

# Satellite Image Classification with Deep Learning: Survey

Roshni Rajendran<sup>1</sup>, Liji Samuel<sup>2</sup>

<sup>1</sup>M. Tech Student, <sup>2</sup>Assistant Professor,

<sup>1,2</sup>Department of Computer Science and Engineering, Sree Buddha College of Engineering, Kerala, India

## ABSTRACT

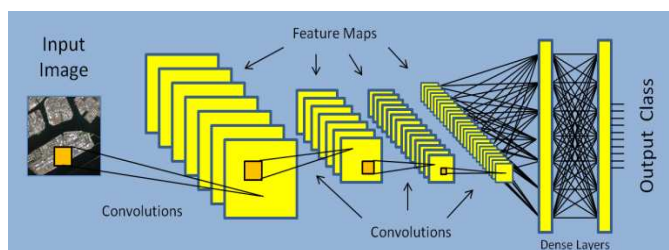
Satellite imagery is important for many applications including disaster response, law enforcement and environmental monitoring etc. These applications require the manual identification of objects in the imagery. Because the geographic area to be covered is very large and the analysts available to conduct the searches are few, thus an automation is required. Yet traditional object detection and classification algorithms are too inaccurate and unreliable to solve the problem. Deep learning is a part of broader family of machine learning methods that have shown promise for the automation of such tasks. It has achieved success in image understanding by means that of convolutional neural networks. The problem of object and facility recognition in satellite imagery is considered. The system consists of an ensemble of convolutional neural networks and additional neural networks that integrate satellite metadata with image features.

**KEYWORDS:** Deep Learning, Convolutional neural network, VGG, GPU

## INTRODUCTION

Deep learning is a class of machine learning models that represent data at different levels of abstraction by means of multiple processing layers. It has achieved astonishing success in object detection and classification by combining large neural network models, called CNN with powerful GPU. CNN-based algorithms have dominated the annual ImageNet Large Scale Visual Recognition Challenge for detecting and classifying objects in photographs. This success has caused a revolution in image understanding, and the major technology companies, including Google, Microsoft and Facebook, have already deployed CNN-based products and services.

A CNN consists of a series of processing layers as shown in Fig. 1.1. Each layer is a family of convolution filters that detect image features [1]. Near the end of the series, the CNN combines the detector outputs in fully connected “dense” layers, finally producing a set of predicted probabilities, one for each class. The network itself learns which features to detect, and how to detect them, as it trains.



**Fig:1 Architecture of satellite image classification**

**How to cite this paper:** Roshni Rajendran | Liji Samuel "Satellite Image Classification with Deep Learning: Survey" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-4 | Issue-2, February 2020, pp.583-587, URL: [www.ijtsrd.com/papers/ijtsrd30031.pdf](http://www.ijtsrd.com/papers/ijtsrd30031.pdf)



Copyright © 2019 by author(s) and International Journal of Trend in Scientific Research and Development Journal. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0) (<http://creativecommons.org/licenses/by/4.0/>)



The earliest successful CNNs consisted of less than 10 layers and were designed for handwritten zip code recognition. LeNet had 5 layers and AlexNet had 8 layers. Since then, the trend has been toward more complexity. VGG appeared with 16 layers followed by Google’s Inception with 22 layers. Subsequent versions of Inception have even more layers, ResNet has 152 layers and DenseNet has 161 layers.

Such large CNNs require computational power, which is provided by advanced GPUs. Open source deep learning software libraries such as Tensor Flow and Keras, along with fast GPUs, have helped fuel continuing advances in deep learning.

## LITERATURE SURVEY

Various methods are used for identifying the objects in satellite images. Some of them are discussed below.

### 1. Rocket Image Classification Based on Deep Convolutional Neural Network

Liang Zhang et al. [2] says that in the field of aerospace measurement and control field, optical equipment generates a large amount of data as image. Thus, it has great research value for how to process a huge number of image data quickly and effectively. With the development of deep learning, great progress has been made in the task of image classification. The task images are generated by optical measurement equipment are classified using the deep learning method. Firstly, based on residual network, a general deep learning image classification framework, a binary image classification network namely rocket image

and other image is built. Secondly, on the basis of the binary cross entropy loss function, the modified loss function is used to achieve a better generalization effect on those images difficult to classify. Then, the visible image data downloaded from optical equipment is randomly divided into training set, validation set and test set. The data augmentation method is used to train the binary classification model on a relatively small training set. The optimal model weight is selected according to the loss value on the validation set. This method has certain value for exploring the application of deep learning method in the intelligent and rapid processing of optical equipment task image in aerospace measurement and control field.

## 2. Super pixel Partitioning of Very High Resolution Satellite Images for Large-Scale Classification Perspectives with Deep Convolutional Neural Networks

T. Postadjiana et al. [3] proposes that supervised classification is the basic task for landcover map generation. From semantic segmentation to speech recognition deep neural networks has outperformed the state-of-the-art classifiers in many machine learning challenges. Such strategies are now commonly employed in the literature for the purpose of land-cover mapping. The system develops the strategy for the use of deep networks to label very high resolution satellite images, with the perspective of mapping regions at country scale. Therefore, a super pixel based method is introduced in order to (i) ensure correct delineation of objects and (ii) perform the classification in a dense way but with decent computing times.

## 3. Cloud Cover Assessment in Satellite Images via Deep Ordinal Classification

Chaomin Shen et al. [4] discuss that the percentage of cloud cover is one of the key indices for satellite imagery analysis. To date, cloud cover assessment has performed manually in most ground stations. To facilitate the process, a deep learning approach for cloud cover assessment in quicklook satellite images is proposed. Same as the manual operation, given a quicklook image, the algorithm returns 8 labels ranging from A to E and \*, indicating the cloud percentages in different areas of the image. This is achieved by constructing 8 improved VGG-16 models, where parameters such as the loss function, learning rate and dropout are tailored for better performance. The procedure of manual assessment can be summarized as follows. First, determine whether there is cloud cover in the scene by visual inspection. Some prior knowledge, e.g., shape, color and shadow, may be used. Second, estimate the percentage of cloud presence. Although in reality, the labels are often determined as follows. If there is no cloud, then A; If a very small amount of clouds exist, then B; C and D are given to escalating levels of clouds; and E is given when the whole part is almost covered by clouds. There is also a label \* for no-data. This mostly happens when the sensor switches, causing no data for several seconds. The disadvantages of manual assessment are obvious. First of all, it is tedious work. Second, results may be inaccurate due to subjective judgement.

A novel deep learning application in remote sensing by assessing the cloud cover in an optical scene is considered. This algorithm uses parallel VGG 16 networks and returns 8 labels indicating the percentage of cloud cover in respective

subscenes. The proposed algorithm combines several state-of-the-art techniques and achieves reasonable results.

## 4. Learning Multiscale Deep Features for High-Resolution Satellite Image Scene Classification

Qingshan Liu et al. [5] discuss about a multiscale deep feature learning method for high-resolution satellite image scene classification. However, satellite images with high spatial resolution pose many challenging issues in image classification. First, the enhanced resolution brings more details; thus, simple lowlevel features (e.g., intensity and textures) widely used in the case of low-resolution images are insufficient in capturing efficiently discriminative information. Second, objects in the same type of scene might have different scales and orientations. Besides, high-resolution satellite images often consist of many different semantic classes, which makes further classification more difficult. Taking the commercial scene comprises roads, buildings, trees, parking lots, and so on. Thus, developing effective feature representations is critical for solving these issues.

Specifically, at first warp the original satellite image into multiple different scales. The images in each scale are employed to train a deep convolutional neural network (DCNN). However, simultaneously training multiple DCNNs is time consuming. To address this issue, a DCNN with spatial pyramid pooling is explored. Since different SPP-nets have the same number of parameters, which share the identical initial values, and only fine-tuning the parameters in fully connected layers ensures the effectiveness of each network, thereby greatly accelerating the training process. Then, the multiscale satellite images are fed into their corresponding SPP-nets, respectively, to extract multiscale deep features. Finally, a multiple kernel learning method is developed to automatically learn the optimal combination of such features.

## 5. Domain Adaptation for Large Scale Classification of Very High Resolution Satellite Images With Deep Convolutional Neural Networks

T. Postadjiana et al. [6] discuss about semantic segmentation of remote sensing images enables in particular land-cover map generation for a given set of classes. Very recent literature has shown the superior performance of DCNN for many tasks, from object recognition to semantic labelling, including the classification of VHR satellite images. A simple yet effective architecture is DCNN. Input images are of size of  $65 \times 65 \times 4$  (number of bands). Only convolutions of size  $3 \times 3$  are used to limit the number of parameters. The second line of the table displays the number of filters per convolution layer. The ReLU activation function is set after each convolution in order to introduce non-linearity and max-pooling layers increase the receptive field (the spatial information taken into account by the filter). Finally, a fully-connected layer on top sums up the information contained in all features in the last convolution layer.

However, while plethora of works aim at improving object delineation on geographically restricted areas, few tend to solve this classification task at very large scales. New issues occur such as intra-class variability, diachrony between surveys, and the appearance of new classes in a specific area, that do not exist in the predefined set of labels. Therefore, this work intends to (i) perform large scale classification and

to (ii) expand a set of land-cover classes, using the off-the-shelf model learnt in a specific area of interest and adapting it to unseen areas.

## 6. Introducing Eurosat: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification

Patrick Helber et al. [7] discuss about the challenge of land use and land cover classification using Sentinel-2 satellite images. The key contributions are as follows. A novel dataset based on Sentinel-2 satellite images covering 13 different spectral bands and consisting of 10 classes with in total 27,000 labeled images are presented. The state-of-the-art CNN on this novel dataset with its different spectral bands are considered. Also evaluate deep CNNs on existing remote sensing datasets. With the proposed novel dataset, a better overall classification accuracy is gained. The classification system resulting from the proposed research opens a gate towards various Earth observation applications.

The challenge of land use and land cover classification is considered. For this task, dataset based on Sentinel-2 satellite images are used. The proposed dataset consists of 10 classes covering 13 different spectral bands with in total 27,000 labeled images. The evaluated state of the art deep CNNs on this novel dataset. Also evaluated deep CNNs on existing remote sensing datasets and compared the obtained results. For the novel dataset, the performance is analysed based on different spectral bands. The proposed research can be leveraged for multiple real-world Earth observation applications. Possible applications are land use and land cover change detection and the improvement of geographical maps.

## 7. Cloud Classification of Satellite Image Based on Convolutional Neural Networks

Keyang Cai et al. [8] proposes about cloud classification of satellite image in meteorological forecast. Traditional machine learning methods need to manually design and extract a large number of image features, while the utilization of satellite image features is not high. CNN is the first truly successful learning algorithm for multi-layer network structure. Compared with the general forward BP algorithm, CNN can reduce the number of parameters needed in learning in spatial relations, so as to improve the training performance. A small piece of local feel in the picture as the bottom of the hierarchical structure of the input, continue to transfer to the next layer, each layer through the digital filter to obtain the characteristics of the data. This method has significant effect on the observed data such as scaling, rotation and so on. The cloud classification process of satellite based on deep convolution neural network includes pre-processing, feature extraction, classification and other steps. A convolution neural network for cloud classification, which can automatically learn features and obtain classification results is constructed. The method has high precision and good robustness.

## 8. Hyperspectral Classification Using Stacked Autoencoders With Deep Learning

A. Okan Bilge Ozdemir et al. [9] says that the stacked autoencoders which are widely utilized in deep learning research are applied to remote sensing domain for hyperspectral classification. High dimensional hyperspectral data is an excellent candidate for deep learning methods.

However, there are no works in literature that focuses on such deep learning approaches for hyperspectral imagery. This aims to fill this gap by utilizing stacked autoencoders. Using stacked autoencoders, intrinsic representations of the data are learned in an unsupervised way. Using labeled data, these representations are fine tuned. Then, using a soft-max activation function, hyperspectral classification is done. Parameter optimization of Stacked Autoencoders (SAE) is done with extensive experiments. It focuses on utilization of representation learning methods for hyperspectral classification task on high definition hyperspectral images. Strong side of representation learning methods are its unsupervised automatic feature learning step which makes it possible to omit the feature extraction step that requires domain knowledge.

Proposed method utilizes SAEs to learn discriminative representations in the data. Using the label information in the training set, representations are fine tuned. Then, the test set is classified using these representations. Effects of different parameters on the performance results are also investigated. Most effective parameter is found as the epoch number of the supervised classification step. For each parameter (learning rate, batch size and epoch numbers), optimal values are approximated with extensive testing. From 432 different parameter configurations, optimal parameter values are found as 25, 60, 100 and 5 for batch size, epoch numbers (representation & classification) and learning rate, respectively. This score is greatly promising when compared to state-of-the-art methods. This study can be enhanced by two important additions, neighbourhood factors and utilization of unlabelled data. With these additions as future study, more general and robust results can be achieved.

## 9. Deep Learning Crop Classification Approach Based on Sparse Coding of Time Series of Satellite Data

Mykola Lavreniuk et al. [10] says that crop classification maps based on high resolution remote sensing data are essential for supporting sustainable land management. The most challenging problems for their producing are collecting of ground-based training and validation datasets, non-regular satellite data acquisition and cloudiness. To increase the efficiency of ground data utilization it is important to develop classifiers able to be trained on the data collected in the previous year. A deep learning method is analyzed for providing crop classification maps using in-situ data that has been collected in the previous year. Main idea is to utilize deep learning approach based on sparse autoencoder. At the first stage it is trained on satellite data only and then neural network fine-tuning is conducted based on in-situ data from the previous year. Taking into account that collecting ground truth data is very time consuming and challenging task, the proposed approach allows us to avoid necessity for annual collecting in-situ data for the same territory.

This approach is based on converting time series of observations into sparse vectors from the unified hyperspace with similar representation. This idea allows further utilization of the autoencoder to learn dependencies in data without labels and fine-tuning final neural network using available in-situ data from previous year. This methodology was utilized to provide crop classification map for central part of Ukraine for 2017 based on multi-temporal Sentinel-1 images with different acquisition dates.



## 10. Deep Learning for Amazon Satellite Image Analysis

Lior Bragilevsky et al. [11] says that satellite image analysis has become very important in a lot of areas such as monitoring floods and other natural disasters, earthquake and tsunami prediction, ship tracking and navigation, monitoring the effects of climate change etc. Large size and inaccessibility of some of the regions of Amazon wants best satellite imaging system. Useful informations are extracted from satellite images will help to observe and understand the changing nature of the Amazon basin, and help better to manage deforestation and its consequences. Some images were labelled and could be used to train various machine learning algorithms. Another set of test image was provided without labels and was used to make predictions. The predicted labels for test images were submitted to the competition for scoring. The hope was that well-trained models arising from this competition would allow for a better understanding of the deforestation process in the Amazon basin and give insight on how to better manage this process.

Machine learning can be the key to saving the world from losing football field-sized forest areas each second. As deforestation in the Amazon basin causes devastating effects both on the ecosystem and the environment, there is urgent need to better understand and manage its changing landscape. A competition was recently conducted to develop algorithms to analyze satellite images of the Amazon. Successful algorithms will be able to detect subtle features in different image scenes, giving us the crucial data needed to be able to manage deforestation and its consequences more effectively.

## 11. Deep Learning-Based Large-Scale Automatic Satellite Crosswalk Classification

Rodrigo F. Berriel et al. [12] proposes about a zebra crossing classification and detection system. It is one of the main tasks for mobility autonomy. Although important, there are few data available on where crosswalks are in the world. The automatic annotation of crosswalks locations worldwide can be very useful for online maps, GPS applications and many others. In addition, the availability of zebra crossing locations on these applications can be of great use to people with disabilities, to road management, and to autonomous vehicles. However, automatically annotating this kind of data is a challenging task. They are often aging (painting fading away), occluded by vehicle and pedestrians, darkened by strong shadows, and many other factors.

High-resolution satellite imagery has been increasingly used on remote sensing classification problems. One of the main factors is the availability of this kind of data. Despite the high availability, very little effort has been placed on the zebra crossing classification problem. Crowd sourcing systems are exploited in order to enable the automatic acquisition and annotation of a large-scale satellite imagery database for crosswalks related tasks.

## CONCLUSION

Various methods to capture and identify the objects in the satellite imageries are analyzed. Each of them works with different schemes. Each scheme varies on the convolution layers used and also their application used.

A deep learning system that classifies objects in the satellite imagery is considered. The system consists of an ensemble of CNNs with post-processing neural networks that combine the predictions from the CNNs with satellite metadata. Combined with a detection component, this system could search large amounts of satellite imagery for objects or facilities of interest. Analyzing satellite imagery has very important role in various applications such as it could help law enforcement officers to detect unlicensed mining operations or illegal fishing vessels, assist natural disaster response teams with the mapping of mud slides or hurricane damage and enable investors to monitor crop growth or oil well development more effectively.

## REFERENCES

- [1] M. Pritt and G. Chern, "Satellite Image Classification with Deep Learning," *2017 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, Washington, DC, 2017, pp. 1-7.
- [2] L. Zhang, Z. Chen, J. Wang and Z. Huang, "Rocket Image Classification Based on Deep Convolutional Neural Network," *2018 10th International Conference on Communications, Circuits and Systems (ICCCAS)*, Chengdu, China, 2018, pp. 383-386.
- [3] C. Shen, C. Zhao, M. Yu and Y. Peng, "Cloud Cover Assessment in Satellite Images Via Deep Ordinal Classification," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia, 2018, pp. 3509-3512.
- [4] T. Postadjian, A. L. Bris, C. Mallet and H. Sahbi, "Superpixel Partitioning of Very High Resolution Satellite Images for Large-Scale Classification Perspectives with Deep Convolutional Neural Networks," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia, 2018, pp. 1328-1331.
- [5] Q. Liu, R. Hang, H. Song and Z. Li, "Learning Multiscale Deep Features for High-Resolution Satellite Image Scene Classification," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 1, pp. 117-126, Jan. 2018.
- [6] T. Postadjian, A. L. Bris, H. Sahbi and C. Malle, "Domain Adaptation for Large Scale Classification of Very High Resolution Satellite Images with Deep Convolutional Neural Networks," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia, 2018, pp. 3623-3626.
- [7] P. Helber, B. Bischke, A. Dengel and D. Borth, "Introducing Eurosat: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia, 2018, pp. 204-207.
- [8] K. Cai and H. Wang, "Cloud classification of satellite image based on convolutional neural networks," *2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, Beijing, 2017, pp. 874-877.
- [9] A. O. B. Özdemir, B. E. Gedik and C. Y. Y. Çetin, "Hyperspectral classification using stacked autoencoders with deep learning," *2014 6th Workshop*

- on *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Lausanne, 2014, pp. 1-4.
- [10] M. Lavreniuk, N. Kussul and A. Novikov, "Deep Learning Crop Classification Approach Based on Sparse Coding of Time Series of Satellite Data," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia, 2018, pp. 4812-4815.
- [11] L. Bragilevsky and I. V. Bajić, "Deep learning for Amazon satellite image analysis," *2017 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, Victoria, BC, 2017, pp. 1-5.
- [12] R. F. Berriel, A. T. Lopes, A. F. de Souza and T. Oliveira-Santos, "Deep Learning-Based Large-Scale Automatic Satellite Crosswalk Classification," in *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 9, pp. 1513-1517, Sept. 2017.

