EAI Endorsed Transactions
on Scalable Information Systems
Research Article EAI.EU

Performance Comparison between SVM and LS-SVM for Rice Leaf Disease detection

Snehaprava Acharya[1], T. Kar[1, *], Umesh Chandra Samal[1], Prasant Kumar Patra[1]

[1] School of Electronics Engineering, KIIT deemed to be University, Bhubaneswar, Odisha, India

Abstract

INTRODUCTION: Automatic detection of rice plant diseases at early stage from its images is quite beneficial over traditional verification methods.
OBJECTIVES: Recent years machine learning (ML) approaches are more efficient in disease classification task. In current generation the statistical machine learning algorithm which shows state-of-arts performance is Support Vector Machine (SVM) and variants of SVM.
METHODS: SVM has an excellent learning performance for linear and non-linear data samples. It works for Quadratic Programming Problems (QPP) due to which it has the drawback of computational complexity. However QPP can be solved linearly with the help of Least Square SVM(LS-SVM) approach. In LS-SVM the epsilon tube and slack variables of SVM are replaced with error variables. The distance is calculated by error square value.
RESULTS: In this research performance comparison is made between SVM and LS-SVM for rice leaf diseases such as Bacterial Leaf Blight (BLB), Brown spot(BS), Leaf smut(LS) and Leaf Blast using two datasets (DS1 and DS2).Accuracy of LS-SVM is found to be 91.3% and 98.87% for DS1 and DS2 respectively whereas accuracy of SVM is 83.3% and 98.75% for DS1 and DS2 respectively.
CONCLUSION: Performance of LS-SVM outperformed than SVM in terms of accuracy.

Keywords: SVM, LS-SVM, rice leaf diseases, QPP, Dual Thresholding

Received on 08 July 2023, accepted on 04 September 2023, published on 21 September 2023

doi: 10.4108/eetsis.3940

*Corresponding author. Email: tkarfet@kiit.ac.in

1. Introduction

India is one of the pioneers in production and export of rice in the world which contributes largely to our country's sustained economic growth and development. Hence the most desirable exponential growth in this sector needs to be paid focused attention in order to ensure the targeted quantity of quality rice production. However, it is experienced that the targeted rice production in terms of quantity and quality is often adversely affected by various plant diseases due to various challenging factors. In this context it is observed that some of the most common diseases like Bacterial leaf blight, Brown spot and Leaf smut which grossly reduce the productivity of the crop [1]. A few sample images of bacterial leaf blight, brown spot and leaf smut is given in Fig 1. (a), (b) and (c) respectively may convince one about how our rice production sector bears with the extent of damaging consequences due to the aforesaid rice plant diseases.

(a)

(b)

(c)

Figure 1. Different rice plant diseases (a) Bacterial leaf blight (b) Brown spot (c) Leaf smut

In the standard procedure of remedial practice for a viable solution to this problem, the plant pathologists detect these diseases using Invasive Methods. However, in recent times the, the newer technologies like Advanced Digital Data Acquisition (ADDA) and the Image Processing of the Plants (IPP) offer a non-invasive solution for the same. Moreover, this image processing method combined with the machine learning also serves as an advantage of automation to the disease detection strategy. At this point it is worth mentioning that the Image based disease detection process proceeds through the stages like image pre-processing, segmentation, feature extraction and classification. Though there are several non-invasive techniques are in current practice for the purpose, the performance of these can still be improvised by incorporating few more steps before applying machine learning based classifier.

Different steps of disease detection are:

Pre-Processing: It involves the preparation of data by transformation of the same into suitable colour space followed by the removal of residual noise.

Segmentation: It involves the separation of the diseased part from the healthy parts of the leaf.

Feature Extraction: This phase deals with the extraction of the suitable features where "Feature" is to be understood as the representation of the image in a different domain. Out of several features only a few significant features of the image are selected for subsequent processing and the series of features arranged in the form of a vector, known as feature vector, should be non-correlated in nature having the lower dimension than its original image. Here the color, the texture and the shape [2] are taken as the three primitive visual features of an image whose image representation can be well improved on the basis of the well-defined feature extraction techniques.

Classification: In this phase the pre-defined patterns of data base need to be categorized properly into different classes using some base classifier such as SVM, ANN, K-NN, Rule based classifiers, K-Mean and Fuzzy c-Mean clustering, Markov Random Field, Decision Tree algorithms or may be some combination or modified version of the base classifier.

The rest of the paper is organised as follows. The Section-II gives a literature survey of the work. Section- III discusses about the materials and the methods of SVM and its Least Square variant (LS-SVM) as a classifier. The section-IV focuses on a credible discussion on the Simulation of the optimal results following conclusions and the future research scopes in section V.

2. Literature Survey

This section gives a brief description of the various justifiable surveys carried out in the last few years for automatic detection of the leaf diseases of different plants and offers a logical comprehension of the entire process. In addition, we have also thrown some light on the works so far done by other researchers for the machine learning based rice plants. B.S. Prajapati et al. [2] focused the work mainly for classification of three diseases like on Bacterial Leaf Blight, Brown Spot, Leaf Smut and implemented various techniques for background removal and segmentation process. Total 88 features are extracted like color, shape and texture in HSV, RGB color plane. Highest testing accuracy of 73.3% was achieved with training accuracy of 93.3% and CV accuracy of 88.5% for K= 10 using SVM classifier. Centroid K-mean segmentation technique showed better performance. Phadikar et al. [3] proposed a morphological feature based automated system (radial hue distribution) to classify brown spot and rice blast diseases using Bayes' and SVM techniques and found accuracy of 79.5% and 68.1%respectively with 1000 images. Pothen et al. [4] used classical methods for detecting diseases for BLB, Leaf Smut and Blast diseases. Otsu's segmentation is followed by pre-processing stage. Then LBP+HOG is used as feature extraction technique. Lastly classification is done with polynomial kernel function of SVM and got accuracy of 94.6%. The Q. Yao et al. [5] used SVM classifier for Bacterial Leaf Blight (BLB), Sheath Blight and rice blast diseases. Rectangular filter used for pre-processing and Otsu's thresholding for segmentation. Total 60 texture features and 4 color features are extracted in 3 models. 97.2% highest accuracy is achieved for model1. Singh et al. [6] proposed a model on rice blast disease and data base of images are collected from IRRI. For pre-processing the wiener filter, the histogram equalization is used followed by K-mean Segmentation along with SVM classifier. Total prediction accuracy of 82% is found which showed better performance than the other classifiers.

Wang et al. [7] done a research survey on SVM. Analysed the equation for quadrature problems, studied parameter optimization for SVM. Different variants of SVM like Multi class SVM, Twin SVM, Fuzzy SVM are studied in detail. Jieping Ye et al. [8] studied theoretically and empirically the relation between SVM and LS-SVM even performance of both is compared at different conditions. Performance of LS-SVM and Hard Margin SVM (Hard M-SVM) based on Mahalanobis distance for binary classification are approximately same which is analysed empirically and theoretically. For multi class classification, performance is approximately same based on Euclidian distance. N V Raja Reddy et al. [9] proposed a model for infectious rice plant diseases using a hybrid deep learning and machine learning approach where they investigated 5 diseases of Oryza Sativa like Brown Spot, Leaf Blast, Leaf Smut, Tungro, Bacterial Leaf Blight. In pre-processing stage, they resized the image followed by histogram equalization and Kuwahara filtering technique. They applied spatial fuzzy Clustering (SFC) for segmentation. Feature extraction is done using LeNet5 structure where LS-SVM is used for classification. To improve the accuracy WSSO (Weighted Sparrow Search Optimization) algorithm is used along with LS-SVM. They got an accuracy of 98.55%.

The above literature review analysis helps us find that although the median filtering method used for pre-processing performs better for detecting the big patches on a leaf, it rarely works for detecting the small spots on rice leaves. Further, it is also worth mentioning that for the segmentation of the affected areas in the leaf image is only achieved by using either the Fuzzy C-Mean, k-means segmentation process in LAB/HSV colour space or by Otsu method and the final disease detection are made via SVM, KNN, LS-SVM or Neural network-based classifiers. However, our literature review studies on the SVM based classifications convincingly demonstrate highly encouraging results for the classification methods which prove as the better methods than the other machine learning based variants. Therefore, having been inspired by these findings based on the invasive methods for disease detection, in this paper, aperformance analysis of the existing learning-based classifiers is made on different data sets of rice leaf. Moreover, a performance comparison of SVM and LS-SVM classifier is also done for 3 class disease classification on the same dataset.

3. Materials and Methods

Current section presents a thorough description of the proposed work through a workflow diagram in fig. (2) for disease detection. The main purpose is to detect different rice plant diseases namely brown spot, bacterial leaf blight, and leaf smut in DS1 and brown spot, bacterial leaf blight, Blast for DS2.

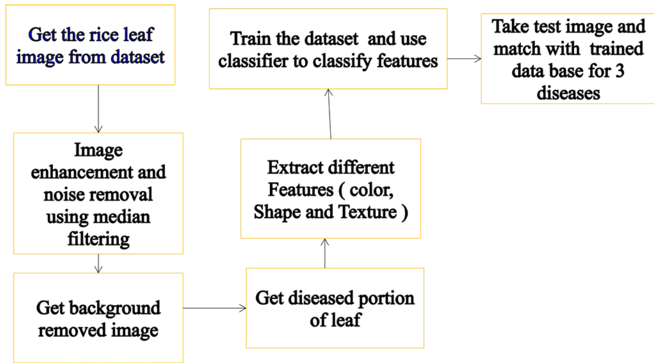
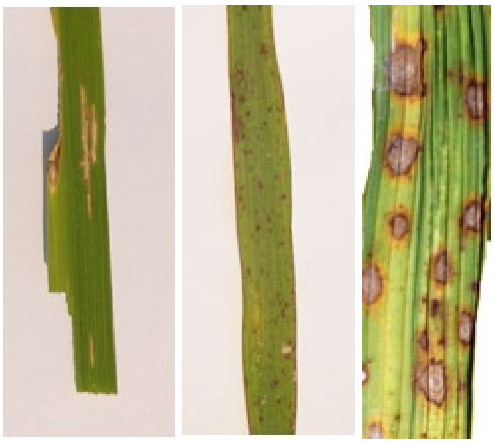


Figure 2. Workflow of the algorithm

The major steps of the algorithm are:

3.1. Image Dataset Collection:

Two rice plant leaf image datasets namely DS1 and DS2 are considered in this paper. DS1 is the first dataset of images collected from Kaggle [2] containing 120 leaf images of 3 different diseased class having 40 images from each class. The sample images of DS1 are shown in Fig. 3. The second dataset DS2[10] contains drone captured 4432 numbers of images for bacterial leaf blight, brown spot and blast diseases. The sample images of DS2 are shown in Fig. 4.



(a) (b) (c)

Figure 3. Sample Diseased image (a)Bacterial Leaf Blight (b) Leaf Smut (c)Brown Spot


(a) (b)

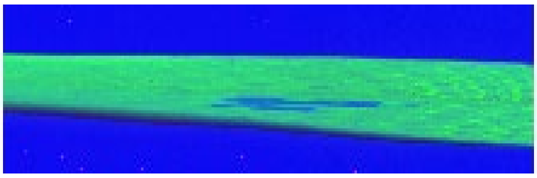
(c)

Figure 4. Sample Diseased image (a)Bacterial Leaf Blight (b) Brown Spot (c) Leaf Blast

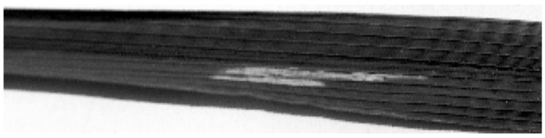
3.2. Image Pre-processing:

Visual quality of an image is enhanced in pre-processing stage by color space transformation, resizing or different filtering techniques. In this work first infected RGB image is converted to HSV color space and saturation plane image is extracted though saturation contains only whiteness then median filtering is applied for noise removal. Median filter is a higher order statistics filter and non-linear in nature.

(a)



(b)



(c)

Figure 5. Pre-processed image (a) Original image (b)HSV plane image (c) S-plane median filtered image

3.3. Segmentation of disease affected image:

Segmentation means clustering of image where each cluster is having some similar kind of features between pixels. In this paper Otsu's thresholding technique is used for segmentation. Normally it is focused to find global optimal value and image pixels are categorized to two groups. Two stage thresholding technique is used for diseased spot detection. First thresholding method is applied in saturation plane to obtain background removed image whereas second threshold is applied to hue plane to get only diseased portion of leaf.

(a)

(b)

(c)

(d)

(e)

Figure 6. (a) S-plane thresholded image (b) Background removed Image (c) Hue-plane thresholded image(d) non-diseased portion of the image (e) Diseased portion segmented image from Bacterial blight disease dataset

3.4. Feature extraction:

Feature extraction plays an important role in this application. mainly it is a dimension reduction method. Color, shape and texture are primary feature analyser. Different features such as mean, standard deviation, kurtosis, skewness, area, percentage of affected area and texture features which is fed to SVM based classifiers.

3.5. Classification:

Classification is the final state for detection of diseases. In this work the Classification is carried out using SVM and LS-SVM.

4. Simulation and Result Analysis

In result analysis different classical supervised classification techniques like ANN, K-NN are compared with SVM result then SVM result is compared with LS-SVM. SVM performance highly depends on the selection of the parameters and kernel functions [11]. From detailed study optimal parameter values and kernel function is considered for the proposed work. Then the performance of least square variant of SVM (LS-SVM) [12] was also studied and the results are compared on the basis of accuracy and time complexity. For simulation and result analysis we have considered two rice plant leaf image datasets namely DS1 and DS2. Both the datasets are divided into an 80:20 ratio for training and testing. Trained dataset is compared with test dataset and classified into a particular class decided by the matched features. All simulations are done in MATLAB-2018 environment using INTEL(R) core (TM)-i5 processor @2.4GHz, 64-bit OS and 6GB RAM configuration.

In the experiment, initially the performances of SVM is compared with two classical existing algorithms i.e. ANN and K-NN based on testing accuracy for both the datasets, DS1 and DS2 and placed in Table 1. The ANN architecture designed for classification has 1 hidden layer having 13 inputs and 3 outputs. KNN has typically good accuracy in low dimensions which is suitable for current applications [13]. The performance of KNN based classifier is reported using optimum distance metric and optimum K-neighbors in each observation. The performance of SVM based classifier is measured with RBF kernel. It is observed from Table 1 that SVM has the best testing accuracy of 83.3% for DS1 and KNN has the best test accuracy of 99.24% for DS2 among all the methods.

SVM outperformed the other classifiers for DS1. As data size increases K-NN outperformed the SVM and ANN which is observed for DS2 as given in Table 1. Again, in terms of execution time K-NN has the highest execution time and ANN has the lowest execution time as shown in the last column of Table 1 for DS1 and DS2 respectively.

Next, we analyze performance of SVM and LS-SVM in terms of accuracy and execution time (ET). Accuracy is found for three class diseases classification and shown in Fig.9.

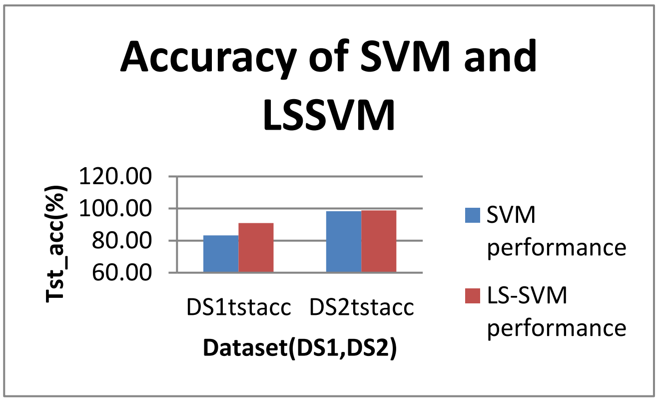


Figure. 9 Accuracy comparison between SVM and LS-SVM.

From Fig:9 it is found that for DS1 and DS2 LS-SVM outperformed SVM.

Table 1. Comparison of Accuracy of all the classifiers for DS1 and DS2

Diseases classes	Data Set	ANN in %	KNN in %	SVM in %	LS-SVM in %
BLB, BS and LS	DS 1	66.25	83	83.3	91
BLB, BS and BLST	DS 2	61	99.24	98.84	98.87

Table 2: Comparison of Execution time (ET) of all the classifiers for DS1 and DS2

Diseases classes	Data Set	ET for ANN in sec	ET for KNN in sec	ET for SVM in sec	ET for LS-SVM in sec
BLB, BS and LS	DS 1	3	136	128	9.25
BLB, BS and BLST	DS 2	12	711	640	666

In the above table performance of all classical classifier along with LS-SVM is evaluated in terms of accuracy and execution time. it is found that accuracy for DS1 LS-SVM outperforms other classifier whereas for DS2 K-NN works best. Similarly, ANN outperformed for both dataset in terms of ET. It is also observed from Table 2 that for small dataset

DS1 LS-SVM has the second lowest ET and for large dataset DS2 LS-SVM has the highest ET.

Table 3: Performance comparison of the proposed method with state-of-the-art methods

Author's Name	Dataset	Methods	Testing Accuracy in % scale
Prajapati et al. [2]	DS1	K-Mean segmentation + SVM	73.3
Azim et al. [14]	DS1	LBP and GLCM Feature+ XGBOOST	86.7
Jhunde chen et al. [15]	DS1	DenseNet-201 +SVM	94.07
P. Sethy et al. [10]	DS2	ResNet50 + SVM	98.3
Nv Rajareddy et al. [9]	DS1	LeNet+LS-SVM+WSSO	98.55
Nv Rajareddy et al. [9]	DS2	LeNet+LS-SVM+WSSO	98.37
Proposed Method	DS1	Dual thresholding+ +SVM	83.3
		Dual thresholding +LS-SVM	91
Proposed Method	DS2	Dual thresholding +SVM	98.4
		Dual thresholding +LS-VM	98.84

Again, the performance of the proposed method is also compared to other state of the art methods in terms of accuracy for same data set images and shown in Table 3. Prajapati et al. [2] used DS1 for their experiments and achieved testing accuracy of 73.3%. Azim et al. [14] applied LBP and GLCM for feature extraction along with SVM classifier with RBF kernel and XGBOOST classifier separately for same dataset DS1 and achieved an accuracy of 81.6% and 86% respectively. Similarly, Chen et al. [15] used DenseNet 201 for feature extraction followed by SVM for DS1 dataset and achieved an accuracy of 94%. Rajareddy et al. [9] used DS1 and DS2 where LeNet network is used for feature extraction following LS-SVM classifier along with Weighted Sparrow Search Optimization (WSSO) for fine tuning of one of the parameters and achieved an improved accuracy of 98.55% for DS1 and 98.37% for DS2 respectively. Sethy et al. [10] applied 13 number of transfer learning models followed by SVM where best performance is achieved by ResNet 50 with an accuracy of 98.3% for DS2 dataset. The proposed method achieved an accuracy of 91% and 98.84% for DS1 and DS2 dataset respectively.

5. Conclusion and Future Scope

Although invasive methods exist for disease detection which can only be handled by expert plant pathologist, automated machine learning based approaches are gaining popularity because of its non-invasive approach of higher detection accuracy at a faster rate. In the current scenario the disease detection is achieved by initially finding disease affected area of the leaf image following the feature extraction of the affected area and final SVM based classification. SVM and least square version of SVM are executed where it is found LS-SVM outperforms in terms of accuracy (91% for DS1 and 98.85% for DS2). Limitation of this work is the use of limited number of datasets for experiment. So, in future this work may be extent to variety of leaf image datasets. This work is limited to machine learning approaches which may be extended to Deep Learning algorithms and hybrid of machine learning and Deep Learning algorithms for improving the performance of the model.

References:

[1] Godwin M. S, Narmadha R, Bernatin T.. A Brief Survey on Diseases of Paddy Plant. Journal of Pharmaceutical Sciences and Research, Vol.11(17),2019,2739-2743.

[2] Prajapati H.B, Shah Jitesh, Dabhi Vipu. . Detection And Classification of Rice plant Disease. International Journal of Intelligent Decision Technologies-2017, vol.11, no.3.

[3] Phadikar S, Sil J, and Das A. K . Classification of Rice Leaf Diseases Based on Morphological Changes.

International Journal of Information and Electronics Engineering, vol. 2, pp. 460-463, May 2012.

[4] Pothen E. Minu, Pai Maya. Detection of Rice Leaf Diseases Using Image Processing: In Proceedings of the fourth International Conference on Computing Methodologies and Communications (ICCMC), 2020

[5] Yao Q, Guan Z, Zhou Y, Tang J, Hu Y and Yang B. Application of Support Vector Machine for Detecting rice Diseases using Shape and Color Texture Feature. Engineering computation 2009, ICEC09, International Conference 2009 pp (79-83).

[6] Singh A. K, Rubiya A, Raja B. S. Classification of Rice Disease Using Digital Image Processing And SVM. International Journal of Electrical and Electronics Engineers, vol. 07, no. 01, 2015.

[7] H Wang, J Xiong, Z Yao, M Lin, J Ren. Research Survey on Support Vector Machine. : Proceedings of EAI International Conference on Multimedia Communications ,Chongging , People's Republic of China ,July 2017

[8] Ye Jieping and Tao Xiong. SVM versus Least Square SVM. Artificial Intelligence and Statistics,2007-Proceedings, mlr press.

[9] Goluguri N. V. R, K Suganya Devi, Prathima C.H. Infectious Diseases of Rice Plants Classified using a Deep Learning-Powered Least Square Support Vector Machine Model. Indian Journal of Computer Science and Engineering Vol. 13No.5 Oct 2022.

[10] Sethy P, Negi B, Barpanda N, Behera S, Dr Rath. A.K. Deep Feature Based Rice Leaf Disease Identification Using Support Vector Machine. Computers and Electronics in Agriculture 175(2020) 105527

[11] Cortes C, Vapnik V. 1995, Support vector networks, Jour. Machine Learning **20** 273-297.

[12] Suykens J.A. K., Vandewale J. 1999 Least squares support vector machine classifiers. Journal of Neural Processing Letters **9(3)** 293-300.

[13] Pestov Vladimir. Is the K-NN Classifier in high Dimensions affected by the course of Dimensionality? Computers and Mathematics with application,2013.

[14] Azim Anwarul. M, Islam Khairul. M.K, Rahman Md.Marufur , Jahan farah. An Effective Feature Extraction method for Rice Leaf Disease. Classification, TELKOMNICA Telecommunication, Computing, Electronics and Control,2021.

[15] Chen J, Zanng D, Nanehkaran Y.A, Li D. Detection of Rice Plant Diseases based on Deep Transfer Learning .Journal of Science of Food and Agriculture March 2020.

.