

Perspective Based Model for Constructing Diverse Ensemble Members in Multi-classifier Systems for Multi-spectral Image Classification

Laxmi Narayana Eeti and Krishna Mohan Buddhiraaju

Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay,
Powai, Mumbai-400076, India

{laxmi.narayana, bkmohan}@iitb.ac.in
<http://www.iitb.ac.in/en>

Abstract. This paper presents a perspective based model for creating diverse ensemble members in a multi-classifier system. With this technique different input feature sets are constructed using standard digital image processing and analyzing techniques viz. Haralick texture features, Gabor texture features, normalized difference vegetation index, standard deviation, spectral signatures, color spaces - CIELAB, HSV. These features are used as descriptors. Input feature sets are created as many as ensemble members. Input feature sets are discrete in nature because there is no common feature shared between any two input feature sets. Each one of these discrete input feature sets is utilized for training a particular ensemble member only. Each ensemble member would identify the classes independently and with completely different set of features. An empirical study for multi-spectral images shows that diverse and independent ensemble members can be constructed through our proposed technique. Results also show that proposed technique outperforms bagging in terms of individual member diversity and classification accuracy.

1 Introduction

In multi-classifier systems, diversity of ensemble members is an important aspect. Each ensemble member/predictor is expected to be independent and possess good generalization ability. By diversity what one means is, that each ensemble member commits error at different inputs and thereby differ in their output values.

Earlier attempts have been made to introduce diversity among ensemble members by random sampling of training data. Bagging [1] was one such technique where different training subsets are created by resampling. However, in this method there is possibility of creating training subsets which are identical or have maximum overlap of representative samples. In such a case correlation between different ensemble members might be high, which is undesirable. Addition of noise in training data to create non-overlapping training subsets was also addressed by [2].

AdaBoost algorithm [3] provides each ensemble member with different distribution of training data by alteration. The errors of previous network (predictor) are taken into account before recalculating training data distribution. Genetic algorithms were also used in selecting best feature subsets for ensemble members [4]. Selecting best group ensemble members through genetic algorithm was explored by [5]. With application to economic forecasting a novel method of input feature grouping was presented by [6]. In this technique all the input variables are aggregated using mutual information. Each ensemble member's input pattern is then formed by selecting a feature from each group. Input Decimation Ensembles [7] reduces correlation between ensemble members by using different subsets of input features. Correlation between category and each feature is taken into consideration for feature selection. In this method members could be specialists in single category or multiple categories depending upon the correlation between category and input feature. Pre-processing of features to change their representation is presented in [8]. These methods are called distortion methods. Techniques mentioned so far are mostly based on manipulation of input patterns. Diversification of ensemble members was also attempted through manipulation of member outputs [9][10].

This paper presents a technique for multi-spectral image classification where each ensemble member is trained in completely different feature space from the rest of the members. A number of different feature spaces are created through some of the available image analysis techniques. Textures, Normalized Difference Vegetation Index (NDVI), standard deviation, color spaces are used as feature descriptors. A number of different groups containing different feature spaces are formed. Each group is then utilized in the training of any one of the ensemble members only. It is expected that by training ensemble members with unique input feature pattern will bring diversity by default. Hence correlation among member outputs will be low. Unlike in methods where random selections are made, there is every possibility of redundant information and overlapping training data between two or more ensemble members, the selection of input feature/training patterns in our method is not random. In addition there is no need of manipulating original data.

Comparative results show that ensemble members constructed through proposed technique have higher difference of diversity values than those members constructed through bagging.

2 Perspective Model

In the real world there are many instances where it is realized that a certain phenomena or object can be identified in more than one perspective. Take for example, a coin. A person with good eye sight can recognize the denomination by merely reading the numerical value embossed on the coin while a visually impaired person can identify the denomination by feeling the texture with his fingers. Sometimes a third person can identify the denomination by looking at the shape and size of coin. So in three different perspectives the coin could

be identified correctly. It can be understood that these three different cases of identifying denomination were independent and diverse. There will also be cases where identification differs.

The same logic is tried here in this paper to construct diverse ensembles by training each of them with a single input feature set that is entirely different from other input feature sets. Doing so enables an ensemble member in learning a different way to identify an object. In this paper we tried to create members each of which is specialist in a specific way of identifying an object.

The different analyzing techniques utilized to create input feature sets are Haralick texture features [11], Gabor texture features [12], Normalized Difference Vegetation Index (NDVI), spectral values of original bands, standard deviation [13] and two color spaces [14][15] viz. Hue-Saturation-Value (HSV), Luminosity(L*)-Chromaticity (a*)-Chromaticity(b*) (CIELAB).

Four different groups of input feature set are formed as shown in Table 1. Here each ensemble is trained only in a particular feature space.

3 Empirical Study

Two multispectral satellite images are chosen as test images i.e. Data1 and Data2. Data1 is a 3-channel 258x410 pixels size and 0.82 m resolution image. Data2 is a 4-channel 800x800 pixels size and 0.6 m resolution image. Scene covered by the Data2 consists of seven categories viz. Built-up, Trees, Roadways, Shadows, Grass, Wetland and Ground. Their respective colors in classified images shown in Figure 1 are Dark blue, light blue, cyan, green, yellow, orange and red.

As combining results of one type of predictor on discrete feature sets is more effective than combining results of disparate predictors on one feature set [16], we adopted neural network for each ensemble member. A group of four neural networks were created. Each was trained by a specific input feature set.

As the test image consisted of three channels, a color coordinate transform (HSV and L*a*b*) was applied directly on this three channel image. Otherwise principal component transform would have to be utilized for arriving at the first three components. Standard deviation feature was computed for each channel as statistical texture descriptor [13]. Four Haralick texture features viz. contrast, correlation, energy/angular second moment, entropy were computed. Similarly two Gabor texture features viz. Gabor Energy and Gabor amplitude were computed.

Entire image and each channel is divided into 20x10 blocks (Data1) and 20x20 blocks (Data2) before processing them for texture features. For Data1, training pattern/input feature pattern is created using 1885 representative samples comprising of all categories. Testing data is of 360 samples. For Data2, number of representative samples are 12540 and testing samples are 6966. For bagging, all the different types of input features are grouped together and then representative samples are randomly picked with replacement to create feature sets for members.

Each neural network and its corresponding input feature training set is shown in Table 1. The evaluation of ensemble is done in terms of confusion matrices (Table 2-9) and in terms of member (neural net-work) diversity $d_i(x)$, member error $\epsilon_i(x)$, ensemble error $e(x)$ and ensemble generalization error (EGE). It was proved in [17] that the ensemble error can be divided into a term measuring the average generalization error of each individual member and a term called diversity that measures the disagreement among the members. Diversity of member, member error, ensemble error and ensemble generalization error are computed by Equation (1), (2), (3) and (4) respectively.

$$d_i(x) = [O_i(x) - EO(x)]^2 . \tag{1}$$

$$\epsilon_i(x) = [O_i(x) - f(x)]^2 . \tag{2}$$

$$e(x) = [EO(x) - f(x)]^2 . \tag{3}$$

$$EGE = \bar{E} - \bar{D} . \tag{4}$$

where x =input, $O_i(x)$ =output of member, $EO(x)$ =ensemble output, $f(x)$ =target value, \bar{E} = average of individual members' generalization error and \bar{D} = average of the diversity among members. The ensemble outputs are computed by the following Equation.

$$EO = \sum_{i \in N} O_i . \tag{5}$$

Some of the classified output images of Data2 are shown in Figure 1. Confusion matrices obtained through proposed technique on Data2 are shown in Table 2-5 and those obtained through bagging in Table 6-9. Evaluation statistics are shown in Table 10. All the average values are computed over all inputs from each category.

Table 1. Input feature set for each ensemble member for proposed technique

Ensemble Member	Input Feature Set
Member 1	Spectral Values
Member 2	L*a*b* + Standard Deviation
Member 3	NDVI + HSV
Member 4	Gabor texture + Haralick texture

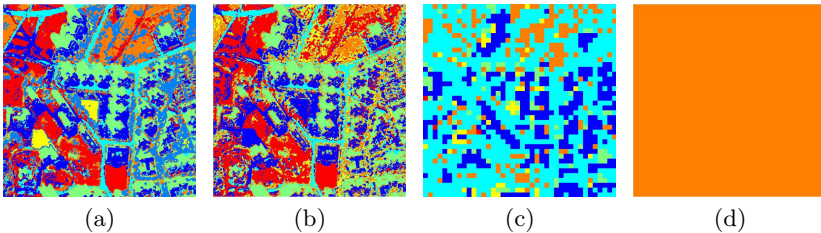


Fig. 1. Classified images (Data2):(a),(c)-perspective model; (b),(d)-bagging

Table 2. Confusion matrix by Member 1 through proposed technique

Category	Target Category							
	Built-up	Trees	Roadways	Shadows	Grass	Wetland	Ground	Total
Built-up	966	0	286	0	0	0	0	1252
Trees	0	1090	0	0	356	0	402	1848
Roadways	41	0	561	0	0	328	0	930
Shadows	1	3	0	944	0	0	0	948
Grass	0	2	0	0	393	0	0	395
Wetland	0	23	0	0	0	783	376	1182
Ground	0	0	0	0	0	10	398	411
Total	1011	1118	847	944	749	1121	1176	

Table 3. Confusion matrix by Member 2 through proposed technique

Category	Target Category							
	Built-up	Trees	Roadways	Shadows	Grass	Wetland	Ground	Total
Built-up	564	123	15	0	19	4	10	735
Trees	205	636	26	0	35	162	108	1172
Roadways	4	0	16	4	4	4	4	36
Shadows	42	0	72	553	59	12	140	752
Grass	0	9	3	127	556	3	3	701
Wetland	27	65	12	0	14	101	80	299
Ground	169	285	703	260	62	835	957	3271
Total	1011	1118	847	944	749	1121	1176	

Table 4. Confusion matrix by Member 3 through proposed technique

Category	Target Category							
	Built-up	Trees	Roadways	Shadows	Grass	Wetland	Ground	Total
Built-up	873	0	139	0	0	0	0	1012
Trees	3	1070	0	0	45	1	0	1119
Roadways	36	0	688	0	0	3	0	727
Shadows	3	0	0	943	0	0	0	946
Grass	0	7	0	0	704	0	0	711
Wetland	1	36	20	1	0	1061	32	1151
Ground	95	5	0	0	0	56	1144	1300
Total	1011	1118	847	944	749	1121	1176	

Table 5. Confusion matrix by Member 4 through proposed technique

Category	Target Category							
	Built-up	Trees	Roadways	Shadows	Grass	Wetland	Ground	Total
Built-up	942	169	0	137	0	10	0	1258
Trees	0	0	0	0	0	0	0	0
Roadways	6	827	549	164	29	306	830	2711
Shadows	0	44	4	402	210	0	104	764
Grass	45	0	70	0	420	0	0	535
Wetland	18	78	224	241	90	805	242	1698
Ground	0	0	0	0	0	0	0	0
Total	1011	1118	847	944	749	1121	1176	

In ensemble design it is required to include members that are characterized by high diversity and accuracy [17]. It can be observed from Figure 1 that classification results by a member constructed through perspective model outperforms those given by members constructed through bagging. For example, trees category are completely absent in Figure 1(b) while they can be seen in Figure 1(a). Similarly, Figure 1(d) shows that member constructed through bagging classified entire image into grass category. Figure 1(b) and Table 6 are result of diverse representative samples obtained through random sampling i.e. bagging. But

Table 6. Confusion matrix by Member 1 through bagging

Category	Target Category						
	Built-up	Trees	Roadways	Shadows	Grass	Wetland	Ground Total
Built-up	929	229	136	0	725	0	2019
Trees	0	0	0	0	0	0	0
Roadways	26	299	600	0	0	1	926
Shadows	0	6	0	932	0	12	950
Grass	10	249	94	0	0	346	699
Wetland	11	159	17	12	0	741	940
Ground	35	176	0	0	24	21	1176
Total	1011	1118	847	944	749	1121	1176

Table 7. Confusion matrix by Member 2 through bagging

Category	Target Category						
	Built-up	Trees	Roadways	Shadows	Grass	Wetland	Ground Total
Built-up	921	1118	623	441	464	462	4579
Trees	0	0	0	0	0	0	0
Roadways	0	0	0	0	0	0	0
Shadows	0	0	0	0	0	0	0
Grass	0	0	0	0	45	0	45
Wetland	0	0	0	0	0	0	0
Ground	90	0	224	503	240	659	2342
Total	1011	1118	847	944	749	1121	1176

Table 8. Confusion matrix by Member 3 through bagging

Category	Target Category						
	Built-up	Trees	Roadways	Shadows	Grass	Wetland	Ground Total
Built-up	0	0	0	0	0	0	0
Trees	18	837	198	24	42	98	1274
Roadways	0	0	0	0	0	0	0
Shadows	0	42	0	919	0	5	966
Grass	108	34	359	0	707	0	1574
Wetland	98	205	164	1	0	1018	2217
Ground	787	0	126	0	0	0	935
Total	1011	1118	847	944	749	1121	1176

Table 9. Confusion matrix by Member 4 through bagging

Category	Target Category						
	Built-up	Trees	Roadways	Shadows	Grass	Wetland	Ground Total
Built-up	0	0	0	0	0	0	0
Trees	0	0	0	0	0	0	0
Roadways	0	0	0	0	0	0	0
Shadows	0	0	0	0	0	0	0
Grass	1011	1118	847	944	749	1121	6966
Wetland	0	0	0	0	0	0	0
Ground	0	0	0	0	0	0	0
Total	1011	1118	847	944	749	1121	1176

Figure 1(d) is obtained because there is not enough diversity among representative samples.

In addition, referring to Table 2-10 it can be understood that individual member diversity is achieved by the proposed technique. As one can see in Table 10 there is considerable high difference between individual diversity values obtained by constructing members through perspective model (proposed).

High error figures by some members might be because of insufficient diversity among representative samples of different categories. In case of input feature sets based on texture, regions adjacent to and at category boundaries have adverse

Table 10. Comparison between proposed technique and bagging in terms of Individual Member Diversity (D) and Quadratic Error (QE)

Ensemble Member	Data1				Data2			
	Bagging		Proposed		Bagging		Proposed	
	D	QE	D	QE	D	QE	D	QE
Member 1	0.823	0.0597	0.356	0.0432	2.7146	0.7386	1.6322	0.3211
Member 2	0.779	0.2321	0.558	0.0582	1.9827	0.3954	1.890	0.493
Member 3	0.418	0.721	0.408	0.0481	1.8191	0.3063	3.023	0.604
Member 4	0.703	0.8109	0.825	0.591	2.6947	0.5054	2.635	0.385
Ensemnle Error	0.601		0.402		5.3022		3.007	
EGE	4.933		3.012		5.0761		3.168	

effects while computing different texture property values in a block. This is the reason behind the dilation kind of phenomenon observed in Figure 1(c). Observe that small categories such as roadways (cyan color) dilated in Figure 1(c). The block size plays the crucial role in texture based analysis of image.

4 Conclusions

An empirical study on two different multi-spectral satellite images has demonstrated that the proposed technique- Perspective based model- is successful in constructing diverse ensemble members and produce accurate members. It also demonstrated that the proposed technique in comparison with bagging technique performed better. Besides these, it also demonstrated that manipulation of training data or output data are not always necessary to construct diverse ensemble members.

References

1. Breiman, L.: Bagging predictors. *Machine Learning* 24(2), 123–140 (1996)
2. Raviv, Y., Intrator, N.: Bootstrapping with noise: An effective regularization technique. *Connection Science* 8, 355–372 (1996)
3. Freund, Y., Schapire, R.E.: Experiments with a new boosting algorithm. In: *Proceedings of the 13th International Conference on Machine Learning*, pp. 148–156. Morgan Kaufmann (1996)
4. Guerra-Salcedo, C., Whitley, D.: Genetic Approach for Feature Selection for Ensemble Creation. In: Banzhaf, W., Daida, J., Eiben, A.E., Garzon, M.H., Honavar, V., Jakiela, M., Smith, R.E. (eds.) *GECCO 1999: Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 236–243. Morgan Kaufmann, San Francisco (1999a)
5. Opitz, D., Shavlik, J.: Generating accurate and diverse members of a neural-network ensemble. In: Touretzky, D., Mozer, M., Hasselmo, M. (eds.) *Advances in Neural Information Processing Systems* 8. MIT Press, Cambridge (1996)
6. Liao, Y., Moody, J.: Constructing heterogeneous committees using input feature grouping. *Advances in Neural Information Processing Systems* 12, 921–927 (1999)
7. Oza, N.C., Tumer, K.: Input decimation ensembles: Decorrelation through dimensionality reduction. In: Kittler, J., Roli, F. (eds.) *MCS 2001. LNCS*, vol. 2096, pp. 238–247. Springer, Heidelberg (2001)

8. Sharkey, A.J.C., Sharkey, N.E., Gerecke, U., Chandroth, G.O.: The test and select approach to ensemble combination. In: Kittler, J., Roli, F. (eds.) MCS 2000. LNCS, vol. 1857, pp. 30–44. Springer, Heidelberg (2000)
9. Breiman, L.: Randomizing outputs to increase prediction accuracy, Technical Report 518, Statistics Department, University of California (1998), <http://www.boosting.org/papers/Bre98.pdf>
10. Dietterich, T., Bakiri, G., Error-correcting, G.: output codes: a general method for improv-ing multiclass inductive learning programs. In: Dean, T.L., McKeown, K. (eds.) Proceedings of the Ninth AAAI National Conference on Artificial Intelligence, pp. 572–577. AAAI Press, Menlo Park (1991)
11. Haralick, R., Shanmugam, M., Dinstein, K., Textural, I.: features for image classification. IEEE Transactions on Systems, Man and Cybernetics SMC-3(6), 610–621 (1973)
12. Manjunath, B.S., Ma, W.Y.: Texture features for browsing and retrieval of image data. IEEE Transactions on Pattern analysis and Machine Intelligence 18(8), 837–842 (1996)
13. Gonzalez, R., Woods, C.: Digital Image Processing, 2nd edn. Prentice Hall, Upper Saddle River 07458 665-667
14. Jain, A.K., Vailaya, A.: Image Retrieval using color and shape. Pattern Recognition 29(8), 1233–1244 (1996)
15. Cheng, H.-D., Sun, Y.: A hierarchical Approach to color image segmentation using homogeneity. IEEE Transactions on Image Processing 9(12) (December 2000)
16. Duin, R.P.W., Tax, D.M.J.: Experiments with classifier combining rules. In: Kittler, J., Roli, F. (eds.) MCS 2000. LNCS, vol. 1857, pp. 16–29. Springer, Heidelberg (2000)
17. Krogh, A., Vedelsby, J.: Neural network ensembles, cross validation, and active learning. In: Tesauro, G., Touretzky, D., Leen, T. (eds.) Advances in Neural Information Processing Systems, vol. 7. MIT Press, Cambridge (1995)