Advances in Geo-Energy Research

Vol. 13, No. 3, p. 161-165, 2024

Perspective

Deep learning in CO_2 geological utilization and storage: Recent advances and perspectives

Yanwei Wang[1,2]*, Hongyang Chu[3], Xiaocong Lyu[4]

[1] *School of Energy Resources, China University of Geosciences (Beijing), Beijing 100083, P. R. China*

[2] *Beijing Key Laboratory of Unconventional Natural Gas Geological Evaluation and Development Engineering, China University of Geosciences (Beijing), Beijing 100083, P. R. China*

[3] *School of Civil and Resources Engineering, University of Science and Technology Beijing, Beijing 100083, P. R. China*

[4] *State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum, Beijing 102249, P. R. China*

Keywords:
Deep learning
CO_2 geological utilization
CO_2 storage
intelligence modeling
practical applications

Cited as:
Wang, Y., Chu, H., Lyu, X. Deep learning in CO_2 geological utilization and storage: Recent advances and perspectives. Advances in Geo-Energy Research, 2024, 13(3): 161-165.
https://doi.org/10.46690/ager.2024.09.01

Abstract:
Deep learning has been widely recognized in the field of CO_2 geological utilization and storage applications. With the development of deep learning algorithms, intelligent models are gradually able to improve multi-source, multi-scale and multi-physicochemical mechanism barriers with high-fidelity solutions in practical applications. In this perspective, an overview of the traditional and state-of-the-art deep learning architectures involved in CO_2 geological utilization and storage is outlined in terms of evolutionary trajectories. Meanwhile, the favorable directions and application scenarios of different deep learning algorithms for geo-energy intelligence modeling are summarized. Moreover, further insights into the future direction of deep learning burgeoning architectures in this field are provided. The physics-guided deep learning, explainable artificial intelligence, and generative artificial intelligence are expected to deliver more accurate solutions for information extraction and decision support within the CO_2 geological utilization and storage communities.

1. Introduction

CO_2 geological utilization and storage (CGUS) technology is expected to realize carbon storage and underground energy development at the same time, which can make CO_2 change from "nuisance in the sky" to "treasure in the ground" (Xu et al., 2022; Noyce et al., 2023). However, there are still serious challenges to improving the efficiency and safety of carbon sequestration in geological structures while ensuring the benefits of hydrocarbon and geothermal resources development.

The process of CGUS involves multi-scale transformations in time and space, multi-dimensional parameter integration of reservoir and engineering, and multi-physical field effects of migration and transformation (Qi et al., 2023). Although numerical simulations have become popular, their prediction and optimization processes require a lot of computational resources, and cannot meet the requirements of integrated prediction and multi-objective optimization (Wang et al., 2023). Therefore, there is an urgent need to develop efficient and accurate decision support tools, considering all techno-economic and environmental aspects of the various segments. The great success of machine learning and deep learning can provide suitable candidates for the development of such decision-support tools, which have gained strong momentum in practical engineering applications (Liu et al., 2023).

Originally inspired by biological models of computation and cognition in the human brain, deep learning can contribute to addressing complex and challenging problems in CGUS, such as dynamic prediction of storage and production, reservoir parameter inversion, and multi-objective optimization (Al-Sakkari et al., 2024). Deep learning models have been outperforming other traditional machine learning methods, es-

Yandy Scientific Press

*Corresponding author.
E-mail address: wangyw@cugb.edu.cn (Y. Wang); hongyangchu@126.com (H. Chu); x.lyu@cup.edu.cn (X. Lyu).

Received May 1, 2024; revised May 11, 2024; accepted May 16, 2024; available online May 18, 2024.

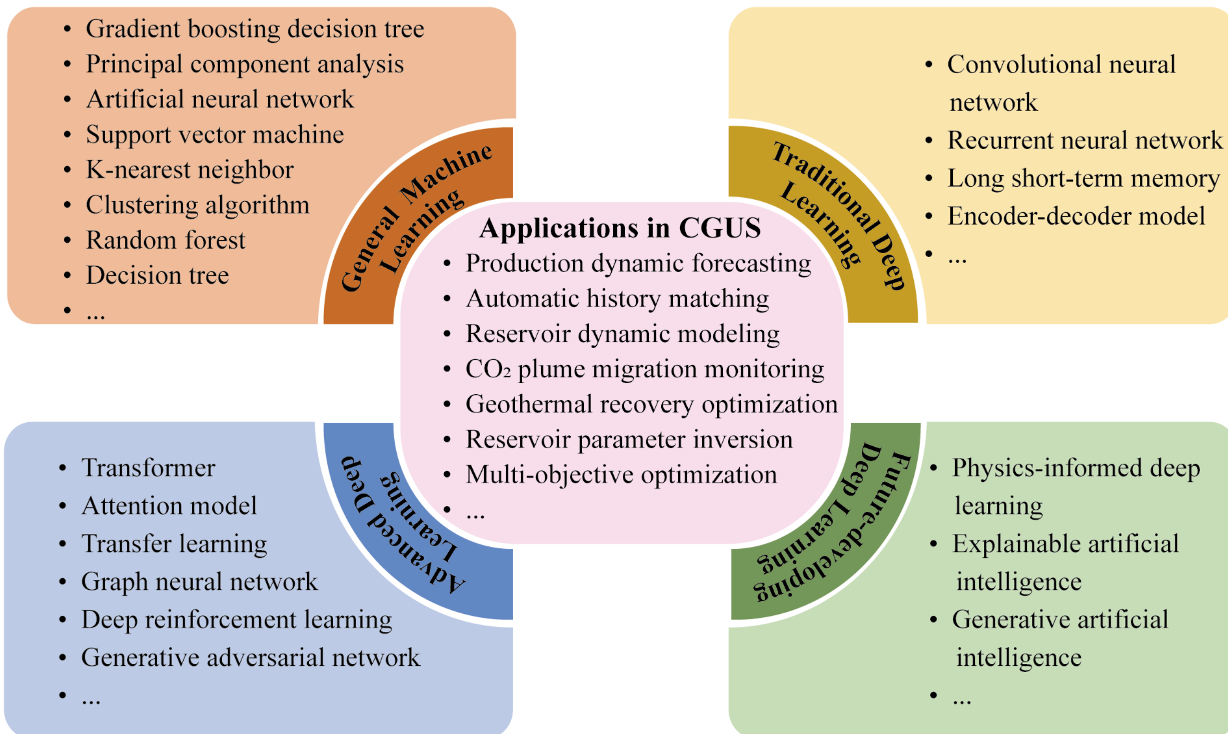

Fig. 1. Advances and perspectives for deep learning in CO_2 geological utilization and storage.

pecially in environments where the variety and amount of data are rich (Xiao et al., 2023). Moreover, the main advantages of deep learning are the ability to improve multi-scale, multi-task and multi-modal barriers with high-fidelity solutions (Wen et al., 2023). Despite their many advantages, the black-box nature of deep learning may hinder physical insights into the phenomena being examined. Evaluating and improving the interpretability and explainability of deep learning models still remains an active field of research (Longo et al., 2024).

To understand the depth and scope of deep learning applications in CGUS, this perspective proposes a conceptual framework delineating the evolutionary trajectory through four distinct stages (Fig. 1): general machine learning, traditional deep learning, advanced deep learning, and future-developing deep learning. This framework not only forecasts a progressive enhancement in the sophistication of CUGS intelligence modeling, but also phases the analysis of deep learning's burgeoning applications in this field.

2. General machine learning models

General machine learning is the initial stage and primarily focused on extracting relevant information from large datasets, using algorithms to identify patterns and trends that may be missed by manual analysis due to their complexity or the time required for human analysis. Key algorithms of general machine learning, including artificial neural network, decision tree, support vector machine, K-nearest neighbor, clustering algorithm, principal component analysis, random forest, and gradient boosting decision tree, have become an integral part of deciphering the complex relationships in reservoir-fluid data (Jerng et al., 2024). In CGUS, general machine learning models are usually used for prediction and optimization tasks in relatively simple scenarios with few data samples, including prediction of CO_2 saturation in multicomponent gas-liquid mixtures, properties of CO_2 under reservoir conditions, CO_2 sequestration potential, and variable analysis for decision-making in engineering projects (Davoodi et al., 2024). However, these general machine learning models are usually limited by human cognitive limitations, often relying on simplifying assumptions to make valid predictions, and struggling to integrate rich and diverse multi-source data.

3. Traditional deep learning models

In the traditional deep learning stage, intelligent frameworks develop the capability to perceive and interpret various multisource data from CGUS systems. This evolution marks a shift from processing only structured data to understanding natural data and recognizing complex temporal and spatial patterns. Traditional deep learning architectures, both for time series prediction and spatial structure analysis, are significantly improved in their ability to handle the type and amount of data in geological resources applications, including the following algorithms and their variants: convolutional neural network (CNN), recurrent neural network (RNN), long short-term memory (LSTM), encoder-decoder (Choudhary et al., 2022). The CNN-based models are capable of handling complex data with high-dimensional nonlinearities, making them suitable for spatial pattern mining tasks in CGUS, such as reservoir fluid distribution prediction, temperature and pressure field prediction, and CO_2 plume transport tracking.

The RNN and LSTM-based models, specialized in processing time series data, are suitable for temporal pattern mining tasks in CGUS, such as production dynamics prediction for hydrocarbon and geothermal development processes, CO_2 leakage monitoring, and automatic history matching (Xue et al., 2023). The encoder-decoder network, providing a way to train time-series models for sequence-to-sequence prediction problems, is remarkably effective in predicting reservoir production dynamics by retaining contextual information over time. Despite these advances, traditional deep learning still relies heavily on correlating inputs and outputs to accurately predict various state variables of the CGUS system through the learning of data features, but these methods do not incorporate a deeper cognitive understanding similar to human thought processes.

4. Advanced deep learning architectures

The prediction and optimization processes of CGUS intelligent models based on traditional machine learning and deep learning algorithms tend to rely on statistical likelihoods, restricting the models to operate within a limited number of statistical parameters and posing a challenge to derive clear explanations from the deep learning process. To efficiently utilize complex multi-source data in geology and engineering, optimize algorithm architecture, and improve computing efficiency, advanced deep learning algorithms have gradually emerged based on traditional deep learning, including: generative adversarial network (GAN), graph neural network (GNN), attention, transformer, transfer learning, and deep reinforcement learning (DRL) (Deng et al., 2024; Meng et al., 2024). CGUS models based on these advanced deep learning architectures are gradually beginning to emulate human cognitive and logical thinking, working to understand and explain complex patterns and relationships in geo-energy systems.

The GAN-based model can refine and adjust the reservoir geological model based on new observations to generate new structures not previously seen in the training samples, enabling more accurate prediction of CO_2 plume migration and evolutionary patterns of reservoir properties without being limited to existing observations. The GNN-based model can aggregate the features of each node with those of the surrounding nodes to obtain more comprehensive structural information, such that it is ideally suited for the prediction of percolation processes in heterogeneous reservoirs based on formation structure and fluid distribution information.

In the domain of geo-energy development forecasting, sequence-based modelling plays a crucial role, among which the attention model is a recent development of deep learning architecture (Mnih et al., 2015). Geo-energy models incorporating attention mechanisms have become an important focus of sequence-to-sequence time series prediction, particularly when simulating long-term production and storage sequences. In addition, the transformer incorporating encoder-decoder and self-attention mechanisms is another emerging deep learning architecture for time-series prediction, allowing for a comprehensive representation of input data from different locations. Therefore, the transformer-based model is conducive to addressing the multi-scale and multi-variable nature of CO_2 sequestration and utilization processes, indicating the potential for exponential growth in this field.

Since labeled sample data are often scarce and difficult to obtain during reservoir development, transfer learning provides an innovative solution to the challenge of training deep learning models on limited datasets and reducing overall training time. Considering that most of the data or tasks in the field are correlated, transfer learning can migrate the parameters of the learned and trained model to the new model to speed up and optimize the learning efficiency (Weiss et al., 2016). In addition, in the case of scarce training data during reservoir parameter inversion, transfer learning can utilize the knowledge gained from deep learning models trained on rich datasets for similar reservoirs to deduce reservoir properties accurately.

DRL is an emerging field of dynamic computing with the attractive nature of being able to find low-dimensional features that accurately represent high-dimensional practical engineering problems and experience-driven autonomous learning (Zhou and Wang, 2022). Tasks such as production system optimization, storage process optimization, and automatic history matching in CGUS require algorithms to make decisions and perform actions at every moment. DRL shows great potential in solving the above bottlenecks by learning directly from the multi-dimensional continuous interaction information between the agent and the environment. Moreover, DRL-based models can mimic human cognition and autonomous learning capabilities, and are dedicated to real-time optimization of injection and production scheme as well as dynamic inversion of reservoir parameters.

5. Future-developing deep learning techniques

The numerical modeling of CGUS is quite complex, considering multiple scales in space and time, multiple physicochemical interaction mechanisms, and dynamic heterogeneity of the formation. While data-driven deep learning models can capture complex and strongly nonlinear CO_2 migration and transformation processes through sophisticated neural networks and patterns, their black-box characterization fails to describe the fundamental laws of physics, which may lead to physically implausible and spurious predictions. By combining physical laws (Navier-Stokes equations, multiphase fluid seepage equations, water-rock reactions) with deep neural networks, Physics-informed deep learning (PIDL) can learn more generalizable models with fewer data samples and can accelerate the training rate (Wang et al., 2024). PIDL is particularly well suited to scientific computational and engineering problems that can be described by explicit physical laws (Qin et al., 2024). In summary, PIDL incorporates data- and physics-driven synergies that offer promising directions for advancing the science of CO_2 utilization and storage modeling.

Despite the success of deep learning in the CGUS field, these models are sometimes criticized as “black boxes” because they do not provide insights to understand how they make predictions and optimizations. CGUS intelligent models based on deep learning involve dynamic optimization of production and storage processes, real-time prediction of high-dimensional field data, reservoir parameter inversion, evaluation of carbon storage safety, multi-objective optimization and so on. The complex network architecture and over-parameterization of the intelligent modeling process hinder the interpretability of the models. In the future, it will be crucial to explain the behavior of deep models in decision making and prediction in CGUS to explain the outputs of the algorithms employed, called explainable artificial intelligence (Saranya and Subhashini, 2023). Successful explanations of deep learning architectures can help us gain domain insights and expand our knowledge about unknown mechanisms, causality, and connections in the CGUS field. Therefore, explainable artificial intelligence, as an important future direction for deep learning in CGUS, motivates us to move beyond deep learning as a knowledge discovery tool rather than a data-fitting model.

In addition, generative artificial intelligence (GenAI), a novel architecture of artificial intelligence, has become an essential research field that realizes the advancement of artificial intelligence from perceiving and understanding the world to generating and creating the world (Gupta et al., 2024). The ability of GenAI models to generate unique data by learning complex patterns in datasets (such as ChatGPT), makes them particularly useful in handling the main challenges in geological and petroleum engineering, which often include handling complexity, nonlinearity, and uncertainty. The large language model for geoscience knowledge understanding and utilization can answer geoscientific questions and follow the instructions of geoscientists with geoscientific professionalism (Deng et al., 2023). As the field of CGUS intelligence continues to evolve, embracing the innovative solutions offered by GenAI will be critical in tackling complex global challenges, thereby paving the way for a more sustainable, efficient, and intelligent CGUS systems.

6. Conclusions

Summarizing the four evolutionary stages in the development of the CGUS intelligent models, it becomes evident that various advanced algorithmic architectures for solving complex geological and engineering problems have been developed based on traditional deep learning architectures. However, current research mainly focuses on the application innovation of methodology, and a comprehensive breakthrough framework has yet to be established as it is still in the preliminary stage.

Several key issues must be urgently addressed to advance the frontiers of intelligent CGUS theory and methodology. First but not least, the datasets used in CGUS are collected from various sources, and the spatial and temporal resolution usually needs to be consistent. High-dimensional data with multi-source, multi-scale, and multi-mechanism make the learning task more difficult. Deep learning frameworks based on the fusion of physical constraints and data-driven have the potential to advance the research of CGUS in limited data by capturing the interaction between high-dimensional learning and different physical processes in complex environments. Then, current research has continuously strived to explain the behavior of deep models in decision-making/prediction to address the limitations and biases of data samples and the black-box nature. Thus, explainable artificial intelligence is considered a promising direction for further development in deep learning. Finally, when dealing with time series and spatio-temporal data and performing complex simulations, optimizations, and decisions, large language models based on GenAI constructed using more than 100 billion parameters show extraordinary potential. Overall, the future development trend of PIDL, explainable artificial intelligence, and GenAI models in CGUS is promising, and their applications are expected to provide more accurate solutions for information extraction and decision support for the integration of CGUS.

Acknowledgements

This work was funded by the National Natural Science Foundation of China (No. 42202292 and 52304054), and Fundamental Research Funds for the Central Universities (No. 2652023052).

Conflict of interest

The authors declare no competing interest.

References

Al-Sakkari, E. G., Ragab, A., Dagdougui, H., et al. Carbon capture, utilization and sequestration systems design and operation optimization: Assessment and perspectives of artificial intelligence opportunities. Science of The Total Environment, 2024, 917: 170085.

Choudhary, K., DeCost, B., Chen, C., et al. Recent advances and applications of deep learning methods in materials science. npj Computational Materials, 2022, 8(1): 59.

Davoodi, S., Vo Thanh, H., Wood, D. A., et al. Machine learning insights to CO_2-EOR and storage simulations through a five-spot pattern-a theoretical study. Expert Systems with Applications, 2024, 250: 123944.

Deng, C., Zhang, T., He, Z., et al. K2: A foundation language model for geoscience knowledge understanding and utilization. Proceedings of the 17th ACM International Conference on Web Search and Data Mining, Merida, Mexico, 4-8 March, 2024.

Deng, Y., Kang, X., Ma, H., et al. Characterization of discrete fracture networks with deep-learning based hydrogeophysical inversion. Journal of Hydrology, 2024, 631: 130819.

Gupta, R., Nair, K., Mishra, M., et al. Adoption and impacts of generative artificial intelligence: Theoretical underpinnings and research agenda. International Journal of Information Management Data Insights, 2024, 4(1): 100232.

Jerng, S. E., Park, Y. J., Li, J. Machine learning for CO_2 capture and conversion: A review. Energy and AI, 2024, 16: 100361.

Liu, P., Zhang, K., Yao, J. Reservoir automatic history matching: Methods, challenges, and future directions. Advances in Geo-Energy Research, 2023, 7(2): 136-140.

Longo, L., Brcic, M., Cabitza, F., et al. Explainable Artificial

Intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions. Information Fusion, 2024, 106: 102301.

Meng, S., Fu, Q., Tao, J., et al. Predicting CO_2-EOR and storage in low-permeability reservoirs with deep learning-based surrogate flow models. Geoenergy Science and Engineering, 2024, 233: 212467.

Mnih, V., Kavukcuoglu, K., Silver, D., et al. Human-level control through deep reinforcement learning. Nature, 2015, 518: 529-533.

Noyce, G. L., Smith, A. J., Kirwan, M. L., et al. Oxygen priming induced by elevated CO_2 reduces carbon accumulation and methane emissions in coastal wetlands. Nature Geoscience, 2023, 16(1): 63-68.

Qi, S., Zheng, B., Wang, Z., et al. Geological evaluation for the carbon dioxide geological utilization and storage (CGUS) site: A review. Science China Earth Sciences, 2023, 66(9): 1917-1936.

Qin, Z., Jiang, A., Faulder, D., et al. Physics-guided deep learning for prediction of energy production from geothermal reservoirs. Geothermics, 2024, 116: 102824.

Saranya, A., Subhashini, R. A systematic review of Explainable Artificial Intelligence models and applications: Recent developments and future trends. Decision Analytics Journal, 2023, 7: 100230.

Wang, Y., Dai, Z., Chen, L., et al. An integrated multi-scale model for CO_2 transport and storage in shale reservoirs. Applied Energy, 2023, 331: 120444.

Wang, Y., Dai, Z., Wang, G., et al. A hybrid physics-informed data-driven neural network for CO_2 storage in depleted shale reservoirs. Petroleum Science, 2024, 21(1): 286-301.

Weiss, K., Khoshgoftaar, T. M., Wang, D. A survey of transfer learning. Journal of Big Data, 2016, 3(1): 9.

Wen, G., Li, Z., Long, Q., et al. Real-time high-resolution CO_2 geological storage prediction using nested Fourier neural operators. Energy & Environmental Science, 2023, 16(4): 1732-1741.

Xiao, C., Zhang, S., Ma, X., et al. Deep-learning-generalized data-space inversion and uncertainty quantification framework for accelerating geological CO_2 plume migration monitoring. Geoenergy Science and Engineering, 2023, 224: 211627.

Xu, T., Tian, H., Zhu, H., et al. China actively promotes CO_2 capture, utilization and storage research to achieve carbon peak and carbon neutrality. Advances in Geo-Energy Research, 2022, 6(1): 1-3.

Xue, L., Li, D., Dou, H. Artificial intelligence methods for oil and gas reservoir development: Current progresses and perspectives. Advances in Geo-Energy Research, 2023, 10(1): 65-70.

Zhou, Y., Wang, Y. An integrated framework based on deep learning algorithm for optimizing thermochemical production in heavy oil reservoirs. Energy, 2022, 253: 124140.