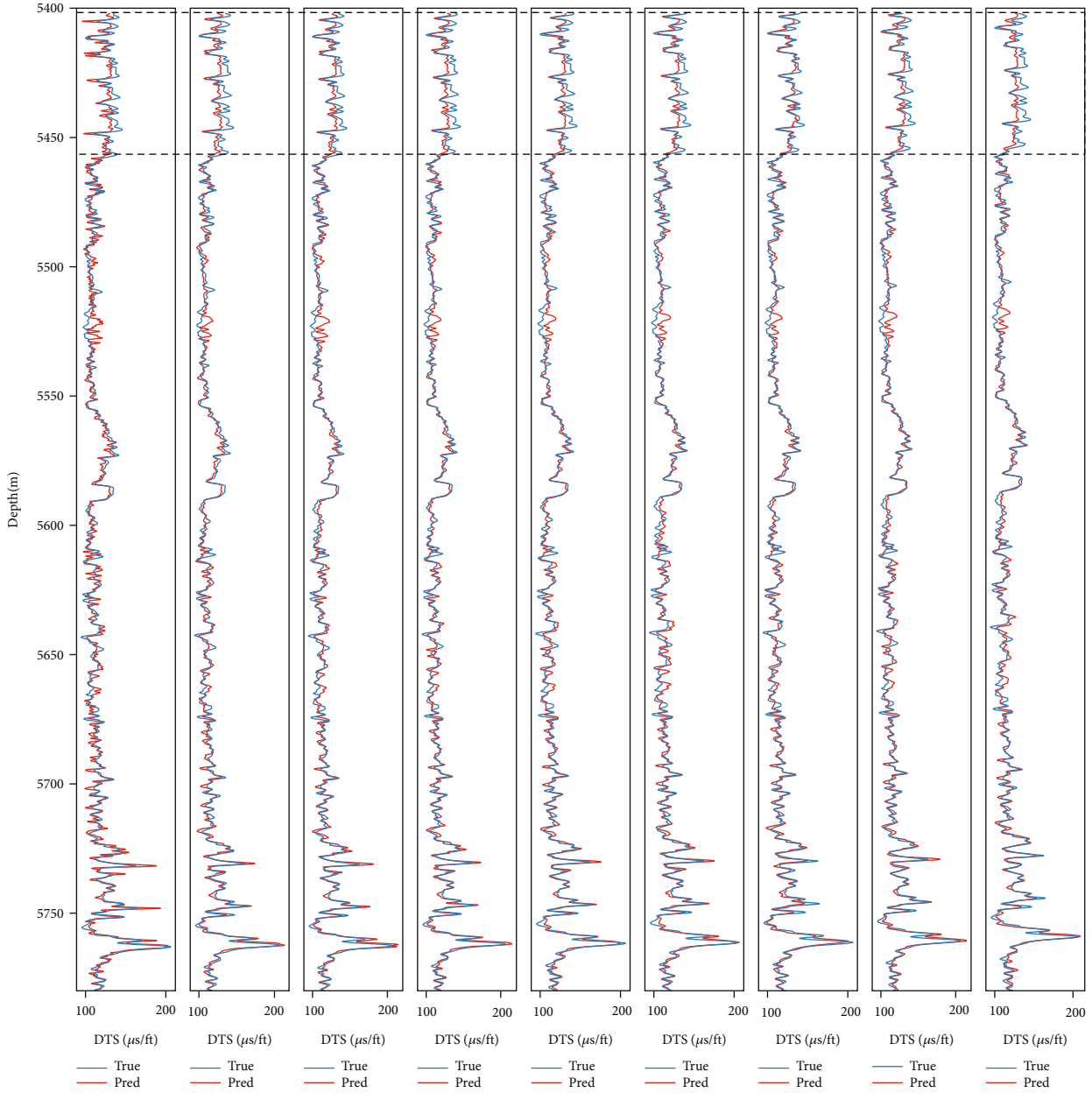


FIGURE 12: Nine experimental results based on the 2DCNN-GRU hybrid networks. The vertical lengths of the convolution kernels from left to right were set to 3, 10, 15, 20, 25, 30, 35, 40, and 65, respectively.

sample lengths, the prediction effects of 2DCN-GRU are different. With the continuous increase of the sample length, the prediction effect of the network becomes better first and then worse. When the input sample length is 35, the prediction effect of the network is the best. Compared with other sample lengths, the predicted values in the dotted box are the closest to the true values, but with the continuous increase of the sample length, the prediction effect of the network changes gets worse. This is because the convolution kernels of 2DCNN can only extract local features but cannot obtain global information as the sample length increases; at the same time, the GRU cannot effectively associate the input at the current moment with the histor-

ical data, which can make the prediction effect much worse than before.

Moreover, all experimental results had been evaluated by the correlation coefficient ( $R^2$ ) and mean absolute error (MAE) (Figure 13). It can be seen clearly that the correlation coefficient ( $R^2$ ) first increased and then decreased with the input sample length, while the trend of MAE was just the opposite during the whole testing process. When the length of the input sample reached 35, the correlation coefficient ( $R^2$ ) reached the highest value at 0.877; at the same time, the MAE reached the lowest value at 0.0160, which indicates that the predicted values were closest to the true values. At the same time, the correlation between the predicted values

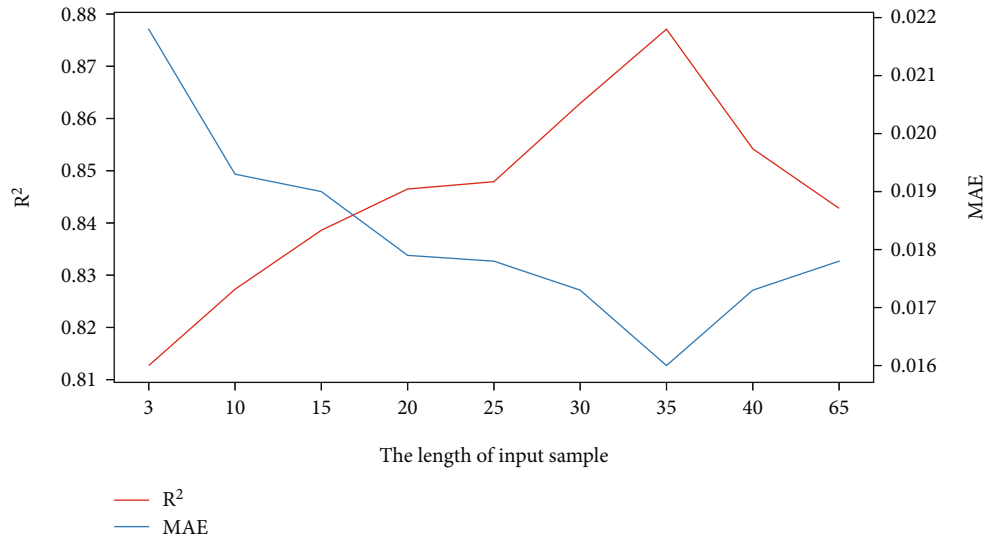


FIGURE 13:  $R^2$  and MAE comparison of nine experimental results based on 2DCNN-GRU hybrid network.

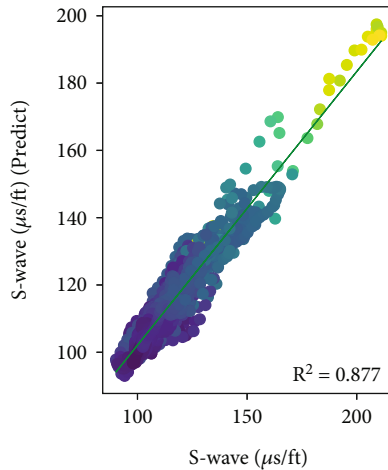


FIGURE 14: The correlation between shear wave velocity and predicted.

and the true values was analyzed in the form of a cross-plot (Figure 14), and the correlation reached 0.877. To sum up, the prediction accuracy of the 2DCNN-GRU hybrid network is affected by the length of the input samples. When the input sample length and convolution kernel length are 35, the prediction effect of the network is the best.

## 5. Conclusion

Considering the sequential features in the depth direction of rocks and spatial features of different physical properties of rocks, a new network 2DCNN-GRU hybrid network was proposed in this study, which can extract the spatial features of logging data from the 2DCNN and input them into the GRU to extract the temporal features, fully considering the temporal and spatial features of the logging data to predict Vs. In the case of I, the correlation coefficient, mean absolute error, and loss function of the evaluation parameters of

2DCNN-GRU were better than those of the separate 2DCNN and GRU, reaching 86.6, 0.165, and  $5.2375e-04$ , respectively; comparing the prediction results of 2DCNN, GRU, and 2DCNN-GRU, the prediction effect of 2DCNN-GRU is better than that of 2DCNN and GRU alone. In the case of II, the prediction accuracy of 2DCNN-GRU was affected by the input sample length. The prediction accuracy of 2DCNN-GRU first increased and then decreased with the input sample length. When the sample length was 35, the prediction accuracy of the network reached the highest. The experimental results show that the newly proposed 2DCNN-GRU hybrid network outperforms other networks in prediction performance. In addition, the 2DCNN-GRU hybrid network proposed in this study was a supervised machine learning whose prediction accuracy was dependent on training sample accuracy.

## Data Availability

The logging data used to support the findings of this study have not been made available because the data are shown in the article graphs.

## Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Authors' Contributions

Each author has contributed to the present paper. Tengfei Chen was responsible for drafting the article, programming, and performing the experiments; Gang Gao was responsible for conceiving the method, directing the experiments, and revising the article; Peng Wang was responsible for analysis of the data; Yonggen Li was responsible for interpretation

of the data; Bin Zhao supervised the experiments; and Zhixian Gui was responsible for revising the article.

## Acknowledgments

This work is jointly supported by the State Key Program of the National Natural Science Foundation of China (Grant No. 42030805) and Scientific Research & Technology Development Project of the China National Petroleum Corporation (Grant No. 2021DJ3704).

## References

- [1] R. Saltzer, C. Finn, and O. Burtz, "Predicting Vshale and porosity using cascaded seismic and rock physics inversion," *The Leading Edge*, vol. 24, no. 7, pp. 732–736, 2005.
- [2] K. Spikes, T. Mukerji, J. Dvorkin, and G. Mavko, "Probabilistic seismic inversion based on rock-physics models," *Geophysics*, vol. 72, no. 5, pp. R87–R97, 2007.
- [3] R. K. Sharma and S. Chopra, "Determination of lithology and brittleness of rocks with a new attribute," *The Leading Edge*, vol. 34, no. 5, pp. 554–564, 2015.
- [4] G. Gao, Y. H. Yang, B. Zhao, H. L. Duan, X. Y. Wang, and Y. Z. Wei, "A method for establishing and directly extracting sensitive identification factors of unconsolidated sandstones," *Oil Geophysical Prospecting*, vol. 54, no. 6, pp. 1329–1347, 2019.
- [5] J. P. Castagna, M. L. Batzle, and R. I. Eastwood, "Relationships between compressional-wave and shear-wave velocities in clastic silicate rocks," *Geophysics*, vol. 50, no. 4, pp. 571–581, 1985.
- [6] M. L. A. Han, A. N. De-hua, and D. Morgan, "Effects of porosity and clay content on wave velocities in sandstones," *Geophysics*, vol. 51, no. 11, pp. 2093–2107, 1986.
- [7] J. Dvorkin, M. Prasad, A. Sakai, and D. Lavoie, "Elasticity of marine sediments: rock physics modeling," *Geophysical Research Letters*, vol. 26, no. 12, pp. 1781–1784, 1999.
- [8] B. H. Russell, K. Hedlin, F. J. Hiltebeitel, and L. R. Lines, "Fluid-property discrimination with AVO: a Biot-Gassmann perspective," *Geophysics*, vol. 68, no. 1, pp. 29–39, 2003.
- [9] Z. Hossain, T. Mukerji, and I. L. Fabricius, "Vp-Vs relationship and amplitude variation with offset modelling of glauconitic greensand," *Geophysical Prospecting*, vol. 60, pp. 117–137, 2012.
- [10] W. M. Mabrouk and W. D. Pennington, "Compressional and shear wave velocity in terms of petrophysical parameters in clean formations," *Journal of Petroleum Science and Engineering*, vol. 65, pp. 62–66, 2009.
- [11] M. Ojha and K. Sain, "Velocity-porosity and velocity-density relationship for shallow sediments in the Kerala-Konkan basin of western Indian margin," *Journal of the Geological Society of India*, vol. 84, no. 2, pp. 187–191, 2014.
- [12] O. Oloruntobi, D. Onalo, S. Adedigba, L. James, and S. Butt, "Data-driven shear wave velocity prediction model for siliclastic rocks," *Journal of Petroleum Science and Engineering*, vol. 183, article 106293, 2019.
- [13] A. Jørstad, T. Mukerji, and G. Mavko, "Model-based shear-wave velocity estimation versus empirical regressions," *Geophysical Prospecting*, vol. 47, no. 5, pp. 785–797, 1999.
- [14] S. Xu and R. E. White, "A new velocity model for clay-sand mixtures1," *Geophysical Prospecting*, vol. 43, no. 1, pp. 91–118, 1995.
- [15] G. T. Kuster and M. N. Toksöz, "Velocity and attenuation of seismic waves in two-phase media: part I. Theoretical formulations," *Geophysics*, vol. 39, no. 5, pp. 587–606, 1974.
- [16] C. H. Chen and M. N. Toksöz, "Inversion of seismic velocities for the pore aspect ratio spectrum of a rock," *Journal of Geophysical Research*, vol. 84, no. B13, pp. 7533–7543, 1979.
- [17] F. Gassmann, "Elastic waves through a packing of spheres," *Geophysics*, vol. 16, no. 4, pp. 673–685, 1951.
- [18] J. Y. Bai, Z. X. Song, L. Su, W. G. Yang, and S. J. Li, "Error analysis of shear-velocity prediction by the Xu-White model," *Chinese Journal of Geophysics*, vol. 55, no. 2, pp. 589–595, 2012.
- [19] J. Y. Bai, C. Q. Yue, Y. Q. Liang et al., "Variable aspect ratio method in the Xu-White model for shear-wave velocity estimation," *Journal of Geophysics and Engineering*, vol. 10, no. 3, p. 035008, 2013.
- [20] Z. S. Liu, S. Sun, and D. Zan, "The differential Kuster–Toksoz rock physics model for predicting S-wave velocity," *Journal of Geophysics and Engineering*, vol. 12, pp. 839–848, 2015.
- [21] Y. Yang, X. Yin, G. Gao, Z. Gui, and B. Zhao, "Shear-wave velocity estimation for calciferous sandy shale formation," *Journal of Geophysics and Engineering*, vol. 16, no. 1, pp. 105–115, 2019.
- [22] S. Xu and M. A. Payne, "Modeling elastic properties in carbonate rocks," *The Leading Edge*, vol. 28, no. 1, pp. 66–74, 2009.
- [23] G. Z. Zhang, C. C. Li, X. Y. Yin, and J. Q. Zhang, "Corrected Xu-White model for carbonate S-wave velocity estimation," *Oil Geophysical Prospecting (in Chinese)*, vol. 32, no. 5, pp. 717–722, 2012.
- [24] L. Huang, R. R. Stewart, S. Samilk, and N. Dyaaur, "Fluid substitution effects on seismic anisotropy," *SEG Technical Program Expanded Abstracts*, vol. 64, pp. 330–335, 2014.
- [25] J. C. Gui, T. S. Ma, and P. Chen, "Rock physics modeling of transversely isotropic shale: an example of the Longmaxi formation in the Sichuan basin," *Chinese Journal of Geophysics*, vol. 63, no. 11, pp. 4188–4204, 2020.
- [26] Z. S. Liu, H. Lu, J. Z. Liu, Q. Z. Bao, and Z. Y. Wang, "Estimation of S-wave velocity from wire-line logs for organic-rich rocks," *Journal of Petroleum Science and Engineering*, vol. 205, p. 108928, 2021.
- [27] M. R. Rezaee, A. K. Ilkhchi, and A. Barabadi, "Prediction of shear wave velocity from petrophysical data utilizing intelligent systems: an example from a sandstone reservoir of Carnarvon Basin, Australia," *Journal of Petroleum Science & Engineering*, vol. 55, no. 3-4, pp. 201–212, 2007.
- [28] M. Asoodeh and P. Bagheripour, "Prediction of compressional, shear and Stoneley wave velocities from conventional well log data using a committee machine with intelligent systems," *Rock Mechanics and Rock Engineering*, vol. 45, no. 1, pp. 45–63, 2012.
- [29] H. Akhundi, M. Ghafoori, and G. R. Lashkaripour, "Prediction of shear wave velocity using artificial neural network technique, multiple regression and petrophysical data: a case study in Asmari reservoir (SW Iran)," *Open Journal of Geology*, vol. 4, no. 7, pp. 303–313, 2014.
- [30] P. Bagheripour, A. Gholami, M. Asoodeh, and M. Vaezzadeh-Asadi, "Support vector regression based determination of shear wave velocity," *Journal of Petroleum Science and Engineering*, vol. 125, pp. 95–99, 2015.
- [31] H. Fattahi, H. Nazari, and A. Molaghab, "Hybrid ANFIS with ant colony optimization algorithm for prediction of shear wave velocity from a carbonate reservoir in Iran,"



- International Journal of Mining and Geological Engineering*, vol. 50, pp. 231–238, 2016.
- [32] N. Weijun, L. Qi, and F. Tao, “Prediction of shear wave velocity in shale reservoir based on support vector machine,” *Journal of Xi’an Shiyou University (Natural Science)*, vol. 32, no. 4, pp. 46–49, 2017.
- [33] M. Anemangely, A. Ramezanzadeh, H. Amiri, and S. Hoseinpour, “Machine learning technique for the prediction of shear wave velocity using petrophysical logs,” *Journal of Petroleum Science and Engineering*, vol. 174, 2019.
- [34] P. Wang and S. Peng, “On a new method of estimating shear wave velocity from conventional well logs,” *Journal of Petroleum Science and Engineering*, vol. 180, pp. 105–123, 2019.
- [35] Q. Y. Ma, X. Zhang, C. L. Zhang, H. Zhou, and Z. Y. Wu, “Shear wave velocity prediction based on one-dimensional convolutional neural network,” *Lithologic Reservoirs*, vol. 33, no. 4, pp. 111–120, 2021.
- [36] W. Xiong, X. Ji, Y. Ma et al., “Seismic fault detection with convolutional neural network,” *Geophysics*, vol. 83, no. 5, pp. O97–O103, 2018.
- [37] Z. Zhong and T. R. Carr, “Geostatistical 3D geological model construction to estimate the capacity of commercial scale injection and storage of CO<sub>2</sub> in Jacksonburg-Stringtown oil field, West Virginia, USA,” *International Journal of Greenhouse Gas Control*, vol. 80, pp. 61–75, 2019.
- [38] R. H. Feng, T. M. Hansen, D. Grana, and N. Balling, “An unsupervised deep-learning method for porosity estimation based on post-stack seismic data,” *Geophysics*, vol. 85, no. 6, pp. M97–M105, 2020.
- [39] V. Das, A. Pollack, U. Wollner, and T. Mukerji, “Convolutional neural network for seismic impedance inversion,” *Geophysics*, vol. 84, no. 6, pp. R869–R880, 2019.
- [40] V. Das and T. Mukerji, “Petrophysical properties prediction from prestack seismic data using convolutional neural networks,” *Geophysics*, vol. 85, no. 5, pp. N41–N55, 2020.
- [41] M. Alfarraj and G. Alregib, “Semi-supervised learning for acoustic impedance inversion,” *SEG Technical Program Expanded Abstracts*, 2019.
- [42] A. Mustafa, M. Alfarraj, and G. Alregib, “Estimation of acoustic impedance from seismic data using temporal convolutional network,” *SEG Technical Program Expanded Abstracts*, 2019.
- [43] K. X. Wang, Q. H. Huang, and S. H. Wu, “Application of long short-term memory neural network in geoelectric field data processing,” *Chinese Journal of Geophysics*, vol. 63, pp. 3015–3024, 2020.
- [44] N. Pham, X. M. Wu, and E. Z. Naeini, “Missing well log prediction using convolutional long short-term memory network,” *Geophysics*, vol. 85, no. 4, pp. WA159–WA171, 2020.
- [45] Y. Zhang, H. Zhong, Z. Wu, H. Zhou, and Q. Ma, “Improvement of petrophysical workflow for shear wave velocity prediction based on machine learning methods for complex carbonate reservoirs,” *Journal of Petroleum Science and Engineering*, vol. 192, article 107234, 2020.
- [46] J. You, J. Cao, X. Wang, and W. Liu, “Shear wave velocity prediction based on LSTM and its application for morphology identification and saturation inversion of gas hydrate,” *Journal of Petroleum Science and Engineering*, vol. 205, p. 109027, 2021.
- [47] A. Graves, A. R. Mohamed, and G. Hinton, “Speech recognition with deep recurrent neural networks,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6645–6649, Vancouver, BC, Canada, 2013.
- [48] Y. H. Sun and Y. Liu, “Prediction of S-wave velocity based on GRU neural network,” *Oil Geophysical Prospecting*, vol. 55, no. 3, pp. 484–492, 2020.
- [49] M. Schuster and K. K. Paliwal, “Bidirectional recurrent neural networks,” *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673–2681, 1997.
- [50] Y. Bengio, P. Simard, and P. Frasconi, “Learning long-term dependencies with gradient descent is difficult,” *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, 1994.
- [51] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [52] K. Cho, B. V. Merriënboer, D. Bahdanau, and Y. Bengio, “On the properties of neural machine translation: encoder-decoder approaches,” *Computer Science*, vol. 1409, p. 1259, 2014.
- [53] H. Yan, G. Chen, Y. Tong, B. Ji, and B. C. Hu, “Human rehabilitation motion recognition based on pose estimation and GRU network,” *Computer Engineering*, vol. 47, no. 1, pp. 12–20, 2021.
- [54] J. Wang, Z. X. Gao, and Y. M. Zhu, “Prediction of Yellow River water quality based on CNN-LSTM model,” *People’s Yellow River*, vol. 43, no. 5, p. 96-99+109, 2021.
- [55] K. B. Zhou, J. Y. Zhang, Y. S. Ren, H. Zhen, and L. X. Zhao, “A gradient boosting decision tree algorithm combining synthetic minority over-sampling technique for lithology identification,” *Geophysics*, vol. 85, no. 4, pp. WA147–WA158, 2020.
- [56] D. Kingma and J. Ba, “Adam: a method for stochastic optimization,” 2014, <http://arxiv.org/abs/1412.6980>.