# A Genetics-Inspired Deep Learning Neural Network Approach to Predicting VAM Emission

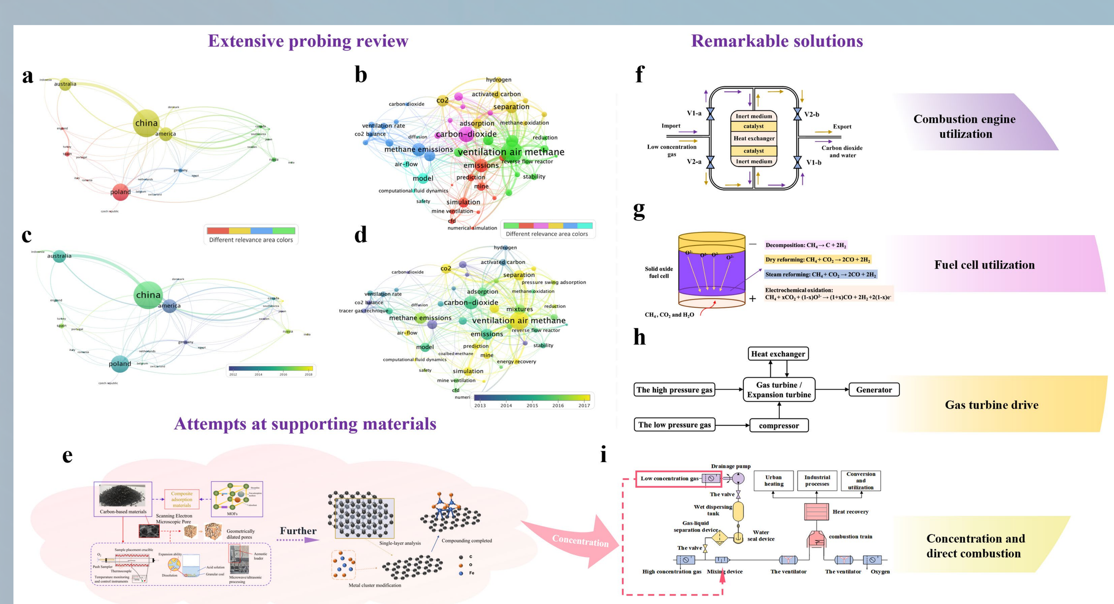
Shenguang (Steve) Fu, Shimin Liu

G3 Center and Energy Institute, The Pennsylvania State University, University Park, PA

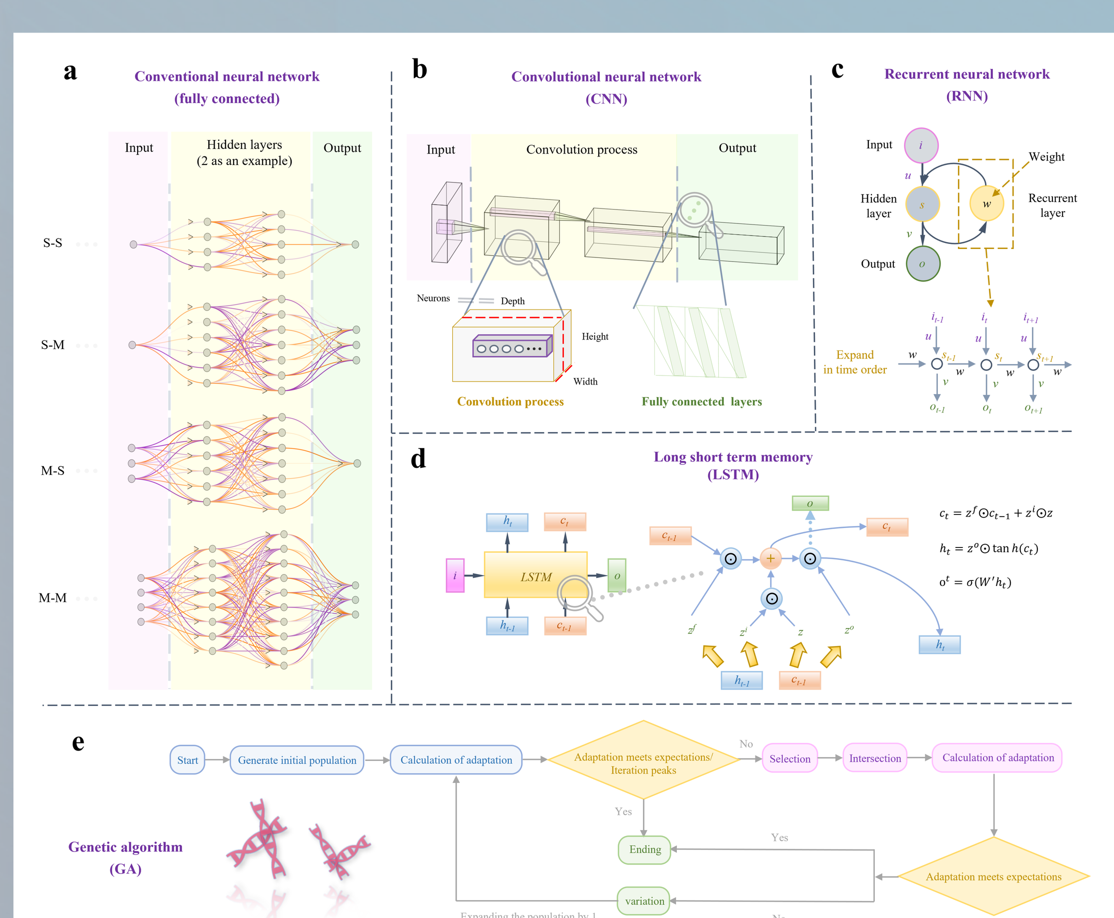
## ABSTRACT

Ventilation air methane (VAM) contributes more than 80% of the total direct emissions of coal mine methane (CMM) in China, with a single mine shaft emitting VAM at a rate of 222 m³/s (or 800,000 m³/h) potentially generating around 1 million tons of $CO_2$-equivalent ($CO_2e$) annually. It is pivotal to quantify the VAM and strategically mitigate the emission towards carbon peak and neutrality. In this study, the Long-Short-Term-Memory (LSTM) recurrent neural network is optimized based on the inspiration generated by the chromosomal inheritance pattern of human genes. Multiple factors with a strong correlation of $CH_4$ emissions are predicted, and future $CH_4$ emissions are predicted using time series and multi-factor series. Combined with the self-determined pattern of VAM emissions, a correlation model is proposed for VAM quantification. Based on the predicted VAM emission, recommendations are summarized for effective VAM mitigation. This work will lay the foundation for the future fugitive gas emissions from coal mining sector for the coal-producing countries.

## LITERATURE REVIEW

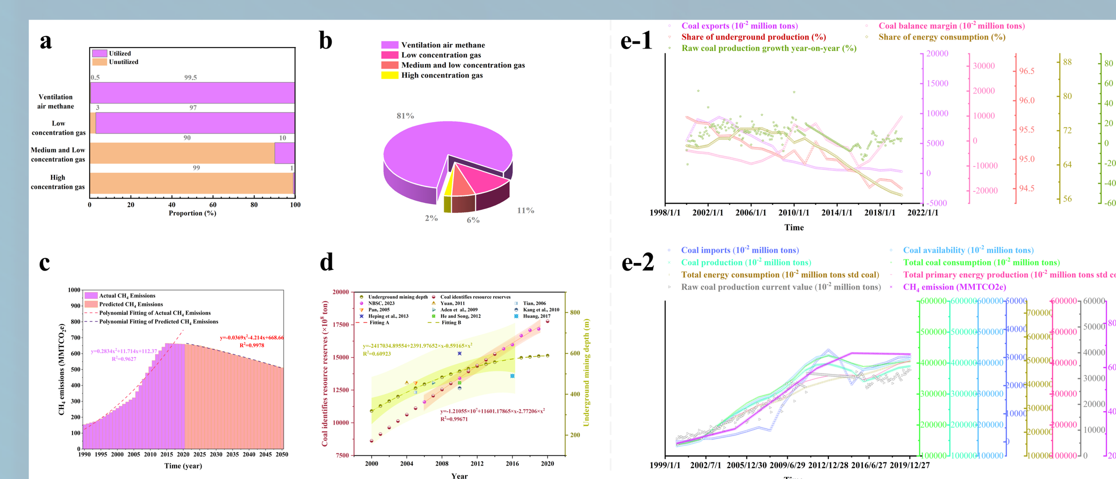


China, Poland, the United States, Australia, Germany, and Canada have been at the forefront of research in this area, Obviously, in China, the status of VAM-related research is gradually gaining prominence. Specifically, those that have been spotlighted include numerical simulation, emissions control, separated components, and pressure swing adsorption [1-3]. It is obvious that we can make related predictions to ensure the strategies are going well.

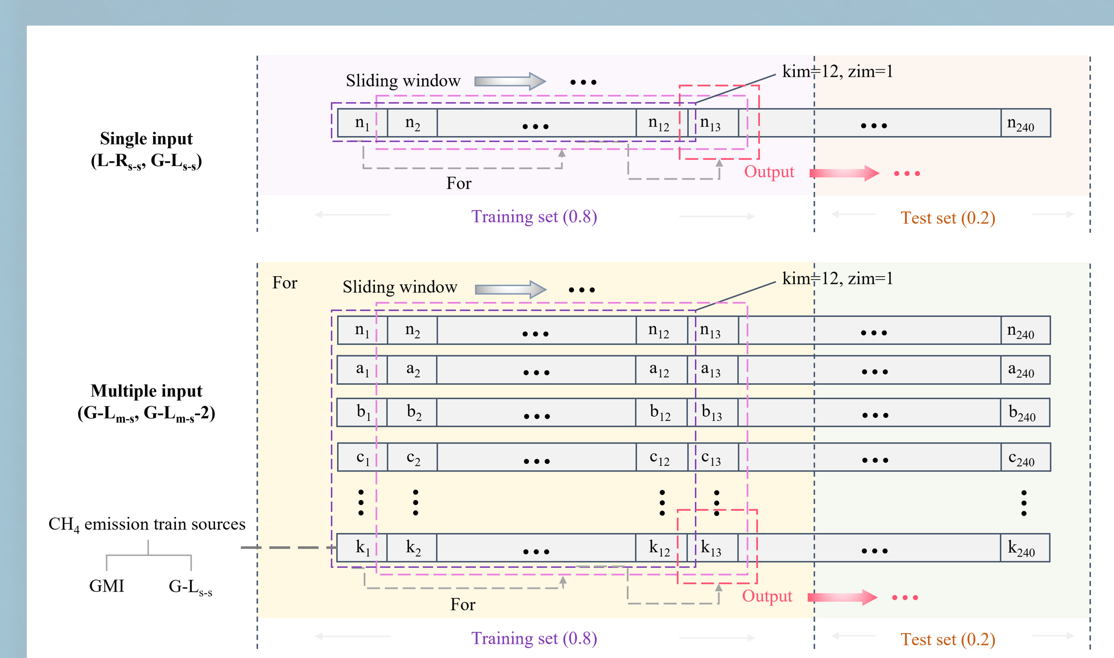


Machine learning recently got a high focus in engineering industries, to utilize such technologies effectively, choosing the best and the most matched model is crucial [4-7]. We put the most general model - traditional fully connected neural networks, convolutional neural networks (CNN) [8, 9], and recurrent neural networks (RNN) [10, 11] here (a, b, and c). This study used the LSTM model, which will be optimized by the Genetic algorithm [12], to point out the prediction accuracy. All the model logic has been shown here.

## DATA & METHODS



As the figure shows, these are the basic data syntheses, which includes various concentrations of gas content (a), utilization rates (b), coal mine methane emissions and predictive analysis (c), a summary of provable coal reserves and underground mining depths (d), and datasets with multiple correlations (e-1 & e-2) [13-17]



This study delves into three categories: the LSTM-RNN, the GA-LSTM (s-s), and the GA-LSTM (m-s). Where "s-s" and "m-s" represent single-input single-output and multiple-input single-output respectively. The sliding window size (denoted as "kim") is 12, with a progressive advancement of 1 unit. It means that a single variable (month data) is input 12 at a time and output 1 at a time. Conversely, the multi-input sliding window covers all distinct monthly variable data from n-k, with the output variable being methane emissions. Specifically, the nomenclature used is as follows: $L\text{-}R_{s\text{-}s}$ denotes single-input single-output for LSTM-RNN, $G\text{-}L_{s\text{-}s}$ denotes single-input single-output for LSTM, $G\text{-}L_{m\text{-}s}$ denotes multi-input single-output based on GMI, and $G\text{-}L_{m\text{-}s}\text{-}2$ denotes multi-input single-output based on the output of $G\text{-}L_{s\text{-}s}$.

$$A=\left(a_{2020.01}\quad a_{2020.02}\quad \cdots \quad a_{2030.01}\right)^T \tag{1}$$

$$B=\left(b_{2020.01}\quad b_{2020.02}\quad \cdots \quad b_{2030.01}\right)^T \tag{2}$$

$$C=\left(c_{2020.01}\quad c_{2020.02}\quad \cdots \quad c_{2030.01}\right)^T \tag{3}$$

$$D=\left(d_{2020.01}\quad d_{2020.02}\quad \cdots \quad d_{2030.01}\right)^T \tag{4}$$

$$X=(A\quad B\quad C\quad D)=\begin{pmatrix} a_{2020.01} & b_{2020.01} & c_{2020.01} & d_{2020.01} \\ a_{2020.02} & b_{2020.02} & c_{2020.02} & d_{2020.02} \\ \vdots & \vdots & \vdots & \vdots \\ a_{2030.01} & b_{2030.01} & c_{2030.01} & d_{2030.01} \end{pmatrix} \tag{5}$$

$$\eta=\left(\eta_1\quad \eta_2\quad \eta_3\quad \eta_4\right)^T \tag{6}$$

$$\zeta=\left(\zeta_1\quad \zeta_2\quad \zeta_3\quad \zeta_4\right)^T \tag{7}$$

$$y_t=\frac{\eta_t\cdot\zeta_t}{\sum_{t=1}^{4}\eta_t\cdot\zeta_t}\times 100\% \tag{8}$$

$$Y=\left(y_1\quad y_2\quad y_3\quad y_4\right)^T \tag{9}$$

$$A'=\eta_1\cdot A \tag{10}$$

$$B'=\eta_2\cdot B \tag{11}$$

$$C'=\eta_3\cdot C \tag{12}$$

$$D'=\eta_4\cdot D \tag{13}$$

$$A':B':C':D'=81:11:6:2 \tag{14}$$

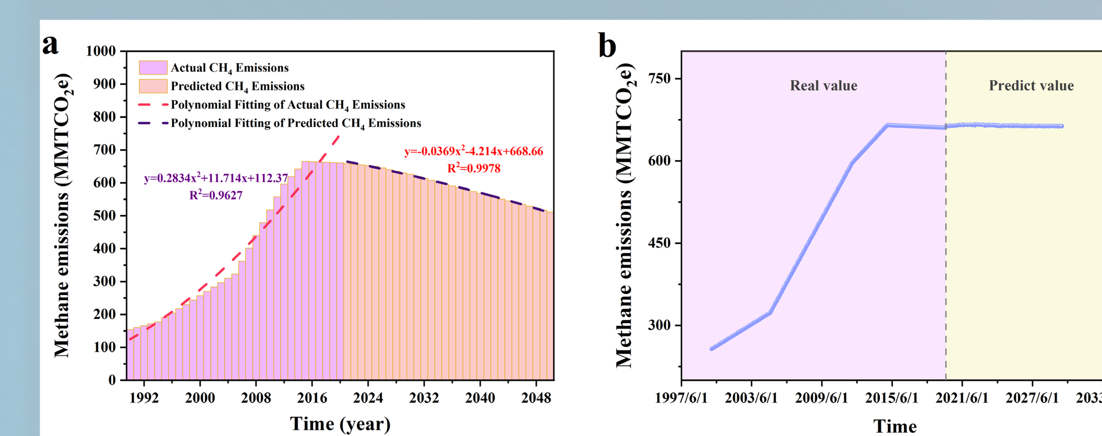
$$N=\left(N_{2020.01}\quad N_{2020.02}\quad \cdots \quad N_{2030.01}\right)^T=X\cdot Y \tag{15}$$

$$\begin{pmatrix} N_{2020.01} \\ N_{2020.02} \\ \vdots \\ N_{2030.01} \end{pmatrix}=\begin{pmatrix} a_{2020.01} & b_{2020.01} & c_{2020.01} & d_{2020.01} \\ a_{2020.02} & b_{2020.02} & c_{2020.02} & d_{2020.02} \\ \vdots & \vdots & \vdots & \vdots \\ a_{2030.01} & b_{2030.01} & c_{2030.01} & d_{2030.01} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \tag{16}$$

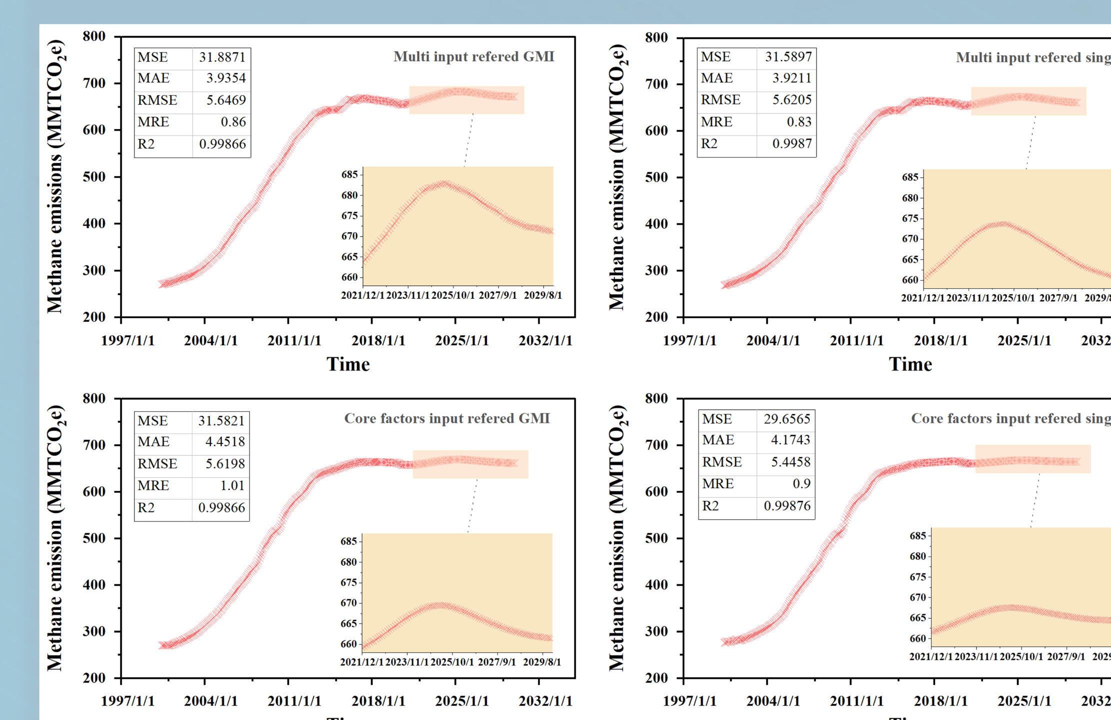
$$D'=\eta_4\cdot d \tag{17}$$

Here, we show the mathematical expression for calculating the predicted VAM emission based on the data results (VAM emission and utilization ratio). Where $N_t$ is the annual methane emission prediction, $MMTCO_2e$. *A*, *B*, *C*, and *D* are the output of high concentration, medium, and low concentration, and spent gas respectively, $MMTCO_2e$. $Y$ is the ratio of methane emitted by various concentrations to the total methane emitted by gas, %. $\eta$ is the unutilized rate, % ($\eta_1$, $\eta_2$, $\eta_3$ and $\eta_4$ are 0.5%, 3%, 90% and 99%, respectively). $\zeta$ is the average coverage of the gas concentration range of each concentration, %.

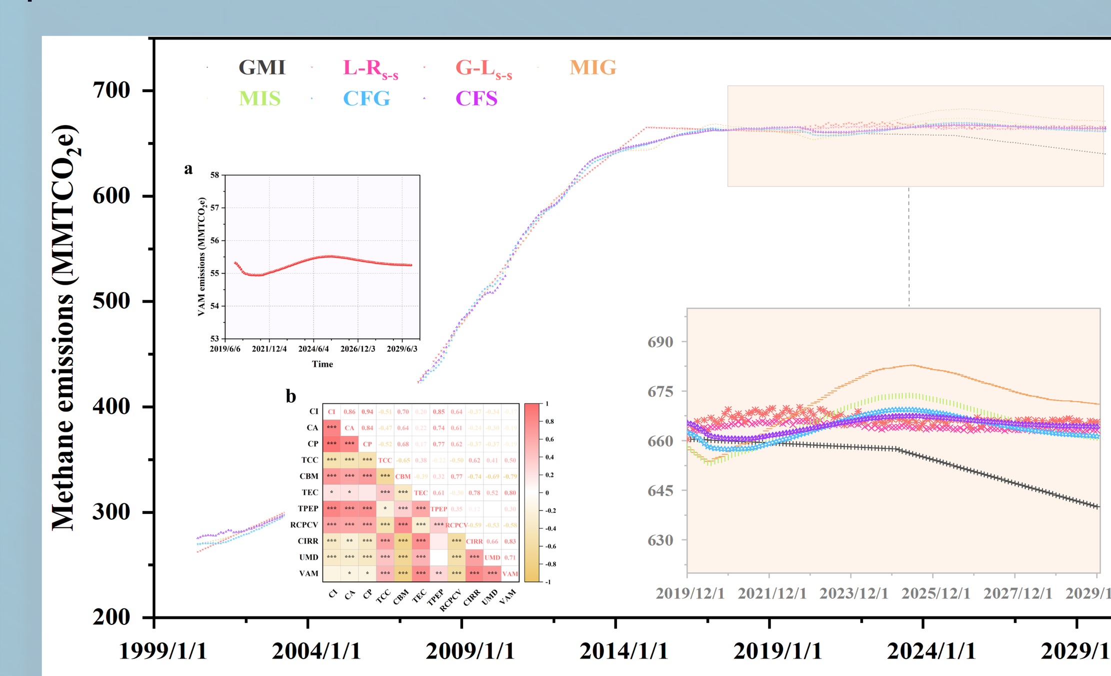
## RESULTS & DISCUSSIONS



The overall trend of $CH_4$ emission in the coal mine based on GMI shows a consistent rise in methane emissions over the past 30 years, coupled with a projected decline in emissions over the subsequent 30 years (a). Using the constructed neural network of LSTM-RNN to predict $CH_4$ emission in the coal mine over the next ten years (2000/01 to 2020/01), the results appear relatively static (hovering consistently around 665 $MMTCO_2e$). The difference between GMI-based and $L\text{-}R_{s\text{-}s}$-based prediction is considerable, indicating that the performance of the constructed LSTM-RNN neural network needs to be optimized, which is significant for refining the predictive accuracy and reliability of the model.



To elucidate the influence of various input indicators and computational data on the model results, the results are compared both horizontally and vertically. They are denoted as Multi-input referred GMI, Multi-input referred single ("single" means $G\text{-}L_{s\text{-}s}$ result), Core factors input referred GMI, and Core factors input referred single. $G\text{-}L_{m\text{-}s}$ results based on different input data are shown in the above figure. Upon closer examination through local zooming, it can be noticed that the results obtained from the input effects of the core factors exhibit superior performance



The big figure shows comparison of total results with local zoom analysis, and the small figures show 10 years VAM predictions (a) and factor correlations (b). Upon scrutinizing the correlation heatmap, CIRR dominates with a correlation of up to 0.83, followed by TEC at up to 0.80. TCC and TPEP are positively correlated with both VAM, but the correlations are only 0.50 and 0.30. Conversely, the remaining factors were all negatively correlated, with CBM as low as -0.79.

## SUMMARY

a) For VAM mitigation and utilization, its future emissions are predicted based on a genetic algorithm-optimized LSTM neural network.

b) We compared the prediction accuracy of VAM emissions for various input indicators and computational data, and the results showed that the results obtained from the input effects of the core factors exhibit superior performance.

c) VAM is strongly related to CIRR and TEC, so we need to control them to effectively manage VAM emissions.

## REFERENCES

[1] K. Karimi, S. Fatemi, Methane capture and nitrogen purification from a nitrogen rich reservoir by pressure swing adsorption; experimental and simulation study, Journal of Environmental Chemical Engineering, 9 (2021) 106210.

[2] H. Xu, Development of a new-type ventilation air methane collection device, Zhongzhou Coal, 40 (2018) 12-16.

[3] S. Ouyang, Separation and Enrichment of Coal Mine Methane by PSA, in, Dalian University of Technology, 2014.

[4] L. Alzubaidi, J. Zhang, A.J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M.A. Fadhel, M. Al-Amidie, L. Farhan, Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions, Journal of big Data, 8 (2021) 1-74.

[5] R. Baraniuk, D. Donoho, M. Gavish, The science of deep learning, Proceedings of the National Academy of Sciences, 117 (2020) 30029-30032.

[6] D. Chen, Z. Wang, D. Guo, V. Orekhov, X. Qu, Review and prospect: deep learning in nuclear magnetic resonance spectroscopy, Chemistry–A European Journal, 26 (2020) 10391-10401.

[7] M. Huang, J. Ninić, Q. Zhang, BIM, machine learning and computer vision techniques in underground construction: Current status and future perspectives, Tunnelling and Underground Space Technology, 108 (2021) 103677.

[8] Z. Geng, Y. Zhang, C. Li, Y. Han, Y. Cui, B. Yu, Energy optimization and prediction modeling of petrochemical industries: An improved convolutional neural network based on cross-feature, Energy, 194 (2020) 116851.

[9] D.-X. Zhou, Universality of deep convolutional neural networks, Applied and computational harmonic analysis, 48 (2020) 787-794.

[10] L. Medsker, L.C. Jain, Recurrent neural networks: design and applications, CRC press, 1999.

[11] M. Schuster, K.K. Paliwal, Bidirectional recurrent neural networks, IEEE transactions on Signal Processing, 45 (1997) 2673-2681.

[12] M. Mitchell, An introduction to genetic algorithms, MIT press, 1998.

[13] Y. Lin, Z. Duan, X. Wang, D. Ma, Analysis on Prospects of Exploration and Development of Coalbed Methane Resources in Shaanxi Province, Coal, 25 (2016) 50-53.

[14] F. Yang, H. Xu, Analysis on the development path of ecological environment protection and resources comprehensive utilization in coal industry during the 14th Five-Year Plan period, China Coal, (2021).

[15] N. Kholod, M. Evans, R.C. Pilcher, V. Roshchanka, F. Ruiz, M. Coté, R. Collings, Global methane emissions from coal mining to continue growing even with declining coal production, Journal of Cleaner Production, 256 (2020) 120489.

[16] L. Zhao, The wind has opened for methane emission reduction and there is a huge opportunity for ultra-low concentration gas utilization, in, China Atmosphere, 2021.

[17] J. Chen, Global Warming Potential calculation and its evolution, Shanghai Shipbuilding, (2014) 27-31.

## ACKNOWLEDGMENTS

Thanks to the authors and relevant official websites for data support.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

PennState College of Earth and Mineral Sciences

John and Willie Leone Family Department of Energy and Mineral Engineering