RESEARCH ARTICLE

A CNNSENet-PSO-SVM approach for multi-classification of mass spectrometry

Liping Wu[1], Jiayao Pan[1], Cheng Xiang[2], Zhou Yang[3], Yong Li[1, *]

[1]Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, Yunnan, China. [2]Faculty of Life Sciences and Technology, Kunming University of Science and Technology, Kunming, Yunnan, China. [3]Zunyi Medical University, Zunyi, Guizhou, China.

Received: May 30, 2024; **accepted:** August 10, 2024.

Mass spectrometry (MS) is a technique used to determine the mass of molecules or atoms within a sample and their relative abundance. It can identify unknown compounds by comparing their mass spectra to known spectra in databases, which helps the determinations of the molecular formula and structure of the compound. However, the current cost of commercially available software is high, and there are relatively few applications of machine learning and deep learning methods in mass spectrometry recognition. To tackle these issues, a convolutional neural network within squeeze and excitation net-particle swarm optimization-support vector machine (CNNSENet-PSO-SVM) was proposed for multi-classification of mass spectrometry data. The method aimed to provide an effective and accurate mass spectra multi-classification method combining machine learning with deep learning to identify unknown mass spectra in mass spectral libraries. The approach involved preprocessing the original molecular compound mass spectrometry dataset using mass-to-charge ratio and intensity values as input and employing CNN for feature extraction. A channel attention mechanism module, SENet, was integrated within CNN to enhance feature extraction capabilities. Subsequently, the extracted feature vectors were fed into SVM with parameters optimized using the PSO algorithm for classification. Compared to traditional methods, the proposed model achieved an accuracy rate of 90.55% on the test set, demonstrating its potential for analyzing small molecular compound MS data.

Keywords: mass spectrometry; classification; CNN; channel attention mechanism; PSO; SVM.

[*]**Corresponding author:** Yong Li, Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, Yunnan 650550, China. Emails: 12309058@kust.edu.cn.

Introduction

Mass spectrometry (MS) is an essential analytical technique in chemistry. It involves ionizing and fragmenting organic molecules, separating and detecting ions based on their mass-to-charge ratio (M/Z), and analyzing the resulting mass spectra [1]. Mass spectrometry identification uses mass spectrometry techniques to identify and characterize compounds. It plays a crucial role in various domains including chemical analysis [2, 3], drug development [4, 5], and biomedical sciences [6, 7]. How to accurately extract meaningful information from large, high-dimensional datasets and, at the same time, effectively utilize this information for mass spectrometry identification and classification is one of the main challenges in mining and analyzing mass spectrometry data. Manual analysis often falls short in meeting these

challenges [8-10]. Conventional commercial software packages such as OpenMS and PEAKS rely on similarity calculations between different mass spectra for identification [11, 12]. However, the use of commercial software generates a high cost, and it is not easy for laymen without specialized knowledge to utilize commercial software for mass spectrometry. Identifying unknown small molecule compounds is a task frequently encountered in scientific research and practical applications. Researchers can search and compare mass spectral libraries to obtain structural information about unknown compounds and further investigate their properties and functions. Accurately and efficiently identifying compounds to be tested is a major challenge.

The application of machine learning and deep learning techniques in analyzing and classifying MS data has been advanced due to computer science. SVM is a typical machine learning algorithm commonly used for the classification of mass spectra. Wu *et al*. proposed a mass spectrometry analysis method with PPCA-SVM applied to the early classification of ovarian cancer with an average prediction accuracy of 90.80% and sensitivity and specificity of 92.98% and 88.97%, respectively [13]. Although the model performed well in the dichotomous problem of health and ovarian cancer, challenges remained when confronted with multi-category categorization problems. Lee *et al*. improved the differentiation of similar species by using an SVM model to classify *Mycobacterium abscessus* and *Mycobacterium fortuitum* by learning the positive and negative markers extracted separately in each group [14]. However, the classification accuracy of their model needed to be further improved. Convolutional neural networks (CNN) are representative algorithms of deep learning and can be used for feature extraction and classification tasks in the analysis of mass spectrometry data. Seddiki *et al*. developed a cumulative learning method using transfer learning in conjunction with a one-dimensional convolutional neural network to classify clinical MS data from rat brain with over 98% accuracy [15]. However, the application of this method to a wider range of clinical scenarios still needed to be improved. Petrovsky *et al*. investigated the effectiveness of 1D-CNN and 3D-CNN neural networks in classifying three types of cancers using metabolomics-proteomics HPLC-MS/MS data. Using histology data in the Mascot Generic Format, they classified various cancer phenotypes such as kidney cancer, ovarian cancer, and healthy individuals with 95% accuracy [16]. Nonetheless, further adaptation to multi-classification tasks with different cancer types and data types was required. Lv *et al*. employed a convolutional neural network to detect multiple compounds from tandem MS data and achieved an excellent accuracy of 98% in detecting mixture MS data from the Human Metabolome Database (HMDB) [17]. The results of these studies demonstrated promise and, in practical applications, the classification of more complex samples such as the rapid identification of one substance among thousands of different substances required further research and methodological improvements. Effective application of mass spectrometry data analysis to a wider range of fields is essential.

Current challenges in mass spectrometry analysis include inefficient manual analysis, high cost of commercial software, and difficulty in achieving multiple classifications by machine learning or deep learning methods. This research proposed a model that combined machine learning and deep learning together using a convolutional neural network within squeeze and excitation net (CNNSENet) that integrated the SENet module into a CNN to enhance feature extraction and particle swarm optimization-support vector machine (PSO-SVM) that PSO algorithm optimized the penalty factor and kernel parameters of SVM to improve the accuracy of SVM in classifying mass spectra of various compounds to overcome the current challenges of mass spectrometry analysis. This study provided an essential technological reference for mass spectrometry in the fields of drug discovery and biomedical research.

Materials and methods

Sample source

The mass spectrometry dataset was provided by the Department of Fundamental Chemistry at Kunming University of Science and Technology, Kunming, Yunnan, China and contained spectral data for 11,532 molecular ions ([M+H]+ and [M+H]-). These molecular ions were from 1,000 different herbal small molecule compounds, and the mass spectra were generated at 14 different voltages ranging from ±10 V to ±70 V. Each mass spectrum was represented as a two-dimensional dataset consisting of M/Z values and corresponding intensity (I) values. The number of M/Z and I data pairs per spectrum ranged from 1 to 100.

Data preprocessing

The mass-to-charge ratios were normalized by instrument calibration. Normalization of the intensity information in the spectra was necessary to minimize the effect of intensity differences on the results of the data analysis and facilitate the comparability of subsequent feature extraction. To address the data imbalance in the dataset, additional data samples were introduced to balance the distribution. Specifically, the dataset with three different intensities were introduced, which included Gaussian white noise with an average intensity of 1 and variance of 0.2, an average intensity of 4 and variance of 0.8, and an average intensity of 8 and variance of 1.6. This augmentation served two main purposes as enhancing the diversity of the training data and improving the robustness and generalization of the model.

Multi-classification model based on CNNSENet - PSO-SVM

To identify unknown mass spectral data in mass spectral libraries, the multiple classification model of CNNSENet and PSO-SVM were utilized to classify the mass spectra of small molecule compounds using the mass-to-charge ratio and intensity data of mass spectral data. The technical route was shown in Figure 1. CNN is a supervised learning model that automatically extracts feature information and exhibits strong generalization capabilities, especially on large-scale datasets [18]. The attention mechanism is commonly used to process sequential or image data [19]. Channel Attention Mechanism, also known as Squeeze-and-Excitation (SENet), is a technique that effectively enhances the performance of CNN feature extraction [20]. This model selected CNN combined with SENet for feature extraction of mass spectrometry data. PSO algorithm that mimics the collective foraging behavior of birds in D-dimensional search space, is a heuristic optimization algorithm usually used to solve a variety of optimization problems [21]. SVM is a supervised machine learning technique commonly used to solve statistical classification and regression analysis problems [22]. The PSO optimization algorithm was used to adaptively find the optimal covariate penalty factor and kernel function parameters of SVM, so that SVM had higher classification accuracy. The flowchart of the method for multiple classification of mass spectrometry data using CNNSENet and PSO-SVM was shown in Figure 2. To make the intensity information of mass spectrometry data convenient for further data processing and analysis, the intensity of each mass spectrum needed to be normalized. In this study, the maximum value linear transformation method was used to ensure that the relative intensity of each mass spectrum was in the range of 0 to 100 as shown in Equation (1) [23].

$$I_N = \frac{I_O}{I_{\max}} \times 100 \quad (1)$$

where I_o was the information I of a mass spectrometer. I_{max} was the maximum value of the relative information I of the mass spectrometer. The normalized data also needed to incorporate different levels of white noise, enhancing the model's generalization ability. After preprocessing, the MS dataset underwent feature extraction using CNNSENet. The MS data in this research only included M/Z and I, which could be viewed as one-dimensional sequential data. There might be patterns and

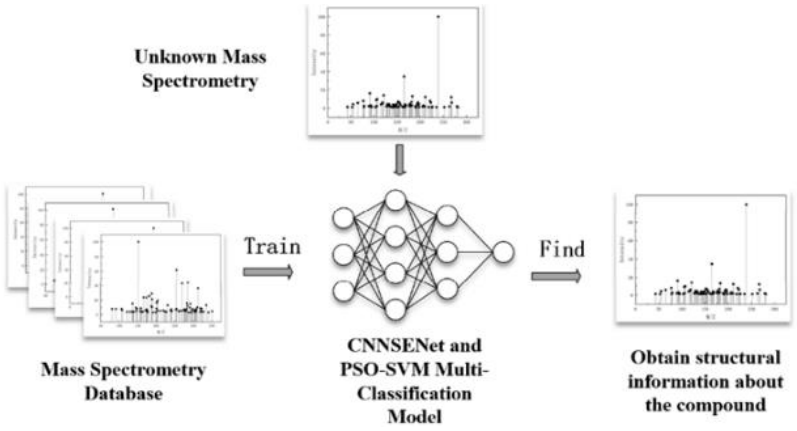

Figure 1. Technical routes for multi-categorization of mass spectrometry data.

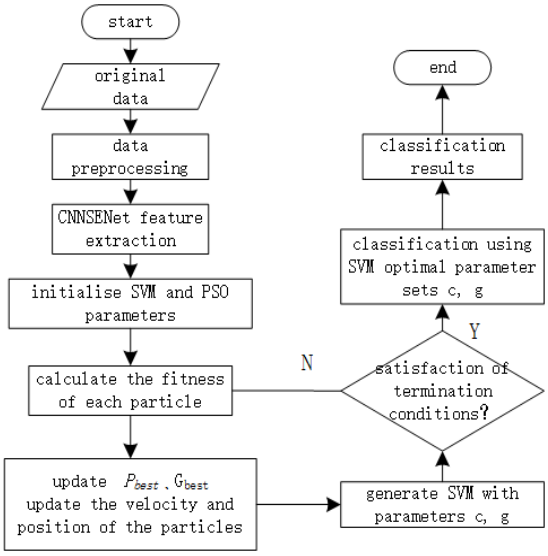

Figure 2. Flowchart of mass spectrometry data recognition based on CNNSENet and PSO-SVM model.

relationships between mass-to-charge ratio and intensity, which a one-dimensional CNN could capture through convolutional operations. Hence, this study employed a one-dimensional CNN to extract features from preprocessed MS data. To enhance the model's ability to represent features and focus on important ones, an SE attention mechanism was incorporated after each convolutional layer. By modeling and weighting the feature maps of each channel, specific feature information could be dynamically learned and emphasized, thereby enhancing the model's performance in feature extraction and classification tasks. The ranges for the penalty parameter *c* and the kernel function parameter *g* in the SVM algorithm were then specified. The parameter *g* regulated the extent of influence of the kernel function, while smaller values corresponded to a broader influence range and larger values indicated a more localized influence range. On the other hand, *c* determined the penalty level for misclassified samples, while higher values resulted in a stricter penalty for misclassification. A particle swarm in the PSO optimization algorithm was initialized by setting the population size and the number of iterations for particles and by initializing parameters such as particle position, velocity, and inertia weight. SVM is a popular choice for binary classification tasks and its classification performance can be improved in multiclassification scenarios by incorporating a kernel function. This function transforms a low-dimensional feature space into a higher-dimensional one, allowing for better separation of previously inseparable classes. In this research, the Radial Basis Function (RBF) was utilized as a selected kernel function, represented by Equation (2).

$$K(x, x_i) = \exp(-\gamma \left\| x - x_i \right\|^2) \quad (2)$$

where x, and x_i were feature vectors. γ was the kernel coefficient that was the parameter controlling the decay rate of the function. The RBF is a nonlinear kernel widely used to map samples to an infinite-dimensional feature space, improving the classifier's performance. The RBF kernel function was integrated into the SVM model, and the classification performance was fine-tuned by adjusting the associated parameters. By leveraging feature transformation and parameter optimization on the training data, the capabilities of SVM in addressing multiclassification problems were effectively extended, enhancing classification performance. The fitness of each particle *max f(c,g)* was calculated as follows.

$$\max f(c,g) = \frac{TP}{TP+FP} \quad s.t. \begin{cases} c \in [c_{\min}, c_{\max}] \\ g \in [g_{\min}, g_{\max}] \end{cases} \tag{3}$$

where c was the penalty factor. g was the kernel function parameter. *TP* was the number of positive samples. *FP* was the number of negative samples identified as positive. The individual best position (P_{best}) and global best position (G_{best}) were than updated. The particle velocity and position were adjusted to maximize the objective function *f(c,g)*. The fitness of each particle with its individual best value ($P_{best(i)}$) and global best value ($G_{best(i)}$) were compared. If the fitness was superior (*max f(c,g)* > $P_{best(i)}$), $P_{best(i)}$ was updated with *max f(c, g)*. If the fitness was the best (*max f(c,g)* > $G_{best(i)}$), $G_{best(i)}$ was updated with *max f(c,g)*. The velocity and position of particles were updated based on Equations (1) and (2). If the final termination condition was satisfied, the process would proceed to commence optimization, otherwise return to calculate the fitness of each particle *max f(c,g)*. After optimization, the optimal parameter sets c and g of the SVM algorithm were obtained and used for classification. The results of classification were output when the conditions were met. This research combined four algorithms. CNN combined with SENet was utilized for feature extraction from original MS data with SVM serving as the classifier for detecting MS types. Presetting parameters empirically might lead to inappropriate parameter settings, affecting the model's performance. Therefore, the model employed the PSO algorithm to optimize SVM parameters.

Experimental equipment

The experiments were conducted using the Keras framework for deep learning in a 12th Gen Intel® Core™ i5-12490F 3.00 GHz CPU, an NVIDIA GeForce RTX 3060 Ti GPU, and the Windows 10 operating system. PyCharm 2021 (https://www.jetbrains.com/zh-cn/pycharm/) was utilized as the development environment.

Parameter setting of CNN network structure

A baseline model called CNN-3L (a convolutional neural network with three convolutional layers) was established. The optimization process primarily focused on tuning hyperparameters such as the choice of optimizer, learning rate, batch size, and number of iterations. The Adam optimizer was selected with a learning rate of 0.001 and a batch size of 50. After establishing the baseline model, the depth of the model was progressively increased by adding more convolutional layers while keeping the hyperparameter settings consistent to enhance the model's feature extraction capability. A convolutional neural network model with 7 layers including a convolutional layer, a standard layer, a global maximum pooling layer, and a fully connected layer were chosen (Table 1). The model gradually increased the number of input channels for each convolutional layer, starting from 16 and progressing to 32, 64, and finally 1,024. The convolutional kernels used had a size of either 5×1 or 7×1 with a step size of 3 and padding set to "same". After each convolutional operation,

Table 1. CNN network structure designed for small molecule compound mass spectrometry datasets.

Layer (type)	Output shape	Param
conv1d (Conv1D)	(None, 100, 16)	96
batch_normalization (Batch Normalization)	(None, 100, 16)	64
conv1d_1 (Conv1D)	(None, 34, 32)	2,592
batch_normalization_1 (Batch Normalization)	(None, 34, 32)	128
conv1d_2 (Conv1D)	(None, 12, 64)	10,304
batch_normalization_2 (Batch Normalization)	(None, 12, 64)	256
conv1d_3 (Conv1D)	(None, 4, 128)	41,088
batch_normalization_3 (Batch Normalization)	(None, 4, 128)	512
conv1d_4 (Conv1D)	(None, 2, 256)	164,096
batch_normalization_4 (Batch Normalization)	(None, 2, 256)	1,024
conv1d_5 (Conv1D)	(None, 1, 512)	918,016
batch_normalization_5 (Batch Normalization)	(None, 1, 512)	2,048
conv1d_6 (Conv1D)	(None, 1, 1,024)	3,671,040
batch_normalization_6 (Batch Normalization)	(None, 1, 1,024)	4,096
global_average_pooling1d (Global Average Pooling 1D)	(None, 1,024)	0
dense (Dense)	(None, 1,000)	1,026,025

feature standardization was applied to normalize the data input, ensuring stability in the model. In the penultimate layer, global average pooling was employed to compress the features of each sample, enhancing the generalization capabilities of the model. Finally, the classification results were obtained through a fully connected layer containing 1,000 neurons. The model took an input with a data shape of 300×1. In the first layer, 16 convolutional kernels of 5×1 were utilized followed by normalization, resulting in an output shape of 100×16. The second layer involved 32 convolutional kernels of 5×1 and normalization, yielding an output shape of 34×32. The third layer employed 64 convolution kernels of 5×1 and normalization, resulting in an output shape of 12×64. In the fourth layer, 128 convolution kernels of 5×1 and normalization were used, generating an output shape of 4×128. The fifth layer comprised 256 convolution kernels of 5×1 and normalization, producing an output shape of 2×256. The sixth layer employed 512 convolution kernels of 7×1 and normalization, resulting in an output shape of 1×512. The seventh layer utilized 1,024 convolution kernels of 7×1 and normalization, leading to an output shape of 1×1,024. The final output was obtained through global average pooling and a fully connected layer.

Network structure of CNNSENet

The SENet attention mechanism was incorporated in CNN model, which helped to enhance the effectiveness of CNN feature extraction. The SENet attention mechanism was seamlessly integrated after the batch normalization layer in each layer of the 7-layer convolutional CNN model. This integration optimized feature extraction across the entire network (Figure 3).

Parameters setting of PSO-SVM

Before using the features extracted by CNNSENet, the critical parameters of the SVM with the radial basis function (RBF) kernel underwent optimization using the PSO. It was experimentally determined that a search range of [1, 100] for the penalty factor (c) and [0.001, 10] for the parameter (g) were more effective in optimizing the parameters. The PSO optimization algorithm selected a population size of 20 and a maximum number of iterations of 80. The termination condition was set as

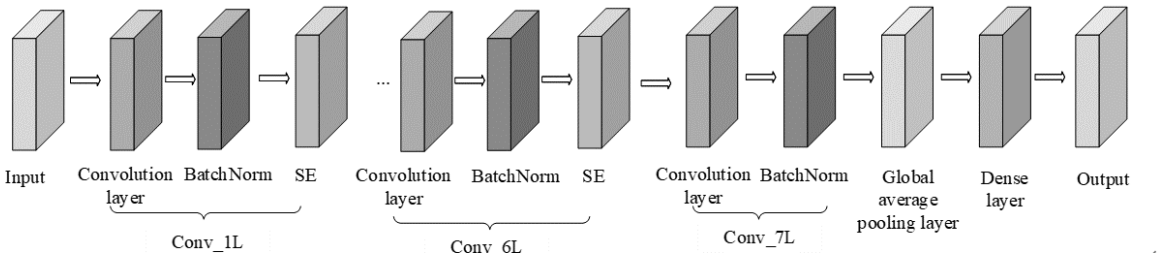

Figure 3. Structure of CNNSENet model.

either achieving a 95% accuracy or completing 80 iterations.

Results and discussion

Data preprocessing results

The intensity values of the raw mass spectrometry data were normalized according to Equation (1), which eliminated the effect of intensity differences on the data analysis results, making the data more comparable and facilitating subsequent feature extraction. After data enhancement, the data were like the normalized data in terms of feature distribution, but with slight differences. The consistency of the data was ensured, which was crucial for training a more stable and reliable model. Although the extended data was very similar to the original data, it introduced new information that helped to expand the feature learning capability of the model. As a result, the generalization performance of the model improved when confronted with new samples.

Feature extraction results for different number of convolutional layers

As the number of layers of MS data recognition model gradually increased from 3 to 7, the convolutional neural network's ability to extract features also improved gradually, resulting in a progressive increase in recognition accuracy. However, once the number of layers exceeded 7, there was no notable enhancement in the model's feature extraction capability, despite the significant increase in parameter count. Furthermore, the model's feature extraction ability leveled off when the number of layers reached 8 and 9. Beyond 9 layers, the model experienced convergence problems due to vanishing or exploding gradients (Figure 4).

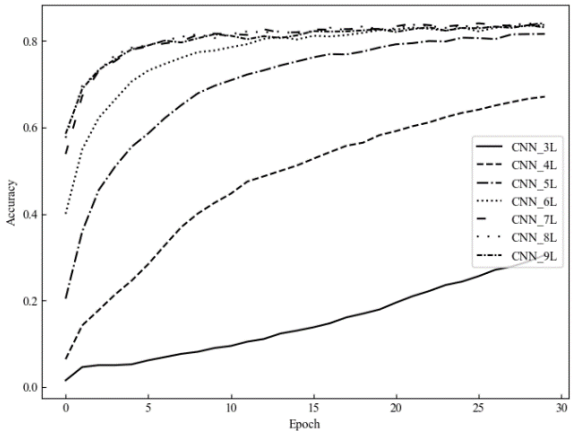

Figure 4. Effect of different number of convolutional layers on feature extraction performance.

CNNSENet-PSO-SVM classification results

Traditional methods typically used Softmax functions to convert the output of the fully connected layer into probabilities. The highest probability out of 1,000 categories was then selected to determine the classification result. Another approach was to make full use of the advanced feature extraction capabilities of CNNs, bypassing the Softmax classifier, and feeding the features extracted by CNNs directly into the fully connected layer, which was then fed into the SVM for classification. The effect of integrating the SENet module into the CNN model compared to the CNN model without SENet was shown in Figure 5. The results showed that integrating the

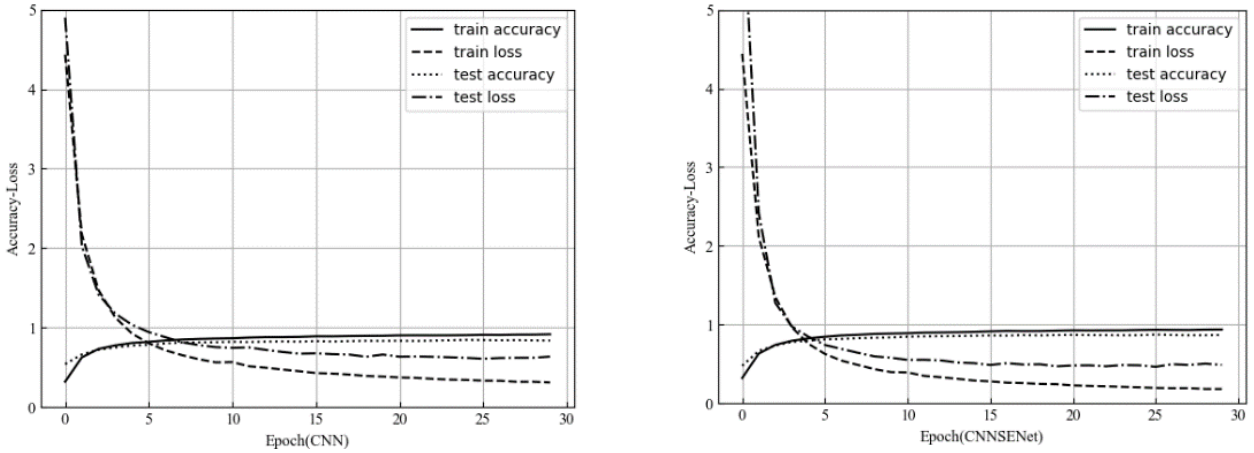

Figure 5. Accuracy and loss function curves for CNN and CNNSENet.

SENet module into the CNN reduced the loss values in both the test and training sets, resulting in faster convergence and a slight increase in classification accuracy. This integration enhanced the generalization ability of the model, made the training process more stable, improved classification accuracy, and reduced the risk of overfitting. Following the PSO optimization process, the final penalty factor for the SVM was found to be 496.146472666083, and the kernel function parameter was determined as 0.0009671318228085051. The features extracted by CNNSENet were then fed into the SVM using the optimized parameter set obtained through PSO, resulting in the ultimate classification (Figure 6). The results indicated that, after optimizing the SVM, by the time the iteration had progressed to the 70th round, the optimal parameters had been found, and the final classification accuracy reached 90.55%. To further illustrate the effectiveness of the proposed multiple classification model for MS data, comparisons were conducted among SVM, CNN, CNNSENet, CNNSENet-SVM, and CNNSENet-PSO-SVM methods. When SVM and CNN were used individually for classifying MS data, the accuracy on the test dataset was 85.45% and 85.82%, respectively, which indicated that both SVM and CNN performed reasonably well in classifying MS data, although there was still room for improvement in accuracy. After incorporating the SENet attention mechanism into CNN for classification of MS data, the percentage of correctly identified samples in the test dataset reached 87.27%, which represented a slight improvement compared to SVM and CNN alone, suggesting that SENet attention enhanced CNN's feature extraction capability, thereby improving classification accuracy. Using the CNNSENet model for feature extraction followed by SVM for classification, the percentage of correctly identified samples in the test dataset was 88.56%, which indicated that the SVM classifier performed better when utilizing features extracted by CNN. However, after optimizing the parameters using the PSO particle swarm algorithm, the proposed model CNNSENet-PSO-SVM achieved the highest classification accuracy on the MS dataset as 90.55% of the spectra classes correctly identified in the test dataset. The result represented a significant improvement over other models, demonstrating a notable increase in classification accuracy.

Mass spectrometry is a vital tool for compound identification. However, manual analysis can be labor-intensive and impractical for the rapid and high-volume identification required in mass spectrometry. This study employed CNNSENet and PSO-SVM for multi-classification of mass spectrometry data. The results demonstrated that the CNNSENet neural network could leverage the intrinsic data features without

extensive prior knowledge, autonomously extracting features effectively. The SVM classifier was incorporated post the convolutional neural network's fully connected layer with the PSO algorithm optimizing key hyperparameters to enhance SVM classification accuracy. The validated classifier successfully categorized samples into 1,000 distinct classes with notably superior accuracy compared to conventional models. This proposed model effectively addressed the challenge of multi-classification in mass spectrometry, facilitating compound analysis, and offering valuable support for research and applications in chemistry, biomedicine, and environmental science.

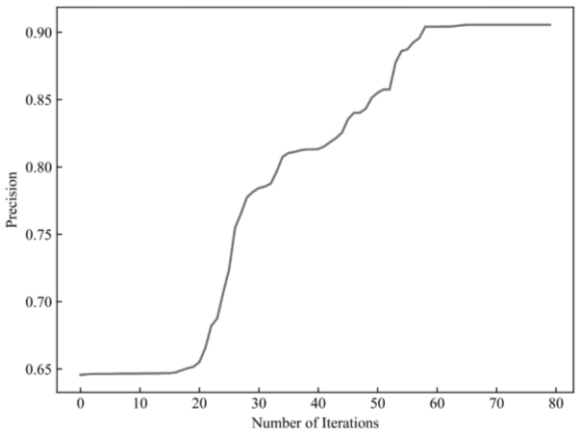

Figure 6. Optimization curves for PSO-SVM.

Acknowledgements

This research was funded by National Natural Science Foundation of China (Grant No. 82160787). Research on Quality Standardization of Chemical Substance Basis of Guizhou Jinshi Slant and Tiepishi (Grant No. 20025800400). The Yunnan Revitalization Talent Support Program “Young Talent” Project (Grant No. YNWR-QNBJ-2020-247).

References

1. De Hoffmann E, Stroobant V: Tandem Mass Spectrometry. Mass spectrometry: principles and applications. 3rd edition. Hoboken: John Wiley & Sons. 2007:189-211.
2. Ahuja V, Singh A, Paul D, Dasgupta D, Urajová P, Sounak Singh SRG, *et al*. 2023. Recent advances in the detection of food toxins using mass spectrometry. Chem Res Toxicol. 36(12):1834-1863.
3. Sun J, Yin Y, Li W, Jin OY, Na N. 2022. Chemical reaction monitoring by ambient mass spectrometry. Mass Spectrom Rev. 41(1):70-99.
4. Meissner F, Geddes-McAlister J, Mann M, Bantscheff M. 2022. The emerging role of mass spectrometry-based proteomics in drug discovery. Nat Rev Drug Discov. 21(9):637-654.
5. Yu JS, Nothias LF, Wang M, Kim DH, Dorrestein PC, Kanget KB, *et al*. 2022. Tandem mass spectrometry molecular networking as a powerful and efficient tool for drug metabolism studies. Anal Chem. 94(2):1456-1464.
6. Neagu AN, Jayathirtha M, Baxyer E, Donnelly M, Alina PB, Costel CD. 2022. Applications of tandem mass spectrometry (MS/MS) in protein analysis for biomedical research. Molecules. 27(8):2411.
7. Messner CB, Demichev V, Wang Z, Hartl J, Kustatscher G, Mülleder M, *et al*. 2023. Mass spectrometry-based high-throughput proteomics and its role in biomedical studies and systems biology. Proteomics. 23(7-8):2200013.
8. Wang S, Zhu H, Zhou H, Cheng JQ, Yang H. 2020. MSpectraAI: a powerful platform for deciphering proteome profiling of multi-tumor mass spectrometry data by using deep neural networks. BMC Bioinformatics. 21:1-15.
9. Seddiki K, Saudemont P, Precioso F, Ogrinc N, Wisztorski M, Salzet M, *et al*. 2020. Cumulative learning enables convolutional neural network representations for small mass spectrometry data classification. Nat Commun. 11(1):5595.
10. Li Y, Gan Z, Zhou X, Chen ZW. 2022. Accurate classification of Listeria species by MALDI-TOF mass spectrometry incorporating denoising autoencoder and machine learning. J Microbiol Meth. 192:106378.
11. Pfeuffer J, Sachsenberg T, Alka O, Walzer M, Fillbrunn A, Nilse L, *et al*. 2017. OpenMS–a platform for reproducible analysis of mass spectrometry data. J Biotechnol. 261:142-148.
12. Ma B, Zhang K, Hendrie C, Liang CZ, Li M, Amanda DK, *et al*. 2003. PEAKS: powerful software for peptide *de novo* sequencing by tandem mass spectrometry. Rapid Commun Mass SP. 17(20):2337-2342.
13. Wu J, Ji Y, Zhao L, Ji MY, Ye Z, Li SY. 2016. A mass spectrometric analysis method based on PPCA and SVM for early detection of ovarian cancer. Comput Math Method M. 2016:6169249.
14. Lee J, Shin Y, Kim S, Rho K, Park KH. 2017. SVM classification model of similar bacteria species using negative marker: based on matrix-assisted laser desorption/ionization time-of-flight mass spectrometry [C]. 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE). IEEE. 2017:145-150.
15. Seddiki K, Saudemont P, Precioso F, Ogrinc N, Wisztorski M, Salzet M, *et al*. 2020. Cumulative learning enables convolutional neural network representations for small mass spectrometry data classification. Nat Commun. 11(1):5595.

16. Petrovsky DV, Kopylov AT, Rudnev VR, Stepanov A, Kulikova L, Malsagova KA, *et al*. 2021. Managing of unassigned mass spectrometric data by neural network for cancer phenotypes classification. J Pers Med. 11(12):1288.
17. Lv J, Wei J, Wang Z, Cao J. 2019. Multiple compounds recognition from the tandem mass spectral data using convolutional neural network. Molecules. 24(24):4590.
18. Alzubaidi L, Zhang J, Humaidi AJ, Al DA, Duan Y, Al SO, *et al*. 2021. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data. 8:1-74.
19. Niu Z, Zhong G, Yu H. 2021. A review on the attention mechanism of deep learning. Neurocomputing. 452:48-62.
20. Wang W, Wang T. 2024. Acoustic emission signal identification of different rocks based on SE-1DCNN-BLSTM network model. Global Geology. 27(1):43-55.
21. Marini F, Walczak B. 2015. Particle swarm optimization (PSO). A tutorial. Chemometr Intell Lab. 149:153-165.
22. Chanhan VK, Dahiya K, Sharma A. 2019. Problem formulations and solvers in linear SVM: a review. Artif Intell Rev. 52(2):803-855.
23. Patro S, Sahu KK. 2015. Normalization: A preprocessing stage. arXiv preprint arXiv:1503.06462.