Scholar Publishing Group
International Journal of Educational Innovation and Science
https://doi.org/10.38007/IJEIS.2022.030403
ISSN 2789-7184 Vol. 3, Issue 4: 26-41

Remote Sensing Image Classification Model Based on Multimedia Network and Its Knowledge Integration Method

Nasser Jaber*

Universiti Teknologi MARA, Malaysia

**corresponding author*

Keywords: Multimedia, Remote Sensing Image, Convolutional Nerve, CNN

Abstract: In recent years, with the rapid development of remote sensing satellite technology, the resolution of images generated by remote sensing satellites has been increasing. Today, high-resolution remote sensing images (referred to as high-resolution remote sensing images) have become the focus of research in the field of remote sensing. The feature information of high-resolution remote sensing images is richer and more accurate than the medium-low resolution remote sensing image information, and it has more practical significance in practical applications: remote sensing application workers and research scholars perform feature extraction and scenes on high-resolution remote sensing images. Analysis and research on classification, feature recognition, and target detection to meet the actual needs of high-resolution remote sensing images in the military and economic fields as well as in the civil and civilian fields. Among them, feature extraction is the basis for understanding and analyzing high-resolution remote sensing images. Through various feature extraction techniques, the information of the essential attributes in the high-resolution remote sensing images is extracted as features, and the features are used to describe the high-resolution remote sensing images, which facilitates accurate and reasonable understanding and analysis of high-resolution remote sensing images, and avoids redundancy. The negative impacts of information, noise, etc. in practical applications, and thus more efficient and accurate completion of the classification, identification and detection of high-resolution remote sensing images. This paper mainly discusses the research progress of deep network model of high-resolution remote sensing image based on the deep network model of multimedia-based Convolutional Neural Network (CNN), and through the experimental results of high-resolution remote sensing image specific problems, Performance comparison and evaluation based on features extracted by the CNN-based deep network model. Finally, the problems of CNS-based deep network model in the extraction of deep semantic features in high-resolution remote sensing images and future development trends are analyzed, which provides reference for the feature classification tasks of high-resolution remote sensing images.

1. Introduction

The rapid development of remote sensing technology has promoted its extensive application in soil and water conservation dynamic monitoring, land and resources surveys, agroforestry surveys and other fields, and played a major role [1-4]. The application of 3S technology to the investigation of land and resources surveys and the dynamic monitoring of vegetation coverage changes began in the 1980s. With the advancement of time, remote sensing technology has made great progress, and the richness of data sources has also extended the remote sensing technology in soil and water conservation. Development in the field of monitoring [5]. The richness of remote sensing data sources also poses a challenge to the rapid application of data sources. Although the accuracy of visual interpretation is high, but the efficiency is low, the supervised classification method can achieve rapid land use classification and provide basic data for subsequent dynamic monitoring of soil and water conservation. However, the existing supervised classification methods are mainly based on traditional pattern recognition techniques, such as maximum likelihood classifier and K-nearest neighbor method. The classification accuracy of these methods is based on samples that tend to infinity, otherwise it is difficult to obtain. More ideal effect. Support Vector Machine (SVM) is one of the machine learning methods based on statistical learning theory. It is based on small sample classification. Based on sample learning, the classification accuracy will be higher and higher [7-9].

Remote sensing image classification and recognition work is of great significance for studying the development process and distribution law of objects or phenomena. The concept of deep learning was first proposed in 2006 [10-12]. In the field of remote sensing, the application of deep learning in classification and recognition has further optimized the classification effect. Previous studies using remote sensing high-resolution and hyperspectral images have proved that deep learning can fully extract remote sensing image features [13,14]]. Currently, there are four types of deep learning mainstream algorithm models: Restricted Boltzmann Machine (RBM), Deep Belief Networks (DBN), Convolutional Neural Networks (CNN), and Automated Encoder (Auto Encoder, AE), etc. Among them, AE and DBN algorithms belong to unsupervised learning, while CNN algorithm is a kind of deep supervised learning [15-19]. At present, deep learning is widely used in remote sensing image classification and recognition. It is mainly composed of DBN, CNN and AE. RBM is a shallow structure that constitutes DBN. Therefore, researchers mainly apply RBM algorithm to DBN algorithm. 20-23].

In the past few decades, experts and scholars at home and abroad have been working on classification techniques and methods to improve the accuracy of remote sensing image classification. Originally adopted various unsupervised and supervised classification techniques, many advanced classification algorithms have been widely used, including neural networks, support vector machines, and expert systems [24]. However, these algorithms have little effect on the classification results, and the classification results have no significant improvement in nature. They are still classified based on pixels and the parameters used are relatively simple [25]. In addition, due to the many factors affecting the quality of remote sensing images and the complexity of geographical phenomena, there are still many uncertain factors in the classification. Improving the algorithm to improve the classification accuracy is still a problem that needs to be solved continuously by the remote sensing community [26,27].

Although domestic and foreign researchers have summarized and sorted out the problems of classification, segmentation and recognition of deep learning in high-resolution remote sensing, feature extraction is an indispensable part of tasks such as classification, segmentation and recognition. Therefore, this paper mainly discusses the research progress of deep neural network based on Convolutional Neural Networks (CNN) on the semantic feature extraction of high-resolution remote sensing images, and the experimental results based on the specific problems of high-resolution remote sensing images. The features extracted by CNN's deep network model are compared and evaluated. Finally, the problems of CNS-based deep network model in the extraction of deep semantic features in high-resolution remote sensing images and future development trends are analyzed, which provides a reference for the feature extraction tasks of high-resolution remote sensing images.

2. Proposed Method

2.1. Combination of Multimedia Technology and Image Classification

Based on the development of multimedia technology for many years, the remote sensing image classification technology has made breakthrough progress, and the application range is expanding, including almost all aspects. CNN-based deep network model is a representative class in deep learning, and has achieved remarkable results in the field of image. A large number of experiments and researches show that in the field of image, CNN with deep network structure and "self-learning" ability can Extract deep hierarchical abstract semantic features with layers. In remote sensing images, the CNN-based deep network model can be successfully applied not only to remote sensing image classification, but also as a feature extractor. The deep semantic features of high-resolution remote sensing images extracted by CNN have achieved good results in high-resolution remote sensing image region classification, scene understanding feature recognition, and ground object extraction. The main difference between deep network model feature extraction and artificial feature extraction is that no complicated human participation is required. Only the training data set needs to be provided. The depth model can “automatically learn” some kind of “mapping” relationship between input and output. This “mapping” relationship is the key to the deep abstraction semantic feature extraction of the data by the depth model. The depth model is automatic. The ability to learn features is the key to a major breakthrough in depth in all areas. The deep abstract semantic features extracted by the model have better stability, invariance and nonlinearity. Such deep abstract semantic features can better utilize and express the feature information of high-resolution remote sensing images, which can be more efficient in practical applications. And accurately analyze and utilize high-resolution remote sensing images. The follow-up will mainly study the development of various CNN-based deep network models in feature extraction of high-resolution remote sensing images.

The structure of the convolutional neural network has a distinct feature: each convolutional layer is usually followed by a pooling layer to form a feature extractor. The extraction of high-level features can be achieved by stacking multiple convolution-pooling structures. However, the convolution-pooling structure outputs a two-dimensional feature map. This complex and highly abstract high-level feature cannot be directly input into the traditional classifier. Therefore, the output of the last pooling layer, that is, the feature extracted by the convolution-pooling structure, needs to be input into a multi-layer fully connected network structure, thereby realizing the mapping and optimization of two-dimensional features to one-dimensional features. . Finally, the one-dimensional features of the fully connected layer output are input into the classifier to obtain the final classification result.

(1) Convolutional layer, as the core of convolutional neural network, is very suitable for efficient extraction of high-level features with the help of local perception and weight sharing

strategies. In the training process of convolutional neural networks, there are two types of parameters in the convolutional layer that need to be trained and optimized, namely the parameters of the convolution kernel and the bias term b of the convolutional layer. The basic principle of the convolutional layer, the convolutional layer performs a dot product operation on the convolution kernel parameter and the local receptive field of the image, and then adds the result of the dot product operation and the offset term b to the activation function n( ) to obtain The feature map of the final output.

(2) In a convolutional neural network, each convolutional layer is often followed by a pooling layer. The role of the pooling layer is to abstract and reduce the output of the convolutional layer. The use of the pooling layer can effectively improve the training efficiency of the convolutional neural network, avoid over-fitting in the training process, and improve the robustness and robustness of the extracted high-level features. The most commonly used pooling layer calculation function is the maximum pooling function MaxPooling(). The maximum pooling function uses the maximum value of the feature values of the feature map in the receptive field region to replace all the feature values in the corresponding region, thereby realizing the dimensionality reduction and abstraction of the image features.

(3) Convolutional neural network The structure of the fully connected layer is the same as that of the traditional neural network. It is used to refine and optimize the features extracted by the convolutional layer. The structure of the fully connected layer is extracted by the last convolutional layer. Large, it is difficult to classify using a classifier. Therefore, in order to meet the needs of the classifier, the features extracted by the convolutional neural network are input into the fully connected layer to reduce the number of features.

Finally, by inputting the output result of the fully connected layer into the classifier, the feature extracted from the input data can be converted into the probability that the input data belongs to each category (for example, the Softmax classifier), wherein the highest probability category is convolution. The final classification result of the neural network.

2.2. Remote Sensing Image Classification Method

The theoretical basis of remote sensing image classification is that similar features in remote sensing images should have similar spectral characteristics under the same external conditions (texture, topography, illumination, season, and vegetation cover, etc.). In the process of transmission, remote sensing images may cause distortion of grayscale features due to various external factors or sensors, which will seriously affect the classification accuracy of images. Usually, the distortion of remote sensing images is mainly represented by noise and blur. At the same time, selecting the appropriate combination of features can make the contrast between different types of features the largest, and the contrast between similar features is the smallest, which is beneficial to the selection of training plots and computer statistics on gray features, greatly improving the classification accuracy.

Training rules are used to determine the principles and methods by which a neural network can output a desired target for a given input. In general, a neural network consists of an input layer, an output layer, and one or more hidden layers (see Figure 1). The input layer is used to obtain information, and the hidden layer is used for information processing and processing. The performance of neural networks.

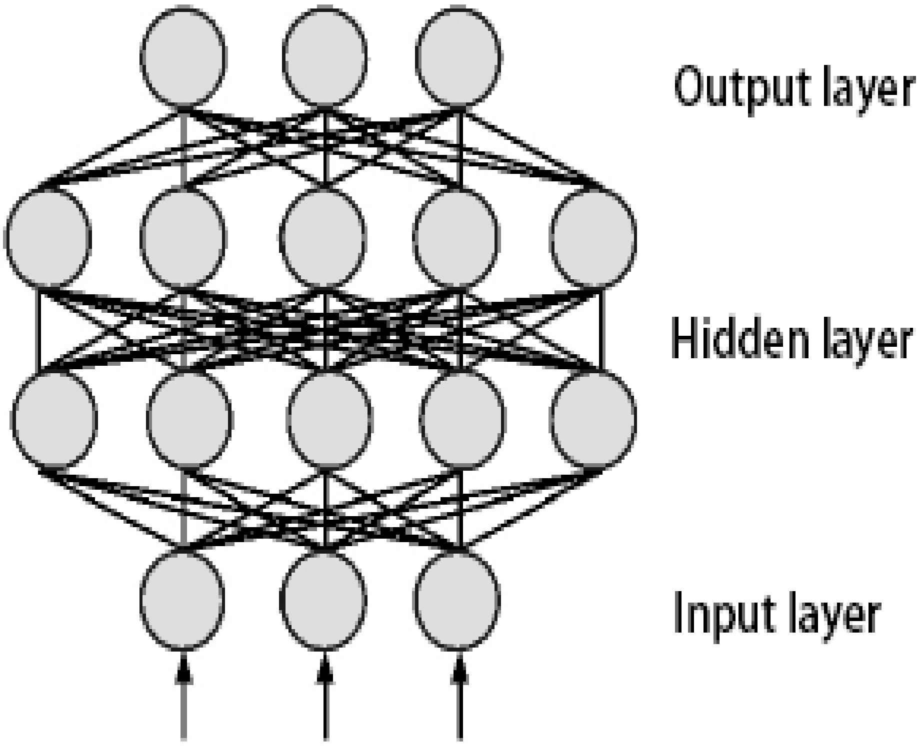

Figure 1. Neural network structure diagram

The strength of the connection between each pair of neurons is represented by a weight Wij, and the output value is changed by the input value of the upper layer. Except for the input layer, the inputs of each node on each layer are the weighted sum of the output values of all nodes in the previous layer. If the i layer is the input layer, the input obtained by the jth layer is

$$net_j = \sum_i o_i \cdot w_{ij} \tag{1}$$

Among them, the input value of the node is obtained by the logarithmic activation function Sigmoid(Log-Sigmoid):

$$o_j = f(net_j) = \frac{1}{1+e^{-net_j}} \tag{2}$$

Its error function is:

$$e_j = \frac{1}{2}\sum (t_j - o_j)^2 \tag{3}$$

After a given input sample is given to the network, it is necessary to learn according to a certain rule, that is, perform weight adjustment to obtain an output sample that satisfies the requirement, and this process is called network training. At present, a variety of neural network models have been developed, including perceptrons, multi-layer perceptrons, backpropagation, radial basis functions and other network models. The most widely used is the Back-Propagation network, referred to as BP network. . The Neural Net provided by ENVI uses standard backpropagation techniques for image classification. Users can select the number of hidden layers and set training parameters for classification.

The network structure design is required before classification, that is, the number of network layers and the number of nodes included in each layer are determined. In general, a complex network can obtain sample training results more accurately, but it may have poor generalization. Second, the more layers, the larger the calculation, and the corresponding increase in computation time. Many studies have shown that the three-layer network can solve most image classification problems. If the number of hidden layers is too large, the network computing time increases, the error does not necessarily decrease, and the network is more likely to fall into paralysis, affecting the classification results. The input nodes of the network are usually aligned with the number of spectral bands of the image to ensure that the spectral information of the image is fully utilized.

Network training is also required after obtaining training samples and determining the structure of the network. The neural network training learning provided by ENVI needs to determine several parameters, including initial weight value, learning rate, momentum factor, training error, and so on. The initial weight value determines the contribution of the internal weight related to the node activation level, which is used to adjust the internal weight change of the node; the learning rate determines the adjustment speed of the weight, and the selection of large parameters can speed up the learning but may cause the disk to vibrate. The result is not concentrated; the momentum factor mainly changes the direction of the weight, causing the weight to change along the current direction. The larger the value, the larger the step size of the training data change; the training error specifies the corresponding error value to determine the training termination condition. The above four parameters have a value range of 0 to 1. When using ENVI's neural network classification tool, the RMS error will be large, but if the training can be performed correctly, the error will gradually decrease and reach a stable lower value; if the error is unstable, the learning rate parameter should be changed or re Select the sampling area.

2.3. Discussion on Knowledge Integration Method

The concept of knowledge integration was first proposed in the context of new product development, which stems from the need to overcome the obstacles in the management of organizational systems and the realization of the value of user services. In 1990, Hendersen and Clark discussed the product innovation practice. They believe that enterprise product development requires two kinds of knowledge, namely, component knowledge and architecture knowledge. The problem-oriented component knowledge is driven by external market demand and promotes the solution-oriented architecture knowledge. The process that is produced is knowledge integration. In-house generation theory, resource-based theory, competence theory and knowledge management theory have jointly spawned knowledge integration theory. The theory of internal generation theory believes that the development of enterprises is the process of continuously expanding production scope through resource collection and knowledge accumulation. Resource-based theory and capability theory are explored from the perspectives of resource classification, management, coordination, development and resource-based capabilities. Based on the mechanism of internal resources to promote enterprise operation and build competitive advantage, knowledge-based theory and knowledge management theory have gradually determined the research perspective and system based on unique knowledge resources for enterprise management issues. Knowledge has been widely recognized as corporate activities. The important foundation.

Grant then formally proposed the concept of knowledge integration. It believes that under the premise that knowledge occupies a large part of the acquisition of organizational value and the importance of knowledge transfer and reuse to the knowledge strategy, the ability of the organization to express the ability to integrate knowledge rather than the knowledge itself. . Kahn also defines knowledge integration as a process for internal and cross-functional operations, linking knowledge and generating new knowledge through interaction and collaboration. When Alavi and Tiwana interacted with the research team, they believed that knowledge integration is the process of synthesizing individual expertise to form a system knowledge that conforms to a particular situation. Therefore, knowledge integration is a goal-oriented knowledge movement process that is reflected at different levels of the organization and re-integration of various knowledge to form new knowledge. The result will be enhanced by the ability of each level within the organization. Focusing on project knowledge integration, most scholars conceptualize the project context based on the definition of organizational knowledge integration. Among them, Tiwana believes that project knowledge integration is a team that combines individual-specific implicit and explicit knowledge into new, team-level project-specific knowledge. From the perspective of social interaction, Huang and Newell believe that project knowledge integration is a collective process of constructing, accumulating and refining shared beliefs through the social interaction of organizational members. Gao Weifang defined knowledge integration as the process of project team to optimize and integrate various innovation resources and knowledge elements under the guidance of system theory. Therefore, from the definition of existing concepts, in project-driven organizations, project knowledge integration has its more prominent connotation, clearer purpose and closer process. The strategy of project knowledge integration is closely related to corporate strategy and project type, and is also a bridge between organizational and personal knowledge integration. In summary, project knowledge integration is a process to complete the project objectives, with the project team members as the main body to gather and reorganize the knowledge required for the project in the project process, and to achieve effective integration of the project knowledge network and the organizational knowledge network.

Based on the above conceptual analysis of knowledge integration and project knowledge integration, project knowledge integration involves the integration of knowledge between individuals, teams, organizations and organizations at the same level and at different levels. Among them, the integration of project knowledge at the individual level is reflected in the integration activities of project members' own knowledge of the project. The integration of project knowledge at the team level is reflected in the integration activities of the project team as a whole, and the integration of project knowledge at the organizational level is reflected in the organizational function. Integration activities for project knowledge. Therefore, this paper will build a systematic research framework for project knowledge integration around knowledge integration research in project context.

3. Expriments

3.1. Experimental Environment and Objects

In this paper, the experimental environment is Windows 10 Pro (64bit), Intel® Core (TIM) i5-8300H, 32GB memory; the experimental equipment is Dell Tower Server T430; the experimental object software is TensorFlow 1.4; use GTX 1060 GPU for acceleration operation. The experimental object is a remote sensing image with a pixel size of 256*256; as shown in Figure 2 (image from the network).

Figure 2. Remote sensing image map

3.2. Experimental Program

The experimental data set contains 45 types of remote sensing scenes: airplanes, airports, baseball fields, basketball courts, beaches, bridges, sparse vegetation, palaces, circular farmland, clouds, commercial areas, dense residential areas, deserts, forests, roads, golf Courses, track and field, port, industrial area, crossroads, islands, lakes, parks, residential areas, mobile homes, mountains, overpasses, churches, parking lots, railways, railway stations, rectangular farmland, rivers, roundabouts, airports Runway, sea ice, boats, wetlands, sparse residential areas, stadiums, oil tanks, tennis courts, terraces, thermal power stations, grassland, each class includes 700 256*256 red, green and blue three-channel color (RGB) images, spatially resolved Rates range from 0.2m to 30m. The images in the dataset are from Google Maps and cover more than 21,000 images in more than 100 countries and territories around the world. The weather, season, illumination, and viewing angles in the dataset are quite varied, which is a test for the scene classification algorithm. Figure 3 shows a representative image in a 45-category scene image.

Figure 3. Scene graph of the data set

3.3. Training CNN

The dataset trained AlexNet, ResNet-50, ResNet-101, ResNet-152, DenseNet-152, DenseNet-121, DenseNet-152, DenseNet-161, DenseNet-169, Inception-v3, Inception-Res-v2, and Xception. 10 CNN networks. In order to compare horizontally with other papers, the same strategy is adopted in dataset partitioning, that is, two types of experiments are set up. The first type of experiment extracts 10% or 70*45 images from each scene of the NWPU-RESISC45 dataset. The training set, the remaining 90% as a test set; the second type of experiment extracts 20% or 140*45 images from each type of scene as the training set, and the remaining 80% as the test set. Before the training, the training set was data-enhanced, and the training set was randomly rotated, flipped, translated, cut, scaled, and stretched, and the final training set was expanded to 10 times.

In the test phase, the unclassified test set is used to test the classification performance of the network, and the test set data of the same category is input to the network as a batch. After the network outputs the prediction result, the classification of each class can be conveniently and accurately calculated. Rate, the average accuracy of classification for all categories can get the classification accuracy of the network as a whole. The classification accuracy of 10 kinds of networks in two types of experiments was statistically analyzed. In the process of screening the network, the running time of the model, the complementary nature between the models and the classification accuracy of the model were comprehensively considered, among which Inception-v3, Inception-Res- The performance of v2 and Xception networks on the remote sensing dataset is not as good as that of Res Net and Dense Net, the accuracy is not high, and the prediction speed is slow, and it is excluded. When choosing ResNet-101, DenseNet-121 and DenseNet-169, the diversity of the network is preserved to the greatest extent, including network depth, network structure, and diversity of convolution kernel parameters. In order to ensure the accuracy of the complexity measurement of BP neural network, only 4 of the 10 models are selected. This is because the input dimension of the designed BP neural network is only 17 dimensions, and the smaller the output dimension, the higher the prediction accuracy. When the output dimension is increased to 5 dimensions, the accuracy of BP network complexity prediction is reduced to less than 50%, which is disadvantageous for the entire integrated network. The parameters and test results of the four types of networks screened in the two types of experiments are shown in Table 1.

Table 1. Training parameters and results for each network

Model	Input size/ (pixel*pixel)	Batch size /frame	Number of cycles	Training accuracy /%	
				Experiment I	Experiment II
AlexNet	224*224	256	300	81.22	85.46
ResNet-50	224*224	256	300	86.52	90.52
ResNet-152	224*224	128	600	85.11	90.11
DenseNet-169	224*224	128	600	82.44	87.44
VGG-16				87.15	90.36
Proposed model				88.47	92.53

It can be seen from Table 1 that the accuracy of the second type of experiment is higher than that of the first type of experiment, that is, the accuracy of all networks increases as the proportion of the training set increases. The most prominent one in the single model is ResNet-50, which has the highest average classification accuracy; AlexNet has the lowest average classification accuracy, but the least number of layers and the fastest. DenseNet-169 has the largest number of layers, but the overall accuracy is reduced. This means that the increase in the number of layers does not necessarily lead to an increase in overall accuracy. DenseNet-169 is chosen for its scenes in trees, sea ice, islands, etc. The recognition accuracy is higher than that of other networks. In both types of experiments, the classification accuracy of the integrated model was higher than that of the other single networks.

4. Discussion

4.1. Experimental Results and Analysis

When using the integrated model for prediction, first calculate the complexity descriptor of the image to be predicted Fcomplex,followed by Fcomplex Input into the BP network, the BP network outputs the complexity level, selects its corresponding CNN network according to the complexity level, and finally predicts the scene image. After classifying and predicting the test set that is not participating in the training, the classification performance is calculated, and the confusion matrix is obtained. As shown in figure. 4, each element fCon(i, j) in the confusion matrix represents the probability that the image with the label i is recognized as j. . The diagonal elements represent the recognition accuracy of each class. It can be seen that the designed integrated network has better classification performance. Among the 45 types of scenes, 33 categories can achieve more than 90% recognition rate and the recognition accuracy of the trees. 95%, the lowest recognition rate is the tennis court, only 67% of the recognition accuracy. Scenes that are easily misjudged by classified networks include moderate residential areas and dense residential areas, railway stations and railways, palaces and churches. The reason for this is that the similarity of these scenes is higher, and on the other hand, the number of data sets is small. In the actual remote sensing interpretation task, the resolution of the network can be improved by increasing the amount of data.

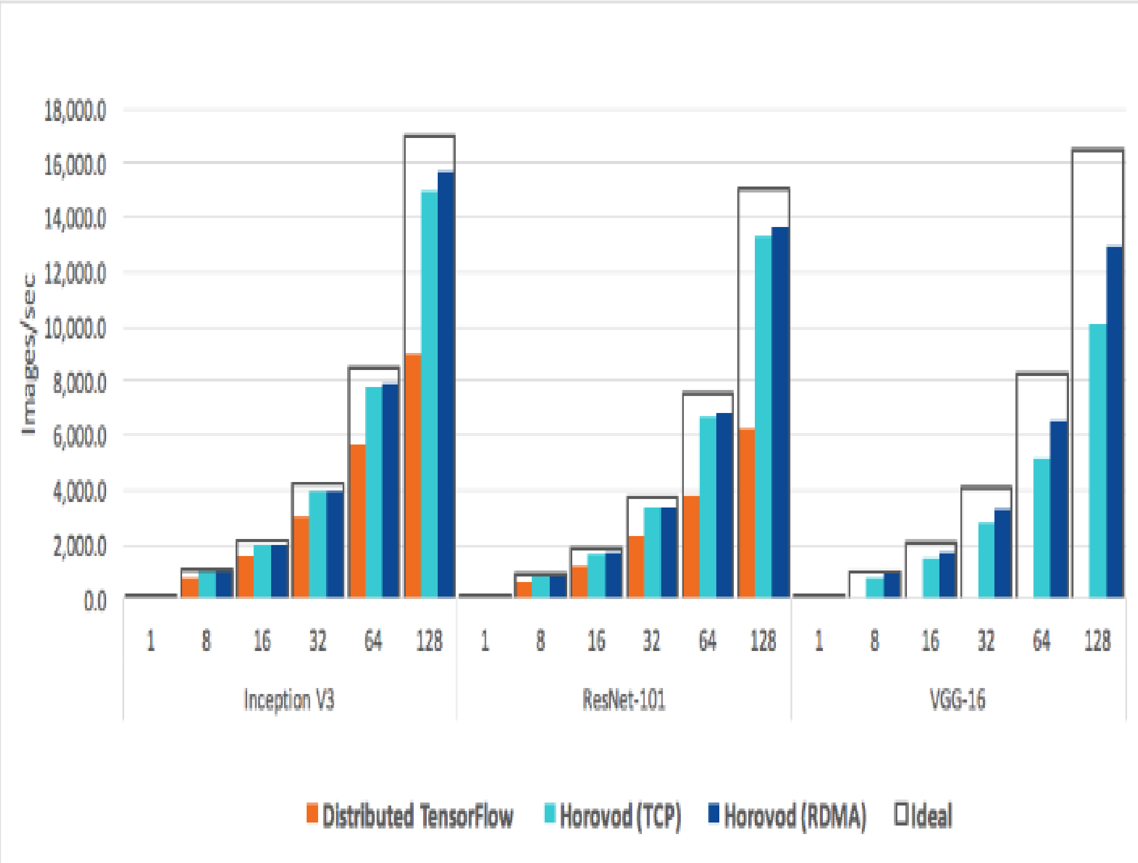

Figure 4. Data obtained by classification and prediction of the data set by the integrated model

In order to compare the performance of the proposed algorithm with other algorithms, according to the confusion matrix of the classification algorithm, the classification accuracy of several representative algorithms, proposed algorithms and network structures integrated by majority voting methods on the NWPU-RESISC45 dataset is drawn. ,As shown in Figure 5. The classification algorithm based on color histogram is selected from the algorithm based on low-level features. The visual poetry (BoVW) model is selected from the unsupervised learning method, and the VGG-16 model is selected from the deep learning-based method. All kinds of algorithms are algorithms with higher accuracy among their similar algorithms. At the same time, in order to test the effect of the proposed integration method, a single ResNet-50 model without integration strategy and a competition model with majority voting method were added. The selection of the training set and the test set of the comparative experiment is the same as that of the second type of experiment described in the third section, that is, randomly selecting 20% as the training set and 80% as the test set. The training and test environment uses the training machine and test machine described in Section 3, and the results are shown in Figure 5. It can be seen that the classification accuracy rate of each algorithm using the deep learning algorithm is higher than that of the low-level feature extraction algorithm and the unsupervised learning method

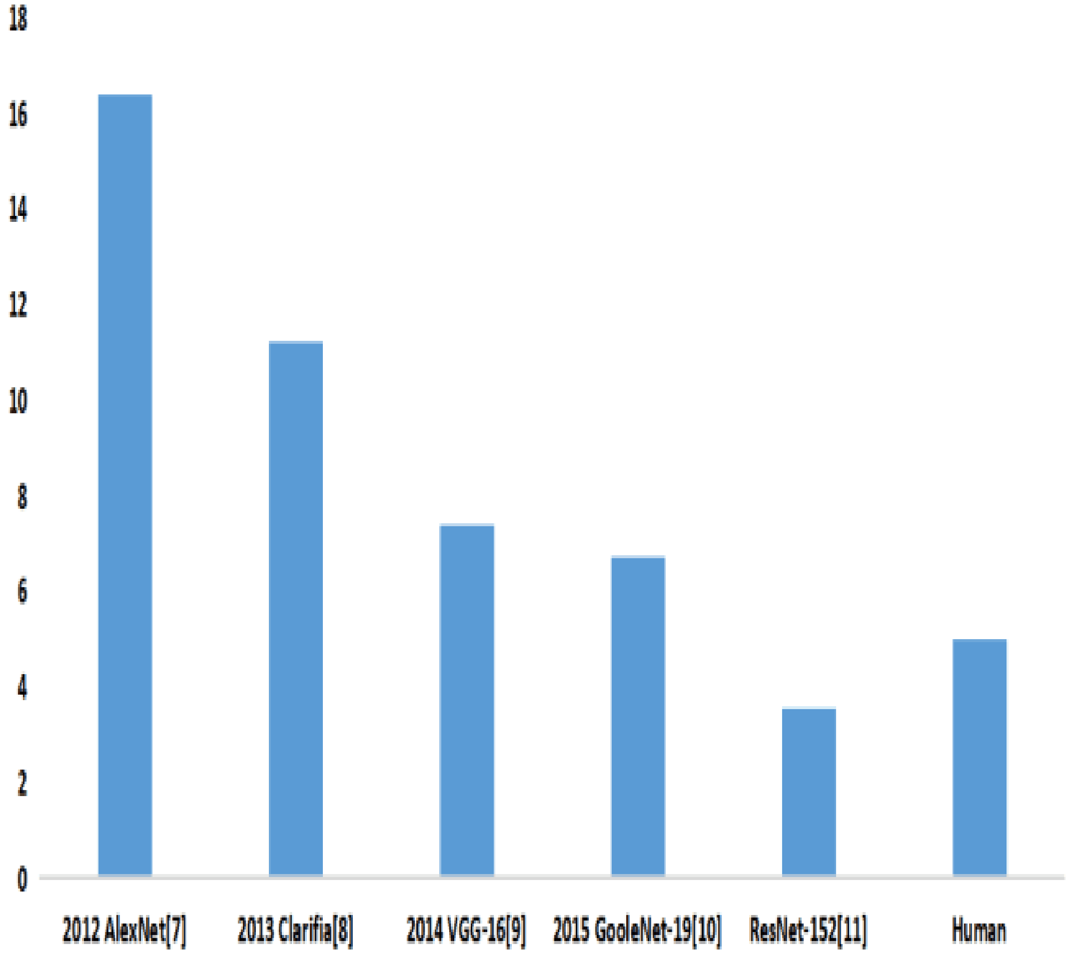

Figure 5. Accuracy comparison with other algorithm single classes

In order to compare the performance of various algorithms numerically, the average accuracy of the algorithm in figure 5, the standard deviation of the accuracy, and the average prediction time of the single scene image are tested and counted. The prediction time is the time taken from the input of a test set remote sensing image to the scene category of the image after the model is trained. In the proposed algorithm, the prediction time includes the time to extract the complexity descriptor, the complexity metric, and the CNN prediction category for the test set image.

The accuracy and prediction time of the integrated model proposed in the experiment are better than those of VGG-16 and ResNet-50, while the competitive network based on majority voting strategy shows very high accuracy in the classification experiment, but the method is The prediction phase needs to use four networks to perform prediction separately. The prediction time is greater than 2s, and the efficiency is low. As the average accuracy increases, the standard deviation of the accuracy decreases, indicating that the robustness of the classification system is improved, and the reduction of the standard deviation is beneficial to the risk assessment of the misclassification behavior in the remote sensing automatic interpretation task.

In addition to the improvement in accuracy in the integrated network proposed in this experiment, the prediction time consumed by the classification is also shorter than that of the single model, as shown in Figure 6.

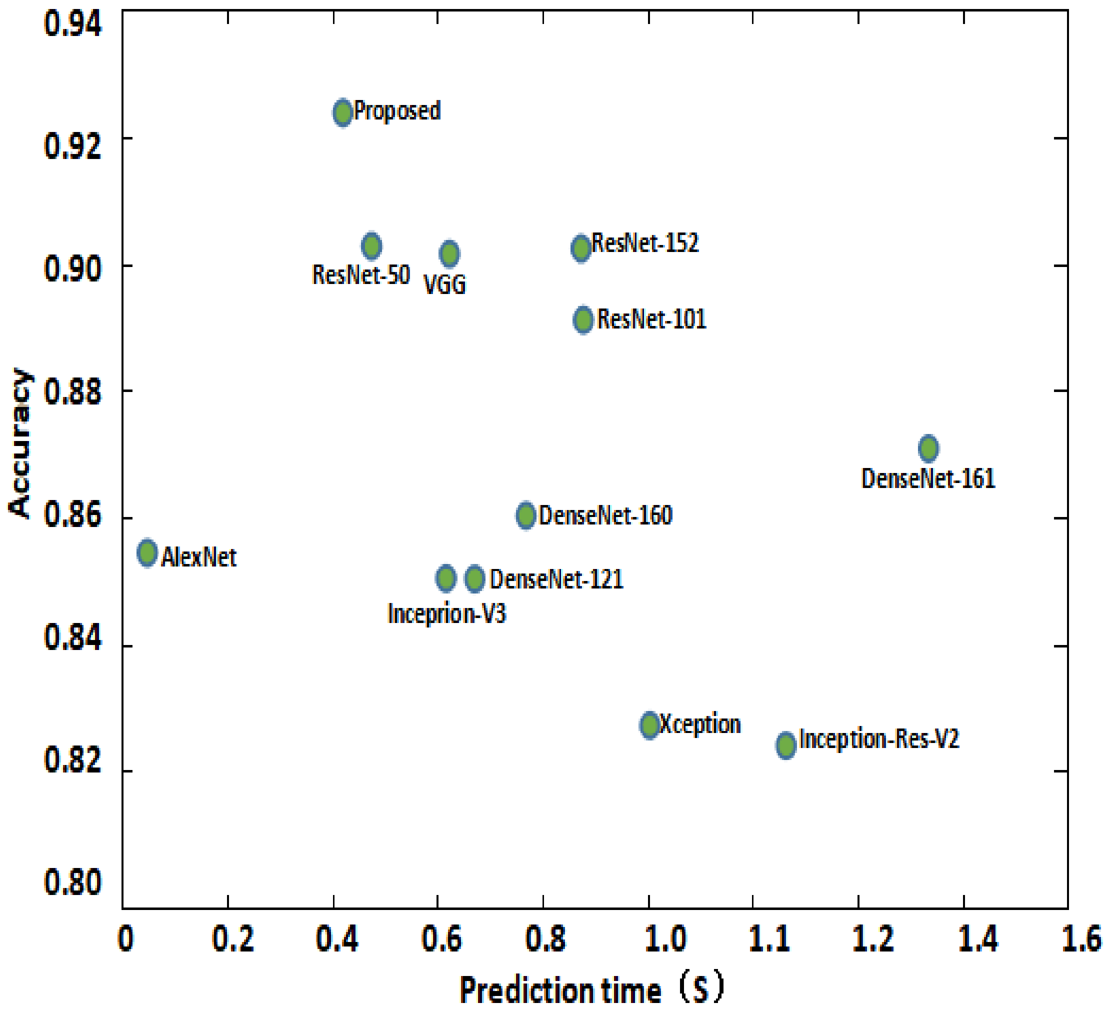

Figure 6. Classification accuracy and prediction time for multiple models

It can be seen from Fig. 6 that the prediction time of AlexNet on the test machine is 0.07s, and the overall classification accuracy is 85.16% (the second type of experiment); the predicted time of the trained VGG-16 model on the test machine is 0.62s, and the classification accuracy is 90.36. %; The predicted time of the trained ResNet-50 model on the test machine is 0.47s, the prediction speed is 12% higher than the VGG-16 model, and the classification accuracy is 90.59%; the integrated network calculates the complexity on the test machine and uses the BP network. The overall running time of the complexity measurement is 0.06s, the average prediction time of the integrated network is 0.41s, and the classification accuracy is 92.53%. Compared with the VGG-16 model, the prediction speed is increased by 33%, and the accuracy is improved. 2.17%.

5. Conclusion

(1)Multiple CNNs are integrated to classify and identify remote sensing scenes, and the complexity of the image is determined before using CNN to predict, and then the CNN that best matches the scene to be predicted is found. The shallow network is used to identify the scene image with low complexity to achieve the purpose of rapid identification; the deep network is used to identify the scene image with high complexity to achieve accurate recognition. The results show that ResNet-50 performs better than VGG-16 on the NWPU-RESISC45 dataset. The integrated model is superior to VGG-16 and ResNet-50 in single model, classification. The overall efficiency is improved. The experimental results demonstrate the efficiency and feasibility of the integrated model. The integrated model needs to train and store multiple models. Compared to a single convolutional neural network, it requires longer training time and model storage space. It can reduce the number of model parameters by appropriate model compression, and can also adopt parallel optimization of multiple graphics cards. Accelerate training methods to reduce training time.

(2)In this paper, based on the fuzzy comprehensive evaluation method, in the image comprehensive evaluation research for image classification, the evaluation model is established according to the specific application, the image classification needs to analyze the image quality, and a comprehensive evaluation model for classification is established. The development of the corresponding program verifies the rationality and applicability of the model, and makes an attempt in the research of remote sensing image comprehensive evaluation. It has certain reference value and reference significance for the comprehensive evaluation of remote sensing image for specific applications in the future.

(3)The use of deep learning methods can significantly improve the speed and accuracy of remote sensing image detection, which has promoted the accurate application of remote sensing images. However, there are still many shortcomings in the detection methods in this paper. The next step will be to improve the segmentation network and more. Experiment with types of data sets to further improve detection accuracy.

References

[1] Romero A, Gatta C, Camps-Valls G. Unsupervised Deep Feature Extraction for Remote Sensing Image Classification. IEEE Transactions on Geoscience & Remote Sensing, 2016, 54(3):1349-1362. DOI:10.1109/TGRS.2015.2478379

[2] Maggiori E, Tarabalka Y, Charpiat G, et al. Convolutional Neural Networks for Large-Scale Remote-Sensing Image Classification. IEEE Transactions on Geoscience & Remote Sensing, 2016, 55(2):645-657. DOI:10.1109/TGRS.2016.2612821

[3] Min H, Liu B. Ensemble of extreme learning machine for remote sensing image classification. Neurocomputing, 2015, 149(PA):65-70. DOI:10.1016/j.neucom.2013.09.070

[4] Liu D, Ling H, Han X. High Spatial Resolution Remote Sensing Image Classification Based on Deep Learning. Acta Optica Sinica, 2016, 36(4):0428001. DOI:10.3788/AOS201636.0428001
[5] Shi Q, Du B, Zhang L. Spatial coherence-based batch-mode active learning for remote sensing image classification.. IEEE Transactions on Image Processing, 2015, 24(7):2037-2050. DOI:10.1109/TIP.2015.2405335
[6] Li W, Fu H, Le Y, et al. Stacked Autoencoder-based deep learning for remote-sensing image classification: a case study of African land-cover mapping. International Journal of Remote Sensing, 2016, 37(23):15.DOI:10.1080/01431161.2016.1246775
[7] Wang L, Zhang J, Peng L, et al. Spectral–spatial multi-feature-based deep learning for hyperspectral remote sensing image classification. Soft Computing, 2017, 21(1):213-221. DOI:10.1007/s00500-016-2246-3
[8] Yue H, Chen W, Wu X, et al. Visualizing bag-of-words for high-resolution remote sensing image classification. Journal of Applied Remote Sensing, 2016, 10(1):015022. DOI:10.1117/1.JRS.10.015022
[9] Li J, Qian D, Li Y. An efficient radial basis function neural network for hyperspectral remote sensing image classification. Soft Computing, 2015, 20(12):1-7. DOI:10.1007/s00500-015-1739-9
[10] Wan L, Tang K, Li M, et al. Collaborative Active and Semisupervised Learning for Hyperspectral Remote Sensing Image Classification. IEEE Transactions on Geoscience & Remote Sensing, 2015, 53(5):2384-2396. DOI:10.1109/TGRS.2014.2359933
[11] Matasci G, Volpi M, Kanevski M, et al. Semisupervised Transfer Component Analysis for Domain Adaptation in Remote Sensing Image Classification. IEEE Transactions on Geoscience & Remote Sensing, 2015, 53(7):3550-3564. DOI:10.1109/TGRS.2014.2377785
[12] Schmidhuber J. Deep learning in neural networks: an overview. Neural Netw, 2015, 61:85-117. https://doi.org/10.1016/j.neunet.2014.09.003
[13] Chan T H, Jia K, Gao S, et al. PCANet: A Simple Deep Learning Baseline for Image Classification?. IEEE Transactions on Image Processing, 2015, 24(12):5017-5032.DOI:10.1109/TIP.2015.2475625
[14] Alipanahi B, Delong A, Weirauch M T, et al. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. Nature Biotechnology, 2015, 33(8):831. DOI:10.1038/nbt.3300
[15] Chen Y, Lin Z, Xing Z, et al. Deep Learning-Based Classification of Hyperspectral Data. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2017, 7(6):2094-2107. DOI:10.1109/JSTARS.2014.2329330
[16] Litjens G, Kooi T, Bejnordi B E, et al. A survey on deep learning in medical image analysis. Medical Image Analysis, 2017, 42(9):60-88. DOI:10.1016/j.media.2017.07.005
[17] Lv Y, Duan Y, Kang W, et al. Traffic Flow Prediction With Big Data: A Deep Learning Approach. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(2):865-873. DOI:10.1109/TITS.2014.2345663
[18] Gulshan V, Peng L, Coram M, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. Jama, 2016, 316(22):2402. DOI:10.1001/jama.2016.17216
[19] Herstad S J, Sandven T, Ebersberger B. Recruitment, knowledge integration and modes of innovation. Research Policy, 2015, 44(1):138-153. DOI:10.1016/j.respol.2014.06.007
[20] Revilla E, Knoppen D. Building knowledge integration in buyer-supplier relationships. International Journal of Operations & Production Management, 2015, 35(10):1408-1436.

DOI:10.1108/IJOPM-01-2014-0030

[21] Papanikolaou N, Pavlopoulos G A, Pafilis E, et al. BioTextQuest(+): a knowledge integration platform for literature mining and concept discovery. Bioinformatics, 2015, 31(6):3249-3256. DOI:10.1093/bioinformatics/btu524

[22] Cao X, Guo X, Liu H, et al. The role of social media in supporting knowledge integration: A social capital analysis. Information Systems Frontiers, 2015, 17(2):351-362. DOI:10.1007/s10796-013-9473-2

[23] Sankowska A, Söderlund J. Trust, reflexivity and knowledge integration: Toward a conceptual framework concerning mobile engineers. Human Relations, 2015, 68(6):973-1000. DOI:10.1177/0018726714549646

[24] Butler J R A. The challenge of knowledge integration in the adaptive co‐management of conflicting ecosystem services provided by seals and salmon. Animal Conservation, 2015, 14(6):599-601. DOI:10.1111/j.1469-1795.2011.00509.x

[25] Schwendimann B A, Linn M C. Comparing two forms of concept map critique activities to facilitate knowledge integration processes in evolution education. Journal of Research in Science Teaching, 2016, 53(1):70-94. DOI:10.1002/tea.21244

[26] Ou L L, Ryoo K, Linn M C, et al. Measuring Knowledge Integration Learning of Energy Topics: A two-year longitudinal study. International Journal of Science Education, 2015, 37(7):1044-1066. DOI:10.1080/09500693.2015.1016470

[27] Tsai K H, Liao Y C, Hsu T T. Does the use of knowledge integration mechanisms enhance product innovativeness? ☆. Industrial Marketing Management, 2015, 46:214-223. DOI:10.1016/j.indmarman.2015.02.030