MAT JOURNALS

Journal of Image Processing and Artificial Intelligence
Volume 4 Issue 2

Two-layer approach for Unsupervised and Semi-Supervised Learning for Satellite Images

***Manish Kumar*[1], *Neha Patil*[2], *Poonam Tiwari*[3], *Pooja Varier*[4]**
[1234]Department of Computer Engineering
K. K. Wagh Institute of Engineering Education & Research
Nashik-03, India
mk14445@gmail.com[1] , 13nehapatil1997@gmail.com[2]
poonamtiwari9028@gmail.com[3], poo.mohan555@gmail.com[4]

Abstract

In our galaxy, there are many advanced satellite. Large distance image can be captured with very high quality. The image provides sufficient information at global level and regional level. The field of satellite imagery is evaluated so much that it has created questions to human and environmental sustainability. It is still a challenge to scale those techniques to very high spatial resolutions. Satellite images are of greater spatial, spectral and high resolution creating large set of information about the image which makes it difficult to identify the features of images. This is because the images are unlabeled. Unsupervised method allows us to organize images into clusters. However, unsupervised method like machine learning uses features for clustering. Those images which are close, are kept in same group. The system uses satellite images datasets which provide aerial shots of different location. Images are grouped into sets of 5 where each image in a set was taken on a different day at a specific location but not necessarily at the same time each day. The images for each set cover the same area but not perfectly align. This dataset is provided as input dataset to proposed system. Feature extraction is done by sending a set of images through a network and extracting features at a certain layer which results in a feature set for a certain network. The process of transfer learning involves sending our own images through the network and extracting features at a certain layer. The process followed in this system is different from fine-tuning because images are not trained and the number of classes is not changed in the SoftMax layer of the network. Rather parameters learned by a pre-trained model to see if it can be used for an unlabeled dataset where fine tuning would not be possible are used.

Key Words**:** satellite imagery, remote sensing, deep learning, convolutional neural networks, transfer learning

INTRODUCTION

Machine Learning has been widely used in various domains to study and learn from patterns in data to make accurate predictions. The use of machine learning can be seen in our daily lives, and in various domains. Satellite images are one of the most powerful and important tools by the meteorologists. They are essentially the eyes in the sky. Satellite imagery are the images of earth or other planets collected by imaging satellites operated by governments and businesses around the world. As the requirement of satellite images is increasing in various fields, it has become vital to do the processing of these images by feature extraction. In machine learning, pattern recognition and image processing, feature extraction involves reducing the amount of resources required to describe a large set of data. Feature extraction is the most important application in spatial data management in the manner of automatic updating of Geographic information system databases from satellite imagery. It has been widely done by manual interpretation or automatic classification. [1]

There are many existing systems which perform classification of satellite images. But if considering the accuracy, most of them had failed. Some systems fail because they are not able to distinguish between the natural and man-made things, differentiate between land cover and water. So, there is need of some system where manual effort should be lower with increase in accuracy and function. The implementation is based on finding similar images from the given trained dataset. This involves trained dataset and testing phase. Trained dataset includes labelled images. Testing phase includes a query image whose features will be extracted and based on those features trained images will be matched and thus similar images will be predicted. Each layer will be extracting and learning some of its feature for future need. It classifies the set of similar images into single group.

Applications

The following departments do not have any system which can tell them about land and water coverage areas. Thus, the proposed system can be applied to overcome the drawback of existing systems.

1. Meteorology: It is a branch of atmospheric science which includes atmospheric chemistry and atmospheric physics with major focus on weather forecasting.
2. Oceanography: It is the study of physical and biological aspects of the ocean.
3. Agriculture: It is the field to monitor farm operations from anywhere and get real time insights. Agricultural applications include crop identification, area estimation, crop condition assessment and soil moisture estimation.
4. Biodiversity: The plant and animal life in the world is of great importance. Capturing images of earth gives us unique information and a perspective of what is going on in the world and thus plays a vital role in biodiversity
5. Forestry: This field includes terrain analysis, forest management, updating of existing forest inventories, forest cover type discrimination, the delineation of burned areas and mapping of cleared areas.
6. Regional planning: In order to measure changes in land use, the need to platform in monitoring, recording, and predicting the changes is necessary for planners and developers. In advance technology of mapping process, remote sensing and GIS as tools for urban planning are already recognized.

LITERATURE SURVEY

While satellite imagery is increasingly accessible, existing techniques for analyzing satellite imagery is increasingly compute intensive as well. Only in the last few years there have been researches that analyze imagery at national and global scales [1], [2], which are keeping up with the volume and resolution of imagery available. Their techniques require many hours of manual classifications, in the case of [2], to validate the analysis, and intensive supercomputing and classification in the case of [1]. To identify the content of images is a severe bottleneck. Due to climatic changes, industrialization, massive migrations and due to many other such causes, the rate of change occurring on the earth's surface increases rapidly. Thus, measuring and monitoring these changes is critical and labelling the content of satellite and aerial scene images is a more difficult task. Automated learning and machine learning gives us the opportunity to track changes of scene imagery which has no dependence on human intervention. Recently deep learning methods for computer vision applications have brought impressive results [3]–[6]. A common theme with deep learning in computer vision is that architectures are based on Convolutional Neural Networks (CNNs) [3] consisting of many computation layers that allow for the learning of high level features.

Authors in [7] showed that Satellite images are unlabeled. Therefore, unsupervised learning allows us to organize images into coherent groups or clusters. Features from pre-trained networks can be used for learning in temporally and spatially dynamic data sources such as satellite imagery. This system explores and evaluates different features and feature combinations extracted from various deep network architectures. The process followed in this system is different from fine-tuning because images are not trained and the number of classes is not changed in the SoftMax layer of the network. Rather parameters learned by a pre-trained model to see if it can be used for an unlabeled dataset where fine tuning would not be possible are used. In addition, authors test the transferability of engineered features and learned features from an unlabeled dataset to a different labeled dataset. The feature engineering and learning are done on the unlabeled Draper Satellite Chronology dataset, and we test on the labeled UC Merced Land dataset to achieve near state-of-the-art classification results. Hence these results give us an idea that even with minimum training, these networks can be generalized well to other datasets. This process can be used in clustering unlabeled images and different other unsupervised machine learning activities

In this system [4], effect of convolutional network depth on its accuracy in the large-scale image recognition setting is investigated. The main contribution is a thorough evaluation of networks of increasing depths using an architecture with very small (3 x 3) convolution filters, which shows that a significant improvement on the prior art configurations can be achieved by pushing depth to 16-19 weight layers. In deep convolutional learning performance of individual ConvNet models is evaluated. In the further experiments, outputs of several models are combined by averaging their SoftMax class posteriors. This improves performance due to complementarity of the models and was used in the top ILSVRC submissions in 2012 and 2013.

Authors in [8] propose a novel method for finger vein presentation attack detection(PAD) by exploring the transfer learning ability of deep convolution neural network(CNN). To this extent this system considers pre-trained Alexnet architecture and augmented the existing architecture with additional seven layers to improve the reliability and reduce over fitting problem. This system the fine-tunes the modified CNN architecture with the finger vein presentation attack samples to make it adaptable to finger vein PAD

Deeper neural networks are more difficult to train. In [6] authors studied, deep residential learning framework presents a residential learning framework to ease the training of networks that are substantially deeper than those used previously. This provides comprehensive empirical evidence showing that these residential networks are easier to optimize and can gain accuracy from considerably increased depth. On the ImageNet dataset, residential nets with a depth of up to 152 layers, 8 x deeper than VGG nets is evaluated but still having lower complexity.

The increase of image data on the internet has the potential to create a set of organized and robust models. Algorithm does the work of indexing, fetching, organizing and interacting with images and multimedia data. But exactly how such data can be used and organized remains a crucial problem. This system introduces here a new database called ImageNet [9], a large-scale ontology of images built upon the backbone of the world net structure. ImageNet aims to populate the majority of 80,000 synsets of WordNet with an average of 500-1000 clean and full resolution images. This system offers a detailed analysis of ImageNet and its current state: 12 subtrees with 5247 synsets and 3.2 million images in total.

MAT JOURNALS

Journal of Image Processing and Artificial Intelligence
Volume 4 Issue 2

In this section we present a deep literature survey on the existing system and studying those system we propose our novel approach for two-layer graph-based learning satellite image classification. The first layer will have unsupervised learning which will do feature transformation and in second layer semi supervised learning will be done for image classification.

PROPOSED SYSTEM

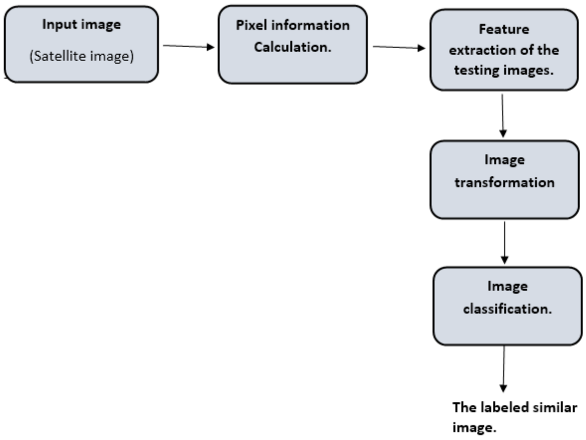

Fig 1. *Block Diagram*

Input: The input to the system is the satellite images. Images are grouped into sets of five where each image in set was taken on a different day at a specific location. The images in each set cover the same area but are not perfectly aligned. Some locations show little evidence of changes from day to day while other locations will have subtle changes such as moving vehicles or changes in shadows. These images are a priority dataset because they are labeled by their location and are grouped together as sets of 5 images which allow us to quantify the performance of a pre-trained networks layer as a feature extractor. The dataset will be split into a validation set and testing set. We denote the rst image in a set of 5 images as the query image for that location. We also use a labeled dataset for additional testing and validation of our methodologies. We use the UC Merced dataset that is composed of 2,100 aerial scene images with 256 x 256 pixels divided into 21 land-use classes.

Pixel Information Calculation: The pixels of the images are calculated and the information is gathered for feature extraction.

Feature Extraction: The method of utilizing pre-trained networks to apply to another learning task is known as transfer learning. Transfer learning allows us to apply the features extracted from large existing already-labeled datasets onto a new task. This approach is motivated by the fact that deep neural networks are supposed to find features in the representation of the task and that features that are learned early in the network are general enough that they can be applied successfully to another image dataset. Feature Extraction is done by sending a set of images through a network and extracting features at a certain layer Lk which results in a feature set fk for a certain Network. The process of transfer learning involves sending our own images through the network and extracting features at a certain layer. The features extracted are from a fully connected layer since the output parameters are more reasonable to work with.

Image transformation: Transformation is a function that maps one set to another set after performing some operations. The input to the system would be an image and output would be an image too. And the system would perform some processing in the input image and gives its output as the processed image.
Consider this equation,
Gx, y = T (fx, y)
In this equation,
Fx, y = Input image on which transformation function has to be applied.
Gx, y = The output image or processed image
T is transformation function. This relation between input image and processed output image can also be represented as,
s = Tr
Where r is actually the pixel value or gray level intensity of fx,y at any point and s is the pixel value or gray level intensity of Gx,y at any point.

Image classification: Based on the features extracted, images of the testing dataset will be classified and hence the similar images will be predicted according to the labeled pre-trained dataset.
Output: Images of the testing dataset which are similar to the query image

CONCLUSIONS

In this system, the process of unsupervised learning will be improved and also accuracy can be increased. This can be done by using another approach which is to fine tune the deeper networks such as residual networks and then take a step to concatenating them which has been shown to have higher performance. An important purpose of this technique is to be able to identify the features of the landscape and be able to map those features back on the earth. The input to the system is a dataset of satellite images. The images are represented as large bi-dimensional array. Array of integer and double will be used to extract the features. The pixel information of the images is gathered. The problem here is divided into two layers-the first layer here is feature transformation which uses unsupervised learning and second layer uses supervised learning for image classification. After processing of the two layers, based on the training dataset and the features classified, the segmented image if the testing dataset will be the output of the system.

REFERENCES

1. M. C. Hansen, P. V. Potapov, R. Moore, M. Hancher, S. Turubanova, A. Tyukavina, D. Thau, S. Stehman, S. Goetz, T. Loveland et al., “High-resolution global maps of 21st-century forest cover change,” Science, vol. 342, no. 6160, pp. 850–853, 2013.
2. R. Goldblatt, W. You, G. Hanson, and A. K. Khandelwal, “Detecting the boundaries of urban areas in india: A dataset for pixel-based image classification in google earth engine,” Remote Sensing, vol. 8, no. 8, p. 634, 2016.
3. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Advances in neural information processing systems, 2012, pp. 1097–1105.
4. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint arXiv:1409.1556, 2014.
5. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.
6. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” arXiv preprint arXiv:1512.03385, 2015.
7. B. Hedayatnia, M. Yazdani, M. Nguyen, J. Block, and I. Altintas, \Determining Feature Extractors for

Unsupervised Learning on Satellite Images," based on big data, 2016, pp. 26552663.
8. J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, \How transferable are features in deep neural networks?" in Advances in neural information processing systems, 2014, pp. 33203328
9. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, \Imagenet: A large-scale hierarchical image database," in Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009, pp. 248255.