Research Progress of Edge Intelligent Class Imbalanced Learning for Gas Well Fluid Accumulation Prediction

Jingguo Du[1], Yan He[2], ManMan Feng[3], Wei Qin[4], Li Wang[5], Qi Dai[6,*]
[1]College of Mining Engineering, North China University of Science and Technology, Tangshan, China
[2]Oil and Gas Field Surface Engineering, Sichuan Changning Natural Gas Development Co., Ltd., Yibin, China
[3]Sichuan Jiacheng petroleum and natural gas pipeline quality inspection testing Co., LTD, Chengdu, China
[4]Chongqing Gas Field, PetroChina Southwest Oil & Gas Field Company, Chongqing, China
[5]Institute of Exploration and Development, Southwest Oil and Gas Field Company, PetroChina, Chengdu, China
[6]College of Sciences, North China University of Science and Technology, Tangshan, China
**Corresponding Author*

Abstract: **Liquid accumulation in gas wells is one of the most common and destructive wellbore flow problems in natural gas development. Since gas cannot be effectively discharged, liquid accumulates in the well, leading to increased bottomhole back pressure, decreased gas production, and increased drainage costs. With the widespread adoption of wellhead sensors, SCADA systems, and digital gas field platforms, edge intelligence is becoming a crucial support for real-time gas well monitoring. However, gas-well liquid accumulation prediction data suffers from problems such as scarce liquid accumulation events, significant operational condition shifts, and obvious label lag, making imbalanced learning a critical step for the successful implementation of liquid accumulation prediction. This paper systematically reviews the gas well liquid accumulation mechanism and prediction task modeling, data-driven modeling, imbalanced learning methods, and edge intelligence system architecture. Analysis results show that considering the imbalanced data distribution problem can improve the accuracy of gas well liquid accumulation analysis and early warning.**

Keywords: **Gas Well Fluid Accumulation; Anomaly Detection; Edge Intelligence; Class Imbalance Learning; Cloud Collaboration**

1. Introduction

In the mid-to-late stages of natural gas development, gas well fluid accumulation has become one of the core issues restricting stable production of single wells and the overall development benefits of gas fields. Well fluid accumulation is not simply the presence of liquid at the bottom of the well but rather an insufficient ability of the gas phase to lift the liquid phase to maintain continuous drainage, leading to a series of chain reactions such as droplet fall, liquid film reversal, liquid column growth, and wellhead production fluctuations. When the above process continues, the bottom pressure rises continuously, and gas production decreases significantly, leading to a passive situation of well shutdown or frequent drainage [1].

For a long time, the identification of liquid accumulation boundaries in engineering sites has mainly relied on critical liquid-carrying theory, with the Turner model, Coleman correction, and Li Min model forming the classic framework for droplet and critical gas velocity analysis [2-4]. Subsequently, researchers introduced liquid film reversal, well inclination effects, transient behavior of gas well and reservoir coupling, and calculation of the temperature and pressure field throughout the well section, shifting the understanding of liquid accumulation from a single critical point judgment to a description of the dynamic evolution process [5]. With the deepening of digital transformation in oil and gas fields and the widespread deployment of IoT sensors, gas well production sites have accumulated massive amounts of high-frequency, multi-dimensional monitoring data. This has prompted a shift in the strategy for dealing with gas well liquid accumulation from traditional post-event diagnosis of mechanisms to pre-event analysis driven by data such as anomaly detection and time-series early warning [6]. Furthermore, on-site labeling largely relies on manual inspection and post-event retrospective analysis, inevitably introducing a large amount of time lag and labeling noise.

The core concept of edge intelligence technology [7] is to bring the massive computing power of the cloud down to the wellhead edge node closest to the data source [8]. Prediction algorithms are deployed directly in the wellhead intelligent terminal equipment, and real-time processing and on-site analysis are performed at the source. This new architecture completely reshapes the traditional data flow path, significantly reduces network latency caused by long-distance communication, effectively alleviates the transmission load pressure on the backbone network, and provides extremely high real-time assurance for the advanced prediction of gas well fluid accumulation [6]. Therefore, deeply integrating the on-site computing architecture of edge intelligence with the class imbalance learning to construct a new gas well fluid accumulation prediction method has become an important research direction for solving this complex engineering problem [9].

2. Modeling the Mechanism-Driven Problem of Predicting Liquid Accumulation in gas Wells

From the perspective of two-phase flow in the gas well, the essence of liquid accumulation in gas wells is the disruption of the balance between the gas phase's upward lifting capacity and the liquid phase's downward trend. In the early stages of production, the reservoir provides sufficient energy, and the formation pressure is high, allowing the gas flow to carry the liquid to the surface in the form of droplets or thin liquid films at a relatively high velocity. In the later stages of production, the formation pressure depletes, and the wellhead pressure drop increases. When the gas velocity in the gas well falls below the critical liquid-carrying capacity, the droplets begin to fall back, and the liquid film on the gas well wall reverses. The liquid gradually accumulates at the bottom of the well, at the low points of the horizontal section, and in local throttling zones [10].

Based on the above mechanism, droplet models, represented by the Turner model, derive the critical liquid-carrying gas velocity through the force balance of droplets. This model is simple in form and widely used [2]. However, it is relatively sensitive to the assumptions regarding droplet morphology and flow pattern and is relatively conservative in its predictions for some actual gas wells. Coleman et al. [3] modified the model for low-pressure wells, making it more applicable to engineering applications under some low-pressure conditions. The Li Min model further modified the droplet morphology assumptions, attempting to explain the conservative nature of the Turner model in some actual gas wells [4]. Furthermore, since liquid accumulation does not necessarily begin from the inability of droplets to rise, it may also begin from the reverse flow of the liquid film on the wellbore and the change in flow pattern [4]. Veeken et al. [11] emphasized the influence of liquid film reversal, well inclination, and flow pattern change on the point of liquid accumulation.

Transient and coupled models no longer treat fluid accumulation as a single critical point problem but instead incorporate temperature and pressure fields, wellbore flow, fluid column growth, and reservoir recharge processes into a unified analytical framework. Pagan and Waltrich [1] analyzed erratic behavior and shutdown or restart processes using simplified transient models. Liu et al. [5] established a coupled prediction model of wellbore temperature and pressure fields and fluid accumulation. Although mechanistic models are irreplaceable in theoretical analysis and engineering interpretation, they have many parameters, and some parameters are difficult to obtain in real time and correct online. Therefore, most models are better at determining whether a gas well has reached the critical fluid-carrying condition but have not yet depicted the severity of fluid accumulation, early warning lead time, and dynamic evolution process under complex operating conditions [12].

3. Research Progress on Data-Driven Gas Well Fluid Accumulation Prediction Methods

3.1 Machine Learning-Based Fluid Accumulation Prediction Models

In early data-driven research on gas well fluid accumulation prediction, linear baseline models such as discriminant analysis and logistic regression were often used as rapid classification tools in the field. Zhu et al. [13] constructed a rapid classification and diagnosis process based on real-time production data using LDA-DA. However, linear baseline models usually implicitly assume that the class boundaries can be linearly separated and the variable relationships are relatively stable. The effective signals of gas well fluid accumulation data appear in the form of threshold transitions, interaction terms, and lag relationships, and the boundaries shift with changes in well type and system. When the training instance size is limited, but there are obvious nonlinear boundaries between pressure, production, and fluid volume, support vector machines can be used as a more robust, moderately complex solution [14].

Compared with linear models, tree models, such as random forest, GBDT, XGBoost, and LightGBM, are more suitable for processing tabular production data in wellhead monitoring. In a recent study, Xia et al. [15] used COA-XGBoost to predict the height of fluid accumulation in gas well tubing in a continuous volume prediction task.

3.2 Deep Learning-Based Fluid Accumulation Prediction Models

Convolutional neural networks (CNNs) automatically extract stable patterns from local windows and are sensitive to local combinations of multiple variables. For gas-well liquid accumulation tasks, pressure, production, and liquid volume sequences often exhibit local fluctuation patterns, abnormal coupling relationships, or short-term peak-valley structures before liquid accumulation. CNNs have good representation capabilities for these patterns. Jiang et al. [6] used CNNs and ensemble learning to collaboratively determine liquid accumulation status in a wellhead edge system, achieving an integrated closed loop of data acquisition, discrimination, uploading, and action recommendations. Zhao et al. [16] converted well-type indicators into two-dimensional representations and input them into a 2D-CNN for gas well classification prediction, verifying the gain of convolutional feature combinations on low-dimensional engineering variables.

In gas well production scenarios, gated recurrent networks such as LSTM and GRU can retain key historical states during sequence updates, thereby better describing wellbore pressure recovery, production decline, stabilization processes after regime switching, and intermittent production rhythms. For intermittent production wells or wells with frequent start-stop cycles, these models are more effective than static classifiers in identifying the difference between early signs of fluid accumulation and short-term disturbances [17].

For multi-well systems, the core idea of anomaly detection and autoencoder representation learning is to first learn the stable pattern of normal operating conditions and then identify potential fluid accumulation based on reconstruction error, characterization deviation, and anomaly score [18]. Anomaly detection [19] is more suitable as a first-stage screening tool, and secondary confirmation still needs to be completed through operating condition rules or supervised models. Recent high-quality research has begun to move from simple binary classification to severity grading and few-label learning. Chen et al. [20] proposed a semi-supervised fluid load severity classification model and introduced a dynamic time relationship self-attention module, achieving earlier warnings than traditional physical models on data from 219 shale gas wells.

Besides state classification, depth regression is another area of interest in recent years. Chen et al. [21] proposed a depth regression method to directly predict the continuous state related to the liquid load of gas wells and verified its timely prediction capability on real well data. Compared with pure classification methods, the regression framework is more suitable for outputting continuous indicators such as liquid column height, risk score, and remaining safety margin, thus providing a more intuitive quantitative basis for process optimization and economic evaluation [22].

In conclusion, the real challenge for future research is not deeper networks, but rather unifying operational stratification, label verification, few-shot learning, edge deployment, and policy decision-making within the same research framework.

4. Edge-Based Intelligent Class Imbalance Learning for Gas Well Fluid Accumulation

Prediction

4.1 Traditional Imbalanced Learning

Traditional class imbalance learning is a methodology system designed for scenarios with significantly skewed class distributions in classification tasks. It aims to alleviate the dominance of majority class samples on the decision boundary and improve the recognition performance and generalization ability of the minority class. In real-world data, class imbalance is prevalent in areas such as fault diagnosis of industrial equipment and traffic analysis of network intrusions.

Based on different processing mechanisms, this type of method can be roughly divided into data-level, algorithm-level, and hybrid methods. Data-level methods optimize the learning performance of the classifier by adjusting the distribution of training instances. Classic data-level methods include oversampling strategies represented by synthetic minority oversampling techniques and their derivatives and undersampling strategies represented by random undersampling and neighborhood cleaning [23]. Algorithm-level methods enhance the model's focus on the minority class by modifying the learning mechanism, including cost-sensitive learning, ensemble learning methods, and single-class learning and anomaly detection paradigms. Hybrid methods combine the advantages of the above two types of strategies, maintaining the integrity of the minority class decision boundary while suppressing the risk of overfitting.

In the field of oil and gas extraction, some researchers have attempted to address the class imbalance problem in the data. Ni et al. [24] proposed an analysis method combining table generative adversarial network data augmentation and meta-learning methods to improve the accuracy of gas-oil ratio prediction under data scarcity conditions. Jamshidi Gohari et al. [25] considered the data distribution imbalance problem caused by the heterogeneity of geological conditions and alleviated the impact of class distribution imbalance on model performance by introducing a deep learning method based on transfer learning. Yi et al. [26] designed an improved label propagation algorithm based on semi-supervised learning, which reduced the imbalance ratio of the original data by assigning high-confidence labels to unlabeled samples. Wood et al. [27] believed that class imbalance and the lack of high correlation between drilling variables and fluid loss categories affected the recognition performance of machine learning models or deep learning models. Kim and Byun [28] used periodically consistent generative adversarial networks to enhance the synthetic data of simulated well pairs, alleviating the problems of insufficient well logging data and class distribution imbalance.

In summary, gas well fluid accumulation prediction is a typical imbalanced class problem. Fluid accumulation instances are sparsely distributed in measured data. Standard classifiers are prone to biased predictions of normal samples due to optimizing global accuracy, resulting in significantly insufficient recall for minority instances. Furthermore, the cost of misclassification is inherently asymmetric. The economic consequences and operational risks caused by missed fluid accumulation reports and false positives differ by orders of magnitude, and traditional evaluation metrics struggle to capture this asymmetric loss. Therefore, exploring imbalanced learning methods in gas well fluid accumulation problems is a necessary methodological foundation for achieving reliable early warning of gas well fluid accumulation.

4.2 Application of Edge Intelligence in Oil and Gas Monitoring

Edge intelligence, as a core paradigm for breaking through the communication bottlenecks of traditional centralized cloud computing, has demonstrated irreplaceable architectural advantages in the field of IoT monitoring of oil and gas fields [21]. The edge node layer, as a local computing power hub of the entire architecture, is deployed directly at the physical boundary close to the data source, greatly shortening the network link for data interaction and fundamentally eliminating the system latency caused by long-distance communication [6].

At the critical edge node level, the physical form of the hardware and the allocation of computing power exhibit highly customized industrial characteristics. Wellhead computing devices, acting as the computing outpost directly connected to the underlying instruments, can stably perform real-time data parsing and primary time-series filtering in extremely harsh field mining environments. Industrial gateways bear the heavy responsibility of heterogeneous network protocol conversion and multi-source sensor data fusion, ensuring that wellbore parameters acquired by different sized acquisition terminals can be uniformly mapped to a standardized feature space.

4.2.1 Federated learning and multi-well collaborative training

Gas well production sites are typically characterized by highly dispersed geographical locations and a severe lack of data on abnormal operating conditions in individual wells. Federated learning breaks the physical limitations of traditional centralized data aggregation, allowing multiple wellhead edge nodes to collaboratively update the global early warning model by interacting with the gradients or weight parameters of local neural networks without sharing the underlying original sensor sequences [29]. This method greatly alleviates the problem of local model overfitting caused by the unbalanced distribution of abnormal data in individual wells, making it possible to construct a global liquid accumulation diagnostic architecture with strong cross-well area generalization capabilities [30].

4.2.2 Online learning for dynamic working conditions

The physical evolution of downhole gas-liquid two-phase flow is a typical non-stationary dynamic time-series process. The natural decay of reservoir energy and the human intervention of production pipeline systems lead to significant conceptual drift in the distribution of bottom-level sensor data. Online learning algorithms allow intelligent early warning models deployed on edge computing gateways to perform real-time fine-tuning and incremental evolution of parameters based on the constantly generated fresh data streams [17].

4.2.3 Lightweight technology for edge hardware

In the deployment of complex deep learning architectures to edge devices, the massive parameter matrix must be slimmed down by deep algorithms due to the extremely stringent power consumption standards and limited computing bandwidth of wellhead instruments. Model lightweighting technology constitutes the core key to bridging this physical deployment gap. At the level of basic network topology construction, researchers usually prefer to choose lightweight backbone networks specifically customized for mobile and embedded scenarios, using deep separable convolution mechanisms to strip away redundant computational parameters [31]. In the field of hardware-software co-processing stacks, the increasing maturity of TinyML technology has made running high-dimensional time series prediction models directly on microwatt-level microcontrollers a new benchmark in the industry [7].

4.3 Edge-Based Intelligent Gas Well Liquid Accumulation Prediction System

The edge-intelligent liquid accumulation prediction system deeply couples a distributed sensing network with cutting-edge lightweight algorithms, thereby establishing a closed-loop dynamic monitoring link at the gas well production site. The typical workflow of the entire prediction system follows a strict data flow and hierarchical progression logic. Sensor acquisition, as the front end of the entire sensing system, relies on various precision instruments deployed in the wellbore and surface pipeline network to continuously and frequently capture multi-dimensional physical parameters such as wellhead pressure, casing pressure, production temperature, and gas-liquid flow rate [6].

The high-throughput feature stream, after standardized preprocessing, is then transmitted to the lightweight prediction network of the edge AI chip, thereby initiating the core real-time prediction and anomaly resolution process. Since the complex neural network reasoning process has been fully decentralized to the production front line, the system can intelligently classify the current fluid state of the gas well and determine the liquid accumulation trend within milliseconds. The novel edge prediction system, which closely aligns with the physical boundaries of oil and gas extraction, exhibits significant advantages that traditional centralized cloud computing architectures cannot match. By adopting a front-end core prediction computing power approach, the entire early warning paradigm achieves extremely high real-time performance, and the on-site dispatch system can accurately capture early, weak signals of liquid accumulation with near-zero communication latency. This system completely abandons the traditional inefficient mode of forcibly uploading the entire underlying sequence to the cloud [8].

5. Research Challenges and Future Works

Due to the extreme complexity of underground fluid evolution and the harsh physical deployment environment of well sites, this field still faces many technical bottlenecks that urgently need to be overcome in its progress towards large-scale industrial application. Future academic exploration and engineering practice can focus on the following five core directions.

5.1 Mining Weak Features under Extreme Long-Tail Distribution

In real oil and gas field production networks, the ratio of normal, stable data to abnormal fluid accumulation data is unbalanced, with some special blocks exhibiting extreme data biases of up to 100 times or more. Traditional boundary synthesis or heuristic undersampling techniques are prone to algorithm failure under such extreme long-tail distributions. Future explorations need to move beyond the traditional shallow physical sampling approach and turn to deep metric learning or zero-sample synthesis techniques based on generative artificial intelligence, striving to accurately anchor weak and hidden abnormal fluid evolution manifolds even under extremely disparate prior probabilities [8].

5.2 Physics- Prior- Driven Few- Shot Generalization Mechanism

For newly commissioned and remote gas wells, the model faces challenges including imbalanced distributions across different categories and an extreme scarcity of absolute outlier data. Conventional deep neural networks, heavily reliant on massive sample training, will completely lose their feature fitting advantages. Future research should prioritize the deep integration of oil and gas field physics mechanisms with data-driven algorithms. By directly embedding fundamental fluid dynamics equations and multiphase flow phase evolution laws as hard penalty terms into the underlying loss function of the neural network, the edge early warning system can be endowed with agile adaptive diagnostic capabilities even with very few observation instances.

5.3 Deep Fusion Sensing of Cross-Modal Multi-Source Heterogeneous Data

Existing gas well fluid accumulation prediction systems heavily rely on numerical time-series sensor sequences, neglecting the rich textual and discrete process event information accumulated over many years in oil and gas fields. Future edge intelligent monitoring systems will inevitably undergo a comprehensive leap towards multimodal holographic perception. By creatively introducing lightweight large language models or dynamic graph neural network architectures, the system can deeply align and fuse high-frequency wellhead physical sensor stream data with low-frequency macroscopic production management logs, regular manual inspection records, and historical maintenance records of critical equipment through spatial semantic alignment and high-dimensional feature fusion. This provides comprehensive and highly three-dimensional causal logic support for underlying anomaly early warnings.

6. Conclusion

This paper systematically reviews the inevitable trend of evolution in the field of intelligent early warning of gas well fluid accumulation from traditional centralized cloud computing to edge intelligent physical architecture. Faced with the increasingly complex oil and gas field development environment and massive, high-frequency IoT sensor flows, traditional monitoring, which heavily relies on remote servers, suffers from communication latency and bandwidth consumption. Deep embedding of class imbalance learning theory constitutes the key to breaking through the bottleneck of gas well anomaly detection algorithms. Because real physical mining sites naturally exhibit the objective law that normal operating conditions constitute the overwhelming majority while fluid accumulation anomaly samples are extremely scarce, conventional data-driven models are prone to falling into the catastrophic majority preference trap. By cleverly integrating lightweight resampling technology, cost-sensitive networks, and edge federated architecture on edge devices, modern intelligent early warning systems can accurately extract subtle early fluid accumulation evolution characteristics within extremely limited hardware power consumption boundaries, thereby completely reversing the passive situation of on-site engineering managers in dealing with sudden production reductions.

Reference

[1] Pagan E V, Waltrich P J. A simplified

model to predict transient liquid loading in gas wells. Journal of Natural Gas Science and Engineering, 2016, 35: 372-381.

[2] Turner R G, Hubbard M G, Dukler A E. Analysis and prediction of minimum flow rate for the continuous removal of liquids from gas wells. Journal of Petroleum technology, 1969, 21(11): 1475-1482.

[3] Coleman S B, Clay H B, McCurdy D G, et al. A new look at predicting gas-well load-up. Journal of petroleum technology, 1991, 43(03): 329-333.

[4] Li M, Li S L, Sun L T. New view on continuous-removal liquids from gas wells. SPE Production & Facilities, 2002, 17(01): 42-46.

[5] Liu H, Lou W, Li H, et al. A modified comprehensive prediction model for wellbore temperature-pressure field and liquid loading of gas wells. Geoenergy Science and Engineering, 2024, 232: 212452.

[6] Jiang J, Li K, Du J, et al. Prediction system for water-producing gas wells using edge intelligence. Expert Systems with Applications, 2024, 247: 123303.

[7] Savaglio C, Mazzei P, Fortino G. Edge intelligence for industrial iot: Opportunities and limitations. Procedia Computer Science, 2024, 232: 397-405.

[8] Liu X, Dong X, Jia N, et al. Federated learning-oriented edge computing framework for the IIoT. Sensors, 2024, 24(13): 4182.

[9] Zhu G, Liu X, Niu J, et al. Learning by imitating the classics: Mitigating class imbalance in federated learning via simulated centralized learning. Expert Systems with Applications, 2024, 255: 124755.

[10] Lin J, Xu S, Liu Y. Experimental and modeling studies on continuous liquid removal in horizontal gas wells. Frontiers in Earth Science, 2023, 11: 1288208

[11] Veeken K, Hu B, Schiferli W. Gas-well liquid-loading-field-data analysis and multiphase-flow modeling. SPE Production & Operations, 2010, 25(03): 275-284.

[12] Wang W, Zhang X, Wei Y, et al. A new mechanistic model to predict liquid loading in inclined natural gas wells. Geoenergy Science and Engineering, 2024, 241: 213163.

[13] Zhu Z, Han G, Liang X, et al. Rapid classification and diagnosis of gas wells driven by production data. Processes, 2024, 12(6): 1254.

[14] Rezvani S, Wang X. A broad review on class imbalance learning techniques. Applied Soft Computing, 2023, 143: 110415.

[15] Xia W, Liu B, Xiang H. Prediction of liquid accumulation height in gas well tubing using integration of crayfish optimization algorithm and xgboost. Processes, 2024, 12(9): 1788.

[16] Zhao C, Jia Y, Qu Y, et al. Forecasting gas well classification based on a two-dimensional convolutional neural network deep learning model. Processes, 2024, 12(5): 878.

[17] Zhou W, Sun W, Ding W, et al. A Deep Learning-Based Model for Real-Time Production Forecasting of Gas Wells with Intermittent Production. SPE Journal, 2025, 30(03): 1377-1395.

[18]Chen Y, Huang Y, Miao B, et al. Adaptive anomaly detection-based liquid loading prediction in shale gas wells. Journal of Petroleum Science and Engineering, 2022, 214: 110522.

[19]Liu Q X, Zhu J J, Wang H B, et al. Deep feature learning for anomaly detection in gas well deliquification using plunger lift: A novel CNN-based approach. Petroleum Science, 2025, 22(9): 3803-3816.

[20]Chen P, Chen Y, Yang C, et al. Gas well production optimization: Classifying liquid loading severity in shale gas wells using semi-supervised learning. Gas Science and Engineering, 2024, 128: 205394.

[21]Chen Y, Miao B, Wang Y, et al. A deep regression method for gas well liquid loading prediction. SPE Journal, 2024, 29(04): 1847-1861.

[22]Kheirollahi H, Zayedi M, Simjoo M, et al. Machine Learning-Based Prediction of Liquid Loading in Gas Wells for Flow Assurance Management and Production Optimization[J]. Results in Engineering, 2025: 29, 108605.

[23] Gao X, Xie D, Zhang Y, et al. A comprehensive survey on imbalanced data learning. Frontiers of Computer Science, 2026, 20(11): 2011622.

[24] Ni Y, Zhang H, Han Y, et al. Study on gas-oil ratio prediction considering the influence of imbalance data. Petroleum

Science and Technology, 2026, 44(7): 1142-1161.

[25]Jamshidi Gohari M S, Niri M E, Sadeghnejad S, et al. Synthetic graphic well log generation using an enhanced deep learning workflow: imbalanced multiclass data, sample size, and scalability challenges. SPE Journal, 2024, 29(01): 1-20.

[26] Yi J, Qi Z L, Li X C Z, et al. A novel class-imbalance learning framework for fluid recognition: Application to Qingshimao-Gaoshawo tight-sand gas reservoirs in the Ordos Basin, China. Computers & Geosciences, 2025: 205, 105993.

[27] Wood D A, Mardanirad S, Zakeri H. Effective prediction of lost circulation from multiple drilling variables: a class imbalance problem for machine and deep learning algorithms. Journal of Petroleum Exploration and Production Technology, 2022, 12(1): 83-98.

[28] Kim D, Byun J. Selection of augmented data for overcoming the imbalance problem in facies classification. IEEE Geoscience and Remote Sensing Letters, 2021, 19: 1-5.

[29] Li H, Ge L, Tian L. Survey: federated learning data security and privacy-preserving in edge-Internet of Things. Artificial Intelligence Review, 2024, 57(5): 130.

[30] Dritsas E, Trigka M. Federated learning for IoT: A survey of techniques, challenges, and applications. Journal of Sensor and Actuator Networks, 2025, 14(1): 9.

[31] Kulkarni U, Meena S M, Gurlahosur S V, et al. Quantization friendly mobilenet (qf-mobilenet) architecture for vision based applications on embedded platforms. Neural Networks, 2021, 136: 28-39.