
Satellite Image Classification based on Ensemble Subspace Discriminant method using Random Subspace Algorithm

*¹K.Radhika, ²S.Varadarajan

¹ECE Dept., JNTUK, Kakinada, AP, ²APSCHE, AP, India.

Email: krkvarma15@gmail.com, varadasouri@gmail.com

Received: 06th October 2017 Accepted: 14th November 2017, Published: 31st December 2017

ABSTRACT:

Classification of images obtained from satellites is a challenge now-a-days in order to know the information contained by them. This requires the adaption of the present state-of-art classifiers. Correct information can be retrieved only with good classification method. But single classifier may not meet the requirement, so that ensemble techniques are preferable. For classification with three or more classes various methods are AdaboostM2, LPBoost, TotalBoost, RUSBoost, Subspace, Discriminant etc etc. It is possible to create an ensemble for classification by combining any of these. Out of these, we have considered random subspace (RS) and discriminant classifiers for ensemble. Here bagging method preferred. This ensemble can be used for weak learners also. To judge whether the classifier is giving correct information or not validation is important. To validate this ensemble technique we have considered 50,100,150 and 200 sample points, examined the accuracy and necessary parameters. The nature of the proposed method for the above said number of sample points is discussed.

Keywords: Classification; Satellite;Image; Random subspace; Ensemble.

I. INTRODUCTION

Classification is playing major role in the extraction of information contained by the satellite images. During this process of classification several problems may come across like accuracy, misclassification etc. There are several classifiers like Support vector machine, K nearest neighbors, neural network, maximum-likelihood [1-3]. But to improve accuracy and to reduce misclassification single classifier is not sufficient. To strengthen the classification efficiency two or more base classifiers need to combine. This is possible by ensemble of several classifiers [4-6].

The prediction capability of the ensemble classifier model can be increased by the attributes. These will support training process of classifier in case of uncertain tuples thereby correct prediction possible. With all these procedure, a high accuracy is possible by ensemble classifiers compared to single classifiers. Ensemble procedure consists of bagging, boosting. Bagging process will generate various training data.

With repeated sampling a new training sets can be derived from original sets. If we repeated N times, N number of new data sets will be generated. After this, boosting is required to enhance the quality of the ensemble model [7]. Random subspace is the majorly using ensemble strategy and this procedure is taken place in feature space. To get final data set majority voting is preferred. For boosting the algorithm used generally is AdaBoost. This will combine a set of weak classifiers into a strong classifier. Weak classifiers are nothing but whose accuracy is lesser than expected. Boosting algorithm will modify the weight instances to make the classifiers stronger. The high accurate classifiers will acquire high weights in the final model. Discriminant Analysis is better method of classification. This will produce more accurate and less complex models [8-9].

II. ENSEMBLE BASED CLASSIFIERS

This can be possible by combining several methods [10]. These are developed to increase the performance of the prediction. All the benefits of individual classifier can be reached with ensemble methodology. The basic steps in ensemble discriminant classifier are data sets which resembles original data set. With repeated training the required data sets can be obtained. This is possible with subspace methodology. One such method considered is Random Subspace. This will results effective training data sets using majority voting rule. After this, boosting is required to convert weak classifiers to strong classifiers. For this AdaBoost algorithm is preferred. These two processes are explained using algorithms.

Random Subspace Classifier

This is one kind of ensemble classifier. This will have several classifiers in a subspace of data feature space. The output of this classifier is depends on individual classifier results by majority voting rule. The classifiers that can be combined to get Random subspace classifier are linear classifier, nearest-neighbour classification, support vector machine etc. The advantage of Random Subspace is, this algorithm subspace the original data size. The obtained training objects are smaller for the original data but larger for the subspace data. The original data size is decreased, whereas the training object size remains same.

RS Algorithm

Random Subspace will build the one base classifier on each subset by taking samples of original feature set. The class label is assigned by ensemble using majority voting or averaging output probabilities. Let $f=\{x_1, \dots, x_n\}$ be the set of 'g' features. To construct an RS ensemble with 'q' classifiers, collect 'q' samples. Every feature subset is a subspace of K and a classifier is trained by base classifiers like, K-nearest neighbour Support Vector machine, Discriminant analysis. The final ensemble decision will be taken by majority vote. This indicates the ability of Random Subspace to handle huge data dimensionality. Because of this, the classifier under training can be trained effectively for small subspaces and the ratio of feature-to-instance is also highly increased. With this procedure there is no disturbance to get more accuracy. The parameters required by Random Subspace ensemble are (i) feature subset cardinality (ii) size of the ensemble.

Algorithm 1:

Input Original data
 For k=1 to K
 Select g(g < q) features from original data and construct new subset

$$\tilde{b} = (b_{i1}, b_{i2}, \dots, b_{ig}) \tag{1}$$

Fit a base algorithm U_m to the new subset
 End

Output ensemble model $U = \text{sign} \sum_{k=1}^K U_m(b)$
 (2)

Algorithm 2:

Input training data
 U_m is the base classifier, $m=1,2,\dots,M$
 Initialize the weight distribution of training data
 $w_i=1/n, i=1,2,\dots,n$
 For m= 1 to M
 Compute the misclassification error rate

$$\varepsilon_m = P(T_m(x_i) \neq y_i) \tag{3}$$

Calculate coefficient of $U_m, U = \text{sign} \sum_{k=1}^K U_m$
 (4)

$$a_m = \frac{1}{2} \log \frac{1 - \varepsilon_m}{\varepsilon_m} \tag{5}$$

Normalization constant is

$$Z_m = \sum_{i=1}^N w_i \exp(-a_m y_i U_m(x_i)) \tag{6}$$

Update the weight

$$w_i = \frac{w_i}{Z_m} \exp(-a_m y_i U_m(x_i)) \tag{7}$$

End

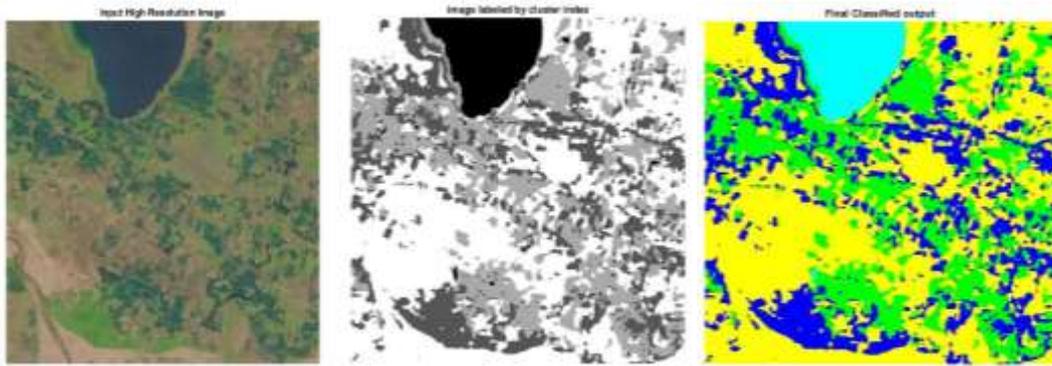
$$\text{Output } U(x) = \text{sign} \left(\sum_{m=1}^M a_m U_m(x) \right) \tag{8}$$

Algorithm 1 is Random Subspace algorithm and Algorithm 2 is AdaBoost algorithm for boosting. Using these algorithms, ensemble discriminant classifier is framed to get more accuracy than single base classifiers. This is efficient technique to extract the information in satellite images [11-12]

III. RESULTS AND DISCUSSIONS

The random subspace technique is preferred for the extraction of information in satellite images. This kind of method improves the accuracy. The classes defined here as A,B,C, and D. The input image is shown in figure1 (a).Figure 1(b) is the clustered output (SOM).The original image is of size 8000x8000 approximately, but this is resized to 512x512 for our convenience (Matlab). Figure 1(a) is a satellite input image. The entire procedure here is of two phases, one is clustering and another is classification. The clustered output is shown in figure 1(b). Figure 1(b) is now applied to ensemble classifier. The classified output is shown in figure 1(c). After getting the classified output we have verified with the ground truth points. Here our intension is how extent our ground truth points are correct. To check this,we have used the algorithm for 50,100,150 and 200 points. Generally for 8000x8000 pixels image the validation points considered are 2000 to4000 approximately. The best case is considering 2000 points for validation,in this case the percentage of pixels that are considered for validation is 0.003125 [(2000/(8000x8000))100=0.003125]. Our image is of size 512x512 pixels, we have checked this for 50,100,150 and 200 points. The percentage of pixels that are considered for validation is 0.019, 0.038, 0.057and 0.0763 respectively. So that the number of validation points considered are good for the size 512x512.

Table 1, is the confusion matrix for 50 validation points. Table 2 consists of quality measurements. In this table, parameters like precision,recall, specificity and f1score are calculated along with overall accuracy. Tables 3,5,7 are the confusion matrix of 100,150 and 200 validation points. Tables 4,6,8 are the different parameters along with overall accuracy. Observing all the tables,there no much change in overall accuracy of different number of validation points.



(a) (b) (c)

Figure1: Input image, clustered image and output image

Table1: Confusion matrix with 50 validation points

		PRIDICTED				
ACTUAL	CLASS	A	B	C	D	Total
	A	10	2	0	0	12
	B	1	13	0	0	14
	C	0	0	15	0	15
	D	0	0	0	9	9
	Total	11	15	15	9	50

Table2: Quality measurements with 50 validation points

	Accuracy	Precision	Recall	Specificity	F1 score
A	0.94	0.909091	0.833333	0.973684	0.869565
B		0.866667	0.928571	0.944444	0.896552
C		1	1	1	1
D		1	1	1	1
Over all		0.943939	0.940476	0.979532	0.941529

Table 3: Confusion matrix with 100 validation points

		PRIDICTED				
ACTUAL	CLASS	A	B	C	D	Total
	A	18	3	0	2	23
	B	2	34	0	0	36
	C	0	0	31	0	31
	D	0	0	0	10	10
	Total	20	37	31	12	100

Table 4: Quality measurements with 100 validation points

	Accuracy	Precision	Recall	Specificity	F1 score
A	0.93	0.9	0.782609	0.974026	0.837209
B		0.918919	0.944444	0.953125	0.931507
C		1	1	1	1
D		0.833333	1	0.977778	0.909091
Over all		0.913063	0.931763	0.976232	0.919452

Table 5: Confusion matrix with 150 validation points

		PRIDICTED				
ACTUAL	CLASS	A	B	C	D	Total
	A	26	6	0	1	33
	B	2	48	0	1	51
	C	0	0	44	0	44
	D	0	0	0	22	22
	Total	28	54	44	24	150

Table 6: Quality measurements with 150 validation points

	Accuracy	Precision	Recall	Specificity	F1 score
A	0.933333	0.928571	0.787879	0.982906	0.852459
B		0.888889	0.941176	0.939394	0.914286
C		1	1	1	1
D		0.916667	1	0.984375	0.956522
Over all		0.933532	0.932264	0.976669	0.930817

Table 7: Confusion matrix with 200 validation points

		PRIDICTED				
ACTUAL	CLASS	A	B	C	D	Total
	A	36	6	1	1	44
	B	3	64	0	1	68
	C	0	0	58	0	58
	D	0	0	0	30	30
	Total	39	70	59	32	200

Table 8: Quality measurements with 200 validation points

	Accuracy	Precision	Recall	Specificity	F1 score
A	0.94	0.923077	0.818182	0.980769	0.86747
B		0.914286	0.941176	0.954545	0.927536
C		0.983051	1	0.992958	0.991453
D		0.9375	1	0.988235	0.967742
Over all		0.939478	0.93984	0.979127	0.93855

IV. CONCLUSION

The algorithm is checked for 50,100,150 and 200 ground truth points and the corresponding overall accuracies and different quality parameters are calculated. The overall accuracy values are obtained are all similar almost in each case. The theme of usage of the different points was to check how far the values are accurate. All the four sets have given similar values in the accuracies and different quality parameters. Generally, with less number of validation points the judgement of a procedure is not correct. In this connection we have verified for different set of validation points. So, it can be concluded that our validation points are correct to use for further procedures.

REFERENCES

[1]S.Nithyakalyani,B.Gopinath.Analysis of Node Clustering Algorithms on Data Aggregation in Wireless Sensor Network.J SciInd Res,vol. 75, 38-42, 2016.

[2]S J. Lee,T R Lin Grace,P H. His.Industrial Cluster Development and its Contribution to Economic Growth in Taiwan - HsinchuScience and Industrial Park (HSIP).J SciIndRes, Vol.76, 273-278, 2017.

[3]K. Netti,Y. Radhika. A Model for Accurate Prediction in GeoRSS Data Using Naive Bayes Classifier.J SciInd Res, vol. 76, 473-476,2017

[4] SA. Mousavi, Z. Hanifeloo, P. Sumari, MRM. Arshad.Enhancing the Diagnosis of Corn Pests using Gabor Wavelet Features and SVM Classification.J SciInd Res, 75, 349-354, 2016.

- [5] Ch. Swapna, VV. Swetha Kumar, JVR. Murthy. A New Approach to Cluster Datasets without Prior Knowledge of Number of Clusters, *J SciInd Res*, vol. 74, 261-264, 2015.
- [6] S. Abdikan, F.B. Sanl, M. Ustuner, F. Calo. Land Cover Mapping using Sentinel-1 SAR Data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLI-B7, pp.12–19, 2016.
- [7] N. Alajlan, N. Ammour, Y. Bazi, H. Hichri. A cluster ensemble method for robust unsupervised classification of VHR remote sensing images. in *Proc. IEEE IGARSS*, pp. 2896–2899, IEEE, 2011.
- [8] G. Suresh, R. Gehrke, T. Wiatr, M. Hovenbitzer. Synthetic Aperture Radar Based Classifiers for Land Applications In Germany. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLI-B1, pp. 12–19, 2016.
- [9] M. Chi, Q. Qian, J. A. Benediktsson. A Survey on SAR Image Classification. *International Journal of Advanced Engineering and Global Technology*, vol. 3, Issue 12, pp. 1461–1465, 2015.
- [10] D. Chen, D. Stow, P. Gong. Examining the effect of spatial resolution and texture window size on classification accuracy: An urban environment case. *International Journal of Remote Sensing*, Vol. 25, Issue. 11, 2177–2192, 2004.
- [11] K. Bahirat, F. Bovolo, L. Bruzzone, S. Chaudhuri. A novel domain adaptation Bayesian classifier for updating land-cover maps with class differences in source and target domains. *IEEE Trans. Geosci. Remote Sens.*, vol. 50, Issue .7, pp. 2810–2826, Jul. 2012.
- [12] I.A. Rizvi, B.K. Mohan. Object-based image analysis of high- resolution satellite images using modified cloud basis function neural network and probabilistic relaxation labeling process. *IEEE Trans. Geosci. Remote Sens.*, vol. 49, Issue .12, pp. 4815–4820, Dec. 2011.