RPE | Radiation Physics & Engineering
A Peer-Reviewed Journal
Published by K.N. Toosi University of Technology
http://rpe.kntu.ac.ir
Research Article

Radiation Physics and Engineering 2026; 7(3):63–73

Comparison of nonlinear autoencoder and linear PCA dimensionality reduction in gamma-ray spectroscopy based radioisotope identification

Kazem Abdolpour, S. Farhad Masoudi*, Atefeh Fathi

Department of Physics, K.N. Toosi University of Technology, P.O. Box 15875-4416, Tehran, Iran

HIGHLIGHTS

- Compared PCA and autoencoder for reducing gamma-ray spectral dimensionality.
- Trained both methods on simulated NaI(Tl) spectra and tested on real measurements.
- Both methods achieved near-perfect isotope ID under ideal conditions.
- Autoencoder showed higher robustness to ±20% gain drift than PCA.
- Autoencoder features transferred better to real spectra than PCA features.

ABSTRACT

Dimensionality reduction can play an important role in radioisotope identification from gamma-ray spectra by compressing spectral information, reducing model complexity, and improving learning efficiency. The importance of this approach becomes more evident under real-world conditions, where spectra are typically characterized by high dimensionality, low counts, peak overlap, and calibration instabilities, all of which make direct analysis more difficult. On this basis, in the present study, two dimensionality-reduction approaches-principal component analysis as a linear method and an autoencoder as a nonlinear method- were compared to evaluate their effectiveness for radioisotope identification. Using a simulated dataset of 1024-channel NaI(Tl) spectra of six common radioisotopes (Co-60, Cs-137, I-131, Ba-133, Am-241, and Tc-99m) for training, we evaluated both approaches with a common multilayer perceptron (MLP) classifier on unseen data, including laboratory-measured spectra and scenarios with gain drift. Under ideal conditions, both approaches achieved nearly perfect identification performance ($F1 \approx 0.98 - 0.99$). In more challenging regimes, including experimental spectra and severe gain drift, the autoencoder's latent features consistently outperform PCA. The autoencoder+MLP model generalized better to real spectra (e.g., achieving $F1 \approx 0.99$ versus 0.91 for PCA) and maintained higher accuracy under a ±20% gain drift ($F1 \approx 0.81$ versus 0.71). These results suggest that the nonlinear latent representation learned by the autoencoder is less sensitive to gain drift and experimental variability than variance-based linear projections, resulting in improved robustness for multi-label radioisotope identification. This insight can support the design of more reliable field-deployable gamma spectroscopy systems, where maintaining high identification performance amid noise and calibration variability is essential.

KEYWORDS

Gamma-ray spectroscopy
Radioisotope identification
Dimensionality reduction
Autoencoder neural networks
Principal component analysis

HISTORY

Received: 11 May 2026
Revised: 25 June 2026
Accepted: 28 June 2026
Published: Summer 2026

1 Introduction

Accurate and rapid identification of radioisotopes from gamma-ray spectra is a critical requirement in nuclear security, environmental monitoring, and emergency response applications. In practical scenarios such as field inspection of unknown radioactive sources, monitoring of contaminated environments, or characterization of radioactive waste measurements are frequently performed under short acquisition times, low counting statistics, and non-ideal measurement geometries. These conditions become particularly challenging when multiple radionuclides are present in a sample, as their spectral signatures may overlap through Compton continua and blended photopeaks. As a result, reliable isotope identification in composite spectra remains a challenging task, especially when us-

*Corresponding author: masoudi@kntu.ac.ir

https://doi.org/10.22034/rpe.2026.580714.1354

ing lowresolution scintillation detectors, which are widely deployed due to their relatively low cost, high efficiency, and portability (Zhang et al., 2022; Galib et al., 2021; Khatiwada et al., 2023).

Conventional gammaspectrum analysis methods typically rely on photopeak detection and spectral library matching. While these techniques can perform well under controlled laboratory conditions, their reliability often degrades in realistic environments. Measured spectra may be influenced by shielding effects, scattering processes, and background radiation, all of which can distort spectral features. In addition, detector instabilities may lead to gain drift and resolution degradation, causing shifts or broadening of spectral peaks and distortions of the continuum. These effects often necessitate frequent recalibration and expert intervention, thereby limiting the robustness of traditional analysis methods in operational settings (Burr and Hamada, 2009).

To address these challenges, datadriven approaches have increasingly been investigated for radioisotope identification. By learning directly from the global spectral shape rather than relying solely on explicit peak detection, machine learning and deep learning methods can capture complex discriminative patterns in the data. Such approaches have demonstrated improved performance in the presence of spectral distortions and complex mixtures. In particular, convolutional neural networks and related deep learning architectures have shown promising results for isotope identification and have demonstrated improved robustness when trained using datasets that incorporate realistic variations such as gain drift (Turner et al., 2021). However, models that process full high-dimensional spectra, such as 1024-channel gamma-ray data, often require large numbers of parameters and extensive training datasets in order to generalize across diverse measurement conditions. Moreover, directly training classifiers on full high-dimensional spectra can increase sensitivity to nuisance variations such as gain drift, since small energy-axis shifts may affect many input channels simultaneously and reduce the stability of the learned representation. These models may also remain sensitive to other factors, including detector resolution variations and background mismatch (Li et al., 2022; Novak et al., 2018).

Dimensionality reduction offers a practical strategy for addressing these challenges by compressing highdimensional spectral data into a lowerdimensional representation that retains the most informative features. By reducing the dimensionality of the input data, this approach can decrease model complexity, improve computational efficiency, and potentially enhance generalization performance. Such strategies are particularly relevant for gamma-ray spectroscopy, where spectra are often characterized by high dimensionality, low count statistics, peak overlap, and calibration instabilities. Nevertheless, it remains an open question whether nonlinear dimensionalityreduction techniques provide a significant advantage over linear methods for radioisotope identification under realistic measurement conditions.

Autoencoder-based nonlinear dimensionality reduction has previously been explored for gamma-ray spectral analysis (Abdolpour et al., 2026). However, its comparative performance relative to classical linear approaches, particularly under challenging and nonlideal measurement conditions, has not been comprehensively evaluated. This study presents a comparative evaluation of principal component analysis (PCA) and an autoencoder (AE) to determine how linear and nonlinear feature compression influence radioisotope identification performance. PCA performs a linear transformation that projects the data onto a lowerdimensional subspace defined by the directions of maximum variance (Jolliffe, 2025; Jolliffe and Cadima, 2016), whereas an autoencoder learns a nonlinear latent representation by training a neural network to reconstruct the input data (Hinton and Salakhutdinov, 2006; Goodfellow et al., 2016). To ensure a fair comparison, both feature representations are coupled with the same multilayer perceptron (MLP) classifier, and their performance is evaluated using unseen test datasets that include both experimentally measured spectra and simulated spectra with different levels of gain drift.

The main contributions of this work are as follows. First, a systematic comparison is performed between linear and nonlinear dimensionalityreduction techniques, specifically PCA and AE, for multilabel radioisotope identification from gamma-ray spectra. To ensure a fair evaluation, both feature representations are coupled with the same multilayer perceptron (MLP) classifier so that the influence of the dimensionalityreduction method can be assessed independently of the classification architecture. Second, the proposed framework is evaluated under realistic and challenging conditions, including experimentally measured spectra and scenarios involving significant energy gain drift, rather than relying solely on idealized simulated datasets. Finally, the study provides quantitative evidence that nonlinear latent representations learned by the autoencoder yield improved robustness and generalization, demonstrating better transferability to real spectra and greater resilience to gain drift compared with linear PCA-based features. These findings provide practical guidance for the design of robust gamma-ray spectroscopy systems intended for realworld radioisotope identification tasks.

2 Methodology

2.1 Dataset Simulation

To ensure a controlled and comprehensive evaluation, a synthetic dataset of gamma-ray spectra was generated using the MCNP6 Monte Carlo simulation code. A cylindrical $2'' \times 2''$ NaI(Tl) scintillation detector was modeled. The pulse-height tally F8:P was used to obtain the deposited-energy spectrum in the NaI(Tl) crystal. To incorporate the finite energy resolution of the detector, Gaussian energy broadening was applied using an energy-dependent resolution function, $(FWHM(E) = a + b\sqrt{E + cE^2})$, where E denotes the photon energy and a, b, and c are empirical broadening parameters. This formulation allows the simulated photopeaks to reproduce the finite-resolution peak shapes expected from a real NaI(Tl) scin-

tillation detector. Gamma emissions from six common radioisotopes, Co-60, Cs-137, I-131, Ba-133, Am-241, and Tc-99m, were simulated. NaI(Tl) was selected because it is widely used in portable gamma-spectroscopy systems and, due to its moderate energy resolution, provides a realistic and challenging case for multi-isotope identification with peak overlap and Compton-continuum effects. For each radionuclide, the photon energy distribution was defined based on its characteristic gamma-ray emission energies and emission probabilities.

The radioactive sources were modeled as point-like photon sources. Photon emission was assumed to be uncollimated and isotropic over 4π, and no directional biasing or collimating angular distribution was imposed. For each simulated isotope configuration, spectra were generated at eleven source-to-detector distances ranging from 1 to 100 cm to represent different measurement geometries. These parameters affect the detector solid angle, full-energy peak efficiency, Compton continuum, peak-to-continuum ratio, and counting statistics, and therefore influence the measured spectral shape without changing the radionuclide composition.

The synthetic dataset was constructed from all non-empty combinations of the six isotopes, yielding 63 distinct isotope mixtures ranging from single-isotope to six-isotope cases. By varying the source-to-detector distance, MCNP6 computational run-time, a total of 4,200 spectra were generated across these combinations. At each distance, six MCNP6 simulation run times, ranging from 30 to 180 s in 30 s increments, were used for generating the simulated spectra. These simulation run times were specified using the CTME card in MCNP6. Increasing the CTME value allows the simulation to proceed for a longer computational time and generally improves the statistical quality of the calculated pulse-height spectra by reducing Monte Carlo statistical uncertainty. Environmental background radiation was modeled using MCNP6's built-in background source capability, which samples location-dependent neutron and photon fluxes based on geographic coordinates, including latitude, longitude, and altitude, from the background.dat dataset (Tutt et al., 2017). This approach provides a physics-based approximation of real-world environmental gamma backgrounds and enhances the realism of the synthetic spectra used for model training and evaluation.

2.2 Dimensionality Reduction Methods

Training a classifier directly on 1024-channel spectra not only increases model size and computational burden but can also heighten sensitivity to spectral variations caused by detector instabilities, such as gain drift. To address these challenges, two dimensionalityreduction strategies are employed: a nonlinear autoencoder (AE) and linear principal component analysis (PCA). The objective of this study is to compare these two approaches as alternative featurecompression methods for radioisotope identification, rather than to construct capacitymatched models. Accordingly, PCA and the autoencoder are treated as distinct representational frameworks, while the same downstream multilayer perceptron (MLP) classifier is used in all experiments to isolate the impact of the learned feature representation on final classification performance.

The dataset consists of 4,200 simulated spectra with 1024 channels. The data are randomly partitioned into an 80/20 train/test split, and an additional 80/20 split within the training set is used for validation during model optimization. PCA components were selected based on cumulative explained variance, which is the standard criterion for linear PCA, whereas the autoencoder latent dimension was selected based on reconstruction MSE, reflecting the natural optimization objective of the autoencoder.

2.2.1 Nonlinear Dimensionality Reduction Using an Autoencoder

The autoencoder is trained in an unsupervised manner to reconstruct the input spectrum, thereby learning a compact latent representation that captures key structural patterns in the data. Because spectra are normalized to the range $[0, 1]$ prior to training, a sigmoid activation is used at the decoder output to ensure reconstructions remain within this interval. The autoencoder employs a fully connected, symmetric encoder-decoder architecture. The encoder transforms the 1024-dimensional input into successive hidden layers of 256 and 128 neurons before mapping it to a 16-dimensional latent space. ReLU activation functions are used throughout the encoder, while the latent layer uses a linear activation. The decoder mirrors this structure in reverse, and a sigmoid activation is applied at the output layer to reconstruct the input spectrum.

Training is performed using mean squared error (MSE) loss and the Adam optimizer with an initial learning rate of 1×10^{-4}. The model is trained for up to 100 epochs with a batch size of 64. To mitigate overfitting and improve convergence stability, EarlyStopping and ReduceLROnPlateau are employed. The full autoencoder contains 595,728 parameters, with 297,360 parameters in the encoder alone.

The latent dimensionality is selected based on the reconstructionerror curve as a function of latent size (Fig. 1-left). The MSE decreases rapidly at small latent dimensions and reaches its minimum at dimension = 16, with larger latent sizes (e.g., 32 or 64) providing no stable or significant improvement. Accordingly, a 16-dimensional latent representation is adopted. After training, only the encoder is retained to generate 16dimensional features for use by the downstream MLP classifier.

2.2.2 Linear Dimensionality Reduction Using PCA

Principal component analysis (PCA) provides a linear and orthogonal projection onto a low-dimensional subspace that preserves maximum variance in the transformed coordinates. The number of principal components is determined using the cumulative explained-variance profile computed from the training set (Fig. 1-right).

Three PCA dimensionalities are examined:

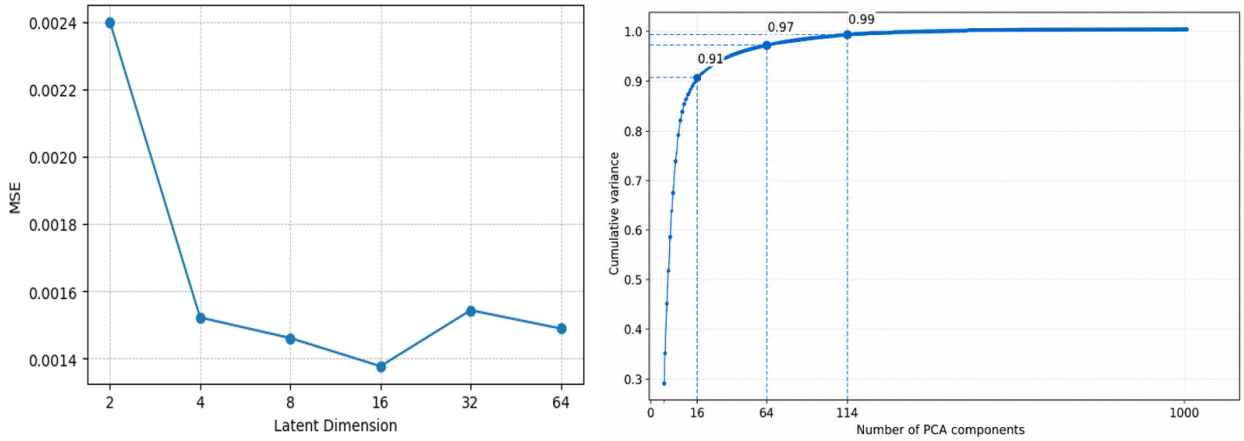


Figure 1: Left: Autoencoder reconstruction MSE versus latent dimensionality, showing minimal error at 16 dimensions. Right: PCA cumulative explained variance from the training set, highlighting 16, 64, and 114 components as key dimensionalities. PCA components were selected based on cumulative explained variance, whereas the autoencoder latent dimension was selected based on reconstruction MSE, reflecting the natural objective of each method.

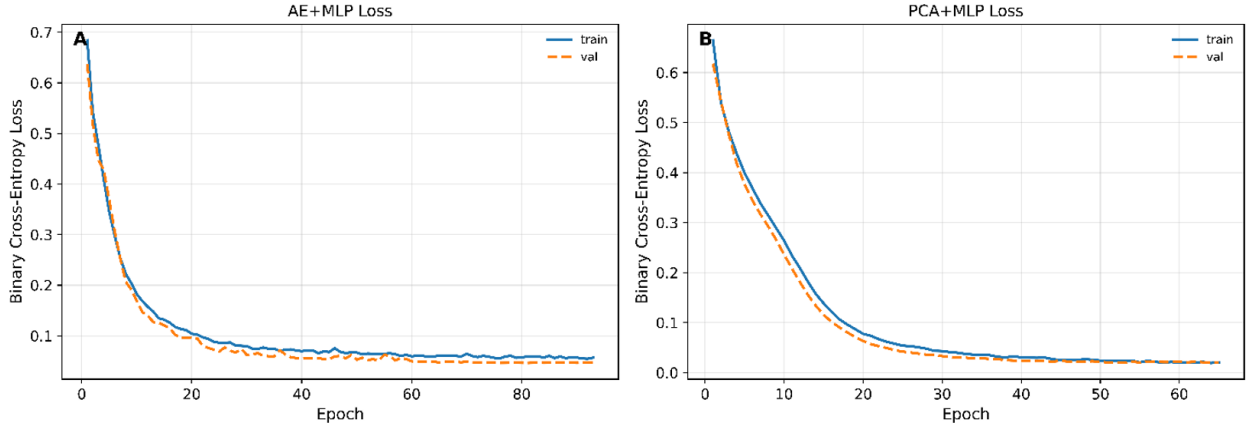


Figure 2: Training and validation binary cross-entropy loss curves for the MLP classifier using (A) autoencoder-based features (AE+MLP) and (B) PCA-based features (PCA+MLP), demonstrating stable convergence and limited overfitting for both models.

- 16 components, matching the dimensionality of the AE latent space and retaining approximately 91% cumulative variance;
- 64 components, representing a higher-capacity linear embedding that retains approximately 98%;
- 114 components, corresponding to nearsaturation with approximately 99% cumulative variance.

In the primary comparison, PCA-16 serves as the dimensionmatched baseline against AE-16. Additional evaluations using 64 and 114 components assess the sensitivity of classification performance to increased linear embedding capacity and the role of cumulative explained variance. It should be noted that the PCA dimensionality was selected according to cumulative explained variance, whereas the AE latent dimensionality was selected according to reconstruction MSE; these criteria reflect the natural objectives of the two dimensionality-reduction methods.

2.3 Multilayer Perceptron Classifier

Both feature-extraction pipelines are followed by the same multilayer perceptron (MLP) classifier to ensure that the performance comparison primarily reflects the quality of the learned feature representations rather than differences in classifier capacity. The compressed features produced by either PCA or the autoencoder encoder are used as inputs to the MLP.

The MLP architecture consists of two fully connected hidden layers with 16 and 8 neurons, respectively. ReLU activation functions are employed in both hidden layers, and batch normalization is applied after each layer to improve training stability. The hiddenlayer sizes are selected to provide sufficient nonlinear modeling capacity while avoiding unnecessary model complexity. The output layer contains six neurons with sigmoid activation, yielding independent probability estimates for the presence of each radioisotope. A fixed decision threshold of 0.5 is applied for multilabel classification.

The model is trained using the binary crossentropy loss function and the Adam optimizer with a learning rate of 0.001 for up to 100 epochs. Early stopping based on the validation loss is employed to reduce overfitting. The training process exhibits stable convergence, as evidenced by the monotonic decrease in both training and validation losses.

Table 1: Comparison of model complexity, training time, and inference time for the evaluated classification approaches.

Model	Feature dimension	Trainable parameters	Total parameters	Total training time (s)	Inference time per spectrum (ms)
AE+MLP	16	596,238	596,286	62.68	0.0267
PCA-16+MLP	16	510	558	15.09	0.0105
Raw-MLP	1024	16,638	16,686	19.52	0.0155

Figure 2 illustrates the training and validation loss curves for the two feature-extraction pipelines. For the AE+MLP model (Fig. 2-A), the loss decreases rapidly during the initial epochs, followed by a more gradual decline, with training and validation curves remaining closely aligned, indicating limited overfitting and stable learning behavior. Similarly, the PCA+MLP model (Fig. 2-B) demonstrates a smooth and steady reduction in loss, with only a small gap between the training and validation curves, confirming stable optimization and convergence.

The computational complexity and timing performance of AE+MLP, PCA-16+MLP, and raw-spectrum MLP were also evaluated. All measurements were performed on the same CPU-only system using Python 3.7.1 and TensorFlow 1.14.0 under Windows 10. The reported training time includes all training steps required for each method, and the inference time is reported as the average prediction time per spectrum. Table 1 summarizes the results. PCA-16+MLP is the most compact and fastest method, while AE+MLP requires more parameters and longer training time due to autoencoder learning. However, the inference time of AE+MLP remains very small, indicating that the improved robustness of the nonlinear representation is obtained with practical computational cost.

2.4 Classification Evaluation Method

The performance of the proposed multilabel radioisotope identification framework is evaluated using a comprehensive set of classification metrics designed to assess both isotopewise detection accuracy and overall system reliability under realistic conditions. A combination of thresholdbased and thresholdindependent evaluation criteria is employed.

2.4.1 Binary Classification Framework

Although the identification task is inherently multilabel, performance evaluation is conducted based on binary classification principles by treating each isotope as an independent detection problem. The neural network outputs perisotope probabilities through sigmoid activation functions in the final layer. A fixed decision threshold of 0.5 is applied, whereby an isotope is classified as present if its predicted probability exceeds the threshold and absent otherwise.

Each prediction falls into one of four standard categories: true positive (TP), false positive (FP), true negative (TN), or false negative (FN), as summarized in Table 2. These outcomes form the basis for computing all subsequent performance metrics.

2.4.2 Precision, Recall, and F1Score

Classification performance is quantified using precision, recall, and $F1$-score. Precision measures the proportion of correctly identified isotopes among all predicted positives, reflecting the model's resistance to false alarms. Recall measures the proportion of correctly detected isotopes among all actual positives, indicating detection sensitivity. The $F1$-score is defined as the harmonic mean of precision and recall, $F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$ and provides a balanced measure of detection accuracy, particularly in multilabel scenarios.

To assess performance across varying levels of spectral complexity, test spectra containing combinations of one to six radioisotopes are generated using MCNP6 simulations under different geometric configurations and environmental background conditions. Under nominal calibration conditions (i.e., without gain drift), fivefold cross-validation is employed on the simulated dataset to obtain statistically robust performance estimates.

2.4.3 ROC Curve and AUC

In addition to thresholddependent metrics, thresholdindependent performance is evaluated using receiver operating characteristic (ROC) curves and precision-recall (PR) curves. The ROC curve illustrates the tradeoff between the true positive rate (TPR) and false positive rate (FPR) as the decision threshold varies, providing insight into the classifier's discriminative capability. The area under the ROC curve (AUC) is used as a scalar summary metric, with values approaching 1 indicate strong class separability.

Precision-recall curves are also examined, as they are particularly informative in scenarios with class imbalance and emphasize the relationship between detection sensitivity and falsealarm control. ROC and PR curves are generated for each isotope class to evaluate classwise performance consistency across the full range of decision thresholds. Figure 3 presents the ROC and PR curves for the AE+MLP and PCA+MLP models, confirming high classwise discrimination across isotope classes.

Table 2: Possible outcomes in binary classification for isotope detection.

Actual Label	Predicted Positive	Predicted Negative
Positive (1)	True Positive (TP)	False Negative (FN)
Negative (0)	False Positive (FP)	True Negative (TN)

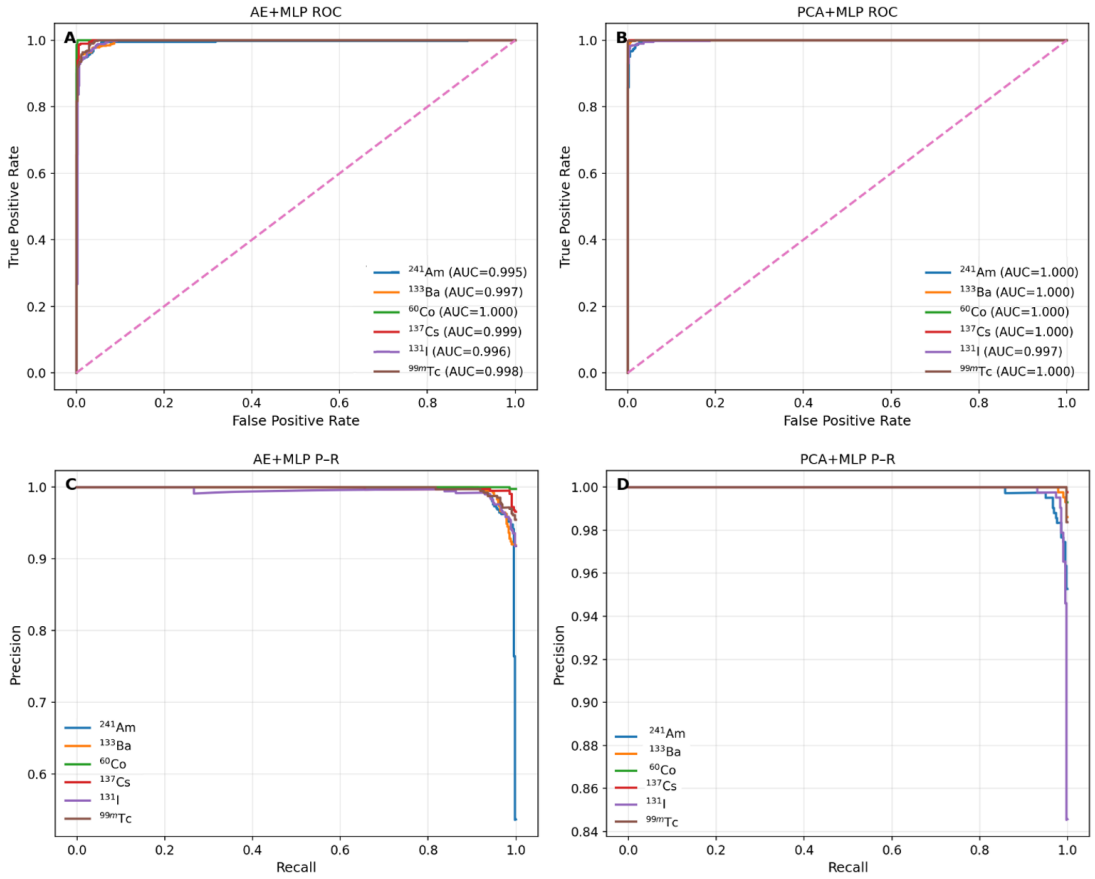


Figure 3: ROC and PR curves for each isotope class showing high classification performance for (A,B) AE+MLP and (C,D) PCA+MLP.

3 Results and Discussion

To provide a clear understanding of the comparative performance of the proposed methods, this section first illustrates representative multiisotope gamma-ray spectra and examines how these spectra are represented in the reduceddimensional feature spaces produced by PCA and the autoencoder. Subsequent subsections analyze the robustness of both models under realistic sources of spectral variation, including gain drift and experimental data acquisition.

3.1 Displaying the spectrum and the compressed dimensions

Figures 4-A1 and 4-A2 present representative simulated gamma-ray spectra with 1024 energy channels for mixtures of Am-241, Ba-133, and Co-60 isotopes under two count-rate conditions: low-count and high-count, respectively. The low-count spectrum is shown in Fig. 4-A1, whereas the high-count spectrum is shown in Fig. 4-A2. The differences between these two conditions are primarily observed in the level of statistical fluctuations, the intensity of the Compton continuum, and the overall signal-to-background ratio.

The corresponding 16-dimensional latent activations obtained from the autoencoder are shown in Fig. 4-B. Each latent dimension reflects the contribution of a learned spectral feature to reconstructing the original spectrum. Notably, even in the lowcount case, the latent vector remains structured rather than random, indicating that the autoencoder has captured stable, recurring spectral patterns instead of noise. This behavior suggests that the nonlinear encoding effectively isolates meaningful spectral factors that persist across variations in counting statistics.

By contrast, Fig. 4-C depicts the reduced PCA representation of the same spectra. As expected, the PCA model retains most of the variance within the first few components, leaving the later components close to zero. However, comparison between low and highcount cases reveals larger changes in the PCA score distributions than in the AE latent representations, due to PCA's inherent sensitivity to global variance. Consequently, PCA embeddings may fluctuate more strongly when statistical noise or minor gain drift affect the measured spectra.

3.2 Evaluating robustness under realistic spectral variations

To assess the robustness and generalization capability of the proposed identification framework, additional evaluations were conducted on spectra containing realistic perturbations that were intentionally excluded from the training data. These experiments aim to determine how well the models maintain their classification performance when

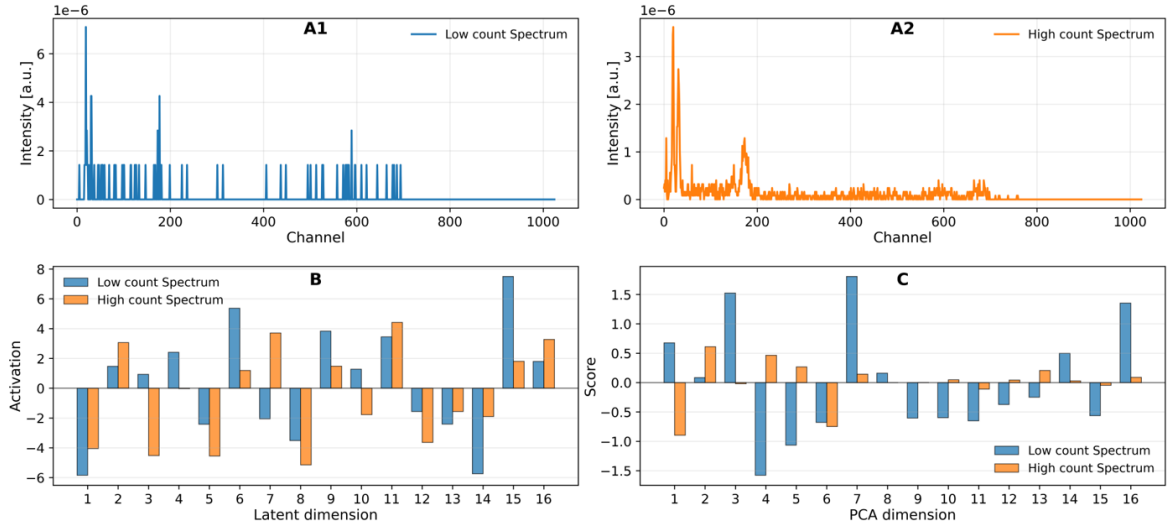


Figure 4: Simulated gamma spectra of Am-241, Ba-133, and Co-60 radioisotopes with 1024 channels at different counting rates. The low-count spectrum is shown separately in panel A1, while the high-count spectrum is shown in panel A2. The corresponding 16-dimensional autoencoder latent representations are presented in panel B, and the reduced 16-dimensional PCA score representations are shown in panel C.

Table 3: Comparison of model performance for isotope identification on experimental spectra and under severe gain drift (±20%).

	Gain- drift spectra			Experimental spectra		
	Precision	Recall	$F1$-score	Precision	Recall	$F1$-score
AE+MLP	0.83	0.82	0.81	1.00	0.98	0.99
PCA-114+MLP	0.75	0.72	0.72	0.90	0.90	0.89
PCA-16+MLP	0.74	0.71	0.71	0.88	0.95	0.91
PCA-64+MLP	0.74	0.70	0.71	0.88	0.89	0.86
MLP	0.77	0.64	0.68	0.89	0.94	0.91

exposed to spectral distortions that commonly arise in practical measurement scenarios.

In this context, particular attention is given to gain drift, which represents one of the most frequent and challenging sources of spectral inconsistency in scintillation-based gamma spectroscopy systems. Such variations alter the effective calibration of the detector, resulting in shifts of the photopeak energies and slight distortions in the Compton continuum. The following subsection details the controlled evaluation setup designed to analyze the influence of these variations on model performance while keeping all other conditions unchanged.

3.2.1 Spectra with gain drift

In practice, gain drift may arise from temperature fluctuations, high-voltage instability, imperfect calibration, or detector aging. To evaluate model robustness under calibration mismatch, controlled perturbations were applied only to the test spectra by performing a multiplicative resampling of the channel axis at drift levels of ±5%, ±10%, and ±20%. Under this perturbation scheme, the displacement of each photopeak is proportional to its energy, such that higherenergy peaks experience larger absolute shifts than lowerenergy peaks. Although real gain drift may involve more complex nonlinear effects, the purpose of this experiment is to compare the relative behavior of the two representations under identical perturbation conditions.

Figure 5 illustrates the variation of Precision, Recall, and $F1$-score as a function of the number of isotopes present in the mixture (from 1 to 6) for the three drift levels. As shown, the AE+MLP model consistently outperforms the PCA+MLP model across nearly all mixture compositions and for all three evaluation metrics at ±5%, ±10%, and ±20% drift levels. This advantage is already observable at ±5%, remains evident at ±10%, and becomes more pronounced under the more severe ±20% drift condition. Although the performance of both models decreases as the drift magnitude increases, the degradation is systematically greater for PCA+MLP, whereas AE+MLP maintains a more stable and balanced behavior. These observations indicate that the nonlinear latent representation learned by the autoencoder provides greater robustness to gain drift than the linear PCA-based representation.

The progressive performance degradation observed for larger mixture sizes and stronger drift levels can be attributed to increasing spectral complexity combined with gaininduced feature displacement. As the number of radioisotopes in a spectrum increases, overlap between photopeaks and Compton continua reduces class separability and makes multilabel identification more challenging. At the same time, gain drift shifts spectral features along the energy axis, weakening the alignment between the observed spectra and the patterns learned during training. These combined effects explain the reduction in precision,

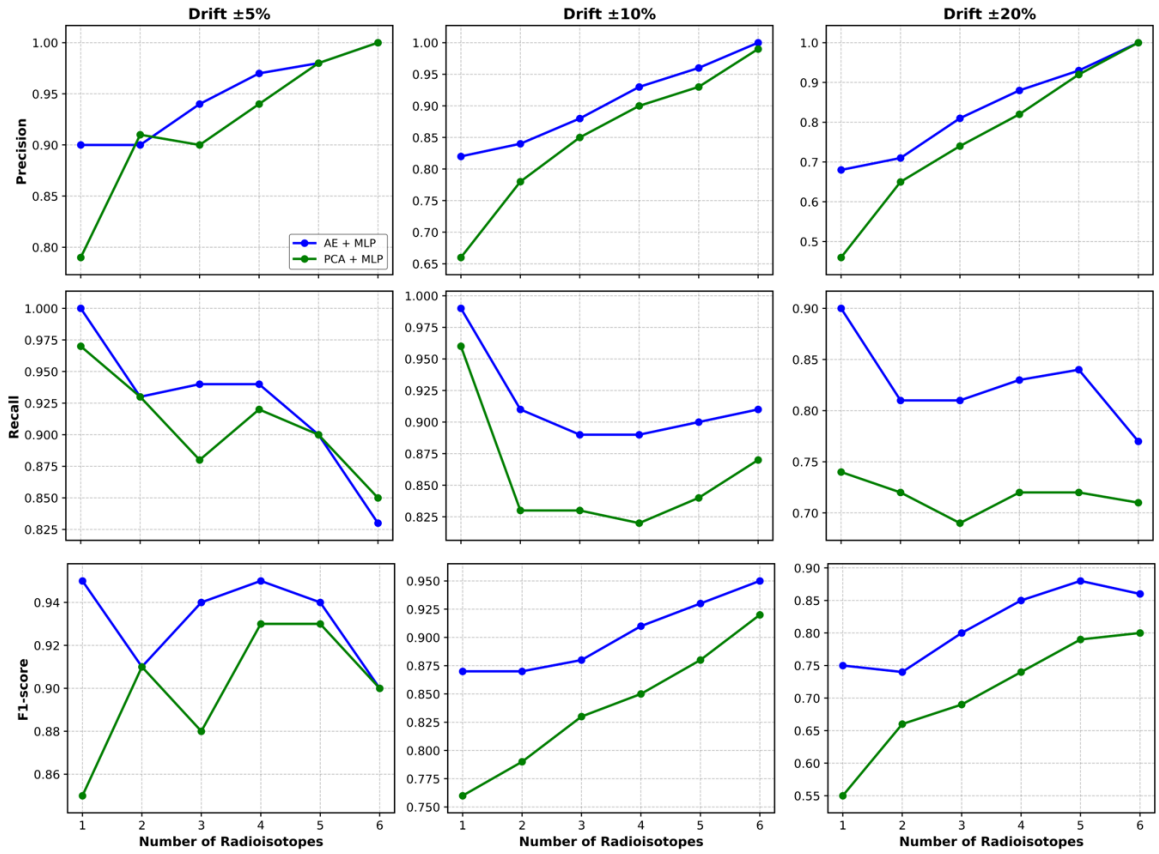


Figure 5: Comparative performance of AE+MLP (blue) and PCA+MLP (green) across different drift levels (±5%, ±10%, ±20%) for 1 to 6 radioisotopes. The three rows represent Precision, Recall, and $F1$-score, respectively.

recall, and $F1$-score under more demanding conditions. Under these circumstances, the AEbased representation appears to preserve a more stable discriminative structure than PCA. Table 3 shows a consistent trend, with AE+MLP maintaining a more balanced precision-recall trade-off than the PCA-based models under both gain-drift and experimental conditions.

3.2.2 Validation with experimental spectra

To evaluate the generalization capability of the proposed models on real measurements, an experimental dataset consisting of 90 gamma-ray spectra was collected using four laboratory radioisotopes: Am-241, Ba-133, Co-60, and Cs-137. The activities of these sources were 107, 115, 101, and 57.7 kBq, respectively. The experimental dataset included only four of the six radioisotopes used during training; consequently, two output classes were absent from the experimental test set, and all reported metrics reflect performance on the available four-isotope subset.The dataset was constructed from all non-empty combinations of the four available isotopes, resulting in 15 source configurations: four single-isotope cases, six two-isotope mixtures, four three-isotope mixtures, and one four-isotope mixture. For multi-isotope cases, the corresponding point sources were measured together under different source-to-detector distances and measurement times. Six spectra were acquired for each source configuration, yielding 90 spectra in total. This corresponds to 24 single-isotope spectra, 36 two-isotope spectra, 24 three-isotope spectra, and 6 four-isotope spectra. Before classification, the experimental spectra were arranged into the same 1024-channel format and normalized using the same preprocessing procedure as the simulated spectra.

Importantly, all models were trained solely on simulated data and then tested directly on the experimental spectra without any retraining or fine-tuning. In addition, to assess robustness to gain drift, the models were also evaluated under a challenging scenario involving severe gain drift of ±20%.

The ability of the models to perform successfully on experimental spectra without finetuning indicates that, despite visible domain differences, the measured spectra preserve the essential spectral structures learned from simulation. As illustrated in Fig. 6 for a representative Co-60 + Cs-137 mixture, the simulated and experimental spectra are not identical in the raw domain, particularly with respect to the continuum shape and relative peak intensities. Nevertheless, the main photopeak locations and the overall spectral structure remain consistent. Their corresponding compressed representations in both the autoencoder (AE) and PCA feature spaces also exhibit broadly similar patterns, suggesting that the reduceddimensional representations capture stable and transferable spectral characteristics. This observation helps explain why inference on real spectra remains effective even when the models are trained exclusively on simulated data. To-

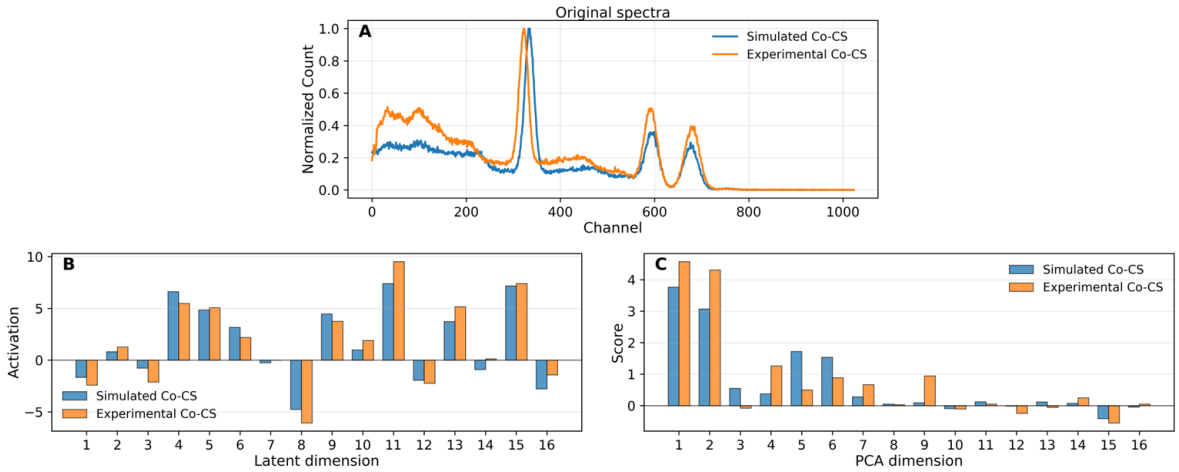


Figure 6: Representative simulated and experimental Co-60 and Cs-137 spectra and their corresponding compressed representations. (A) shows the original spectra in the raw 1024-channel domain. (B) shows the 16-dimensional AE latent representation. (C) shows the 16-dimensional PCA-compressed representation.

gether with the quantitative results, it indicates that the AEbased representation is not limited to simulationspecific details, but instead encodes spectral structure that remains informative in real measurements.

Quantitative results for the experimental spectra are summarized in Table 3. The best overall performance was achieved by the AE+MLP model, which obtained Precision=1.00, Recall=0.98, and $F1$-score=0.99, indicating that nearly all present isotopes were correctly identified with excellent selectivity. For comparison, an MLP trained directly on the raw 1024-channel spectra, without dimensionality reduction, was also evaluated. This rawspectrum MLP achieved Precision=0.89, Recall=0.94, and $F1$-score=0.91, demonstrating good performance but still falling short of the AEbased pipeline.

Among the PCA-based approaches, PCA-16+MLP achieved Precision=0.88, Recall=0.95, and $F1$-score=0.91, yielding performance comparable to that of the rawspectrum MLP. This highlights an important advantage of PCA-based models, which achieve comparable performance with significantly fewer input features and a much smaller number of model parameters. Increasing the number of retained PCA components did not lead to a consistent improvement in experimental performance. Specifically, PCA-64+MLP achieved Precision=0.88, Recall=0.89, and $F1$-score=0.86, while PCA-114+MLP achieved Precision=0.90, Recall=0.90, and $F1$-score=0.89. These results indicate that simply preserving more variance in PCA does not necessarily yield features that generalize better to real measured spectra, and that changes in PCA dimensionality primarily affect the precision-recall balance rather than producing a clear gain in overall performance.

The differences between the models become more pronounced under the severe gainshift condition (±20%), as also reported in Table 3. Under this challenging scenario, AE+MLP again achieved the best overall performance, with Precision = 0.83, Recall = 0.82, and $F1$-score = 0.81, demonstrating superior robustness to substantial peak displacement. In contrast, the rawspectrum MLP experienced a marked degradation in performance, particularly in recall (Precision = 0.77, Recall = 0.64, $F1$-score = 0.68), indicating that a significant fraction of present isotopes were missed. The PCA-based models also showed reduced performance under ±20% gain drift. PCA-16+MLP achieved Precision = 0.74, Recall = 0.71, and $F1$-score = 0.71, while PCA-64+MLP and PCA-114+MLP yielded $F1$-scores of 0.71 and 0.72, respectively. These results suggest that retaining additional PCA components and therefore preserving more global variance does not necessarily produce feature representations that remain stable and discriminative under severe gain drift. Overall, the experimental validation confirms that AE+MLP provides superior robustness to gain drift and better transferability from simulated to experimental spectra compared with PCA-based and rawspectrum approaches. The present results can be compared with recent deep-learning-based radionuclide identification studies. For example, Li et al. (Li et al., 2022) used a feature-enhancement strategy combined with a one-dimensional neural network for low-count spectra with multiple radionuclides, while Wang et al. (Wang et al., 2022) employed a CNN with a channel attention module for multiple-radionuclide identification. Although direct numerical comparison is limited by differences in isotope libraries, datasets, detector models, and evaluation conditions, these studies generally rely on more elaborate feature-enhancement or deep-learning architectures. In contrast, the present AE+MLP framework uses a compact 16-dimensional latent representation and a small common MLP classifier, yet achieves high performance on experimental spectra and maintains better robustness under severe gain-drift conditions than PCA-based and raw-spectrum MLP models. This suggests that nonlinear autoencoder-based feature compression can provide a lightweight and robust alternative for practical gamma-ray spectral identification.

It must be mentioned that this study was limited to six common radionuclides; therefore, the generalizability of the findings to a broader isotope library requires fur-

ther investigation. In addition, the simulations and experimental validation were performed using NaI(Tl) spectra, and future studies should evaluate the proposed framework under different detector types, detector resolutions, and measurement conditions.

4 Conclusions

The results of this study demonstrate that while both PCA and autoencoder-based dimensionality-reduction methods are capable of supporting highly accurate radioisotope identification under ideal and wellcalibrated conditions, their behavior diverges substantially when evaluated under more realistic and challenging scenarios. Under nominal conditions without gain drift, both approaches achieved near-perfect multi-label classification performance, indicating that either linear or nonlinear compression can preserve the dominant spectral structures required for accurate isotope identification.

However, clearer and more meaningful differences emerged when the models were evaluated outside the training domain. Two scenarios were investigated in detail: validation on real experimental spectra and evaluation under severe gain shift ($\pm 20\%$). In both cases, the autoencoder-based pipeline consistently outperformed the PCA-based approaches and the rawspectrum MLP baseline. Specifically, AE+MLP achieved an $F1$-score of 0.99 on experimental spectra, demonstrating excellent transferability from simulation to real measurements. Under severe gain drift, where spectral peaks are substantially displaced, AE+MLP maintained an $F1$-score of 0.81, whereas the performance of PCA-based models fell to 0.71-0.72, and the rawspectrum MLP showed a more significant degradation ($F1 = 0.68$).

These findings indicate that the nonlinear latent representation learned by the autoencoder captures a more stable and physically meaningful structure of gamma-ray spectra-one that remains informative even when the spectra deviate from the conditions seen during training. In contrast, PCA's variance-based linear representation is more sensitive to calibration errors, mixture complexity, and domain mismatch. Taken together, the results show that autoencoder-based feature extraction provides superior robustness, better generalization to real measured spectra, and greater resilience to severe gain drift, making it a more reliable choice for practical radioisotope identification systems operating under non-ideal conditions. Increasing the number of principal components did not lead to a consistent improvement in performance under either experimental conditions or severe gain drift. Instead, higher-dimensional PCA variants mainly produced modest changes in the balance between precision and recall, without yielding a clear gain in overall robustness. Overall, the results demonstrate that nonlinear dimensionality reduction using an autoencoder provides more stable and transferable features for multi-label radioisotope identification than linear PCA-based compression, particularly when the data deviate from ideal simulation conditions. In future work, the isotope library will be expanded to include a wider range of sources and mixture compositions. In addition, the proposed framework will be evaluated using spectra acquired under different detector resolutions and measurement conditions in order to further assess its reliability in practical applications.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work.

Funding

The authors declare that no funds, grants, or other financial support were received during the preparation of this manuscript.

References

Abdolpour, K., Masoudi, S. F., and Fathi, A. (2026). Using autoencoder feature compression combined with deep neural networks for precise radioisotope identification in composite spectra. *Results in Physics*, page 108608.

Burr, T. and Hamada, M. (2009). Radio-isotope identification algorithms for NaI γ spectra. *Algorithms*, 2(1):339–360.

Galib, S., Bhowmik, P., Avachat, A., et al. (2021). A comparative study of machine learning methods for automated identification of radioisotopes using NaI gamma-ray spectra. *Nuclear Engineering and Technology*, 53(12):4072–4079.

Goodfellow, I., Bengio, Y., Courville, A., et al. (2016). *Deep learning*, volume 1. MIT press Cambridge.

Hinton, G. E. and Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507.

Jolliffe, I. (2025). Principal component analysis. In *International Encyclopedia of Statistical Science*, pages 1945–1948. Springer.

Jolliffe, I. T. and Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202.

Khatiwada, A., Klasky, M., Lombardi, M., et al. (2023). Machine Learning technique for isotopic determination of radioisotopes using HPGe γ-ray spectra. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1054:168409.

Li, C., Liu, S., Wang, C., et al. (2022). A new radionuclide identification method for low-count energy spectra with multiple radionuclides. *Applied Radiation and Isotopes*, 185:110219.

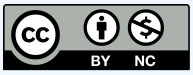
Novak, R., Bahri, Y., Abolafia, D. A., et al. (2018). Sensitivity and generalization in neural networks: an empirical study. *arXiv preprint arXiv:1802.08760*.

Turner, A. N., Wheldon, C., Wheldon, T. K., et al. (2021). Convolutional neural networks for challenges in automated nuclide identification. *Sensors*, 21(15):5238.

Tutt, J., Anderson, C., and McKinney, G. (2017). Background-Source Cosmic-Photon Elevation Scaling and Cosmic-Neutron/Photon Date Scaling in MCNP6. *Physics Procedia*, 90:237–241.

Wang, Y., Zhang, Q., Yao, Q., et al. (2022). Multiple radionuclide identification using deep learning with channel attention module and visual explanation. *Frontiers in Physics*, 10:1036557.

Zhang, S., Marsden, E., and Goulermas, J. Y. (2022). Isotope identification using artificial neural network ensembles and bin-ratios. *IEEE Transactions on Nuclear Science*, 69(6):1194–1202.



To cite this article:

Abdolpour, K, Masoudi, S F and Fathi, A (2026). Comparison of nonlinear autoencoder and linear PCA dimensionality reduction in gamma-ray spectroscopy based radioisotope identification. *Radiation Physics and Engineering*, 7(3), 63–73. doi: 10.22034/rpe.2026.580714.1354

DOI: 10.22034/rpe.2026.580714.1354

To link to this article: https://doi.org/10.22034/rpe.2026.580714.1354