Enhanced accident diagnosis for molten salt reactors using an adaptive residual CNN with Bayesian optimization

Authors: Chaoqun Wang, Kai Wang, Qun Yang, Rongjian Liang, Xin Yue, Liangjie Sun, Wentao Jiang, Zhaozhong He, Naxiu Wang, Qun Yang, Naxiu Wang

Date: 2026-03-19T10:08:41+00:00

Abstract

Safety is of paramount importance in nuclear power plants. Accurate and reliable accident diagnosis is essen#2;tial for ensuring operational safety in reactor systems. The convergence of Industry 4.0 technologies and deeplearning methods has emerged as a promising approach for improving the operational safety of nuclear energysystems, particularly in fault detection and diagnosis (FDD) applications. This study proposes a novel adaptiveaccident diagnosis framework tailored for molten salt reactors (MSRs) based on an enhanced residual convo#2;lutional neural network (AM-RCNN). The AM-RCNN incorporates an anti-noise module implemented usingthe Soft Threshold method, together with an attention mechanism, to improve robustness. Datasets representingeight distinct operational scenarios were generated using the RELAP5-TMSR simulation tool. An appropriatesubset of input features for MSR accident diagnosis was selected using Pearson correlation analysis and randomforest importance ranking. The models were subsequently trained, validated, optimized, and tested. Compar#2;ative analyses with conventional RCNN and CNN architectures demonstrate the diagnostic advantages of theproposed approach. In addition, the integration of Bayesian optimization further enhances the performance ofthe AM-RCNN. As a contribution to intelligent monitoring research for MSRs, the proposed method providesreliable decision support for nuclear system operation, particularly in autonomous scenarios.

Full Text

Preamble

Enhanced accident diagnosis for molten salt reactors using an adaptive residual CNN with Bayesian optimization Chao-Qun Wang, Kai Wang, Qun Yang, 1, 2, Rong-Jian Liang, Xin Yue, Liang-Jie Sun, Wen-Tao Jiang, Zhao-Zhong He, and Na-Xiu Wang 1, 2, 1 Shanghai Institute of Applied Physics, Chinese Academy of Sciences, 201800 Shanghai, China University of the Chinese Academy of Sciences, 100049 Beijing, China Safety is of paramount importance in nuclear power plants. Accurate and reliable accident diagnosis is essen- tial for ensuring operational safety in reactor systems. The convergence of Industry 4.0 technologies and deep learning methods has emerged as a promising approach for improving the operational safety of nuclear energy systems, particularly in fault detection and diagnosis (FDD) applications. This study proposes a novel adaptive accident diagnosis framework tailored for molten salt reactors (MSRs) based on an enhanced residual convo- lutional neural network (AM-RCNN). The AM-RCNN incorporates an anti-noise module implemented using the Soft Threshold method, together with an attention mechanism, to improve robustness. Datasets representing eight distinct operational scenarios were generated using the RELAP5-TMSR simulation tool. An appropriate subset of input features for MSR accident diagnosis was selected using Pearson correlation analysis and random forest importance ranking. The models were subsequently trained, validated, optimized, and tested. Compar- ative analyses with conventional RCNN and CNN architectures demonstrate the diagnostic advantages of the proposed approach. In addition, the integration of Bayesian optimization further enhances the performance of the AM-RCNN. As a contribution to intelligent monitoring research for MSRs, the proposed method provides reliable decision support for nuclear system operation, particularly in autonomous scenarios.

Keywords

Accident diagnosis, Feature selection, Deep residual CNN, Bayesian optimization, Molten salt reactors

INTRODUCTION

Escalating global energy consumption exacerbates climate change, underscoring the critical need for carbon emission mitigation strategies to address environmental challenges [As a nearly carbon-free power source, nuclear energy plays a significant role in mitigating the global greenhouse effect and addressing global warming [ ]. The widespread imple- mentation of nuclear energy relies predominantly on nuclear power plants (NPPs) that harness nuclear fission reactions.

Compared with traditional energy sources, NPPs have more stringent safety requirements owing to their complex nature and the potential hazards associated with radioactive materi- als [In pursuit of nuclear reactor designs with enhanced safety and economic viability, the Generation IV International Forum (GIF) proposed six reactor technologies for future development, among which molten salt reactors (MSRs) are considered a promising alternative [ ]. MSRs exhibit exceptional inherent safety characteristics owing to their unique design and operating principles. In these reactors, liquid molten salt serves a dual function as both the nuclear fuel carrier and primary coolant, efficiently transferring heat from the core [The safety advantages of MSRs are reflected in several aspects: (1) operation at near-atmospheric pressure, (2) strong negative reactivity temperature feedback combined with a large void coefficient [ ], (3) low proliferation risk, and (4) immunity to fuel irradiation damage or mechanical failure [ ]. Despite these inherent safety features, potential risks associated with operational anomalies and accidents remain significant concerns. Therefore, accurate and timely accident diagnosis is essential to prevent the escalation of abnormal events and mitigate their consequences, serving as an external safeguard for reactor operation.

This work was supported by the Youth Innovation Promotion Association (YIPA) of the Chinese Academy of Sciences (Grant No. E329290101).

Currently, accident diagnosis in NPPs relies primarily on conventional threshold-monitoring approaches and human expertise [ ]. Although operators are typically well trained, they may experience significant psychological stress during NPP accidents, making it challenging to accurately assess the nature of an accident within a limited time frame. Consequently, further research on accident diagnosis technologies for nuclear energy systems is urgently needed. Developing systems capable of accurately and efficiently identifying accidents is essential for supporting prompt decision-making during emergency response.

Accident diagnosis strategies for NPPs are generally classified into two methodological approaches: model-driven and data-driven. Model-based methods require the formulation of precise mathematical representations that capture the fundamental physical principles governing system behavior [The continuous evolution of industrial technologies and computational capabilities has propelled modern nuclear systems toward increasingly sophisticated architectures characterized by large-scale designs, complex structures, and distributed configurations [ ]. However, the inherent complexity of nuclear phenomena, coupled with uncertainties in input parameters [ ], poses significant challenges for developing high-fidelity models capable of accurately representing the behavior of these advanced systems.

Data-driven methodologies, in contrast, establish diagnostic frameworks through statistical analysis of operational data, thereby eliminating the requirement for explicit model-

ing of complex parameter interactions [ ]. Such approaches exploit the inherent correlations among system parameters by developing data-driven models in which complex interdependencies are automatically captured during the training process.

This paradigm has led to extensive research and practical applications of machine learning techniques in nu- clear system diagnostics, including support vector machines (SVMs) [ ], ensemble decision trees (DTs) [ ], adaptive nearest-neighbor algorithms (KNN) [ ], artificial neural net- works (ANNs) [ ], and multivariate statistical analy- sis [However, with the rapid advancement of artificial intelli- gence (AI), along with the deployment of advanced sensors and high-speed data transmission mechanisms, operational datasets reflecting reactor status have begun to exhibit big data characteristics. These developments have rendered con- ventional data-driven methods potentially insufficient for the demands of the emerging nuclear big data era. To address these limitations, deep learning architectures have emerged as a viable solution with remarkable capabilities for diagnos- ing anomalies and accidents in nuclear systems [Among deep learning architectures, convolutional neural networks (CNNs) have gained significant attention in fault diagnostic applications owing to their strong capability for extracting high-dimensional and complex features [ ]. For example, researchers have developed an enhanced CNN ar- chitecture incorporating mini-batch processing optimization, which has been successfully implemented in nuclear system diagnostics, achieving higher accuracy and providing valu- able guidance for operators [ ]. Additionally, a multiclas- sification framework based on a CNN architecture was de- veloped to detect core disturbances and abnormal operational signatures, demonstrating high robustness and practical po- tential [ ]. In a recent study, an SSA-CNN-LSTM model for NPP fault diagnosis was proposed, achieving a classifi- cation accuracy of 98.24% through CNN-LSTM parameter optimization using the sparrow search algorithm, thereby en- hancing nuclear safety and reducing hazards caused by hu- man error [However, a persistent limitation of conventional deep learning architectures is performance degradation as the net- work depth increases, which can adversely affect classification accuracy [ ]. This issue has been addressed by intro- ducing deep residual CNNs (ResNets) [ ]. Compared with traditional CNNs, ResNets employ identity shortcut connec- tions that reduce the difficulty of training deep architectures, thereby improving the model' s discriminative capability for complex classification tasks. As a result, ResNets have gradu- ally gained popularity in fault diagnosis applications [For example, adaptive wavelet transform features have been integrated with the ResNet architecture to enhance the fault detection capability of planetary gear transmission sys- tems [ ]. Sun and Wang developed several diagnostic mod- els by integrating multilayer perceptrons (MLPs), CNNs, re- current neural networks (RNNs), and ResNets, and compared their performance using a dataset from Case Western Reserve University [ ]. Zhao et al. adopted an adaptive paramet- ric rectified linear unit (ReLU) in ResNet to improve feature

learning capability, which was validated through fault diag- nosis applications [40]. Yao et al. proposed an innovative diagnostic framework incorporating adaptive residual con- volutional networks specifically designed for anomaly iden- tification in small modular reactor (SMR) operational sys- tems [41].

These studies provide useful references for applying the residual CNN (RCNN) method to diagnostic tasks in the nu- clear energy field. However, several issues still require further consideration and improvement:

(1) Nuclear energy systems consist of numerous subsys- tems and a large number of monitoring parameters.

These datasets often exhibit high-dimensional charac- teristics, posing significant challenges for data analysis and decision-making. To improve model learning ef- ficiency and performance, feature selection techniques should be implemented to identify the optimal feature subset for classification.

(2) In practical applications, data collected in NPPs are of- ten contaminated by noise. Therefore, it is necessary to combine the RCNN structure with an attention mecha- nism to enhance the model' s robustness to noise.

(3) In NPPs, different accidents may exhibit similar sig- nals during their early stages. RCNNs typically use ReLUs as activation functions, which have a constant slope and zero output in the negative region. This prop- erty may reduce effective feature learning capability, thereby limiting model performance. Therefore, an ac- tivation function with more dynamic parameters should be incorporated into the model to enhance classification capability.

(4) Hyperparameter optimization is essential for improv- ing the performance of deep learning models. An in- depth investigation of hyperparameters should there- fore be conducted for the proposed network to maxi- mize its effectiveness.

To address the aforementioned challenges and improve ac- cident diagnosis efficiency in nuclear energy systems, this study proposes a three-step methodology. First, Pearson cor- relation analysis and random forest importance ranking are used for feature selection. Next, a modified RCNN frame- work, termed an adaptive antinoise module RCNN (AM- RCNN), is developed for MSR accident diagnosis by inte- grating an antinoise module and adaptive parametric ReLUs.

Finally, Bayesian optimization (BO) is applied to adjust the model' s hyperparameters.

The remainder of this paper is organized as follows. Sec- outlines the methodological framework of the pro- posed approach. Section presents the case study for the nuclear accident diagnosis model. Section compares the results with those of alternative models and describes the op- timization process. Finally, Sect. concludes the paper with key findings and potential directions for future research.

METHODOLOGY

A. Feature selection technique

Feature selection is a methodological approach for extract- ing the most relevant subset of attributes from a comprehen- sive feature space based on predefined optimization criteria and statistical evaluation metrics. This process aims to retain only the most relevant features from a dataset, thereby reduc- ing the scale of data processing by eliminating redundant and irrelevant variables.

1. Pearson correlation coefficient

The Pearson correlation coefficient is a widely used sta- tistical measure for feature selection tasks, quantifying the degree of linear dependence between paired variables. This metric evaluates the strength of the association between two random variables by analyzing their joint variability, which is represented through their covariance structure. Mathemati- cally, the Pearson correlation coefficient for variables can be expressed as Eq. ( ], where cov represents the joint variability between variables , and var and var denote the variances of the respective variables.

$\mathrm{Corr}\ (\ \alpha,\ \beta\ ) = \mathrm{cov}\ (\ \alpha,\ \beta\ )$

The Pearson correlation coefficient ranges from A correlation value of zero (Corr) = 0) indicates no linear dependence between variables . Negative val- ues (Corr) indicate an inverse linear relationship, whereas positive values (Corr) indicate a direct linear relationship. The strength of the linear association is reflected by the absolute magnitude of the coefficient, with values closer to unity indicating stronger statistical relation- ships between the variables.

2. Random forest variable importance

Random forests are a class of nonlinear predictive models that aggregate multiple decision trees for regression or classi- fication tasks. Individual trees within the ensemble are gen- erated using a bootstrap sampling process, in which each tree is constructed from an independently drawn random subset of the training data.

Random forest models provide variable importance mea- sures that help identify features contributing significantly to class separation. Among these measures, permutation impor- tance is one of the most commonly used approaches. This method sequentially shuffles the values of each predictor vari- able and evaluates the corresponding reduction in the predic- tive performance of the model. Shuffling a variable disrupts its relationship with the outcome; therefore, a substantial de- crease in classification accuracy after permutation indicates a strong predictive relationship between that feature and the target variable.

The permutation-based importance score for each variable is computed using the following formula [where represents the model accuracy, denotes a distinct bootstrap sample, and corresponds to the out- of-bag (oob) samples with the th feature randomly shuffled.

For tree ensembles, the importance scores from individual trees can be averaged to obtain an overall measure of feature importance [

ω j = 1

k =1 ω j T L (θ k) (3)

Residual convolutional neural networks (RCNNs)

CNNs represent a class of deep-learning architectures that employ multiple filters to hierarchically process input sam- ples through convolution and pooling operations across suc- cessive layers [27]. Building on this concept, RCNNs intro- duce skip connections to effectively address the model degra- dation problem [35 , 36]. Instead of relying solely on deeper network structures to approximate complex hidden nonlin- earities, RCNNs facilitate the learning of residual represen- tations, which simplifies the optimization process. The archi- tectural configuration of RCNNs is illustrated in Fig. 1 [Figure 1: see original paper] .

The transformation from input to output can be expressed as follows [

y = H (x) = F (x) + x (4)

where represents the composite mapping learned by the network layers, corresponds to the residual function, denotes the input features. The residual function is often parameterized as to explicitly indicate its dependence on both the input and the learnable weights.

Accordingly, the th residual block can be formulated as fol- lows [

y n = H (x n) = F (x n , { W n }) (5)

x n +1 = f ReLU (y n) (6)

In this formulation, represent the input and output feature maps of the th residual module in the net- work architecture, respectively. The identity mapping is cap- tured by , whereas denotes the ReLU activation function. According to Eq. ( ), the feature propagation from th to the th layer can be expressed as [

x N = x n +

i = n F (x i , W i) (7)

In deep network architectures, the backpropagated gradient can be expressed as [

= ∂loss

∂loss

∂x n = ∂loss

i = n F (x i , W i)

Here, ∂loss

∂x n denotes the gradient propagated through the n th layer. The identity term 1 in this equation enables direct gradient propagation without attenuation. While the resid- ual component must pass through weighted layers, poten- tially modifying the gradient, the constant identity term en- sures stable gradient flow. Even when the residual gradient approaches zero, the presence of this identity term prevents the gradients from vanishing.

Anti-noise module

In this study, an anti-noise (AN) architecture that combines a soft threshold function with an attention mechanism inte- grated into an RCNN is proposed. The resulting framework, termed AM-RCNN, enhances noise robustness in MSR acci- dent diagnosis.

The soft thresholding (ST) operator is a widely adopted de- noising technique in signal processing applications. Its fun- damental principle involves threshold contraction: when the input exceeds a predefined threshold value, the output ap- proaches this boundary; otherwise, it is set to zero.

This mechanism effectively suppresses noise interference while preserving signal characteristics. The mathematical represen- tation is as follows [44]:

x $> \tau$,

Soft (x, τ) =

In conventional implementations, the threshold parameter τ typically requires manual configuration by system engi- neers prior to network deployment. In this study, the thresh- old is computed using an attention mechanism [45] that can

be trained within the network and updated adaptively. The architectural configuration of the AN module is illustrated in The global average pooling (GAP) operation in this sub- module computes channel-wise mean values while preserving essential feature information. Parallel to this path, a compact network processes the GAP outputs through two sequential fully connected layers, culminating in a sigmoid activation function that generates scaling factors within the range.

This auxiliary network incorporates batch normalization and PReLU layers for feature normalization and nonlinear activa- tion, respectively. The resulting adaptive threshold is ob- tained by multiplying the GAP-derived averages by the scal- ing coefficients produced by the compact network.

Bayesian theory-based optimization Hyperparameters play a crucial role in machine learning al- gorithms because they directly influence the training process and significantly affect model performance. Tuning these hy- perparameters often requires professional knowledge and ex- pert experience. Consequently, hyperparameter tuning rep- resents a challenging optimization problem with an implicit objective function. BO is a powerful approach for address- ing such optimization problems [ ]. Figure presents the workflow of BO.

The BO process begins by defining the hyperparameter space, which encompasses the feasible value ranges for each tunable parameter. The optimization procedure then begins by sampling an initial set of configurations for model training and validation. Following performance evaluation, an acqui- sition strategy determines the next parameter combination for evaluation, guided by probabilistic surrogate models.

This process iterates through successive cycles, refining the parameter selection based on the performance observed in previous iterations. The iterative procedure continues un- til a predefined termination condition is satisfied. Once this condition is met, the optimization process concludes by iden- tifying the hyperparameter configuration that yields the best model performance.

To provide a comprehensive and intuitive visualization of the computational procedure, a flowchart summarizing the in-

tegrated methodology is presented in Fig. . The flowchart il- lustrates the step-by-step procedure of the proposed method- ology. The process begins with feature selection using Pear- son correlation analysis and random forest importance rank- ing, followed by the construction and training of the AM- RCNN model. BO is then applied to fine-tune the model's hyperparameters, ultimately yielding an optimized diagnostic model for MSR accident scenarios.

CASE STUDY A designed MSR description In 2011, the Chinese Academy of Sciences (CAS) initi- ated a strategic research program focusing on advanced nu- clear fission energy systems, in which MSR technology rep- resented a key development pathway. As part of this initiative, the TMSR program [ ], a pioneering effort in MSR technol- ogy development, was conducted by the Shanghai Institute of Applied Physics (SINAP), CAS. The MSR investigated in this study corresponds to the design version proposed in the TMSR program [ ]. Figure illustrates the conceptual design of the reactor, and Table summarizes its key techni- cal specifications [The reactor adopts a dual-loop molten salt configuration.

The primary loop, designated as the fuel-salt system, com- prises a core vessel, primary heat exchanger, circulation pump, and connecting piping. The fuel salt circulates upward Main parameter Design value Thermal power

2 MWt

Fuel salt flow rate 50 kg Fuel salt temperature (inlet/outlet) Coolant salt flow rate 42 kg Coolant salt temperature (inlet/outlet)

from the bottom plenum through the core and then flows into the shell side of the heat exchanger via the circulation pump.

Thermal energy is subsequently transferred to the secondary loop, after which the cooled fuel salt completes the cycle by returning to the bottom plenum.

The secondary loop, referred to as the coolant salt system, contains a circulation pump, air-cooled heat exchanger, and associated piping. The coolant salt flows through the tube side of the primary heat exchanger, absorbs thermal energy from the primary loop, and releases heat to the atmosphere through the air-cooled exchanger. Furthermore, the MSR de- sign incorporates a passive reactor vessel air cooling system (RVACS) for effective decay heat removal.

Operation scenario data collection Given the limited availability of operational data for MSRs, accident diagnosis datasets were generated using the RELAP5-TMSR simulation code. This specialized tool, de- veloped at SINAP [ ], extends RELAP5/MOD4.0 with enhanced capabilities for liquid-fueled reactor analysis. The accuracy of the code has been verified through experiments conducted on a high-temperature molten-salt natural circu- lation test bench [ ], and it is currently applied in the design and analysis of MSR systems. Figure illustrates the nodal configuration, where the primary circuit components are labeled 0XX, while 1XX and 2XX denote the secondary loop and the air-cooling system, respectively. The RVACS is modeled using nodes 301–306.

In this study, eight typical operational scenarios are consid- ered: one steady-state condition and seven transient events.

By training the model using data that include both steady- state and transient conditions, the model' s generalization ca- pability is improved, enabling comprehensive accident detec- tion across a wide range of operating states. Table lists these operational conditions.

Given the varying transient characteristics at different power levels, each scenario was generated at reactor power levels ranging from 10% to 120% of full power (FP) in in- crements of 10% FP. The resulting dataset comprised more than 550,000 data samples, with each sample consisting of 21 monitoring parameters aligned with the architecture of the instrumentation and control system. The data-partitioning scheme allocated 70% of the dataset for model training, while equal proportions of 15% were reserved for validation and testing.

Additionally, to prevent the diagnostic system from ignor- ing variables with small absolute values caused by differences in measurement scales, the Z-score normalization method was adopted [Feature selection The 21 monitored parameters are listed in Table Assuming that all monitoring parameters are used as in- put features would impose a substantial computational bur- den, thereby compromising the efficiency and performance of the model. Therefore, feature selection was performed to reduce redundant variables among these parameters.

Reactor monitoring systems typically generate multiple in- terdependent parameters in which strong correlations may exist (e.g., perfectly correlated variables contain redundant information and can be represented by a single parameter).

Therefore, the Pearson correlation metric was used as the pri- mary statistical tool to quantify the linear dependence among system variables.

The Pearson correlation coefficients be- tween variables are shown in Fig. , where features 0–20 cor- respond to the parameters listed in Table

A Pearson correlation coefficient with an absolute value ex- ceeding 0.8 is commonly considered to indicate a strong cor- relation between two parameters [56]. However, this thresh- old should be evaluated within the specific context of the ap- plication. In practical scenarios, when two parameters exhibit strong correlation, preference is typically given to retaining the parameter that is easier to measure and has higher mea- surement accuracy.

I (normal) Steady state II (accident) Uncontrolled withdrawal of a control rod (shim rod) III (accident) Uncontrolled withdrawal of a control rod (regulating rod) IV (accident) Coolant circulating pump rotor seizure V (accident) Loss of air flow of the molten salt-air heat exchanger VI (accident) Fuel circulating pump speed increases unexpectedly VII (accident) Coolant circulating pump speed increases unexpectedly VIII (accident) Loss of offsite power

Monitoring parameter Unit 0-Reactivity $ 1-Reactor power MW 2-Fuel circulating pump speed (FCPS) rpm 3-Coolant circulating pump speed (CCPS) rpm 4-Fuel salt mass flow rate (FSMFR) kg / s 5-Coolant salt mass flow rate (CSMFR) kg / s 6-Air mass flow rate of the molten salt-air heat exchanger (AMFR) kg / s 7-Core inlet temperature (CIT) ∘ C 8-Middle core temperature (MCT) ∘ C 9-Core outlet temperature (COT) ∘ C 10-Upper plenum temperature (UPT) ∘ C 11-Tube side inlet temperature of the primary heat exchanger (TITHE) ∘ C 12-Tube side outlet temperature of the primary heat exchanger (TOTHE) ∘ C 13-Tube side inlet temperature of the molten salt-air heat exchanger (TITSHE) ∘ C 14-Tube side outlet temperature of the molten salt-air heat exchanger (TOTSHE) ∘ C 15-Air side inlet temperature of the molten salt-air heat exchanger (AITSHE) ∘ C 16-Air side outlet temperature of the molten salt-air heat exchanger (AOTSHE) ∘ C 17-Power of PRVACS kW 18-Air mass flow rate of the PRVACS kg / s 19-Air inlet temperature of PRVACS ∘ C 20-Air outlet temperature of PRVACS ∘ C

Following this principle, a parameter-screening process was conducted to evaluate parameter significance across three correlation thresholds: , and perfect linear correlation ( ). The results indicate that selecting a threshold greater than 0.9 for parameter retention yields opti- mal model performance. Accordingly, the input feature sub- set filtered using the Pearson correlation coefficient comprises seven parameters: reactor power; fuel circulating pump speed (FCPS); coolant circulating pump speed (CCPS); inlet, mid- dle, and outlet temperatures of the reactor core (CIT, MCT, and COT); and the inlet temperature of the primary heat ex- changer tube side (TITHE).

Uncorrelated features in machine learning models may act as noise sources, potentially reducing model accuracy by in- troducing classification bias. The random forest variable im- portance method was therefore applied to identify the rel- evance between variables and classes. The evaluation and ranking of feature importance are shown in Fig. . To ensure that the selected features effectively support the model’ s diag- nostic capability, scenarios with relative importance thresh- olds of 0.05 and 0.10 were examined during the parameter-screening process. Rigorous testing indicated that the model achieved the best performance with a threshold of 0.05. Con- sequently, all seven parameters were retained in the final fea- ture set to preserve diagnostic accuracy and ensure robust sys-

tem performance. Model structure and hyperparameter setting The structure of the AM-RCNN model for MSR accident diagnosis, which integrates advanced deep-learning method- ologies, is shown in Fig. , and Table provides details of the AM-RCNN architecture.

The AM-RCNN adopts a four-convolutional-layer archi- tecture with batch normalization layers and activation func- tions incorporated throughout the network to enhance per- formance. It integrates several key components, including a residual mapping module for feature transformation, an anti- noise (AN) processing unit for signal enhancement, and a GAP layer for feature aggregation.

The subsequent layers include a tensor-flattening operation for dimensionality reduction, a dropout regularization layer to mitigate overfitting [ ], and a fully connected transforma- tion layer for feature classification. The network employs a softmax activation function for final classification decisions, while utilizing parametric ReLUs (PReLUs) throughout the model to improve feature representation learning [ ]. The shortcut connection in the residual architecture from Layers 3 to 14 is automatically adapted to either identity mapping or a convolutional layer to maintain dimensional consistency with the main network structure. The AN processing module span- ning network layers 9–14 is described in detail in Sect. and illustrated in Fig.

Shortcut AN module Evaluation index Model performance was evaluated using confusion matri- ces [ ], which provide a detailed analysis of classification performance. As illustrated in Fig. , the matrix structure for multiclass problems assigns distinct labels to each cate- gory, with diagonal elements indicating correct predictions.

The row sums correspond to the number of samples belong- ing to each actual class.

In addition, the evaluation incorporated accuracy and the F1-score metric. Classification accuracy is a fundamental per- formance indicator that quantifies the ratio of correctly pre- dicted instances to the total number of samples. The F1-score represents the harmonic mean of precision and recall and is particularly sensitive to the lower value between these two measures. Higher F1-scores indicate improved model perfor- mance, especially in handling imbalanced classification tasks.

Detailed definitions of precision, recall, F1-score, and diag- nostic accuracy are provided in Eqs. (

Precision = TP TP + FP $\times$100% (10)

Recall = TP TP + FN $\times$100% (11)

F1-score = 2 $\times$Precision $\times$Recall Precision + Recall $\times$100 (12)

Diagnosis accuracy = TP + TN TP + FP + TN + FN ×100% (13)

where (true positive) and (true negative) represent the numbers of correctly classified positive and negative in- stances, respectively, while (false positive) and (false negative) represent the numbers of incorrectly classified neg- ative and positive instances.

RESULTS AND DISCUSSION Model performance and comparison The case study primarily evaluated the proposed AM- RCNN against conventional CNN and RCNN architectures, focusing on comparative training and testing performance within similar network frameworks.

Figure illustrates the architectural differences among the baseline CNN, standard RCNN, and the proposed AM- RCNN approach. To ensure consistency, the number of con- volutional layers in both the RCNN and CNN models was fixed at four, matching that of the AM-RCNN. In addition, the other layers in the main structures of the RCNN and CNN were identical to those of the AM-RCNN. The shortcut struc- ture in the RCNN was configured as identity mapping.

The training procedure employed cross-entropy as the op- timization objective function. Model parameter optimization was performed using stochastic gradient descent (SGD) with momentum acceleration and L2 regularization. The convolu- tional and fully connected layers were initialized using Kaim- ing and Xavier initialization methods [61 , 62], respectively.

Based on prior training experience, the model hyperparame- ters were configured as follows: a learning rate of 0.005 for

gradient updates, a dropout probability of 0.5 for regulariza- tion, a mini-batch size of 128 samples per iteration, and 200 training epochs.

To prevent overfitting and determine the optimal model configuration, an early stopping strategy was adopted. The validation loss function served as the primary convergence criterion, triggering early termination when no improvement was observed over ten consecutive training epochs. In such cases, the model parameters corresponding to the minimum validation loss, typically occurring ten epochs prior to termi- nation, were retained as the optimal configuration.

Figure 11 [Figure 11: see original paper] presents the epoch-wise progression of training loss and validation accuracy across the different models, re- vealing several key observations.

(1) According to the training loss curves, the proposed AM-RCNN model converged faster than the RCNN and CNN models. This indicates that the AM-RCNN model exhibits enhanced generalization capability and operational robustness, making it particularly suitable for deployment in noisy environments and transfer- learning scenarios.

(2) The proposed AM-RCNN model triggered early stop- ping at the 28th epoch, compared with the 62nd epoch for RCNN and the 90th epoch for

CNN. This demon- strates that the AM-RCNN model has stronger learning capability and stability, reducing training time and computational cost and making it suitable for flexible or real-time application environments.

(a) Training loss (b) Validation accuracy

All three models achieved high validation accuracy, ex- ceeding 90%. Specifically, the validation accuracy of the proposed AM-RCNN model exceeded 95%, reaching 97%, while the RCNN achieved 95% and the CNN reached 94%.

Operation scenario CNN RCNN AM-RCNN I (normal) II (accident) III (accident) IV (accident) V (accident) VI (accident) VII (accident) VIII (accident) 100 Average F1-score (%) Accuracy (%) F1-score (%) present the classification results for all evaluated models on the test dataset, including the con- fusion matrices and performance metrics (accuracy and F1- score). This analysis yielded several significant findings.

The accuracy of the AM-RCNN model on the test set is consistent with that on the validation set, further demonstrating the model' s generalization capability.

Analysis of the true positive and true negative rates in the confusion matrices indicates accurate identi- fication of all operational scenarios across all models. These findings demonstrate precise identification of critical system parameters associated with different operational scenarios. Comparative analysis revealed mean F1-scores of 93.9% for conventional CNN archi- tectures, 95.4% for standard RCNN implementations, and 97.2% for the proposed AM-RCNN framework.

(3) The relatively lower F1-scores observed for accident II, accident III, and steady-state conditions can be at- tributed to the inherent difficulty of distinguishing sce- narios with overlapping feature distributions during certain operational phases.

Specifically, accidents II and III, both reactivity insertion accidents, share the same triggering mechanism and exhibit similar devel- opment trends, making accurate classification partic- ularly challenging.

Similarly, certain transient acci- dent conditions were misclassified as steady-state con- ditions because the dynamic evolution of key parame- ters during the early stages of accidents often resembles steady-state behavior.

(4) The proposed model demonstrates superior perfor- mance in differentiating accidents II and III compared with conventional CNN and RCNN architectures, reflecting enhanced feature extraction and pattern- recognition capabilities.

Robustness test Liquid-fueled MSR operational environments inherently contain multiple noise sources that affect detector readings, necessitating a rigorous evaluation of the robustness of the diagnostic model.

Real-world operational conditions typically exhibit com- posite noise characteristics arising from multiple sources. For analytical purposes, the mixed noise was approximated as Gaussian white noise. The noise impact was quantified using the signal-to-noise ratio (SNR), which evaluates the relative strength between the target signal and background interfer- ence. The SNR is mathematically defined as follows [signal noise

SNR (dB) = 10 log 10

(a) CNN (b) RCNN (c) AM-RCNN (a) CNN (b) RCNN (c) AM-RCNN (a) Accuracy (b) F1-score where signal represents the spectral energy of the target sig- nal component and noise denotes the corresponding noise energy within the measurement bandwidth.

Monitoring signals in NPPs generally maintain relatively low noise levels, with vibration measurements typically ex- ceeding 30 dB SNR [64]. This threshold was therefore used as the baseline for assessing model robustness, and the corresponding test results are presented in Fig. 13 [Figure 13: see original paper] and Table 6 .

The results indicate that all three models performed well in

noisy environments, achieving diagnostic accuracy exceeding 90%. In particular, the accuracy and F1-score of the proposed model remained close to 97%.

As shown in Fig. 14 [Figure 14: see original paper] , the performance of all three mod- els decreased in noisy environments, with the conventional CNN model experiencing the most significant decline. Per- formance evaluation under noisy conditions revealed reduc- tions of 2.4% in classification accuracy and 2.6% in the F1 metric for the conventional CNN architecture. Similarly, the

Operation scenario CNN RCNN AM-RCNN I (normal) II (accident) III (accident) IV (accident) V (accident) VI (accident) VII (accident) VIII (accident) 100 Average F1-score (%) Accuracy (%) F1-score (%)

standard RCNN implementation exhibited an equivalent de- crease of 2.4% in both performance metrics.

In contrast, the proposed AM-RCNN framework demonstrated superior noise resilience, with only a 0.4% reduction in both metrics.

This result highlights the effectiveness of the integrated anti- noise processing module in maintaining model stability.

Model optimization To further enhance the performance of the AM-RCNN model, hyperparameter optimization was conducted using the BO method. In training deep neural networks, the learning rate, batch size, and dropout rate are critical hyperparameters that significantly influence learning efficiency, convergence speed, and model generalization capability. Consequently, these hyperparameters were selected as the primary optimiza- tion variables in this study.

The hyperparameter search space was defined as follows: the learning rate ranged from ; the batch size was randomly selected from the discrete set {32, 64, 128, 256}; and the dropout rate ranged from 0.1 to 0.6.

CONCLUSIONS

To satisfy the safety and reliability requirements of ad- vanced nuclear energy systems and reduce excessive manual intervention during operation, research on intelligent accident diagnosis systems is crucial for the future development of MSRs. This paper proposes AM-RCNN with an anti-noise module and a parametric activation function, and validates it using a designed MSR. The key contributions and distinc- tive features of the developed methodology are summarized as follows:

Rather than being limited to individual components or isolated subsystems, the selected input features com- prehensively represent the operational state of the en- tire reactor system architecture. These parameters were systematically filtered for MSR accident diagnosis us- ing Pearson correlation analysis and random forest im- portance ranking methods, thereby reducing data dimensionality and improving model efficiency.

Compared with traditional RCNN and CNN models, the proposed approach demonstrates improved diag- nostic performance, particularly in noisy environments.

(3) A BO framework was implemented for automated hy- perparameter tuning, specifically targeting the learn- ing rate, mini-batch size, and regularization parame- ters. This optimization strategy further improved the diagnostic performance of the AM-RCNN architecture, yielding a classification accuracy of over 98% when us- ing the optimized model configuration.

The developed model provides real-time operational deci- sion support for MSRs and has potential applications in other energy systems. Its architecture also offers a framework for developing deep-learning-based diagnostic approaches.

Future research directions include: (1) advanced data pre- processing techniques, (2) heuristic optimization algorithms for model refinement, and (3) adaptive learning rate strate- gies to enhance training efficiency. Furthermore, sequential deep learning architectures, including LSTM networks and gated recurrent units, will be explored and integrated into a predictive analytics framework for MSR anomaly detection and early warning systems.

AUTHOR CONTRIBUTIONS All authors contributed to the study conception and design.

Material preparation, data collection and analysis were per- formed by Wang Chaoqun. The first draft of the manuscript was written by Wang Chaoqun and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

DATA AVAILABILITY support findings study openly available Science Geng, Song, Static namic energy structure

analysis

world source optimization using total factor productivity method based on slacks-based measure integrating data envelopment analysis. Energy Convers. Manage. , 113713 (2021).

G. Gungor, R. Sari, Nuclear power and climate policy integra- tion in developed and developing countries. Renew. Sust. En- erg. Rev. , 112839 (2022).

A. Azam, M. Rafiq, M. Shafique et al., Analyzing the effect of natural gas, nuclear energy and renewable energy on GDP and carbon emissions: A multivariate panel data analysis. Energy , 119592 (2021).

P. Singh, K.L. Singh, Design of safety critical and control sys- tems of Nuclear Power Plants using Petri nets. Nucl. Eng. Tech- (5), 1289-1296 (2019).

M.H. Jiang, H.J. Xu, Z.M. Dai et al., Advanced fission en- ergy program—TMSR nuclear energy system. Bull. Chin.

Acad. Sci. (3), 366-374 (2012). Z.M. Dai, Thorium molten salt reactor nuclear energy system (TMSR), in Molten Salt Reactors and Thorium Energy , ed. by T.J. Dolan (Woodhead Publishing, Cambridge, 2017), pp. 531-H.J. Xu, Z.M. Dai, X.Z. Cai et al., Thorium based molten salt reactor and utilization of nuclear energy. Mod. Phys. (4), 25-34 (2018).

K. Wang, C.Q. Wang, Q. Yang et al., Uncertainty and sen- sibility analysis of reactivity insertion transient accident of a

150 MWt

molten salt reactor (SM-MSR). Nucl. Tech. (11), 110602 (2024).

R.J. Liang, L.J. Sun, X.W. Jiao et al., Reliability analysis and optimization of core fuel salt emergency drain system for the molten salt reactor experiment. Nucl. Tech. (3), 030604 (2023).

G.F. Zhu, R. Yan, H.H. Peng et al., Application of Monte Carlo method to calculate the effective delayed neutron frac- tion in molten salt reactor. Nucl. Sci. Tech. , 34 (2019).

M.L. Tan, G.F. Zhu, Z.D. Zhang et al., Burnup optimization of once-through molten salt reactors using enriched uranium and thorium. Nucl. Sci. Tech. , 5 (2022).

J.X. Zuo, C.M. Zhang, The Introduction of the Safety of Molten Salt Reactor. Nucl. Saf. 73-78 (2011).

X.Z. Cai, Z.M. Dai, H.J. Xu, Thorium molten salt reac- tor nuclear energy system. Physics (9), 578-590 (2016).

CONFLICT OF INTEREST The authors declare that they have no competing interests.

BIBLIOGRAPHY S.Z. Qiu, D.L. Zhang, W.X. Tai et al., Research on in- herent safety and relative key issues of a molten salt reactor. At. Energy Sci. Technol. (S1), 64-75 (2009).

B. Zhou, X.H. Yu, Y. Zou et al., Study on dynamic characteris- tics of fission products in molten salt reactor. Nucl. Sci.

Tech. , 17 (2020). R. Isermann, Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance (Springer, Berlin, 2006).

J.P. Ma, J. Jiang, Applications of fault detection and diagnosis methods in nuclear power plants: A review. Prog. Nucl. Energy (3), 255-266 (2011).

Y.C. Wu, Development and application of virtual nuclear power plant in digital society environment. Int. J. Energy Res. , 1521-1533 (2019).

J.Q. Zeng, H.X. Zhang, H.L. Gong et al., Ensemble Bayesian method for parameter distribution inference: Ap- plication to reactor physics. Nucl. Sci. Tech. , 199 (2023).

P. Costamagna, A.D. Giorgi, G. Moser et al., Data-driven tech- niques for fault diagnosis in power generation plants based on solid oxide fuel cells. Energy Convers. Manage. , 281-291 (2019).

A. Ayodeji, Y.K. Liu, Support vector ensemble for incipient fault diagnosis in nuclear plant components. Nucl. Eng. Tech- (8), 1306-1313 (2018).

Y. Zhao, F.D. Maio, E. Zio et al., Optimization of a dynamic uncertain causality graph for fault diagnosis in nuclear power plant. Nucl. Sci. Tech. , 34 (2017).

T.Z. Zhou, K.C. Yu, M.S. Cheng et al., Development and analysis of a K-nearest-neighbor-based transient identification model for molten salt reactor systems. Nucl. Tech. (11), 110604 (2023).

H.L. Zheng, X.G. Tuo, S.M. Peng et al., Determina- tion of Gamma point source efficiency based on a back- propagation neural network. Nucl. Sci. Tech. , 61 (2018).

A. Aboshosha and H. A. Hamad, Employing adaptive fuzzy computing for RCP intelligent control and fault diagnosis.

Nucl. Sci. Tech. , 138 (2023). 01288-y Y. Yao, J. Wang, M. Xie et al., A new approach for fault diagno- sis with full-scope simulator based on state information imag- ing in nuclear power plant. Ann. Nucl. Energy , 107274 (2020).

H. Wang, M. Peng, A. Ayodeji et al., Advanced fault diagnosis method for nuclear power plant based on con- volutional gated recurrent network and enhanced particle swarm optimization. Ann. Nucl. Energy , 107934 (2021).

H.A. Saeed, M.J. Peng, H. Wang et al., Novel fault diagnosis scheme utilizing deep learning networks. Prog. Nucl. Energy , 103066 (2020). convolutional neural network model for abnormality diagnosis in a nuclear power plant.

Appl. Comput. (2021). Z. Zhang, W.L. Liang, X.L. Tang et al., Loss of coolant ac- cident diagnosis for large pressurized water reactors based on long short-term memory network. Ann. Nucl. Energy 111167 (2025).

S.Y. Zhang, Z.X. Wang, H.B. Yang et al., Hformer: highly effi- cient vision transformer for low-dose CT denoising. Nucl. Sci.

Tech. , 61 (2023). Y.T. Yao, J. Wang, P.C. Long et al., Small-batch-size convolutional neural network based fault diagnosis system for nuclear energy production safety with big-data environment. Int. J. En- ergy Res. , 5841-5855 (2020).

G. Ioannou, T. Tagaris, G. Alexandridis et al., Intelligent tech- niques for anomaly detection in nuclear reactors. EPJ Web Conf. , 21011 (2021).

C.Y. Zhang, P.Y. Chen, F.L. Jiang et al., Fault diagnosis of nuclear power plant based on sparrow search algorithm op- timized CNN-LSTM neural network. Energies (6), 2934 (2023).

K. He, X. Zhang, S. Ren et al., Identity mappings in deep resid- ual networks. In Computer Vision -ECCV 2016 , ed. by B.

Leibe, J. Matas, N. Sebe, M. Welling. Lecture notes in com- puter science, vol. 9908 (Springer, Cham, 2016), pp.630-645.

K. He, X. Zhang, S. Ren et al., Deep residual learning for image recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) , Las Vegas, NV, USA, 2016, pp.

M. Zhao, M. Kang, B. Tang et al., Deep residual networks with dynamically weighted wavelet coefficients for fault diagnosis of planetary gearboxes. IEEE Trans. Ind. Electron. , 4290-4300 (2018).

Y. L. Sun, H. Wang, Study of diagnosis for rotating machinery in advanced nuclear reactor based on deep learning model. Front. Energy Res. , 1210703 (2023).

Y.J. Ma, Y. Ren, P. Feng et al., Sinogram denoising via at- tention residual dense convolutional neural network for low- dose computed tomography. Nucl. Sci. Tech. , 41 (2021).

M. Zhao, S. Zhong, X. Fu et al., Deep residual networks with adaptively parametric rectifier linear units for fault di- agnosis. IEEE Trans. Ind. Electron. , 2587-2597 (2021).

Wang, Adaptive residual based fault detection and diagnosis system of small mod- ular reactors. Appl. Soft Comput. 108064 (2022).

Z.Z. Zhang, H.L. Wang, H.Y. Meng et al., Uncertainty evalua- tion and correlation analysis of single-particle energies in phe- nomenological nuclear mean field: an investigation into prop- agating uncertainties for independent model parameters. Nucl.

Sci. Tech. , 16 (2021). C. Aldrich, L. Auret, Fault detection and diagnosis with random forest feature extraction and variable importance methods. IFAC Proceedings Volumes , 79-86 (2010).

D.L. Donoho, De-noising by soft-thresholding. IEEE Trans.

Inf. Theory , 613-627 (1995). A. Vaswani, N. Shazeer, N. Parmar et al., Attention is all you need. In NIPS' 17:

Proceedings of the 31st In- ternational Conference on Neural Information Processing Systems , Red Hook, NY, USA, 2017, pp. 5998-6008.

A.M. Hernández, I.V. Nieuwenhuyse, S.R. Gonzalez, A survey on multi-objective hyperparameter optimization algorithms.

Artif. Intell. Rev. , 8043-8093 (2023). K. Wang, X.W. Jiao, Q. Yang et al., The effect of scram rod drop time on the consequences of molten salt reactor reactivity insertion transient. Nucl. Tech. (9), 090606 (2020).

X.W. Jiao, K. Wang, C.Q. Wang et al., Study on sensitivity of initial conditions of reactivity initiated accident under low power conditions of molten salt reactor. Nucl. Tech. 060602 (2021).

C.Q. Wang, Q. Yang, K. Wang et al., Sensitivity anal- ysis of power related parameters in a reactivity-initiated accident of a molten salt reactor. In Proceedings of the 2021 28th International Conference on Nuclear Engineering, vol. 1 , Virtual, Online, 4-6 August 2021, V001T03A006.

C.Q. Wang, K. Wang, Q. Yang, Prediction of multi- parameters of molten salt reactor based on deep learning net- works. Nucl. Tech. (9), 090601 (2025).

R. Jian, B. Xu, M.H. Li et al., A specialized code for opera- tion transient analysis and its application in fluoride salt-cooled high-temperature reactors. Nucl. Sci. Tech. , 119 (2017).

Y.S. Song, M.S. Cheng, M. Lin et al., Development and appli- cation of multiscale thermal fluid coupling program for molten salt cooled fast reactor based on RELAP5 and sub-channel pro- gram. Nucl. Tech. (7), 070602 (2022).

Y.S. Chen, Experimental and simulation research on heat trans- fer characteristic of high temperature molten salt and molten salt heat exchanger. University of Chinese Academy of Sci- ences, 2021.

K. Wang, C.X. Cai, Z.Z. He et al., Characteristics analysis of nitrate salt natural circulation loop. Nucl. Tech. (6), 060601 (2015).

L.A. Shalabi, Z. Shaaban, B. Kasasbeh, Data mining: preprocessing engine. J. Comput. Sci. (9), 735-739 (2006).

J.D. Evans, Straightforward statistics for the behavioral sci- ences (Thomson Brooks/Cole Publishing Co., 1996).

H.Y. Chen, P.Z. Gao, S.C. Tan et al., Prediction of auto- matic scram during abnormal conditions of nuclear power plants based on long short-term memory (LSTM) and dropout. Sci. Technol. Nucl. Install. , 226376 (2023).

K. He, X. Zhang, S. Ren et al., Delving deep into recti- fiers: Surpassing human-level performance on imagenet clas- sification. In 2015 IEEE International Conference on Com-

puter Vision (ICCV) , Santiago, Chile, 2015, pp. 1026-1034.

T. Landgrebe, R. Duin, Efficient multiclass ROC approxima- tion by decomposition via confusion matrix perturbation anal- ysis. IEEE Trans. Pattern Anal. Mach. Intell. , 810-822 (2008).

H. Huang, H. Xu, X. Wang et al., Maximum F1-score discrimi- native training criterion for automatic mispronunciation detec- tion. IEEE/ACM Trans. Audio Speech Lang. Process. 787-797 (2015).

K. He, X. Zhang, S. Ren et al., Delving deep into recti- fiers: Surpassing human-level performance on ImageNet clas- sification. In 2015 IEEE International Conference on Com- puter Vision (ICCV) , Santiago, Chile, 2015, pp. 1026-1034.

X. Glorot, Y. Bengio, Understanding the difficulty of train- ing deep feedforward neural networks. In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics (AISTATS) , vol. 9, Chia Laguna Resort, Sar- dinia, Italy, 2010, pp. 249-256.

H.P. Suryawan, Gaussian white noise analysis and its applica- tion to Feynman path integral. AIP Conf. Proc. , 030001 (2016).

B. Yang, H. Xia, M. Annor-Nyarko et al., Application of total variation denoising in nuclear power plant sig- nal pre-processing. Ann. Nucl. Energy , 106981 (2020).

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

Source: ChinaXiv —Machine translation. Verify with original.