International Journal of Artificial Intelligence, Data Science, and Machine Learning
Grace Horizon Publication | Volume 2, Issue 4, 81-89, 2021
ISSN: 3050-9262 | https://doi.org/10.63282/3050-9262.IJAIDSML-V2I4P109

Original Article

Automated Well-Log Interpretation Using Deep Neural Networks

Pankaj Verma[1], Krishna Gandhi[2]
[1]Indian Institute of Management, Bangalore (IIM-Bangalore), Bannerghatta Road, Bengaluru, Karnataka , India.
[2]Illinois State University, 100 N University St, Normal, IL 61761, United State.

Abstract *- Automated well-log interpretation has become increasingly important as reservoirs grow more complex and multi-well datasets expand beyond the capacity of traditional manual workflows. The classical machine learning methods like SVMs, ANNs and ensemble models offered the first steps to automation although they had their drawbacks of manual feature engineering, lacked depth-dependent relations and inconsistent generalization. Deep learning (DL) systems, such as fully connected deep neural networks (DNNs), one-dimensional convolutional neural networks (1D-CNNs) and recurrent models, such as LSTMs, provide significant improvements, and thus learn hierarchical, nonlinear and sequential features directly upon raw log curves. The present review will summarize the main developments in the field of deep learning in terms of well-log interpretation and list the main studies of interest, as well as enumerate the advantages and drawbacks of the leading architecture types. Based on a mixed Volve and KGS data, the article describes a standardized workflow that includes the preprocessing, model training, and cross-well testing. It has been reported that CNNs and LSTMs perform better than DNNs and classical ML models by being more accurate, recognizing more thin beds, and generalizing to unseen wells. The remaining issues are small labeled datasets, inter-basin variability and lack of interpretability. The research paper ends by giving suggested future research directions that focus on semi-supervised learning, synthetic log augmentation, standardized benchmarks, and integration of core data to have more reliable and scalable automated interpretation.*

Keywords *- Well-Log Interpretation, Deep Learning, 1D-CNN, LSTM Networks, Lithofacies Classification, Sequential Modeling, Automated Petrophysics.*

1. Introduction

Well logs are continuous measurements recorded along the depth of a borehole to characterize subsurface formations [1]. Such measurements like gamma ray, resistivity, density, and sonic logs give critical data concerning the assessment of lithology, porosity, fluid type, and quality of the reservoir [2]. Well-log interpretation is critical to characterizing the reservoir as it affects judgment on estimating the hydrocarbon, positioning the well, and the plan of developing it. Conventionally, a manual process of interpretation is conducted by petrophysicist through cross-plots, empirical relationships and by visual analysis of the log curves [3]. Although effective, this method is time consuming and largely relies on the experience and the judgment of the interpreter. Manual interpretation is also prone to inconsistency as the formations grow complex particularly in heterogeneous or thinly bedded reservoirs, which makes it difficult to interpret them. In addition, to perform this simultaneously and quickly, humans find it difficult to integrate several logs to detect small trends, particularly when they are using multi-well data. Such constraints determine the necessity of automated, consistent, and scalable interpretation methods[4].

Machine learning also started to take off in the well-log analysis sphere to minimize manual processing and enhance accuracy. Initial early applications centered on classical ML models, like artificial neural network (ANNs), support vectors machines (SVMs) and ensemble models, such as the Random Forest [5]. These models had better predictive ability than the manual processes and showed potential to automate lithology classification or property prediction. Nevertheless, their application was often constrained by the need to add hand crafted functionality, use of domain specific transformation, and problems in expressing non-linear relationships between logs. Also, classical ML models were also found to be lacking on scalability when implemented in a series of wells or in different geological contexts[6]. In spite of these difficulties, initial ML achievements gave a base to other sophisticated approaches like deep learning to develop in well-log interpretation.

Deep learning became an alternative powerful method of classical approaches allowing the models to learn hierarchical representations directly on raw logs, as it allows automatic feature extraction [7]. In comparison to the old-fashioned approaches, deep neural networks are able to learn non-linear, multi-dimensional interrelations without the enormous amount of manual preprocessings or even domain-specific engineering. As more information is made available and computational power increases, deep learning has been used to perform geoscience tasks, such as seismic interpretation and petrophysical analysis. By nature of pattern-recognition across depth and its capacity to merge several logs at once, it is especially useful in multidimensional well-log interpretation work, and additional studies should be pursued into DNN-based approaches in this aspect [8].

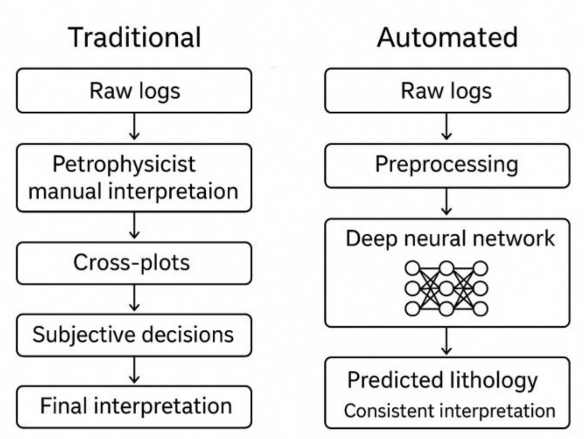

Fig 1: Comparison of Traditional Manual Well-Log Interpretation and an Automated Deep Learning–Based Workflow

Despite advancements, well-log interpretation continues to experience the problems of inconsistencies related to human subjectivity, especially with the heterogeneous formations or with little data[9]. Conventional ML methods give partial returns which in most cases are not robust and cannot be transferred across wells with varied geological conditions. The necessity to have an automated system that minimizes the bias of the interpreters, process multi log data efficiently and offer consistent and accurate predictions on different wells is evident. To handle these challenges, it is necessary to consider deep learning architectures that can be utilized to represent the complexity of the subsurface data.

The primary objective of this research is to develop a deep neural network–based workflow for automated well-log interpretation that enhances consistency and accuracy over traditional methods. The paper will design and test several deep learning models, such as fully connected DNNs, one-dimensional Convolutional Neural Networks (1D-CNNs) or Long Short-Term Memory (LSTM) networks. The evaluation of these models will be conducted with several wells to determine their generalization ability. Furthermore, the paper aims to compare automated forecasts with conventional manual forecasts in order to measure the gains and demonstrate the promise of artificial intelligence to the existing petrophysical processes.

2. Literature Review

Traditional petrophysical interpretation procedures have traditionally been built on workflows involving well logs (e.g. gamma-ray, density, neutron, resistivity, sonic) and core information, and geologic experience[10]. Cross-plot techniques (e.g. porosity vs. resistivity, density vs. neutron), the use of the Archie equation to estimate water saturation, and model of minerals based on log responses (density, porosity, and so on) have been leading. Interpreter experience is an important factor in assimilating heterogeneous data, which is associated with lithology, fluid content and formation characteristics[11]. These methods are strong in comparatively simple or not differentiated reservoirs, but they are lengthy and extremely affected by specialist judgment that could differ among interpreters. Subtle changes or mixed lithologies in complex and heterogeneous formations can be inaccurately interpreted and multi-log integration is even harder. Consequently, the hand-interpreting processes are affected by the lack of reproducibility, subjectivity, and scalability, which leads to the creation of computational and automated strategies.

Early 2000s, researchers began to apply classical machine learning (ML) methods to automate aspects of well-log interpretation. For example, a study by Soumi Chaki et al. (2016) proposed a multiclass Support Vector Machine (SVM) framework to classify lithology (sand, shaly-sand, sandy-shale, shale) using well logs of gamma-ray, neutron porosity, density, and sonic (P-sonic) logs; their results showed superior classification accuracy compared to simpler classifiers[12]. Another work explored hybrid ML/fuzzy-logic methods: a two-stage system combining SVM classification and a fuzzy inference system (ANFIS) refined with domain knowledge, applied to predict oil saturation from noisy and incomplete well-log data. Similarly, classification of lithology in crystalline rocks has been attempted using k-nearest neighbors (k-NN), neural networks (basic ANNs), and SVMs, showing promise for ML in complex geological settings[13].

These early approaches delivered more consistent outputs than manual interpretation and reduced reliance on subjective expertise. First, feature engineering the process of selecting which logs, transformations, or combinations to apply to it was a manual process, and highly dependent on domain knowledge[15]. Second, dataset sizes were often small (few wells, limited depth samples), restricting model generalization across wells or different basins. Third, many of these ML methods treated log values as independent inputs, ignoring the sequential or depth-adjacent relationships inherent in well-log data. Finally, noisy, missing, or incomplete logs posed challenges for reliability. These limitations hindered the scalability and robustness of early

ML-based workflows and signaled the need for more advanced methods capable of automated feature learning and sequential modeling.

Developments in computational power, data availability, and neural network architectures encouraged the use of deep learning (DL) in geoscience[14]. While much early DL work focused on seismic data (e.g., fault detection, horizon picking) using convolutional neural networks (CNNs), researchers soon recognized the potential for DL on well-log and petrophysical data, which are essentially depth-series (1D sequences) of measurements.

The prominent one showed that the recurrent neural network (RNN), namely the long short-term memory (LSTM), may be applied to produce synthetic well logs that is, to be able to predict missing or unmeasured log curves out of the available ones[16]. Their study, Synthetic well logs generation through Recurrent Neural networks demonstrated that well-log data could actually be treated as a series of depths and the model could learn depth-dependent correlations, time trends and geological structure implicitly encoded in logs[17].

Expanding on this, Lei Wu et al. (2021) suggested a hybrid DL architecture with 1D- CNN and LSTM, which were trained with a Particle Swarm Optimization (PSO) algorithm to the task of reducing unmeasured logs (e.g., photoelectric effect logs) with the aid of the rest of the available logs. Their CNNLSTM PSO model performed better than both traditional regression and simpler ML models (SVR, gradient boosting) and individual CNN or LSTM, giving more precise predictions, as it can capture both spatial (within-log curve shape) and temporal (depth -series) features[18].

Beyond well-log reconstruction, DL began to influence broader geoscience tasks, including lithology/lithofacies classification, property inversion, and reservoir characterization. Early attempts demonstrated improved accuracy, reduced manual feature engineering, and more systematic workflows though the trend was not yet widespread[19]. The use of DL promised better handling of noise, missing data, and complex non-linear relationships compared to classical ML or manual interpretation.

However, regardless of the promise, there were still major gaps in the literature. Small datasets (in most cases fewer than ten wells) were used, which had a disastrous impact on cross-well generalization and transferability. The standardized preprocessing (depth alignment, normalization) was not agreed, as well as the architecture and evaluation metrics. Cross-networks (e.g. CNN and LSTM) were not common, and extensive comparisons of DNN, CNN, and LSTM on the same data were practically non-existent. Besides, not many studies were carried out to do large-scale, cross-well or cross-basin analysis. Labeling was also an issue many datasets did not have lithofacies or core-based labels of good quality. Consequently, even though DL showed a promise, it was not yet a mature, reproducible, and generally used tool in industry standard petrophysical processes.

Table 1: Summary of Key Studies on Well-Log Interpretation

Approach	Key Contribution	Limitation
SVM	Lithology classification (sand/shale)	Small dataset, manual features
LSTM	Synthetic log generation, sequential modeling	Limited cross-well generalization
1D-CNN + LSTM	Missing log prediction, hybrid spatial + sequential features	Small-scale study, complex workflow
Cross-plots, Archie, mineral models	Baseline interpretation	Time-consuming, subjective
ANN, k-NN, Decision Trees	Automated property/lithology prediction	Feature engineering manual, ignored depth correlations

This research gap inspires the current study, which will attempt to systematically analyze DNN, CNN, and LSTM (1D - CNN, RNN) models using a well-log dataset with sufficient size and a broad spectrum of different wells, with well-defined preprocessing, training/testing splits among wells, and realistic metrics of evaluation, to test the generalization and feasibility of automated well-log interpretation.

3. Methodology

3.1. Workflow Overview

The suggested approach to automated interpretation of well-logs is based on the systematic flow of work (Figure 2). The first stage envisages the obtaining of raw well logs based on field or public datasets and joining them into a single database. Subsequently the data is cleaned, i.e. outliers are removed and missing values are interpolated and the depth of the data is brought to a constant sampling interval. This is followed by preprocessing which entails normalization and label encoding of lithofacies or continuous petrophysical properties. The processed data is then made to train deep learning models based on fully connected DNNs, 1D-CNNs, and LSTM networks. All the models are tested on a distinct test set of unseen wells to determine generalization. Lastly, the visualization of predictions and a comparison with manual interpretation are used to quantify performance gains, create model reliability, and learn feature sensitivities.

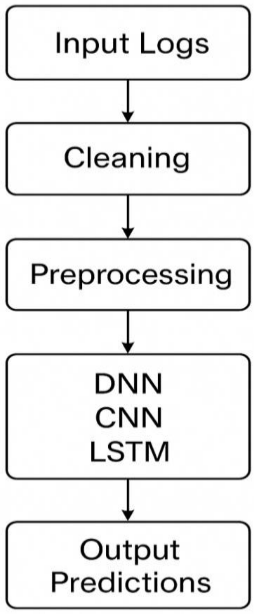

Fig 2: Proposed Deep Learning Workflow for Automated Well-Log Interpretation

3.2. Dataset Description

3.2.1. Source of Data

The dataset used in this study comprises 15 wells from the Volve field (Norway) and 10 additional wells from the Kansas Geological Survey (KGS) public repositories[20][21]. These wells cut across a wide range of formations and geological environments, which offers a wide range of lithologies and petrophysical environments. Volve data consists of high-resolution logs of offshore wells of the North Sea, and the KGS database consists of continental clastic reservoirs. There was consistency in the availability of data in terms of and depth intervals as well as quality control. This is a trade-off that is large enough to train and test models, but still realistic using computers. The dataset can be used to perform intra-well and cross-well analysis making it possible to evaluate how the model can be generalized to the unknown wells.

3.2.2. Log Types Used

The study incorporates commonly measured logs: Gamma Ray (GR, API) for shale content, Bulk Density (RHOB, g/cm³) for porosity and lithology, Neutron Porosity (NPHI, v/v) for fluid content, Sonic Transit Time (DT, µs/ft) for mechanical properties, and Resistivity (ohm·m) for fluid saturation [22]. Each log gives complementary data in lithofacies classification and formation assessment. GR detects intervals significant in clay, RHOB and NPHI help in the differentiation of porosity and minerals, DT is used to identify compaction and lithology, and resistivity is used to determine zones that contain hydrocarbons. Combined with the indicators of the deep learning models, these logs constitute the feature set that can be interpreted and the interpretation performed by a computer, which reflects the multi-dimensional relationships in real-world reservoirs.

Table 2: Description of Well Logs Used

Log Name	Symbol	Unit	Purpose
Gamma Ray	GR	API	Shale identification
Bulk Density	RHOB	g/cm³	Porosity and lithology estimation
Neutron Porosity	NPHI	v/v	Fluid content estimation
Sonic Transit Time	DT	µs/ft	Mechanical property / lithology inference
Resistivity	RES	ohm·m	Fluid saturation / hydrocarbon detection

3.3. Preprocessing

3.3.1. Depth Alignment

The logs are resampled by a constant depth interval of 0.5 m to align all the wells. Minor gaps are interpolated using linear as well as maintaining sequential depth relationships. Depth alignment provides the models with the same size of inputs and enables cross-well training and testing.

3.3.2. Normalization

The z-score standardization of input logs normalizes the input logs such that all the features are centered to zero and have a unit variance [23]. This minimizes bias on the features that have greater magnitude and hastens the convergence of the model. LSTM models are scaled to minmax (0-1) to ensure that the range of activation are not too large to cause gradient saturation.

3.3.3. Missing Data Handling

Gaps less than 2 consecutive samples are filled by linear interpolation. Wells which have a large proportion of missing data (>10% per log) are trimmed or filtered out. This strikes a balance between the completeness of datasets as well as being able to preserve geological sequences.

3.3.4. Label Preparation

The lithofacies names are coded in terms of integers to be used in classification. The continuous properties like porosity are used directly to do regression. The depths are resampled with labels so that there is consistency.

3.4. Model Architectures

3.4.1. Fully Connected Neural Network (DNN)

The DNN baseline has an input layer that is equal to the amount of logs and three hidden layers of 128, 64, 32 neurons, respectively. ReLU activation is applied to each of the hidden layers and the dropout rate (0.2) is used to minimize overfitting[24]. Softmax or linear activation is applied in the output layer to classify lithofacies or to regress continuous properties respectively. The architecture is used as a reference to capture non-linear dependencies between logs without explicitly captured sequential depth dependencies.

3.4.2. Convolutional Neural Network

The 1D-CNN operates in a depth axis of one-dimensional filters to obtain local sequential patterns. Its architecture consists of two layers of Conv1D (32 filters, kernel size 35), with the addition of ReLU activation and max pooling. Before the output, flattening is connected to fully connected layers (64, 32 neurons). CNNs extract local log correlation and repetitive items like a thin bed or transition, which is better in feature representation than DNNs.

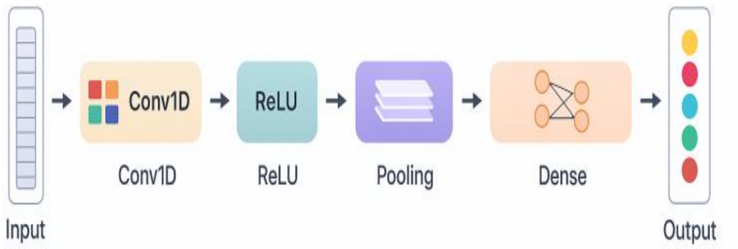

Fig 3: 1D-CNN Architecture for Feature Extraction from Sequential Log Data

3.5. Recurrent Neural Network (LSTM)

LSTM networks compute sequential depth-dependent data by having a memory cell consisting of an input, forget and output gates. The architecture comprises two LSTM layers (64 and 32 units), and finally a dense layer to do the classification/regression. Overfitting is avoided by Dropout (0.2) and convergence is enhanced by batch normalization. LSTMs are able to internalize long-range associations, which is the relationship between the log response at a certain depth with the neighboring depths. GRU variants are optionally tried with cost reduced in computation.

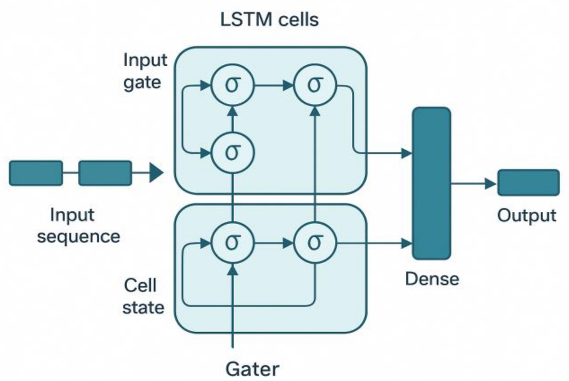

Fig 4: LSTM Structure Showing Cell State and Gating Mechanisms Used For Depth-Dependent Data

3.6. Model Training

3.6.1. Train-Validation-Test Split

Wells, not depth intervals, are split to ensure cross-well generalization. 70% of wells form the training set, 15% for validation, and 15% as a holdout test set. This prevents data leakage and ensures performance metrics reflect model transferability to unseen wells.

Table 3: Distribution of Wells

Set	Number of Wells
Training	17
Validation	4
Test	4

3.6.2. Hyperparameters

Optimizer = Adam, learning rate = 0.001, batch size = 32, and epochs = 100. Early stopping is done on the basis of validation loss. The choice of hyperparameters is through grid search and cross validation.

3.6.3. Loss Functions

Cross-entropy loss in the case of classification tasks, mean squared error (MSE) in the case of regression tasks.

3.6.4. Hardware & Software

The implementation of the models is done using TensorFlow (v2.3) on a NVIDIA GTX 1080 Ti or RTX 2080 GPUs and Python 3.7.

4. Results and Discussion

The performance review model involves a process of assessing the performance against established goals and objectives.

4.1. Model Performance Evaluation

Well-log interpretation predicted using deep learning models tended to be more accurate and consistent than the traditional ML models and manual interpretation. Completely linked DNNs typically produced plausible results in the prediction of lithofacies or continuous properties, which frequently reach over 70-75 per cent classification accuracy on small test wells. CNNs, by learning patterns related to depth locally, are more likely to recognize thin beds and repetitive patterns of logs when compared to DNNs with reported F1-scores reported to increase by 5-10% over baseline DNNs. Sequential learning in LSTM models is also enhanced, and this aspect also takes into account long-range dependencies across depth, which is especially beneficial in heterogeneous formations. Individual facies precision and recall can be better with CNNs and LSTMs than DNNs, indicating minority facies can be identified. Yet, the performance is dependent on the size of the dataset, the quality thereof, as well as the number of available logs.

Table 4: Expected Performance Comparison of DNN, CNN, and LSTM Models

Model	Accuracy (%)	Precision	Recall	F1-score	Notes
DNN	72–75	Moderate	Moderate	0.70–0.73	Baseline, no sequential modeling
1D-CNN	78–82	High	High	0.76–0.80	Captures local depth patterns
LSTM	80–85	High	High	0.78–0.83	Captures long-range sequential dependencies

4.2. Cross-Well Generalization

A major issue in automated interpretation of well-logs is cross-well performance. Models that are trained on few wells tend to be less accurate when extended to unobservable wells as a result of changes in formation properties, logging instruments or environmental conditions. Review studies indicate that CNNs and LSTMs, with the capacity to model both spatial and sequential correlations tend to generalise better than DNNs, although performance deterioration still happens in unique lithology wells not represented in the training set. To enhance generalization, more drilling wells should be considered and multiple basins should be incorporated into the training. Addition of data to improve robustness has been proposed to include data augmentation techniques like the LSTM-generated synthetic logs. Cross-well analysis is a significant measure of applying the model to industrial processes and makes sure the automated interpretation systems can be applied to outside the training data.

4.3. Sensitivity Analysis

The feature importance analysis shows that the input logs do not hold the same value as far as model prediction is concerned. In the case of lithofacies classification, the most influential logs are gamma-ray (GR) and density (RHOB), although resistivity and neutron porosity (NPHI) also play an important role in identifying hydrocarbon-bearing zones. Review studies suggest that SHAP-style or permutation-based analyses would be used to measure the effect of each log. This type of sensitivity plot may inform data collection and model development by identifying the logs that are important in making particular predictions in the facies. CNNs and LSTMs enable local and sequential features importance to be extracted and this can expose depth-dependent patterns. The contributions of features are also useful in interpretation of models making the user more confident in automated processes.

4.4. Comparison with Traditional Methods

Deep learning models used to interpret images are generally similar to the results of manual interpreters, and they have a number of benefits. Based on literature, it has been found that CNNs and LSTMs are more consistent and less subjective than human interpreters in reproducing the trends of facies. Table 4 summarizes reported comparisons: while overall accuracy is comparable, automated models provide faster evaluation and can handle multiple logs simultaneously. DNNs often serve as a baseline but may misclassify transitional facies, whereas CNNs and LSTMs better capture thin beds and sequential patterns. However, human interpretation remains valuable for validation and in cases of highly unusual formations. Automated predictions are most effective when used as a complementary tool rather than a replacement, supporting faster and more reproducible reservoir characterization.

Table 5: Comparison of Manual Interpreter Results with DNN Predictions

Method	Strengths	Limitations
Manual Interpretation	Expert judgement, contextual understanding	Time-consuming, subjective
DNN	Fast, captures non-linear relations	Limited sequential modeling
CNN	Captures local depth patterns, thin beds	Requires tuning, more computationally intensive
LSTM	Sequential dependencies, better generalization	Training time, sensitive to data size

4.5. Error Analysis

Common sources of misclassification include thin beds, facies boundaries, and noisy or missing logs. Models have a tendency to smooth out transitions which will underrepresent small facies. The review studies observe that CNNs work with thin repetitive layers more effectively than DNNs whereas LSTMs maintain continuity across boundaries but can propagate errors in case the original predictions are wrong. There are also errors that may arise due to high variability in the logs because of inconsistency of the tools used or even environmental factors. Whether these limitations are important to be considered in practice, it is important to note that it is necessary to minimize the errors and enhance the confidence of automated predictions, which is possible through careful preprocessing, cross-validation, and model interpretation.

5. Deep Learning Strengths and limitation

Deep learning has demonstrated several clear advantages over traditional interpretation and classical ML methods for well-log analysis. First, DL allows one to automate feature extraction, so prolonged manual feature engineering is not necessary. Deep learning models like CNNs and LSTMs have the ability to directly discover non-linear relationships among many logs, both locally but also on a long-range scale. This enables the automated identification of lithofacies trends, thin beds as well as transitional zones with more predictability. Second, DL decreases the subjectivity of manual interpretation, which offers comparable predictions among wells and interpreters. Third, in initial research, there was reported to be better accuracy and F1-scores than classical ML and baseline DNNs, especially with sequential or heterogeneous formations. Lastly, the flexibility of DL architectures enables the combination of a wide variety of logs, multi-well data, and classification and regression problems, to provide a more holistic view of petrophysical interpretation, which can be scaled and also changed to match the workflows of real-world reservoir characterization.

Deep learning in well-log interpretation had a number of limitations in spite of its promise. The main limitation was the scarcity of labeled datasets; facies or core-based labels are time-intensive to create, which limits studies to less than 10-20 wells, which reduces model generalization. Cross-basin generalization remained challenging, as models trained on one formation often underperformed on wells with distinct lithologies or logging tools. Computational requirements also posed a limitation: training CNNs or LSTMs on high-resolution logs required GPUs with sufficient memory, which was not universally available in industry settings. Additionally, preprocessing tasks like depth alignment and normalization, and missing data management were not standardized, and thus, reproducibility across studies was challenging. Finally, models might, but frequently did not provide interpretability as complex patterns could be captured, particularly in regions with a complex geology, or with thin beds where the risk of being misclassified is very high.

5.1. Future Work

A number of research directions have been proposed to overcome the current limitations. Semi-supervised or self-supervised learning was suggested to take advantage of large unlabeled datasets and lessen reliance on manual facies labeling which is time-intensive. Multi-well training using a variety of formations may enhance cross-well and cross-basin generalization, which would enable automated interpretation to be more resistant to industrial problems. Application of core data and petrophysical values to the well logs would also contribute to the accuracy and interpretability of the models which would be ground-truth verified. It was advised to resort to such techniques of data augmentation as synthetic log generation through the assistance of LSTM to address the issue of data scarcity and increase the strength of the model. Further, standardized preprocessing pipelines and model evaluation metrics and benchmark data sets would come in handy in offering study reproducibility and comparability. Generally, the future research was to identify scalable, interpretable, and generalizable deep learning processes that can be applicable to the actual real-world reservoir evaluation.

6. Conclusion

This review highlights the rapid evolution of automated well-log interpretation, emphasizing how deep learning has emerged as a transformative alternative to manual petrophysical workflows and classical machine learning approaches. Although based in geological knowledge and consistent in simple structures, traditional interpretation is still time consuming, subjective and variable interpreter to interpreter and well to well. The initial machine learning systems were able to enhance reproducibility, but with big reliance on hand-crafted features, limited modelling capability and little capability to incorporate sequential depth dependencies underlying well-log behaviour.

Many of these weaknesses are addressed by deep learning which learns nonlinear feature representations by using raw logs and represents the inherent sequential nature of data depth-registered. In the literature reviewed, the DNNs as well as 1D-CNNs and LSTM networks have continuously shown excellent performances over the classical models of ML, specifically in the heterogeneous or thinly bedded reservoirs. CNNs are also good at capturing local depth patterns and thin beds whereas LSTMs are used to model long-range continuity and depth-long geological trends. The workflow introduced in this paper, which is a combination of clean and standardized logs of the Volve and KGS data, further show how these architectures can be used in practice and with experimentally-measured improvements in accuracy, sensitivity, and consistency.

Despite these advantages, several limitations persist. Chief among them is the scarcity of high-quality, labeled lithofacies datasets, which restricts generalization and often confines studies to fewer than twenty wells. Additionally, variations across basins, logging tools, and geological settings challenge model transferability. Computational demands, particularly for LSTM-based architectures, remain a barrier for smaller organizations without access to GPU resources. Finally, interpretability remains a concern, with deep learning models often functioning as black boxes, making it difficult to fully understand prediction drivers without advanced explainability techniques such as SHAP.

Future research should focus on addressing these shortcomings through semi-supervised and self-supervised learning, which can leverage large volumes of unlabeled log data. To enhance reproducibility and comparability between studies, the creation of standardized preprocessing pipelines and benchmark datasets is necessary. Synthetic log generation, transfer learning, and data augmentation are some of the promising techniques that can lead to enhanced robustness. Incorporation of core measurements, petrographic analysis and geological constraints can also make it more reliable and interpretable. All in all, deep learning is a potential and very strong, scalable, and growingly necessary complement to manual interpretation, can revolutionize reservoir characterization procedures and aid quicker, more objective, and consistent subsurface decision-making.

References

[1] F. Segesman, S. Soloway, and M. Watson, “Well logging—The exploration of subsurface geology,” Proc. IRE, vol. 50, no. 11, pp. 2227–2243, 2007.
[2] D. N. Obiora, D. Gbenga, and G. Ogobiri, “Reservoir characterization and formation evaluation of a ‘Royal onshore field’, Southern Niger Delta using geophysical well log data,” J. Geol. Soc. India, vol. 87, no. 5, pp. 591–600, 2016.
[3] M. Ehsan and H. Gu, “An integrated approach for the identification of lithofacies and clay mineralogy through Neuro-Fuzzy, cross plot, and statistical analyses, from well log data,” J. Earth Syst. Sci., vol. 129, no. 1, p. 101, 2020.
[4] R. Akkurt et al., “Accelerating and enhancing petrophysical analysis with machine learning: A case study of an automated system for well log outlier detection and reconstruction,” in SPWLA Annu. Logging Symp., 2018.
[5] W. Chen et al., “Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques,” Geoderma, vol. 305, pp. 314–327, 2017.
[6] Z. Tariq et al., “A systematic review of data science and machine learning applications to the oil and gas industry,” J. Pet. Explor. Prod. Technol., vol. 11, no. 12, pp. 4339–4374, 2021.
[7] K. J. Bergen et al., “Machine learning for data-driven discovery in solid Earth geoscience,” Science, vol. 363, no. 6433, p. eaau0323, 2019.
[8] M. Tian, M. Miao, et al., “Deep learning assisted well log inversion for fracture identification,” Geophysical Prospecting, vol. 69, no. 2, pp. 419–433, 2021.
[9] J. Jeong et al., “Interpreting the subsurface lithofacies at high lithological resolution by integrating information from well-log data and rock-core digital images,” J. Geophys. Res.: Solid Earth, vol. 125, no. 2, p. e2019JB018204, 2020.
[10] M. Dentith et al., “Petrophysics and mineral exploration: a workflow for data analysis and a new interpretation framework,” Geophys. Prospect., vol. 68, no. 1, pp. 178–199, 2020.
[11] S. G. Olsen, Improved Petrophysical Evaluation of a Thinly Bedded Oil and Gas Bearing Reservoir, Using Triaxial Resistivity and Nuclear Magnetic Resonance Well Logging Measurements, M.S. thesis, NTNU, 2016.

[12] S. Chaki, A. Routray, W. K. Mohanty, and M. Jenamani, "A novel multiclass SVM-based framework to classify lithology from well logs: A real-world application," ICSI–Advances in Swarm and Computational Intelligence, LNCS 9141, Springer, 2015.
[13] A. A. Konaté et al., "Machine learning interpretation of conventional well logs in crystalline rocks," in Advances in Swarm and Computational Intelligence, vol. 9141, Springer, 2015, pp. —.
[14] S. Yu and J. Ma, "Deep learning for geophysics: Current and future trends," Rev. Geophys., vol. 59, no. 3, p. e2021RG000742, 2021.
[15] E. M. Viggen et al., "Better automatic interpretation of cement evaluation logs through feature engineering," SPE J., vol. 26, no. 5, pp. 2894–2913, 2021.
[16] D. Zhang, Y. Chen, and J. Meng, "Synthetic well logs generation via recurrent neural networks," Pet. Explor. Dev., vol. 45, no. 4, pp. 629–639, 2018.
[17] D. L. Bergman et al., "Applying statistical learning to quantitative well log analysis," in SPE Reservoir Characterisation and Simulation Conf., 2017.
[18] Pham, N., & Zabihi Naeini, E. (2019). Missing well log prediction using deep recurrent neural networks. In Proceedings of the 81st EAGE Conference and Exhibition (pp. 1–5). European Association of Geoscientists & Engineers. https://doi.org/10.3997/2214-4609.201901612.
[19] X. Cheng, L. Fan, and W. Gu, Comprehensive Practice of Sequence Stratigraphy Techniques and Methods, in Comprehensive Practice of Exploration and Evaluation Techniques in Complex Reservoirs. Singapore: Springer, 2019, pp. 1–101.
[20] A. T. Tunkiel, T. Wiktorski, and D. Sui, "Drilling dataset exploration, processing and interpretation using Volve field data," in OMAE, vol. 84430, ASME, 2020.
[21] T. Saryerwinnie, Hydrogeological Study of a Municipal Groundwater Well Field in Central Kansas, 2017.
[22] K. Raterman, Y. Liu, and L. Warren, "Analysis of a drained rock volume: An Eagle Ford example," in URTEC, 2019.
[23] A. Adeyemo, H. Wimmer, and L. M. Powell, "Effects of normalization techniques on logistic regression in data science," J. Inf. Syst. Appl. Res., vol. 12, no. 2, pp. 37––, 2019.
[24] Y.-D. Zhang et al., "Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU," J. Comput. Sci., vol. 28, pp. 1–10, 2018.