International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Remote Sensing Scene Classification using Deep Learning

P. Bhargavi [a], S. Raju [b], D. Mohitha [c], P. Vijay Prakash [d], P. Sai Ram [e]

[a] *Department of Computer Science and Engineering, GMR Institute of Technology, Rajam, 532127, India. pedadabhargavi98@gmail.com*
[b] *Department of Computer Science and Engineering, GMR Institute of Technology, Rajam, 532127, India. sambangiraju752@gmail.com*
[c] *Department of Computer Science and Engineering, GMR Institute of Technology, Rajam, 532127, India. mohithadasari@gmail.com*
[d] *Department of Computer Science and Engineering, GMR Institute of Technology, Rajam, 532127, India. vijayprakashpenugonda@gmail.com*
[e] *Department of Computer Science and Engineering, GMR Institute of Technology, Rajam, 532127, India. Sairampatti43@gmail.com*

ABSTRACT:

In many applications, including land cover mapping, urban planning, environmental monitoring, and emergency management, remote sensing scene classification is crucial. Traditional scene classification techniques frequently rely on custom features and subject-matter expertise, which may not be able to adequately represent the complex and high-dimensional nature of remote sensing data. Deep learning approaches have recently proven to be incredibly effective in a variety of computer vision applications, including image classification. In-depth research on the use of deep learning for remote sensing scene classification is presented in this publication. Convolution neural network (CNN), VGG16, and AlexNet are three deep learning methods that are used to identify remote sensing images. Here the algorithms are compared based on accuracy and other metrics.

KEYWORDS: Remote sensing, Scene Classification, Deep learning, CNN, VGG16.

INTRODUCTION:

Deep learning-based remote sensing scene classification has become a ground-breaking method for geospatial analysis. It is crucial to be able to automatically and reliably classify these images in an era when enormous volumes of Earth observation data are being gathered by satellites, drones, and other sensors. This method has been revolutionized by deep learning, a subset of artificial intelligence, which enables computers to learn complex patterns and features straight from the data. As a result, they are now able to discriminate between various landscapes, urban regions, agricultural fields, water bodies, and more. Deep learning methods like CNN, VGG16, and AlexNet were used to overcome this issue and produce an accurate outcome. CNN did well and outperformed the other techniques in terms of evaluation metrics.

LITERATURE SURVEY:

The work of the current techniques for remote sensing image scene classification is presented in the study. The disadvantages of conventional approaches and the benefits of deep learning-based approaches, in particular CNNs, are discussed by the authors. They also draw attention to the difficulties in interpreting the learnt features and the requirements for a lot of training data when utilizing CNNs for the categorization of remote sensing images. The idea of utilizing MLPs as a deep classifier is then put out by the authors as a means of overcoming these difficulties and enhancing classification performance. Finally, they demonstrate the superior performance of the new CNN-MLP model by comparing it to modern techniques on three open remote sensing datasets. [1].

The paper proposes an Improved Deep Recursive Residual Network (IDRRN) model for single-frame super-resolution of remote sensing images. The IDRRN model uses a deep recursive structure and short-path recursive connections to control the model parameter number, alleviate gradient disappearance, and enhance feature propagation. The model is designed to reduce the difficulty of network training and improve the details of rebuilt images. The paper evaluates the performance of the IDRRN model using peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS) as reference evaluation indicators for image quality. The results show that the IDRRN model outperforms other state-of-the-art models in both quantitative and visual perception evaluations. The paper also suggests several future works to extend the applicability and effectiveness of the proposed IDRRN model in remote sensing and other related fields. [2].

The paper provides a literature review of different approaches for remote sensing image classification. The approaches are categorized into three groups based on the type and usage of visual clues: low-level visual features, mid-level features, and high-level feature extraction approaches. The paper hand-picks recent state-of-the-art approaches from these categories that have reported results on similar image benchmarks[3].

The paper proposes a novel remote sensing scene classification method based on high-order graph convolutional network (H-GCN) that uses the attention mechanism to focus on the key components inside the image during CNN feature learning. The proposed method investigates the class dependencies using the graph structure built where each node is described by the mean of attentive CNN features from each semantic class. The semantic class dependencies are propagated with mixing neighbour information of nodes at different orders, and thus the more informative representation of nodes can be gained. The node representations of H-GCN and attention CNN features are finally integrated as the discriminative feature representation for scene classification. Comprehensive experiments were conducted to evaluate the proposed method and compare it with related methods on four public remote sensing image datasets (UCM, RSSCN7, AID, and NWPU-RESISC45), and the results demonstrate the feasibility and effectiveness of the proposed method for remote sensing scene classification. As future work, the paper suggests introducing the attention mechanism into H-GCN for automatic neighbor selection at different orders to improve the classification capability. [4].

The paper proposes a system for efficient classification of high-resolution remote sensing images by extracting features using a deep convolutional neural network. The proposed system combines deep features with other features like Gabor features and novel reformed local binary pattern features to improve the efficiency of the extracted features. The system introduces two novel ideas in its feature extraction implementation, namely initialisation of filter values for the CNN and change in local binary pattern feature extraction process. The experimental results show that the proposed system produces good results when compared with other existing methods. The paper suggests that potential future works could include further improving the efficiency and accuracy of the system by exploring different combinations of features and experimenting with different datasets. [5].

The paper proposes a systematic analysis of the threat of adversarial examples on deep neural networks for remote sensing scene classification, and proposes an adversarial training strategy to increase the resistibility of deep models towards adversarial examples. The experiments were conducted on three benchmark high-resolution remote sensing image datasets, and the results may not be generalizable to other types of remote sensing data such as hyperspectral image and LiDAR data. The experimental results demonstrate that the adversarial training is a simple but effective way to increase the resistibility of deep models toward adversarial examples for the remote sensing scene classification task. The paper provides insights into the generation of adversarial examples and their impact on the accuracy of deep neural networks for remote sensing scene classification[6].

The paper proposes a novel feature representation method for scene classification, named bag of convolutional features (BoCF), which generates visual words from deep convolutional features using off-the-shelf convolutional neural networks. The paper aims to demonstrate the effectiveness of the proposed BoCF method through comprehensive evaluations on a publicly available remote sensing image scene classification benchmark and comparison with state-of-the-art methods. The visual words generated by BoCF have more semantic properties as they are obtained from deep convolutional features using off-the-shelf CNN models. Even compared to deep CNN features, the proposed BoCF methods with GoogleNet and VGGNet-16 still obtained better accuracies, which show the effectiveness of the proposed method. The best result of the proposed BoCF method almost doubles the accuracy of BoVW (82.65%)[7].

The paper proposes a classification method based on multi-structure deep features fusion (MSDFF) to address this limitation. The proposed method uses data augmentation based on random-scale cropping to expand limited data, three popular CNNs as feature extractors to capture deep features from the image, and a deep feature fusion network. The proposed method uses data augmentation based on random-scale cropping to expand limited data, which can improve the classification accuracy. The main contributions of this study are the introduction of three types of CNNs, CaffeNet, VGG-VD16, and GoogLeNet, to extract complementary features, and the proposal of an improved network including a concatenation layer and four fully connected layers with a softmax to integrate different features adaptively[8].

The paper proposes a few-shot classification method for remote sensing scene data using a novel attention module called channel attention and spatial attention module (CSAM) to fuse multi-scale feature representations from classification targets in different sizes. The paper mentions that the existing methods for few-shot scene classification have limitations in terms of using few annotated data, which results in poor representations. The proposed method in the paper, called DLA-MatchNet, uses a novel attention module called channel attention and spatial attention module (CSAM) to fuse multi-scale feature representations from classification targets in different sizes, which can further improve the model's performance. The paper presents the experimental results of nine few-shot learning methods and SCL-MLNet on three remote sensing scene datasets. All methods are trained based on meta-learning[9].

The paper proposes a novel deep-learning-based feature selection method for remote sensing scene classification, which can significantly boost the final performance of the classification. The method aims to select the most discriminative features by formulating the feature selection problem as a feature reconstruction problem. The proposed method is evaluated on 2800 remote sensing scene images of seven categories to demonstrate its effectiveness. The proposed method selects features that are more constructible as the discriminative features, which helps to improve the final performance of the scene classification. The experimental results demonstrate the effectiveness of the proposed method in improving the performance of remote sensing scene classification [10].

Methodology:

1. Convolutional Neural Networks:

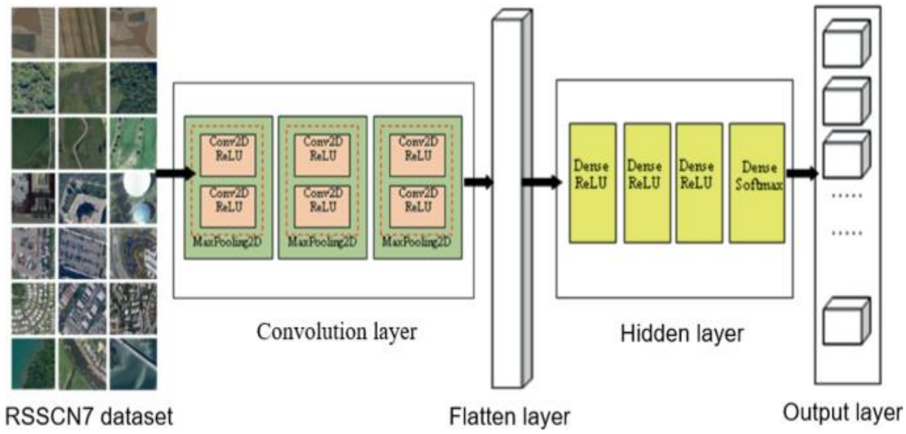

Figure (1): CNN on RSSCN7 dataset

The ability of Convolutional Neural Networks (CNNs) to extract spatial patterns from satellite and aerial data makes them essential for remote sensing. They enable automatic feature learning, assisting with tasks like classifying land cover, detecting objects, and identifying changes. CNNs use pooling layers for dimensionality reduction and convolutional layers to detect hierarchical features. Large datasets and transfer learning using pre-trained CNNs have sped up development in remote sensing applications, which is advantageous for areas like environmental monitoring, disaster assessment, and urban planning.

2 VGG16:

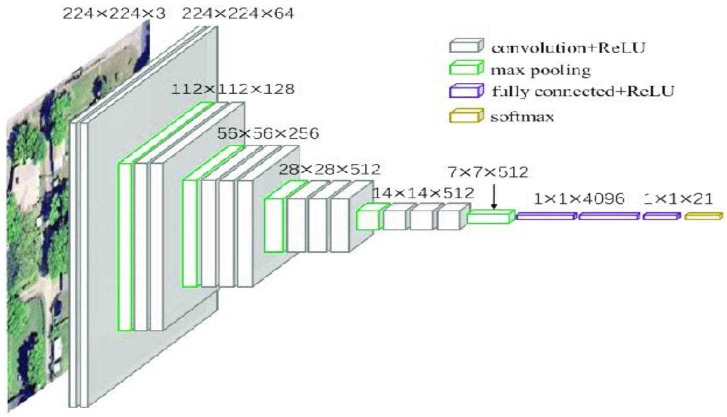

Figure (2): VGG16 on remote image

In order to extract features from high-resolution imagery, the well-known deep CNN architecture VGG16 uses its depth and standardized structure. It is excellent at identifying complex spatial patterns in land cover thanks to its 16 weight layers, which helps with tasks like classification and segmentation. Transfer learning benefits for remote sensing applications are provided by VGG16's pre-trained models on large datasets, increasing effectiveness and precision. Although advantageous, its complexity necessitates a significant investment in processing power, which affects the deployment option based on the size of the work and the infrastructure that is available.

3. AlexNet:

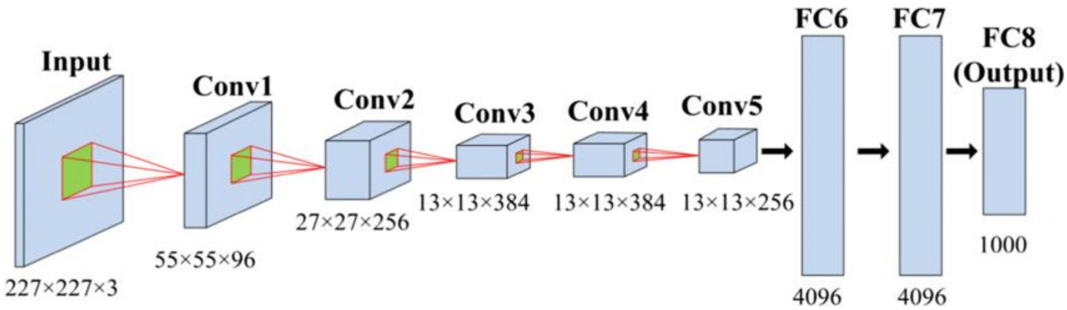

Figure (3): AlexNet

AlexNet, the world's first deep CNN, adds value to remote sensing applications by extracting significant information from airborne and satellite imagery. Its cutting-edge design, with multiple convolutions and aggregation, makes it easy to classify land cover and detect objects. The pre-trained models, trained from a variety of datasets, provide transfer learning capabilities that help speed up remote sensing application development. AlexNet's introduction marked a game-changer in image analysis for applications such as environmental monitoring and agriculture, as well as disaster management. Nevertheless, its depth may require optimization for effective deployment on devices or platforms with limited resources.

RESULTS:

Remote sensing images can be classified using various deep learning algorithms such as Convolutional neural network (CNN), VGG16, and AlexNet. Among the above mentioned algorithms, CNN performs well at 91% accuracy on the benchmark datasets, while Alexnet and VGG16 performed at 89% and 90%, respectively.

DISCUSSION:

Among the proposed papers, Gladima,N.T., and Rajesh,S. discussed how the CNN algorithm performed well on the proposed dataset in comparison to the proposed deep learning algorithms.

CONCLUSION:

There is a need for remote sensing image classification for the analysis of earth features. As mentioned in the previous post on remote images classification, the comparison of the other remote images classification algorithms shows that Convolutional Neural Networks (CNN) outperformed with an accuracy higher than the other remote image classification algorithms. CNN achieved a 91.02% accuracy. This CNN technique performed better in all the applied datasets. This technique performed better in the challenge dataset as well. There is scope for future work on the remote sensing scene classifications in the environmental control area.

References:

[1] Shawky, O. A., Hagag, A., El-Dahshan, E. S. A., & Ismail, M. A. (2020). Remote sensing image scene classification using CNN-MLP with data augmentation. *Optik, 221*, 165356.

[2] Tang, J., Zhang, J., Chen, D., Al-Nabhan, N., & Huang, C. (2021). Single-frame super-resolution for remote sensing images based on improved deep recursive residual network. *EURASIP Journal on Image and Video Processing*, *2021*(1), 1-19.

[3] Shabbir, A., Ali, N., Ahmed, J., Zafar, B., Rasheed, A., Sajid, M., ... & Dar, S. H. (2021). Satellite and scene image classification based on transfer learning and fine tuning of ResNet50. *Mathematical Problems in Engineering*, *2021*, 1-18.

[4] Gao, Y., Shi, J., Li, J., & Wang, R. (2021). Remote sensing scene classification based on high-order graph convolutional network. *European Journal of Remote Sensing*, *54*(sup1), 141-155.

[5] Gladima, N. T., & Rajesh, S. (2022). Ensemble of features for efficient classification of high-resolution remote sensing image. *European Journal of Remote Sensing*, *55*(1), 326-337.

[6]Xu, Y., Du, B., & Zhang, L. (2020). Assessing the threat of adversarial examples on deep neural networks for remote sensing scene classification: Attacks and defenses. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(2), 1604-1617.

[7]Cheng, G., Li, Z., Yao, X., Guo, L., & Wei, Z. (2017). Remote sensing image scene classification using bag of convolutional features. *IEEE Geoscience and Remote Sensing Letters*, *14*(10), 1735-1739.

[8] Xue, W., Dai, X., & Liu, L. (2020). Remote sensing scene classification based on multi-structure deep features fusion. *IEEE Access*, *8*, 28746-28755

[9] Li, X., Shi, D., Diao, X., & Xu, H. (2021). SCL-MLNet: Boosting few-shot remote sensing scene classification via self-supervised contrastive learning. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-12.

[10] Zou, Q., Ni, L., Zhang, T., & Wang, Q. (2015). Deep learning based feature selection for remote sensing scene classification. IEEE Geoscience and remote sensing letters, 12(11), 2321-2325.