International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 08 | Aug-2016 www.irjet.net p-ISSN: 2395-0072

A TUTORIAL ON CLASSIFICATION OF REMOTE SENSING DATA

[1]K.Radhika* and [2]Dr. S. Varadarajan

[1]Research Scholar, Dept. of ECE, JNTUK, Kakinada, AP, India.
[2] Secretary, APSCHE, AP, India.
[1]krkvarma15@gmail.com

Abstract: Image classification is a well known of the significant tools used to recognize and examine most sharp information in satellite images. In any remote sensing research, the decision-making way mainly rely on the efficiency of the classification process. There are disparate classification algorithms on the large satellite imagery: Multilayer perceptron back propagation neural network (MLP BPNN), Support vector machine (SVM), k-means, Cluster ensemble based (CEB) method, Unsupervised Deep Feature learning (UDFL), Semi-supervised Ensemble Projection (SSEP). We discussed different performance measures such as classification accuracy, root-mean-square error, kappa statistic, true positive rate, false positive rate, to know the performance of each classifier.

Keywords: Classification, Image, Satellite, kappa statistics and accuracy.

1. INTRODUCTION:

Image classification plays a significant role in the interpretation of remote sensing imagery. Classification is important which involves extracting interesting patterns representing knowledge from real-world databases. Different classification techniques are used to analyze remote sensing images and set up convenient classifications and patterns in the images for scientific research and commercial purposes. Remote sensing as a field of study has attracted a lot of research for greater than two decades. This research work has used various computer-assisted techniques for image examination, including artificial neural network, k-means, SVM, statistical machine learning and other analysis methods.

Remotely sensed imagery can be used for multiple applications including land-use monitoring, undercover operation missions, estimate of environmental figure, civil planning, aurora particle monitoring, growth regulation, soil inspection, and crop production assessment. The classification of this imagery is essentially an indispensable case of these applications.

The steps fascinated in Image classification are (1) The first step builds a classifier or model depicting a predefined collection of classes. During the training phase the classification method constructs a classifier by learning from the training dataset and its corresponding class label. (2) testing phase, estimates the classification accuracy of the model for classification via a test set independent of the training data set. Popular classification techniques can be consider to label samples for training the classifiers, where the classification accuracy seriously depends on the amount and the quality of the training samples. However, the annotations of a well known labeled samples are repeatedly time consuming and routinely impossible to develop in several real-world problems.

In order to enrich the information given as input to the classification algorithm, elective methods like active learning and semi supervised learning have been proposed, which strictly exploit labeled and unlabeled samples for training classifiers to improve classification accuracy. In the remote sensing, multiple efforts have been made to develop semi supervised methods to improve the classification performance are majorly based on support vector machines (SVMs). In addition to SVM-based techniques, graph-based methods have further been effectively applied to the semi supervised classification of remote sensing images. In graph-based methods, each sample spreads its label information to its neighbors until a global stable state is reached on the complete data set. SVM has been found to be encouraging when used for pattern classification problems. Applying the Support Vector approach to a peculiar practical problem involves resolving a plenty of questions based on the problem definition and the design involved in it.

2. METHODOLOGIES:

2.1. Multilayer Perceptron Back propagation Neural Network:

One fair and functional method of data classification from the artificial neural network domain is the multilayer perceptron neural network based on the back propagation algorithm. It is a neural network model that employs a supervised learning technique as it converts a group of input patterns into a group of predefined output patterns. A typical MLP BPNN model has a single input layer, at uttermost one hidden layer, and a single output layer. In the network architecture, the input units are connected in a feed forward fashion, by generally told of input layer units completely connected to the hidden layer units, and hidden units completely connected to the output layer units. The nodes in the input layer and in the hidden layer(s) are associated with weights, or varying connection strengths. Initially, earlier training, all weights and biases are selected randomly.

A BPNN exhibits two modes of operation: feed forward and back propagation. When a BPNN is cycled in feed forward fashion, the input data pattern is propagated forward to the output layer nodes across all the input-to-hidden and hidden-to-output weights for producing output. The network's actual (activation) output value is then compared by all of the target output, and the fault value is estimated for each node in the output layer. In back propagation fashion, the errors, as well as the learning procedure (updating the weights and biases), transmit in the backward direction starting from the output layer unit to the inner nodes. This process is continued plenty of times to reduce the root-mean-square error (RMSE) between the network‘s predicted and target values until all the training samples are processed or the stopping criterion has been reached.

2.2. Support Vector Machine: Support Vector Machine (SVM) is a classification and regression prediction tool that uses machine learning theory to maximize predictive accuracy at the same time automatically avoiding over-fit to the data. Support Vector machines uses hypothesis space of a linear functions in a large dimensional feature space, trained by all of a learning algorithm from optimization theory that implements a learning bias obtained from statistical learning theory. SVM performs better in terms of not over generalization when the neural networks might end up over generalizing easily.

The goals of SVM are separating the data with hyper plane and approach this to non-linear boundaries by kernel trick. The Kernel trick allows SVM to construct nonlinear boundaries. The selection of kernel function will judge the performance of SVM. SVM is powerful tool for approximating any training data and generalizes better on given datasets. The difficulty in terms of kernel affects the performance on new datasets.SVM supports parameters for controlling the complexity, SVM is a useful technique for data classification.

Neural Networks are easier to use than SVM, but sometimes unsatisfactory results are obtained. A classification task usually involves by the whole of training and testing data which consist of some data instances. Each instance in the training set contains one target values and several attributes. The idea of SVM is to produce a model which predicts target value of data instances in the testing set which are given only the attributes.

Choosing an appropriate kernel is the main challenge for a given task. There are standard choices such as a Gaussian or polynomial kernel that are the default options, but if these prove ineffective or if the inputs are discrete structures more elaborate kernels will be needed. For high dimensionality of the data normal classification approaches are not suitable for better performance, but Support Vector Machines can eliminate the pitfalls of very high dimensional representations.

The main advantages of SVM are the training is relatively easy and No local optimal, unlike in neural networks. It scales relatively well to high dimensional data. In SVM , there is a trade-off between classifier complexity and error can be controlled explicitly. The failure includes the need for a good kernel function.

2.3. K-Means: The k-means algorithm is a parametric technique i.e., the several clusters are known priori. The unsupervised k-means technique is used to know the prior view of the image site that will helps us to elect the first-class training area for SVM classification system. When the clusters are globular, k-means can produce tighter clusters than other clustering methods. However, certain types of lands have similar spectral responses which makes the constructed k-means cluster non globular (centres are not well defined).

This is for that cause a regularization process (using SVM) is needed to correct the k-means results in order to gain a suitable structure. Indeed, SVM is a kernel based approach that contains non linear transformation which deals by all of data (pixels) non-linearly separable. The desire of k-means segmentation is to gain a first knowledge about the satellite image regions (groups).

K-means inherently considers that the clusters are hyper-spherical or hyper-ellipsoidal in shape. k-nearest neighbor is an instance-based learning technique for classifying objects founded on the principle of closest training examples in the feature space. The k-NN classifier is based solely on the principle of learning by concrete illustration, specifically, by comparing a given test instance by all of training instances that are similar to it.

The training instances, are represented by n attributes. Each instance represents a data point within the n-D pattern space or feature space. Given a query-instance or unknown tuple, a k-NN classifier searches the feature space by all of the intent of finding the k training instances nearest to it. These newly found k training tuples are surely the k ―nearest neighbors of the given query-instance. The term ―nearest‖ can be defined in terms of Euclidean distance as the distance metric.

2.4. Unsupervised Deep Feature Learning: single-layer convolutional networks can extract powerful discriminative features only when the receptive field accounts for neighboring pixels and are selected when the classification requires steep resolution and detailed results. Deep neural networks are models that capture hierarchical representations of data. In the case of CNNs, weights are assigned locally, i.e., the alike weights are

applied at every location of the input. The weights wired to the alike output unit construct a filter.

CNN layers comprise the following:
1) a convolution of the input with a set of learnable filters to extract local features;
2) a point-wise nonlinearity, e.g., the logistic function, to support deep architectures to learn nonlinear representations of the input data; and
3) a pooling operator, which aggregates the statistics of the features at adjoining locations, to minimize the computational charge (by decrease the spatial size of the image), while providing a local translational invariance in the earlier extracted features. CNN architectures have a significant number of meta parameters.

The closely relevant ones may be the following:
1) the number of layers;
2) the number of outputs per layer;
3) the size of the filters, which is further called receptive field; and
4) the size and type of spatial pooling. Another applicable aspect is how to train such architectures.

Deep convolutional networks can be trained in a supervised way, e.g., by manner of standard back-propagation or in an unsupervised manner, by means of greedy layer wise pre-training. Unsupervised greedy layer wise pre-training has been effectively used to train deep CNN. Supervised methods usually demand a large amount of reliable labeled data, which is difficult to gain in remote sensing classification problems.

Therefore, in the case of MS and HS images, it is recommended to consider an unsupervised learning strategy given the typically few available labeled pixels per class. Patch-based training is the most routinely used approach to learn the convolutional layers' parameters by means of unsupervised criteria. It consists in using a set of randomly extracted patches from input images (or feature maps) to train each layer. After that, the layer weights are applied to each input location to gain output convolutional feature maps that will serve as input to the next layer. Greedy layer-wise unsupervised pre-training is based on the idea that a local (layer-wise) unsupervised criterion can be applied to pre-train the network's parameters, allowing the use of large amounts of unlabeled data. deep networks have been further trained in a purely unsupervised way.

2.5. Unsupervised Learning Criteria With Sparsity:

Sparsity is among the properties of a good feature representation. Sparsity can be defined in terms of population sparsity and lifetime sparsity. On one hand, population sparsity ensures simple representations of the data by allowing only a tiny subsets of outputs to be active at the alike time. On the other hand, lifetime sparsity controls the frequency of activation of each output throughout the data set, ensuring rare but valuable activation of each output. soft-threshold encoding is a popular choice, which involves a tunable meta parameter to control the desired degree of sparsity. If the network's depth increases, the dead outputs effect becomes more significant and impacts the performance.

To validate this approach, the following are the experimental pipeline procedure:
1) Extract random patches from raw images and normalize them for contrast and brightness.
2) Train a network in a patch-based way by means of an unsupervised criterion, in our case, EPLS with logistic activation.
3) Use the trained network parameters and an encoding strategy to retrieve sparse representations.
4) Pool the uppermost feature map into four quadrants via sum-pooling.
5) Feed the pooled features to a linear SVM classifier.

Including spatial information is essential in order to avoid poor performance in single-layer networks;
1) combining high numbers of output features and max-pooling steps in deep architectures is crucial to achieve fine results; and
2) adding new layers to the deep architecture improves the classification score substantially, until the repeated max-pooling steps heavily minimize the features spatial resolution and/or the number of parameters becomes too large, thus inducing a form of over-fitting.

2.6. Semi-supervised ensemble projection (SSEP):

Classification of satellite images when limited labeled data is available together with a large amount of unlabeled data simple semi-supervised classifier are not adequate, these can be solved by learning a high-level features, called semi-supervised ensemble projection (SSEP). The method is based on the idea of transferring a small precise training set to an ensemble of WT sets and projecting images onto these WT sets for the final features. To create an ensemble of diverse WT sets, a novel sampling algorithm is designed and performed in different feature spaces.

As data distributed in a high-dimensional feature space can be approximated by the covariance matrix of the data, GNA can be used to find neighbours in the sampling algorithm. Then discriminative classifiers are learned on these WT sets, and images are represented by concatenating their projection values on the WT sets. By doing so, images are represented with their affinities to a rich set of discovered image attributes for classification. Furthermore, the potential redundant information contained in SSEP and conduct experiments to verify this

assumption. With sparse coding, the classification performance can be further improved. SEEP is superior in feature leaning.

3. CLASSIFICATION PARAMETERS

1) A confusion matrix helpful in the assessment of the overall accuracy of classes and the kappa coefficient per class.

2) False Positive Rate (FPR): For a good classifier this will be close to zero.

$$FPR= FP/(FP+TN)$$

True Positive Rate (TPR): For a good classifier this will be closer to one.

$$TPR=TP/(TP + FN)$$

confusion matrix is used to calculate the overall accuracy and the Kappa coefficient. The accuracy assessment also includes the producer's and user's accuracy in order to evaluate the omission and commission errors for each class.

4. CONCLUSION: BPNN is a neural network which reduce Mean square error while classifying the classes. Even though it is considered that neural networks are easier but sometimes unsatisfactory results are obtained. The unsupervised k-means technique is used to get an initial view of the image site that will helps us to choose the best training area for SVM classification system. Certain types of lands have similar spectral responses which makes the constructed k-means cluster non-globular. This is why regularization processes i.e., SVM is needed to correct the k- means results. SVM is powerful to approximate any training data and generalizes better on given datasets.

SVM performs better in terms of not over generalization. The kernel trick allows SVM to form non linear boundaries. The complexity in terms of kernel affects the performance of SVM. In SVM training is easy than neural network. Supervised methods usually require a large amount of reliable labeled data, which is difficult to obtain in remote sensing classification problems. Unsupervised greedy layer wise pre-training has been successfully used to train deep CNN. Population sparsity ensures simple representations of the data by allowing only a small subsets of outputs to be active at the same time.

Lifetime sparsity controls the frequency of activation of each output throughout the data set, ensuring rare but high activation of each output. Classification of satellite images when limited labeled data is available together with a large amount of unlabeled data simple semi-supervised classifier are not adequate, these can be solved by learning a high-level features, called semi-supervised ensemble projection (SSEP). With sparse coding, the classification performance can be further improved. SEEP is superior in feature leaning.

REFERENCES:

[1] Biplab Banerjee, Francesca Bovolo, Avik Bhattacharya, Lorenzo Bruzzone, Subhasis Chaudhuri and B. Krishna Mohan, "A New Self-Training-Based Unsupervised Satellite Image Classification Technique Using Cluster Ensemble Strategy", IEEE Geoscience and Remote Sensing Letters, Vol. 12, No. 4, 741-745, April 2015.

[2] Minh-Tan Pham, Grégoire Mercier and Julien Michel, "Pointwise Graph-Based Local Texture Characterization for Very High Resolution Multispectral Image Classification" IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 8, No. 5, 1962- 1973, May 2015.

[3] Wen Yang, Xiaoshuang Yin, and Gui-Song Xia, "Learning High-Level Features for Satellite Image Classification With Limited Labeled Samples" IEEE Transactions On Geoscience And Remote Sensing, Vol. 53, No. 8, 4472-4482, August 2015.

[4] Adriana Romero, Carlo Gatta, and Gustau Camps-Valls, " Unsupervised Deep Feature Extraction For Remote Sensing Image Classification" IEEE Transactions on Geoscience and Remote Sensing, Vol. 54, No. 3, 1349-1362, March 2016.

[5] Akkacha Bekaddour, Abdelhafid Bessaid & Fethi Tarek Bendimerad, "Multi Spectral Satellite Image Ensembles Classification Combining K-Means, LVQ and SVM Classification Techniques" J Indian Soc Remote Sensing, Vol.43, Issue4, 671–686, 2015.

[6] Y.Murali Mohan Babu, M.V. Subramanyam & M.N.Giriprasad "A Modified BM3D Algorithm for SAR Image Despeckling", Procedia Computer Science (Elsevier), Volume 70, Issue 1, 69 - 75, December-2015.

[7] M. Tyagi, F. Bovolo, A. K. Mehra, S. Chaudhuri, And L. Bruzzone, "A Context-Sensitive Clustering Technique Based on Graph-Cut Initialization and Expectation–Maximization Algorithm," IEEE Geosci. Remote Sens. Lett., Vol. 5, No. 1, 21–25, Jan. 2008.

[8] Z.H. Zhou and W. Tang, "Clusterer Ensemble," Knowledge Based Systems, Vol. 19, No. 1, 77–83, March 2006.

[9] Polikar, R. Ensemble based systems in decision making IEEE Circuits and Systems Magazine, 6(3), 21–45, 2006.

[10] T.Blaschke "Object based image analysis for remote sensing", ISPRS Journal of Photogrammetry and Remote sensing, vol. 65, 2- 16, 2010.

[11] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," ISPRS J. Photogramm. emote Sens., vol. 66, no. 3, pp. 247–259, May 2011.

[12] L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive SVM for semisupervised classification of remote-sensing images," IEEE Trans.Geosci. Remote Sens., vol. 44, no. 11, pp. 3363–3373, Nov. 2006.