applied sciences

MDPI

Article

Lithofacies Prediction from Well Log Data Based on Deep Learning: A Case Study from Southern Sichuan, China

Yu Shi [1,2], Junqiao Liao [2], Lu Gan [1,*] and Rongjiang Tang [1]

[1] Yangtze Region Delta Institute, University of Electronic Science and Technology of China, Huzhou 313002, China; geo.shiyu@foxmail.com (Y.S.); rongjiang.igp@hotmail.com (R.T.)
[2] Sixth Geological Brigade of Sichuan Province, Luzhou 611731, China; liaojunqiao99@163.com
* Correspondence: ganlu@pku.edu.cn

Abstract: This paper utilizes prevalent deep learning techniques, such as Convolutional Neural Networks (CNNs) and Residual Neural Networks (ResNets), along with the well-established machine learning technique, Random Forest, to efficiently distinguish between common lithologies including coal, sandstone, limestone, and others. This approach is highly significant for resource extraction—such as coal, oil, natural gas, and groundwater—by streamlining the process and minimizing the need for the time-consuming manual interpretation of geophysical logging data. The natural gamma ray, density, and resistivity log data were collected from 22 wells in the mountainous region of Southern Sichuan, China, yielding approximately 70,000 samples for developing lithofacies prediction models. All the models achieved around 80% accuracy in classifying carbonaceous lithologies and up to 88% accuracy in predicting other lithologies. The trained models were applied to the logging data in the validation dataset, and the outputs were validated against geological core data, showing overall consistency, although variations in the classification results were observed across different wells. These findings suggest that deep learning techniques have the potential to develop a general model for effectively handling lithology classification with well logging data.

Keywords: lithofacies prediction; well log; deep learning

check for updates

Citation: Shi, Y.; Liao, J.; Gan, L.; Tang, R. Lithofacies Prediction from Well Log Data Based on Deep Learning: A Case Study from Southern Sichuan, China. *Appl. Sci.* **2024**, *14*, 8195. https://doi.org/10.3390/app14188195

Academic Editor: Tiago Miranda

Received: 4 August 2024
Revised: 10 September 2024
Accepted: 10 September 2024
Published: 12 September 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In geological exploration and resource development, identifying target horizons based on the well logging data is a critical step [1]. Target horizons refer to specific geological layers, such as sandstone, coal seams, and limestone, which often contain valuable resources like oil, natural gas, coal, and minerals [2]. Accurately defining the extent of coal is important for safety, resource estimation, and mine design, and helps optimize mining and oil extraction plans and minimize resource waste. In addition, sandstone and limestone are common reservoir rocks that often contain abundant oil and natural gas resources.

The lithology identification from well logging is to establish the relationship between the petrological characteristics and the logging curves. Typical lithologies are supposed to have their own specific logging responses [3]. For example, coal seams typically have low mass density and exhibit natural gamma density because they are rich in organic matter, resulting in low levels of radioactive elements. Sandstone, on the other hand, has high resistivity and, if it is free of impurities or water, generally shows low natural gamma content and higher density. Pure limestone typically has low natural gamma values because it contains few radioactive clays, and it exhibits high resistivity and mass density [4–6].

However, traditional logging interpretation relies extensively on expert knowledge and experience, making it labor-intensive and time-consuming. Additionally, it is often affected by subjectivity and inconsistency in expert judgment [7]. Sometimes, well logging interpretation is based on the mathematical and physical modeling of the processes under study, statistical methods (correlation and discrimination), solving systems of non-linear

equations petrophysics (inverse problem of geophysics), and some other linear statistical methods. The complexities inherent in the geological conditions of a combination of different rock types, such as muddy sandstone, muddy limestone, carbonate, tight sandstone, or sandy conglomerate reservoirs [8,9], the growing diversity, and volume of logging data highlight the significant limitations of traditional interpretation methods.

In recent years, in the face of massive geophysical data processing and interpretation, deep learning (DL) methods have provided new approaches for utilizing and analyzing data [10–13]. Due to the data-driven nature of deep learning methods, which do not rely on physical models, once the model is trained well, it can establish a rapid mapping from the observed data to the predicted parameters. As a result, automatic interpretation techniques of well logging curves using deep learning methods exhibit extremely high efficiency. In the past decades, different statistical and machine learning approaches have been used to automatedly predict lithotypes directly using a combination of geophysical well log data in petroleum wells [14–17]. Generally, the identification of coal intervals or other different lithology is not a difficult issue, and several studies have already revealed that different machine learning approaches can accurately differentiate coal from other non-coal lithologies due to its distinguishing low density and gamma ray ranges [18–22]. Recently, more studies have attempted to compare the performance of machine learning methods for lithology identification [23–25]. The results showed that Random Forest has the best score among considered algorithms.

However, most of these studies rely on a limited number of well log curves in the same geological area, with sample sizes ranging from hundreds to a few thousand, which are insufficient to establish a generalized lithofacies prediction model. Typically, well log data in these studies originate from the same region, raising concerns about the transferability of trained models to other areas. Furthermore, there are many types of well log curves, including gamma ray (GR), gamma-gamma density log (GG), self-potential (SP), caliper log (CALI), normal resistivity log (NR), neutron porosity log (PHIN), bulk density log (RHOB), and interval transit time (DT). In practical applications, however, the data are often incomplete, preventing trained models from being effectively applied to incomplete test datasets.

In this study, we collected 70,572 well log samples from 22 wells in the Southern Sichuan region. The well log curves include the most commonly used normal resistivity logging, natural gamma logging, and gamma-gamma density logging. Our objective is to establish a generalized lithology classification model that can be transferred to different regions. We aim to explore the extent to which machine learning models can achieve accurate lithology classification with limited information. Additionally, we intend to conduct a systematic and comprehensive comparison of machine learning methods (Random Forest, RF) and deep learning methods (Convolutional Neural Network (CNN) and Residual Neural Network (RNN)) for predicting lithofacies using well log data. This study also seeks to determine whether neural network models, which are more adept at image recognition, outperform machine learning models (e.g., Random Forest) in well log curve identification.

2. Study Area and Dataset

The study area is located at the border between Junlian Town in the Qianbei–South Sichuan subregion of the Yangtze Block and Luzhou in the Sichuan Basin (Figure 1). The regional tectonic setting lies on the northern edge of the Loushan fold-thrust belt in Southern Sichuan and Northern Guizhou. The fault structures in the area are relatively well developed, predominantly consisting of strike-slip faults with orientations similar to the fold axis. The fault dip angles are generally gentler on the northern limb of the anticline and steeper on the southern limb.

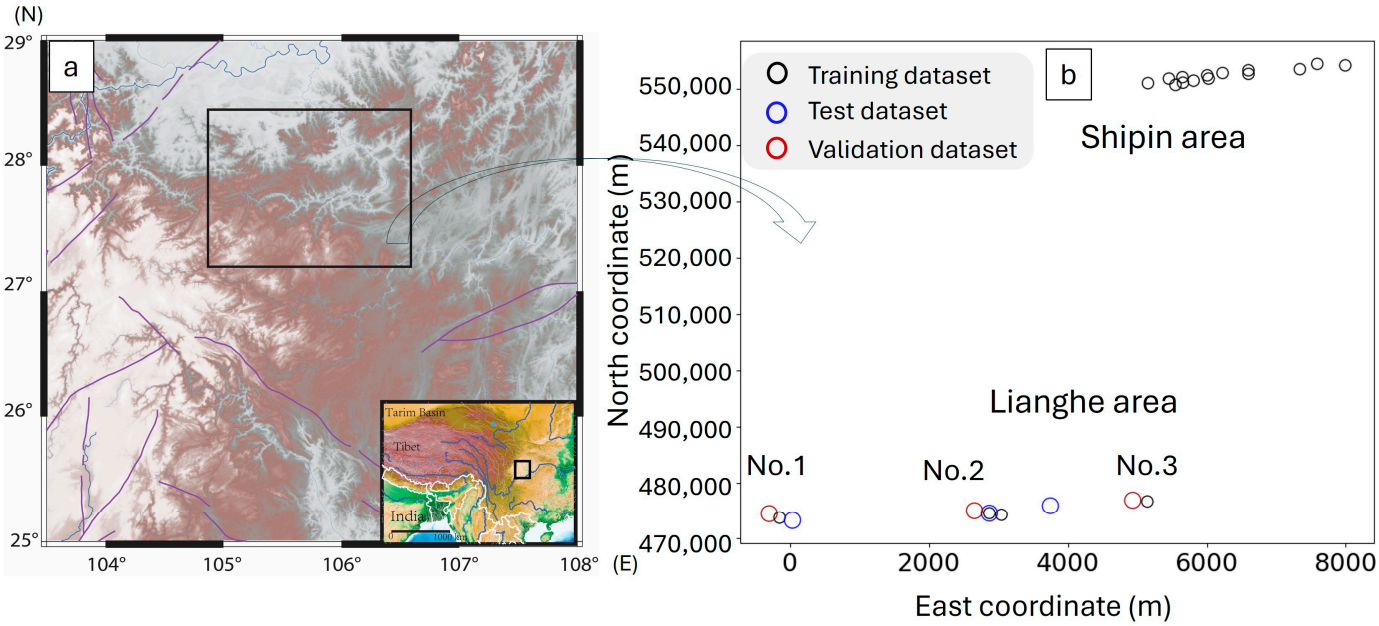


Figure 1. Location of the study area (**a**) and well distribution (**b**), with coordinates based on the Modified China Geodetic Coordinate System 2000.

The lithology is predominantly composed of argillaceous stones, followed by siltstone, fine-grained sandstone, and limestone, with medium-grained sandstone being relatively rare (Table 1). Coal and carbonaceous mudstone are present in smaller proportions, while marlstone and residual breccia are locally distributed. The oldest exposed strata in the study area are the Cambrian Gaotai Formation, and the youngest are the Quaternary System. The Triassic strata are widely exposed, while Jurassic and Cretaceous strata are absent in this region. Existing drilling targets the coal seams, primarily in the Permian Longtan Formation. Due to cost considerations, drilling only penetrates the Silurian strata (Table 1). To simplify, we assume that rocks of the same type from all geological ages have similar physical properties, so we categorize them as one type for deep learning training.

Table 1. Strata thickness and lithology from different geological age in the study area.

Geological Timechart				Thickness (m)	Lithology
System	**Series**	**Formation**	**Code**		
	Quaternary			0~50	Clay, Soil strata, Alluvial deposits
Triassic	Neogene	Xujiahe	T3xj	250~680	The upper part: sandstone and feldspathic sandstone, the lower part: mudstone, sandy mudstone, and sandstone interbedded with coal seams.
	Middle	Leikoupo	T_2l	303~337	Dolomitic limestone and limestone interbedded with thin layers of sandy mudstone
	Lower	Jialingjiang	T_1j	456~463	Limestone, argillaceous dolomite, Mudstone, Calcareous mudstone
		Feixianguan	T_1f	288	Mudstone, Sandy mudstone, Interbedded with thin layers of limestone and calcareous mudstone, and siltstone.

Table 1. *Cont.*

Geological Timechart				Thickness (m)	Lithology
System	**Series**	**Formation**	**Code**		
Permian	upper	Changxing	P_3c	32~58	Limestone
		Longtan	P_3l	86~112	Mudstone, sandy mudstone, fine sandstone, and claystone interbedded with coal seams
	Middle and Lower	Maokou, Xixia	P_2m, P_2q	196~295 130~175 2~11	Limestone, containing flint, with a base of sandy mudstone and claystone
Silurian	Middle	Hanjiadian	S_2h	300~650	Mudstone and siltstone interbedded with sandstone.
	Lower	Shiniupeng	S_1s	180~501	Limestone, argillaceous limestone
		Longmaxi	S_1l	180~460	Mudstone

Ordovician and older stratas are not recorded.

At the end of the Middle Permian, influenced by the Dongwu Movement, the Lianghe mining area, along with the vast region of Southern Sichuan, was uplifted as a whole, leading to a large-scale retreat of seawater. The Maokou Formation limestone was extensively exposed on the surface, undergoing prolonged weathering and erosion, which gradually leveled the terrain, forming a slightly undulating karst residual plain. From the Late Permian onward, the crust continued to slowly subside, accompanied by frequent oscillatory movements within the overall downward trend. This created favorable conditions for the deposition of the coal-bearing rock series of the Longtan Formation, marking the beginning of a prolonged coal accumulation period in the area. Based on various genetic indicators of the coal-bearing strata, the characteristics of the coal seams, coal quality, and metallogenic markers, the depositional environment can be divided into the following five sedimentary systems from bottom to top: weathered residual, lagoon, lake, river, and lagoon–tidal flat. Among these, the river sedimentary system, characterized by terrestrial deposition, is the most developed, accounting for 77% of the total thickness of the Longtan Formation.

The training dataset for this study was obtained from the Sichuan Provincial Sixth Geological Brigade, encompassing data from 22 wells in the Shiping and Lianghe areas of Southwestern Sichuan (Figure 1). Logging data from 16 wells were allocated to the training set, 3 wells to the test set, and the remaining 3 wells to the validation set. The sample statistics for the training and test datasets are illustrated in Figure 2. Majority of the training data are from the Shiping area, which is situated 100 km from the Lianghe area, allowing the Lianghe data to be used for evaluating the model's transferability.

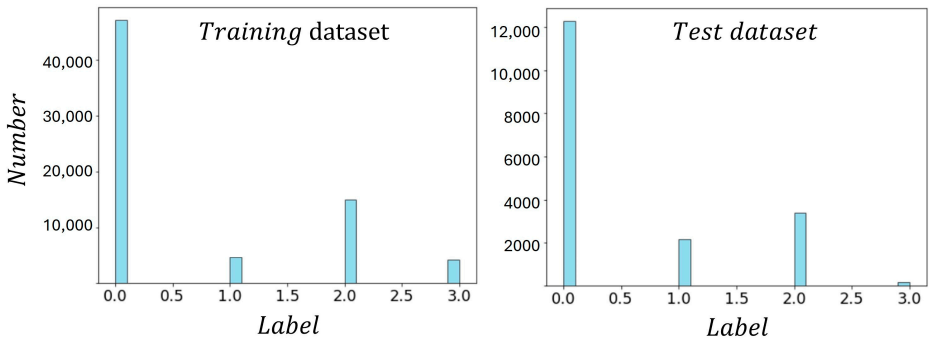


Figure 2. Distribution of sample numbers in the training and test sets: Categories 0–3 correspond to other stones, coal, sandstone, and limestone, respectively.

For classification purposes, we divided the rocks in this region into the following four categories: coal, sandstone, limestone, and others. The majority in others are argillaceous

complex rock, such as argillaceous limestone, calcareous mudstone, argillaceous sandstone, and carbonaceous mudstone (Table 1). In addition, there are also small amounts of basalt and fault breccia. Therefore, the others category primarily reflects the properties of mudstone. For convenience, we can call it others in the subsequent text. Coal seams, sandstone, and limestone generally exhibit distinct logging curve characteristics and can be classified individually. In contrast, mud-bearing sandstone and limestone contain more complex information, making them challenging to differentiate using well logging curves alone; hence, they were combined into a single category.

Although the different wells have varying types of well logging curves, most include normal resistivity (NR), natural gamma (GR), and gamma-gamma (GG) density logging data. Thus, we utilized these three types of logging data as inputs for our model. This approach offers the advantage of making the trained model applicable to a broader range of scenarios.

Normal resistivity logging involves measuring the electrical resistivity of the formations by sending an electric current through the rock and measuring the resulting voltage. The resistivity is inversely related to the conductivity of the rock, which reflects the amount of water and the type of minerals present. The resistivity measurements can indicate the presence of hydrocarbons and the salinity of the formation water. High resistivity typically suggests low water content or the presence of hydrocarbons, while low resistivity indicates high water content or clay-rich formations.

Natural gamma ray logging detects gamma rays emitted spontaneously by radioactive elements (such as clays, uranium, potassium, and thorium) in rock formations. It is often measured in picoamperes (PA) per kilogram of rock (Unit: 10 pA/kg). Thus, "10 pA/kg" indicates that the gamma ray intensity produces a current of 10 picoamperes per kilogram of rock. Variations in gamma ray intensity help determine the fractional volume of radioactive elements such as clay. Therefore, the formations containing mudstone typically exhibit higher natural gamma ray intensity.

In addition, gamma-gamma logging is used to measure the mass density of subsurface rock formations by detecting gamma rays counts per second (CPS). It is also known as "gamma-gamma density logging" or simply "density logging". The detected gamma ray counts per second are inversely proportional to the density of the rock. Higher CPS indicates lower density, while lower CPS indicates higher density.

The depth interval for the logging curves is 0.05 m. Previous studies typically treat information from different logging curves at the same depth as a single input sample, which is a straightforward and natural approach. However, this method only considers numerical values without accounting for their trends. For instance, coal beds, which have a lower density, often correspond to minimum values in gamma-gamma logging. Therefore, we obtained data from every 40 adjacent sampling points (2 m in length) as a whole to form a single input sample. After the local interception of all logging curves, a total of 70,756 samples were generated, of which 52,756 samples were used for training and 18,000 samples were used for testing.

Additionally, we normalized the data to ensure that the values fall within a reasonable range, which is crucial for the convergence and stability of neural network training. To preserve the information of the data trends, normalization is carried out in two steps. First, we normalize the raw logging data for each well (Figure 3b), and then we normalize each sample (i.e., 40 sampling points) (Figure 3c). It is crucial to perform regularization in two steps. First, if only the first step is performed, it can cause data with rapid local variations to become slow changes. For example, when there are very large local values in a logging curve, after normalization, the other data may become very small, failing to reflect the data variations. If only the second step is performed, it may cause significant changes in data with slow local variations. As long as the maximum and minimum values of the data are not equal, their range will always be normalized to the 0–1 interval. By combining the data from both normalizations in each well and each sample, we generated a total of six input

curves (three types of curves and two normalizations), each with 40 sampling points, to serve as the input for each sample (Figure 3).

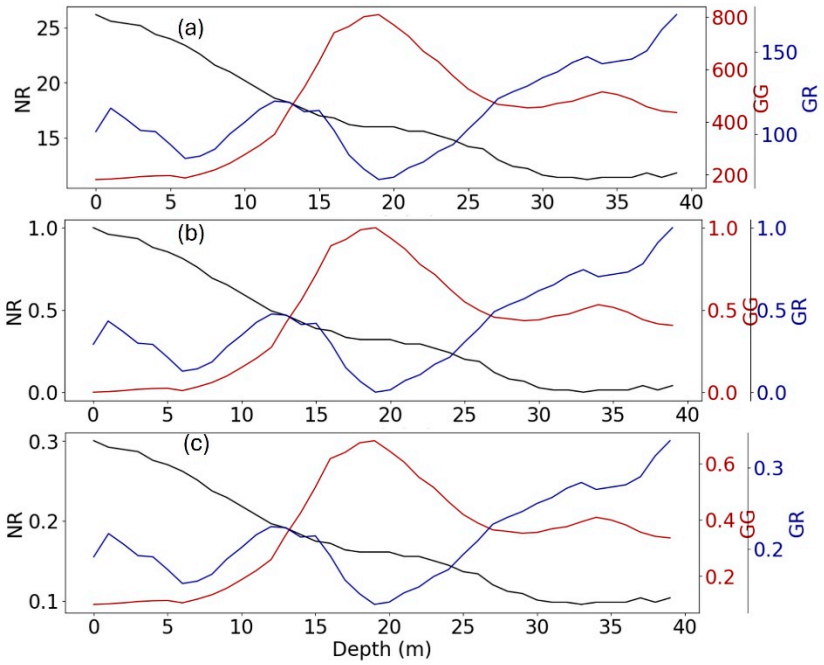


Figure 3. Sample data processing procedure: (**a**) raw data, (**b**) normalization of the individual samples, (**c**) normalization of well logging data. NR, GG, and GR represent normal resistivity (NR, unit: Ω·m), natural gamma (GR, unit: 10 pA/kg), and gamma-gamma (GG, unit: CPS) density logging data, respectively.

3. Method for Lithology Classification

Logging curves reflect the physical properties of rocks. For the same area, we assume the rocks with similar physical properties to belong to the same type. However, the physical properties of rocks are influenced by various factors, such as purity, borehole size, and fluids, making it very challenging to identify lithology based on logging curves manually. This process is subject to significant uncertainty and subjectivity. Observing core samples is the most direct and accurate method, but it is time-consuming and costly.

The essence of machine learning (including traditional machine learning algorithms and deep learning) is to establish a nonlinear relationship between the observed data and the predicted outcomes from a statistical perspective, based on large amounts of data. Unlike physical models, once training is complete, machine learning can achieve rapid predictions within seconds, making it suitable for well log curve identification. Recently, more studies have attempted to compare the performance of machine learning methods for lithology identification [23–25]. The results showed that Random Forest had the best score among the algorithms considered. Additionally, CNN and RNN are classic sample models suited for small sample sets. Therefore, we tried to determine whether the neural network models, which are more adept at image recognition, outperform Random Forest models in well log curve identification.

3.1. Convolutional Neural Networks

CNNs use convolution operations to capture local features, which is especially useful for image processing as it can effectively identify edges, textures, and other local structures [26]. The weight sharing mechanism in convolution layers significantly reduces the number of parameters, lowering the model's complexity and computational cost. Furthermore, convolution operations allow the model to be invariant to translations in the input data, meaning it can recognize the same features even if the image is shifted spatially. As a consequence, CNNs excel at handling spatial information related to images [27]. The well

logging signals used in this study can be considered as multiple one-dimensional images stacked together, making them suitable for CNNs.

The CNN model utilized in this study for predicting lithology is depicted in Figure 4. The model's input comprises six normalized well logging curves, each containing 40 continuous sampling points. The input features are processed through three alternating convolutional and pooling layers for effective feature extraction. Following each convolution operation, batch normalization is applied to standardize the data and mitigate the risk of gradient vanishing, while the ReLU activation function defined as $\mathrm{ReLU}(x) = \max(x, 0)$ is employed to enhance the model's non-linearity [28]. Following the convolutional layers, two fully connected networks are employed to perform dimensionality reduction on the extracted features, aligning them with the output categories. The softmax function defined as $\mathrm{Softmax}(x) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_i}}$ is used in the output layer of neural networks, which not only provides the probability for each class but also ensures that these probabilities sum to 1, making it easy to make decisions based on the model's confidence. At last, the model's output utilizes the cross-entropy loss function [26,29]:

$$Loss = -\frac{1}{m}\sum_{k=1}^{m}\sum_{i=0}^{n} p'_{i,k}\log(p_{i,k})$$

where $p'_{i,k}$ is the reference catalogue of hot vector, and each element in the vector is a binary value (0 or 1), with exactly one element set to 1 and the rest set to 0. The $p_{i,k}$ is the prediction of CNN in the range of [0, 1] for k th sample. Here, $n = 4$ denotes the number of output neurons, m is the trained number of samples for each training loop.

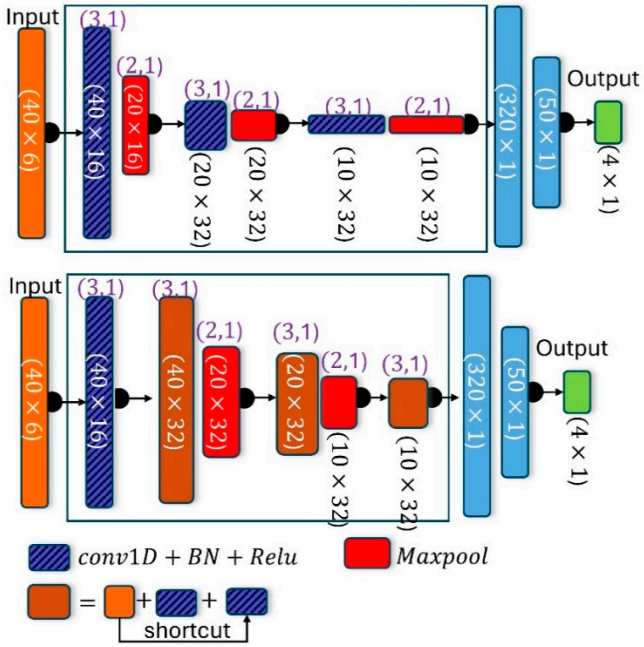


Figure 4. Architecture of CNN and ResNet. Blue blocks denote the convolutional layers, red blocks denote the max-pooling layers and blue-gray blocks the fully connected layers. Brown represents residual module. The input (orange) and output (green) parameters are 400 points of logging data and probability distribution over four labels. The number of neurons and filters in each layer is shown on the bottom or inner part of each block, and kernel sizes and stride in convolutional or max-pooling layers are shown on the top of each block.

3.2. Residual Neural Network

A ResNet is essentially deep CNNs with a specific architecture modification. It introduces residual blocks that allow information to flow more easily through the network,

addressing the issue of vanishing gradients in deep networks (Figure 4). Skip connections in ResNets make it possible to build much deeper neural networks, capturing more complex features and patterns [30]. Therefore, ResNets perform exceptionally well in many tasks, such as image classification, object detection, and image segmentation, due to their ability to learn more effective feature representations. The input and output layers, as well as loss functions of the residual network are consistent with those of CNNs. This study utilizes a residual network to explore whether a more advanced model can improve prediction performance under the given training set conditions.

3.3. Random Forest

Random Forest is an ensemble learning method used primarily for classification and regression tasks [31]. It builds multiple decision trees during training and merges their outputs to improve the predictive performance and control overfitting. Decision trees make splits based on specific feature values, and these splits are independent of the scale of the features. Each tree considers each feature separately when making splits. As a consequence, Random Forests can handle high-dimensional data without the need for feature scaling or normalization.

To investigate the predictive performance of different DL or machine learning models, all inputs in three models underwent the same data preprocessing procedure. Since the Random Forest can only handle one-dimensional data, the inputs for the neural networks needed to be flattened. The output and loss function of the Random Forest are identical to those of the first two models.

4. Model Training

Training a model is an optimization problem equivalent to finding the minima of the loss function, which measures how 'good' the architecture is. During the training process, all neuron weights in each layer are updated using the backpropagation algorithm, Adam [32]. An L2 regularization technique is employed in all the convolutional and fully connected layers to prevent overfitting, limit model complexity, and thereby improve the model's generalization ability [33]. The random batch gradient descent method is used to improve the generalization and reduce the risk of overfitting [34]. A total of 300 samples are randomly selected for each training epoch, and 50 iterations are required. Each model was trained using a GPU (NVIDIA GeForce RTX 3070, NVIDIA Corporation, Santa Clara, CA, USA) of desktop workstation (Processor: 12th Gen Intel(R) Core(TM) i9-12900K 3.20 GHz, Memory: 64 GB, Intel Corporation, Santa Clara, CA, USA), with the 200 epoch training process taking approximately 5 min.

Figure 5 shows the evolution of accuracy and recall rate with the number of iterations for the training/test datasets. As the number of epochs increased, the loss on the training set gradually decreased, while the loss on the test set initially decreased but then gradually increased, indicating the occurrence of overfitting. This overfitting is attributed to the small and homogeneous sample size and, for each depth sampling point, we extracted 40 m of data, leading to many samples with similar information. A viable solution is to employ early stopping, as indicated by the gray vertical line in Figure 5, to enhance the model's predictive accuracy and generalization ability.

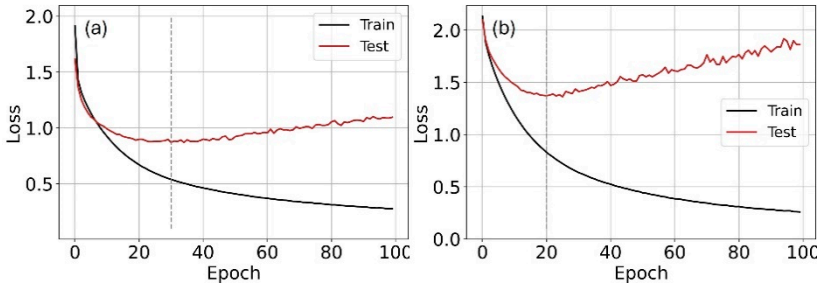


Figure 5. The loss evolution of CNN (**a**) and RESNET (**b**).

In the Random Forest algorithm, the parameter estimators and random states are essential [30]. The number of estimators defines the number of trees in the forest. Generally, increasing the number of trees enhances model performance; however, beyond a certain threshold, adding more trees results in diminishing returns. The random state parameter sets the seed for the random number generator, ensuring the reproducibility of the results. The selection of the number of estimators and random states has a significant impact on the model's performance. In this study, optimal values for these parameters were determined through grid search based on their performance on the test set [35] (Figure 6).

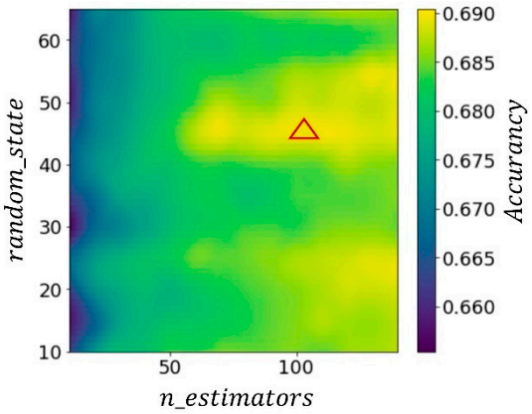


Figure 6. The distribution of test set accuracy under different estimators and random states, with the red triangle indicating the location of the highest accuracy.

5. Result and Discussion

The common metrics used to evaluate a model's performance include precision, recall, and F1-Score. Precision is the proportion of correctly predicted positive samples to the total number of samples predicted as positive, defined as follows [36]:

$$Precision = \frac{TP}{TP + FP}$$

where TP and FP are true positive and false positive. TP is the number of instances correctly classified as positive, while FP is the number of instances incorrectly classified as positive.

Recall, also known as sensitivity, is the proportion of correctly predicted positive samples to the total number of actual positive samples.

$$Recall = \frac{TP}{TP + FN}$$

The F1-Score is the harmonic mean of precision and recall, providing a single metric that balances the two [37].

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

We evaluated the performance of various trained models on the test set, with the statistical distributions presented in Figure 7. It is clear that the predictions of other types exhibit the best performance due to the substantial samples, with both the precision and recall for all three models approaching 90%. Following this, the precision and recall of coal seams are approximately 80% across the three models. Sandstone typically achieves precision levels of above 70%, but its lower recall suggests that the models may overlook sandstone instances. Despite the greater number of sandstone samples compared to coal, its performance is inferior, indicating that the three types of well logs are more sensitive to coal seams. Limestone shows the poorest performance, characterized by lower precision due to the limited number of limestone samples, which account for only about one-tenth of the total sample size. Nevertheless, limestone exhibits a high recall, suggesting that all

three models are capable of identifying limestone structures of sufficient thickness (e.g., Figure 8).

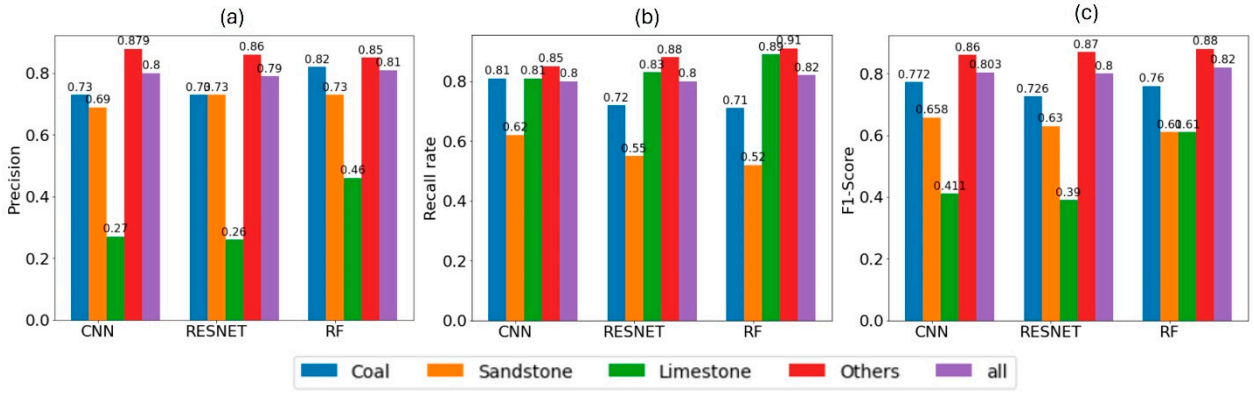


Figure 7. Lithofacies prediction performance on test dataset. (**a**) CNN (**b**) RESNET (**c**) RF.

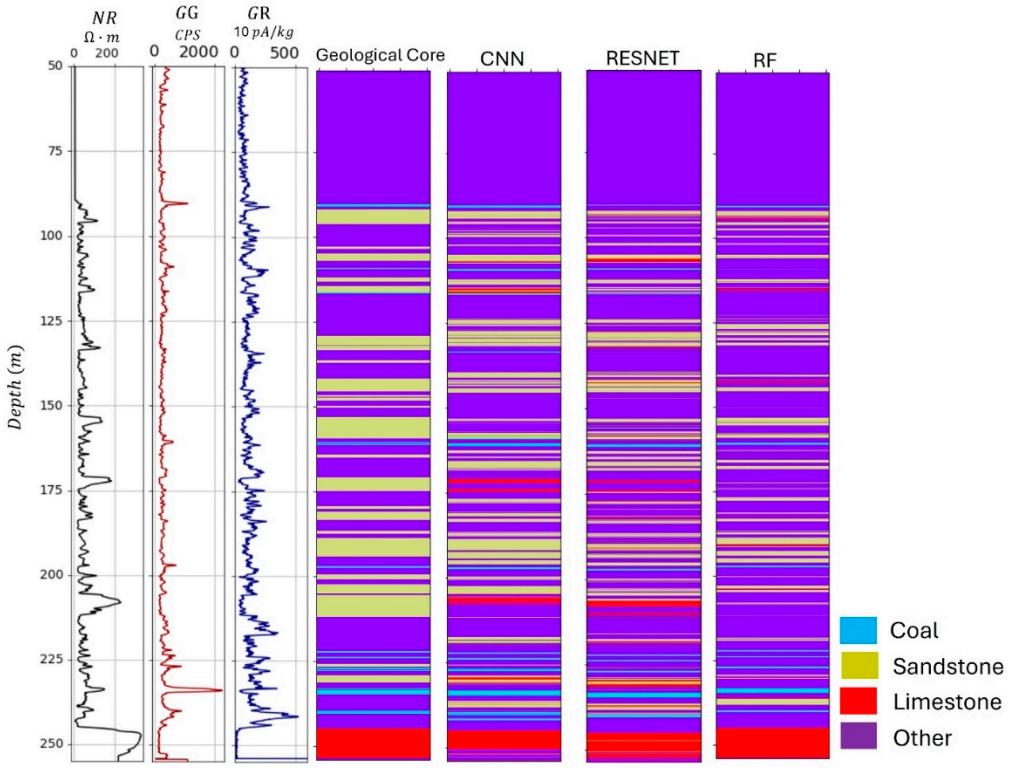


Figure 8. Comparison of predictions of different models and input logs in well No. 1.

To further assess the performance of the models, we employed additional data from three wells (No. 1, No. 2, No. 3) for lithology classification. These data are independent of both the training and test sets, allowing us to evaluate the models' generalization capabilities.

In Well No. 1 (Figure 8), the predictions from all three models are generally consistent. Coal seams are characterized by intermediate resistivity and low density; sandstone is associated with high resistivity, high density, and low gamma values; and limestone, while similar to sandstone, exhibits higher resistivity. Argillaceous' features are more complex, with a natural gamma content varying with clay content, generally higher density, and lower resistivity. All three models demonstrate high accuracy in predicting coal. Among the predictions, the primary sandstone interfaces are largely aligned. However, some sandstone layers are mistakenly identified as limestone due to the similarities in high

resistivity, high density, and low radioactivity between sandstone and limestone in well logs, despite limestone's higher resistivity. The models effectively identify thick limestone layers in deeper sections, indicating a high recall for limestone detection. Nevertheless, the CNN and ResNet exhibit lower precision in limestone identification, while RF achieves higher precision for limestone but shows relatively lower precision and recall for sandstone compared to the neural networks (Table 2).

Table 2. Precision, Recall, and F1-Score of different lithofacies predictions for Well No. 101. Black, blue and red correspond to CNN, RESNET and RF respectively.

Type	Coal	Sandstone	Limestone	Others
Number	131	1101	173	2565
Precision	0.70/0.85/0.94	0.57/0.61/0.59	0.46/0.46/0.74	0.81/0.80/0.76
Recall rate	0.88/0.75/0.67	0.46/0.41/0.33	0.66/0.78/0.97	0.84/0.87/0.90
F1-Score	0.78/0.80/0.78	0.51/0.49/0.42	0.54/0.57/0.84	0.82/0.83/0.82

The prediction results for Well No. 2 are comparable to those for Well No. 1 (Figure 9). Within the limestone and sandstone intervals above 650 m, there are notable errors in predicting limestone and sandstone. The presence of interbedded argillaceous stone with low resistivity causes significant fluctuations in the resistivity curve for this region, while the GG and GR curves remain relatively stable, consistent with the characteristics of limestone. Compared to Well No. 1, Random Forest demonstrates lower accuracy in predicting limestone in Well No. 2, with considerable discrepancies from the test set results (Table 3). This highlights the uncertainty in RF's predictions across different wells, whereas the neural network predictions exhibit greater consistency.

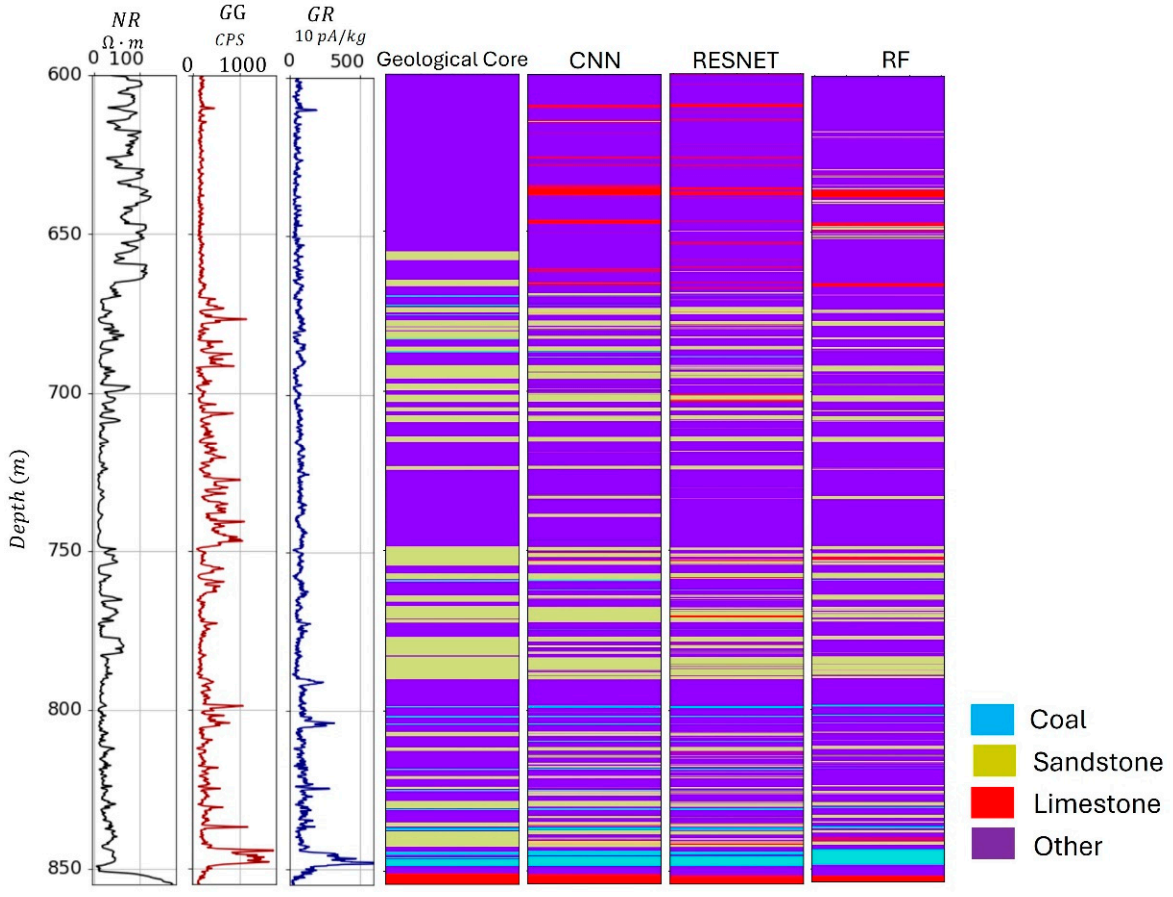


Figure 9. Comparison of predictions of different models and input logs in well No. 2.

Table 3. The same as Table 1 but for Well No. 2.

Type	Coal	Sandstone	Limestone	Others
Number	177	1363	67	3454
Precision	0.74/0.75/0.79	0.82/0.89/0.79	0.25/0.2/0.23	0.87/0.84/0.81
Recall rate	0.83/0.68/0.66	0.68/0.57/0.5	0.86/0.74/0.56	0.88/0.92/0.91
F1-Score	0.79/0.71/0.72	0.74/0.7/0.61	0.39/0.32/0.33	0.88/0.88/0.86

The top section of Well No. 3 exhibits characteristics similar to those of Well No. 2 (Figure 10, Table 4); however, the overall prediction performance for Well No. 3 is superior to that of Wells No. 1 and No. 2. The accuracy, recall, and F1-Score for Well No. 3 are generally higher than those for Wells No. 1 and No. 2. In addition, there is no obvious correlation between geological age and logging data. This indicates a significant variability in our models' prediction performance across the different wells, highlighting the need for further data expansion to achieve a more generalized model. Additionally, the performances of the CNN and ResNet are comparable, suggesting that more complex models do not provide a significant advantage for feature extraction from the current training set. While RF performs similarly to the neural networks in predicting sandstone, others, and coal, it outperforms the neural networks in predicting limestone, but its performance is less stable.

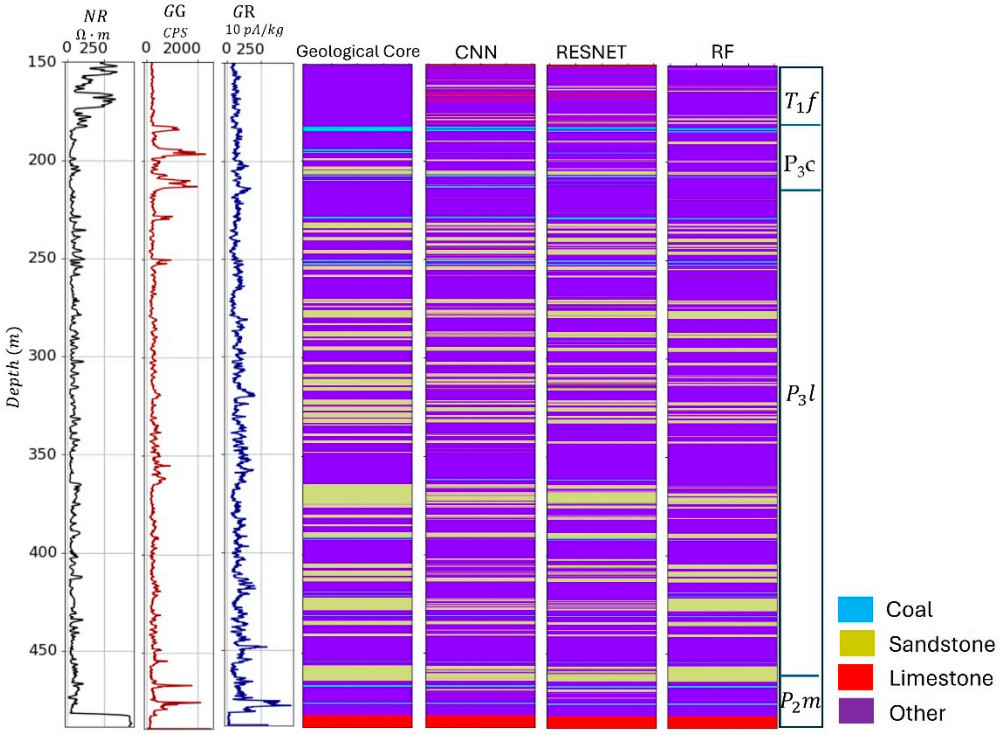


Figure 10. Comparison of predictions of different models and input logs in well No. 3.

Table 4. The same as Table 1 but for Well No. 3.

Type	Coal	Sandstone	Limestone	Others
Number	221	1835	123	4557
Precision	0.91/0.77/1.0	0.64/0.85/0.89	0.62/0.65/1.0	0.87/0.85/0.87
Recall rate	0.65/0.63/0.63	0.51/0.65/0.69	0.81/0.97/1.0	0.93/0.93/0.96
F1-Score	0.75/0.70/0.78	0.57/0.74/0.78	0.70/0.78/1.0	0.90/0.89/0.91

It is important to note that borehole diameter and mud density have a significant impact on the count rate (CPS) of a single detector, and resistivity logging reflects the

water content or mineral composition of the rock and has little information about lithology. However, we believe that resistivity and count rate still have a correlation with lithology, at least for the area where the data was trained. Although there is significant uncertainty in the correspondence between the data and lithology, this is precisely the kind of problem that deep learning aims to solve. Due to the considerable uncertainty and the limited types of logging data available, our prediction accuracy for sandstone and limestone is not very high. Nevertheless, our deep learning model provides a preliminary lithology classification model, which is meaningful for identifying target layers using logging curves.

Figure 9 shows the bottom coal with a high gamma ray, but the coal seams typically exhibit low natural gamma values. In our work area, which is mainly characterized by Permian deltaic to coastal facies coal, the most significant difference compared to coal seams in most other regions is that the bottom coal seam is underlain by a layer of clay. This clay is a weathered product of igneous rock, which affects the overlying coal seam, causing it to exhibit a certain level of natural radioactivity, leading to higher gamma values in the coal seam. Additionally, coal seams usually show high resistivity; however, when the coal seam is impure and contains moist clay, such as bottom coal in Figure 9, it can also result in lower resistivity.

6. Conclusions

In this study, we collected 70,756 samples from 22 wells in Lianghe Town and Shiping Town in the Southern Sichuan region to establish a generalized lithofacies classification model. We used three machine learning algorithms to classify stratigraphy based on NR, GG, and GR logging data. All the algorithms demonstrated high accuracy in identifying coal and other types mainly composed of argillaceous rocks, with precision and recall exceeding 80%, and showed good transferability to the well log data from different regions.

Among the three models, RF performed slightly better on the test dataset compared to the two neural network models but exhibited greater uncertainty on the validation set. The precision of all three models in identifying sandstone ranged from 60% to 80% or higher, but the recall was relatively low at between 50% and 70%. When sandstone was misclassified, it was usually identified as limestone. This is primarily because sandstone and limestone share similar logging characteristics—high resistivity, low gamma intensity, and high density, although the limestones have relatively higher resistivity and density than sandstone. The CNN and ResNet demonstrated a low precision in identifying limestone but a high recall. This is due to the similar logging curve characteristics between limestone and sandstone, and the insufficient representation of limestone samples, which constitute less than 10% of the total training samples, leading to the inadequate recognition of limestone features.

Overall, while deep learning models have shown great potential in image recognition, they do not currently exhibit a significant advantage over RF in identifying one-dimensional well log data. This may be due to the insufficiency of the current training samples and the simplicity of the input, which consists of only three curves. In theory, as the complexity of input data features increases and the number of samples grows, deep learning will gradually demonstrate its advantages [35]. It is foreseeable that, with an increasing number of samples, we can achieve a more accurate and generalized lithofacies classification model, which will be the focus of future work.

Author Contributions: Conceptualization, Y.S. and L.G.; methodology, L.G., Y.S., and R.T.; validation, Y.S., J.L., and R.T.; writing—original draft preparation, Y.S. and L.G.; writing—review and editing, J.L. and R.T.; visualization, J.L.; supervision, J.L.; funding acquisition, L.G. and Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Sichuan Provincial Science and Technology Innovation and Entrepreneurship Seedling Project (2024JDRC0023) and Huzhou Public Welfare Research Project (2023GZ17).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The open-source Python code (Python version 3.9) and data can be obtained from the following GitHub repository: https://github.com/Rongjiang007/pylitho.

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

References

1. Buryakovsky, L.; Chilingar, G.V.; Rieke, H.H.; Shin, S. *Fundamentals of the Petrophysics of Oil and Gas Reservoirs*; Wiley: Hoboken, NJ, USA, 2012; p. 400.
2. Xie, Y.; Zhu, C.; Lu, Y.; Zhu, Z. Towards Optimization of Boosting Models for Formation Lithology Identification. *Math. Probl. Eng.* **2019**, *2019*, 5309852. [CrossRef]
3. Sun, Z.; Jiang, B.; Li, X.; Li, J.; Xiao, K. A Data-Driven Approach for Lithology Identification Based on Parameter-Optimized Ensemble Learning. *Energies* **2020**, *13*, 3903. [CrossRef]
4. Chaki, S.; Takarli, M.; Agbodjan, W.P. Influence of thermal damage on physical properties of a granite rock: Porosity, permeability and ultrasonic wave evolutions. *Constr. Build. Mater.* **2008**, *22*, 1456–1461. [CrossRef]
5. Borsaru, M.; Zhou, B.; Aizawa, T.; Karashima, H.; Hashimoto, T. Automated lithology prediction from PGNAA and other geophysical logs. *Appl. Radiat. Isot.* **2006**, *64*, 272–282. [CrossRef] [PubMed]
6. Oyler, D.C.; Mark, C.; Molinda, G.M. In situ estimation of roof rock strength using sonic logging. *Int. J. Coal Geol.* **2010**, *83*, 484–490. [CrossRef]
7. Kumar, T.; Seelam, N.K.; Rao, G.S. Lithology prediction from well log data using machine learning techniques: A case study from Talcher coalfield, Eastern India. *J. Appl. Geophys.* **2022**, *199*, 104605. [CrossRef]
8. Zhao, Z.; He, Y.; Huang, X. Study on Fracture Characteristics and Controlling Factors of Tight Sandstone Reservoir: A Case Study on the Huagang Formation in the Xihu Depression, East China Sea Shelf Basin, China. *Lithosphere* **2021**, *2021*, 3310886. [CrossRef]
9. Lu, X.; Sun, D.; Xie, X.; Chen, X.; Zhang, S.; Zhang, S.; Sun, G.; Shi, J. Microfacies characteristics and reservoir potential of Triassic Baikouquan Formation, northern Mahu Sag, Junggar Basin, NW China. *J. Nat. Gas. Geosci.* **2019**, *4*, 47–62. [CrossRef]
10. Colombo, D.; Turkoglu, E.; Li, W.; Sandoval-Curiel, E.; Rovetta, D. Physics-driven deep-learning inversion with application to transient electromagnetics. *Geophysics* **2021**, *86*, E209–E224. [CrossRef]
11. Noakoasteen, O.; Wang, S.; Peng, Z.; Christodoulou, C. Physics-informed deep neural networks for transient electromagnetic analysis. *IEEE Open J. Antennas Propag.* **2020**, *1*, 404–412. [CrossRef]
12. Tang, R.; Li, F.; Shen, F.; Gan, L.; Shi, Y. Fast Forecasting of water-filled bodies position using transient electromagnetic method based on deep learning. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 4502013. [CrossRef]
13. Gan, L.; Wu, Q.; Huang, Q.; Tang, R. Quality classification and inversion of receiver functions using convolutional neural network. *Geophys. J. Int.* **2023**, *232*, 1833–1848. [CrossRef]
14. Al-Anazi, A.; Gates, I.D. A support vector machine algorithm to classify lithofacies and model permeability in heterogeneous reservoirs. *Eng. Geol.* **2010**, *114*, 267–277. [CrossRef]
15. Corina, A.N.; Hovda, S. Automatic lithology prediction from well logging using kernel density estimation. *J. Pet. Sci. Eng.* **2018**, *170*, 664–674. [CrossRef]
16. El Sharawy, M.S.; Nabawy, B.S. Determination of electrofacies using wireline logs based on multivariate statistical analysis for the Kareem Formation, Gulf of Suez, Egypt. *Environ. Earth Sci.* **2016**, *75*, 1–15. [CrossRef]
17. Kuroda, M.C.; Vidal, A.C.; Leite, E.P.; Drummond, R.D. Electrofacies characterization using self-organizing maps. *Braz. J. Geophys.* **2012**, *30*, 287–299. [CrossRef]
18. Horrocks, T.; Holden, E.J.; Wedge, D. Evaluation of automated lithology classification architectures using highly-sampled wireline logs for coal exploration. *Comput. Geosci.* **2015**, *83*, 209–218. [CrossRef]
19. Roslin, A.; Esterle, J.S. Electrofacies analysis using high-resolution wireline geophysical data as a proxy for inertinite-rich coal distribution in Late Permian Coal Seams, Bowen Basin. *Int. J. Coal Geol.* **2015**, *152*, 10–18. [CrossRef]
20. Schmitt, P.; Veronez, M.R.; Tognoli, F.M.W.; Todt, V.; Lopes, R.D.C.; Silva, C.A.U.D. Electrofacies modelling and lithological classification of coals and mud-bearing fine-grained siliciclastic rocks based on neural networks. *Earth Sci. Res.* **2013**, *2*, 193–208. [CrossRef]
21. Xie, Y.; Zhu, C.; Zhou, W.; Li, Z.; Liu, X.; Tu, M. Evaluation of machine learning methods for formation lithology identification: A comparison of tuning processes and model performances. *J. Pet. Sci. Eng.* **2018**, *160*, 182–193. [CrossRef]
22. Zhou, B.; O'Brien, G. Improving coal quality estimation through multiple geophysical log analysis. *Int. J. Coal Geol.* **2016**, *167*, 75–92. [CrossRef]
23. Bhattacharya, S.; Mishra, S. Applications of machine learning for facies and fracture prediction using Bayesian Network Theory and Random Forest: Case studies from the Appalachian basin, USA. *J. Pet. Sci. Eng.* **2018**, *170*, 1005–1017. [CrossRef]
24. Gu, Y.; Bao, Z.; Rui, Z. Complex lithofacies identification using improved probabilistic neural networks. *Petrophysics* **2018**, *59*, 245–267. [CrossRef]
25. Maxwell, K.; Rajabi, M.; Esterle, J. Automated classification of metamorphosed coal from geophysical log data using supervised machine learning techniques. *Int. J. Coal Geol.* **2019**, *214*, 103284. [CrossRef]
26. Buduma, N.; Buduma, N.; Papa, J. *Fundamentals of Deep Learning*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2022.

27. Gupta, J.; Pathak, S.; Kumar, G. Deep learning (CNN) and transfer learning: A review. *J. Phys. Conf. Ser.* **2022**, *2273*, 012029. [CrossRef]
28. Banerjee, C.; Mukherjee, T.; Pasiliao, E., Jr. An empirical study on generalizations of the ReLU activation function. In Proceedings of the 2019 ACM Southeast Conference, Kennesaw, GA, USA, 18–20 April 2019; pp. 164–167.
29. Gan, L.; Tang, R.; Li, F.; Shen, F. A Deep learning estimation for probing Depth of Transient Electromagnetic Observation. *Appl. Sci.* **2024**, *14*, 7123. [CrossRef]
30. Targ, S.; Almeida, D.; Lyman, K. Resnet in resnet: Generalizing residual architectures. *arXiv* **2016**, arXiv:1603.08029.
31. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
32. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
33. Cortes, C.; Mohri, M.; Rostamizadeh, A. L2 regularization for learning kernels. *arXiv* **2012**, arXiv:1205.2653.
34. Keskar, N.S.; Mudigere, D.; Nocedal, J.; Smelyanskiy, M.; Tang, P.T.P. On large-batch training for deep learning: Generalization gap and sharp minima. *arXiv* **2016**, arXiv:1609.04836.
35. Ramadhan, M.M.; Sitanggang, I.S.; Nasution, F.R.; Ghifari, A. Parameter tuning in random forest based on grid search method for gender classification based on voice frequency. *DEStech Trans. Comput. Sci. Eng.* **2017**, *10*, 625–629. [CrossRef] [PubMed]
36. Janiesch, C.; Zschech, P.; Heinrich, K. Machine learning and deep learning. *Electron. Mark.* **2021**, *31*, 685–695. [CrossRef]
37. Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **2020**, *21*, 1–13. [CrossRef]

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