



A novel approach for feature extraction from a gamma-ray energy spectrum based on image descriptor transferring for radionuclide identification

Hao-Lin Liu^{1,2} · Hai-Bo Ji¹ · Jiang-Mei Zhang³ · Cao-Lin Zhang² · Jing Lu² · Xing-Hua Feng²

Received: 18 July 2022 / Revised: 17 October 2022 / Accepted: 17 October 2022 / Published online: 8 December 2022

© The Author(s), under exclusive licence to China Science Publishing & Media Ltd. (Science Press), Shanghai Institute of Applied Physics, the Chinese Academy of Sciences, Chinese Nuclear Society 2022

Abstract

This study proposes a novel feature extraction approach for radionuclide identification to increase the precision of identification of the gamma-ray energy spectrum set. For easier utilization of the information contained in the spectra, the vectors of the gamma-ray energy spectra from Euclidean space, which are fingerprints of the different types of radionuclides, were mapped to matrices in the Banach space. Subsequently, to make the spectra in matrix form easier to apply to image-based deep learning frameworks, the matrices of the gamma-ray energy spectra were mapped to images in the RGB color space. A deep convolutional neural network (DCNN) model was constructed and trained on the ImageNet dataset. The mapped gamma-ray energy spectrum images were applied as inputs to the DCNN model, and the corresponding outputs of the convolution layers and fully connected layers were transferred as descriptors of the images to construct a new classification model for radionuclide identification. The transferred image descriptors consist of global and local features, where the activation vectors of fully connected layers are global features, and activations from convolution layers are local features. A series of comparative experiments between the transferred image descriptors, peak information, features extracted by the histogram of the oriented gradients (HOG), and scale-invariant feature transform (SIFT) using both synthetic and measured data were applied to 11 classical classifiers. The results demonstrate that although the gamma-ray energy spectrum images are completely unfamiliar to the DCNN model and have not been used in the pre-training process, the transferred image descriptors achieved good classification results. The global features have strong semantic information, which achieves an average accuracy of 92.76% and 94.86% on the synthetic dataset and measured dataset, respectively. The results of the statistical comparison of features demonstrate that the proposed approach outperforms the peak-searching-based method, HOG, and SIFT on the synthetic and measured datasets.

Keywords Radionuclide identification · Feature extraction · Transfer learning · Gamma energy spectrum analysis · Image descriptor

This work was supported by the National Defense Fundamental Research Project (No. JCKY2020404C004) and Sichuan Science and Technology Program (No. 22NSFSC0044).

✉ Hao-Lin Liu
lhl0502@mail.ustc.edu.cn

¹ Department of Automation, University of Science and Technology of China, Hefei 230026, China

² School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China

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

1 Introduction

Nuclear science and technology are rapidly developing and have been applied in various fields, having increasingly important roles in the sphere of scientific research and production [1–3]. Simultaneously, menacing nuclear weapons and radioactive contamination by nuclear industrial accidents present long-term and significant consequences for the environment, ecology, and biological health [4–6]. The detection and identification of radionuclides are crucial tasks under such circumstances [1, 7]. It is important to develop effective algorithms for the detection and identification of radionuclides with stronger discrimination and

high accuracy. One of the most critical steps in radionuclide identification is the feature extraction from gamma-ray energy spectra, which is a complicated task owing to the background conditions, energy resolution of the radiation detector, calibration shift, characteristic peak overlap, source strength, and shielding status [8, 9].

Feature extraction methods of traditional radionuclide identification algorithms can be summarized as searching for characteristic energy peaks from the gamma-ray energy spectra and matching them with peaks in the radionuclide library [2, 4, 8–11], which are usually based on physical rules and do not require a training process. Another classical approach is the template matching method, for which the main idea is establishing a template library of the gamma-ray energy spectra in advance and matching the entire spectrum or transformation of the spectrum with the template in the library [12, 13]. These methods are highly operator-dependent, and their limitations are magnified when characteristic peaks are overwhelmed by background noise or several interfering peaks are extracted.

With the development of artificial intelligence, radionuclide identification has gradually become a widely studied classification problem [14–16]. The main idea of classification methods applied in radionuclide identification is to extract features from gamma-ray energy spectra of known types, and train classification models using extracted features; then, the trained model is applied to estimate the probability of existing radionuclides of unknown types [17]. Numerous feature extraction algorithms have been used for radionuclide identification, such as the Karhunen-Loeve transform (K-L transform) [18], principal component analysis (PCA) [13, 19], singular value decomposition (SVD) [20], wavelet [21, 22], discrete cosine transform (DCT) [23], and sparse representation [24]. Subsequently, classification methods such as Bayesian [23, 25], extreme gradient boosting tree [26], back propagation neural network [17], artificial neural network (ANN) [27, 28], fuzzy logic [29], long short-term memory (LSTM) [30], convolutional neural network (CNN) [31], and deep convolutional neural network (DCNN) [32, 33] were applied for radionuclide identification. The key factor for the success of these methods is the extraction of strong discriminative features. The limitation of the aforementioned features is that only the relative relationship between the corresponding counts of the front and rear energy addresses is considered, and errors may increase owing to the non-smoothness of the low-count spectra [34–36].

Recent studies have shown that image descriptors transferred by CNNs and DCNNs provide a stable and reliable performance for image classification problems [37–44]. Hu et al. [37] reported that features transferred from CNNs were sufficiently generalized to high-resolution remote sensing image datasets and were more expressive

than low-level and mid-level features. Babenko et al. [38] experimentally determined that the activation of the top layers of CNNs is competitive despite being trained for unrelated classification tasks such as ImageNet. Moreover, Gong et al. [39] transferred the outputs of the last fully connected layer of a DCNN as an image descriptor, and Razavian et al. demonstrated that features transferred from convolution layers can provide useful global descriptors of specific image regions [40–44]. Liu et al. [44] demonstrated that convolution layers have excellent generalization and efficiency and that transferring convolution layer features can achieve an advanced performance.

This study constructed a novel feature extraction method from gamma-ray energy spectra for radionuclide identification. First, the gamma-ray energy spectra are transformed from a vector to matrix and then to image form. Feature transferring is then performed using a DCNN model. The transferred image descriptors consist of the activations from the convolution layers and activation vectors of the fully connected layers. To verify the effectiveness of the proposed method, 11 classical classification methods were employed to perform a statistical comparison, and the results demonstrate that the proposed method significantly outperforms the peak-searching-based method, histogram of oriented gradients (HOG), and scale-invariant feature transform (SIFT).

The following two main contributions are presented in this study:

- A novel pre-process method of the gamma-ray energy spectra is proposed. The vectors of the gamma-ray energy spectra are mapped to matrix form, and further mapped to image form. This form conversion can improve the utilization of spectral information and serve as the basis for extracting essential features and constructing a more discriminative classifier.
- Exploring and verifying the application of image descriptors transferring from a DCNN model in the field of radionuclide identification. Experimental results demonstrate that image descriptors can effectively extract the essential features of the gamma-ray energy spectrum images. Local image descriptors transferred from higher convolution layers can provide more discriminative descriptors, and global image descriptors transferred from the first fully connected layer has the strongest semantic information among the fully connected layer.

The remainder of this study is organized as follows. In Sect. 2, we introduce the proposed feature extraction approach for radionuclide identification. Section 3 presents a series of experiments using both synthetic datasets and measured datasets from real laboratory environments and offers a comparative analysis. Section 4 concludes the study.

2 Method

The proposed method consists of the following two major steps: (a) Mapping the vectors of the gamma-ray energy spectra from a Euclidean space to matrices in a Banach space, and then mapping the matrices to images in the RGB color space. (b) Constructing a DCNN model trained on ImageNet and transferring the corresponding activation vectors of fully connected layers and activations of convolution layers as global and local features of the gamma-ray energy spectrum images. Fig. 1 presents the procedure of the proposed method.

2.1 Data mapping

In this subsection, the essential features of the different radionuclides are extracted from a novel perspective.

Gamma rays are the products of the de-excitation process of atomic excitation and manifest as short-wavelength electromagnetic radiation. The essence of gamma rays is a stream of particles, which are gamma photons. Gamma photons are uncharged particles, and their interaction with matter is a random event. Collecting gamma photons using a general signal acquisition device is complex, whereas collecting electrical signals is relatively easy; the amplitude of the electrical signal is proportional to the photon energy value. Gamma photons can be converted into electrical signals for signal processing to be collected by a nuclear radiation detector. The principle of this process is that photons emitted by the radioactive source interact with the atoms of the medium in the detector to produce charged particles. The detector collects the particles, converts them into electrical signals, and detects nuclear signals by measuring the electrical signals. The counts distributed with the energy value of the particles can be obtained by scaling the pulse amplitude by the energy; the energy spectrum is the curve of

the distribution of the counts with the energy of the particles. As the fingerprints of radionuclides, the energy spectrum contains the distinguishable information of different radionuclides [2–4, 6].

For nuclear events, the count and time of events are random within a certain time interval. In radiation detection, the number of nuclear events measured over a certain period (e.g., detector counts) is also random. Because radioactive decay is a random process, each observation can be considered as a random experiment, and the count per unit time can be regarded as a random variable that obeys the Poisson or Gaussian distributions. For a spherical space with a single point source as the center of the sphere and a certain radius, the process of generating gamma photons by radioactive decay is random and continuous, and photons are uniformly emitted in all directions in space.

The gamma-ray energy spectrum s is a broad stationary random vector in Euclidean space, $s = \{s_k\} \in \mathbf{H}^l$. For the convenience of subsequent expressions, let $k = 0, 1, \dots, l-1$, where s_k is the $k+1$ -th count of photons distributed over the energy value, $s_k \in \mathbb{N}$. As previously indicated, the photons generated by radioactive decay are uniformly emitted in all directions in spherical space with a single point source as the center of the sphere and a certain radius. Therefore, s_k ideally positively correlates with the duration of the measurement, as formulated by Eq. (1).

$$s_k = \alpha_k \cdot t, \quad (1)$$

where t is the duration of the measurement and α_k is a parameter over the $k+1$ -th count value affected by the background noise of the environment, the measuring angle and distance, and the intensity of the radiation source.

Considering s as a Markov chain or Markov process, which is a stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event, s_k is only

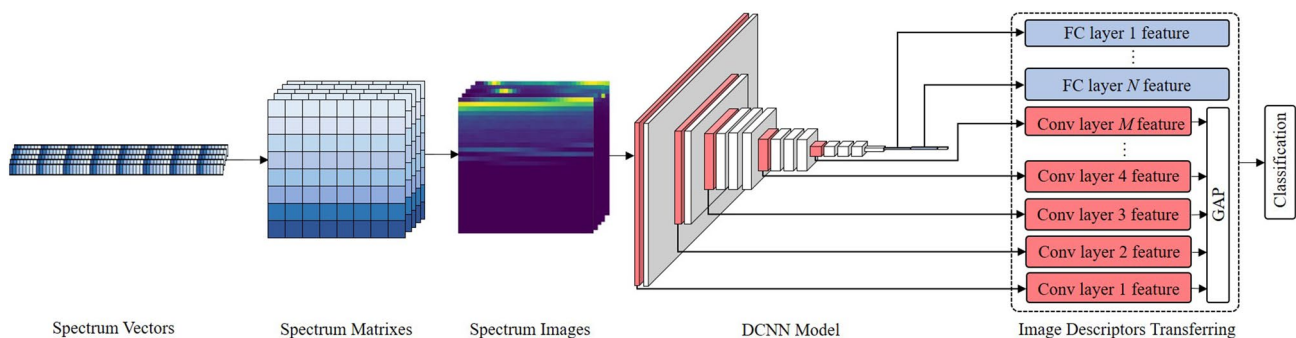


Fig. 1 (Color online) Block diagram of the proposed method. Mapping the vectors of the gamma-ray energy spectra from a Euclidean space to matrices in a Banach space, then to images in an RGB color space. Constructing a DCNN model trained on ImageNet and trans-

ferring the corresponding activation vectors of the fully connected layers and activations of the convolution layers as global and local features of the gamma-ray energy spectrum images, which are subsequently used for classification

related to s_{k-1} . In vector-based feature extraction methods, only part of the information of s is considered, that is, the relative relationship between the corresponding counts of the front and rear energy addresses. This is not conducive to mining the mutual relationship between non-adjacent s_k , which will cause difficulties in extracting discriminative and effective features.

Every radioactive decay produces photons of different energy values, and the photon counts of a certain energy obey the Poisson or Gaussian distributions. Thus, the total count of photons generated by radioactive decay per unit time also obeys the Poisson or Gaussian distributions. Let \mathbf{a} and \mathbf{b} be the gamma-ray energy spectra obtained by using different measurement durations in a measurement scenario with a fixed background noise, using a specified detector to measure the specified radioactive source at a fixed location and orientation. Therefore, as previously indicated, \mathbf{a} and \mathbf{b} are vectors from the same \mathbf{H} , $\mathbf{a} \in \mathbf{H}^l$ and $\mathbf{b} \in \mathbf{H}^l$. The corresponding durations of the measurement are t_a and t_b , respectively.

$$a_k = \alpha_k^{(a)} \cdot t_a, \quad (2)$$

$$b_k = \alpha_k^{(b)} \cdot t_b, \quad (3)$$

where the measuring conditions, that is, the background noise of the environment, measuring angle and distance, and intensity of the radiation source, are unchanged during the measurement, and we can assume that $\alpha_k^{(a)} = \alpha_k^{(b)}$. Assuming that the measurement time is sufficient, $\frac{a_k}{b_k}$ converges in probability to $\frac{t_a}{t_b}$; that is, for a_k and b_k , $\forall \varepsilon > 0$, $P\left(\left|\frac{a_k}{b_k} - \frac{t_a}{t_b}\right| \geq \varepsilon\right) = 0$ (when photons are uniformly emitted in all directions).

To efficiently transfer discriminative information for identification from the gamma-ray energy spectra, mapping the gamma-ray energy spectra from a vector form in \mathbf{H} to a matrix form in Banach space \mathbf{B} , the mapping f can be formulated by Eq. (4).

$$\mathbf{P} = f(\mathbf{s}) \quad (4)$$

where $\mathbf{P} = (p_{ij})_{m \times n} \in \mathbf{B}^{m \times n}$, $m \times n = l$. For the convenience of subsequent expressions, let $i = 0, 1, \dots, m-1$ and $j = 0, 1, \dots, n-1$, where p_{ij} represents the element in the $i+1$ -th row and $j+1$ -th column in the matrix. The relationship between s_k and p_{ij} is represented by Eq. (5).

$$p_{ij} = P_{[k/n], \text{mod}(k,n)} = s_k \quad (5)$$

where $k = 0, 1, \dots, l-1$.

From the matrix perspective, there are more elements adjacent to p_{ij} . When extracting features from gamma-ray energy spectra, not only is the relative relationship

between the corresponding counts of the front and rear energy addresses considered, but also the relative relationship between the upper and lower counts and the diagonal counts, which can be easily used and is more conducive to mining the internal and mutual relationships between elements in \mathbf{P} .

To apply \mathbf{P} as the input of a DCNN model and transfer image descriptors as features of the gamma-ray energy spectra, it is essential to map \mathbf{P} to the image form. Mapping g maps \mathbf{P} from the matrix form in \mathbf{B} to the image form in the RGB color space \mathbf{J} , can be formulated by Eq. (6).

$$\mathbf{Q} = g(\mathbf{P}) \quad (6)$$

where $\mathbf{Q} = (q_{ij})_{m \times n} \in \mathbf{J}^{m \times n}$, $m \times n = L$. For the convenience of subsequent expressions, let $i = 0, 1, \dots, m-1$, $j = 0, 1, \dots, n-1$, where q_{ij} represents the element in the $i+1$ -th row and $j+1$ -th column in the matrix. \mathbf{P} is normalized before the mapping process, which can be formulated using Eq. (7).

$$q_{ij} = \frac{p_{ij} - p_{\min}}{p_{\max} - p_{\min}}, \quad (7)$$

where $p_{\max} = \max_{i,j} \{p_{ij}\}$, $p_{\min} = \min_{i,j} \{p_{ij}\}$.

Equation (6) maps the element values of \mathbf{Q} onto the corresponding pixels of an image with specified colors. Each q_{ij} corresponds to a rectangular area in the image, and the values of q_{ij} are indices in the Parula color map [45] that determine the color of each patch. Equation (6) maps the smallest value in \mathbf{Q} to the first entry in the Parula color map and maps the largest value in \mathbf{Q} to the last entry in the Parula color map. All intermediate values of \mathbf{Q} are linearly scaled to the Parula color map in the ascending order. The relationship between the values of the elements in \mathbf{Q} and the colors of the corresponding pixels in the Parula color map is shown in Fig. 2.

\mathbf{a} and \mathbf{b} are two gamma-ray energy spectra obtained for different measurement durations under the same measurement conditions, and $\mathbf{Q}_a = g(f(\mathbf{a}))$ and $\mathbf{Q}_b = g(f(\mathbf{b}))$ are the images of \mathbf{a} and \mathbf{b} , respectively, under mapping $f * g$ which are defined as previously mentioned. For $\forall q_{ij}^{(a)} \in \mathbf{Q}_a$, $q_{ij}^{(a)}$ and t_a are related, as shown in Eq. (8).

$$q_{ij}^{(a)} = \frac{p_{ij}^{(a)} - p_{\min}^{(a)}}{p_{\max}^{(a)} - p_{\min}^{(a)}} = \frac{s_k^{(a)} - s_{\min}^{(a)}}{s_{\max}^{(a)} - s_{\min}^{(a)}}, \quad (8)$$

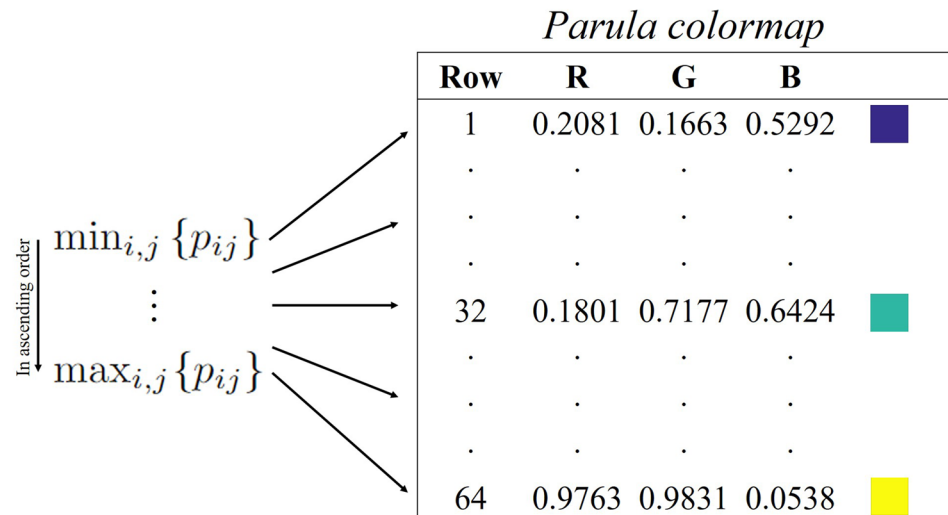
where $k = i \times m + j$ and

$$s_{\max}^{(a)} = \max_{k'=0,1,\dots,l-1} \{s_{k'}^{(a)}\} = \max_{k'=0,1,\dots,l-1} \{\alpha_{k'}^{(a)}\} \times t_a,$$

$$s_{\min}^{(a)} = \min_{k'=0,1,\dots,l-1} \{s_{k'}^{(a)}\} = \min_{k'=0,1,\dots,l-1} \{\alpha_{k'}^{(a)}\} \times t_a.$$

Obviously,

Fig. 2 (Color online) Relationship between the elements and the colormap. The Parula colormap is a three-column array with 64 rows, where each row in the array defines one color using an RGB triplet, i.e., contains the red, green, and blue intensities for a specific color. Each row in the matrix defines one color using an RGB triplet. The intensities are in the range [0,1], where a value of 0 indicates no color and a value of 1 indicates full intensity



$$q_{ij}^{(a)} = \frac{\alpha_{i \times m + j}^{(a)} - \min_{k'=0,1,\dots,l-1} \{\alpha_{k'}^{(a)}\}}{\max_{k'=0,1,\dots,l-1} \{\alpha_{k'}^{(a)}\} - \min_{k'=0,1,\dots,l-1} \{\alpha_{k'}^{(a)}\}}.$$

Similarly, for $\forall q_{ij}^{(b)} \in Q_b$, we have

$$q_{ij}^{(b)} = \frac{\alpha_{i \times m + j}^{(b)} - \min_{k'=0,1,\dots,l-1} \{\alpha_{k'}^{(b)}\}}{\max_{k'=0,1,\dots,l-1} \{\alpha_{k'}^{(b)}\} - \min_{k'=0,1,\dots,l-1} \{\alpha_{k'}^{(b)}\}}.$$

This is because the measuring conditions, that is, the background noise of the environment, the measuring angle, and distance are unchanged between Q_a and Q_b . If Q_a and Q_b were from the same radioactive source, then $\alpha_k^{(a)} = \alpha_k^{(b)}$.

Based on the aforementioned analysis, the effects of different measurement durations on Q_a and Q_b can be ignored under the same measurement conditions. More explicitly, the mappings of the gamma-ray energy spectra in the RGB color space from the same radionuclide would present nearly identical images in an ideal situation, which can reduce the intraclass differences caused by different measuring durations. Even in a real measuring situation, the information processing of mapping the gamma-ray energy spectra from a Euclidean space to a Banach space and then to an RGB color space remains to have a strong applicability and can reduce intraclass differences, thereby helping to extract the essential features from the gamma-ray energy spectra of different radionuclides.

2.2 Feature transferring

In this subsection, a DCNN model was constructed and trained on the ImageNet dataset. The mapped gamma-ray energy spectrum images were applied as inputs to the DCNN

model, and the corresponding activation vectors of fully connected layers and activations from convolution layers were transferred as descriptors of images to construct a new classification model for radionuclide identification.

VGG is a widely used convolutional neural network (CNN) model proposed by Karen Simonyan and Andrew Zisserman at the University of Oxford [46]. The VGG has various configurations; instance, VGG-11, VGG-16, VGG-19, etc. Of all the configurations, VGG-16 was identified as the best-performing model on the ImageNet dataset. The basic building block of VGG can be summarized as a stack of multiple (usually one, two, or three) convolution layers with a filter size of 3×3 , one stride, and one padding, followed by a max-pooling layer of size 2×2 . Different configurations of this stack were repeated in the network to achieve various depths. The number associated with each configuration is the number of layers with the weight parameters. The convolution stacks are followed by three fully connected layers, two with a size of 4096, and the last one with a size of 1000. The last layer is the output layer with a Softmax activation. A size of 1000 refers to the total number of possible classes in ImageNet. In the method proposed in this study, the structure of the VGG consists of U convolution layers, V max-pooling layers, and N fully connected layers. The convolution layers use convolution kernels to convolve the input image, and the results of the convolution constitute the feature maps of the input image; thus, the local features of the image are extracted after the convolution layers. The max-pooling layers are arranged after the convolution layers, and the maximum values of the corresponding positions in the feature maps are calculated. This can reduce the dimension of the extracted feature information to make the feature maps smaller, simplify the computing complexity of the network, and avoid overfitting. The U convolution layers can be divided into M groups using the max-pooling

layers as separations. The fully connected layers map the distributed feature representation learned from the convolution layers and max-pooling layers to the sample marker space, the essence of which is to perform a weighted sum of the features to integrate the local features and output them as a value to reduce the influence of the local feature position on the classification.

The convolution layers and max-pooling layers have parameters such as the height of the convolution kernel, width of the convolution kernel, number of input channels, number of output channels (number of convolution kernels), padding, and stride. The padding parameter refers to the boundary padding of the original matrix with '0' before the convolution; thus, the convolution kernel can extend to the pseudo-pixels beyond the edge when scanning the input image, thereby avoiding the loss of edge information. The stride parameter refers to the length of each movement of the convolution kernel, and the stride size affects the efficiency of the model. Each convolution kernel strictly has a bias parameter. To simplify the calculation, the bias is omitted in the following calculation.

For the convolution layers, the height and width of the convolution kernel are represented as h_k^{conv} and w_k^{conv} , respectively; the number of input and output channels are represented as c_i^{conv} and c_o^{conv} , respectively; and the parameters of the padding and stride are represented as pad^{conv} and str^{conv} . The number of neurons in the fully connected layer is represented by c_o^{fc} .

The size of the input image is represented as $c_i^{\text{im}} \times h_i \times w_i$, and the size of the output image is represented as $c_o^{\text{im}} \times h_o \times w_o$, where c_i^{im} , h_i , w_i , and c_o^{im} , h_o , w_o represent the channel, height, and width of the input and output images, respectively. The height and width of the output image of the convolution and max-pooling layers can be formulated as Eqs. (9) and (10).

$$h_o = \lfloor h_i - h_k + 2 \times \text{pad} \rfloor / \text{str} + 1, \quad (9)$$

$$w_o = \lfloor w_i - w_k + 2 \times \text{pad} \rfloor / \text{str} + 1, \quad (10)$$

where Eqs. (9) and (10) can be generalized to the max-pooling layers to calculate the height and width of the convolution kernels. c_i^{conv} of a convolution layer depends on the channel of the input image, and c_o^{conv} of a convolution layer depends on the number of convolution kernels.

VGG is training on the ImageNet dataset to obtain the classification model PM. ImageNet is an image database organized according to the WordNet hierarchy, in which hundreds or thousands of images depict each node of the hierarchy. The dataset has been instrumental in advancing computer vision and deep learning research [47]. Let S be a set of the gamma-ray energy spectra in Euclidean space, $S \subseteq \mathbf{H}^l$, s be the gamma-ray energy spectrum, $s \in S$. s is converted into Q

using Eqs. (4) and (6), and Q is applied as an input to the PM to transfer features. Generally, PM will only have one final output, and the corresponding activation vectors of the fully connected layers and activations are transferred from the convolution layers in the process as the image descriptors of Q .

The convolution kernel of PM corresponds to a receptive field, and a small part of the image (local receptive area) is used as the input of the lowest convolution layer, which makes each neuron output by the convolution layer only experience the local image area and does not need to experience the global image. This operation is equivalent to passing through a digital filter to obtain the most salient features of the observed data. In the fully connected layers, different local features from the convolution layers are synthesized through the weight matrix to form a representation of the global information. Therefore, the transferred activation vectors of the fully connected layers are regarded as global feature representations for the gamma-ray energy spectrum images, and the transferred activation maps from the convolution layers are regarded as local features describing particular gamma-ray energy spectrum image regions.

Specifically, the activated maps transferred from the convolution layers, that is, the set of feature maps transferred from the i -th group of convolution layers of PM, is denoted as c_i , $i \in [1, 2, \dots, M]$, which is essentially a set of multiple square matrices. c_i contains $d^{(i)}$ feature maps, where $d^{(i)}$ depends on the number of convolution kernels in one of the convolution layers in the i -th group of convolution layers. The size of the feature map is $h_o \times w_o$, where h_o and w_o can be calculated using Eqs. (9) and (10). c_i can be written as:

$$c_i = \{c_1^{(i)}, c_2^{(i)}, \dots, c_{d^{(i)}}^{(i)}\}.$$

To convert the set of feature maps into a set of feature vectors, which can be further used for classifier training, the conversion operation can be formulated as follows:

$$\text{GAP}(c_i) = \left(\text{ave}(c_1^{(i)}), \text{ave}(c_2^{(i)}), \dots, \text{ave}(c_{d^{(i)}}^{(i)}) \right),$$

where $\text{GAP}(c_i)$ is the global average pooling on the set of feature maps, and $\text{ave}(c^{(i)})$ is the average operation on all elements of the feature maps, where

$$\text{ave}(c^{(i)}) = \text{ave} \left(\begin{bmatrix} c_{1,1}^{(i)} & \dots & c_{1,w_o}^{(i)} \\ \vdots & \ddots & \vdots \\ c_{h_o,1}^{(i)} & \dots & c_{h_o,w_o}^{(i)} \end{bmatrix} \right).$$

The transferred local features of Q are expressed by Eq. (11).

$$F_L^{(Q)} = \{\text{GAP}(c_i)\}, i \in [1, 2, \dots, M] \quad (11)$$

where $GAP(c_i)$ is a vector of length $d^{(i)}$ and $F_L^{(Q)}$ is a set of multiple $GAP(c_i)$, $i \in [1, 2, \dots, M]$, which are essentially the local image descriptors of Q transferred from PM.

The transferred activation vectors of the fully connected layers, that is, the feature vectors transferred from the j -th fully connected layer of PM, are denoted as f_j , $j \in [1, 2, \dots, N]$, whose length is the same as the number of neurons in the j -th fully connected layer c_o^{fc} . The transferred global features of Q are expressed by Eq. (12).

$$F_G^{(Q)} = \{f_j\}, j \in [1, 2, \dots, N]. \quad (12)$$

where $F_G^{(Q)}$ is a set of multiple f_j , which are the global image descriptors of Q transferred from PM. Table 1 summarizes the overall flow of the algorithm.

2.3 Illustration

In this subsection, we present a simple example of the proposed method. For illustration, we selected two synthetic gamma-ray energy spectra s_1 and s_2 of ^{60}Co . The essence of s_1 and s_2 are two vectors in Euclidean space H^l , where $l = 4096$. The simulation settings of s_1 and s_2 are essentially the same, and the only difference is the total number of simulated particles, which is equivalent to the difference in the measurement duration. The number of simulated particles correlates with the duration of the measurement because the process of generating gamma photons by radioactive decay is random and continuous, and photons are uniformly emitted in all directions in space. The simulated particle numbers corresponding to s_1 and s_2 are represented as t_1 and t_2 , where t_1 is five hundred thousand and t_2 is two hundred and fifty thousand. To illustrate the difference between the two spectra more vividly, we plot s_1 and s_2 in Fig. 3, where the blue spectrum is the description of s_1 , and the red spectrum is the description of s_2 . Figure 3 clearly shows differences between s_1 and s_2 due to the duration of measurement, where the counts distributed over the energy addresses of the entire spectrum is significantly more prominent in s_1 than s_2 .

To extract more discriminative image descriptors from the gamma-ray energy spectra for identification, s_1 and s_2 are mapped to the matrix form using Eq. (4). The mapped matrices are represented as P_1 and P_2 , which are in the Banach space $B^{m \times n}$, where $m = n = 64$. To reduce the discrepancy between P_1 and P_2 caused by the difference in the measurement duration, and to facilitate the matrices as the inputs of the DCNN for the transfer of image descriptors, P_1 and P_2 are mapped to the image form using Eq. (6), and the results are represented by Q_1 and Q_2 . The gamma-ray energy spectra of ^{137}Cs and ^{152}Eu were randomly selected for comparison, and their corresponding Q_{Cs} and Q_{Eu} values were obtained using Eq. (4). The intraclass similarity and interclass distinction of the different gamma-ray energy

Table 1 The overall flow of the proposed method

Algorithm	Feature extraction approach based on image Descriptor transferring
<i>Input</i>	<p>s: Gamma-ray energy spectrum; n: Number of columns of P and Q; cm: Parula colormap, a 64×3 matrix; VGG: A DCNN framework with M groups of convolution Layers and N fully connected layers; ImageNet: Public image data set;</p>
<i>Begin</i>	<p>1. s is mapped to matrix form by:</p> $P_{ij} = P_{\lfloor k/n \rfloor, \text{mod}(k,n)} = s_k$ <p>where $i = 0, 1, \dots, m-1$, $j = 0, 1, \dots, n-1$; 2. p_{ij} is normalized by:</p> $q_{ij} = \frac{p_{ij} - p_{\min}}{p_{\max} - p_{\min}}$ <p>where $i = 0, 1, \dots, m-1$, $j = 0, 1, \dots, n-1$; 3. q_{ij} is mapped to image form by:</p> $q_{ij} = \begin{cases} cm(1), & q_{ij} == q_{\min} \\ cm(64), & q_{ij} == q_{\max} \\ \xrightarrow{ls} cm(2 : 63), & q_{\min} < q_{ij} < q_{\max} \end{cases}$ <p>where $q_{\max} = \max_{i,j} \{q_{ij}\}$, $q_{\min} = \min_{i,j} \{q_{ij}\}$, ls means linearly scaled; 4. $Q = (q_{ij})_{m \times n}$ is the image form of s after mapping $f * g$; 5. Pre-trained model is obtained based on Imagenet: $PM = VGG(\text{ImageNet})$ 6. c_i is output of the i-th group of convolution layers of PM, where $i \in [1, 2, \dots, M]$; 7. Converting the set of feature maps into set of feature vectors by $GAP(c_i)$; 8. f_j is output of the i-th fully connected layer of PM, where $j \in [1, 2, \dots, N]$;</p>
<i>End</i>	
<i>Output</i>	<p>Transferred local features of s:</p> $F_L^{(Q)} = \{GAP(c_i)\}, i \in [1, 2, \dots, M]$ <p>Transferred global features of s:</p> $F_G^{(Q)} = \{f_j\}, j \in [1, 2, \dots, N]$

spectra are shown in Fig. 4 vividly. Specifically, Q_1 and Q_2 exhibited nearly identical images in Fig. 4a and b, which indicates that the difference between s_1 and s_2 in Fig. 3 is significantly reduced when s_1 and s_2 are mapped to RGB color space. By comparing Fig. 4c, d with Fig. 4a–d clearly have completely different characteristics than Fig. 4a, b. Namely, the mapping of the gamma-ray energy spectra of the randomly selected ^{137}Cs and ^{152}Eu in the RGB color space exhibits completely different characteristics from those of ^{60}Co . Through the aforementioned comparison and analysis

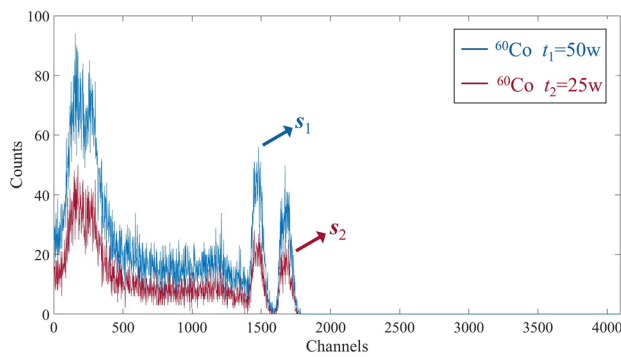


Fig. 3 (Color online) Two original spectra of ^{60}Co . The simulation settings of s_1 and s_2 are essentially the same, and the only difference is the total number of simulated particles t_1 and t_2 , which is equivalent to the difference in the measurement duration

in Sect. 2, the effects of different measuring durations on Q_1 and Q_2 can apparently be ignored under the same measuring conditions. Therefore, the information processing of mapping the gamma-ray energy spectra from Euclidean space to

Banach space and then to the RGB color space remains to have a strong applicability and can reduce intraclass differences, thereby helping to extract the essential features from the gamma-ray energy spectra of different radionuclides. A further quantitative analysis is presented in Sect. 3.

In the feature transfer phase, the structure of VGG and the corresponding feature sizes are listed in Table 2, which consists of five groups of convolution layers and two fully connected layers. The layer number only considers the number of convolution layers, and the last layer of each group is the max-pooling layer. The number of channels in the feature size is not changed in the convolution layers, but changes in the max-pooling layer owing to the size of the stride parameter [the number of output channels in different groups can be calculated by Eqs. (9) and (10)]. For fully connected layers, the number of output channels is equal to the number of neurons in the fully connected layers. The partial parameters of the convolution, max-pooling, and fully connected layers are shown in Table 3.

The VGG framework was trained on the ImageNet dataset to obtain the classification model PM. Q_1 and Q_2 are

Fig. 4 (Color online) Gamma-ray energy spectrum images. **a** and **b** are the images of Q_1 and Q_2 , **c** and **d** are the images of Q_{Cs} and Q_{Eu}

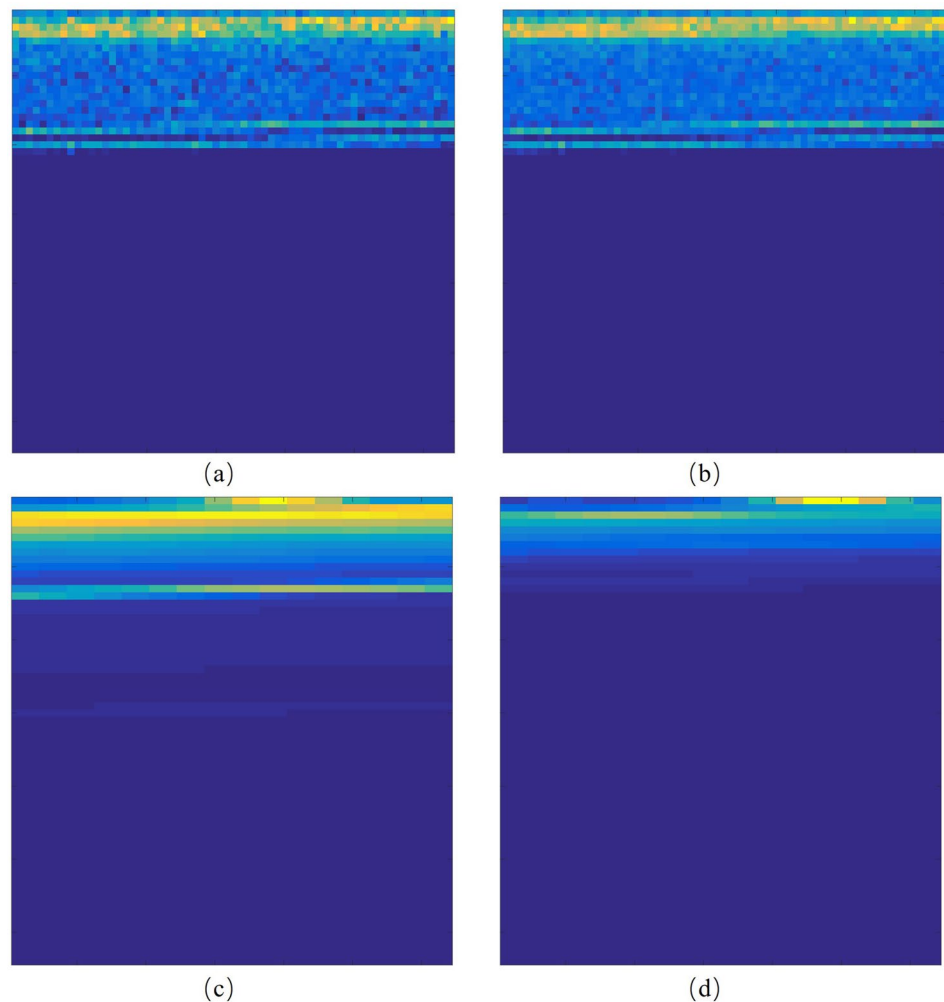


Table 2 The structure of VGG

Layer group	Layer number	Feature size
Conv1	2	$64 \times 64 \times 64$
Conv2	2	$128 \times 32 \times 32$
Conv3	3	$256 \times 16 \times 16$
Conv4	3	$512 \times 8 \times 8$
Conv5	3	$512 \times 4 \times 4$
FC	2	4096
Output	1	1000

Table 3 Partial parameters of VGG

Layer name	Kernel size/neurons number	Stride	Padding
Convolution	3×3	1	1
Max pooling	2×2	2	0
Fully connected	4096	–	–

applied as input images for the PM to transfer features, and the activation vectors of the fully connected layers and activations from the convolution layers in the process are transferred as the image descriptors of Q_a and Q_b . Image descriptors transferred from the fifth group of convolution layers c_5 and the first fully connected layer f_1 were selected as illustrations. c_5 and f_1 from the synthetic and measured datasets are shown in a low-dimensional space through the t-SNE [48] in Fig. 7. Each color in the figure represents one type of radionuclide, which intuitively reflects the difference in the expressive and discriminative abilities between the different features. The transferred image descriptors have a strong intraclass similarity and interclass differentiation, and a further analysis of the feature performance is presented in the next section.

3 Experiments and analysis

This section introduces the acquisition and preprocessing of the synthetic data and the measured data and establishes a series of experiments using 28 classification methods based on the Weka machine learning toolkit to verify the feasibility of image descriptors transferred from the gamma-ray energy spectrum for radionuclide identification. Based on the previous experiments, statistical comparisons of features through nonparametric and Friedman tests were conducted to verify whether image descriptors transferred from DCNNs can be used as an essential feature representation for gamma-ray energy spectrum images.

3.1 Data preparation

The production of the synthetic dataset consisted of the following three steps: (1) Background data acquisition. A self-made 3-inch NaI detector was used to measure the ambient background. The measurement was conducted in a laboratory environment without the presence of a separate radioactive source. The detector was then placed at a fixed position for 12 h. Two measurements were performed, one with lead bricks placed around the detector and the other without; two sets of background data were obtained. (2) Single-nuclide energy spectrum acquisition. Based on the Geant4 platform, the transport process of the radioactive gamma-ray particles of 26 single nuclides was simulated using the Monte Carlo method. We constructed simulated scenarios in Geant4 containing only different types of single radioactive point sources and a 3-inch NaI detector. The positions and relative distances of the radioactive point sources and detector were fixed. A total of one million photons were simulated, and their trajectories were recorded by the detector and converted into the gamma-ray energy spectrum. (3) Data synthesis. To simulate real measurements as closely as possible, two sets of background data and 26 sets of synthetic spectra were linearly superimposed with a random signal-to-noise ratio (SNR). $SNR = N_{nc}/N_{bg}$, where N_{nc} is the sum of the photon counts emitted by the radioactive point source and N_{bg} is the sum of the photon counts of the background. In the linear superposition process, the SNR value is a random number ranging between 0.3 and 1. A total of 2080 synthetic gamma-ray energy spectra of 26 radionuclides were obtained and named as dataset 1. Table 4 demonstrates a list of common radionuclides, which includes a total of 26 radionuclides in the following four categories: SNM, industrial, medical, and NORM.

The measured gamma-ray energy spectra were obtained from radioactive sources in a laboratory environment with lead brick shielding. The measurements took advantage of a cadmium zinc telluride (CZT) gamma-ray spectrometer from Kromek. The spectrometer has 4096 measuring channels; the measurable energy ranges between 25 keV and 3.0 MeV, and the electronic noise is lower than 10 keV. Two types of γ radiation sources, ^{137}Cs and ^{60}Co ,

Table 4 Radionuclide library of synthetic sample set

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 , ^{99m}Tc , ^{201}Tl , ^{133}Xe
NORM	^{40}K , ^{226}Ra , ^{232}Th

Table 5 Grouping of gamma-ray energy spectrum samples

Set name	Nuclide type	Capacity
Data set 1	Radionuclides listed in Table 4	2080
Data set 2	^{60}Co	50
	^{137}Cs	50
	^{152}Eu	50
	^{137}Cs , ^{60}Co	50
	^{137}Cs , ^{152}Eu	50
	^{60}Co , ^{152}Eu	50
	^{137}Cs , ^{60}Co , ^{152}Eu	50

and one type of IV radiation source, ^{152}Eu were used in the measurement. The spectrometer was carried on the Turtlebot robot to quantitatively control the measuring distance and reduce radiation damage to the experimental

operators. As shown in Table 5, seven groups of samples were established; the measuring object, measuring distance, and measuring duration were varied in the process. A total of 150 measured gamma-ray energy spectra of three single radiation sources and a total of 200 measured gamma-ray energy spectra of four mixed radiation sources were obtained and named dataset 2. Figure 5 presents the spectrum of ^{60}Co in different sample sets. Figure 6 presents the gamma-ray energy spectrum images in different sample sets.

3.2 Feature performance comparison and analysis

Research have shown that image descriptors transferred from DCNNs can provide a reliable performance for image classification problems. However, choosing different features for a specific classification domain remains worth discussing. For the domain of the radioactive gamma-ray

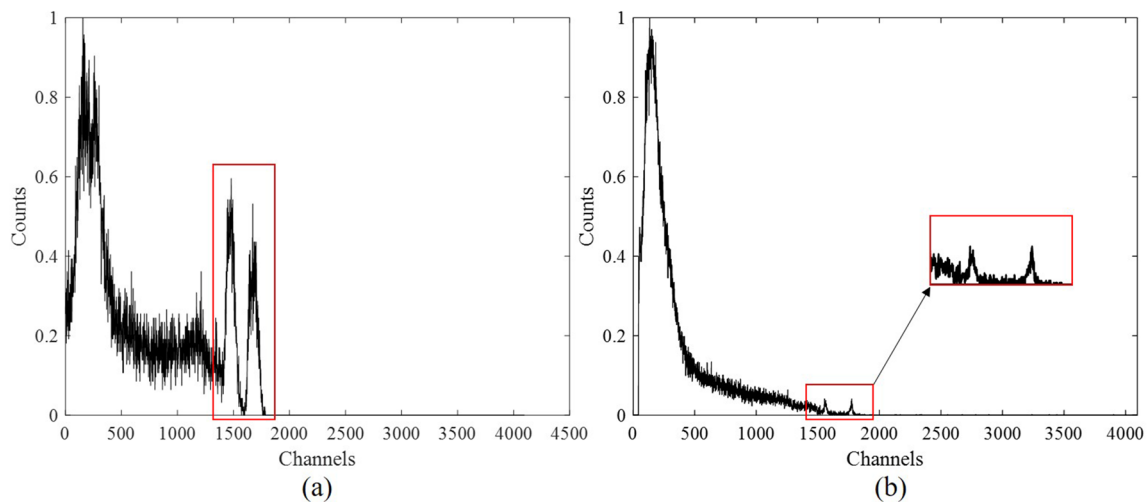


Fig. 5 Examples of synthetic and measured spectra. **a** displays a ^{60}Co synthetic spectrum, and **b** displays a real measured ^{60}Co spectrum

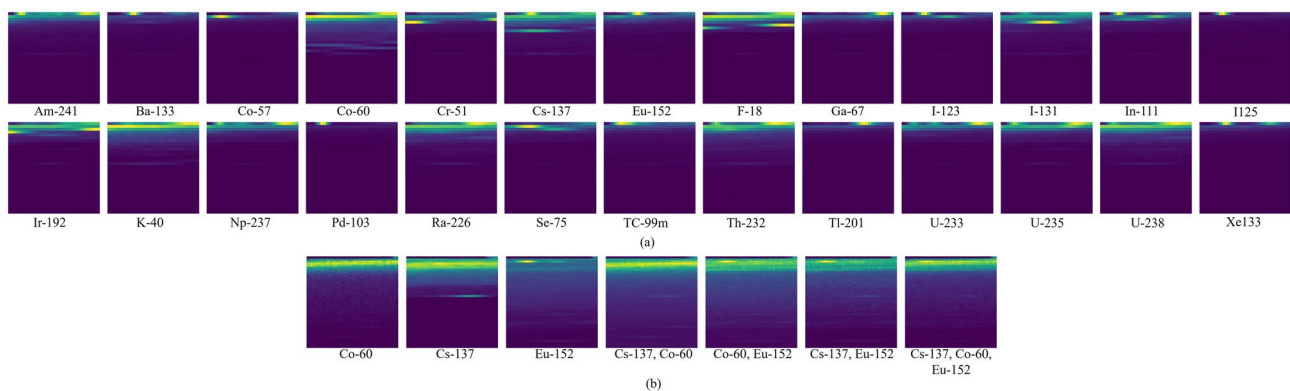


Fig. 6 (Color online) Examples of synthetic and measured gamma-ray energy spectrum images. Group **a** are synthetic gamma-ray energy spectrum images of 26 radionuclides, group **b** are measured gamma-ray energy spectrum images of 7 single and mixed radionuclides

energy spectral classification presented in this study, and considering the limitation of the computing resources and the scale of parameters, VGG-16 was chosen as the DCNN framework owing to the characteristic that VGG-16 applies a significantly small 3×3 receptive field (filters) throughout the entire network with the stride of 1 pixel. A combination of multiple 3×3 filters and nonlinear activation layers can replace a receptive area of a larger size, which makes the decision functions more discriminative to the characteristics of the spectrum. This imparts the ability of the network to converge faster [46]. In addition, the consistent use of 3×3 convolutions across the network makes the network significantly simple, elegant, and conveniently transfers image descriptors.

In this subsection, 28 classification methods based on the Weka machine learning toolkit [49] were applied in a series of experiments to verify whether image descriptors transferred from a DCNN model can be used to construct a classification model with a strong discrimination and advanced accuracy. Owing to the large number of transferred features (five sets of local features and two sets of global features for a single energy spectrum), in an actual training process, using only a certain set of features is sufficient to build a satisfactory classifier. Therefore, in the experiments described in this subsection, single sets of local and global features and combinations of high-level features were used to conduct the experiments. Multiple groups of transferred image descriptors from the synthetic and measured datasets were applied in the training and testing processes of the 28 classification models. All classification models were trained with the default parameters and settings specified in the toolkit, and the experiments applied the tenfold cross-validation method to avoid the imbalance caused by random data segmentation. The percentage of misclassified cases for transferred features on the synthetic dataset and measured dataset are listed in Tables 6 and 7. The significance of bold in the tables is the percentage of misclassified cases that performed well in each experiment.

The results demonstrate that although the gamma-ray energy spectrum images are completely unfamiliar to the DCNN model and have not been used in the pre-training process, the transferred image descriptors achieved good classification results.

For local image descriptors, a higher group of convolution layers can provide more discriminative descriptors than a lower group of convolution layers. Part of the multi-group union image descriptors can provide better resolution, and different arrangements of image descriptors from the same groups also have an impact on the results. The best average classification effect of the local image descriptors on the synthetic dataset was the multi-group union image descriptors composed of c_3 , c_4 , and c_5 , which was 93.08%. Local image descriptors achieved the best average classification

effect on the measured dataset c_5 , which was 93.51%. The best classification effects of the local image descriptor on the synthetic and measured datasets were 100.00% and 99.71%, respectively.

For the global image descriptors, f_1 , which is transferred from the first fully connected layer, has strong semantic information, achieving an average accuracy of 92.76% and 94.86% on the synthetic dataset and measured dataset, respectively. The best classification effects of the global image descriptors on the synthetic and measured datasets were both 100.00%. In contrast, the classification model trained on the global image descriptors transferred from the second fully connected layer achieved a poor classification performance. The aforementioned experiments have preliminarily proved that image descriptors transferred from a DCNN model can be applied as the image descriptors of radionuclide gamma-ray energy spectrum images for classification; however, further comparisons of the features of gamma-ray energy spectra and different image descriptors of the gamma-ray energy spectrum images are needed.

3.3 Statistical comparison

Characteristic peaks are the most important features of traditional radionuclide identification methods. Four groups of characteristic peaks were extracted from the gamma-ray energy spectra by changing the peak properties, that is, distance, prominence, width, and threshold. Scale-invariant feature transform (SIFT) [50] and a histogram of oriented gradients (HOG) [51] are two classical feature extraction algorithms used in computer vision. Two groups of features extracted from the gamma-ray energy spectrum images were obtained using the HOG and SIFT algorithms. Fig. 7 demonstrates various features of the synthetic and measured data sets in a low-dimensional space through the t-SNE [48], which including characteristic peaks, features extracted by HOG and SIFT, present local image descriptors c_5 and global image descriptors f_1 . As shown in Fig. 7, the distribution of the characteristic peaks is chaotic, and no apparent clustering center is observed for the different radionuclides. While HOG and SIFT features can form apparent clustering centers in certain radionuclides, the features of these radionuclides exhibit a significant crossover. Simultaneously, the transferred local and global image descriptors were significantly enhanced in feature discrimination, and the aggregation between similar feature points was also stronger.

The aforementioned features from one- and two-dimensional spectral images were applied in 21 classification methods using the Weka machine learning toolkit [49]. All experiments applied the tenfold cross-validation method to avoid the imbalance caused by random data segmentation. The proportions of misclassified samples in the experiments are listed in Tables 8 and 9. These results provide a basis

Table 6 Percentage of misclassified cases for transferred image descriptors on synthetic data set

Classification method	c_1	c_2	c_3	c_4	c_5	$c_{4,5}$	$c_{5,4}$	$c_{3,4,5}$	$c_{5,4,3}$	f_1	f_2	$f_{1,2}$
Bayes Net	16.68	5.62	5.62	1.20	45.14	0.43	0.43	0.87	0.82	0.43	68.80	0.48
Lib SVM	56.87	39.04	39.04	32.36	27.40	29.47	29.57	30.43	30.48	15.38	97.12	34.62
RBF Classifier	70.38	69.37	69.37	75.34	71.39	72.07	72.12	72.98	72.60	69.23	80.12	–
RBF Network	11.01	3.03	3.03	0.38	0.00	0.00	0.00	0.00	0.00	0.00	85.34	32.45
Simple Logistic	2.07	0.05	0.05	0.10	0.00	0.24	0.24	0.00	0.24	0.00	69.95	0.00
SMO	15.53	0.05	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	57.31	0.24
IBk	1.73	0.05	0.05	0.00	0.10	0.00	0.00	0.00	0.00	0.00	81.35	1.44
LWL	38.37	24.71	24.71	15.72	14.76	33.41	13.65	14.47	14.47	8.37	85.58	34.86
Attribute Selected Classifier	9.62	1.54	1.54	1.25	1.15	0.91	0.96	0.96	0.72	1.43	85.10	1.43
Bagging	4.37	1.06	1.06	0.29	3.70	0.38	0.19	0.38	0.14	0.24	82.69	0.48
Classification Via Regression	3.27	1.59	1.59	0.53	1.15	0.38	0.82	4.00	0.82	0.96	82.69	2.88
Filtered Classifier	8.89	6.59	6.59	0.53	6.54	6.39	7.26	6.01	5.72	6.63	86.92	7.74
Iterative Classifier Optimizer	2.98	0.72	0.38	0.24	0.00	0.00	34.86	0.00	0.00	0.72	78.85	0.72
Logit Boost	2.98	0.38	0.38	0.24	0.19	0.14	0.00	0.00	0.00	0.48	80.71	0.00
MultiClass Classifier	2.31	0.05	0.05	0.10	0.24	0.00	0.00	0.00	0.00	0.24	86.54	0.00
MultiClass Classifier Updateable	0.00	0.34	0.34	0.00	0.00	0.00	0.00	0.00	0.00	0.24	82.21	3.12
Random Committee	2.50	0.24	0.24	0.05	0.19	0.00	0.00	0.00	0.00	0.00	87.02	0.00
Randomizable Filtered Classifier	4.86	0.34	0.34	0.24	0.10	0.10	0.19	0.14	0.19	0.24	93.51	90.19
Random Sub Space	3.51	0.77	0.77	0.05	0.19	0.10	0.05	0.10	0.19	0.24	83.89	0.19
Decision Table	20.38	17.93	17.93	12.93	0.34	12.31	12.36	15.91	18.27	21.87	88.70	21.87
JRip	8.46	5.34	5.34	3.70	3.75	12.31	4.57	0.00	6.01	0.19	88.46	0.00
OneR	41.44	41.39	41.39	45.14	48.12	43.17	43.17	41.54	42.12	64.66	94.23	64.71
PART	4.95	1.39	1.39	1.15	1.59	1.30	1.15	1.25	1.06	2.36	88.46	1.59
J48	4.86	1.30	1.30	0.62	1.25	0.62	0.62	0.87	0.53	1.20	91.11	1.44
LMT	0.00	0.00	0.05	0.00	0.48	0.48	0.48	0.24	0.24	0.00	68.92	0.00
Random Forest	2.07	0.24	0.24	1.73	0.24	0.00	0.00	0.00	0.00	0.00	79.52	0.05
Random Tree	6.73	2.93	2.93	1.73	3.12	2.84	2.26	2.21	2.45	4.66	92.55	5.19
REP Tree	6.39	2.79	2.79	2.16	2.16	1.87	1.35	1.49	1.44	3.08	86.30	3.08
Mean	12.61	8.17	8.16	7.06	8.33	7.82	8.08	6.92	7.09	7.24	83.36	11.44

for evaluating the performance of the transferred features. However, from a statistical point of view, these results do not provide strong support for identifying a group of features. For a statistical comparison of features, the nonparametric test method recommended by Demšar [52] was considered for comparing multiple features using several classification methods. First, we tested whether there were significant differences among the seven groups of features. The Friedman test was then applied to compare several features of the multiple classification methods.

Assuming that Num types of classification methods and k groups of features are involved in this experiment, the implementation of the Friedman test assigns a value r_i^j to each group of features, which indicates the rank of the feature j on the i -th classification method. The average rank R of feature j was computed using Eq. (13).

$$R_j = \frac{1}{\text{Num}} \sum_i r_i^j. \quad (13)$$

The null hypothesis was established, which states that all features have the same statistical performance. The Friedman statistic is asymptotically χ^2 distributed with $k-1$ degrees of freedom and is computed using Eq. (14).

$$\chi_F^2 = \frac{12\text{Num}}{k(k+1)} \left(\sum_j R_j^2 - \frac{k(k+1)^2}{4} \right) \quad (14)$$

The null hypothesis is rejected if the size of χ_F^2 exceeds the critical value, which indicates that there is a statistically significant difference between the classifiers. Conversely, the null hypothesis is accepted if the size of χ_F^2 does not exceed the critical value. A post hoc test was applied to determine the nature of the differences and compare the relative performances of different features when the null hypothesis was rejected.

Table 7 Percentage of misclassified cases for transferred image descriptors on measured data set

Classification method	c_1	c_2	c_3	c_4	c_5	$c_{4,5}$	$c_{5,4}$	$c_{3,4,5}$	$c_{5,4,3}$	f_1	f_2	$f_{1,2}$
Bayes net	6.57	5.71	6.00	8.29	8.29	8.57	8.57	8.00	8.00	4.86	32.86	32.86
Lib SVM	54.29	54.29	52.57	45.14	40.00	41.43	41.43	41.14	41.14	30.00	36.00	30.57
RBF classifier	38.29	45.71	33.71	30.29	26.00	36.29	27.14	33.43	34.29	22.86	61.43	—
RBF network	18.86	22.57	9.71	6.29	2.86	3.43	3.43	3.71	3.71	1.43	65.71	33.14
Simple logistic	9.71	4.86	2.57	2.29	3.14	1.14	3.43	1.14	1.14	0.57	34.29	0.57
SMO	24.00	14.57	1.43	0.86	0.57	0.57	0.57	0.86	0.86	0.57	22.57	22.57
IBk	2.57	2.00	1.14	0.86	0.86	0.57	0.57	0.86	0.57	0.00	42.86	42.86
LWL	39.43	30.00	35.43	18.86	16.00	18.29	18.29	18.29	18.29	11.43	62.86	63.71
Attribute selected classifier	5.14	7.43	5.71	6.57	4.29	6.86	6.57	5.43	6.00	6.00	65.71	1.43
Bagging	3.71	2.86	4.86	3.14	2.86	3.43	2.00	4.57	2.29	2.00	47.14	34.29
Classification via regression	4.00	4.00	3.71	4.00	4.86	4.57	4.29	4.00	4.57	2.00	52.74	2.00
Filtered classifier	5.14	5.71	5.43	6.29	9.43	5.71	6.00	6.00	6.00	62.57	62.57	62.86
Iterative classifier optimizer	3.43	2.86	2.86	2.86	2.57	2.00	4.00	2.57	2.57	1.71	51.14	0.00
Logit boost	3.43	2.57	3.43	2.86	2.29	2.00	2.00	2.57	2.57	0.57	42.86	0.29
MultiClass classifier	9.71	5.43	3.14	2.29	1.43	1.43	1.14	1.71	2.00	2.00	54.28	8.57
MultiClass classifier updateable	38.29	14.29	4.00	4.00	1.43	1.71	1.71	2.29	1.71	4.00	35.71	4.29
Random committee	2.00	2.57	1.71	2.29	2.00	2.00	1.14	1.14	1.14	1.14	55.71	0.57
Randomizable filtered classifier	4.57	5.14	3.71	3.71	2.57	2.57	3.14	3.14	3.14	2.00	79.14	61.14
Random sub Space	3.43	2.00	1.71	2.00	2.00	2.00	4.00	3.14	3.14	1.71	48.86	79.71
Decision table	5.43	8.29	8.86	9.43	6.57	9.71	8.00	10.29	9.14	8.00	66.86	8.00
JRip	6.86	6.57	5.14	5.71	9.43	4.29	6.29	6.86	6.57	6.00	63.43	4.29
OneR	8.86	11.72	9.14	8.57	12.57	9.71	9.71	9.71	9.71	14.29	79.71	14.29
PART	3.71	5.72	5.43	5.14	4.57	4.00	4.29	6.00	5.14	2.00	63.71	2.00
J48	3.14	4.86	4.86	4.00	4.29	3.43	3.43	4.86	4.57	2.57	72.00	2.57
LMT	2.57	4.57	2.57	2.29	0.57	1.14	0.86	1.14	1.14	0.57	32.86	0.57
Random forest	2.29	2.29	1.14	0.86	1.14	0.29	0.57	0.57	0.57	0.57	40.57	0.57
Random tree	4.29	4.57	4.57	5.71	4.00	4.57	5.14	4.57	3.71	2.86	74.57	4.00
REP tree	5.14	6.86	7.71	5.43	5.14	4.86	5.14	5.71	5.14	5.71	63.14	5.71
Mean	11.39	10.36	8.29	7.14	6.49	6.66	6.53	6.92	6.74	5.14	53.97	19.39

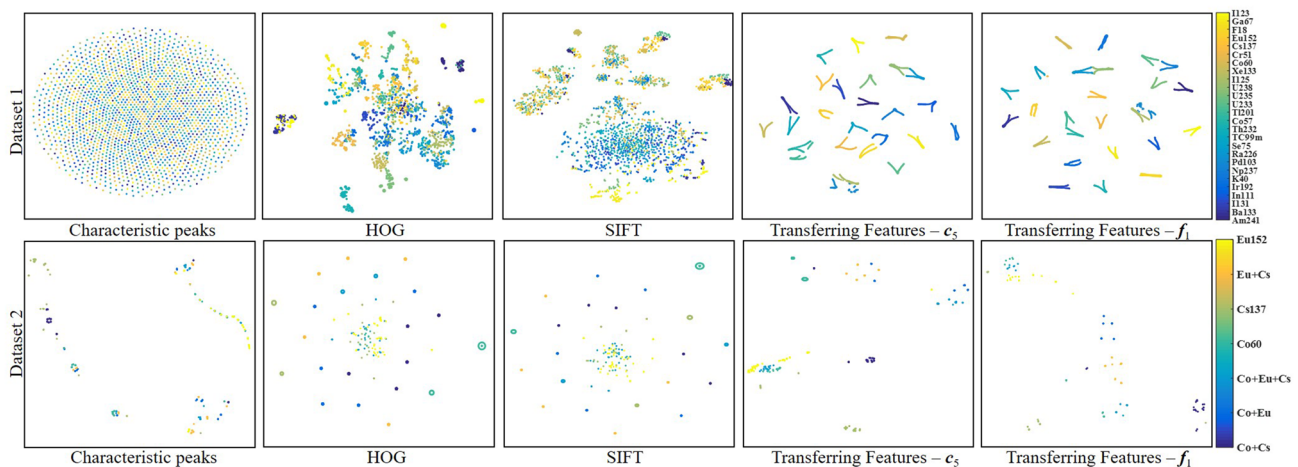
**Fig. 7** (Color online) Visualization of various features by t-SNE. Each color in the figure represents one type of radionuclide, which intuitively reflects the difference in the expressive ability and discrimination ability between different features

Table 8 Percentage of misclassified cases for various features on synthetic data set

Classification method	peaks-1	peaks-2	peaks-3	peaks-4	f_1	HOG	SIFT
Bayes net	0.0111	0.0236	0.0125	0.0332	0.0043	0.1091	0.4663
RBF network	0.0625	0.0409	0.0721	0.0841	0.0000	0.2404	0.4976
Simple logistic	0.0163	0.0899	0.1212	0.1986	0.0000	0.1562	0.4952
SMO	0.1370	0.3389	0.3144	0.3649	0.0000	0.1562	0.4663
IBk	0.0163	0.3389	0.0264	0.0538	0.0000	0.1774	0.4832
Bagging	0.0139	0.0139	0.0298	0.0269	0.0024	0.1457	0.3918
Multi class classifier	0.0308	0.2043	0.1625	0.2904	0.0024	0.4452	0.7572
JRip	0.0163	0.0130	0.0274	0.0341	0.0019	0.2144	0.5673
J48	0.0087	0.0091	0.0216	0.0192	0.0120	0.1712	0.4784
LMT	0.0101	0.0423	0.0221	0.0596	0.0000	0.1683	0.4952
Random forest	0.0034	0.0043	0.0091	0.0087	0.0000	0.1187	0.3822
Mean	0.0308	0.0726	0.0745	0.1067	0.0021	0.1912	0.4982

Table 9 Percentage of misclassified cases for various features on measured data set

Classification method	peaks-1	peaks-2	peaks-3	peaks-4	f_1	HOG	SIFT
Bayes net	0.0943	0.1343	0.1114	0.1457	0.0486	0.1257	0.1257
RBF network	0.2143	0.2857	0.1771	0.2600	0.0143	0.1314	0.1429
Simple logistic	0.1057	0.1057	0.1343	0.1343	0.0057	0.1286	0.1629
SMO	0.1971	0.1114	0.1457	0.1114	0.0057	0.1286	0.1286
IBk	0.0914	0.1143	0.1371	0.0971	0.0000	0.1571	0.1343
Bagging	0.1286	0.1143	0.1343	0.1343	0.0200	0.1457	0.1429
Multi class Classifier	0.0943	0.1514	0.1571	0.1343	0.0200	0.1629	0.1600
JRip	0.1514	0.1200	0.1229	0.1286	0.0600	0.1571	0.1400
J48	0.1314	0.0743	0.1171	0.1143	0.0257	0.1857	0.1714
LMT	0.0857	0.1057	0.1371	0.1257	0.0057	0.1286	0.1714
Random forest	0.0829	0.0800	0.0914	0.1057	0.0057	0.1171	0.1171
Mean	0.0943	0.1557	0.1354	0.1356	0.0192	0.1426	0.1452

In our experiments, the significance level α value was set to 0.05, the corresponding critical value of dataset 1 was calculated as 2.2541 while $\chi_F^2 = 44.8442$, and the corresponding critical value of dataset 2 was calculated as 2.2541 while $\chi_F^2 = 2.6218$. The result of the Friedman test indicated that the null hypothesis was rejected; therefore, further Holm tests [52] were required to compare the performance of the different features.

The test statistics of the Holm test for comparing features in a pair-wise manner are formulated in Eq. (15).

$$z = (R_i - R_j) / \sqrt{\frac{k(k+1)}{6\text{Num}}} \quad (15)$$

Using the z value obtained from Eq. (15), the corresponding probability p can be determined from the normal distribution table and is compared with the appropriate α to determine whether the hypothesis is rejected, that is, whether the proposed model is better than a certain classifier. Let p_1, p_2, \dots , denote the ordered corresponding probability p

values; thus, p from all the features have the following relationship: $p_1 \leq p_2 \leq \dots \leq p_{k-1}$. The step-down procedure of the Holm test begins with the most significant p value and compares each p_i with the adjusted α , which is calculated using $\alpha/(k-i)$.

If the adjusted α corresponding to the feature, $\alpha/(k-1)$, is higher than p_1 , the corresponding null hypothesis is rejected and indicates that both features have the same performance, and subsequently proceeds to compare p_2 with the corresponding adjusted α , $\alpha/(k-2)$. If the second null hypothesis remains rejected, the experiment continues to test the third null hypothesis, and so on. Once a certain null hypothesis is accepted, all remaining null hypotheses are preserved.

In our case, f_1 was chosen as the representative of the transferred image descriptors and compared with other groups of features. For data set 1, $R_{f_1} = 1.1818$, $R_{\text{peaks-1}} = 2.0909$, $R_{\text{peaks-2}} = 3.3636$, $R_{\text{peaks-3}} = 3.9091$, $R_{\text{peaks-4}} = 4.7727$, $R_{\text{HOG}} = 5.5455$, $R_{\text{SIFT}} = 7.0000$. For data set 2, R_{f_1}

$= 1.0000$, $R_{\text{peaks-1}} = 3.5455$, $R_{\text{peaks-2}} = 3.2273$, $R_{\text{peaks-3}} = 4.5000$, $R_{\text{peaks-4}} = 4.1818$, $R_{\text{HOG}} = 5.3636$, $R_{\text{SIFT}} = 5.2273$. $\alpha = 0.05$, $\text{Num} = 11$, and $k = 7$, and the standard error was $SE = \sqrt{\frac{7 \times (7+1)}{6 \times 11}} = 0.9211$. The results of the Holm test are listed in Table 10.

For Dataset 1, the Holm procedure rejects the first, second, third, fourth, and fifth hypotheses because the corresponding p values are smaller than the adjusted α . Thus, the final hypothesis cannot be rejected. This indicates that the transferred image descriptors perform significantly better than the characteristic peaks searched by prominence, width, and threshold, and features extracted by HOG and SIFT at the significance level of $\alpha=0.05$. The transferred image descriptors were not significantly better than the characteristic peaks searched for by distance. For Dataset 2, the Holm procedure rejects all the hypotheses, which indicates that the transferred image descriptors perform significantly better than the characteristic peaks and features extracted by HOG and SIFT at the significance level $\alpha = 0.05$. The filtering basis of searching the characteristic peaks by using the distance removes peaks with smaller distances by defining the minimum horizontal distance between adjacent peaks until all remaining peaks satisfy the distance condition. This peak-finding method is effective for the synthesized gamma-ray energy spectra using the NaI detector and has a larger classification error in the measured gamma-ray energy spectra using the CZT detector. Because the resolution of the NaI detector is not excellent, and the synthetic energy spectra comes from an ideal environmental simulation, the simulated energy spectrum has fewer interference peaks, while the measured gamma-ray energy spectra fluctuate significantly.

The results of the aforementioned comparative experiments prove that image descriptors transferred from DCNNs are better than the characteristic peaks extracted from the gamma-ray energy spectra and the HOG and SIFT shallow

features extracted from the gamma-ray energy spectrum images. These transferred image descriptors are essential, discriminative, and can provide a reliable performance for gamma-ray energy spectrum image classification problems.

4 Conclusion

This study proposes a novel feature extraction approach for radionuclide identification to facilitate the extraction of structural and essential features and increase the precision of identification on the gamma-ray energy spectrum set.

The results of a series of comparative experiments between the proposed method, peak-searching-based method, HOG, and SIFT using both synthetic and measured data demonstrate the following conclusions. (1) The information preprocessing of the proposed method, that is, mapping the gamma-ray energy spectra from a Euclidean space to a Banach space and then to an RGB color space, is significant for extracting the essential features, which can reduce the intraclass differences caused by different measuring durations. (2) The feature transfer process of the proposed method, that is, transferring the corresponding activation vectors of fully connected layers and activations from convolution layers in the process from DCNNs as image descriptors, can effectively extract the essential features of gamma-ray energy spectrum images. (3) Local image descriptors transferred from higher convolution layers provide more discriminative descriptors. (4) The global image descriptors transferred from the first fully connected layer had the strongest semantic information among the fully connected layers. (5) The proposed method outperforms the peak-searching-based method, HOG, and SIFT on synthetic and measured datasets.

Future studies will focus on exploring the available value of other DCNNs in the field of radionuclide identification, exploring more feature fusion methods and aggregation

Table 10 Results of statistical comparison

Dataset name	i	Feature name	$\mathbf{z} = (\mathbf{R}_i - \mathbf{R}_{\text{TF}}) / SE$	p	$\alpha / (k - i)$
Dataset1	1	SIFT	$(7.0000 - 1.1818) / 0.9211 = 6.3163$	0.0000	0.0083
	2	HOG	$(5.5455 - 1.1818) / 0.9211 = 4.7373$	0.0000	0.0100
	3	peaks-4	$(4.7727 - 1.1818) / 0.9211 = 3.8984$	0.0000	0.0125
	4	peaks-3	$(3.9091 - 1.1818) / 0.9211 = 2.9608$	0.0031	0.0167
	5	peaks-2	$(3.3636 - 1.1818) / 0.9211 = 2.3686$	0.0179	0.0250
	6	peaks-1	$(2.0909 - 1.1818) / 0.9211 = 0.9869$	0.3237	0.0500
Dataset2	1	HOG	$(5.3636 - 1.0000) / 0.9211 = 4.7373$	0.0000	0.0083
	2	SIFT	$(5.2273 - 1.0000) / 0.9211 = 4.5892$	0.0000	0.0100
	3	peaks-3	$(4.5000 - 1.0000) / 0.9211 = 3.7997$	0.0001	0.0125
	4	peaks-4	$(4.1818 - 1.0000) / 0.9211 = 3.4542$	0.0006	0.0167
	5	peaks-1	$(3.5455 - 1.0000) / 0.9211 = 2.7634$	0.0057	0.0250
	6	peaks-2	$(3.2273 - 1.0000) / 0.9211 = 2.4180$	0.0156	0.0500

approaches to develop more powerful descriptors, and establishing more universal datasets to further advance the research process of radionuclide identification.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Hao-Lin Liu, Cao-Lin Zhang and Xing-Hua Feng. 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.

References

1. J.M. Ghawaly, A.D. Nicholson, D.E. Peplow et al., Data for training and testing radiation detection algorithms in an urban environment. *Sci. Data*. **7**(1), 1–6 (2020). <https://doi.org/10.6084/m9.figshare.12654065>
2. M.A. Mariscotti, A method for automatic identification of peaks in the presence of background and its application to spectrum analysis. *Nucl. Instrum. Methods*. **50**, 309–320 (1967). [https://doi.org/10.1016/0029-554X\(67\)90058-4](https://doi.org/10.1016/0029-554X(67)90058-4)
3. 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>
4. R. Shi, X.G. Tuo, H.L. Li et al., Unfolding analysis of the LaBr₃: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>
5. D.K. Fagan, S.M. Robinson, R.C. Runkle, Statistical methods applied to gamma-ray spectroscopy algorithms in nuclear security missions. *Appl. Radiat. Isot.* **70**, 2428–2439 (2012). <https://doi.org/10.1016/j.apradiso.2012.06.016>
6. 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>
7. 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>
8. E. Yoshida, K. Shizuma, S. Endo et al., The application of neural networks for the analysis of gamma-ray spectra was measured with a Ge spectrometer. *Nucl. Instrum. Meth. A*. **484**(1–3), 557–563 (2002). [https://doi.org/10.1016/S0168-9002\(01\)01962-3](https://doi.org/10.1016/S0168-9002(01)01962-3)
9. 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>
10. J. Wang, J. Jiang, Determination of net area for 92.6 keV peak of ²³⁸U by the spectrum-stripping method. *Nucl. Tech.* **15**(4), 205–207 (1992). (in Chinese)
11. J.G. Lu, F.X. Zhao, D.M. Lu, Study on the Analysis Method of γ -Spectral Peaks of Environmental Samples. *J. Ningde Nor. Univ. (Nat. Sci.)* **13**(2), 15–17 (1995). (in Chinese)
12. 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>
13. K.T. Nelson, J.R. Romo, M. Monterial et al., Feature extraction and design from gamma-ray spectra for radionuclide identification, in *Paper Presented at the INMM and ESDARSA Joint Virtual Annual Meeting* (Livermore, 2021) <https://www.osti.gov/servlets/purl/1818419>
14. C. Tsabaris, E.G. Androulakaki, S. Alexakis et al., In-situ gamma-ray spectrometer for the deep ocean. *Appl. Radiat. Isot.* **142**, 120–127 (2018). <https://doi.org/10.1016/j.apradiso.2018.08.024>
15. C. Bobin, O. Bichler, V. Lourenço et al., Real-time radionuclide identification in c-emitter mixtures based on spiking neural network. *Appl. Radiat. Isot.* **109**, 405–409 (2016). <https://doi.org/10.1016/j.apradiso.2015.12.029>
16. 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>
17. 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>
18. L. Chen, Y.X. Wei, Nuclide identification algorithm based on K-L transform and neural networks. *Nucl. Instrum. Meth. A*. **598**(2), 450–453 (2009). <https://doi.org/10.1016/j.nima.2008.09.035>
19. Y. Liu, W. Wei, D. Niu, 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>
20. 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)
21. C.J. Sullivan, S.E. Garner, K.B. Butterfield, Wavelet analysis of gamma-ray spectral, in *Paper Presented at the 13th IEEE Symposium Conference Record on Nuclear Science* (Rome, Italy, 2004) <https://doi.org/10.1109/NSSMIC.2004.1462198>
22. C.J. Sullivan, M.E. Martinez, S.E. Garner, Wavelet analysis of sodium iodide spectral. *IEEE T. Nucl. Sci.* **53**(5), 2916–2922 (2006). <https://doi.org/10.1109/TNS.2006.881909>
23. 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>
24. 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>
25. C.J. Sullivan, J. Stinnett, Validation of a Bayesian-based isotope identification algorithm. *Nucl. Instrum. Meth. Phys. Res. Sect. A*. **784**, 298–305 (2015). <https://doi.org/10.1016/j.nima.2014.11.113>
26. 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>
27. 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.qhdx.2020.22.033>
28. M. Kamuda, J. Stinnett, C.J. Sullivan, Automated isotope identification algorithm using artificial neural networks. *IEEE T. Nucl. Sci.* **64**, 1858–1864 (2017). <https://doi.org/10.1109/TNS.2017.2693152>
29. C.J. Wang, D.M. Bao, C. Song et al., Investigation of fuzzy recognition mechanism for c-ray fingerprints of nuclear materials. *Acta Phys. Sinica*. **57**(9), 5361–5365 (2008). <https://doi.org/10.7498/aps.57.5361>
30. Y. Wang, Z.M. Liu, Y.P. Wan et al., Energy spectrum nuclide recognition method based on long short-term memory neural networks. *High Power Laser Part Beams* **32**(10), 106001 (2020). (in Chinese)
31. B.T. Koo, H.C. Lee, K. Bae et al., Development of a radionuclide identification algorithm based on a convolutional neural network

- for a radiation portal monitoring system. *Radiat. Phys. Chem.* **180**, 109300 (2021). <https://doi.org/10.1016/j.radphyschem.2020.109300>
32. 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>
33. A.G.L. Otero, J.T. Marumo, A.J. Potiens Junior, Applying deep-learning in gamma-spectroscopy for radionuclide identification, in *Paper Presented at the 3rd International Conference on Dosimetry and its Applications*. (Lisbon, Portugal, 2019) <http://repositorio.ipen.br/handle/123456789/31012>
34. M. Alamaniotis, J. Mattingly, L.H. Tsoukalas, 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>
35. M. Alamaniotis, S. Lee, T. Jevremovic, 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>
36. S.Y. Wen, B.R. Wang, G. Xiao et al., The study of nuclide identification algorithm 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>
37. F. Hu, G.S. Xia, J. Hu et al., Transferring DCNNs for scene classification of high-resolution remote sensing imagery. *Remote Sens.-Basel.* **7**(11), 14680–14707 (2015). <https://doi.org/10.3390/rs71114680>
38. A. Babenko, A. Slesarev, A. Chigorin et al., Neural codes for image retrieval, in *Paper Presented at the 13th European Conference on Computer Vision*. (Zurich, Switzerland, 2014) https://doi.org/10.1007/978-3-319-10590-1_38
39. Y. Gong, L. Wang, R. Guo et al., Multiscale orderless pooling of deep convolution activation features, in *Paper Presented at the 13th European Conference on Computer Vision*. (Zurich, Switzerland, 2014) https://doi.org/10.1007/978-3-319-10584-0_26
40. A.S. Razavian, J. Sullivan, S. Carlsson et al., visual instance retrieval using deep convolutional networks. *ITE Trans. Media Technol. Appl.* **4**(3), 251–258 (2016). <https://doi.org/10.3169/mta.4.251>
41. M. Cimpoi, S. Maji, A. Vedaldi, Deep filter banks for texture recognition and segmentation, in *Paper Presented at the 28th IEEE Conference on Computer Vision and Pattern Recognition*. (Boston, USA, 2015) <https://doi.org/10.1109/CVPR.2015.7299007>
42. A. Babenko, V. Lempitsky, Aggregating local deep features for image retrieval, in *Paper Presented at the 15th IEEE International Conference on Computer Vision*. (Santiago, Chile, 2015) <https://doi.org/10.1109/ICCV.2015.150>
43. A. Babenko, V. Lempitsky, Aggregating deep convolution features for image retrieval (2015). [arXiv:1510.07493](https://arxiv.org/abs/1510.07493)
44. L. Liu, C. Shen, A. Van den Hengel, The treasure beneath convolution layers: Cross-convolution-layer pooling for image classification, in *Paper Presented at the 28th IEEE Conference on Computer Vision and Pattern Recognition*. (Boston, USA, 2015) <https://doi.org/10.1109/CVPR.2015.7299107>
45. S. Eddins, A New Colormap for MATLAB – Part 1 – Introduction, <https://blogs.mathworks.com/steve/2014/10/13/a-new-color-map-for-matlab-part-1-introduction/>; 2014 [accessed October 13 2014]
46. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in *Paper Presented at the 3rd International Conference on Learning Representations* (San Diego, USA, 2015) <https://doi.org/10.48550/arXiv.1409.1556>
47. Stanford Vision Lab, The homepage of ImageNet, <https://ImageNet.org/>; 2021 [accessed Mar 11 2021]
48. L. Van der Maaten, G. Hinton, Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**(11), 2579–2605 (2008)
49. I.H. Witten, E. Frank, M.A. Hall et al., *Data mining: Practical machine learning tools and techniques*, 3rd ed. (New Zealand, 2014), pp. 403–406
50. D.G. Lowe, Object recognition from local scale-invariant features, in *Paper Presented at the 7th IEEE International Conference on Computer Vision*. (Kerkyra, Greece, 1999). <https://doi.org/10.1109/ICCV.1999.790410>
51. N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in *Paper Presented at the 18th IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. (San Diego, USA, 2005) <https://doi.org/10.1109/CVPR.2005.177>
52. J. Demšar, Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.* **7**, 1–30 (2006)

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.