Eng Int Syst (2024) 4: 309–317
© 2024 CRL Publishing Ltd

Engineering
Intelligent Systems

Automatic Classification and Recognition of Spatiotemporal High-resolution Image Data Based on Deep Neural Networks

Yetong Wang[1,a], Guozhang Li[1,b*], Kongduo Xing[1,c] and Rayner Alfred[2,d]

[1] *College of Information Engineering, Hainan Vocational University of Science and Technology, Haikou 571126, Hainan, China*
[2] *Fakulti Komputeran Dan Informatik, Universiti Malaysia Sabah, Sabah, Malaysia*

Due to the huge volume of data obtained from spatiotemporal high-resolution images, traditional data classification methods cannot meet the requirements. This paper applied an automatic classification and recognition method to such data based on a deep convolutional neural network (DCNN). This method used the convolutional neural network (CNN) to extract the spatial and temporal features of the image, and combined the full connection layer for classification and recognition. Finally, the Softmax classifier was used to complete the classification of high-resolution images. For the experiments, 21 high-resolution remote sensing image (RS) datasets were selected as the research subjects. For the experimental analysis, the classification accuracy, precision, and recall of different classification algorithms were tested. Finally, the recognition rate of the neural network was adjusted by setting different expected errors. The data showed that the DCNN algorithm achieved a classification accuracy of 100% for the spatiotemporal high-resolution RS of shrubs and intersections, and the average precision and recall of DCNN were between 0.90 and 0.95, respectively. The image classification time after training was also the shortest, at 0.06ms. The experimental results show that for the classification and recognition of spatiotemporal high-resolution image data, the use of DCNN can achieve better performance and effect, and improve the efficiency and accuracy of image classification.

Keywords: Spatiotemporal High-resolution Image Data, Remote Sensing Images, Deep Convolutional Neural Network, Image Classification and Recognition, Feature Extraction

1. INTRODUCTION

As computer vision technology continues to develop, the application of spatiotemporal high-resolution image data is becoming more and more widespread, and the accurate classification and recognition of such data is crucial to providing solutions to specific problems. However, due to the complex features of this data and its large volume, traditional classification and recognition methods are difficult to meet the requirements. Target classification and recognition is an essential information-extraction technology used for high-resolution images. Because high-accuracy methods usually have high time complexity, massive amounts of high-resolution images make real-time applications extremely difficult. With the continuous development of image processing and artificial intelligence, various methods based on computer vision technology, especially deep learning, have been rapidly and widely applied in the field of image analysis [1–2]. In recent years, the emergence of deep learning

*Address for correspondence: Guozhang Li, College of Information Engineering, Hainan Vocational University of Science and Technology, Haikou 571126, Hainan, China, E-mail: 15203649997@163.com.
[a]E-mail: yhq@hainanu.edu.cn, [b]E-mail: 15203649997@163.com, [c]E-mail: 181741@hainanu.edu.cn, [d]E-mail: ralfred@ums.edu.my

technology has improved the classification and recognition of high-resolution RS. CNN is a new method represented by deep learning, which has important application value in image classification and recognition. CNN can automate the feature extraction and pattern recognition, and greatly improve the accuracy and efficiency of classification and recognition.

The classification and recognition of spatiotemporal high-resolution image data is one of the important research directions in the field of computer vision. In order to predict many different types of stars still undetected in space, Antipin et al. (2019) automatically classified the newly-discovered variable stars based on the random forest algorithm. Experiments showed that automatic classification can also be carried out under noisy photographic data [3]. In the medical field, Rashid and Kaur (2018) implemented a non-auxiliary ovarian cancer classification and recognition system, which made cancer detection on ultrasound images difficult to classify based on clustering or segmentation [4]. To enhance the classification accuracy of high-resolution images, Wang et al. (2018) combined association rule method and object-oriented method to automatically and intelligently collect classified images of different modes. The association rule object-oriented method broadens the information acquisition in terms of logic reasoning classification methods, makes classification more intelligent, and improves classification speed and algorithm reliability [5]. To improve the ability to derive detailed information from low resolution images, Ge et al. (2020) proposed a learning algorithm based on hierarchical relationship for knowledge extraction, which has been validated in fields such as low-resolution image classification and face recognition [6]. After analyzing the characteristics of low-resolution data, Li et al. (2018) proposed an aircraft target classification and recognition algorithm that integrates multiple features, and applied these features to classification using support vector machines. Experiments were conducted on measured data, proving the effectiveness of the algorithm [7]. These studies provide references for the processing of image data.

In recent years, deep learning has also made new progress in the classification and recognition of image data processing [8–9]. For example, in the industrial field, to identify and classify defects in sewage, Xie et al. (2019) collected a large number of videos and images of sewage, and studied the use of deep learning to automatically extract feature expressions of defects in sewage. He also built a complete automatic system for defect classification in sewers based on two-level DCNN, which has a high classification accuracy [10]. To solve the classification problem, Timofeev and Denisov (2018) proposed an automatic object classification task solution based on single photon counting technology data, and used machine learning methods and multi-layer neural network methods to solve the problem. The experimental results can be applied to the design of mobile single photon counting technology, LIDAR, which can very reliably detect and classify objects on the Earth's surface in real-time [11]. In multi-vision images, compared to convolution computation of a single image, dynamic gesture multi-frame images have more computational complexity and feature extraction, which affects recognition efficiency and real-time performance. Liu et al. (2021) proposed a dynamic gesture recognition algorithm based on 3D (three dimensional) CNN attention mechanism, which improved data redundancy and format inconsistency [12]. These studies are valuable for research on automatic classification and recognition of high-resolution image data, although they are not aligned with the current situation.

The traditional RS classification and recognition methods require the manual setting of features. To address this issue, this current study explored a new method of RS classification and recognition based on DCNN. The performance of DCNN was tested on RS datasets. The data showed that for the RS of shrubs and intersections, the accuracy of the DCNN algorithm can reach 100%. For other types of RS, the accuracy was not less than 80%, with at the majority being over 90%. The results demonstrated that DCNN can be applied to RS, effectively improving the accuracy and stability of classification and recognition, and outperforming other classification and recognition methods. This method has great potential in the field of RS processing and has many practical applications.

2. CLASSIFICATION PRINCIPLES OF DEEP NEURAL NETWORKS

Deep neural networks are complex models consisting of multiple neural network layers. In classification tasks, deep neural networks are used to divide input data into predefined categories. Classification is the main technique applied to hyperspectral images, as it assigns labels to each pixel based on its characteristics [13]. In the training process, the deep neural network continuously adjusts the weight and deviation of the network through backpropagation, so that the category labels output by the neural network are as consistent as possible with the real labels. When new and unfamiliar data is input, the deep neural network passes it onto the neural network and places it into a certain category based on the learned patterns and features. Due to the ability of deep neural networks to learn high-level and abstract features, they achieve excellent performance in classification tasks.

2.1 DCNN

CNN has become a research hotspot in hyperspectral RS processing because of its strong feature-representation ability and high computing efficiency [14–15]. In the traditional CNN, features are extracted from the convolution layer first, and then the convolution core and network layer are modified to enable multiple applications. However, this traditional method has problems: it requires a long training time and has low accuracy. DCNN is a neural network model that combines CNN with multiple convolution layers. This method uses the characteristics of CNN to recognize and classify high-resolution images. The function of convolutional layers is to convert low-dimensional convolutional images to high-dimensional ones. Generating a multidimensional image expression can better express the intrinsic clarity of the original image [16]. In the depth convolution neural model,

Table 1 Comparison of image sampling training time for convolutional kernels of different sizes.

Data sample	$3 \times 3 \times 64$	$5 \times 5 \times 64$	$7 \times 7 \times 64$
Sample 1	0.4s	0.41s	0.4s
Sample 2	0.39s	0.4s	0.42s
Sample 3	0.38s	0.39s	0.43s

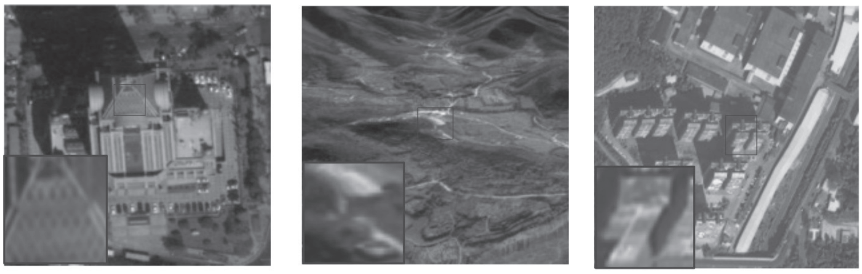

Figure 1 Sample of a deep convolutional network image.

image sampling is conducted through the convolution layer. The principle of sampling can be expressed with this formula:

$$G(Y) = MAX(W \times Y + B, 0) + \alpha_j MIN(0, W \times Y + B) \quad (1)$$

where the parameter W represents the convolutional kernel. $G(Y)$ represents the upsampled image, and the Y table represents the input image. By setting convolutions of three different sizes, the image sampling runtime is tested, as illustrated in Table 1:

As shown in Table 1, overall, under the same depth structure of the network model, smaller convolutional kernels have shorter sampling training times. Therefore, in high-resolution images, the sampling method can choose images with a size of $3 \times 3 \times 64$. When the moving step of the convolutional kernel in DCNN is smaller than the edge length, uneven phenomena occur. To solve this problem, deep mapping can be used. The main purpose of depth mapping is to convert the original image into a vector in the high-dimensional feature space. Deep mapping is implemented using the principle of CNN, in which each convolution layer and pooling layer can convert the image from the original pixel representation to high-dimensional feature representation. Traditional neural network pooling methods have the problem of oversimplification, making it difficult to extract effective features [17]. Deep mapping utilizes a network structure with residual networks; the residual function is:

$$F_i(Y) = H(Y) - F_1(Y) \quad (2)$$

When $F_i(Y)$ equals 0, an identical mapping relationship can be obtained. The use of deep residual network can promote the rate of convergence of the network and optimize the network model. This principle is applied to deep mapping and is represented by the following formula:

$$F(Y) = MAX(w_i \times F_i(Y) + \beta, 0) + \alpha_j MIN(0, w_i \times F_i(Y) + \beta) \quad (3)$$

where w_i represents the convolutional kernel of deep mapping. β represents bias, and $F_i(Y)$ represents the output image of each layer. Depth mapping is used to eliminate image noise. It is inevitable that the original dataset will contain noise in the images and labels, indicating that there is a strong need for a robust classification system [18]. Figure 1 is a sample of a deep convolutional network image.

2.2 Activation Function

In the CNN model, the Rectified Linear Unit (ReLU) is a commonly-used activation function. The ReLU Activation function is expressed as:

$$F(X) = MAX(0, X) \quad (4)$$

In Formula (4), when the input X value is negative, 0 is returned; otherwise, the original value is returned. Compared with other Activation function, the ReLU function avoids a certain amount of calculation, which can effectively avoid the phenomenon of small gradient and improve the rate of convergence of CNN model. This paper uses an improved ReLU activation function, expressed as:

$$g(\theta) = \frac{1}{N} \sum_{i=1}^{N} ||F(Y_i \theta) - X_i||^2 \quad (5)$$

In Formula (5), the relationship between F and X is the processed image and the original image. The parameters of each layer are modified by minimizing $g(\theta)$ back-propagation, and the best depth CNN mapping model is obtained.

Spatiotemporal high-resolution images can be preprocessed first by, for instance, cropping and scaling; then they are input into DCNN for recognition and classification. The network model is trained by means of supervised learning, and the model parameters are optimized through a large number of labeled image data so as to improve the accuracy and efficiency of recognition and classification. Practical applications, it can achieve automated processing and analysis of large-scale spatiotemporal high-resolution image data.

3. AUTOMATIC CLASSIFICATION AND RECOGNITION METHODS FOR IMAGE DATA

3.1 Spatiotemporal High-resolution Image Data Processing

The recognition and classification of high spatial resolution RS is an automatic classification and recognition method

Buildings Mediumresidential River

Figure 2 Remote sensing sample images.

based on high spatial resolution RS. High-resolution RS, videos, etc., are important data in high-resolution RS that have a large amount of data and high dimensionality. It is essential to use machine learning and other technologies for feature extraction and classification. In deep learning, features are part of the CNN and are automatically extracted [19]. To automatically recognize and classify image data, sufficient preparation must be made for sample data. In order to improve the recognizability and distinguishability of high-resolution images, representative training and detection samples can be selected from existing images to preprocess high-resolution images. Figure 2 shows sample data from spatiotemporal high-resolution RS.

3.1.1 Image Data Augmentation Processing

Data augmentation is a commonly-used, deep-learning technique that improves the generalizability of models and reduces overfitting by transforming or expanding data. Before image data is recognized and classified, it must be augmented and preprocessed. Data augmentation can effectively improve the model's generalizability and robustness, and reduce the risk of overfitting. Data augmentation can be achieved by these operations: vertical and horizontal flipping, transpose transformation, brightness transformation, contrast transformation, and hue and saturation transformation. The formula for vertical and horizontal flip transformation is:

$$G(X_1, Y_1) = G(W - X_0, Y_0) \tag{6}$$

where X_1, Y_1 represents the coordinates after the vertical horizontal flip transformation. X_0, Y_0 represents the coordinates before the transformation, and W represents the width of the image. The mathematical expression for transpose transformation is as follows:

$$F(X_1, Y_1) = F(X_0, Y_0) \tag{7}$$

Image transposition transformation involves flipping an image along a diagonal, which can also be understood as the swapping of rows and columns of the image. The brightness transformation operation of an image can change the brightness level of the image, making it brighter or darker. The contrast transformation operation can adjust the degree of brightness difference in the image, making the image more vivid. These are a series of operations for image data augmentation. Data augmentation can be carried out in real time during the training process of the model, or data can be preprocessed before training.

3.1.2 Feature Extraction and Classification

In the deep convolutional neural model mentioned earlier, DCNN is used to process the convolutional layers of spatiotemporal high-resolution images. DCNN usually consists of multiple convolutional and pooling layers. During the convolutional process, the feature map is convolved to restore the details and spatial position information of the image. In DCNN, the last layer is usually a fully-connected layer, used to map the feature map of the convolutional layer onto an output vector for classification or regression tasks. This classification approach is known as a fully-connected layer classification method.

The core idea of this method is to embed a fully connected layer classifier into a DCNN, use the feature maps of the convolutional layer as input, and map the feature maps to a predefined set of categories through the fully connected layer. By using both spatial and temporal deep mapping methods to reduce overfitting problems, the generalizability of DCNN can be improved. The selection of image classifiers is crucial in spatiotemporal high-resolution image recognition and classification. Traditional classifiers distinguish objects in the underlying feature space, which cannot adapt to the complexity of different features with different imaging mechanisms, resulting in significant intra-class differences and imbalances between different classes [20]. In this study, the Softmax classifier was chosen for the classification of high-resolution images.

3.2 Image Classifier

The Softmax classifier is a multi-class classifier, commonly used for the classification problem associated with output layers in neural networks. The classifier normalizes the output values through the Softmax function to obtain a probability distribution that represents the confidence level of each category. The formula for the Softmax function is:

$$\mathit{softmax}\,(Z_i) = \frac{1}{e^{-z}+1} \tag{8}$$

In Formula (8), Z_i represents a value of the i-th item in the output vector. In softmax classifier, the output vector Z is obtained by multiplying a weighted matrix moment by an input vector x and adding it to a compensation vector b:

$$Z = wx + b \tag{9}$$

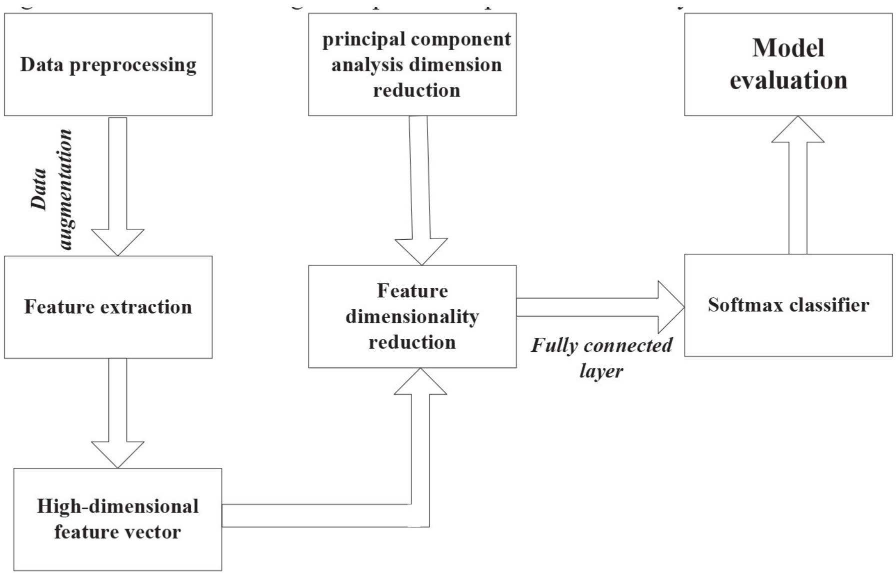

Figure 3 Framework diagram of high-resolution image classification and recognition based on DCNN.

After executing Formula (9), the output vector Z is normalized using the Softmax function to obtain a probability distribution y, which represents the confidence level of each category:

$$y = softmax\ (Z) \tag{10}$$

In Formula (10), the category with the highest probability can be selected as the prediction result. The Softmax maximum classifier uses cross entropy to measure the difference between the predicted results and the actual results, and uses the backward propagation method to update the weighting matrix and deviation vector, so as to improve the classification result. This method extracts high-resolution image samples in space and time, inputs their feature vectors into the entire connected layer, and classifies them. In practical applications, the image to be tested is input, and the category of the image to be tested is obtained by feature extraction and classification. For the recognition and classification of high-resolution RS based on DCNN, image enhancement, data expansion and other pre-processing operations are required to improve image clarity and reduce noise. Secondly, in terms of feature extraction, multiple levels of deep learning networks such as convolutional layers, deconvolution layers, and pooling layers are utilized to extract more information from the image. Finally, classification recognition is carried out. The fused features are input into the classifier for classification and recognition. Figure 3 depicts the spatiotemporal high-resolution image classification and recognition process explored in this study.

4. DEEP CONVOLUTIONAL NEURAL RS CLASSIFICATION EXPERIMENT

4.1 Introduction to Experimental Datasets

For this study, 21 levels of RS data were taken as a sample to test the accuracy and effectiveness of image classification. The 21 level RS UCM (UC Merced Land Use Dataset) is currently the most commonly used RS classification method. This dataset comprised a total of 21 types of ground object images, with approximately 100 types of each image. These images were extracted from high-resolution satellite images, each with a size of 256 × 256 pixels. Each image in the UCM dataset has a corresponding label that represents the category of features it belongs to, such as airports, forests, parks, etc. The data set includes the following 21 categories: agriculture, aircraft, Ballpark, bathing beach, buildings, bushes, dense housing, forests, highways, golf courses, ports, intersections, medium-sized housing, mobile home parks, overpasses, parking lots, rivers, runways, sparse housing, tanks, tennis courts, etc. Figure 4 shows the images of the UCM dataset, which includes 20 spatiotemporal high-resolution RS.

4.2 Experimental Results

To evaluate the performance of the RS classification and recognition methods used in this article, experiments were conducted using the UCM public dataset. Firstly, the images were preprocessed, which involved data augmentation, including mirroring, brightness adjustment, hue adjustment, saturation adjustment, and other operations. Figure 5 shows the RS after image preprocessing.

The preprocessed image was input into the DCNN for training to obtain a classification model. The images in the test set were input into the classification model for classification, and the classification results were obtained, as shown in Figure 6.

The accuracy data of DCNN for high-resolution RS classification were counted to determine the effectiveness of the algorithm proposed in this study. The data is presented in Table 2.

Table 2 shows the classification accuracy of high-resolution RS from the UCM dataset mentioned above, which includes

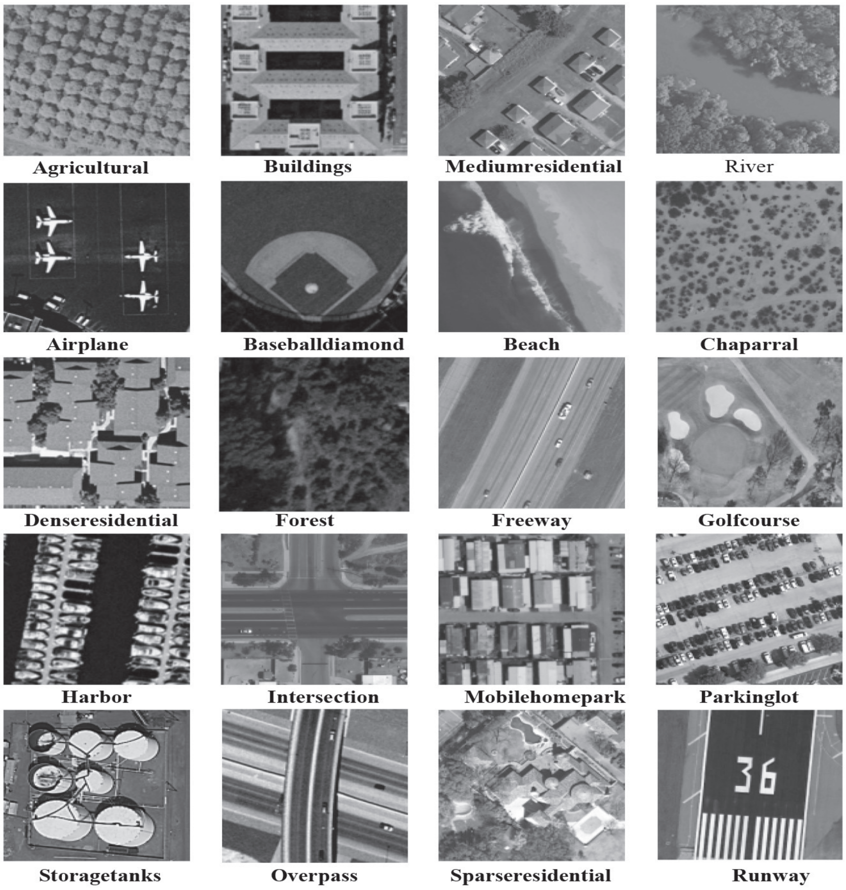

Figure 4 Spatiotemporal high-resolution image dataset.

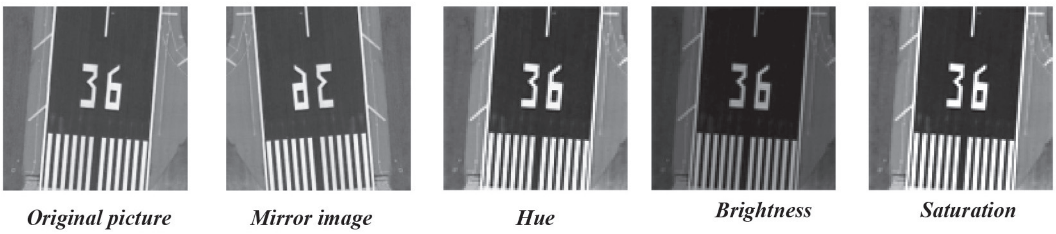

Figure 5 RS data preprocessing image.

20 types of RS. As shown in the table, the DCNN algorithm achieves a classification accuracy of 100% in RS of bushes and intersections. The classification accuracy of all types of RS is not less than 80%, with most concentrated being over 90%, indicating that its classification accuracy is high. To compare the performance of the DCNN algorithm, the classification precision and recall of other classification algorithms were tested, including CNN, Support Vector Machine (SVM), Back Propagation (BP), and K-Nearest Neighbor (KNN), as shown in Figure 7.

Figure 7 shows the precision and recall of several splitting algorithms tested separately. The data presented in Figures 7a and 7b indicate that of the five algorithms, DCNN had the highest level of classification accuracy and recall. The data of the mean variance plot shows the average precision of DCNN was 0.90-0.95, and the highest reached was above 0.95. However, the average precision of CNN, SVM, BP, and KNN were all between 0.85-0.90, with SVM having the lowest precision. Looking at the recall rate chart again, the average recall rate of DCNN was between 0.90 and 0.95. The average recall rate of CNN, BP, and KNN was 0.85 to 0.90, and the average recall rate of SVM was below 0.85. The data show that RS classification and recognition methods based on DCNN have excellent performance in terms of accuracy and robustness. For high-resolution satellite RS, this method can better preserve the detailed information of the image and improve the accuracy of classification.

Figure 8 shows the training time of the five algorithms and the classification time of the RS obtained from the experiments. After 10 tests, the average training time of the algorithm Line chart and the average test classification time Line chart were obtained. The data in Figures 8a and 8b show that during sample training of the algorithm, DCNN took the longest training time, averaging between 3.2s

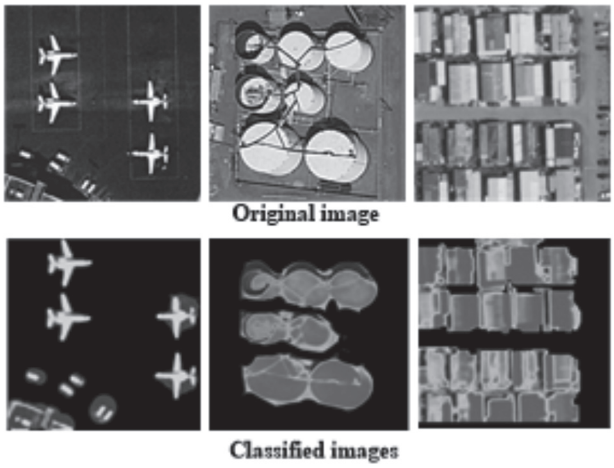

Figure 6 Classification effect of deep convolutional neural RS.

Table 2 DCNN classification accuracy table.

Image category	Classification accuracy	Image category	Classification accuracy
Agricultural	86%	Harbor	93%
Airplane	98%	Intersection	100%
Baseball diamond	85%	Medium residential	94%
Beach	96%	Mobile home park	95%
Buildings	86%	Overpass	87%
Bushes	100%	Parking lot	97%
Dense residential	93%	River	98%
Forest	92%	Runway	88%
Freeway	92%	Sparse residential	97%
Golf course	94%	Storage tanks	95%

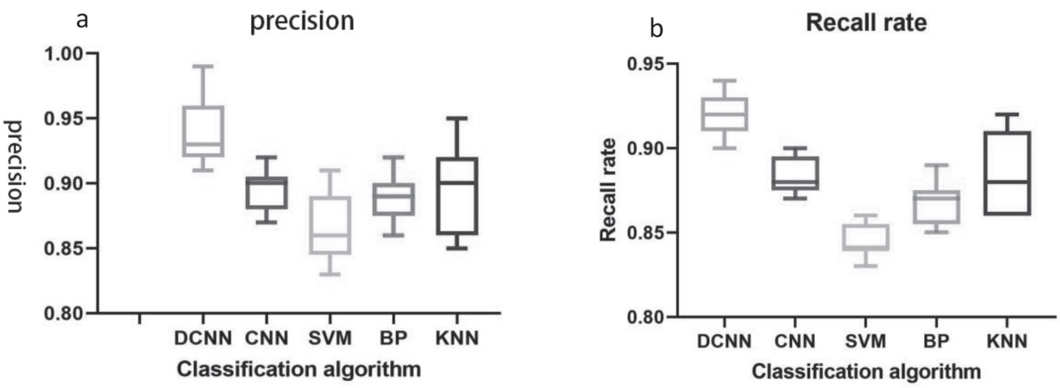

Figure 7 Classification precision and recall of RS using different classification algorithms. (a) Comparison of precision. (b) Comparison of recall rates.

and 4s. The average training time of the SVM algorithm was the shortest, between 2.6s and 2.8s. The longer training time is due to the complexity of the DCNN model and the high complexity of RS data. However, the image classification time presented in Figure 8b shows that DCNN had the shortest classification time of 0.06ms, while other algorithms did not produce this level of efficiency. DCNN greatly improves the efficiency of RS classification and recognition.

In the classification and recognition of RS, increasing the depth and number of nodes of the algorithm network can help improve the fitting ability of the network, thereby better approaching the training data and reducing training errors. However, when the network depth and number of nodes are too large, overfitting may occur. That is to say, the network can fit the training data well, but its generalizability is relatively poor and therefore cannot adapt well to the test data, leading to an increase in test errors. Therefore, it is necessary to adjust the number of network nodes to reduce errors, while ensuring the efficiency of the algorithm and improving the recognition rate of RS. Table 3 shows the relationship between DCNN error and recognition rate.

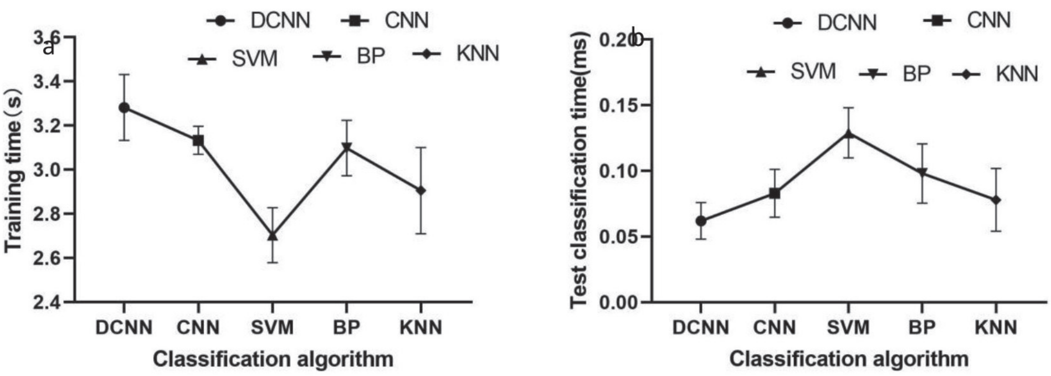

Figure 8 Comparison of training time and testing classification time for different algorithms. (a) Training time. (b) Test classification time.

Table 3 Error and recognition rate of DCNN.

Expected error	Hidden layer depth	Reconstruction error	Recognition rate
1	1	0.725	97%
0.5	2	0.406	99%
0.1	4	0.026	75%
0.05	4	0.026	74%
0.01	5	0.0025	98%

Table 4 Comparison of image recognition rates between different neural network algorithms before and after noise addition.

Neural network algorithm	Recognition rate before noise addition (%)	Recognition rate after noise addition (%)
DCNN	99.8%	98.6%
CNN	99.09%	97.98%
BP	98.54%	93.76%

Table 3 shows the image recognition rate obtained by reconstructing the errors of DCNN under different expected errors and hidden layer depths. These results were derived from multiple experiments. The data in the table shows that when the expected error of the target is set to 0.5, the image recognition rate of the DCNN network is the highest at 99%. Then, when the expected error of the target is set to 0.01, the image recognition rate is 98%. Therefore, it can be concluded that the setting of the expected error value of the target has a significant impact on the recognition rate of the algorithm. In this study, to improve the image recognition rate of the algorithm, the target expected error was set to a threshold of 0.5. Several tests were conducted to compare the recognition rates of three neural networks on RS before and after noise addition, as shown in Table 4.

The experimental data in Table 4 shows that the RS of DCNN had the highest recognition rates before and after noise addition, with an image recognition rate of 99.8% before noise addition. The image recognition rates of CNN and BP neural network models before noise addition are 99.09% and 98.54%, respectively. The DCNN image recognition rate of the denoised image can still reach 98.6%. Therefore, it can be concluded that DCNN is more suitable for processing high-resolution RS data and has more obvious advantages. Its recognition rate is higher than that of other neural networks, which also indicates that DCNN has stronger noise robustness.

5. CONCLUSIONS

This study adopted a RS classification and recognition method based on DCNN. This method can classify and recognize spatiotemporal high-resolution image images by learning the features of each image. Firstly, preprocessing and data augmentation operations including mirroring, brightness adjustment, hue adjustment, and saturation adjustment, were performed on images with high temporal and spatial resolution. Then, the DCNN was used to learn the image features. DCNN can extract feature maps from fully-connected convolutional layers and then use feature classifiers to classify images. In this way, RS can be mapped to the feature space and classified and recognized using specific features. The proposed approach was tested using the UCM RS classification and recognition dataset, and satisfactory results were obtained. The DCNN network based on deep learning can achieve very high accuracy, outperforming the traditional classification and recognition methods. By comparing the accuracy and recall of several algorithms in terms of high-resolution image segmentation and recognition, the analysis results showed that the DCNN algorithm had efficient and high-accuracy performance, and good robustness in RS classification. The conclusions obtained also indicate promising prospects for the application of the model in RS processing.

FUNDING

This work was supported by Key Laboratory of Philosophy and social Science in Hainan Province of Hainan Free Trade Port International Shipping Development and Property Digitization, Hainan Vocational University of Science and Technology. This work was supported by Hainan Provincial Natural Science Foundation of China in 2021 under Grant No. 421RC609.

AVAILABILITY OF DATA AND MATERIALS

No data were used to support this study.

COMPETING INTERESTS

The authors declare that they have no competing interests.

REFERENCES

1. Liu, Y. Xie, Y., Yang, J. (2020). "Target Classification and Recognition for High-Resolution Remote Sensing Images: Using the Parallel Cross-modal Neural Cognitive Computing Algorithm". *IEEE Geoscience and Remote Sensing Magazine, 8*(3):50–62.
2. Soomro, T. A., Afifi, A. J., Zheng L. "Deep Learning Models for Retinal Blood Vessels Segmentation: A Review". *IEEE Access 7.1*(2019):71696–71717.
3. Antipin, Sergei, V., Becker, I, Belinski, A. A., Kolesnikova, D. M., Pichara, K., Samus, N. N., et al. "New variable stars from the photographic archive: semi-automated discoveries, attempts at automatic classification and the new field 104 Her". *Research in Astronomy and Astrophysics 18.8*(2019):92–92.
4. Rashid, S. and R. Kaur. "Automatic detection of ovarian cancer based on improved DWT transformation". *International Journal of Engineering & Technology 7.2*(2018):104–108.
5. Wang, F., Xia L., Chen, Z. "Remote sensing identification of coastal zone mariculture modes based on association-rules object-oriented method". *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering 34.12*(2018):210–217.
6. Ge, S., Zhang, K., Liu, H. "Look One and More: Distilling Hybrid Order Relational Knowledge for Cross-Resolution Image Recognition". *Proceedings of the AAAI Conference on Artificial Intelligence 34.7*(2020):10845–10852.
7. Li, M., Wu, J., Zuo, L. "Aircraft Target Classification and Recognition Algorithm Based on Measured Data". *Dianzi Yu Xinxi Xuebao/Journal of Electronics and Information Technology 40.11*(2018):2606–2613.
8. Umma Kavita. "Visual Intelligent Recognition System based on Visual Thinking". *Kinetic Mechanical Engineering 2.1*(2021): 46–54.
9. Imran A. Khan. "Video Monitoring System for Natural Environment Protection Area Supporting Image Recognition". *Nature Environmental Protection 2.3*(2021): 31–39.
10. Xie, Q., Li, D., Xu, J. "Automatic Detection and Classification of Sewer Defects via Hierarchical Deep Learning". *IEEE Transactions on Automation Science and Engineering 16.4*(2019):1836–1847.
11. Timofeev, A., and V. M. Denisov. "Automatic object classification according to 3d-lidar data based on single-photon counting technology". *Scientific and Technical Journal of Information Technologies Mechanics and Optics 18.5*(2018):709–718.
12. Liu, Y., Jiang D., Duan, H. "Dynamic Gesture Recognition Algorithm Based on 3D Convolutional Neural Network". *Computational Intelligence and Neuroscience 2021.12*(2021):1–12.
13. Shamsolmoali, Pourya, Zareapoor. "Convolutional neural network in network (CNNiN): hyperspectral image classification and dimensionality reduction". *IET Image Processing 13.2*(2019):246–253.
14. Zou, L., Zhu, X. L., Wu, C. "Spectral–Spatial Exploration for Hyperspectral Image Classification via the Fusion of Fully Convolutional Networks". *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13.99*(2020):659–674.
15. Petkov Christopher. Ship Image Classification Based on the Architecture of Imaging Spectral Data Distributed Processing System. *Distributed Processing System 3.1*(2022): 1–9.
16. Zhang, Z. "Joint classification of multiresolution representations with discrimination analysis for SAR ATR". *Journal of electronic imaging 27.4*(2018):043030.1-043030.8.
17. Lai, X., Zhou, L., Fu, Z. "Enhanced pooling method for convolutional neural networks based on optimal search theory". *Image Processing, IET 13.12*(2019):2152–2161.
18. Zhang, Z., Liu, M., and Zhu, D. "Automatic recognition and classification of diabetic retinopathy images by combining an attention mechanism and an efficient network". *Journal of Image and Graphics 25.8*(2020):1708–1718.
19. Kurek, J., Swiderski, B., Osowski, S. "Deep learning versus classical neural approach to mammogram recognition". *Bulletin of the Polish Academy of Sciences, Technical Sciences 66.6*(2018):831–840.
20. Li, Y., Shao, H., Jiang, N. "Classification of land cover in high-resolution remote sensing images based on Space-LDA model". *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering 34.8*(2018):177–183.