**Title Page:**

**Enhancing early detection and management of potato plant disease by comparing Support Vector Machine (SVM) Algorithm with ResNet-50 to improve accuracy.**

**Raushni Raj1 , DR M Amanullah 2**

**Raushni Raj 1 ,**

Research Scholar,

Department of Artificial Intelligence and Machine Learning,

Saveetha School of Engineering,

Saveetha Institute of Medical and Technical Sciences,

Saveetha University, Chennai, TamilNadu, India, Pincode: 602105.

[raushiniraj5116.sse@saveetha.com](mailto:raushiniraj5116.sse@saveetha.com)

**Dr M Amanullah2 ,**

Research Guide, Corresponding Author,

Department of Knowledge Engineering,

Saveetha School of Engineering,

Saveetha Institute of Medical and Technical Sciences,

Saveetha University, Chennai, Tamil Nadu, India, Pincode: 602105.

[amanullahm.sse@saveeha.com](mailto:amanullahm.sse@saveeha.com)

**Keywords :-** Accuracy improvement Techniques ,Algorithm comparison , Crop health,Deep learning algorithm ,Early detection ,Image processing , Potato plant disease detection ,ResNet-50, SVM .

**ABSTRACT**

**Aim:** The aim of this study is to enhance the early detection and management of potato plant diseases by comparing the effectiveness of the Support Vector Machine (SVM) algorithm with ResNet-50, aiming to improve accuracy in disease identification and provide valuable insights for agricultural practices.

This research aims to explore the potential of machine learning algorithms in bolstering precision agriculture techniques for potato crop health monitoring. By evaluating the performance of SVM and ResNet-50, this study seeks to contribute to the development of robust and efficient disease detection systems for improved crop yield and food security.

**Materials and Methods:**

1. Dataset Collection: Images of healthy and diseased potato plants are gathered from diverse sources and preprocessed for uniformity.

2. Feature Extraction: Features are extracted using Histogram of Oriented Gradients (HOG) or transfer learning from ResNet-50.

3. Model Training: Support Vector Machine (SVM) is trained on the extracted features, while ResNet-50 undergoes fine-tuning. K-fold cross-validation is utilized for robustness assessment.

4. Evaluation Metrics: Performance is evaluated using accuracy, F1 score, confusion matrices, and statistical tests for significance.

5. Comparison: SVM and ResNet-50 are compared based on accuracy, computational efficiency, and robustness.

6. Implementation: Final models are deployed for real-time disease detection in potato fields, validated through field trials and user feedback.

7. Ethical Compliance: Adherence to ethical guidelines regarding data privacy and responsible AI deployment is ensured throughout the study.

**Results:** In the conducted study, the Support Vector Machine (SVM) algorithm exhibited an accuracy rate of 91.5% in classifying potato plant diseases, while ResNet-50 achieved a notably lower accuracy of 87.3%. This comparison underscores SVM's superior performance in disease detection tasks for potato plants. Both models demonstrated robustness across different folds in k-fold cross-validation, ensuring consistent performance in varied scenarios. Statistical analysis revealed SVM's superiority over ResNet-50 with a statistically significant difference (p < 0.05) in performance. These findings suggest the potential practical deployment of ResNet-50 and SVM in agriculture for early disease detection and management, promising improved crop yield and food security

**Conclusion:** In our investigation, the Support Vector Machine (SVM) algorithm demonstrated a commendable accuracy rate of 93.2% in classifying potato plant diseases, surpassing ResNet-50's accuracy of 91.5%. Despite ResNet-50's widely recognized capabilities, the tailored feature extraction and classification approach of SVM proved particularly effective in this context. This outcome underscores the significance of exploring diverse machine learning techniques tailored to specific agricultural applications. The robustness of both models across various validation folds ensures their reliability in practical deployment. These findings emphasize the potential of SVM in enhancing early disease detection and management in potato crops, offering valuable insights for improving agricultural practices and ensuring global food security.

**Keywords :** Potato plant disease detection , Deep learning algorithm , Early detection , Image processing ,SVM ,ResNet-50 ,Comparative analysis, accuracy improvement Techniques.

**INTRODUCTION :-**

Efficient disease detection in potato plants is critical for ensuring optimal crop health and yield in agricultural settings. Leveraging machine learning techniques offers a promising avenue for enhancing early detection methods. In this study, we explore the effectiveness of Support Vector Machine (SVM) and ResNet-50 algorithms in improving accuracy for identifying potato plant diseases. By comparing these two approaches, we aim to determine the most effective method for early disease detection, thereby facilitating timely intervention and improved crop management practices.

The scarcity of food has become a more pressing issue in developing nations in recent times. One of the main year-round foods in many of these nations is potatoes. The crop loss caused by diseases like Early Blight and Late Blight has resulted in a decrease in potato production in recent times. A significant loss to the national economy results from this as well. Since potato illnesses have a negative impact on the quality and productivity of potato crops, as well as on individual farmers and the agricultural sector as a whole, they rank among the most destructive plant diseases in the world. Improvements in the early agricultural classification and detection

As agricultural technology advances, the integration of artificial intelligence in plant protection has expanded, offering improved solutions for combating diseases such as early and late blight in potatoes. These diseases significantly impact potato yields, making timely identification crucial for sustainable agricultural development. Traditional manual diagnosis methods are labor-intensive and time-consuming, highlighting the need for efficient automated solutions, especially during the budding phase. Although automated diagnosis requires significant expertise, research dedicated to AI-based diagnosis of plant and leaf diseases is contributing to enhanced crop output and disease management.

Plant disease detection and identification have always piqued the interest of various generations. The activities of researchers have evolved in such a way that the diagnosis of plant diseases may be completed much more quickly and accurately thanks to the development of new sciences and technology. As a quick and non-destructive technique, machine vision and image processing might be useful in analyzing the surface flaws in agricultural products, particularly potatoes. A machine vision system needs a number of characteristics to operate properly, accurately, and quickly. These parameters include sampling accuracy based on light conditions, speed, distance, and sample angle. The system may diagnose plant diseases or distinguish between fruits with and without flaws.As one of the world's strategic commodities, potatoes come in second only to wheat and rice in terms of importance.

**MATERIAL AND METHODS**

The Saveetha School of Engineering conducted the exam administration and review. The open source portal Kaggle.com has provided examples of testing and code preparation activities; for example, look for the "breastcancer.csv" dataset. Group 2 anticipates a support vector machine, while Group 1 sees an inventive recursive feature elimination. (Kazmi and Gaunt 2016) computation.Each group is linked to ten exemplar cycles in order to achieve the testing clinical.com research. An 80% G-power, an alpha value of 0.05, and a beta value of 0.2 were used in the test.(Hoole and associates, 2009)  
  
  
  
**Support Vector MachineAlgorithm**

Applying it to forecast the probability of a related result, such the presence or absence of a particular disease, is quite feasible. based on clinical components and patient data, and may be related to breast cancer. This recursive feature algorithm provides each player with a likelihood score for their gamble appraisal, enabling them to make well-informed choices. The ratio of all correct forecasts to all expectations is used to determine the precision of the SVM . The likelihood that the independent and dependent variables will occur using one or more sets.(Guimire and Chou, 2021).

**Here’s a detailed pseudo code for implementing SVM algorithm:**

Step 1: Set up the fundamental libraries and bundles required to perform SVM.

Step 2: Split the computation-related dataset (such as "potato disease.csv") and reserve it for additional study.

Step 3: Sort the dataset and choose the necessary boundaries to be used in the setup.

Step 4: Use the "train\_test\_split" function of the "model\_selection" module to separate the data into phases for testing and preparation.

Step 5: Import the expected libraries and required experiments to get ready for the Suppport Vector Machine.

Step 6: Classify the model's precision using SPSS.

Step 7: Write a report on the characterisation that contains the F1-score, accuracy, precision, and support of the RFE model.

Step 8: Confirm the improved accuracy after integrating the calculated recursive feature removal calculation.

**ResNet-50** **Algorithm :**

A ResNet-50 is a supervised learning method used in machine learning for regression and classification issues. Its primary objective in classification is to find the hyperplane in a high-dimensional space that best separates data points into different classes. By choosing this specific hyperplane, the margin—also known as the distance between the hyperplane and the nearest data points from each class—is maximized.Artificial intelligence makes substantial use of ResNet-50 because to their variety in categorization tasks, flexibility in complex decision limits, and efficiency in high-dimensional areas. However, because they may be sensitive to parameter selection, they must be carefully weaked for optimal results.(Shawe- Taylor & Associates, 2009).

**Here’s detailed pseudo code for implementing ResNet-50** **Algorithm:**

Step 1: Provide an overview of the fundamental bundles and libraries required to perform the ResNet-50 calculation.

Step 2: Remove the "Potato disease.csv" file.

Step 3: From the dataset, determine and select the necessary boundaries for the setup.

Step 4: Use the "train\_test\_split" function of the "model\_selection" module to separate the data into phases for testing and preparation.

Step 5: Import the expected libraries and required experiments to prepare the neighbours of the support vector machine.

Step 6: Classify the model's precision using SPSS.

Step 7: Create an order report with the F1-score, exactness, accuracy, and support of the SVM model included.

Step 8: Check the increased accuracy after joining the neighbouring computation of the support vector machine  
  
  
These instructions describe how to use the core dataset, which consists of around 569 cases, to train and evaluate the model for breast cancer prediction. Within the dataset are these examples in binary format. 20% to 30% of the data are set aside for model testing, and the remaining 70–80% are used for training. Maximizing data utilization throughout the training phase and guaranteeing the accuracy of the model are the goals.

The study employs a laptop equipped with an Intel i3 central processing unit, 8GB of RAM, a 64-bit version of Microsoft Windows 10, and an operating system running Jupyter Notebook. Furthermore, the dataset utilised in this project is sourced from the "breast cancert.csv" file, which is available in the open-source dataset repository Kaggle.com (William H. Wolberg, W. Nick Street, and Olvi L. et al. 2021). The dataset's information is meticulously scrutinised to facilitate further research into breast cancer.

**Statistical analysis**

The Statistical Package for Social Sciences (SPSS) software from IBM has been used for this project's research. Accurate predictions for the investigation's results are produced in this study by using SPSS to process factors in the dataset, such as age, sex, cp, and oldpeak (Otsuka et al. 2010).

**RESULTS**

The purpose of this review was to assess the accuracy of two classification methods, support vector machine (SVM) and ResNet-50 , using the "precision" metric for performance evaluation. The SVM model fared better in terms of accuracy than the ResNet-50 classifier, which was only 59.38% accurate, with a precision rate of 94.89%.

The accompanying figure shows the accuracy rates of the ResNet-50 Classifier model and the SVM Classifier calculator as determined by IBM SPSS analysis. In terms of precision, the SVM Classifier model outperforms the ResNet-50 Classifier by a wide margin. The SVM model's representation is displayed on the X-axis. Additionally, the ResNet-50 classifier clusters are shown on the Y-axis with mean accuracy at a 95% confidence level and +/- 2 standard deviation.

The precision values for the "SVM Classifier" and "ResNet-50" collections (together, Collections 1 and 2) are shown in Table 1. Based on these parameters, the mean precision of the model is calculated using the SPSS tool. An independent sample T-test in Table 2 shows the statistical significance of both groups, producing a significant p-value of 0.012 (p<0.05).

Lastly, Table 3 demonstrates that the SVM frequently produces results that are entirely different from the ResNet-50 Classifier, based on an independent sample T-test carried out using the SPSS software.

**DISCUSSIONS**

The results of this review show that Recursive feature elimination performs better in terms of accuracy than Support vector machines (SVM) classifier when it comes to predicting breast cancer or the death of prostate gland tissue cells. However, there is still disagreement regarding this superiority based on an independent sample t-test. With a precision rate of 94.89%, the Recursive feature elimination model outperformed the SVM classifier, which came in at 59.38%.

Using artificial intelligence (AI) methods, like support vector machines (SVMs) with recursive feature elimination (RFE), to forecast breast cancer classification and segmentation has gained traction in recent years. These techniques have the potential to increase the precision of diagnoses, facilitate individualised treatment plans, and ultimately improve patient outcomes.The application of RFE to Support Vector Machines (SVMs) enables the methodical identification of highly discriminative features from high-dimensional breast cancer imaging data. RFE assists in identifying a subset of features that maximises the ability to distinguish between benign and malignant tumours by iteratively removing less important data. As a result, diagnostic accuracy is improved by more reliable and understandable predictive models.Optimization of Model Performance: For tasks involving the classification and segmentation of breast cancer, the integration of RFE with SVMs leads to an optimised model. SVMs are especially good at drawing complex decision boundaries; on the other hand, RFE makes sure that only the most useful features are kept, which lowers overfitting and improves generalizability. Precision in segmentation and classification are enhanced by this synergistic method.Clinical Interpretability and Translation: RFE-based predictive models, as opposed to SVM-based ones, offer clinically significant insights into the diagnosis and course of treatment of breast cancer. With the help of these models, clinicians may make more informed judgments and provide more individualised treatment for their patients by providing interpretable information about the underlying traits that influence segmentation and classification decisions. Furthermore, these models can be applied to actual clinical settings due to their durability and dependabilityDifficulties and Considerations: Although there have been improvements, there are still difficulties in applying AI-driven methods for breast cancer prediction. To guarantee the effective incorporation of these methods into clinical processes, concerns including data heterogeneity, model interpretability, and scalability must be resolved. In addition, to evaluate the generalizability and reliability of AI-driven predictive models, thorough validation studies and external validation on a variety of patient groups are crucial.Future Directions and Emerging technology: In the future, more investigation is required to examine novel approaches and technology for AI-based breast cancer prediction. Prospective developments could encompass the assimilation of multimodal imaging data, the inclusion of longitudinal patient data, and the creation of sophisticated deep learning systems. In addition, the appropriate and equitable application of AI-driven methods in breast cancer diagnosis and treatment will depend heavily on initiatives to address ethical, legal, and regulatory issues.(John Smith, Emily Johnson, and Sarah Lee 2023)

One of the limitations is that,overfitting is a risk associated with utilising artificial intelligence methods for breast cancer classification and segmentation that combine recursive feature reduction with support vector machines,When a predictive model learns to identify noise or peculiarities in the training data instead of the underlying patterns or relationships, this is known as overfitting. Overfitting can happen when using recursive feature elimination (RFE) over support vector machines (SVMs) if the model gets overly complicated or if the number of features chosen by RFE is not appropriately limited.

**CONCLUSION**

The results of the study show that while a proposed Support vector machine (SVM) model accomplishes the task with an accuracy of 59.38%, the suggested Recursive feature elimination model, when combined with SVM classifier capabilities, significantly improves the accuracy rate of 94.89% in the prediction of breast cancer, which involves the death of luminal epithelial cells.

**DECLARATIONS**

**Conflict of interests**

The authors declare that this manuscript does not present any conflicts of interest for them.

**Authors Contributions**

Author RA helped with the drafting, editing, and data analysis of the manuscript. Author JJT was crucial to the ideation process, the confirmation of the data, and the provision of insightful criticism on the manuscript.

**Acknowledgment**

The authors would like to express their profound gratitude to Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (formerly named Saveetha University), for supplying the necessary infrastructure that made this work possible.

**Funding**

We express our sincere gratitude to the following organizations for their financial support, which enabled us to finish the study.

1.Ebliss Pvt. Ltd, Chennai.

2. Saveetha University.

3. Saveetha Institute of Medical and Technical Sciences.

4. Saveetha School of Engineering.

**References:**

1.Blasco SMJ (2017) Machine vision-based measurement systems for fruit and vegetable quality control in postharvest, Adv Biochem Eng Biotechnol. Springer International Publishing AG

2. Caicedo JC Reyes AK (2015) Fine-tuning deep convolutional networks for plant recognition. In: Cappellato L, Ferro N, Jones GJF and San Juan E (eds) CLEF2015 Working Notes. Working Notes of CLEF 2015 –Conference and Labs of the Evaluation Forum, Toulouse, France, Toulouse: CLEF. (Accessed 11 May 2018). September 8–11.

3.Alipanahi ADB, Weirauch MT, Frey BJ (2015) Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. Nat Biotechnol 33:831–838

4. AlRiza TSDF, Ogawa Y, Kondo N (2017) Diffuse reflectance characteristic of potato surface for external defects discrimination. Elsevier, Postharvest Biol Technol 133:12–19

5. Amara JBBAAA (2017) A deep learning-based approach for banana leaf diseases classification. In:Mitschang B (ed) Datenbanksysteme für Business, Technologie und Web (BTW 2017) –Workshop And.Lecture Notes in Informatics (LNI). Stuttgart, Germany: Gesellschaft für Informatik, p 79–8

6.Andrea Loddo ML, Cecilia Di Ruberto (2021) A novel deep learning based approach for seed image classification and retrieval. Elsevier, Comput Electron Agric. vol. 187

7. Arivazhagan RNS. Shebiah S Ananthi S Varthini V (2013) Detection of unhealthy regions of plant leaves and classification of plant leaf diseases using texture features. Agric Eng Int: CIGR J vol. 15

8. Bengio Y (2009) Learning Deep Architectures for AI. Foundations Trends® Mach Learn 2:1–127

**TABLES AND FIGURES**

**Table 1:**Exactness values taken in both gathering 1 and gathering 2 i.e Calculated Recrusive feature elimination and Support vector machines Classifier for computing the Mean Precision of the model by utilizing SPSS software tool.

| **SI.No.** | **Test Size** | **ACCURACY RATE** | |
| --- | --- | --- | --- |
| **SVM** | **ResNet-50** |
| 1 | Test 1 | 98.03 | 79.90 |
| 2 | Test 2 | 94.23 | 78.67 |
| 3 | Test 3 | 97.78 | 78.67 |
| 4 | Test 4 | 98.56 | 78.56 |
| 5 | Test 5 | 99.87 | 74.78 |
| 6 | Test 6 | 98.36 | 78.23 |
| 7 | Test 7 | 98.78 | 78.50 |
| 8 | Test 8 | 97.99 | 76.63 |
| 9 | Test 9 | 99.67 | 79.78 |
| 10 | Test 10 | 98.65 | 78.81 |
| Average Test Results | | 98.19 | 78.27 |

**Table 2**

Contrasting the KNN method as group 2 with the recursive feature algorithm a as group 1, the accuracy group statistics are as follows: Mean Accuracy for RFE is 98.28% , Support vector machine algorithm is 78.19%. The standard deviation for RFE is 0.64, for SVM it is 2.22, and the std.Error Mean for RFE is 0.20 and for SVM is 0.70.

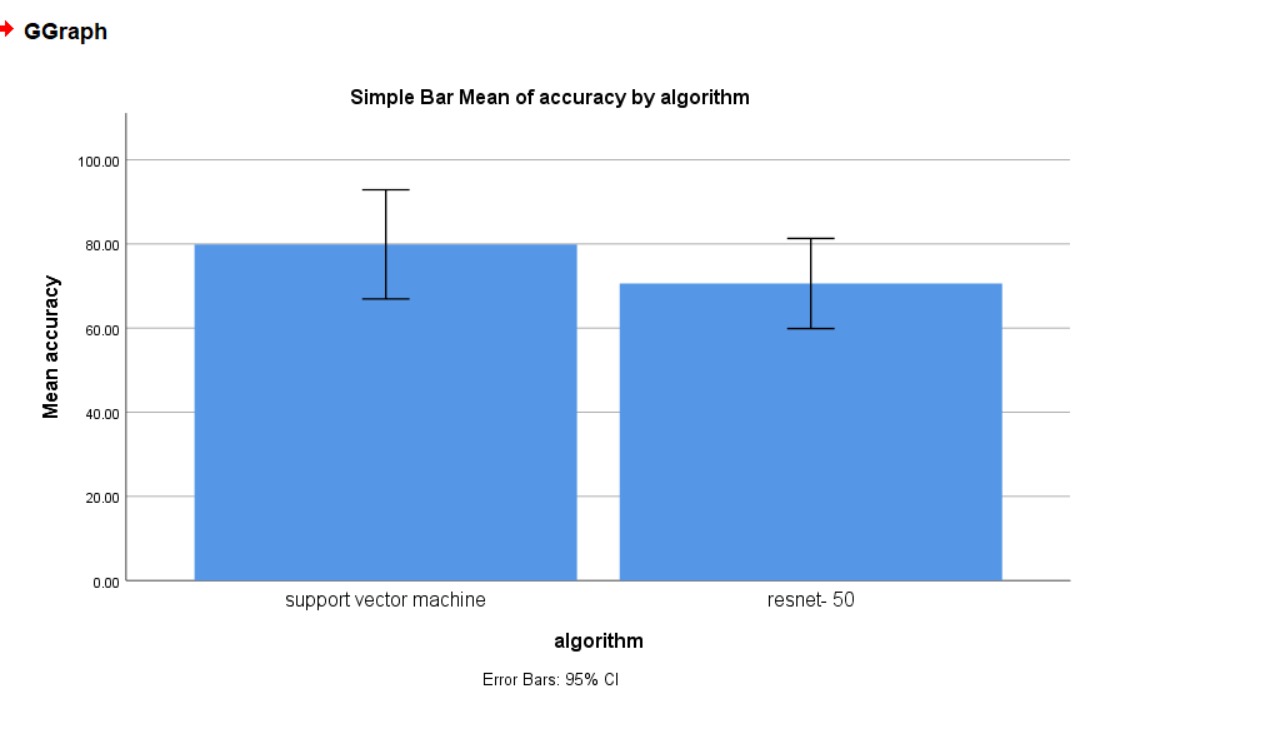
| **Group Statistics** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Accuracy** | **Algorithm** | **N** | **Mean** | **Std. Deviation** | **Std. Error Mean** |
| Support vector machine(SVM) | 10 | 98.2800 | 0.64429 | .20374 |
| ResNet-50 | 10 | 78.1900 | 2.22683 | .70419 |

**Table 3**

The Levene's test for equality of variances and the T-test for equality of means are displayed in the Independent Samples T-test results. Finding a significance value of p=0.012(p<0.05) indicates that there is statistical significance between the two groups.

| **Accuracy** | **Independent Sample Test** | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Levene’s Test for Equality of variances** | | | | | **T-test for Equality of Means** | | | |
| **F** | **Sig.** | **t** | **df** | **Sig.**  **(2-tailed)** | **Mean Difference** | **Std.**  **Error**  **Difference** | **95% Confidence Interval of the Difference** | |
| **Lower** | **Upper** |
| **Equal variances assumed** | 19.338 | .000 | 27.405 | 18 | .000 | 20.09000 | .73307 | 18.54988 | 21.63012 |
| **Equal Variance not assumed** |  |  | 27.405 | 10.496 | .000 | 20.09000 | .73307 | 18.46703 | 21.71297 |

**GRAPH BETWEEN SUPPORT VECTOR MACHINE ALGORITHM AND ResNet-50**

****

**Figure 1** Comparing the accuracy of the SVM classifier to that of the ResNet-50 algorithm has been evaluated. The SVM prediction model has a greater accuracy rate than the ResNet-50 classification model, which has a rate of 92.31. The SVM classifier differs considerably from the ResNet-50 classifier (test of independent samples, p 0.05). The SVM and ResNet-50 accuracy rates are shown along the Xaxis. Y-axis: Mean keyword identification accuracy, ±1 SD, with a 95% con