ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

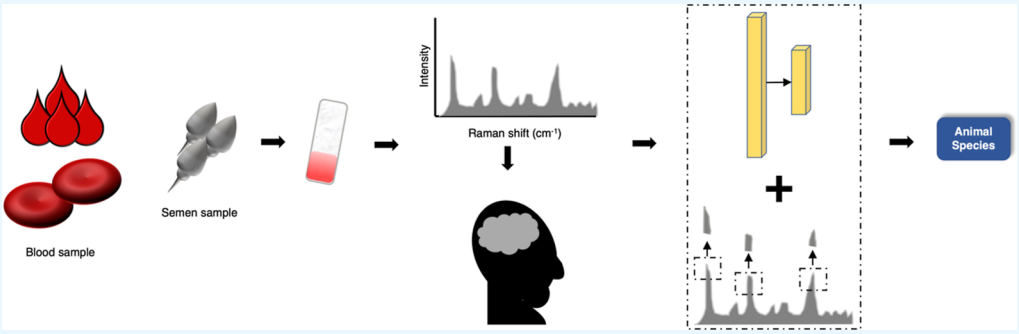
Raman ConvMSANet: A High-Accuracy Neural Network for Raman Spectroscopy Blood and Semen Identification

Pengju Ren, Ri-gui Zhou,* Yaochong Li, Shengjun Xiong, and Bing Han

Cite This: *ACS Omega* 2023, 8, 30421−30431

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Animal blood and semen analysis plays a significant role in national biological resource management, wildlife conservation, and customs security quarantine. Traditional blood analysis methods have disadvantages, such as complex sample preparation, time consumption, and false positives. Therefore, proposing a rapid and highly accurate analysis method is highly valuable. Raman spectroscopy has been widely used in blood analysis, and efficient and accurate analysis results can be obtained through the machine learning algorithm feature extraction. Recently, the transformer network structure was applied to Raman spectroscopy recognition. However, the multihead self-attention mechanism does not perform well in extracting local feature peaks, although it obtains global feature relations. This paper proposes a neural network based on the combination of one-dimensional convolution and multihead self-attention mechanism (Raman ConvMSANet) to identify 52 species of blood and semen Raman spectra. The network can achieve reliable identification effects in multiclassification and sample imbalance situations, and the average identification accuracy of blood and semen can reach more than 98.5%. The proposed network model can be applied not only to blood and semen identification but also to other biological fields.

INTRODUCTION

Interspecies blood and semen identification has played a crucial role in analytical chemistry, forensics, biochemistry, customs inspection, and wildlife preservation. Illegal hunting and the slaughter of wild animals seriously damage biodiversity and cause the mass extinction of rare animals.[1] Animal blood and semen contain a large amount of genetic material and genetic resources, and the blood of some precious species has great medicinal value.[2] Animal semen contains genetic material that can be used for artificial insemination, breed improvement,[3] and other purposes. The theft and loss of biological resources will cause serious consequences, such as invading alien species and losing genetic property rights. A rapid, accurate, and nondestructive detection method is the key to preventing the loss of biological resources. The present techniques for interspecies blood identification are high-performance liquid chromatography (HPLC),[4] mass spectrometric analysis (MS),[5,6] quantitative PCR,[7] and the genome profiling method (GP method).[8] Inouel et al.[4] used the HPLC method to identify human, primate, and nonprimate blood. Espinoza et al.[5] used the MS method to identify differences in blood and blood composition in 62 different species. Five body fluids were differentially analyzed using the quantitative PCR method by Sauer et al.[7] Human blood samples were distinguished from the other nine animal blood samples using the GP method by Suwa et al.[8] However, all of these methods require a complex sample preparation process, which not only destroys the sample but also poses a risk of infection to the experiment operator due to the need to contact the sample, and the experiment is a time-consuming process. Hence, it is of great practical value to propose a new method of blood species identification. Vibrational spectroscopy is a rapid and noninvasive method for identifying the composition of chemicals and functional groups in solid, liquid, and gaseous forms. Raman spectroscopy is one of them that is frequently

Received: May 22, 2023
Accepted: August 1, 2023
Published: August 11, 2023

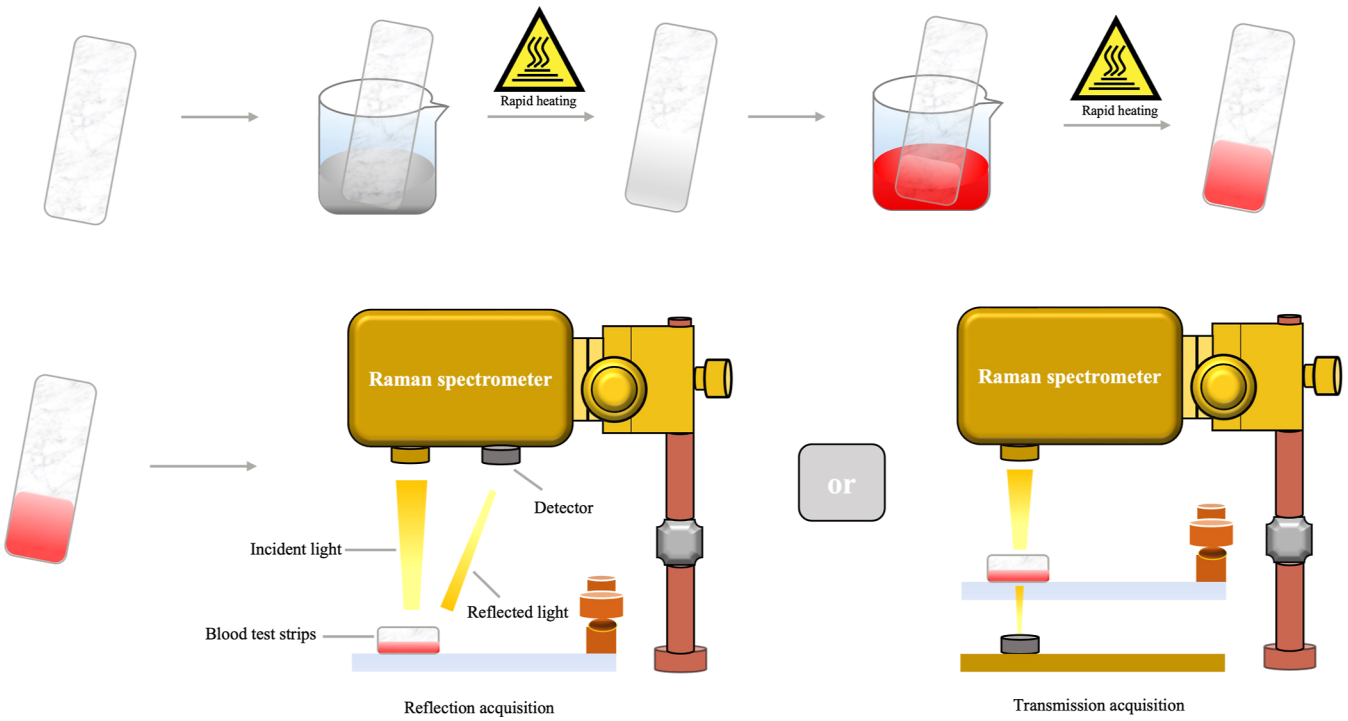


Figure 1. Schematic diagram of Raman spectra measurements.

used to analyze structural information about substances. Raman spectroscopy has the advantages of nondestructive, rapid and label-free. At present, Raman spectroscopy has been applied to many technical fields, including food safety,[9,10] clinical laboratory,[11−13] petrochemicals,[14] etc.

Raman spectroscopy is currently being used in blood species identification studies. The interaction of laser photons with molecules in the blood sample creates vibrational modes that lead to Raman scattering. The information provided by this scattering includes vibrational information about molecules such as hemoglobin, proteins, lipids, and sugar molecules in the blood, which can be used to determine the concentration and composition of these molecules in the blood and to analyze and identify the blood. With the development of machine learning, Raman spectroscopy has achieved better identification results with machine learning algorithms. Partial least-squares discriminant analysis (PLS-DA)[15−17] and principal component analysis (PCA)[18,19] are widely used in blood spectral analysis. The first successful study of blood identification using near-infrared Raman spectroscopy was carried out by Virkler and Lednev[18] in 2009. They used PCA to extract three principal component features and plotted 3D maps to distinguish human, feline, and canine blood visually. To distinguish between human and nonhuman blood from the Raman spectra of 10 blood species, Bian et al.[16] constructed a pairwise double PLS-DA model, significantly improving single PLS-DA identification. Wang et al.[20] used the support vector machine (SVM) method to identify four poultry species' blood and analyze the presence of food additives in the blood. However, while classical machine learning methods provide better identification results for small data samples, they still have certain limitations for large data samples. As a branch of machine learning, artificial neural networks extract data features through multilayer neural networks. Deep convolutional neural networks (CNNs) have achieved excellent results in image classification.[21,22] In contrast, recurrent neural networks (RNNs) have shown promising effects in feature extraction in audio and other sequence data.[23] Therefore, applying deep neural networks to Raman spectroscopy feature extraction is worth exploring, and multilayer artificial neural networks have been used for spectral blood identification analysis. Dong et al.[24] constructed a two-layer CNN for Raman spectra to perform denoising and baseline correction preprocessing, and two fully connected layers were used for human and animal blood identification. The method achieved better results than the SVM and PLS-DA methods. Huang et al.[25] constructed a multilayer CNN based on Raman spectroscopy to identify 20 animal blood species, and the average accuracy of blind detection is more than 97%. The network models proposed by the scholars mentioned above have shown excellent performance in Raman spectroscopy identification. However, CNN models have limitations in Raman spectroscopy identification. First, they can only extract local feature information and not capture the relationship between local and global feature information, leading to suboptimal identification results. Second, CNN models have a large number of training parameters and high model complexity, resulting in slow training and prediction speeds. In order to solve the problem of insufficient global feature extraction by the convolutional network, the transformer model structure has been proposed.[26] It computes the correlation between local features to obtain global feature relations. The first Raman spectroscopy classification network based on the transformer network structure has been proposed to identify deep-sea cold seep microorganisms at the single-cell level.[27] However, the transformer network structure also has shortcomings. It focuses too much on the correlation between local features and lacks detailed extraction of local feature information. Therefore, designing a network that can both extract local spectral peak features and effectively capture the relationship among global features is the key to our research.

https://doi.org/10.1021/acsomega.3c03572
ACS Omega 2023, 8, 30421−30431

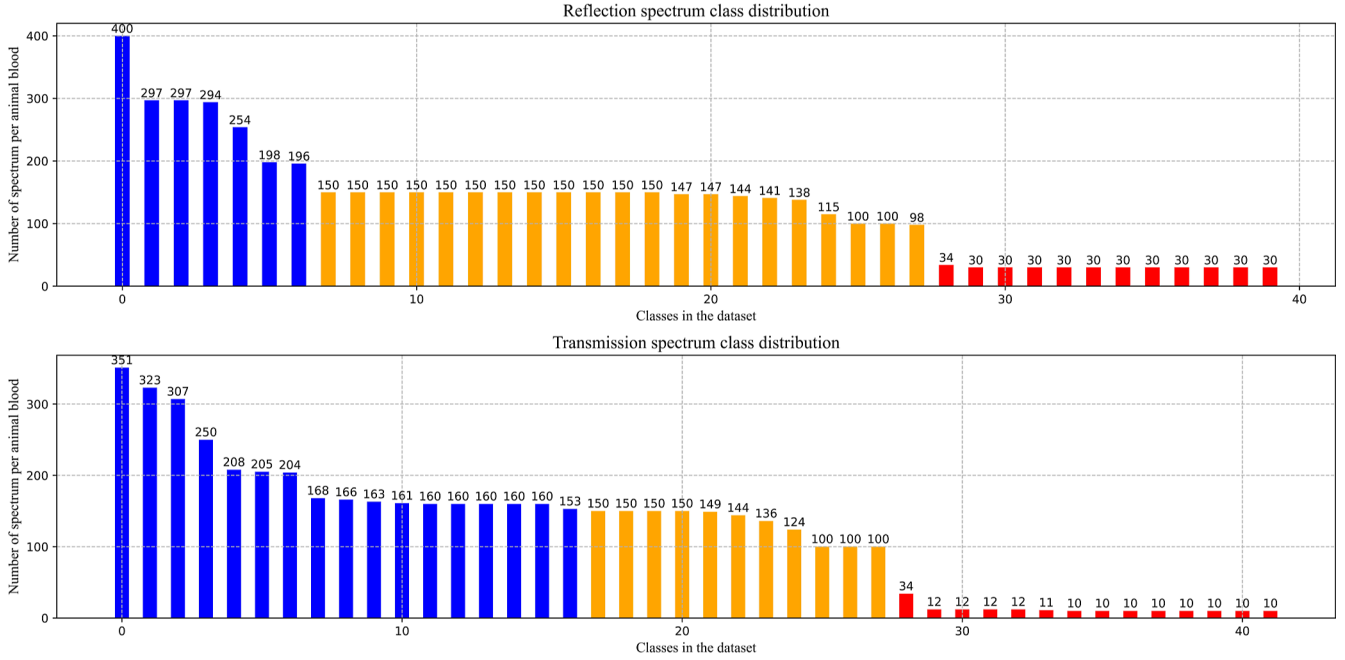


Figure 2. Distribution of animal blood and semen Raman spectral data sets. The top subgraph shows the distribution of the reflection spectrum data set, and the bottom subgraph shows the distribution of the transmission spectrum data set. The blue histograms are the classes where the number of samples is not less than 150, the orange histograms are the classes where the number of samples is less than 150 and not less than 50, and the red histograms are the classes where the number of samples is less than 50.

This paper proposes a Raman spectra feature extraction neural network called Raman ConvMSANet, which combines one-dimensional convolution and multihead self-attention mechanisms. The network is trained on both reflection and transmission Raman spectra data sets to verify its classification performance, which contains information on two types of biological resources, blood and semen of 52 national precious animal species such as panda, kangaroo, golden monkey, plum deer, and animal species common to our daily lives, with a total of 10,314 Raman spectra data. Experimental results show that Raman ConvMSANet has significantly improved identification accuracy and more stable classification performance than pure one-dimensional convolution and multihead self-attention mechanism neural networks. This confirms that our proposed network model can extract local and global features effectively. Our research is divided into three parts. First, we introduce the process of sample data collection, the training data set, and data preprocessing procedures and propose the network structure and training process. Second, we discuss the network's performance in identification, comparing it with existing networks that exhibit excellent identification performance and discussing the effectiveness of multihead self-attention. Finally, we summarize the research conclusions presented in this paper.

MATERIALS AND METHODS

Equipment and Instrument. Blood and semen samples were collected and diluted ten times with 0.9% NaCl water solution. The composite AgNP and AuNP test strips were used for surface-enhanced Raman spectroscopy. In this study, the surface-enhanced composite test paper was selected, which has the material applicability of both AgNPs and AuNPs and can produce significant enhancements to materials that can be enhanced by either AgNPs or AuNPs, which improves the convenience of surface-enhanced Raman detection. The detailed process for preparing the test strips was described in a previous work.[28] The test strips were taken out of the enhancement reagent and rapidly dried in a microheater at 65 °C. After drying, the test strip is dipped into the blood and semen samples for 5 s and then rapidly dried in a microheater at 65 °C. Raman spectra data of the blood and semen samples on the test strips were collected using a Raman spectrometer (HT-NOVA Co., Ltd.). Both reflection and transmission Raman spectra were acquired using a wavelength of 785 nm, 5 mW micropower laser with an integration time of 1 s to avoid destroying the internal components of the sample. The spectral range of the reflection Raman signals was 200−2998 cm^{-1} and that of the transmission Raman signals was 166−2084 cm^{-1}, with a spectral resolution of 4−6 cm^{-1}. The schematic diagram of the Raman spectral data collection process is shown in Figure 1. All collected blood and semen samples were stored in a −20 °C refrigerator with trisodium citrate as the anticoagulant. During the sample collection process, it was ensured that the animals were healthy and that they had not taken any medication. All of the sources of the spectra and the collection methods were approved and complied with the safety quarantine standards.

Data Set. In this study, two sets of Raman spectra data of blood and semen were constructed based on different collection methods, namely, reflection and transmission. The data sets included 10,314 Raman spectra from 52 species of animals, including rare animals such as pandas, kangaroos, golden monkeys, spotted deer, and common animals in our daily lives. The reflection data set consisted of 5229 spectra from 40 species while the transmission data set contained 5085 spectra from 42 species, with some species being inconsistent between the two data sets. Due to the difficulty of sample collection for rare and protected animal species with limited population, some species had only 10 spectra. In contrast, common animals were easier to obtain, with some having up to

https://doi.org/10.1021/acsomega.3c03572
ACS Omega 2023, 8, 30421−30431

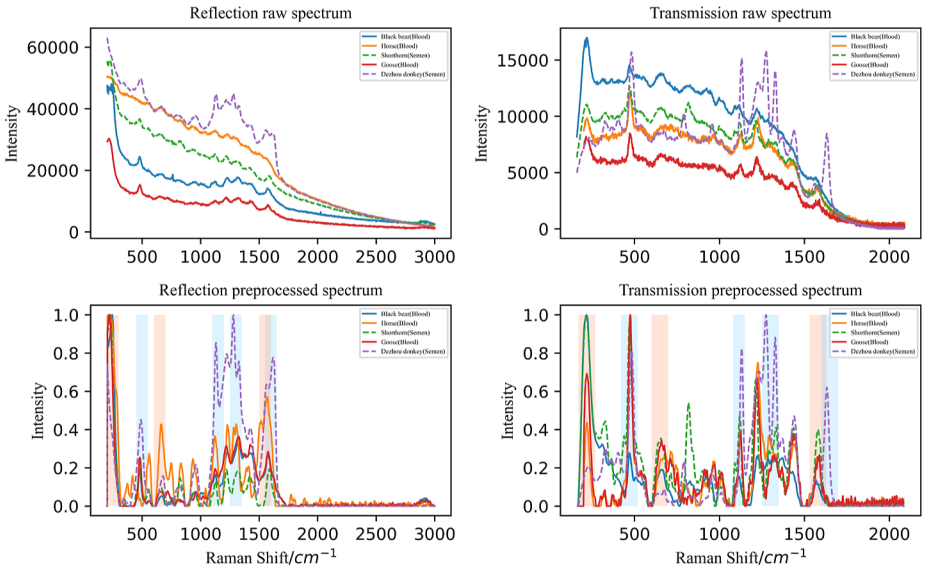


Figure 3. Comparison of raw and preprocessed Raman spectra; the two subgraphs on the left are reflection spectral data and the two subgraphs on the right are transmission spectral data. The solid line indicates blood, the characteristic peaks are marked in light salmon, the dashed line indicates semen, and the characteristic peaks are marked in light sky blue.

400 spectra. As a result, the data sets were imbalanced, presenting a challenge to the classification network. The data sets were divided into two independent sets to test the proposed network model's classification performance under different Raman spectral information collection methods. The distribution of the reflection and transmission blood Raman spectral data sets is shown in Figure 2.

Data Preprocessing. Effective preprocessing of Raman spectra data can remove invalid information, such as instrument noise and fluorescence background, during data acquisition and mitigate the impacts on the spectral data. The preprocessing operations applied to the spectral data include denoising, baseline correction, removing negative values, normalization, and down-sampling. Denoising can remove sharp spectral peaks caused by noise during acquisition, resulting in smoother data. The Savitzky–Golay[29] filter with a 3rd order polynomial is used for data smoothing to achieve denoising. The drift of the baseline can blur the signal and deteriorate the analysis results, affecting model training. To correct the baseline, the airPLS[30] algorithm is used. Negative data values are replaced with 0 to avoid the impact of normalization in subsequent processing steps. Normalization of the data can increase the speed of the model computation. Finally, due to differences in the equipment and measurement methods used to acquire the data, down-sampling is necessary to ensure that the Raman shifts at the same position in different spectra are the same and that the dimensions of all Raman spectra in the same data set are consistent. From 200 to 2998 cm^{-1}, sampling every 2 cm^{-1}, a total of 1400 spectral points in the reflection spectrum data set. From 166 to 2084 cm^{-1}, sampling every 2 cm^{-1}, a total of 960 spectral points in the transmission spectrum data set. The missing spectral points are filled in with 0. A comparison of the raw and preprocessed spectral data is shown in Figure 3, where the solid line represents the blood Raman spectra and the dashed line represents the semen Raman spectra. The characteristic peaks that are visible to the naked eye are marked prominently, with the blood and semen characteristic peak regions indicated in light salmon and light sky blue, respectively. The reflection Raman spectra of the blood samples show several specific peaks, such as at 250, 698, and 1568 cm^{-1}, and the semen samples show several specific peaks, such as at 490, 1132, 1280, and 1628 cm^{-1}. The transmission Raman spectra of the blood samples show several specific peaks, such as at 246, 703, 1578 cm^{-1}, and the semen samples show several specific peaks, such as at 496, 1130, 1274, and 1632 cm^{-1}. The data set is divided into a training set, a validation set, and a test set in an 8:1:1 ratio, and 10-fold cross-validation is used to ensure that all spectra are used for training and testing, evaluating the model's generalization performance.

One-Dimensional Convolution Combined with Multihead Self-Attention Mechanism (Raman ConvMSANet). Raman spectroscopy is one-dimensional data, and it is typically used to extract features from spectra using one-dimensional convolution. Convolution operations have outstanding performance in extracting local feature peaks in spectra but due to the limitation of convolution kernel size increasing the network layer is necessary to expand the receptive field when the kernel is small. When the kernel is large, the number of parameters increases and the computational complexity increases. The shortcomings of the convolutional calculation are thus evident. Although local information is extracted well, it is not sensitive to global information and it is insufficient to use convolutional calculations for feature extraction in the network. The transformer structure has been proposed recently and has achieved excellent results in natural language processing (NLP) and computer vision (CV). Liu et al.[27] have proposed the first Raman spectroscopy classification network based on the transformer structure. However, although the transformer structure relies on internal multihead self-attention mechanisms to solve the problem of convolutional networks being insensitive to global information, it is still weaker than convolutional networks in extracting local information. Therefore, we consider combining these two structures, which can finely extract spectral feature peaks while also linking the correlation between each feature peak. We propose a Raman spectroscopy classification network (Raman ConvMSANet)

https://doi.org/10.1021/acsomega.3c03572
ACS Omega 2023, 8, 30421−30431

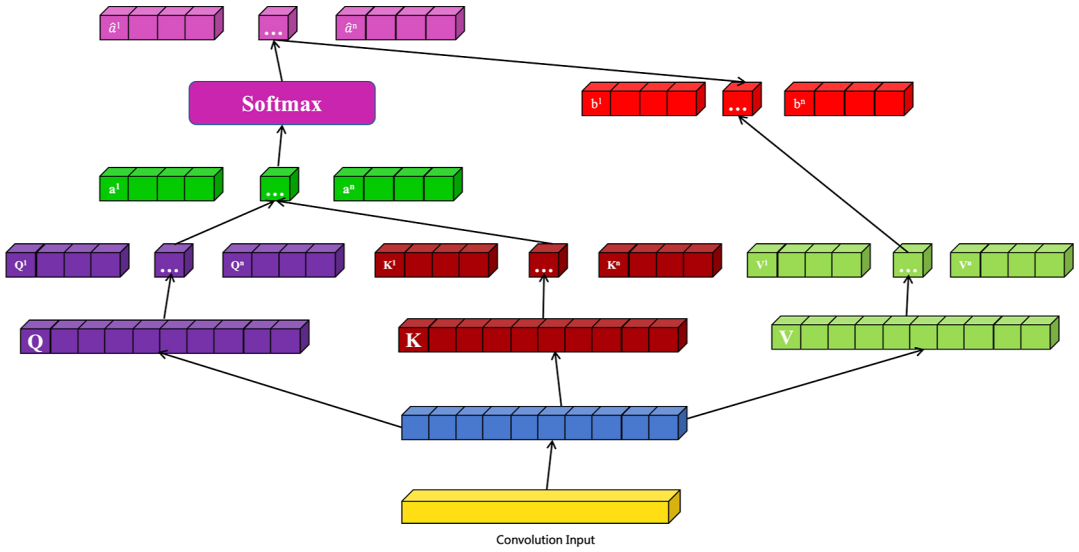


Figure 4. Schematic diagram of computing the multihead self-attention mechanism.

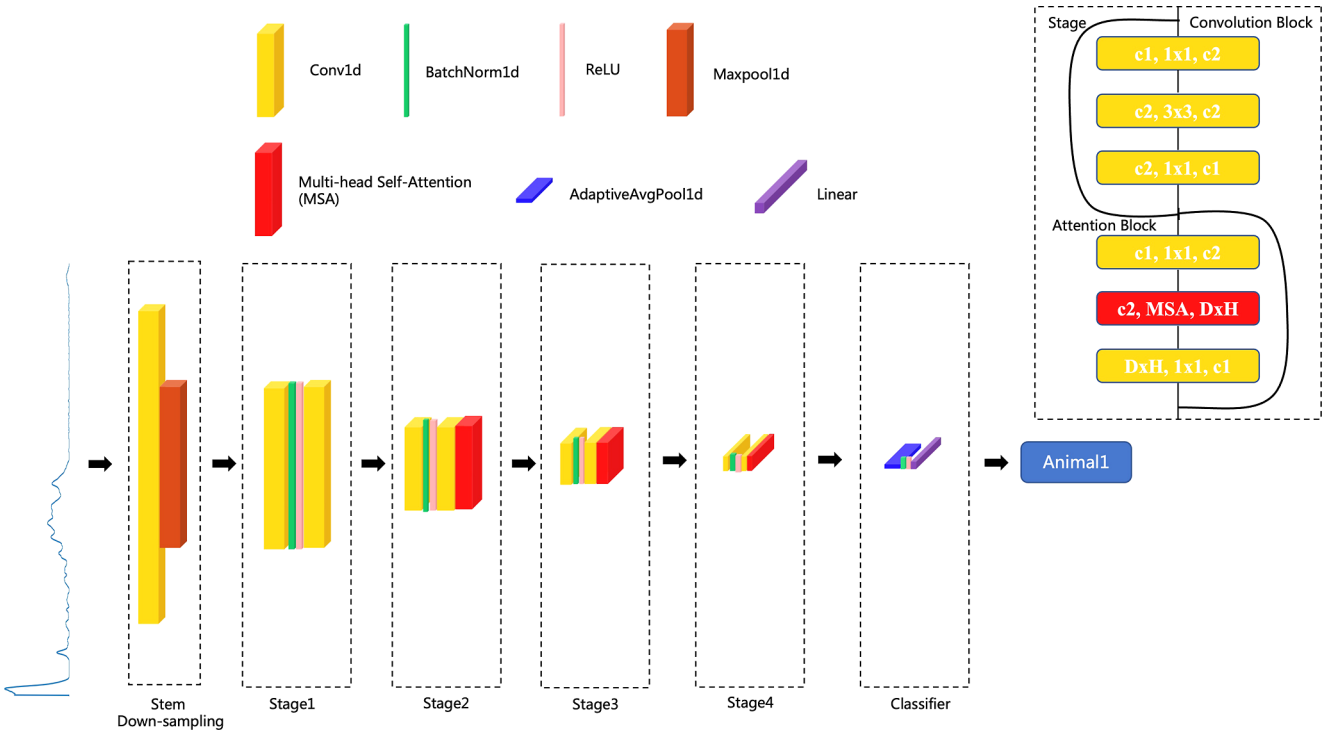


Figure 5. Schematic diagram of Raman ConvMSANet.

that combines one-dimensional convolution and multihead self-attention mechanisms.

The design of the network structure was inspired by Park and Kim.[31] The one-dimensional convolution and multihead self-attention mechanisms are complementary, and it is proposed to place the multihead self-attention after each convolution calculation. First, the input spectrum is down-sampled to reduce the data dimension. The network is mainly divided into four stages. The first stage consists of two layers of one-dimensional convolution, which does not change the length of the spectrum. The latter three stages consist of two layers of one-dimensional convolution and one layer of multihead self-attention. In each stage, the length of the spectrum is reduced by half and the channel number is doubled. A residual connection is used between each stage to prevent information loss and gradient explosion. A normalization layer and ReLU activation function are added between every two layers of one-dimensional convolution. In the multihead self-attention mechanism, the convolutional features extracted from the spectra are divided into several patches of equal size. Each patch is added with relative position encoding to record the position information on each spectral block (eq 1). Here, X_P^N represents the N-th convolutional feature vector of size P, E_{pos} is the corresponding position information, and D is the number of convolutional feature channels. For integrating the feature patches to obtain Z and dot-multiplying with the initialized Uqkv weight vector based on the number of attention heads, the resulting vector is divided into a number

https://doi.org/10.1021/acsomega.3c03572
ACS Omega 2023, 8, 30421−30431

of Q, K, and V pairs, with Q, K, and V being the trainable weight vectors for each feature patch, and the network is adjusted according to the loss function after training iterations (eq 2). The dot product between different Q and K vectors is calculated to obtain their correlation. The correlation is scaled to the range of (0, 1) using Softmax (eq 3) and then the final weights of the correlation between different patches are calculated by a dot product with different V vectors (eq 4). The higher the correlation between the two patches, the larger the weight. The calculation process of multihead self-attention is shown in Figure 4, where "b" is the final calculated correlation weight (eq 5). Finally, adaptive average pooling is used to reduce the dimensions of the spectral data, and a fully connected layer is used to output the final classification result. The overview of the network structure is shown in Figure 5, and the training parameter details for the reflection spectra are shown in Table 1. The training parameter details for the transmission spectra are shown in the mentary file. The network is built using PyTorch (version: 2.0.0) and Python (version: 3.9.16) in the NVIDIA 2080Ti graphics card environment for training acceleration.

Table 1. Parameter Details of Raman ConvMSANet during Training on Reflection Raman Spectra[a]

layers	type	numbers	output shape	parameters
input	input	1	(batch, 1, 1400)	
stem	Conv1D	1	(batch, 1, 700)	(16, 7, 2)
	Maxpool1D	1	(batch, 16, 350)	(3, 2)
stage 1	Conv1D	2	(batch, 16, 350)	(32, 3, 1/1)
	BatchNorm1D	1		
	ReLU	1		
stage 2	Conv1D	2	(batch, 32, 350)	(64, 3, 2/1)
	MSA	1	(batch, 64, 175)	(1, 35)
	BatchNorm1D	1		
	ReLU	1		
stage 3	Conv1D	2	(batch, 64, 175)	(128, 3, 2/1)
	MSA	1	(batch, 128, 88)	(2, 44)
	BatchNorm1D	1		
	ReLU	1		
stage 4	Conv1D	2	(batch, 128, 88)	(256, 3, 2/1)
	MSA	1	(batch, 256, 44)	(4, 22)
	BatchNorm1D	1		
	ReLU	1		
classifier	AdaptiveAvgpool1D	1	(batch, 256, 44)	(7)
	BatchNorm1D	1		
	ReLU	1		
	flatten	1	(batch, 256*7)	
output	linear	1	(256*7, #classes)	(#classes)

[a]The parameters of the Conv1D represent (filters, kernel_size, stride), the Maxpool1D represent (kernel_size, stride), the MSA represent (head_numbers, patch_size), the AdaptiveAvgpool1D represent (output_size), and the Linear represent (out_features).

$$Z = [X_P^1; X_P^2; \ldots; X_P^N] + E_{\text{pos}}\ X \in R^{P\times D},\quad E_{\text{pos}} \in R^{N\times D} \tag{1}$$

$$[Q, K, V] = ZU_{QKV}\ U_{QKV} \in R^{D\times 3D_h} \tag{2}$$

$$A = \text{Softmax}(QK^T/\sqrt{D_h}) \tag{3}$$

$$SA(z) = AV \tag{4}$$

$$MSA(z) = [SA_1(z); SA_2(z); \ldots; SA_h(z)] \tag{5}$$

Network Model Training. The network was trained on reflection and transmission Raman spectra data sets using AdamW as the optimizer. The learning rate decreased gradually with training epochs using an exponential strategy with γ = 0.98 and an initial learning rate 0.0003. The CrossEntropyLoss function evaluated the difference between the predicted and true values. The maximum number of training epochs was set to 200, and when the validation set classification accuracy did not improve within 10 epochs, training was stopped early to prevent overfitting. On average, the network was trained for 80 epochs. The loss and accuracy curves of the trained models on both data sets are shown in Figure 6.

RESULTS AND DISCUSSION

Model Evaluation and Classification Performance. The network model evaluates classification performance and generalization ability using 10-fold cross-validation. In the classification network model, the most intuitive evaluation metric is the accuracy of classification, which is the ratio of the number of correctly classified samples to the total number of test samples (eq 6). Precision, recall, and $F1$ score are also used as evaluation metrics for the model. In this study, because the two data sets have an imbalanced sample size, with more spectral data of common animal species than of rare national animals, we used weighted average precision, recall, and $F1$ score. Assuming that for class i in the test set, the number of samples is N_i, then the weight is W_i (eq 7). Precision evaluates the proportion of true positive data among all data predicted as positive (eq 8). Recall evaluates the proportion of positive data correctly predicted among all positive data (eq 9). $F1$ score evaluates the overall performance of model precision and recall (eq 10). The model accuracy, precision, recall, and $F1$ score are shown in Figure 7, where it can be seen that the model achieves an average accuracy of 98.6 and 98.5% on the two data sets, demonstrating excellent performance in the case of multiclassification and imbalanced samples. The average weighted precision is 99.7 and 99.8%, indicating the model's high accuracy in predicting positive samples. The average weighted recall is 98.6 and 98.5%, showing that the model correctly identifies a high proportion of positive samples among all samples. The model's average weighted $F1$ score is 99.1 and 99.2%, indicating that the model has achieved a favorable balance between precision and recall.

After training the network, confusion matrices were plotted for the predicted results of the test samples in the two data sets, as shown in Figure 8, providing a clear and intuitive visualization of the model's classification performance on each class. In the matrices, the vertical axis represents the correct class labels in the data set, while the horizontal axis represents the predicted classes by the model. The receiver operating

https://doi.org/10.1021/acsomega.3c03572
ACS Omega 2023, 8, 30421−30431

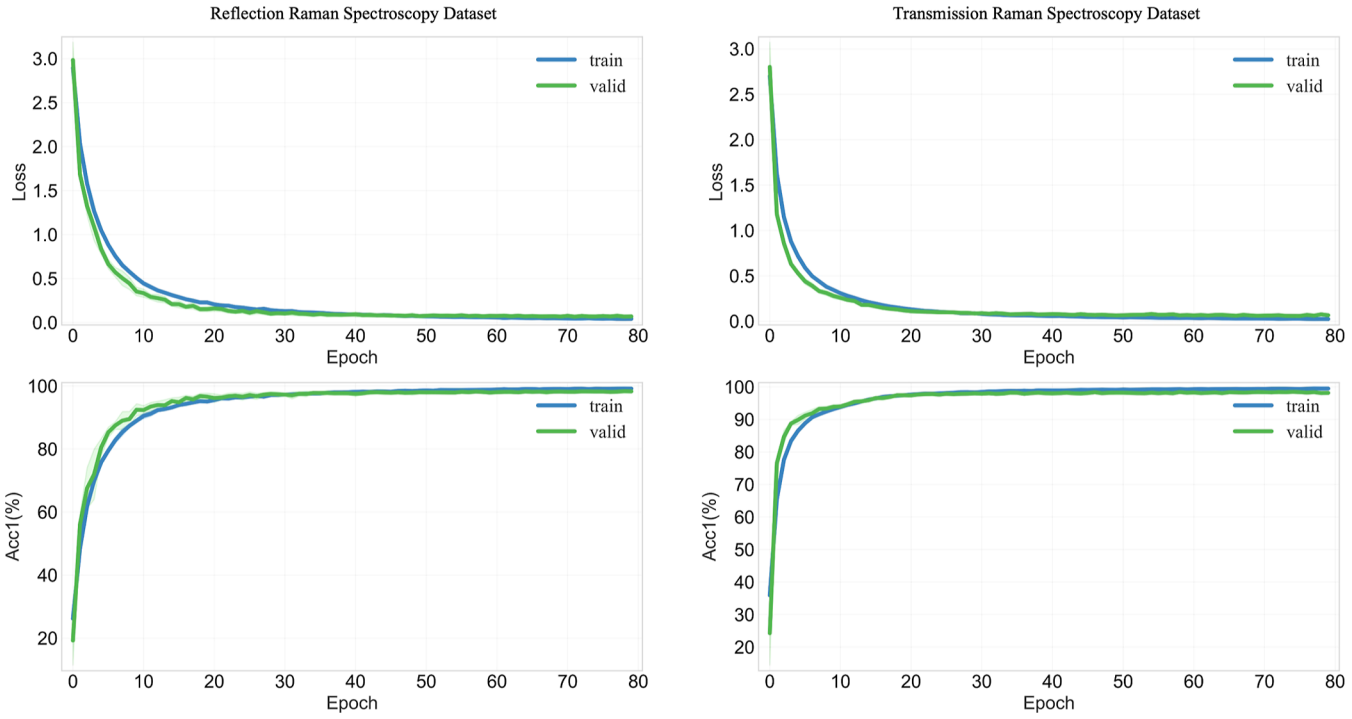


Figure 6. Training loss and accuracy curves of the network are shown in the two left subplots for the reflection spectral data set and in the right two subplots for the transmission spectral data set.

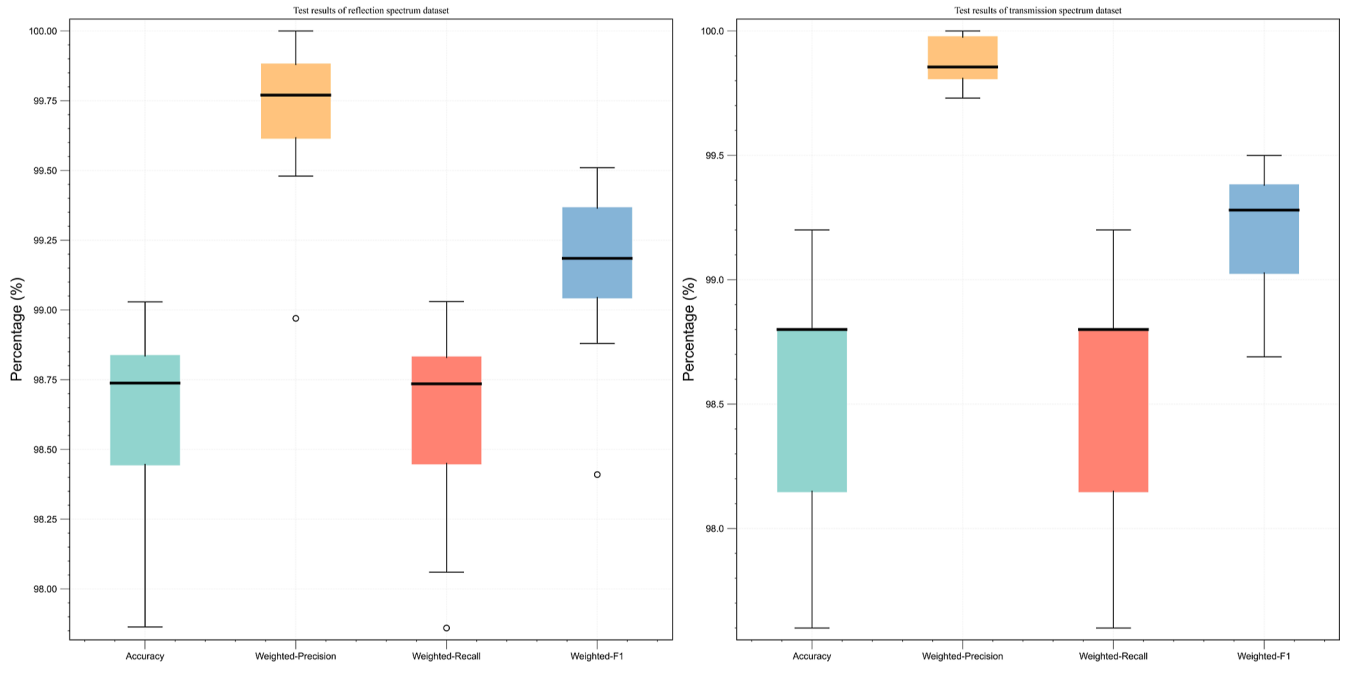


Figure 7. Boxplots of evaluation metrics for Raman ConvMSANet are presented, with the left subplot displaying the 10-fold cross-validation results for the reflection spectral data set and the right subplot displaying the 10-fold cross-validation results for the transmission spectral data set.

characteristic (ROC) curve was utilized to assess the sensitivity and specificity of classification results, and the area under the curve (AUC) is a common evaluation metric for models. The vertical axis of the ROC curve represents the true positive rate (TPR), and the horizontal axis represents the false positive rate (FPR). An effective model is demonstrated by a ROC curve that is closer to the upper-left corner (0, 1) of the graph, which implies a larger area under the curve (AUC value) formed by the curve and the horizontal axis and the vertical line at FPR = 1, signifying higher TPR and lower FPR. The ROC curves for each class in the test samples of the two data sets are shown in Figure 9. These model evaluation metrics demonstrate that our proposed Raman ConvMSANet performs well in multi-classification tasks of Raman spectra and has high credibility.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} = \frac{\text{TP}_{\text{total}} + \text{TN}_{\text{total}}}{N_{\text{total}}} \quad (6)$$

https://doi.org/10.1021/acsomega.3c03572
ACS Omega 2023, 8, 30421−30431

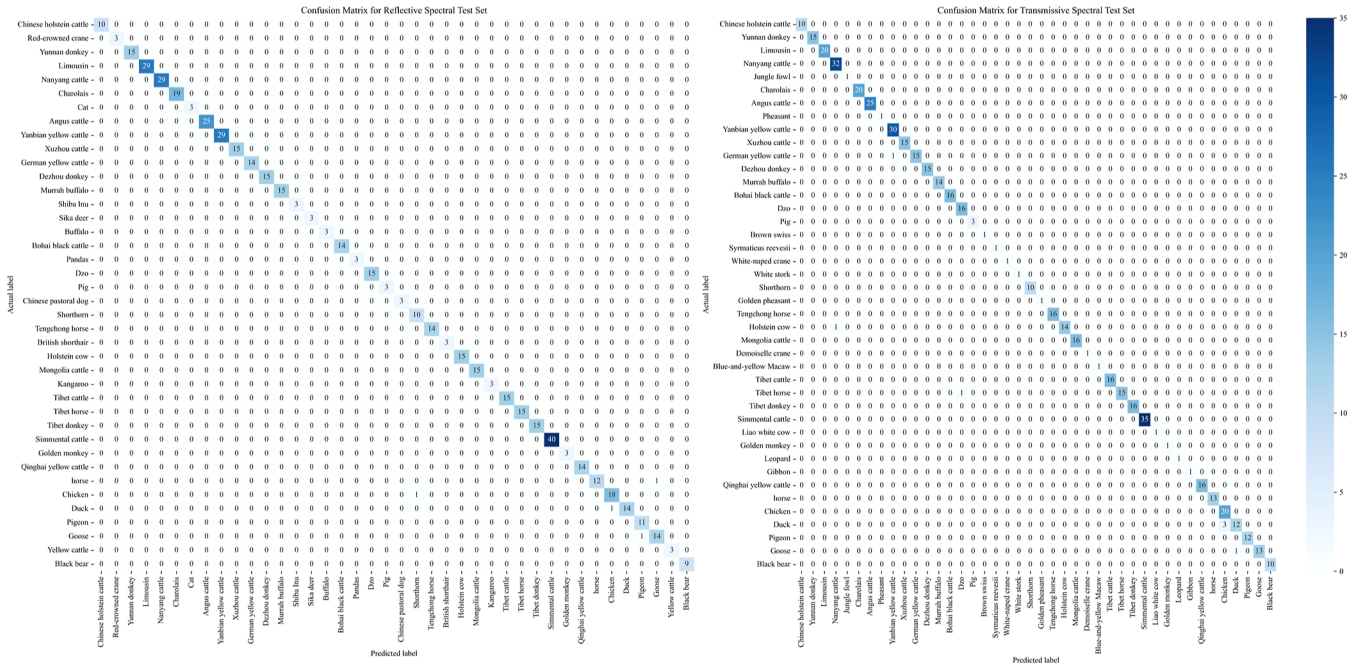


Figure 8. Confusion matrix results for Raman ConvMSANet are depicted, with the left subplot illustrating the confusion matrix for the reflection spectral data set and the right subplot illustrating the confusion matrix for the transmission spectral data set.

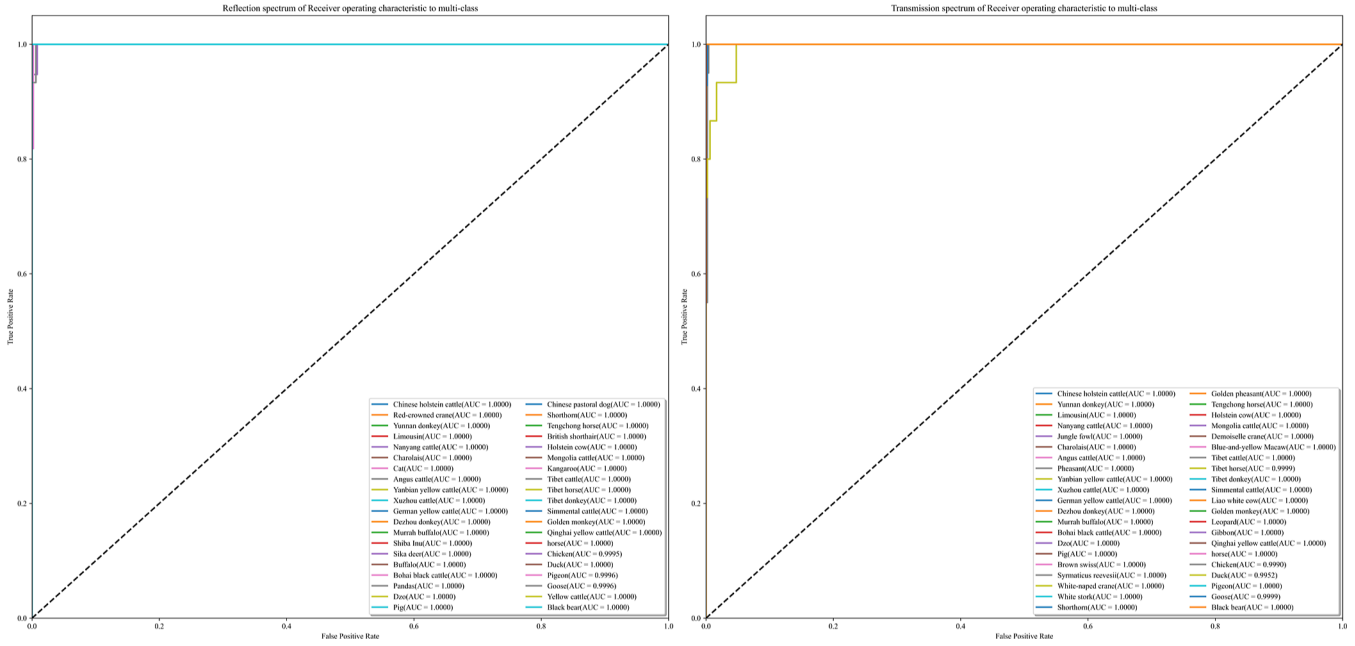


Figure 9. Multiclass ROC curves of Raman ConvMSANet are shown, with the left subplot displaying the ROC curve for the reflection spectral data set and the right subplot displaying the ROC curve for the transmission spectral data set.

$$W_i = \frac{N_i}{N_{\text{total}}} \tag{7}$$

$$\text{Weighted-precision} = \sum_{i=1}^{n} W_i \times \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i} \tag{8}$$

$$\text{Weighted-recall} = \sum_{i=1}^{n} W_i \times \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i} \tag{9}$$

$$\text{Weighted-}F1 = \sum_{i=1}^{n} W_i \times \frac{2 \times \text{precision}_i \times \text{recall}_i}{\text{precision}_i + \text{recall}_i} \tag{10}$$

Model Performance Comparison. In order to verify the better feature extraction capability of Raman ConvMSANet, comparative experiments were conducted to compare the classification performance with other state-of-the-art Raman spectroscopy classification models proposed by other scholars, including pure 1D convolutional networks, pure multihead self-attention networks, and recurrent neural networks. These network structures proposed by Huang et al.,[25] Bratchenko et

https://doi.org/10.1021/acsomega.3c03572
ACS Omega 2023, 8, 30421−30431

Table 2. Comparative Results between the Methods Proposed by Previous Scholars and Raman ConvMSANet[a]

	Top_1 Acc.	Top_5 Acc.	weighted precision	weighted recall	weighted F1 score	parameters (M)	FLOPs (G)
Huang et al.[25b]	0.9691 (0.9553−0.9805)	0.9990 (0.9980−1.0000)	0.9966 (0.9904−0.9990)	0.9691 (0.9515−0.9806)	0.9824 (0.9703−0.9896)	14.5767	2.211255
Bratchenko et al.[13b]	0.9712 (0.9611−0.9786)	0.9996 (0.9980−1.0000)	0.9966 (0.9929−0.9990)	0.9712 (0.9612−0.9786)	0.9836 (0.9770−0.9875)	2.4716	0.0304
Al-Shaebi et al.[32b]	0.9871 (0.9844−0.9902)	0.9992 (0.9961−1.0000)	0.9976 (0.9949−1.0000)	0.9871 (0.9825−0.9903)	0.9923 (0.9898−0.9951)	1.5840	0.1220
Chen et al.[33b]	0.9778 (0.9728−0.9825)	0.9992 (0.9980−1.0000)	0.9964 (0.9934−0.9990)	0.9778 (0.9728−0.9825)	0.9869 (0.9830−0.9895)	1.1433	0.1320
Sang et al.[34b]	0.9798 (0.9747−0.9844)	0.9992 (0.9980−1.0000)	0.9970 (0.9946−1.0000)	0.9798 (0.9748−0.9845)	0.9882 (0.9850−0.9921)	18.2450	1.0819
Liu et al.[27b]	0.9825 (0.9708−0.9922)	0.9984 (0.9941−1.0000)	0.9975 (0.9915−1.0000)	0.9825 (0.9709−0.9922)	0.9898 (0.9842−0.9941)	8.0162	0.0878
ConvMSANet[b]	0.9860 (0.9786−0.9902)	0.9990 (0.9961−1.0000)	0.9969 (0.9897−1.0000)	0.9860 (0.9786−0.9903)	0.9914 (0.9841−0.9951)	0.9476	0.0553
Huang et al.[25c]	0.9794 (0.9740−0.9920)	0.9976 (0.9920−1.0000)	0.9982 (0.9970−1.0000)	0.9794 (0.9700−0.9920)	0.9886 (0.9841−0.9950)	10.9674	1.5142
Bratchenko et al.[13c]	0.9824 (0.9700−0.9900)	0.9986 (0.9940−1.0000)	0.9987 (0.9980−1.0000)	0.9824 (0.9700−0.9940)	0.9904 (0.9841−0.9970)	2.4432	0.0268
Al-Shaebi et al.[32c]	0.9870 (0.9740−0.9940)	0.9992 (0.9960−1.0000)	0.9991 (0.9952−1.0000)	0.9870 (0.9740−0.9940)	0.9929 (0.9866−0.9960)	1.5601	0.0899
Chen et al.[33c]	0.9756 (0.9560−0.9880)	0.9980 (0.9960−1.0000)	0.9988 (0.9970−1.0000)	0.9756 (0.9560−0.9880)	0.9869 (0.9758−0.9934)	1.0261	0.0910
Sang et al.[34c]	0.9824 (0.9700−0.9920)	0.9990 (0.9960−1.0000)	0.9995 (0.9971−1.0000)	0.9824 (0.9700−0.9920)	0.9907 (0.9845−0.9960)	14.8381	0.7432
Liu et al.[27c]	0.9840 (0.9760−0.9920)	0.9992 (0.9960−1.0000)	0.9986 (0.9953−1.0000)	0.9840 (0.9760−0.9920)	0.9912 (0.9876−0.9960)	7.9839	0.0874
ConvMSANet[c]	0.9850 (0.9760−0.9920)	0.9986 (0.9940−1.0000)	0.9987 (0.9973−1.0000)	0.9850 (0.9760−0.9920)	0.9917 (0.9869−0.9939)	1.2952	0.0515

[a]The error range of the 10-fold cross-validation is indicated in brackets. [b]Reflection data sets. [c]Transmission data sets.

al.,[13] Al-Shaebi et al.,[32] Chen et al.,[33] Sang et al.,[34] and Liu et al.[27] were constructed and trained and tested on two data sets to evaluate their classification performance using 10-fold cross-validation. Each network was pretuned based on the two data sets to ensure the network was trained with the optimal parameters. The comparison results are shown in Table 2. The results show that Raman ConvMSANet classification performance was slightly improved in both data sets under the same model parameter size and computational complexity level. Compared to the UNet Raman spectroscopy classification network proposed by Al-Shaebi et al., Raman ConvMSANet exhibited slightly lower accuracy on both data sets but with a favorable advantage in terms of model parameters and computational complexity.

Multihead Self-Attention Classification Performance Analysis. After the Raman spectra are fed into the network, a series of one-dimensional convolutions are applied to extract local spectral peak features. In order to verify the effectiveness of adding the multihead attention mechanism after the convolution operation, the multihead self-attention is removed, and only the convolution operation is retained. The model classification performance is shown in Table 3. Although adding multihead self-attention increases the model parameters and computational complexity, it also improves the classification performance compared to the pure convolutional network. Regarding the parameters of patch size and number of heads in the multihead self-attention, their impact on the final classification performance is relatively small. When the patch size is too small and the number of heads is too large, the model parameters and computational complexity increase due to the generation of more (Q, K, V) pairs. In this study, for the reflection Raman spectra data set, the patch sizes and numbers of heads in the last three stages of the network were chosen as

Table 3. Effectiveness of the MSA Structure on the Classification Performance in Both Data Sets[a]

case	Top_1 Acc.	parameters (M)	FLOPs (G)
MSA w/[b]	0.9860 (0.9786−0.9902)	0.9476	0.0533
MSA w/o[b]	0.9827 (0.9708−0.9941)	0.5153	0.0298
MSA w/[c]	0.9850 (0.9760−0.9920)	1.2952	0.0515
MSA w/o[c]	0.9802 (0.9720−0.9880)	0.5188	0.0204

[a]The error range of the 10-fold cross-validation is indicated in brackets. [b]Reflection data sets. [c]Transmission data sets.

(35, 44, 22) and (1, 2, 4), respectively, while for the transmission Raman spectra data set, the patch sizes and numbers of heads were chosen as (15, 10, 5) and (2, 4, 8), respectively. The influence of patch size and the number of heads' parameter selection on the classification performance is shown in Table 4.

CONCLUSIONS

The illegal hunting of endangered species seriously endangers ecological balance and biodiversity, and the analysis of the blood and semen of these species plays an essential role in customs quarantine. Deep learning methods can effectively identify the corresponding species of blood and semen by their Raman spectra, reducing the phenomenon of illegal smuggling of endangered species. This paper proposes a high-accuracy neural network model for Raman spectra identification of species in blood and semen. To address the problem that one-dimensional convolution is insensitive to global spectral features and that multihead self-attention mechanisms are insufficient for extracting local feature peaks in spectra, we propose a solution that effectively combines one-dimensional convolution and multihead self-attention mechanisms. The

https://doi.org/10.1021/acsomega.3c03572
ACS Omega 2023, 8, 30421−30431

Table 4. Influence of Patch Size and Number of Heads on the Classification Performance in Both Data Sets[a]

patch_size	head	Top_1 Acc.	parameters (M)	FLOPs (G)
(35, 44, 22)[b]	(1, 2, 4)	0.9860 (0.9786−0.9902)	0.9476	0.0533
(35, 22, 11)[b]	(2, 4, 8)	0.9844 (0.9766−0.9922)	1.2916	0.0754
(25, 8, 4)[b]	(3, 6, 12)	0.9856 (0.9728−0.9922)	1.6357	0.0956
(5, 4, 2)[b]	(4, 8, 16)	0.9831 (0.9786−0.9883)	1.9798	0.1158
(30, 20, 10)[c]	(1, 2, 4)	0.9848 (0.9780−0.9900)	0.9512	0.0378
(15, 10, 5)[c]	(2, 4, 8)	0.9850 (0.9760−0.9920)	1.2952	0.0515
(12, 6, 3)[c]	(3, 6, 12)	0.9834 (0.9740−0.9920)	1.6393	0.0653
(5, 3, 2)[c]	(4, 8, 16)	0.9816 (0.9760−0.9900)	1.9834	0.0790

[a]The error range of the 10-fold cross-validation is indicated in brackets. [b]Reflection data sets. [c]Transmission data sets.

proposed method achieves better classification performance and more comprehensive feature extraction by comparing pure one-dimensional convolution and multihead self-attention neural networks. The average identification accuracy of blood and semen reflection Raman spectra reaches 98.6%, and the average identification accuracy of blood and semen transmission Raman spectra reaches 98.5%. Finally, the Raman ConvMSANet network structure can be applied not only to species blood and semen identification analysis but also to other biological fields.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.3c03572.

Sample size of the data set, training parameter details for Raman ConvMSANet on the transmission data set, and classification report of the test samples for Raman ConvMSANet on both the reflection and transmission data sets. Code: https://github.com/pengjuRen99/Raman-ConvMSANet (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Ri-gui Zhou − *College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China*; Email: rgzhou@shmtu.edu.cn

Authors

Pengju Ren − *College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China*; orcid.org/0000-0003-1041-762X

Yaochong Li − *College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China*

Shengjun Xiong − *HT-NOVA Co.,Ltd, Beijing 100621, China*

Bing Han − *National Engineering Research Center of Ship & Shipping Control System, Shanghai Ship and Shipping Research Institute Co.,Ltd, Shanghai 200135, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.3c03572

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Key R&D Plan under Grant No. 2021YFF0601200, 2021YFF0601204 and 2021YFF0601202.

■ REFERENCES

(1) Ripple, W. J.; Abernethy, K.; Betts, M. G.; Chapron, G.; Dirzo, R.; Galetti, M.; Levi, T.; Lindsey, P. A.; Macdonald, D. W.; Machovina, B.; Newsome, T. M.; Peres, C. A.; Wallach, A. D.; Wolf, C.; Young, H. Bushmeat hunting and extinction risk to the world's mammals. *R. Soc. Open Sci.* **2016**, *3*, 160498.

(2) Maloney, T.; Phelan, R.; Simmons, N. Saving the horseshoe crab: A synthetic alternative to horseshoe crab blood for endotoxin detection. *PLoS Biol.* **2018**, *16*, No. e2006607.

(3) Gui, J.; Zhu, Z. Molecular basis and genetic improvement of economically important traits in aquaculture animals. *Chin. Sci. Bull.* **2012**, *57*, 1751−1760.

(4) Inouel, H.; Takabe, F.; Takenaka, O.; Iwasa, M.; Maeno, Y. Species identification of blood and bloodstains by high-performance liquid chromatography. *J. Leg. Med.* **1990**, *104*, 9−12.

(5) Espinoza, E. O.; Lindley, N. C.; Gordon, K. M.; Ekhoff, J. A.; Kirms, M. A. Electrospray Ionization Mass Spectrometric Analysis of Blood for Differentiation of Species. *Anal. Biochem.* **1999**, *268*, 252−261.

(6) Seok, A. E.; Lee, J.; Lee, Y.-R.; Lee, Y.-J.; Kim, H.-J.; Ihm, C.; Sung, H. J.; Hyun, S. H.; Kang, H.-G. Estimation of Age of Bloodstains by Mass-Spectrometry: A Metabolomic Approach. *Anal. Chem.* **2018**, *90*, 12431−12441.

(7) Sauer, E.; Reinke, A.-K.; Courts, C. Differentiation of five body fluids from forensic samples by expression analysis of four microRNAs using quantitative PCR. *Forensic Sci. Int.: Genet.* **2016**, *22*, 89−99.

(8) Suwa, N.; Ikegaya, H.; Takasaka, T.; Nishigaki, K.; Sakurada, K. Human blood identification using the genome profiling method. *Leg. Med.* **2012**, *14*, 121−125.

(9) Xu, M.-L.; Gao, Y.; Han, X. X.; Zhao, B. Detection of Pesticide Residues in Food Using Surface-Enhanced Raman Spectroscopy: A Review. *J. Agric. Food Chem.* **2017**, *65*, 6719−6726.

(10) Craig, A. P.; Franca, A. S.; Irudayaraj, J. Surface-Enhanced Raman Spectroscopy Applied to Food Safety. *Annu. Rev. Food Sci. Technol.* **2013**, *4*, 369−380.

(11) Araújo, D. C.; Veloso, A. A.; de Oliveira Filho, R. S.; Giraud, M.-N.; Raniero, L. J.; Ferreira, L. M.; Bitar, R. A. Finding reduced Raman spectroscopy fingerprint of skin samples for melanoma diagnosis through machine learning. *Artif. Intell. Med.* **2021**, *120*, 102161.

(12) Ma, D.; Shang, L.; Tang, J.; Bao, Y.; Fu, J.; Yin, J. Classifying breast cancer tissue by Raman spectroscopy with one-dimensional convolutional neural network. *Spectrochim. Acta, Part A* **2021**, *256*, 119732.

(13) Bratchenko, I. A.; Bratchenko, L. A.; Khristoforova, Y. A.; Moryatov, A. A.; Kozlov, S. V.; Zakharov, V. P. Classification of skin cancer using convolutional neural networks analysis of Raman spectra. *Comput. Methods Progr. Biomed.* **2022**, *219*, 106755.

(14) Chen, K.; Zhang, H.; Ibrahim, U.-K.; Xue, W.; Liu, H.; Guo, A. The quantitative assessment of coke morphology based on the Raman spectroscopic characterization of serial petroleum cokes. *Fuel* **2019**, *246*, 60−68.

(15) Doty, K. C.; Lednev, I. K. Differentiation of human blood from animal blood using Raman spectroscopy: A survey of forensically relevant species. *Forensic Sci. Int.* **2018**, *282*, 204−210.

(16) Bian, H.; Wang, P.; Wang, N.; Tian, Y.; Bai, P.; Jiang, H.; Gao, J. Dual-model analysis for improving the discrimination performance of human and nonhuman blood based on Raman spectroscopy. *Biomed. Opt Express* **2018**, *9*, 3512.

(17) Bai, P.; Wang, J.; Yin, H.; Tian, Y.; Yao, W.; Gao, J. Discrimination of Human and Nonhuman Blood by Raman Spectroscopy and Partial Least Squares Discriminant Analysis. *Anal. Lett.* **2017**, *50*, 379−388.
(18) Virkler, K.; Lednev, I. K. Blood Species Identification for Forensic Purposes Using Raman Spectroscopy Combined with Advanced Statistical Analysis. *Anal. Chem.* **2009**, *81*, 7773−7777.
(19) Fujihara, J.; Fujita, Y.; Yamamoto, T.; Nishimoto, N.; Kimura-Kataoka, K.; Kurata, S.; Takinami, Y.; Yasuda, T.; Takeshita, H. Blood identification and discrimination between human and nonhuman blood using portable Raman spectroscopy. *J. Leg. Med.* **2017**, *131*, 319−322.
(20) Wang, H.; Fang, P.; Yan, X.; Zhou, Y.; Cheng, Y.; Yao, L.; Jia, J.; He, J.; Wan, X. Study on the Raman spectral characteristics of dynamic and static blood and its application in species identification. *J. Photochem. Photobiol., B* **2022**, *232*, 112478.
(21) He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. 2015, arXiv:1512.03385. http://arxiv.org/abs/1512.03385.
(22) Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. 2015, arXiv:1409.1556. http://arxiv.org/abs/1409.1556.
(23) Graves, A. Generating Sequences With Recurrent Neural Networks. 2014, arXiv:1308.0850. http://arxiv.org/abs/1308.0850.
(24) Dong, J.; Hong, M.; Xu, Y.; Zheng, X. A practical convolutional neural network model for discriminating Raman spectra of human and animal blood. *J. Chemom.* **2019**, *33*, No. e3184.
(25) Huang, S.; Wang, P.; Tian, Y.; Bai, P.; Chen, D.; Wang, C.; Chen, J.; Liu, Z.; Zheng, J.; Yao, W.; Li, J.; Gao, J. Blood species identification based on deep learning analysis of Raman spectra. *Biomed. Opt Express* **2019**, *10*, 6129.
(26) Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. 2017, arXiv:1706.03762. http://arxiv.org/abs/1706.03762.
(27) Liu, B.; Liu, K.; Qi, X.; Zhang, W.; Li, B. Classification of deep-sea cold seep bacteria by transformer combined with Raman spectroscopy. *Sci. Rep.* **2023**, *13*, 3240.
(28) Wang, C.; Wan, Y.; Su, Y.; Cai, Y.; Xiong, S.; Yuan, D.; Xia, Z.; Zhu, J. An Expedient SERS Strip Tactic for Rapid On-Site Detection with Long-Time Sensitivity and Repeatability. *Adv. Mater. Sci. Eng.* **2021**, *2021*, 1−8.
(29) Gorry, P. A. General least-squares smoothing and differentiation by the convolution (Savitzky-Golay) method. *Anal. Chem.* **1990**, *62*, 570−573.
(30) Zhang, Z.-M.; Chen, S.; Liang, Y.-Z. Baseline correction using adaptive iteratively reweighted penalized least squares. *Analyst* **2010**, *135*, 1138.
(31) Park, N.; Kim, S. How Do Vision Transformers Work?. 2022, arXiv:2202.06709. http://arxiv.org/abs/2202.06709.
(32) Al-Shaebi, Z.; Uysal Ciloglu, F.; Nasser, M.; Aydin, O. Highly Accurate Identification of Bacteria's Antibiotic Resistance Based on Raman Spectroscopy and U-Net Deep Learning Algorithms. *ACS Omega* **2022**, *7*, 29443−29451.
(33) Chen, X.; Xie, L.; He, Y.; Guan, T.; Zhou, X.; Wang, B.; Feng, G.; Yu, H.; Ji, Y. Fast and accurate decoding of Raman spectra-encoded suspension arrays using deep learning. *Analyst* **2019**, *144*, 4312−4319.
(34) Sang, X.; Zhou, R.-G.; Li, Y.; Xiong, S. One-Dimensional Deep Convolutional Neural Network for Mineral Classification from Raman Spectroscopy. *Neural Process. Lett.* **2022**, *54*, 677−690.

https://doi.org/10.1021/acsomega.3c03572
ACS Omega 2023, 8, 30421−30431