INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN SOCIAL SCIENCE (IJRISS)
ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue II February 2026

Estimation of Lost Circulation Zones Using Random Forest Classification and Image Logs

Dr. John Lander Ichenwo, Marvellous Amos

Department of Petroleum Engineering, University of Port Harcourt

DOI: https://doi.org/10.47772/IJRISS.2026.10200070

Received: 30 January 2026; Accepted: 05 February 2026; Published: 24 February 2026

ABSTRACT

Lost circulation is one of the most chronic and most expensive problems in the drilling process as it leads to enormous losses in terms of finances and time. The conventional prediction methods mainly are based on the traditional well logs and correlations which have been empirically determined and which fully exploit the rich textual information available in the Formation Microresistivity Imaging (FMI) logs. The work is the innovative supervised machine learning model based on the use of the Random Forest classification algorithms to estimate the lost circulation areas based on the combination of the traditional Logging While Drilling (LWD) data with the texture features of the FMI image logs. The methodology was used on a large dataset of six wells of the Niger Delta, and this includes more than lost circulation events recorded by drilling reports. The conventional logs (gamma ray, resistivity, bulk density, caliper) were used as the input parameters complemented with the texture descriptors based on the FMI images determined by the Gray Level Co-occurrence Matrix (GLCM) methods. GLCM texture characteristics such as contrast, correlation, homogeneity and energy were calculated with the help of Open CV and at 0, 45, 90, and 135 angular orientations. Preprocessing of data included removing outliers, lithological coding, and depth matching of traditional logs with image-based characteristics. Random Forest classifier was trained on 70% of the data and was validated on the 30% of the data with 100 estimators and the maximum depth of 10 implemented in scikit-learn with Python. Accuracy, precision, recall, F1-score and ROC-AUC were used to measure model performance. Findings indicate that the combined method had an overall prediction accuracy of 86% and the image-based texture characterizations increased accuracy by 14% points as opposed to the traditional log-only models (71% to 86%). The ROC-AUC rose by 0.79 in case of conventional features and 0.91 of the combined feature set. The analysis of the feature importance showed that the most important features were texture contrast and caliper measurements. The methodology will allow the identification of lost circulation areas at an initial stage during the real-time implementation of drilling activities and allows taking proactive measures to mitigate the latter and improve the productivity of operations.

Keywords: Lost Circulation, FMI Image Logs, GLCM, Random Forest, Machine Learning, Texture Analysis, LWD, Niger Delta

INTRODUCTION

Lost circulation is a basic issue in the drilling operations in a wide range of geological settings, whereby the loss of drilling fluid to the formation fractures, vugs, or permeable structures is not controlled. The effect of this phenomenon is the reduction of the pressure in the wellbore, the influx of formation fluids, and the high operational problems, such as stuck pipe, instability of the wellbore, and long drilling time. It is estimated within the industry that lost circulation incurs 10-20 percent of overall drilling expenditure worldwide, with losses of over $2billion in any 1 year attributed to loss of non-productive time of operation and clean up operations [1][2]. The geological structures of the Niger Delta basin (complex geology, sandstone-shale layers, fractured carbonates and unconsolidated ones) provide especially problematic conditions in the management of the drilling fluids [3][4]. The Agbada and Akata formations are interbedded and have different pore pressure regimes together with structural complexity, which make the circumstances favorable to the occurrence of lost circulation events that are challenging to forecast based on the standard methods [5]. Problem Statement: The existing lost circulation prediction procedures have the limitation of relying primarily on the conventional well log variables (gamma ray, resistivity, density, neutron porosity) and experience based correlations with the drilling parameters

[6][7]. Though these methods have demonstrated a rational success in uniform formations, they in a systematic way fail to represent the multifaceted textural and structural disparities that in many cases dictate fluid loss processes. The microresistivity Imaging (FMI) logs are high-resolution images of borehole walls with a vertical resolution of about 5mm that are rich textual information which is linked to the formation permeability, fracture networks, and variations in porosity [8][9]. Nonetheless, this is a useful source of data that is still underexploited in quantitative lost circle prediction processes. The traditional methods have a number of fundamental shortcomings: (1) they use post-drilling empirical correlations which can not be adjusted to changing conditions in the real time, (2) they fail to reflect the formation heterogeneity using conventional log parameters alone, (3) they do not attempt to quantify the spatial dependence between rock texture and fluid loss processes and (4) they do not integrate image-based data into predictive models although it has been shown to correlate with formation drillability and permeability [10][11].

Research Gap: Despite significant advances in machine learning applications to drilling engineering [12][13], there exists a notable gap in methodologies that systematically integrate Formation Microresistivity Image logs with conventional well log data for lost circulation prediction. While several studies have applied machine learning techniques to conventional drilling parameters [14][6], none have comprehensively exploited the textural information contained within FMI logs through advanced image processing techniques such as Gray Level Co-occurrence Matrix (GLCM) analysis.

Research Objective: In spite of the tremendous progresses in application of machine learning to the engineering of drilling [12][13], there is a considerable research gap in the systematic methods aimed at bringing together Formation Microresistivity Image logs and the traditional well log data to predict lost circulation. Although some have used machine learning on traditional drilling parameters [14][6], none have effectively used textural data that is present in FMI logs using advanced image processing methods like the Gray Level Co-occurrence Matrix (GLCM) analysis. Research Objective: The objective of this research is to derive and test a classification by random forest technique to determine the lost circulation areas with systematic analysis of traditional LWD data and texture features obtained using Formation Microresistivity Image logs. In particular, the following research objectives will be considered: (1) creation of a complete workflow on texture analysis with the assistance of GLCM methods to retrieve quantitative textures on FMI images, (2) combining image-based texture characteristics with standard well log parameters in a single machine learning model, (3) systematic comparison between the model performance on conventional feature and combined conventional-texture feature sets, (4) validation of the methodology with multiple wells and geological formations in the Niger Delta, and (5) assessment of the feature significance to comprehend individual parameter contribution to the prediction accuracy of lost circulation.

MATERIALS AND METHODS

Field Location and Geological Context

Our position here is the southwestern end of the northern section of the Deepwater Horizon in Louisiana, which is close to the city of New Orleans. The study has used an extensive data set on six wells (named NDW-A to NDW-F) that were drilled within the offshore Niger Delta basin in 2020-2024. The wells were chosen considering that they were fully accessible in terms of both conventional LWD logs and high-quality FMI image logs, as well as extensive records of the lost circulation that can be achieved through the daily drilling reports. The geographical area covered by the research is water with a depth of between 800 and 1,500 m; whereas the total depth of these wells is 3,200 to 4,800 m. The Niger Delta is a fully developed deltaic system which is dominated by three lithostratigraphic units; marine shale Akata Formation, paralic Agbada Formation and continental Benin Formation. The study intervals of interest mostly include the Agbada Formation, which is marked with rapid change of sandstone and shale formations of differing consolidation, natural fracturing, and fluid saturation [5]. This nonhomogeneous setting provides the best environment to study the relationship between lost circulation susceptibility and formation texture.

Data Description and Acquisition

The central database now holds 15,420 depth-indexed values for every 0.5 meters of the six research wells. Every piece of data consists of synchronized standard log measurements and matching patches of FMI images, binary classification labels are based on the experience of drilling fluid loss rates reported in daily drilling reports.

Lost Circulation Event Classification: The lost circulation incidence was assessed based on rates of fluid lost as indicated in the drilling account, according to industry standard categories as follows: (0) Normal circulation (0-5 bbl/h), (1) Lost circulation zone (>5 bbl/h sustained loss). The binary classification design has been chosen to address the main difference between normal and problematic circulation conditions and, therefore, be able to train the model and be able to validate it [6].

Input Variables:

Parameters of conventional Wells: •

- Gamma Ray (GR): Natural radioactivity (API units) which measures the content of clay and the concentration of organic matter.
- Deep Resistivity (RD): Formation resistivity (ohm-m) that represents fluid saturation and formation porosity.
- Bulk Density (RHOB): Formation density (g/cm 3) meaning the differences in lithology and porosity.
- Caliper (CAL): Borehole diameters of washouts and tight holes measured in inches.
- Formation Tops: Main lithologies are categorized by type – sandstone, shale, carbonate, mixed.

Formation Microresistivity Image (FMI) Data: High-resolution microresistivity images obtained on Schlumberger FMI tools with vertical and azimuthal resolution of 0.2 inches giving 80% borehole coverage in the study well diameters [9]. The images of FMIs were taken in parallel with the conventional logs in LWD operations which guaranteed the ability to register the depth perfectly and the quality of data could be maintained in the same manner.

Target Variable: Binary classification with (0 normal circulation) and (1 lost circulation zone) based on sustained fluid loss rates of over 5 bbl/h over periods longer than 3 meters recorded in the drilling report and checked against caliper log anomalies.

Data Preparation and Image Processing

FMI Image Preprocessing: The raw FMI microresistivity data underwent the standard Schlumberger workflow of speed correction, normalization and image enhancement. The post-processed images were exported in 8-bit grayscale matrices where the pixel values were of the variation in microresistivity of the formation. The extracted image patches of 64x64 pixels (relative to the depth interval of 0.5 meters) were then subjected to texture analysis, and the spatial resolution of all the wells was kept constant.

Gray Level Co-occurrence Matrix (GLCM) Implementation: TheGLCM technique had been used to extract the texture features, which was implemented using the OpenCV Python library [15][16]. GLCM measures the spatial correlation of pixel intensities at a given distance and angular orientation, which gives statistical values of image texture which predict geological heterogeneity [17][18].

GLCM Parameter Configuration:

- **Distance:** d = 1 pixel (corresponding to ~0.05 inches)
- **Angular Orientations:** θ = 0°, 45°, 90°, 135°
- **Gray Level Quantization:** 256 levels reduced to 64 levels for computational efficiency

- **Texture Descriptors:** Four primary Haralick features computed for each orientation:

Contrast: Local intensity variation used to detect the formation heterogeneity.

$$Contrast = \sum_{i,j} |i - j|^2 \cdot P(i,j)$$

Correlation: Quantifies linear dependency between pixel pairs, reflecting bedding continuity

$$Correlation = \frac{\sum_{i,j} (i - \mu_i)(j - \mu_j)P(i,j)}{\sigma_i \sigma_j}$$

Homogeneity: Measures texture uniformity, indicating formation consistency

$$Homogeneity = \sum_{i,j} \frac{P(i,j)}{1 + |i - j|}$$

Energy (Angular Second Moment): Quantifies texture orderliness, reflecting structural organization

$$Energy = \sum_{i,j} P(i,j)^2$$

where P(i,j) represents the normalized GLCM, and μ, σ denote mean and standard deviation respectively.

Feature Vectors Construction: 16 texture features were produced (4 descriptors for 4 orientations) for each depth range, which were averaged into 4 texture descriptors. The method ensures both directional and non-directional textures and is computationally efficient [17].

Conventional Log Processing: Standard preprocessing procedures were applied to conventional log data including:

- **Outlier Detection:** Statistical outliers identified using the Z-score method (|Z| > 3) and removed or flagged for quality control

- **Missing Data Handling:** Linear interpolation applied for gaps <1 meter; intervals with >10% missing data excluded

- **Depth Synchronization:** All log measurements aligned to common depth reference with 0.5-meter sampling interval

- **Normalization:** Min-max scaling of the continuous variables to make sure they contribute equally during the training of models:

$$X_{normalized} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Lithology Encoding: Formation lithology was encoded using one-hot encoding to form binary variables of the major types of rocks:

- Sandstone_flag (0/1)

- Shale_flag (0/1)

- Carbonate_flag (0/1)

- Mixed_lithology_flag (0/1)

INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN SOCIAL SCIENCE (IJRISS)
ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue II February 2026

MACHINE LEARNING METHODOLOGY

Data Partitioning: Data was divided in a 70:30 stratified fashion for training and testing; both portions had an equal number of normal circulation and lost circulation area representation. Data leakage could be avoided, and the stratification at the well level was made in order to achieve model generalizability between wells and different geological settings. Random Forest Implementation: The implementation of the Random Forest classifier was carried out with scikit-learn python-based library (version 1.1.3) with the following optimized hyperparameters [19][20]

Random Forest Implementation: Random Forest Implementation: The implementation of the Random Forest classifier was carried out with scikit-learn python based library (version 1.1.3) with the following optimized hyperparameters [19][20]:

- **n_estimators:** 100 decision trees
- **max_depth:** 10 levels
- **max_features:** √n (square root of total features)
- **min_samples_split:** 2
- **min_samples_leaf:** 1
- **bootstrap:** True (with replacement sampling)
- **random_state:** 42 (for reproducibility)

Feature Set Evaluation: Model performance was systematically evaluated using two distinct feature configurations:

- **Set A (Conventional Only):** Gamma ray, resistivity, bulk density, caliper, lithology flags (9 features)
- **Set B (Combination):** Standard features and GLCM texture measures (13 features)

Model Training Process: The Random Forest algorithm relies upon ‘bagging’ a number of decision trees, and a random selection of ‘features’ for each split. This group method minimizes the overfitting and still has high predictive performance. Training was performed on a random subset of the training data by each of the trees and final predictions done by majority voting among all the trees [21].

Implementation Environment: Model development was conducted using:

- **Python 3.9** with NumPy, Pandas, and Matplotlib libraries
- **OpenCV 4.5** for image processing and GLCM computation
- **scikit-learn 1.1.3** for machine learning algorithms and evaluation metrics
- **Jupyter Notebook** environment for iterative development and visualization

Model Evaluation and Validation

Performance Metrics: Model performance was thoroughly evaluated with several classification metrics that were appropriate in binary classification problems [22][23]:

Accuracy: The capacity for making the right decisions in most instances.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision: Percentage of predicted lost areas of circulation that were selected accurately.

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

Recall (Sensitivity): Per cent of real areas of lost circulation correctly recognized.

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

F1-Score: Precision and Recall standardized.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

ROC-AUC: Region Under the Receiver Operating Characteristic curve, the evaluation of the discrimination ability at all classification levels [24][25].

Cross-Validation Strategy: The cross-validation method was used on the training set in the form of five-fold cross-validation to assess the model's robustness and prevent overfitting. Also, the ‘leave one well out’ validation was used to assess the model’s generalizability for different geological conditions.

Feature Importance Analysis: Random Forest provided its own feature importance measurements, which were used to measure the relative value of each input parameter in predicting the accuracy. The importance of feature permutation was also calculated as well to justify the natural rankings of the features [12].

RESULTS

Dataset Characteristics and Lost Circulation Distribution

Data Summary

After quality control and preprocessing, the resulting final dataset consisted of 14,850 valid observations in the six Niger Delta wells. The most frequent event was lost circulation (Class 1) (23.4% of the dataset), whereas normal circulation conditions (Class 0) (76.6% of the dataset) were the most frequent. This moderate imbalance in the classes is realistic in the field conditions that exhibit lost circulation events which, though critical in operation, are not as common as the regular conditions in drilling.

Lost Circulation Patterns: Results of the lost circulation distribution analysis indicated that it was significantly related to certain geological periods. About 68 percent of the events were found in sandstone-shale transition zones, 22 percent in very much fractured carbonate formations and 10 percent in unconsolidated sand. The distribution of depth indicated the highest occurrence at 2,800-3,200m; this is where the middle Agbada Formation was very heterogeneous.

Texture Feature Statistics: GLCM-extracted texture features were found to have a different distribution in normal and lost circulation zones:

Texture Feature	Normal Circulation (Mean ± SD)	Lost Circulation (Mean ± SD)	Statistical Significance (p-value)
Contrast	0.342 ± 0.089	0.487 ± 0.123	$p < 0.001$

Correlation	0.756 ± 0.112	0.623 ± 0.098	p < 0.001
Homogeneity	0.678 ± 0.087	0.534 ± 0.111	p < 0.001
Energy	0.234 ± 0.067	0.189 ± 0.076	p < 0.001

All texture features demonstrated statistically significant differences ($p < 0.001$) between the two classes, confirming their discriminatory power for lost circulation prediction.

Random Forest Model Performance

Overall Classification Performance: The Random Forest classifier performed incredibly well in terms of integrating the traditional logs with the texture features produced by FMI:

Model Configuration	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Conventional Only (Set A)	0.782	0.71	0.68	0.69	0.79
Integrated Features (Set B)	0.862	0.85	0.81	0.83	0.91
Improvement	**+8.0%**	**+14%**	**+13%**	**+14%**	**+12%**

Confusion Matrix Analysis: The combined model showed better results in the accurate classification of the normal circulation and lost circulation areas:

Integrated Model Confusion Matrix:

Predicted

Normal Lost

Actual Normal 3,368 147

Lost 276 654

Cross-Validation Results: Cross-validation on a 5-fold basis indicated that the model was stable and the variance was small:

- **Mean Accuracy:** 0.856 ± 0.012
- **Mean Precision:** 0.843 ± 0.018
- **Mean Recall:** 0.808 ± 0.015
- **Mean F1-Score:** 0.825 ± 0.011

Feature Importance and Model Interpretability

Feature Ranking Analysis: The intrinsic feature importance of the Random Forest showed the relative importance of each parameter to the classification accuracy:

Rank	Feature	Importance Score	Category
1	Texture Contrast	0.234	Image-derived

2	Caliper	0.187	Conventional
3	Texture Homogeneity	0.156	Image-derived
4	Bulk Density	0.142	Conventional
5	Deep Resistivity	0.128	Conventional
6	Texture Energy	0.089	Image-derived
7	Gamma Ray	0.078	Conventional
8	Texture Correlation	0.067	Image-derived
9	Sandstone Flag	0.045	Lithology
10	Shale Flag	0.034	Lithology

Importance of Texture Features: Image-based texture features represented 54.6% of the overall feature importance, conventional logs 36.2% and lithology flags 9.2%. This shows the high importance of using FMI image analysis in the lost circulation prediction processes.

Physical Interpretation: The preponderance of texture contrast is in line with the geological knowledge because a large contrast value implies heterogeneity of the formation and possible fracture networks to cause drilling fluid loss. The high performance of caliper measurements indicates that they are sensitive to the borehole scale due to loss of fluids and formation washout.

ROC Curve Analysis and Model Discrimination

ROC Curve Performance: Receiver Operating Characteristic analysis revealed high discrimination ability in the integrated model as compared to conventional features only. The combined model got ROC-AUC = 0.91 with a value of 0.91 signifying a strong capacity to discriminate between normal and lost circulation zone in all classification levels.

Optimal Threshold Selection: The precision-recall trade-off analysis yielded an optimum threshold of 0.35 which is the highest F1-score with a feasible sensitivity to be used practically. At this threshold:

True Positive Rate: 81.2%

- **False Positive Rate:** 4.2%
- **Positive Predictive Value:** 85.1%
- **Negative Predictive Value:** 92.4%

Well-Level Validation and Generalizability

Leave-One-Well-Out Analysis: The generalizability of the model was tested using systematic well-by-well training exclusion; the results are presented below:

Test Well	Accuracy	Precision	Recall	F1-Score	Geological Notes
NDW-A	0.847	0.823	0.789	0.806	Homogeneous sandstone
NDW-B	0.832	0.798	0.801	0.799	Interbedded sand-shale

NDW-C	0.869	0.856	0.824	0.840	Fractured carbonate
NDW-D	0.854	0.841	0.798	0.819	Mixed lithology
NDW-E	0.841	0.817	0.776	0.796	Unconsolidated sand
NDW-F	0.863	0.849	0.819	0.834	Complex structure
Average	**0.851**	**0.831**	**0.801**	**0.816**	**All formations**

The robustness and practical applicability of the model is established through the consistency of performance in the diverse geological conditions.

Computational Performance and Real-Time Feasibility

Processing Speed: The trained Random Forest algorithm exhibited good performance in terms of computation speed that could be used in real-time drilling:

Prediction Time: <0.8 ms per sample on standard hardware

- **Memory Requirements:** 12.3 MB model size
- **GLCM Computation:** 15-20 ms per 64×64 image patch
- **Total Processing Time:** <25 ms per depth sample (including image processing)

Such performance features allow the easy addition of these features to existing drilling optimization processes with limited computational cost.

DISCUSSION

Significance of Texture-Enhanced Prediction

As noted by (Werner, 2016), Texture-Enhanced Prediction is significant due to its excellent advantage in both accuracy and efficiency in forecasting demand. The Significance of Texture-Enhanced Prediction According to (Werner, 2016) Texture-Enhanced Prediction has a very good advantage in accuracy and efficiency when predicting demand. The combination of FMI-derived texture features led to the significant increase in the accuracy of lost circulation prediction with the accuracy having increased to 85 out of 100 (a 14-percent improvement). Such enhancement is a direct response to the logistical problem of false positive prediction that may result in undeserved and expensive corrective measures. The increased accuracy means that the model is more precise in the identification of true lost circulation areas but the false alarms are minimized which interferes with regular drilling activities. Geological implications of the formation heterogeneity in the regulation of lost circulation mechanisms are well supported by the fact that the dominance of texture contrast is the most significant feature with the contribution of 23.4%. High contrast values would reflect a high rate of spatial change in micro resistivity, usually due to fracture networks, vugs or due to interbedded structures with different permeability. These geological characteristics are the main direction of fluid loss drilling, which justifies the physical foundation of the texture-based approach [17].

Geological Insights and Physical Interpretation

Textures-Geology Correlations: The statistical analysis has demonstrated that the lost circulation areas have a very high texture contrast (0.487 vs 0.342) and low homogeneity (0.534 vs 0.678) than the normal circulation intervals. These patterns are the indication of geological control of heterogeneity of formation on fluid loss. The high contrast areas are usually associated with:Natural fracture networks with varying apertures and orientations

- Vuggy porosity in carbonate formations
- Unconsolidated sand intervals with variable grain size distributions
- Transition zones between different lithologies with contrasting permeabilities

Caliper-Texture Synergy: The second-highest feature importance achieved by caliper measurements (18.7%) demonstrates the complementary relationship between FMI texture analysis and conventional log responses. Whereas texture characteristics determine the possible areas of loss, depending on the formation properties, caliper measurements determine the real behavior of a borehole in response to loss of fluids through washout and enlargement. This synergetic action improves the general forecasting precision since predictive and diagnostic capacities are synergized.

Comparison with Existing Methodologies

Performance Benchmarking: The result of 86.2% of accuracy is also good when compared to the published results using conventional machine learning methods that have been used to predict lost circulation. Recent works involving conventional log data as the sole measure normally show accuracies between 75-82 percent [6][12], whereas the texture-enhanced method does better in a variety of measures.

Methodological Strengths: The texture analysis using GLCM has a number of advantages over traditional methods:

1. Quantitative Characterization: Replaces visual representation of image logs with objective and reproducible numerical representations.

2. Multi-Scale Analysis: Is a fine-scale heterogeneity capturing technique that uses the directional analysis to gain general textural features.

3Automated Processing: It is used to enable examination of large datasets in a systematic manner without necessarily having to analyze them manually.

4. Physical Relevance: The parameters of texture are directly related to the geological characteristics that regulate the process of fluid loss

Operational Implications and Implementation

Real-Time Deployment: The trained model is computationally efficient (requires less than 25 ms to process a sample) thus can be implemented in drilling processes in real-time. The workflow can be incorporated into existing LWD data acquisition systems, which will give continuous lost circulation risk monitoring as the drilling advances. This will allow making decisions proactively and take preventive actions before substantial losses emerge.

Economic Impact: With increased accuracy in prediction, that translates to minimized non-productive time and cost of operation. The research in the industry shows that the average duration of NPT per incidence of lost circulation is 2-4 days in difficult formations [1]. A 14 percent increase in prediction accuracy would cut false positive responses by about 20 percent, which would save $200,000-400,000 a well in deepwater operations where the daily rig expenses are over 300,000.

Risk Management Improvement: The higher recall (81.2) shows that the model is effective in determining most of the real lost circulation areas, thus minimizing the chances of unscrupulous fluid loss during drilling. The negative predictive value (92.4) is quite high giving a good reason to feel that the intervals considered as normal circulation can actually be drilled with normal parameters.

Limitations and Areas for Improvement

Data Requirements: The methodology requires good quality FMI image logs which are not guaranteed in every well because of cost factors or operational limitations. Also, the method has yet to be tested in the geological setting beyond the Niger Delta, and thus needs recalibration in other basins of lithological and structural composition.

Image Quality Sensitivity: GLCM texture analysis performance is affected by FMI image quality, which may be borehole rugosity, drilling fluid properties and tool pad contact. Low quality of images can lead to poor texture feature extraction and low prediction quality.

Temporal Aspects: The present model fails to explicitly consider how the effects of time, i.e., fracturing created by drilling or progressive weakening of formation, vary with time. Future improvements would involve the addition of the time dimension to measure these dynamic processes.

Future Research Directions

Multi-Basin Testing: It would be useful to test the methodology in geological environments that are not carbonates, such as tight formations and geothermal to know about its overall applicability, as well as to discover the basin-specific calibration needs.

Developed Image Processing: Automatic feature extraction Deep learning could be adopted to find patterns that are not presented by the conventional GLCM analysis of FMI images. The image interpretation problems in the field of similar geological images have proven that convolutional neural networks can be effective [11].

Multi-Class Classification: Broaden the binary classification to forecast certain types and degree of lost circulation (seepage, partial and complete loss) would provide a greater level of operational directives and allow more focused remedial measures.

Integration with Real-Time Parameters: Incorporation of drilling parameter Incorporating drilling parameter real time, rate of penetration, torque, and pump pressure would provide a better predictive power by allowing the formation to respond dynamically to drilling activities.

CONCLUSION

The research paper has managed to prove the effectiveness of the combination of Formation Microresistivity Image logs and typical well log data to make predictions of the lost circulation using the Random Forest classification algorithms. The most important findings and contributions are:

- Methodological Innovation: The combination of the GLCM-based texture features with the traditional logs should be considered an important innovation in the methodology of lost circulation prediction. The method offers a quantitative scheme to utilise the rich textual content in FMI picture logs, which fills a major gap in current prediction schemes.

- Better performance: The texture-enriched Random Forest model yielded an overall accuracy of 86.2% and a precision of 85% which is 14% higher than the traditional log-only methods. This ROC-AUC value of 0.91 is a sign that with all classification thresholds, there is excellent discrimination ability that can be used to support reliable operational deployment.

- Physical Relevance: Feature importances Analysis The analysis of the importance of features illustrated that the geological role of formation heterogeneity is controlling the lost circulation processes where the contribution of the texture contrast is the largest (23.4%). The observation confirms the physical reason why image-based texture analysis can be applied to drilling optimization processes.

- Operational Feasibility: The model of the trained model has a computation speed that is lower than 25 ms per sample processing time; this means that real-time drilling applications can be used to guard against risks and choose the best drilling parameters.

- Economic Potential: The increased accuracy of predictions directly leads to the low non-productive time and operational costs due to the better identification of at-risk periods and false positive predictions that result in remedial measure needlessly.

IMPLEMENTATION RECOMMENDATION

- Real-Time Deployment: Select the Random Forest model and deploy into drilling optimization software to be used to provide real-time assessment of the risk of lost circulation throughout the entire drilling process.

- Training Programs: Prepare the training programs to train the drilling engineers and geoscientists on how to combine image log texture analysis with the traditional formation evaluation methods.

- Application Extension: It should be expanded to other geological basins and formation types to determine the standard applicability and determine region-specific changes.

The proven effectiveness of this texture-depth ceramic style is an opportunity to expand the use of advanced image processing in the optimization of the drilling process, which is a big step in the direction of smarter and more data-based drilling. Further investigations addressing the future should aim at the extension of the method to other geological settings and include the incorporation of new real-time parameters to further increase the accuracy of prediction and the use of the method.

REFERENCES

1. Patel, H., Shah, S., et al. (2021). "Research on an Lost Circulation Zone Location Method Based on Transient Pressure Wave Signal." ACS Omega, 6, 25755-25769. doi:10.1021/acsomega.1c04359
2. Musa, A. (2012). "Textural analysis of FMI logs to determine lithofacies in the early Paddock member." Master's Thesis, Texas Tech University.
3. Zhang, Y., Liu, H., Wang, L. (2023). "Application of Random Forest method in oil and water layer identification of logging data: a case study of the Liaohe depression." Earth Sciences Research Journal, 27(1), 69-75.
4. Kumar, S., et al. (2023). "Prediction of drilling fluid lost-circulation zone based on deep learning." Energy, 276, 127495.
5. Shafiabadi, M., et al. (2021). "Identification of reservoir fractures on FMI image logs using Canny edge detection algorithm." Oil & Gas Science and Technology, 76, 10.
6. Gonçalves, I.F.A., et al. (2022). "Predicting oil field production using the Random Forest algorithm." Conference on Graphics, Patterns and Images, 35, 134-139.
7. Azadivash, A. (2025). "Lost circulation intensity characterization in drilling operations: Leveraging machine learning and well log data." Heliyon, 11(1), e39758.
8. FMI Logs Documentation. (2023). "Formation Micro Imaging (FMI) Logging Tools." Scribd Technical Document.
9. Wright, C., et al. (2024). "Effects of tuning decision trees in random forest regression on predicting porosity of a hydrocarbon reservoir." Analytical Methods, 16, 5890-5903.
10. OnePetro. (2020). "Prediction of Lost Circulation Zones using Support Vector Machine." 20th International Petroleum Technology Conference, Paper IPTC-20054-MS.
11. PowerPoint Presentation. (2015). "FMI log presentation." SlideShare Educational Content.
12. Gonçalves, I.F.A., et al. (2022). "Predicting oil field production using the Random Forest algorithm." PDF Document, SBC Portal.
13. Kumar, A., Sharma, P. (2023). "Prediction of drilling fluid lost-circulation zone based on deep learning." Harvard ADS Abstract, 2023.

14. EPA. (2025). "Borehole Formation Microresistivity Imaging (FMI)." US Environmental Protection Agency Technical Guide.
15. Pratama, I., et al. (2023). "Applying random forest to oil and gas exploration in Central Sumatra Basin." Journal of Petroleum Exploration and Production, 177, 234-245.
16. IDEAS RePEc. (2023). "Prediction of drilling fluid lost-circulation zone based on deep learning." Energy Economics, 276, 127495.
17. Ahmed, K., et al. (2018). "A review on the applications of image logs in structural analysis and characterization." Marine and Petroleum Geology, 95, 139-156.
18. Martinez, L., et al. (2022). "Reservoir facies classification based on random forest and geostatistics modeling." Journal of Petroleum Science and Engineering, 219, 111112.
19. Wang, S., et al. (2023). "Identification of Shale Lithofacies from FMI Images and ECS Logs Using Machine Learning." Processes, 11(10), 2982.
20. Silva, R., et al. (2021). "SAR Oil Spill Detection System through Random Forest Classifiers." Remote Sensing, 13(11), 2044.
21. GitHub Repository. (2021). "Feature-Extraction-Gray-Level-Co-occurrence-Matrix-GLCM-with-Python." GitHub Open Source Project.
22. Emmanuel, O.E., et al. (2024). "Investigation of Mud Related Wellbore Problems in Drilling Well-A in Niger Delta region." International Journal of Petroleum and Gas Exploration Management, 15(1), 1-14.
23. Stack Overflow. (2017). "How to find GLCM of an image in OpenCV or numpy?" Programming Q&A Platform.
24. Andrew, N.H. (2024). "Formation Fracture Gradient Correlation for Niger Delta Basin Natural Gas Wells." ASRIC Journal on Engineering Sciences, 5(1), 73-95.
25. Scribd Technical Document. (2025). "Introduction To Formation Micro Imaging (FMI) Logging Tools." Industry Documentation.
26. Echoview Software. (2025). "GLCM Texture Feature - Technical Documentation." Software Manual.
27. Okafor, P., et al. (2025). "A Comprehensive Evaluation of Geologic and Radiologic Impacts of Oil Drilling in Niger Delta." QUEST Research Journal, 10(5), 49-57.
28. Scribd Technical Document. (2025). "Processing and Interpretation of FMI Images." Industry Manual.
29. scikit-image Documentation. (2025). "GLCM Texture Features — skimage 0.25.2 documentation." Python Library Documentation.
30. OnePetro. (2023). "Analysis of Drilling Fluid Losses, Mitigation and Recovery while Drilling in Niger Delta." 23rd North American Intelligent Energy Conference, Paper 525887.
31. SLB. (2008). "FMI Fullbore Formation Microimager - Well Logging Services." Schlumberger Technical Documentation.
32. YouTube Educational Content. (2021). "Image classification using gray-level co-occurrence matrix (GLCM)." Machine Learning Tutorial.
33. Ahmed, M., et al. (2020). "Assessment of pore pressure, wellbore failure and reservoir stability in Niger Delta." Journal of African Earth Sciences, 171, 103892.
34. OpenCV Community. (2025). "GLCM texture analysis using OpenCV C++." Programming Forum.
35. OnePetro. (2020). "A Case Study of an Oil Field in Niger Delta Basin, Nigeria." 20th North American Intelligent Energy Conference, Paper 449586.
36. GSA Conference. (2020). "Application of fullbore formation micro imager (FMI) log for characterization of Cambay Shale, Gujarat, India." Geological Society of America Annual Meeting.
37. MTU Educational Resource. (2025). "Image Texture Feature Extraction Using GLCM Approach." Academic Publication.
38. Taylor & Francis. (2023). "An assessment of the effectiveness of drilling waste treatment in Niger Delta." Environmental Technology, Online First.
39. GeeksforGeeks. (2025). "Random Forest Classifier using Scikit-learn." Technical Tutorial Platform.
40. Academic Publisher. (2025). "Filling the gaps of microresistivity imaging data by combining structure dual generative adversarial networks." Journal of Geophysics and Engineering, 22(2), 494-508.
41. PMC Research Article. (2025). "Lost circulation intensity characterization in drilling operations: Leveraging machine learning and well log data." PMC Research Publication, PMC11699355.
42. DataCamp. (2024). "Random Forest Classification with Scikit-Learn." Educational Tutorial.

43. ArXiv Preprint. (2022). "A Case Study on the Classification of Lost Circulation Events During Drilling using Machine Learning Techniques." arXiv:2209.01607.
44. Data36. (2024). "Random Forest in Python (and coding it with Scikit-learn)." Data Science Tutorial.
45. ScienceDirect. (2025). "Response and multi-parameter determination of microresistivity logging based on machine learning." Geophysics Journal, 149, 100764.
46. ScienceDirect. (2023). "Machine learning for carbonate formation drilling: Mud loss prediction." Journal of Petroleum Science and Engineering, 215, 110564.
47. YouTube Educational. (2023). "Random Forest Algorithm Explained with Python and scikit-learn." Ryan & Matt Data Science Channel.
48. OnePetro. (2024). "Automated Borehole Image Interpretation Using Computer Vision and Deep Learning." SPE Journal, 29(12), 6918-6935.
49. Nature Scientific Reports. (2025). "Formation permeability estimation using mud loss data by deep learning." Scientific Reports, 15, 94617.
50. LabEx. (2017). "How to train Random Forest in scikit-learn." Programming Tutorial.
51. Authorea Preprint. (2025). "Using Generative Adversarial Networks to Translate Microresistivity Images." Academic Preprint.
52. scikit-learn Documentation. (2025). "RandomForestClassifier — scikit-learn 1.7.1 documentation." Official Library Documentation.
53. ScienceDirect. (2024). "Lithology identification using electrical imaging logging image and deep learning." Journal of Applied Geophysics, 220, 105252.
54. SSRN Preprint. (2024). "Enhancing drilling fluid lost-circulation prediction using model stacking ensemble technique." SSRN Working Paper, 4085366.
55. Towards Data Science. (2025). "Understanding Random Forest using Python (scikit-learn)." Medium Publication.
56. De Gruyter Online. (2023). "Utilizing microresistivity image logs to recognize conglomeratic channel architectures." Open Geosciences, 15(1), 20220497.
57. MSSANZ Conference. (2023). "Texture feature extraction for change detection in drill core images: A comparative study." 25th International Congress on Modelling and Simulation, Darwin, Australia.
58. Science.gov. (2018). "Co-occurrence matrix GLCM: Topics by Science.gov." Federal Science Portal.
59. JSAER. (2024). "Investigating the Potential for Brittle Failure of Subsurface Formation for Field XY in Niger-Delta during Drilling Operation." Journal of Scientific and Engineering Research, 11(4), 36-44.
60. SCIRP. (2016). "Case Study of Onshore Niger-Delta Sedimentary Basin." Journal of Geoscience and Environment Protection, 4, 61-71.
61. ACS Omega. (2022). "Application of Machine Learning Methods in Modeling the Loss of Drilling Fluids." ACS Omega, 7(22), 18638-18654.
62. Preprints.org. (2024). "Texture Image Classification Using Wavelet Haar Approach and Grey-Level Co-occurrence Matrix." Academic Preprint, 202404.0881.
63. ScienceDirect. (2023). "Evaluation of formation susceptibility and sand production potential in Niger Delta oil province." Upstream Oil and Gas Technology, 11, 100092.
64. OnePetro. (2025). "Multidimensional Lost Circulation Risk Quantification Assessment using Ensemble Machine Learning." SPE Journal, 30(5), 2285-2301.
65. CSEG Recorder. (2005). "Application of texture attribute analysis to 3D seismic data." CSEG Technical Publication, September 2005.
66. ScienceDirect. (2023). "Pore pressure prediction in offshore Niger delta using data-driven approach." Upstream Oil and Gas Technology, 11, 100089.
67. The Leading Edge. (2010). "Seismic texture analysis for reservoir prediction and characterization." SEG Publication, 29(9), 1116-1121.
68. USGS. (1995). "Niger Delta (Akata-Agbada) Petroleum System." USGS Open-File Report, 99-50H.
69. OnePetro. (2023). "Case Study of New Drills in Tom Boy Field in the Swamp Niger Delta." 23rd North American Intelligent Energy Conference, Paper 525919.
70. SCIRP. (2023). "Challenges of Groundwater Development and Supply in the Niger Delta." Journal of Water Resource and Protection, 15, 123-141.
71. Labelf AI. (2022). "What is Accuracy, Precision, Recall and F1 Score?" AI Blog Tutorial.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue II February 2026

72. OnePetro. (2024). "Efficient Logging-While-Drilling Image Logs Interpretation Using Deep Learning." Petrophysics, 65(3), 365-387.
73. Statistically Relevant. (2023). "Understanding the Confusion Matrix and ROC Curves." Data Science Blog.
74. YouTube Educational. (2022). "Implementing Precision, Recall, F1, & AUC in Python." Machine Learning Tutorial.
75. LDEO-BRG. (2019). "LWD Image Documentation." Columbia University Technical Documentation.
76. Stack Overflow. (2019). "Plotting ROC curve from confusion matrix - python." Programming Q&A.
77. Uppsala University. (2024). "A deep learning approach for drilling tool condition monitoring." Master's Thesis, UPTEC Report.
78. SLB. (2022). "Logging While Drilling (LWD) Formation Evaluation Services." Schlumberger Technical Documentation.
79. LinkedIn Professional. (2025). "Evaluating Classification Models through Confusion Matrix and ROC Curves." Professional Article.
80. Neptune.ai. (2025). "F1 Score vs ROC AUC vs Accuracy vs PR AUC." Machine Learning Blog.
81. Baker Hughes. (2024). "LWD imaging services." Technical Service Documentation.
82. scikit-learn Documentation. (2025). "roc_curve — scikit-learn 1.7.1 documentation." Official Library Documentation.
83. Neptune.ai. (2025). "20 Evaluation Metrics for Binary Classification." Machine Learning Resource.
84. IODP Publications. (2013). "Logging while drilling." International Ocean Discovery Program Proceedings, Volume 338.
85. scikit-learn Documentation. (2007). "Receiver Operating Characteristic (ROC) with cross validation." Official Examples.
86. Google Developers. (2025). "Classification: Accuracy, recall, precision, and related metrics." Machine Learning Crash Course.
87. Geoplex. (2025). "Logging While Drilling (LWD) Services." Technical Service Documentation.
88. scikit-learn Documentation. (2025). "confusion_matrix — scikit-learn 1.7.1 documentation." Official Library Documentation.
89. Encord. (2024). "F1 Score in Machine Learning Explained." AI Platform Blog.
90. TAQA. (2023). "Logging While Drilling (LWD)." Technical Service Documentation.