Nuclear Science and Techniques (2025) 36:42
https://doi.org/10.1007/s41365-024-01595-y

Identification algorithm of low-count energy spectra under short-duration measurement based on heterogeneous sample transfer

Hao-Lin Liu[1,2] · Hai-Bo Ji[2] · Jiang-Mei Zhang[3] · Jing Lu[4]

Received: 30 January 2024 / Revised: 14 April 2024 / Accepted: 22 April 2024 / Published online: 28 January 2025

Abstract

In scenarios such as vehicle radiation monitoring and unmanned aerial vehicle radiation detection, rapid measurements using a NaI(Tl) detector often result in low photon counts, weak characteristic peaks, and significant statistical fluctuations. These issues can lead to potential failures in peak-searching-based identification methods. To address the low precision associated with short-duration measurements of radionuclides, this paper proposes an identification algorithm that leverages heterogeneous spectral transfer to develop a low-count energy spectral identification model. Comparative experiments demonstrated that transferring samples from 26 classes of simulated heterogeneous gamma spectra aids in creating a reliable model for measured gamma spectra. With only 10% of target domain samples used for training, the accuracy on real low-count spectral samples was 95.56%. This performance shows a significant improvement over widely employed full-spectrum analysis methods trained on target domain samples. The proposed method also exhibits strong generalization capabilities, effectively mitigating overfitting issues in low-count energy spectral classification under short-duration measurements.

Keywords Radionuclide identification · Low-count · Gamma energy spectral analysis · Heterogeneous · Transfer learning

1 Introduction

In recent years, rapid advancements in nuclear science and technology have led to diverse applications across various domains [1–4]. However, radioactive contamination from nuclear weapons and accidents in the nuclear industry has long-term and significant impacts on the environment, ecology, and biological health [5–8]. Therefore, detecting and identifying radionuclides is crucial, particularly in scenarios such as vehicle radiation monitoring and unmanned aerial vehicle radiation detection [9, 10]. Developing more powerful and effective identification algorithms is essential for detecting and identifying low-count energy spectra in short-duration measurements.

Traditional nuclide identification methods typically search for characteristic peaks and then match these peaks' energy information with a standard nuclide library [11]. In contrast, radioactive nuclide identification methods based on supervised machine learning focus on pattern recognition [12–15]. Machine learning often assumes that all samples in the input space follow an underlying unknown distribution, with both training and test samples independently

This work was supported by the National Defense Fundamental Research Project (No. JCKY2022404C005), the Nuclear Energy Development Project (No. 23ZG6106), the Sichuan Scientific and Technological Achievements Transfer and Transformation Demonstration Project (No. 2023ZHCG0026), and the Mianyang Applied Technology Research and Development Project (No. 2021ZYZF1005).

✉ Jing Lu
lujing_017@live.cn

[1] School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China

[2] Department of Automation, University of Science and Technology of China, Hefei 230026, China

[3] Fundamental Science on Nuclear Wastes and Environment Safety Laboratory, Southwest University of Science and Technology, Mianyang 621010, China

[4] School of Automation and Information Engineering, Sichuan University of Science and Engineering, Zigong 643000, China

drawn from this distribution (i.e., independently and identically distributed (IID)). Effective learning usually requires a large amount of labeled data [16].

However, this assumption often proves challenging to meet in practical scenarios. It is difficult to obtain radionuclide energy spectra with exactly the same data distribution. Real measured spectra are influenced by various factors such as environmental conditions, equipment, and the radiation source. Additionally, strict national regulations on the management of radioactive sources pose risks of radiation exposure to experimenters and require long-duration measurements for effective sample acquisition [17]. The high threshold for obtaining data through detector measurements of radioactive sources results in a limited number of labeled samples available for research. Additionally, the limited availability of publicly accessible datasets for energy spectra of radioactive nuclides measured using detectors in real-world scenarios, coupled with the expensive and time-consuming process of labeling vast amounts of unlabeled data, further compounds these challenges [18]. Variations in measurement environments, equipment, target radioactive sources, and measurement durations introduce potential spectral distortions due to background interference and statistical fluctuations [19]. These variations make it difficult to ensure consistent distributions, especially in cases of obscure gamma spectral characteristics under low-count conditions, which can render conventional methods ineffective [20, 21].

In recent years, transfer learning has relaxed two conventional constraints of pattern recognition, enabling the use of existing knowledge to address challenges in new tasks. It facilitates the recognition and prediction of data from different heterogeneous domains [22].

To address the challenge of identifying low-count gamma energy spectra in rapid measurement scenarios using NaI(Tl) detectors, this paper proposes a novel identification algorithm for low-count energy spectra under short-duration measurement based on heterogeneous sample transfer. To mitigate transfer difficulties arising from distinct domain distributions, this method aligns samples from both the source and target domains into a unified subspace characterized by discriminative feature representation. This alignment function leverages the inherent distribution attributes of energy spectrum data and emphasizes the aggregation of multiple regions of interest ratios. Additionally, to effectively augment target domain samples, this method aligns source domain labels by calculating the distances between source domain samples and the centroids of target domain categories using class mean measurements. A decision tree model, known for its resilience to outliers, is used as the classification architecture. The tree structure vividly and intuitively illustrates the significance of spectral features and the decision-making process.

The remainder of this paper is organized as follows. In Sect. 2, we introduce a novel identification algorithm for low-count energy spectra under short-duration measurements based on heterogeneous sample transfer. Section 3 presents a series of comparative experiments and provides the corresponding analysis. Section 4 summarizes the findings and conclusions drawn from this study.

2 Method

The proposed method consists of three major steps: (a) To address the challenge of negative transfer due to significant disparities between the source and target domain distributions, the method uses an alignment function to project both domains into a unified subspace. (b) To effectively augment target domain samples, the method employs distance measurement techniques along with clustering methods to reassign class labels to source domain samples. This reassignment is based on the aligned characteristics and true labels of a small number of target domain samples in the unified subspace. (c) Employing a decision tree model, known for its robustness against outliers, for classification. The hierarchical structure of the decision tree provides a clear and intuitive representation of feature importance within spectra and the decision-making process. Figure 1 illustrates the procedure of the proposed method.

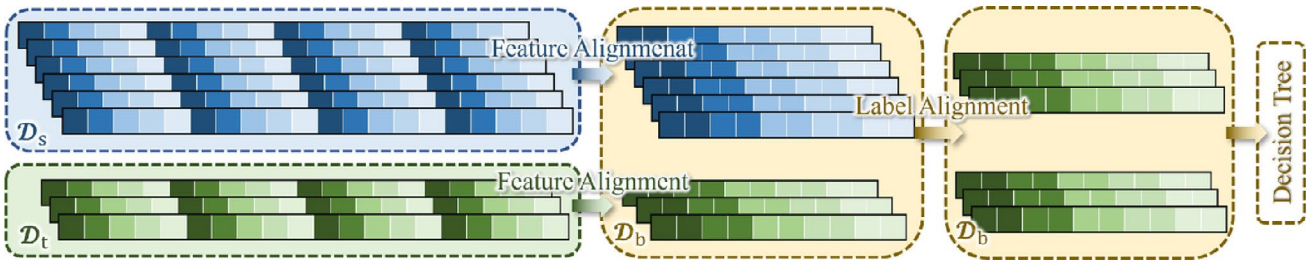

Fig. 1 (Color online) Overall process of the proposed method. $\mathcal{D}_s$: source domain, $\mathcal{D}_t$: target domain, $\mathcal{D}_b$: unified subspace. The figure uses a color palette with distinct schemes and arrangements of squares to represent samples from different domains. Each color scheme indicates a unique label space, while the varied color arrangements denote different feature distributions

In the context described in this study, the problem can be abstracted as follows: Given a source domain $\mathcal{D}_s = \{\boldsymbol{x}_i, y_i\}_{i=1}^{N_s}$ and a target domain $\mathcal{D}_t = \{\boldsymbol{x}_j, y_j\}_{j=1}^{N_t}$, where $\boldsymbol{x} \in \mathcal{X}, y \in \mathcal{Y}$. The source and target domains have the same feature spaces and different label spaces, that is, $\mathcal{X}_s = \mathcal{X}_t$ and $\mathcal{Y}_s \neq \mathcal{Y}_t$. As the total number of samples N_t in the target domain is limited, achieving a high classification accuracy using only $\mathcal{D}_t$ is unattainable. The goal of this method is to use the source and target domain data to learn a predictive function f that can achieve the minimum prediction risk in the target domain (evaluated by $\ell(f(\boldsymbol{x}), y)$).

To achieve this, it is essential to select a subset of the samples from the source domain. This selection aims to ensure that the probability distribution formed by the selected samples aligns closely with the probability distribution of the target domain. However, it is not possible to intuitively ascertain in advance which data within $\mathcal{D}_s$ will be beneficial for training the target model, when the probability distributions of $\mathcal{D}_s$ and $\mathcal{D}_t$ were different, i.e., $P_s(\boldsymbol{x}, y) \neq P_t(\boldsymbol{x}, y)$.

For clarity in subsequent discussions, the aforementioned method was further abstracted as a problem related to model training. The training dataset is denoted as $\boldsymbol{T} = \boldsymbol{T}_s \cup \boldsymbol{T}_t$, where $\boldsymbol{T}_s \subseteq \mathcal{D}_s$ represents the labeled training samples from the source domain, and $\boldsymbol{T}_t \subseteq \mathcal{D}_t$ represents the labeled training samples from the target domain. Specifically, $\boldsymbol{T}_s = \{\boldsymbol{x}_i^s, y_i^s\}_{i=1}^n$, where $\boldsymbol{x}^s \in \mathcal{X}_s$ and $y^s \in \mathcal{Y}_s$; $\boldsymbol{T}_t = \{\boldsymbol{x}_j^t, y_j^t\}_{j=1}^m$, where $\boldsymbol{x}^t \in \mathcal{X}_t$ and $y^t \in \mathcal{Y}_t$, with $m \ll n$. The test dataset is denoted as $\boldsymbol{S} \subseteq \mathcal{D}_t$. It should be noted that $\boldsymbol{S} \cap \boldsymbol{T}_t = \varnothing$, indicating that the target domain samples used for training and those used for testing are entirely distinct.

This method addresses a scenario involving a small set of labeled training data $\boldsymbol{T}_t$ from a specific distribution, a larger set of differently distributed labeled training data $\boldsymbol{T}_s$, and some unlabeled test data $\boldsymbol{S}$. The objective of this method is to train a classifier $\hat{f} : \boldsymbol{x} \mapsto y^t$ to minimize prediction errors on the unlabeled dataset $\boldsymbol{S}$.

2.1 Feature alignment

To address the issue of poor transferability caused by differing distributions between heterogeneous domains in sample transfer learning, i.e., $P_s(\boldsymbol{x}, y) \neq P_t(\boldsymbol{x}, y)$, directly using source domain data for target domain model training may degrade model performance. The proposed method utilizes the distribution characteristics of gamma energy spectrum data by employing linear feature transformation to align samples from different domains into a shared subspace with discriminative feature representation. The focus is on aggregating the ratios of multiple regions of interest to represent energy spectra from different domains, thereby reducing distribution differences and effectively enhancing sample transferability across domains. Figure 2 depicts the schematic diagram of feature alignment.

A random sample from the source and target domains is represented as $\{\boldsymbol{x}, y\}$. The energy information of characteristic peaks and auxiliary peaks of the target domain serves as prior knowledge. The feature alignment strategy of the proposed method is defined in Table 1.

The projection of a sample $\boldsymbol{x}$ from both the source and target domains onto a lower-dimensional shared subspace (denoted as $\mathcal{D}_b$) is represented as $\boldsymbol{x}^b$ (abbreviated as SC, for scaled photon counts). The formula for the projection is given in Eq. 1. Here, d is the feature dimensionality of $\boldsymbol{x}$, with $d > 0$. The scale parameter c acts as a scaling factor to adjust the magnitude of dimension changes. Optimizing energy spectrum samples involves adjusting the scale parameter c, which not only reduces the number of features

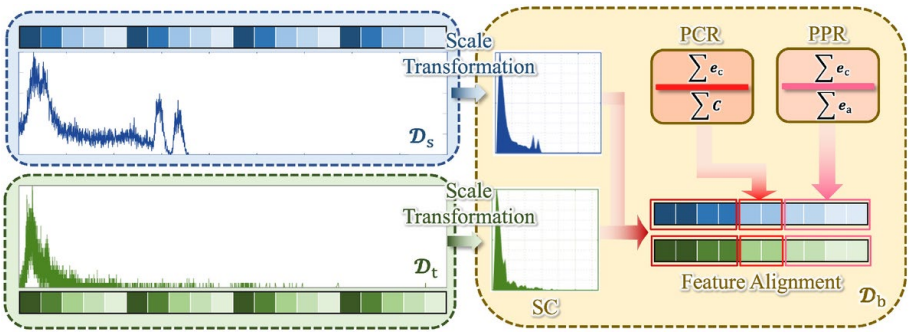

Fig. 2 (Color online) Schematic of feature alignment. Distribution characteristics of gamma energy spectrum data, including SC, PCR and PPR, were leveraged to align samples from different domains into a common subspace with discriminative feature expression, where, SC represents the projection of samples from both source and target domains into a shared lower-dimensional subspace. PCR denotes the ratios of characteristic energy values to the corresponding areas of the Compton continuum. In contrast, PPR refers to the ratios of characteristic energy values to energy values in an auxiliary set. The different colors of gamma energy spectra in this figure indicate samples originating from distinct domains. By applying scale transformations and aggregating ratios from multiple regions of interest, samples from different domains show comparable feature distributions

Table 1 Pseudo-code for feature alignment

Algorithm 1	*Feature alignment method*
Input	
	$\{\boldsymbol{x}, y\}$: Sample from $\mathcal{D}_s$ and $\mathcal{D}_t$
	c: Energy scale parameter
	$\boldsymbol{e}_c$: Characteristic energies of $\mathcal{D}_t$
	$\boldsymbol{e}_a$: Auxiliary energies of $\mathcal{D}_t$
Begin	
	1. $\boldsymbol{x} \longrightarrow \boldsymbol{x}^b = \{x_1^b, x_2^b, \cdots, x_{d/2^c}^b\}$
	2. **For** e in $\boldsymbol{e}_c$
	$r^c = \frac{x_i^a}{\sum_{j\in[c_l,c_r]}(x_j^a)}$
	end
	3. **For** e' in $\boldsymbol{e}_a$
	For e in e_c
	$r^p = \frac{x_i^a}{x_j^a}$
	end
	end
End	
Output	
	Sample in $\mathcal{D}_b$: $\{\boldsymbol{x}^b, \boldsymbol{r}^c, \boldsymbol{r}^p, y\}$

while preserving essential information within the energy spectrum, facilitating dimensionality reduction, but also mitigates issues arising from statistical fluctuations in low-count energy spectra. The range of c is $0 \leqslant c \leqslant \lfloor \log_2 d \rfloor$.

$$x_i^b = \sum_{i=1}^{\lfloor d/2^c \rfloor} \left[x_{2^c \cdot i - 2^c + 1}, x_{2^c \cdot i} \right] \tag{1}$$

significant statistical fluctuations were common in the characteristics of instrumental spectra due to limited photon counts within the full-energy peak during short-duration measurements. The channel with the highest counts may not necessarily correspond to the expected value of a Gaussian distribution [23, 24]. To mitigate the impact of these statistical fluctuations, spectra were transformed into multiple energy bins along the energy axis. Each bin represented an energy interval, and counts within each interval were aggregated to form a new feature vector. Fine-tuning the scaling parameter effectively reduced the feature dimensionality of energy spectra, thereby decreasing computational complexity and avoiding overfitting.

Multiple specific regions of interest in $\boldsymbol{x}^b$ were selected based on the decay properties of radionuclide materials in the target domain. These regions corresponded to the theoretical Compton continuum, characteristic peaks, and auxiliary peaks. The energy values and corresponding emissivities of photons emitted during the decay process of radionuclides in the target domain were considered essential prior knowledge. Photon energy values with emissivities greater than 20% were defined as characteristic energies, while those with emissivities greater than 1% but less than 20% were defined as auxiliary energies. Photon energy values with emissivities less than 1% were not considered due to their low emissivity, which resulted in low count values at corresponding points in the energy spectra.

Assuming that the characteristic energy values of nuclides in the target domain are $\boldsymbol{e}_c$ and the auxiliary energy values are $\boldsymbol{e}_a$, the peak-to-continuum ratio (PCR) for each energy value e in $\boldsymbol{e}_c$ was calculated using Eq. (2). Here, $i = \left\lfloor \frac{e}{E} \times \frac{d}{2^c} \right\rfloor$. The boundaries of the Compton continuum are denoted by $c_l = 1$ and $c_r = \left\lfloor \frac{e}{E} \times \frac{d}{2^c} \right\rfloor - 1$. E represents the energy range of the spectra. It should be noted that if e is too low to determine the boundary of the Compton continuum, r^c is set to 1 to avoid errors.

$$r^c = \frac{x_i^b}{\sum_{j\in[c_l,c_r]}(x_j^b)} \tag{2}$$

For each energy value e in $\boldsymbol{e}_c$ and for each energy value e' in $\boldsymbol{e}_a$, PPR (abbreviation of peak-to-peak ratio) was calculated by Eq. (3), where $i = \left\lfloor \frac{e}{E} \cdot \frac{d}{2^c} \right\rfloor$, and $j = \left\lfloor \frac{e'}{E} \cdot \frac{d}{2^c} \right\rfloor$.

$$r^a = \frac{x_i^b}{x_j^b} \tag{3}$$

In general methods, energy values of photoelectric peaks are commonly used as primary characteristics for identifying radionuclides, while the Compton continuum is rarely considered [25, 26]. However, identifying radionuclides solely through characteristic peaks may be ineffective under complex background conditions [27, 28]. From a macroscopic perspective, PCR characterizes the ability to detect low-energy weak peaks amid complex backgrounds, while PPR quantifies the likelihood of gamma rays undergoing various interactions within the detector and ultimately contributing to the full-energy peaks.

The formation mechanism and measurement principles of the gamma-ray instrument spectrum served as the basis for feature alignment. The alignment function utilized the inherent distribution characteristics of energy spectra, focusing on aggregating ratios from multiple regions of interest. By converting both source and target domain samples into two types of ratios, the distribution gap between them was reduced. This function was beneficial for analyzing and interpreting the physical significance of energy spectra.

2.2 Label alignment

Due to the challenge of expanding target domain samples caused by the heterogeneous label space in sample transfer

learning (i.e., $\mathcal{Y}_s \neq \mathcal{Y}_t$), directly using source domain data for target domain model training was not feasible. This paper employed discriminative features in a shared subspace to measure the distance between source domain samples and the centroids of target domain categories using class mean measurement. Additionally, source domain labels were updated through label propagation, which leveraged the discriminative structural information within the source domain data. Figure 3 illustrates the schematic diagram of label alignment.

Let $\boldsymbol{T}_t = \{\boldsymbol{x}_j^t, y_j^t\}_{j=1}^m$ denote a set of m target domain samples, where $\boldsymbol{x}_j^t$ represents the j-th sample and y_j^t denotes its corresponding label. $\boldsymbol{U}$ represents the set of non-repeating labels in $\mathcal{Y}_t$, as defined by Eq. 4.

$$\boldsymbol{U} = \text{unique}(\mathcal{Y}_t) \tag{4}$$

For each non-repeating label $U_k \in \boldsymbol{U}$, $\boldsymbol{T}_k$ represents the index set of all the samples in the target domain with label U_k.

$$\boldsymbol{T}_k = \{\{\boldsymbol{x}_j^t, y_j^t\} \mid y_j^t = U_k\} \tag{5}$$

Compute the centroid vector $\boldsymbol{\mu}_k$ corresponding to label U_k using Eq. 6, where $|\boldsymbol{T}_k|$ denotes the number of samples with label U_k.

$$\boldsymbol{\mu}_k = \frac{1}{|\boldsymbol{T}_k|} \sum_{y_j^t = U_k} \boldsymbol{x}_j^t \tag{6}$$

Furthermore, $\boldsymbol{T}_s = \{\boldsymbol{x}_i^s, y_i^s\}_{i=1}^n$ denotes a set of n source domain samples, where $\boldsymbol{x}_i^s$ represents the i-th sample and y_i^s denotes its corresponding label. To effectively align source domain labels, efficient clustering of these n source domain samples is crucial. The samples are partitioned into k clusters, each centered around centroids $\boldsymbol{\mu}_k$ ($k \ll n$), to minimize the Within-Cluster Sum of Squares (WCSS). The key to this process is identifying clusters $\boldsymbol{O}_k$ that satisfy Eq. 7.

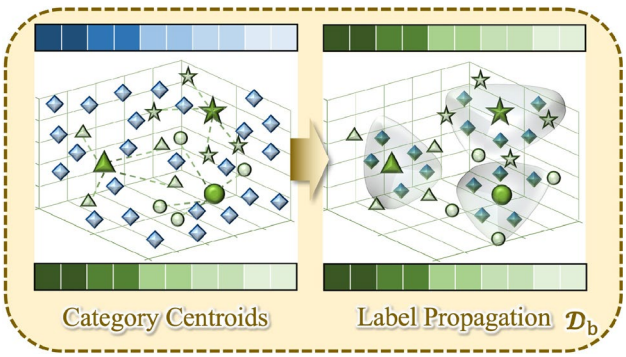

Fig. 3 (Color online) Schematic of label alignment. The various markers colors denote samples originating from distinct domains. The source domain labels were updated via label propagation by computing centroids for target domain categories and assessing the distance between source domain samples and these centroids

$$\arg\min_{\boldsymbol{O}} \sum_{k=1}^{|U|} \sum_{\boldsymbol{x}^s \in \boldsymbol{O}_k} \left\|\boldsymbol{x}^s - \boldsymbol{\mu}_k\right\|^2 \tag{7}$$

For each $\boldsymbol{x}_i^s$, a distance threshold θ_k was used to ensure compliance with Eq. 8. Samples that did not meet this criterion were excluded from consideration.

$$\left\|\boldsymbol{x}_i^s - \boldsymbol{\mu}_k\right\|_{x^s \in \boldsymbol{O}_k}^2 \leq \theta_k \tag{8}$$

Assigning a novel class label to each retained source domain sample was crucial for its role in training the target model. The reallocation of source domain labels was based on the cluster to which each sample belonged, as determined by class mean measurement. Specifically, the label was set to k if $\boldsymbol{x}_i^s \in \boldsymbol{O}_k$.

The purpose of this process was to extract meaningful samples from the source domain to aid in training the target model through clustering and distance measurement methods. This approach improved the effectiveness of sample transfer. Specifically, samples from the target domain training set were used as initial centroid points for clustering. The similarity between each sample in the source domain training set and these centroid points was measured using distance metrics, such as Euclidean or Mahalanobis distance, depending on the problem's requirements. Samples with larger distances, indicating significant differences in feature distribution from the target task, were deemed less suitable for training and were discarded. After identifying the retained samples, new class labels were assigned based on their distances to the centroid points.

A decision tree architecture was used for classifying low-count energy spectra with small sample sizes. Specific parameter settings were implemented to control the tree's structure. Notably, the maximum depth of the tree was limited to 20, establishing the maximum number of nodes from the root to the farthest leaf. This constraint helped manage the tree's complexity and reduce overfitting, particularly beneficial for small sample sizes [29]. The Gini diversity index was used as the splitting criterion for node partitioning. This choice aimed to minimize the presence of interfering radionuclides in each node, enhancing the classification purity of the resulting subsets. By focusing on creating nodes with instances predominantly from a single class, the model's accuracy in classifying low-count gamma-ray spectra was improved [30]. This parameter configuration effectively balances the model's complexity and performance, leading to better classification accuracy, especially in scenarios with low-count energy spectra under short-duration measurements.

3 Experiments and analysis

This section describes two primary components of spectral data: synthetic and measured data. The source domain corresponded to the synthetic data generated through Monte Carlo simulations, whereas the target domain corresponded to the measured spectral data obtained from the detectors.

This section is comprised of six sections that demonstrate the feasibility of the proposed method for radionuclide identification through a series of experiments. Part A introduces the acquisition and preprocessing of both synthetic and measured data. Random examples are presented to visually illustrate the comparison of the gamma energy spectra of various radionuclides originating from distinct domains. Part B involved a comparative analysis of the gamma energy spectra of different radionuclides from diverse domains under various scale parameters. Part C conducted calculations and analyses of PCR and PPR individually using authentic samples in the target domain. Part D compares the various effects resulting from employing different features as the alignment basis and adjusting distinct thresholds during the label alignment process. Section E compares the classification effects of several commonly used methods with varying proportions of actual target domain samples participating in the training process. Part F conducted classification experiments on multiple sets of individual features and their aggregations, with a subsequent comparison of their classification effects.

3.1 Data preparation

This section introduces the acquisition and preprocessing of both synthetic and measured data. The synthetic data generated through the Monte Carlo simulations served as the source domain, whereas the measured spectral data obtained from the detectors were employed as the target domain.

First, the generation of synthetic data encompasses a three-step process: acquisition of background data, generation of single-nuclide energy spectra, and synthesis of data.

Step 1. Acquisition of background data: A custom 3-inch NaI detector was employed to measure ambient background radiation in a laboratory setting devoid of external radioactive sources. The detector remained stationary for 12 h, and two measurements were conducted: one with lead bricks surrounding the detector and the other without lead bricks, yielding two distinct sets of background data.

Step 2. Generation of single-nuclide energy spectra The Monte Carlo method was used to simulate the transport of gamma-ray particles emitted from 26 individual radionuclides. The simulated scenarios involved various single radioactive point sources and a stationary 3-inch NaI detector. One million photons were simulated, and their trajectories were tracked by the detector, ultimately producing gamma-ray energy spectra.

Step 3. Synthesis of data: Linear combinations were performed on two sets of background spectra and 26 synthetic radioactive spectra using randomly assigned signal-to-noise ratios (SNR). The SNR was computed as $\mathrm{SNR} = N_{\mathrm{nc}}/N_{\mathrm{bg}}$, where N_{nc} represents the sum of the photon counts emitted by the radioactive source and N_{bg} denotes the sum of the background photon counts. For linear superposition, SNR values were randomly generated within a range of 0.3 to 1.

This process generated a total of 15,600 synthetic gamma-ray energy spectra for 26 radionuclides. Table 2 lists common radionuclides, including special nuclear materials (SNM), commonly used nuclear materials in industry and medicine, and naturally occurring radioactive materials (NORM).

Measurement experiments in this study used the AT6104DM spectrometer from Atomtex. The detection unit, housed in a seismic-resistant and waterproof stainless steel container, featured a NaI(Tl) crystal measuring Φ 63 mm × 63 mm to record gamma radiation from controlled radionuclides, with a typical resolution of 7.5% at 662 keV (^{137}Cs). The spectrometer's detection energy range extended from 70 keV to 3 MeV.

Measurement experiments were conducted in a laboratory setting. A specific open area within the laboratory was chosen as a fixed reference point. A ruler was placed on the ground with the radiation source positioned at the origin of the ruler. The spectrometer was then moved to vary parameters such as the relative distance from the source and the measurement angle. Two V-type radiation sources, ^{137}Cs and ^{60}Co, as well as one IV-type radiation source, ^{152}Eu, were used in the measurements, with activities of 3.7×10^5 Bq, 3.7×10^5 Bq, and 3.7×10^8 Bq, respectively. To evaluate the proposed method's accuracy in identifying isotopes under shorter measurement times, the measurement duration was limited to between 1 and 5 s.

Table 2 Radionuclide library of synthetic sample

Type	Radionuclide
SNM	^{237}NP, ^{233}U, ^{235}U, ^{238}U
Industrial	^{241}Am, ^{133}Ba, ^{57}Co, ^{60}Co, ^{137}Cs ^{152}Eu, ^{192}Ir, ^{75}Se
Medical	^{51}Cr, ^{18}F, ^{67}Ga, ^{123}I, ^{125}I, ^{131}I ^{111}In, ^{103}Pd, $^{99\,m}$Tc, ^{201}Tl, ^{133}Xe
NORM	^{40}K, ^{226}Ra, ^{232}Th

Table 3 Radionuclide library of measured sample

Nuclide type	Size	Capacity	Label
^{137}Cs	1024	100	1
^{152}Eu	1024	100	2
^{60}Co	1024	100	3

As depicted in Table 3, we acquired 300 gamma-ray energy spectra from three individual radiation sources. The dataset included 300 spectral samples encompassing 1024 discrete energy channels along with their respective event counts.

Random samples were selected from both the source and target domains, as illustrated in Fig. 4. All energy spectral data presented were normalized. For the simulated energy spectra, the total number of photon counts was 20,000, while for the measured energy spectrum, it ranged from approximately 400 to 700 counts. In the simulated experiments, the ample number of photons led to distinct features in the gamma energy spectrum (highlighted by red boxes). However, in practical measurements lasting 1 to 5 s, only the ^{137}Cs spectrum exhibited approximate characteristic peaks, while other gamma energy spectra lacked distinct features. This posed a challenge in identifying different peaks in the measured spectra.

3.2 Scale parameter

The scale parameter, denoted as c, was essential for optimizing the accuracy of energy spectrum transformation and played a crucial role in feature alignment. Its fine-tuning had a dual purpose: it enabled effective dimensionality reduction, providing a streamlined representation of spectral characteristics, and it reduced the impact of statistical fluctuations, enhancing the robustness and reliability of the alignment process.

To visually illustrate the effects of different scale parameter settings on spectral features, the scaling parameter was set to values of 0, 2, 4, and 6. Random samples were selected from both the source and target domains, and the scaled spectra are presented in Fig. 5. This graphical representation offers a comparative view of the variations induced by different scale parameter settings, serving as a valuable resource for researchers seeking a deeper understanding of energy spectrum transformations.

The observed spectra provided valuable insights into the nuanced effects of the scale parameter on energy spectrum transformation. At a scale parameter of 0, the original spectra showed pronounced statistical fluctuations. As the scale parameter increased, these fluctuations were notably suppressed while preserving prominent features. This improvement highlights an enhancement in the stability and reliability of the spectral data.

Short-duration measurements result in heightened statistical fluctuations in the target domain, making fine-tuning the scale parameter crucial for achieving meaningful alignment. Gradually increasing the scale parameter leads to a smoother representation of the energy spectra, reducing the impact of uncertainties introduced by short-duration measurements, particularly in the target domain.

However, observations at a scale parameter value of six revealed a potential risk. At this value, there was a significant loss of detailed features, retaining only broad trends. Although a larger scale parameter effectively suppressed statistical fluctuations, it also led to the loss of finer details in the spectra. This highlights the need for a careful balance in selecting the scale parameter, aiming to mitigate statistical fluctuations while preserving essential feature information. This finding is crucial for practical applications, especially in scenarios involving short-duration measurements in the target domain.

3.3 PCR and PPR

To achieve improved classification performance, prior knowledge was essential. It provided a better understanding of the changes and peak-valley features that radionuclides may exhibit during measurements. This understanding helped optimize classification algorithms, enhancing the accurate differentiation of radionuclides. By considering radiation characteristics, energy spectral shapes, and potential sources of noise outlined in prior knowledge, the classification model was finely tuned to adapt better to real measurement conditions. As shown in Table 4, the table includes the half-life, primary characteristic energies, and corresponding emission rates for the radionuclides of interest. This information formed the basis for further analysis and interpretation in the classification process.

Characteristic and auxiliary energy values were crucial for calculating the PCR and PPR. Leveraging the characteristic information of these target radionuclides, the characteristic energy values were defined as 122 keV, 344 keV, 662 keV, 1173 keV, 1332 keV, and 1408 keV, while auxiliary energy values were set at 245 keV, 411 keV, 444 keV, 779 keV, 867 keV, 964 keV, 1086 keV, 1090 keV, 1112 keV, 1213 keV, and 1299 keV. Taking the measured sample of ^{60}Co as an example, the characteristic peak was selected as 662 keV, with an auxiliary peak chosen as 245 keV. The energy range was set to 3000 keV, and the scale parameter was set to 0. Through calculations, the PCR value was determined to be 0.0016, and the PPR value was 0.3636.

To further illustrate the discriminability of PCR and PPR, we compared the intra-class and inter-class distances between the baseline samples and aligned samples based on

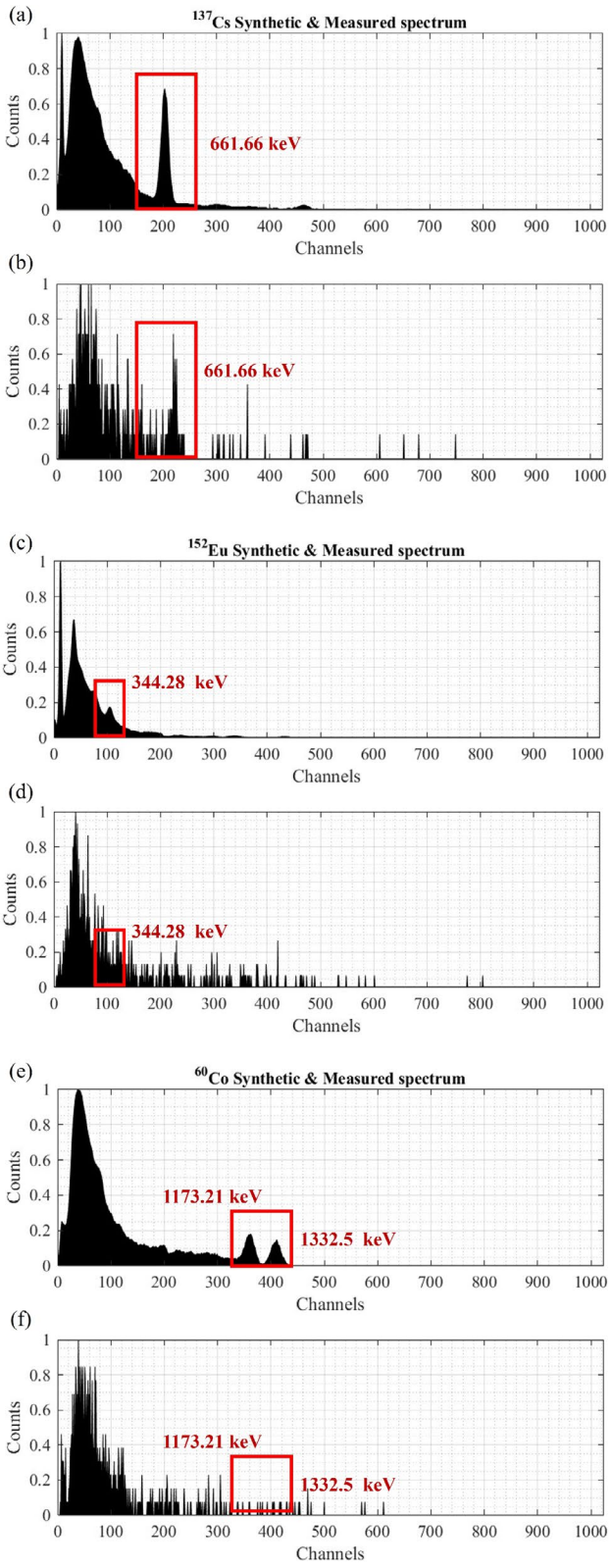

Fig. 4 (Color online) Random samples of three different radionuclides from the source (**a, c, e**) and target domains (**b, d, f**). Red boxes highlight the positions of characteristic peaks and their corresponding energy values. Distinct characteristic peaks are observable in (**a**), (**c**), and (**e**). However, due to the short-duration measurement, only the approximate shape of the characteristic peak can be discerned in (**b**), and no evident characteristic peaks are observable in (**d**) and (**f**)

features from the aforementioned ^{60}Co example, for both the synthetic and measured sample sets. However, because of the different dimensions of the baseline samples, PCR and PPR, it was not feasible to directly compare them using distance calculations. Therefore, we computed the intra- and inter-class distances between the randomly synthesized and measured samples. We then calculated the ratios of intra- and inter-class distances in the baseline samples, PCR, and PPR to make horizontal comparisons. The results are presented in Table 5.

Ideally, after feature alignment, intra-class distances should be minimized while inter-class distances should be maximized. Consequently, the ratio of intra-class to inter-class distances was minimized. The results in Table 5 show that the ratios of intra-class to inter-class distances for the initial samples were 0.7708 and 1.4092, respectively. For PCR, these ratios were 0.6426 and 0.8796, respectively, and for PPR, they were 0.0821 and 0.1655, respectively. All these values were smaller than those of the baseline samples. These results indicate that after feature alignment, both PCR and PPR improved the ratio of intra-class to inter-class distances. This suggests that PCR and PPR were more effective in reducing intra-class differences and enhancing inter-class differences, thereby improving the discriminative power of the model. Notably, PPR performed better than PCR in terms of the ratio of intra-class to inter-class distances. The smaller ratio for PPR may imply that it more effectively preserves key information during dimensionality reduction. This likely reflects that PPR better captures the intrinsic structure among samples during feature selection and projection, aiding in distinguishing between different sample categories.

3.4 Label alignment basis

In the label alignment process, using different features as alignment criteria and varying thresholds can produce distinct outcomes. In this part, each of the three feature sets-SC, PCR, and PPR-was used individually as alignment criteria, as well as their combined features, referred to as AGG.

Experiments were conducted by adjusting threshold intervals based on each feature's characteristics. The scaling parameter was fixed at 2 to ensure that 10% of target domain samples were included in training, and the classification feature was set to AGG. Figure 6 shows how the number of retained samples changed after aligning source

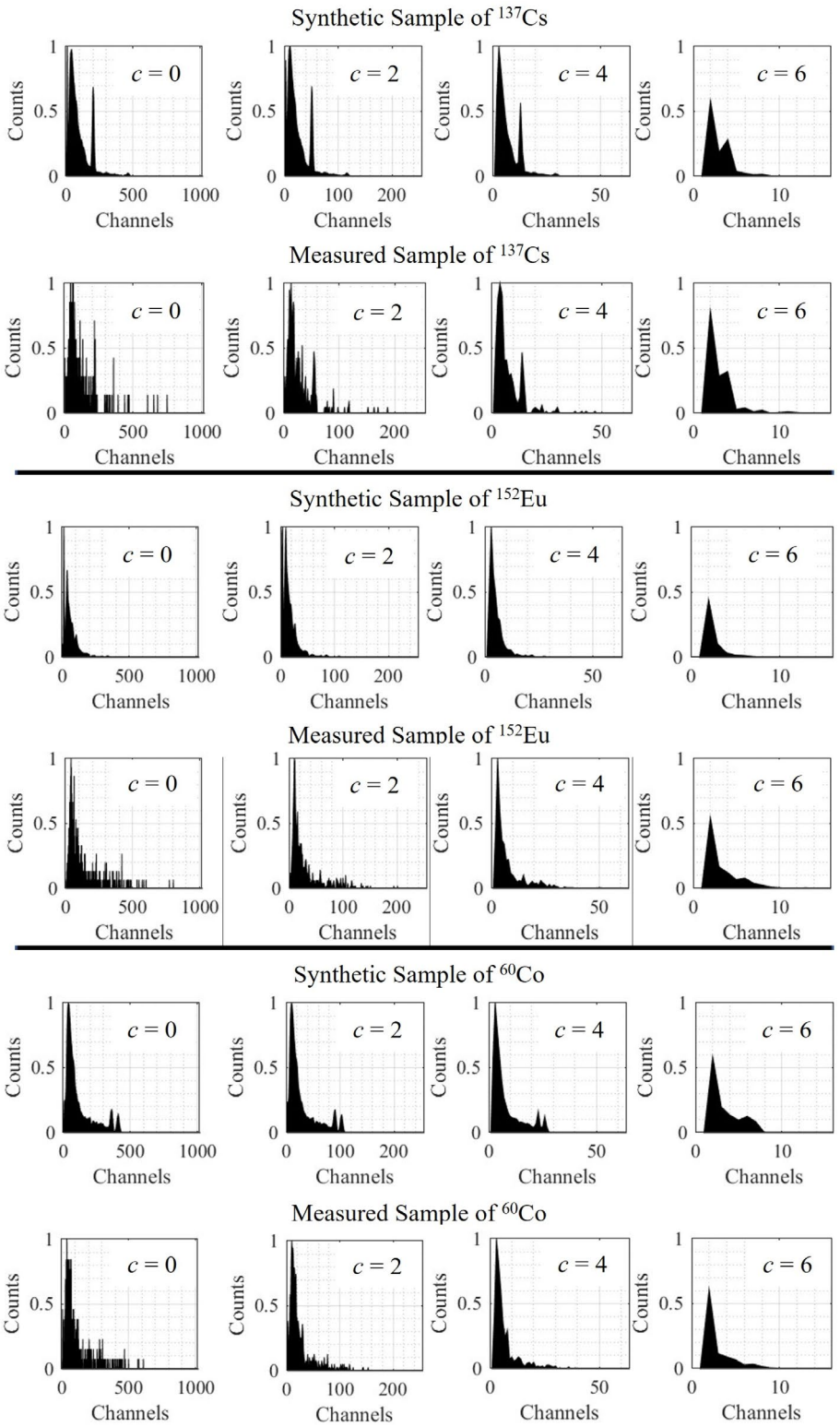

Fig. 5 (Color online) Synthetic and measured samples with scale parameters. When $c = 0$, the spectra were untreated, containing 1024 sample feature channels, exhibiting pronounced noise and strong statistical fluctuations. At $c = 2$, noise diminished, and the number of sample feature channels reduced to 256. With $c = 4$, the spectral noise was suppressed, revealing distinct characteristic peaks, while the number of sample feature channels decreased to 64. At $c = 6$, a significant loss of detailed spectral features occurred, leaving only the approximate shapes of the spectra, with the number of sample feature channels reduced to 16

domain samples using different features as alignment criteria. Figure 7 illustrates the trend in classification accuracy as thresholds varied, depending on which features were used as alignment criteria.

It is important to note that the x-axis in both figures represents variations in thresholds. For the SC feature, it covers 100 evenly spaced values within the interval [0,3]. For the PCR feature, it covers 100 evenly spaced values within

Table 4 Energy and emissivity of measured radionuclide

Nuclide	Half-life (a)	Energy (keV)	Emissivity (%)
^{60}Co	5.27	1173.21	99.98
		1332.5	99.87
^{137}Cs	30.17	661.66	85.21
		32.19	3.61
		31.82	1.96
		36.4	1.31
^{152}Eu	13.5	121.78	28.58
		344.28	26.5
		1408.01	21.005
		964.08	14.605
		1112.07	13.644
		778.90	12.942
		1085.87	10.207

Table 5 Comparison between intra-class and inter-class distance

Distance	Intra Co-Co	Inter Co-Cs	Inter Co-Eu	Intra/Inter Co-Co/Co-Cs	Intra/Inter Co-Co/Co-Eu
Baseline	2.9792	3.8653	2.7430	0.7708	1.4092
PCR	0.0169	0.0263	0.0299	**0.6426**	**0.8796**
PPR	8.8039	107.1756	647.5620	**0.0821**	**0.1655**

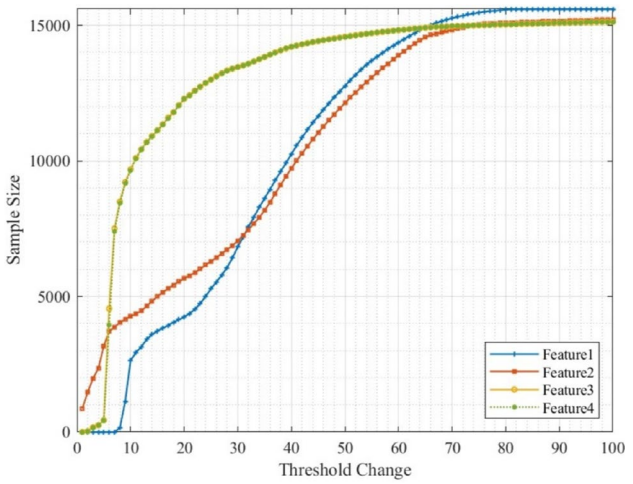

Fig. 6 (Color online) Variation in remaining samples after aligning source domain samples using different features as alignment criteria. Feature1 through Feature4 correspond to SC, PCR, PPR, and AGG, respectively

the interval [0,0.75]. This experimental design aimed to thoroughly understand how different features affect label alignment and classification performance, offering valuable insights for further model parameter optimization.

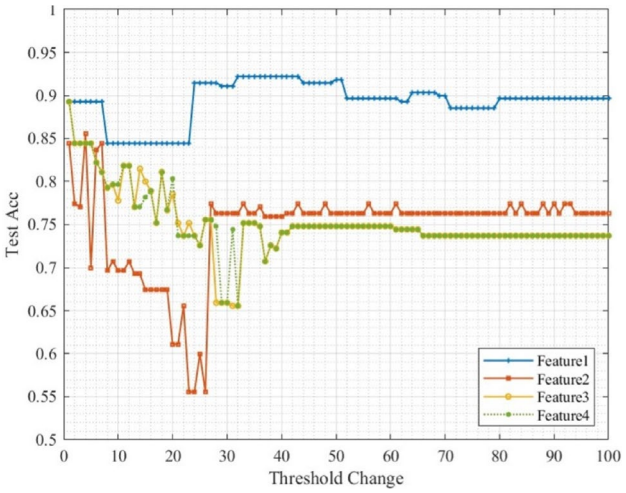

Fig. 7 (Color online) Classification testing accuracy at varying thresholds when different features were used as alignment criteria. Feature1 through Feature4 correspond to SC, PCR, PPR, and AGG, respectively

From Fig. 6, it is evident that as the threshold gradually increased, the number of source domain samples remaining after alignment also increased, eventually approaching the total number of samples in the source domain. Specifically, for the SC curve, the number of retained samples was approximately 2.57% to 3.11% higher compared to other scenarios once the number of retained samples stabilized.

The trends in the SC and PCR curves showed a similar pattern. The number of retained samples increased rapidly in the early stage (threshold change 1–10), leveled off in the mid-stage (threshold change 11–30), accelerated again in the late stage (threshold change 31–100), and eventually saturated. A noticeable intersection between the SC and PCR curves occurred around a threshold change of 30.

Conversely, the curves for PPR and AGG nearly overlap, showing a similar trend. The number of retained samples increased rapidly in the early stage (threshold change 1–10). Throughout most of the threshold changes, the quantity of retained samples consistently exceeded that in scenarios where SC and PCR were used as alignment references. This likely negatively impacted alignment performance due to the sharp increase in heterogeneous samples from the source domain.

Examining Fig. 7, a clear trend emerged showing that using SC as the alignment criterion resulted in overall superior test accuracy, consistently ranging between 84% and 92%. PCR, despite significant fluctuations in the early stages (threshold variation 1–10), eventually achieved a high classification accuracy of about 86%. However, during the mid-stage (threshold variation 11–30), accuracy gradually declined to around 56% and stabilized at approximately 76% in the later stages (threshold variation 31–100). In contrast,

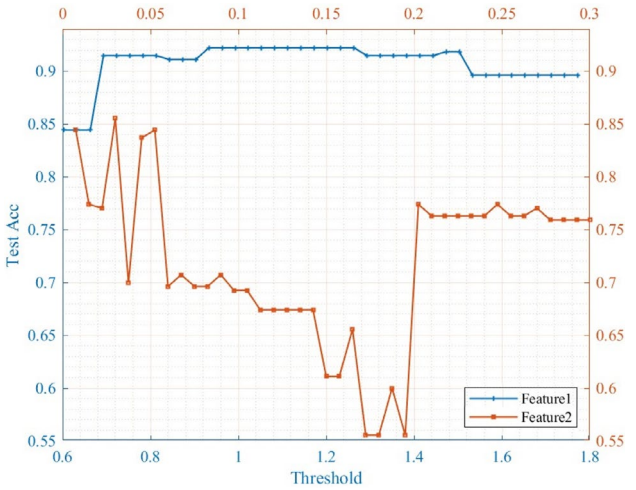

Fig. 8 (Color online) Local classification testing accuracy curves at varying thresholds using different features employed as alignment criteria. Feature1 and Feature2 correspond to SC and PCR, respectively. The blue (horizontal) and orange (vertical) axes represent the threshold intervals and the corresponding classification test accuracy for the two feature sets used as label alignment standards

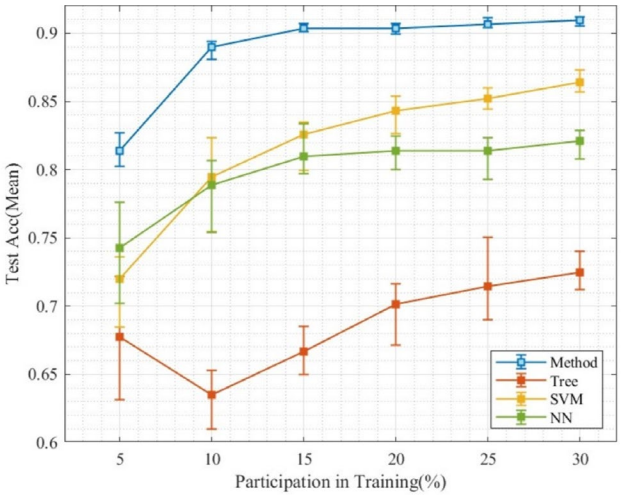

Fig. 9 (Color online) Average classification test accuracy by participation in training. Method: the proposed method in this study; tree: decision tree; SVM: support vector machine; NN: neural network

PPR and AGG showed overall declining trends with varying thresholds, indicating their limited effectiveness as discriminative criteria in the label alignment process.

Figure 8 highlights intervals where the test accuracy curves for SC and PCR show exceptional performance. Notably, SC demonstrated a significant advantage in the threshold range [0.96, 1.24], achieving an impressive test accuracy of 92.22%. In comparison, PCR performed best at a threshold value of 0.03, with a notable test accuracy of 85.56%. Clearly, SC was the most suitable alignment criterion.

3.5 Impact of proportion of target samples participating in training

To validate the advancements of the proposed method, we compared it with several widely used approaches [31–36]. All comparative methods used target domain data, starting with 30% of the target dataset samples in the training set. This proportion was gradually reduced from 30% to 5% to evaluate the classification performance of each method with varying amounts of target dataset samples used in training.

In the experimental procedure, the proportions of target dataset samples used in training were set at 5%, 10%, 15%, 20%, 25%, and 30% for 10 rounds of experiments. Each round consisted of 100 iterations, and the average test accuracy values across these iterations were calculated and analyzed. Additionally, each test involved five-fold cross-validation. The key classification feature was set to AGG, the scaling parameter was fixed at 2, and the feature used for filtering source domain samples was SC, with a threshold set to 1.

Three machine learning classification methods based on full-spectrum information were selected for comparison: (1) decision tree, (2) support vector machine (SVM), and (3) neural network. The parameter settings for each method were as follows: (1) decision tree: The maximum depth was set to 20, and the splitting criterion used was the Gini diversity index. (2) SVM: A polynomial kernel function was employed with a Box Constraint level of 1 to limit the penalty on observations that violate the margin, thereby reducing the risk of overfitting. Data normalization was applied to maintain consistent scales across different features. (3) Neural network: The network used a fully connected layer of size 25 as the first layer with a ReLU activation function. The maximum number of iterations was set to 1000, regularization term strength was set to 0, and data normalization was performed.

The experimental results, depicted in Fig. 9, comprehensively evaluated the classification effectiveness of the proposed method compared to traditional methods across varying proportions of target dataset samples in the training set. The iterative nature of the experiments ensured a thorough assessment of each method's performance. Statistical analysis of the average test accuracy values offered insights into the stability and peak performance of each method under different experimental conditions.

Figure 9 shows the average classification accuracy across multiple iterations as the proportion of target domain training samples varies from 5% to 30%. The results clearly demonstrate the significant effectiveness of the proposed method across different proportions of training sample involvement.

When the training sample size was limited to 5%, the average classification accuracy across all four methods showed only minor variations and remained low. In this context, the proposed method achieved an average accuracy of 81.39% in 1000 random tests, outperforming the other three methods by 13.62%, 9.41%, and 7.15%, respectively. As the proportion of training samples increased to 10%, the proposed method's average accuracy improved significantly to 88.95%, demonstrating its effectiveness in handling sample scarcity. With a larger training sample proportion of 30%, the proposed method achieved an average accuracy of 90.95% in 1000 random tests, surpassing the other methods by 18.46%, 4.57%, and 8.85%, respectively.

The figure also shows the maximum and minimum errors from 1000 random tests, represented by short lines above and below each data point. It is clear that the proposed method's errors were significantly lower than those of the comparison methods, as supported in Table 6. This indicates its overall high stability and minimal susceptibility to uncontrollable data fluctuations. While full-spectrum machine learning classification methods achieve high test accuracy at certain stages, their performance notably declines as the proportion of training data decreases, highlighting their limitations compared to the proposed method. These results underscore the exceptional performance of the proposed method in comparative experiments, emphasizing its robustness and ability to handle the challenges of identifying low-count spectra in short-term measurements.

When the proportion of training samples varies, the time required for each phase may differ. To assess these differences, experiments were conducted where the time spent on alignment, training, and testing was systematically recorded. The proportions of target dataset samples used for training were set at 5%, 10%, 15%, 20%, 25%, and 30%. The key classification feature was set as AGG, and the scaling parameter was fixed at 2. SC was used for filtering source domain samples, with a threshold of 1. The total number of source domain samples was 15,600, while the target domain samples numbered 300. The original feature dimension of the samples was 1,024. The results are presented in Table 7.

Table 7 lists the durations for each algorithmic phase, the total algorithm duration, and the average duration of each phase for various sample participation ratios. This comparison highlights variations in resource allocation and efficiency throughout the training process. In the alignment phase, the average duration was 1 s, suggesting some complexity in data preparation or feature extraction. The training phase averaged 0.3 s, while the testing phase averaged 0.01 s, demonstrating rapid inference capability. Despite the differences in duration across phases, the overall efficiency in model training and testing was commendable.

Table 6 Average classification accuracy

Proportion (%)	Tree (%)	SVM (%)	NN (%)	Method (%)
5	67.77±4.64	71.98±3.52	74.24±4.03	**81.39**±1.28
10	63.49±2.48	79.47±4.02	78.86±3.44	**88.95**±0.85
15	66.69±1.86	82.58±2.64	80.94±2.40	**90.36**±0.35
20	70.15±3.02	84.31±1.67	81.41±1.38	**90.38**±0.47
25	71.48±3.55	85.23±0.78	81.37±2.11	**90.67**±0.45
30	72.49±1.56	86.38±0.93	82.10±1.31	**90.95**±0.44

3.6 Impact of features on classification performance

To assess the impact of individual features and their combinations on classification performance, we conducted four sets of experiments. Our goal was to determine the specific contributions of each feature group to classification tasks. In these experiments, we used a fixed scaling parameter of 2

Table 8 Validation accuracy of features

	Feature1	Feature2	Feature3	Feature4
1	**0.9936**	0.9925	0.9922	**0.9936**
2	0.9913	0.9919	0.9933	**0.9945**
3	0.9919	**0.9948**	0.9933	0.9936
4	0.9933	0.9942	0.9945	**0.9962**
5	0.9928	0.9945	0.9933	**0.9957**
6	**0.9948**	0.9913	0.9910	0.9939
7	**0.9942**	0.9928	0.9936	0.9936
8	0.9928	0.9942	0.9925	**0.9951**
9	0.9925	0.9931	0.9942	**0.9959**
10	**0.9936**	0.9928	0.9933	0.9933
11	0.9945	**0.9962**	0.9922	0.9933
12	0.9933	0.9939	0.9928	**0.9939**
Mean	0.9932	0.9935	0.9930	**0.9944**

Table 7 Temporal dynamics of efficiency across varying sample participation ratios

Proportion (%)	5	10	15	20	25	30	Mean
Align(s)	0.9974	1.0292	0.9977	1.0416	0.9806	1.0333	1.0133
Train(s)	0.3456	0.3112	0.3392	0.3388	0.3287	0.3512	0.3358
Test(s)	0.0144	0.0129	0.0125	0.0198	0.0186	0.0159	0.0157
Sum(s)	1.3574	1.3534	1.3494	1.4001	1.3279	1.4004	1.3648

Table 9 Testing accuracy of features

	Feature1	Feature2	Feature3	Feature4
1	0.5963	0.6593	0.6889	**0.8037**
2	0.5630	0.6111	0.7481	**0.8074**
3	0.5963	0.7111	0.7778	**0.8111**
4	0.6741	0.6926	0.7185	**0.8148**
5	0.6444	0.6889	0.7259	**0.8185**
6	0.6222	0.6074	0.6778	**0.8222**
7	0.6333	0.6815	0.7259	**0.8259**
8	0.6111	0.7111	0.7296	**0.8296**
9	0.7333	0.6481	0.6704	**0.8333**
10	0.6630	0.6333	0.6259	**0.8370**
11	0.6407	0.5481	0.6185	**0.8407**
12	0.6889	0.6630	0.7481	**0.8444**
Mean	0.6389	0.6546	0.7046	**0.8241**

and set the target domain sample participation ratio in training to 10%. We employed PCR for filtering source domain samples, with a threshold of 0.04. This standardized design ensured a fair comparison of each feature group's performance, providing a consistent baseline for evaluating the effects of both individual and combined features on classification outcomes.

Table 8 shows the validation accuracy of the four features across 12 experiments. Notably, the validation accuracy remained consistently high due to the substantial proportion of source domain samples in both the training and validation sets. This trend was observed for all feature groups, indicating high accuracy on the validation set for each group. However, this high validation accuracy should be interpreted with caution, as it may not accurately reflect the model's performance on target domain samples or reveal potential variations in the generalization of different feature groups to the target domain.

Table 9 presents the testing accuracy of the four features across 12 experiments. The results show that the joint features significantly outperformed other scenarios, indicating that this feature group effectively captured and integrated information from the preceding three feature groups. This integration not only improved classification performance but also highlighted the synergistic effects among the features. The findings emphasize the importance of feature fusion in enhancing model generalization and classification accuracy.

4 Conclusion

This study proposed a novel algorithm for identifying low-count energy spectra in short-duration measurements based on heterogeneous sample transfer. The proposed method tackles challenges associated with negative transfer due to significant differences between source and target domain distributions. It utilized an alignment function to project both domains into a unified subspace. To effectively enhance the target domain samples, distance measurement methodologies and clustering techniques were employed to reassign class labels to the source domain samples. A decision tree model, known for its resilience to outliers, was employed, offering a hierarchical structure that clearly highlights the importance of spectral features and provides an intuitive representation of the classification process in spectral analysis.

The results of comparative experiments between the proposed method and several widely used approaches using only target domain data led to the following conclusions:

(1) The proposed method leveraged the inherent distribution properties of energy spectrum data and focuses on aggregating ratios from multiple regions of interest. This approach effectively mitigates transfer difficulties arising from different domain distributions in the transfer learning of short-duration energy spectrum samples.

(2) The proposed method utilized class mean measurements to calculate the distance between source domain samples and target domain category centroids. This alignment of source domain labels effectively reduces transfer difficulties caused by different label spatial distributions in low-count heterogeneous energy spectrum sample transfer learning, thereby increasing the total sample space in the target domain.

It should be noted that this study focused on static radiation sources. The effects of source motion or dissolution in water have not yet been considered. Although this method performs well in low-count energy spectrum classification, its reliance on prior knowledge remains a limitation. Future research should aim to enhance the applicability of this method to address more complex and diverse scenarios. Through an in-depth exploration of novel theoretical frameworks and technical approaches, we seek to overcome these limitations and make the method more versatile and adaptive.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Hao-Lin Liu and Cao-Lin Zhang. The first draft of the manuscript was written by Hao-Lin Liu, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data availability statement The data that support the findings of this study are openly available in Science Data Bank at https://cstr.cn/31253.11.sciencedb.j00186.00354 and https://doi.org/10.57760/sciencedb.j00186.00354.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. M.G. Paff, A. Di Fulvio, S.D. Clarke et al., Radionuclide identification algorithm for an organic scintillator-based radiation portal monitor. Nucl. Instrum. Meth. A. **849**, 41–48 (2017). https://doi.org/10.1016/j.nima.2017.01.009
2. Y. Altmann, A. Di Fulvio, M.G. Paff et al., Expectation propagation for weak radionuclide identification at radiation portal monitors. Sci. Rep.-Uk. **10**(1), 1–12 (2020). https://doi.org/10.1038/s41598-020-62947-3
3. X. Shi, T. Liu, P. Pei et al., Radionuclide-labeled antisilencing function 1a inhibitory peptides for tumor identification and individualized therapy. ACS Nano. **18**(12), 9114–9127 (2024). https://doi.org/10.1021/acsnano.4c00081
4. Y. Shi, Y. Zhang, X. Feng et al., A new method for the correction of spectrum drift caused by temperature changes when using a NaI (Tl) seawater radioactivity sensor. J. Mar. Sci. Eng. **12**(4), 546 (2024). https://doi.org/10.3390/jmse12040546
5. Y. Chen, L.P. Zhang, S. Xiao et al., Identification of unknown shielding parameters with gamma-ray spectra using a derivative-free inverse radiation transport model. Nucl. Sci. Tech. **29**(5), 70 (2018). https://doi.org/10.1007/s41365-018-0401-5
6. R. Shi, X.G. Tuo, H.L. Li et al., Unfolding analysis of the $LaBr_3$:Ce gamma spectrum with a detector response matrix constructing algorithm based on energy-resolution calibration. Nucl. Sci. Tech. **29**, 1 (2018). https://doi.org/10.1007/s41365-017-0340-6
7. X.Z. Li, Q.X. Zhang, H.Y. Tan et al., fast nuclide identification based on sequential Bayesian method. Nucl. Sci. Tech. **32**(12), 143 (2021). https://doi.org/10.1007/s41365-021-00982-z
8. D.U. Xiaochuang, T.U. Hongbing, L.I. Ke et al., Radionuclide identification method based on a gamma-spectra template library synthetic by radial basis function neural networks. J. Tsinghua Univ (Sci. and Tec.) **61**(11), 1308–1315 (2021). https://doi.org/10.16511/j.cnki.qhdxxb.2020.22.033
9. S. Qi, S. Wang, Y. Chen et al., Radionuclide identification method for NaI low-count gamma-ray spectra using an ANN. Nucl. Eng. Technol. **54**(1), 269–274 (2022). https://doi.org/10.1016/j.net.2021.07.025
10. Y. Liu, W. Wei, D. Niu et al., Nuclide identification and analysis using artificial neural networks. Ord. Indu. Auto. **34**(11), 86–91 (2015). https://doi.org/10.7690/bgzdh.2015.11.022
11. J. Wang, J. Jiang, Determination of net area for 92.6 keV peak of 238 U by the spectrum-stripping method. Nucl. Tech. **15**(4), 205–207 (1992). (**in Chinese**)
12. J.S. Ren, J.M. Zhang, K.P. Wang, Radioactive nuclide identification method based on SVD and SVM. Ord. Indu. Auto. **36**(5), 50–53 (2017). (**in Chinese**)
13. J.M. Zhang, H.B. Ji, X.H. Feng et al., nuclide spectrum feature extraction, and nuclide identification based on a sparse representation. High Power Laser Part. Beams. **30**(04), 046003 (2018). https://doi.org/10.11884/HPLPB201830.170435
14. S. Wen, B. Wang, C. Shen et al., Nuclide identification algorithms based on sequential Bayesian analysis. Nucl. Elec. Det. Tech. **36**(2), 179–183 (2016). https://doi.org/10.3969/j.issn.0258-0934.2016.02.015
15. Y. Wang, Z.M. Liu, Y.P. Wan et al., Energy-spectrum nuclide recognition methods based on long short-term memory neural networks. High Power Laser Part. Beams. **32**(10), 106001 (2020). https://doi.org/10.11884/HPLPB202032.200118 (in Chinese)
16. Z.F. Zhuang, P. Luo, Q. He et al., Survey on transfer learning research. J. Softw. **26**(1), 26–39 (2015). https://doi.org/10.13328/j.cnki.jos.004631
17. H.Q. Huang, X.F. Yang, W.C. Ding et al., Estimation method for parameters of overlapping nuclear pulse signals. Nucl. Sci. Tech. **28**(1), 12 (2017). https://doi.org/10.1007/s41365-016-0161-z
18. X. Li, Q. Zhang, H. Tan et al., Research of nuclide identification method based on background comparison method. Appl. Radiat. Isot. **192**, 110596 (2023). https://doi.org/10.1016/j.apradiso.2022.110596
19. L. Wang, Q. Li, J. Tang et al., Experimental studies on nuclide identification radiography with a CMOS camera at Back-n white neutron source. Nucl. Instrum. Methods. Phys. Res. A **1048**, 167892 (2023). https://doi.org/10.1016/j.nima.2022.167892
20. M. Alamaniotis, J. Mattingly et al., Kernel-based machine learning for background estimation of NaI low-count gamma-ray spectra. IEEE. T. Nucl. Sci. **60**(3), 2209–2221 (2013). https://doi.org/10.1109/TNS.2013.2260868
21. M. Alamaniotis, S. Lee et al., Intelligent analysis of low-count scintillation spectra using support vector regression and fuzzy logic. Nucl. Technol. **191**(1), 41–57 (2015). https://doi.org/10.13182/NT14-75
22. S.J. Pan, Q. Yang, A survey on transfer learning. IEEE T. Knowl. Data. En. **22**(10), 1345–1359 (2009). https://doi.org/10.1109/10.1109/TKDE.2009.191
23. E.G. Androulakaki, M. Kokkoris, C. Tsabaris et al., In situ spectrometry in a marine environment using full spectrum analysis for natural radionuclides. Appl. Radiat. Isot. **114**, 76–86 (2016). https://doi.org/10.1016/j.apradiso.2016.05.008
24. X. Zhang, Y. Yang, M. Shi et al., Novel energy identification method for shallow cracked rotor system. Mech. Syst. Signal. Pr. **186**, 109886 (2023). https://doi.org/10.1016/j.ymssp.2022.109886
25. Y.L. Song, F.Q. Zhou, Y. Li et al., Methods for obtaining the characteristic c-ray net peak count from the interlaced overlap peak in the HPGe c-ray spectrometer system. Nucl. Sci. Tech. **30**, 11 (2019). https://doi.org/10.1007/s41365-018-0525-7
26. H.D. Wang, J.B. Lu, R.P. Li et al., A phoswich design using real-time rise time discrimination for Compton suppression of $LaBr_3$:Ce detector. Nucl. Instrum. Method Phys. Res. A. **1048**, 16792 (2019). https://doi.org/10.1016/j.ymssp.2022.109886
27. T. Lim, S. Song, S. Kim et al., Monte Carlo simulations for gamma-ray spectroscopy using bismuth nanoparticle-containing plastic scintillators with spectral subtraction. Nucl. Eng. Technol. **55**(9), 3401–3408 (2023). https://doi.org/10.1016/j.net.2023.05.030
28. C.K. Qiao, J.W. Wei, L. Chen et al., An overview of the Compton scattering calculation. Crystals **11**(5), 525 (2021). https://doi.org/10.3390/cryst11050525
29. Y.Y. Song, L.U. Ying et al., Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry **27**(2), 130 (2015). https://doi.org/10.11919/j.issn.1002-0829.215044
30. B. Charbuty, A. Abdulazeez, L. Chen et al., Classification based on decision tree algorithm for machine learning. J. Appl. Sci. Tech. Tre. **2**(01), 20–28 (2021). https://doi.org/10.38094/jastt20165
31. T. Lan, H. Hu, C. Jiang et al., A comparative study of decision tree, random forest, and convolutional neural network for spread-F identification. Adv. Space. Res. **65**(8), 2052–2061 (2020). https://doi.org/10.1016/j.asr.2020.01.036
32. S. Qi, W. Zhao, Y. Chen et al., Comparison of machine learning approaches for radioisotope identification using NaI (TI) gamma-ray spectrum. Appl. Radiat. Isot. **186**, 110212 (2022). https://doi.org/10.1016/j.apradiso.2022.110212
33. L. Chen, Y.X. Wei, Nuclide identification algorithm based on K-L transform and neural networks. Nucl. Instrum. Method Phys. Res. A **598**(2), 450–453 (2009). https://doi.org/10.1016/j.nima.2008.09.035
34. S.M. Galib, P.K. Bhowmik, A.V. Avachat et al., Comparative study of machine learning methods for automated identification of radionuclides using NaI gamma-ray spectra. Nucl. Eng. Technol. **53**(12), 4072–4079 (2021). https://doi.org/10.1016/j.net.2021.06.020

35. J. He, X. Tang, P. Gong et al., A rapid radionuclide identification algorithm based on the discrete cosine transform and a BP neural network. Ann. Nucl. Energy. **112**, 1–8 (2018). https://doi.org/10.1016/j.anucene.2017.09.032
36. W. Zhao, R. Shi, X.G. Tuo et al., Novel radionuclides identification method based on Hilbert-Huang transform and convolutional neural network with gamma-ray pulse signal. Nucl. Instrum. Methods. Phys. Res. A **1051**, 168232 (2023). https://doi.org/10.1016/j.nima.2023.168232