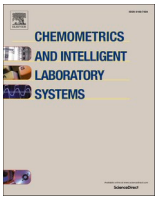
Chemometrics and Intelligent Laboratory Systems 234 (2023) 104757

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemometrics



ConInceDeep: A novel deep learning method for component identification of mixture based on Raman spectroscopy

Ziyan Zhao, Zhenfang Liu, Mingqiang Ji, Xin Zhao, Qibing Zhu, Min Huang *

Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi, 214122, China

ARTICLE INFO

Keywords:
Raman spectroscopy
Component identification of mixture
Convolutional neural network
Continuous wavelet transform
Inception module

ABSTRACT

For mixture component identification, the methods based on deep learning are becoming prevalent due to their end-to-end characteristic, being completely data-driven and reducing the demand for prior knowledge. The most used CNN model currently, however, relies only on the original one-dimensional (1D) spectral data. When faced with complex mixtures, the features of overlapping peaks and weak peaks are probably not sufficient mined by 1D CNN model, which would reduce the identification accuracy. Thus, a novel deep learning method entitled ConInceDeep for component identification of mixture by Raman spectroscopy was proposed in this study. The ConInceDeep firstly performed continuous wavelet transform on 1D spectral data using the Lorentz4 wavelet to fully reveal the detailed characteristics of weak and overlapping spectral peaks in the mixture and then employed Inception modules consisting of multiple-size kernels to construct two-dimensional CNN model to improve the model's adaptability to different Raman peaks. The proposed ConInceDeep model was trained and validated entirely by the spectra of virtual mixtures which were generated by simple manipulation of substances' spectra in the database. Three mixture datasets were used for verifying the identification performance of the ConInceDeep, including 191 liquid mixtures, 33 powder mixtures and two kinds of real samples. In liquid mixture set, ConInceDeep was compared with other three experimental methods, and achieved an increase of *Acc* from 90.32% to 96.60% compared with the traditional 1D-CNN. In powder mixtures set, ConInceDeep reached 99.66% *Acc* and 0.35% *FPR* by 54 models. Additionally, it was also applied to real samples to demonstrate its value. Summarily, the ConInceDeep provides a new reference for mixture identification based on Raman spectroscopy.

1. Introduction

The peaks in Raman spectrum of the measured substance corresponds to specific substance molecules, so Raman spectrum has fingerprint characteristic. Based on it, Raman spectroscopy can be used for the identification of specific component in mixtures. Currently, Raman technology has become a powerful and important tool in the field of mixture component identification due to its advantages of simple sample preparation, quick analysis speed, and non-destructive operation. For mixture component identification by Raman spectroscopy, database searching and matching algorithm was one of the most common methods [1]. This method extracts the typical Raman spectral peak features of the mixture to be identified, calculates the similarity between the extracted features and the spectral peak features of the substances stored in the database, and determines the substances contained in the mixture through the similarity value [2–5]. While the performance of this method relies on the detection of spectral peaks in the mixture to be identified and the similarity calculation of features between the detected mixture and pure substances in the database. The overlapping and weak peaks in the mixture spectrum will bring difficulties to the detection of typical peaks in the mixture, thus affecting the accuracy of the subsequent similarity calculation. In addition, the interaction of some coexisting substances in mixtures and the variation of the measurement environment possibly bring spectral distortion, such as changes in peak position and peak shape, which will also cause errors in the similarity calculation result.

With the development of deep learning theory, the method introducing deep learning in the field of Raman spectroscopy data analysis has received extensive attention [6–8]. Deep learning gradually transforms the initial low-level feature representation into high-level feature representation through multi-layer processing, then completes complex classification, regression and other learning tasks based on the transformed high-level features. In the field of Raman spectroscopy, deep learning is an end-to-end learning method without pre-processing

* Corresponding author. School of Internet of Things, Jiangnan University, 1800 Lihu Avenue, Wuxi, Jiangsu Province, 214122, China.
E-mail address: huangmzqb@163.com (M. Huang).

https://doi.org/10.1016/j.chemolab.2023.104757
Received 5 July 2022; Received in revised form 10 December 2022; Accepted 9 January 2023
Available online 11 January 2023

methods and feature extraction based on numerical calculation, and it can learn features relevant to the task from a large amount of spectral data automatically only relying on input and output, and can solve problems which are difficult to deal with by traditional methods such as high dimension. Fully-connected neural networks and convolutional neural networks are two types of models in deep neural networks, which are typical structures for implementing deep learning. The existing research results show that the fully-connected neural network can be used to detect the components of the mixture. Isaex [9] et al. used multilayer perceptron in the artificial neural network system to detect the ethanol concentration in a water-ethanol mixed solution and determine the concentration of harmful impurities in vodka. Zheng [10] et al. proposed a method of screening and diagnosing echinococcosis with back propagation (BP) neural network, and the simulation results showed that the overall accuracy of echinococcosis diagnosis reached 94.69%. The method based on the fully-connected neural network gets rid of the peak feature extraction stage based on numerical calculation, but a large number of parameters need to be trained in the network, and there is a risk of overfitting in the model when there are few training samples.

Currently, convolutional neural network (CNN) has been an accepted architecture displacing fully connected neural network due to its sparse connectivity and parameter sharing [11]. The introduction of convolutional kernels in CNN helps to extract complex spectral features automatically at a low cost. Huang [12] et al. proposed a deep learning model based on a 1D convolutional neural network to detect the species of blood containing 20 kinds, and the blood samples to be identified showed that the average recognition accuracy was up to 97.33%. Liu [13] et al. proposed an improved 1D convolutional neural network based on LeNet for Raman spectrum recognition, which used a pyramid convolutional layer and two fully connected layers. The model was trained without data preprocessing, and the RRUFF spectral database containing various mineral data was used to predict the performance of the model, which achieved superior classification results. Fan [14] et al. proposed a new approach entitled deep learning-based component identification (DeepCID). 1D convolutional neural network models were established to identify one single component in the mixture to be tested. DeepCID showed higher accuracy and lower false positive rate compared with logistic regression, k nearest neighbor, random forest, and back propagation artificial neural network models with L1 regularization. At present, the existing mixture component identification methods based on CNN mainly rely on 1D data, which is spectral data itself. However, when the multicomponent mixture has similar molecules or functional groups, there will be serious overlapping phenomena in the Raman peaks region, and it is quite difficult for 1D CNN to extract and mine features of covered peaks, as well as some weak peaks, which possibly has a negative effect on component identification. Moreover, few types of research related to CNN in Raman spectroscopy consider the local characteristics of Raman peaks and the adaptability of networks to different substances. Therefore, the ability of the models to analyze spectra remains to be improved. How to design a neural network structure with high stability and accuracy according to the characteristics of Raman spectrum is still a problem to be solved.

Wavelet analysis has been proven to be a successful tool in multiscale analysis of spectral peaks and widely used in the field of Raman spectroscopy, such as feature extraction [15], peak detection [16,17], deconvolution of overlapping peaks [18,19], identifying and separating the signal from noise [20]. And continuous wavelet transform (CWT) is one of the most important methods to decompose 1D spectral signals into different scales, so as to facilitate the separation and recognition of weak and overlapping spectral peaks. Inspired by this, CWT was introduced in this study to generate two-dimensional (2D) CWT coefficient matrixes to carry our multiscale analysis of spectral peaks, and then combined with the convolutional operation to extract features further, which potentially improve the identification result when faced with complex mixtures. By setting a large range of scale, the 2D CWT coefficient matrix could reflect both morphologic and intensity characteristics of Raman spectral peaks. Small scale contributes to the identification of overlapping and low-intensity peaks, and large scale pays more attention to the distinct peaks with wider full width at half maximum (FWHM) and greater intensity. Correspondingly, 2D CNN models were established by applying the CWT coefficient matrixes as input. Besides, the Inception module with parallel convolutional kernels in different sizes was used to diversify the receptive field and improve the adaptability of the network to different local spectral features. Moreover, 1*1 convolutional kernel cascading other kernels with different sizes has a positive impact on cross-channel feature fusion and reduces the number of parameters in the network, which helps to save computing resources and decrease the risk of overfitting. This proposed method combining CWT and Inception was named as ConInceDeep. To verify the effectiveness of the ConInceDeep method, we designed and compared the performance of four CNN models as the ablation experiments on liquid mixture set. Solid powder mixture set and real sample set were also set up to test ConInceDeep and provide widely applicable and feasible proof further. Besides, we also used Grad-CAM to visualize the key spectral regions concerned by model, which directly demonstrated the improvement of the ConInceDeep.

2. Materials and methods

2.1. Dataset description

2.1.1. Experimental data acquisition

In this study, Raman spectra of 100 pure substances, 191 liquid mixtures, 33 powder mixtures and two kinds of real samples were collected by a handheld Raman spectrometer (Zolix Finder Edge, Beijing Zolix optical instrument company, China). This handheld Raman spectrometer adopts a 785 nm laser as excitation. The collected spectra span a Raman shift range of 0–2700 cm^{-1} and the spectral resolution is 8–10 cm^{-1}. During spectra acquisition, the laser power was adjustable within 0–350 mW, and the integration time of the spectrometer was set from 1.5 s to 2.5 s. The collected Raman spectra contain the most Raman peaks in the Raman shift range of 200 cm^{-1}–1800 cm^{-1}, so the spectral data in this region were selected.

Additionally, the spectra of mixtures and substances were obtained at uncontrolled room temperature and the selection of instrument parameters was diverse possibly according to different samples. Therefore, the spectra of all samples in this study were preprocessed by maximum normalization and interpolation to standardize the spectra. The spectra of these 100 substances, which are common chemicals, were used to construct virtual mixture spectra which were divided into training set and validation set to build models (see the detailed procedure in section 2.1.4).

2.1.2. Liquid mixture set

191 liquid mixtures were used to evaluate the performance of the models as the first test set. These mixtures were mixed with six substances (namely, diethyl malonate, acetone, acetonitrile, ethanol, diacetone alcohol, and cyclohexane) and included binary mixtures, ternary mixtures, quaternary and quinary mixtures, and the volume ratio of each component varied from 0.1 to 0.9. The detailed components of the mixtures above are shown in Table 1. The Raman spectra of six substances to make up the mixture are shown in Fig. 1. As can be seen from Fig. 1, diethyl malonate and diacetone alcohol have a large number of Raman peaks and their peaks overlap seriously, while acetone, acetonitrile, ethanol, and cyclohexane have fewer Raman peaks, which are distinct relatively. Moreover, the Raman shift of some peaks in different substances is close so that their local characteristics are similar, which increases the difficulty of identifying the component in mixtures(for example, diethyl malonate, ethanol, diacetone alcohol and cyclohexane, all have a characteristic Raman peak around 1450 cm^{-1}). Therefore, the substances to be identified in the mixtures are representative to cover

Table 1
The composition of liquid mixtures to be identified.

	Ethanol	Acetonitrile	Acetone	Cyclohexane	Diacetone alcohol	Diethyl malonate
Binary (36)	○	○	●	○	○	●
	●	○	○	○	●	○
	○	●	●	○	○	○
	○	●	○	●	○	○
Ternary (81)	○	○	●	○	●	●
	●	○	●	○	●	○
	●	○	●	●	○	○
	●	○	○	●	○	●
	○	●	●	○	●	○
	○	●	●	●	○	○
Quaternary (54)	●	○	●	○	●	●
	●	●	●	○	○	●
	●	●	●	○	●	○
	●	●	●	●	○	○
Quinary (20)	●	●	●	●	●	○

(The number following 'Binary' represents the quantity of binary mixtures, and others are the same; ● means containing the corresponding substance; ○ means not containing the corresponding substance.).

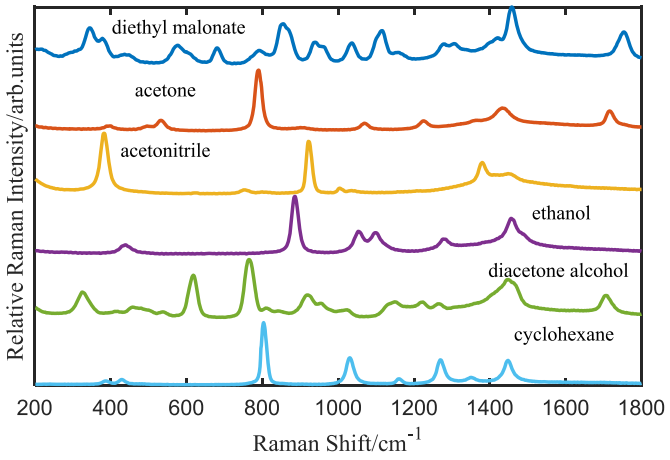


Fig. 1. The spectra of six liquid substances that need to be identified from the mixtures.

more cases in the real world, and the task of identifying these six substances from mixtures helps to validate the effectiveness of the proposed method.

2.1.3. Powder mixture set and real sample set

There were 33 powder mixture samples as the second test set. These powder mixture samples were mixed by four pure substances, namely, glutamic acid, aspartic acid, histidine, and glycine. These solid amino acids are composed of similar basic amino and acid carboxyl, so that some of their spectral peaks have semblable positions and shapes which led to a low degree of differentiation for these solid mixture samples in this set. The powder mixtures also included 2-component mixtures, 3-component mixtures and 4-component mixtures, and the volume ratio of each component varied from 0.1 to 0.8. The detailed components and volume ratios of the powder mixtures are shown in Table 2.

Besides, the mouthwash and baking powder, as the representative of liquid mixture and powder mixture respectively, were selected to demonstrate the feasibility of the models established by ConInceDeep applied to real samples. Mouthwash, also known as oral rinses, is mainly used to clean the oral cavity and freshen the breath. As a common ingredient in traditional mouthwashes, ethanol helps to dissolve certain organic flavor components and also contributes to cleaning and bactericidal effects. However, ethanol is slightly irritating and not suitable for people with sensitive mouths. At present, there are both traditional mouthwashes available in market which contain ethanol, and mouthwashes that do not add ethanol under consideration of mouthfeel. The

Table 2
The compositions and volume ratios of powder mixture samples.

	Volume ratios	Components
2-component (18)	0.6:0.4/ 0.7:0.3/ 0.8:0.2	1. glycine; aspartic acid 2. glycine; histidine 3. histidine; aspartic acid 4. glutamic acid; aspartic acid 5. glutamic acid; glycine 6. glutamic acid; histidine
3-component (12)	0.1:0.2:0.7/ 0.2:0.3:0.5/ 0.3:0.3:0.4	7. aspartic acid; glycine; glutamic acid 8. histidine; aspartic acid; glycine 9. histidine; aspartic acid; glutamic acid 10. histidine; glutamic acid; glycine
4-component (3)	0.25:0.25:0.25:0.25/ 0.2:0.2:0.2:0.4/ 0.2:0.2:0.3:0.3/	11. histidine; aspartic acid; glycine; glutamic acid

(The 2-component mixtures have three volume ratios and six kinds of compositions, hence there are a total of 3*6 = 18 2-component mixtures. Similarly, there are a total of 3*4 = 12 3-component mixtures and 3*1 = 3 4-component mixtures).

two kinds of mouthwashes above from LISTERINE were purchased from the local supermarket, which were used as real samples to test the ability of ConInceDeep method to identify the presence of ethanol and distinguish two types of mouthwash. In addition, baking powder is a common compound leavening agent, which can make food achieve fermenting effect and rise with a lot of air pockets after baking, and often used in the production of cake, bread, steamed bread, etc. It is mainly through the reaction of acid and alkaline substances to produce carbon dioxide to play the role of expanding and softening food. The common baking powders choose disodium pyrophosphate as the acidic substance and sodium bicarbonate as the alkaline substance. We also bought Angel brand's aluminum-free double acting baking powder from the local supermarket. According to the ingredient list, the content of dihydrogen disodium pyrophosphate is less than or equal to 40%, and the content of sodium bicarbonate is less than or equal to 35%. Sodium bicarbonate is the one with lower relative content. Thus, we also built a ConInceDeep model of sodium bicarbonate to see if it could detect the presence of sodium bicarbonate.

2.1.4. Data augmentation for training and validation

The model based on deep learning requires a large number of training samples to meet the requirements of feature learning. In view of the fact that it is difficult to collect quite a number of mixture samples in practical application, data augmentation was introduced to reduce sampling costs and eliminate random errors during spectrum acquisition.

In this study, a total of 20000 virtual mixture samples were generated for each of the specific pure substances, including 10,000 positive samples(containing the chosen substance) and 10,000 negative samples (not containing the chosen substance). With reference to the literature [21], which discussed the impact on the recognition results choosing different number of components in augmented mixture spectra, it is reasonable to set the max number of components as three in virtual mixture samples. Among them, there were 3750 binary mixtures and 6250 ternary mixtures in positive and negative samples, respectively. As for components selection and concentration design of virtual mixtures, in positive samples, the components except the chosen substance were selected at random in the database of 100 substances and the concentration of the chosen substance was changed from 2% to 98% with a span of 4%, while the components of negative samples were randomly determined during database with random concentration ratios from 2% to 98%. As for the spectra generation of virtual mixtures, they were obtained by linear superposition of pure components' spectra and then processed by maximum normalization. Finally, the virtual mixtures were added labels, shuffled and saved as training set and validation set in the ratio of 8:2. It is worth noting that the training and validation samples in this study were composed of virtual mixtures completely.

2.2. ConInceDeep model architecture

Overall, the ConInceDeep in this study for identifying one specific component in the mixture, firstly expanded the 1D Raman spectrum of the mixture to a 2D coefficient matrix as input data using CWT. Then, 2D data was processed by two Inception modules and two Max-pools for feature extraction. Finally, fully connected layers made a judgment whether the input mixture contained the specific substance concerned by the model or not. It is remarkable that in model training and validation, virtual mixtures obtained by data augmentation were used as the input data, while liquid mixtures, powder mixtures and real mixtures were utilized in model testing. And the flowchart of ConInceDeep is shown in Fig. 2.

2.2.1. Continuous wavelet transform

Raman spectrum has significantly localized features with characteristics of peak signal, which is a typical non-stationary signal. Common multiresolution signal analysis methods such as Fourier transform perform poorly in Raman technology due to fixed-size windows and invariable frequency resolution, and they are only suitable for stationary signals with small frequency fluctuation [22–25]. Continuous wavelet transform (CWT) is a common non-stationary signal processing method, which was introduced in this study to provide as detailed spectral information as possible, including weak and overlapping peaks which are difficult to reflect in 1D spectra signals. Different from the Fourier transform, CWT provides a localized window in the spectrum domain that varies with scale. The transformed signal obtains frequency information and locates the corresponding spectrum domain position at the same time, which corresponds to the FWHM of Raman peaks as well as the Raman shift.

CWT can be regarded as the convolutional operation of the Raman signal and the wavelet function in essence. When the scale of wavelet function is determined, wavelet coefficients can be obtained by sliding the wavelet function over the spectrum domain. A higher coefficient represents a better match between the Raman signal and the mother wavelet. In addition, the width of wavelet function changes each time the scale is adjusted, so the information of the Raman signal with different frequencies can be detected in CWT. The above means that

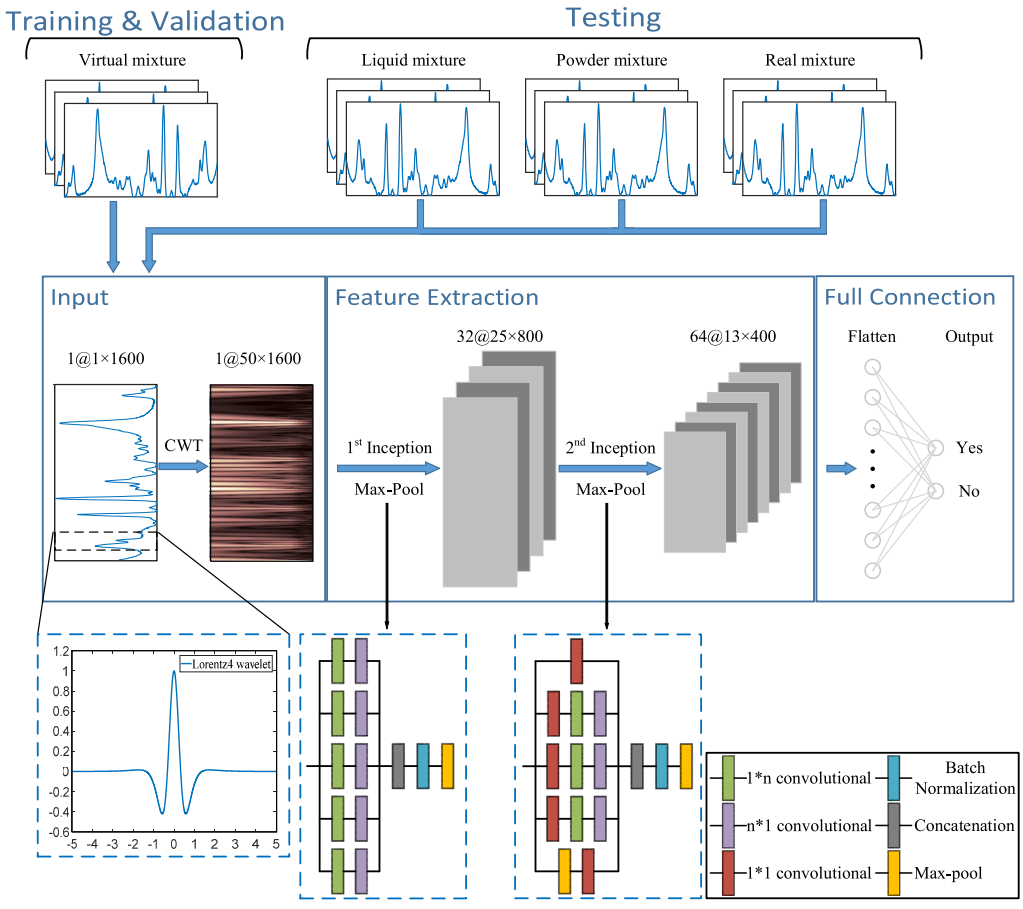


Fig. 2. The flowchart of ConInceDeep and the architecture of the corresponding 2D CNN model.

CWT coefficients could reflect more details of Raman spectrum peaks with different FWHM. The formula of CWT is as follows:

$$WT(a,\tau)=\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} f(t)*\varphi\left(\frac{t-\tau}{a}\right)dt \quad (1)$$

where *WT* is the 2D matrix of CWT coefficient, $f(t)$ is the Raman signal, $\varphi(t)$ is the mother wavelet function, a represents the scale and τ represents the translation amount. Wavelet has the functions of expansion and translation, which are determined by a and τ respectively.

During the process of CWT, it is very important to select the appropriate wavelet base function for different signals, which has a positive impact on the accurate analysis of signals. In general, the Raman raw spectrum is represented by a Voigt function profile with smaller linewidth [19]. Traditional wavelets like Gaus4 [19] and Mexican hat [16] were often selected as the wavelet functions in previous research on Raman spectra for peak detection, which could be attributed to their similar line profile with Raman peaks, including symmetrical shape and one major peak. The linewidths of the two functions above, however, are larger than the Voigt function, so they are slightly different from the raw spectrum. Compared with the traditional Gaus4 and Mexican hat wavelet functions, the linewidth of the Lorentz4 [19] function is smaller, and the line shape is closer to the intrinsic line shape contour of the Raman spectral peak. Therefore, the Lorentz4 wavelet base function is more suitable for the analysis of Raman spectral signals. It is defined as:

$$\varphi(t)=\frac{24\cdot(5\cdot t^4-10\cdot t^2+1)}{(t^2+1)^5} \quad (2)$$

To illustrate the suitability of Lorentz4 wavelet in Raman spectrum, we simulated two peaks with different width and intensity using Voigt function profile, as shown in Fig. 3(a), and contrasted the shape difference between the Raman peak and each wavelet in Fig. 3(b), (c) and (d). Compared with the other two wavelets, the FWHM of Lorentz4 was smaller at the same scale, which was more closed to the intrinsic line of the Raman peak, so it resulted in a good performance in the analysis of Raman spectra. In addition, we could also draw the following conclusion from Fig. 3, which stated that the different scales of wavelet helped to adapt itself to Raman peaks with different width and intensity. This is mainly because the half-width of wavelet function will change with the selection of scale, small scale provides a small half-width to wavelet which is closer to the weak peak with small FHWM and intensity, and vice versa.

In this study, the 1D spectral data of virtual mixtures and all mixtures in test sets were subjected to CWT using Lorentz4 as wavelet function, and the wavelet scale was set from 5 to 54. The wavelet coefficients corresponding to small scales, such as 5–20, mainly play a role in detecting weak peaks, which have high coefficients at small scale, and the coefficients obtained by large scale detect strong peaks crucially [17]. In conclusion, 50*1600-size 2D CWT coefficient matrixes were generated(1600 is set as the range of spectral wavenumber) for model training and testing.

2.2.2. Inception module

In general, the convolutional layers are mainly used for feature extraction of input data, which is the most important part of CNN model, and the most effective feature for prediction tasks will be extracted after training and learning. The convolutional operation in these layers is a kind of similarity calculation between convolutional kernel and data essentially. The weights and biases of the convolution kernel are parameters that need to be learned, while the size of kernels is determined manually, which is often fixed in general operations. However, in Raman spectrum, the FWHM of spectral peaks is often diverse, and the distance between peaks is also different, so a single-size convolutional kernel may not fully mine the features in the spectrum. To ease this problem, the Inception module [26], which is a sparse network structure, was introduced in this study to improve the adaptation of neural networks to local Raman spectra features with the precondition of high efficiency using computing resources and eliminate the influence of kernels' size on the prediction result.

Referring to two key ideas of the Inception module, two similar Inception modules were designed in ConInceDeep to replace traditional convolutional layers in 2D CNN models, the schematic structures of

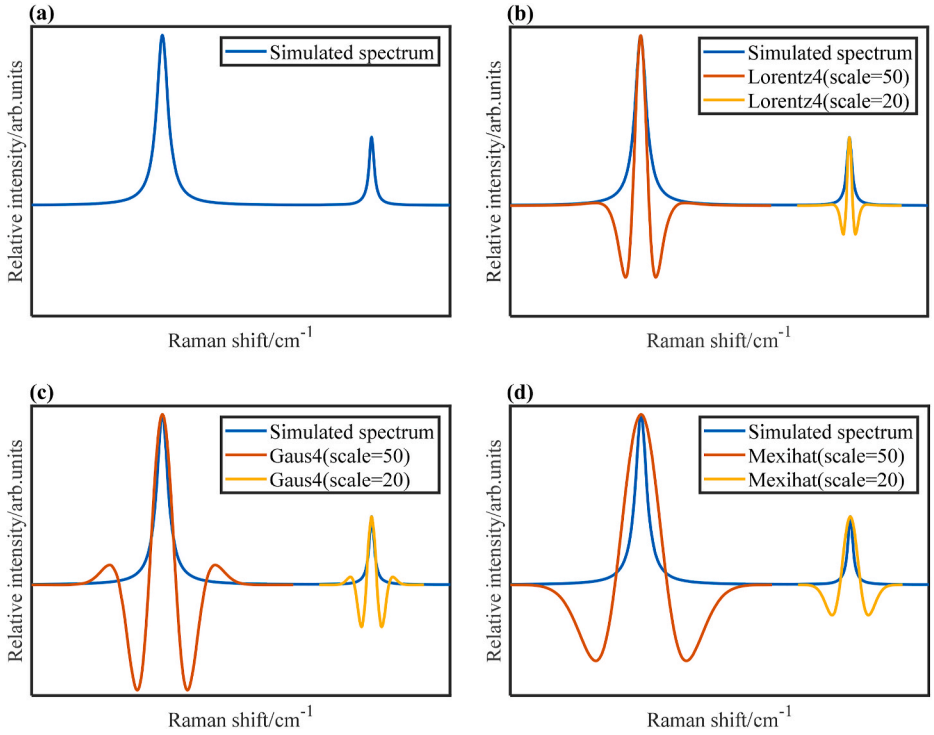


Fig. 3. The contrast diagrams of three wavelets and simulated spectrum. (a)The simulated spectrum with two Raman peaks; (b)Lorentz4 wavelet and simulated spectrum; (c)Gaus4 wavelet and simulated spectrum; (d)Mexican hat wavelet and simulated spectrum.

which are shown in Fig. 4.

One of the two ideas attributes to the introduction of a 1*1 convolutional kernel. It allows a cascade of 1*1 convolutional kernel and different kernels in another size like 1*3, 1*5, such as the stacking of the first and the second convolutional operation in the middle three branches connected after input in Fig. 4(b). And the number of 1*1 convolutional kernel is set smaller than the previous channels. Therefore, it helps to reduce network parameters [27] and realize cross-channel feature extraction [28] under the premise of low cost, and the risk of overfitting could be reduced, so as to improve the expression ability of the models.

The other key idea is that convolutional structures are designed to possess parallel kernels with different sizes replacing fixed-size kernels, so after the input data is processed by those Inception layers, the feature maps obtained will have different sizes of receptive fields [29]. The difference between spectra is mainly reflected in the Raman spectral peaks region, which is a local characteristic with a different range of wavenumber. Therefore, diverse receptive fields contribute to obtaining spectral features of the regions of different sizes from the original spectra, which serves the characteristic of Raman technology and increases the adaptability of networks to different spectral peak widths.

In ConInceDeep, since the input of the 1st Inception module was original 2D data, 1*1 convolutional operation could not change the shape of the data, as well as the number of channels, thus it was not used as seen in Fig. 4(a). And the 1st Inception module adopted five convolution kernels of different sizes. In the 2nd Inception module, 1*1 convolutional kernel came into play and four different convolutional kernels were employed as seen in Fig. 4(b). Moreover, as for the selection of the size of kernels in ConInceDeep, in the literature [30] proposing Inception-v3, convolutional kernels of 3, 5, and 7 were used. However, considering that Inception modules here were mainly used to extract the features of Raman spectral data, the spectral peaks in multiple regions were of our attention, and the widths of spectral peaks were different. So, we added two more groups of convolutional kernels, the sizes of which were 9 and 11 respectively, in an attempt to increase the adaptability of the network to different spectral peaks. Additionally, ConInceDeep also used asymmetric convolutions referring to Inception-v3, which is that single 1*n convolutional operation is followed by single n*1 convolutional operation.

2.2.3. 2D Convolutional neural network

After the CWT, which transformed the 1D spectral data to 2D wavelet coefficients, the input of CNN changed accordingly to matrixes with a size of 50*1600, so that 2D CNN was constructed. The network consisted of an input layer, two convolutional blocks, the 1st Inception and the 2nd Inception respectively, two Max pool layers, a flatten layer and an output layer. Fig. 2 shows the detailed graphic description of the 2D CNN model in ConInceDeep.

In this experiment, the stride of each Max-pool operation was set to 2, except in the 2nd Inception module set to 1. The rectified linear unit (ReLU) was used as the activation function in all convolutional operations. SoftMax function was set as the output of the last layer of the classifier, and the cross-entropy function was applied as the loss function of the network. Dropout, Batch Normalization (BN) were introduced to act on the training process of the network. Adaptive Moment Estimation (ADAM) was chosen as the optimization algorithm for the weights of network. The Dropout rate was 0.5, the training rounds were set to 200, the Batch size of training samples was 64, and the learning rate was set to 0.0001 after optimization. The model reaching higher accuracy on validation set was saved as the final prediction model. This process occurred in each epoch, so the model might be constantly updated until the end of training. Furthermore, to prevent overfitting due to excess epochs and save computing resources, the training would end when the model achieved same accuracy on the validation set for 5 epochs in a row. Detailed information about the parameters of the ConInceDeep model including the number of filters in the convolutional operation, the size of pool operation and neuron number in flatten layer and dense are listed in Table 3.

2.3. Model evaluation

Accuracy(Acc), Precision(Pre), Recall(Rec), F1_Score and *FPR* were used to evaluate the performance of the models constructed in this study, which are shown in Eq. (3), Eq. (4), Eq. (5), Eq. (6) and Eq. (7).

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$F1\_Score = \frac{2 \times Pre \times Rec}{Pre + Rec} \tag{6}$$

$$FPR = \frac{FP}{TN + FP} \tag{7}$$

where *TP*, *FP*, *TN*, *FN* are true positive (the samples with positive prediction and actual label), false positive (the samples with positive pre-

Table 3
The list of the detailed parameters of ConInceDeep model.

	Size	Numbers	Stride	Output shape
Input	50*1600	–	–	(50,1600,1)
1st Inception	3/5/7/9/11	32	1	(50,1600,32)
MaxPool 1	3*3	–	2	(25,800, 32)
2nd Inception	1/3/5/7	64	1	(25,800, 64)
MaxPool 2	3*3	–	2	(13,400,64)
Flatten	–	332800	–	(332800)
Output	–	2	–	(2)

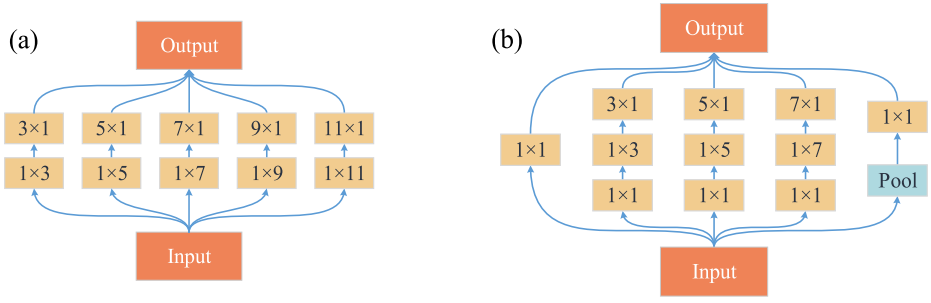


Fig. 4. The schematic structures of Inception modules used in ConInceDeep. (a) The 1st Inception module in 2D CNN; (b) The 2nd Inception module in 2D CNN.(The blocks in this figure represent convolutional operations except input, output and a pooling operation).

diction label and negative actual label), true negative (the samples with negative prediction and actual label) and false negative (the samples with negative prediction and positive actual label) respectively.

Generally, *Acc* illustrates the proportion of samples that are properly identified and a higher value represents the better performance of the model. But in some cases, the evaluation result of *Acc* is possibly affected by unreasonable sample distribution. Hence, *Pre*, *Rec* and *F1_Score* were also selected as indices to increase the comprehensiveness of model performance evaluation and they were used to evaluate the ability to identify a specific substance and distinguish similar components. Besides, *FPR* stands for false positive rate and it will be close to 0% when models achieve good results, which was also selected to appraise the ability of ConInceDeep to exclude the presence of some substances facing unknown mixtures.

2.4. Computing facilities and model training environment

The hardware environment used in this experiment was Intel Core™ i9-10900X @3.70 GHz CPU, 64 GB running memory, NVIDIA GeForce RTX 3090 GPU, and the operating system was Windows10 (64-bit). In the stage of data augmentation, 1D virtual mixture spectra and corresponding 2D CWT coefficient data were created in MATLAB 2020b. Then, Python (version: 3.7) was used to build and test the model, and Tensorflow-gpu (version: 2.4.0) was chosen as the deep learning framework. The codes above are available at https://github.com/Zhao-zzzzyy/ConInceDeep.

3. Results and discussion

3.1. Comparison of four convolutional neural networks

In order to evaluate the positive effect on the introduction of CWT and Inception module separately, the ablation experiments were used so that four CNN structures were designed in this study, which were named as 1D-CNN (see detailed information about the parameters in Table S1), 1D-all-Ince-CNN (Table S2), 2D-Wav-CNN (Table S3) and ConInceDeep respectively. The difference between the first two and the latter two lay in the dimension of data, which meant that 2D-Wav-CNN and ConInceDeep applied 2D CWT coefficient data obtained by CWT to train 2D CNN models instead of 1D spectral data getting 1D CNN models. And the difference between 1D-all-Ince-CNN and 1D-CNN, ConInceDeep and 2D-Wav-CNN both involved that two traditional convolutional layers with the kernels of 1*5 or 5*5 were substituted by Inception modules. The design differences between the above four CNN structures are shown in Table 4. Moreover, the network structure parameters of the other three models were set similarly with ConInceDeep for the fairness of the comparison, as well as the utilization of validation sets for optimization, the selection of activation function, optimization algorithm and so on. It is supposed to be noted that due to time consuming of CWT operation and 2D convolution, ConInceDeep required an average of around 2 h for training in this study, which is much longer than 4 min for 1D-CNN. While for inference, it took less than 16 ms for ConInceDeep model to make prediction on one mixture sample, which was about three times that of 1D-CNN. Although the training and inference time of ConInceDeep is higher than that of 1D-CNN, in practical application, people tend to pay more attention to the inference time of the model, and we think that the inference time of 16 ms of ConIncDeep can meet the demand.

The above four different networks were used to build the identification models and the liquid mixture set was utilized as the test set for evaluation of the four networks for the detection of six specific substances, which is relatively more adequate and meaningful containing most samples. A total of 24 models were established for the six liquid substances. The identification results for 191 mixtures are shown in Table 5.

The results suggested that in a general way, compared with traditional 1D-CNN achieving 90.32% average *Acc* and 90.30% average *F1_Score*, the average *Acc* of 1D-all-Ince-CNN and 2D-Wav-CNN were improved by 1.91% and 5.14% respectively, with *F1_Score* improved by 2.40% and 5.62%. It indicated that the model introducing CWT and Inception separately performed better than the traditional 1D-CNN, proving the positive effects of two methods to some extent.

Additionally, ConInceDeep, as the model introducing both Inception and CWT, provided the best identification result in four CNN structures, and four evaluation indices were all promoted. The average *Acc* and *F1_Score* of ConInceDeep achieved 96.60% and 97.05%, 6.28% and 6.75% higher than that of 1D-CNN. The average *Rec* dramatically increased from 84.66% in 1D-CNN to 94.57% in ConInceDeep, which is the index people concern more in reality, that is, the models are required to be more sensitive to the presence of a specific pure substance in the test mixture samples. And ConInceDeep achieved the highest average *Pre* at even 100%.

Especially, the *Acc* achieved by ConInceDeep model of ethanol was 85.86%, which was much higher than that by 1D-CNN at only 69.11%. *Rec* and *F1_Score* were also increased in ConInceDeep. To discuss the performance improvement of ConInceDeep model in detail, samples identified incorrectly in 1D-CNN model of ethanol while identified correctly in ConInceDeep were statistically analyzed.

The result showed that these samples were all positive mixture samples (containing ethanol) with a low volume of ethanol. When the volume ratio of ethanol was 0.1, all 39 mixture samples were not identified by 1D-CNN, as well as 15 and 4 samples with volume ratios of 0.2 and 0.3. In ConInceDeep, however, the model of ethanol achieved better performance at low concentration detection, and 17 mixtures with 0.1 vol ratio, 11 mixtures with 0.2 and all 4 mixtures with 0.3 were successfully identified. Although there were still some test samples that failed when the volume ratio of ethanol was 0.1, 11 of these were quinary with quite complicated spectral lines, as well as 8 quaternary mixtures, which were tough for models to perceive the peaks of ethanol. In fact, the error samples in 1D-CNN were mainly attributed to low concentration and little peak number of ethanol, hence, the peaks of ethanol were easily overlapped by other substances' peaks and it was difficult for the 1D-CNN model to perceive information about ethanol. The following Fig. 5 were diagrams of several liquid mixtures as examples which illustrated this phenomenon concretely. As can be seen, among the spectra of the mixture, the spectral peaks of ethanol were almost completely covered except the strongest peak near 885 cm^{-1} which could be observed. There was little information in the mixture spectra that contributed to detecting ethanol. Therefore, it is extremely difficult for models to detect ethanol from these mixtures, and the 1D-CNN model for ethanol did not successfully predict the presence of ethanol from these four mixtures. While with the help of CWT and Inception modules, ConInceDeep model of ethanol successfully identified ethanol from 4 mixtures above, which was possibly attributed to

Table 4
Differences in experimental design between the four CNNs.

	1D-CNN	1D-all-Ince-CNN	2D-Wav-CNN	ConInceDeep
CWT	×	×	✓	✓
Inception module	×	✓	×	✓

(✓ represents that CNN introduced the corresponding method; × represents that CNN did not introduce the corresponding method).

Table 5
Identification results of the models obtained based on four CNN structures.

	1D-CNN				1D-all-Ince-CNN				2D-Wav-CNN				ConInceDeep			
	Acc	*Pre*	*Rec*	*F1_Score*	*Acc*	*Pre*	*Rec*	*F1_Score*	*Acc*	*Pre*	*Rec*	*F1_Score*	*Acc*	*Pre*	*Rec*	*F1_Score*
Diethyl malonate	95.29	100	82.00	90.11	97.38	100	90.00	94.74	97.38	92.45	98.00	95.15	**98.43**	**100**	**94.00**	**96.91**
Acetone	89.53	99.32	88.41	93.55	90.58	100	89.02	94.19	94.24	98.11	95.12	96.59	**95.29**	**100**	**94.51**	**97.18**
Acetonitrile	96.86	100	94.44	97.14	97.38	98.13	97.22	97.67	98.95	99.07	99.07	99.07	**100**	**100**	**100**	**100**
Ethanol	69.11	100	53.91	70.05	72.25	100	58.59	73.89	82.20	100	73.44	84.68	**85.86**	**100**	**78.91**	**88.21**
Diacetone alcohol	92.67	92.78	92.78	92.78	95.81	100	91.75	95.70	100	100	100	100	**100**	**100**	**100**	**100**
Cyclohexane	98.43	100	96.43	98.18	100	100	100	100	100	100	100	100	**100**	**100**	**100**	**100**
Average	90.32	98.68	84.66	90.30	92.23	99.69	87.76	92.70	95.46	98.27	94.27	95.92	**96.60**	**100**	**94.57**	**97.05**

(Units of all data above are %).

improving its sensitivity to overlapping and weak peaks of ethanol due to low concentration to some intent. Therefore, compared with the 1D-CNN model, the model of ConInceDeep method performed better in the detection of ethanol with low concentration in complex mixtures.

3.2. Prediction results of the powder mixture set and real sample set

The spectra of the mixtures in the powder mixture set and real sample set were used to demonstrate the reliability of the models of ConInceDeep further, which are more difficult to be identified than liquid mixtures. In the powder mixture set, the ConInceDeep models of each four pure powder components (glutamic acid, aspartic acid, histidine, and glycine) were firstly built. Then the Raman spectra of 33 powder mixture samples were acquired and fed into the four models respectively for component prediction. ConInceDeep models of four key components achieved good identification results, the models of glutamic acid, aspartic acid and histidine reached 100% *Acc*, as well as *Pre*, *Rec* and *F1_Score*. Only one mixture sample with negative actual label was of false detection by glycine model as positive prediction label. Even so, glycine model based on ConInceDeep still achieved 96.97% *Acc*, 95.45% *Pre*, 100% *Rec* and 97.67% *F1_Score*.

Above, the detection of specific substance and the distinction of similar components by ConInceDeep models has been analyzed stressfully. Herein, *FPR* was used to further verify the feasibility of ConInceDeep models on more pure substances and explore the possibility of ConInceDeep applied to unknown mixtures. 50 out of 100 pure substances were randomly selected and ConInceDeep models of these 50 substances were subsequently built, which were used to make predictions for the 33 powder mixtures. Prediction results of these 50 models of other substances and 4 models of key components were combined and displayed in Table 6. It could be seen that ConInceDeep acquired a high *Acc* of 99.66% and a low *FPR* of 0.35% by 54 models testing 33 powder mixtures, which illustrated that ConInceDeep could identify and exclude the existence of some substances successfully.

In addition, two kinds of mixture samples (mouthwash and baking powder) existing in the real world were selected to demonstrate the application field and value of ConInceDeep. For the experiment of mouthwashes, the main thing was to detect the presence or absence of ethanol in two different types of mouthwash, so the ethanol model based on ConInceDeep needed to be established first. Since in the test of previous liquid mixture dataset, ethanol model had been built which was used directly right here. Then, the two types of mouthwash were sampled and acquired spectra. Finally, through the prediction of the model, mouthwash containing ethanol was identified successfully with positive prediction label, and the presence of ethanol was also successfully excluded with negative prediction label for mouthwash that did not contain ethanol. With regard to the experiment of baking powder, sodium bicarbonate was exactly involved in the 100 pure substances used in this study, thus its model was built in the same way as described in this research. The experimental results showed that the ConInceDeep method could still detect sodium bicarbonate even it was of the lower content of the two main ingredients in baking powder. This proved the feasibility of the ConInceDeep method for the detection of real mixtures.

3.3. Advantages of the introduction of CWT and Inception modules

The above mentioned verifies the effectiveness of ConInceDeep from the perspective of performance indices indirectly. For the CNN model itself, it is a black-box model which brings deficient explanation of the connection between the input data and the output classification. To identify the important spectral region in input data and add interpretation to the CNN models in this study, Gradient-weighted Class Activation Mapping(Grad-CAM) was implemented to visualize the critical input variables for prediction output using gradient information flowing into the final convolution layer [31], which was used to illustrate the advantages of introducing CWT and Inception respectively.

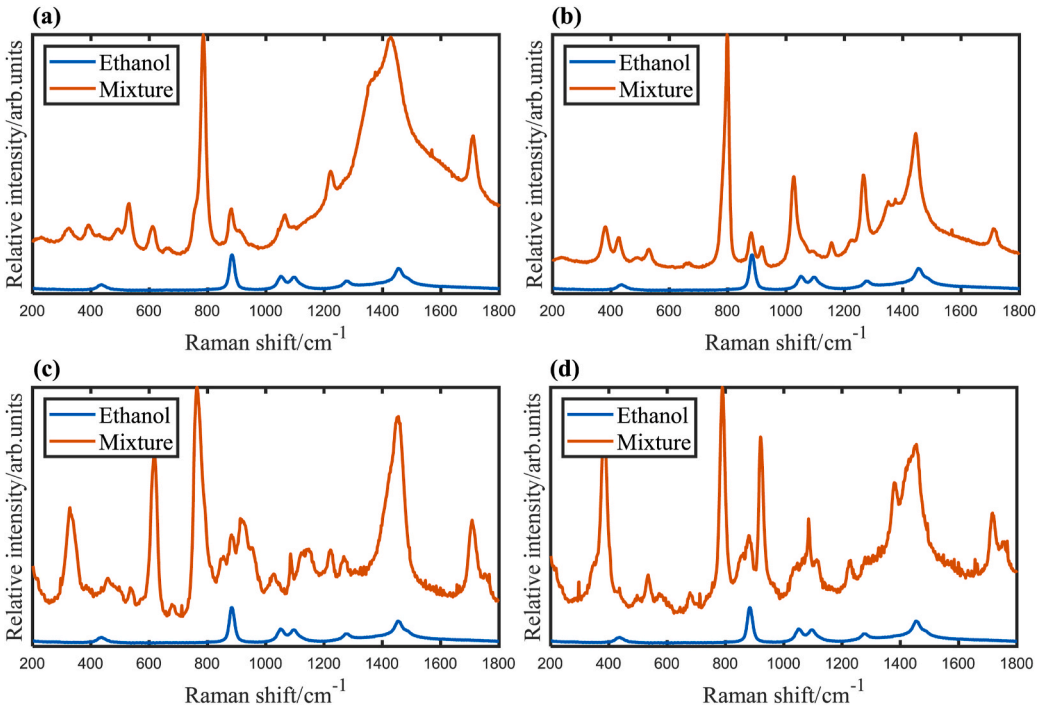


Fig. 5. Spectra of ethanol and the mixtures with low concentration of ethanol (The spectra of ethanol were all multiplied by its volume ratio corresponding to mixture).
(a) mixture composed of acetone, ethanol and diacetone alcohol (the volume ratio of 0.7:0.1:0.2); (b) mixture composed of acetonitrile, acetone, cyclohexane and ethanol (0.1:0.4:0.4:0.1); (c) mixture composed of ethanol, acetonitrile, diacetone alcohol and diethyl malonate (0.1:0.1:0.6:0.2); (d) mixture composed of ethanol, acetonitrile, acetone and diethyl malonate (0.1:0.4:0.3:0.2).

Table 6
Prediction results of the powder mixture set based on 50 models of other substances and 4 models of key components.

	TP	TN	FP	FN	*Acc*(%)	*FPR*(%)
2-component	36	933	3	0	99.69	0.32
3-component	36	610	2	0	99.69	0.33
4-component	12	149	1	0	99.38	0.67
Total	84	1692	6	0	99.66	0.35

Herein, the acetone models obtained based on 1D-CNN, 1D-all-Ince-CNN and 2D-Wav-CNN respectively were taken as examples to detect the mixture composed of acetonitrile, acetone and diacetone alcohol. Fig. 6 shows Grad-CAM heatmaps for the three models and the Raman spectrum of this mixture, as well as acetone.

In this example, the result of Grad-CAM showed that compared with the 1D-CNN model, there were more regions of input spectra focused on by 1D-all-Ince-CNN model, which also mainly distributed in the spectral peak area of the mixture. As shown in Fig. 6(a) and (b), only two peaks around 383 cm^{-1} and 921 cm^{-1} in the mixture were activated clearly in 1D-CNN, while 1D-all-Ince-CNN also paid attention to other three additional regions, around 617 cm^{-1}, 764 cm^{-1} and 789 cm^{-1} respectively, which could correspond to three peaks of the mixture. It demonstrated that for 1D-all-Ince-CNN model, the addition of Inception modules enriched the receptive field and increased the adaptability of the network to different spectral peaks.

As for the 2D-Wav-CNN model with 2D CWT coefficient matrix as input, because the wavelet coefficients on both sides of the spectral peaks were negative with large absolute values, it was more concerned by the model. With this phenomenon, the position of the spectral peak could be located indirectly. In other words, the range of Raman shift in the middle should be the region of Raman spectral peak. And the distance between two sides could reflect the width of peak to a certain extent. Further distance meant larger width of Raman characteristic peaks and vice versa. It was noticed that there were overlapping peaks in the spectral range of 720 cm^{-1}–820 cm^{-1}, the 2D-Wav-CNN model could identify the existence of two peaks in this region, where the activated region in the middle was the left and right side of two peaks respectively in Fig. 6(c). Moreover, compared with the 1D-CNN model and the 1D-all-Ince model, the Grad-CAM result of 2D-Wav-CNN model also reflected the existence of two weak peaks around 1069 cm^{-1} and 1224 cm^{-1} more clearly.

Moreover, the three models in Fig. 6 are all models to predict the presence of acetone in the mixture, so it is also necessary to pay attention to the activation of acetone peaks in the three models respectively. In 2D-Wav-CNN, all peaks belonging to acetone were activated at different degrees, located in 391 cm^{-1}, 493 cm^{-1}, 532 cm^{-1}, 789 cm^{-1}, 1069 cm^{-1}, 1224 cm^{-1}, 1434 cm^{-1} and 1715 cm^{-1} respectively. However, in 1D-CNN and 1D-all-Ince-CNN, the activation of some peaks was not as obvious and clear as that in 2D-Wav-CNN, like the peaks in 1069 cm^{-1}, 1224 cm^{-1}, 1434 cm^{-1} and 1715 cm^{-1}. And the amount of attention to the peak in 789 cm^{-1}, the strongest peak in acetone, was also not as high as that in 2D-Wav-CNN, but more attention was paid to the peak not related to acetone at 921 cm^{-1} in the mixture. Hence, the introduction of CWT to 2D-Wav-CNN model enhanced the ability to detect overlapping and weak peaks.

3.4. Different sizes of convolutional kernels

To illustrate the identification effectiveness of Inception module adding multiple sizes of parallel convolutional kernels, convolutional kernels in different fixed sizes were used to establish multiple 1D-CNN models, which were compared with 1D-all-Ince model. The network structure of these 1D-CNN models was as same as that mentioned in section 2.2.3, but had two more sizes, 1*7, 1*9 convolutional kernels respectively. *Acc* and *F1_Score* were used in model evaluation. A total of 24 models were compared corresponding to 6 liquid pure substances with 4 kinds of CNN structures. 191 liquid mixtures were still used as test set. The identification results of these models are shown in Fig. 7.

As can be seen, for the 1D-CNN models with the idea of fixed convolutional kernel size, the three models of the same substance had different identification effects with different convolutional kernel sizes. For some pure substances, the convolutional kernel with larger size was better, such as acetone, cyclohexane and so on. But in the three models of diacetone alcohol, the one with a smaller filter size performed better. This indicates that different substances have different adaptability to

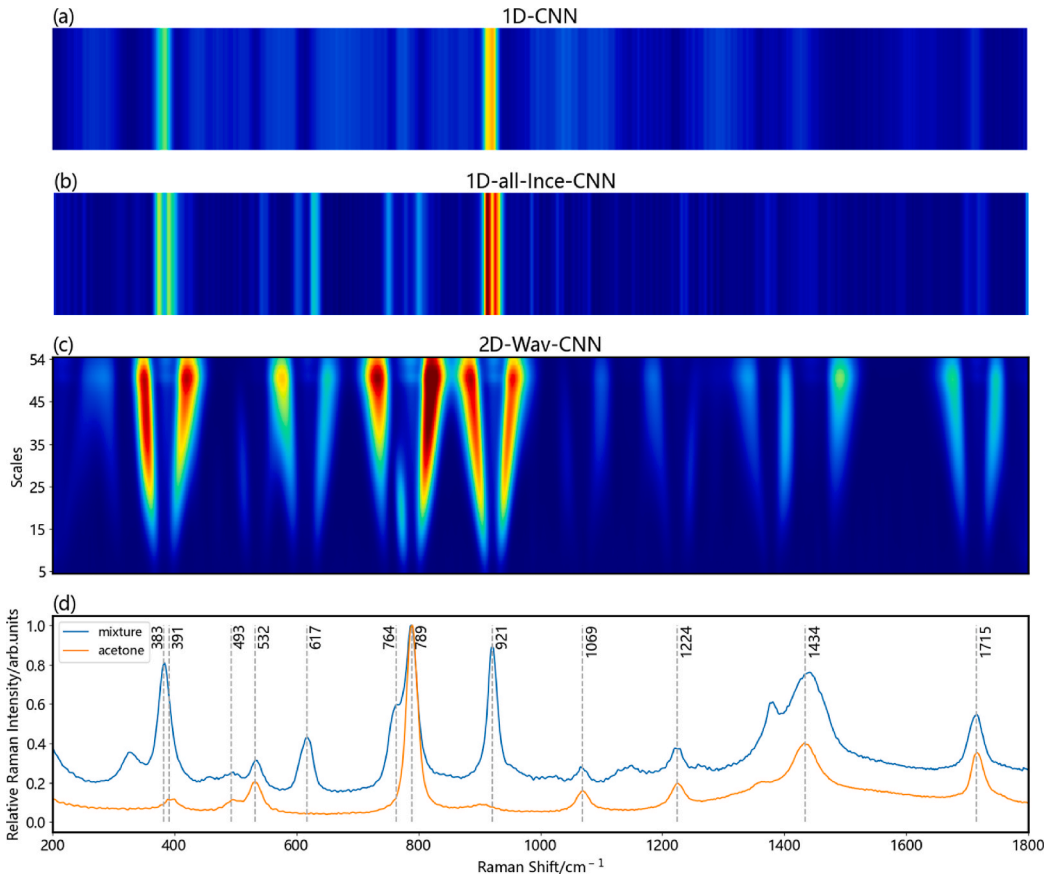


Fig. 6. Grad-CAM result of three models testing the mixture composed of acetonitrile, acetone and diacetone alcohol. (a)1D-CNN model; (b)1D-all-Ince-CNN model; (c)2D-Wav-CNN model; (d)the Raman spectrum of the mixture and acetone.

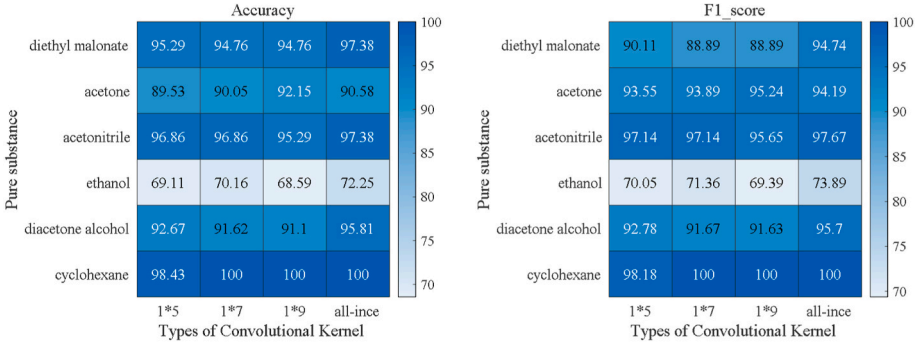


Fig. 7. Model performances using different convolutional layers. (a) *Acc*; (b) *F1_Score*.

convolutional kernels in different sizes, so it is necessary to diversify the size of convolutional kernels and improve the network performance facing various samples. At the same time, it could also be noted that among the four models corresponding to each pure substance, all models established with Inception modules achieved the best results, which further proves the positive significance of Inception module.

3.5. Different wavelet functions of CWT

To discuss the effects of the wavelet functions in this study, we used Gaus4, Mexican hat and Lorentz4 as wavelet function respectively in the ConInceDeep method to carry out CWT on 1D spectral data of virtual mixtures, which were used as the training set and the validation set of the three models respectively. A total of 18 models corresponding to 6 liquid pure substances were established. Then, the corresponding

Table 7
The *Acc* and *F1_Score* of models obtained using three wavelet functions.

	Gaus4		Mexican hat		Lorentz4	
	Acc (%)	*F1_Score* (%)	*Acc* (%)	*F1_Score* (%)	*Acc* (%)	*F1_Score* (%)
Diethyl malonate	98.43	97.03	96.86	93.62	**98.43**	**96.91**
Acetone	92.67	95.54	90.58	94.27	**95.29**	**97.18**
Acetonitrile	94.24	95.15	98.43	98.59	**100**	**100**
Ethanol	84.29	87.07	80.63	83.11	**85.86**	**88.21**
Diacetone alcohol	96.34	96.52	100	100	**100**	**100**
Cyclohexane	97.38	97.11	98.95	98.80	**100**	**100**
Average	93.89	94.74	94.24	94.73	**96.60**	**97.05**

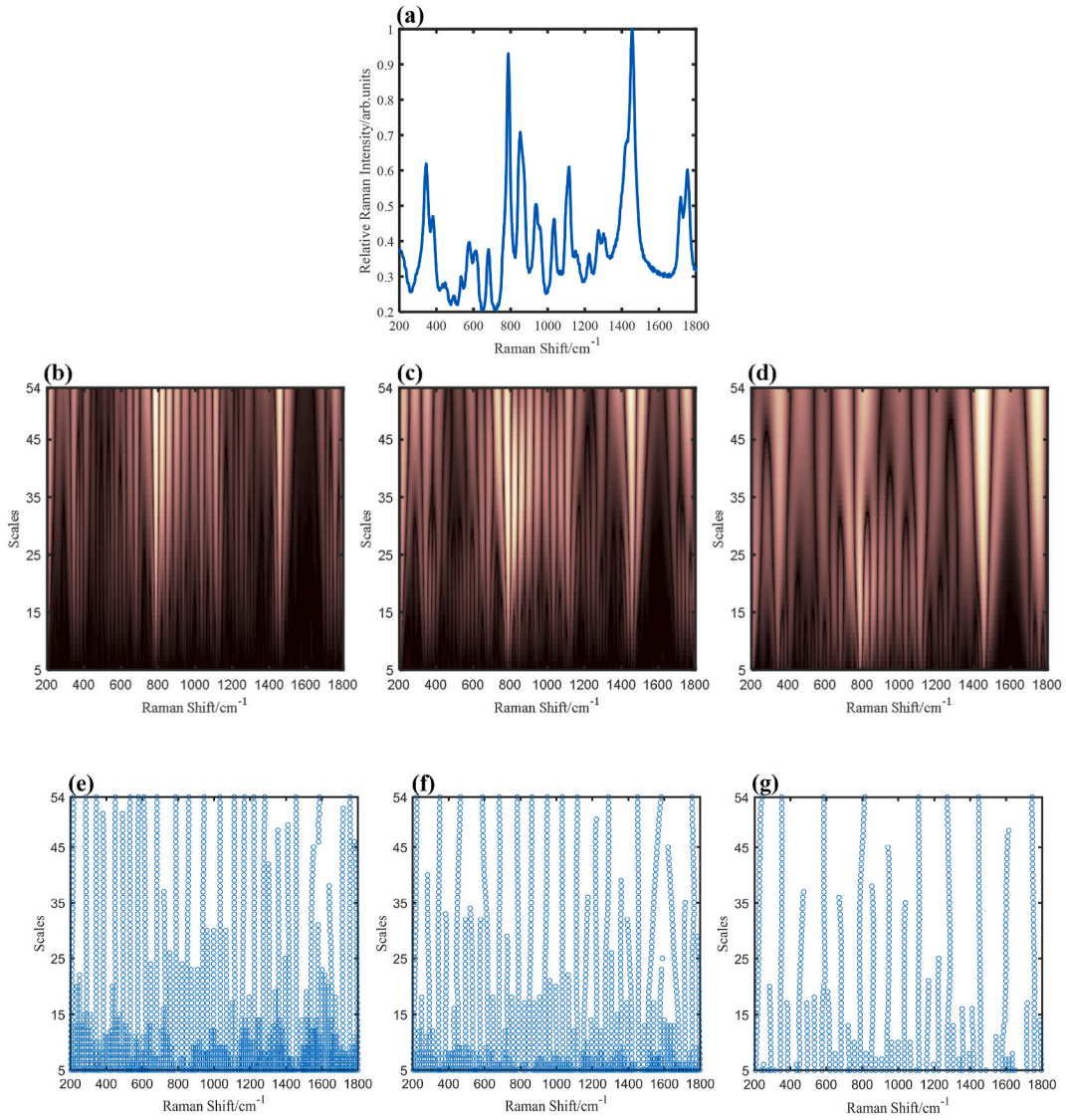


Fig. 8. The result of CWT using Lorentz4, Gaus4 and Mexican hat wavelet acting on the mixture composed of diethyl malonate, acetone and diacetone alcohol. (a) Original spectrogram of the mixture; (b)–(d) Schematic diagrams of CWT coefficients matrixes using Lorentz4, Gaus4 and Mexican hat as wavelet function respectively; (e)–(g) The ridge line diagrams obtained by coefficients matrixes using Lorentz4, Gaus4 and Mexican hat respectively.

wavelet functions were used in CWT to 191 liquid mixtures as the input of test samples, and the evaluation result of models was finally obtained, as shown in Table 7. The result showed that the models corresponding to Lorentz4 achieved the highest average *Acc* and *F1_Score*. The average *F1_Score* of Lorentz4 models was 2.31% and 2.32% higher than the one of Gaus4 and Mexican hat respectively.

Fig. 8 is schematic diagrams of the result of CWT on the spectrum of the mixture composed of diethyl malonate, acetone and diacetone alcohol using Lorentz4 wavelet, Gaus4 wavelet and Mexican hat wavelet. The original spectrogram of the mixture is shown in Fig. 8(a), then Fig. 8(b), (c) and (d) are schematic diagrams of wavelet coefficient matrixes using three wavelet functions respectively. The wavelet coefficients represent the transform values of the original spectrum at different scales. The local maximum values at each scale can be circled and linked to form ridge lines, and each ridge line corresponds to a spectral peak. Fig. 8(e) and (f) and (g) are ridge line diagrams corresponding to the wavelet coefficients matrixes using three different wavelets. As can be observed, the coefficient matrix obtained by Lorentz4 wavelet was more concentrated and showed richer spectral details compared with Gaus4 and Mexican hat wavelets. For some hidden overlapping peaks, Lorentz4 wavelet could still detect them relying on its better shape.

4. Conclusion and future work

In this study, a novel deep learning model method: ConInceDeep introducing CWT and Inception module was proposed to identify one specific component in mixture, which considers the intrinsic Raman spectral characteristics and solves the problems of inadequate expression of 1D spectral data. Ablation experiments were firstly applied in this study to illustrate the positive effects of CWT and Inception module respectively by testing 191 liquid mixtures, and the model of each specific pure substance based on ConInceDeep achieved better performance than that of other three structures. Then, the identification performance of ConInceDeep was further verified on 33 powder mixtures and two kinds of real samples. The prediction results proved that it could be used to detect the existence of specific pure substance and distinguish the pure substances with similar structure successfully, as well as to exclude some pure substances dealing with unknown mixtures. Grad-CAM was also introduced to visualize the key spectral regions and understand the advantages of CWT and Inception modules respectively. At the same time, we also discussed the impact on different sizes of convolutional kernels and the choice of wavelet functions, and proved the rationality of Inception structure and Lorentz4 wavelet.

Model repeatability and stability are important for further application of the learning-based ConInceDeep method due to random initialization of parameters and dropouts. More models obtained from multiple trainings will help to verify the robustness of this method. Meantime, ConInceDeep as a general method is possibly expected to play a role in classification and identification tasks using other spectroscopy, such as near-infrared spectroscopy and terahertz spectroscopy.

Credit author statement

Ziyan Zhao: Conceptualization, Methodology, Software, Writing - Original Draft **Zhenfang Liu:** Formal analysis, Writing - Review & Editing **Mingqiang Ji:** Methodology, Software **Xin Zhao:** Validation, Writing - Review & Editing **Qibing Zhu:** Supervision, Writing - Review & Editing **Min Huang:** Resources, Data Curation, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgments

The authors would like to thank Beijing Zolix optical instrument company, China for their assistance in sample preparation and data acquisition, and gratefully acknowledge the financial support from the National Natural Science Foundation of China (Grant no.61775086). Besides, we also would like to express our gratitude to the two anonymous reviewers for their helpful and constructive comments on our work, which improved this paper greatly.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemolab.2023.104757.

References

[1] M. Benedict, H. Husain, T. Michael, et al., Synthetic cannabinoid receptor agonists detection using fluorescence spectral fingerprinting, Anal. Chem. 91 (20) (2019) 12971–12979, https://doi.org/10.1021/acs.analchem.9b03037.
[2] S.S. Khan, M.G. Madden, New similarity metrics for Raman spectroscopy, Chemometr. Intell. Lab. Syst. 114 (2012) 99–108, https://doi.org/10.1016/j.chemolab.2012.03.007.
[3] D. Pankin, I. Kolesnikov, A. Vasileva, et al., Raman fingerprints for unambiguous identification of organotin compounds, Spectrochim. Acta Mol. Biomol. Spectrosc. 204 (2018) 158–163, https://doi.org/10.1016/j.saa.2018.06.044.
[4] X. Zhao, C.Z. Liu, Z.Y. Zhao, et al., Performance improvement of handheld Raman spectrometer for mixture components identification using fuzzy membership and sparse non-negative least squares, Appl. Spectrosc. 76 (5) (2022) 548–558, https://doi.org/10.1177/00037028221080205.
[5] Z.M. Zhang, X.Q. Chen, H.M. Lu, et al., Mixture analysis using reverse searching and non-negative least squares, Chemometr. Intell. Lab. Syst. 137 (2014) 10–20, https://doi.org/10.1016/j.chemolab.2014.06.002.
[6] J.J. Zhu, A.S. Sharma, J. Xu, et al., Rapid on-site identification of pesticide residues in tea by one-dimensional convolutional neural network coupled with surface-enhanced Raman scattering, Spectrochim. Acta Mol. Biomol. Spectrosc. 246 (2021), 118994, https://doi.org/10.1016/j.saa.2020.118994.
[7] F. Lussier, V. Thibault, B. Charron, et al., Deep learning and artificial intelligence methods for Raman and surface-enhanced Raman scattering, Trends Anal. Chem. 124 (2020), 115796, https://doi.org/10.1016/j.trac.2019.115796.
[8] X. Yan, S. Zhang, H. Fu, et al., Combining convolutional neural networks and online Raman spectroscopy for monitoring the Cornu Caprae Hircus hydrolysis process, Spectrochim. Acta Mol. Biomol. Spectrosc. 226 (2020), 117589, https://doi.org/10.1016/j.saa.2019.117589.
[9] I. Isaex, S. Burikov, T. Dolenko, et al., Artificial neural networks for diagnostics of water-ethanol solutions by Raman spectra, Stud. Comput. Intell. 799 (2019) 167–175, https://doi.org/10.1007/978-3-030-01328-8_18.
[10] X.X. Zheng, G. Lu, G. Du, et al., Raman spectroscopy for rapid and inexpensive diagnosis of echinococcosis using the adaptive iteratively reweighted penalized least squares-Kennard-stone-back propagation neural network, Laser Phys. Lett. 15 (8) (2018), 085702, https://doi.org/10.1088/1612-202X/aac29f.
[11] X.C. Sang, R.G. Zhou, Y.C. Li, et al., One-dimensional deep convolutional neural network for mineral classification from Raman spectroscopy, Neural Process. Lett. 54 (1) (2021) 677–690, https://doi.org/10.1007/s11063-021-10652-1.
[12] S. Huang, P. Wang, Y b Tian, et al., Blood species identification based on deep learning analysis of Raman spectra, Biomed. Opt Express 10 (12) (2019) 6129–6144, https://doi.org/10.1364/BOE.10.006129.
[13] J. Liu, M. Osadchy, L. Ashton, et al., Deep convolutional neural networks for Raman spectrum recognition: a unified solution, Analyst 142 (21) (2017) 4067–4074, https://doi.org/10.1039/c7an01371j.
[14] X. Fan, W. Ming, H. Zeng, et al., Deep learning-based component identification for the Raman spectra of mixtures, Analyst 144 (5) (2019) 1789–1798, https://doi.org/10.1039/c8an02212g.
[15] C.A. Smith, E.J. Want, G. O'Maille, et al., XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification, Anal. Chem. 78 (3) (2006) 779–787, https://doi.org/10.1021/ac051437y.
[16] Y. Zheng, R.L. Fan, C.L. Qiu, et al., An improved algorithm for peak detection in mass spectra based on continuous wavelet transform, Int. J. Mass Spectrom. 409 (2016) 53–58, https://doi.org/10.1016/j.ijms.2016.09.020.
[17] Z.M. Zhang, X. Tong, Y. Peng, et al., Multiscale peak detection in wavelet space, Analyst 140 (23) (2015) 7955–7964, https://doi.org/10.1039/c5an01816a.
[18] L. Jiao, S. Gao, F. Zhang, et al., Quantification of components in overlapping peaks from capillary electrophoresis by using continues wavelet transform method, Talanta 75 (4) (2008) 1061–1067, https://doi.org/10.1016/j.talanta.2008.01.016.
[19] M.H. Liu, Z.R. Dong, G.F. Xin, et al., An improved method based on a new wavelet transform for overlapped peak detection on spectrum obtained by portable Raman system, Chemometr. Intell. Lab. Syst. 182 (2018) 1–8, https://doi.org/10.1016/j.chemolab.2018.08.002.
[20] H.Y. Guo, Q.G. He, B. Jiang, The application of Mexican hat wavelet filtering and averaging algorithm in Raman spectra denoising, Proc. Congr. Image Signal Process (2008) 321–326, https://doi.org/10.1109/CISP.2008.191.
[21] Y. Wang, X.Q. Fan, S. Tian, et al., EasyCID: make component identification easy in Raman spectroscopy, Chemometr. Intell. Lab. Syst. 231 (2022), 104657, https://doi.org/10.1016/j.chemolab.2022.104657.
[22] P.F. Qi, Y.C. Wang, Seismic time-frequency spectrum analysis based on local polynomial Fourier transform, Acta Geophys. 68 (1) (2019) 1–17, https://doi.org/10.1007/s11600-019-00377-0.
[23] D. Komorowski, S. Pietraszek, The use of continuous wavelet transform based on the fast fourier transform in the analysis of multi-channel electrogastrography recordings, J. Med. Chem. 40 (1) (2016) 10, https://doi.org/10.1007/s10916-015-0358-4.
[24] D. Zhen, Z.L. Wang, H.Y. Li, et al., An improved cyclic modulation spectral analysis based on the CWT and its Application on broken rotor bar fault diagnosis for induction motors, Appl. Sci-basel 9 (18) (2019) 3902, https://doi.org/10.3390/app9183902.
[25] K.L. Chong, S.H. Lai, A. El-Shafie, Wavelet transform based method for river stream flow time series frequency analysis and assessment in tropical environment, Water Resour. Manag. 33 (6) (2019) 2015–2032, https://doi.org/10.1007/s11269-019-02226-7.
[26] C. Szegedy, W. Liu, Y. Jia, et al., Going Deeper with Convolutions, *IEEE Conf. Comput. Vis. Pattern Recognit*, Boston, USA, 2015, pp. 1–9, https://doi.org/10.1109/CVPR.2015.7298594.
[27] M.Q. Qiu, S.G. Zheng, L. Tang, et al., Raman spectroscopy and improved Inception network for determination of FHB-infected wheat kernels, Foods 11 (4) (2022) 578, https://doi.org/10.3390/foods11040578.
[28] T.U. Rehman, D.D. Ma, L.J. Wang, et al., Predictive spectral analysis using an end-to-end deep model from hyperspectral images for high-throughput plant phenotyping, Comput. Electron. Agric. 177 (2020), 105713, https://doi.org/10.1016/j.compag.2020.105713.
[29] X.L. Zhang, T. Lin, J.F. Xu, et al., DeepSpectra: an end-to-end deep learning approach for quantitative spectral analysis, Anal. Chim. Acta 1058 (2019) 48–57, https://doi.org/10.1016/j.aca.2019.01.002.
[30] C. Szegedy, V. Vanhoucke, S. Lodde, et al., Rethinking the inception architecture for computer vision, *IEEE Conf. Comput. Vis. Pattern Recognit.* Las Vegas (2016) 2818–2826, https://doi.org/10.1109/CVPR.2016.308. USA.
[31] C. Selvaraju, M. Cogswell, A. Das, et al., Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization, *IEEE Int. Conf. Comput. Vis*, Venice, Italy, 2017, pp. 618–626, https://doi.org/10.1109/ICCV.2017.74.