applied sciences

MDPI

Article

Deep Learning Logging Sedimentary Microfacies via Improved U-Net

Hanpeng Cai [1,*], Yongxiang Hu [1], Liyu Zhang [1], Mingjun Su [2], Cheng Yuan [2] and Yuting Zhao [1]

[1] School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu 611731, China; huyongxiang2019@outlook.com (Y.H.); zhangliyu0414@163.com (L.Z.); uestc_zhao@foxmail.com (Y.Z.)

[2] Research Institute of Petroleum Exploration and Development-Northwest (NWGI), PetroChina, Lanzhou 730020, China; sumj802@sina.com (M.S.); yuancheng0124@139.com (C.Y.)

* Correspondence: hanpengcai@uestc.edu.cn; Tel.: +86-13880725308

Abstract: Well logging data contain abundant information on stratigraphic sedimentology. Artificial identification is usually strongly subjective and time-consuming. Pattern recognition algorithms like SVM may not adequately capture the depth-related variations in logging curve shape. This paper defines logging sedimentary microfacies as unidirectional 2D image segmentation and builds an improved U-net model to meet the requirements of logging sedimentary microfacies acquaintance. The proposed model contains three characteristics: (1) It removes pooling layers to avoid the loss of spatial features; (2) it utilizes multi-scale convolution blocks for mining multi-scale spatial features in logging data; (3) a one dimensional convolution layer is added to achieve deep single-direction segmentation. In this model, a 2D image composed of several standardized logging curves is used as the network's input. In addition, we propose an effective data enhancement method and calculate the geometric feature attributes of well logging curves to reduce the complexity of the data characteristics. We tested the model on manually annotated validation datasets. Our method automatically measures fine sedimentary microfacies characteristics, improving the accuracy of sedimentary microfacies identification and achieving the desired result. Additionally, the model was tested on unlabeled actual logging data, which shows the generalizability of this deep learning method on different datasets.

Keywords: sedimentary microfacies; logging curves; deep learning; U-net

check for updates

Citation: Cai, H.; Hu, Y.; Zhang, L.; Su, M.; Yuan, C.; Zhao, Y. Deep Learning Logging Sedimentary Microfacies via Improved U-Net. *Appl. Sci.* **2023**, *13*, 10862. https://doi.org/10.3390/app131910862

Academic Editor: Andrea Prati

Received: 12 September 2023
Revised: 25 September 2023
Accepted: 27 September 2023
Published: 29 September 2023

1. Introduction

The study of sedimentary microfacies is beneficial for revealing the properties of reservoirs and providing practical guidance for oil and gas exploration. As core acquisition is difficult and expensive, because specialized equipment and technology are required, as well as a series of experiments and analyses. In addition, the sampling process may encounter some difficulties, such as difficulty drilling into the rock, core fracture and other problems, which will also increase the cost. So far, logging data are the only geological data that can realize continuous, high-precision and high-resolution formation analysis and interpretation of the whole well interval. The values of different types of logging curves vary with the sedimentary characteristics of the rock mass, such as source, lithology, grain size and structure, and these geological characteristics will be reflected in the morphological characteristics of the logging curves. Therefore, well logging data have become an important source of information for sedimentary microfacies analysis. The amplitude, shape and other features of logging data can be utilized to identify related lithofacies and sedimentary environments. Traditionally, geologists have relied on subjective interpretations of logging curves for sedimentary microfacies identification [1,2]. For example, Chen et al. [3] used the morphological change trend of a natural gamma logging curve to analyze the sedimentary microfacies of the Taiwan Hill fan delta. This method is highly subjective, inaccurate, time-consuming and laborious, and consumes a great deal of human and material resources. To define sedimentary microfacies division criteria more accurately and conveniently, identification of sedimentary microfacies has gradually transitioned from traditional qualitative identification to quantitative identification. Quantitative analysis is more efficient and accurate than qualitative analysis, as it applies complex mathematical operations to sedimentary facies in logging data and constructs specific parameters as the basis for judging sedimentary microfacies. Numerous scholars have mathematically described variations in logging curve characteristics and constructed specific parameter indicators as the basis for identifying sedimentary microfacies [4]. For example, Deng [5] used the analysis of variance to automatically stratify the curve initially, then carried out fine stratification through further calculations, and then obtained parameter values related to log curve morphology in each layer as the basis for distinguishing sedimentary microfacies. Bayesian criteria [5,6], linear discriminant analysis [6], fuzzy logic [7], K-nearest neighbor algorithms (KNN) [8] and support vector machines (SVM) [9,10], as well as other mathematical statistics or machine learning methods, have been successively applied to sedimentary microfacies identification based on logging curves. These methods involve the pre-layering of logging curves and manual selection of criteria and parameters to extract curve features. This process is complex and subjective, demanding a high level of expertise from layering personnel.

The volume of well logging data is large and has multi-source heterogeneity [11]. The use of logging curves for identifying sedimentary microfacies in the logging process faces difficulties such as ambiguity and uncertainty. In recent years, deep learning composed of neural networks that can autonomously learn and extract features between data has provided a new way to identify sedimentary microfacies in logging. The deep neural network autonomously learns curve features, which maximizes the characteristics of the logging data themselves. Networks such as SOM [12,13] and ANN [14,15] have already been preliminarily applied in the identification of sedimentary microfacies. These methods still require the logging curves to be layered in advance, so that the single-layer logging data and feature parameters can be input in vector form. Although these traditional neural networks can explore complex nonlinear mapping relationships between logging data, they only construct mapping relationships between logging depth sampling points and sampling points, ignoring the local curve structure and morphological characteristics of logging data with changes in formation depth.

Sedimentary microfacies reflect the progressive sedimentation process of geological periods. Through well logging, the characteristics of rock strata such as resistivity and radioactivity are obtained and the corresponding changes in the values of these characteristics at the longitudinal depth constitute the logging curves. Logging phase labeling is obtained according to the morphological characteristics of the logging curve, then the sedimentary microfacies information can be obtained by analyzing the changes in the lithologic combination, composition, particle size, mud content, porosity and other characteristics of the corresponding sediment at the longitudinal depth. As a response to the progressive sedimentation process, the local curve structure and morphological characteristics of logging curves with depth changes are key to identifying sedimentary microfacies. By constructing a logging dataset as a two-dimensional image with depth as the vertical axis and different types of logging parameters (such as longitudinal and transverse wave velocities and density) as the horizontal axis [16], the complex rules of sedimentary microfacies in logging curves can be excavated by employing deep learning methods on the strength of image semantic segmentation. Fully Convolutional Networks (FCNs) are one of the most commonly used deep learning architectures for image semantic segmentation. Using FCNs for the identification of sedimentary microfacies in logging curves can not only consider the local or overall curve structure and morphological characteristics of logging curves with depth changes, but also perform layering, feature extraction and classification simultaneously. The U-net network is a network model developed on the basis of the FCN structure. Compared with other network models, its major feature is that it fuses the up-sampled output with the corresponding down-sampled output by means of "skip connection", so that the network can spread the context information to the layer with higher resolution, and the segmentation effect is often very good. Its unique network architecture and data augmentation method have alleviated to some extent the requirement for a large number of sample inputs in deep learning, which is beneficial for solving the problem that logging data samples are insufficient due to the limitation of drilling costs and cannot be applied to fully convolutional neural networks. The U-net network presents great advantages in processing image segmentation tasks of small datasets, and has been widely applied in remote sensing image analysis [17], medical image segmentation [18,19] and other tasks. Unlike the analysis of medical images and remote sensing images, identifying sedimentary microfacies based on logging curves only requires image segmentation in the depth direction. The traditional sedimentary microfacies identification methods often ignore the change in the morphological characteristics of the logging curve in the depth direction. The shape of the logging curve is "vectorial" because of the time sequence of sedimentation. Rotation, translation and inversion of the logging curve image will destroy the meaning of the sedimentary relationship represented by the curve itself. Therefore, deep learning network algorithms applied to sedimentary microfacies division in logging data need to maintain scale invariance of feature extraction while discarding rotation invariance and flip invariance of feature extraction.

We propose a U-net framework-based deep learning logging sedimentary microfacies identification method based on the ideas of deep learning and image semantic segmentation. According to the U-net framework, the pooling layer is removed, the multi-scale convolutional block [20] is introduced, and the one-dimensional convolutional layer is added to construct a deep learning model that can realize multi-scale feature extraction and unidirectional segmentation and meet the requirements of logging sedimentary microfacies identification. We then constructed a 2D image from multiple well logs reflecting sedimentary facies as inputted to the constructed deep learning network. To further improve the performance of the constructed network model for the identification of logging sedimentary microfacies, we transformed the input data in two ways: (1) Affected by the length of the sedimentary cycle and the shape of the sedimentary basin, and the formation of the same sedimentary environment, there are differences in the stratum thickness. According to the time-series sedimentation characteristics contained in the well logging data, the structural shape of the well logging curve is "vectorial" and "multi-scale", so the curve stretching transformation was used to achieve data augmentation. (2) Considering that some salient features of well logging curves can be extracted quickly and effectively based on the physical model, by calculating the slope value, azimuth value and variance features of the original logging data, the complexity of the features contained in the data is reduced and the classification is improved. We applied our method to actual logging data, and confirmed that the proposed method can enhance the accuracy of sedimentary microfacies identification from logging data and obtain expected results. We tested the adjustable parameters in the network model constructed in this paper from qualitative and quantitative perspectives, and compared it with an Artificial Neural Network (ANN), Support Vector Machine (SVM) and original U-net.

2. Materials and Methods

2.1. U-Net Architecture

The U-net network consists of two parts: the compression path and the expansion path. The two parts are symmetrically distributed to form a U shape. Among them, the compression path is called the downsampling process, which is composed of a convolutional layer, a ReLU corrected linear unit and a maximum pooling, which mainly performs the role of feature extraction. The convolution layer is composed of convolution kernel and convolution layer parameters, which have the function of feature extraction. After every 3×3 convolution, ReLU is used for de-linearization, so that later convolution can effectively extract features. The pooling layer achieves feature dimension reduction and field of view expansion at the cost of spatial information loss, which is generally carried out after the convolution layer. The maximum pooling layer is used for feature selection and information filtering of the input image to retain the most important features of the image. The expansion path is also called the upsampling process, which consists of the deconvolution layer, splicing layer and convolution layer, which mainly play the role of restoring image information. The U-net network architecture is shown in Figure 1. The architecture replicates the encoder layers and connects to the corresponding layers in the decoder. This type of connection that reuses the network properties of the encoder is often referred to as a "skip connection". The upsampled feature map has a low resolution, while the corresponding downsampled feature map has a high resolution. Connecting the two can enhance the local features of the image and make the recovered image have a higher resolution. The advantage of "skip connection" in image segmentation is that it can provide multi-scale and multi-level information, so as to obtain finer segmentation results.

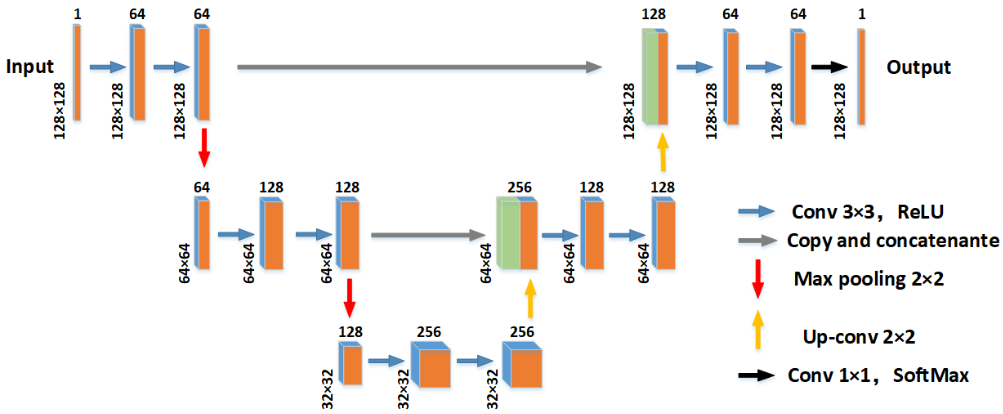


Figure 1. The network architecture of U-net.

2.2. Improved U-Net Architecture

The two-dimensional images constructed from well logging data may have the same texture characteristics after mirror inversion in the formation depth direction, but the sedimentary characteristics they represent are completely different. For example, the value of the GR curve decreases first and then gradually increases from bottom to top, forming a bell shape. The sediment grain size has a combination of fine–coarse–fine characteristics, indicating that the hydrodynamic force changes from weak to strong and then weak. If the GR curve is reversed and becomes funnel-shaped, it represents a completely different regressive depositional environment. The shapes are similar, but the direction is different, as shown in Figure 2. Since the sedimentation is time-sequential, the logging data have a certain time-series feature, as the geophysical response of the sedimentary body. As the shape of the logging curve has "vectoriality", log curves with similar shapes but different directions represent different sedimentary facies features, so the morphological and structural features of the log curves are the key to identifying sedimentary microfacies. However, the continuous pooling layer in the U-net architecture increases the receptive field at the cost of losing spatial information, which may lead to the loss of the structural feature information of the log curve. In addition, the maximum pooling can maintain the scale invariance, rotation invariance and translation invariance of the features to a certain extent, which do not match the actual features of the logging data.

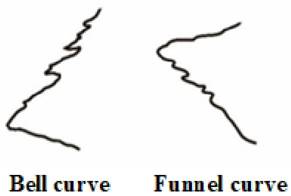


Figure 2. The bell shape (**left**) and funnel shape (**right**) of the GR curve.

Thanks to the particularity and complexity of the formation of geological bodies, influenced by the length of sedimentary cycle and the shape of sedimentary basins, the thickness of geological bodies formed in the same depositional environment can be different. This is due to periodic changes in the depositional environment caused by changes in crustal movement, sea level, climate and sediment sources. Some periods are longer and some are shorter, resulting in variations in sediment thickness in depositional cycles. In addition, in different geographical locations and under the same depositional environment, affected by depositional events and the shape of depositional basins, there are certain differences in the depositional thickness. In the form of well logging curves, this difference is manifested as the difference in the length and amplitude of the curves, but the overall morphological characteristics are similar, which means they are "multi-scale". However, in the original U-net architecture, the convolution operation in each layer of the downsampling and upsampling parts is composed of two 3 × 3 convolutions, which cannot better extract important features of different scales on the logging curve information. In addition, the identification of sedimentary microfacies based on images composed of well logging curves only requires image segmentation in the depth direction. However, the resulting map of semantic segmentation using the U-net network is the result of the classification of each pixel in the image, so it cannot be guaranteed that the values of different logging types corresponding to the same depth have the same category.

Therefore, while considering the characteristics of logging curves for sedimentary microfacies identification, we constructed a deep learning model that is more suitable for logging sedimentary microfacies identification on the basis of the original U-net framework. The characteristics of building a deep learning model are as follows:

(1) Considering that the rotation invariance and inversion invariance of feature extraction should be discarded when logging data are applied to deep learning network algorithms. So as to reduce the loss of spatial feature while retaining feature dimensionality reduction, expanding the field of vision and decreasing network parameters, we removed all the pooling layer from the U-net architecture, and replaced it with a 3 × 3 convolution operation with a step size of 2 to form a compression layer, as shown in Figure 3. The compression layer can reduce the feature dimension of the image and cut down the loss of spatial characteristics.

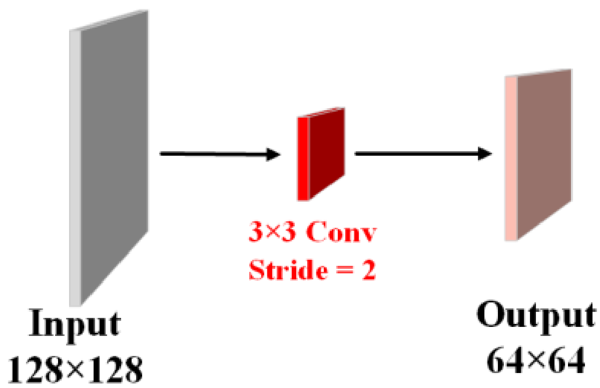


Figure 3. The schematic diagram of the compress layer.

(2) To better learn and dig out more useful features from curves of different scales, a multiscale convolution block is added to the network, as shown in Figure 4. The multiscale convolution block can work 3 × 3, 5 × 5 and 7 × 7 convolution operations in parallel to extract these spatial features of different scales of the logging data, and splice the feature maps obtained from different convolution operations. In addition to this, other connection and a 1 × 1 convolution are aggrandized to the network to obtain more spatial features and improve the performance of network feature extraction [21].

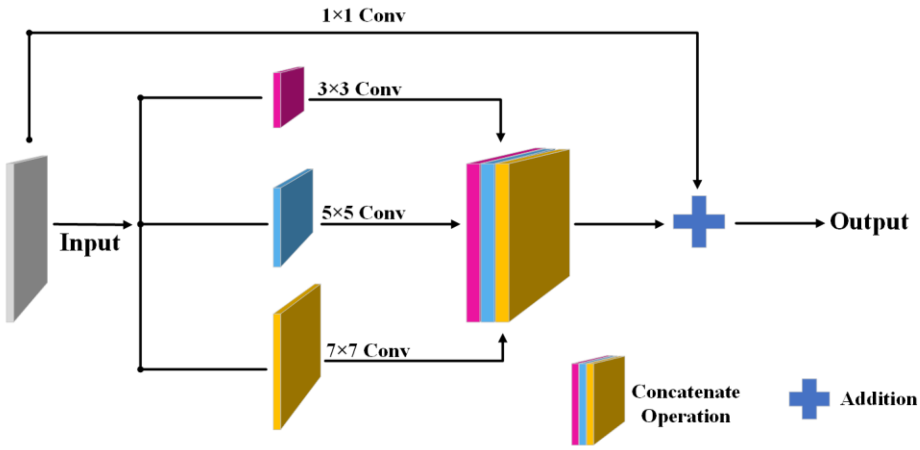


Figure 4. The schematic diagram of the multi-scale convolution block.

(3) To achieve unidirectional segmentation of logging sedimentary microfacies only in the depth direction, we must avoid different classification results for different logging curve pixel points from U-net at the same depth. We added a one-dimensional convolutional layer before the network output to comprehensively analyze the eigenvalues of various curves at the same depth, and transform 2D image features into 1D feature vectors in the depth direction, as shown in Figure 5. Finally, we classified each eigenvalue of the vector, so as to achieved the one-way segmentation of sedimentary microfacies.

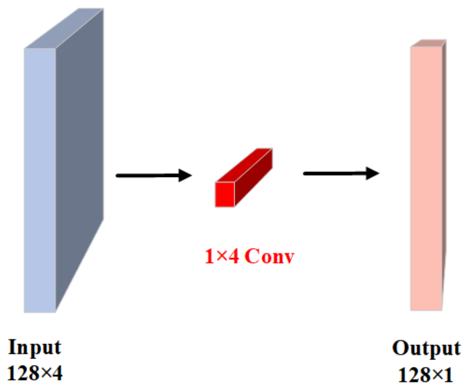


Figure 5. The schematic diagram of the 1D convolutional layer.

Based on the above method, the model used for sedimentary microfacies identification is shown in Figure 6. The input log image, formed from various types of logging curves, is fed back into a network consisting of shrinking and expanded paths.

Shrinking path: Each step consists of a multi-scale convolutional block and compression layer, which plays the role of feature extraction. The multiscale convolution block comprises 1×1, 3×3, 5×5 and 7×7 convolution operations and other connections, which can extract feature information from images of different scales. After each convolution operation, there is a ReLU unit to delinearize, so that the subsequent convolution operators can effectively extract features. A 3×3 convolution operation with a step size of 2 is applied to the compression layer, which can reduce the size of the image while extracting features, expand the receptive field, and reduce the loss of spatial information.

Expanded path: Each step includes a 2×2 deconvolution operation, a multi-scale convolution block, and a concatenate layer, which play the role of restoring image. Additionally, a 1D convolutional layer is added before the final output. The deconvolution layer is used for image restoration to achieve upsampling of feature maps. Each 2×2 deconvolution operation doubles the image size, so that layer-by-layer upsampling can restore the original input image. The concatenate layer stitches the upsampled low-resolution feature map with the corresponding downsampled high-resolution feature map to heighten the local features of the image. The multi-scale convolution block also plays an important role in feature extraction. The 1D convolutional layer converts the 2D image features into 1D feature vectors in the depth direction before the output layer, and comprehensively analyzes the characteristic values of various curves at the same depth. Finally, we used a 1×1 convolution and SoftMax function to classify each feature of the 1D feature vector to obtain the final recognition result of logging sedimentary microfacies.

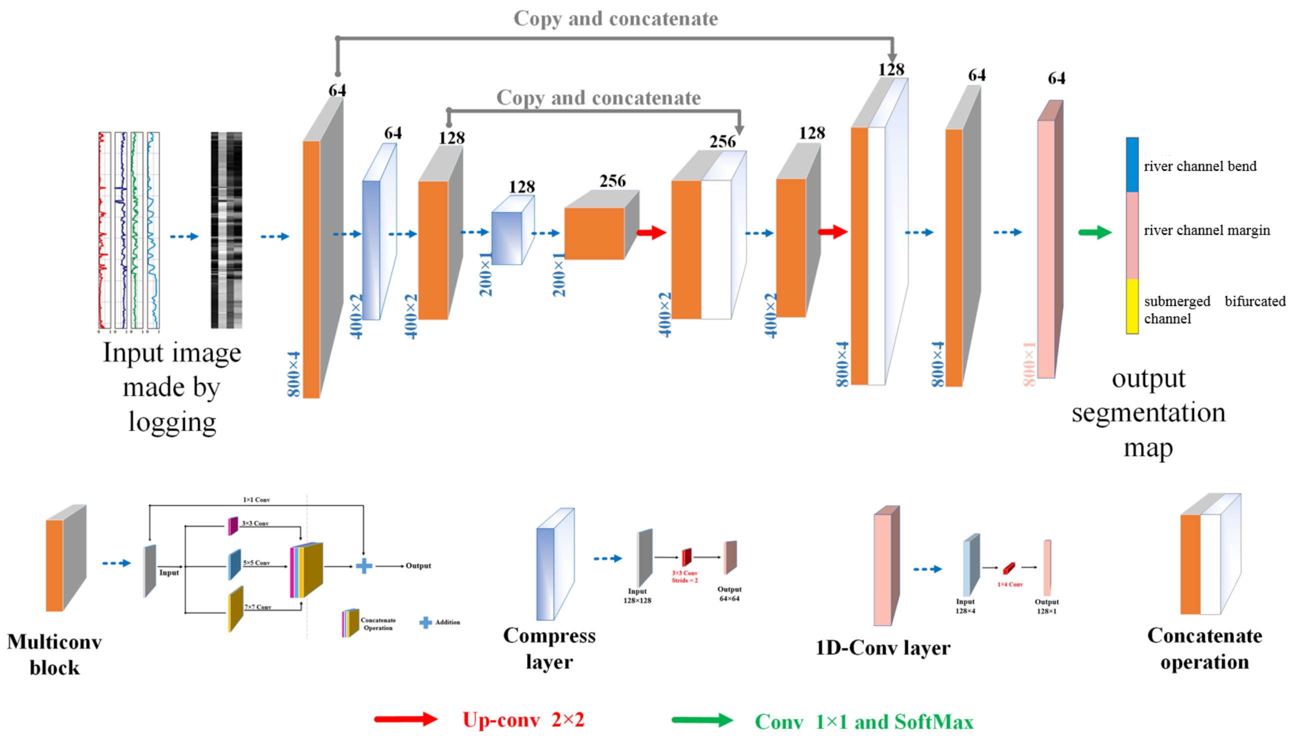


Figure 6. The network architecture of improved U-net.

The advantage of the model is that it can realize the identification of sedimentary microfacies in logging curves, and can basically identify each sedimentary microfacies correctly. Compared with the traditional algorithm, this method has higher accuracy and universality in identifying sedimentary microfacies. The disadvantage is that the model is more sensitive to the data and will further divide the places where the data changes. Therefore, compared with the real label, the model will divide the sedimentary microfacies more carefully, and the local location of a certain sedimentary microphase will be divided into other types, resulting in misclassification.

2.3. Overall Implementation Process

In terms of the geological background characteristics of the survey area and logging core data, gamma-ray (GR), spontaneous potential (SP), density (DEN) and other logging curves that can reflect sediment lithology and porosity, which are beneficial to the division of sedimentary facies. SOM can use well logging data including compensated neutron log (CNL), DEN, GR and other sample values for lithofacies analysis [22]. Firstly, the logging data are analyzed using SOM clustering to initially determine the lithologic characteristics of each depth sampling. The numerical representation is used to characterize a certain lithological characteristic with a specific numerical value, and the lithological characteristic curve of the logging curve is formed. They are then arranged in columns to form a 2D matrix. The logging image is shown in Figure 7. Then, according to the experience and knowledge of experts, manually mark the corresponding labels on the logging images to form a one-dimensional vector label map that changes with depth, as shown in Figure 8. Perform data enhancement processing and attribute extension on the logging images, and cut these images into 2D matrices of the same size as the network input. The built model is trained with different parameters to obtain the optimal model and realize the segmentation of sedimentary microfacies.

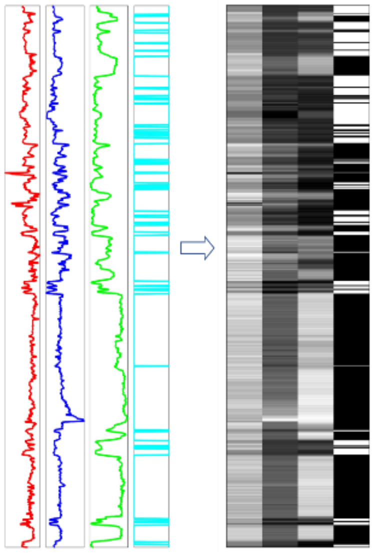


Figure 7. Logging curve data values are arranged in columns to form a logging image. GR, SP, DEN and Lithofacies curves in the left image are in red, blue, green and light blue, respectively. Different colors in the right image Indicates the relative value of the curve amplitude.

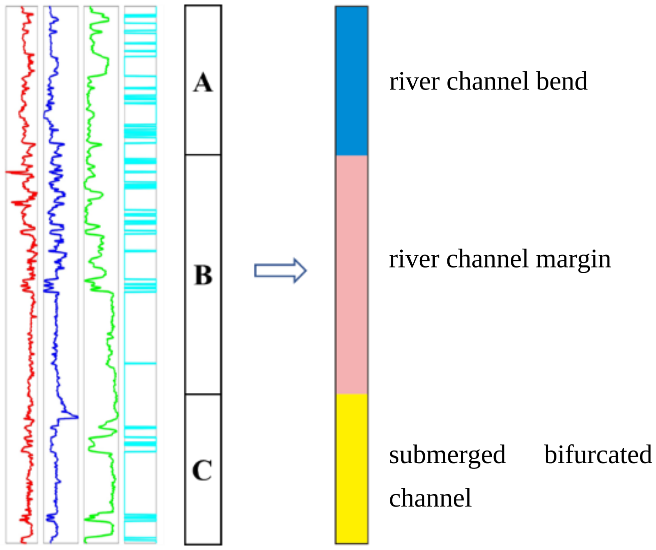


Figure 8. Tagging labels for well log images. The meaning of the curve color is the same as in Figure 7.

2.3.1. Data Preprocessing

The input data consist of various types of log curve values arranged in columns. It is necessary to normalize the logging data separately, column by column, to reduce the influence of the data dimension and facilitate the network solution. The normalization method can maintain the morphological characteristics of the log curve changing with the depth direction, as well as the relationship between the relative amplitude changing with the depth, while having no effect on the classification task.

The minimum–maximum (Min–Max) method is a data normalization method commonly used in machine learning which can scale data to the [0, 1] range. Each log can be viewed as a 1D vector composed of N values $X = (x_1, x_2, \ldots, x_n)$. The normalized data can be calculated with the following equation:

$$x_i{}' = \frac{x_i - \min(X)}{\max(X) - \min(X)} \tag{1}$$

here, $x_i{}'$ represents the normalized result of x_i, $\max(X)$ and $\min(X)$ represent the maximum and minimum values of each column of well log curves in turn, respectively.

2.3.2. Data Enhancement

An important feature of using deep learning algorithms to train neural networks is the need for a large number of samples. However, logging data are difficult to obtain and the sample size is small. In the traditional image segmentation process, data enhancement methods such as rotation, translation, scaling, random occlusion, horizontal flip and noise perturbation are commonly utilized to increase the number of samples, so that the constructed network can learn features more fully [23]. Allowing for the timing and vector feature of well logging curves, rotating, translating, flipping and other transformations on the well logging curve images will destroy the meaning of the sedimentary relationship, so traditional data enhancement methods are not suitable for well logging data.

At the same time, due to the complexity and particularity of the formation of geological bodies, the identification of sedimentary microfacies using the shape of well logging curves must satisfy scale invariance. If the length and amplitude of well logging curves are different, but the overall shape is similar, they still represent the same sedimentary microfacies. Therefore, a data enhancement method of curve stretching transformation is proposed in this paper to make full use of existing logging data and increase the size of training samples. This method is principally used to stretch the curve segment representing a certain depositional environment in the entire logging curve, reflecting the variation characteristics of sediments with different thicknesses formed in the same depositional environment. Stretch transformation inserts the average value of two adjacent points to stretch the length of the curve. Aiming at the existing well logging curves, one curve segment belonging to the same type is selected for stretching with the permutation and combination method under the constraints of sedimentary laws, so as to construct new samples. On analyzing the sample data, the maximum thickness of the three sedimentary microfacies of the submerged bifurcated channel, river channel margin and river channel bend in this area are 104.7 m, 108.2 m and 124.4 m, respectively. The thickness will not exceed the maximum value so as to avoid violating the sedimentary law. The well logging curve contains many classes, so the proposed method can significantly augment the set of samples, which can make up for the insufficient number of samples of the well logging data in deep learning.

2.3.3. Attribute Extension

Extracting some salient features of logging curves based on physical drive helps to reduce the complexity of data feature extraction and achieve the purpose of improving classification accuracy. The log curve is the geophysical response of the sedimentary body, and the abrupt change point of its shape is the key to dividing different sedimentary microfacies. Feature augmentation combined with domain knowledge can further improve machine learning capabilities [24]. Based on this, this study defined three attributes of the logging curve slope value, azimuth value and variance value, and used them as characteristic curves extracted via physical drives to be involved in the identification of subsequent sedimentary microfacies. When selecting attributes, it is necessary to select attributes that can describe the morphological characteristics of curves such as lithology curves and resistivity curves in each layer section [25]. These three attributes are used to enhance the morphological characteristics of the curve, enhance the “directivity” of the logging data and highlight the sudden change points of the logging curve, thus reducing the complexity of data feature extraction.

(1) Slope D refers to the difference between the values of two adjacent logging samples, indicating the rate of change of the curve, which is conducive to highlighting the division point between different sedimentary microfacies and enhancing the morphological characteristics of the curve. This attribute essentially reflects the rate of depositional cycles and expands the characteristics of depositional rhythms. The simplest difference equation can find the change between two points of the curve. By calculating the difference value corresponding to each measurement point, a new slope value property curve can be extended.
(2) Azimuth G: Well logging curves have time-series characteristics, and their overall directionality is the key to identifying sedimentary microfacies. A square position is assigned to each measurement point, as a supplement to the attribute "slope", which reflects the characteristics of the data value becoming larger or smaller, and can enhance the attribute of the "directivity" of the logging data. The log curve composed of n values $X = (x_1, x_2, \ldots, x_n)$ and the corresponding azimuth value curve $G = (g_1, g_2, \ldots, g_n)$ can be obtained with the following formula:

$$g_i = \begin{cases} 0.1, & if\ x_{i+1} - x_i > 0.01; \\ -0.1, & if\ x_{i+1} - x_i < -0.01; \\ 0, & Otherwise; \end{cases} \qquad (2)$$
$$g_n = g_{n-1}.$$

where g_i represents the azimuth value corresponding to the value of log data.
(3) Variance A can reflect the degree of dispersion of each measurement point relative to the overall data. This can highlight the sudden change point of the well log curve and better realize its stratification. It can be obtained with a simple variance equation.

2.3.4. Evaluation Metrics

The confusion matrix is widely used in machine learning to verify the quality of classification results. It is a table of relationships between predicted and true values. The rows represent the true categories, and the columns of the matrix represent the predicted categories. This paper sets the value of the confusion matrix as the probability of the category being correctly identified or misjudged, rather than the number of samples. It is a concentrated expression of the classification results of each pixel of the input image. For example, the three-category confusion matrix is shown in Table 1, where Pxx represents the probability that the X category is correctly identified, and P_{X1X2} represents the probability that the X1 category is misidentified as the X2 category.

Table 1. Three-category confusion matrix.

		Predictive Label		
		A	**B**	**C**
True label	A	PAA	PAB	PAC
	B	PBA	PBB	PBC
	C	PCA	PCB	PCC

According to the confusion matrix, we can clearly know the accuracy rate of each category and analyze which categories are easily misjudged by the others.

3. Results and Analysis

3.1. Dataset

The logging data used in this paper come from the Triassic formation in the LN area of the northern Tarim Basin, which is a very important oil–gas-bearing formation [24]. It is mainly composed of clastic rocks deposited in fluvial lacustrine facies, including conglomerate, sandstone, siltstone, mudstone and their intermediate products. It has the characteristics of a positive rhythmic grain sequence combination, and is a set of inland fluvial lake delta facies deposits. Drilling and well logging data show that these sedimentary facies can be divided into three sedimentary microfacies: submerged bifurcated channel, river channel margin and river channel bend [26].

Sedimentary microfacies, as a combination of lithofacies, have certain characteristic rules in terms of sediment lithology, grain size distribution, shale content and porosity, which are caused by hydrodynamic factors in the depositional environment. In addition, the difference in hydrodynamic conditions will cause changes in the petrophysical properties of sediments, such as electrical conductivity, natural radioactivity and acoustic wave transmission velocity. As the response of petrophysical properties around the borehole wall, the logging curve is of great significance for the fine division of lithofacies or sedimentary microfacies. The morphology of well logging curves can help to classify sedimentary microfacies. The sediments in the submerged bifurcated channel are mainly sandstone and pebble sandstone, and their spontaneous potential curves and natural gamma ray curves are characterized by low to medium microdentate or toothed bell shapes, and sometimes box shapes, or a combination of both. The sediments of the microfacies at the lateral edge of the channel are mainly fine sandstone or siltstone, with argillaceous sandstone interbeds, medium to high resistivity, medium negative spontaneous potential, and obvious zigzag curves. River channel bends have high micro-facies shale content, low apparent resistivity amplitude and generally low amplitudes of natural gamma rays and spontaneous potential curves, which are flat or micro-toothed with a low amplitude [27]. Figure 9 is an example of the division of three sedimentary microfacies.

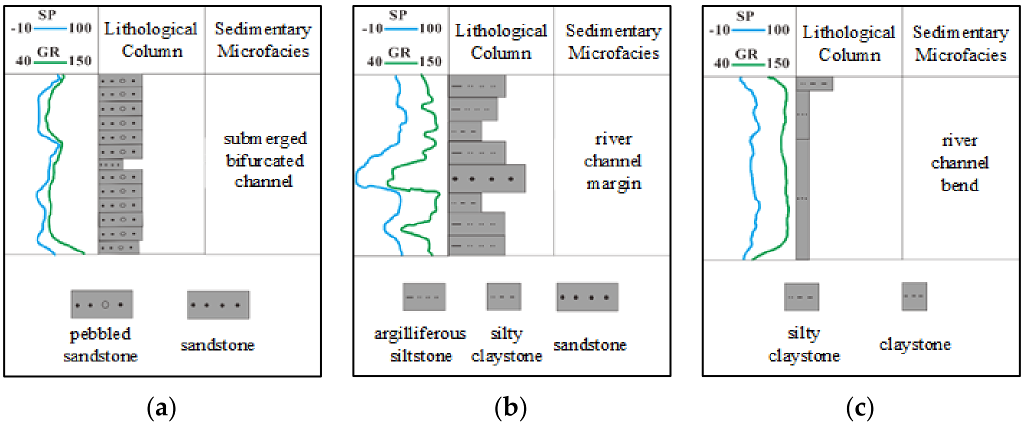


Figure 9. Log characteristics of different sedimentary microfacies [26]: (**a**) Submerged bifurcated channel; (**b**) river channel margin; (**c**) river channel bend.

In the process of identifying sedimentary microfacies using logging curves, logging curves that can reflect rock lithology, grain size, shale content and porosity are usually selected. The abnormal amplitude of (SP) and the amplitude of GR can reflect the grain size and shale content in the formation. For clastic rock reservoirs, GR and SP are the most commonly used logging curves to indicate lithology, and permeable layers can be displayed in sand–mudstone sections [28]. DEN can indirectly describe the porosity of the sedimentary body and reflect the grain size and lithological characteristics of the sedimentary body [29]. Since a sedimentary facies is a combination of lithofacies, and a lithofacies is a combination of lithological characteristics, the change curve L of lithological characteristics reflected by logging curves is an important basis for identifying sedimentary microfacies. According to the logging data, we selected SP, GR, DEN logging curves and lithology curve L as the basic data for identifying sedimentary microfacies in this area.

The available dataset contains well log data from 19 wells. The four curves of each well are arranged to form a two-dimensional matrix, forming 19 original images. We took the results of the manual identification of sedimentary microfacies by geologists by observing lithological columns and characteristics of logging curves, and made them into the actual labels of the corresponding wells. Using the enhancement method proposed, the label data were expanded from the original 19 logging images to 258. Starting from the first line of the well logging image, we cut it with 800 lines and saved the last 800 lines directly when the remaining images were smaller than 800 lines. Thus, the logging image was cut into 1286 sample sets of 800×4 size as the input data, among which there were 1125 training datasets, 125 validation sets and 36 test sets.

We separately obtained the slope curves (DEN-D, GR-D, SP-D), azimuth curves (DEN-G, GR-G, SP-G) and variance curves (DEN-A, GR-A, SP-A) corresponding to the DEN, GR and SP curves. The effectiveness and correctness of the proposed method was verified by comparing the results of different input data features in different network architectures.

3.2. Network Performance

We used MATLAB for data processing, and used the Python language to build and train the model on PyCharm based on the Keras framework. To obtain the parameter settings of the optimal model, we conducted a lot of experiments to analyze relevant parameters, mainly including learning rate, batch size, convolution kernel size and downsampling times. The accuracy rate Acc is the standard for judging the effect of the model. The changes in network model accuracy under different parameter configurations are shown in Table 2. We mainly set the learning rate to the negative power of 10, and set the batch size to the power of 2 for testing. In order to obtain a more accurate parameter value, according to the results, we further narrowed down the range of optimal parameter values by testing the intermediate values of 5×10^{-4} and 6 for the learning rate within the range of $[10^{-3}, 10^{-4}]$ and the batch size within the range of [4, 8]. When the input data size is 800×8, a maximum of three downsampling operations can be performed. Based on the test results, we obtained the optimal model parameters: a batch size of 6, a learning rate of 10^{-4}, and two downsampling operations.

Table 2. The accuracy rate of the network model under different parameter configurations.

Learning Rate	10−1	10−2	10−3	5 × 10−4	10−4	10−5	10−6
Acc (%)	37.82	27.18	84.02	83.10	86.42	80.24	78.80
Batch size	1	2	4	6	8	16	32
Acc (%)	82.23	80.69	83.42	86.98	76.24	79.18	76.18
Downsampling Times	1		2			3	
Acc (%)	73.95		84.69			81.86	

3.3. Result Analysis

To validate the superiority of our proposed method over conventional methods, we conducted experiments on eight different input datasets and compared the performance of our Improved U-net model with U-net architecture, as well as the ANN and SVM models used in traditional methods for well log lithology identification. The results demonstrate that our Improved U-net model performs better in identifying sedimentary microfacies. The following results are based on test datasets that were not involved in the training process.

Table 3 shows the accuracy obtained by the four models on eight test datasets with different input data characteristics. The eight datasets were mainly divided into four categories:

Category 1: Consists of only three well log curves and one lithology curve;
Category 2: Consists of one well log curve and its corresponding three extension curves (slope, azimuth, variance);
Category 3: Consists of two well log curves and their corresponding three extension curves (slope, azimuth, variance);
Category 4: Consists of three well log curves and their corresponding three extension curves (slope, azimuth, variance).

Table 3. The accuracy obtained by the four models on eight test datasets with different input data characteristics.

Input Data \ Classification Model	SVM	ANN	Raw U-Net	Improved U-Net
DEN, GR, SP, L	60.26%	65.20%	69.47%	73.34%
DEN, DEN-D, DEN-G, DEN-A	48.04%	49.83%	58.95%	64.71%
GR, GR-D, GR-G, GR-A	55.11%	60.63%	68.12%	74.31%
SP, SP-D, SP-G, SP-A	55.11%	58.65%	63.52%	68.09%
GR, GR-D, GR-G, GR-A, DEN, DEN-D, DEN-G, DEN-A	62.85%	60.30%	76.59%	80.93%
SP, SP-D, SP-G, SP-A, DEN, DEN-D, DEN-G, DEN-A	63.01%	63.18%	76.13%	78.88%
SP, SP-D, SP-G, SP-A, GR, GR-D, GR-G, GR-A	62.27%	66.56%	76.81%	82.90%
DEN, DEN-D, DEN-G, DEN-A, GR, GR-D, GR-G, GR-A, SP, SP-D, SP-G, SP-A	60.13%	65.44%	81.53%	86.49%

Table 3 shows that our proposed method has significant advantages over traditional ANN and SVM classification algorithms in identifying sedimentary microfacies. Compared with the Raw U-net architecture, our Improved U-net model can significantly improve the accuracy of sedimentary microfacies identification. In addition, compared with the DEN curve, GR and SP curves make a greater contribution to the characterization of sedimentary microfacies in this study area, and adding data features through attribute extension can significantly improve the test results. Among the input data consisting of DEN, GR, SP, DEN-D, DEN-G, DEN-A, GR-D, GR-G, GR-A, SP-D, SP-G and SP-A, the Improved U-net model has the highest accuracy at 86.49%.

Noise was added to the input data and the influence of data with different signal-to-noise ratios on the deposition level was analyzed. Table 4 shows the accuracy of input data with different SNR in Improved U-net and Raw U-net models, and the highest accuracy of data without added noise in the two models. It can be seen that the accuracy of Improved U-net is higher than that of Raw U-net when the SNR is the same. Moreover, the higher the signal-to-noise ratio of the input data, the higher the accuracy of the Improved U-net for deposition micro-recognition.

Table 4. The accuracy rate of the network model under different SNR.

S/N	Data without Noise	12	10	8	5	3	2
Improved U-net	86.49%	78.36%	72.68%	62.44%	36.43%	10.53%	11.23%
Raw U-net	81.53%	70.46%	71.34%	54.95%	22.35%	12.61%	9.65%

We also analyzed the impact of sample size on different network architectures. For the eighth input dataset (composed of 12 curves: three well log curves and their corresponding slope, azimuth and variance extension curves), we used a training dataset of $1250 \times 800 \times 12$ well log curves and randomly selected 30%, 50%, 70%, 90% and 100% of the sample size to train different network models. The same test dataset was employed to achieve the recognition accuracy of each trained model and analyze the training effect. The accuracy of Raw U-net slowly increases with the increase in sample size, and the performance is poor (Figure 10). In contrast, the Improved U-net model performs well overall and the accuracy gradually increases with the increase in sample size.

The confusion matrix of the prediction results for various methods is shown in Figure 11. It can be seen that the ANN and SVM methods have inaccurate feature extraction for sedimentary microfacies and are prone to misclassification, while Raw U-net and Improved U-net show relatively good classification results. Compared with Raw U-net, Improved

U-net can be used to more accurately characterize submarine distributary channels and the river channel bay, but the identification of the river channel margin is not as good. This may result in the misclassification of sandstone interbeds at the edge of the river channel as distributary channels mainly composed of sandstone, and the misclassification of mudstone interbeds as river channel bay mainly composed of mudstone.

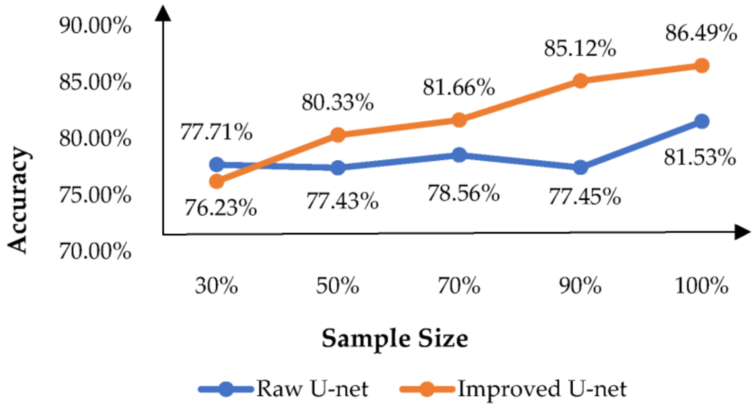


Figure 10. Variation in accuracy with sample size under different network architectures.

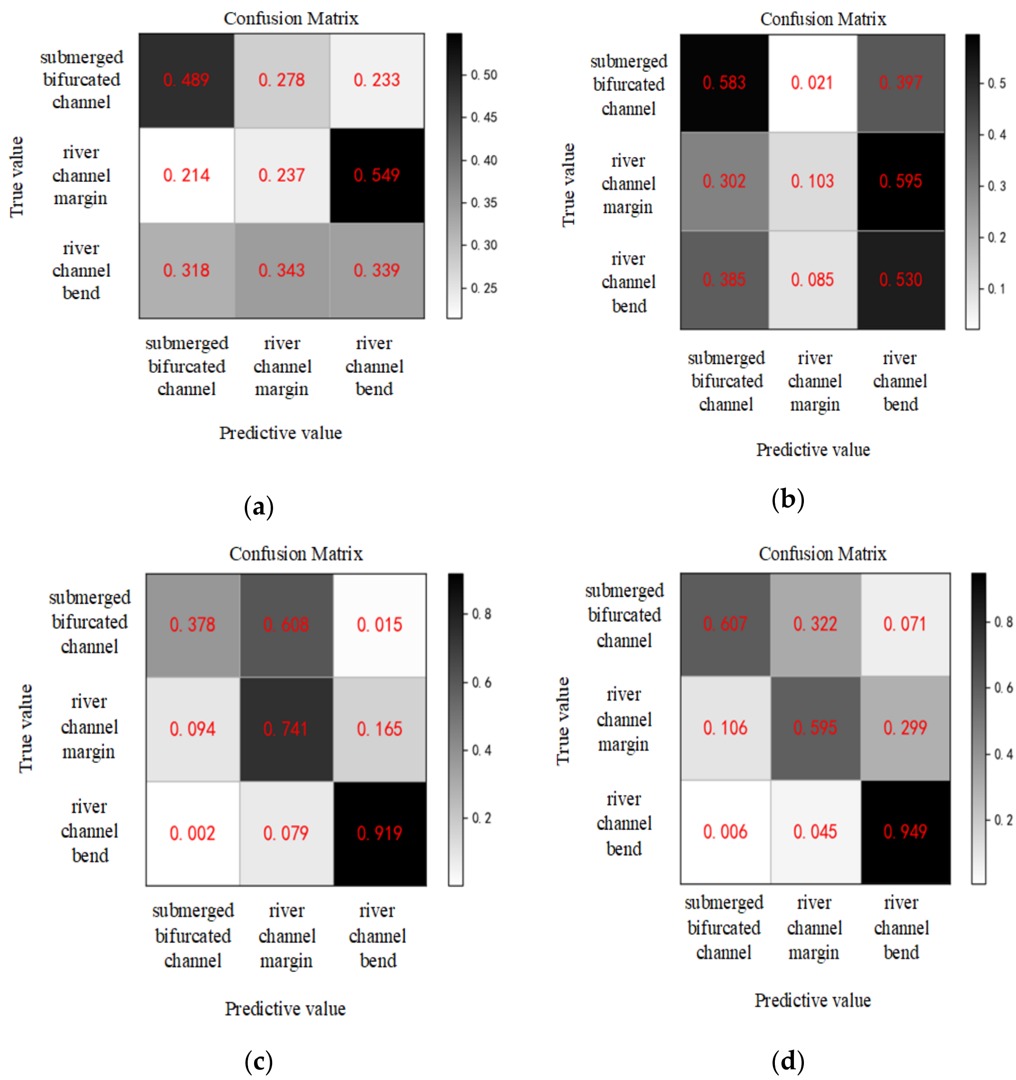


Figure 11. Confusion matrix of the prediction results (**a**) ANN; (**b**) SVM; (**c**) Raw U-net; (**d**) Improved U-net.

The segmentation and identification results of the test wells is shown in Figure 12. The experimental results show that our method can be applied to the description of sedimentary microfacies in well logging data. Compared with the manually identified original labels, this method can accurately segment different categories of sedimentary microfacies and identify them. For the results in Figure 12a,b, the segmentation results of the ANN and SVM methods show fragmented characteristics, which are due to the fact that the input samples only consider the well logging curve values at the depth points without considering the morphological changes of the well logging curves. On the other hand, both Raw U-net and Improved U-net segmentation results perform better, with good continuity and block. Most of all, Improved U-net shows the best performance. In local areas, due to sudden changes in the structure of well logging curves, Improved U-net can perform more detailed identification of sedimentary microfacies. Compared with manual identification, this method has higher sensitivity and can further refine the results of manual classification.

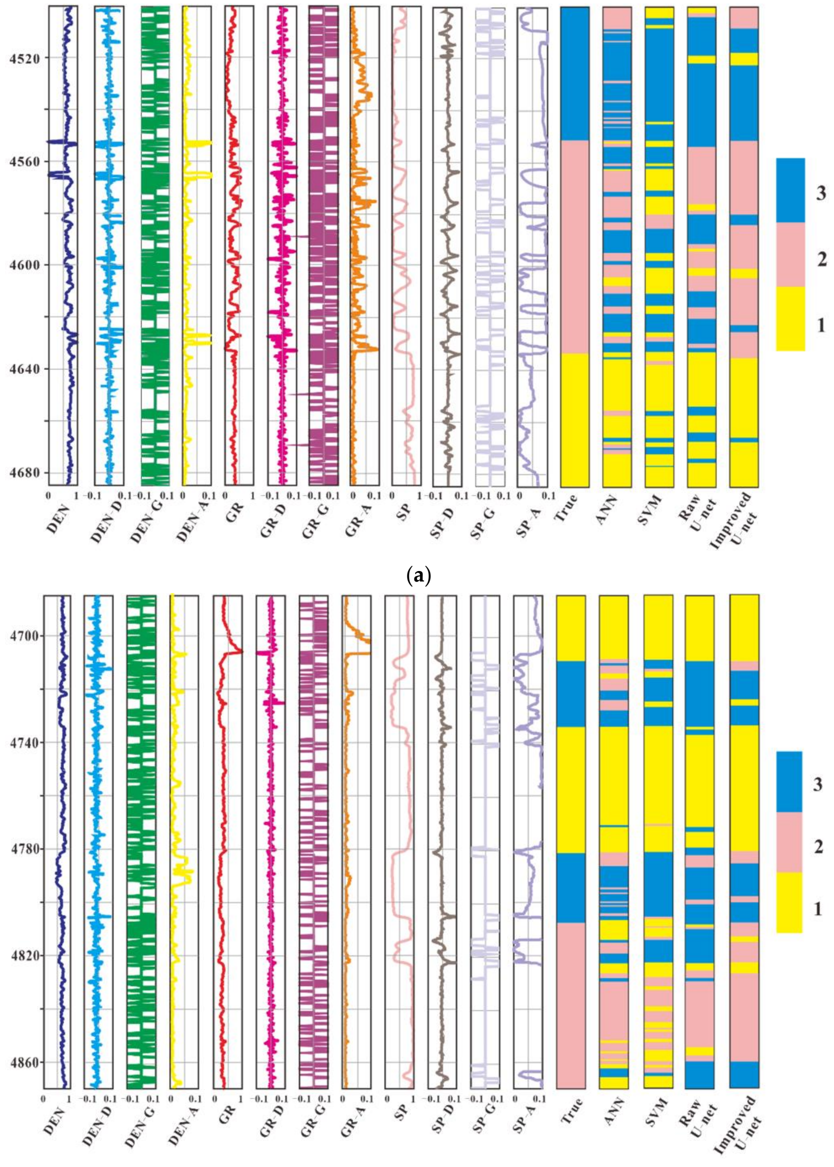
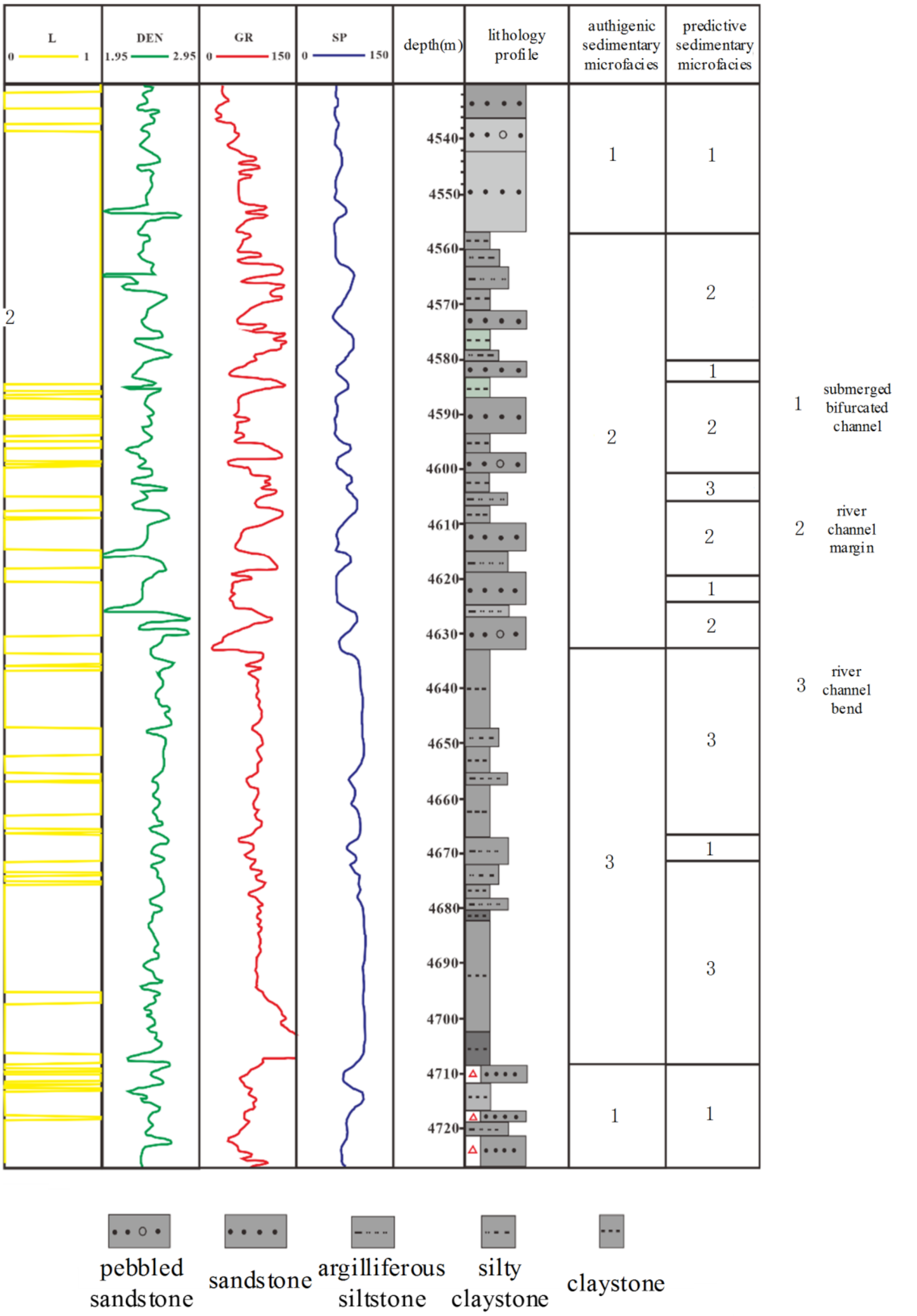


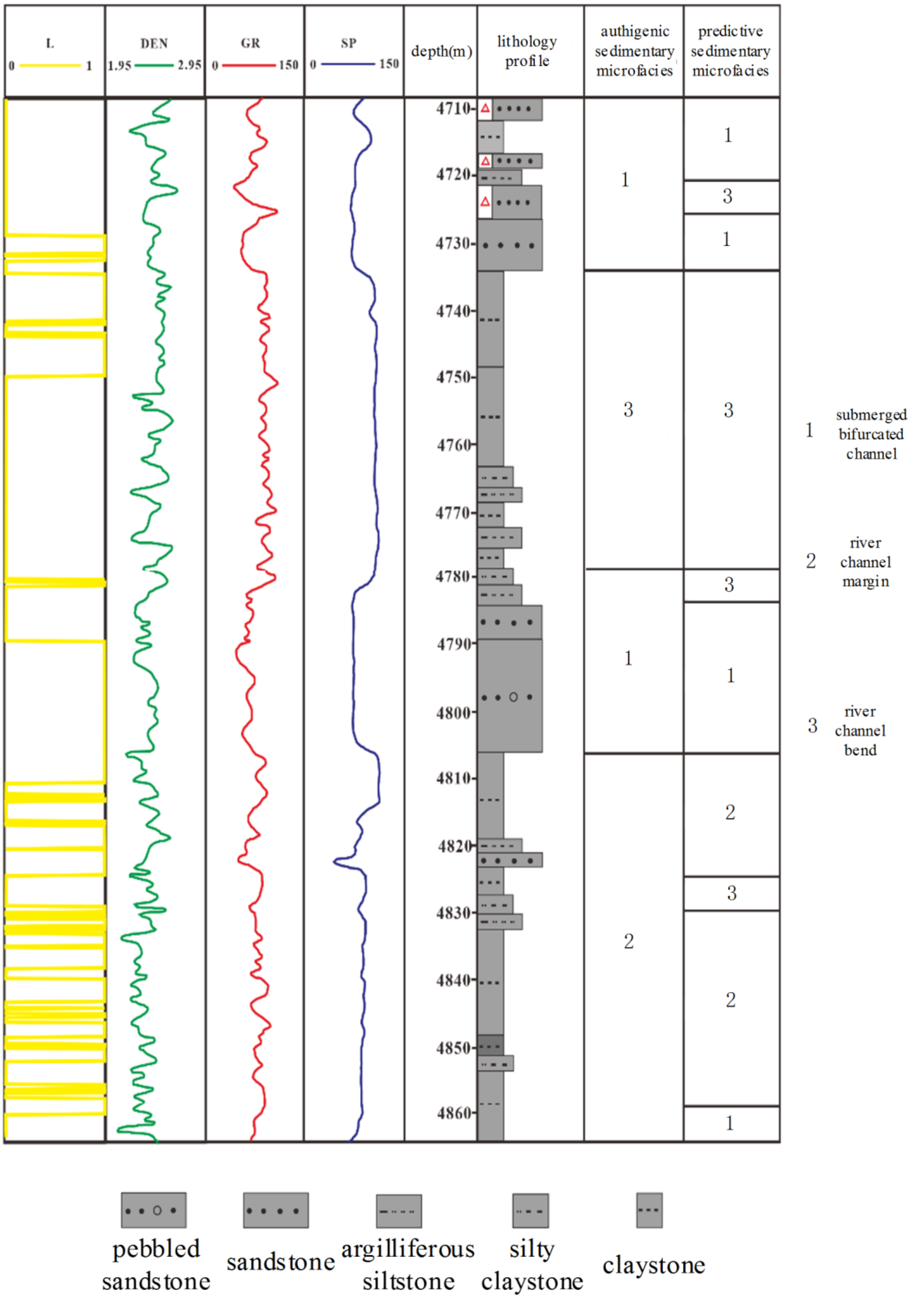
Figure 12. The testing results of different methods for Triassic formation in the LN area of Northern Tarim Basin, 1, 2 and 3 represent three sedimentary microfacies: submerged bifurcated channel, river channel margin and river channel bend, respectively; (**a**) depth range from 4500 m to 4700 m; (**b**) depth range from 4700 m to 4880 m.

The lithologic profile of the test well is shown in Figure 13, along with a comparison of the results obtained using our method and geological experts (labels) [30]. This figure shows that our method can effectively segment and precisely describe sedimentary microfacies. Our method is more sensitive to the change in data characteristics than the identification results of geological experts, so the identification of sedimentary microfacies is more detailed.



(a)

Figure 13. *Cont.*



(b)

Figure 13. The lithology profile of the test wells and a comparison of the identification results obtained with Improved U-net and labels; (**a**) well LN-1; (**b**) well LN-2. The red triangles in the figure represents oil producing intervals.

The trained model was also tested on unlabeled actual well logging data (Figure 14). It can be inferred that the model can also accurately identify and classify sedimentary microfacies in new actual well logging data, and the overall identification results are ideal. However, the segmentation in local areas is too detailed, mainly due to the model's sensitivity to data changes. The results indicate that this deep learning well logging sedimentary microfacies identification method based on the U-net framework has good generalization ability on different datasets, and no specific adjustment is needed. For logging data from other areas, considering the corresponding curve shape characteristics, the proposed model can also be used to identify sedimentary microfacies. According to the background characteristics of the area where the data are used, the corresponding parameters can be fine-tuned to obtain a better recognition effect.

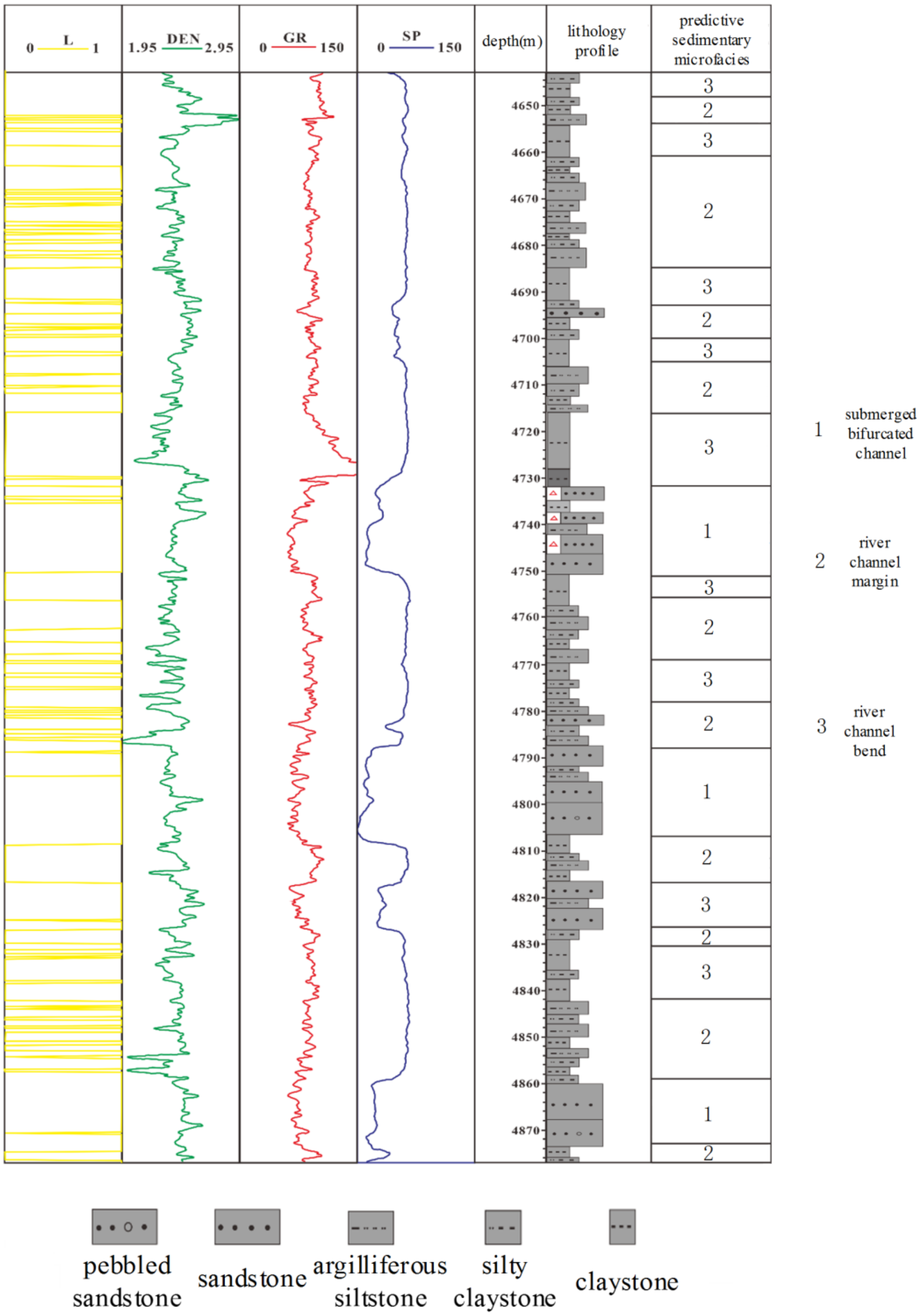


Figure 14. The identification results of Improved U-net in unlabeled actual well logging data.

4. Conclusions

Our research realizes the intelligent identification of logging sedimentary microfacies by using the ideas of deep learning and semantic segmentation. Considering that the time-series deposition reflected by the logging curve determines that it has the characteristics of "vectoriality" and "multiscale" in the depth direction, we removed the pooling layer and introduced a multiscale convolution block to the traditional U-net. The last layer was also improved by adding a 1D convolutional operation. A deep learning network architecture that can realize multi-scale feature extraction and unidirectional segmentation of logging sedimentary microfacies identification is constructed. At the same time, a unique data enhancement method for well logging curves is proposed, and attribute extension was used to extract more intuitive curve features to alleviate the complexity of feature extraction, thereby significantly improving the training efficiency and application effect of the model. The experimental results of applying our method to the sedimentary microfacies identification of the Triassic formation in the LN area of the Northern Tarim Basin show that our method shows advantages compared with the SVM, ANN algorithm and the original U-net. It can realize the fine identification of sedimentary microfacies using log data on the premise of meeting the requirement of geologists for the accuracy of sedimentary microfacies.

Author Contributions: Conceptualization, H.C.; methodology, H.C. and Y.H.; validation, C.Y.; formal analysis, M.S.; data analysis, L.Z.; data curation, Y.H.; writing—original draft preparation. Y.Z.; visualization, L.Z. All authors have read and agreed to the published version of the manuscript.

Funding: The National Natural Science Foundation of China (Grant No. 42130812) and PetroChina Science and Technology Major Project (Grant No. 2021DJ0403).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: We would like to acknowledge the anonymous reviewers for their thoughtful comments and valuable suggestions, which greatly improved the manuscript.

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

References

1. Cao, G.H.; Hu, Y.H.; Zhang, Q.W. Study on the method of identification of sedimentary microfacies using well logging data. *Sci. Technol. Eng.* **2007**, *7*, 3674–3680. [CrossRef]
2. Ding, F.; Zhang, J.L.; Xie, J. Fine description of structure and sedimentary microfacies of Li32 block of Lijin oilfield, Dongying depression, China. *Arab. J. Geosci.* **2014**, *7*, 1693–1704. Available online: https://link.springer.com/article/10.1007/s12517-013-0972-8 (accessed on 8 September 2023).
3. Chen, W.F. Gamma log trend facies in the Choshui Fan-delta, Taiwan. *Terr. Atmos. Ocean Sci.* **1998**, *9*, 633–642. [CrossRef]
4. Deng, R.; Meng, F. On logging curves fine delamination to identify sedimentary microfacies. *Well Logging Technol.* **2010**, *34*, 554–558. [CrossRef]
5. Chen, Y.F.; Peng, S.M. Quantitative study on sedimentary microfacies. *Petrol. Explor. Dev.* **2003**, *30*, 51–53. [CrossRef]
6. Li, Y.M.; Anderson-Sprecher, R. Facies identification from well logs: A comparison of discriminant analysis and naïve Bayes classifier. *J. Petrol. Sci. Eng.* **2006**, *53*, 149–157. [CrossRef]
7. Saggaf, M.M.; Nebrija, E.L. A fuzzy logic approach for the estimation of facies from wire-line logs. *AAPG Bull.* **2003**, *87*, 1223–1240. [CrossRef]
8. Moradi, M.; Tokhmechi, B.; Masoudi, P. Inversion of well logs into rock types, lithofacies and environmental facies, using pattern recognition, a case study of carbonate Sarvak Formation. *Carbonates Evaporites* **2019**, *34*, 335–347. Available online: https://link.springer.com/article/10.1007/s13146-017-0388-8 (accessed on 8 September 2023). [CrossRef]
9. Ai, X.; Wang, H.Y.; Sun, B.T. Automatic identification of sedimentary facies based on a support vector machine in the Aryskum Graben, Kazakhstan. *Appl. Sci.* **2019**, *9*, 4489. [CrossRef]
10. Wang, D.H.; Peng, J.; Yu, Q.; Chen, Y.Y.; Yu, H.H. Support vector machine algorithm for automatically identifying depositional microfacies using well logs. *Sustainability* **2019**, *11*, 1919. [CrossRef]

11. Kuang, L.C.; Liu, H.; Ren, Y.L.; Luo, K.; Shi, L.Y.; Su, J.; Li, X. Application and development trend of artificial intelligence in petroleum exploration and development. *Petrol. Explor. Dev.* **2021**, *48*, 1–11. [CrossRef]
12. Saggaf, M.M.; Nebrija, E.L. Estimation of lithologies and depositional facies from wire-line logs. *AAPG Bull.* **2000**, *84*, 1633–1646. [CrossRef]
13. Li, W.B.; Yu, Y.L.; Wang, J.Q.; Ye, B.; Xin, W. Application of self-organizing neural network method in logging sedimentary microfacies identification. *Adv. Mater. Res.* **2013**, *616*, 38–42. [CrossRef]
14. Bhatt, A.; Helle, H.B. Determination of facies from well logs using modular neural networks. *Petrol. Geosci.* **2002**, *8*, 217–228. [CrossRef]
15. Zhang, J.; Liu, S.; Li, J.; Liu, L.; Liu, H.; Sun, Z. Identification of sedimentary facies with well logs: An indirect approach with multinomial logistic regression and artificial neural network. *Arab. J. Geosci.* **2017**, *10*, 247–253. Available online: https://link.springer.com/article/10.1007/s12517-017-3045-6 (accessed on 8 September 2023). [CrossRef]
16. He, J.H.; La-Croix, A.D.; Wang, J.H.; Ding, W.L.; Underschultz, J.R. Using neural networks and the Markov Chain approach for facies analysis and prediction from well logs in the Precipice Sandstone and Evergreen Formation, Surat Basin, Australia. *Mar. Petrol. Geol.* **2019**, *101*, 410–427. [CrossRef]
17. John, D.; Zhang, C. An attention-based U-Net for detecting deforestation within satellite sensor imagery. *Int. J. Appl. Earth Obs.* **2022**, *107*, 102685. [CrossRef]
18. Xu, Q.; Ma, Z.C.; Na, H.; Duan, W.T. DCSAU-Net: A deeper and more compact split-attention U-Net for medical image segmentation. *Comput. Biol. Med.* **2023**, *154*, 106626. [CrossRef]
19. Wang, H.Y.; Xie, S.A.; Lin, L.F.; Lwamoto, Y.; Han, X.H.; Chen, Y.W.; Tong, R.F. Mixed Transformer U-Net for Medical Image Segmentation. In Proceedings of the ICASSP 2022–2022 IEEE International Conference on Acoustics, Speech and Signal Processing, Singapore, 27 April 2022. [CrossRef]
20. Ibtehaz, N.; Rahman, M.S. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural Netw.* **2020**, *121*, 74–87. [CrossRef]
21. He, K.M.; Zhang, X.Y.; Ren, S.Q.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016. [CrossRef]
22. Liu, H.; Chen, P.; Xia, H. Automatic Identification of Sedimentary Microfacies with Log Data and Its Application. *Well Logging Technol.* **2006**, *30*, 233–236. [CrossRef]
23. Ren, G.F.; Tian, Z.M. Application of Self-Organizing Competitive Network in Lithologic Identification of the Logging Data. In Proceedings of the 2012 International Conference on Computing, Measurement, Taiyuan, China, 7–9 July 2012. [CrossRef]
24. Simard, P.Y.; Steinkraus, D.; Platt, J.C. Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis. In Proceedings of the Seventh International Conference on Document Analysis and Recognition, Edinburgh, UK, 6 August 2003. [CrossRef]
25. Chen, J.; Zeng, Y. Application of Machine Learning in Rock Facies Classification with Physics-Motivated Feature Augmentation. In Proceedings of the ICML 2018: International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018. [CrossRef]
26. Ma, A.; Zhang, S.C.; Zhang, D.J. Ruthenium-ion-catalyzed oxidation of asphaltenes of heavy oils in Lunnan and Tahe oilfields in Tarim Basin, NW China. *Org. Geochem.* **2008**, *39*, 1502–1511. [CrossRef]
27. Wang, Z.D.; Wang, X.H.; Wang, J.F.; He, Y.B.; Hu, Z.H.; Fang, H.F.; Liu, H. Triassic sedimentary microfacies in the second and third well fields in the Lunnan Oil Field, Tarim Basin. *Sediment. Geol. Tethyan Geol.* **2008**, *28*, 66–74. Available online: http://en.cgsjournals.com/article/id/cjyttsdz_20080312 (accessed on 8 September 2023).
28. Zhong, L.X.Z.; Fu, H.; Liu, Y.T.; Chen, J. Sedimentary facies of Triassic-Jurassic formation in Yuqi area, Tarim Basin. *Lithol. Reserv.* **2013**, *25*, 29–34. [CrossRef]
29. Lai, J.; Fan, X.C.; Liu, B.C.; Pang, X.J.; Zhu, S.F.; Xie, W.B.; Wang, G.W. Qualitative and quantitative prediction of diagenetic facies via well logs. *Mar. Petrol. Geol.* **2020**, *120*, 104486. [CrossRef]
30. Zhao, P.Q.; Ostadhassan, M.; Shen, B.; Liu, W.H.; Abarghani, A.; Liu, K.Q.; Luo, M.; Cai, J.C. Estimating thermal maturity of organic-rich shale from well logs: Case studies of two shale plays. *Fuel* **2019**, *235*, 1195–1206. [CrossRef]