ARGA-3DCNN-Based Three-Dimensional Neutron Flux Prediction Method for Lead-Cooled Fast Reactors

Authors: Mo Ziwen, Yang Zihui, Li Zhongyang, Sun Guomin, Li Zhaodong

Date: 2025-04-24T00:00:00+00:00

Abstract

Three-dimensional prediction of neutron flux is crucial for the design, optimization, and safety analysis of reactor cores; however, due to the compact space of micro-scale lead-cooled fast reactors and the difficulty in detector placement, existing methods have mostly focused on the two-dimensional level, with less attention paid to three-dimensional flux prediction. This paper proposes a three-dimensional convolutional neural network model that integrates Residual Network (ResNet) and Multi-head Self-Attention mechanism (MSA), termed Genetic Algorithm-Enhanced 3D CNN with Multi-Head Self-Attention and Residual Connections (ARGA-3D CNN). This model can effectively capture the spatial distribution characteristics of core neutron flux and address spatial dependency issues. ResNet is employed to mitigate gradient vanishing and explosion, thereby enhancing training stability, while MSA is utilized to strengthen the identification of key regions. Furthermore, a genetic algorithm is adopted to optimize hyperparameters, further improving the prediction accuracy of core neutron flux. Experiments construct a dataset based on calculation results from the Monte Carlo particle transport simulation software SuperMC, and use this dataset to train and optimize the ARGA-3D CNN model for prediction. Results show that, compared with SuperMC calculation results, the model achieves Mean Absolute Error (MAE), Mean Squared Error (MSE), and coefficient of determination (R^2) metrics of 3.19$\times 10-6, 2.14\times$10-11, and 0.9735, respectively. Computational efficiency is significantly improved, with a single prediction taking only seconds. Compared with traditional network models such as Convolutional Neural Network (CNN), Artificial Neural Network (ANN), and Long Short-Term Memory (LSTM), the prediction performance is superior. The ARGA-3D CNN model demonstrates high accuracy and computational efficiency in three-dimensional neutron flux prediction, providing a new method for rapid prediction of nuclear reactor core parameters, which holds certain practical value and significance.

Full Text

Preamble

Three-Dimensional Neutron Flux Prediction for Lead-Cooled Fast Reactors Based on ARGA-3DCNN

Ziwen Mo[1],[2], Zihui Yang[1], Zhongyang Li[1], Guomin Sun[1], Zhaodong Li[3]

[1](Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China)
[2](School of Physics and Electronic Information, Huaibei Normal University, Huaibei 235000, China)
[3](School of Artificial Intelligence and Big Data, Hefei University, Hefei 230027, China)

Abstract

[**Background**] Three-dimensional prediction of neutron flux is crucial for reactor core design, optimization, and safety analysis. However, due to the compact spatial configuration of small lead-cooled fast reactors and the difficulty in detector placement, existing methods primarily focus on two-dimensional analysis, with limited attention to three-dimensional flux prediction. [**Purpose**] This paper proposes a Genetic Algorithm-Enhanced 3D CNN with Multi-Head Self-Attention and Residual Connections (ARGA-3D CNN) model that integrates Residual Networks (ResNet) and Multi-Head Self-Attention (MSA) mechanisms. This model effectively captures the spatial distribution characteristics of neutron flux in reactor cores and addresses spatial dependency issues. [**Methods**] The ARGA-3D CNN model employs 3D convolutional neural networks to capture spatial distributions of neutron flux in the reactor core. ResNet alleviates vanishing and exploding gradient problems, enhancing training stability, while MSA strengthens identification of critical regions. Additionally, a genetic algorithm optimizes hyperparameters to further improve prediction accuracy. The model is trained and optimized using a dataset constructed from Monte Carlo particle transport simulation results obtained with SuperMC software. [**Results**] Compared with SuperMC calculations, the model achieves Mean Absolute Error (MAE), Mean Squared Error (MSE), and R^2 values of 3.19$×10^{{-6}, 2.14×10}${-11}$, and 0.9735, respectively. Computational efficiency is significantly improved, with single predictions requiring only seconds. Compared to traditional models such as Convolutional Neural Networks (CNN), Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM) networks, the proposed method demonstrates superior prediction performance. **Conclusions** The ARGA-3D CNN model provides a high-precision, computationally efficient method for three-dimensional neutron flux prediction in reactor cores. It offers a novel approach for rapid prediction of nuclear reactor core parameters

with significant practical value and implications for reactor design, operation, and safety analysis.

Keywords: Lead-cooled fast reactor, Neutron flux, 3D convolutional neural network, Multi-head self-attention, Residual network, Genetic algorithm

Introduction

Lead-cooled fast reactors have attracted considerable attention in recent years as an advanced reactor type with unique advantages in nuclear energy technology development. These reactors exhibit excellent neutron physics characteristics, high fuel utilization efficiency, and favorable safety performance, demonstrating broad application prospects in the nuclear energy field. However, in core design, operational optimization, and safety analysis, precise prediction of core neutron flux distribution is particularly critical as it serves as an essential parameter for evaluating core performance. As the primary energy carrier in nuclear reactions, neutron flux distribution directly affects reactor power output and fuel burnup, and is closely related to key factors such as reactivity control, thermal-hydraulic characteristics, and irradiation behavior.

Traditional neutron flux prediction methods, such as numerical calculations based on deterministic transport theory and empirical models, can provide effective results to a certain extent but often suffer from high computational costs, slow calculation speeds, and insufficient adaptability to complex operating conditions when facing intricate core structures and multi-physics coupling problems. In recent years, artificial intelligence technologies, particularly deep learning algorithms, have been widely applied to nuclear reactor physics parameter prediction due to their advantages in processing large-scale data and efficient computation.

For instance, Wang et al. proposed a core parameter prediction method based on adaptive backpropagation neural networks (BPNN) that effectively improved prediction accuracy while reducing computational resource consumption. Palmi et al. developed a method based on artificial neural networks (ANN) for predicting core loading reactivity and fuel cycle length, accurately forecasting reactivity changes and fuel cycle duration under different core loading configurations. Additionally, Bei et al. developed a three-dimensional surrogate model based on BPNN that significantly improved computational efficiency, particularly demonstrating advantages in neutron flux prediction for molten salt reactors. Chen et al. further enhanced prediction accuracy for neutron flux and keff by optimizing the Extreme Learning Machine (ELM) model combined with whale optimization algorithms, achieving remarkable improvements in training efficiency.

Despite these significant advances in reactor neutron flux and key parameter prediction, existing methods often face challenges in balancing computational accuracy and efficiency when dealing with spatially compact and structurally complex cores, particularly in handling three-dimensional neutron flux distributions. To improve prediction accuracy for three-dimensional core neutron flux distribution and address limitations of existing methods when processing complex spatial structures and nonlinear relationships, this paper proposes an ARGA-3D CNN model that combines residual networks and self-attention mechanisms with three-dimensional convolutional neural networks. This model effectively avoids gradient vanishing problems through residual structures while utilizing self-attention mechanisms to automatically identify and focus on key features in input data, thereby enhancing prediction accuracy for three-dimensional core neutron flux distribution. Genetic algorithm optimization of network hyperparameters significantly improves model performance. Experimental results demonstrate that the ARGA-3D CNN model can efficiently process three-dimensional complex data, showing remarkable advantages particularly in three-dimensional neutron flux prediction.

1.1 Three-Dimensional Convolution Module

Three-Dimensional Convolutional Neural Networks (3D CNN) represent an extension of Convolutional Neural Networks (CNN) for three-dimensional data processing. By introducing three-dimensional convolution kernels, 3D CNN can simultaneously extract spatial and temporal features, overcoming the limitation of 2D CNN that can only capture two-dimensional spatial characteristics. As shown in Figure 1: see original paper, when 2D CNN is applied to a single image, it slides only in spatial dimensions, producing a two-dimensional output. In contrast, Figure 1: see original paper shows that 3D CNN slides in both spatial and temporal dimensions, preserving temporal information in a three-dimensional volume output. This 3D convolution operation is particularly suitable for tasks such as video analysis, action recognition, and medical image processing, demonstrating significant advantages over 2D convolution.

In three-dimensional convolution, the depth of the convolution kernel is typically smaller than the depth of the input data (i.e., kernel size smaller than channel size), as shown in [Figure 2: see original paper], allowing the filter to slide in three directions: height, width, and channels. Each time the convolution kernel moves to a new position, element-wise multiplication and addition operations are performed to produce an output value. Since the filter slides throughout the three-dimensional space, the output data is also arranged in a three-dimensional structure. The three-dimensional convolution formula is as follows:

$$Y_t(i) = \sum \sum \sum j$$

where $Y_t(i)$ represents the output feature map value at time step t, position i (where height, width, and depth dimensions are considered). The input data at time step $t - d+$ and position (d, m, n) is convolved with weights, and b

represents the bias term. The convolution kernel size is $D_k \times k \times k$ (temporal depth, height, width).

1.2 Multi-Head Self-Attention Module

Multi-Head Self-Attention (MSA) enhances model capability in processing different information by computing multiple attention heads in parallel. Each attention head calculates its own attention scores through linear transformations of input queries (Q), keys (K), and values (V), then performs weighted summation of the results to capture different subspace features of the input. Multiple attention heads working in parallel enable the model to focus on information across different subspaces simultaneously, thereby improving expressive capability and the ability to capture long-range dependencies. [Figure 3: see original paper] illustrates the implementation process of the multi-head attention mechanism: each head extracts features from input Q, K, and V through linear transformations, then concatenates (Concat) the outputs from each head, and finally obtains the output through a linear transformation.

The specific calculation process is as follows: For each attention head, attention scores are first computed:

$$\text{Attention}(K, V, Q) = \text{Softmax}(\text{score}(Q, K, V))$$

where $\text{score}(Q, K)$ is typically calculated through dot product. Then, outputs from multiple heads are combined through concatenation and finally transformed linearly to obtain the result:

$$\text{MultiHeadSelfAttention}(K, V, Q) = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h)W_0$$

where W_0 is the final weight matrix. This mechanism enables the model to capture multi-level feature information across different subspaces simultaneously, thereby enhancing overall representation capability and generalization performance.

1.3 Residual Connection Module

Residual connections represent an effective technique for addressing gradient vanishing problems in deep network training. By directly passing input data to deeper layers, they prevent gradient attenuation during backpropagation, thereby accelerating gradient propagation. The core of ResNet, proposed by He et al. in 2015, is the introduction of residual connections. [Figure 4: see original paper] depicts two module structures in neural networks: a normal block is the basic building unit in neural networks, typically consisting of weight layers and activation functions, where input x undergoes linear transformation and nonlinear activation to output $f(x)$. A residual connection block is a key component of ResNet that addresses degradation problems in deep networks through residual learning. The residual connection block includes weight layers, activation functions, and residual connections. Input x first undergoes linear transformation through weight layers, then passes through activation functions to obtain $f(x)$. Subsequently, $f(x)$ is added to the original input x to obtain the final output. This structure can be expressed as $y = f(x) + x$, where y is the output, $f(x)$ is the result processed by weight layers and activation functions, and x is the input.

1.4 AR-3D CNN Model

The innovation of the AR-3D CNN model lies in its combination of residual networks, multi-head self-attention mechanisms, and three-dimensional convolutional neural networks to address the complex requirements of core neutron flux distribution prediction tasks. This integration enhances prediction accuracy, computational efficiency, and robustness while improving the capability to process three-dimensional spatial data. The overall model flow is shown in [Figure 5: see original paper]. The model comprises the following main steps: (1) In the input layer, 3D CNN extracts three-dimensional spatial features from core neutron flux data, combined with Batch Normalization (BatchNorm) and ReLU activation functions to enhance nonlinear expressive capability; (2) Multiple residual 3D convolution blocks (Residual 3DConvBlock) extract deep features, avoiding gradient problems and learning complex spatial dependencies; (3) The multi-head attention layer (Multi-Head Attention Layer) dynamically focuses on key regions, enhancing feature learning and noise robustness; (4) Fully connected layers (Fully Connected Layer) further integrate features to complete nonlinear mapping and final prediction; (5) Average pooling layers (Average Pooling Layer) aggregate feature maps to optimize output results.

1.5 ARGA-3D CNN Model

This paper combines deep learning with genetic algorithm optimization methods, employing a hyperparameter optimization strategy based on validation set loss minimization. In this approach, the deep learning model first extracts features from input data and performs multi-task prediction modeling. The network architecture includes 3D convolutional layers, residual connection layers, and multi-head attention mechanisms to capture spatiotemporal features and global dependencies. Subsequently, the genetic algorithm defines the negative value of the loss function (Huber loss) on the validation set as the fitness function:

$$\text{Fitness} = -\text{Huber}(y_{\text{true}}, y_{\text{pred}})$$

where N represents the number of validation set samples, $y_{\text{pred}}(i)$ denotes the predicted value for the i-th sample, and $y_{\text{true}}(i)$ denotes the true value for the i-th sample. The Huber loss function combines the advantages of Mean Squared Error (MSE) and Mean Absolute Error (MAE), improving robustness to outliers while ensuring smoothness, thereby optimizing model stability and convergence.

The ARGA-3D CNN model extracts three-dimensional features through 3D CNN and utilizes genetic algorithms to optimize model parameters, improving prediction accuracy and generalization capability. The overall model flow is shown in [Figure 6: see original paper], involving key steps such as data preprocessing, model initialization, model optimization, and genetic algorithm search for optimal parameters, enabling the model to achieve superior performance in complex three-dimensional data scenarios.

2.1 Core Model

This experiment employs the Monte Carlo particle transport simulation software SuperMC to construct a lead-cooled fast reactor model. The number and materials of core components are summarized in . As shown in [Figure 7: see original paper], the core components adopt the hexagonal prism shape commonly used in fast reactors, with the core divided into three regions from inner to outer: active zone, reflector zone, and shielding zone. The active zone consists of 101 assemblies, including 92 fuel assemblies (primarily 19.75% enriched UO_2) and 9 control rod assemblies. The reflector zone comprises 98 assemblies using lead-bismuth eutectic material, while the shielding zone contains 54 assemblies made of boron carbide. Additionally, there are 4 safety rods, 3 control rods, and 2 compensation rods made of 90%^{10}B boron carbide, boron carbide, and 90%^{10}B boron carbide, respectively.

SuperMC is a large-scale nuclear neutronics design and radiation safety evaluation software system independently developed by the Institute of Nuclear Energy Safety Technology (INEST), Hefei Institutes of Physical Science, Chinese Academy of Sciences. Based on the Monte Carlo method, this software can perform accurate neutron transport calculations for complex nuclear systems and is widely applied in nuclear reactor design, radiation safety analysis, and various aspects of nuclear engineering. SuperMC can handle reactor models with complex geometries and effectively simulate multi-material interfaces, providing high-precision neutron flux calculations. Its main functions cover comprehensive neutron safety analysis throughout the entire process, including radiation transport, fuel depletion, radiation activation, and dose assessment. [Figure 8: see original paper] shows the overall core structure of the lead-cooled fast reactor constructed using SuperMC.

2.2 Experimental Dataset and Evaluation Metrics

The dataset was obtained through core simulation programs that calculate core neutron flux density under steady-state conditions. Input data includes positions of three control rods, positions of two compensation rods, and 30,000 spatial coordinate information entries, which collectively describe the physical state of the core and its influence on neutron flux distribution. Output data consists of 30,000 neutron flux values corresponding to the spatial coordinates, reflecting neutron flux intensity at various core positions under different operating conditions. The data is shown in . To ensure stability and efficiency of model training, all input and output data were normalized. Through this dataset, the model learns the complex nonlinear relationship between input parameters and neutron flux distribution, thereby achieving accurate prediction of core neutron flux.

This experiment employs four widely used metrics to evaluate model accuracy and performance: MAE, MSE, Mean Absolute Percentage Error (MAPE), and coefficient of determination (R^2). These metrics reflect the model' s fitting effectiveness and are calculated as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\%$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

where y_i is the true value, $\hat{y}_i$ is the predicted value, n is the sample size, and $\bar{y}$ is the mean of true values. The R^2 value ranges from 0 to 1, with values closer to 1 indicating better model fit.

2.3 Hyperparameter Optimization

This experiment employs genetic algorithms to adjust neural network hyperparameters. Genetic algorithms can perform global optimization in high-dimensional search spaces to gradually find optimal hyperparameter combinations. Given the relatively complex structure and large data volume of this experiment' s dataset, a hybrid strategy is adopted: the learning rate is treated as a continuous hyperparameter searched on a logarithmic scale, while other hyperparameters (such as batch size, number of attention heads, and number of fully connected layer nodes) use discrete selection strategies. This hybrid approach balances computational efficiency and search precision, allowing both fine-tuning of the learning rate and optimal selection from finite sets of discrete hyperparameters.

After multiple iterations and optimization through the genetic algorithm, a final hyperparameter configuration was determined, as shown in . The specific hyperparameters are: learning rate 1.17$×10^{-4}$, batch size 16, number of attention heads 4, and number of fully connected layer nodes 1024, representing the optimal solution for model performance within the current search space. During training, both training loss and validation loss decreased significantly, as shown in [Figure 9: see original paper], indicating that the model gradually converged and achieved good performance during the learning process. The loss curves show that loss values stabilized as training progressed, validating the effectiveness of the optimization process.

2.4 Comparative Experiments

To comprehensively evaluate the effectiveness of the proposed ARGA-3D CNN model, comparative experiments were conducted with multiple advanced deep learning prediction models, including traditional CNN, ANN, and Long Short-Term Memory (LSTM) networks, which are widely used in prediction tasks. All models were trained and tested on the lead-cooled fast reactor three-dimensional neutron flux dataset.

The comparative experimental results are presented in . In terms of MAE and MSE, ARGA-3D CNN achieves errors of only 3.19$×10^{-6}$ and 2.14$×10^{-11}, respectively, reducing errors by more than half compared to CNN, ANN, and LSTM. This demonstrates that the method po
ARGA-3D CNN attains the best fit of 0.9735, significantly higher than CNN' s 0.9118, ANN' s 0.9134, and LSTM' s 0.9212. This indicates that ARGA-3D CNN' s powerful capability in three-dimensional spatial modeling enables it to effectively learn the nonlinear features of neutron flux and construct prediction models that better conform to physical laws.

To further visually demonstrate the performance of different models, visualization results are provided. The neutron flux distribution predicted by the ARGA-3D CNN model was compared with the distribution calculated by the Monte Carlo program on the test dataset. Considering the relative error between predicted and calculated values, one instance was randomly selected from the test dataset to evaluate the surrogate model' s performance. [Figure 10: see original paper] shows a two-dimensional comparison of predicted versus actual neutron flux distribution, while [Figure 11: see original paper] provides a three-dimensional comparison. Overall, the predicted neutron flux distribution matches well with the actual distribution, with the highest neutron flux observed in the reactor core, which is completely consistent with the expected behavior of reactor core neutron flux.

In summary, ARGA-3D CNN significantly outperforms traditional methods in neutron flux prediction tasks through deep three-dimensional convolutional feature extraction, optimized training stability from residual networks, and global information modeling from self-attention mechanisms. Compared to CNN, ANN, and LSTM, ARGA-3D CNN exhibits lower errors, more accurate predictions, and greater model stability while maintaining robustness in complex data environments, providing a new efficient and accurate prediction method for nuclear reactor design and safety assessment with promising application prospects.

2.5 Ablation Experiments

To deeply analyze the contribution of each module in ARGA-3D CNN to overall performance, ablation experiments were conducted by removing the genetic algorithm (AR-3D CNN), residual network (A-3D CNN), multi-head self-attention mechanism (R-3D CNN), and global average pooling layer (AR-3D CNN(NAP)), respectively, to verify the impact of different modules on final prediction results. The experimental results are shown in .

The results demonstrate that each module in the proposed ARGA-3D CNN network effectively enhances performance on three-dimensional neutron flux prediction tasks. Introducing genetic algorithms (GA) for hyperparameter optimization significantly enhances global search capability, reduces the risk of local optima from manual tuning, and further improves model accuracy and stability. The ResNet module alleviates gradient vanishing and explosion problems through skip connections in deep structures while enhancing feature transmission efficiency, making a particularly prominent contribution to network performance. MSA accurately captures global dependencies of input data in three-dimensional space, helping the model focus on key regions and suppress interference, resulting in more accurate predictions. The Global Average Pooling (GAP) layer compresses channel dimensions and reduces redundant parameters while maintaining the integrity of main features, thereby improving network generalization capability. When GAP is removed, the network' s utilization of detailed features decreases, leading to slightly degraded reconstruction performance. Overall, the synergistic effect of these modules enables ARGA-3D CNN to exhibit stronger adaptability and higher prediction accuracy for complex three-dimensional neutron flux data, establishing a solid foundation for subsequent prediction and analysis of key parameters in nuclear reactor cores.

2.6 Timeliness Experiment

This experiment aims to evaluate the timeliness advantages of the proposed ARGA-3D CNN model, particularly in comparison with the nuclear energy simulation software SuperMC. All timeliness experiments were conducted under identical hardware environments to ensure comparability. The experimental computer configuration includes: AMD Ryzen 7 5800 CPU, 32GB memory, and NVIDIA GeForce RTX 3060 GPU. Timeliness was evaluated by comparing the time required for both methods to predict three-dimensional neutron flux. When using SuperMC with 12-core parallel computing, completing one full prediction requires minutes. In contrast, the ARGA-3D CNN model completes the same prediction in only 0.4011 seconds. This demonstrates that the model possesses significant advantages in computational timeliness, especially when handling complex prediction tasks, as it can complete calculations in extremely short time and substantially improve processing efficiency.

Conclusions

This paper proposes a three-dimensional neutron flux prediction model based on ARGA-3D CNN to address the complex spatial feature extraction challenges in lead-cooled fast reactor core neutron flux distribution prediction, significantly improving prediction accuracy and computational efficiency. The main conclusions are as follows:

1) By introducing ResNet and MSA, the model effectively alleviates gradient vanishing problems and enhances global information modeling capability, thereby improving prediction accuracy and reducing errors common in traditional methods. Compared to CNN, ANN, and LSTM, the ARGA-3D CNN model demonstrates superiority across evaluation metrics including MAE, MSE, MAPE, and R^2, requiring only 0.4011 seconds to predict 50,000 data points—a substantial improvement in computational efficiency compared to traditional calculation software.

2) Combining genetic algorithm optimization for hyperparameters further improves model robustness and accuracy. The optimized hyperparameter configuration enables more efficient performance in large-scale core data prediction tasks while exhibiting stronger generalization capability.

In summary, the proposed model provides a new approach for efficient three-dimensional neutron flux prediction, particularly showing promising application prospects in nuclear reactor design and safety assessment. Future work will introduce more diverse datasets covering different operating conditions and scenarios to further enhance model adaptability and robustness in practical applications.

Author Contributions

Ziwen Mo: Overall paper design, neural network programming, manuscript drafting.
Zihui Yang: Critical review of the article, research funding support.
Zhongyang Li, Guomin Sun: Technical support and guidance.
Zhaodong Li: Core data acquisition and data analysis.

References

1. Wei S Y, Wang C L, Tian W X, et al. Research progress in key thermal-hydraulic issue of lead-based fast reactor [J]. *Atomic Energy Science and Technology*, 2019, 53(2): 326–336. DOI: 10.7538/yzk.2018.youxian.0335.

2. Song T X, Zhang L, Gu Z X, et al. Transient analyses on overpower transients of small modular natural circulation Lead-cooled fast reactor SNCLFR-100[J]. *Kerntechnik*, 2025, 90(2): 157–171. DOI: 10.1515/kern-2024-0076.

3. Li J, Li W, Zhang X, et al. Research Status and Suggestions of Nuclear Reactor Online Core Monitoring System [J]. *Atomic Energy Science and Technology*, 2024, 58(7): 1393–1405. DOI: 10.7538/yzk.2024.youxian.0147.

4. Yu Y, Park H, Novak A, et al. High fidelity multiphysics tightly coupled model for a lead cooled fast reactor concept and application to statistical calculation of hot channel factors[J]. *Nuclear Engineering and Design*, 2025, 435: 113915. DOI: 10.1016/j.nucengdes.2025.113915.

5. Zare S G, Ehsan Z. Neutronic investigation of ALFRED lead-Cooled fast reactor core using monte carlo modeling[J]. *Atomic Energy*, 2024, 134(3-4): 264–274. DOI: 10.1007/S10512-024-01054-2.

6. Zhang D K, Tang Z H, Wei Y J, et al. Uncertainty analysis method and application of neutron multisphere spectrometer measurement based on Monte Carlo method[J]. *Nuclear Techniques*, 2024, 47(12): 65–74. DOI: 10.11889/j.0253-3219.2024.hjs.47.120403.

7. Wang T, Liu Z J, Zhao P C, et al. Design method of high-flux lead-bismuth cooled reactor neutron flux maximization based on BP neural network[J]. *Nuclear Techniques*, 2024, 47(10): 100602. DOI: 10.11889/j.0253-3219.2024.hjs.47.100602.

8. Zhang H M, Zhang H C. Prediction method of effective multiplication factor of reactor core based on convolution neural network model[J]. *Modern Applied Physics*, 2022, 13(2): 61–66. DOI: 10.12061/j.issn.2095-6223.2022.020209.

9. Zhou T Z, Yu K C, Cheng M S, et al. Development and analysis of a transient identification model for molten salt reactor systems based on

the KNN method[J]. *Nuclear Techniques*, 2023, 46(11): 125–135. DOI: 10.11889/j.0253-3219.2023.hjs.46.110604.

10. Ji N, Yi J H, Zhao P C, et al. Research on adaptive RBF neural network prediction method for core thermal-hydraulic parameters of fast reactor[J]. *Nuclear Techniques*, 2022, 45(9): 090601. DOI: 10.11889/j.0253-3219.2022.hjs.45.090601.

11. Wang D, Wang W C, Pan C J, et al. Prediction method for key parameters of core refueling in pressurized water reactors based on an adaptive BP neural network[J]. *Atomic Energy Science and Technology*, 2020, 54(01): 112–118. DOI: 10.7538/yzk.2019.youxian.0016.

12. Palmi K, Kubinski W, Darnowski P. Prediction of the evolution of the nuclear reactor core parameters using artificial neural network[J]. *Annals of Nuclear Energy*, 2025, 211: 110891. DOI: 10.1016/j.anucene.2024.110891.

13. Bei X Y, Dai Y Q, Yu K C, et al. Three-dimensional surrogate model based on back-propagation neural network for key neutronics parameters prediction in molten salt reactor[J]. *Energies*, 2023, 16(10): 4044. DOI: 10.3390/en16104044.

14. Chen J Y, Liu X Y, Zhao P C, et al. Study on the prediction method of neutron flux and keff in reactors based on the optimized extreme learning machine model[J]. *Nuclear Techniques*, 2024, 47(10): 178–187. DOI: 10.11889/j.0253-3219.2024.hjs.47.100604.

15. Liu S Q, Wang X, Zhao L, et al. 3DCANN: A spatio-temporal convolution attention neural network for EEG emotion recognition[J]. *IEEE Journal of Biomedical and Health Informatics*, 2022, 26(11): 5321–5331. DOI: 10.1109/JBHI.2022.3171330.

16. Bai J, Yang Z Y, Peng B, Li W J. Research on 3D convolutional neural network and its application to video understanding[J]. *Journal of Electronics & Information Technology*, 2023, 45(6): 2273–2283. DOI: 10.11999/JEIT220596.

17. Li Z H, Zhang L, Shan X Y. Review of improvements in three-dimensional convolutional neural network methods and their applications[J]. *Journal of Computer Engineering & Applications*, 2025, 61(3): 48–61. DOI: 10.3778/j.issn.1002-8331.2407-0031.

18. Gao G S. Survey on attention mechanisms in deep learning recommendation models[J]. *Computer Engineering and Applications*, 2022, 58(9): 9–18. DOI: 10.3778/j.issn.1002-8331.2112-0382.

19. Alaeddine H, Jihene M. Deep residual network in network[J]. *Computational Intelligence and Neuroscience*, 2021, 2021(1): 6659083. DOI: 10.1155/2021/6659083.

20. He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770–778. DOI: 10.1109/CVPR.2016.90.

21. Bouktif S, Fiaz A, Ouni A L, et al. Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: comparison with machine learning approaches[J]. *Energies*, 2018, 11(7): 1636. DOI: 10.3390/en11071636.

22. Ozturk N, Celik E. Application of genetic algorithms to core loss coefficient extraction[J]. *Progress in Electromagnetics Research M*, 2011, 19: 133–146. DOI: 10.2528/pierm11051310.

23. Aszemi N M, Dominic P D D. Hyperparameter optimization in convolutional neural network using genetic algorithms[J]. *International Journal of Advanced Computer Science and Applications*, 2019, 10(6): 269–278. DOI: 10.14569/ijacsa.2019.0100638.

Note: Figure translations are in progress. See original paper for figures.