A Deep Learning model for the identification of Potato leaf diseases using Wrapper Feature Selection and Concatenation

Muhammad Ahtsam Naeem
University of Electronics Science and Technology of China
Muhammad Asim Saleem
Ripah International University
Muhammad Imran Sharif
Kansas State University
Shahzad Akbar
Ripah International University
Muhammad Zaheer Sajid
27mcs@student.nust.edu.pk

National University of Science and Technology

Research Article

Keywords: Classification, Deep learning, Equilibrium Optimization, Late Blight, SVM, Potato Leaf Disease

Posted Date: March 28th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4155580/v1

Additional Declarations: No competing interests reported.

Abstract

Potato is a popular crop that is cultivated in many different climates. Potato farming has recently gained incredible traction, increasing relevance in international agricultural production. Potatoes are susceptible to several illnesses that stunt their development. This plant has significant leaf disease. Early blight (EB) and late blight (LB) are the two devastating leaf diseases for potato plants. The early detection of these diseases would be beneficial for enhancing the yield of this crop. The ideal solution is image processing to identify and analyze these disorders. Using image processing and machine learning, we detail a method that requires no outside help to detect late-blight potato leaf in this article. The pro- posed method comprises four different phases: (1) Histogram input images may improve from equalization to boost their overall quality; (2) feature extraction is performed using a Deep CNN model, then these extracted features are concatenated; (3) feature selection is performed using wrapper-based feature selection; (4) classification is performed using an SVM classifier and its variants. By utilizing SVM and a meticulously selected set of 550 characteristics, the suggested technique achieves an unprecedented 99% accuracy.

1. Introduction

In areas where potatoes are planted, late blight diseases are frequent. According to Arora et al., Potato blight is a widespread disease that affects green leaves. It starts as small, irregular, light green lesions around the leaf's tip and edges, and eventually turns into huge, dark brown or even purple necrotic areas (1),(2). Its production of over 329 million metric tons of non-grain agricultural products in 2009 serves as an essential element of over 1.5 billion people's daily diets (3). Potato is a multipurpose and widely available crop that has played a critical role in China's economic growth. Pests and diseases, on the other hand, significantly limit potato production. Late blight, a disease that affects potatoes over the globe, is the most destructive (4). The agriculture industry is vital to the world economy. It emphasizes the need to provide adequate care for plants from when they are seedlings until they produce the desired crop. The crop has to undergo several stages throughout this procedure (5). If sufficient measures are taken to protect the playing area, this issue may be remedied. People won't be able to change some parts of the weather; their only option is to wish for better conditions. Finally, the significance of preventing infections that might stunt the crop's growth and reduce its yield cannot be overstated (6). The crop will be able to be protected with the right nutrients if these illnesses can be detected early enough. Providing a method for recognizing and categorizing illnesses that are amenable to digitalization would be beneficial for agriculturists (7). Agriculture nowadays, management is increasing complexity due to many factors (7). Regarding farming, in particular, the business has evolved into an increasingly competitive and worldwide one, where producers must consider local climatic conditions, environmental issues, and worldwide environmental and economic elements (8). Compared to grains, the potato is a more productive crop in terms of the amount of protein, dry matter, and minerals it generates per unit of land. On the other hand, the production of potatoes is susceptible to a wide variety of illnesses, which may cause yield losses and a decline in tuber quality, ultimately leading to an increase in the price of potatoes (9). Potato crops are susceptible to various illnesses, particularly parasitic infections, which may significantly decrease crop output and cause severe financial losses for farmers and producers. Enzymes (10),(11) play an essential role in developing potato leaf disease caused by the pathogenic fungus Phytophthora infectans. Farmers who want to minimize annual losses must first identify the many illnesses that harm potato plants. Blights, both early and late, are among the most prevalent diseases. Economic loss and waste reduction prevention may be achieved by identifying these illnesses early and applying effective therapy. Because early and late blight treatments are distinct, every potato plant must be correctly identified. Certain strains of microalgae (12) (a diverse group of tiny, aquatic microorganisms (13)), such as Chlorella pyrenoidosa, protect potato plants from infection. Here, a convolutional neural network model is trained using deep learning to categorize illnesses affecting potato plants. For farmers to select the proper treatment, the model should provide a way to identify whether their potato plants have early or late blight diseases (14). Environmental factors have an impact on the occurrence of late blight. Excessive humidity and specific temperature conditions may wreak havoc on an entire field in a short period (15). Leaf yellowing, wilting, and withering will be seen, as well as a shift in spectral characteristics (16). The approach suggested by Islam, Dinh, Wahid et al. (2017) utilizes imaging technology and AI to identify plant diseases on leaves. Potato plant diseases and natural leaves have been identified from public-domain images using an automated approach validated using so-called ’Plant Village’ data. The accuracy of the proposed model is around 95%, and it is achieved via the use of support vector machines for the segmentation and classification out- lined below. The proposed technology may now be used for widespread plant disease diagnosis. An easy-to-use and automated method was developed using multiclass support vector machine picture segmentation. A small amount of computational work is required to identify basic potato illnesses such as Late Blight and Early Blight. Using this strategy, farmers could identify diseases more quickly, accurately, and efficiently (17). Agriculture nowadays, management is increasing complexity due to many factors (7). Regarding farming, in particular, the business has evolved into an increasingly competitive and worldwide one, where producers must consider local climatic conditions, environmental issues, and worldwide environmental and economic elements (8). Compared to grains, the potato is a more productive crop in terms of the amount of protein, dry matter, and minerals it generates per unit of land. On the other hand, the production of potatoes is susceptible to a wide variety of illnesses, which may cause yield losses and a decline in tuber quality, ultimately leading to an increase in the price of potatoes (9). Potato crops are susceptible to various illnesses, particularly parasitic infections, which may significantly decrease crop output and cause severe financial losses for farmers and producers. Enzymes (10), (11) play an essential role in developing potato leaf disease caused by the pathogenic fungus Phytophthora infestans. The inter-class resemblance and intra-class variation of potato leaf images affect the recognition accuracy. Variation in shape, texture, and size is a challenging task for the segmentation of late blight. The low contrast of potato leaf disease infection area in agriculture images is difficult to detect. It will shorten the time needed to identify diseases and improve the accuracy with which they are classified. The authors compared two CNN approaches for diagnosing 26 diseases in 14 crops using the (18) Plant Village Dataset.

1.1. Research Contribution

The main contributions of this study are:

1. Histogram Equalization is used for image enhancement.
2. Deep CNN model is used for feature extraction.
3. Feature Concatenation and Wrapper Based Feature Selection is used for selecting core features.
4. SVM classifier and its Variants are used for Classification.

1.2. Paper Organization

We will expand on our findings in the parts that follow. Section 2 will go over previous research in the field. The materials and procedures we used are described in Section 3, and our findings are presented and thoroughly discussed in Section 4. The summary and conclusions of our investigation will be presented in Section 5 to conclude.

2. Related Work

A modest convolutional neural network was developed by Kang et al (34), using the Django framework, they developed a system to tell the difference between early blight, late blight, and healthy leaves in potatoes. The parameters have been drastically reduced to guarantee a Top-1 identification accuracy of 93% or higher for the model. The Django framework includes this trained model to build a website for spot- ting early potato leaf diseases. This technical assistance helps develop a mobile-based diagnosis and warning system for potato leaf disease. (19) presented a CANet-based context-aware 3D convolutional neural network model for autonomously segmenting leaf lesions and identifying plant diseases. In order to identify the subtype of lesions in the segmented area, a deep CNN is used. Finally, a hybrid technique using CNNs and linear regression is used to forecast the plant's prognosis for survival. They then evaluated the Plant Village Benchmark Dataset with DICE and IoU matrices for lesion segmentation to assess detection strategy and forecast accuracy. Average efficiency for the suggested lesion segmentation model was 92%, with an IoU of 90%. For tomato, potato, and pepper plants, the lesion subtype identification model's precision was 91.11%, 93.01%, and 99.04%, correspondingly. Consequently, the most recent lesion segmentation algorithm is ideal for real-time disease diagnosis using UAVs and for offline provision of crop health updates to cut the risk of reduced yields. Mahum et. al [20] potato vertex wilt, potato leaf roll, and potato late blight are the five states of potato leaves that may be identified using this improved deep learning method. The given approach, which is trained on a subset of "The Plant Village" dataset, can differentiate among Early Blight and Late Blight on potato departs as well as a Healthy class. Manual data collection is also performed for PLR, PVw, and PH categories. To effectively categorize potato leaf diseases, we use an Efficient DenseNet model that includes a transition layer. Given the asymmetry in the training data, Reweighting the cross-entropy loss function is what really gave the proposed technique some oomph. Small datasets of potato leaf samples may be trained with less overfitting thanks to the strong connections with regularization. The suggested method is an innovative and ground-breaking strategy for diagnosing four illnesses in potato leaves. The accuracy of the algorithm was determined to be 97.2% based on evaluations on the test set. Multiple trials proved the method's superiority over previously used models for identifying and diagnosing potato leaf diseases. Chakraborty et al. (21) evaluated the current state of the art in deep learning algorithms for autonomously detecting late and early blight diseases in potato leaves using optical imagery. Using the PlantVillage Dataset, four deep learning models were tested: VGG16, VGG19, MobileNet, and ResNet50. The findings showed that VGG16 achieved the maximum accuracy, at 92.69 percent. By tweaking the model's parameters, we were able to improve performance to an accuracy of 97.89% while trying to disclose a healthy potato leaf from one affected by late blight or early blight. Comparisons of validation accuracy and losses between the current study's enhanced VGG16 model and those of previously utilized methods. Kumar et al. (22) introduced a Hierarchical Convolutional Neural Network for Deep Learning to identify diseases in leaves. First, median filtering is used to lessen the visual disturbances. The Intuitionistic Fuzzy Local Bi-nary pattern (IFLBP) is then used to retrieve the leaf's individual properties. HDLCNN is then utilized for dis- ease classification, and Decision Support Systems allow farmers to manage treatment programs without increasing risk effectively. Matlab Simulink software comparisons show that this method outperforms the VGG-INCEP, Variations on the Deep Convolutional Neural Network, Random Forest, and Spiking Neural Network. The proposed HDLCNN yields an accuracy, precision, recall, and F-score that is roughly 4%, 6%, 3%, and 3.5% better than existing methods. Additionally, this method has a specificity, sensitivity, and PSNR that outmatches existing methods by 4.5%, 1%, and 2%. Utilizing this modified HDLCNN improves the method's performance with practical implications for former studies. This research serves as an alert to prevent leaf diseases worldwide which will, in turn, enhance potato crop yield globally.

3. Methodology

In this proposed four phases, the first is preprocessing for image enhancements. In preprocessing, we resize the image into 300 X 300 X 3. Histogram equalization is utilized for image enhancement. In the second phase, we used feature engineering using Deep CNN models. Darknet-53 (23), AlexNet (24), and Vgg-19 (25) are used for feature extrac-tion. Moreover, in the third phase, feature concatenation is performed to concatenate the in-depth CNN features. Equilibrium optimization (EO) (26) is performed for feature selec-tion. These selected features are given to different SVM (27) for Classification. In Fig. 1, we see a schematic depiction of the proposed procedure. We used plant village dataset for detection of late blight disease (28). The total number of images in dataset is 1760, 1000 images in late blight folder and 760 images in healthy leaf folder.

3.1. Preprocessing

Our proposed method begins with the preprocessing step of resizing (29) and histogram equalization. The resizing method displays each matrix slice as a separate, unconnected pixel. On the first, more substantial plane, we see how bright the red pixels are, on the second, how bright the green ones are, and on the third, how bright the blue ones are. With this method, we can demonstrate two behaviors in the column while maintaining the same image quality regardless of the width. It is important to note that the supplied image's first two dimensions must be modified. Due to its ease of use and high level of performance, Histogram Equalization (HE) (30) has become one of the most used algorithms for contrast enhancement. As a rule, the HE creates a better picture with a linear cumulative histogram because of the even distribution of pixel values. Histogram modification is often used with HE enhancement for applications (31), including medical image processing, object recognition, and texture. The whole discussion of the image-enhancing method known as histogram equalization (HE), which is described in Eq. 1.

$$P\left(vi\right) = nvi/N /\!/ i = 0, \ldots, L-1\,(1)$$

In this case, a digital image with grayscale values between $[0, L-1]$ and examine the probability distribution of those values. Eq. 1 is utilized to calculate the image's function. If $n'i$ is the fraction of a picture's pixels that are shades of grey vi, and vi is the equivalent grey level. The CDF may also be calculated in the following ways:

$$\mathrm{C}\left(\mathrm{vi}\right) = \sum_{\mathrm{j}=0}^{j=i} \mathrm{P}\left(\mathrm{r}\right), \mathrm{i} = 0, \mathrm{L}-1, 0 \leq \mathrm{C}\left(\mathrm{vi}\right) \leq 1\,(2)$$

To rectify this, Histogram Equalization (HE) uses the following equation to adjust the input image's grayscale from level yi to level vi in Eq. 2. As a result:

$$yi = (L-1)XC\left(vi\right)(3)$$

$$\varDelta yi = (L-1)XP\left(vi\right)(4)$$

Gray level Δyi Changes can be compute in the usual histogram equalization method as shown in the above equation (3 & 4). Figure 2 displays the results of the histogram equalization of late blight leaf disease.

3.2. Feature Engineering

In this phase, we implemented feature engineering using a deep neural network. Feature extraction is accomplished with the help of Darknet-53, AlexNet, and Vgg-19. Features are extracted on the darknet using the Global Average Pooling (GAP) layer. The 1760 X 1024 are total features extracted using the GAP layer of Darknet-53. Moreover, the FC7 layer of AlexNet and Vgg-19 is used for feature extraction. A total of 4096 features are selected individually and extracted. Following the completion of feature extraction, the next step is feature concatenation, which is used to combine the deep CNN features. It is mathematically discussed in Eq. 5. The 1760 X 9216 are total features after feature con-catenation as described in Eq. 6.

For features optimization, equilibrium optimization (EO) is often carried out. These core selected features are sent to several SVM variants for the classification. Figure 3 shows the mesh representation of feature engineering.

$$Fc = \sum_{N=1}^{1024} Darknet53U \sum_{N=1}^{1024} AlexnetU \sum_{N=1}^{1024} Vgg19 \ (5)$$

$$Fc = \sum_{N=1}^{9216} (TotalFeatures) \ (6)$$

4. Results and Discussion

Extensive tests are carried out in this part to test the proposed structure. Here are the processes that were used to get those outcomes: Features are extracted using a pre-trained deep CNN model, the top features from the model are enhanced using equilibrium optimization, SVM variations are used to integrate the most useful information for classification. An image is upgraded using histogram equalization (32). In this work, the Plant Village dataset is used. Moreover, only binary Classification (potato healthy and late potato blight, as shown in Fig. 4) is performed using deep learning. The implied method is evaluated using a variety of measures, including accuracy, total cost, prediction speed, training speed, precision, F1 Rate, and recall rate. In our research, MATLAB served as a useful tool for data analysis and modeling. Classifiers and the best outcomes in each test case are compared, along with a confusion matrix. Five folds are used in the experiments.

Several measures of performance are used to determine how well the proposed research might work. The confusion matrix created during identification task testing is often used in procedures. These protocols are simply calculated as Accuracy (Ac), Sensitivity (Se) & Specificity (Sp) shown in equations 7, 8 and 9.

$$Ac = \frac{TruePositiveValues + TrueNegValues}{TruePosValues + TrueNegValues + FalsePosValues + FalseNegValues} (7)$$

$$Se = \frac{TruePositiveValues}{TruePosValues + FalseNegValues} (8)$$

$$Sp = \frac{TrueNegValues}{TrueNegValues + FalsePosValues} (9)$$

4.1. Experiment setup 1 (150 Features)

In this study, we test a novel approach using a realistic dataset consisting of 150 features. The selected feature vector has dimensions of 1760 by 150. SVM classifiers use a feature set to distinguish between occurrences with late blight and those without. The C-SVM & Q-SVM classifier performed the best, with a success rate of 98.1%. The accuracy of the L-SVM classifier is second best. Table 1 displays a summary of the results of this experiment using 150 attributes. Figure 4.16 below depicts the training durations and prediction speeds of several true classifiers. Confusion matrices for the classifier were shown in Fig. 5.

Table 1
Experiment Results for 150 Feature Approach with Different SVM Classifiers

Classifier	Features	Accuracy	Sensitivity	Specificity
Linear SVM	150	97.00	98.66	94.92
Quadratic SVM	150	98.10	99.09	96.78
Cubic SVM	150	98.10	98.69	96.26
Fine Gaussian SVM	150	56.82	56.82	0
Medium Gaussian SVM	150	97.61	99.08	95.79
Corse-Gaussian SVM	150	95.63	98.53	54.01

4.2. Experiment setup 2 (250 Features)

An innovative approach is tested through its trials in this experiment by being applied to a representative sample of 250 features. The dimension chosen for the selected feature vector is 1760 by 250. SVM classifiers use the feature set to distinguish between late blight occurrences and normal occurrences. In this test, the C-SVM classifier had the highest average performance, at 98.9 percent. The Q-SVM classifier has second- best accuracy in the business. The results of this trial, as measured by 250 criteria, are summarized in Table 2. In the graph below, we see a real classifier, and in Figs. 4.2, 8, we see how long it takes to train the model and how quickly it can make predictions. The confusion matrices with regard to the classifier are given in Fig. 7.

Table 2
Experiment Results for 250 Feature Approach with Different SVM Classifiers

Classifier	Features	Accuracy	Sensitivity	Specificity
Linear SVM	250	97.61	98.68	94.14
Quadratic SVM	250	98.52	99.29	97.54
Cubic SVM	250	98.94	99.60	98.04
Fine-Gaussian	250	56.83	56.82	0
Medium-Gaussian	250	98.42	99.39	97.04
SVM Coarse-Gaussian SVM	250	96.31	98.85	93.28

4.3. Experiment setup 3 (550 Features)

This trial paved the way for a whole new strategy through its tests by applying it to 550 features chosen at random. The selected feature vector has dimensions of 1760 by 550. SVM classifiers are used in the feature set to distinguish late blight outbreaks from normal events. The C-SVM classifier came out on top in this test, with an impres- sive total accuracy of 99.6 percent. When compared to other commercial classifiers, Q-SVM's Accuracy ranks second-best. The results of this trial, as measured by 550 criteria, are summarized in Table 3. Figure 4.3; 10,demonstrate the training and pre- diction times for one such classifier, respectively. As shown in Fig. 9, the confusion matrices for the classifier were provided.

Table 3
Experiment Results for 550 Feature Approach with Different SVM Classifiers

Classifier	Features	Sensitivity	Specificity	Accuracy
Linear SVM	550	98.21	99.19	97.03
Quadratic SVM	550	98.57	99.29	97.54
Cubic SVM	550	99.62	99.50	97.42
Fine-Gaussian	550	56.81	56.82	0
Medium-Gaussian	550	98.24	99.19	96.91
SVM Coarse-Gaussian SVM	550	97	99.17	93.28

4.4. Discussion

This study evaluates a novel method using a substantial data set of 150 features. Dimensions of the selected feature vector are 1760 by 150. SVM classifiers employ the feature set to separate late blight from normal occurrences. The C-SVM & Q-SVM classifier (33) performed the best in this evaluation with a success rate of 98.1%. The Accuracy of the L-SVM classifier is second best. Table 1 provides a brief overview of the results of this experiment based on a set of 150 features. In setup 2 of the experiment, a novel method is put through its paces by being applied to 250 features meant to indicate the whole. The selected feature vector will have dimensions of 1760 by 250. SVM classifiers are used in the feature set to distinguish late blight occurrences from typical occurrences. The C-SVM classifier came out on top in this test, with an impressive total accuracy of 98.9 percent. The Q-SVM classifier is second only to the gold standard in terms of Accuracy—the 250-criterion outcomes of this study. Using 550 randomly selected characteristics, prepare a new technique for testing in experiment setup 3. The feature vector size that was ultimately decided upon is 1760 by 550. Late blight outbreaks may be distinguished from everyday occurrences using SVM classifiers in the feature set. The C-SVM classifier is shown to be the most effective in this test, with an accuracy of 99.6 percent. Q-SVM is a classifier that has the second-best Accuracy in the market. Table 4 provides a concise summary of the 550 criteria-based trial findings.

Table 4
Comparison of proposed method with approaches.

References	Accuracy
(34)	93%
(35)	92%
(36)	97.2%
(37)	97.89%
Proposed	99%

The proposed approach for identifying late blight illness in potato leaves was evaluated in comparison to numerous recent research publications in the area. Kang et al. (34) used the Django framework to create a compact convolutional neural network with a Top-1 detection accuracy of over 93% for diagnosing several potato plant diseases, including late blight. Shoib et al. built a 3D CANet-based Convolutional Neural Network model for disease diagnosis and lesion segmentation using contextual information, with an average accuracy of 92% for lesion segmentation and good accuracy for distinguishing lesion subtypes in various plants. Mahum et al. (20) developed an improved deep learning technique that used an Efficient DenseNet model with re- weighted cross-entropy loss to successfully classify five separate classes of potato leaves, reaching illness detection accuracy of 97.2%. Chakraborty et al. explored deep-learning models for identifying blight infections at different stages; tweaked VGG16 achieved 97.89% accuracy. Kumar et al. suggested a Hierarchical Deep Learning Convolutional Neural Network that outperforms existing illness detection methods by roughly 4%. Our results demonstrate that the proposed model is robust and effective.

5. Conclusion

In this research paper, we aim to classify potato leaf diseases. Specifically, there are two types of leaf diseases that can affect potato plants: early blight (EB) and late blight (LB). In this article we only focus on potato late blight disease. Detecting this disease early on can significantly increase the yield of this crop. As a result, the application of image processing for diagnosing and evaluating these conditions is very desirable. Here, we de-scribe a self-sufficient approach using image processing and machine learning to identify potato leaf late blight. The beginning of the pipeline's four stages, preprocessing, focuses on enhancing images. During the first processing stage, the image is shrunk to 300 X 300 X 3—image enhancement using histogram equalization. Second, we implemented feature engineering using Deep CNN models. Features are extracted using Darknet-53, AlexNet, and Vgg-19. In addition, the deep CNN features are con- catenated in the third step. Feature selection is an EO (equilibrium optimization) task. We next employ many supports vector machine (SVM) variations for classification based on these characteristics. Using image processing and machine learning techniques, the research also proposes a unique autonomous strategy for the early identification of Late Blight disease in potato leaves. The suggested technique, which incorporates Histogram Equalization, Deep CNN feature ex-traction, wrapper-based feature selection, and SVM classification, achieves a remarkable 99% accuracy. This method has significant potential for increasing potato farming productivity by allow- ing for the prompt diagnosis and study of leaf diseases, ultimately leading to increased crop yield and agricultural production globally.

Declarations

Conflicts of Interest:

The authors declare no conflict of interests.

Funding Statement:

No external funding is available for the research.

Author Contribution

Muhammad Ahtsam Naeem; Methodology, Muhammad Ahtsam Naeem; Software, Muhammad Ahtsam Naeem; Validation, Muhammad Ahtsam Naeem,; Formal analysis, Muhammad Imran Sharif; Investigation,; Data curation, Muhammad Imran Sharif and Muhammad Zaheer Sajid; Writing original draft, Muhammad Imran Sharif and Muhammad Zaheer Sajid; Writing review and editing, Muhammad Imran Sharif, Muhammad Zaheer Sajid; Visualization, Shahzad Akber, Muhammad Asim Saleem; Supervision, Shahzad Akber, Muhammad Asim Saleem.

Data Availability Statement:

The PlantVillage Dataset can be freely obtained from:

https://www.kaggle.com/datasets/emmarex/plantdisease.

References

1. Arora, R., Sharma, S., Singh, B.: Late blight disease of potato and its management. Potato J., **41**(1). (2014)
2. Arnaud, S.E., Rehema, N., Aoki, S., Kananu, M.L.: Comparison of Deep Learning Architectures for Late Blight and Early Blight Disease Detection on Potatoes. Open. J. Appl. Sci. **12**(5), 723–743 (2022)

3. Muthoni, J., Nyamongo, D.: A review of constraints to ware Irish potatoes production in Kenya. J. Hortic. forestry. **1**(7), 98–102 (2009)
4. Ugonna, C., Jolaoso, M., Onwualu, A.: A technical appraisal of potato value chain in Nigeria. Int. Res. J. Agricultural Sci. Soil. Sci. **3**(8), 291–301 (2013)
5. Khalifa, N.E.M., Taha, M.H.N., El-Maged, L.M.A., Hassanien, A.E.: Artifi- cial intelligence in potato leaf disease classification: a deep learning approach. Machine Learning and Big Data Analytics Paradigms: Analysis, Applications and Challenges, : p. 63–79. (2021)
6. Hou, C., Zhuang, J., Tang, Y., He, Y., Miao, A., Huang, H., Luo, S.: Recognition of early blight and late blight diseases on potato leaves based on graph cut segmentation. J. Agric. Food Res. **5**, 100154 (2021)
7. Fayyaz, A.M., Al-Dhlan, K.A., Rehman, S.U., Raza, M., Mehmood, W., Shafiq, E., Choi, J.-G.: Leaf Blights Detection and Classifi-cation in Large Scale Applications. Intell. Autom. Soft Comput., **31**(1). (2022)
8. Tsiafouli, M.A., et al.: Optimizing the delivery of multiple ecosystem goods and services in agricultural systems. Front. Media SA p. 97. (2017)
9. Sj¨olin, K., Rayner, M., Purhagen, J., Sj¨oholm, I.: A Review of Cooking of Potatoes (Solanum tuberosum L.) Served in Large-Scale Food-Service Sys- tems, Including Industrial Pre-Treatments. J. Food Eng. Technol. **11**(1), 22–35 (2022)
10. Zuorro, A., Malavasi, V., Cao, G., Lavecchia, R.: Use of cell wall degrading enzymes to improve the recovery of lipids from Chlorella sorokiniana. Chem. Eng. J. **377**, 120325 (2019)
11. Maffei, G., Bracciale, M.P., Broggi, A., Zuorro, A., Santarelli, M.L., Lavecchia, R.: Effect of an enzymatic treatment with cellulase and mannanase on the struc- tural properties of Nannochloropsis microalgae. Bioresour. Technol. **249**, 592–598 (2018)
12. Mehariya, S., Fratini, F., Lavecchia, R., Zuorro, A.: Green extraction of value-added compounds form microalgae: A short review on natural deep eutec- tic solvents (NaDES) and related pre-treatments. J. Environ. Chem. Eng. **9**(5), 105989 (2021)
13. Rani, A., Saini, K.C., Bast, F., Mehariya, S., Bhatia, S.K., Lavecchia, R., Zuorro, A.: Microorganisms: A potential source of bioactive molecules for antioxidant applications. Molecules. **26**(4), 1142 (2021)
14. El Moutaouakil, K., Jabir, B., Falih, N.: A convolutional neural networks- based approach for potato disease classification. in International Conference on Business Intelligence. Springer. (2022)
15. Hou, B., Hu, Y., Zhang, P., Hou, L.: Potato Late Blight Severity and Epidemic Period Prediction Based on Vis/NIR Spectroscopy. Agriculture. **12**(7), 897 (2022)
16. Mora-Romero, G.A., F´elix-Gast´elum, R., Bomberger, R.A., Romero-Ur´ıas, C., Tanaka, K.: Common potato disease symptoms: ambiguity of symptom-based identification of causal pathogens and value of on-site molecular diagnostics. J. Gen. Plant Pathol. **88**(2), 89–104 (2022)
17. Islam, M., Dinh, A., Wahid, K., Bhowmik, P.: Detection of potato diseases using image segmentation and multiclass support vector machine. in 2017 IEEE 30th canadian conference on electrical and computer engineering (CCECE). IEEE. (2017)

18. Zhou, B., Lapedriza, A., Xiao, J., Torralba, A., Oliva, A.: Learning deep fea- tures for scene recognition using places database. Adv. Neural. Inf. Process. Syst., 27. (2014)
19. Shoaib, M., Shah, B., Hussain, T., Ali, A., Ullah, A., Alenezi, F., Gechev, T., Ali, F., Syed, I.: A deep learning-based model for plant lesion segmentation, subtype identification, and survival probability estimation. Front. Plant Sci. **13**, 1095547 (2022)
20. Mahum, R., Munir, H., Mughal, Z.-U.-N., Awais, M., Khan, F.S., Saqlain, M., Mahamad, S., Tlili, I.: A novel framework for potato leaf disease detection using an efficient deep learning model. Hum. Ecol. Risk Assessment: Int. J. **29**(2), 303–326 (2023)
21. Chakraborty, K.K., Mukherjee, R., Chakroborty, C., Bora, K.: Automated recog- nition of optical image based potato leaf blight dis-eases using deep learning. Physiol. Mol. Plant Pathol. **117**, 101781 (2022)
22. Kumar, A., Patel, V.K.: Classification and identification of disease in potato leaf using hierarchical based deep learning convolutional neural network. Multimedia Tools Appl., : p. 1–27. (2023)
23. Deng, J., Dong, W., Socher, R., Li, L., Li, K., Li, F., Imagenet: A large-scale hierarchical image database. in IEEE conference on computer vision and pattern recognition. 2009. Ieee. (2009)
24. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet Classification with Deep Convolutional Neural Networks (AlexNet).
25. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, (2014)
26. Too, J., Mirjalili, S.: General learning equilibrium optimizer: a new feature selection method for biological data classification. Appl. Artif. Intell. **35**(3), 247–263 (2021)
27. Keerthi, S.S., Shevade, S.K., Bhattacharyya, C., Murthy, K.R.K.: Improvements to Platt's SMO algorithm for SVM classifier design. Neural Comput. **13**(3), 637–649 (2001)
28. Mohanty, S.: PlantVillage Dataset
29. Wang, Y.-S., Tai, C.-L., Sorkine, O., Lee, T.-Y.: Optimized scale-and-stretch for image resizing, in ACM SIGGRAPH Asia 2008 pa-pers. p. 1–8. (2008)
30. Tekalp, A.M.: Digital video processing. Prentice Hall (2015)
31. De La Torre, A., Peinado, A.M., Segura, J.C., Perez-Cordoba, J.L., Benitez, M.C., Rubio, A.J.: Histogram equalization of speech representation for robust speech recognition. IEEE Trans. Speech Audio Process. **13**(3), 355–366 (2005)
32. Tasnim, Z., Chakraborty, S., Shamrat, F.M.J.M., Chowdhury, A.N., Nuha, H.A., Karim, A., Binte, S.: Deep learning predictive model for colon cancer patient using CNN-based classification. Int. J. Adv. Comput. Sci. Appl. **12**, 687–696 (2021)
33. Johari, S.N.A., Mohd, S., Khairunniza-Bejo, A.R.M., Shariff, N.A., Husin: Mohamed Mazmira Mohd Masri, and Noorhazwani Kamarudin. Detection of Bagworm Infestation Area in Oil Palm Plantation Based on UAV Remote Sensing Using Machine Learning Approach. Agriculture. **13**(10), 1886 (2023)

34. Kang, F., Li, J., Wang, C., Wang, F.: A Lightweight Neural Network-Based Method for Identifying Early-Blight and Late-Blight Leaves of Potato. Appl. Sci. **13**(3), 1487 (2023)
35. Shoaib, M., Shah, B., Hussain, T., Ali, A., Ullah, A., Alenezi, F., Gechev, T.: Farman Ali, and Ikram Syed. A deep learning- based model for plant lesion segmentation, subtype identification, and survival probability estimation. Front. Plant Sci. **13**, 1095547 (2022)
36. Mahum, R., Munir, H., Mughal, Z.-U.-N., Awais, M., Khan, F.S., Saqlain, M.: Saipuni-dzam Mahamad, and Iskander Tlili. A novel framework for potato leaf disease detection using an efficient deep learning model. Hum. Ecol. Risk Assessment: Int. J. **29**(2), 303–326 (2023)
37. Chakraborty, K., Kashyap, R., Mukherjee, C., Chakroborty, Bora, K.: Automated recognition of optical image based potato leaf blight diseases using deep learning. Physiol. Mol. Plant Pathol. **117**, 101781 (2022)

Figures

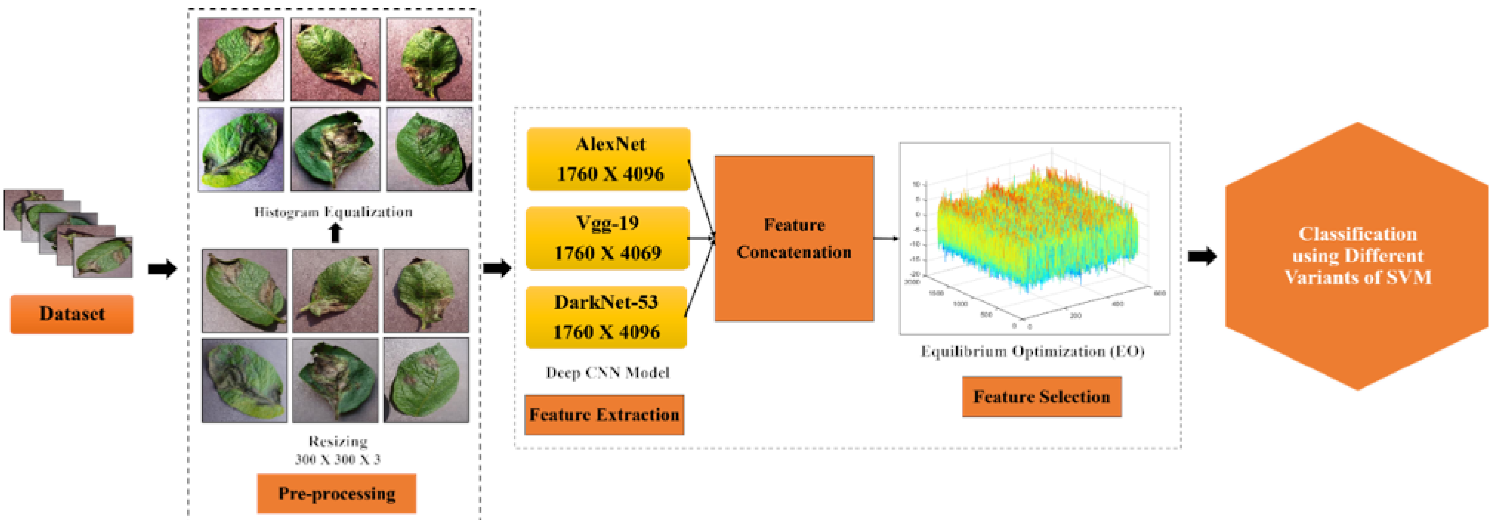


Figure 1

Proposed Deep Learning model for Potato leaf disease identification

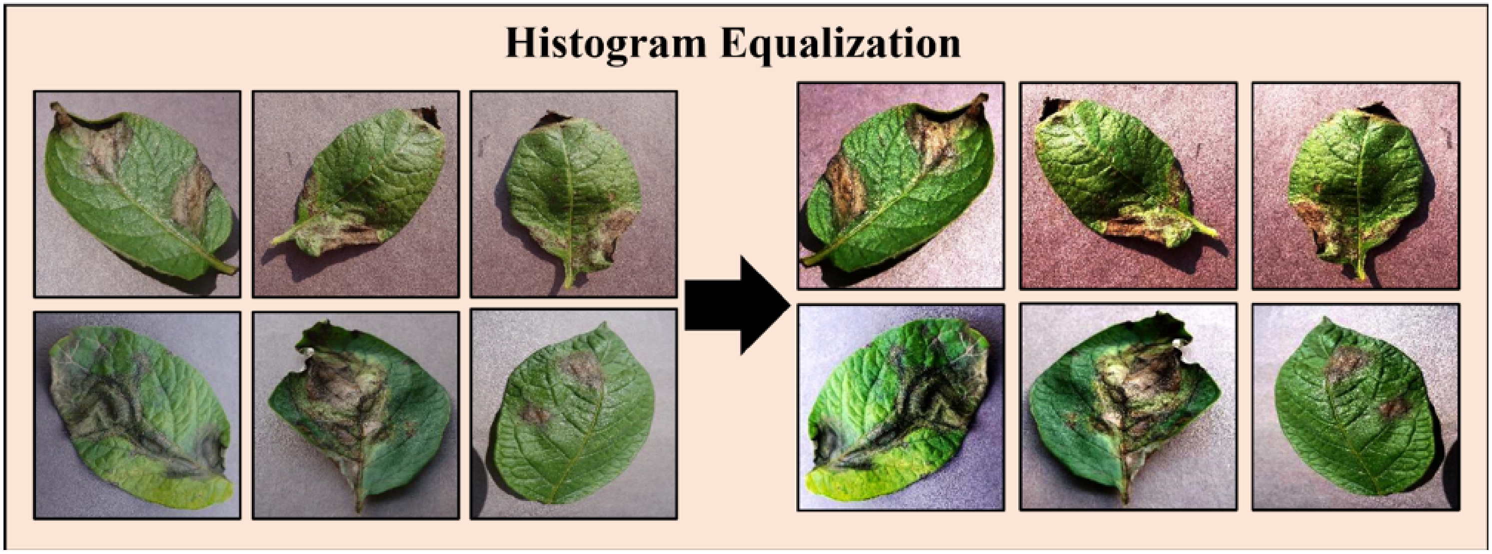


Figure 2

Leaves late blight images after histogram equalization for enhanced contrast

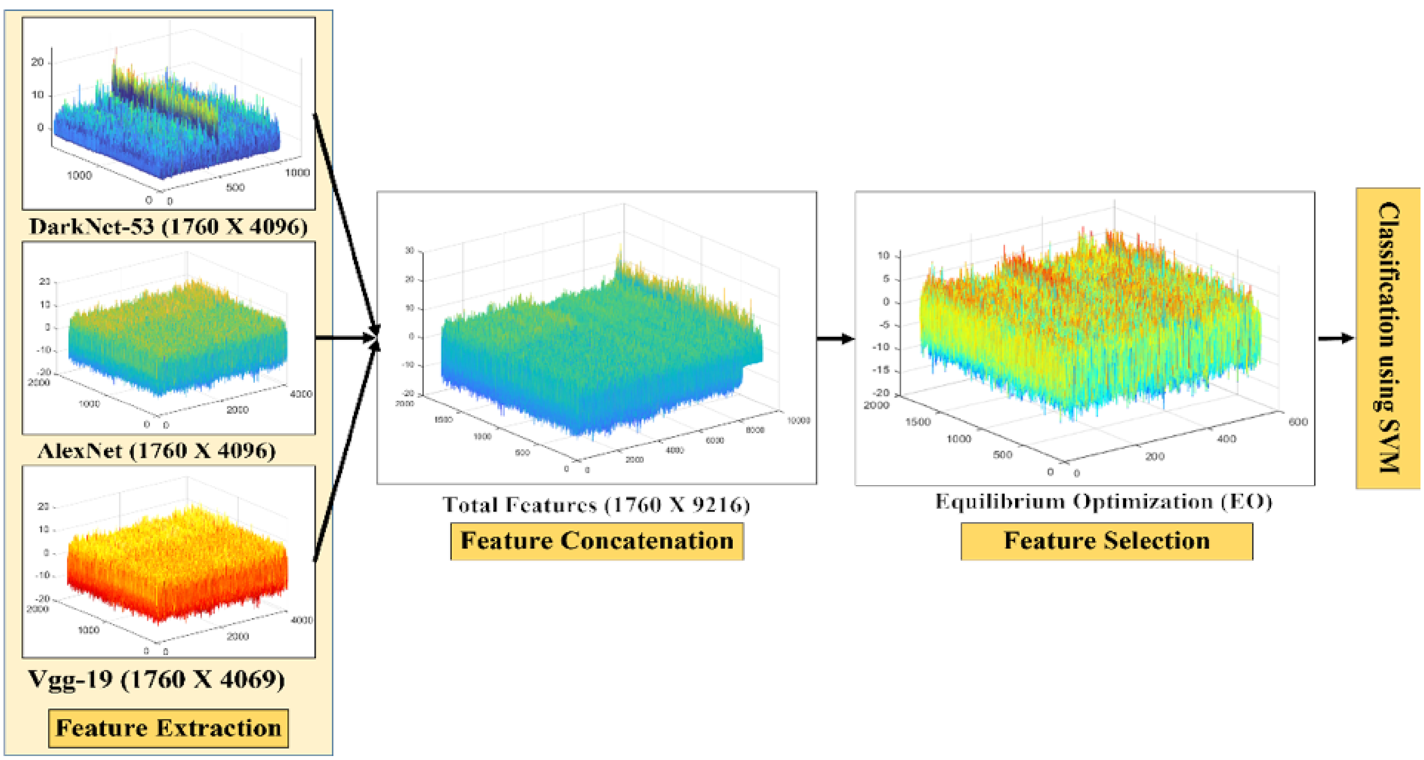


Figure 3

Mesh representation illustrating the process of feature engineering.

Figure 4

(a) Healthy Potato Leaves Images, (b) Late Blight Potato Leaves Images

True Class	Potato Late Blight	960	14	True Class	Potato Late Blight	975	25	True Class	Potato Late Blight	979	29
	Potato Healthy	13	747		Potato Healthy	9	751		Potato Healthy	13	747
Test Case 1 L-SVM		Potato Late Blight	Potato Healthy	Test Case 2 Q-SVM		Potato Late Blight	Potato Healthy	Test Case 3 C-SVM		Potato Late Blight	Potato Healthy
		Predicted Class				Predicted Class				Predicted Class	
True Class	Potato Late Blight	1000	0	True Class	Potato Late Blight	967	33	True Class	Potato Late Blight	937	63
	Potato Healthy	760	0		Potato Healthy	9	751		Potato Healthy	14	74
Test Case 4 FG-SVM		Potato Late Blight	Potato Healthy	Test Case 5 MG-SVM		Potato Late Blight	Potato Healthy	Test Case 6 CG-SVM		Potato Late Blight	Potato Healthy
		Predicted Class				Predicted Class				Predicted Class	

Figure 5

Displayed confusion matrices (150 Features) with respect to classifier

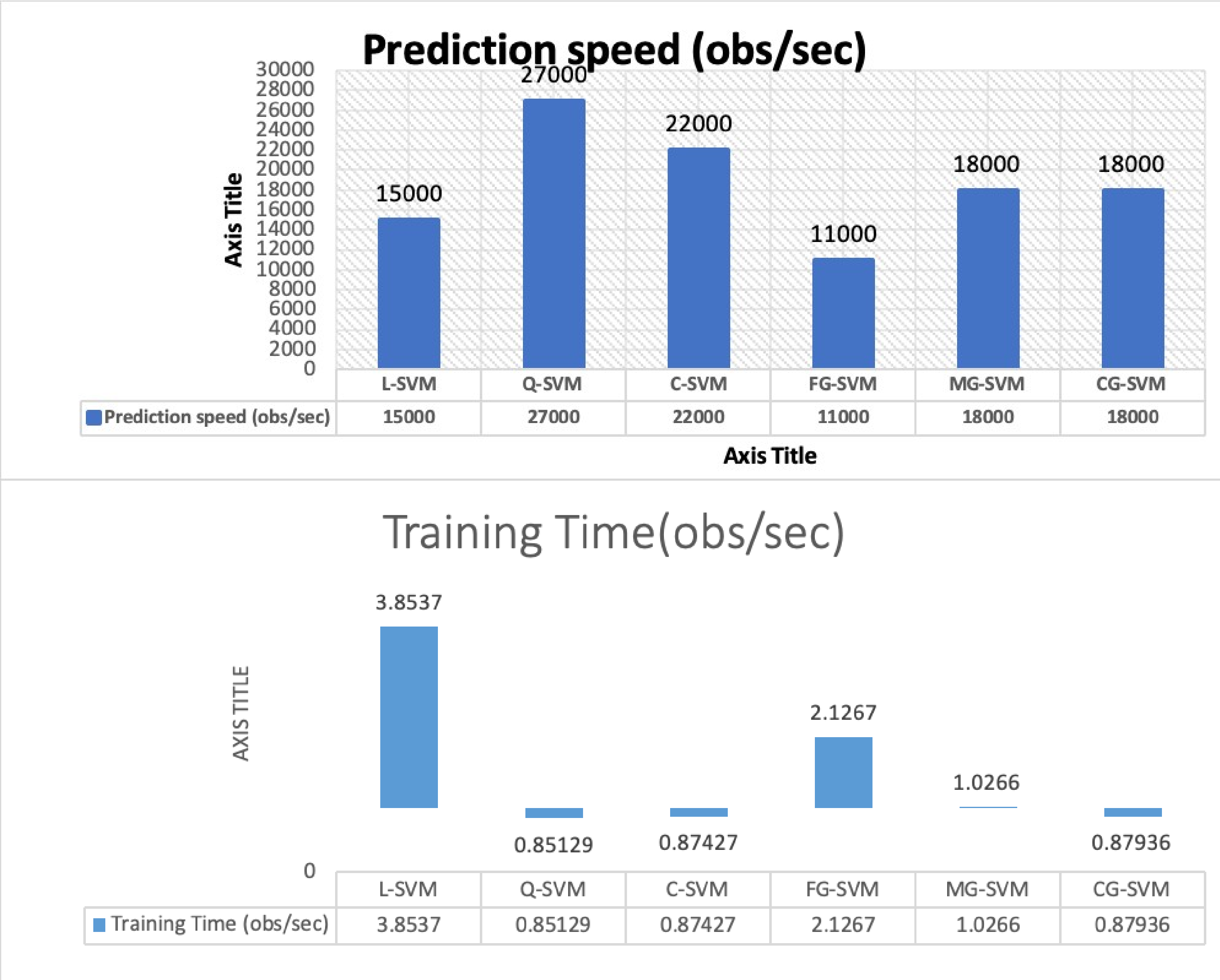
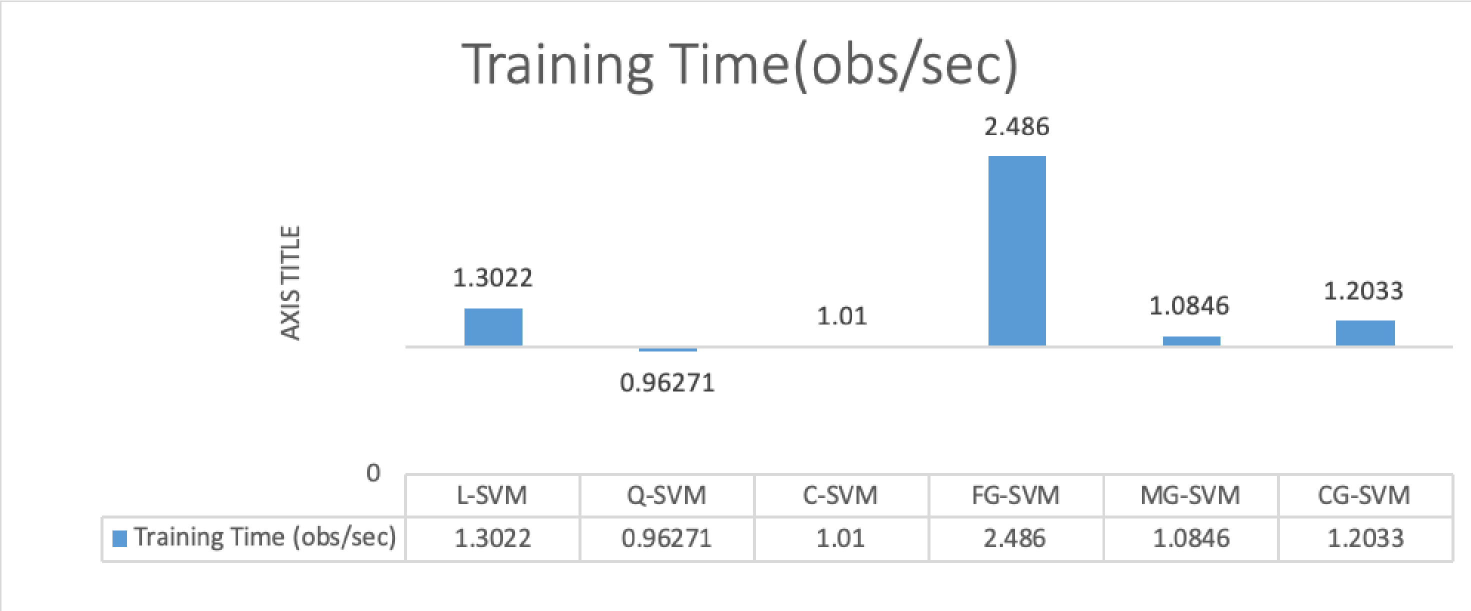
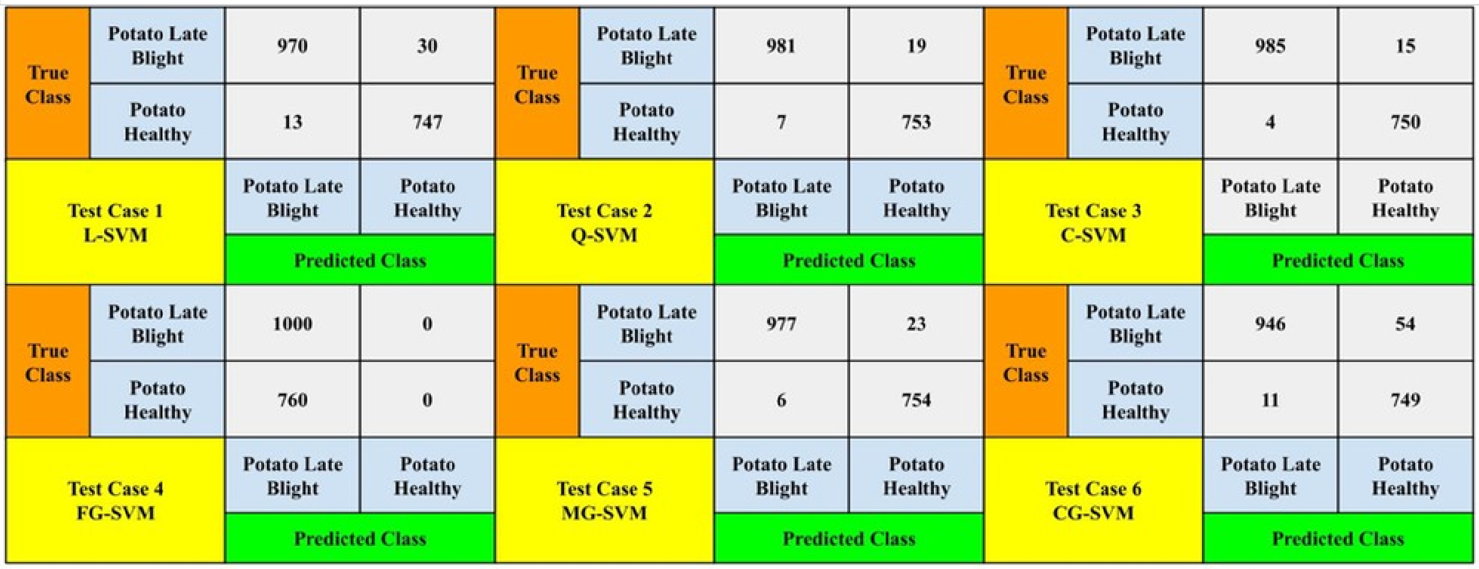


Figure 6

Comparison of training and prediction times for 150 features across classifiers





Test Case 1 L-SVM

True Class	Potato Late Blight	Potato Healthy
Potato Late Blight	970	30
Potato Healthy	13	747
	Predicted Class	

Test Case 2 Q-SVM

True Class	Potato Late Blight	Potato Healthy
Potato Late Blight	981	19
Potato Healthy	7	753
	Predicted Class	

Test Case 3 C-SVM

True Class	Potato Late Blight	Potato Healthy
Potato Late Blight	985	15
Potato Healthy	4	750
	Predicted Class	

Test Case 4 FG-SVM

True Class	Potato Late Blight	Potato Healthy
Potato Late Blight	1000	0
Potato Healthy	760	0
	Predicted Class	

Test Case 5 MG-SVM

True Class	Potato Late Blight	Potato Healthy
Potato Late Blight	977	23
Potato Healthy	6	754
	Predicted Class	

Test Case 6 CG-SVM

True Class	Potato Late Blight	Potato Healthy
Potato Late Blight	946	54
Potato Healthy	11	749
	Predicted Class	

Figure 7

Displayed confusion matrices (250 Features) with respect to classifier

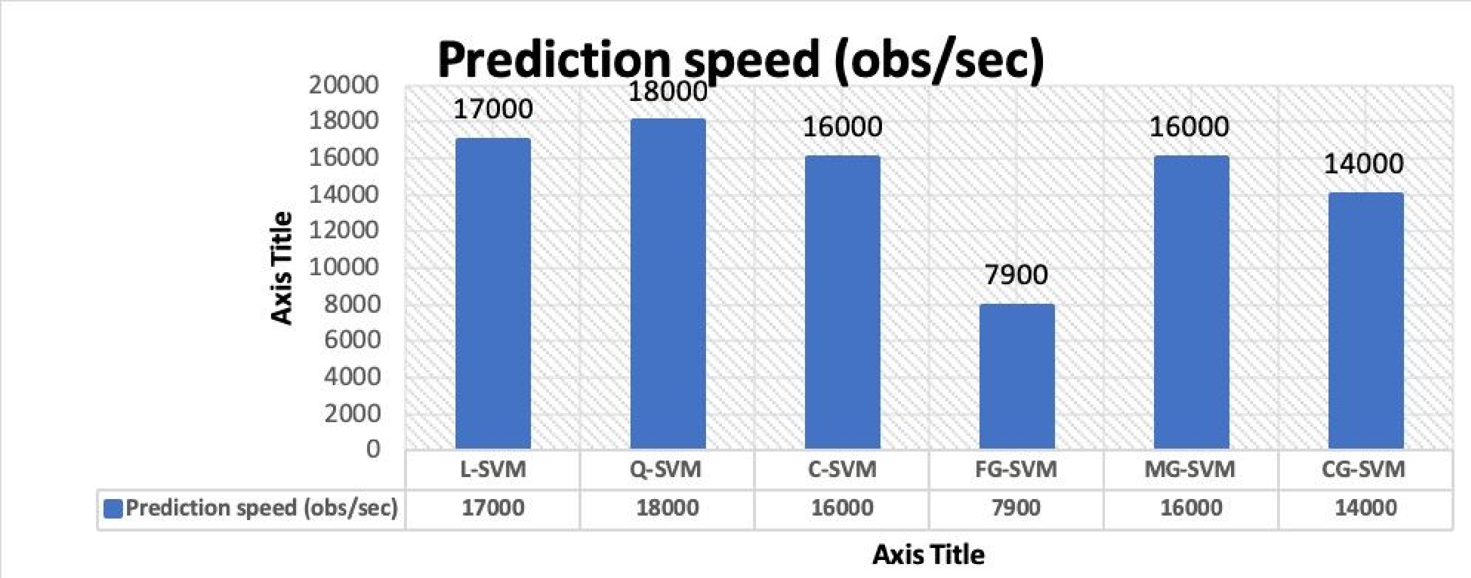


Figure 8

Comparison of training and prediction times for 250 features across classifiers

Test Case 1 L-SVM

True Class	Predicted Class: Potato Late Blight	Predicted Class: Potato Healthy
Potato Late Blight	977	23
Potato Healthy	8	752

Test Case 2 Q-SVM

True Class	Predicted Class: Potato Late Blight	Predicted Class: Potato Healthy
Potato Late Blight	981	19
Potato Healthy	7	753

Test Case 3 C-SVM

True Class	Predicted Class: Potato Late Blight	Predicted Class: Potato Healthy
Potato Late Blight	990	20
Potato Healthy	5	755

Test Case 4 FG-SVM

True Class	Predicted Class: Potato Late Blight	Predicted Class: Potato Healthy
Potato Late Blight	1000	0
Potato Healthy	760	0

Test Case 5 MG-SVM

True Class	Predicted Class: Potato Late Blight	Predicted Class: Potato Healthy
Potato Late Blight	976	24
Potato Healthy	8	752

Test Case 6 CG-SVM

True Class	Predicted Class: Potato Late Blight	Predicted Class: Potato Healthy
Potato Late Blight	955	45
Potato Healthy	8	752

Figure 9

Displayed confusion matrices (550 Features) with respect to classifier

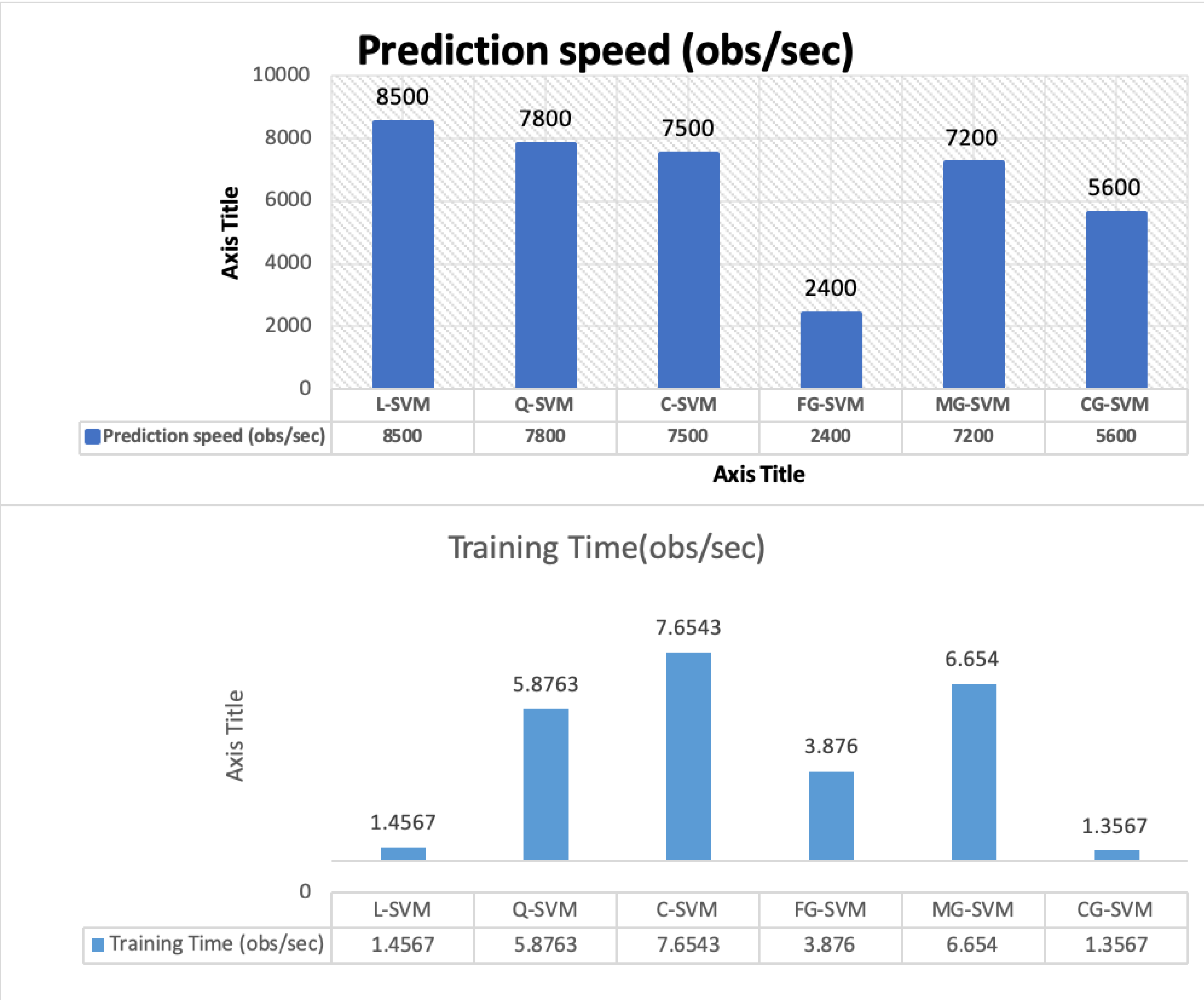


Figure 10

Comparison of training and prediction times for 550 features across classifiers