A Hybrid CNN-LSTM Deep Learning Framework for Porosity Prediction in Carbonate Reservoirs

Amirreza Mehrabi[1], Majid Bagheri[2], Majid Nabi Bidhendi[2], Ebrahim Biniaz Delijani[1], Mohammad Behnoud[1]

[1]*Department of Petroleum Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.*
[2]*Institute of Geophysics, University of Tehran, Tehran, Iran.*

ABSTRACT

Porosity estimation in carbonate reservoirs is crucial for hydrocarbon exploration and production. Traditional methods for porosity estimation, such as empirical correlations and geostatistical techniques, often rely on limited geological knowledge and may not capture the complex relationships between porosity and various geological features. This research aims to develop a novel hybrid Convolutional Neural Network (CNN)-Long Short-Term Memory (LSTM) deep learning framework for porosity prediction in carbonate reservoirs. The proposed framework aims to leverage the strengths of both CNNs and LSTMs to capture both spatial and temporal patterns in well log data, thereby improving the accuracy and reliability of porosity estimation. The combination of CNNs and LSTMs in a deep learning architecture allows for the efficient extraction of both local spatial features and long-term temporal dependencies in well log data. The model prediction performance improved from 0.67 (for MLP) to 0.98 for LSTM, indicating the accuracy of the model. The results suggest that the CNN-LSTM model can accurately estimate porosity in heterogeneous carbonate reservoirs, and its ability to capture spatial and temporal dependencies makes it well-suited for modeling complex geological systems.

Keywords: Carbonate, porosity estimation, deep learning, CNN, LSTM, well data

INTRODUCTION

The reservoir characterization in carbonate reservoirs is challenging due to their complex pore structures and inherent heterogeneities. Porosity is determined by core analysis, well logs, and seismic data[1, 2]. Core analysis is considered the most reliable method for determining porosity because it directly measures pore space. However, core analysis is expensive and time-consuming. Well logs indirectly measure porosity by analyzing rock's electrical, acoustic, or nuclear properties. The most common well logs for porosity estimation are neutron, density, and sonic logs[3, 4].

Numerous studies have utilized machine learning algorithms, such as support vector machines, random forests, and artificial neural networks, to predict porosity in carbonate reservoirs[5]. These methods have demonstrated promising results by incorporating well-log data, seismic attributes, and geologic information[6]. Although machine learning has been used in reservoir characterization, the complexity of the data in carbonate reservoirs requires more advanced techniques to extract meaningful insight. Therefore, recent studies have focused on hybrid machine learning and deep learning to improve the prediction accuracy of machine learning[3, 7].

This work presents the hybrid deep learning model for porosity estimation in heterogeneous carbonate reservoirs. This hybrid architecture harnesses the strengths of both CNN and LSTM networks, enabling the capture of spatial and temporal dependencies in the input well data. The proposed method lies in its ability to learn a representation of the complex data pattern and structure that is specific to the heterogeneous carbonate reservoirs. The porosity estimation problem is formulated as a regression problem, where the input is a well log, and the output is a core analysis porosity. The input well log can be preprocessed using CNN layers to extract relevant features, followed by LSTM layers to model the spatial dependencies within the dataset. The final output can then be generated using a fully connected layer. The input to the CNN-LSTM model is a sequence of log measurements, and the output is the predicted porosity value. The

input data includes Gamma ray (GR), travel-time (DT), NPHI, ROHB, LLB logs, and core analysis data from a carbonate reservoir. The CNN-LSTM model was implemented using the TensorFlow platform in Python language programming. The model's hyperparameters such as the number of hidden layers and epochs, learning rate, and batch size were optimized using the Keras Tuner library's optimization framework.

Methodology

The study utilizes a comprehensive dataset comprising well log data and core data from carbonate reservoirs. The hybrid CNN-LSTM model developed for porosity estimation in carbonate reservoirs consists of two main components: a CNN for feature extraction and an LSTM for capturing temporal dependencies. The porosity estimation with hybrid CNN-LSTM Deep Learning Framework can be summarized as follows

- Data Preprocessing: collect and normalize the data.
- Feature Engineering: Extract relevant features from the well-log data.
- Data Splitting: Split the dataset into training, validation, and test sets.
- Training Initialization: Initialize batch size and the number of epochs.
- Training and Validation: Train the hybrid CNN-LSTM model on the training set.
- Model Evaluation: Calculate porosity estimates using the trained model.
- Result Analysis: Analyze the model's predictions compared to actual porosity values

The CNN is designed to learn and extract relevant features from the input well log data, capturing spatial relationships and patterns indicative of porosity variations within the reservoir.

The outputs of the CNN are then fed into the LSTM, which is responsible for capturing the temporal dependencies and generating the final porosity estimations. The LSTM network is designed to effectively model long-term dependencies and sequential patterns within the input data, enabling the model to learn and predict porosity values based on the complex interactions among the petrophysical properties derived from the well logs.

The CNN-LSTM architecture (as shown in Figure 1) was specifically formulated for the regression prediction problem, where the goal is to estimate continuous porosity values based on the input well data.

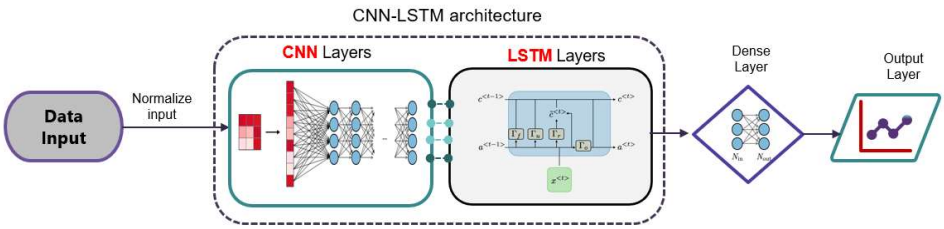

Figure 1- CNN-LSTM model architecture

The architecture consists of a 1D convolutional layer followed by a max pooling layer, an LSTM layer, a dropout layer to prevent overfitting, a time-distributed dense layer with a sigmoid activation function, and a flatten layer followed by a dense layer with a linear activation function. We compile the model using the Adam optimizer and mean squared error as the loss function, and train it on the training set for a specified number of epochs. Also, we use Multi-Layer Perceptron (MLP) neural network as a conventional machine learning model to compare the results of the proposed hybrid deep learning model. MPL is an artificial neural network consisting of multiple layers of interconnected nodes, each representing a mathematical function.

Prior to model training, extensive preprocessing is performed to clean and normalize the well log and core data. This includes handling missing values, scaling features, and ensuring data consistency to facilitate effective learning by the CNN-LSTM model. The well-log data includes

Gamma-ray (GR), compressional sonic travel-time (DT), bulk density (RHOB), neutron porosity (NPHI), and deep resistivity (LLD), as shown in Figure 2.

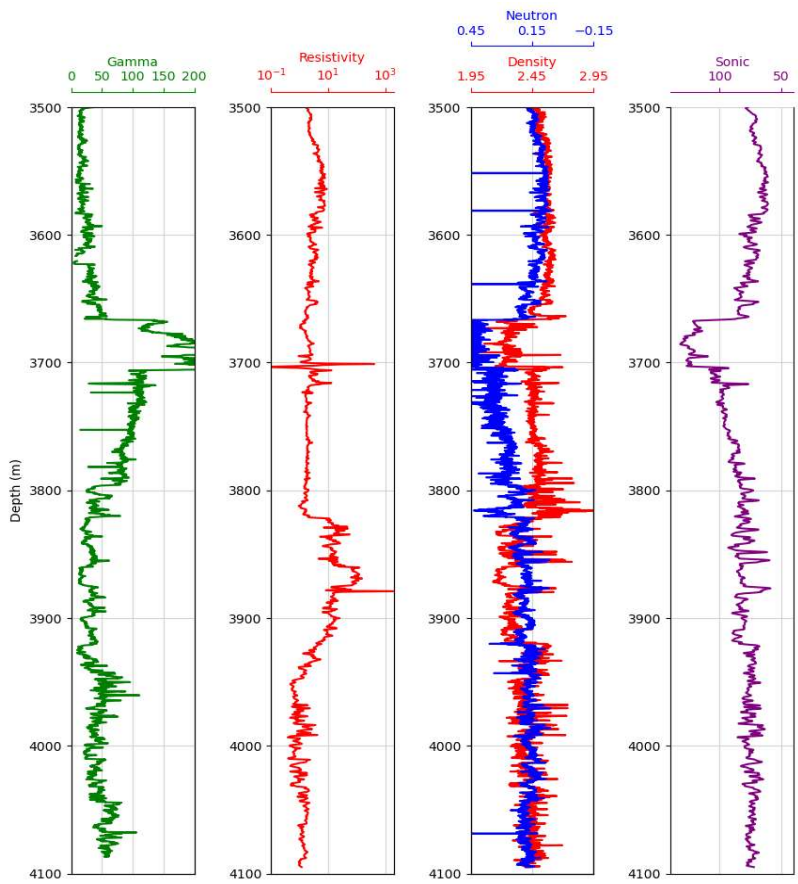

Figure 2-Well log data

The hybrid CNN-LSTM model was trained and evaluated using well dataset, including well logs and core data.

Results and Discussion

We first used the MPL neural network to train a dataset of well-log data and corresponding porosity values to learn the relationship between these variables. The input layer of the network receives the well log data as inputs, which include measurements of gamma-ray, sonic, neutron, resistivity, and density logs data. The output layer of the network produces a predicted porosity value. During the training of this model, the weights of the connections between nodes are adjusted to minimize the difference between the predicted porosity values and the actual porosity values in the training dataset. Once trained, the MPL neural network was used to predict the porosity of new well-log data that it has not seen before. The performance of the MLP model is shown in Figure 3.

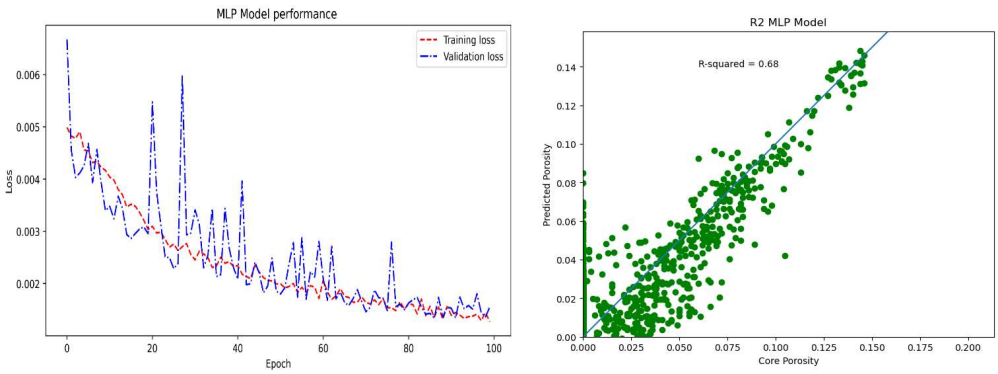

Figure 3- MLP model performance

The results of MLP model are poor as can be seen from the correlation coefficient between the model and core data. The performance evaluation of this model shows MSE is 0.0039 and the correction coefficient is 0.67.

We evaluate the model's performance on the testing set using mean squared error as the evaluation metric. Figure 4 shows the performance of the CNN-LSTM model for the training and validation dataset.

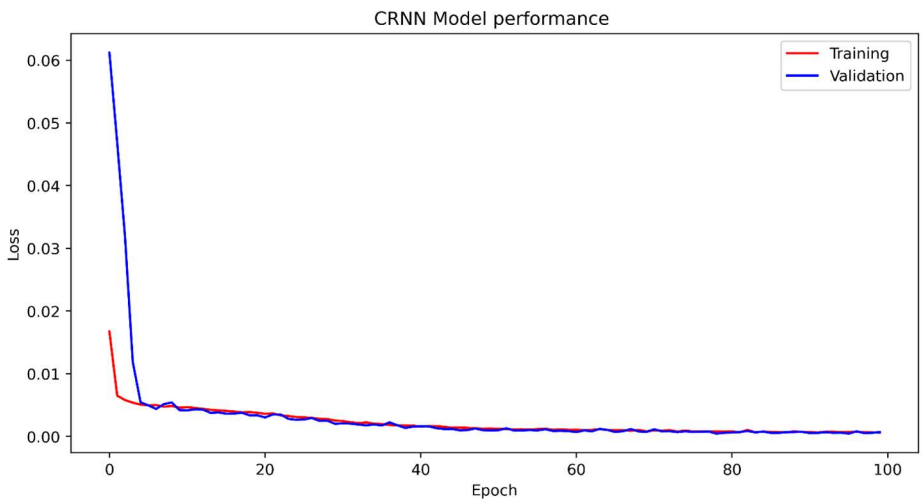

Figure 4- CNN-LSTM model performance

As can be seen, A good fit is identified by a training and validation loss that decreases to a point of stability with a minimal gap between the two final loss values. The coefficient of determination (R-squared) to measure how well the model can predict the output values for CRNN is shown in Figure 5.

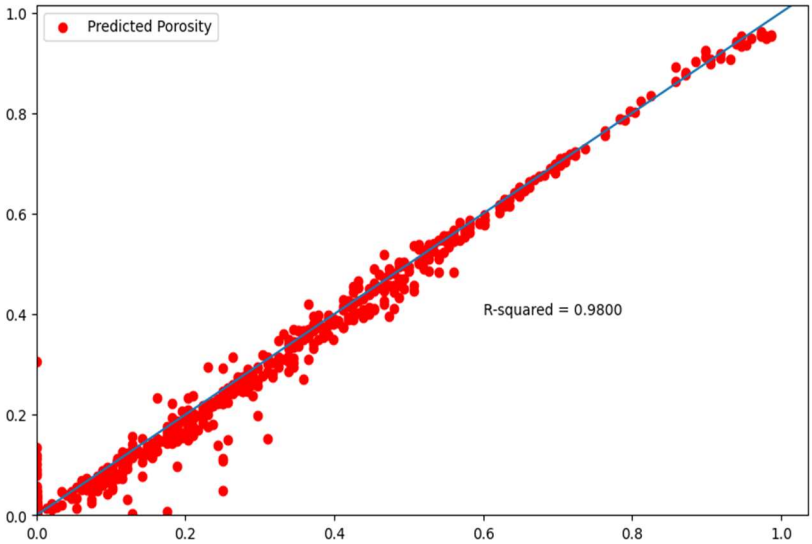

Figure 5- CNN-LSTM model prediction evaluation

The correlation coefficient between the CRNN model and the core data obtained was 0.98, which shows excellent performance. Table 1 presents the comparison of the performance of both tuned models.

Table 1- Performance comparison of MLP and CRNN model

Calculated parameter	MSE	R-Square
MPL model	0.00039	0.67
CRNN model	0.01342	0.98

The results showed that the CRNN model had a MAS of 0.01342, while the MLP model had an average MSE of 0.00039. This indicates that the CRNN model was able to predict porosity with higher accuracy than the MLP model.

CONCLUSION(S)

A hybrid CNN-LSTM model is proposed in this work to predict porosity on real data from a carbonate oil field in Iran. The model is a combination of CNN and LSTM, where well-logs are used as input, and core porosity is estimated as output. The CNN is used to extract features of high dimensional data, while LSTM can remember past inputs and model sequential data. The performance of the CNN-LSTM model was compared with the MLP model in terms of mean square error (MSE) and determination coefficient (R2) values. The hybrid model proved to be highly effective in predicting porosity accurately in multidimensional spaces, achieving approximately 98% porosity estimation in this study. The model could also handle complex predictions of reservoir porosity by extracting relevant features and using gate control, such as a forgetting gate. The correlation coefficient between the model prediction and core data improved from 0.67 to 0.98, indicating the accuracy of the model. Compared to traditional MLP algorithms, the CNN-LSTM model performed better, achieving higher accuracy rates with fewer model parameters. The ability of the CNN-LSTM model to effectively capture and analyze spatial and temporal patterns in well-log data makes it an ideal choice for porosity prediction tasks.

REFERENCE

1. Bagrintseva, K.I., *Carbonate Reservoir Rocks*. 2015: Scrivener Publishing LLC.
2. Bust, V.K., J.U. Oletu, and P.F. Worthington, *The Challenges for Carbonate Petrophysics in Petroleum Resource Estimation.* SPE Reservoir Evaluation & Engineering, 2011. **14**(01): p. 25-34.
3. Wang, J. and J. Cao, *Deep Learning Reservoir Porosity Prediction Using Integrated Neural Network.* Arabian Journal for Science and Engineering, 2022. **47**(9): p. 11313-11327.
4. Tiab, D. and E.C. Donaldson, *Chapter 3 - Porosity and Permeability*, in *Petrophysics (Fourth Edition)*, D. Tiab and E.C. Donaldson, Editors. 2016, Gulf Professional Publishing: Boston. p. 67-186.
5. Moosavi, N., et al., *Fuzzy support vector regression for permeability estimation of petroleum reservoir using well logs.* Acta Geophysica, 2022. **70**(1): p. 161-172.
6. Moosavi, N., et al., *Porosity prediction using Fuzzy SVR and FCM SVR from well logs of an oil field in south of Iran.* Acta Geophysica, 2022.
7. Xu, C., et al., *Machine learning in petrophysics: Advantages and limitations.* Artificial Intelligence in Geosciences, 2022. **3**: p. 157-161.