



Artificial intelligence classification of wetland vegetation morphology based on deep convolutional neural network

Ping Lin^{1,2} | Qun Lu^{1,2} | Du Li^{1,2} | Yongming Chen^{1,2} |
Zhiyong Zou³ | Shanchao Jiang^{1,2}

¹College of Electrical Engineering,
Yancheng Institute of Technology,
Yancheng, China

²Key Laboratory for Advanced
Technology in Environmental Protection
of Jiangsu Province, Yancheng Institute
of Technology, Yancheng, China

³College of Mechanical and Electrical
Engineering, Sichuan Agricultural
University, Ya'an, China

Correspondence

Yongming Chen, College of Electrical
Engineering, Yancheng Institute of
Technology, No. 1 Middle Road Hope
Avenue, Yancheng, 224051 Jiangsu,
China.

Email: billrange007@gmail.com

Funding information

National Natural Science Foundation of
China, Grant/Award Numbers: 31501221,
31601227, 61803325; Jiangsu Government
Scholarship for Overseas Studies, Grant/
Award Number: JS-2015-065; Natural
Science Foundation of Jiangsu Province,
Grant/Award Numbers: BK20161310,
BK20181049; Natural Science Foundation
of the Jiangsu Higher Education
Institutions of China, Grant/Award
Number: 18KJB510046

Abstract

In real-world wetland vegetation morphology (WVM) detection, large scene variations such as those due to landform, vegetation, sunlight, weather, and sky, as well as camera parameter settings such as focal length and shooting angle, require systematic and complicated artificial intelligence technology to accurately discriminate inter and intra-class wetland objections. To deal with these challenges, we introduced a deep-level discriminative model based on convolutional neural networks (CNN) for classifying the images of DongZhai Harbor intertidal, Lashi Lake alpine, Yancheng coastal and Zoige plateau wetlands in China. A 96-dimensional convolution operation with kernel sizes of $11 \times 11 \times 3$ first applied to the resized 227×227 WVM input pictures to acquire the effective morphologic features. The perceptron layers of the rectified linear unit and the batch normalization were used in the middle layer to achieve the better gradient propagation property during the training process. The WVM features were down-sampled by the pooling networks to reduce the neuron dimensions. The fully connected layer linked to the output of the convolutional and pooling layers to obtain the high-level WVM species information for the final WVM classification purpose. The deep-level CNN-based method was compared with



the traditional shallow-level feature-designed algorithms of conditional maximum entropy regression, multilayer perceptron and support vector machine. The deep-level algorithm showed the superior performance for detecting the WVM species, which provided a superior alternative routine for accurate artificial intelligence classification of WVM in the ecological engineering application.

Recommendations for Resource Managers

- This proposed artificial intelligence model can be used to effectively help people understand the macroscopic scene changes of ecological elements from the wetland vegetation morphology images.
- The detection results show the robustness of the discriminative model, for it is not affected by the complex natural factors such as landform, vegetation, sunlight, weather, and sky, as well as human factors such as camera parameter settings of focal length and shooting angle to some extent.
- The deep-level model based on the cascade convolution framework can be used to extract more distinguishing detailed ecological features from the original wetland vegetation morphology images compared with the shallow-level algorithms using a single convolutional layer.

KEY WORDS

classification, convolutional neural network, ecological engineering, vegetation morphology, wetland

1 | INTRODUCTION

Wetlands are areas where the presence of water determines or influences most, if not all, of an area's biogeochemistry. Wetland plays very important roles in the environment, principally water purification (Saggaï, Ainouche, Nelson, Cattin, & El Amrani, 2017; Weller & Bossart, 2017), flood control (Dube et al., 2017; James, Capon, & Quinn, 2015), shoreline stability (Davies-Vollum & West, 2015; Patrick, Weller, & Ryder, 2016), and so forth. Wetland is also considered as the most biologically diverse of all ecosystems, serving as home to thousands of animals and plants (Brandt, Petersen, Grossman, Allen, & Benzing, 2015; Sirami, Jacobs, & Cumming, 2013). The wetland covers a total area of 6,600 million hectares in Chinese territory. The regions account for more than

10% of the world's wetland and rank first in Asia. Therefore, the study and conservation of the ecological environment, balance, and biological diversity of Chinese wetlands make remarkable



impacts on global ecosystems (T. Sun, Lin, Chen, Guo, & Zeng, 2016; X. Sun, Xiong, & Zhu, 2015; R. Yin & Yin, 2010). Despite the huge impact on our everyday production and life, loss and damage to wetlands and their biodiversity continue. Various ecological regions have been seriously destroyed by the rapid growth of the social economy through the occupations of wetlands for agricultural land (Bai et al., 2013; Murray, Clemens, Phinn, Possingham, & Fuller, 2014), community facilities (Kirwan & Megonigal, 2013; Zhang et al., 2015), tourism development (Lee & Hsieh, 2016; Wu, Gao, Wang, Wang, & Xu, 2015), and so forth in recent decades all over the world. Moreover, the uncontrollable uses of wetland water resources for those developments lead to a further threat (G. Liu, Tian, Sun, Xiao, & Yuan, 2016; Lyu, Chen, Zhang, Fan, & Jiao, 2016). More than 65% of water resources are extracted from the wetlands in industrial and agricultural production, which seriously affects the physical and chemical properties of soil (Leung et al., 2017; Meng et al., 2017) and alters the wetland vegetation drastically (Feng, Han, Hu, & Chen, 2016; Roberts, Hunt, Arroyo-Kalin, Evans, & Boivin, 2017). The study of the classification of wetland vegetation morphology (WVM) images can provide a new way of measuring the macroscopic scene changes of ecological elements. It will help reduce chemical pollution and limit invasive species in the ecological regions. The technology of remote sensing is generally applied to analyzing the status of WVM, in which the sensors are too far away from the objects. Therefore, the captured pictures can only reflect a rather rough view for the survey of the wetland vegetation (Aslan, Rahman, Warren, & Robeson, 2016; Jetz et al., 2016; H. Yin, Khamzina, Pflugmacher, & Martius, 2017; Zheng et al., 2017). Even though the WVM pictures are taken by near-ground digital cameras where richer structure and texture information from the WVM scenes are obtained, effective methods to analyze such fine-scale morphologic images are rarely reported.

To deal with these challenges, we introduced a deep-level cascade discriminative model based on convolutional neural networks (CNN) architecture (Esteva et al., 2017; LeCun, Bengio, & Hinton, 2015; Wang, Peng, Ma, & Xu, 2016) for identifying the WVM in the Ramsar sites list. There were four different classes of wetlands in our investigation, including DongZhai Harbor (DZH) intertidal mangrove wetland (Tang et al., 2014; Xi, Li, Xia, & Qu, 2016), Lashi Lake (LSL) alpine peat wetland (Y. C. Huang, Tian, Yue, Liu, & Lai, 2012; Liao, Ye, Huang, & Peng, 2017), Yancheng (YC) coastal saline wetland (L. Huang et al., 2015; Zang et al., 2017), and Zoige plateau freshwater lake wetland (Jiang, Lv, Wang, Chen, & Liu, 2017; Li et al., 2014) in China. We followed deep-level AlexNet CNN network structures (Krizhevsky, Sutskever, & Hinton, 2012; Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014) to construct the WVM identification model. A total of 25 deep layers of neuron architecture was set up. It mainly includes the convolutional, spatial pooling, ReLU, dropout, and fully connected neurons. An anti-oscillatory stochastic gradient descent (SGD) method was used to modify the weight parameters to avoid the formidable overfitting problem during the training procedure. The deep-level model was compared with several traditional shallow-level detection models of Gabor-kernel-based convolution architecture combined with the conditional maximum entropy regression (CMER; Yu, Huang, & Lin, 2011), scaled conjugate gradient (SCG) multilayer perceptron (MLP; Schmidhuber, 2015), and multiclass support vector machine (SVM; Chang & Lin, 2011). The deep-level algorithm showed the superior performance for distinguishing the WVM species.

We aimed to create a robust deep-level CNN-based architecture for accurately predicting the WVM species. The main contribution is that we first develop the deep-level learning model of CNNs for the robust identification of WVM images in the ecological application. The follow-up

article structure is as follows. The experimental data and devices are introduced in Section 2. The experimental methods are described in Section 3. Then, the experimental results are analyzed and discussed in Section 4. Finally, the conclusions are drawn in Section 5.



2 | EXPERIMENTAL DATA AND DEVICES

2.1 | Characteristics of wetlands

Four classes of WVM listed in the Ramsar sites including DZH, LSL, Zoige, and YC wetlands are investigated. A geographical map of four typical wetland can be seen in Figure 1. Several main characteristics of each wetland are described as follows:

- DZH wetland is located at the northern end of Hainan island and the geographical coordinates are $19^{\circ}51' \sim 20^{\circ}01'N$, $110^{\circ}30' \sim 110^{\circ}37'E$, 0 m. It consists principally of one ecological unit—tropical, intertidal mangrove forest with associated mud- and sand-flats and shallow water subzones. Of these subzones, more than 2,000 hectares are covered by the mangrove forest and more than 1,330 hectares are intertidal mud- and sand-flats. The ecological elements of mangrove forest, mud, sand, river, and sea mainly become the discriminative morphologic characteristics (Tang et al., 2014; Xi et al., 2016).
- LSL wetland is located on the west side of Lijiang County in northwestern Yunnan Province and the geographical coordinates are $26^{\circ}44' \sim 27^{\circ}00'N$; $100^{\circ}05' \sim 100^{\circ}13'E$; 2,441–3,100 m. With an area of 5,330 hectares, it lies on the southern slope of the Jade Dragon Snow Mountain. The freshwater comes from the Meiquan and Keluokang rivers and the melted snow from the nearby mountains, which forms about 256 m^2 catchments. It belongs to the plateau freshwater lake type. The crisscrossed mountain, basin, and lake form the complex landforms. The ecological elements of mountain, lake, shrub, grassland, and forest mainly become the discriminative morphologic characteristics (Y. C. Huang et al., 2012; Liao et al., 2017).
- YC wetland is the largest intertidal coastal saline wetland in the world. It is located in the east part of Jiangsu Province and the geographical coordinates are $32^{\circ}34' \sim 34^{\circ}28'N$, $119^{\circ}48' \sim 121^{\circ}56'E$, 0–4 m. With an area covers 453,000 hectares, it faces the south of the Yellow Sea. The rivers

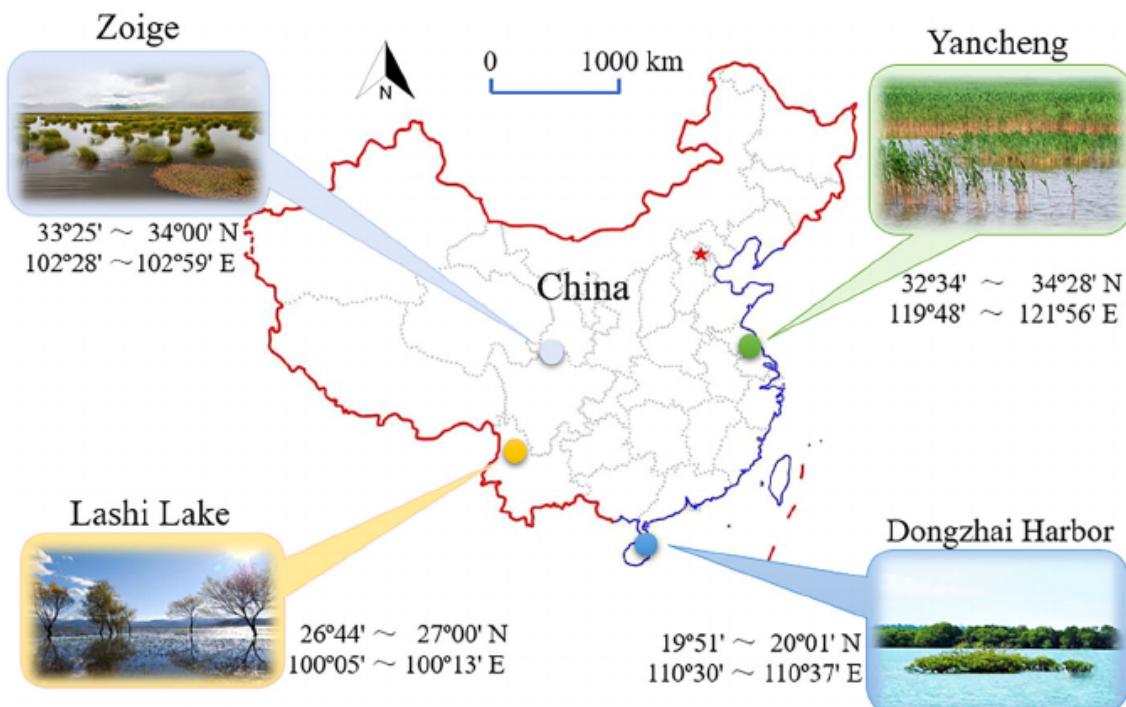


FIGURE 1 A geographical map of four typical wetlands in China. Dongzhai Harbor wetland consists principally of the typical ecological unit—tropical, intertidal mangrove forest. Lashi Lake wetland belongs to the plateau freshwater lake type. Yancheng wetland is the largest intertidal coastal saline wetland in the world. Zoige wetland contains the largest high-altitude peat-substrate swamp area in the world



and catchments are scattered around the areas. The tidal vegetation areas can be divided into four categories of the mud, salt living wormwood, prevailing lalang grass, and dominated reed belts. The ecological elements of sea, river, mud, catchment, shrub, and grassland mainly become the discriminative morphologic characteristics (L. Huang et al., 2015; Zang et al., 2017).

- Zoige wetland contains the largest high-altitude peat-substrate swamp area in the world. It is located at the upstream of the Yellow River in the northern Sichuan Province and extends into southern Gansu Province and southeastern Qinghai Province and the geographical coordinates are $33^{\circ}25' - 34^{\circ}00'N$; $102^{\circ}28' - 102^{\circ}59'E$; 3,422–3,704 m. With an area of 260,000 hectares, it lies between the west of the Min Mountains and the east of the Amne Machin mountain. It is distributed on flat and wide beaches, lakes and terraces. The mountain, river, lake, marsh, semi-marsh, shrub, and grassland form the landscape patterns (Jiang et al., 2017; Li et al., 2014). The attributes of WVM are mainly influenced by the plant species and geographic characteristics and form the discriminated characteristics of the captured ecological vegetation images (Chen, Lin, He, & He, 2015; Junk et al., 2014; Marani, Da Lio, & D'Alpaos, 2013).

2.2 | Experiment data

These WVM pictures were taken from digital cameras. The captured work was carried out in the daytime (7.00 a.m.–5.00 p.m.) throughout all the seasons when the sunlight was sufficient. There was a vast absolute distance between the observer and the fixated objects, usually greater than 6 m. The real distance and the aperture of the cameras were varied for the used images. Most of the images were in focus and with realistic colors, thereby, there were not any blur images used for measurements. The images are randomly chosen for the modeling target and the rest of 60 images are used for the test purpose. A total of 1,200 captured WVM images were used for modeling the relationship between the WVM features and the corresponding logical ecological attributes. The representative WVM images can be seen in Figure 2. Each class of experimental WVM data set has 300 instances, where 240 instances are used for training tasks and the rest of 60 instances for test aims for each type.

2.3 | Experimental devices

The test machine is a Dell Alienware 17 R4 notebook with the hardware solution of Intel Core (TM) i7-6700H CPU and the software solution of the Microsoft Windows 10 system. The implementation of CNN architecture for validating the performance of the discrimination algorithm is based on the Matlab (The Mathworks, Inc., Natick, MA) application platform. The deep-level cascade discriminative framework is based on Caffe explored by Berkeley vision and learning center (Jia et al., 2014; Vedaldi & Lenc, 2015). To boost the efficiency for training the CNN model, the graphics processing units of NVIDIA GPU with 8 GB memory and 1,024 kernels (NVIDIA GeForce GTX 1070) with application programming interface of CUDA 8 is used to implement the algorithms. It also makes use of the NVIDIA library of CUDNN 7 installed to gain significant speed and space benefits.

3 | METHODS

3.1 | Image data preprocessing

3.1 | Image data preprocessing

Because the number of units and the operation amount of the CNN directly correcting with the number of image pixels, the aspect ratios of all original WVM images are adjusted to the



FIGURE 2 The snapshots of four rows of vegetation morphology represent the four different categories of wetlands including DongZhai Harbor mangrove wetland (a), Lashi Lake plateau freshwater Lake wetland (b), Yancheng coastal saline wetland (c), and Zoige alpine peat wetland (d), in China. The integrated complex ecological components mainly consisting of beaches, muds, mires, grasslands, shrubs, flowers, lakes, rivers, mountains, flowers, trees, forests, and so on, which form the current independent wetland vegetation systems

uniform square sizes of 227×227 pixels. The similar image resolutions are also applied to other traditional feature-designed based algorithms of CMER, SCG-MLP, and multiclass SVM. Due to the collected WVM images having the complex objects and backgrounds, the data may distribute in an arbitrary large space. To extract the useful feature information from the WVM image data set and enhance the learning ability of our model, it is extremely important to deal with these images before feeding them into our model. To adapt to data variation in the learning models, the techniques of removing image mean and standardization are carried out. The first one removed the mean by subtracting the mean for each pixel across all training examples. The latter controls the spread of pixel values by dividing the standard deviation. The implementation of data preprocessing will be helpful for further improving the performance of detector (Esteva et al., 2017; LeCun et al., 2015; Wang et al., 2016).

3.2 | CMER

Given a training data set of N points of the form $\{(x_i, y_i) | x_i \in \mathbb{R}^p, y_i \in \{1, 2, \dots, M\}\}_{i=1}^N$, CMER (Yu et al., 2011) is used to estimate the likelihood $p(y | x)$ by discriminating among the different possible values of the logical ecological class y from the given observation of WVM features x :

$$p(y | x) = \frac{\exp\left(\sum_{i=1}^N w_i \cdot g_i(x, y)\right)}{N}, \quad (1)$$

$$\textstyle \sum_{y' \in Y} \exp\Bigl(\sum_{i=1}^I w^i \cdot g^i(x,y')\Bigr)$$



where, \exp denotes the exponential function, g_i is the feature function, and w_i is the weight. CMER will find the parameter \hat{w} that maximizes the following probability:

$$\begin{aligned}\hat{w} &= \arg \max_w \sum_{j=1}^M \log p(y_j | x_j) \\ &= \arg \max_w \sum_{j=1}^M \log \exp\left(\sum_{i=1}^N w_i \cdot g_i(x_j, y_j)\right) - \sum_{j=1}^M \log \sum_{y' \in Y} \exp\left(\sum_{i=1}^N w_i \cdot g_i(x, y'_j)\right).\end{aligned}\quad (2)$$

Then, the class can be chosen in terms of the maximum probability:

$$\hat{y} = \arg \max_{y \in Y} \frac{\exp\left(\sum_{i=1}^N \hat{w}_i \cdot g_i(x, y)\right)}{\sum_{y' \in Y} \exp\left(\sum_{i=1}^N \hat{w}_i \cdot g_i(x, y')\right)}. \quad (3)$$

3.3 | SCG-MLP

MLP method utilizes the SCG algorithm to predict the unknown WVM features to a logical ecological category (Schmidhuber, 2015). The SCG algorithm can avoid the traditional time-consuming line-search per learning iteration by using a step size scaling mechanism. The SCG-MLP generates a search direction that is mutually conjugate to the previous search directions, with respect to a given positive definite matrix Θ and looks for the optimal location in that direction using a line searching technique. Two search directions δ_m and δ_n are mutually conjugate to Θ if the following condition is satisfied:

$$\delta_m^T \Theta \delta_n \text{ where } n \neq m. \quad (4)$$

The new search direction can be determined by using the previous direction and current gradient ξ_i :

$$\delta_i = \begin{cases} -\xi_i & , i = 0, \\ -\xi_i + \lambda_i \delta_{i-1} & , i > 0, \end{cases} \quad (5)$$

where δ is the i th conjugate direction and λ is computed as follows:

$$\lambda_i = \frac{\xi_i^T (\xi_i - \xi_{i-1})}{\delta_{i-1}^T (\xi_i - \xi_{i-1})}. \quad (6)$$

3.4 | Multiclass SVM

Given a training data set of N points of the form $\{(x_i, y_i) | x_i \in \mathbb{R}^p, y_i \in \{1, 2, \dots, M\}\}_{i=1}^N$, multiclass SVM (Chang & Lin, 2011) trains the i th class with all the training instances with positive labels, and the other classes with negative labels. The decision hyperplane $f_i(x) = \omega_i^T \varphi(x) + b_i$ is used to divide the sample set by solving the following optimization tasks:



$$\begin{aligned} \text{minimize: } & \Omega(\omega, \zeta_j^i) = \frac{1}{2} \|\omega_i\|^2 + C \sum \zeta_j^i, \\ \text{subject to: } & 1 - \zeta_j^i \leq \hat{y}(\omega_i^T \varphi(x) + b_i), 0 \leq \zeta_j^i, \end{aligned} \quad (7)$$

where C is the tuning parameter and ζ_j^i is the slack variable. If y_j belongs to the i th class, $\hat{y}_j = 1$, otherwise $\hat{y}_j = -1$. Finally, the i th class to which an unknown instance x belongs can be determined according to the corresponding largest value of $f_i(x)$:

$$\hat{i} = \arg \max_{i=1,2,\dots,M} f_i(x) = \arg \max_{i=1,2,\dots,M} \omega_i^T \varphi(x) + b_i. \quad (8)$$

3.5 | CNN structures

The CNN detector consists of a variety of layers such as the input, convolution, rectified linear units, batch normalization, pooling, dropout, and fully connected layers (Krizhevsky et al., 2012; Srivastava et al., 2014). The input layer receives a square, zero-mean and standardized WVM input pictures. The perceptron layers of the rectified linear unit and batch normalization are used in the middle layer to achieve better gradient propagation property during the training process. The WVM features are down-sampled by the pooling networks to reduce the neuron dimensions. The dropout layers give up nearly 50% of hidden nodes disposed before the fully connected layer to avoid the overfitting occurrence. The fully connected layers connect to the output of the convolutional and pooling layers to obtain the high-level WVM species information for the final WVM classification purpose. Each node in the fully connected layer will be linked to all the activations in the previous nodes. The proposed overall CNN structures with input, feature extraction and output for discriminating the WVM images can be seen in Figure 3.

3.6 | Anti-oscillatory SGD method

The standard gradient descent algorithm computes the cost and gradient for the entire training set. The SGD simply computes the gradient of the parameters using only a few stochastic training examples, which updates the parameter vector ξ of the loss function $L(\xi)$ given by

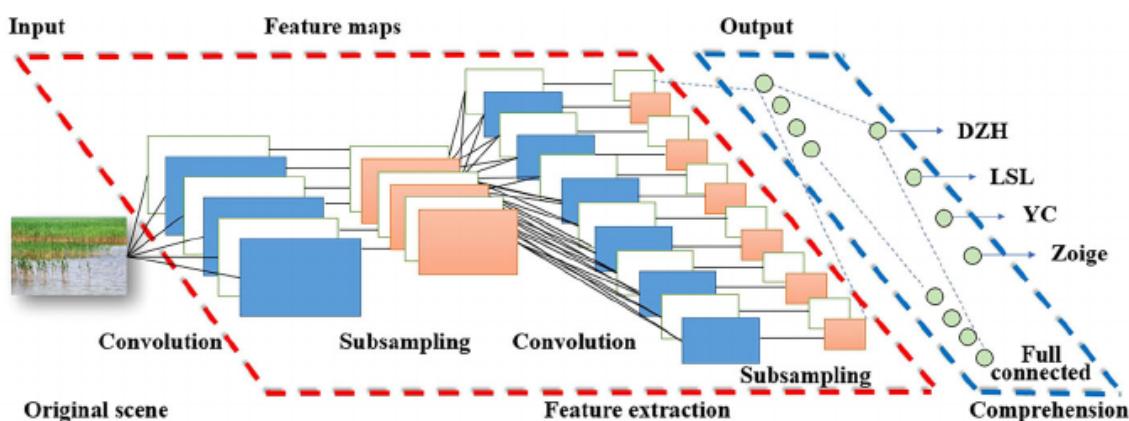


FIGURE 3 The proposed overall convolutional neural networks structures with input, feature extraction and output for discriminating the wetland vegetation morphology (WVM) images. Data flow is from left to right:

raw image pixels are used for analysis at the input level on the left, the feature extraction stage is performed in the middle red dashed rectangle and the classified WVM attributes are completed in the right blue dashed rectangle. DZH, DongZhai Harbor; LSL, Lashi Lake; YC, Yancheng



$$\xi_{i+1} = \xi_i - \mu \nabla L(\xi_i), \quad (9)$$

where μ is the learning rate and i denotes the iteration number. In this paper, a batch of 20 stochastic examples is used for training the CNN model. The classical algorithm of SGD tends to oscillate along the steep edges when the optimum is located at the bottom of the narrow steep ravine. The convergence speed of the CNN algorithm with the standard SGD becomes worse and sometimes it is hard to attain the global optima. To avoid this phenomenon, the anti-oscillatory coefficient of momentum η is added to the previous SGD function and renewed as follows:

$$\xi_{i+1} = \xi_i - \mu \nabla L(\xi_i) + \eta (\xi_i - \xi_{i-1}), \quad (10)$$

The anti-oscillatory SGD method can overcome the high cost of running backpropagation over the full training set, thus leads to fast convergence in the CNN setting (Andrychowicz et al., 2016).

3.7 | Training networks

Their size of the filters in the first layer is usually considered to be the most sensitive parameter. The routine determination of the parameters is in the region from 7×7 to 15×15 when the image sizes are around 200×200 to 800×800 (Z. Liu, Gao, Yang, Zhang, & He, 2016). Generally, the small convolutional kernels might catch the texture properties, while the large ones attain structure properties. Intuitively, the middle size kernels show the tradeoff between these two tendencies. In this investigation, the first 96 kernel features of the size of $11 \times 11 \times 3$ are employed to acquire the effective morphologic information from the WVM pictures, which exhibits the superior ability in discriminating the WVM species. Figure 4 illustrates the first 96 kernel features of size $11 \times 11 \times 3$ learned by the first convolutional layer on the input $227 \times 227 \times 3$ true-color images of WVM. These images mostly contain a variety of frequency- and orientation-selective kernels, as well as various colored blobs, which indicates



FIGURE 4 Illustrate the first 96 kernel features of size $11 \times 11 \times 3$ learned by the first convolutional layer

on the input $227 \times 227 \times 3$ true-color images of wetland vegetation morphology. These images mostly consist of edges and colors in the network architecture, which indicates that the convolutional filters at the second convolutional layer are edge detectors and color filters



that the filters at the first convolutional layer are edge detectors and color filters. The edge detectors are at different angles, which allows the network to construct the more complex structure and texture features of WVM in the later layers.

3.8 | Evaluating momentum coefficient

Figure 5 illustrates the results of the anti-oscillatory SGD method with five different momentum coefficients of $\eta = 0.0, 0.3, 0.6, 0.9$, and 1.0 affecting the convergence of the CNN model in the training optimization process on the WVM data set. The momentum term takes into account the historical contribution of gradient changes. While the momentum coefficients increase from $0.0, 0.3, 0.6-0.9$, the convergence of the CNN training model is improved. It accounts that the added momentum item helps improve the prediction performance of the CNN model by searching the global optimal solution. When the momentum coefficient is set as 1 , the convergence of the loss function continues to improve at the beginning stage where the iteration numbers are less than 200 . However, the loss values become larger after the 200 iterations. Thereafter, the anti-oscillatory SGD with the momentum coefficient $\eta = 0.9$ will be used to train the CNN model.

4 | RESULTS AND DISCUSSION

4.1 | Accuracy by CNN

The training process itself performed species classification on 960 WVM images and the rest of 240 WVM images are used for the validation purpose. The confusion matrix diagram (Deng, Liu, Deng, & Mahadevan, 2016) is a metric used for the visualization of the performance of the CNN classification model. The confusion matrix diagrams are illustrated in Figure 6 for classifying four types of WVM instances based on the deep-level architecture of CNN. Each column and row represent the target and output classes, respectively. The parameters in the

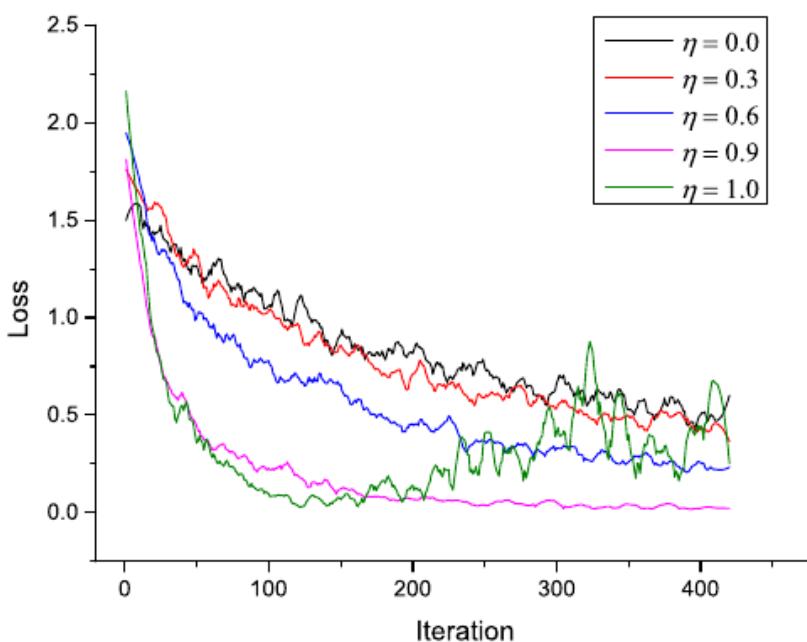


FIGURE 5 The anti-oscillatory stochastic gradient descent method with five different momentum coefficients of $\eta = 0.0, 0.3, 0.6, 0.9$, and 1.0 affect the convergence of convolutional neural networks model in the training optimization process on the wetland vegetation morphology data set

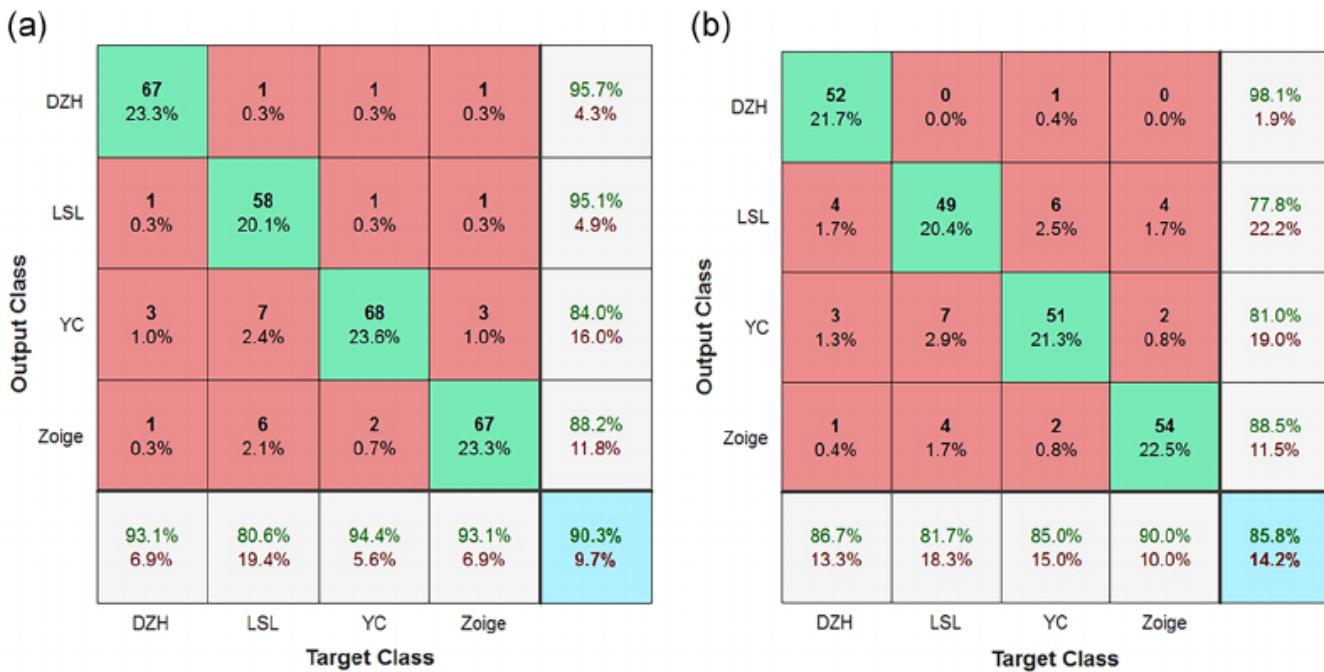


FIGURE 6 Confusion matrix diagrams of classification of wetland vegetation morphology images based on deep learning tricks of convolutional neural networks on the training (a) and test (b) data set, respectively

green diagonal box indicate the correct numbers and accuracy predicted by the CNN algorithm. For the training set, 67, 58, 68, and 67 objects are correctly identified as the WVM of DZH, LSL, YC, and Zoige, respectively. These correspond to 23.3%, 20.1%, 23.6%, and 23.3% of all 288 WVM, respectively. Similarly, for the test set, 52, 49, 51, and 54 classes are correctly detected as WVM of DZH, LSL, YC, and Zoige, respectively. These correspond to 21.7%, 20.4%, 21.3%, and 22.5% of all 240 WVM, respectively. The parameters in the red non-diagonal box indicate the false predicted numbers and rates by the CNN algorithm. The parameters in the white box on the far right of the diagram indicate the accuracy for each output class by the CNN algorithm. The parameters in the white box at the bottom of the diagram indicate the accuracy for each target species by the CNN algorithm. The overall accuracy is shown in the bright blue diagonal box. Overall, the accuracy of 90.3% and 85.8% of the training classes are true and the rest of 9.7% and 14.2% test classes are false.

4.2 | Comparing the performance of algorithms

The precision-recall curves are used for evaluating the prediction performance of scoring classifiers. The recall is a performance measure of the whole positive part of a data set, whereas precision is a performance measure of positive predictions. The precision-recall curve attempts to reach the balance between precision and recall for different thresholds. Figure 7 shows the precision-recall curves of the classification of WVM images based on deep-level learning architecture of CNN and three different kinds of shallow-level learning models of Gabor-kernel-based convolution architecture combined with the CMER, SCG-MLP, and multiclass SVM on the training and test data set. As shown in Figure 7, as the threshold of recall rates increases, the corresponding precision rates of CNNs are much higher than the other three algorithms of

CMER, SCG-MLP, and multiclass SVM. The overall performance of the algorithms is measured with the mean average precision (mAP) score, which is the average precision at the ranks where recall changes. The geometric interpretation of the mAP score is the area below the curve. A

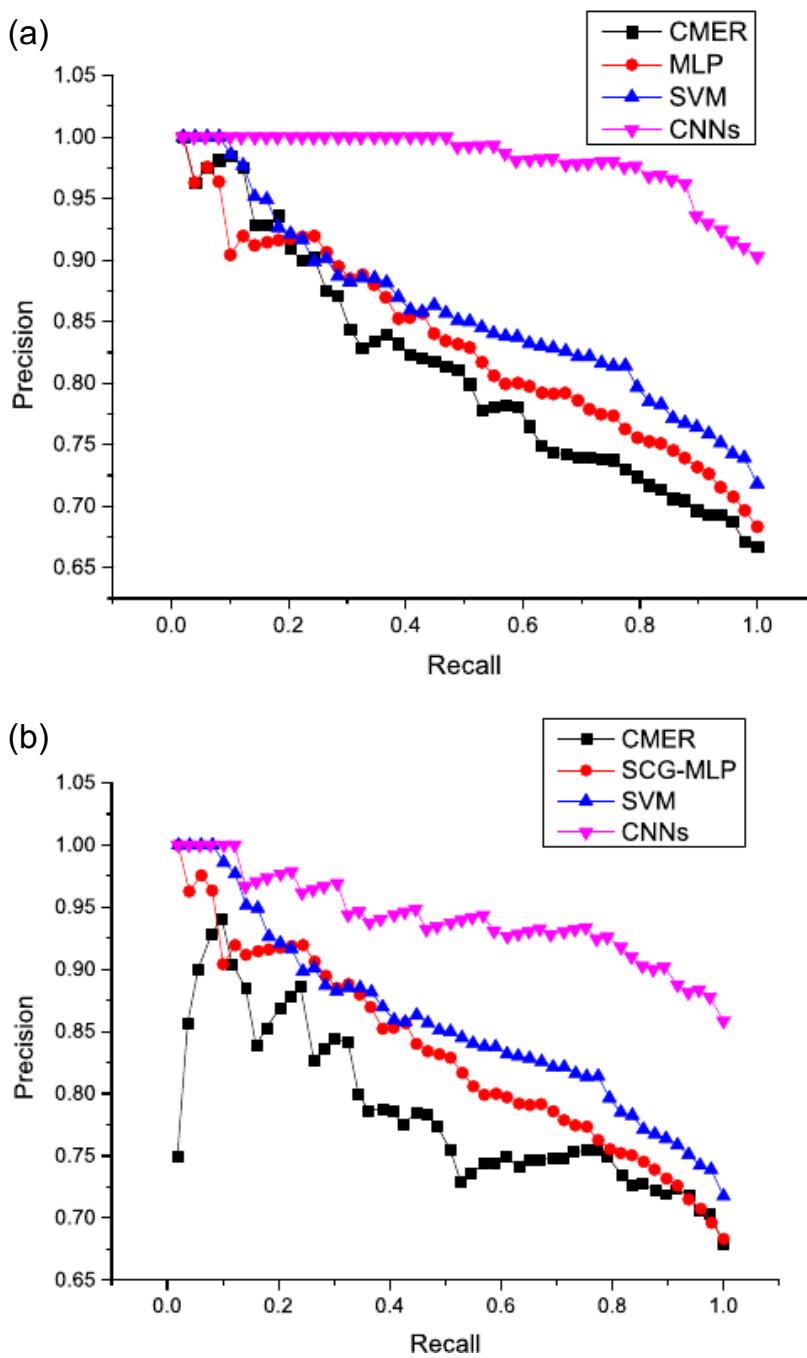


FIGURE 7 Precision-recall curves of the classification of wetland vegetation morphology images based on deep-level learning architecture of convolutional neural networks (CNN) and three different kinds of shallow-level learning architectures of Gabor-kernel-based convolutional architecture combined with the conditional maximum entropy regression (CMER), scaled conjugate gradient (SCG) multilayer perceptron (MLP), and multiclass support vector machine (SVM) on the training (a) and test (b) data set, respectively

large area under the precision-recall curve denotes the overall superior performance of the algorithm with a high mAP score. The CNN-based model achieves the highest mAP scores of 0.963 and 0.923 on the training and test WVM data set, respectively (See Table 1). The compared results illustrated that the improvement of the proposed model for the classification of the WVM images in China on both of the training and test data set is substantial. It appears that more detailed features are abstracted effectively from the original images of WVM by using the

deep-level learning tricks of CNNs compared with the other three algorithms.

In the natural image processing task, Gabor-based features are ideal for identifying a class from a large-scale image data set. Gabor filters with different frequencies and with



TABLE 1 Accuracy and mean average precision (mAP) scores of classification of wetland vegetation morphology on the training and test data set, respectively

Method	Training set		Test set	
	Accuracy (%)	mAP	Accuracy (%)	mAP
CMER	66.7	0.792	67.9	0.772
SCG-MLP	68.3	0.814	67.5	0.763
Multiclass SVM	71.8	0.839	70.0	0.791
CNNs	90.3	0.963	85.8	0.923

Note: The highest accuracy and maximum mean average precision (mAP) scores of classification of wetland vegetation morphology on the training and test data set are both bold.

orientations in different directions have been used to extract interest of regions from complex scene images since the natural scene pictures sometimes consist of rich frequency and texture components. In this study, the optimal configuration of multiresolution Gabor functions (Chen et al., 2015) is implemented to extract the global feature variables of the WVM. The optimal configuration indicates that the multiresolution spectral functions are rearranged to be approximately orthogonal in the spatial-frequency domain. These extracted features will be finally processed by following the three nonlinear learning algorithms of CMER, SCG-MLR, and multiclass SVM. The detection accuracy of three kinds of methods can be seen in Table 1. The Gabor-based CMER classifier reaches the 66.7 percentage of training accuracy and 67.9 percentage of test accuracy. The Gabor-based SCG-MLP classifier output the 68.3% and 67.5% detection accuracy in the training and test sets, respectively. In the training and test stage, the Gabor-based multiclass SVM classifier respectively output 71.8% and 70.0% accuracy. The Gabor-based algorithms can be viewed as a single convolutional layer in the CNN architecture. It is notable that the classification performance of degrades if the other convolutional layers are removed the networks in the least loss of about 18.5% and 15.8% on the training and test set, respectively. The traditional classification strategies based on CMER, SCG-MLR, and multiclass SVM depend on feature engineering such as the Gabor detector for mapping the WVM image data to feature vector. It is difficult to determine the valid ecological feature variables for feeding the classification algorithms. Moreover, they are restrained by the curse of data dimensionality and unable to organize the discriminative ecological features from the WVM image data set. The CNN model makes progress towards learning the high-dimensional data and reusing the features. The deep-level architecture allows the algorithm to learn progressively more abstract ecological features at higher semantic layers and removes the most invalid elements by embedding the machine learning predictor in the multiple-layer nonlinear units. The identification accuracy of the CNN classifier outputs 90.3% and 85.8% in the training and test process, respectively. The final result shows that the CNN model obtains the optimum measure performance. The deep-level model provided a superior alternative routine for accurate artificial intelligence classification of WVM. Thereafter, the deep-level convolutional structure is really important for attaining more discriminative information from images and boosting the accuracy of detection models.



5 | CONCLUSIONS

In this study, we presented a deep-level cascade discriminative model based on CNN for classifying the images of DZH intertidal, LSL alpine, YC coastal, and Zoige plateau wetlands in China. The detection results showed the robustness of the proposed algorithm, for it was not affected by the complex natural factors such as landform, vegetation, sunlight, weather, and sky, as well as human factors such as camera parameter settings of focal length and shooting angle to some extent. The training set achieved a high prediction accuracy of 90.3% by using the cross-validation approach, and the accuracy of the test set is 85.8%. The deep-level detection algorithm significantly outperformed the shallow-level such as Gabor-kernel-based convolutional architecture combined with the CMER, SCG-MLP, and multiclass SVM. The model provided a new way of analyzing the macroscopic ecological elements, which had the potential application to perceiving the changes of environment and species in the wetland regions. The CNN model could also be transplanted to the smartphone, which provided low-cost universal access to the real-time application of classification of WVM images and monitored the state changes of wetland vegetation. Further research is to use the technology of fine-tuning an existing CNNs (Nogueira, Penatti, & dos Santos, 2017) as an effective strategy to limit the amount of training data and improve accuracy.

ACKNOWLEDGMENTS

This study was supported by the National Natural Science Foundation of China (Grant Nos 31601227, 31501221, 61803325), the Natural Science Foundation of Jiangsu Province (Grant Nos. BK20161310, BK20181049), the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant No. 18KJB510046), and the Jiangsu Government Scholarship for Overseas Studies (Grant No. JS-2015-065).

AUTHOR CONTRIBUTIONS

P. L., S. J., and Z. Z. wrote the manuscript; Q. L. and D. L. prepared figures and tables. Y. C. revised the manuscript. All authors reviewed the manuscript.

ORCID

Yongming Chen  <http://orcid.org/0000-0002-7943-7541>

REFERENCES

- Andrychowicz, M., Denil, M., Gomez, S., Hoffman, M. W., Pfau, D., Schaul, T., ... de Freitas, N. (2016). Learning to learn by gradient descent by gradient descent. *Advances in Neural Information Processing Systems*, 3981–3989.
- Aslan, A., Rahman, A. F., Warren, M. W., & Robeson, S. M. (2016). Mapping spatial distribution and biomass of coastal wetland vegetation in Indonesian Papua by combining active and passive remotely sensed data. *Remote Sensing of Environment*, 183, 65–81.
- Bai, J., Xiao, R., Zhang, K., Gao, H., Cui, B., & Liu, X. (2013). Soil organic carbon as affected by land use in young

and old reclaimed regions of a coastal estuary wetland, China. *Soil Use Manage*, 29, 57–64.
Brandt, E. C., Petersen, J. E., Grossman, J. J., Allen, G. A., & Benzing, D. H. (2015). Relationships between spatial metrics and plant diversity in constructed freshwater wetlands. *PLOS One*, 10, e0135917.



- Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2, 1–27.
- Chen, Y., Lin, P., He, Y., & He, J. (2015). A new method for perceiving origins of international important Ramsar wetland ecological habitat scenes in China. *Computers and Electronics in Agriculture*, 118, 237–246.
- Davies-Vollum, K. S., & West, M. (2015). Shoreline change and sea level rise at the Muni-Pomadze coastal wetland (Ramsar site), Ghana. *Journal of Coastal Conservation*, 19, 515–525.
- Deng, X., Liu, Q., Deng, Y., & Mahadevan, S. (2016). An improved method to construct basic probability assignment based on the confusion matrix for classification problem. *Information Sciences*, 340-341, 250–261.
- Dube, T., DeNecker, L., Van Vuren, J. H., Wepener, V., Smit, N. J., & Brendonck, L. (2017). Spatial and temporal variation of invertebrate community structure in flood-controlled tropical floodplain wetlands. *Journal of Freshwater Ecology*, 32, 1–15.
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542, 115–118.
- Feng, L., Han, X., Hu, C., & Chen, X. (2016). Four decades of wetland changes of the largest freshwater lake in China: Possible linkage to the Three Gorges Dam? *Remote Sensing of Environment*, 176, 43–55.
- Huang, L., Zhang, Y., Shi, Y., Liu, Y., Wang, L., & Yan, N. (2015). Comparison of phosphorus fractions and phosphatase activities in coastal wetland soils along vegetation zones of Yancheng National Nature Reserve, China. *Estuarine, Coastal and Shelf Science*, 157, 93–98.
- Huang, Y. C., Tian, K., Yue, H. T., Liu, Z. P., & Lai, J. D. (2012). Effects of dam impoundment on lakeside vegetation succession in the plateau wetlands of the Lashi Lake. *Res. Environ. Yangtze Basin*, 21, 1197–1203.
- James, C. S., Capon, S. J., & Quinn, G. P. (2015). Nurse plant effects of a dominant shrub (*Duma florulenta*) on understorey vegetation in a large, semi-arid wetland in relation to flood frequency and drying. *Journal of Vegetation Science*, 26, 985–994.
- Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F. W., Asner, G. P., ... Ustin, S. L. (2016). Monitoring plant functional diversity from space. *Nature Plants*, 2, 16024.
- Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., ... Darrell, T. (2014). Caffe: Convolutional architecture for fast feature embedding. *ACM International Conference on Multimedia*, 675–678.
- Jiang, W., Lv, J., Wang, C., Chen, Z., & Liu, Y. (2017). Marsh wetland degradation risk assessment and change analysis: A case study in the Zoige Plateau, China. *Ecological Indicators*, 82, 316–326.
- Junk, W. J., Piedade, M. T. F., Lourival, R., Wittmann, F., Kandus, P., Lacerda, L. D., ... Agostinho, A. A. (2014). Brazilian wetlands: Their definition, delineation, and classification for research, sustainable management, and protection. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 24, 5–22.
- Kirwan, M. L., & Megonigal, J. P. (2013). Tidal wetland stability in the face of human impacts and sea-level rise. *Nature*, 504, 53–60.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 1097–1105.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444.
- Lee, T. H., & Hsieh, H. P. (2016). Indicators of sustainable tourism: A case study from a Taiwan's wetland. *Ecological Indicators*, 67, 779–787.
- Leung, H. M., Duzgoren-Aydin, N. S., Au, C. K., Krupanidhi, S., Fung, K. Y., Cheung, K. C., ... Tsui, M. T. K. (2017). Monitoring and assessment of heavy metal contamination in a constructed wetland in Shaoguan (Guangdong Province, China): Bioaccumulation of Pb, Zn, Cu and Cd in aquatic and terrestrial components. *Environmental Science and Pollution Research*, 24, 9079–9088.
- Li, B., Yu, Z., Liang, Z., Song, K., Li, H., Wang, Y., ... Acharya, K. (2014). Effects of climate variations and human activities on runoff in the Zoige alpine wetland in the eastern edge of the Tibetan Plateau. *Journal of Hydrologic Engineering*, 19, 1026–1035.
- Liao, J. T., Ye, H., Huang, T. F., & Peng, G. H. (2017). Seasonal waterbird population changes in Lashihai Lake in northwest Yunnan, China. *Journal of Mountain Science*, 14, 1–11.
- Liu, G., Tian, K., Sun, J., Xiao, D., & Yuan, X. (2016). Evaluating the effects of wetland restoration at the watershed scale in Northwest Yunnan Plateau, China. *Wetlands*, 36, 169–183.

Liu, Z., Gao, J., Yang, G., Zhang, H., & He, Y. (2016). Localization and classification of paddy field pests using a saliency map and deep convolutional neural network. *Scientific Reports*, 6, 20410.



- Lyu, S., Chen, W., Zhang, W., Fan, Y., & Jiao, W. (2016). Wastewater reclamation and reuse in China: Opportunities and challenges. *Journal of Environmental Sciences*, 39, 86–96.
- Marani, M., Da Lio, C., & D'Alpaos, A. (2013). Vegetation engineers marsh morphology through multiple competing stable states. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 3259–3263.
- Meng, W., He, M., Hu, B., Mo, X., Li, H., Liu, B., & Wang, Z. (2017). Status of wetlands in China: A review of extent, degradation, issues and recommendations for improvement. *Ocean & Coastal Management*, 146, 50–59.
- Murray, N. J., Clemens, R. S., Phinn, S. R., Possingham, H. P., & Fuller, R. A. (2014). Tracking the rapid loss of tidal wetlands in the Yellow Sea. *Frontiers in Ecology and the Environment*, 12, 267–272.
- Nogueira, K., Penatti, O. A., & dos Santos, J. A. (2017). Towards better exploiting convolutional neural networks for remote sensing scene classification. *Pattern Recognition*, 61, 539–556.
- Patrick, C. J., Weller, D. E., & Ryder, M. (2016). The relationship between shoreline armoring and adjacent submerged aquatic vegetation in Chesapeake Bay and nearby Atlantic Coastal Bays. *Estuaries and Coasts*, 39, 158–170.
- Roberts, P., Hunt, C. O., Arroyo-Kalin, M., Evans, D., & Boivin, N. (2017). The deep human prehistory of global tropical forests and its relevance for modern conservation. *Nature Plants*, 3, 17093.
- Saggaï, M. M., Ainouche, A., Nelson, M., Cattin, F., & El Amrani, A. (2017). Long-term investigation of constructed wetland wastewater treatment and reuse: Selection of adapted plant species for metaremediation. *Journal of Environmental Management*, 201, 120–128.
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85–117.
- Sirami, C., Jacobs, D. S., & Cumming, G. S. (2013). Artificial wetlands and surrounding habitats provide important foraging habitat for bats in agricultural landscapes in the Western Cape, South Africa. *Biological Conservation*, 164, 30–38.
- Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15, 1929–1958.
- Sun, T., Lin, W., Chen, G., Guo, P., & Zeng, Y. (2016). Wetland ecosystem health assessment through integrating remote sensing and inventory data with an assessment model for the Hangzhou Bay, China. *Science of The Total Environment*, 566–567, 627–640.
- Sun, X., Xiong, S., Zhu, X., Zhu, X., Li, Y., & Li, B. L. (2015). A new indices system for evaluating ecological-economic-social performances of wetland restorations and its application to Taihu Lake Basin, China. *Ecological Modelling*, 295, 216–226.
- Tang, W., Shan, B., Zhang, H., Zhang, W., Zhao, Y., Ding, Y., ... Zhu, X. (2014). Heavy metal contamination in the surface sediments of representative limnetic ecosystems in eastern China. *Scientific Reports*, 4, 7152.
- Vedaldi, A., & Lenc, K. (2015). Matconvnet: Convolutional neural networks for matlab. *ACM International Conference on Multimedia*, 689–692.
- Wang, S., Peng, J., Ma, J., & Xu, J. (2016). Protein secondary structure prediction using deep convolutional neural fields. *Scientific Reports*, 6, 18962.
- Weller, M. O., & Bossart, J. L. (2017). Insect community diversity tracks degradation and recovery of a wastewater assimilation marsh in Southeast Louisiana. *Wetlands*, 1–13.
- Wu, G., Gao, Y., Wang, Y., Wang, Y., & Xu, D. (2015). Land-use/land cover changes and their driving forces around wetlands in Shangri-La County, Yunnan Province, China. *International Journal of Sustainable Development & World Ecology*, 22, 110–116.
- Xi, L., Li, H., Xia, Y., & Qu, W. (2016). Comparison of heavy metal concentrations in groundwater in a mangrove wetland and a bald beach in Dongzhaigang National Nature Reserve (DNNR), China. *Environmental Earth Sciences*, 75, 726.
- Yin, H., Khamzina, A., Pflugmacher, D., & Martius, C. (2017). Forest cover mapping in post-Soviet Central Asia using multi-resolution remote sensing imagery. *Scientific Reports*, 7, 1375.
- Yin, R., & Yin, G. (2010). China's primary programs of terrestrial ecosystem restoration: Initiation, implementation, and challenges. *Environmental Management*, 45, 429–441.

Yu, H. F., Huang, F. L., & Lin, C. J. (2011). Dual coordinate descent methods for logistic regression and maximum entropy models. *Machine Learning*, 85, 41–75.



- Zang, Z., Zou, X., Zuo, P., Song, Q., Wang, C., & Wang, J. (2017). Impact of landscape patterns on ecological vulnerability and ecosystem service values: An empirical analysis of Yancheng Nature Reserve in China. *Ecological Indicators*, 72, 142–152.
- Zhang, L., Wu, B., Yin, K., Li, X., Kia, K., & Zhu, L. (2015). Impacts of human activities on the evolution of estuarine wetland in the Yangtze Delta from 2000 to 2010. *Environmental Earth Sciences*, 73, 435–447.
- Zheng, Y., Niu, Z., Gong, P., Li, M., Hu, L., Wang, L., ... Dang, Z. (2017). A method for alpine wetland delineation and features of border: Zoigê Plateau, China. *Chinese Geographical Science*, 27, 784–799.

How to cite this article: Lin P, Lu Q, Li D, Chen Y, Zou Z, Jiang S. Artificial intelligence classification of wetland vegetation morphology based on deep convolutional neural network. *Natural Resource Modeling*. 2020;33:e12248. <https://doi.org/10.1111/nrm.12248>

