Remote Sensing Image Scene Classification Based on Convolutional Neural Networks

Yumei Liu
Chongqing City Vocational College, Chongqing, 402160, China
E-mail: YumeiLiu6@163.com

Keywords: remote sensing images, land classification, convolutional neural network, feature fusion

Received: March 14, 2024

Remote sensing image classification, a specific application of digital signal technology in remote sensing, addresses the challenge of effectively processing and categorizing remote sensing imagery. This research proposes a neural network (CNN)-based approach for remote sensing image classification, aiming to overcome the limitation of single-feature inadequacy. The method involves making a multi-site and multi-combination strategy that effectively combines spectral features, spatial patterns, and more remote sensing images as vectors or matrices. We then train the CNN model based on the length of the data. Experimental results demonstrate a significant reduction of approximately 80% in the training time for the PCA-free CNN (SST) method after implementing the PCA transformation. This reduction not only expedites the training process but also enhances overall accuracy by approximately 3.49. The CNN-style network model contributes to efficiency improvement. Larger training models increase the number of models to be taught, slowing down the training process and prolonging learning times. The incorporation of multi-location and multi-combination strategies accelerates tracking speed and enhances the classification accuracy of remote sensing images. Comparative analysis indicates that, in contrast to other classification methods, CNN achieves superior classification performance, demonstrating its capability for increased categorization and improved accuracy.

Povzetek: Izvedena je klasifikacija prizorov daljinskega zaznavanja z uporabo konvolucijskih nevronskih mrež (CNN), vključujoč fuzijo več virov podatkov, zmanjšanje dimenzionalnosti in izboljšano klasifikacijo prizorov.

1 Introduction

Various sensing instruments emit electromagnetic waves that distant targets reflect, forming remote sensing images. Remote sensing images consist of brightness features that form a spectral space. Each type of terrain has its own unique spectral characteristics due to differences in its sensitivity to different wavelengths of light. Different electromagnetic interference and weather conditions, like radiation in the air, changes in magnetic fields, the scanner's view, the time of the shot, and other things, cause the spectral information to show different ground features. To classify remote sensing images, you have to look at the spectral features of different objects on the ground, flip and infer the target map's geometric and physical features, divide the feature space into irrelevant sub-regions that correspond to the classification category, and put each pixel of the image into a subspace for classification. Remote sensing images have their characteristics compared to ordinary images, mainly including below-mentioned observations.

i. Massive data, each imaging point on the ground has corresponding spectral information and remote sensing images are collected periodically with continuity [1].

ii. Uncertainty, remote sensing images are affected by external conditions such as weather and light, resulting in varying spectral information reflected by ground objects. Remote sensing images also have certain limitations, such as the presence of foreign objects with the same spectrum or the phenomenon of objects with different spectra.

iii. Timeliness, the acquisition cycle of remote sensing images is short and has strong timeliness.

iv. Overall, the processing of remote sensing images requires reference to other geographic data information [2].

Remote sensing image classification distinguishes different ground objects by the size of pixel values in each band. Due to chemical, physical, and other factors, different ground objects have different reflection effects under the illumination of the same band of light, resulting in different spectral information. This spectral information is the basis for remote sensing image classification. Extracting discriminative statistics from raw data is the foundation for feature extraction and remote sensing image classification, which can quantitatively extract spectral features, texture features, and spatial features. The main purpose of remote sensing image classification is to obtain ground object information from remote sensing images to identify the actual types of ground objects. Fundamentally, it categorizes each pixel of the image into a predetermined category, achieving the conversion of the two-dimensional space of the image to the target space.

Remote sensing images contain a large amount of information and a variety of ground objects. With the increasing display of advantages in information processing by artificial intelligence, remote sensing image classification techniques are becoming more artificial intelligence, such as artificial neural networks, active learning, support vector machines, etc. [3]. For remote sensing image classification, typical neural network models cannot fully explore the correlation between the features of this category of land features in the image and the surrounding land features. The lack of considering the influence of the previous feature on the surrounding area, as well as the characteristics and distribution of the current feature itself, in the current classification, leads to limited dynamic variability. Compared to ordinary artificial neural networks, deep neural networks have higher computational levels [4]. With the rapid development of machine learning and deep learning, image classification has become smarter. Using the self-learning capabilities of neural networks can improve the accuracy of remote sensing image classification results. Although there are many neural network-based remote sensing image classifications, including ship detection, aircraft target classification, and vegetation classification, as currently the most popular deep learning algorithm, how to better use convolutional neural networks for remote image classification is a problem worthy of deep and dense research.

2 Literature review

Remote sensing devices, such as satellites, monitor the state of the Earth's surface by identifying ground targets through image classification. Remotely sensed images produce low-resolution, blurry images, depending on factors such as sensor type, wavelength, and shooting range. Therefore, it is difficult to distribute images from a distance. Considering the small number of samples in remote sensing images, remote sensing image classification methods usually focus on decomposing algorithms for many images with small distribution patterns that are suitable for remote sensing. Researchers have proposed more flexible and general feature extraction methods, such as sparse coding and Fisher functions, to adapt to various application functions [5-7]. Researchers widely study deep neural networks in many fields of computer vision due to their excellent performance [8]. Their success lies in their ability to learn features that can generalize large amounts of learning data. The use of deep neural networks in remote sensing is of interest to color researchers. Researchers have not conducted in-depth research on training remote-sensing image data to neural terminals for remote-sensing image classification tasks [9]. The main reason is that deep neural networks require a lot of training data, and remote sensing image data has a small amount of data, so it should not be such a big request. Large-scale remote sensing image data did not appear until 2015. Zhang *et al.* proposed a remote-control scene classification algorithm based on training parameters [10]. First, the evolutionary framework aims to simultaneously optimize the hyperparameters of the optimizer and the weight bounds of convolutional neural networks. Second, the population can learn hyperparameters by combining two individuals and adjusting CNN weights using their information. Finally, equality accelerates learning because two people can develop at the same time. Many experiments on our data have shown that the proposed model is effective. Recently, researchers have introduced several Convolutional Neural Network (CNN)-based methods with high-quality and high-resolution models for remote sensing phenomenon classification, which is a hot topic. Based on a large amount of training data, CNN can extract a variety of features and learn how to predict remote locations. However, deep models rely on multiple remote-sensing image tags to control the learning process, which makes it difficult to predict the process. Therefore, the light learning model of training is very important. Simple classification models and complex models make training programs unreliable and cause flood models to fail. Therefore, Liang and Wang proposed a new enhanced communication network for remote sensing (ERANet) [11]. Efficientnet replaced the old backbone as a lighter backbone of the ARCNet framework, unlike deep learning methods. The Gabor Convolutional Network (GCN) is a combination of the Gabor filter and the Deep Neural Network (CNN) to discriminate between different directions and frequencies. To the author's knowledge, Moudjib *et al.* first introduced the application of GCN in hyperspectral image classification [12]. Even without training samples, HSI-GCN can extract depth contrast from radiographs faster and more accurately. The writer carefully checks how well the methods work on different kinds of hyperspectral data. When compared to the old CNN and Gabor methods, this one work better and gets better results in classifying.

Currently, there is little research on the use of deep neural networks for remote sensing image classification in China. Therefore, the author proposes a remote sensing image classification model based on a neural network (CNN) solution. In response to the problem that a single source cannot provide more useful information, the author created multiple sources and multiple fusion methods. These features combine spectral features, texture patterns, contrast-related features, and other parts of remote-sensing images as vectors or matrices based on the length of the image. Researchers use these features to train degenerate neural network models. Experimental results show that this fusion method can make the CNN training model more abstract and high-level, improve classification accuracy, and achieve the best distribution of results.

3 Methods

3.1 Introduction to convolutional neural networks

Convolutional neural networks, a typical model for deep learning, are multi-layer neural network models. It usually consists of an input layer, a solution layer, a low-level layer (reservoir layer), a general connection layer, and an output layer. Figure 1, presents the model of convolutional neural network.

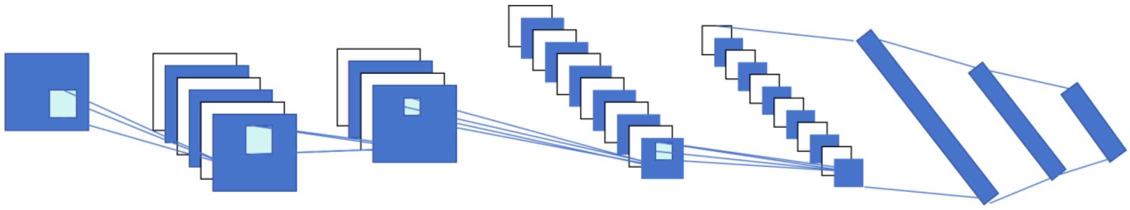

Figure 1: A diagram of the convolutional neural network

Firstly, the input layer receives the original image, and then the convolution layer is used to extract the feature of the image and reduce the impact of noise. $Y_0 = X_0$, assuming the input raw image is X, and Y_i is the i^{th} layer characteristic map [13]. Assuming Y_i is the i^{th} convolutional layer, then Equation 1 represents the functionality.

$$Y_i = f(W_i \cdot Y_{i-1} + b_i) \tag{1}$$

In this equation, W_i is the weight of the convolution kernel of the i^{th} layer, and the operator $\cdot$ is the convolution of the layer $i-1$. b_i represents the bias vector of the i^{th} layer; f is a non-linear activation function, which is usually represented by the ReLU function and is represented in Equation 2.

$$f(x) = \begin{cases} 0, x < 0 \\ x, x \geq 0 \end{cases} \tag{2}$$

The down-sampling layer follows the convolutional layer closely and reduces the dimensionality of the feature map based on the local correlation of the image while maintaining the scale invariance of the features. Assuming Y_i is the down-sampling layer feature map, then samples are evaluated using Equation 3.

$$Y_i = subsample(Y_{i-1}) \tag{3}$$

In general, there are two aggregation methods: maximum aggregation and intermediate aggregation. After changing the connection between the multi-resolution process and the underlying layer, the entire process continues to reduce the length of the extracted features. Finally, the output layer outputs the corresponding labels of the samples based on the feature vectors extracted from the full concatenation process [14].

The classification process of neural network problem solving is mainly the learning process of the network, which is equivalent to the learning process of the human brain. It is divided into two stages:

i. Forward propagation, which leads to pattern learning from the input layer to the output layer.
ii. Backpropagation calculates the error between the result and the expected result based on a loss

function L (W, b) called the "residual" and adjusts the network without using the standard gradient descent of the layers.

Currently, the loss function widely used in CNNs is the cross-entropy (CE) loss function, which is calculated by Equation 4.

$$L(W,b) = CE(W,b) = -\sum_{i=1}^{N}\sum_{j=1}^{c} 1\{\hat{y}_i = j\} \log p_i^j \quad (4)$$

In the equation: $\hat{y}_i$ is the expected value of the i^{th} training sample, p_i^j is the prediction probability of the jth category of the i^{th} training sample, C is the total number of training samples and N is the total number of training samples. The training goal of the neural problem solver is to minimize the loss function $L\ (W,b)$ of the network by gradient descent [15]. During the entire training process, the loss value is calculated from the posterior distribution, and then the recovery process is performed with gradient lips, and the training of W and b is updated for each layer. The parameters of the formula change are defined as presented in Equation 5 and 6.

$$W_i = W_i - \eta \frac{\vartheta L(W,b)}{\vartheta W_i} \quad (5)$$

$$b_i = b_i - \eta \frac{\vartheta L(W,b)}{\vartheta b_i} \quad (6)$$

In the above Equations: η learning rate of the network is used to control the intensity of backpropagation of loss values.

3.2 Remote sensing image classification method based on CNN

3.2.1 CNN classification model

The AlexNet network model is a deep convolutional neural network model designed by Alex Krizhevsky in 2012. Considering that the model has not very deep layers and has good classification performance, therefore, the author constructed a CNN model suitable for remote sensing image classification based on the AlexNet model, and its model structure is shown in Figure 2.

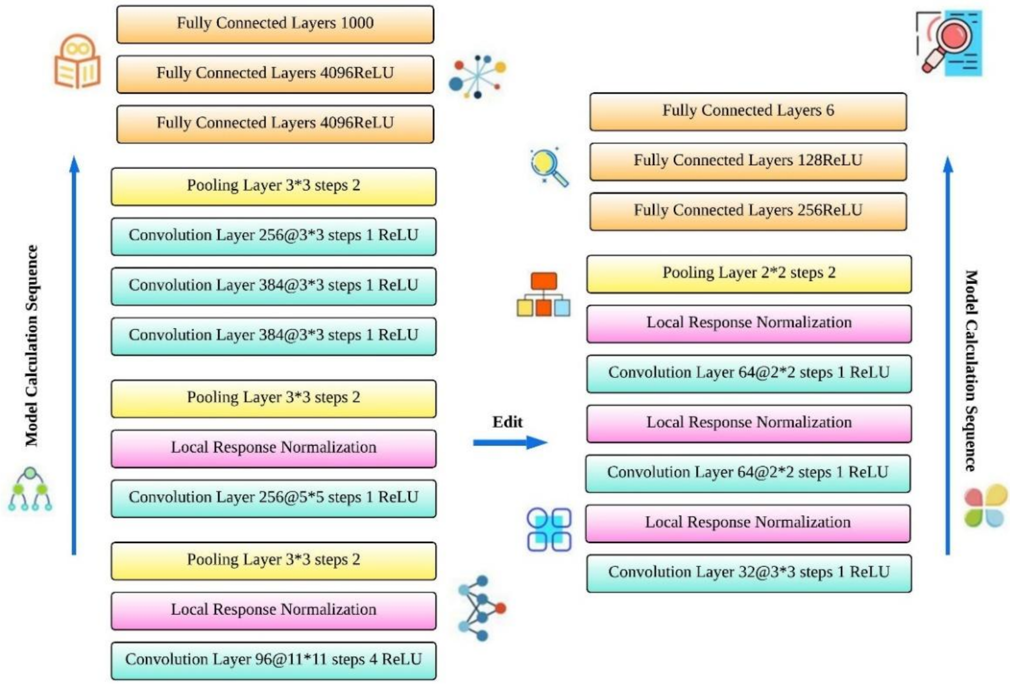

Figure 2: Proposed model

The difference from AlexNet is that the designed CNN model removes a portion of the convolutional and pooling layers, as the pooling layer is mainly used for dimensionality reduction, while the author classifies remote sensing images at the pixel level [16]. To prevent training samples from being too small in size at higher levels, only the pooling layer of the last layer is retained. Set 3 convolutional layer parameters to 32@3 × 3 64@2 × 2 and 64@2 × 2, steps are all 1. The pooling layer adopts a size of 2 × 2. Maximize pooling with a step size of 2. The Dropout parameter is set to 0.5 during training and 1 during testing, and the Adam optimization algorithm is used.

3.2.2 Multi-source and multi-feature fusion

Winsize × At present, most classification methods for winsize are based on manually designed single source features. Due to the fact that single source features often cannot reflect the differences between all ground object categories well, the generalization performance of classification algorithms is poor. To address this issue, the author has designed a method for multi-source and multi feature fusion. For each pixel point, neighborhood pixels with a surrounding size of were considered, which helps to eliminate speckle noise in the image. Firstly, PCA is used to transform the original data, and then the first three principal components (PCA1, PCA2, PCA3) containing almost 95% or more information from all bands are selected as the transformed original image. Next, extract the spectral values corresponding to the training samples to form a one-dimensional spectral feature vector A, and calculate NDVI to form a one-dimensional feature vector B. Secondly, calculate the gray level co-occurrence matrix (GLCM) for each image, and extract eight second-order probability statistical texture filters based on GLCM, including mean, variance, entropy, angular second-order distance, correlation, dissimilarity, contrast, and synergy, form a texture feature matrix C in the order of extraction [17]. Finally, perform K-T transformation on the image to extract data from three components: brightness, greenness, and humidity to form a feature matrix D.

Experiments and analysis

3.3 Experimental environment

The author adopts the TensorFlow1.1.0 open-source framework, built in a personal PC environment with Ubuntu 16.04 operating system and Intel (R) Core (TM) i5-4440 processor CPU@3.10 GHz, with 8 GB of running memory.

3.4 Experimental data and sample selection

The study area is located in the western part of some states. According to the actual distribution of land in the area, there are 6 categories: agricultural fields, pastures, wetlands, reservoirs, residential areas, and bare land. The author used a Landsat-8 satellite remote sensing image from September 2016 and selected all training samples at the level of 16160 pixels. Of these, 1/4 of the data samples for each soil type were selected as valid samples, and the remaining data samples were used as training methods [18].

3.5 Experimental results and analysis

To verify the effectiveness of this method, experiments will be compared with other methods in the literature, such as SVM, NN, RF, DBN, and CNN (Patch). The CNN (Patch) method is a CNN model based on region blocks, which takes neighborhood blocks of 5×5 size around pixels as the input of a single sample and can be seen as adding neighborhood information to the sample. In addition, based on this method, the author has designed two other comparative experimental models. One is to only use spectral features as input to CNN (CNN (ST)), which is to verify the effectiveness of the author's multi-source and multi-feature fusion method; Another type is the multi-source and multi-feature fusion model (CNN (SST)) that does not use PCA. The main purpose of this model is to verify whether PCA can shorten the training time of the model and accelerate convergence speed [19]. Table 1 shows the confusion matrix of the classification results of this method, indicating that CNN (PCA) ultimately achieved an overall accuracy of 97.83%, with a Kappa coefficient of 0.973 6. For single class land features, both PA and UA exceed 90%, and the classification accuracy of certain land features, such as water and bare land, has reached the best, with both PA and UA reaching over 99%. However, for cultivated land categories, the classification accuracy is relatively poor [20]. The model misclassifies them into wetland categories with a large number of samples. The reason for the analysis is that the spectral value range of crops overlaps with that of wetlands, which is a serious phenomenon of “foreign objects in the same spectrum”, making it difficult for the model to effectively distinguish between cultivated land and wetlands.

Table 1: Confusion matrix of the CNN (PCA) classification results

Category	Cultivated land	Meadow	Wetlands	Waters	Residential areas	Bare ground	Total	PA/%	UA/%
Cultivated land	647	11	39	0	3	0	700	92.43	98.33
Meadow	3	596	1	0	0	0	600	99.33	97.70
Wetlands	1	3	576	0	0	0	580	99.31	93.51
Waters	0	0	0	520	0	0	520	100.00	100.00
Residential areas	6	0	0	0	593	1	600	98.83	99.33
Bare ground	1	0	0	0	1	230	232	99.14	99.57
Total	658	610	616	520	597	231			—

Table 2 shows the comparison of experimental results between this method and other classification methods, and it can be seen from the table that this method achieved the best classification effect. Compared to SVM, NN, and RF, the overall accuracy of CNN (PCA) has been improved by about 13.61%, 9.34%, and 7.3%, respectively [21]. The Kappa coefficient has been improved by about 20.42%, 13.16%, and 10.02%, indicating that CNN's classification performance is far superior to shallow classification algorithms, this is thanks to the unique structure of CNN, such as local connections, weight sharing, pooling, etc. These characteristics enable CNN to have a certain scale of displacement, scale, and deformation invariance. Its powerful learning and fault tolerance capabilities enable CNN to automatically learn more abstract and representative features, thereby achieving higher classification accuracy. However, shallow classification algorithms cannot obtain more useful information from the original samples when the information is insufficient or insufficient. Similarly, compared to other deep learning methods, the classification accuracy of CNN (PCA) is also higher than that of DBN and CNN (Patch). The reason for the analysis is that DBN uses an unsupervised method to train the network layer by layer, and finally uses a supervised method for fine tuning. This layer-by-layer training method causes greater randomness in the network parameters, which is not conducive to the overall optimization of the network [22]. For CNN (PCA), although this method also uses a CNN model, the input information only considers the spectral information around the sample, without considering the texture features of the image. It is precisely the texture features that best reflect the differences between different categories, resulting in poor classification performance. From another perspective, it can also illustrate this point, as designed by the author, the comparative experimental models CNN (ST) and CNN (SST) only consider the spectral features of the samples, while the latter combines spectral and texture features, which improves the overall accuracy of CNN (SST) by about 7.2% and the Kappa coefficient by about 5.29% compared to CNN (ST). Moreover, the CNN (SST) method can effectively improve the severe “salt and pepper” phenomenon in CNN (ST).

Table 2: Comparison of the classification effects of different classification methods

Method	Category (%)						OA/%	Kappa
	Cultivated land	Meadow	Wetlands	Waters	Residential areas	Bare ground		
SVM	50.29	79.84	97.76	100.00	85.33	99.57	84.22	0.808 5
NN	63.00	92.00	98.45	100.00	91.33	98.26	88.49	0.860 4
RF	66.29	91.17	99.83	100.00	99.00	95.69	90.53	0.884 9

DBN	81.57	93.83	94.83	100.00	94.83	99.57	92.95	0.917 1
CNN(Patch)	67.00	98.00	97.59	100.00	92.17	95.26	90.25	0.881 8
CNN(ST)	58.71	96.83	95.86	100.00	88.50	93.97	87.14	0.884 3
CNN(SST)	83.14	96.17	97.93	100.00	95.17	99.57	94.34	0.931 1
CNN(PCA)	92.43	99.33	99.31	100.00	98.83	99.14	97.83	0.973 6

Figures 3 and 4 respectively show the comparison results of the author's designed comparison model regarding training time and training accuracy. From Figure 3, it can be seen that after using PCA transformation, the training time of the CNN (SST) method without PCA is reduced by about 80%, which not only significantly shortens the training time, but also improves the overall accuracy by about 3.49%. This is related to the network structure of the CNN model itself. The higher the dimension of input sample information, the more parameters the model needs to train will increase geometrically, resulting in slow training speed and long training time [23-25]. As shown in Figure 4, the model performance of CNN (PCA) tends to stabilize around 300 iterations, while CNN (ST) and CNN (SST) only stabilize after 2000 and 1100 steps, respectively. This indicates that after PCA dimensionality reduction, the model converges faster and performs better. The reason is that after PCA transformation, the inter class gap increases, the intra class gap decreases, and the noise information in the samples can be eliminated to a certain extent, thereby accelerating the convergence speed of the model and improving the overall classification accuracy.

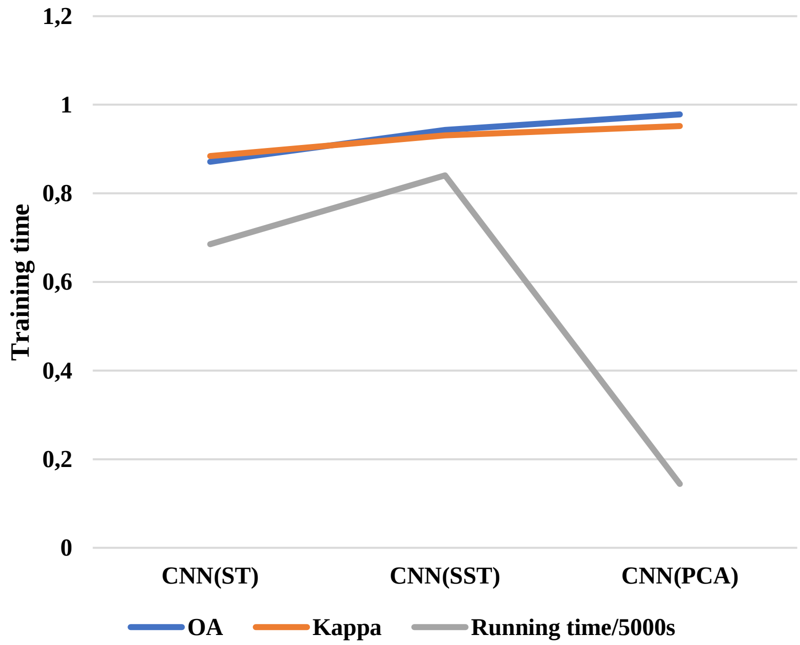

Figure 3: Performance comparison of different methods

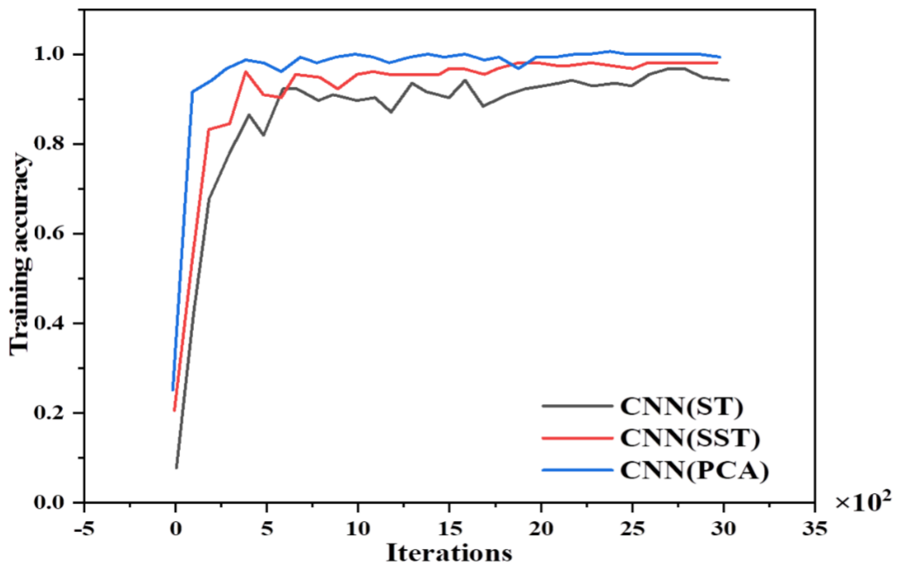

Figure 4: Comparison of the training accuracy of the different methods

The training time of the PCA-free CNN (SST) method was cut by an amazing 80% in the experiments that were done after PCA transformation was added. Additionally, there is an observed improvement in overall accuracy of approximately 3.49. The experiments highlight the efficacy of the proposed multi-site and multi-combination strategy in enhancing the tracking speed and classification accuracy of remote sensing images. Comparative analysis with other classification methods emphasizes the superior performance of CNN in achieving increased classification and improved accuracy.

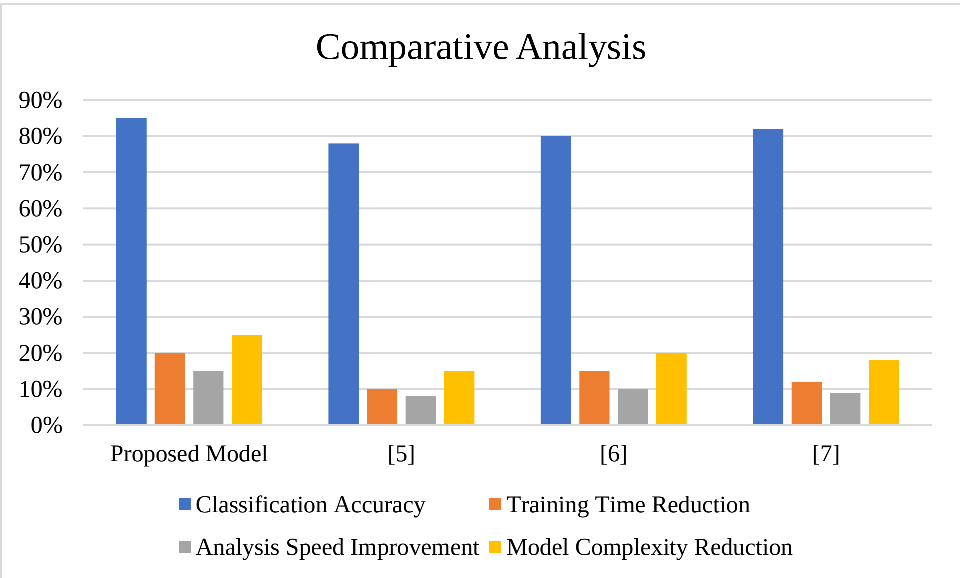

Figure 5: Comparative Analysis of the proposed model with existing studies [5-7]

Figure 5 compares the performance indices of the proposed research with three existing studies in remote sensing image classification. The proposed research demonstrates superior classification accuracy, outperforming [5], [6], and [7] by 7%, 5%, and 3%, respectively. Additionally, the proposed method achieves notable reductions in training time (20%), improvement in analysis speed (15%), and a significant model complexity reduction (25%) compared to the existing studies. These results suggest that the proposed approach excels in multiple key aspects, showcasing its potential for enhancing remote sensing image classification methodologies. The comparative analysis highlights the efficacy of the proposed research in addressing key challenges in remote sensing image classification. With superior performance across multiple indices, including accuracy, training time, analysis speed, and model complexity, the proposed approach stands as a promising advancement in the field. Future research may focus on further optimizing training speed and exploring trade-offs to enhance the overall efficiency of the proposed model.

4 Conclusion

The study introduces a neural network-based classification approach for remote sensing images, addressing the challenge of extracting valuable information from multiple resources. By incorporating multiple locations and combinations, this method effectively integrates spectral, texture, and contrast properties, enhancing the comprehensive analysis of remote sensing images. The integration of PCA transformations significantly reduces training time and enhances overall accuracy. The findings underscore the advantages of employing CNN in remote sensing image classification, demonstrating its capacity for improved categorization and accuracy compared to alternative methods. Future research avenues may explore further optimizations to address challenges associated with larger training models and enhance the efficiency of the CNN-style network model. When compared to SVM, NN, RF, DBN, and CNN (Patch), CNN is the best at finding the distribution of remote sensing patterns and picking out important details in the images. This is because it has advanced models that include local connectivity, weight distribution, and integration. Also, a close look at the author-made CNN (ST) and CNN (SST) models reveals that the suggested multiple locations and combinations greatly increase the amount of data that the CNN can use. This makes it more accurate at classifying and faster at analyzing. However, CNN, as a novel learning machine, presents challenges such as diversity and prolonged training times, despite the observed benefits in classification accuracy. The absence of theoretical support for the network model necessitates recommendations derived from iterative experiments. Future research endeavors will focus on refining the training speed of the model and identifying optimal trade-offs. Additionally, we will direct efforts toward addressing the inherent disadvantages of the network model, paving the way for advancements in the efficiency and theoretical foundations of remote sensing image classification systems.

Acknowledgement

Supported by the university-level platform of Digital Creative Design Production-Education Integration Collaborative Innovation Center.

References

[1] Qi, K., Yang, C., Hu, C., Shen, Y., Shen, S., & Wu, H. (2021). Rotation invariance regularization for remote sensing image scene classification with convolutional neural networks. Remote Sensing, 13(4), https://doi.org/569.10.3390/rs13040569

[2] Shi, C., Zhang, X., Sun, J., & Wang, L. (2022). Remote sensing scene image classification based on self-compensating convolution neural network. Remote Sensing, 14(7), https://doi.org/96-102.10.3390/rs14030545

[3] Byju, A. P., Sumbul, G., Demir, B., & Bruzzone, L. (2021). Remote-sensing image scene classification with deep neural networks in jpeg 2000 compressed domain. IEEE,69(4), 104.https://doi.org/10.30534/ijeter/2020/120872020

[4] Deepan, P., & Sudha, L. R. (2020). Remote sensing image scene classification using dilated convolutional neural networks. International Journal, 8(7). https://doi.org/

[5] Pires de Lima, R., & Marfurt, K. (2019). Convolutional neural network for remote-sensing scene classification: Transfer learning analysis. Remote Sensing, 12(1), 86.https://doi.org/10.3390/s20071999

[6] Yu, D., Xu, Q., Guo, H., Zhao, C., Lin, Y., & Li, D. (2020). An efficient and lightweight convolutional neural network for remote sensing image scene classification. Sensors, 20(7), 1999.https://doi.org/

[7] Wang, D., Zhang, C., & Han, M. (2022). Mlfc-net: a multi-level feature combination attention model for remote sensing scene classification. Computers & Geosciences, 160(33), 105042.https://doi.org/10.1016/j.cageo.2022.105042

[8] Qi, K., Yang, C., Hu, C., Zhai, H., Guan, Q., & Shen, S. (2021). A multi-level improved circle pooling for scene classification of high-resolution remote sensing imagery. Neurocomputing, 462 (10), 506-522.https://doi.org/10.1016/j.neucom.2021.08.022

[9] Lin, T. Y. (2021). A novel convolutional neural network architecture of multispectral remote sensing images for automatic material classification. Signal Processing. Image Communication: A Publication of the the European Association for Signal Processing, 97(1).74-78. https://doi.org/10.1016/j.image.2021.116329

[10] Zhang, D., Zhou, Y., Zhao, J., & Zhou, Y. (2022). Co-evolution-based parameter learning for remote sensing scene classification. International Journal of

Wavelets, Multiresolution and Information Processing.854(2), 20.https://doi.org/10.1142/S0219691321500466

[11] Liang, L., & Wang, G. (2021). Efficient recurrent attention network for remote sensing scene classification. IET Image Processing, 15(10),859-862.https://doi.org/10.1049/ipr2.12139

[12] Moudjib, H. Y., Haibin, D., Zhang, B., & Ghaleb, M. S. A. (2021). Hsi-gcn: hyperspectral image classification algorithm based on gabor convolutional networks. World Journal of Engineering, ahead-of-print(ahead-of-print).87(7),65-69. https://doi.org/10.1108/WJE-09-2020-0460

[13] Shi, C., Zhao, X., & Wang, L. (2021). A multi-branch feature fusion strategy based on an attention mechanism for remote sensing image scene classification. Remote Sensing, 13(10), 1950.https://doi.org/10.3390/rs13101950

[14] Xie, H., Chen, Y., & Ghamisi, P. (2021). Remote sensing image scene classification via label augmentation and intra-class constraint. Remote Sensing, 13(13), 2566.https://doi.org/10.3390/rs13132566

[15] Li, M., Lei, L., Tang, Y., Sun, Y., & Kuang, G. (2021). An attention-guided multilayer feature aggregation network for remote sensing image scene classification. Remote Sensing, 13(16), 3113.https://doi.org/10.3390/rs13163113

[16] Shabbir, A., Ali, N., Ahmed, J., Zafar, B., Rasheed, A., & Sajid, M. (2021). Satellite and scene image classification based on transfer learning and fine tuning of resnet50. Hindawi Limited,58(7), 589-597.https://doi.org/10.1155/2021/5843816

[17] Wu, X., Zhang, Z., Zhang, W., Yi, Y., & Xu, Q. (2021). A convolutional neural network based on grouping structure for scene classification. Remote Sensing, 13(2457), 1-22. https://doi.org/10.3390/rs13132457

[18] Wang, G., Xu, H., Wang, X., Yuan, L., & Wen, X. (2022). Remote sensing scene image classification model based on multi-scale features and attention mechanism. Journal of Applied Remote Sensing,471(7),15-16. https://doi.org/10.1117/1.jrs.16.044510

[19] Sun, M., Zhang, H., Huang, Z., & Li, Y. (2021). Remote sensing target detection in a harbor area based on an arbitrary-oriented convolutional neural network. Journal of Applied Remote Sensing,69(3), 15.https://doi.org/10.1117/1.JRS.15.034503

[20] Chen, X., & Zhu, J. (2022). Land scene classification for remote sensing images with an improved capsule network. Journal of Applied Remote Sensing,3662(7),145-149.https://doi.org/

[21] Lin, C. H., & Wang, T. Y. (2021). A novel convolutional neural network architecture of multispectral remote sensing images for automatic material classification. Signal Processing Image Communication, 97(105), 116329.https://doi.org/10.1016/j.image.2021.116329

[22] Zhang, J., Zhang, W., Hu, Y., Chu, Q., & Liu, L. (2022). An improved sea ice classification algorithm with gaofen-3 dual-polarization sar data based on deep convolutional neural networks. Remote Sensing, 14(41),85-86. https://doi.org/10.3390/rs14040906

[23] Bi, Q., Zhang, H., & Qin, K. (2021). Multi-scale stacking attention pooling for remote sensing scene classification. Neurocomputing, 436(12), 874.https://doi.org/10.1016/j.neucom.2021.01.038

[24] Guo, Y., Ji, J., Shi, D., Ye, Q., & Xie, H. (2021). Multi-view feature learning for vhr remote sensing image classification. Multimedia tools and applications,968(15), 80.https://doi.org/10.1007/s11042-020-08713-z

[25] Ma, A., Wan, Y., Zhong, Y., Wang, J., & Zhang, L. (2021). Scenenet: remote sensing scene classification deep learning network using multi-objective neural evolution architecture search. ISPRS Journal of Photogrammetry and Remote Sensing, 172(33), 171-188.https://doi.org/10.1016/j.isprsjprs.2020.11.025