

*Scientific Paper*Doi: <http://dx.doi.org/10.1590/1809-4430-Eng.Agric.v43n6e20230127/2023>**INTELLIGENT IDENTIFICATION OF RICE GROWTH PERIOD (GP) BASED ON RAMAN SPECTROSCOPY AND IMPROVED CNN IN HEILONGJIANG PROVINCE OF CHINA****Rui Liu¹, Feng Tan^{2*}, Bo Ma³**

^{2*}Corresponding author. College of Electrical and Information, Heilongjiang Bayi Agricultural University, Daqing 163000, China.
E-mail: tf1972@163.com | ORCID ID: <https://orcid.org/0000-0002-0724-7132>

KEYWORDS

rice growth period,
Raman spectroscopy,
low temperature,
chilling injury,
CNN-CBAM.

ABSTRACT

The fertile land in Heilongjiang Province of China is suitable for rice cultivation, but this area is susceptible to low temperature and chilling injury, which is prevented by planting rice varieties suitable for GP that is an important measure. However, selection based on rice traits is vulnerable to environmental influences and takes a long time, and selection based on molecular markers may result in progeny recombination and lack of reliability. Therefore, an efficient accurate and intelligent identification method for rice growth period is urgently needed. In this study, machine learning and deep learning methods in Python were used to analyze the Raman spectra of 6 rice varieties in three accumulated temperature region of Heilongjiang Province. 1) In machine learning, Principal Component Analysis (PCA) was adopted for feature extraction, in combination with Support Vector Machine (SVM) classification models suitable for nonlinear data sets for identification, the identification rate was 93.33% and the type of this experimental data set was determined to be discrete. 2) In deep learning, Continuous Wavelet Transform (CWT) methods was adopted for data preprocessing, combined with the Convolutional Neural Networks (CNN) model with its own feature extraction, with the highest accuracy of 94.82%, which was higher than the PCA+SVM identification model. 3) Based on the method mentioned in 2), in order to improve the feature extraction ability of the model as a whole, Convolutional Block Attention Module (CBAM) was used to improve the CNN identification model for the first time for one-dimensional data sets, and the highest identification rate was 98.28%, which was better than the PCA+SVM identification model. 3) In the verification test, Raman spectral information of 4 rice varieties was brought into the constructed CWT+CNN-CBAM identification model for identification, and the identification results were as high as 94.79%. The experimental results showed that the CWT visualization data processing method based on Raman technology combined with the CNN identification model of CBAM with improved feature extraction ability in deep learning achieved the best identification results, which could provide an efficient, accurate and intelligent method for the identification of different growth period of rice varieties in Heilongjiang Province.

INTRODUCTION

The black soil cultivated land area of Heilongjiang Province accounts for 50.6% of the total black soil area of China. The soil quality is fertile, the surface layer is deep, the cultivated land is flat, and the organic matter content ranks first in China, which is the natural advantage of

high-quality rice production (Wang et al., 2022; Xie, 2022). Rice is sensitive to temperature and light, and it often encounters low temperature from sowing to ripening, which is easy to form chilling injury, resulting in varying degrees of loss to rice yield (Sun et al., 2020; Liu et al., 2022; Cai & Zhang, 2018). Heilongjiang Province is the northernmost province of China, bordering cold Siberia in the north,

¹ College of Agricultural Engineering, Heilongjiang Bayi Agricultural University, Daqing 163000, China.

³ Qiqihar Branch of Heilongjiang Academy of Agricultural Sciences, Qiqihar 161006, China.

Area Editor: Teresa Cristina Tarlé Pissarra

Received in: 9-13-2023

Accepted in: 11-4-2023



straddling the middle and cold temperate zones in the north and south. According to the climatic characteristics of Heilongjiang Province and the accumulation temperature required for rice growth, GP of the rice could be divided into three categories. Therefore, selecting rice varieties suitable for GP in Heilongjiang Province is an important measure to prevent low temperature and chilling injury.

Through the analysis of phenotypic data of rice varieties in Heilongjiang province, the researchers selected rice varieties suitable for accumulated temperature regions according to their trait indicators (Sun et al., 2019; Xue, 2016; Lv, 2017). The process of phenotypic data collection requires a large number of long-term field investigations, depends on the long-term practical experience of researchers, and is subject to time and human interference. Other researchers made molecular markers for different GP of the rice varieties based on the identification results of rice phenotypic traits, and finally determined the molecules that affect the GP of rice (Tian et al., 2019; Liang et al., 2021; Wang et al., 2023). Molecular marking method is not only expensive, but also has the disadvantages of progeny recombination, cost of material and lack of reliability. With the rapid development of science and technology, it is of great significance to design and construct an efficient, accurate and intelligent identification method for GP of the rice to prevent low temperature and chilling injury.

Raman spectroscopy was first applied to physics and chemistry research by Hibben J H and Teller E. It has the advantages of fast, efficient and accurate measurement, and can realize fast, accurate and nondestructive analysis of samples (Hibben & Teller, 1939; Herzberg, 1945), providing a new idea for realizing fast and real-time nondestructive sensing. Ling et al. (2018) collected a large number of rice Raman spectra to conduct high-precision detection of adulterated rice, laying a foundation for an advanced rice quality identification technology system. Tian et al. (2020) established a rapid non-destructive detection method for distinguishing rice-producing areas using Raman spectroscopy. Min et al. (2020) analyzed 72 Raman spectra of 3 rice varieties and found that PCA, window analysis and HCA combined with SVM could be used as an effective feature extraction method to improve the efficiency of rice variety identification. Zhu (2021) took 33 kinds of japonica seeds mainly planted in Heilongjiang Province as the research object, collected spectral information of seeds by Raman spectroscopy technology, extracted CARS feature based on AIRPLS+1-Der pretreatment method, and finally combined with cosine similarity algorithm discrimination model, the similarity discrimination effect of 33 kinds of rice seeds was better than others. At present, the application of Raman

spectroscopy in the identification of rice varieties is widespread, but there are few reports on the identification of rice characteristics.

With the rapid development of detection technology and artificial intelligence, smart agriculture faces new opportunities (Li, 2022). Python is a computer programming language with strong operability, easy-to-use, and full-featured tools, which is widely used in data analysis (Vankeirsbilck et al., 2002). Among them, SVM is a typical nonlinear calibration method in machine learning (Giang et al., 2020; Ma et al., 2022), and CNN is a representative method in deep learning, therefore, SVM and CNN were selected in this paper for differential analysis of Raman spectral data. However, when SVM and CNN were used to discriminate spectral data, due to the interference of machine noise, stray light and fluorescent background, the accuracy of the discrimination analysis was affected, some researchers used the visualization tools Matplotlib and Seaborn (Pezzotti et al., 2021) in Python library to analyze the data and in order to further improve the accuracy of identification.

In recent years, Raman spectroscopy, which is famous for its fast, efficient, accurate and nondestructive testing technology, has been widely used in agricultural areas. In view of this, this study used Raman spectrometer to obtain rice variety information, combined with Python programming software, established an efficient, accurate and intelligent identification model of the rice GP, and selected rice varieties suitable for GP in Heilongjiang Province, which is an important measure to prevent chilling injury from low temperature.

MATERIAL AND METHODS

Test Material

The experimental materials were collected from the rice experimental field of Qiqihar Branch of Heilongjiang Academy of Agricultural Sciences on September 20, 2021. There were 6 rice varieties in total, and there were 3 main accumulated temperature region in Heilongjiang Province, with 2 rice varieties in each accumulated temperature region. After 5 holes of each rice variety were placed in the laboratory at 23°C to air naturally for 10 days, 10 ears were taken from each hole and 10 grains were taken from different positions of each ear, a total of 500 grains were taken from each variety. Shanghai Chaoxing LJJM Rice Mopping machine was used to conduct a one-time 50-second dehulling treatment on the seeds of each variety, and then 48 grains with complete appearance were selected as samples for each rice variety, a total of 288 grains were obtained for 6 varieties, as shown in Table 1.

TABLE 1. Details of sample acquisition and preparation. (ATZH:accumulated temperate region in Heilongjiang Province.)

Material	ATZH	Number of holes	Ear number	Number of grains	Number of samples
LJ31	First	5	10	10	48 (5*10*10)
LJ46	First	5	10	10	48 (5*10*10)
LD18	Second	5	10	10	48 (5*10*10)
QJ10	Second	5	10	10	48 (5*10*10)
QJ15	Third	5	10	10	48 (5*10*10)
SJ9	Third	5	10	10	48 (5*10*10)
Total		30	300	3000	288 (3000)

Test Instrument

The sample information was collected by Raman spectrometer (OPTOSKY, China) equipped with a 785 nm laser, which was an ideal light source for resonance Raman research. The measurement range was 200–3400 cm^{-1} under optimal measurement conditions, and the

displacement deviation of the measured standard peak is zero, which conforms to the requirement that the displacement accuracy does not exceed $\pm 4 \text{ cm}^{-1}$. The data processing software was Python. The full band information of the original Raman spectra of 288 samples is shown in Figure 1.

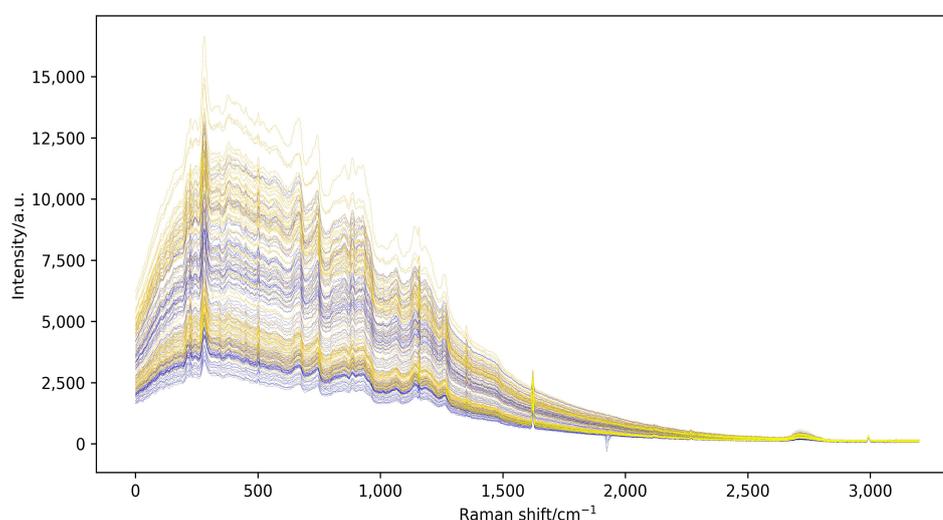


FIGURE 1. Full band information of 288 original Raman spectra. Raman shift is the reciprocal of wavelength, and its range is 200–3400 cm^{-1} . Intensity is the intensity of Raman scattering.

Test Methods

Feature Extraction Method

The original Raman spectra (200–3400 cm^{-1}) is not only data-intensive, but also contains a large amount of redundant collinear information, which affects the accuracy and computational speed of the model, making it necessary to reduce the interference of the original Raman spectrum data (Tian et al., 2020; Zhu, 2021). PCA, also known as principal component analysis, aims to convert multiple indicators into a few comprehensive indicators (i.e. principal components) by using the idea of dimensionality reduction, in which each principal component can reflect most of the information of the original variable without repeating the information. PCA method is a mathematical transformation

method, which converts a given set of related variables into another set of unrelated variables through linear transformation, and these new variables are arranged in the order of decreasing variance. In the mathematical transformation, keep the total variance of variables unchanged, so that the first variable has the largest variance, called the first principal component, the second variable has the second largest variance, and is unrelated to the first variable, called the second principal component. By analogy, I variables have I principal components. In this study, PCA was run through Python 3.8.5 to extract features from the full spectrum of 200–3400 cm^{-1} . Type numbers of rice spectral data samples (0,1,2) were entered, the number of feature variables was set within the range of 3–5, and 4 main features are obtained for subsequent modeling analysis (Figure2).

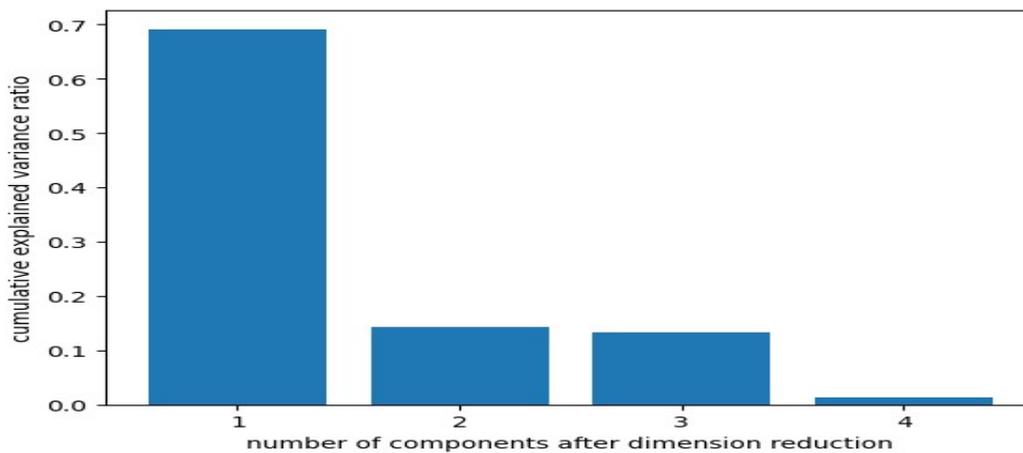


FIGURE 2. Four main components of Raman spectral information of rice GP were extracted by PCA characteristic method.

Preprocessing Method

The data of Raman spectrum are easily disturbed by internal causes of changes in the intensity of instrument excitation light and external causes of excitation of fluorescent material in the sample to be tested in the fluorescent background, which affect the accuracy of the data, it is therefore necessary to preprocess the original Raman spectral data (Tian et al., 2020). CWT is a new signal processing tool developed in recent years, which can translate forward and backward along the time axis, as well as stretch and compress proportionally to get low and high frequency wavelet. The constructed wavelet function can be used to filter or compress signals, so as to extract useful signals without interference. The implementation of the CWT function is shown in Formula (1).

$$Wf(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x) \psi^* \left(\frac{x-b}{a} \right) dx \quad (1)$$

Where:

- a (a>0) is the scale parameter used by the variable to control expansion;
- b(b∈R) is the position translation parameter, and
- x is the Raman shift of the input function.

Scipy.signal.cwt in Python uses a wavelet function to convolve the data, and its characteristic width parameter N (ndarray) and length parameter M (sequence), N returns the number of points that the vector will have, and M defines the size of the wavelet for the transform width, as shown in Figure 3.

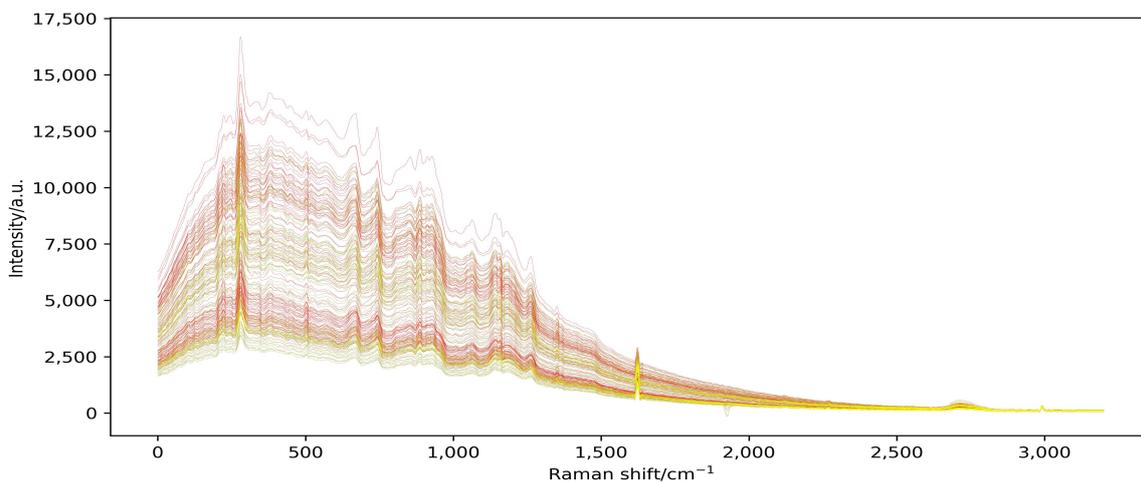


FIGURE 3. Raman spectral information was processed by CWT method.

Identification Methods

In this study, SVM identification model was used to identify the GP of 6 rice varieties. Meanwhile, CNN identification model was used to establish the identification model of rice GP based on the CWT data pre-processing of full spectrum (200-3400cm⁻¹).

When the data set is linearly indivisible, SVM uses the nonlinear mapping algorithm to transform the linearly

indivisible samples in the low-dimensional input space into the high-dimensional feature space to make them linearly separable, thus making it possible to use the linear algorithm in the high-dimensional feature space to perform linear analysis on the nonlinear features of the sample (Vladimir & Michiel, 2009). The solution is to map them to a higher dimensional space, but the difficulty with this approach is the increase in computational complexity, and the kernel neatly solves this problem. Therefore, kernel

function is the key to SVM model. The kernel list parameter of the SVM model is radial basis function kernel (rbf), which is applicable to linear non-fractional data sets, as shown in Formula (2) and parameter setting (3):

$$RBF : \exp\left(-\frac{1}{2\delta^2}\|X - X_i\|^2\right) \quad (2)$$

$$\text{kernel} = 'rbf', \text{random_state} = 1, \text{max_iter} = -1, \text{tol} = 1e - 4 \quad (3)$$

RBF can map samples to higher-dimensional Spaces and process samples when the relationship between class labels and features is nonlinear. In this study, RBF was used as the kernel function of SVM, and PCA+SVM and SWT+SVM models were established on the basis of PCA feature extraction and SWT preprocessing of the original data, respectively.

CNN is a kind of Feedforward Neural Networks with deep structure including convolution computation, and it is one of the representative algorithms of deep learning. Its structure is mainly divided into input layer, hidden layer and output layer. The input layer can process multidimensional data. The input layer of this experiment was two-dimensional data. Since the input features of CNN need to be standardized, the input data is normalized and brought into the input layer of CNN. Hidden layers include convolutional layer, pooling layer and full connection layer.

Upstream of the output layer is the fully connected layer, which directly outputs the classification results. In this experiment, there were 5 convolutional layers and 3 full connection layers, namely, three types of identification were carried out on the data, and combined with the CWT pretreatment method, the rapid identification of rice GP in Heilongjiang Province, China based on Raman spectroscopy was carried out.

Improved Method of Identification Model

In neural network learning, generally speaking, the more parameters of the model, the stronger the expressive ability of the model and the more information stored by the model, but this will bring the problem of information overload. Then the problem of information overload can be solved by introducing the Attention Mechanism. The Attention Mechanism can be divided into Channel Attention Mechanism, Spatial Attention Mechanism and the combination of the two (CBAM). Figure 4 shown that CBAM contains two independent sub-modules, Channel Attention Module (CAM) and Spatial Attention Module (SAM), which perform channel and spatial attention respectively. This not only saves parameters and computing power, but also ensures that it can be integrated into the existing network architecture as a plug and play module (arXiv et al., 2018). In this study, a lightweight structure CBAM based on attention mechanism was proposed, which was added to the convolutional layer of CNN to perform intelligent identification of rice GP in Heilongjiang Province, China based on Raman spectroscopy.

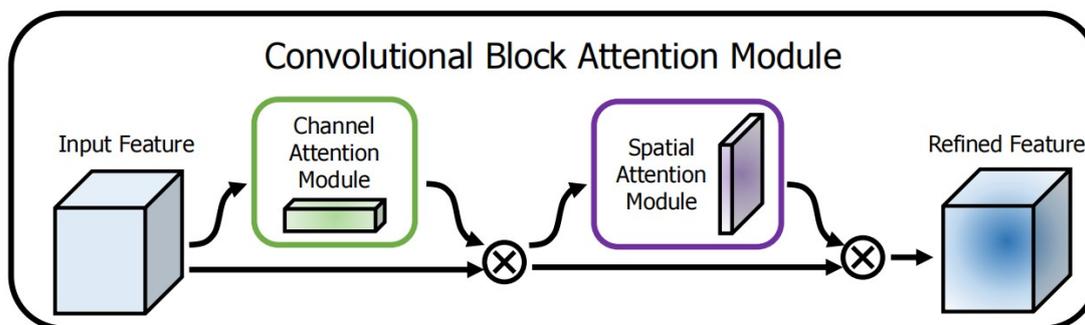


FIGURE 4. Schematic diagram of CBAM implementation.

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn.

RESULTS AND DISCUSSION

Identification Models for Machine Learning Methods

The Figure 1 shown that the spectral information of 288 samples was cluttered and interleaved and difficult to distinguish. Therefore, it is very necessary to reduce dimensions and preprocess the original spectral data. Four principal component variables were obtained by PCA dimension reduction method (Figure 2), among which the

contribution rate of the first three principal components reached more than 98%. The original Raman spectral information was filtered through the CWT data preprocessing method (Figure 3), the filtered Raman spectral information was smoother than the original information, and the interference caused by instrument noise and stray light during the collection of spectral data information was removed (Figure5). After the reduction and processing of 288 original Raman spectrum samples, the training set and test set were randomly divided according to 8:2, and the SVM identification model was introduced. The identification results of PCA+SVM and CWT+SVM classification models were 55% and 95%, respectively.

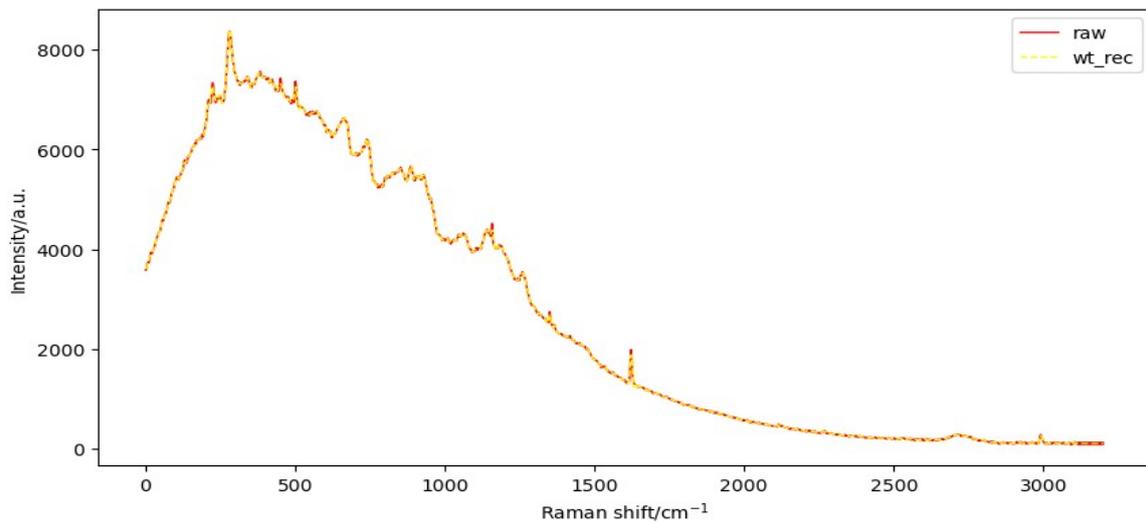


FIGURE 5. Comparison of the Raman spectral information processed by CWT method with the original Raman spectral information.

In the identification method of the machine learning method based on Raman spectrum in Heilongjiang Province of China, the identification result of PCA method with feature extraction and nonlinear SVM classification model was 95%, much higher than that of data processing CWT method with SVM classification model (55%). In the feature extraction PCA method, 3,201 variables of Raman spectrum were introduced and complex factors were reduced to 4 principal components (Figure 2). The 3 principal components with the highest contribution rate were selected and 288 sample data were visualized as three-dimensional maps (Figure 6). It was known that the spectral data set of this test was a discrete and linear non-fractional data set. It was put into the SVM

discrimination model which used the nonlinear mapping algorithm to transform the linear indivisible samples in the low-dimensional input space into the high-dimensional feature space, making it linearly separable and achieving 95% identification results. Although the CWT method could effectively visualize the full-band data of Raman spectra (Figure 3 and Figure 5), filter and compress the Raman spectral data, and extract the useful Raman spectral information containing interference information, the identification effect was not good when combined with the SVM classification model in machine learning. The three-dimensional maps of 288 samples were constructed by extracting the first three principal component characteristics of contribution rate through PCA (Figure 6).

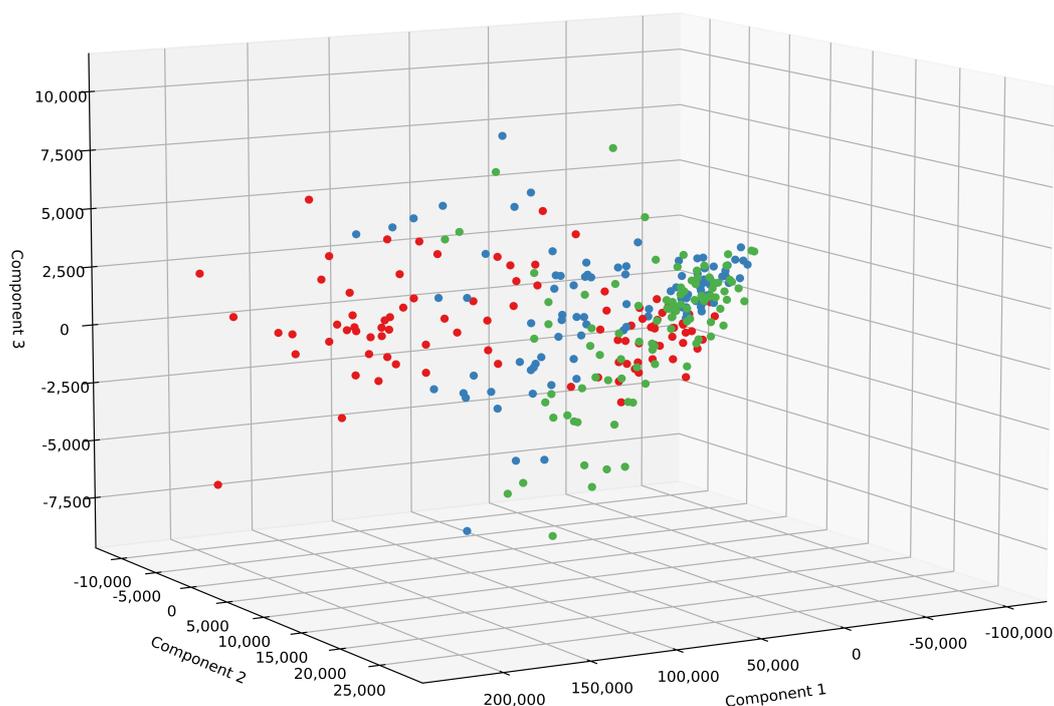


FIGURE 6. The three-dimensional maps of 288 samples were constructed by extracting the first three principal component characteristics of contribution rate through PCA.

Comparison between Machine Learning Identification Model and Deep Learning Identification Model

For the identification methods of rice growth stage based on Raman spectrum in Heilongjiang Province of China, the identification of CWT method combined with deep learning CNN classification model could reach up to 94.82%, the identification result of CWT+CNN was higher than that of PCA+SVM based on machine learning (The right of the figure 7).

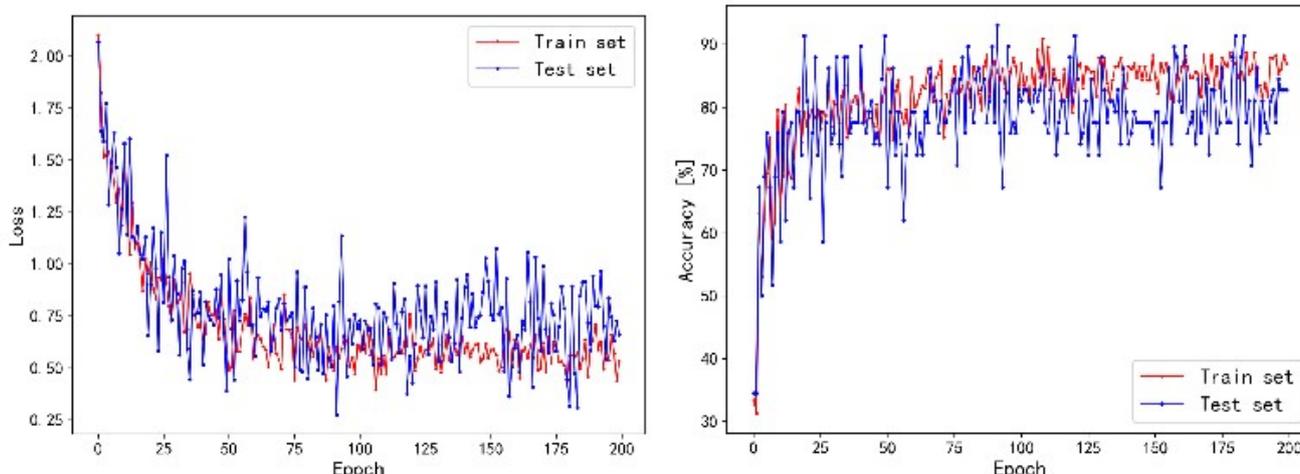


FIGURE 7. Loss value and accuracy value of the CWT+CNN identification model.

The CWT data processing method was translated forward and backward along the axis of Raman shift, and stretched and compressed proportionally to effectively avoid the phase difference of Raman spectral information and obtained useful Raman spectral information after noise removal and disturbance reduction. Each convolution layer in CNN was followed by a computing layer for local average and secondary extraction, and the unique two-feature extraction structure reduced feature resolution (Cui & Tan, 2023), so the feature extraction effect was better than the PCA method in machine learning. In this experiment, although CWT+SVM classification model was far inferior to PCA+SVM classification model, the combination of data processing CWT method and CNN identification model including feature extraction has achieved better identification results.

Comparison between Improved CNN Identification Model and CNN Identification Model

For the identification method of rice GP in Heilongjiang Province of China based on Raman spectroscopy, a lightweight structure CBAM based on Attention Mechanism was proposed, which was added to the convolutional layer of CNN to improve the CNN identification model, and obtained the best accuracy of 98.28%, which was higher than that of SWT+CNN identification model (The right of the figure8). As shown in the light side of Figure 7 and the light side of Figure 8, when epoche ranges from 0 to 200, the loss value of CNN identification model ranges from 2 to 0.5, and the loss value of CNN-CBAM identification model ranges from 1 to 0.2, the loss value of CNN-CBAM identification model is lower than that of CNN identification model.

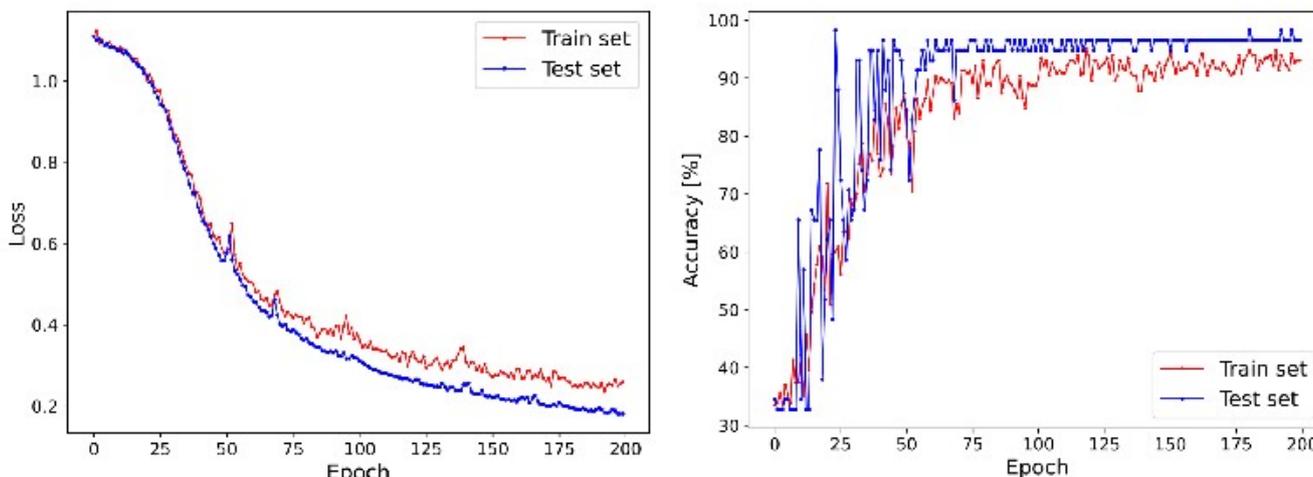


FIGURE 8. Loss value and accuracy value of the CNN-CBAM identification model.

In neural network learning, generally speaking, the more parameters of the model, the stronger the expressive ability of the model and the more information stored by the model, but this will bring the problem of information overload. Then, by introducing an Attention Mechanism to focus on the information that is more critical to the current task, reduce the attention to other information, and even filter out irrelevant information, the problem of information overload can be solved and the efficiency and accuracy of task processing can be improved. CBAM is a combination of Channel Attention Mechanism and Spatial Attention Mechanism, which perform channel and spatial attention respectively. This not only saves parameters and computational power, but also ensures that it can be integrated into existing network architectures as a plug and play module (Cui & Tan, 2023). Therefore, among the identification methods of rice GP based on Raman spectroscopy in Heilongjiang Province, China, the CNN identification model of CBAM with improved feature

extraction ability obtained the best identification results and stability, which could provide an efficient, accurate, and intelligent method for the identification of rice varieties with different GP in Heilongjiang Province.

Verification Test

According to the above results, four rice varieties in Heilongjiang Province were identified for GP. The verification test materials were collected from the rice experimental field of Qiqihar Branch of Heilongjiang Academy of Agricultural Sciences on September 25, 2021. There were 4 rice varieties in total, 3 holes of each rice variety were placed in the laboratory at 23°C for 10 days to air naturally, 5 ears were taken from each hole, 5 grains were taken from different positions of each ear, and a total of 75 grains of each rice variety were shelled. Finally, 48 complete grains of each rice variety were selected as verification test samples, and 192 grains of 4 varieties were obtained, as shown in Table 2.

TABLE 2. Details of the collection and preparation of verify test samples and results. (ATZH: accumulated temperate region in Heilongjiang Province.)

Material	ATZH	Number of holes	Ear number	Number of grains	Number of samples	Results of CNN (first/second/third)	Results of CNN-CBAM (first/second/third)
QJ24	second	3	5	5	48 (3*5*5)	05/32/11(66.67%)	06/42/00(87.50%)
LJ21	third	3	5	5	48 (3*5*5)	00/00/48(100%)	00/00/48(100%)
LJ47	third	3	5	5	48 (3*5*5)	00/01/47(97.92%)	00/04/44(91.67%)
Q2112	third	3	5	5	48 (3*5*5)	00/00/48(100%)	00/00/48(100%)
Total		12	60	300	192 (300)	91.15% (accuracy rate)	94.79% (accuracy rate)

The spectral data information of the verification test samples was obtained according to the method of obtaining sample spectral data information of the above experiment, and the identification models CWT+CNN and CWT+CNN-CBAM were introduced respectively for validation tests. 1) The results of CNN were shown in Table 2. Among 48 Raman spectrum samples from rice variety QJ24 in the second accumulated temperature region, 5 was divided into the first accumulated temperature region, 32 were divided into the second accumulated temperature region and 11 were divided into the third accumulated temperature region, the identification result was 66.67%. Among 48 Raman spectrum samples from rice variety LJ21 in the third accumulated temperature region, 48 were divided into the third accumulated temperature region, the identification result was 100%. Among 48 Raman spectrum samples from rice variety LJ47 in the third accumulated temperature region, 1 was divided into the second accumulated temperature region and 47 into the third accumulated temperature region, the identification result was 97.92%. All samples in 48 Raman spectrum samples from rice variety QJ24 in the third accumulated temperature region were divided into the third accumulated temperature region, and the identification result was 100%. According to the CWT+CNN identification method of rice GP in Heilongjiang Province of China based on Raman

spectroscopy, the final identification result of verification test was 91.15%. 2) The results of CNN-CBAM were shown in Table 2. Among 48 Raman spectrum samples from rice variety QJ24 in the second accumulated temperature region, 6 was divided into the first accumulated temperature region and 42 were divided into the second accumulated temperature region, the identification result was 87.50%. Among 48 Raman spectrum samples from rice variety LJ21 in the third accumulated temperature region, 48 were divided into the third accumulated temperature region, the identification result was 100%. Among 48 Raman spectrum samples from rice variety LJ47 in the third accumulated temperature region, 4 was divided into the second accumulated temperature region, 44 into the third accumulated temperature region, the identification result was 91.67%. All samples in 48 Raman spectrum samples from rice variety QJ24 in the third accumulated temperature region were divided into the third accumulated temperature region, and the identification result was 100%. According to the CWT+CNN-CBAM identification method of rice GP in Heilongjiang Province of China based on Raman spectroscopy, the final identification result of verification test was 94.79%. The verification test shown that the best identification model of rice GP in Heilongjiang Province of China based on Raman spectroscopy, namely CWT+CNN-CBAM, was introduced.

CONCLUSIONS

The Raman spectra of 6 rice varieties were measured in the identification method of rice GP in Heilongjiang Province. In the machine learning method, the identification rate of PCA feature extraction combined with hyperplane SVM classification method (93.33%) was 38% higher than that of CWT data processing combined with SVM method (55%), indicating that feature extraction was more important than data preprocessing, and feature extraction combined with hyperplane classification (a classification model suitable for discrete data sets) could obtain higher identification results (93.33%). In the deep learning method, the CWT+CNN identification model (94.82%) was better than the CWT+SVM identification model of machine learning, because the convolutional layer in the CNN classification method contains feature extraction function combined with data preprocessing and the identification effect of the CWT method was improved. In the improved identification model method, CBAM with improved feature extraction capability was added to the convolutional layer of CNN to improve the CNN structure, and efficient and intelligent identification results (98.28%) were obtained. In the verification test, Raman spectra of the remaining 4 rice varieties were measured and brought into the CWT+CNN-CBAM identification model to obtain 94.79% efficient, accurate and intelligent identification results. Therefore, the CWT visualization data preprocessing method based on Raman technology combined with the CNN identification model of CBAM, which has the ability to enhance feature extraction in deep learning, obtained the best identification results and could provide an efficient, accurate, and intelligent method for the identification of rice varieties at different GP in Heilongjiang Province, China.

Funding: Research Expenses of Provincial Research Institutes of Heilongjiang Province (CZKYF2023-1-B013), Heilongjiang Province Higher Education Teaching Reform Research Project (SJGY20210622), Science and Technology Project of Heilongjiang Academy of Agricultural Sciences (2021YYF011).

REFERENCES

Cai Z, Zhang G (2018) Research progress on low temperature chilling injury of rice. *Crop Research* 32 (03): 249-255.

<https://doi.org/10.16848/j.cnki.issn.1001-5280.218.03.18>

Cui J, Tan F (2023) Rice plaque detection and identification based on an improved convolutional neural network. *Agriculture* 13:1.

Giang LT, Trung PQ, Yen D (2020) Identification of rice varieties speciaties in Vietnam using Raman spectroscopy. *Vietnam Journal of Chemistry* 58:711–718.

Herzberg G (1945) Molecular spectra and molecular structure. In *Infrared and Raman Spectra of Polyatomic Molecules*; Van Nostrand, R., Ed.; American Journal of Physics: New York, NY, USA, Volume 2.

Hibben JH, Teller E (1939) The Raman effect and its chemical applications and physical research. *Industrial and Engineering Chemistry. News Ed.* 17:556.

Li D (2022) Research on countermeasures for the development of smart agriculture in counties of Heilongjiang province - Taking the construction of Zhao Guang digital farm of Bei'an administration bureau as an example. *China Market* 30: 89-91.

<https://doi.org/10.13939/j.cnki.zgsc.202230.089>

Liang H, Pan Y, Liu K (2021) Development of a specific molecular marker for the cold tolerance gene CTB4a in rice based on primer amplification blocked mutation technology. *Genomics and Applied Biology* 40 (04): 1719-1724.

<https://doi.org/10.13417/j.gab.040.001719>

Ling Z, Juan S, Gangcheng W, Yanan W, Hui Z, Li W, Haifeng Q, XiGuang Q (2018) Identification of rice varieties and determination of their geographical origin in China using Raman spectroscopy. *Journal of Cereal Science* 82.

Liu L, Bian J, Sun X, Shao K, Liu K, Lai Y, Jiang S (2022) Research progress on low temperature chilling injury of rice. *Jiangsu Agricultural Science* 50 (24): 9-15.

<https://doi.org/10.15889/j.issn.1002-1302.2022.24.002>

Lv Z (2017) Rice variety selection experiment in Jiwen district of Bei'an city. *Modern Agricultural Science and Technology* 691 (05): 52+57.

Ma B, Liu C, Hu J, Liu K, Zhao F, Wang J, Zhao X, Guo Z, Song L, Lai Y, Tan K (2022) Intelligent identification and features attribution of saline-alkali-tolerant rice varieties based on Raman spectroscopy. *Plants* 11(9).

Min S, Ding Z, Zhengyong Z, Jinhong W, Yuan C, Mengtian W, Jun L (2020) Improving Raman spectroscopic identification of rice varieties by feature extraction. *Journal of Raman Spectroscopy* 51:4.

Pezzotti G, Zhu W, Chikaguchi H, Marin E, Boschetto F, Masumura T, Sato Y, Nakazaki T (2021) Raman molecular fingerprints of rice nutritional quality and the concept of Raman barcode. *Frontiers in Nutrition* 8: 663569.

Sun X, Dong G, Wei C, Hua S, Guan W, Xu W, Chen L, Jiang X, Chu X, Fan T, Wang Q, Ren C (2020) Low temperature and cold damage of rice in the northeast cold region and its defense measures. *Northern Rice* 50(06):64-65+69.

<https://doi.org/10.16170/j.cnki.1673-6737.2020.06.022>

Sun Y, Liu D, Wang B (2019) Study on rice variety selection and irrigation mode in the first accumulated temperature region in cold regions. *Water Conservancy Science and Cold Region Engineering* 2 (06): 1-6.

Tian FM, Tan F, Li H (2020) A rapid nondestructive testing method for distinguishing rice producing areas based on Raman spectroscopy and support vector machine. *Vibrational Spectroscopy* 107:103017.

Tian M, Zhang S, He Y (2019) Development and validation of a molecular marker for the low temperature tolerance gene bZIP73 in rice. *Jiangsu Agricultural Journal* 35 (06): 1265-1270.

Vankeirsbilck T, Vercauteren A, Baeyens W, Van DWD (2002) Applications of Raman spectroscopy in pharmaceutical analysis. *TrAC Trends in Analytical Chemistry* 21: 869–877.

Vladimir N, Michiel K (2009) On stochastic optimization and statistical learning in reproducing Kernel Hilbert spaces by support vector machines (SVM). *Informatica* 20:273–292. arXiv:1807.06521v2 [cs.CV]

Wang G, Liu S (2022) Heilongjiang communication industry successfully completed the communication guarantee work for the 2022 International Green Expo and Rice Festival. *Communication Management and Technology* 06: 4

Wang X, Zhang F, Wan X (2023) Study on genetic diversity of rice local germplasm based on molecular markers and phenotypic traits. *Journal of Plant Genetic Resources*: 1-19.

<https://doi.org/10.13430/j.cnki.jpgr.20221018002>

Xie L (2022) International Green Expo and Rice Festival Build a Global Cooperation Bridge. *China Trade Journal* 11(15):005.

<https://doi.org/10.28113/n.cnki.ncmyb.2022.001316>

Xue J (2016) Classification of source and sink types of rice populations in the third accumulated temperature region of Heilongjiang province. *Heilongjiang Agricultural Science* 259 (01): 1-7.

Zhu P (2021) Study on the identification method of northern japonica rice seed varieties based on Raman spectroscopy. Heilongjiang Bayi Agricultural University.

<https://doi.org/10.27122/d.cnki.ghlnu.2021.000256>