A 3D Fluorescence Classification and Component Prediction Method Based on VGG Convolutional Neural Network and PARAFAC Analysis Method

Kun Ruan 1, Shun Zhao 1, Xueqin Jiang 1, Yixuan Li 2, Jianbo Fei 3, Dinghua Ou 3, Qiang Tang 1, Zhiwei Lu 4, Tao Liu 1 and Jianguo Xia 3,*

1 College of Information Engineering, Sichuan Agricultural University, Ya’an 625000, China; qkmc@outlook.com (K.R.); 2018319005@stu.sicau.edu.cn (S.Z.); 2021219002@stu.sicau.edu.cn (X.J.); 13599@sicau.edu.cn (Q.T.); 13148@sicau.edu.cn (T.L.)
2 College of Environmental Sciences, Sichuan Agricultural University, Ya’an 625000, China; 202002722@stu.sicau.edu.cn
3 College of Resources, Sichuan Agricultural University, Chengdu 611130, China; 14674@sicau.edu.cn (J.F.); 14340@sicau.edu.cn (D.O.)
4 College of Science, Sichuan Agricultural University, Ya’an 625000, China; zhiweilu@sicau.edu.cn
* Correspondence: xiajianguo@126.com; Tel.: +86-187-8221-0536

Abstract: Three-dimensional fluorescence is currently studied by methods such as parallel factor analysis (PARAFAC), fluorescence regional integration (FRI), and principal component analysis (PCA). There are also many studies combining convolutional neural networks at present, but there is no one method recognized as the most effective among the methods combining convolutional neural networks and 3D fluorescence analysis. Based on this, we took some samples from the actual environment for measuring 3D fluorescence data and obtained a batch of public datasets from the internet species. Firstly, we preprocessed the data (including two steps of PARAFAC analysis and CNN dataset generation), and then we proposed a 3D fluorescence classification method and a components fitting method based on VGG16 and VGG11 convolutional neural networks. The VGG16 network is used for the classification of 3D fluorescence data with a training accuracy of 99.6% (as same as the PCA + SVM method (99.6%)). Among the component maps fitting networks, we comprehensively compared the improved LeNet network, the improved AlexNet network, and the improved VGG11 network, and finally selected the improved VGG11 network as the component maps fitting network. In the improved VGG11 network training, we used the MSE loss function and cosine similarity to judge the merit of the model, and the MSE loss of the network training reached $4.6 \times 10^{-4}$ (characterizing the variability of the training results and the actual results), and we used the cosine similarity as the accuracy criterion, and the cosine similarity of the training results reached 0.99 (comparison of the training results and the actual results). The network performance is excellent. The experiments demonstrate that the convolutional neural network has a great application in 3D fluorescence analysis.

Keywords: 3D fluorescence; 3D-EEMs; deep learning; convolutional neural network; PARAFAC analysis; VGG neural network

1. Introduction

The treatment of organic pollution in rivers and drinking water is one of the social problems that needs to be dealt with for the development of human society, and the water pollution situation is related to people’s health and is widely present in natural water bodies [1–3]. In recent years, more and more attention has been paid to organic pollution in rivers and drinking water [4], and therefore to water quality monitoring methods [5]. For example, the use of three-dimensional fluorescence and PARAFAC analysis methods to
track the treatment of natural organics in drinking water treatment plants [6], combined with three-dimensional fluorescence using open set identification methods to identify organic contamination in water bodies [7]. In addition, there are still many methods that use PCA, SVM, and deep learning algorithms combined with 3D fluorescence for analysis [8,9].

According to previous studies, three-dimensional fluorescence has been widely used in the analysis of water pollution [10] and various foods [11]. Most of these organic substances such as humic acids, lignin, aromatic amino acids, and DNA can be analyzed by fluorescence [12]. When these substances are present in solution and different types and concentrations, their fluorescence characteristics follow the changes. Therefore, three-dimensional fluorescence can be used to distinguish and detect different organic solutions [13–15]. In addition, 3D fluorescence differs from 2D fluorescence in that it consists of a large number of excitation–emission fluorescence matrices, so an efficient 3D fluorescence interpretation method is needed to extract useful components information from 3D fluorescence data [16].

There have been many analytical methods for 3D fluorescence. Wenchen et al. used the FRI method for the effective analysis of DOM components [17]. Yamashita Y et al. [18] used PARAFAC analysis to characterize the interactions between trace metals and dissolved organic matter. Ruiyu Xu and others [9] used PCA combined with SVM to classify and identify Chinese strong aroma-type liquors. Francesca Guimet et al. [19] used the expanded principal component analysis combined with PARAFAC analysis to distinguish pure olive oil and virgin olive oil, founding three classes of possible compounds. Ramila H. Peiris et al. [20] used the PCA method to track pollutants in the drinking water treatment process. C. W. cuss et al. [21] used four machine learning algorithms (i.e., decision tree, k-nearest neighbor, multilayer perceptron, and support vector machine) combined with the PARAFAC analysis method to classify a large number of 3D-EEMs data. In addition, there are also many research papers on 3D-EEMs data combined with deep learning algorithms. For example, Shi Fei [7] understood the 3D-EEMs matrix as an image channel, added X and Y coordinate channels, and established an organic pollutants evaluation model based on a position convolution network (CoordConv) and cosine similarity. Wu Xijun [22] used the contour map generated by 3D-EEMs as the training dataset of AlexNet, combined with SVM and PLS (least square method) to judge whether there is forgery in sesame oil, and obtained a 100% recognition accuracy. Xu Runze [8], combined with the analysis results of PARAFAC analysis, predicted the number of components of different kinds of 3D-EEMs data by using the NCNN network (a revolutionary neural network model for the number of fluorescent components) and fitted the components with MCNN network (a revolutionary neural network model for the map of fluorescent components). In addition, Lu Zhiwei [23] used yolo-v3 algorithm to classify and recognize the fluorescent image and used the algorithm to segment the fluorescent region. They analyzed the relationship between heavy metal ions in agricultural residues and fluorescence intensity from the perspective of RGB value, HSV value, and the concentration of agricultural residues. Liu Tao [24] also used the yolo-v3 algorithm combined with the wechat applet to detect and classify fluorescence images and realized a scheme to quickly detect GSH and ADA on site based on an image processing algorithm.

Based on the above research, nowadays, in addition to the traditional FRI, PCA, and PARAFAC, etc., there have been many machine learning algorithms and deep learning algorithms to prove the feasibility of the study of 3D fluorescence data. Based on these studies, combined with the analysis results of PARAFAC analysis and VGG16 and VGG11 convolutional neural network, in this paper, we developed a new 3D fluorescence data classification model and 3D fluorescence component maps fitting model. Among them, VGG16 is used for the classification of 3D fluorescence data, which is used to classify the 3D fluorescence from different sources. VGG11 is used for components map fitting, and compared with the MCNN model of Xu et al., this model does not need to create a submodel for each component map but determines the number of submodels based on the maximum number of components of 3D fluorescence, which makes the components fitting model
more robust and prevents the overfitting effect of the network. In addition, VGG11 has a better fitting effect by comparing the model results with AlexNet and LeNet (the backbone network used by MCNN). Experiments show that according to the PARAFAC analysis results, the convolutional neural network can predict the category of 3D fluorescence data and fit its specific components map only by inputting a single 3D fluorescence datum. In this way, we don’t need to perform complex PARAFAC analysis on 3D-EEMs and can quickly obtain the category and components map of 3D fluorescence data using CNN models only, which provides a reliable technical means in river monitoring, drinking water monitoring, and other scenes that need rapid results.

2. Materials and Methods

2.1. Samples Collection

We collected 3D fluorescence data from different sources. The first part is 45 3D fluorescence samples (FU) collected from Fu River (the reach of Huanglongxi town and Yong’an town) in Shuangliu District, Chengdu City, China. The second part is the data published in the article (Murphy et al. (2013)) [25] including 206 3D fluorescence samples (P). The third part is the “Pure Tutorial” from drEEM toolbox, including 60 3D fluorescence data samples (PU). The fourth part is 105 3D fluorescence data samples (F) extracted from fish molecules (Andersen et al.) [11]. Despite that the datasets are from different regions and using different raw materials, they all have similar characteristics of 3D fluorescence data. The details of these datasets are shown in Table 1.

Table 1. Catalogues of collected 3D-EEMs.

<table>
<thead>
<tr>
<th>Catalogues</th>
<th>Label</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fu river</td>
<td>FU</td>
<td>45</td>
</tr>
<tr>
<td>Fish muscle (Andersen et al.)</td>
<td>F</td>
<td>105</td>
</tr>
<tr>
<td>PortSurvey (Murphy et al.)</td>
<td>P</td>
<td>206</td>
</tr>
<tr>
<td>Pure (drEEM Tutorial)</td>
<td>PU</td>
<td>60</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>416 samples</td>
</tr>
</tbody>
</table>

The FU dataset contains a total of 45 samples, all of which were collected from the Fu River and its surroundings. A total of 10 sites were collected (samples in site 8 and site 9 were not located in Fu River. For meticulous samples collection sites, please refer to Figure 1). Among them, 10 samples were collected from site 1 and site 10, respectively, 8 samples were collected from site 7, 5 samples were collected from site 2, and 3 samples were collected from site 3, site 4, site 5, and site 6, respectively. The collected river water samples were refrigerated in an incubator at 5 °C, and the measurement shall be completed within 12 h of being kept at a low temperature. Before measuring 3D fluorescence, the samples were filtered using a sterile syringe with a 0.45 µm microporous filter membrane (the syringe and filter membrane and the quartz dish were each repeatedly washed three times with 3 mL of sample solution before filtration). The P dataset was collected from 26 different sites and 4 different cruises. After removing the abnormal samples, the remaining 206 samples were used in this experiment. The PU dataset consists of pure fluorescence from the “Pure Tutorial” in the drEEM V0.6, which contains 60 samples. The F dataset is from frozen cod fillets, which were stored at 2 °C for 21 days. To expand the variation of products, different frozen storage temperatures, frozen storage periods, and refrigerated storage periods were used. Therefore, the 3D fluorescence data come from different periods with some quantitative differences.
Figure 1. FU dataset sampling sites. The sequence of sampling sites shall be sorted according to the sampling time. The number of samples collected from each site are: Site 1 collected 10 samples, Site 2 collected 5 samples, Site 3 collected 3 samples, Site 4 collected 3 samples, Site 5 collected 3 samples, Site 6 collected 3 samples, Site 7 collected 8 samples, Site 8 collected 2 samples, Site 9 collected 2 samples, and Site 10 collected 10 samples. The number of samples shall be determined according to the actual situation of the sampling sites.

2.2. 3D-EEMs Measurement

For the FU dataset, the Hitachi F-4500 fluorescence spectrometer was used to measure its three-dimensional fluorescence data (the concentration of organic matter in the river water is relatively low, in this study, we selected several samples for dilution and repeated measurements to ensure that no fluorescence in-filtering effect occurred in the river water [26]). During the measurement, the excitation wavelength was set to 200–500 nm with an interval of 10 nm; the emission wavelength was set to 250–550 nm with an interval of 5 nm, the excitation and emission slits were both set to 10 nm, the PMT voltage was set to 700 V, and the scanning speed was set to 1200 nm/min. The measurement of F data and P data can be found in published articles (Andersen et al., Murphy et al.), while the PU data are accurately described in the official drEEM toolbox V0.6 documentation.

To use convolutional neural networks for classification and component maps fitting of 3D fluorescence data, the prerequisite is to obtain 3D fluorescence data. During the 3D fluorescence measurement experiment of the FU dataset, in total we measured and obtained 45 samples (except outlier samples), each 3D-EEM consists of 31 excitation spectra and 61 emission spectra, therefore, the 3D-EEM size is $61 \times 31$.

In the other three types of datasets, the F dataset was measured by fluorescence spectrometry using a Perkin Elmer LS50B fluorescence spectrometer in 10 nm $\times$ 10 nm quartz cuvettes with excitation spectra ranging from 250–370 nm with an excitation interval of 10 nm and emission spectra ranging from 270–600 nm with an emission interval of 1 nm. Each emission–excitation matrix consists of 13 excitation spectra and 331 emission spectra, and the 3D-EEM size is $331 \times 13$. In the P dataset, the excitation spectra range from 230–455 nm with an excitation interval of 5 nm, the emission spectra range from 290–682 nm with an emission interval of 4 nm, and each excitation–emission matrix consists of 46 excitation spectra and 99 emission spectra, and the 3D-EEM size is $99 \times 46$. In the PU dataset, the excitation spectra range from 243–399 nm with an interval of 4 nm, the
emission spectra range from 303–499 nm with an emission interval of 4 nm, and each emission-excitation matrix consists of 40 excitation spectra and 50 emission spectra, and the 3D-EEM size is 50 × 40.

2.3. Data Preprocessing for PARAFAC and CNN

The pre-processing of 3D-EEMs data includes two parts. The first part is to process the original 3D-EEMs data to be suitable for the PARAFAC analysis model’s construction, and the second part is to convert 3D-EEMs data into grayscale images for training and testing of the convolution neural network (CNN) model.

Since these four kinds of data are obtained from different regions and measured with different instruments under different conditions and parameters, they have different excitation and emission wavelength ranges, as well as different excitation and emission intervals. Therefore, it is necessary to adjust the excitation wavelength range and emission wavelength range of the data and normalize the data by using the Raman normalization method [27]. The Pandas and Numpy tools in the Python toolkits were used in this study to unify the excitation spectra interval to 200–500 nm and the emission spectra interval to 250–610 nm (The reason is that this interval contains the common interval of all different datasets to ensure that the fluorescence characteristics of all samples are not lost). In addition, Raman normalization of the data after the interval unification is required. The reference formula for Raman normalization information is as follows:

\[
A_{\lambda_{exc}} = \int_{\lambda_{em}}^{\lambda_{em}'} I_{\lambda_{em}} d\lambda_{em}
\]

(1)

\[
F_{\lambda_{exc},\lambda_{em}}(R.U.) = \frac{I_{\lambda_{exc},\lambda_{em}}(A.U.)}{A_{rp}}
\]

(2)

In the formula, \(I_{\lambda}\) is the spectral correction intensity of the Raman peak at the emission wavelength \(\lambda\). In this paper, \(A_{rp}\) is the integral value of the fluorescence intensity of emission ranges from 365–430 nm (excitation = 350 nm). \(F\) is the fluorescence intensity after Raman normalization. \(A.U.\) stands for the arbitrary unit, \(R.U.\) stands for Raman unit.

The pre-processed 3D fluorescence data can be used for PARAFAC analysis, while the processed data and the results of the PARAFAC analysis are normalized and converted into grayscale images, as is shown in the following Figure 2.

![Figure 2](image-url)

**Figure 2.** (a) Corrected data by Raman normalization; (b) Adjust the excitation ranges to: 200–550 nm, emission: 250–610 nm; (c) FU components (PARAFAC results); (d) F component; (e) PU components; (f) F components; (g) Grayscale image of 3D-EEMs components; and (h) Original 3D-EEMs grayscale image. This figure shows the data processing method in this paper, in which the grayscale images are generated by the samples EEMs and components EEMs, rather than the contour image, and pixels represent fluorescence intensity. In (b), the reason why the range is set as this range is to preserve the characteristic information of all 3D-EEMs. By using the bilinear interpolation method to interpolate 3D-EEMs, the size of the gray image is 60 × 60.
2.4. PARAFAC Method

We performed a PARAFAC analysis of the 3D fluorescence data using the drEEM toolbox v0.6. The principle of the PARAFAC analysis method is to decompose the 3D fluorescence data into a set of trilinear terms and a residual array structure [28] (as is shown in the following Figure 3), which is modeled by the process of minimizing the sum of squares of the residuals. Its mathematical expression is as follows:

\[ x_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} + \epsilon_{ijk} \]  \hspace{1cm} (3)

\[ i = 1, \ldots, I; j = 1, \ldots, J; k = 1, \ldots, K \]  \hspace{1cm} (4)

where, \( x_{ijk} \) is the fluorescence intensity of sample \( i \) at emission wavelength \( j \) and excitation wavelength \( k \). \( a_{if} \) is directly proportional to the concentration of the \( f \)—the analyte in sample \( i \). \( b_{jf} \) is linearly related to the fluorescence quantum efficiency of the \( f \)—the analyte in sample \( i \). \( c_{kf} \) is linearly proportional to the specific absorption coefficient at the excitation wavelength \( k \). \( F \) is defined as the number of components of the model and the residual matrix \( \epsilon_{ijk} \) represents the variability not considered in the model.

![Figure 3](image-url)  
Figure 3. 3D-EEMs are decomposed into PARAFAC components.

2.5. VGG16 Classification Network

In this study, VGG16 was chosen as the backbone of the 3D fluorescence samples classification network model, and the model diagram is shown below in Figure 4.

![Figure 4](image-url)  
Figure 4. 3D-EEMs classification network based on VGG16. The size of the image is interpolated from size (60 × 60) to size (224 × 224) using the bilinear interpolation method. Grayscale images are converted into 3-channel RGB images using the Python Pillow function: Image.convert (“RGB”), which uses the Floyd–Steinberg dither method [29].

The VGG16 network consists of 13 convolution layers and 3 fully connected layers. In this paper, the VGG16 model’s input size is 1 × 224 × 224 pixels (grayscale image), and
Through the previous data preprocessing section, the four types of 3D fluorescence datasets are adjusted to a unified excitation–emission interval, and the Raman normalization operation has been performed on samples. Although this operation was completed, since the excitation interval and emission interval of fluorescence measurement datasets have been adjusted to a unified excitation–emission interval, and the Raman normalization operation has been performed on samples. Although this operation was completed, since the excitation interval and emission interval of fluorescence measurement equipment were different during the measurement process of 3D-EEMs, we need to adjust the interval of four types of 3D-EEMs data to be consistent by using interpolation. After interpolation, the size of 3D-EEMs is uniformly transformed to 60 × 60. Figure 5 shows the changes in 3D-EEMs data before and after interpolation. The interpolation method uses bilinear interpolation and the mathematical expression is as follows:

\[
f(x) = \frac{x_2 - x}{x_2 - x_1} f(Q_{11}) + \frac{x - x_1}{x_2 - x_1} f(Q_{21})
\]

(6)

\[
f(x) = \frac{x_2 - x}{x_2 - x_1} f(Q_{12}) + \frac{x - x_1}{x_2 - x_1} f(Q_{22})
\]

(7)

where \( f(Q_{11}) \), \( f(Q_{21}) \), \( f(Q_{12}) \) and \( f(Q_{22}) \) represent the values of function \( f \) at 4 points: \( Q_{11} = (x_1, y_1), Q_{12} = (x_1, y_2), Q_{21} = (x_2, y_1), \) and \( Q_{22} = (x_2, y_2) \). In our experiment, \( x \) represents the position of emission–excitation wavelength, and \( f \) represents the fluorescence intensity at a certain position.

Figure 5. Comparison before and after 3D-EEMs interpolation. (a) Before interpolation; and (b) After interpolation. In (b), emission and excitation from 1 to 60 represent the number of wavelengths and excitation wavelengths, and the range of wavelengths is still consistent with that in (a).
After data normalization, use Python code to convert 3D-EEMs into grayscale images $(1 \times 60 \times 60)$. The emission interval is the length of the grayscale image, and the excitation interval is the width of the grayscale image, and the fluorescence intensity is the grayscale value. The transformation relationship is as follows:

$$y = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \times 255$$  \hspace{1cm} (8)

where $x$ represents the intensity value of fluorescence in 3D-EEMs, and $y$ represents the grayscale value after conversion to a grayscale image. The data are linearly mapped to the $[0, 1]$ interval using dispersion standardization. $x_{\text{min}}$ is the minimum fluorescence intensity, $x_{\text{max}}$ is the maximum fluorescence intensity.

2.6. CF-VGG11 Network

After using the VGG16 network to classify the 3D fluorescence, we can know the type and number of components of 3D-EEMs. Then we used the improved components fitting network CF-VGG11 (Components Fitting VGG11 network) as a model for 3D fluorescence components mapping analysis, which is used to fit the component maps of 3D fluorescence samples. The single model structure of the CF-VGG11 network used in this study is shown below in Figure 6a,b, which shows the structure of the entire network model in this paper. CF-VGG11 network contains six improved sub-VGG11 network structures (determined by the largest components of each type of 3D-EEMs), in which each sub-model is responsible for fitting one component of various 3D fluorescence data. Compared with the original VGG11 network structure, the CF-VGG11 network structure is optimized and adjusted in this paper by adding a dropout layer after the first fully connected layer and the second fully connected layer to enhance the stability of the model and prevent overfitting.

The same as the VGG16 used in this study, CF-VGG11 receives the input as a $1 \times 60 \times 60$ grayscale image and outputs a vector of size 3600 (it can be understood that the channel $\times$ length $\times$ width: $1 \times 60 \times 60$), and the vector represents the fluorescence intensity value of the component map. In this study, the four datasets used for CF-VGG11 are analyzed by PARAFAC analysis. FU, F, PU, and P have 4 components, 4 components, 5 components, and 6 components, respectively, then the components of each type of 3D-EEMs are numbered, respectively, from C1–C6, using components maps (C1–C6) as the target (label) values for CF-VGG11 training. In the CF-VGG11 model, the mean square error MSE was used as the loss function, the cosine similarity was used as the judgment standard of the model fitting results, and the Adam optimizer was used to optimize the model training [30].

In summary, the network model, and (3) training CF-VGG11 component fitting network model. A 3D-EEM, after removing Raman scattering, removing Rayleigh scattering, and using Raman normalization corrected, as a grayscale image inputting into the VGG16 classification experiment of using a convolutional neural network to analyze 3D fluorescence in this study includes three parts: (1) data preprocessing (PARAFAC analysis and preparation of input data for the CNN model), (2) training VGG16 classification network, combining with the PARAFAC analysis results, so the classification network can derive the category and the number of components. Based on the number of components obtained from the classification, the 3D-EEM, which is previously transmitted to the VGG16 network, is now transmitted to each sub-network of the CF-VGG11 network again, so that the component maps of the 3D-EEM can be obtained.
The process of analysis and component maps are shown in Figure 7. Components (FU), four components (F), five components (PU), and six components (P) were obtained, respectively. PARAFAC analysis is conducted separately for each type of dataset.

3. Results and Discussion

3.1. The Results of the PARAFAC Analysis

In this study, 3D-EEMs were analyzed by the PARAFAC method, which will obtain the correct component number and component maps of 3D fluorescence data [31]. We collected a total of four types of 3D fluorescence data online as well as through field sampling, including a total of 416 samples, with the least amount of FU datasets containing 45 samples. PARAFAC analysis requires a split-half test analysis, so a data volume greater than 21 is more appropriate [32].

After the analysis of these four types of datasets by PARAFAC analysis, four components (FU), four components (F), five components (PU), and six components (P) were obtained, respectively. PARAFAC analysis is conducted separately for each type of dataset. The process of analysis and component maps are shown in Figure 7.
In the experiment we performed PARAFAC analysis on all datasets (Figure 8 and Table 2 show the results of the PARAFAC analysis on FU dataset). Then, we uploaded the PARAFAC results to the OpenFluor database [33] for comparison and obtained the possible matter of these four 3D fluorescence components. Among them, the comparison results show that the similarity scores of components exceed 95% except component-three (94%).

Figure 7. PARAFAC analysis process and result (taking FU dataset as an example).

![PARAFAC analysis process and result](image)

Figure 8. The results of PARAFAC analysis on FU dataset. The loading maps on the right show the result of split-half validation.

Table 2. Spectral characteristics of EEMs from comparing result of PARAFAC components and OpenFluor database.

<table>
<thead>
<tr>
<th>Component</th>
<th>EX&lt;sub&gt;max&lt;/sub&gt;/Em&lt;sub&gt;max&lt;/sub&gt;</th>
<th>Description</th>
<th>Number of OpenFluor Matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>350 nm/440 nm</td>
<td>terrestrial humic-like, high relative aromaticity and molecular weight [34]</td>
<td>21</td>
</tr>
<tr>
<td>C2</td>
<td>300 nm/400 nm</td>
<td>UV-A humic-like, low molecular weight [35]</td>
<td>14</td>
</tr>
<tr>
<td>C3</td>
<td>290 nm, 400 nm/480 nm</td>
<td>humic-like [35]</td>
<td>4</td>
</tr>
<tr>
<td>C4</td>
<td>280 nm/320 nm</td>
<td>autochthonous protein-like, sensitive to microbial degradation [36,37]</td>
<td>11</td>
</tr>
</tbody>
</table>

The characteristics of C1 were characterized by a peak at 350 nm excitation with 440 nm emission wavelengths and associated with a group of high molecular weight and aromatic molecules of terrestrial origin [34]. The characteristics of C2 were characterized by a peak at 300 nm excitation with 400 nm emission wavelengths, which is sensitive to the UV-A
humic-like and low molecular weight [35]. The characteristics of C3 were characterized by a peak at 290 nm and 400 nm excitations with 480 nm emission wavelengths, which is similar to the humic-like [35]. The characteristics of C4 were characterized by a peak at 280 nm excitation with 320 nm emission wavelengths, which is similar to autochthonous protein-like, and sensitive to microbial degradation [36,37].

3.2. Training and Validation of the VGG16 Network and the CF-VGG11 Network

The training set and validation set used for VGG16 contains 333 samples and the test set contains 83 samples, at a ratio of 6:2:2 (Table 3). In the actual training process, we expanded the images by gamut distortion and mirror flipping. Although the transformed images cannot represent the 3D fluorescence characteristics, it can improve the training efficiency of the model and prevent the poor training effect due to the lack of data. The VGG16 classification network was evaluated by using the cross-entropy loss function and training accuracy. We trained the network with 100 epochs and set the initial learning rate of the first 50 epochs to $1.0 \times 10^{-3}$ and the last 50 epochs to $1.0 \times 10^{-4}$. We use the stepLR method to automatically update the learning rate. The optimizer adopts the Adam optimizer, with the step parameter set to one and the gamma parameter set to 0.94. In addition, the input size of the network is 224 × 224, the output is 4 classes, the batch_size of the first 50 epochs is set to 8, and the batch_size of the last 50 epochs is set to 16. During the training process, the training loss and validation loss decrease gradually, and the training loss value tends to be stable after 80 epochs. The training loss and validation loss curve is shown in Figure 9. After 5 repeat experiments, as the training logs, we take the best epoch of all as the training result. The loss value is: $2.39 \times 10^{-2}$ (training loss) and $1.03 \times 10^{-3}$ (validation loss). Then, we tested the accuracy of the model using the test set samples and the top-5 accuracy is 100%. In the dataset, the test set has a small amount and few types, so the appearance of the test accuracy is good. This is a normal training result since the network can always distinguish the 3D-EEM types well. At the same time, we used the PCA + SVM method (Figure 10) to reduce the dimension and classify 3D-EEMs with an accuracy of 99.69%. However, the PCA + SVM approach has the same requirements as the PARAFAC analysis method, which requires a large amount of data as inputs, so the convolutional neural network has greater advantages (Table 4). The training results show that there are significant differences between different classes of 3D-EEMs and using the convolutional neural network can easily identify the type of 3D-EEMs and the number of their components.

Table 3. Samples division table.

<table>
<thead>
<tr>
<th>Samples</th>
<th>Number</th>
<th>Train</th>
<th>Validate</th>
<th>Test</th>
<th>Total Samples after Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>FU</td>
<td>45</td>
<td>27</td>
<td>9</td>
<td>9</td>
<td>135</td>
</tr>
<tr>
<td>F</td>
<td>105</td>
<td>63</td>
<td>21</td>
<td>21</td>
<td>315</td>
</tr>
<tr>
<td>P</td>
<td>206</td>
<td>124</td>
<td>41</td>
<td>41</td>
<td>618</td>
</tr>
<tr>
<td>PU</td>
<td>60</td>
<td>36</td>
<td>12</td>
<td>12</td>
<td>180</td>
</tr>
<tr>
<td>Total</td>
<td>416</td>
<td>250</td>
<td>83</td>
<td>83</td>
<td>1248</td>
</tr>
</tbody>
</table>

Table 4. Comparison results between VGG16 and PCA + SVM.

<table>
<thead>
<tr>
<th>Method</th>
<th>Training Accuracy (%)</th>
<th>Sample Demand</th>
<th>Speed</th>
<th>Operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGG16</td>
<td>99.60</td>
<td>≥1</td>
<td>Fast</td>
<td>Easy</td>
</tr>
<tr>
<td>PCA + SVM</td>
<td>99.69</td>
<td>Many</td>
<td>Slow</td>
<td>Complex</td>
</tr>
</tbody>
</table>
Figure 9. Loss curve and training accuracy curve of VGG16. In (a), when epoch > 80, the training loss and validation loss are both stable. In (b), the training accuracy reaches 99.6% at epoch 97 (the accuracy also reaches 99.6% at epoch 80 and epoch 81).

Figure 10. In the PCA experiment, the cumulative of the first three PCs is 87.77%, and the cumulative of the first four PCs is 91.73%. The classification accuracy of SVM is 99.69. The reason for the error is that there are some similar characteristics between PC1 and PC2 of F and PU.

The CF-VGG11 network structure contains six VGG11 network model structures (Figure 5), where each submodel gets a corresponding one component map of each of four datasets based on 3D-EEMs. We used the VGG11 model to fit a component map, which is a vector of size $60 \times 60$ (it is not an output class), we could not judge the results by the method of traditional target detection and object recognition algorithms, so we used cosine similarity [38] to replace the accuracy of the model fitting results. We divided the data in a ratio of 3:1:1 (250 samples for training and 83 samples for validation and 83 samples for test) and input each 3D-EEM datum into each sub-model while the fitting component map of each model is given by the PARAFAC analysis results. The results of model training are shown in Figure 11.
Figure 11. (a,b) show the loss curves, (c,d) show the cosine similarity of model training. In (a,b), the comparison of loss values, the VGG11 model is better than AlexNet model, while the training loss of LeNet is smaller than that of VGG11 model, but the validation loss of VGG11 model is smaller. A smaller loss value means better network performance. Although these 3 models perform well, the VGG11 model has better and more stable performance. (c) shows the change of cosine similarity during model training. VGG11 network shows the best performance of the 3 models. Through the histogram (d) we can see the training change law in the process of 30–40 epochs.

3.3. Model Performance

To show that the trained VGG16 and CF-VGG11 models have a good fitting performance, we selected some 3D fluorescence data for testing, and the components fitted by the CF-VGG11 model were compared with those analyzed by PARAFAC (Figure 12). It can be seen from the figure that the results fitted by the model are very similar to those obtained by the PARAFAC analysis method, with the difference that there are a few noise points in the results fitted by the model, and these noise points do not have much effect on the components of 3D-EEMs. Therefore, we can conclude that VGG16 and CF-VGG11 can be well classified and fitted according to 3D-EEMs.
In recent years, many people have tried to use 3D fluorescence data for the rapid analysis of river water \cite{39,40} to characterize the DOM pollution of river water. The model based on this study can achieve the purpose of rapid analysis of river water.

3.4. Comparison between the VGG Network and the PARAFAC Analysis Method

Our convolutional neural networks are based on the PARAFAC analysis method. We need the PARAFAC analysis method to provide the correct number of components and component maps, and the network is trained on these correct data for the final purpose of 3D-EEMs classification and component map fitting.

In addition, our models of VGG16 and CF-VGG11 have some advantages over the PARAFAC analysis in that the models require only one 3D-EEM input to obtain the correct sample category and its component maps. This is of great application value in scenarios where rapid 3D-EEMs analysis needs to be established. The specific comparison between PARAFAC and our model is shown in Table 5.

<table>
<thead>
<tr>
<th>Model</th>
<th>Train</th>
<th>Progress</th>
<th>Data</th>
<th>Skills</th>
<th>Integrate to Apps</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARAFAC</td>
<td>No</td>
<td>Complex</td>
<td>≥21</td>
<td>Yes</td>
<td>Complex</td>
</tr>
<tr>
<td>VGG16+CF-VGG11</td>
<td>Yes</td>
<td>Easy</td>
<td>≥1</td>
<td>No</td>
<td>Easy</td>
</tr>
</tbody>
</table>

In conclusion, we use VGG16 to accurately classify 3D-EEMs and use the CF-VGG11 model to realize the component maps fitting different kinds of 3D-EEMs. The model relies on PARAFAC analysis for preliminary data preparation, but it can overcome the problem of strict data requirements of PARAFAC in the process of use and provides a new analytical tool for the real-time 3D fluorescence analysis of small samples.

**Author Contributions:** Methodology, S.Z. and X.J.; Project administration, K.R.; Resources, Y.L. and T.L.; Supervision, J.F., D.O., Q.T. and J.X.; Writing—original draft, K.R.; Writing—review & editing, K.R., X.J. and Z.L. All authors have read and agreed to the published version of the manuscript.
References


3. Duan, P.; Wei, M.; Yao, L.; Li, M. Relationship between non-point source pollution and fluorescence fingerprint of riverine dissolved organic matter is season dependent. Sci. Total Environ. 2022, 823, 153617. [CrossRef] [PubMed]


