A Lightweight CNN Architecture for Land Classification on Satellite Images

Prof. D.S. Hirolikar, Omkar Mankar, Parag Shelke, Akash Waman, Dhiraj Dixit, Tushar Shinde

Department of Information Technology, PDEA's College of Engineering, Pune, India

ABSTRACT

Land cover classification using satellite images is an important tool in the study of terrestrial resources. Satellite based information is presently available as huge sets of high resolution images from a large number of satellites like Sentinel, Landsat8, etc. Land cover classification from these images is a difficult task because of very large sized data and high variation types.

Deep Neural Networks can play a vital role in this regard and can perform classification on these large sized data. Related works in this field have used lighter models and included a large number of handcrafted parameters which requires domain knowledge on the subject. It is realised that most models are too shallow for such a complicated image. In this paper, a deeper Convolutional Neural Network (CNN) model without any satellite image specific parameters is proposed.

On SAT4 and SAT6 images, our 13- layered network has achieved better accuracy upto 99.84% and 99.47% which is state-of-the-art. It is still called lightweight model because most models in Artificial Intelligence(AI)-CNN are much deeper and larger than ours.

Key Words: SimpleNet, Satellite Imagery, Deep Neural Networks, SAT4, SAT6, Remote Sensing, Scene Classification. CNN

1.INTRODUCTION

Classification of large satellite imagery is a challenging task for understanding and portraying land cover information. Land cover is the physical land which includes trees, crop fields, barren lands, rivers, forests, etc. Information about land cover is an input for classifying, planning, monitoring and devising ways to use earth resources potentially in greater interest of the human race. This classification is important for various geospatial application like agriculture, environmental and urban management. Accurate and up-to-date information about land cover goes a long way in helping various government and other agencies to update their plans on regular basis. Traditional methods of gathering land cover information are field surveys that are time consuming as well as include much physical labour. The data collected is also out of date it is not possible to obtain information in short intervals time.

Remote sensing satellite imagery is a viable resource gathering effective land cover information due to their large view and repetitive coverage area. Raw images taken by the user satellites cannot be processed directly because they contain more bands with a large volume of data.

Because of the involvement of such huge data and higher variability of land cover classes, it is not easy to determine land cover types. Land cover types the most useful are:
1) Forests / trees
2) Pasture
3) Barrenlands
4) Water bodies

There are two approaches to classify land cover from them images that include:
• Supervised learning
• Unsupervised learning

Supervised learning involves training using data tags with the correct answer known as labelled data. This helps in predicting unforeseen data. Various techniques in satellite image processing can be classified as supervised learning.

The disadvantage of these techniques is that they cannot can easily scale to large amounts of data. Getting a label data in supervised learning for satellite data is therefore difficult unsupervised learning algorithms are gaining popularity. Unsupervised learning deals with unlabeled data and a model it is left alone to discover useful functions from the input. Deep Learning and Convolutional Neural Networks (CNN) are a class of unsupervised learning techniques that have showed promising results in land cover classification. Deep neural network has the ability to classify unlabeled data in a hierarchical manner. Recent advances show that deeper functions are learned from simple level to higher level patterns in an orderly manner. In recent years, deep learning has been used to satellite images have become a permanent topic of research. This work presents the CNN model, which is a modification of the Simplenet model for performing land cover classification on SAT4 and SAT6 datasets. Special features our proposed model is:

• Lightweight model with fewer parameters.
• Effective in terms of saving time.
• Requires less storage space.
• Achieves good accuracy in fewer epochs.

Objective:

1. To assist infrastructure development and urban expansion planning by designating areas suitable for roads, utilities and other public services.

2. To Improve agricultural planning and land use for farming, taking into account soil quality, climate and topography to optimize crop selection and land management practices.

3. To guide the growth of urban areas, ensuring efficient land use, minimizing urban sprawl and promoting sustainable urban development.

2. METHODOLOGY

The spatial resolution of earlier satellites was so low that most objects of interest were pixel in size. Therefore during early period of remote sensing, per-pixel analysis was standard. Thanks to the availability of high-resolution images objects of interest such as buildings, roads, etc. cover several pixels. A classic paper therefore questions the use of the pixel level classification in the field of high-resolution images and from here the journey of advanced scene classification begins. Some authors call it Object Based Image Analysis (OBIA). Because of their utility, these object techniques has dominated the field of remote sensing for the past two decades.

However, they show their limitations when it comes to classification pastures from forest land, residential buildings from other buildings and the like. Traditional techniques used in object classification in remote sensing images is based on Supervised learning. Popular among them are Support Vector Machines (SVM) and Random Forests (RF). But Unsupervised Learning helps learn features from large amounts of data and proved to be beneficial in the study of satellite imagery.

Between different learning models, Deep Belief Network (DBN) was the first to produce better than traditional results models. They are also credited with creating SAT4 and SAT6 datasets which are now almost universally used. They uses a combination of supervised and unsupervised learning algorithms. Various statistical and satellite specific features the input image is extracted and normalized before serving to the DBN model.

Specifically, an image was passed through a restricted Boltzman machine and a contrast Divergence algorithm. It was seen that with the addition more layers of the neural network, the accuracy decreased. Even The CNN they use is no more than 6 layers. The accuracy obtained on SAT4 reached 87% compared to 82% DBN. To further improve the accuracy, it was designed use the traditional features used previously in satellite imagery processing. Examples are statistical mean, various central moments, entropy and such other 50 descriptors.

They were evaluated according to the distribution separability criterion and the best 22 were selected. They were normalized to be within the range [0,1] and fed to DBN. Accuracy on SAT4 has improved to 98%, which was the state of the art. Later, Ma used the form A CNN derived from the initial GoogleNet module. It was they found they had a large set of hyperparameters to choose from. A genetic algorithm has been found to be useful in retrieving best hyperparameters.

They also used data augmentation increase the size of the data set. Model with optimized hyper-parameters achieved an accuracy of 98.4%. In this embodiment it should be noted the absence of features related to the satellite image. One of the most cited articles in this area introduces a new dataset named NWPU-RESISC45 with 45 land cover classes. A number of standard CNN models were and the resulting accuracies were in the 75-85 range %.

DATASETS

Our experimental data includes the SAT4 and SAT6 datasets. SAT4 has 500,000 images with four landscape types. SAT6 has 6 landscape classes with 405,000 images. Different landscapes include trees, bodies of water, agricultural fields, countryside areas, urban areas, etc. These 1m resolution images were taken from US database. Each image is 28×28 in size of pixels covering an area of approximately 28 × 28 m. Size images are chosen to be 28 m because they represent the size of a typical plot of land or urban building. These pictures have four channels namely red, green, blue and near infrared (NIR).

PROPOSED MODEL

The architecture proposed by us is almost similar to the original one SimpleNet architecture, which consists of 13 layers convolutional layer with 3 × 3 kernels except 11. a 12. layers that have a 1 × 1 core. The kernel is 1 × 1 not used for early layers as it avoids local information input, although it increases the nonlinearity of the model. 2 × 2 cores are used for pooling operations.

Our model differs from SimpleNet in the number of cores used in each of them layer as shown in Fig. Cores of the First Hidden layer computes vast feature information from the input. So the feature map of the first hidden layer is richer in information compared to the input satellite image which consists of R,G,B and NIR bands with pixel values ranging from (0, 0, 0, 0) to (255, 255, 255, 255). Since the given layer is powered directly from

fig. 1: Modified Simplenet Architecture

In order to create new kernels by combining the properties of the previous layer, the information richness increases per layer. To perfectly extract richer and richer information from the previous hidden layers, the number of cores used in each layer has been increased. By using the activation function in multiple layers, the loss function converges to zero, making the network impossible to train.

Batch normalization was used to overcome this problem. The experimental dataset contains compressed images and may contain a lot of noise. Therefore, the model learns noisy information along with details that negatively affect the performance of the model, and this is called over-fitting.

A Dropout technique with a probability of 0.2 was used to solve this overlap problem so that the upper layers can be more informative. The results show that our model outperforms other models in fewer epoch.

EXPECTED RESULTS

The experiment will be performed on an NVIDIA GPU-accelerated GPU machine with python 3.6 and the tensorflow framework. The test and training accuracy and model loss are calculated to determine the performance of the proposed architecture and the results have been compared with related works. FIG. 2 shows the variation of training and testing (classification) accuracy for each epoch for the SAT4 Dataset. It can be concluded that after training for 15 epoch numbers with a batch size of 50, our model achieves an accuracy of 99.84%, which is better than the previous results.

Fig. 3 shows that the loss gradually decreases as the number of epochs increases, but the small increase in the loss function at some points may be due to noise in the data set. FIG.4 shows the variation in testing and training accuracy for the SAT6 dataset. Our model is trained on 30 epochs

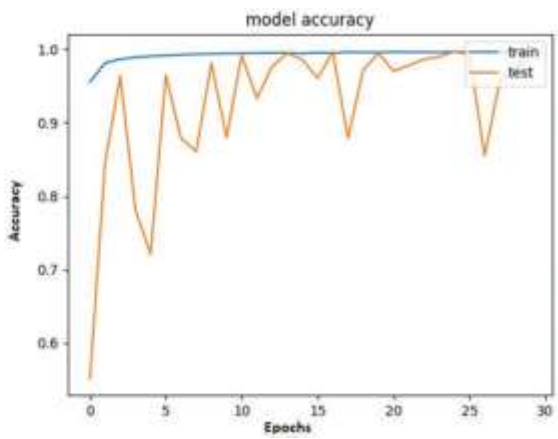

Fig. 2: Accuracy on SAT 4 dataset

batch size 50 and the highest accuracy achieved is 99.47%. It can be observed from Fig. 5 that as the number of epochs increases, the loss approaches zero. Table I shows the comparison of the training accuracy of our proposed SimpleNet model with related works. From the results, it can be concluded that our proposed model was able to achieve 99.84% and 99.47% accuracy on the SAT 4 and SAT 6 datasets in just 30 epochs.

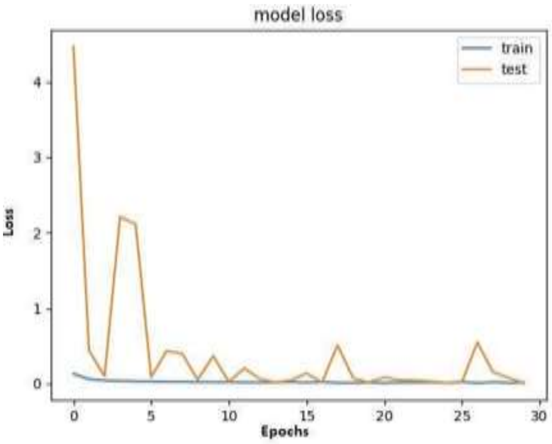

Fig. 3: Loss on SAT 4 dataset

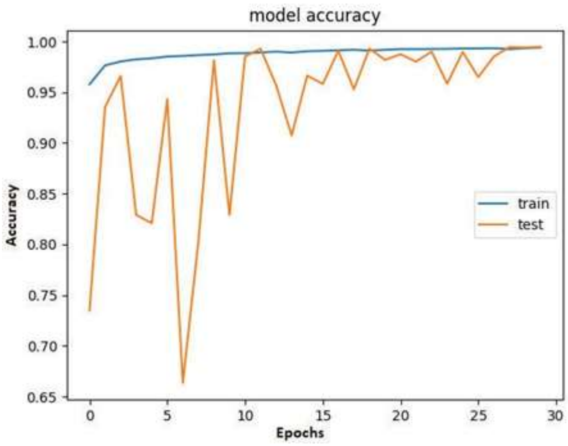

Fig. 4: Accuracy on SAT 6 dataset

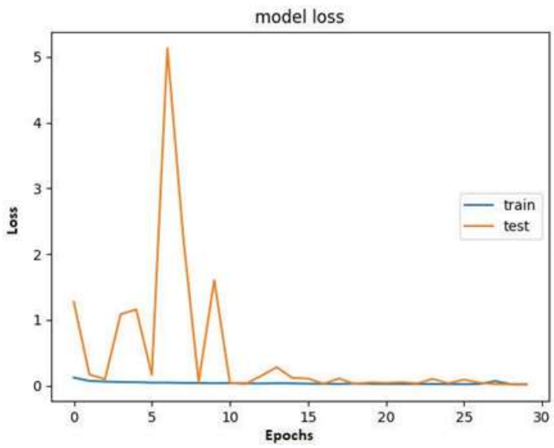

Fig. 5: Loss on SAT 6 dataset

3. CONCLUSION & FUTURE SCOPE

This paper proposes a CNN architecture for extracting scene information from satellite imagery. Most of the other models proposed in the literature use the lightweight CNN model and extend it with features specific to satellite imagery. Such domain knowledge is not generally available in the AI community.

Another advantage is that the images do not need to be pre-processed, so it can be useful to batch process large numbers of images. It is planned to use this model for processing the data of the entire state and determining its soil fund.

The main scope of our project is that we classify the land as per various feature for example: Forest/trees, crop, grasslands, water bodies and we make sure how it will help the environment.

4. REFERENCES

[1] Krizhevsky, Alex, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Advances in Neural Information Processing Systems, pp. 1097–1105, 2012.

[2] M.J. Canty, Image Analysis, Classification and Change Detection in Remote Sensing with Algorithms for ENVI/IDL and Python. CRC Press, 3rd edition, 2014.

[3] S. H. Hasanpour, M. Rouhani, M. Fayyaz, and M. Sabokrou, "Lets keep it simple, Using Simple Architectures to overcome deeper and more complex Architectures", Computer Vision and Pattern Recognition, vol. arXiv: 1608.06037, February 2018.

[4] S. Basu, S. Ganguly, S. Mukhopadhyay, R. DiBiano, M. Karki, and R. Nemani, "Deepsat-A Learning Framework for Satellite Imagery", arXiv preprint, vol. arXiv:1509.03602, September 2015.

[5] T. Blaschke and J. Strobl: “What's wrong with pixels? Some recent developments in remote sensing and GIS,” GeoBIT/GIS, vol. 6, No. 1, pp. 12–17, 2001.

[6] T. Blaschke, "Object based Image Analysis for Remote Sensing", ISPRS J. Photogramm. Remote Sens, Vol. 65, no. 1, pp. 2–16, 2010.

[7] C. L. Zhong Ma, Zhuping Wang, and X. Liu, "Deep Convolution Network Based Satellite Image Classification," International Journal of Computer and Information Engineering, vol. 10, No. 6, pp. 1156–1159, 2016.

[8] G. Cheng and J. Han, "Remote Sensing Scene Classification: Benchmark and State-of-the-art," Proceedings of the IEEE, vol. DOI:10.1109/JPROC.2017.2675998, 2017.

[9] Y. Zhong, F. Fei, Y. Liu, B. Zhao, H. Jiao, and L. Zhang, "SatCNN: Satellite Image Dataset Classification using Agile Convolutional Neural Networks," Remote Sensing Letters, vol. 8, No. 2, pp. 136–145, 2017.

[10]Z. Gong, P. Zhong, Y. Yu, and W. Hu, "Diversity-Promoting Deep Structural Metric Learning for Remote Sensing Scene Classification," IEEE Transactions on Geoscience and Remote Sensing, January 2018.

[11]J. Wang, "Learning Fine-Grained Image Similarity with Deep Ranking", IEEE Conf. on Computer Visual Pattern Recognition, June 2014.

[12] Q. Liu, S. Basu, S. Ganguly, S. Mukhopadhyay, R. DiBiano, M. Karki, and R. Nemani, "DeepSat V2: Feature Augmented Convolutional Neural Nets for Satellite Image Classification," Remote Sensing Letters , November 2019.

[13]Y. Heryadi and E. Miranda, "Land cover classification based on Sentinel-2 satellite imagery using a convolutional neural network model: A case study in Semarang, Indonesia", Asian Conference on Intelligent Information and Database Systems (ACIIDS), p .191–206, 2019.