Journal of Systems Engineering and Electronics
Vol. 37, No. 3, June 2026, pp.800 – 815

Cascaded ensemble learning for efficient and high-accuracy direction of arrival estimation

ZHENG Guimei[1,*], XIAO Liyuan[1,2], ZHENG Yu[1,2], and ZHANG Saiyu[1,2]

1. Air and Missile Defense College, Air Force Engineering University, Xi'an 710038, China;
2. Graduate School, Air Force Engineering University, Xi'an 710038, China

Abstract: Aiming at the issues where traditional direction-of-arrival (DOA) estimation algorithms experience substantial performance degradation in low signal-to-noise ratio environments, and deep learning-based DOA estimation methods rely on massive training data with prolonged model training cycles, this paper proposes two efficient and high-precision DOA estimation methods based on ensemble learning. By formulating DOA estimation as a multi-label classification problem and leveraging the classification chain paradigm, data-driven models, classification chain-random forest (CC-RF) and classification chain-eXtreme gradient boosting (CC-XGBoost), are constructed, which are capable of handling multi-label classification tasks. To verify the effectiveness of the proposed methods, a multi-dimensional comparative experiment is designed to benchmark their performance against the traditional multiple signal classification (MUSIC) algorithm and a convolutional neural network (CNN) model. Experimental results indicate that in both single-source and multi-source scenarios, the proposed CC-RF algorithm exhibits excellent performance, achieving DOA estimation accuracy comparable to the MUSIC algorithm; in multi-source estimation scenarios, both proposed models demonstrate strong noise adaptability. Compared with the traditional MUSIC and CNN algorithms, the estimation error of the CC-XGBoost and CC-RF models is reduced by up to nearly 30 times while maintaining low time complexity, with the single estimation time reduced by approximately 90% compared to traditional methods. This study provides a technical pathway for DOA estimation in complex environments and holds significant application value in fields such as radar detection and wireless communication.

Keywords: direction of arrival estimation (DOA), multi-label classification, eXtreme gradient boosting (XGBoost), random forest, classification chain.

DOI: 10.23919/JSEE.2026.000101

1. Introduction

Direction-of-arrival (DOA) estimation is a core issue in array signal processing, with wide applications in radar, sonar, wireless communication, medical imaging, and other fields. The objective of DOA estimation is to determine the spatial direction of signal sources through processing signals received by the array. With the advancement of sensor array technology and the continuous evolution of signal processing algorithms, the accuracy and robustness of DOA estimation have been significantly enhanced. Currently, DOA estimation algorithms are primarily categorized into traditional model-driven methods and emerging data-driven methods.

The traditional DOA estimation methods are mainly model driven methods represented by subspace class methods [1−5], represented by multiple signal classification (MUSIC), estimation of signal via rotational techniques (ESPRIT) and their extended algorithms, and sparse representation class methods [6−10] (Bayesian learning, greedy algorithm, etc.). The MUSIC algorithm [1] and ESPRIT algorithm [2] are representative of subspace methods, which achieve high-resolution DOA estimation by exploiting the orthogonality between the signal subspace and the noise subspace. The maximum likelihood method [11] is theoretically optimal. However, in practical applications, due to complex and variable conditions including noise, multipath effects, signal coherence, and array errors, the performance of traditional DOA estimation methods degrades drastically or even becomes ineffective, making it difficult to meet the requirements of high precision and robustness. Meanwhile, traditional model-driven algorithms suffer from high computational complexity, hindering their real-time application. The sparse representation framework assumes that all true directions fall exactly on the pre-selected grid. If the unknown source's incoming wave direction is not in the angle grid, the algorithm almost fails. The solution of sparse representation problems usually has high computational complexity, especially when dealing with large-

Manuscript received January 07, 2026.
*Corresponding author.

scale data.

With the development of machine learning technology, more and more researchers have begun to explore the application of machine learning algorithms in DOA estimation [12]. In 2005, the application of support vector regression (SVR) in DOA estimation was first proposed, and experiments showed that this method has high precision and generalization ability [13]. Later, scholars successively applied the support vector machine (SVM) algorithm to DOA estimation in different scenarios [14–16]. To address the great difficulties brought by high computational complexity and complex spatial structure in large-scale multiple input multiple output (MIMO) systems to the utilization of channel characteristics and sparsity, Huang et al. [17] constructed a deep neural network (DNN) for DOA estimation, proving the effectiveness of adopting a deep learning framework in large-scale MIMO systems. In [18], the authors transformed the multi-source signal angle estimation problem into a multi-label classification task and proposed using convolutional neural networks (CNN) to predict DOA at all signal-to-noise ratio (SNR). Experiments proved that this method has significant noise resistance and strong robustness in low SNR and small snapshot scenarios. The scholar also proposed using CNN to infer the number of source signals and predict DOA with high confidence, and experiments also proved the feasibility of this idea. In addition to using CNN alone for DOA estimation, most scholars combine CNN with other neural networks [19−21]. In [19], the authors proposed a combination of CNN and long short-term memory (LSTM) for DOA estimation, where LSTM is mainly used to suppress impulsive noise, and CNN is used to learn spatial features of signals and perform high-level feature extraction. This algorithm can greatly improve the accuracy and robustness of DOA estimation in impulsive noise environments. As early as 2021, scholar Alexander et al. proposed extending CNN with LSTM to solve the sound source localization problem of microphone arrays, and also proposed a deep learning algorithm integrating CNN and temporal convolutional network (TCN) [20]. Experiments showed that the performance of CNN combined with LSTM is superior to other algorithms, including ordinary CNN and TCN extension. In [21], to solve the problem of DOA of two closely spaced sources, the authors transformed this problem into a multi-label classification task and proposed a new network integrating deep AutoEncoder (DAE) and CNN, named DAE-CNN-MUSIC. Simulation experiments showed that compared with state-of-the-art algorithms, this method has a significant performance gain under low SNR conditions, and can estimate angles more accurately under high SNR conditions. In addition, Zheng et al. proposed an efficient decomposed CNN network for 2D DOA estimation based on sub-Nyquist tensors, derived an enhanced co-array tensor as network input, and this method saves system resources while maintaining competitive performance [22]. Wang et al. proposed a one-shot architecture search and transformation for robust DOA estimation (OAST-DOA) framework, introducing LSTM network as a controller to guide the architecture search and optimal unit selection process, solving the problem of poor DOA estimation performance under low SNR [23]. With the rise of Transformer models in the field of computer vision, some scholars have introduced them into DOA estimation [24–26]. In [24], the authors proposed a dual class token vision transformer (DCT-ViT), which contains two class tokens located at the beginning and end of the latent vector sequence. Experiments show that this framework can adapt to DOA estimation with different numbers of signal sources, and under low SNR, the performance of this method is superior to classical model-based methods and other deep learning-based methods. To address the limitations of subspace-based algorithms, Ji et al. [26] proposed a method using Transformer to assist MUSIC. This method can process multiple snapshots in parallel, thereby capturing global correlations across them. When most scholars use deep learning frameworks for DOA estimation, they convert the problem into a multi-label classification problem, which leads to insufficient utilization of signal source features. In [27], Fan et al. defined the DOA estimation problem as an object detection problem, using the You Only Look Once v3 (YOLOv3) to directly predict the DOA of sources with confidence scores from spectral proxies and achieve end-to-end estimation. Experiments show that compared with several advanced algorithms, it has more advantages in terms of network scale, computational cost, prediction time and DOA estimation accuracy.

However, these methods usually require a large amount of training data and are highly sensitive to model parameters. In recent years, ensemble learning, as a powerful machine learning paradigm, has demonstrated superior performance in multiple fields [28,29]. By combining the prediction results of multiple base learners, ensemble learning can effectively improve the generalization ability and robustness of the model. Compared with the popular neural network algorithms in recent years, ensemble learning algorithms have stronger generalization ability, higher training efficiency, lower data dependence and stronger model interpretability, and are widely used in classification and regression tasks [30]. In 2001, Breiman [31] formally proposed the random forest (RF) algorithm, which is a milestone in the field of ensemble learning. Its advantages in robustness, interpretability and engineering convenience have kept it irreplaceable in industry to this day. In 2016, Chen et al. [32] proposed the eXtreme gradient boosting (XGBoost) algorithm at the SIGKDD conference. This algorithm quickly became a common algorithm in various industries and competitions, and is one of the most influential and widely used algorithms in the field of ensemble learning, known as the "ultimate weapon for machine learning competitions". In DOA estimation, ensemble learning methods can integrate the advantages of various features and algorithms to improve estimation accuracy and adaptability. However, currently, the application of ensemble learning algorithms in the specific signal processing field of DOA estimation is relatively limited.

In this context, drawing on the idea of classification chains (CC), this paper proposes two DOA estimation methods based on RF and XGBoost, CC-RF, and CC-XGBoost, which transform the DOA estimation problem into a multi-label classification problem. Through simulation experiments, the performance of the proposed methods is compared and analyzed with the classical MUSIC algorithm and the currently most commonly used deep learning network-based DOA estimation method, exploring the effectiveness and robustness of the two algorithms in estimating uniform linear array (ULA) far-field narrowband single-source and dual-source signals.

2. Fundamentals of DOA estimation

2.1 Array signal model

The antenna receiving array in this paper adopts a ULA. A ULA with M array elements, receiving K far-field narrowband signals and with an element spacing of d is shown in Fig. 1.

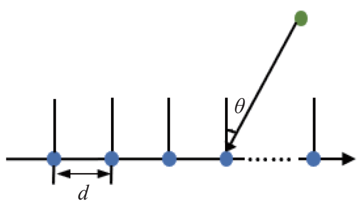


Fig. 1 ULA

The array signal model is expressed as

$$\boldsymbol{x}(t)=\boldsymbol{A}(\theta)\boldsymbol{s}(t)+\boldsymbol{n}(t) \tag{1}$$

where $\boldsymbol{\theta}=[\theta_1,\theta_2,\cdots,\theta_K]$ represents the angle between the DOA and the normal direction of the array elements, θ_k is the estimated value of the k signal source, $\boldsymbol{A}(\theta)$ is the array manifold matrix, $s(t)$ is the signal model, and $\boldsymbol{n}(t)$ represents zero-mean stationary Gaussian white noise with variance σ^2. The noise between each array element is uncorrelated, and uncorrelated with the signal source.

The array steering vector of a ULA with element spacing d is

$$\boldsymbol{a}(\theta_k)=\left[1,\mathrm{e}^{-2\pi\mathrm{j}d\sin\theta_k/\lambda},\cdots,\mathrm{e}^{-2\pi\mathrm{j}d(M-1)\sin\theta_k/\lambda}\right]. \tag{2}$$

The array manifold matrix is defined as $\boldsymbol{A}(\theta)=[\boldsymbol{a}(\theta_1),\boldsymbol{a}(\theta_2),\cdots,\boldsymbol{a}(\theta_K)]$.

A signal whose bandwidth is much smaller than its center frequency is called a narrowband signal, which is expressed in this paper as

$$\boldsymbol{s}(t)=a(t)\mathrm{e}^{\mathrm{j}[w_0 t+\theta(t)]} \tag{3}$$

where $a(t)$ is the amplitude modulation function, $\theta(t)$ is the phase modulation function, and $w_0=2\pi f_0$ is the carrier frequency.

When the signals are incoherent, the received signal covariance matrix is

$$\boldsymbol{R}_{xx}=\boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^{\mathrm{H}}+\sigma^2\boldsymbol{I} \tag{4}$$

where $\boldsymbol{P}$ is the signal power diagonal matrix, and $\boldsymbol{R}_{xx}$ is full-rank.

In practical application environments, due to the fixed phase relationship between multiple signal sources in time or space, that is, the waveforms are statistically correlated, such as multipath effects, there is a linear relationship between the steering vectors, and the signals are called coherent signals. At this time, the received signal covariance matrix $\boldsymbol{R}_{xx}$ is expressed as

$$\boldsymbol{R}_{xx}=\boldsymbol{A}\boldsymbol{\Sigma}_s\boldsymbol{A}^{\mathrm{H}}+\sigma^2\boldsymbol{I} \tag{5}$$

The rank of $\boldsymbol{\Sigma}_s=\mathrm{E}[\boldsymbol{s}(t)\ \boldsymbol{s}^{\mathrm{H}}(t)]$ ranges from 1 to K.

2.2 Classification and regression tree (CART)

The core principle of ensemble learning is to build multiple models (called base learners) on the dataset, and combine the prediction results of multiple base learners in a specific way to obtain a more robust prediction result than a single learner. According to different integration methods, common ensemble learning methods are divided into three categories: bagging, boosting, and stacking. The methods adopted in this paper are RF, a representative method of bagging, and XGBoost, a representative method of boosting. When performing multi-label classification tasks, the base learner used by both algorithms is usually the CART algorithm.

Decision tree is a widely used supervised learning algorithm in the fields of machine learning and data mining. It is a binary tree structure, consisting of nodes and edges, as shown in Fig. 2. Nodes include a root node, several child nodes, and leaf nodes. Each internal node represents an attribute or feature, each branch represents an attribute test, and each leaf node represents a category label.

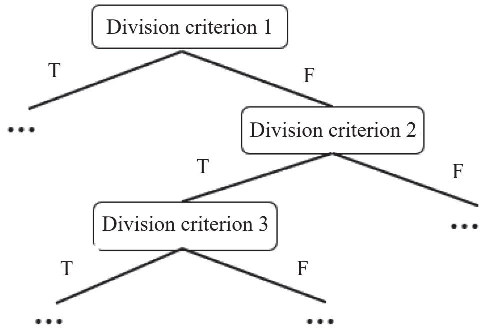


Fig. 2 Classification structure of decision tree

The basic principles of CART are:

(i) Feature selection. CART uses the Gini index to select split attributes. The purity of dataset D can be expressed by the Gini coefficient as follows:

$$\mathrm{Gini}(D) = 1 - \sum_{k=1}^{|Y|} p_k^2 \tag{6}$$

$p_k(k = 1,2,\cdots,|Y\rangle)$ is the proportion of the kth class of samples in the current sample set D, and Y is the total number of sample classes divided by the decision tree. Features with smaller Gini index have higher correlation with classification results. Screen out attributes or features with higher correlation with classification results.

(ii) Split the dataset. Split the dataset into multiple subsets according to the selected attributes.

(iii) Recursively generate subtrees. Repeat the above process for each subset until the stopping conditions are met, such as the sample nodes belonging to the same class, the depth of the tree reaching a preset value, the number of samples being less than a preset threshold.

To ensure that the generated tree model has strong generalization ability, pruning operations can be performed. During model generation, pre-pruning can be performed by limiting the minimum number of samples and the depth of the tree, or post-pruning strategy can be performed on the generated tree from bottom to top.

2.3 Principle of RF

RF is an ensemble learning algorithm improved on the basis of bagging. It not only retains the bootstrap sampling operation of training samples in bagging, but also adds the random selection operation of attributes when training CART. Specifically, on the basis of building the model using the bagging integration principle, before selecting the optimal attribute to split the node during the training of the decision tree, a subset containing m attributes is randomly selected from the attribute set of the node (assuming there are d attributes), and then the optimal attribute is selected from this subset for splitting. k controls the degree of randomness, and it is generally recommended to take $k = \lfloor \log_2 d \rfloor$. The specific principle of the RF algorithm is shown in Fig. 3.

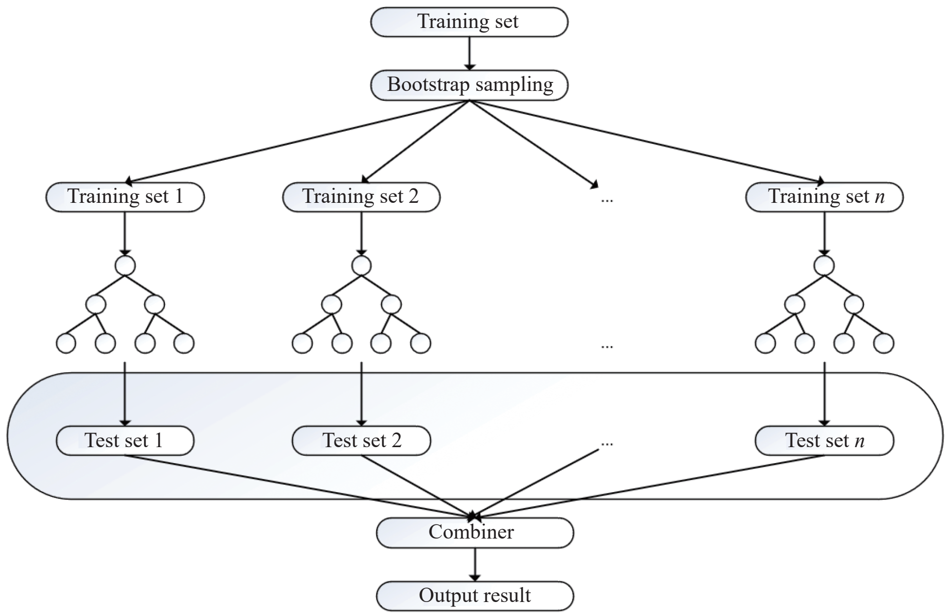


Fig. 3 RF algorithm

The random selection of attributes makes the diversity of base learners in the RF come not only from data sample disturbance but also from attribute disturbance, which further improves the generalization performance of the final ensemble. At the same time, since only the attribute subset is considered when training the base learner, the training efficiency of the RF is better than that of bagging.

For classification tasks, RF selects voting as the combination strategy for prediction results, mainly including hard voting and soft voting. Hard voting means that each tree votes for the predicted class of the sample, and finally selects the class with the most votes. The final prediction result can be expressed as

$$\hat{y} = \text{mode}\{f_1(x), f_2(x), \cdots, f_T(x)\} \tag{7}$$

where $f_T(x)$ represents the prediction result of the T tree. Soft voting means that each tree outputs the probability distribution of the class, and finally selects the class with the largest average probability. The final prediction result is expressed as

$$\hat{y} = \arg\max_k \left(\frac{1}{T} \sum_{t=1}^{T} P_t(y = k \mid x) \right) \tag{8}$$

$P_t(y = k \mid x)$ represents the probability that the t tree predicts the sample x belongs to class k. Combining the prediction results of multiple base learners through voting not only enhances the generalization ability of the model but also improves robustness.

2.4 Principle of XGBoost

XGBoost belongs to the boosting class of ensemble algorithms. The core idea of the XGBoost algorithm is to iteratively train weak learners (usually decision trees), gradually correct the prediction residuals of the previous models, and finally build a strong learner through weighted combination. Compared with the gradient boosting decision tree (GBDT) algorithm, XGBoost has been improved in objective function design, regularization strategy and engineering optimization, and is significantly ahead in accuracy and efficiency.

2.4.1 Mathematical principle

Assume the data sample set is $D = (x_i, y_i)_{i=1}^N$, x_i represents the feature of the i sample, and y_i represents the label of the i sample. The objective function of the XGBoost algorithm is defined as

$$L(\phi) = \sum_{i=1}^{N} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k). \tag{9}$$

It consists of a loss function and a regularization term. l is the loss function, $\hat{y}_i$ is the model prediction value, and f_k represents the k complete CART decision tree. For the multi-label classification problem of DOA estimation in this paper, binary cross-entropy loss is adopted:

$$l = -\sum_{h=1}^{H} [y_{i,h} \ln(p_{i,h}) + (1 - y_{i,h}) \ln(1 - p_{i,h})] \tag{10}$$

where H is the number of labels after angle discretization, and $p_{i,h}$ represents the predicted probability of the i sample on the h angle label.

The second term in (9) is a regularization term, with K representing the number of trees. For each tree f, there is

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^{T} \omega_j^2 \tag{11}$$

where T represents the number of leaf nodes, and $\boldsymbol{\omega}$ is the weight vector of leaf nodes, and parameters γ and λ are hyperparameters, used to control the complexity of the tree and the magnitude of weights, respectively.

XGBoost obtains the k tree through iterative training. When training to the t round, the training result $\hat{y}_i^{t-1}$ of the $t-1$ round is regarded as known, and only the objective function of this round is considered for optimization. To minimize the loss function, XGBoost approximates the loss function using a second-order Taylor expansion, then the loss function obtained in the t iteration can be expressed as

$$L^t = \sum_{i=1}^{N} l\left(y_i, \hat{y}_i^{t-1} + f_t(x_i)\right) + \Omega(f_i) \approx \\ \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_i) \tag{12}$$

where

$$g_i = \frac{\partial l(y_i, \hat{y}_i^{t-1})}{\partial \hat{y}_i^{t-1}}$$

is the first-order derivative, and

$$h_i = \frac{\partial^2 l(y_i, \hat{y}_i^{t-1})}{\partial (\hat{y}_i^{t-1})^2}$$

is the second-order derivative. Let the set of all samples x_i belonging to the j leaf node be $I_j = \{i \mid q(x_i) = j\}$, then the optimal weight of the leaf node can be derived as

$$\omega_j^* = \frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda}. \tag{13}$$

Substituting into the objective function, the total loss before splitting is

$$L_{\text{before}} = -\frac{1}{2}\frac{\left(\sum_{i\in I} g_i\right)^2}{\sum_{i\in I} h_i + \lambda} + \gamma. \tag{14}$$

Split the node into left child node I_L and right child node I_R, the loss after splitting is recorded as

$$L_{\text{after}} = -\frac{1}{2}\left(\frac{\left(\sum_{i\in I_L} g_i\right)^2}{\sum_{i\in I_L} h_i + \lambda} + \frac{\left(\sum_{i\in I_R} g_i\right)^2}{\sum_{i\in I_R} h_i + \lambda}\right) + 2\gamma. \tag{15}$$

The loss value before and after splitting is the split gain, which can be expressed as

$$\text{Gain} = \frac{\left(\sum_{i\in I_L} g_i\right)^2}{\sum_{i\in I_L} h_i + \lambda} + \frac{\left(\sum_{i\in I_R} g_i\right)^2}{\sum_{i\in I_R} h_i + \lambda} - \frac{\left(\sum_{i\in I} g_i\right)^2}{\sum_{i\in I} h_i + \lambda} - \gamma. \tag{16}$$

This gain formula considers both the first-order gradient direction and the second-order curvature information. The larger the Gain value, the less the loss after splitting, ensuring the global optimality of the splitting strategy.

2.4.2 Algorithm flow

The specific flow of the XGBoost algorithm is shown in Fig. 4.

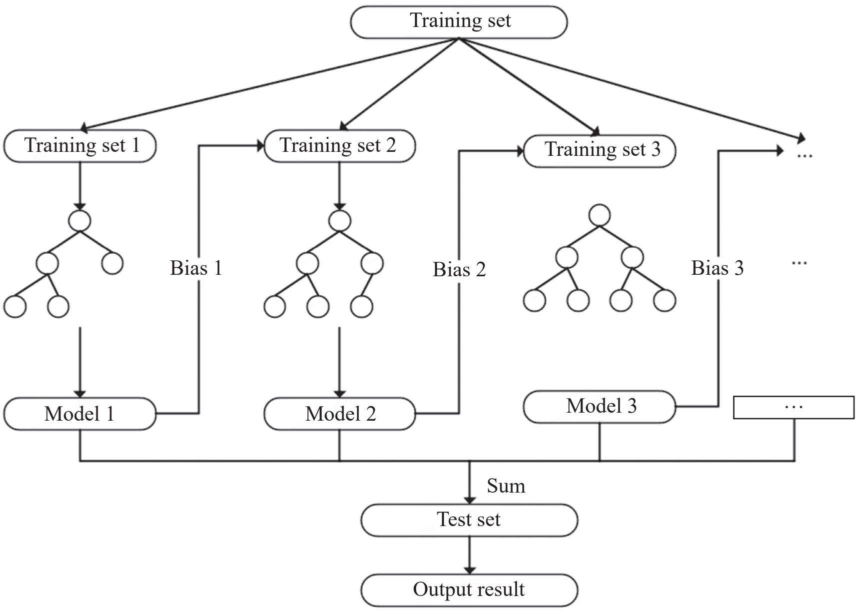


Fig. 4 XGBoost framework

The main steps are as follows.

Step 1 Initialization. The XGBoost algorithm uses the mean of the training labels as the initial prediction value for all samples, usually the mode (for classification problems) or the average value (for regression problems).

Step 2 Calculating residuals. Calculate the difference between the actual value and the predicted value of the current model. This residual reflects the prediction error of the model in the current round.

Step 3 Building a new tree. Take the residual as the new training target, build a new decision tree, and fit the residual of the current round.

Step 4 Calculating the objective function and determine the leaf node weights. Obtaining the optimal weight of each leaf node by solving the minimization problem of the objective function. The weight represents the prediction contribution of the leaf node to all samples.

Step 5 Updating the model. Add the newly built decision tree and its leaf node weights to the model, and update the predicted value of the model. The new predicted value is equal to the predicted value of the previous round plus the predicted value of the new tree for the current sample.

Step 6 Regularization processing. It usually includes setting the learning rate, and limiting parameters such as the maximum depth of the tree and the number of leaf nodes.

Step 7 Determining whether the termination condition is met. If not, repeat Steps 2−6; if yes, the iteration terminates. Termination conditions include a preset number of iterations or convergence of the loss function.

Step 8 Model output. Output the final model.

XGBoost’s special integration method and regularization strategy ensure that it can automatically screen key features, suppress redundant information when performing DOA estimation. At the same time, the parallel tree construction method increases the DOA estimation delay of a single sample to the millisecond level.

3. Design of DOA estimation model

This paper mainly addresses the real-time and noise resistance requirements of DOA estimation, and transforms the DOA estimation problem into a multi-label estimation problem. The DOA estimation framework proposed in this paper is shown in Fig. 5.

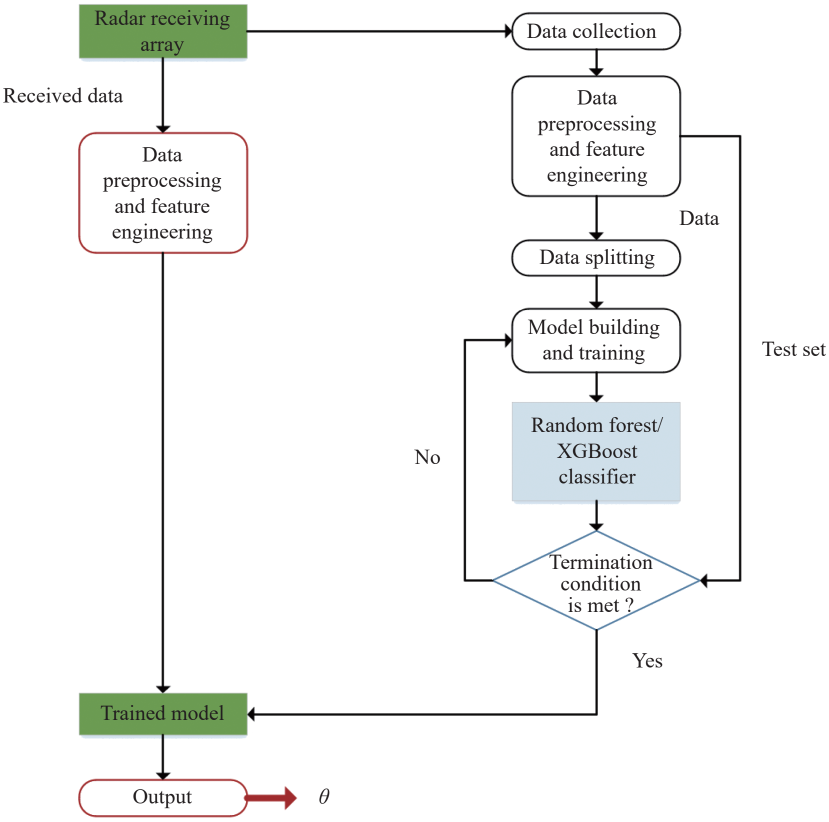


Fig. 5 DOA estimation framework

3.1 Data preprocessing

Data annotation. Consider that the range of arrival angles in practical situations is generally $[-60°, 60°]$. When performing single-source DOA prediction, the angle is annotated directly as a label, with a label every 0.1°, totaling 1 201 labels.

For dual-source DOA estimation, the angle range $[-60°, 60°]$ is discretized into 121 labels (one label every 1°), and angle combinations are annotated, such as $\theta_1 = -30°, \theta_2 = 58°$. The two algorithms proposed in this section use angle combinations as labels. Multi-labels are encoded using multi-hot encoding and mapped to binary vectors, where the corresponding angle label is 1 and the rest are 0. For example, the angle combination $[-60°, -57°]$ is mapped to $[1, 0, 0, 1, 0, \cdots, 0]$. Under the experimental settings of this chapter, the label vector length of each sample is 121, which is used as the training label of the comparative algorithm CNN. Remove data with missing values and duplicates from the dataset.

Data cleaning. Remove data with missing values and duplicates from the dataset.

3.2 Feature extraction

The real and imaginary parts of the covariance matrix of the array received signals respectively contain the amplitude and phase relationships of the signals, which are information compression of the original signals. Since the original RF and XGBoost algorithms can only process two-dimensional data, this chapter flattens the real and imaginary parts of each signal covariance matrix into one-dimensional vectors and combines them into a two-dimensional array as the model feature input. The specific operation methods are as follows.

(i) Calculate the covariance matrix.

The array steering vector of a ULA with element spacing d is

$$\boldsymbol{a}(\theta_k) = \left[1, \mathrm{e}^{-2\pi \mathrm{j} d \sin\theta_k/\lambda}, \cdots, \mathrm{e}^{-2\pi \mathrm{j} d(M-1)\sin\theta_k/\lambda}\right]. \tag{17}$$

The array manifold matrix is defined as $\boldsymbol{A}(\theta) = [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \cdots, \boldsymbol{a}(\theta_k)]$.

A signal whose bandwidth is much smaller than its center frequency is called a narrowband signal, which is expressed in this section as

$$\boldsymbol{s}(t) = a(t)\mathrm{e}^{\mathrm{j}[w_0 t + \theta(t)]} \tag{18}$$

where $a(t)$ is the amplitude modulation function, $\theta(t)$ is the phase modulation function, and $w_0 = 2\pi f_0$ is the carrier frequency.

When the signals are incoherent, the received signal covariance matrix $\boldsymbol{R}_{xx}$ is

$$\boldsymbol{R}_{xx} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^{\mathrm{H}} + \sigma^2 \boldsymbol{I} \tag{19}$$

where $\boldsymbol{P}$ is the signal power diagonal matrix, and $\boldsymbol{R}_{xx}$ is full-rank.

(ii) Extract the upper triangular data of the covariance matrix.

(iii) Flatten the real and imaginary parts into two vectors row by row.

(iv) Concatenate the vectors $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ into one vector $\boldsymbol{\beta}$.

The upper triangular part has a total of $M(M-1)/2$ elements. The first $M(M-1)/2$ elements in the $M(M-1)$-dimensional vector $\boldsymbol{\beta}$ are the real parts of the upper triangular elements of the covariance matrix, while the last $M(M-1)/2$ elements are the imaginary parts of the upper triangle. $\boldsymbol{\beta}$ is the vector of effective eigenvalues ultimately extracted.

The model input feature matrix with N samples is denoted as $\boldsymbol{B}$, expressed as

$$\boldsymbol{B} = [\beta_1, \beta_2, \cdots, \beta_N]. \tag{20}$$

3.3 Model training

3.3.1 Single-source DOA estimation

For the single-source DOA estimation problem, RF and XGBoost classifiers are directly created, the annotated data and the extracted covariance matrix features are used as model inputs, and model parameters are set. Parameter tuning is an important step in model training, which is crucial for improving model performance. Parameter tuning needs to understand the meaning of parameters and be carried out according to specific tasks and data, mainly relying on the experimenter's experience and automatic parameter tuning algorithms. This paper adopts the grid search method for parameter optimization.

The core parameters of the RF model mainly include n_estimators (number of trees), max_depth (maximum depth of the tree), min_samples_split (minimum number of samples for leaf node splitting), and min_samples_leaf (minimum number of samples for leaf nodes).

The core parameters of the XGBoost model include objective (objective function, default is binary:logistic in this paper), n_estimators (number of trees), max_depth (maximum depth of the tree), learning_rate (learning rate), and subsample (subsampling ratio).

Set the value range of each parameter, use the grid search method for automatic parameter tuning, the loss function of the RF model is the Gini index, and the loss function of the XGBoost model is the binary cross-entropy loss function.

3.3.2 Multi-source DOA estimation

Multi-source DOA estimation corresponds to a multi-label classification task, while the original RF and XGBoost algorithms cannot directly perform multi-label classification tasks.

There are three common traditional machine learning methods for handling multi-label classification problems: binary relevance (BR), CC, and label powerset (LP). The BR method transforms the multi-label classification problem into multiple independent binary classification problems, each label corresponds to a classifier, and finally summarizes the prediction results of all classifiers as the final result. The multi-output-classifier wrapper of the Scikit-learn library is consistent with the idea of the BR method, but since it cannot guarantee the number of predicted labels, and the number of signal sources is known in the DOA estimation experiment simulated in this paper, this method cannot obtain high estimation performance. The idea of CC is to connect multiple binary classifiers in a chain. The input of each classifier includes not only the original features but also the prediction results of the previous classifiers. LP regards all label combinations as an independent category, transforming the multi-label classification problem into a multi-classification problem. The number of categories of this method grows exponentially with the increase of label categories, and the algorithm complexity is high. XGBoost has poor performance when processing such data.

Based on the idea of CC, this section constructs chained XGBoost and RF models, namely proposing the CC-RF and CC-XGBoost algorithms. When performing multi-source DOA estimation, the number of classifiers corresponds to the number of signal sources. For training a DOA estimation model with K signal sources, the labels are split column-wise into K label arrays: Label 1, Label 2, $\cdots$, Label K, and multiple XGBoost and RF models are trained sequentially. The specific steps are as follows.

Step 1 When training the first model, the original features and Label 1 are used as the training set.

Step 2 When training the second model, the array formed by concatenating the original features and Label 1 is used as the feature input, and Label 2 is used as the label for training.

Step 3 Following this method, when training the kth model, the original features concatenated with the first $k-1$ label arrays are used as the feature input, and the kth label is used as the label for training. This process continues until all K models are trained.

The models are linked in a chain through the concatenation of the original features and the labels from the previous models. For DOA estimation, the kth model uses the original features and the prediction results from the first $k-1$ models as input. Finally, the predictions from all classifiers are combined to form the final DOA estimation result. Fig. 6 shows the structure of the dual-source DOA estimation model in this section. The combination operation is performed column-wise.

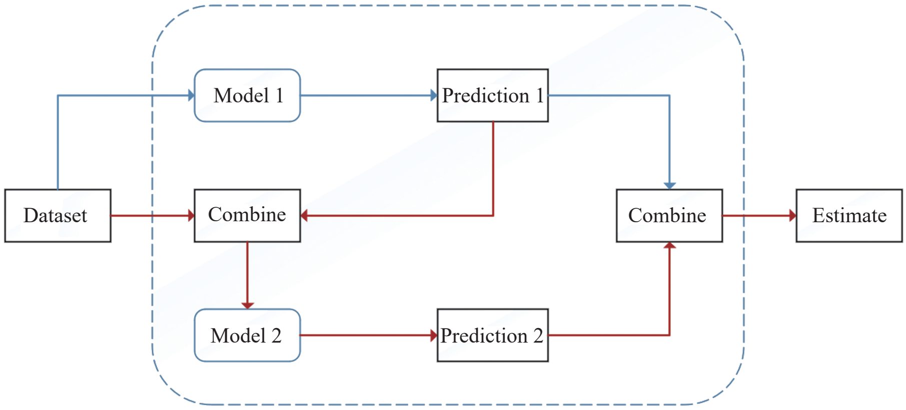


Fig. 6 Estimation model

The parameters are consistent with the single-source DOA estimation model. The parameter value range is set, and the grid search method is used for automatic parameter tuning. The loss functions of the two models remain unchanged.

After multi-source DOA estimation, when converting vector labels into angle labels, it is necessary to note whether sorting is needed to eliminate the impact of order on evaluation indicators.

3.4 Model evaluation

Model evaluation is an important part of machine learning. The quality of a model not only depends on its performance on the training set but also, more importantly, on its performance on new data. Common evaluation methods include hold-out method, cross-validation method, and bootstrap method. The hold-out method divides the dataset into two mutually exclusive sets, denoted as the training set and the test set. After training the model on the training set, the test set is used for prediction to evaluate the generalization error of the model. The cross-validation method divides the dataset into m mutually exclusive subsets of similar size by stratified sampling. Each time, $m-1$ subsets are used as the training set, and the remaining subset is used as the test set, so that m groups of training and test sets can be obtained. The bootstrap method is based on bootstrap sampling. In a dataset containing s samples, random sampling with replacement is performed. After s samplings, a sampling set containing s training samples is obtained. Repeat the sampling operation R times to obtain R sampling sets containing s training samples, and R model trainings can be performed. Moreover, the probability that a sample is never sampled after s samplings is

$$\lim_{s\to\infty}\left(1-\frac{1}{s}\right)^{s}=\frac{1}{\mathrm{e}}. \tag{21}$$

About 36.8% of the samples do not appear in the training set and can be used to test the model. The evaluation method adopted in this paper is the hold-out method, where 80% of the dataset is used for model training and the remaining 20% is used for testing the model.

Although discrete labels are used in this subsection, in the DOA estimation problem, the accuracy of estimation cannot be judged by right or wrong. The root mean square error (RMSE) is the most commonly used evaluation indicator for the performance of DOA estimation algorithms, and it is also a typical evaluation indicator in machine learning for measuring the performance of models in performing continuous value prediction tasks. Its calculation formula is written as

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\theta}_i-\theta_i\right)^2} \tag{22}$$

where $\hat{\theta}_i$ represents the predicted value of the DOA estimation model, and n is the number of test samples.

In this section, the predicted discrete interval is converted into a continuous angle, and its RMSE with the true angle is calculated to evaluate the model. For multi-source estimation, angle resolution is also needed as a measure to evaluate the signal source identification ability of the model.

4. Experiments

4.1 Dataset generation and experimental settings

The number of array elements is set to 8, the radar operating frequency is 77e9, the number of snapshots is 500, the system is Windows 10, and the Python version is 3.8. When performing DOA estimation with two signal sources, the two signal sources are randomly chosen from a set of angles, resulting in a total of $C_{121}^{2} = 7\,260$ possible angle combinations.

Single-source. The target incident range is $[-60°, 60°]$, discretized into 1 201 labels with an interval of 0.1°. In the SNR interval of [−10,15] dB, with an SNR interval of 5 dB, 10 samples are generated for each label. 12 010 samples are generated under each SNR condition, totaling 72 060 samples. The training set and test set under each SNR are randomly selected in a ratio of 8:2. The total number of training sets is 57 648 samples, and the test set is 14 412 samples.

Dual-source. The target incident range is $[-60°, 60°]$, discretized into 121 basic angles with an interval of 1°. In the SNR interval of [−10,15] dB, with an SNR interval of 5 dB, 10 samples are generated for each DOA combination, resulting in a total of 435 600 samples in the sample set. The training set and test set under each SNR are randomly selected in a ratio of 8:2. The total number of training sets is 348 480 samples, and the test set is 87 120 samples.

In all experiments, the methods proposed in this paper are compared and analyzed with the classic traditional algorithm MUSIC [1] and the deep learning algorithm CNN, and the RMSE is used as the model evaluation indicator.

4.2 Single-source DOA estimation experiments

4.2.1 Experiment 1: single-source estimation of the proposed algorithm under ideal conditions

To verify the effectiveness of the angle estimation of the proposed algorithm, this experiment is carried out under the conditions of SNR = 10 dB, number of snapshots = 500, and other conditions unchanged. The amount of training set data is 9 608, and the number of test sets is 2 402. Fig. 7 shows the comparison between the DOA estimation values and the true values of 2 402 randomly selected test samples. Table 1 shows the RMSE and time consumption of different algorithms for estimating the test samples.

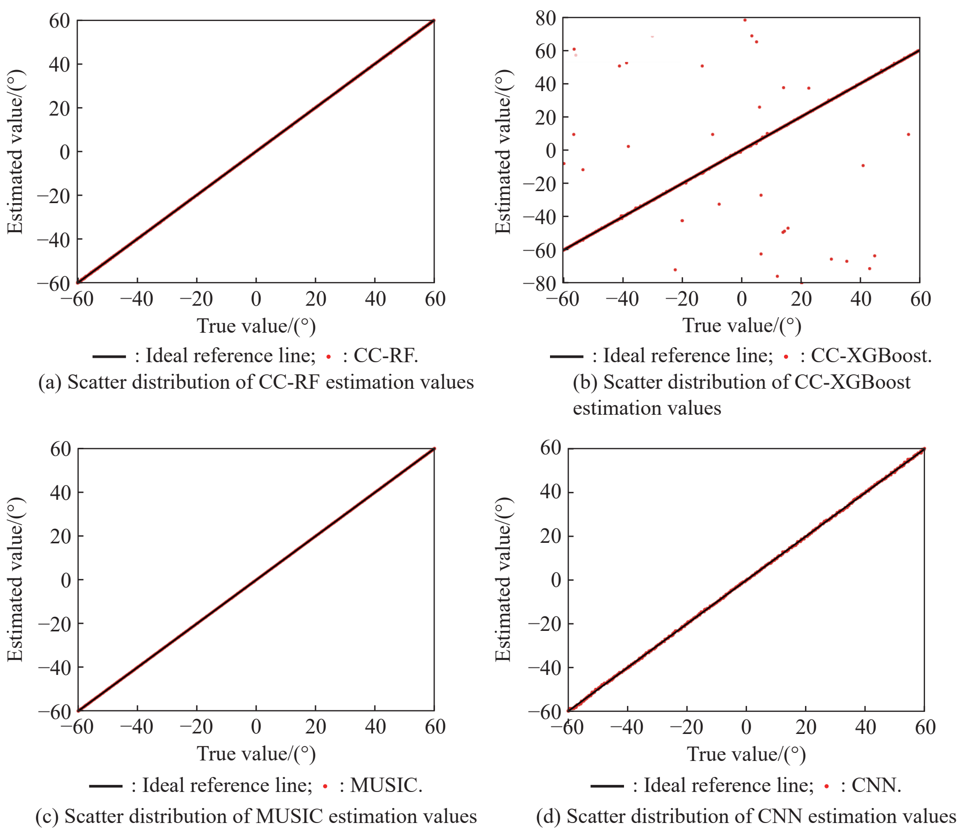


Fig. 7 Scatter distribution of true values and estimated values of each algorithm

Table 1 Time consumption and RMSE of each algorithm s

Algorithm	RMSE/(°)	Times/s
CC-XGBoost	8.1246	0.572195
CC-RF	0.0823	2.361395
MUSIC	0.0396	104.343696
CNN	0.2761	0.5029

As shown in Fig. 7, when the label angle range of the test samples is $[-60°, 60°]$ and the angle interval is 0.1°, the scatter distribution diagrams of the estimated values and true values of the three DOA estimation algorithms shown in Fig. 7(a), Fig. 7(c), and Fig. 7(d) all show the characteristics of being densely around the ideal reference line. Different from the other three algorithms, the DOA estimation algorithm based on CC-XGBoost in Fig. 7(b) has a large number of points completely falling on the ideal reference line within the entire angle range, and some discrete points are randomly distributed around the reference line. Combined with the RMSE of different algorithms in Table 1, among the four algorithms, the MUSIC algorithm has the lowest estimation RMSE, and the CC-XGBoost algorithm has the highest estimation error, which is consistent with the information shown in the scatter distribution of the four algorithms in Fig. 7. The CC-RF algorithm proposed in this paper achieves the same order of estimation accuracy as the MUSIC algorithm. In addition, the time spent by the traditional MUSIC algorithm to estimate 2402 samples in Table 1 is much longer than the other three algorithms, followed by the CC-RF algorithm, and the CC-XGBoost and CNN data are close, with the least time consumption. After calculation, the time spent by the CC-XGBoost, CC-RF, and CNN algorithms for single sample estimation is less than 1 ms, which can meet the real-time estimation requirements.

Under the experimental conditions, the CC-RF algorithm achieves the same order of estimation accuracy as the MUSIC algorithm, which is higher than the CNN and CC-XGBoost algorithms. The reason is that XGBoost itself is more suitable for large-scale and high-dimensional features. Under the condition of eight array elements, the upper triangular function is taken as the feature input, and the length of the feature vector is only 56, while the number of label categories is 1 201. The model cannot fully learn the mapping rules between features and labels, resulting in large estimation errors. The main reason why the performance of the CNN-based DOA estimation model is worse than that of CC-RF under this dataset is that neural network training requires a large amount of data. In this experiment, there are only 10 samples per label, and after dividing the dataset into training set and test set, there are about eight samples per label in the training set. Too few samples mean that the neural network cannot fully learn the rules of the samples, leading to performance degradation.

Based on the above experimental results, under the conditions of SNR = 10 and number of snapshots = 500, the CC-RF algorithm shows significant advantages in both estimation accuracy and real-time performance, and the overall performance is better than the other three algorithms.

4.2.2 Experiment 2: impact of SNR on single-source estimation performance of the proposed algorithm

To verify the anti-interference ability of the proposed algorithm, the experiment is carried out under different SNRs. The SNR value range is $[-10, 15]$, with an interval of 5 dB. Fig. 8 shows the values of DOA estimation RMSE and prediction time as a function of SNR.

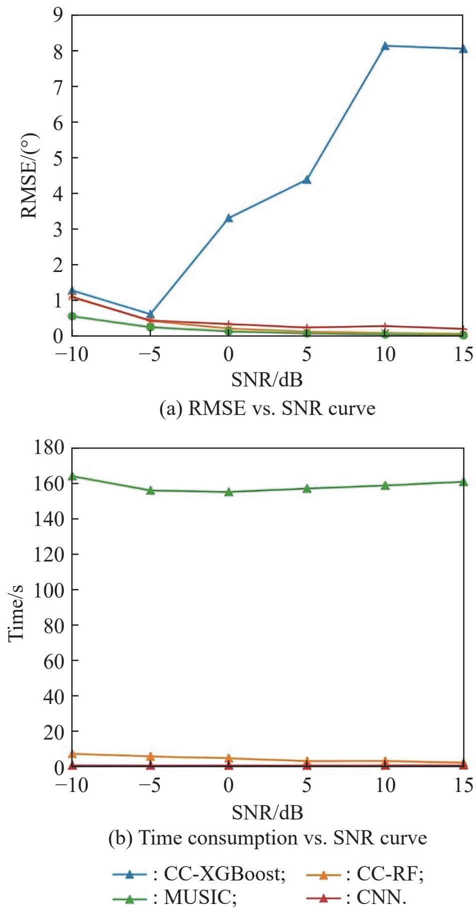


Fig. 8 RMSE and time consumption vs. SNR curves

Fig. 8 shows the changes in RMSE and time consumption of the four algorithms under different SNRs. As shown in Fig. 8(a), the RMSE curves of the CC-RF, MUSIC, and CNN algorithms all show a slow downward trend as the SNR increases, and the CC-RF has the most drastic downward trend. At SNR=−10 dB, the estimation accuracy is reduced by 0.4° compared with MUSIC. When SNR⩽−5 dB, the estimation accuracy of CC-RF is equivalent to that of the CNN algorithm; when SNR increases from −5−5 dB, the RMSE decline rate of CC-RF is higher than that of CNN, and the estimation accuracy gradually maintains the same level as MUSIC. After SNR>5 dB, the estimation accuracy tends to be stable, and CC-RF achieves DOA estimation accuracy comparable to MUSIC, which is higher than the CNN algorithm. Compared with CNN, when SNR<−5 dB, the estimation errors of both are equivalent; when SNR>−5 dB, the estimation accuracy of CC-RF is always higher than that of CNN. The estimation error of the CC-XGBoost algorithm is always higher than the other three algorithms, and after SNR>−5 dB, the RMSE gradually increases, and the gap in estimation accuracy with the other three algorithms also gradually increases. In Fig. 8(b), the estimation time of the four algorithms changes little overall, but under any SNR, the time cost of MUSIC is much higher than that of XGBoost, CC-RF, and CNN. Among the other three algorithms, CC-RF is slightly higher than CC-XGBoost and CNN, and the curves of the latter two are almost overlapping.

Although both CC-XGBoost and CC-RF are ensemble algorithms with CART trees as base learners, the difference in their integration methods leads to such a large difference in DOA estimation performance under the same conditions. In the ULA, the real and imaginary parts of the covariance matrix have strong linear correlation. Under high SNR, the covariance is dominated by signals, and the features extracted from adjacent angle labels are similar. Moreover, the splitting of XGBoost leaf nodes is based on a greedy algorithm, and the selection of local optimality leads XGBoost to easily select redundant features, resulting in the model learning wrong rules. The random selection of feature subsets before training of the RF algorithm avoids the impact of redundant features. In addition, when the angle interval is 0.1°, there are 1201 labels. XGBoost needs to train a large number of base learners. When the covariance matrices of adjacent angles are close, the label boundaries are blurred and cannot be distinguished clearly. The voting system of RF is more robust to the problem of blurred boundaries.

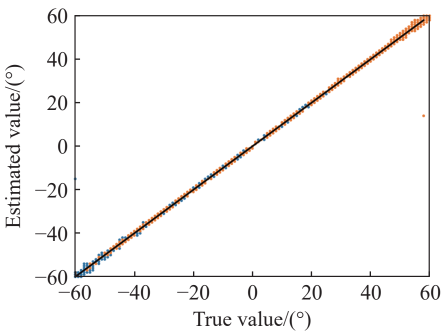
MUSIC relies on subspace orthogonality and essentially uses the statistical independence of signals and noise, while the RF algorithm learns the statistical characteristics of data through a voting system. Therefore, under high SNR conditions, the principles of the two algorithms are similar, and the estimation accuracy is always comparable. Compared with the XGBoost algorithm, it is more suitable for DOA estimation problems under high SNR.

The results of Experiment 2 show that among the algorithms proposed in this section, the CC-RF algorithm has higher estimation accuracy under any SNR condition, and the time cost is similar to that of CC-XGBoost, which is much lower than that of the MUSIC algorithm. Therefore, the CC-RF algorithm can maintain the same estimation accuracy as the MUSIC algorithm while maintaining real-time estimation, and its performance is the best.

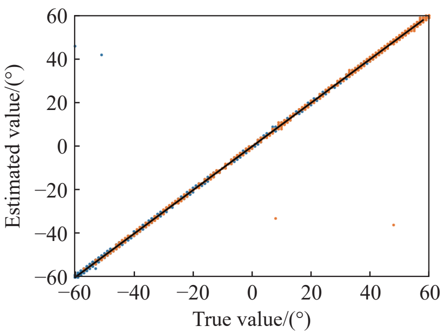
4.3 Dual-source DOA estimation experiments

4.3.1 Experiment 1: dual-source estimation of the proposed algorithm under ideal conditions

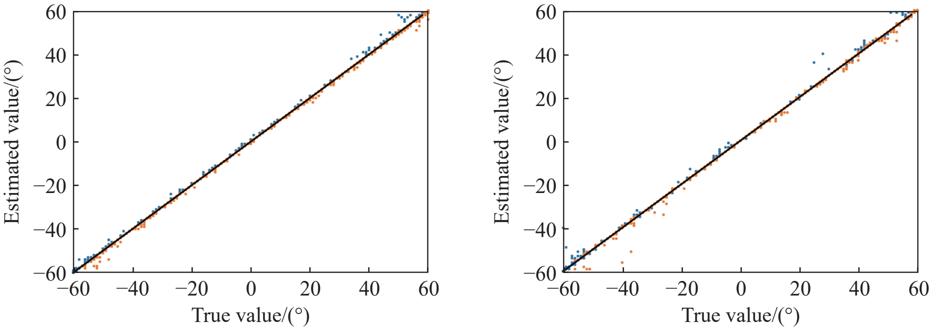
To verify the effectiveness of the angle estimation of the proposed algorithm, this experiment is carried out under the conditions of SNR = 10 dB, number of snapshots=500, and other conditions unchanged. The amount of training set data is 58080, and the number of test sets is 14 520. Fig. 9 shows the comparison between the two DOA estimation values and the true values of the test samples. Table 2 shows the time spent by different algorithms to estimate the test samples under the current conditions.



(a) Scatter distribution of CC-RF estimation values



(b) Scatter distribution of CC-XGBoost estimation values



(c) Scatter distribution of MUSIC estimation values (d) Scatter distribution of CNN estimation values

—— : Ideal reference line; • : First target; • : Second target.

Fig. 9 Comparison of true values and estimated values of dual-sources under ideal conditions

Table 2 RMSE and time consumption of each dual-sources

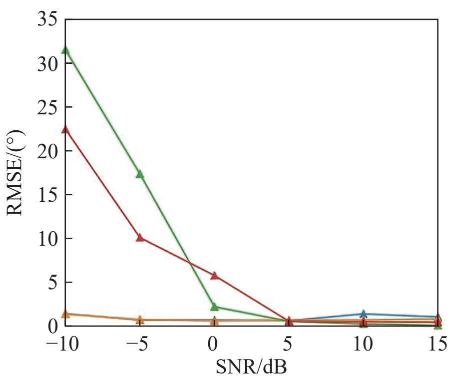
Algorithm	RMSE/(°)	Time/s
CC-XGBoost	1.4297	1.9995
CC-RF	0.7375	1.3473
MUSIC	0.2767	107.5744
CNN	0.5174	2.1647

As shown in Fig. 9, all points of the four algorithms are around the diagonal. Among the four algorithms, the CNN has the largest degree of dispersion, and all show the phenomenon of being concentrated in the middle and more dispersed at the edges. In the images of the CC-XGBoost algorithm and the CC-RF algorithm, the distribution of the first signal source points is more discrete at low angles, and the estimation error is larger. With the increase of the angle, the estimation error of the second signal source increases, and the estimation accuracy of signal source 1 increases. This is because when generating angle combinations, to ensure the accuracy of RMSE calculation, the angle combinations are artificially sorted from small to large, so that the second angle is always larger than the first angle, leading to the distribution of algorithm points. In general, the CC-XGBoost algorithm and the CC-RF algorithm have no directional deviation under the two angles. The scatter plot of the MUSIC algorithm shows that the estimated value of the first signal source is too large, and the estimated value of the second signal source is too small. Although there are many outlier points in the scatter plot of the CNN algorithm, the distribution is roughly uniform without directional deviation. Combined with Table 2, the MUSIC algorithm has the highest estimation accuracy, followed by the CNN algorithm, which seems inconsistent with Fig. 9. However, compared with 14520 test samples, the number of outlier points of the MUSIC algorithm and the CNN algorithm is very small, which has little impact on the overall error. Although the images of the CC-XGBoost algorithm and the CC-RF algorithm seem dense, most of the points are not completely strictly distributed on the diagonal, and the estimation error of a single sample is similar to the error of all samples. In terms of time cost, the estimation time of a single sample of the three data-driven algorithms is less than 1 ms, which meets the real-time DOA estimation, while the time cost of MUSIC is nearly 100 times that of the CC-XGBoost and CC-RF algorithms.

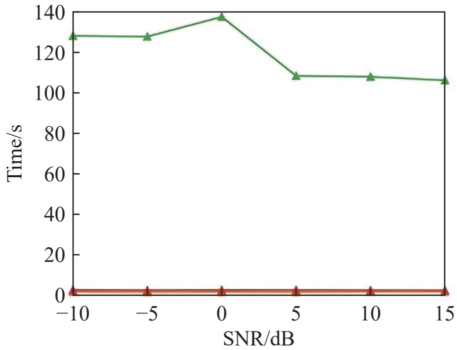
In general, under the experimental conditions, the CNN has the best performance. The CC-RF has the lowest time cost, but the estimation accuracy is slightly lower than that of MUSIC and CNN.

4.3.2 Experiment 2: impact of SNR on dual-source estimation performance of the proposed algorithm

To verify the anti-interference ability of the proposed algorithm, the experiment is carried out under different SNRs. The SNR value range is [−10,15] dB, with an interval of 5 dB. Fig. 10 shows the values of DOA estimation RMSE and prediction time as a function of SNR.



(a) RMSE vs. SNR curve



(b) Time consumption vs. SNR curve

: CC-XGBoost; : CC-RF; : MUSIC; : CNN.

Fig. 10 Dual-source RMSE and time consumption vs. SNR curves

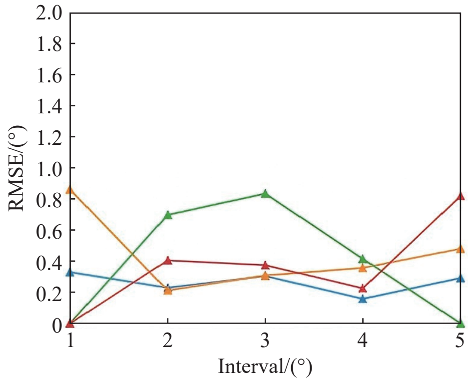


Fig. 11 RMSE change curves of various algorithms under different angle intervals

The experimental results in Fig. 10(a) show that the RMSE curves of the CC-RF and CC-XGBoost algorithms remain almost stable with the change of SNR, and the two only have an error change of nearly 0.7° from −10—−5 dB. Overall, the CC-RF algorithm has significant noise robustness and can always maintain high-precision DOA estimation under low SNR. The RMSE curves of the MUSIC and CNN algorithms show a downward trend as the SNR increases, and start to stabilize when the SNR is equal to 5 dB. The reason for the large estimation error of the CNN algorithm under low SNR is that the model splits the real and imaginary parts of the covariance matrix into two 8×8 matrices as inputs. Under low SNR, it is mainly dominated by noise, and the CNN cannot completely learn accurate rules when the number of data training samples is small. From the curve of DOA estimation time as a function of SNR in Fig. 10(b), it can be seen that only the time spent by the MUSIC algorithm is sensitive to SNR changes, and the time consumption decreases as the SNR increases until it stabilizes when the SNR is equal to 5 dB.

In summary, the CC-XGBoost and CC-RF algorithms always maintain real-time and accurate DOA estimation under different SNRs, and have excellent noise suppression performance. The CC-RF algorithm has the best overall performance among the four algorithms and strong robustness.

4.3.3 Experiment 3: angle resolution test of multi-source estimation

To verify the ability of the model to accurately identify the direction of each signal source in a multi-source environment, this experiment generates 1 000 sets of data with true angle intervals of [1,2,3,4,5] to test the angle resolution of the model at different angle intervals. Fig. 11 shows the RMSE change curves of the four algorithms under different angle intervals.

As can be seen from Fig. 11, with the increase of the angle interval, the estimation error curves of each algorithm show a non-monotonic change. When the angle interval is 1°, the CC-RF algorithm has the largest estimation error, reaching 0.86°, followed by CC-XGBoost (0.33°), and the CNN and MUSIC algorithms achieve unbiased estimation in this scenario (RMSE = 0°). With the increase of the angle interval, the RMSE of CNN generally shows an upward trend, the CC-RF first decreases and then increases, the MUSIC algorithm first increases and then decreases, and the CC-XGBoost is relatively stable. It is worth noting that when the angle interval ⩾ 1°, the traditional MUSIC algorithm can completely resolve dual signal sources by virtue of the subspace orthogonality principle, while the three algorithms CNN, CC-XGBoost, and CC-RF also show angular resolution comparable to that of the classical algorithm by directly learning rules from the covariance matrix.

5. Conclusions

This paper mainly proposes two DOA estimation models based on XGBoost and RF. By connecting multiple XGBoost and RF through learner chains, they are used for multi-label classification tasks. This paper also conducts multiple experiments on ULA far-field narrowband single-source and dual-source signals under single-source and dual-source conditions, comparing with MUSIC and CNN algorithms to analyze the effectiveness of the proposed models.

Simulation experimental results show that in the DOA estimation task, the CC-RF model shows comprehensive advantages: in the single-source scenario, its estimation accuracy is comparable to that of the MUSIC algorithm, but the computational efficiency is improved by 50 times, and in the dual-source task with low SNR, the error is reduced by an order of magnitude compared with the comparative algorithms (MUSIC/CNN), while maintaining high precision and strong generalization in small-sample and multi-label data; after label simplification, CC-XGBoost can achieve performance comparable to CC-RF in the dual-source task, but it is sensitive to SNR and fails in the single-source small-sample scenario, with limited applicability.

The statistical performance reflected by the voting system of CC-RF is essentially consistent with DOA estimation under high SNR. The anti-interference ability under low SNR also makes the performance of CC-RF superior to other algorithms. At the same time, the millisecond-level DOA estimation rate meets the requirements of current applications for real-time estimation. In summary, under small samples and low SNR, the CC-RF proposed in this chapter is a high-precision, real-time, and robust DOA estimation algorithm, which is superior to CC-XGBoost, CNN, and MUSIC algorithms.

Although this study explores the great potential of ensemble learning in DOA estimation, it inevitably faces some challenges. A large number of label categories mean high computational complexity. Transforming the DOA estimation problem into a multi-label estimation problem, if it is necessary to estimate angles with smaller grids, the model training time and estimation time cost will increase sharply, making it difficult to apply in real time. In future research, methods to reduce the time complexity of ensemble learning can be explored by combining deep learning ideas.

References

[1] ROY R, KAILATH T. ESPRIT-estimation of signal parameters via rotational invariance techniques. IEEE Trans. on Acoustics, Speech, and signal Processing, 1988, 37(7): 984–995.

[2] SCHMIDT R. Multiple emitter location and signal parameter estimation. IEEE Trans. on Antennas and Propagation, 1986, 34(3): 276–280.

[3] LAN C F, CHEN H, ZHANG L, et al. Underwater acoustic DOA estimation of incoherent signal based on improved GA-MUSIC. IEEE Access, 2023, 11(1): 69474–69485.

[4] HERZOG A, HABETS E. Eigenbeam-ESPRIT for DOA-vector estimation. IEEE Signal Processing Letters, 2019, 26(4): 572–576.

[5] LIU S, ZHAO J, WU D C, et al. 2D DOA estimation of coherent signals with a separated linear acoustic vector-sensor array. China Communications, 2024, 21(2): 155–165.

[6] WANG L, ZHAO L F, BI G A, et al. Novel wideband DOA estimation based on sparse Bayesian learning with dirichlet process priors. IEEE Trans. on Signal Processing, 2016, 64(2): 275–289.

[7] WANG Q S, YU H, LI J, et al. Adaptive grid refinement method for DOA estimation via sparse bayesian learning. IEEE Journal of Oceanic Engineering, 2023, 48(3): 806–819.

[8] OLLILA E. Multichannel sparse recovery of complex-valued signals using Huber's criterion. Proc. of the 3rd International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing, 2015. DOI: 10.1109/CoSeRa.2015.7330257.

[9] TRONG-DAI H, HUANG X J, QIN P Y, et al. Low-complexity direction-of-arrival estimation with orthogonal matching pursuit for large-scale lens antenna array. IEEE Trans. on Communications, 2025, 73(7): 3924–3939.

[10] DONG Y J, XU Y Y, LIU S, et al. DOA estimation based on sparse Bayesian learning under amplitude-phase error and position error. Journal of Systems Engineering and Electronics, 2025, 36(5): 1122–1131.

[11] ELBIR A M, CELIK A, ELTAWIL A M. NEAT-MUSIC: auto-calibration of DOA estimation for terahertz-band massive MIMO systems. IEEE Wireless Communications Letters, 2024, 13(2): 451–455.

[12] ZHU H G, CHEN X X, MA T, et al. Deep unfolded amplitude-phase error self-calibration network for DOA estimation. Journal of Systems Engineering and Electronics, 2025, 36(2): 353–361.

[13] PASTORINO M , RANDAZZO A. A smart antenna system for direction of arrival estimation based on a support vector regression. IEEE Trans. on Antennas & Propagation, 2005, 53(7): 2161−2168.

[14] TARKOWSKI M, KULAS L. RSS-based DoA estimation for ESPAR antennas using support vector machine. IEEE Antennas and Wireless Propagation Letters, 2019, 18(4): 561–565.

[15] GAO Y L, HU D S, CHEN Y P, et al. Gridless 1-b DOA estimation exploiting SVM approach. IEEE Communications Letters, 2017, 21(10): 2210–2213.

[16] WU L L, HUANG Z T. Coherent SVR learning for wideband direction-of-arrival estimation. IEEE Signal Processing Letters, 2019, 26(4): 642–646.

[17] HUANG H J, YANG J, HUANG H, et al. Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system. IEEE Trans. on Vehicular Technology, 2018, 67(9): 8549–8560.

[18] PAPAGEORGIOU K, SELLATHURAI E, YONINA C. Deep networks for direction-of-arrival estimation in low SNR. IEEE Trans. on Signal Processing, 2021, 69: 3714–3729.

[19] TIAN Q, CAI R Y, LUO Y, et al. DOA estimation: LSTM and CNN learning algorithms. Circuits, Systems, and Signal Processing, 2025, 44(1): 652–669.

[20] ALEXANDER B, ANN S, WOUTER T, et al. Exploiting temporal context in CNN based multisource DOA estimation. IEEE/ACM Trans. on Audio, Speech, and Language Processing, 2021, 29: 1594–1608.

[21] THEJA D, PULI, KISHORE K. Deep learning approach for high-resolution DOA estimation. International Journal of Ad Hoc and Ubiquitous Computing, 2024, 46(2): 90–103.

[22] ZHENG H, ZHOU C W, SERGIY A, et al. Decomposed CNN for sub-Nyquist tensor-based 2-D DOA estimation. IEEE Signal Processing Letters, 2023, 30: 708–712.

[23] WANG Q, LI S, GUO R Z, et al. One-shot architecture search and transformation for robust DOA estimation. IEEE Trans. on Aerospace and Electronic Systems, 2025, 61(2): 3642–3653.

[24] GUO Y, ZHANG Z, HUANG Y Z. Dual class token vision transformer for direction of arrival estimation in low SNR. IEEE Signal Processing Letters, 2024, 31: 76–80.

[25] LIU J C, WANG T Y, LI Y X, et al. A Transformer-based signal denoising network for AoA estimation in NLoS environments. IEEE Communications Letters, 2022, 26(10):

2336–2339.
[26] JI J K, MAO W, XI F, et al. TransMUSIC: a Transformer-aided subspace METhod for DOA estimation with low-resolution ADCS. Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2024. DOI: 10.1109/ICASSP48485.2024.10446722.
[27] FAN R, SI C K, YI W C, et al. YOLO-DoA: a new data-driven method of DoA estimation based on YOLO neural network framework. IEEE Sensors Letters, 2023, 7(2): 1–4.
[28] CHEN Y, WANG C, XIONG K L. Synchronized perturbation elimination and DOA estimation via signal selection mechanism and parallel deep capsule networks in multipath environment. Chinese Journal of Aeronautics, 2021, 34(12): 158–170.
[29] ZHANG K Z, XU L, FENG Z Y. A novel automatic modulation classification method based on dictionary learning. China Communications, 2019, 16(1): 176–192.
[30] SUFYAN D, ASFANDYAR K, DANG M L, et al. Metaverse applications in bioinformatics: a machine learning framework for the discrimination of anti-cancer peptides. Information, 2024, 15(1): 48.
[31] BREIMAN L. Random forests. Machine Learning, 2001, 45(1): 5–32.
[32] CHEN T, GUESTRIN C. XGBoost: a scalable tree boosting system. Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016. DOI: 10.1145/2939672.2939785.