International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 4 (2019) pp. 881-887

A Study on Deep Learning Models for Satellite Imagery

M. P. Vaishnnave[1], K. Suganya Devi [*], P. Srinivasan[2]

[1]*Dept of CSE, University College of Engineering, Panruti, Tamil Nadu 607106, India.*

[*] *Assistant Professor Grade I, National Institute of Technology Silchar Cachar, Assam 788010, India. (*Corresponding Author)*

[2]*Assistant Professor Grade I, National Institute of Technology Silchar Cachar, Assam 788010, India.*

Abstract

Satellite Imagery Scene classification has been receiving a remarkable attention as it plays a vital role in a wide range of applications. Also, there has been a massive growth in Deep learning in many fields such as computer vision and natural language processing. But, still there exists a lack of deep review for the datasets and methods available for scene classification from the satellite imagery. This paper focuses on enlightening the concept and evolution of Deep Learning. Also, this paper provides a comprehensive review of the recent progress on various datasets and methods available for scene classification and the same has been summarized in a table. The state-of-art quantitative evaluation metrics for various approaches are discussed concluding by identifying some research gaps in the deep learning for satellite imagery.

Keywords: Satellite Imagery scene classification, Machine Learning, Deep learning models, Deep Learning

I. INTRODUCTION

Earth Observation (EO) is a process of gathering the information about planet Earth through remote sensing. The location, where we can collect the most data about our planet, is in space. Deep learning is a subgroup of machine learning, and refers to the application of a set of algorithms called neural networks, and their variants. In such techniques, one offers the network with a set of labeled examples which it learns, or trains on [2].Labeling these examples is done in many ways. Machine learning feature extraction is done by manually and classification is done by machine as shown in fig2. However, in deep learning both the feature extraction and the classification are done by machine as given in fig1. So Deep Learning neural network is more efficient to identify the Satellite Imagery.

Fig1. Deep Learning neural networks

II. REVIEW OF DEEP LEARNING METHODS FOR SATELLITE IMAGERY CLASSIFICATION

A. *Auto-Encoders*

Auto-Encoders consists of three layers viz inputs layer, output layer and hidden layer. In the training mapping of the input $V \epsilon R^d$ to the hidden layer takes place first and produces the output $K \epsilon R^d$ This process is done by encoder. After this, mapping of hidden layer to output layer is done by decoder [3]. A decoder produces the output layer having the same number of nodes as the input layer. Output layer is called as "reconstruction." The values in reconstructed layer are denoted as $Z \epsilon R^d$. Figure 6 shows the workflow of AutoEncoder .The architecture determines a hidden feature "s" from input "v" by reconstructing it on "k".Auto-Encoder Workflow shows in fig2.

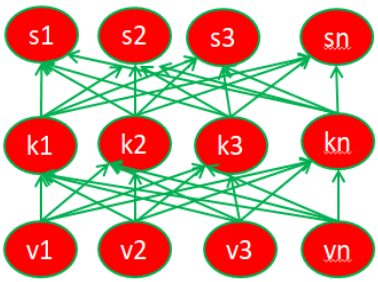

Fig2. AutoEncoder workflow

$$s = f(N_s V + a_s) \quad (1)$$

$$k = f(N_k s + a_k) \quad (2)$$

where N_s and N_k denote the input-to-hidden and the hidden-to output weights, respectively, a_s and a_k denote the bias of hidden and output units, f(.)denotes the activation function. Conventionally, the nonlinearity is provided in f(.)such as sigmoid function.

$$N_s = N_k' = N \quad (3)$$

The goal of training is to minimize the "error" between input and reconstruction, i.e.,

$$\arg\min_{N,b_s,b_k}[J(V,K)] \tag{4}$$

Where K is dependent on parameter N, a_s and a_k while v is given J(V,K) stands for the error and ή is learning rate.

$$N = N - \eta \frac{\partial \cos t(V,K)}{\partial N} \tag{5}$$

$$a_s = a_s - \eta \frac{\partial \cos t(V,K)}{\partial a_s} \tag{6}$$

$$a_k = a_K - \eta \frac{\partial \cos t(V,K)}{\partial a_k} \tag{7}$$

After training the network, the reconstruction layer composed with its bounds are removed and the learned feature lies in the hidden layer,which can successively be used for classification or used as the input of an upper layer to produce a deeper feature[4].

B. *Stacked Auto-Encoder*

Stacked Auto-encoder consist of multiple Autoencoders together. If we stacked input and output layers of Auto-encoders layer by layer,then get a stacked Autoencoder[6]. During training procedure every single Auto-encoder of stacked auto-encoder deals with the same technique of autoencoder discussed above. Every time training of subsequent layers of Auto-encoder takes place with the help of the outputs of the previous layer[5]. After this layer of training, the decoder of the third layer Auto-Encoder is useless and we are considering input to hidden parameters as weights between second & third layer. If the following classifier is implemented as a neural network, the parameters throughout the whole network can be adjusted slightly while training the classifier.Fig3. shows the work flow of Stacked Auto –Encoder.

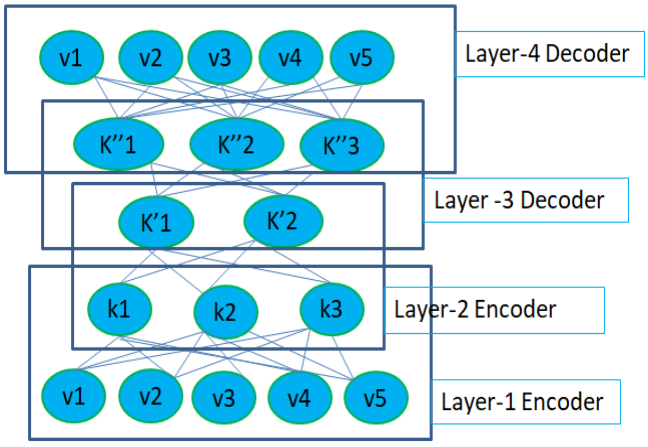

Fig3. show the workflow

This step is called fine-tuning[7]. For logistic regression, the training is simply back propagation, searching for a minimum in a peripheral region of parameters initialized by the former step.

C. *Restricted Boltzmann machine:*

An Restricted Boltzmann Machine is a unique type of Markov random field contains one layer of stochastic hidden units and one layer of stochastic visible or observable units [9]. Fig.4 shows the Restricted Boltzmann machine graph .In that graph all visible units are joined to all hidden units, and there are visible-visible or hidden-hidden connections not available [8].

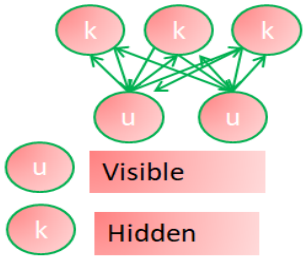

Fig 4. Resticted Boltzmann graph

In an Restricted Boltzmann Machine, the joint distribution B(u,k;θ) over the visible units u and hidden units k, model parameters i ,is defined in the term of an energy function F(u,k;θ) of

$$B(u,k;\theta) = \frac{\exp(-F(u,k;\theta))}{k} \tag{8}$$

Normalization

$$k = \sum_u \sum_k \exp(-F(u,k;\theta)) \tag{9}$$

$$B(u;\theta) = \sum_k \frac{\exp(-F(u,k;\theta))}{k} \tag{10}$$

Energy function

$$F(u,k;\theta) = -\sum_{s=1}^{S}\sum_{r=1}^{R} N_{sr}u_s k_r - \sum_{s=1}^{S} l_s u_s - \sum_{r=1}^{R} m_r k_r \tag{11}$$

Where $\mathbf{N_{sr}}$ represents the symmetric communication term between visible units u and hidden units k, l_s and m_r are the bias terms. m,n are the number of hidden and visible units.

$$B(k_r = 1 \mid u;\theta) = \sigma(\sum_{s=1}^{S} N_{sr}u_s + m_r) \tag{12}$$

$$B(u_r \mid k;\theta) = \sigma(\sum_{r=1}^{R} N_{sr}k_r + l_s, 1) \tag{13}$$

Gaussian distributed mean

$$\sum_{r=1}^{R} N_{sr}k_r + l_s \tag{14}$$

Gradient of the log likelihood **log B(u, θ)**

$$\Delta N_{sr} = F_{data}(u_s k_r) - F_{model}(u_s k_r) \quad (15)$$

Where $F_{data}(u_s k_r)$ observed in the training set F_{model} $(u_s k_r)$ expectation ubder the distribution model.

Gradient ascent training set c

$$\partial \log B(c) = \sum_{y \in c} \frac{\partial \log F(y)}{\partial N_{sr}} \quad (16)$$

$$F_{data}\left[\frac{\partial F(y,\theta)}{\partial N_{sr}}\right] - F_{model}\left[\frac{\partial F(i,\theta)}{\partial N_{sr}}\right] \quad (17)$$

Where the first term is the expectation of $\frac{\partial F(y,\theta)}{\partial N_{sr}}$

when the input variables are set to an input vector y and according to the conditional distribution B (K|y) the hidden units are sampled. The second term $\frac{\partial F(i,\theta)}{\partial N_{sr}}$ shows an expectation of when i and o are sampled similarly to the joint distribution of the RBM.

D. DEEP BELIEF NETWORK

The Stacking a number of the Restricted Boltzmann Machines (RBMs) learned layer by layer from bottom-up gives rise to a Deep Belief Network[9]. The Deep Belife Network (DBN) is consisting of many layer neural networks made up of several stacked Restricted Boltzmann Machines.For the structure blocks of, the Deep Belief Network, and Restricted Boltzmann machine, contains layer of visible units u, and a layer of hidden units k, connected by symmetrically weighted connections[10].Fig5. shows the Deep Belife Network Assuming binary units, the RBM defines the energy of joint configuration of the visible and hidden units (u, k) as

$$F(u,k;\theta) = -\sum_{s=1}^{S}\sum_{r=1}^{R} N_{sr} u_s k_r - \sum_{s=1}^{S} l_s u_s - \sum_{r=1}^{R} m_r k_r \quad (18)$$

Where $\mathbf{N_{sr}}$ represents the symmetric communication term between visible units u and hidden units k, l_s and m_r are the bias terms. M,n are the number of hidden and visible units[61].

$$B(u,k) = e^{-F(u,k)} / \sum_u \sum_k e^{-F(u,k)} \quad (19)$$

Visible layer vector u the probability hidden node k_r is calculated

$$B(k_r = 1/u) = \sigma\left[l_r + \sum_{s=1}^{S} u_s N_{rs}\right] \quad (20)$$

Where σ (y)=1/(1+exp(y)) visible layer node u_r is activated given hidden layer k.

$$B(k_r = 1/u) = \sigma\left[m_s + \sum_{r=1}^{R} k_r N_{rs}\right] \quad (21)$$

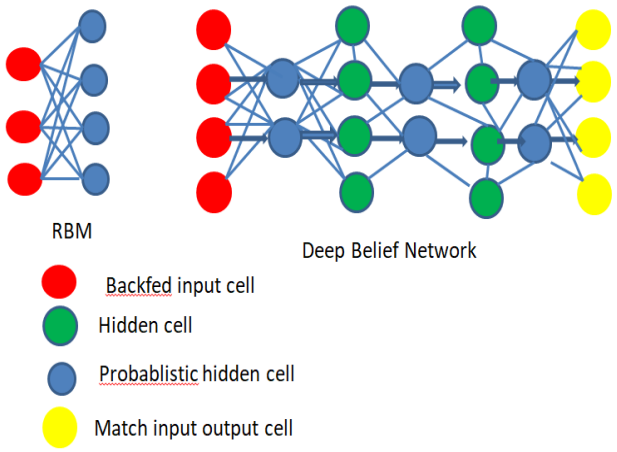

Fig5. Deep Belife Network

Given the training data on the visible nodes u_s and hidden k_r sample equation (20), is called as positive phase.A reconstruction of the visible node u_s is a sample equation (21) is called as negative phase.

Contrastive Divergence algorithm

$$\Delta N_{sr} = \in((uk_r) - (u_s' k_r')) \quad (22)$$

$$\Delta m_s = \in((u_s) - (u_s')) \quad (23)$$

$$\Delta l_r = \in((k_r) - (k_r')) \quad (24)$$

Where $\in$ denote learning rate and f(.) refers to the expectation of the tested states. The DBN takes a layer wise learning approach, in which RBMs are individually trained one after another in a bottom up fashion. For supervised classification, a softmax neuron network is to be found on the top of the last RBM as a multiclass classifier[11]. As a result, the softmax classifier learns a joint approach of the features extracted by the RBMs and the corresponding label of the samples. Once training is done data pass from the lowest level visible layer through multi BMS output layer.

E. Deep CNN models

1) AlexNet

AlexNet, stated that by A. Krizhevsky et al. [12] is a innovative deep convolutional neural network architecture and a winning model in the 2012 ImageNet Huge ScaleVisual Recognition Task. AlexNet consists of five convolution layers. The first layer, second layer and fifth with pooling layers, and three layers consist fully-connected layers. The convolution layer is implemented by the input which is convoluted with a set of filters. The non-linear function is

used to generate a group of feature. The max-pooling layer is used to maximal values of spatially successive local regions on the feature maps to extract the output.

2) CaffeNet

The Convolutional Architecture for Fast Feature Embedding also called Caffe. Caffe provides a complete toolkit for training, testing, finetuning, and organizing models, with well-documented. At the same time, it's likely the quickest available implementation of these algorithms, making it instantly valuable for industrial deployment[13].

3) GoogleNet

Going Deeper with Convolutions also called GoogleNet[14]. i) by employing filters of different sizes at each layer, it recalls more accurate spatial data. ii) it significantly decreases the number of free parameters of the network, making it fewer disposed to overfitting and allowing it to be deeper. Inception modules are used inside of the all convolutions layer by using rectified linear activation.

4) VGGNet-16

VGGNet has two well-known architectures: VGGNet-16 and VGGNet-19. In this estimation[15], its used the former- one because of its simpler design and somewhat better performance. This model had three phases first phase 13 convolutional layers, second phase 5 pooling layers, and third phase 3 fully connected layers[15]. The VGGNet-16 CNN feature was also take out from the second fully connected layer to get a feature vector of 4,096 dimensions.

5) PlacesNet

PlacesNet, which was established by Zhou et al. [16], has a similar architecture with CaffeNet. The deep features from PlacesNet are more effective in recognizing natural scenes than deep features from convolution neural network trained on ImageNe[59]t. We will calculate PlacesNet to confirm whether it results in outstanding performance.

6) VGGNet

To estimate the performance of different deep convolution neural network framework and associate them on a common ground, N. Srivastava et al. [17] developed three convolution neural network models based on the Caffe toolkit, each of which discovers a not the same speed/accuracy trade-off[18]:

- ***VGG-F:*** The fast CNN framework is identical to AlexNet[42]. The major changes from AlexNet are the smaller number of filters and small stride in some convolutional layers.
- ***VGG-M:*** The medium CNN architecture is indentical to the one presented by Zeiler et al. [59]. A lesser number of filters in the 4th convolutional layer is discovered as complementary to the computational speed.
- ***VGG-S***: The slow CNN framework is a simplified version of the accurate model in the OverFeat model [19,20], which retains the first five convolutional layers of the six layers in the original accurate OverFeat model and has a lesser number of filters in the 5th layer.

7) Inception-v3

Inception-v3 is additional ImageNet-optimized framework [21]. It is established by Google and has a strong importance on making scaling to deep networks computationally efficient. The Inception model takes in 299 x 299 images for this approach.

8) ResNet-50

ResNet-50 framework is established by Microsoft Research using a structure that uses residual functions to which help to add considerable stability to deep networks[22].It evaluate 18-layer and 34-layer residual nets (ResNets). The baseline designs connections are added to each pair of 3X3 filters.ResNet decreases the top-1 error by, resulting from the effectively reduced training error.ResNet-50 is the 50-layer version of ResNet. ResNet uses 224 x 224 images for this framework.

III. LITERATURE GAP IN DEEP LEARNING FOR SATELLITE IMAGERY

Urban surroundings have been examined using other types of imagery information that have become available in recent times. O. A. Penatti et al[23] has proposed to use the same type of imagery from Google Street View to measure the relationship between urban appearance and quality of life measures such as perceived safety[24]. For this, they used hand-craft standard image features that are widely used in the computer vision community,and train a shallow machine learning classifier (a Support Vector Machine). In a parallel fashion, it has trained a convolutional neural network on ground-level Street View imagery paired with a crowd-sourced mechanism for collecting ground truth labels to predict subjective perceptions of urban environments.

F. Hu et al [25] has presented Deep Convolutional Neural Network model to perform multi-label classification of Amazon satellite images. This approach identifies the weather conditions and natural terrain features in the images as well as man-made developments. The Deep Residual Learning for image Recognition (ResNet) model is used to identify artisanal mines which makes it extremely useful to solving the overarching problem of illegal human activity in the Amazon rainforest. Using deep Convolutional Neural Network framework designed for the ImageNet Challenge combined with task-specific refining layers produces good results on multi-label classification of satellite images.

C. Tao et al , X. Yao et al [26,27] has proposed that Deep convolutional neural networks (DCNNs) are used to achieve state-of-the-art performance on many tasks, such as object detection and object recognition. This method used RIT-18 dataset which contains very-high resolution multispectral imagery (MSI) collected by an unmanned aircraft system.followed by capturing synthetic aerial images of the scene with a MSI sensor model. By using the synthetic data to initialize a DCNN for object recognition, and then combine the pre-trained DCNN with two different fully-convolutional semantic segmentation models using real MSI. The

disadvantage is RIT-18 is difficult to perform well because of the high-spatial variability and unbalanced class distribution.

K. He, X. Zhang et al [28] has stated the challenge of land use and land cover classification using remote sensing satellite images which is freely accessible Sentinel-2 satellite images of the Earth observation. It used the state-of-the-art deep Convolutional Neural Network (CNNs) on this novel dataset with its diff erent spectral bands. In this framework, the classification performance using single-band images as well as images based on common band combinations are evaluated with accuracy of 98.57%. Possible applications in the area of land use and land cover are change detection and improving geographical maps.

M. Papadomanolak et al[1] has presented deep-learning frameworks based on Convolutional Neural Networks for the accurate classification of multispectral remote sensing data. This models used SAT-4 and SAT-6 high resolution satellite multispectral datasets .

O.A.Penatti et al [27] has proposed an unsupervised deep feature extraction for remote sensing image classification. They suggest the use of layer-wise unsupervised pre-training coupled with a greedy algorithm for unsupervised learning of sparse features[55]. The average accuracy of this model is 74.34%, which is a decent accuracy for an unsupervised classifier. Pre-training methodologies for unsupervised deep networks is currently active and an essential.

Zhong Ma, [2] has stated that Satellite Imagery classification based on Deep Convolution Neural Network (DCNN). It describes the Inception module as the building block to construct our DCNN model to cope with the large-scale variance in the satellite images. It also used the genetic algorithm based hyper-parameters optimization technique to report this issue and efficiently examined a big group of DCNN models. This model used DeepSat dataset which is a standard database of satellite imagery classification. It includes two divisions: SAT-4 and SAT-6. The proposed technique is guided towards better regions of the parameter space, and can find a better configuration within limited trails.

Dimitrios Marmanis[28] has stated Deep learning approaches such as convolutional neural networks (CNNs) can provide highly accurate classification results when delivered with large enough data sets and respective labels. Through this two-stage structure, effectively deal with the limited information tricky in an end-to-end processing pattern.Absolute results over the UC Merced Land use standard prove that our method significantly outperforms the previously greatest stated results, refining the complete accuracy up to 92.%.

Table 1: COMPARISON OF STATE-OF-ART

S.. No.	Method	Year	Approach	DataSet	Accuracy %
1	UFL(Unspervised Feature Learning)[51]	2014	Dense SIFT and Unspervised Feature Learning	STL-10 unlabeled dataset,CIFAR10,Cattech-101	81.6
2	COPD[52]	2014	Classification Based on Collection of Part Detectors	High –Spatial resolution colour image	91.33
3	FV(Fisher Vector)[53]	2014	HOG+RGB, Fisher Vector,Vectors of Locally Aggregated Descriptors(VLAD) and Vectors of Locally Aggregated Tensors(VLAT)	UC Merced Land Use dataset	93.80
4	UFL-SC[54]	2015	Unsupervised Feature learning with Spectral Clustering and Bag –of-Visual words	UC Merced Land –use dataset,WHU-RS dataset	90.26
5	CNN(Convolutional neural network)[40]	2015	Pre-trained ConvNet (CaffeNet)with SVM Classifier	UC-Merced dataset	93.42
6	CNN[57]	2015	Pre-trained ConvNet (GoogleNet)with fine-tuning of target data	UC-Merced dataset	97.10
7	DeConvolution Neural Network[46]	2016	Baseline classification approach	Sentinel-2 dataset and Hollstein pixel-level dataset	90.25
8	CNN[39]	2016	ResNet-50, ImageNet –optmized model,Baseline model	UC-Merced aerial image dataset	96.01
9	Deep-Learning Frameworks[1]	2016	AlexNet,AlexNet-small and VGG approach	SAT-4 and SAT-6 high resolution satellite multispectral dataset	99.9
10	Deep-Learning Frameworks[56]	2017	Deep Learning based large scale automatic satellite crosswalk classification	Satellite images using google static maps Cross walks-positive sample Zebra crossing-Negative samples	97.11

IV. QUANTITATIVE EVALUATION METRICS FOR DEEP LEARNING

a) Accuracy

Accuracy is the greatest intuitive performance measure and it is basically a ratio of correctly predicted observation to the total observations[48]. If the accuracy value is high then it is the best model. Accuracy is a great measure, but only when you have symmetric datasets where values of false positive and false negatives are almost. Jesse Davis et al [49] stated the accuracy as from total number of predictions number of correct predictions

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

where TP is True Positive, TN is True Negative,FN is False Negative,FP is False Positive.

b) Precision

Precision is the percentage of properly predicted positive observations to the total predicted positive observations[48]. Datasets are unbalanced in information retrieval.

$$\Pr ecision = \frac{TP}{TP + FP}$$

where TP is True Positive,FP is False Positive.

c) Recall (Sensitivity)

Recall is the percentage of correctly predicted positive observations to the all observations in the actual class[49].

$$\mathrm{Re}call = \frac{TP}{TP + FN}$$

CONCLUSION

An extensive comparative study has been given on various methods available in deep learning. Then, by analyzing the existing literature for deep learning, this paper discusses some literature gaps to enable the research community to develop new data-driven algorithms. Finally this paper provides some state-of-art quantitative metrics such as accuracy, precision, recall and F1 score for evaluating the satellite Imagery scene classification. All the additional information will be very much useful for image classification and recognition of satellite Imagery as it could help the users to learn important feature representations. Consequently, in the future we need to explore new data-driven algorithms and deploy it to promote the state-of-art of satellite Imagery scene classification.

REFERENCES

[1] Y. Yang and S. Newsam (2010) Bag-of-visual-words and spatial extensions for land-use classification.in Proc. ACM SIGSPATIAL Int. Conf. Adv. Geogr. Inform. Syst:270-279.

[2] Y. Wang, L. Zhang, X. Tong, L. Zhang, Z. Zhang, H. Liu, X. Xing, and P. T. Mathiopoulos (2016) A Three-Layered Graph-Based Learning Approach for Remote Sensing Image Retrieval. IEEE Trans. Geosci. Remote Sens.:54(10),6020-6034.

[3] L. Gueguen (2015)Classifying compound structures in satellite images: A compressed representation for fast queries. IEEE Trans. Geosci. Remote Sens:53(4),1803-1818.

[4] G. Sheng, W. Yang, T. Xu, and H. Sun (2012)High-resolution satellite scene classification using a sparse coding based multiple feature combination.int. J. Remote Sens:33(8), 2395-2412.

[5] B. Zhao, Y. Zhong, L. Zhang, and B. Huan (2016)The Fisher Kernel coding framework for high spatial resolution scene classification.Remote Sensing:8(2), 157.

[6] B. Zhao, Y. Zhong, G.-S. Xia, and L. Zhang (2016)Dirichlet-derived multiple topic scene classification model for high spatial resolution remote sensing imagery. IEEE Trans. Geosci. Remote Sens:54(4),2108-2123.

[7] H. Wu, B. Liu, W. Su, W. Zhang, and J. Sun (2016) Hierarchical Coding Vectors for Scene Level Land-Use Classification. Remote Sensing:8(5), 436.

[8] Q. Zou, L. Ni, T. Zhang, and Q. Wang (2015) Deep learning based feature selection for remote sensing scene classification.IEEE Geosci. Remote Sens:12(11), 2321-2325.

[9] 9.Powers,D.M.W (2011)Evaluation: From Precision, Recall And F-Measure To Roc, Informedness, Markedness & Correlation.Journal Of Machine Learning Technologies:2(1),37-63.

[10] Jesse Davis,Mark Goadrich(2006) The Relationship Between Precision Recall And ROC Curves. ICML '06 proceddings of the 23rd International Conference on Machine Learning :233-240.

[11] C. Farabet, C. Couprie, L. Najman, Y. LeCun (2013)Learning hierarchical features for scene labelling.IEEE transactions on pattern analysis and machine intelligence:35 (8),1915–1929.

[12] A. M. Cheriyadat (2014)Unsupervised feature learning for aerial scene classification.IEEE Trans. Geosci. Remote Sens:52(1), 439–451.

[13] D. Koester, B. Lunt, and R. Stiefelhagen (2016) Zebra crossing detection from aerial imagery across countries.in International Conference on Computers Helping People with Special Needs:27–34.

[14] T. N. Sainath, B. Kingsbury, G. Saon, H. Soltau, A. Mohamed, G. E. Dahl, and B. Ramabhadran (2015) Deep convolutional neural networks for large-scale speech tasks. Neural Networks: 64, 39–48.

[15] Y. Bengi, A. Courville, and P. Vincent (2013) Representation learning: A review and new perspectives.IEEE Trans. Pattern Anal. Mach. Intell.:35(8),1798–1828.

[16] N. Kruger et al.(2013) Deep hierarchies in primate visual cortex what can we learn for computer vision.IEEETrans.PatternAnal.Mach.Intell.,vol.35, no. 8, pp. 1847–1871, Aug. 2013.

[17] W. Caesarendra, A. Widodo, B. S. Yang (2010) Application of relevance vector machine and logistic regression for machine degradation assessment.*Mechanical Systems and Signal Processing*:24(4), 1161-1171.

[18] C. J. Schuler, H. Christopher Burger, S. Harmeling, B. Scholkopf (2013) A machine learning approach for non-blind image deconvolution.*Proceedings of the IEEE Conference* on Computer Vision and Pattern Recognition:1067-1074.

[19] Chen,Y.;Zhao,X.;Jia,X (2015) Spectral—Spatial classification of hyper spectral data based on deep belief network. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens: 8, 2381–2392.

[20] M.E. Midhun, S.R. Nair, V.T.N. Prabhakar, and S.S. Kumar (2014) Deep model for classification of hyperspectral image using restricted boltzmann machine.in International Conference on Interdisciplinary Advances in Applied Computing (ICONIAAC):35:1–35:7.

[21] Gao, B., Montes, M. J., Davis, C. O. and Goetz, A. F. H. (2009) Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean. Remote Sensing Environ:113, 17– 24.

[22] Dong, C., Loy, C. C., He, K. and Tang, X.(2016) Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) :38(2), 295–307.

[23] L.hao,p.Tang, and L.Huo (2016) Feature significance-based multibag-of-visual-wordsmodel for remote sensing image scene classification.Appl Remote Sens:10(3),35004-35009.

[24] P.Sermanet ,D.Eigen,X.hand ,M.Mathieu (2014) Ovrfeat:intergrated recognition ,localization and detection using cnn.in proc.in :1-16.

[25] Zhou,B,Lapedriza (2014) A Learning Deep Learning Features for Scene Recongnition using database .neural network information processing system:8-14

[26] Wang, J. and Jebara, T. and Chan (2012)S.-F. Semi-supervised learning using greedy max- cut.Journal of Machine Learning Research :14(1), 771-800.

[27] Chapelle, O. and Sindhwani, V. and Keerthi, S. S. (2013)Optimization Techniques for Semi- Supervised Support Vector Machines, Journal of Machine Learning Research : 9, 203–233.

[28] Komal M. Sukre,Imdad A. Rizvi and Mahesh.\ (2015) M. Kadam DEEP LEARNING METHOD FOR SATELLITE IMAGE CLASSIFICATION:A LITERATURE REVIEW . International Journal of Technical Research and Applications: 111-116.

[29] Komal M. Sukre1, Imdad A. Rizvi2, M. M. Kadam3 (2017) Image Classification of High Resolution Satellite Imagery Using Deep Learning Approach. International Research Journal of Advanced Engineering and Science: Vo 2(1), 84-86.

[30] L. Liebel∗, M. K¨orner (2016) SINGLE-IMAGE SUPER RESOLUTION FOR MULTI SPECTRAL REMOTE SENSING DATA USING CONVOLUTIONAL NEURAL NETWORKS. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences:12-19.