

## Machine Learning Application of Two-Dimensional Fracture Properties Estimation

Ardian Nurcahya, Aldenia Alexandra, Satria Zidane Zainuddin, Fatimah Az-Zahra, M. I. Khoirul Haq, Irwan Ary Dharmawan\*

Department of Geophysics, Faculty of Mathematics and Natural Science, Universitas Padjadjaran, Jatinangor, Indonesia

\* Corresponding author : [iad@geophys.unpad.ac.id](mailto:iad@geophys.unpad.ac.id)

Received: May 20, 2023. Revised : May 31, 2023, Accepted: June 10, 2023, Published: July 31, 2023

DOI: 10.25299/jgeet.2023.8.02-2.13874

### Abstract

Fractures are substantial contributors to solute transport sedimentary systems that form pathways. The pathway formed in a fracture has two physical parameters, there are mean aperture and surface roughness. Mean aperture is the thickness of the pathway that the fluid will pass through, and surface roughness is the roughness of the fracture pathway. The two physical parameters of the fracture are important to determine since they affect the permeability value in petroleum reservoir analysis. We developed a machine learning algorithm based on the Convolutional Neural Network (CNN) to predict those two parameters. Furthermore, image processing analysis is performed to generate the datasets. The results show that the CNN algorithm shows good agreement with the reference results. In addition, the algorithms showed efficient performance in terms of computational time. CNN is a type of deep neural designed to perform analysis on multi-channel images that can classify fracture geometry. The best model was determined using a benchmark dataset with a CNN model provided by Keras. The results of experiments conducted on fracture geometry images show that the machine learning model created is able to predict the mean aperture and surface roughness values.

**Keywords:** Fracture, Mean aperture, Surfaces roughness, Machine learning, CNN

### 1. Introduction

Energy is an essential part of daily life, which one of the main energies used is fossil energy. Fossil energy is the main energy source and source of foreign exchange for Indonesia, but fossil energy sources have negative impacts on the environment, such as air pollution, greenhouse gas emissions, and global warming. In addition, high demand for fossil energy is another issue. Increasing demand for fossil energy such as oil is accompanied by rising prices, which leads to diminishing reserves of fossil energy. Oil production has decreased over 10 years from 346 million barrels or 949 thousand barrels per day (BPD) in 2009 to around 283 million barrels or 778 thousand bpd in 2018. This is because the majority of wells are older, while the production of new wells is relatively limited (ESDM, 2019).

Fractures are important objects or structures, especially in oil and gas exploration, because fractures are one of the secondary petroleum reservoirs (Koesoemadinata, 1980). This is also supported by Herdiansyah who stated that volcanoclastic reservoirs in Indonesia are reservoirs with significant production where one of the most important factors is natural fractures that determine the quality and quantity of the reservoir (Herdiansyah et al., 2020). At first glance, the fracture shape only looks like a line, but in reality the fracture geometry has many variations because there are parameters that can affect the fracture geometry such as surface roughness and mean aperture.

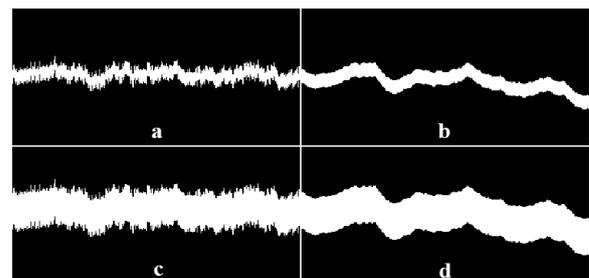


Fig. 1. Physical parameters mean aperture and surface roughness in fracture (a) Mean aperture 15 *lu* and surface roughness 0.1, (b) Mean aperture 15 *lu* and surface roughness 0.9, (c) Mean aperture 40 *lu* and surface roughness 0.1, and (d) Mean aperture 40 *lu* and surface roughness 0.9

Mean aperture is the relative mean heights between the two surfaces used to define the fracture aperture. The unit used in the mean aperture used is the lattice unit (*lu*). Surface roughness is commonly used to indicate the roughness level of the surface in the fracture, which has a range of values from 0 to 1. Fig. 1 shows that when the surface roughness value is close to 0, the surface will be more rough, while when the surface roughness value is close to 1, the surface will be smoother. The surface roughness value that is often found in the field is between 0.45 - 0.85 (Wang et al., 2021). The complex and irregular shape of the fracture surface geometry is another major factor that impacts the permeability value in petroleum reservoir analysis, therefore it is necessary to further investigate the fluid flow within the fracture medium. As it is known that the experiments at the laboratory especially

in oil and gas exploration have high operational costs and require a significant amount of time. In addition, the physical parameter values of the fracture are obtained through numerical simulation of fluid dynamics, but that method also requires a considerable amount of computation time (Dharmawan et al., 2016). Therefore, in this research, machine learning algorithms is used as one of the alternative to solve cases or problems in the oil and gas field.

Machine learning has been widely developed as a solution in science and technology fields such as production optimization and hydrocarbon drilling. It is also being applied to simplify and speed up the computational process in estimating fracture physical parameters. Artificial neural network (ANN) are a popular type of machine learning model used to solve complex problems. They are based on the structure of human nerves that adaptively solve tasks. A typical ANN consists of multiple layers with multiple perceptrons in each layer. The basic building block of an ANN is the perceptron, which is modeled after neurons in the human brain. In an ANN, the input for one layer serves as the output for the next layer. One ANN algorithm that is often applied to solve image recognition problems is the CNN. CNN uses a convolution method that applies filters of a certain size to various locations of the input data. From the convolution of the input data and filters, the machine obtains new representative information. The output of the convolution is then used as the input for the neural network layer below. Because the feature extraction and training process in the CNN algorithm is done by the computer simultaneously, CNN is a good solution for estimating the physical parameters of fractures with complex patterns. It is also beneficial because it does not require testing, such as fluid dynamics modeling, and can calculate physical parameter values in less time.

A CNN is a deep learning algorithm that can take in an input image, assign importance (learnable weights and biases) to various aspects or objects in the image, and differentiate one from the other. It is a promising tool for solving pattern recognition problems (Gao and Mosalam, 2018). CNNs are a specialized type of ANN that use a mathematical operation called convolution in place of general matrix multiplication in at least one of their layers. They are designed to process pixel data and are used in image recognition and processing. A CNN consists of an input layer, hidden layers, and an output layer, with the hidden layers including layers that perform convolutions. A CNN can have tens or hundreds of layers, each learning to detect different features of an image. Filters are applied to each training image at different resolutions, and the output of each convolved image is used as the input to the next layer. The general layout of the layers of the CNN architecture is shown in Fig. 2.

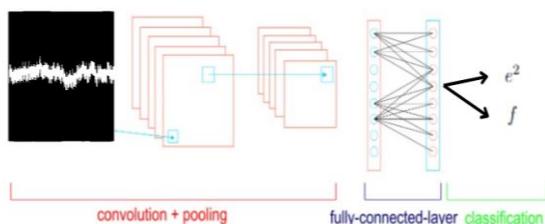


Fig. 2. CNN scheme to predict an image

Convolution puts the input images through a set of convolutional filters, each of which activates certain features from the images. The rectified linear unit (ReLU)

allows for faster and more effective training by mapping negative values to zero and maintaining positive values. Pooling simplifies the output by performing nonlinear downsampling, reducing the number of parameters that the network needs to learn. After learning features in many layers, the architecture of a CNN shifts to classification. The final layer of the CNN architecture uses a classification layer to provide the final classification output. CNNs provide an optimal architecture for uncovering and learning key features in image and time series data, and are key technology in applications such as object detection, audio processing, and synthetic data generation (Talo, 2019). In this research, the problem to be solved is image recognition. The pre-trained model architectures used in this research are those available in the Keras library which have been tested and have good performance. The three types of pre-trained models used in this research are DenseNet201, DenseNet169, and Xception.

## 2. Material and Methods

The dataset was generated using the SmartFract application, which is based on Matlab. This software uses a fractional Brownian motion algorithm, which is a random movement of a value with a Gaussian process in continuous time starting from zero and centered at a mean of zero based on a covariance function. The generated data was used for training with a total of 45,000 data points, including 36,000 for training and 9,000 for validation. The variations in this data consist of two classes, mean aperture and surface roughness, as shown in Table 1.

Table 1. Variation of data used

No	Mean Aperture (lu)	Surfaces Roughness	Number of data
1	5	0.1 to 0.9	4,500
2	10	0.1 to 0.9	4,500
3	15	0.1 to 0.9	4,500
4	20	0.1 to 0.9	4,500
5	25	0.1 to 0.9	4,500
6	30	0.1 to 0.9	4,500
7	35	0.1 to 0.9	4,500
8	40	0.1 to 0.9	4,500
9	45	0.1 to 0.9	4,500
10	50	0.1 to 0.9	4,500
TOTAL			45,000

The method in this study employs a CNN to identify object parameters in fracture geometry. This research utilizes Transfer Learning, a machine learning approach that leverages previously acquired knowledge to solve related problems in different classes. Transfer learning can significantly speed up the training process by using a pre-trained model, eliminating the need for trial and error. It also has the potential to produce more accurate predictions with higher success rates and faster training times using fewer training data points, as it builds upon prior knowledge. A schematic comparison of transfer learning and traditional machine learning is shown in Fig. 3.

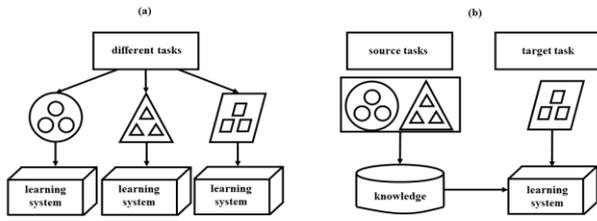


Fig. 3. The comparison between traditional machine learning and transfer learning (a) traditional machine learning, (b) transfer learning.

The pre-trained DenseNet201, DenseNet169, and Xception architectures used in this study are provided by Keras. These models use feedforward connections between every layer to mitigate vanishing-gradient problems, enhance feature propagation, encourage feature reuse, and significantly reduce the number of parameters. Activation functions in the model should also be carefully considered as they can impact the parameter values of the resulting model. In this study, the rectified linear activation function (ReLU) was used to predict fracture parameter values. The DenseNet201 architecture takes advantage of a compact network, allowing for easy training and highly efficient models due to the ability of different layers to reuse features, which increases the diversity of inputs to subsequent layers and improves performance (Huang et al., 2017). The DenseNet201 architecture has been widely used in image recognition tasks, such as determining the type of weather based on available weather datasets, and has shown good performance in estimating object parameter values. This architecture can be used for object classification or value estimation (Naufal and Kusuma, 2022).

Mean relative error (MRE) is used to determine the value of data deviation relative to the actual data. The relative error value is obtained by subtracting the predicted data  $y_i$  from the actual data  $y$  and then dividing by the actual data  $y$ . Relative error is generally used in determining the error in percentage form so that the error can be more easily read. The equation of the relative error can be written as follows.

$$MRE = \sum_{i=1}^n \frac{y_i - y}{y} \quad (1)$$

Mean squared error (MSE) is a value used to estimate a certain quantity in data. RSE is commonly used to estimate unobserved quantities in a training model. The MSE value can be used to see how far the estimated value is from the predicted value with the estimated value  $\hat{y}_i$  and the predicted value  $y_i$ . MSE can be written in the following equation.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

Mean absolute error (MAE) is a function used for regression models. MAE is the sum of absolute differences between the target and independent variables. It measures the average of the residuals, where  $n$  represents the number of observations,  $y_i$  is the predicted price at the point of sale  $i$  and  $y_a$  is the actual value.

$$MAE = \sum_{i=1}^n |y_i - y_a| \quad (3)$$

Root mean square error (RMSE) is another commonly used metric to evaluate the accuracy of predictions obtained by a model. It takes the residuals between actual and predicted values and compares the prediction errors of

different models for particular data. The primary advantage of using RMSE is that it penalizes large errors and scales the results in the same units as the forecast values.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (A_i - F_i)^2} \quad (4)$$

$R^2$  is a widely used statistical measure in regression-based machine learning. It indicates the percentage of the variance in the dependent variable that the independent variables explain collectively. The closer the value of  $R^2$  to 1, the better the model is fitted.

$$R^2 = 1 - \frac{\sum_{i=1}^n (A_i - F_i)^2}{\sum_{i=1}^n A_i^2} \quad (5)$$

### 3. Results and Discussion

As a result of this research, two models using a CNN were developed to predict mean aperture and surface roughness values. These models were able to predict the values using linear regression. The performance of each model can be seen in Fig. 4, which is presented as a histogram.

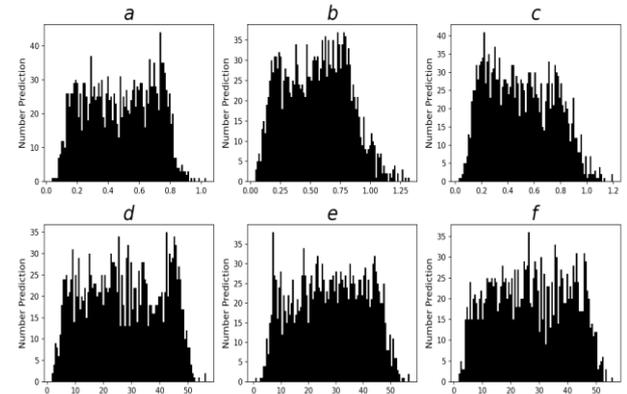


Fig. 4. Histogram prediction surface roughness and mean aperture, (a) base on surfaces roughness with DenseNet201 architecture, (b) base on surfaces roughness with Dense169 architecture, (c) base on surfaces roughness with Xception architecture, (d) base on mean aperture with DenseNet201 architecture, (e) base on mean aperture with Dense169 architecture, (f) base on mean aperture with Xception architecture

The surface roughness can be predicted using a model. Based on Fig. 4, the model performs fairly well in estimating the roughness values in the range of 0.1 to 0.7, as indicated by the high number of accurate predictions in that range. However, the model is weaker in predicting roughness values greater than 0.7, as shown by the decrease in data readings in that range in the histogram in Fig. 4. The performance of the mean aperture model can also be seen in Fig. 4, where the shape of the model's graph shows that the predictions are not too far off. In addition, Fig. 4 shows that the highest prediction results for mean aperture data occur in the range of 1 to 10, and then remain fairly stable from 15 to 45. However, the model is weaker in predicting low mean aperture values, as shown in Fig. 4. The distribution of the data for each model can also affect the prediction results and errors in the test data process. The actual and predicted data can be plotted to show the distribution of the experimental results, as seen in Fig. 5.

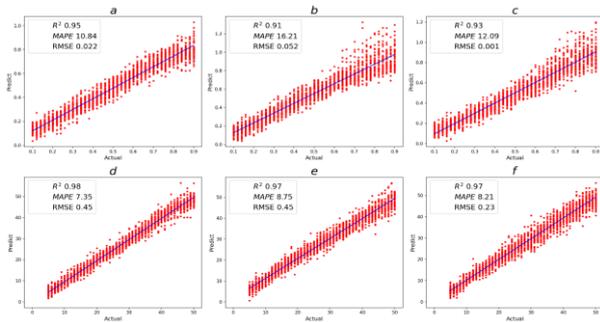


Fig. 5. Linear regression predict and actual value base on surfaces roughness and base on mean aperture, (a) base on mean aperture with DenseNet201 architecture, (b) base on mean aperture with Dense169 architecture, (c) base on mean aperture with Xception architecture, (d) base on surfaces roughness with DenseNet201 architecture, (e) base on surfaces roughness with Dense169 architecture, (f) base on surfaces roughness with Xception architecture.

The performance of the model can be evaluated by plotting the distribution of data points and creating a linear regression line. According to Fig. 5, some data points have a large deviation from the regression line, indicating poor accuracy of the trained model in certain ranges. For instance, Fig. 5 shows poor performance in predicting surface roughness in the range of 0.8 to 0.9, as indicated by the missing data in the range of 0.95 and the large distance between the data and the linear regression line, resulting in a high standard deviation in that range. A high standard deviation can significantly affect the error generated by the model, as a larger distance between data points leads to a larger error. On the other hand, the mean aperture model training results shown in Fig. 5 demonstrate better performance, with data points not deviating significantly from the regression line, resulting in better image prediction. However, it should be noted that the mean aperture training model clearly shows poor performance in the range of 1 to 10, as shown by the clustering of data points at one point in this range. As a result, the test performance results for the mean aperture range of 1 to 10 are not recommended due to their poor performance. In contrast, the mean aperture training model shows quite good performance in the range of 10 to 50, as demonstrated by the data distribution shown in Fig. 6 and the boxplot in Fig. 6.

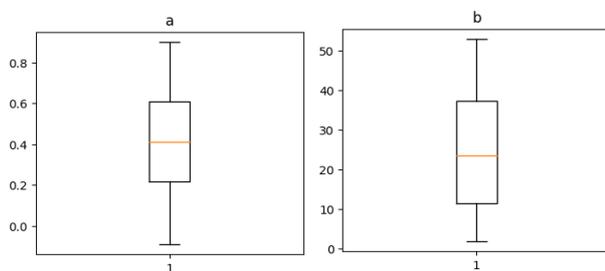


Fig. 6. Boxplot of dataset DenseNet201 (a) base on surfaces roughness and (b) base on mean aperture

Boxplots are used to show quartiles or boundaries at the top and bottom. The boundary at the top is the third quartile (Q3), which means that 75% of the predicted data is below the other quartiles. In Fig. 6, it can be seen that Q3 for the surface roughness model has a value of 0.6, and Q3 for the mean aperture training model has a value of 38. The bottom boundary is the first quartile (Q1), which represents the minimum value with 75% of the data below it. In Fig. 6, it

can be seen that the Q1 value for the predicted data has a value of 0.2, while the Q1 value for the mean aperture training model is in the range of 10. The median value or second quartile (Q2) of the data can be seen in Fig. 6, which is at a value of 0.4 for the surface roughness model and 23 for the mean aperture model. The boxplot created does not show any outliers, meaning that the predicted data is within the range of the maximum and minimum observation data. In addition, the performance of the model can be evaluated by directly determining the error, as shown in Table 2 below.

Table 2. Error value DenseNet201

No	Type Error	Surfaces Roughness Error	Mean Aperture Error
1	MRE	0.097	1.86
2	MAE	0.16	0.087
3	RMSE	0.12	2.53
4	MSE	0.015	6.42
5	R <sup>2</sup>	0.95	0.97

Table 2 shows the errors between the predicted and actual data from the model created with the DenseNet201 architecture. The calculation of the MRE shows a value of 0.18 for the surface roughness model and 0.14 for the mean aperture model. MRE has a relationship with MSE, so based on the obtained errors, the mean aperture model performs better than the surface roughness model. However, in the MAE and RMSE values, the error value for the mean aperture model is higher compared to the error generated by the surface roughness model. This is a common occurrence due to the different range of values in the classes. The range of values in the surface roughness class is from 0.1 to 0.9, while the range of values in the mean aperture class is from 5 to 50, resulting in a larger standard deviation in the class with a higher range of values. In the coefficient of determination (R<sup>2</sup>) results, it can be seen that both models have good performance, with a value of 0.93 for the surface roughness model and 0.97 for the mean aperture model.

#### 4. Conclusion

Based on the results and discussion, the machine learning model can accurately estimate the surface roughness and mean aperture values. The model, built using a CNN, performs reasonably well, although it does show a drop in performance over a certain range. Two models were created, one for surface roughness and the other for mean aperture. The best model in the results of this research is the model created with the DenseNet201 architecture. The results of the model show good performance at Q1 and Q2 values, but decreased performance at Q3 and above, as shown in Fig. 6. This is in contrast to the average aperture training model, which has the highest performance at Q2 and Q3, while its performance at Q1 is low, as shown in Fig. 6(b). In addition, the model shows good performance in the calculation of the coefficient of determination (R<sup>2</sup>) with values of 0.93 for surface roughness and 0.97 for average aperture. Therefore, the training model can be used to estimate the surface roughness and mean aperture values in fracture geometry images. Further research is needed to improve the performance of the model, particularly in predicting real data from a single fracture, so that the model can be applied in industry.

## Acknowledgements

The authors acknowledge the Department of Geophysics Universitas Padjadjaran supercomputing resources "RockExplorer" made available for conducting the research reported in this paper.

## References

- Dharmawan, I. A., Ulhaq, R. Z., Endyana, C., AUFARISTAMA, M., 2016. Numerical simulation of non-Newtonian fluid flows through fracture network. In IOP conference series: earth and environmental science 29(1), 012030. doi: 10.1088/1755-1315/29/1/012030.
- ESDM., 2019. Outlook Energi Indonesia (OEI) 2019. Ministry of Energy and Mineral Resources Republic of Indonesia. Retrieved from <https://www.esdm.go.id/assets/media/content/content-outlook-energi-indonesia-2019-bahasa-indonesia.pdf>
- Gao, Y., Mosalam, K. M., 2018. Deep transfer learning for image based structural damage recognition. Computer-Aided Civil and Infrastructure Engineering 33(9), 748–768.
- Herdiansyah, H., Negoro, H. A., Rusdayanti, N., Shara, S., 2020. Palm oil plantation and cultivation: Prosperity and productivity of smallholders. Open Agriculture 5(1), 617-630.
- Huang, G., Liu, Z., Der, V., Weinberger, K. Q., 2017. Densely connected convolutional networks. 2261–2269. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Doi: 10.1109/CVPR.2017.243
- Koesoemadinata, R.P., 1980. Geologi Minyak dan Gas Bumi (1<sup>st</sup> Ed). Penerbit ITB, Bandung.
- Naufal, M.F., Kusuma, S.F., 2022. Weather image classification using convolutional neural network with transfer learning. AIP Conference Proceedings 2470(1), 050004.
- Talo, M., 2019. Convolutional neural networks for multi-class histopathology image classification. arXiv:1903.10035.
- Wang, D., de Boer, G., Neville, A., Ghanbarzadeh, A., 2021. A new numerical model for investigating the effect of surface roughness on the stick and slip of contacting surfaces with identical materials. Tribology International 159, 106947. doi: 10.1016/j.triboint.2021.106947



© 2023 Journal of Geoscience, Engineering, Environment and Technology. All rights reserved. This is an open access article distributed under the terms of the CC BY-SA License (<http://creativecommons.org/licenses/by-sa/4.0/>).