International Journal of Engineering &Technology, 7 (1.8) (2018) 6-9

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Satellite image classification and quality parameters using ML classifier

K. Radhika [1] *, S. Varadarajan [2]

[1] *Department of ECE, JNTUK, Kakinada, AP, India*
[2] *Secretary, APSCHE, AP, India*
**Corresponding author E-mail: krkvarma15@gmail.com*

Abstract

Remote sensing images are an important source of information regarding the Earth surface. For many applications like geology, urban planning, forest and land cover/land use, the underlying information from such images is needed. Extraction of this information is usually achieved through a classification process which is one of the most powerful tools in digital image processing. Good classifier is required to extract the information in satellite images. Latest methods used for classification of pixels in multispectral satellite images are supervised classifiers such as Support Vector Machines (SVM), k-Nearest Number (K-NN) and Maximum Likelihood (ML) classifier. SVM may be one-class SVM or multi-class SVM. K-NN is simple technique in high-dimensional feature space. In ML classifier, classification is based on the maximum likelihood of the pixel. The performance metrics for these classifiers are calculated and compared. Totally 200 points have been considered for validation purpose.

Keywords: *Classification; Classifier; Image; Multispectral; Satellite.*

1. Introduction

Classification of Multispectral satellite images is required to extract information related to earth surface. The methods used for the classification of pixels incase of multispectral images are Support Vector Machine, K-NN, Random forest etc. Support vector machine is a set of related supervised learning techniques applied for classification. There are many publications on mathematical formulation and algorithm development of the SVM [2], [3].

Support Vector Machine works on the concept of machine learning to increase the accuracy. A classification procedure typically includes training data and testing information which would Hosting a portion of information instances. The kernel selection plays major role to decide the performance of SVM on new datasets.SVM controls the complexity of kernel by introducing parameters. To set these parameters, Cross-Validation is used.

The traditional classification approaches perform weakly on high dimensionality of the data, but Support Vector Machines can eliminate the pitfalls of very large dimensional representations. The Training phase is easy in SVM. The 4-most important features of SVM are duality, kernels, convexity and sparseness.

To classify objects k-nearest neighbor [4] uses an instance-based learning technique that works on the principle of closest training examples in the feature space. It is quite simple technique. Unknown sample can be classified correctly by comparing distance with respect to known samples. So, unknown samples are identified based on nearest neighbors.

Decision tree represent another kind of classification algorithms. Decision trees are simple in handling the non-normal, non-homogeneous and noisy data [5]. Decision tree classifiers are more efficient than single-stage classifiers. Decisions can be made at multiple levels in case of decision tree classifier [6].

2. Classification techniques

2.1. Support vector machines

The principle employed in SVM [7] is Structural Risk Minimization. SVM is a very good tool for regression and classification problems [8]. It exhibits good generalization performance. The SVM is a linear machine that increases the margin by developing a model for transforming low dimensional feature space to high dimensional feature space. SVM have applications in text categorization, time series analysis, database mining and face identification.

SVM is a method for the classification of both linear and nonlinear data. It uses a nonlinear mapping to transform the original data into higher dimension. Within this dimension, it searches for the linear optimal separating hyperplane.SVM finds this hyper plane using support vectors or margins. Consider an example based on two input attributes of two class SVM, A1 and A2 as shown in figure 1.

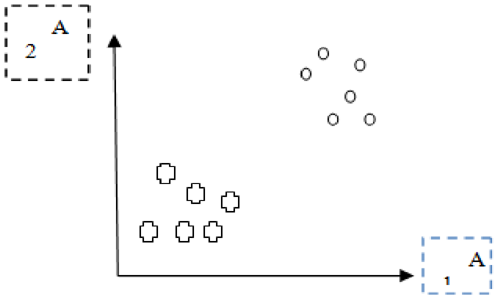

Fig. 1: The 2-D Training Data.

In multi-class classification more than one hyperplane is required to separate data. In case of linear classification, each hyper plane Hi separates the classes Ci from the tuples of all other classes as shown in figure 2. To achieve this, the classes should be linearly separable. If the classes are not linearly separable as shown in figure 3 a complex nonlinear data broken into smaller sets.

Generally used multi-class classifiers are, one-against-rest SVM consists of K two-class classifiers to obtain K-class classification. When the number of classes is large, classification is effected by the unbalance of the training samples.It also suffers from the problem of large computation and error accumulation. Hierarchical SVM use (K-I) two-class classifiers to obtain a K-class classification.

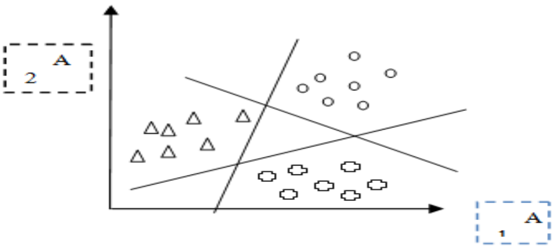

Fig. 2: Hyper Planes for Linearly Separable Data.

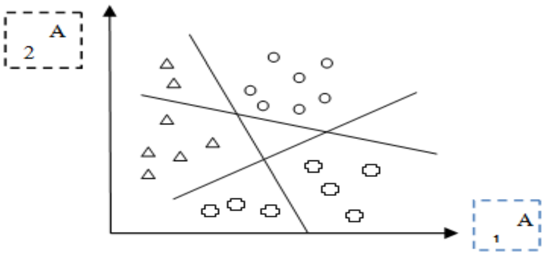

Fig. 3: Hyper Planes for Linearly Inseparable Data.

In this method, there is no sub-regional and the whole number of classifiers is less than one-against-one SVM. This is advantageous. Directed acyclic graph (DAG) constructs a multi-class classifier by few two-class classifiers having the similar training processing with one against-one SVM. DAG consists K(K-l)/2 two-class classifiers, which utilizes the guidance from the root node acyclic graph as shown in Figure. 4. In this case, error depends on the class number K and the intervals of the node, and it has no relationship with the dimension of the input space.

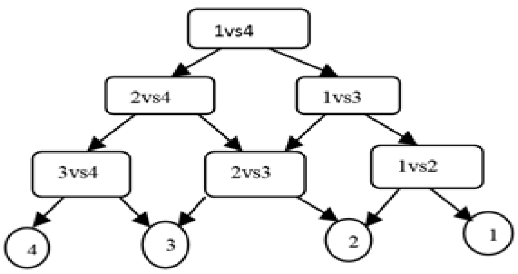

Fig. 4: Directed Acyclic Graph.

2.2. K-Nearest neighbors

K-Nearest Neighbor uses the principle of closest training tuples in the feature space to classify the objects. The training tuples are indicated with n-attributes. A tuple may be a point in a n-dimensional space. In case of an unknown tuple, k-Nearest Neighbor classifier searches in the space for a tuple which will be closest of the unknown tuple. The metric used here is Euclidean distance. The Euclidean distance between two tuples, say, X1=(x11, x12... x1n) and X2 =(x21, x22... x2n), is

$$dist(x_1, x_2) = \sqrt{\sum_{i=1}^{n} (x_{1i} - x_{2i})^2} \quad (1)$$

2.3. Maximum likelihood (ML) classifier

This is useful for classifying the under lying data in remote sensing images, where classification is based on the maximum likelihood of the pixel. For kth class, the posterior probability of a pixel is denoted by Yk.

$$Y_k = P(k/X) = P(k)*P(X/k) / \Sigma P(i)*P(X/i) \quad (2)$$

Where P (k) is the prior probability of class k P(X/k) is the conditional probability to observe X from class k

For normal distributions, the likelihood can be expressed as follows.

$$Y_k(X) = \frac{1}{(2\pi)\frac{n}{2}|\Sigma_k|^{\frac{1}{2}}} \exp\left|-\frac{1}{2}(X-\mu_k)\Sigma_k^{-1}(X-\mu_k)^k\right| \quad (3)$$

Where,

'n' is the number of bands, 'X' is the image data of 'n' bands 'Yk (X)' is the likelihood of X belonging to class k;'μk' is the mean vector of class k;

∑k is the variance-covariance matrix of class k; |∑k| is the determinant of ∑k;

For symmetric variance-covariance matrix, the Yk is the same as the Euclidian distance, whereas when the determinants are same, the Yk becomes the same as the Mahalanobis distances.

3. Results and conclusion

In this paper different state of art classifiers have been explained. The propose method is validated with 200 points of ground truth data. The classified image and ground truth data has been verified in validation process [9]. The ground truth points are shown in table 3. The confusion matrix and accuracy values have been measured for different methods and ML classifier outputs have been given in table 1 and table 2. Table 4 shows all possible quality parameters that can be evaluated with given classifier. A, B, C and D are the four different classes (shown in table 1) possible in the satellite image. Table 2 represents both user and producer's accuracy along with over accuracy.

Landsat-8 data is considered for analysis. 4 classes have been considered in the confusion matrix. Figure 5, Figure6, Figure7 and Figure8 show the comparison of quality parameters for different methods. More number of ground truth data points give correct values also [10]. That is reason to consider 200 points. More number of points lead to accurate measurement. It may be extended to even more number of ground truth values. But it is very difficult to make more ground truth points. The only drawback is that it is not that much of easy to take ground truth data values as said earlier. Validation process gave confusion matrix of four or more classes. This paper contains the explanation of four class classification using ML classification model.

Table 1: Confusion Matrix for ML Classifier

	PRIDICTED					
	CLASS	A	B	C	D	Total
ACTUAL	A	41	1	2	1	45
	B	4	44	3	2	53
	C	2	4	40	4	50
	D	0	5	1	46	52
	Total	47	54	46	53	200

Table 2: Accuracy Parameters for ML Classifier

Type of land cover	Reference Pixels	Classified Pixels	Matching	Accuracy type	
				Procedures	Users
A	47	45	41	87.23%	91.11%
B	54	53	44	81.48%	83.02%
C	46	50	40	86.96%	80.00%
D	53	52	46	86.79%	88.46%
Total	200	200	171		
Overall Classification Accuracy				85.5	

Table 3: 200 Points of Ground Truth Values Used for Validation

.SNo	X	Y	C	S.No	X	Y	C	S.No	X	Y	C	S.No	X	Y	C
1	247	440	1	51	19	138	1	101	147	84	2	151	426	136	3
2	408	389	1	52	39	157	1	102	172	111	2	152	280	317	3
3	417	230	1	53	174	171	1	103	166	89	2	153	317	271	3
4	38	473	1	54	173	173	1	104	103	46	2	154	342	269	3
5	47	479	1	55	175	171	1	105	135	41	2	155	380	231	3
6	456	246	1	56	136	236	1	106	145	56	2	156	486	128	3
7	323	261	1	57	101	305	1	107	145	65	2	157	335	226	3
8	31	477	1	58	89	382	1	108	151	54	2	158	60	84	3
9	198	491	1	59	87	438	1	109	107	80	2	159	63	87	3
10	246	359	1	60	91	437	1	110	168	85	2	160	248	312	3
11	380	141	1	61	35	465	1	111	121	108	2	161	426	180	3
12	155	183	2	62	40	472	1	112	146	164	2	162	339	228	3
13	117	169	2	63	43	475	1	113	120	143	2	163	224	315	4
14	160	65	2	64	45	474	1	114	186	69	2	164	223	306	4
15	154	61	2	65	48	480	1	115	167	83	2	165	215	291	4
16	124	183	2	66	44	482	1	116	61	186	2	166	239	277	4
17	168	87	2	67	40	476	1	117	139	64	2	167	254	285	4
18	166	90	2	68	38	476	1	118	178	298	3	168	251	273	4
19	109	51	2	69	30	476	1	119	182	298	3	169	300	219	4
20	62	190	2	70	25	480	1	120	183	300	3	170	301	212	4
21	186	70	2	71	29	483	1	121	177	301	3	171	230	204	4
22	65	184	2	72	38	477	1	122	181	298	3	172	231	210	4
23	128	139	2	73	33	479	1	123	180	306	3	173	294	206	4
24	378	87	3	74	180	478	1	124	181	304	3	174	351	238	4
25	425	27	3	75	391	442	1	125	202	338	3	175	373	141	4
26	57	81	3	76	423	256	1	126	207	338	3	176	377	130	4
27	25	450	3	77	197	306	1	127	204	338	3	177	389	115	4
28	145	450	3	78	204	388	1	128	206	342	3	178	419	177	4
29	161	444	3	79	209	387	1	129	245	314	3	179	299	123	4
30	425	15	3	80	250	355	1	130	280	327	3	180	310	118	4
31	61	79	3	81	269	336	1	131	247	343	3	181	297	115	4
32	25	447	3	82	183	336	1	132	253	339	3	182	90	288	4
33	413	135	3	83	327	273	1	133	202	372	3	183	91	320	4
34	236	314	3	84	334	243	1	134	209	381	3	184	115	245	4
35	370	227	3	85	344	315	1	135	327	213	3	185	312	100	4
36	25	449	3	86	380	146	1	136	313	210	3	186	449	105	4
37	124	454	3	87	389	127	1	137	312	202	3	187	454	92	4
38	270	100	4	88	392	127	1	138	340	176	3	188	428	126	4
39	322	41	4	89	155	183	2	139	323	193	3	189	282	96	4
40	251	275	4	90	120	174	2	140	345	160	3	190	255	143	4
41	320	44	4	91	125	166	2	141	345	138	3	191	252	139	4
42	244	248	4	92	117	174	2	142	313	177	3	192	427	71	4
43	350	43	4	93	167	84	2	143	375	224	3	193	123	246	4
44	427	85	4	94	118	136	2	144	372	220	3	194	225	375	4
45	242	79	4	95	146	105	2	145	412	143	3	195	381	113	4
46	238	281	4	96	161	87	2	146	341	136	3	196	220	70	4
47	241	191	4	97	169	88	2	147	372	89	3	197	330	40	4
48	297	121	4	98	165	113	2	148	380	101	3	198	193	28	4
49	426	87	4	99	157	81	2	149	394	99	3	199	293	59	4
50	339	50	4	100	143	96	2	150	418	149	3	200	267	72	4

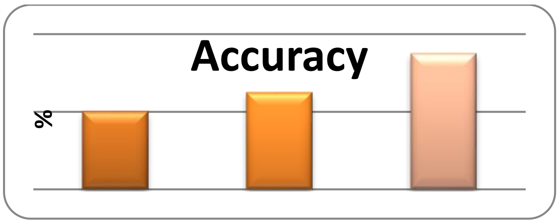

Fig. 5: Comparison of Accuracy Values.

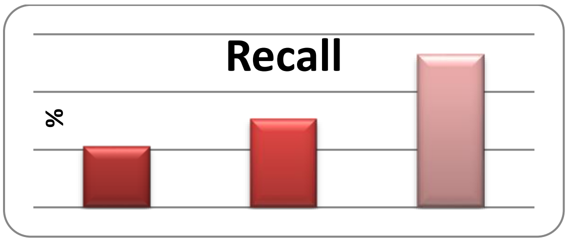

Fig. 6: Comparison of Accuracy Values.

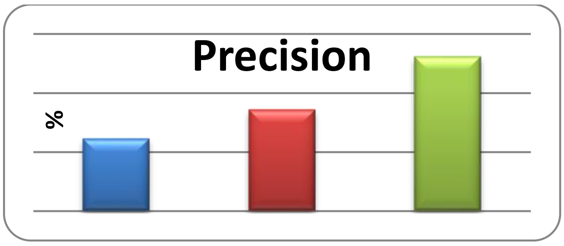

Fig. 7: Comparison of Accuracy Values.

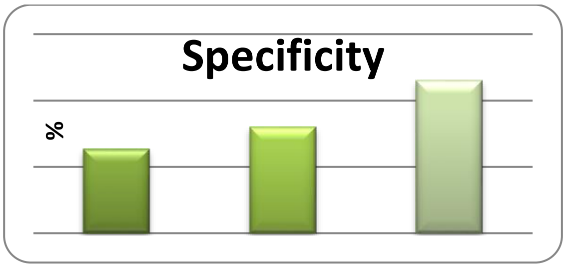

Fig. 8: Comparison of Accuracy Values.

The following are the familiar classification quality parameters that say about a model's efficiency. Especially the overall accuracy and kappa values will explore the quality of the model. The kappa value is more for a very good classification. It is obtained that over all kappa coefficient is 0.806. Individually third class (C) is having less efficiency in obtaining the information. But still it is very good for a classifier. Since it is greater than 0.7 (A classifier that is giving kappa greater than 0.7 is a very good classification model). Even other parameters also it is giving very useful values. The following values are the continuation of table 1 and table 2. Even it can be extended to more classes also. But the problem is only with validation points.

Table 4: Quality Parameters of Four Class ML Classifier

	Accuracy	Precision	Recall	Specificity	F1 score	Kappa
A		0.87234	0.911111	0.96129	0.891304	0.883
B		0.814815	0.830189	0.931973	0.82243	0.767
C	0.855	0.869565	0.8	0.96	0.833333	0.740
D		0.867925	0.884615	0.952703	0.87619	0.843
Over all		0.856161	0.856479	0.951491	0.855815	0.806

References

[1] Burges, C.J.C. (1988), A tutorial on support vector machines for pattern recognition, Data Mining and Knowledge Discovery, 2 (2), 121–167. https://doi.org/10.1023/A:1009715923555.

[2] Swapna, Ch., Swetha K V V and Murthy J V R. (2015). A new approach to cluster datasets without prior knowledge of number of clusters, J SciInd Res, 74, 261-264.

[3] Osornio Rios, R.A. (2017). Identification of positioning system for industrial applications using neural network, J SciInd Res, 76, 141-144.

[4] Bakshi, S. Thakur, T. and Khanna, R. (2017). Performance analysis of a novel hybrid optimization technique for sizing and placement of DG units, J SciInd Res, 76, 203-207.

[5] Bakshi S, Thakur T, Khanna R and Gupta A. (2017). DG placement and capacity calculation using an improved human opinion dynamics based optimization technique, J SciInd Res, 76, 427-432.

[6] Mousavi, S.A., Hanifeloo, Z., Sumari, P. and Arshad, M. R. M. (2016) Enhancing The Diagnosis of Corn Pests Using Gabor Wavelet Features and SVM classification, J SciInd Res, 75, 349-354.

[7] Sivakumar, S. and Venkatesan, R. (2016). Performance evaluation of hybrid evolutionary algorithms in minimizing localization error for wireless sensor networks, J SciInd Res, 75, 289-295.

[8] Huang, W. Y and Lippmann, R.P. (1988), Neural Net and Traditional Classifiers, in Neural Information Processing Systems, 387–396.

[9] Murali Mohan Babu, Y., Subramanyam, M.V., and Giriprasad, M.N. (2015). Fusion and Texture based Classification of Indian Microwave Data - A Comparative Study", International Journal of Applied Engineering Research, 10, 1003-1010.

[10] Murali Mohan Babu, Y.andRadhika, K. (2016).A new approach for microwave imagery de-noising, I J Image, Graph & Sig Proc, 52-60.