International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-5, January 2020

SVM Parameter Optimization using ALO for Object Based Land Cover Classification

K.Jayanthi, L.R. Sudha

Abstract*: Machine Learning algorithms are often used to solve various kinds of data classification task. Support Vector Machine (SVM) performs better for object oriented classification of high dimensional remote sensing datasets even with minimum training samples. In order to obtain improved performance in classification, the generalization and learning ability of SVM can be enriched by proper tuning of kernel and penalizing parameters of SVM. In this methodology ALO optimizer performs the optimal searching of SVM parameter in the direction of reducing misclassification rate. The proposed approach results better SVM parameters for the significant feature sub set which characterize the Landsat image objects of the study area. Performance of ALO is compared with GA based SVM parameter optimization. Accurate thematic classification map of land cover classes of the area of study also resulted in this module.

Keywords: SVM, Parameter tuning, Land cover classification, ALO algorithm, Object based classification.

I. INTRODUCTION

Applications like urban planning, environmental monitoring, land cover/land use etc. can be effectively analyzed based on the information extracted from the satellite images of the area of interest. These images also deliver detailed spatial, structural and textural information for interpretation regarding the state of land condition. It is quite common that the conventional "pixel-based" approaches simply concerned about the spectral characteristics of the image points that are not opt for better resolution images classification. They can be effectively performed by segmenting the image in to image objects and application of object-oriented classification. It is accepted as superior alternative for analyzing the high resolution land cover image with accurate objected oriented image classification.

Optimal selection of essential features to exhibit the image objects improves the classification accuracy and also reduces the complexity in classification process. It is well known that SVM has been proved as a capable pattern classifier for various classification tasks [1]. SVM works with significant feature set for achieving better classification performance. Parameter tuning of SVM enhances the generalization ability in the training stage for the image data set. Non-linear separable data can be effectively mapped using suitable kernel parameter like RBF kernel and penalizing parameters [2]. In the literature, empirical selection, grid search and soft computing optimization approaches are proposed for SVM parameter tuning [3]-[7].

Manuscript published on January 30, 2020.
* Correspondence Author
K.Jayanthi*, Assistant Professor, Department of Computer Application, Govt. Arts College, Chidambaram, India, jayanthirab@gmail.com
L.R.Sudha, Associate Professor**,** Department of Computer Science & Engineering, Annamalai University, Annamalainagar.

This paper attempts to search suitable SVM parameters with the objective of reducing misclassification rate based on ALO algorithm. Governing equations related SVM classifier are modeled for non-linear separable problem. Lagrangian multipliers are incorporated in the model to get dual of the objective. This model tries to improve the classification performance by achieving less misclassification among class samples.

The paper is structured as follows: a brief introduction to the SVM for multi-class problem and state-of-the-art of parameter tuning is presented. Section -2 formulates the important governing equations related to SVM classifier. The proposed ALO methodology for parameter tuning is discussed in Section-3. Suitability of the proposed ALO method is demonstrated in Section-4 and summary of the work is presented in Section-5.

II. PREAMBLE AND GOVERNING EQUATIONS OF SVM

SVM is initially formed to classify the linear two-class problem based on the statistical theory by Vapnik [1]. It is extended to multi-class problem by separating the input training vectors using hyper plane in the problem space with most possible margin between them. Land cover classification is a non-linear separable problem. SVM also addressing the non-linear separable classification problem with the help of kernel parameter mapping for generalizing the classifier performance. The governing equations of SVM model is presented in this section.

Let the training inputs be $\{(p_1, q_1), (p_2, q_2),\ldots(p_x, q_x)\}$ and $p \in R^d$ denotes the training samples related to the class $p_i \in [1,-1]$. Kernel parameter function of SVM maps the input space of the training vectors into high dimensional problem space. Among the popular kernel functions, Radial Basis Function (RBF) kernel performs mapping efficiently. It is modeled as

$$K_{RBF}(p_i, p_j) = \exp\left(-\lambda \left\|p_i - p_j\right\|^2\right) \tag{2.1}$$

Where λ is the kernel parameter that modifies mapping between input space and high dimensional problem space. Hyper plane distance in the decision problem space is given by

$$D(p) = w^T * p + b$$

(2.2)
Where w^T is weight vector of dimension x and b is the bias vector. The target mapping function based on reducing the structural risk for achieving proper hyper plane separation is formulated with minimizing objective is given by

$$Min\ \ Q_m(w,b,\varepsilon) = \frac{1}{2} * w^T * w + C * \sum_{i=1}^{m} \varepsilon_i \tag{2.3}$$

Subjected to satisfy the inequality

$$q_i(w^T p_i + b)\ \geq 1 - \varepsilon_i\ \ for\ i = 1 \ldots x \tag{2.4}$$

Where ε_i is a variable having non-negative numeric value expresses the misclassification in class 'i' and the constant C is a penalizing parameter for misclassification.
Target mapping function is represented using Lagrangian multiplier as

$$\min_{(w,b,\varepsilon)} \max_{a,b} \left\{ \frac{1}{2} w * w^T + C \sum_{i=1}^{m} \varepsilon_i - \right.$$
i=1maiqiw.pi−b−1+εi−i=1mbiεi

Subjected to $\quad a_i, b_i \geq 0$ (2.5)

Where a_i, b_i are the Lagrangian multipliers with non-negative numerical values.
Dual form of the target mapping function with Lagrangian multipliers is given by

$$\max_{a} \left\{ \sum_{i=1}^{m} a_i - \frac{1}{2} \sum_{i,j=1}^{m} a_i a_f q_i q_f\, K_{RBF}(p_i, p_j) \right\} \tag{2.6}$$

Subject to the equality, bound constraints

$$\sum_{i=1}^{m} a_i b_i = 0 \quad C \geq a_i \geq 0 \quad \text{for i} = 1, \ldots, \text{x} \tag{2.7}$$

Distance of the hyper plane in the decision problem space is achieved with optimum C and K_{RBF}

$$D(p) = \sum_{i \in L} a_i y_i K_{RBF}(p_i, p_j) + b \tag{2.8}$$

Where $K_{RBF}(p_i, p_j)$ is the radial basis kernel function of the SVM classifier. The set 'L' signifies a vector related to Lagrangian multipliers which denote the Support Vectors.

The performance of the SVM classifier for a given data set depends upon the tuning of kernel parameter λ and penalizing parameter C for construct a separating hyperplane with maximum possible margin between classes. SVM classification model is governed by several important relations like Equ. (2.3), (2.5), (2.6) and (2.8). Satisfying the governing relations with best possible λ and C will result an SVM model with improved accuracy for that data set. Optimal values of λ and C are searched in the direction of obtaining better classification accuracy. The objective to be realized in the SVM parameter tuning is based on classification accuracy in 'k-fold' validation of the given data set i.e misclassification rate.

$$Minimize\ ZZ = \frac{1}{k} \sum_{i=1}^{k} E_{MCR_i} \tag{2.9}$$

Where E_{MCRi} is the misclassification rate occurred in 'i^{th}' validation.

$$E_{MCR_i} = f(\lambda, C) = \frac{False\ Positives + False\ Negatives}{Total\ instances} \tag{2.10}$$

SVM parameters turning model for optimal kernel and penalizing parameters are searched using ALO algorithm which is presented in the next section.

III. ALO BASED OPTIMAL SVM PARAMETER TUNING

The best possible solution in the problem space is effectively obtained by ALO algorithm due to better travelling across the problem plane during unintended movement & arbitrary selection of ants and antlions, considered as search agents [8]. The local optima are highly prevented during the unintended movement and arbitrary selection of antlions based on their health fitness. During optimization, antlions take the position of healthy ants and this location is saved. The healthier antlion in that generation is stored, named Elite.

Step 1: Initialization

Total ants & antlions, maximum planned generation, stopping criteria are initialized. The initial population of the search agents; ants and antlions are generated randomly generated. Decision variables, kernel parameter and penalizing parameter are randomly generated. ALO belongs to the group of natural process mimicking optimization procedure which replicates the capturing of food process in Doodlebugs insects (larve of antlions).

$$X_i^0 = A_i + rand * (B_i - A_i) \tag{3.1}$$

X_i^0 is the initial position of search agent i for each dimension whose value is within the bounds of (A_i, B_i).

Step 2: Evaluating fitness value

Evaluating the health of the searchers, a fitness value is found using the objective given by Equ.(2.9). In this methodlogy SVM misclassification rate is estimated using 'k' fold validation. The antlion with better health from the initial generation can be stored as elite.

Step 3: Trapping in antlions' pits

The unintended movements of ants in the problem plane are affected via Ant lions' created pits. It directs the prey insects towards unknown search regions. This can be observed by the following equations

$$C_i^t = Antlion_j^t + C^t \tag{3.2}$$

$$D_i^t = Antlion_j^t + D^t \tag{3.3}$$

Where C_i^t - the least value of 'i' in't', D_i^t - highest value of 'k' in 't', C^t – the least of all variables in 't' and D^t is the highest value of all 'k' in t, k denotes variables and t denotes iteration.

Step 4: Sliding ant in the direction of antlion

During the process of building traps, traps are built based on antlions' strength and ants are necessary to shift their position randomly. The roulette wheel is used to assign antlion depends on the fitness or simply health. When they found an ant in the trap, they started throwing sand from the middle to over the pit until the ant tumble into the created pit. This step is modelled as the range of ant's unintended circular movement that changed adaptively depending on present iteration level.

$$C^t = \frac{C^t}{I}; D^t = \frac{D^t}{I}; \ \ I = f(w, t, It) = 10^w \frac{t}{It} \tag{3.4}$$

Step 5: Normalize the Random walks of ants

Ant's movement is random in nature during the food searching process. This unintended movement may be imitated to the next positional change as

$$X(t) = [0, cums(2s(t\_1) - 1), cums(2s(t\_2) - 1, \\ .., cums(2s(t\_(N\_I)) - 1)] \quad (3.5)$$

Where cums calculates the cumulative sum and r(t) is defined as follows:

$$s(t) = \begin{cases} \mathrm{I} \;\; if \quad rand > 0.5 \\ 0 \;\; if \quad rand \leq 0.5 \end{cases} \quad (3.6)$$

To retain the unintended movement within the problem plane, normalization is carried based on Equ. (3.5)

$$X_i^t = \frac{(X_i^t - A_i) \times (D_i^t - C_i^t)}{(B_i - A_i)} + C_i \quad (3.7)$$

Where $X_{i,k}^t$ - k[th] variable position in that generation of a particular ant i.

Step 6: Catching ants for food and recreate the pit

Every ants' health in its present position is found using objective selected. Modify antlion's location to the ant's present location if the hunted ant has a better health otherwise keep the original antlion position for the next iteration.

$$Antlion_j^t = Ant_i^t \quad if \; f(Ant_i^t) > f(Antlion_j^t) \quad (3.8)$$

Step 7: Elitism

Healthiest antlion acquired to be preserved, named elite. Superior elite antlion can able to disturb the ants location thereby updating of ants position in the problem plane. Ants should perform unintended movements based on the antlions fixed and elite. It is performed as

$$Ant_i^t = \frac{Antlion_a^t + Antlion_e^t}{2} \quad (3.9)$$

Step 8: Convergence

If the convergence measures are not satisfied then the ants update their position for further exploration of the searching space otherwise the searching procedure is terminated. The position of elite antlion gives the best possible result. Convergence criteria may be either an acceptable solution found or no betterment in health is possible or a chosen number of unintended movement finished.

Step 9: Optimal SVM parameters and Classification

The elite antlion is the best possible SVM parameter for the given dataset with minimal misclassification rate. The data samples are classified with the converged SVM model for correct classification of objects in the test image.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

Landsat-8 image (fig.1) representing Trichirappali (10^0 39'N 78^0 33'E 10^0 51'N 78^0 48'E) study area recorded at 2019 is used as a test system for analyzing the SVM classifier performance. Landsat-8 image has resolution of 30m with 11 bands.

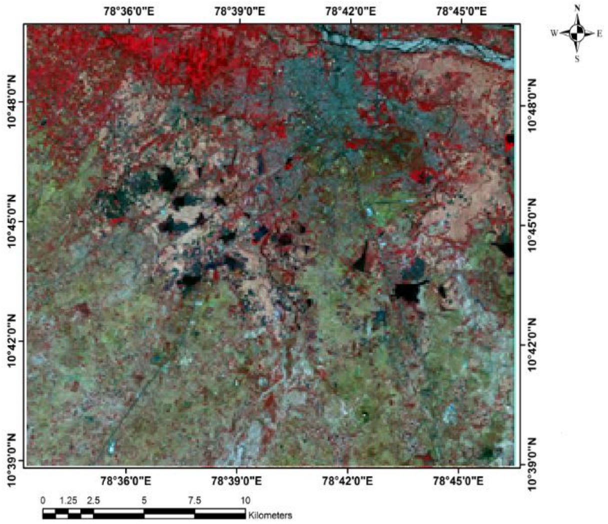

Fig. 1: Landsat FCC of the study area -2019

Landsat image is preprocessed with atmospheric and radiometric correction steps. Improvement in Information interpretability and structural preservation of the image is taken care during contrast enhancement process. Enhanced image is segmented in to image objects by the segmentation algorithm.

Special, spectral and textural features quantify the image objects are extracted. The extracted feature set has 70 features for characterizing the image objects. Field survey in the study area and ancillary data from the Google Earth are used to generate ground truth reference data for the Landsat image.

Predominant land cover components found in the study area are agriculture, forest, buildup urban, buildup rural, water body and barren land. Based on the ground truth reference, the image-objects in Landsat image is classified in to 6 land cover classes. SVM classifier with fixed RBF kernel parameter of 0.003 and penalizing parameter of 100 results an overall accuracy of 88.96%. This result is presented in Table-1.

Table 1: Accuracy Performance (Confusion matrix) - Complete 70-features with fixed SVM parameters

	Agriculture	Forest	Buildup Urban	Buildup Rural	Water body	Barren land	Total no. of samples	Producer's Accuracy (%)
Agriculture	1341	9	18	49	3	18	1438	93.26
Forest	2	147	16	15	2	10	192	76.56
Buildup Urban	3	7	264	24	1	6	305	86.55
Buildup Rural	6	5	11	187	3	12	224	83.48
Water body	7	5	1	4	137	3	157	87.26
Barren land	5	2	12	13	4	148	184	80.44

Total column samples	1364	175	322	292	150	197	2500	-----
User's Accuracy (%)	98.31	84.01	81.98	64.04	91.33	75.13	------	-----

Overall Accuracy = 88.96%; Kappa = 0.829

Table 2: Accuracy Performance (Confusion matrix) - Significant 24-features with fixed SVM parameters

	Agriculture	Forest	Buildup Urban	Buildup Rural	Water body	Barren land	Total no. of samples	Producer's Accuracy (%)
Agriculture	1353	16	11	44	2	12	1438	94.09
Forest	0	158	6	10	1	7	192	82.29
Buildup Urban	2	6	271	19	2	5	305	88.85
Buildup Rural	6	4	10	195	1	8	224	87.05
Water body	7	3	1	4	140	2	157	89.17
Barren land	4	3	12	11	2	152	184	82.61
Total column samples	1382	190	311	283	148	186	2500	-----
User's Accuracy (%)	97.90	83.16	87.14	68.91	94.59	81.72	------	-----

Overall Accuracy = 90.76%; Kappa = 0.856

Improvement in classification accuracy is guaranteed if the redundant and erroneous features are removed from feature set. Feature selection (FS) is carried out to achieve progress in the classification and reduction in feature dimension with fixed SVM parameters results 24 significant features at 90.76% overall accuracy. Confusion matrix elaborating the classification performance achieved in FS is shown in Table-2.

The proposed ALO methodology is assigned to improve the classification accuracy further by optimally tuning the kernel and penalizing parameters with the objective of best possible misclassification rate convergence. ALO parameters initialized for parameter tuning are ants: 30, antlions: 20, iterations $Iter_{max}$: 100, stopping criteria: $Iter_{max}$ completed or no change in fitness for consecutive 25 iterations. Bounds imposed on SVM parameters are RBF Kernel: 0.001 to 10; Penalizing parameter: 10 to 1000. Fitness function is estimated based on 5-fold cross-validation with 2500 samples. Proposed ALO methodology converges to kernel parameter of 0.2140 and penalizing parameter 96.3524 with the misclassification error of 7.12%.

In order to compare the evolution of classification improvement with SVM classifier for the Landsat 2019 data set, classification process with 2500 random samples using complete feature set of 70 features with fixed SVM parameters, significant feature subset of 24 features with fixed SVM parameters and significant feature set with optimum SVM parameters are presented in Tables 1,2 and 3.

It is from tables, SVM with optimal parameter (Significant feature data set) achieves the best overall classification accuracy of 92.68% than with fixed SVM parameters. It also predicts the individual classes accurately compare to other cases. Precision in classification and recalling ability is depicted in fig.2 & 3, that exhibits the effect of optimal SVM parameter in Land cover Classification.

Table 3: Accuracy Performance (Confusion matrix) - Significant 24-features with Optimal SVM parameters Converged using Proposed ALO methodology

	Agriculture	Forest	Buildup Urban	Buildup Rural	Water body	Barren land	Total no. of samples	Producer's Accuracy (%)
Agriculture	1368	8	11	36	2	13	1438	95.13
Forest	8	165	5	9	3	2	192	85.94
Buildup Urban	3	4	276	17	1	4	305	90.49
Buildup Rural	2	1	7	207	0	7	224	92.41
Water body	6	4	1	2	142	2	157	90.45
Barren land	4	2	8	9	2	159	184	86.41

Total no. of samples	1391	184	308	280	150	187	2500	-----
User's Accuracy (%)	98.35	89.67	89.61	73.93	94.67	85.03	-----	-----

Overall Accuracy = 92.68%; Kappa = 0.886

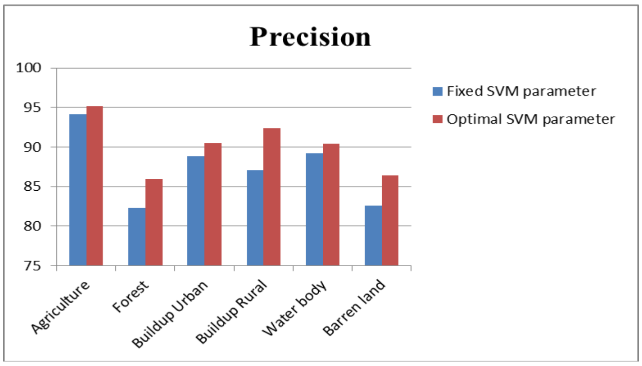

Fig.2 Precision Performance Metric Analysis with fixed and optimal SVM parameter

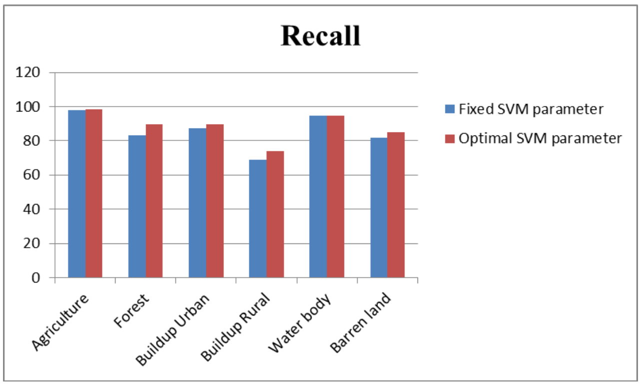

Fig. 3 Recall Performance Metric Analysis with fixed and optimal SVM parameter

SVM model with optimal λ and C is employed to the significant feature set of Landsat 2019 for the preparation of classification thematic map which is presented in fig.4.

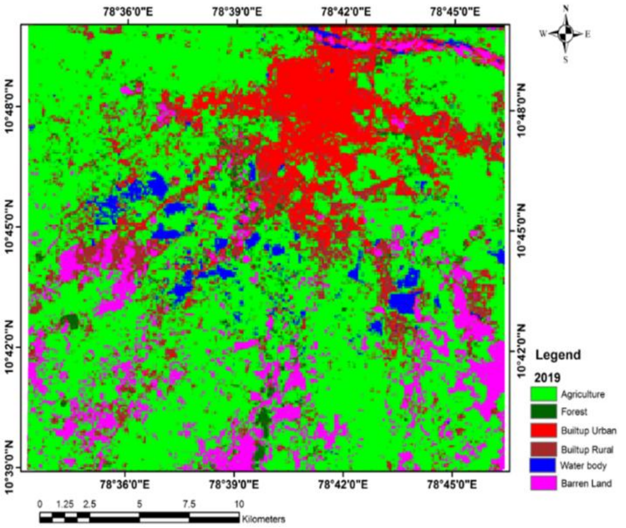

Fig.4 Land Cover Thematic Classification Map

V. CONCLUSION

This methodology decodes the best possible SVM parameters for land cover classification based on the agenda of optimal reduction of misclassification error for Landsat image data of the area of study. ALO algorithm explores various possibilities in problem search and converges to the best parameters for achieving improved classification accuracy. Results to exhibit the evolution of methodologies for improved SVM performance are also presented for the test case. SVM classifier outputs presented for different steps in the evolution process revealed the proposed methodologies enhance the classification process. With tuned SVM model, classification thematic maps are generated for further socio-economic analysis.

REFERENCES

1. Vapnik Vladimir N. The Nature of Statistical Learning Theory. New York: Springer, 1995.
2. Huang C L , Dun J F . A distributed PSO–SVM hybrid system with feature selection and parameter optimization. Applied Soft Computing, 2008, vol. 8(4), pp1381-1391.
3. D. H. Damian Eads, Sean Davis, "Genetic Algorithms and Support Vector Machines for Time Series Classification," In SPIE, Seattle, WA, USA 2002, pp. 74-85.
4. C.-H. Wu, G.-H. Tzeng, Y.-J. Goo and W.-C. Fang, "A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy," Expert Systems with Applications,2007, vol.32, pp. 397-408.
5. C.-L. Huang and C.-J. Wang, "A GA-based feature selection and parameters optimization for support vector machines," Expert Systems with Applications, 2006, vol. 31, pp. 231-240.
6. C.-H. Wu, G.-H. Tzeng, Y.-J. Goo and W.-C. Fang, "A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy," Expert Systems with Applications, 2007, vol. 32, pp. 397-408.
7. Y. Bazi and F. Melgani, "Toward an optimal SVM classification system for hyperspectral remote sensing images," Geoscience and Remote Sensing, IEEE Transactions on, 2006, vol. 44, pp. 3374-3385.
8. Seyedali Mirjalili, "The Ant Lion Optimizer", Advances in Engineering Software, Vol. 83, pp.80-98, 2015.

AUTHORS PROFILE

K.Jayanthi, working as Assistant Professor in the Department of computer application, Government Arts College, Chidambaram, India. She completed M.C.A from Bharathidhasan University and M.Phil [Computer science] in Annamalai University. She is pursuing Ph.D in Annamalai University. She is doing research in the area of image processing application to remote sensing. She has published 5 research articles in international journals and presented research papers in 6 international / national conferences.

Dr. L R. Sudha, working as Associate Professor in the Department of computer Science & Engineering, Annamalai University, Chidambaram, India. She completed M.E and Ph.D in Computer Science & Engineering from Annamalai University during 2006 & 2014 receptively. She has 20 years of teaching and 16 years of research experience. In her credit, she published 16 research articles in international journals and presented papers in 20 international/national conferences. Her area of interest includes image processing and machine learning.