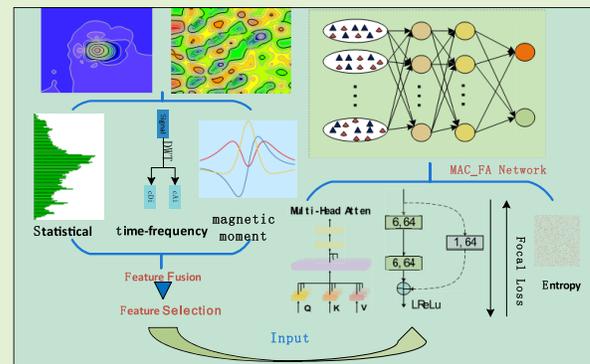


Feature Entropy Adaptive Network for Weak Magnetic Signal Classification

Ruiping Liu¹, Qing Chang, Yaoli Wang¹, and Lipo Wang², *Senior Member, IEEE*

Abstract—Magnetic anomaly signals are composed of anomaly signal and the geomagnetic field. Due to the similarity in magnitude between these two types of signals and the difficulty in acquiring magnetic field data, distinguishing between them is challenging, and the available dataset is small. This article aims to address the classification of weak magnetic signals with limited samples obtained from actual measurements, and a novel neural network-based approach for magnetic anomaly classification is proposed. First, the feature selection is performed on the fused magnetic field signal features. The measured magnetic signals are decomposed using the standard orthogonal basis functions (OBFs), and the coefficients of the basis functions are utilized as magnetic moment features. The wavelet transform is employed to calculate the coefficients as the time–frequency features of the magnetic field data. Statistical features are extracted based on the characteristics of the magnetic anomaly data. Using the statistical feature mean as a benchmark, selection is conducted considering the characteristics of the feature dataset, resulting in improved classification results. Afterward, a lightweight magnetic anomaly classification model, MAD_FA, was designed, resulting in an average reduction of 41.67% in training time. Focal loss was employed as the loss function during training, leading to an improvement of 3.86% in classification accuracy. A multifeature adaptive entropy weighting (MFAEW) method is proposed to extract magnetic signal features, which adaptively determines feature weights and effectively utilizes the mutual information and complementarity between features. This approach accelerates network convergence and improves the classification accuracy by 2.53%. Finally, a comparison was made between the MAD_FA model and classical signal classification models, and a series of ablation experiments were conducted to evaluate the model. The suggested technique performs well in the task of classifying weak magnetic signals, with a classification accuracy of 99.96%, an F1 score of 96.38%, and an AUC score of 99.12%. The higher classification accuracy and stronger robustness compared to the traditional methods demonstrate the potential application of the magnetic anomaly classification model with feature adaptation (MAC_FA) model in weak magnetic signal classification tasks.

Index Terms—Magnetic anomaly classification model with feature adaptation (MAC_FA) network, multifeature adaptive entropy weighting (MFAEW), orthogonal basis function (OBF), weak magnetic classification.



I. INTRODUCTION

MAGNETIC anomalies refer to magnetic field disturbances that exist in subsurface or underwater media in the field of magnetism. These disturbances can originate

Manuscript received 10 August 2023; revised 9 October 2023; accepted 17 October 2023. Date of publication 26 October 2023; date of current version 12 January 2024. This work was supported in part by the Overseas Chinese Scholars and Scholars in Hong Kong and Macao under Grant 61828601, in part by the Natural Science Foundation of Shanxi Province under Grant 201801D121141, and in part by the Provincial Program on Key Research Projects of Shanxi (Social Development Area) under Grant 201903D321003. The associate editor coordinating the review of this article and approving it for publication was Prof. Pai-Yen Chen. (Corresponding author: Yaoli Wang.)

Ruiping Liu, Qing Chang, and Yaoli Wang are with the College of Information and Computer, Taiyuan University of Technology, Taiyuan 030024, China (e-mail: 2423912204@qq.com; tychangqing@126.com; wangyaoli@tyut.edu.cn).

Lipo Wang is with the Electrical and Electronic Engineering Department, Nanyang Technological University, Singapore 639798 (e-mail: elpwang@ntu.edu.sg).

Digital Object Identifier 10.1109/JSEN.2023.3326138

from underground mineral deposits, rock structures, geological changes, and other factors. Different types of magnetic anomalies exhibit distinct characteristics and generation mechanisms. Magnetic anomaly classification finds extensive applications in mineral exploration, geological hazard monitoring, and other related domains [1], [2], [3]. Automating the classification of various magnetic anomalies can enhance exploration efficiency, reduce exploration costs, and mitigate geological hazard risks, among other benefits.

In the past few years, significant progress has been made in the classification of magnetic anomalies, primarily in the following areas.

- 1) *Optimization of Network Architecture* [4], [5]: Researchers have continuously improved the basic neural network architecture based on the characteristics of magnetic anomaly data, resulting in enhanced feature extraction capabilities and classification performance.

- 2) *Use of Feature Fusion Techniques [6], [7]*: In order to improve the classification performance and address the issue of multidimensional features in magnetic anomaly data, researchers have proposed a number of feature fusion techniques. Convolutional neural networks (CNNs) [8] are a neural network structure capable of automatically extracting features from data. Through computational load on neural networks, multilayer convolution, and pooling operations, higher level features can be extracted from the original data. When processing magnetic anomaly data, variables, such as magnetic field components in various directions, geological lithology, and topography, are used as separate inputs to the CNN network, and their corresponding features are extracted using convolution and pooling operations. Finally, these features are fused to produce a more complete and accurate feature representation of magnetic anomaly data.
- 3) *Use of Data Enhancement Techniques [9]*: Researchers addressed the issue of limited magnetic anomaly data by employing a transfer learning approach. They fine-tuned a pretrained model on a magnetic anomaly dataset, leading to a significant improvement in classification performance.

Significant progress has been made in using neural networks for magnetic anomaly classification. However, several challenges and issues persist, including data imbalance and limited sample sizes, necessitating further research and exploration. Fan et al. [10] proposed an adaptive magnetic anomaly detection method based on support vector machine (SVM). Although this method demonstrated satisfactory performance in experiments, its classification accuracy may be limited due to the sole application of SVM in magnetic anomaly detection without fully leveraging existing deep learning techniques. This might result in an insufficient exploration of latent information within magnetic anomaly data. Jiaqi et al. [11] employed the wavelet packet method to denoise the measured gradient signals and utilized the orthogonal basis function (OBF) approach to decompose the gradient signals into seven components. This method effectively enhances the signal-to-noise ratio. However, it exhibits significant performance fluctuations across different datasets, possibly due to the algorithm's inherent instability. Du et al. [12] proposed a novel feature extraction method based on singular spectrum analysis (SSA) and OBF, which enables a better representation of both local and global characteristics of magnetic anomaly signals. However, this approach involves significant computational overhead due to the requirement of performing SSA and OBF decomposition during the data preprocessing stage. Zhang et al. [13] proposed a magnetic anomaly detection method based on feature fusion and the isolation forest algorithm. This approach considers multiscale features, allowing for more effective capture of magnetic anomaly signal characteristics. However, in the process of anomaly detection using the isolation forest algorithm, it is necessary to adjust certain parameters, which could potentially impact the final detection results. Hu et al. [14] employed a CNN for detecting magnetic dipole target signals, but it was only

applicable to the detection of magnetic dipole target signals and not suitable for other types of magnetic anomaly data.

In this article, small ferromagnetic blocks are regarded as the source of magnetic anomalies. Magnetic anomaly data are obtained using a three-axis fluxgate sensor and measured in an environment with minimal interference. These data are then used as a positive sample for magnetic anomaly classification. Geomagnetic field measurements made in the same environment are used as negative samples. In order to provide a more comprehensive and multiperspective characterization of the data, to overcome data limitations, to reduce the impact of human intervention, and to reduce the computational effort of the neural network, this article makes the following contributions.

- 1) Multiangle extraction of magnetic field signal features, integrating magnetic moment characteristics, time-frequency characteristics, and statistical characteristics, and using the selected features as inputs to the magnetic anomaly classification network to provide comprehensive data representation, overcoming data limitations.
- 2) To adaptively determine network weights and fully utilize the mutual information and complementarity among features, a multifeature adaptive entropy weighting (MFAEW) method is proposed for network weight update based on fused features.
- 3) Proposing a lightweight MAD_FA magnetic anomaly classification model that combines residual modules with the multihead attention mechanism from Transformer to enhance the model's feature extraction capability. Furthermore, introducing the focal loss function to address the distinctive characteristics of magnetic anomaly data and improve the network's discriminative power, thereby enhancing the accuracy and robustness of the model.

II. FEATURE EXTRACTION

After preprocessing the collected magnetic field data, this section focuses on the initial extraction of magnetic signal characteristics from three aspects.

A. Dataset Construction

The field measurement system for magnetic field detection consists of a data acquisition module, a magnetic sensor module, and upper level software. The data acquisition module collects data and communicates with the upper level software via Ethernet communication, allowing for visual and intuitive observation of the actual magnetic field measurement situation. In addition, the upper level software integrates data storage functionality, facilitating subsequent processing of the measured magnetic field data.

Finding a location with relatively low interference, a three-axis magnetic flux gate sensor is employed to acquire magnetic field data at a sampling frequency of 200 Hz. The magnetic anomaly data are processed using point sampling, while the geomagnetic field is measured using a fixed-point measurement approach. The measured data are organized into groups

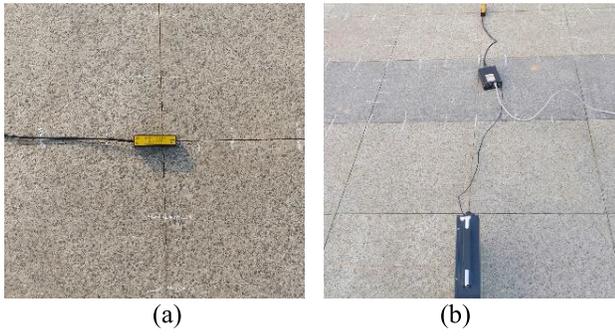


Fig. 1. (a) Three-axis fluxgate sensor. (b) Measurement device.

of 200 points each in chronological order, and corresponding features are computed. Subsequently, these datasets are partitioned into a training set and a validation set using a 7:3 ratio. Fig. 1 shows a three-axis fluxgate sensor and the measured field site.

B. Magnetic Moment Characteristics

The OBF decomposition [15] technique can be used to extract magnetic anomaly signals from geomagnetic field data by decomposing magnetic anomaly data. The algorithm's primary concept is to describe the geomagnetic field signal as a linear combination of a number of OBFs, where each basis function denotes a particular type of geomagnetic field structure. When there is a great distance between the sensor and the target, a straightforward dipole model is typically adequate. The target under test can be regarded as a magnetic dipole when the observation distance is more than 2.5 times the magnetic target size [16], and the magnetic field B created by the dipole at a distance r from the sensor with a moment M is

$$B = \frac{\mu_0}{4\pi r^3} \left[\frac{3(M \cdot r)r}{r^2} - M \right]. \quad (1)$$

$\mu_0 = 4\pi \times 10^{-7} H/m$ is the magnetic permeability of space.

When the magnetic anomaly target satisfies the magnetic dipole model and the relative motion between the target and the sensor is uniform or uniformly variable speed, the OBF magnetic anomaly detection model is established as shown in Fig. 2, and the model parameters are shown in Table I.

When the local magnetic field G is much larger than the ferromagnetic field B , the signal S measured by the fluxgate sensor is considered as the projection of B onto G

$$S = \frac{B \cdot G}{|G|}. \quad (2)$$

Bringing (1) into (2) gives

$$S = \frac{\mu_0 M}{4\pi R_0^3} \sum_{n=1}^4 b_n \varphi_n(w) \quad (3)$$

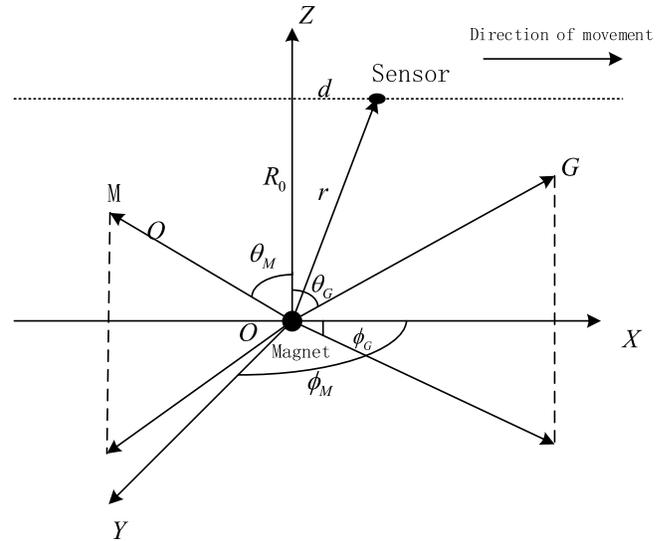


Fig. 2. Magnetic anomaly model detection map.

TABLE I
PARAMETER DESCRIPTION

Parameters	Explain
M	Target magnetic moment
G	Geomagnetic field
R_0	Minimum distance between the sensor and the target
φ_M	The Angle between the X-axis and the projection of M onto the XY plane
φ_G	The Angle between the X-axis and the projection of G onto the XY plane
θ_M	The Angle between M the Z axis
θ_G	The Angle between G the Z axis
τ	Distance between target and sensor
d	Distance between sensor and YOZ plane

where

$$\begin{cases} \varphi_1(w) = \frac{w^2}{[1+w^2]^{5/2}} \\ \varphi_2(w) = \frac{1}{[1+w^2]^{5/2}} \\ \varphi_3(w) = \frac{1}{[1+w^2]^{5/2}} \\ \varphi_4(w) = \frac{1}{[1+w^2]^{3/2}} \end{cases} \quad (4)$$

$$\begin{cases} b_1 = 3 \sin \theta_M \cos \phi_M \sin \theta_G \cos \phi_G \\ b_2 = 3(\cos \theta_M \sin \theta_G \cos \phi_G + \sin \theta_M \cos \phi_M \cos \theta_G) \\ b_3 = 3 \cos \theta_M \cos \theta_G \\ b_4 = -\sin \theta_M \sin \theta_G \cos \phi_M - \phi_G - \cos \theta_M \cos \theta_G \end{cases} \quad (5)$$

b_j is the coefficient of the basis function and w is defined as $w = (D/R_0)$. It can be easily demonstrated that $\varphi_4(w) = \varphi_1(w) + \varphi_3(w)$. The remaining three functions $\varphi_1(w)$ – $\varphi_3(w)$ are linearly independent, which is used as a triplet of standard

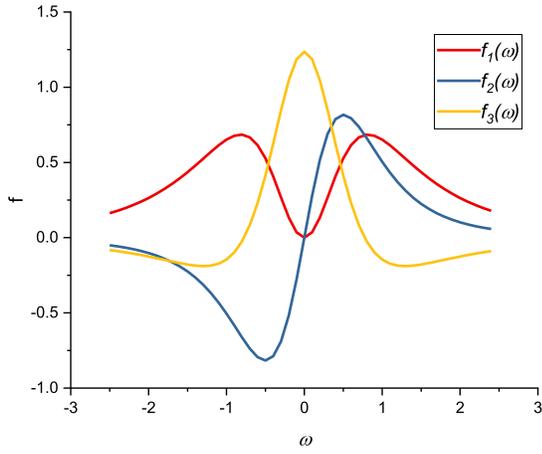


Fig. 3. Standard OBFs.

orthogonal bases to represent S

$$\begin{cases} f_1(w) = \varphi_1(w) \sqrt{\frac{128}{3\pi}} \\ f_2(w) = \varphi_2(w) \sqrt{\frac{128}{5\pi}} \\ f_3(w) = \left[\varphi_3(w) - \left(\frac{5}{3}\right) \varphi_1(w) \right] \sqrt{\frac{24}{5\pi}} \end{cases}$$

and meets

$$\begin{cases} \int_{-\infty}^{+\infty} f_i(w) f_j(w) dw = 0, i \neq j \\ \int_{-\infty}^{+\infty} f_j^2(w) dw = 1, i, j = 1, 2, 3. \end{cases}$$

In this case, $S = (\mu_0 M / 4\pi R_0^3) \sum_{n=1}^3 a_n f_n(w)$, where

$$\begin{cases} a_1 = \sqrt{\frac{3\pi}{128}} \left[(b_1 + b_4) + \frac{5}{3}(b_3 + b_4) \right] \\ a_2 = b_2 \sqrt{\frac{5\pi}{128}} \\ a_3 = (b_3 + b_4) \sqrt{\frac{5\pi}{24}}. \end{cases}$$

The standard orthogonal bases $f_1(w)$ – $f_3(w)$ are shown in Fig. 3.

The coefficients of basis functions a_1 – a_3 calculated by the OBF algorithm are input to the neural network for anomaly classification as magnetic moment of the magnetic field data.

C. Time–Frequency Characteristics

A signal can be divided by the wavelet transform into a number of smaller signals of various dimensions, each of which retains elements of the original signal at various frequencies and can be used to extract features. The signal is gradually downsampled and divided into a low-frequency and a high-frequency signal in the discrete wavelet transform (DWT) [17], [18]. The difference between these two signals is referred to as the wavelet coefficients. In this article, the magnetic field signal is decomposed into multiple frequency bands by db4 wavelet transform, and then, the features of

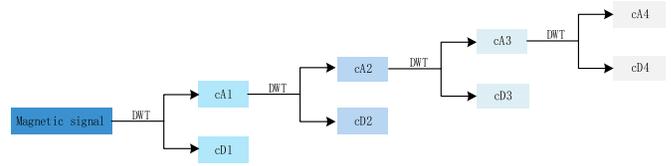


Fig. 4. Extraction of magnetic field features using wavelet transform.

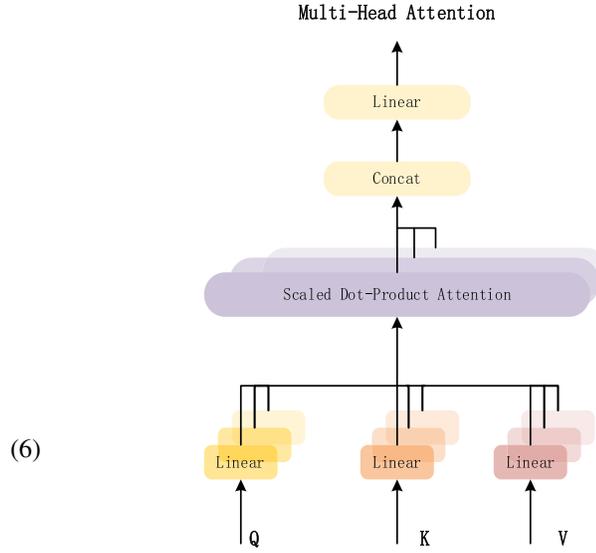


Fig. 5. Multiheaded attention mechanism architecture.

(6) each band are extracted to describe the local properties of the magnetic field signal. The process of extracting magnetic field features using wavelet transform is shown in Fig. 4, where cAi denotes the approximate coefficients of layer i and cDi denotes the detail coefficients of layer i . The extracted wavelet coefficients are input to the neural network as time–frequency features for anomaly classification.

D. Statistical Characteristics

Statistical characterization of magnetic field data is a method used to describe the distribution of data and can provide information about the dataset, including the central location of the data, the dispersion of the data, and the skewness of the data. For the characteristics of the measured data, the mean, standard deviation, root mean square, and variance are mainly selected as statistical features in this article

$$\begin{cases} \beta_1 = \frac{1}{N} \sum_{i=1}^N x_i \\ \beta_2 = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \\ \beta_3 = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_n^2}{n}} \\ \beta_4 = \frac{1}{N-1} \sum_{i=1}^N x_i^2 - N\beta_1^2. \end{cases}$$

The statistical features obtained from the above four calculations are input to the neural network for anomaly classification.

III. NETWORK FOR CLASSIFICATION OF MAGNETIC ANOMALIES DESIGNED

A. Implementation Details

1) *MFAEW Method*: The proposed method in this article is the MFAEW [19], [20], which evaluates the information content and importance of each feature and comprehensively considers the relative importance of different features. This approach provides a more accurate reflection of the contributions of different features in the magnetic field data. Specifically, the MFAEW method specifically turns the problem of allocating weights to features into a problem of minimizing entropy by using entropy as an uncertainty measure for multiple features. The entropy value of each indicator is determined throughout the neural network training iteration using MFAEW, which also updates the weight assignment until the convergence condition is met. The MFAEW calculation process is shown in Algorithm 1.

Algorithm 1 Computing Weights Using MFAEW

Require: Characteristics y_i

Ensure: Model weights w_i

for each feature y_i **do**

 Use y_i to compute model parameters

 Compute the information entropy of model parameters using Eq. (10)

 Calculate the utility value of the model parameters using Eq.(11)

 Normalize the information to obtain the weight using Eq.(12)

end for

Return w_i

The formulas involved in the algorithm are given as follows:

$$f_i = \frac{\sum_{i=1}^n y_i \ln y_i}{\ln n} \quad (10)$$

$$E_i = 1 - f_i \quad (11)$$

$$w_i = \frac{E_i}{\sum_{i=1}^n E_i}. \quad (12)$$

2) *Focal Loss Function*: In the magnetic anomaly classification task, normal samples tend to be in the majority and magnetic anomaly samples in the minority, which leads to a tendency for the network to favor the majority category and ignore the minority category during training. In addition, because the weak magnetic anomaly signal is insignificant in comparison to the geomagnetic field and the magnetic anomaly signal is a combination of both, it is more challenging to distinguish the anomaly signal from the geomagnetic field signal. Based on the above two considerations, the selection of the network loss function becomes crucial for the classification results. Through repeated screening and comparison, focal loss [21], [22] is introduced into the classification task as the loss function of the model in this article.

Focal loss aims to enhance the neural network's focus on challenging minority classes that are difficult to classify while reducing the loss associated with normal samples, thereby enabling more accurate learning and classification of rare

abnormal samples. It extends the conventional cross-entropy loss by introducing a modulating factor, which downweights easily classifiable samples, thus elevating the attention given to misclassified instances and ultimately improving the model's classification capability

$$FL(p_i) = -\alpha_i(1 - p_i)^\gamma \log(p_i). \quad (13)$$

The loss contribution of easily discernible samples is decreased using the modulation factor $(1 - p_i)^\gamma$. The modulation factor is smaller and the sample is easier to distinguish the bigger the value of p_i . The $[0, 5]$ parameter γ is used to regulate the sample imbalance issue. The focal loss changes to a cross-entropy loss function when γ is 0. The ratio of positive to negative sample loss is managed by the constant α_i . The values of these two parameters affect each other and should be used in combination during practical applications.

3) *Multiheaded Attention Mechanism*: One of the most important features in the Transformer [23] design is the multi-headed attention mechanism, which is an enhanced algorithm of the self-attentive mechanism. By including it into the magnetic anomaly classification network, the model is able to learn several feature representations, which enhances the model's representational capabilities and helps the model learn the relationships between geological features. Furthermore, the model's ability to collect geological features is improved by the multihead attention mechanism, which can simultaneously learn features at many scales and locations. The multiheaded attention mechanism is shown in Fig. 5. The expression is given as follows:

$$\text{Attention}(Q_i, K, V) = \text{softmax} (v_i * [v_1^T, v_2^T, \dots, v_n^T]) * \begin{bmatrix} v_1^T \\ v_2^T \\ \dots \\ v_n^T \end{bmatrix} = \text{softmax} (Q_i K^T) V. \quad (14)$$

Softmax is a normalized exponential function. Q , K , and V are query, key, and value, respectively. They can be obtained from the same sequence or they can be different sequences with practical meaning. In Pytorch, the multihead attention method requires two parameters, namely, the embedding dimension `embed_dim` and the number of heads `num_heads`. The settings of these two parameters in this article are given as follows:

$$\begin{aligned} \text{self.attention} &= \text{nn.MultiheadAttention} \\ (\text{embed_dim} = 16, \text{num_heads} = 4). \end{aligned} \quad (15)$$

In general, the multihead attention mechanism can be divided into the following three steps.

- 1) *Linear Transformation*: The input data are transformed linearly to produce various feature subspaces.
- 2) *Attention Calculation*: To acquire various attention weights, attention computation is carried out on various feature subspaces.
- 3) *Weighted Summation*: To produce the final multiheaded attention output, the outputs from various feature subspaces are weighted and added in accordance with various attention weights.

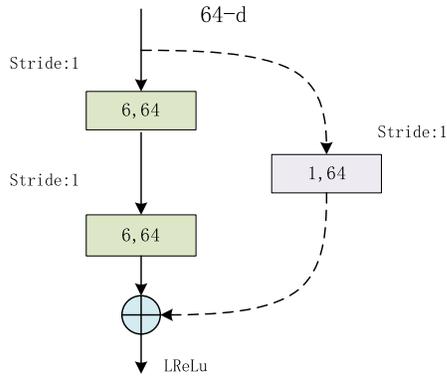


Fig. 6. Residual structure.

4) *Residuals Module*: The residual module is a commonly used component in CNNs [24], comprising a main pathway and a skip connection. The main path consists of a series of operations such as convolution, batch normalization, and activation functions for extracting the features of the input. Direct information transfer from input to output is therefore made possible by the cross-layer link, which transmits the input data directly to the main path's output. The residual term can be added directly to the output of the main path, thus enabling the process of residual learning. In this study, the residual structure is created as shown in Fig. 6 in accordance with the properties of the measured data.

B. Magnetic Anomaly Classification Model With Feature Adaptation (MAC_FA) Network for Magnetic Anomaly Classification

Based on the aforementioned design details, we propose the MAC_FA architecture for magnetic anomaly classification. The MAC_FA model consists of three crucial modules and an important method for updating weights, including the multi-head attention module, residual module, focal loss function, and MFAEW method. The MAC_FA architecture is shown in Fig. 7.

IV. EXPERIMENTAL RESULTS

In this section, we present the experimental results of the MAD_FA model on a real magnetic anomaly dataset. The performance of the MAD_FA model is compared against other state-of-the-art models, including 1DCNN [25], FCDNET [26], and CNN-LSTM [27]. In addition, we compare MFAEW with other commonly used weight initialization methods. To provide a more comprehensive and in-depth insight, we also conduct ablation experiments to analyze the contributions of each module and method within the MAD_FA model.

A. Experimental Procedure

Based on the discussion in Section IV, the process of magnetic anomaly classification is summarized as Algorithm 2.

B. Data Acquisition and Preprocessing

When the measurement environment satisfies the magnetic dipole model requirements, utilize the Kriging [28] interpolation algorithm to generate a magnetic field distribution map.

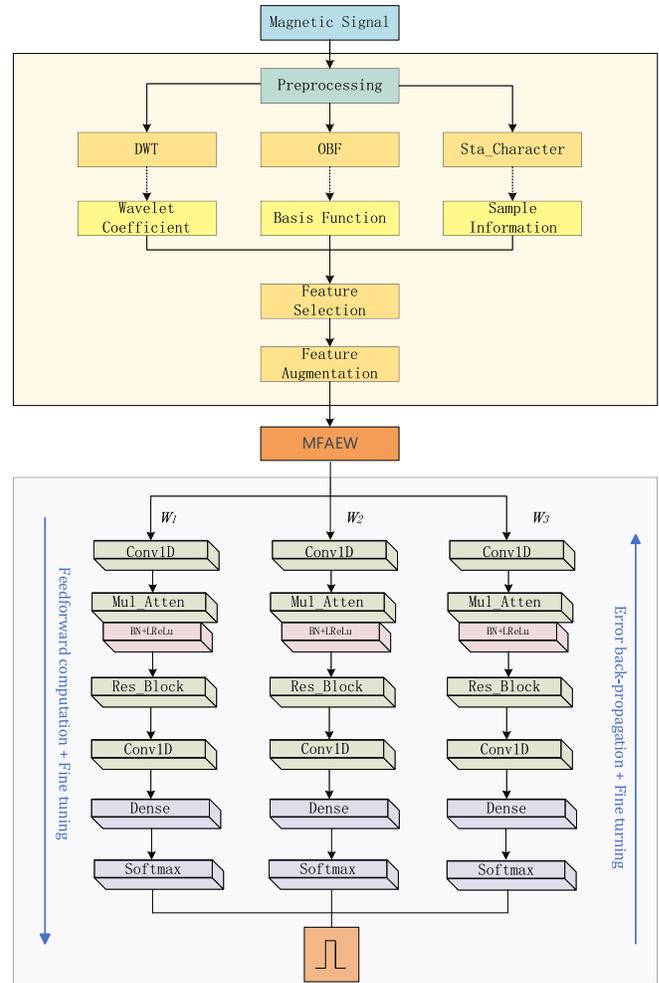


Fig. 7. Overall architecture of the MAC_FA model.

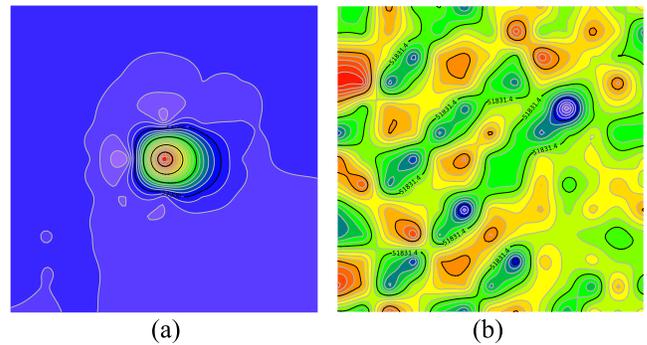


Fig. 8. (a) Magnetic anomaly. (b) Geomagnetic field.

Fig. 8 shows the magnetic field distribution map obtained using the Kriging interpolation algorithm. This map provides a visual representation of the spatial distribution of the magnetic field in the measurement environment, allowing us to observe the intensity variations of the magnetic field.

To gain a more detailed understanding of the distribution of measurement data, a subset of data points was visually presented in Fig. 9. These data points represent the actual measured results, showcasing their positions and corresponding magnetic field values within the measurement environment.

Algorithm 2 Experimental Procedure

Step 1: Measure anomalous data and geomagnetic field data using fluxgate magnetometer sensors.

```
data = collect_magnetic_data()
```

Step 2: Compute the corresponding features of the preprocessed magnetic field data.

```
data = preprocessed(data)
```

```
magnetic_moment_fea = calculate_magnetic_moment_fea(data)
```

```
time_frequency_fea = calculate_time_frequency_fea(data)
```

```
statistical_fea = calculate_statistical_fea(data)
```

Step 3: Use statistical features as filtering criteria to clean the fused features.

```
base = statistical_features(average_value)
```

```
claened_fea = fea_select(base)
```

Step 4: Design a magnetic anomaly classification network based on the magnetic field data features and use the cleaned characteristic data as input to complete the magnetic anomaly classification.

```
model = magnetic_anomaly_classification_net(claened_fea)
```

Step 5: Analyze experimental results and conduct ablation experiments on the magnetic anomaly classification network.

```
analysis_results(predicted_labels)
```

```
perform_ablation_experiment(model)
```

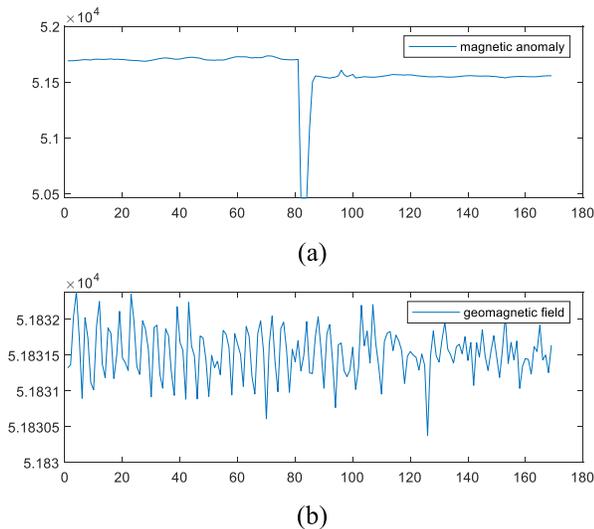


Fig. 9. (a) Measured magnetic anomaly signals. (b) Measured geomagnetic field signals.

The horizontal axis represents the measurement points of the magnetic field, while the vertical axis represents the magnetic field strength at those points. Such visual representation enables observation and analysis of the spatial characteristics and magnetic field variations.

1) *Feature Selection:* Considering the inherent difficulty in acquiring weak magnetic signals through direct measurement, while the acquisition of geomagnetic signals is relatively unconstrained, the data obtained from the three-axis fluxgate magnetometer sensors primarily consists of geomagnetic data,

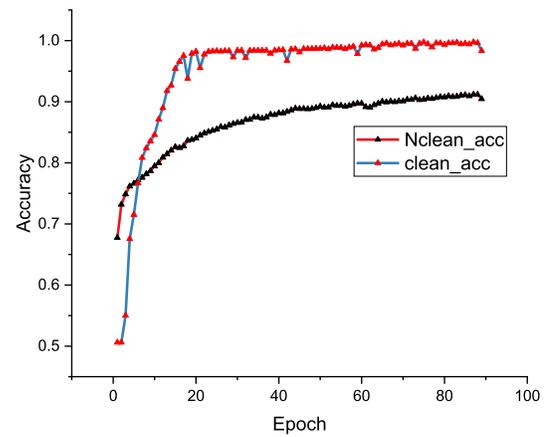


Fig. 10. Comparison of classification results before and after feature selection.

far exceeding the amount of anomalous data. The classification results before and after feature selection are shown in Fig. 10.

Filtering Criterion: The average value of statistical features is used as the cleaning criterion. The measured geomagnetic field results are predominantly in the range of 50 000–53 000 nT. The anomaly data vary in magnitude, with the highest occurrence observed around 52 000 nT.

Feature Filtering: Remove $k\%$ of the average values in the geomagnetic field that are centered around 52 000 nT, as well as $(1 - k)\%$ of the average values in the anomaly data ranging from 50 000 to 51 000 nT and 53 000 nT.

Reintegration: Utilize the cleaned data as input for the magnetic anomaly classification model.

Adjusting the feature selection ratio to 7:3, i.e., $k = 7$. From the comparison graph, it is evident that the model after feature selection exhibits faster convergence and higher classification accuracy. The selected features effectively capture the distinctive characteristics of magnetic anomaly signals, reducing the interference of redundant information. These results demonstrate the crucial role of feature selection.

2) *Feature Augmentation:* Feature enhancement has a positive impact on the robustness and overfitting resistance of the model. Considering the characteristics of the real-world measured data, it can be observed that the measured data already contain a certain level of noise. Therefore, the focus of data enhancement can be directed toward enhancing the extracted magnetic field features rather than the original measured data. The results of adding noise to the features are shown in Fig. 11.

The experimental results shown in Fig. 11 indicate that although the classification accuracy remains relatively similar after adding Gaussian noise with different variances, the model exhibits the fastest convergence speed and shows no oscillation during training when the extracted features are augmented with Gaussian noise having a variance of 0.1 and a mean of 0. Based on this observation, we decide to enhance the extracted features before conducting anomaly classification.

C. Comparative Experiments and Results

To ensure comparability, we conducted the experiments under the same experimental conditions. We used three

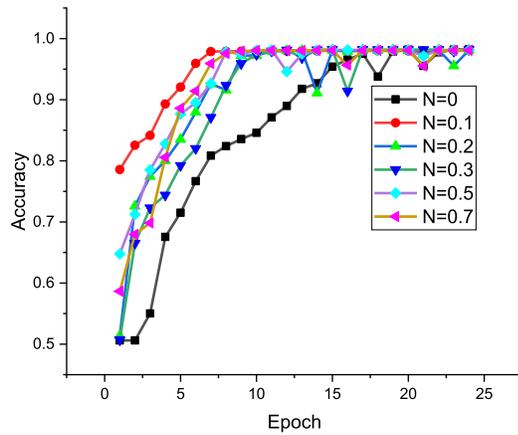


Fig. 11. Addition of Gaussian noise to features.

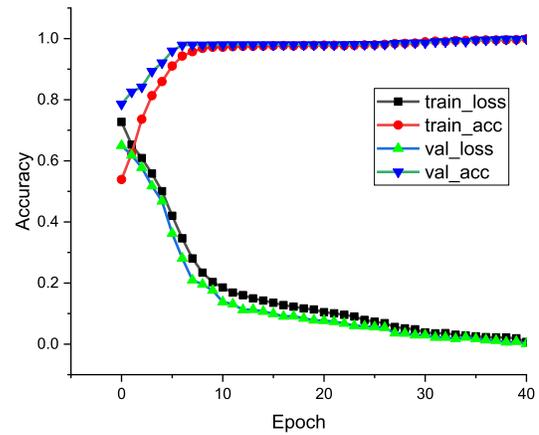


Fig. 12. Training results of the MAD_FA model.

TABLE II
COMPARISON OF EXPERIMENTAL RESULTS

Method	Accuracy	Model size(bit)	Total Parameter	Training time(s)
IDCNN	92.17	81032	20258	405
FCDNET	93.08	298760	74690	524
CNN-LSTM	95.45	401416	100354	857
MAD_FA(ours)	99.96	26184	6546	317

TABLE III
MODEL EVALUATION RESULTS UNDER DIFFERENT NOISE CONDITIONS

Noise Level	Accuracy	F1	AUC
0	98.83	95.24	99.02
0.1	99.96	96.38	99.12
0.2	98.12	89.18	95.43
0.3	98.20	91.37	96.56
0.5	98.14	89.40	94.25
0.7	98.08	92.75	96.61

crucial features as inputs for the magnetic anomaly classification network and evaluated the results of different classification methods. All experiments are performed on a computer equipped with Intel Xeon Gold 6230R CPU*2, RTX 3090 GPU*2, and 256 GB RAM. The experimental results, presented in Table II, illustrate the performance of each classification method in magnetic anomaly classification tasks.

Based on the comparative results, the proposed MAD_FA magnetic anomaly classification model achieves a classification accuracy of 99.96%, with only 6546 parameters in the entire network model. The average training time has been reduced by 41.67%. In contrast, other models achieve classification accuracies ranging from 90% to 96%. CNN-LSTM is one of the relatively better-performing network architectures in the comparative experiments and shows some effectiveness in handling magnetic anomaly classification. However, compared to our MAD_FA model, it has 93.48% more parameters, resulting in higher computational resource requirements. Next, we employ the early stopping mechanism to monitor the classification metrics on the validation set. We set the stopping condition as when the metric change on the validation set for ten consecutive epochs is less than 0.001, indicating model convergence. The magnetic anomaly data training results using the MAD_FA model are shown in Fig. 12.

The gradually increasing accuracy and decreasing loss in the training results indicate that the model is learning to make more accurate predictions. As the model trains, the MFAEW method adjusts the weights of the parameters to minimize the loss function and improve accuracy, and the convergence speed is very fast, reaching about 97% classification accuracy in the first ten epochs. With increasing epochs, the MAD_FA

model did not exhibit overfitting, oscillation, or other issues, which fully affirm the stability and reliability of the model. However, it should be noted that improving accuracy does not necessarily guarantee good model performance, and other metrics such as F1 score, area under the curve (AUC) [29], or receiver operating characteristic (ROC) [30], should be used to evaluate the overall effectiveness of the model

$$\begin{cases} \text{Precision} = \frac{TP}{TP + FP} \\ \text{Recall} = \frac{TP}{TP + FN} \\ \frac{2}{F_1} = \frac{1}{\text{Precision}} + \frac{1}{\text{Recall}} \end{cases} \quad (16)$$

in which TP represents the number of true positives, TN represents the number of true negatives, FP represents the number of false positives, and FN represents the number of false negatives. Precision is the ratio of true positives to all predicted positives, and recall is the ratio of true positives to all actual positives. The F1 score is the harmonic mean of precision and recall, and a higher value indicates better performance. The AUC is a measure of the classifier's ability to rank and distinguish, while the ROC curve is a curve plotted with FPR as the horizontal axis and TPR as the vertical axis. The closer the ROC curve is to the top-left corner, the better the performance of the classifier, which can help us balance the classifier's performance and threshold. Therefore, these metrics are very useful in evaluating the performance of classifiers. Model evaluation results under different noise conditions are presented in Table III.

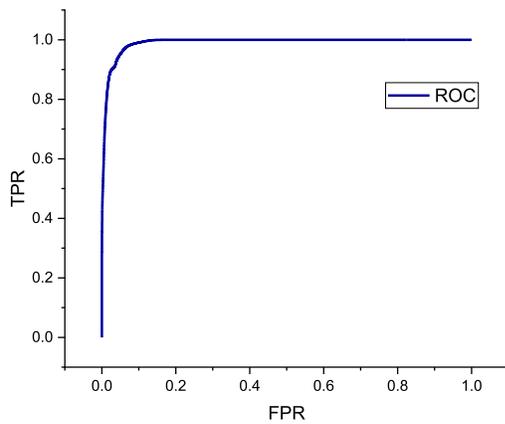


Fig. 13. ROC curve of the MAC_FA network.

Based on the results in Table III, it can be observed that the model performs best when the features are augmented with Gaussian noise with a variance of 0.1. This indicates that the model achieves optimal performance in the classification task at this noise level. Therefore, in this study, the set of features that exhibit the fastest convergence, the most stable model, and the best evaluation results is selected as the input for the magnetic anomaly classification network. The corresponding ROC results of the MAD_FA model are shown in Fig. 13.

According to the model evaluation results, the F1 score and AUC score have achieved good performance, indicating that the classifier has high accuracy in predicting positive and negative samples, and the model has high stability and reliability, with good performance in the classification task. The ROC curve is very close to the top-left corner, indicating that the MAC_FA model has good classification performance, and the probability of misclassifying negative samples while correctly classifying positive samples is small, indicating that the model has high accuracy and low error rate. Therefore, it can be concluded that the MAC_FA model established in this study has achieved excellent results in magnetic anomaly classification tasks.

D. Sensitivity Analysis

Rigorous experiments to evaluate the impact of individual components on the performance of a model are essential. These ablation experiments involve systematically removing specific components of the model to evaluate their impact on overall performance. In the present study of magnetic anomaly classification, we performed ablation experiments to evaluate the contribution of individual components to the performance of the model. Table IV presents the results of the ablation experiments.

Through ablation experiments, Fig. 14 visually shows the significant impact of certain components of the model on the final classification performance. In the third group of ablation experiments in Table IV, we discovered that using the focal loss as the loss function for backpropagation, while simultaneously increasing the weight of difficult samples, led to a remarkable 3.86% improvement in the classification accuracy of the MAD_FA model. This highlights the crucial importance of the choice of the loss function for the magnetic

TABLE IV
RESULTS OF ABLATION EXPERIMENTS

Test	Accuracy	F1 Score	AUC
Atten+FL+Res	97.43	92.99	97.03
MFAEW+Atten+FL	97.62	90.75	95.86
MFAEW+Atten+Res	96.10	93.84	97.26
MFAEW+Res+FL	97.87	92.17	96.63
Ours	98.83	95.24	99.02

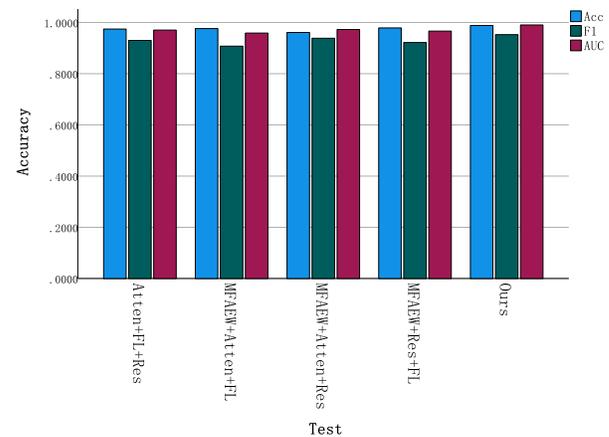


Fig. 14. Comparison of ablation experiments.

TABLE V
COMPARISON OF DIFFERENT WEIGHT INITIALIZATIONS

Module	Accuracy	F1 Score	AUC Score
Xavier[31]	97.70	91.27	94.45
He[32]	98.20	88.05	93.86
Lecun[33]	97.85	89.29	94.33
MFAEW(ours)	99.96	96.38	99.12

anomaly classification task. In addition, in the second group of ablation experiments, it is evident that the residual module has a substantial overall influence on the model, resulting in an average increase of 4.57% in F1 score and AUC score, which significantly impacts the stability of the model. Furthermore, this article compares the proposed MFAEW method with other weight initialization techniques. The results are shown in Table V.

The text evaluated three commonly used weight initialization methods, namely, Xavier, He, and Lecun, along with MFAEW for a magnetic anomaly classification task. In terms of accuracy, MFAEW initialization achieved the highest result with 99.96%, followed closely by He and Lecun initializations with accuracies of 98.20% and 97.85%, respectively. However, when considering the F1 score, MFAEW initialization significantly outperformed the other methods, obtaining a high score of 96.38%. In addition, in terms of the AUC score, MFAEW initialization again exhibited outstanding performance with a score of 99.12%, while the AUC scores of the other three

TABLE VI
MFAEW WEIGHT CALCULATION RESULTS

Layer	Conv1	Conv2	Residual	Attention	Dense
Entropy	26.98	37.98	68.06	91.20	26.91
MFAEW	0.11	0.15	0.27	0.36	0.11

methods did not exceed 95%. This indicates that MFAEW initialization is better suited for the magnetic anomaly classification task. Table VI presents the final weight results obtained using the MFAEW approach.

V. CONCLUSION

In this study, a multifeature fusion method is employed for magnetic anomaly classification tasks, utilizing filtered magnetic moment features, time–frequency features, and statistical features to obtain a comprehensive and accurate data representation. The lightweight magnetic anomaly classification model, MAD_FA, is designed specifically for data feature selection, resulting in an average reduction of 41.67% in training time. By analyzing the role of each module and making targeted adjustments, outstanding classification results are achieved.

Furthermore, the combination of focal loss and MFAEW methods effectively captures salient features in magnetic anomaly data. Increasing the network’s focus on challenging samples and allocating model parameter weights based on feature entropy not only enhances the accuracy of magnetic anomaly classification to 99.96% but also plays a crucial role in stabilizing the model with 96.38% F1 score and 99.12% AUC score.

Overall, the proposed method in this study demonstrates excellent practicality and scalability for magnetic anomaly classification tasks. Future research directions may focus on exploring more effective ways of acquiring magnetic anomaly data, enhancing cross-region and cross-dataset generalization abilities, and enabling the model to exhibit strong classification performance across different regions and datasets.

REFERENCES

- W. Wang et al., “Novel coil transducer induced thermoacoustic detection of rail internal defects towards intelligent processing,” *IEEE Trans. Ind. Electron.*, vol. 71, no. 2, pp. 2100–2111, Feb. 2024.
- S. Zhao, C. Zhang, and Y. Wang, “Lithium-ion battery capacity and remaining useful life prediction using board learning system and long short-term memory neural network,” *J. Energy Storage*, vol. 52, Aug. 2022, Art. no. 104901.
- C. Zhang, S. Zhao, Z. Yang, and Y. He, “A multi-fault diagnosis method for lithium-ion battery pack using curvilinear Manhattan distance evaluation and voltage difference analysis,” *J. Energy Storage*, vol. 67, Sep. 2023, Art. no. 107575.
- Q. Chang, R. Liu, Y. Wang, and L. Wang, “High-precision magnetic field reconstruction and anomaly classification,” *IEEE Sensors J.*, vol. 23, no. 17, pp. 19163–19175, Sep. 2023.
- C. Zhang, S. Zhao, Z. Yang, and Y. Chen, “A reliable data-driven state-of-health estimation model for lithium-ion batteries in electric vehicles,” *Frontiers Energy Res.*, vol. 10, Sep. 2022, Art. no. 1013800.
- L. Wang, Y. Wang, and Q. Chang, “Feature selection methods for big data bioinformatics: A survey from the search perspective,” *Methods*, vol. 111, pp. 21–31, Dec. 2016.
- W. Wang et al., “Laser-induced surface acoustic wave sensing-based malaria parasite detection and analysis,” *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–9, 2022.
- Y. Wang, L. Wang, Q. Chang, and C. Yang, “Effects of direct input–output connections on multilayer perceptron neural networks for time series prediction: L. Wang,” *Soft Comput.*, vol. 24, no. 7, pp. 4729–4738, Apr. 2020.
- Z. Zhu, K. Lin, A. K. Jain, and J. Zhou, “Transfer learning in deep reinforcement learning: A survey,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 11, pp. 13344–13362, Nov. 2023.
- L. Fan, C. Kang, H. Wang, H. Hu, X. Zhang, and X. Liu, “Adaptive magnetic anomaly detection method using support vector machine,” *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- Z. Jiaqi, P. Genzhai, W. Chengdong, Y. Huan, Z. Zhihong, and C. Yong, “Magnetic anomaly detection based on magnetic gradient orthonormal basis function and wavelet packet,” in *Proc. 19th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process. (ICCWAMTIP)*, Dec. 2022, pp. 1–9.
- C. Du, C. Zhang, X. Peng, and H. Guo, “Magnetic anomaly detection based on singular spectrum analysis and orthonormal basis functions,” in *Proc. 7th Int. Conf. Commun., Image Signal Process. (CCISP)*, Nov. 2022, pp. 342–346.
- N. Zhang, Y. Liu, L. Xu, P. Lin, H. Zhao, and M. Chang, “Magnetic anomaly detection method based on feature fusion and isolation forest algorithm,” *IEEE Access*, vol. 10, pp. 84444–84457, 2022.
- M. Hu, S. Jing, C. Du, M. Xia, X. Peng, and H. Guo, “Magnetic dipole target signal detection via convolutional neural network,” *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- B. Ginzburg, L. Frumkis, and B.-Z. Kaplan, “Processing of magnetic scalar gradiometer signals using orthonormalized functions,” *Sens. Actuators A, Phys.*, vol. 102, nos. 1–2, pp. 67–75, Dec. 2002.
- J. J. Holmes, “Modeling a ship’s ferromagnetic signatures,” *Synth. Lectures Comput. Electromagn.*, vol. 2, no. 1, pp. 1–75, Jan. 2007.
- J. A. Aghamaleki and A. Ghorbani, “Image fusion using dual tree discrete wavelet transform and weights optimization,” *Vis. Comput.*, vol. 39, no. 3, pp. 1181–1191, 2023.
- M. Sibgatullin, L. Gilyazov, R. Achmerov, I. Plotnikova, and M. K. Salakhov, “Extraction of a seismic signal of hydraulic fracturing using discrete wavelet transform,” *Chem. Technol. Fuels Oils*, vol. 59, pp. 51–57, Apr. 2023.
- H. Yang et al., “Optimizing the cutting edge geometry of micro drill based on the entropy weight method,” *Int. J. Adv. Manuf. Technol.*, vol. 125, nos. 5–6, pp. 2673–2689, Mar. 2023.
- L. Jiayi, C. Anqi, and L. Xiangyuan, “Power load feature identification and prediction based on structural entropy weight method and improved Bayesian algorithm,” in *Proc. IEEE 5th Int. Conf. Autom., Electron. Electr. Eng. (AUTEEE)*, Nov. 2022, pp. 555–559.
- Q. Xu, Q. Cheng, C. Jiang, and X. Wang, “Industrial false signal recognition based on deep learning with focal loss,” in *Proc. Int. Conf. Cyber Secur. Intell. Anal.* Cham, Switzerland: Springer, 2022, pp. 192–199.
- J. Li, X. Wang, X. Wang, S. Qiao, and Y. Zhou, “Improving the ResNet-based respiratory sound classification systems with focal loss,” in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, Oct. 2022, pp. 223–227.
- A. Vaswani et al., “Attention is all you need,” in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–12.
- S. Zhang, Z. Liu, Y. Chen, Y. Jin, and G. Bai, “Selective kernel convolution deep residual network based on channel-spatial attention mechanism and feature fusion for mechanical fault diagnosis,” *ISA Trans.*, vol. 133, pp. 369–383, Feb. 2023.
- X. Xu, X. Li, W. Ming, and M. Chen, “A novel multi-scale CNN and attention mechanism method with multi-sensor signal for remaining useful life prediction,” *Comput. Ind. Eng.*, vol. 169, Jul. 2022, Art. no. 108204.
- H. S. Elsayed, O. M. Saad, M. S. Soliman, Y. Chen, and H. A. Youness, “Attention-based fully convolutional DenseNet for earthquake detection,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5918610.
- L. Huang, X. Hong, Z. Yang, Y. Liu, and B. Zhang, “CNN-LSTM network-based damage detection approach for copper pipeline using laser ultrasonic scanning,” *Ultrasonics*, vol. 121, Apr. 2022, Art. no. 106685.
- L. Belkhir, A. Tiri, and L. Mouni, “Spatial distribution of the groundwater quality using Kriging and co-kriging interpolations,” *Groundwater Sustain. Develop.*, vol. 11, Oct. 2020, Art. no. 100473.
- J. Myerson, L. Green, and M. Warusawitharana, “Area under the curve as a measure of discounting,” *J. Experim. Anal. Behav.*, vol. 76, no. 2, pp. 235–243, Sep. 2001.

- [30] M. S. Pepe, G. Longton, and H. Janes, "Estimation and comparison of receiver operating characteristic curves," *Stata J.*, vol. 9, no. 1, pp. 1–16, 2009.
- [31] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1026–1034.
- [32] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. 13th Int. Conf. Artif. Intell. Statist.*, 2010, pp. 249–256.
- [33] Y. LeCun et al., "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989.



Ruiping Liu is pursuing the M.Sc. degree with the Taiyuan University of Technology, Shanxi, China.

Her research interests include the detection and classification of weak magnetic signals.



Qing Chang received the M.S. and Ph.D. degrees from the Taiyuan University of Technology, Shanxi, China, in 2005 and 2011, respectively.

He is a (co-)author of over 30 papers, including six papers indexed in EI. He has received one national invention patent and has authored two monographs. His research interests include embedded systems, computer vision, and information system design theory.

Dr. Chang has served as the Principal Investigator for the following projects: "Theoretical Research on Embedded Lean Servers" funded by the Shanxi Provincial Fund and "Key Technologies of WSN-based Mine Pressure Monitoring System."



Yaoli Wang received the B.S. and M.Sc. degrees from Northwestern Polytechnical University, Shanxi, China, in 1985 and 1988, respectively, and the Ph.D. degree from the Taiyuan University of Technology, Shanxi, in 2012.

He is a (co-)author of over 70 papers. He holds two Chinese patents. His research interests include intelligent techniques with applications to communications, image/video processing, and robots.

Dr. Wang is a member of the Expert Committee on Embedded Systems of the China Electronic Society, a member of the Seventh Forest and Grassland Fire Prevention Committee of the China Forestry Society, and a member of the Seventh Forest Fire Control Branch of the China Fire Control Association.



Lipo Wang (Senior Member, IEEE) received the B.Sc. degree from the National University of Defense Technology, Changsha, China, in 1983, and the Ph.D. degree from Louisiana State University, Baton Rouge, LA, USA, in 1988.

He is a (co-)author of over 270 papers, of which more than 90 are in journals. He holds a U.S. patent in neural networks and a Chinese patent in VLSI. He has coauthored two monographs and (co-)edited 15 books. His research interests include intelligent techniques

with applications to communications, image/video processing, biomedical engineering, and data mining.

Dr. Wang was/will be a keynote/panel speaker for 21 international conferences. He is/was an associate editor/editorial board member of 30 international journals, including three IEEE Transactions, and the guest editor for ten journal special issues. He was an AdCom Member (2010–2015) of the IEEE Computational Intelligence Society (CIS) and served as a CIS Vice President for Technical Activities and the Chair of the Emergent Technologies Technical Committee. He was a member of the Board of Governors of the International Neural Network Society (2011–2016) and an AdCom Member of the IEEE Biometrics Council. He served as the Chair of the Education Committee, IEEE Engineering in Medicine and Biology Society (EMBS). He was the President of the Asia-Pacific Neural Network Assembly (APNNA) and received the APNNA Excellent Service Award. He was the Founding Chair of the EMBS Singapore Chapter and the CIS Singapore Chapter. He serves/served as a chair/committee member for over 200 international conferences.