ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

CNN FOR SATELLITE IMAGE ANALYSIS

Gautam Kumar

Assistant Professor, Dept. Of CSE, Chandigarh University, Mohali Punjab, India

Piyush Kumar

Dept. of CSE, Chandigarh University, Mohali Punjab, India

Lovish Bansal

Dept. of CSE, Chandigarh University Mohali Punjab, India

Shivam Aanand

Dept. of CSE, Chandigarh University Mohali, India

ABSTRACT—

The field of multimedia applications and processing is very promising and plays an essential role in many type of artificial intelligence applications like in video summarization, image retrieval, image classification, etc. Multimedia methods using Convolutional neural networks, convinced researchers to build systems that can classify images using data without human input. For feature extraction based on deep learning with convolutional neural networks using pre-trained models such as Alexnet, we renew the measures for satellite image classification. On all three datasets, SAT4, SAT6, and UC, the Resnet50 model outperforms the other models. Merced Land. As depicted in Figure 5, after applying the developed model, the accuracy of UC Merced Land dataset achieves 98%, SAT4 achieves 95.8%, and SAT6 achieves 94.1%.

Keywords—Deep Learning method, Image Compression, Autoencoders, (C_N_N), (G_A_N), (P_S_N_R).

1. INTRODUCTION

Development of remote sensing technologies has exploded over the last few years. This is, obtaining large datasets of high-resolution remote sensing images has become much easier. Many researchers of remote sensing recognition and classifications have been transitioning from traditional methods to recent techniques relying on this notion. The classical approaches are based on the pixel level information contents and the contemporary methods camp on the semantic understanding of the scene. The semantic understanding strives to divide the information into several sets of semantic groupings and a set of classes determined by remote detecting picture content [1–3]. Generally, image classification can be categorized into three broad categories based on its characteristics [1]. The ‘handcrafted feature-based method’ focusing on color-based as well as shape information is a possible property of sense images [2–4], whereas ‘unsupervised feature learning-based methods work to learn a basic function set such as a bag of words model that is used for features encoding. Quantization is the most common encoding method, and a more sophisticated approach is fisher encoding, where the input of the Fisher method is a collection of handcrafted features and the output is a collection of learned characteristics [5–7]. Lastly, the ‘ deep feature learning-based methods referred to Deep Learning (DL) [8–10]. Deep learning of remote sensing image features in recent years demonstrated an incredible power in classification through characteristic selection of the specific problems that arise in remote sensing image classification. The choice of new deep learning is a part of device learning based on several levels of learning. The architecture of deep learning extends NN by adding more hidden layers. Deep learning has many architectures, one of which is a Convolutional Neural Network (CNN). The CNN is popular and has been used in recent years for addressing different and complex issues like image classification and recognition just using a sequence of feed-forward layers. The CNN is analogous to conventional neural networks and it is composed of neurons that contain adjustable weights and biases. The input is processed by the neurons through a set of inputs followed by some non-linear processing and it can be seen as multi-layer feed-forward artificial neural network [11]. These images serve as inputs to different convolutional network architectures that enable the encoding of specific properties into the architecture. CNN usually consists of a part with several layers that include a convolutional layer, a pooling layer, and full connection layers [1]. It is a special case of the neural network containing one or more convolutional layers responsible for extracting low-level features like edges, corners, and lines; pooling/subsampling layers that protect the features against distortion and noise non-linear layers that work as a trigger function to signal differing identification of likely features on each hidden layer; and fully connected layers that mathematically sum weighting of the previous features layer, and full connection layers [12]. In this paper we proposed four approaches and architecture of CNN to enhance the performance of satellite image classification, four approaches of CNN (AlexNet, VGG19, GoogLeNet, and Resnet50) are used as are-trained for feature extraction, each of them trained on image dataset. We then assess our methods by concatenating the previous features with deeper features in a dense layer and compare our entire model outputs with a few innovative approaches on three datasets: SAT 4, SAT6, and UCMD. This paper is structured as follows: Sect. In Sec. 2, we introduce the related works with CNN for image classification and recognition. Section 3 summarizes the datasets we employed in our system. Sect. 4. Finally, Sects. Sections 5 and 6 contain Sections5and6contain the results of the experiment sand the conclusions of this work, respectively. Satellite Image Classification with Convolutional Neural Network.

2. BACKGROUND AND LITERATURE REVIEW

The classification of the satellite image is a process to classify the images depending on the object or semantic meaning of the images, thus the classification can be decomposed into three major parts: methods based on low features, or other methods based on high scene features [13]. The first method of classification is based on low features and uses a simple type of texture feature or shape feature or shape features. one of the most used techniques for low features are local binary pattern or histogram–based features like [14], the author in that paper using texture with LBP as classification tool. The based methods are applicable for a co middle based appearances and structure types [5]. The methods based on high features are the most efficient methods for complex images compared to other. CNN — It is one of the most well-known and widely used in deep learning algorithms with image processing. Saikat Basu, Sangram Ganguly et [28] have proposed that the method is a learning framework for satellite imagery "DeepSat", they focus on classification based on deep unsupervised learning "Deep Belief Network for classification" with Convolutional Neural Networks and achieve accuracy result 97.946 for SAT4-dataset[25] and 93. 916 for SAT6-dataset [10]. Ju et al. The produced research paper is based on a widely used ensemble approach for image classification and recognition tasks using deep convolutional neural networks, mentioned above [15]. Some of these methods are majority voting, the Bayes Optimal Classifier, and super learner. Albert et al. [16] analyse patterns in land use in urban neighborhoods, using large-scale satellite imagery data and state-of-the-art computer vision techniques based on deep CNN. They use ground truth and: class labels are carefully sampled from open-source surveys, namely, the Urban Atlas land classification dataset (20 land use classes in 300 European cities). They also demonstrate that deep representations learned from satellite images of urban areas could be used to compare neighborhoods across multiple cities. Robinson et al. Proposed a deep learning convolutional neural networks model for creating high-resolution population estimations from satellite imagery [17]. The proposed CNN model has been trained to predict the population in the USA at 0.01 ×0.01resolution grid using 1-year composite Landsat imagery. The CNN model was assessed and evaluated in two ways: quantitatively and qualitatively. In quantitative validation, we compared the proposed model's grid cell estimates aggregated at the account level with a number of U.S. Census county-level population projections, as well as qualitatively by directly interpreting the model's predictions as functions of the satellite image inputs. Overall, the proposed model exemplifies how machine learning techniques can serve as a valuable tool for obtaining information from naturally unstructured, remotely sensed data to provide real-world solutions to social issues. Harry Pratt [16] have propose a Convolutional Neural networks-based algorithm that work for diabetic retinopathy detection for the fundus image. They create cnn architecture and augmentation data which able to detect the tiny detail that relates to the classification task like as micro-aneurysms, exudates, and haemorrhages in the retain, and ultimately provide the diagnosis without human input. They trained their proposed CNN model on the GPU with the Kaggle dataset and they showed some exciting results. In this paper, we will consider CNN as a classification method. Shamsolmoali et al. have the new classification pipeline based on a unified deep CNN and a modified residual network to integrate them with the other feed-forward network style in an endwise training fashion for a high-dimensional multimedia data analysis [19].

3. Datasets

We will be using three datasets SAT4, SAT 6, and UCMD in the proposed work. The first two "SAT4 and SAT6" types of images are extracted from the NAIP program, These data set a total of 330,000 scenes of all United States images. Each of the images is made up of 4 layers red, green, blue, and near-infrared (NIR). The 3rd data sets UC Merced Land Use Dataset includes "tif" file image format.

3.1 SAT 4

The dataset comprises 500000 image patches from four landforms, such as barren land, trees, grassland, and a class that contains all land cover classes. For the training set, we take 400,000 classes and the 100,000 remaining are used for a testing dataset. We normalize all images into 28 ×28 pixels [10].

3.2 SAT 6

This new version of the dataset has 405,000 images of 28 ×28 pixels each, distributed across six land classes - barren land, trees, grassland, roads, buildings, and water bodies. Choose 324,000 images as a training dataset and another 81,000 images as an a test1 dataset [10]. Also, Fig. 1 shows sample images from SAT 4 and SAT 6 datasets.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

3.3 UC Merced Land

This data set contains a total of classes of 21 land use image datasets each class includes 100image each image size is 256 × 256 pixels. The images were manually extracted from large dataset images for the USGS National Map Urban Area Imagery collection. Selected samples of the 20 class images from 20[20]are shown in Figure 2.

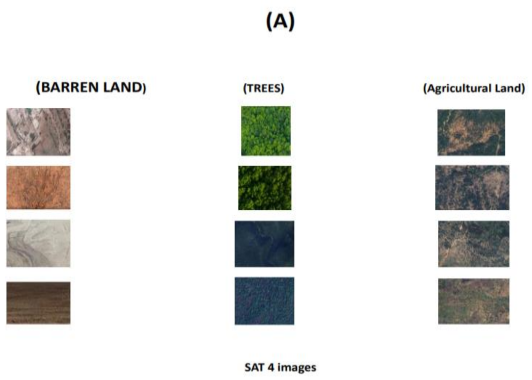

Fig.1 SAT 4 image samples

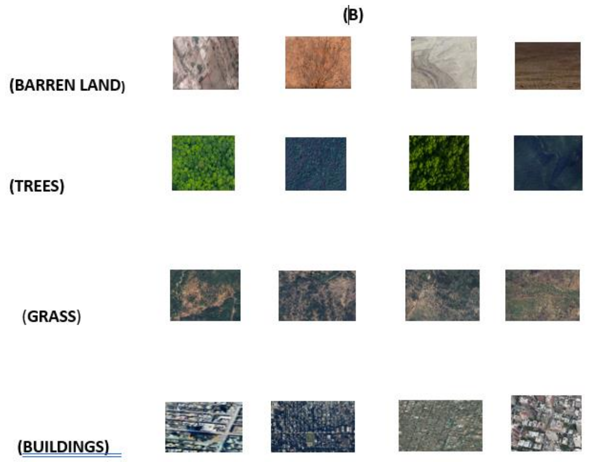

Fig.2 SAT 5 image samples

4. Proposed work

The architecture of the proposed work was designed by reviewing literature work on satellite image classification as in Fig. 3which provides an overview of the proposed model of image classification of the satellite imagery based on CNN. The work has two stages. The first stage is the Training Stage, and the second stage is the Testing Stage. The datasets are being split into two initials. The first is the training image and the second is the testing of our models. The SAT dataset includes SAT4 and SAT6, which respectively contain 400,000 and 324,000 images selected sequentially as a training set and 100,000 and 81,000 images selected as a testing set. All the datasets UC Merced Land Use which are 21 classes each one has 100 images, of these we have selected 70 images as training sets and 30 images as testing set for all the classes. Furthermore, due to the model adopted and evaluated on two separate datasets, the preprocessing step is an important step to make the input image the same characteristics.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

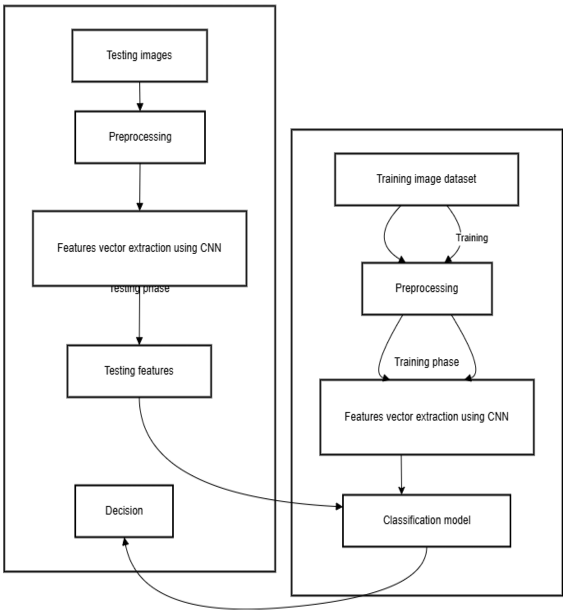

Fig.3 The diagram of the proposed system

4.1 Training Phase

Stage one in our model is the training stage. In this section, the trained images selected from both of the datasets are processed in steps starting from feature computation, followed by image pre-processing through CNN.

Preprocessing. The datasets used for our model are different; the colour images of SAT airborne datasets consist of four bands 28 × 28 uint8, and the other dataset, UCMD 256*256 uint8, has three bands for red, green, and blue. So, to create a classification model, we must preprocess all images by color normalization, remove the invisible band NIR of the SAT datasets, and change the images into grayscale format. After that all satellite images are ready for the next stage to extract the feature vector that represents each image in the training set.

Features extraction based on CNN. The efficiency of satellite image classification depends on the strength of the features extracted from the training data. The power of those features will be reflected in the testing phase. So we propose off-the a Convolutional Neural Network for Satellite Image Classification. Thus, we are offering the high-level features to be training the CNN from Satellite for Images Classifying proposed off-the-shelf features extraction from the Convolutional Neural Network the set of training data. Here, we utilize a pre-trained network with all networks pre-trained on the ImageNet dataset available at the link http://www.image-net.org which has a thousand object categories. The characteristics of each of the one used and the fully connected layer that we have considered as a features vector. and can be seen in Table 1.

	AlexNet	VGGNet-19	GoogleNet	Resnet50
Input data	Input image 227 x 227 x 3	Input image 224 x 224 x 3	Input image 224 x 224 x 3	Input image 224 x 224 x 3
Layers	25 x 1 nnet.cnn.layer.Layer	47 x 1 nnet.cnn.lay er.Layer	144 x 1 nnet.cnn.layer.Layer	177 x 1 nnet.cnn.layer.Layer
Features layer	fc8	fc8	loss3-classifier	fc1000

Table.1 Pretrained network, layers and features layers

4.2 Testing phase

The second stage of the satellite picture classification model is the testing stage. In the test, the 30% left of every dataset will be tried to test and quantify the precision of the classifier strategy. This will happen for the testing images and with the preparation the input data for the training phase it will also happen for the testing images initiate the preprocessing image from all categories in the datasets and continue extracting features and save it as two-dimensional matrices each row belongs to the one image. we will publish the outcomes of the satellite image classification based on CNN.

Table.2 Configuration of the pre-trained models

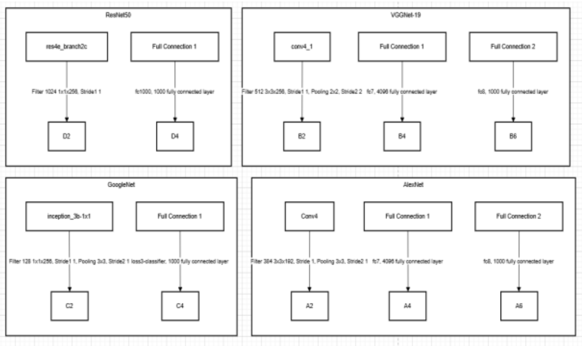

Table.3 Configuration of the pretrained models

Dataset	Subset of images		Images resolution	
	Training	Testing	Width	Height
SAT 4	400,000	100,000	28	28
SAT 6	324,000	81,000	28	28
UC Merced Land	21 class × 70	21 class × 30	256	256

Table. 3 Dataset setting for experimental results

5. Experimental Outcomes

We then directly write down the sentence of the datasets which discuss in the datasets segment above. This division covers the experimental re-sults that are generated by combining deep features with the previous features of CNN by implementing four different networks, namely AlexNet, VGGNet-19, GoogleNet, andResnet50 pretrained on the imageNet dataset. Depending on the model type and in the full connection layers appearing in Table1, the attributes are derived from various layers. Because we used different datasets and different dimensions, we kept the size of an image and normalized the four bands to only visible layers: red, green and blue. The features layer are chosen from the last pooling full connection layer on four models: layer number 23 "fc8" for AlexNet, layer 45 for VGGNet-19 "fc8", layer number 142 for GoogLeNet "loss3-classifier" and layer number 175 for Resnet50 is "fc1000". Table 2shows the configuration of the four models on UCMD dataset. This work is an evaluation of four pretrained CNN using the configuration as shownin Table 2 on three different datasets SAT 4, SAT 6 and UC Merced Land. Eachdataset has been randomly divided into two part, a training set of images and a testing subset of images,Table 3presents the dataset configuration in our experimental results. The proposed method based on the deep and shallow features combining usingResnet50 surpass with the earlier works the result thatdetermine are extract from "fc1000" layer that are better than the feature extracted from first convolution or deep convolution, also it get the better performance than any pre-trained CNN like AlexNet, VGG-19 and GoogleNet since they use only 70% percentage of training as configured in Table 2. Also, Fig. 4 shows that the accuracy and achievement of the Resnet50 model have a higher result than the other models, also it supports the finding of Fig. As shown for example, fig. 5 shows the training loss on the same model on UC Merced Land Datasets with 250 epochs. Fig. 6 Presents the comparison among the models used for feature extraction, it's visible that theResnet50 model used for feature extraction has a better result of classification than other models and the loss function is less than others. So Table 4 shows the accuracy of all datasets that were used with different models and algorithms. have seems in Table 4 above that the Research Paper [10] achieved a classification ratio on SAT4 and SAT6 of 97.946 and 93. respectively. 916 respectively. They proposed a way of extracting data and features of an input image, and based on this features normalization as a vector in Deep Belief Network for classification. They have shown two datasets SAT4 and SAT6 and the Proposed work has been not evaluated on UC Merced Land. The first paper is listed in Table 4[21] same as it is investigated in our experiments the architecture is agile CNN architecture called Sat CNN for HSR-RS image scene classification. They are built with less kernel with the modern CNN architectures that are also based on the recent improvements of CNNs. The achievement ratio on SAT4 and SAT6 was 99.65 and 99.54 respectively Convolutional Neural Network for Satellite Image Classification. yet it doesn't get tested on UC Merced Land. The accuracy of the third research paper [22] is 99.72, 99.65, and 97.99 on SAT4, SAT6, and UC Merced Land datasets respectively. The researchers propose a new classification method based on triple networks. Further analysis on the performance of proposed method based feature extraction depends on Resnet50 best model achievement for classifying image set of UC Merced

Land dataset in our experiment results. The performance of our proposed model (Resent50) is better than results yielded from research paper [10] for SAT6dataset and it is the worst for SAT4 dataset.

Algorithms	Classifier accuracy % SAT4	Classifier accuracy % SAT6	Classifier accuracy % UC Merced Land
DeepSat [10]	97.946	93.916	–
Agile CNN SatCNN [21]	99.65	99.54	–
Triplet networks [22]	99.72	99.65	97.99
Features extraction based on AlexNet	84	82	87
Features extraction based on VGG19	89	84	89
Features extraction based on GoogleNet	91	89	90
Features extraction based on Resnet50 (proposed method)	95.8	94.1	98

Table 4: Accuracy Comparison of Different Models and Algorithms on SAT4, SAT6, and UC Merced Land Datasets

6. Conclusion

This paper presentsbeneciary models for satellite image classification basedon convolutional neural network, the classifying-features extracted usingfour pretrained CNN models: AlexNet, VGG19, GoogleNet andResnet50 and compares the result among them. The features are extracted from a combination layer or full connection layer of earlier and deep layers. The experiment results of the datasets and the pre-trained models. The Resnet50 model achieves a better result than other models for all datasets used 'SAT4, SAT6, and UC Merced Land'. The feature extraction using Resnet50-based classification has better accuracy and minimum loss value than other approaches and works in different data sets. The performance of the Resnet50-based method we proposed is superior to that of the approach in the above paper [10] for image classifying of SAT6, it is also a better than research paper [22] for classify UC Merced Land data set.

REFERENCES

1. Cheng, G., Li, Z., Yao, X., Guo, L., & Wei, V. (2017). Scene classification of remote sensing images using a combination of convolutional features. IEEE Geoscience and Remote Sensing Letters, 14(10).
2. Bian, X., Chen, C., Tian, L., & Du, Q. (2017). High-resolution scene classification through the fusion of local and global features. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10(6), 2889–2901.
3. Huang, L., Chen, C., Li, W., & Du, Q. (2016). A multi-scale approach for remote sensing image classification utilizing completed local binary patterns and Fisher vectors. Remote Sensing, 8(10).
4. Chen, C., Zhang, B., Su, H., Li, W., & Wang, L. (2016). Land-use classification based on multi-scale completed local binary patterns. Signal, Image and Video Processing, 10(4), 745–752.
5. Zhang, F., Du, B., & Zhang, L. (2015). Unsupervised feature learning for scene classification driven by saliency guidance. IEEE Transactions on Geoscience and Remote Sensing, 53(4), 2175–2184.
6. Li, Y., Tao, C., Tan, Y., Shang, K., & Tian, J. (2016). Automated satellite image classification using a multi-layer unsupervised feature learning approach. IEEE Geoscience and Remote Sensing Letters, 13(2), 157–161.
7. Yuan, Y., Wan, J., & Wang, Q. (2016). Classification of congested scenes through efficient unsupervised feature extraction and density estimation. Pattern Recognition, 56, 159–169.
8. Yao, X., Han, J., Cheng, G., Qian, X., & Guo, L. (2016). High-resolution satellite image annotation using weakly supervised learning. IEEE Transactions on Geoscience and Remote Sensing, 54(6), 3660–3671.
9. Zou, Q., Ni, L., Zhang, T., & Wang, Q. (2015). Remote sensing scene classification enhanced by deep learning-based feature selection. IEEE Geoscience and Remote Sensing Letters, 12(11), 2321–2325.
10. Basu, S., Ganguly, S., Mukhopadhyay, S., DiBiano, R., Karki, M., & Nemani, R. (2015). DeepSat: A framework for satellite image classification using deep learning. ACM SIGSPATIAL'15, Bellevue, WA, USA.
11. Yu, X., Wu, X., Luo, C., & Ren, P. (2017). A convolutional neural network framework with data augmentation for remote sensing scene classification. GIScience & Remote Sensing.

12. Hijazi, S., Kumar, R., & Rowen, C. (Year unknown). Image recognition utilizing convolutional neural networks. IP Group, Cadence.
13. Yu, Y., & Liu, F. (2018). Two-stream deep feature fusion with dense connectivity for aerial scene classification. Remote Sensing.
14. Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Local binary patterns for texture classification with multi-resolution gray-scale and rotation invariance. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24, 971–987.
15. Ju, C., Bibaut, A., & van der Laan, M.J. (2017). Comparative analysis of ensemble methods in convolutional neural networks for image classification. ArXiv e-prints.
16. Albert, A., Kaur, J., & Gonzalez, M. (2017). Identifying urban patterns using convolutional networks and satellite imagery at scale. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1357–1366).
17. Robinson, C., Hohman, F., & Dilkina, B. (2017). Population estimation from satellite imagery using deep learning. In Proceedings of the 1st ACM SIGSPATIAL Workshop on Geospatial Humanities (pp. 47–54).
18. Pratt, H., Coenen, F., Broadbent, D.M., Harding, S.P., & Zheng, Y. (2016). Diabetic retinopathy detection with convolutional neural networks. In Medical Imaging Understanding and Analysis Conference, Loughborough, UK.
19. Shamsolmoali, P., Jain, D.K., Zareapoor, M., Yan, J., & Alam, M.A. (2018). Deep CNN with extended residual units for multimedia classification. Multimedia Tools and Applications. https://doi.org/10.1007/s11042-018-6146-7 .
20. Yang, Y., & Newsam, S. (2010). Enhancing land-use classification with bag-of-visual-words and spatial extensions. In ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS).
21. Zhong, Y., Fei, F., Liu, Y., Zhao, B., Jiao, H., & Zhang, L. (2017). Sat CNN: Agile convolutional neural networks for satellite image classification. Remote Sensing Letters, 8(2), 136–145.
22. Liu, Y., & Huang, C. (2018). Scene classification via triplet networks. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(1), 220–237.