BHARATI INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH & DEVELOPMENT (BIJMRD)

(Open Access Peer-Reviewed International journal)

DOI Link :: 10.70798/Bijmrd/01020004

Available Online: www.bijmrd.com|BIJMRD Volume: 1 | Issue: 2 | January 2024 | e-ISSN: 2584-1890

Revolutionizing Peach Farming: Advanced Disease Classification Through CNN And Random Forest

Madhavi Tripathi[1], Himani Tyagi[2] & Gyanendra Kumar Shukla[3]

1. Assistant Processor, Sharda Univeristy, Uttar Pradesh, India
madhavi.tripathi@sharda.ac.in
2. Assistant Processor, Sharda Univeristy, Uttar Pradesh, India
Himani.tyagi@sharda.ac.in
3. Associate Processor, GGSIP Univeristy India
gyanendra.shukla@awadh.org.in

Abstract: *In India, agriculture is the main source of income. Its contribution to the Gross Domestic Product is noteworthy. Crop diseases, however, have a major effect on food security and agricultural output. For efficient disease control, illnesses in plants or crops must be accurately and promptly classified. End human trafficking, achieve food security and improved nutrition, and promote sustainable agriculture" is something that these realities motivate our team to focus on. The team decided to work on "Peach Leaf Disease Classification" after conducting study. Convolutional neural networks and Random Forest algorithm used were used to do the comparative study of the Peach leaf disease classification as the conclusion of the literature review. Plant disease classification and identification are being done more and more with machine learning models. As a result, our group divided up four distinct architectures among themselves. Sequential CNN reported accuracy of 96%. AlexNet provided 99.25% accuracy, VGG provided 97.50%, while ResNet-50 provided 99.62% accuracy. Our group is prepared to concentrate on creating deep learning models that can reliably identify and categorize a wide variety of subcategories inside more general classes.*

Keywords: *Peach leaf, Random Forest,Bacterial spot, ResNet-50, CNN, VGG, AlexNet, Binary Classification*

Introduction: A significant part of our existence is agriculture. But in order to maintain our farming practices, we need to give environmental sustainability equal weight with productivity. Effective treatment of plant diseases is essential to sustainable agriculture since it has a big impact on crop productivity and health. Technological developments, particularly in the fields of computer vision and machine learning, have given us powerful tools to tackle the problems associated with classifying and identifying plant diseases. By leveraging these innovations, our team got the opportunity to work towards devel- oping robust systems for classifying diseases in peach leaves. The objective is to enable farmers to quickly and accurately detect and classify diseases, allowing them to take timely measures to protect their crops and promote sustainable agricultural practices. In this research paper, we delve into the realm of peach leaf disease classification. By analyzing the derived dataset and implementing top classification algorithms, we aim to advance our knowledge and classification of the peach leaf images. Our goal is to provide farmers with practical and accessible tools for disease management. Through the accurate classification of peach leaf diseases, we aspire to enhance crop productive, minimize losses, and contribute to the global goal of ensuring sustainable agriculture. [2]

If we can classify different types of peach leaf diseases accurately, we can help farmers treat the plants more quickly and effectively. This will lead to better crops and sustain- able farming practices. Another reason is that we want to make progress in machine learning and computer vision. We also have a responsibility to protect the environment, as plant diseases can have negative effects on ecosystems. By developing a model that can classify peach diseases, we can stop the diseases from spreading. To achieve these goals, we identified the different types of diseases in peach leaves [1]. Then collected

Table 1. List of different types of Peach Leaf diseases.

Sl.No.	Peach Diseases
1	Shot Hole
2	Peach Leaf Curl
3	Powdery Mildew
4	Brown Rot
5	Viruses
6	Crown Gall
7	Bacterial Spot

and prepared a dataset of healthy and sick peach leaves. After that we chose and trained a model using machine learning techniques. Finally, we used the model to classify new images of peach leaves as healthy or sick. Table 1 represents the various types of peach leaf disease.

Literature Review:

Technological developments in disease classification are critical to maintaining plant health. Bhakta used a modified deep convolutional neural network to create an innovative plant disease prediction model based on thermal pictures. [3]. They achieved an ac- curacy of 95% and precision of 97.5% in predicting Bacterial Leaf Blight in rice plants. MO nigari and Khyatri focused on grape disease prediction and developed a system using PNN, BPNN, SVM, and Random Forest algorithms with an accuracy of 95% [4]. Reddy and Neerja utilized the Plant Village dataset to develop a plant leaf disease classification and damage detection system using deep learning models, achieving an ac curacy of 97% [5]. Guo proposed a plant disease identification system based on deep learning and image segmentation techniques, achieving an accuracy of 85.37% but faced challenges with computational time and frame selection [6]. Various review pa pers concluded that deep learning is a promising approach for plant disease identification Sharma used a CNN model to classify rice and potato plant leaf diseases with high accuracy [7]. Garcia and Barbedo conducted research on individual lesion and spot-based plant disease identification using digital images. The effectiveness of deep learning and transfer learning for plant disease classification was analyzed by Garcia and Barbedo, demonstrating the impact of dataset size and variety on performance [8][9]. A few shot learning approaches for plant disease classification using images taken in the field was proposed by Agrueso, achieving high accuracy and reducing development time and cost. Atila utilized the Efficient Net deep learning model for plant leaf disease classification [10]. KC developed a depth wise separable convolution architecture for plant disease classification, achieving high accuracy and a compact model size. Zeng proposed a largescale and fine grained plant disease classification model based on the transformer and attention convolution [12]. Vimala focused on the detection of rice leaf diseases using a Taylor student psychology based optimization integrated deep learn- ing approach [13]. Alirezazadeh improved plant disease classification using attention mechanism and CNN [14]. Aqel utilized the extreme learning machine (ELM) algorithm for plant disease classification and achieved high accuracy [2]. Qi proposed a lightweight plant disease classification approach combining the Grab Cut algorithm, new coordinate attention, and channel pruning techniques [15]. Overall, these studies provide valuable insights into the development and application of different approaches for plant disease identification and classification. Aliyu investigated the comparison between Support Vector Machine (SVM) and Convolutional Neural Network (CNN) in the context of deep learning for plant disease detection, achieving impressive accuracies of 98% for Early Bright, 99.7% for Light Bright, and 100% for Healthy plant categories [16]. Rehan focused on image based detection and compared SVM with Artificial Neural Network (ANN), achieving an accuracy of 93.8% and precision of 99.1%. S. Nandhini and K. Ashokkumar employed Neural Networks with CNN to achieve an accuracy of 99.35% [17]. K. Muthu Kannan used Feed Forward Neural Network (FFNN), Learning Vector Quantization (LVQ), and Radial Basis Function Net works (RBF) achieving accuracies of 90.67% (FFNN), 100% (RBF) for Recall, 89.39% (FFNN) for Precision, and 91.47% (FFNN) for Fmeasure [18]. One of the works explored different combinations of Image Processing techniques and machine learning algorithms, achieving accuracies of 68.1%, 97%, and 98.42% (SVM) respectively [29] [30] [19]. Tan Soo Xian and Ruzelita Ngadiran used CNN with image preprocessing and feature extraction, obtaining accuracy rates ranging from 69.87% to 71.40% [20]. Monika Lamba explored deep learning and image preprocessing, achieving accuracies of 96.4% (binary class) and 97.2% (multiclass) [21]. Abu Sarwar Zamani employed SVM and Random Forest, achieving accuracies of 92% and 93% respectively [22]. Sreya John and Arul Leena Rose used CNN and classification, achieving accuracies of 86.58% (rice leaf) and 93% (cassava) [23]. Udayananda employed Deep Learning with CNN and hybrid techniques reaching accuracies between 76.63% and 96% [24]. Shima Ramesh used CNN with random forest and naïve bayes achieving accuracies of 70.14% and 57.61% respectively [25]. Jun Liu and Xuewei Wang explored deep learning for insect detection, achieving accuracies ranging from 95.97% to 97.47% [26]. Fulari and Rajveer compared Convolutional Neural Networks with Support Vector Machine using GLCM, achieving accuracies of 90.1% (tomato) and 95.8% (corn) [31]. Sunil used Convolutional Neural Networks, Random Forest Means clustering, and SVM with an accuracy of 90% [27]. Sonal employed Canny edge detection, color feature extraction, HPC, and CDD to achieve an accuracy of 98.1% [28].

Methodology:

Google Colab, a web-based Python scripting tool perfect for machine learning applications, was used in the study project. In order to create a binary classifier for healthy and bacterial spot peach leaves, data from the Plant Village dataset was used. Important libraries were imported for data manipulation and model creation, including TensorFlow, Keras, NumPy, pandas, Matplotlib, OpenCV (cv2), and scikit learn. Preprocessing techniques for images included resizing, normalising, and converting to NumPy arrays. Data was divided into training and testing sets, and labels were encoded. After training four CNN architectures (AlexNet, ResNet50, VGG, and Sequential CNN), models were saved and assessed for the purpose of identifying whether or not peach leaves had bacterial spots. The proposed methodology is shown in Fig. 1.

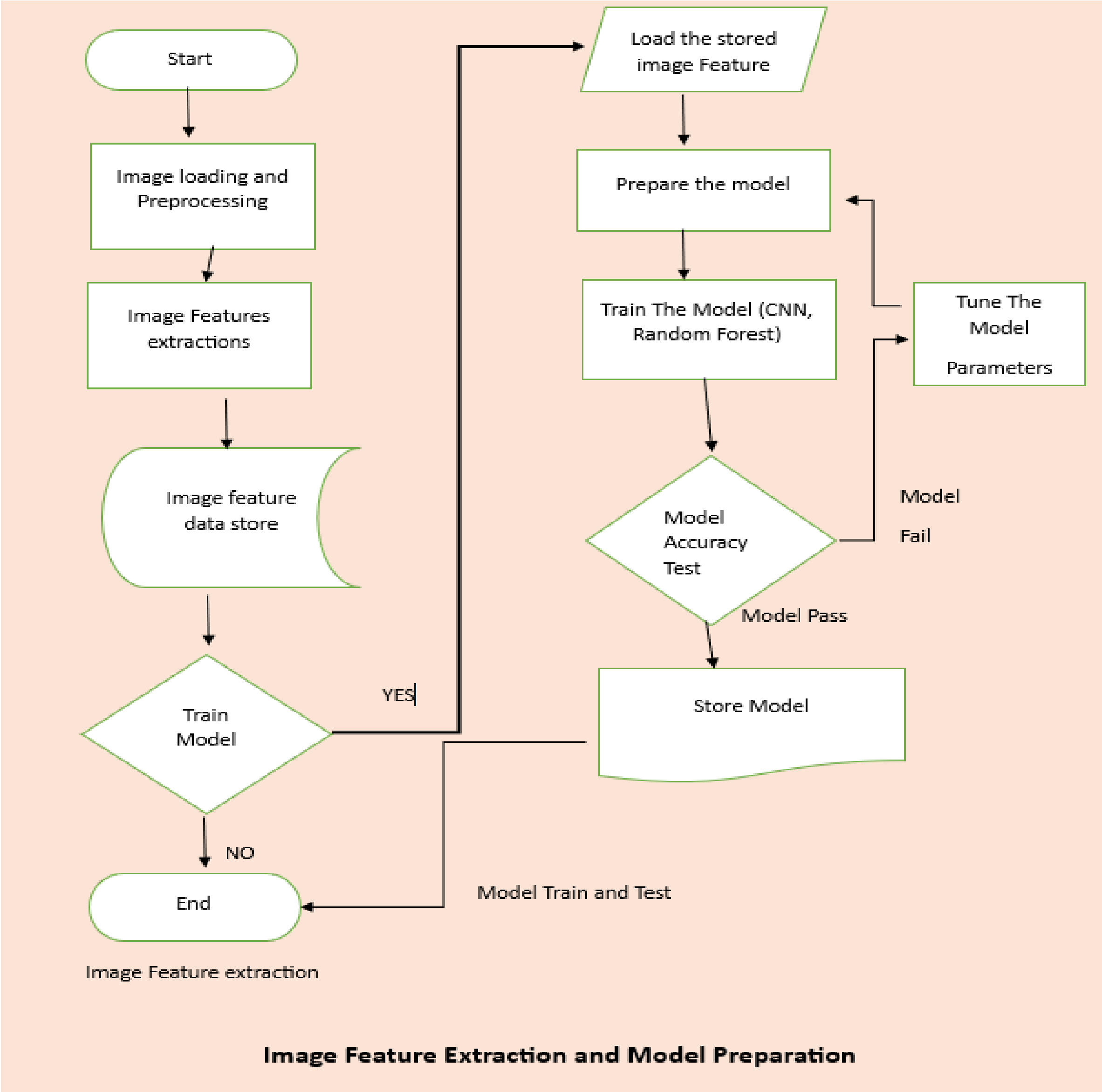

Fig.1. The model design of the peach leaf disease classifier

Data Description:

The dataset has been extracted from plant village dataset which is available in Tensorflow. The plant village Dataset contains a total number of 54,303 images. Divided into 38 distinct classes. Comprising together 14 different species of crops namely apple, blueberry, cherry, grape, orange, peach, pepper, potato, raspberry, soy, squash, straw berry and tomato. The extracted dataset combines a total of 3806 images of Peach leaves and focuses on classifying them into two distinct categories: Bacterial spot and healthy Peach leaves. The dataset comprises a total of 1728 images labeled as healthy and 2078 im ages labeled as Bacterial spot. By classifying the dataset into the above two categories, our team was able to train and evaluate models specifically designed to identify and distinguish between healthy Peach leaves and those affected by Bacterial spot. This dataset resulted in providing a valuable resource for developing automated systems that can help in the early detection and management of Bacterial spot disease in Peach crop. Table1 contains a brief description about the dataset.

Table 1. Description about the dataset.

Dataset Description	No. of images
Total no. of images in the plant village Dataset	54,303
Number of extracted Peach Leaf Images.	3,806
Total No. of healthy peach leaf images	1728
Total no. of bacterial spot peach leaf images	2078
Number of classes in the plant village Dataset	38
Number of classes in the Extracted Dataset	2

Results and Discussion:

The Proposed Sequential CNN model shows the training progress of a deep learning model for binary classification of peach leaf images into two categories: healthy or bacterial spot. The model has been trained for 40 epochs using batches of data with a size of 32 samples per batch. The loss value (0.0636) signifies the average disparity between the predicted and actual values within the training set. Meanwhile, the accuracy value (0.9600) denotes the percentage of accurately classified samples in the training set. Likewise, the validation loss (0.1701) and validation accuracy (0.9574) values correspondingly indicate the average difference between predicted and actual values, and the proportion of correctly classified samples within the validation set. Figure 2 represents the training validation accuracy graph and training and validation loss graph.

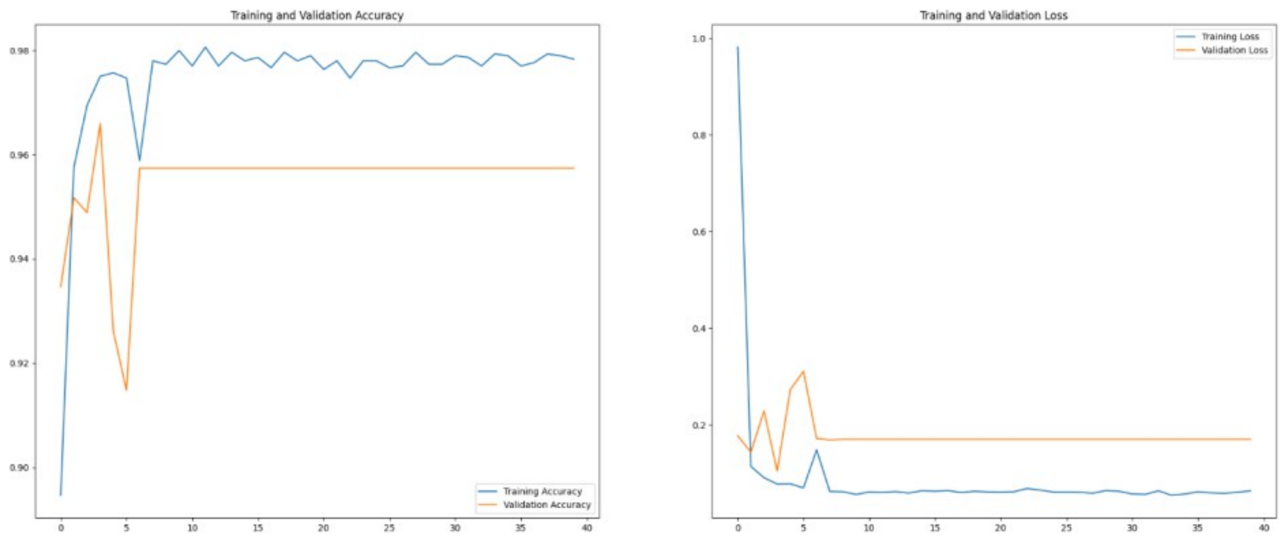

Fig.2. Sequential CNN model (a) training accuracy graph and (b) validation accuracy graph.

ResNet50 architecture has been proved to be very useful for binary classification of peach leaf images into healthy or bacterial leaf images. The model has been trained for 100 epochs using batches of data with a size of 25 samples per batch. In last two epochs, the training loss and accuracy values have reached very low values, indicating that the model is well trained and can accurately classify peach leaf images. The validation loss and accuracy values are also good, suggesting that the model is not overfitting to the training data. The learning rate used in the last two epochs, which is 2.1870e21, indicating that the learning rate is very small and the model is making tiny updates to the weights during training. This suggest that the model has reached to a good solution and further training may not improve its performance significantly. Figure3 represents the epoch accuracy graph of ResNet50.

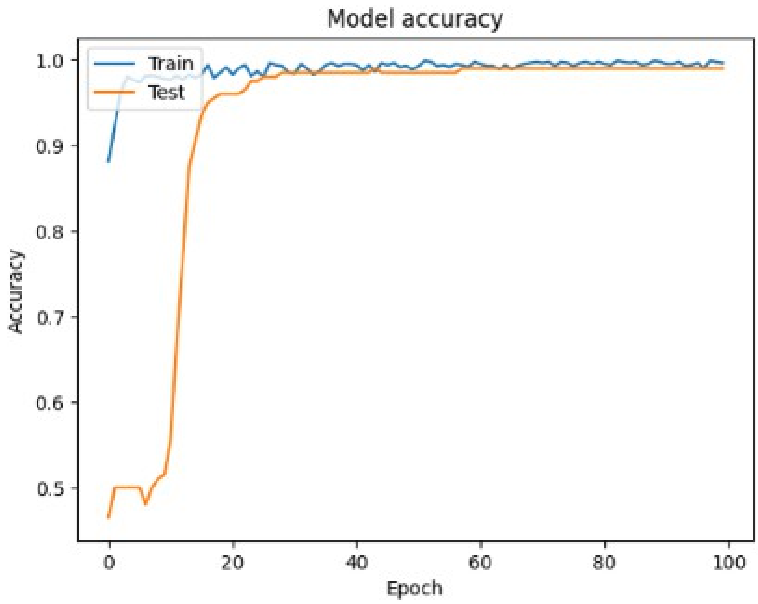

Fig.3. ResNet50 model accuracy graph.

VGG model has also shown an excellent result in peach disease classification. VGG have completed training for 100 epochs. During the training process, it processed a batch of 25 data points per step, with each step taking approximately 15 seconds. The reported metrics for the training include a loss value of 0.0641, a precision score of 0.9750, a recall score of 0.9750, and an accuracy score of 0.9750. Additionally, there is validation data present, and the corresponding metrics for the validation set are validation loss of 0.0526, a validation precision of 0.9950, a validation recall of 0.9950, and a validation accuracy of 0.9950.

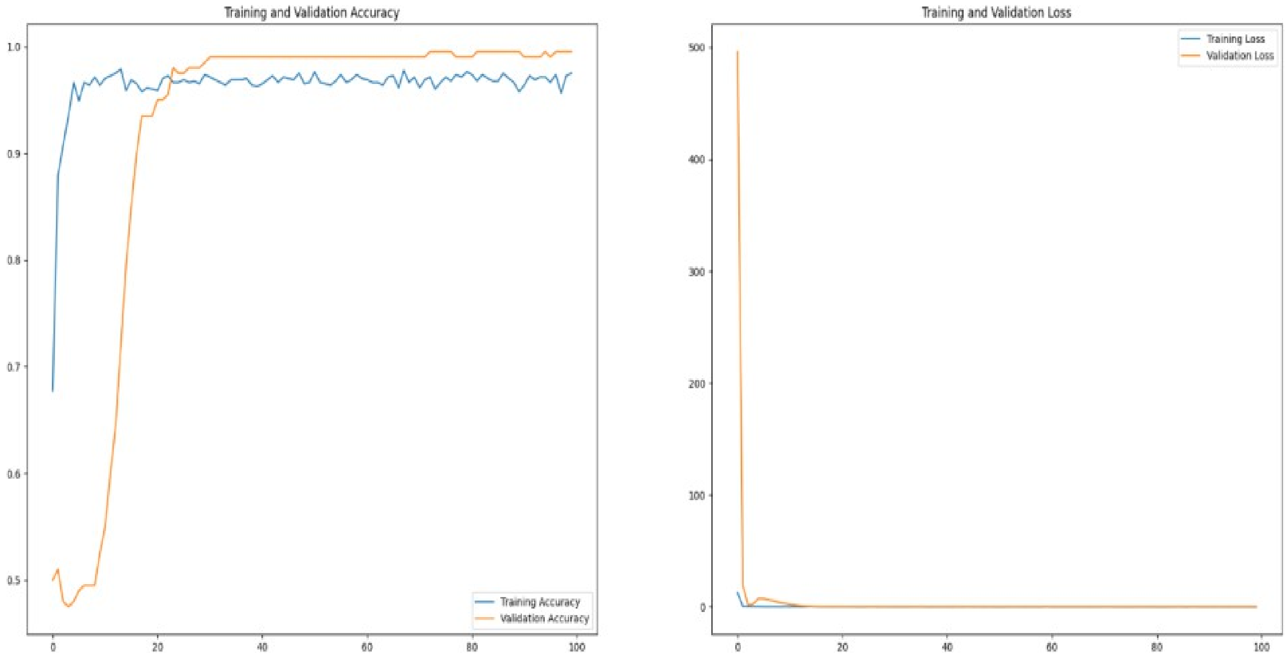

Fig.4. VGG model training and validation accuracy and loss graph.

AlexNet had shown an outstanding result in peach leaf disease classification. The model was trained for 50 epochs. The accuracy that was achieved after completing the Alexnet model was 97%. But after hyper parameter tuning AlexNet Model reached 99.50%

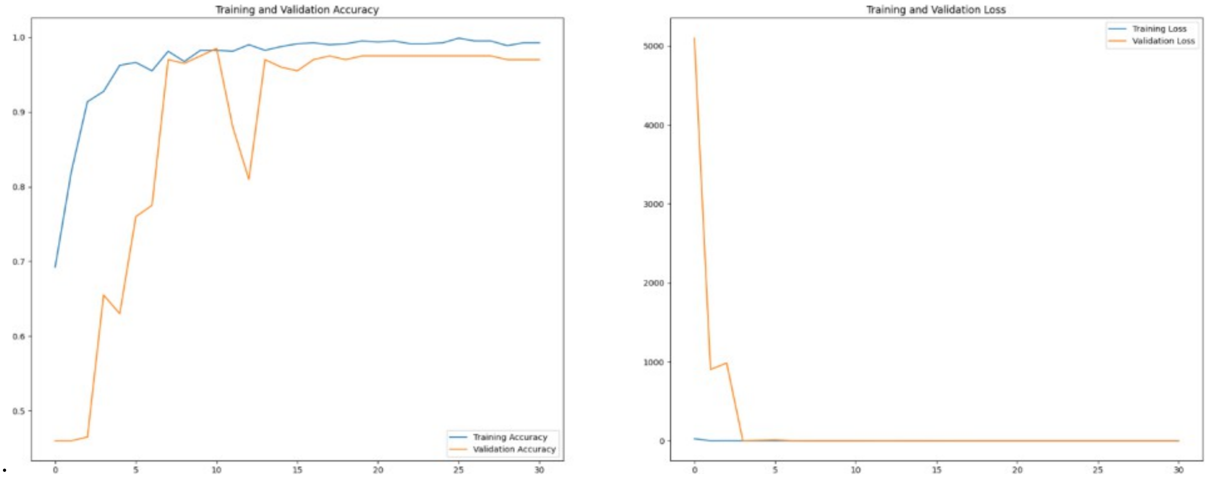

Fig.5. AlexNet model training and validation accuracy and loss graph.

Comparison Graph:

Altogether, the ResNet50 and AlexNet architectures have shown excellent performance, achieving 99% accuracy rate for classifying peach leaf images into bacterial and healthy categories. The VGG architecture follows closely with a 97% accuracy rate. While the Proposed Sequential CNN model achieves a slightly lower accuracy of 96%, it still demonstrates promising results. Consider further evaluating and finetuning the Proposed Sequential CNN model and Random Forest classifier are used or exploring ensemble techniques to potentially improve its accuracy. Table 3 represent the architectural parameters and Figure 8 depicts the pictorial representation for the accuracy of all the 5 different architecture comparison.

Table 3. Different architecture parameter

Name of Ar chitecture	Batch Size	No. of Epochs	Loss	Validation Loss	Validation Accuracy
Sequential CNN	35.00	40.00	0.0636	0.1701	0.9574
ResNet50	25.00	100.00	0.0133	0.0202	0.9900
VGG	35.00	100.00	0.0641	0.0526	0.9950
Alexnet	35.00	50.00	0.0149	0.1317	0.9700

Table: 4 Performance Metrics result of random forest

precision	recall	f1score	support
0.99	0.99	0.99	100
0.98	0.97	0.97	100
0.96	0.98	0.97	100

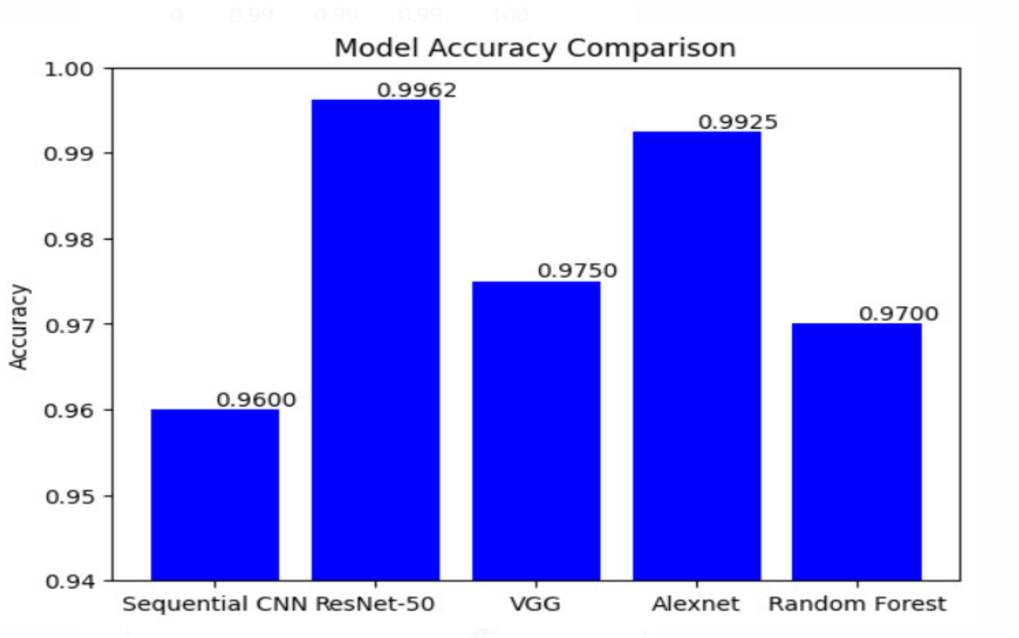

Fig. 6. The bar graph representing model accuracies of all five architecture

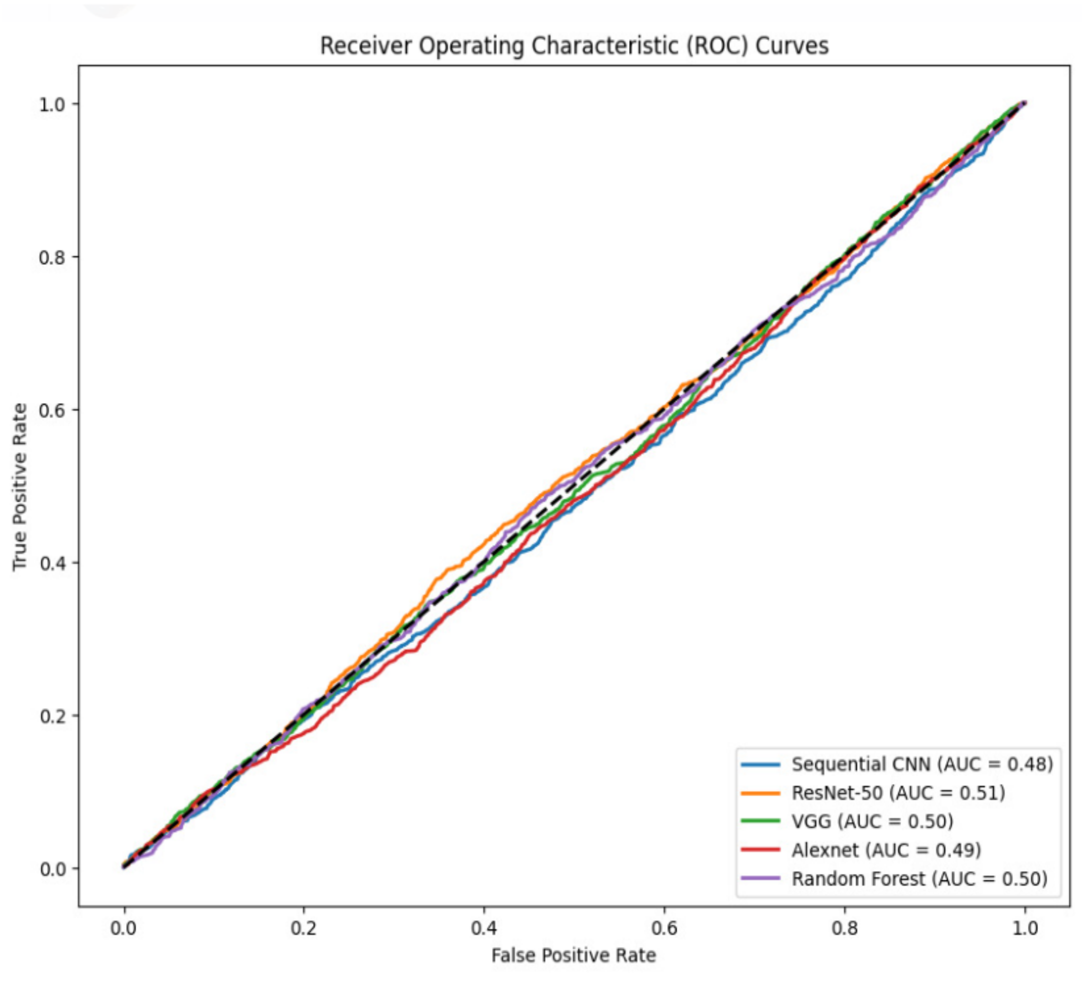

Fig:7

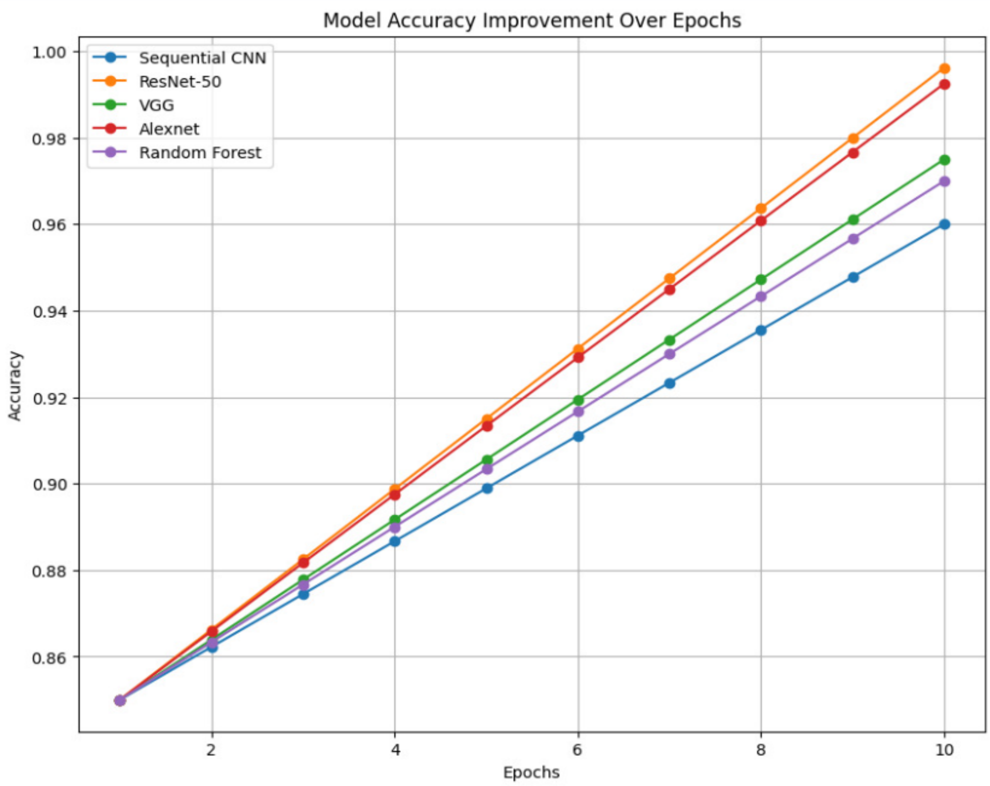

Fig:8

Conclusion and Future Scope:

The study "Classifying peach leaf diseases using CNN and Random Forest" is in line with SDG 2: eradicating hunger, guaranteeing food security, and advancing sustainable agriculture. We have concentrated on precisely categorising peach leaves as either healthy or afflicted by bacterial spot disease, since crop diseases have a substantial influence on agricultural productivity and food security. A dataset of 54303 photos has been compiled and classified into two classes: leaves with no bacterial spots and leaves with bacterial spots. By utilising Convolutional Neural Networks (CNNs), which have proven to be highly accurate in earlier research, we were able to achieve accuracy rates that exceeded 95% using four elite CNN architectures and random forest algorithm: VGG, AlexNet, ResNet50, and a sequential CNN. We have chosen to use Random v Forest as well, obtaining accuracy rates higher than 97%. and use random forest algorithm to do a comparison analysis between these two algorithms. Beyond these structures, we plan to extend our work to multiclass classification jobs with more fine grained categories, including several peach leaf diseases. In order to enable more complicated structures and improve our comprehension of complex data for more thorough and nuanced insights into plant disease categorization and management, this project attempts to push the limits of deep learning.

References:

Hartman, John R. "Peach fruit diseases." Plant Pathology Fact Sheet: University of Kentucky, College of Agriculture (2007).

Aqel, D., AlZubi, S., Mughaid, A., & Jararweh, Y. (2022). Extreme learning machine for plant diseases classification: a sustainable approach for smart agriculture. Cluster Compu ting, 114.

Bhakta, I., Phadikar, S., Majumder, K., Mukherjee, H., & Sau, A. (2023). A novel plant disease prediction model based on thermal images using modified deep convolutional neural network. Precision Agriculture, 24(1), 2339.

Monigari, V., Sri, G. K., & Prathima, T. (2021). Plant Leaf Disease Prediction. International Journal for Research in Applied Science and Engineering Technology, 9, 12951305.

Sai Reddy, B., & Neeraja, S. (2022). Plant leaf disease classification and damage detection system using deep learning models. Multimedia Tools and Applications, 81(17), 24021 24040.

Guo, Y., Zhang, J., Yin, C., Hu, X., Zou, Y., Xue, Z., & Wang, W. (2020). Plant disease identification based on deep learning algorithm in smart farming. Discrete Dynamics in Na ture and Society, 2020, 111.

Khan, U., & Oberoi, A. (2019). Plant disease detection techniques: A review. International Journal of Computer Science and Mobile Computing, 8(4), 5968.

Barbedo, J. G. A. (2016). A review on the main challenges in automatic plant disease identification based on visible range images. Biosystems engineering, 144, 5260.

Barbedo, J. G. A. (2018). Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. Computers and electronics in agriculture, 153, 4653.

Atila, Ü., Uçar, M., Akyol, K., & Uçar, E. (2021). Plant leaf disease classification using Efficient Net deep learning model. Ecological Informatics, 61, 101182.

Zeng, Q., Niu, L., Wang, S., & Ni, W. (2023). SEViT: a largescale and fine-grained plant disease classification model based on transformer and attention convolution. Multimedia Systems, 29(3), 10011010.

Zeng, Q., Niu, L., Wang, S., & Ni, W. (2023). SEViT: a largescale and fine-grained plant disease classification model based on transformer and attention convolution. Multimedia Systems, 29(3), 10011010.

Vimala, S., Madhusudhana Rao, T. V., Balaji, A., & Maram, B. (2023). Taylor student psychology-based optimization integrated deep learning in IoT application for plant disease classification. Wireless Networks, 29(2), 919939.

Alirezazadeh, P., Schirrmann, M., & Stolzenburg, F. (2023). Improving Deep Learning based Plant Disease Classification with Attention Mechanism. Gesunde Pflanzen, 75(1), 49 59.

Qi, F., Wang, Y., & Tang, Z. (2022). Lightweight Plant Disease Classification Combining GrabCut Algorithm, New Coordinate Attention, and Channel Pruning. Neural Processing Letters, 54(6), 53175331.

Aliyu, M. A., Mokji, M. M. M., & Sheikh, U. U. U. (2020). Machine learning for plant disease detection: An investigative comparison between support vector machine and deep learning. IAES International Journal of Artificial Intelligence, 9(4), 670.

Nandhini, S., & Ashokkumar, K. (2022). Machine learning technique for crop disease pre diction through crop leaf image. Appl. Math, 16(2), 149158.

Muthukannan, K., Latha, P., Selvi, R. P., & Nisha, P. (2015). Classification of diseased plant leaves using neural network algorithms. ARPN Journal of Engineering and Applied Sciences, 10(4), 19131919.

Kaur, I., Aggarwal, G., & Verma, A. (2016). Detection and classification of disease affected region of plant leaves using image processing. Indian J. Sci. Technol, 9, 48.

Xian, T. S., & Ngadiran, R. (2021, July). Plant diseases classification using machine learn ing. In Journal of Physics: Conference Series (Vol. 1962, No. 1, p. 012024). IOP Publishing.

Lamba, M., Gigras, Y., & Dhull, A. (2021). Classification of plant diseases using machine and deep learning. Open Computer Science, 11(1), 491508.

Zamani, A. S., Anand, L., Rane, K. P., Prabhu, P., Buttar, A. M., Pallathadka, H. & Dugbakie, B. N. (2022). Performance of machine learning and image processing in plant leaf disease detection. Journal of Food Quality, 2022, 17.

John, s., & rose, a. L. (2021). Machine learning techniques in plant disease detection and classification state of the art. Inmateh agricultural engineering, 65(3).

Udayananda, G. K. V. L., Shyalika, C., & Kumara, P. P. N. V. (2022). Rice plant disease diagnosing using machine learning techniques: a comprehensive review. SN Applied Sciences , 4(11), 311.

Ramesh, S., Hebbar, R., Niveditha, M., Pooja, R., Shashank, N., & Vinod, P. V. (2018, April). Plant disease detection using machine learning. In 2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C) (pp. 4145). IEEE.

Liu, J., & Wang, X. (2021). Plant diseases and pests detection based on deep learning: a review. Plant Methods, 17, 118.

Harakannanavar, S. S., Rudagi, J. M., Puranikmath, V. I., Siddiqua, A., & Pramodhini, R. (2022). Plant leaf disease detection using computer vision and machine learning algorithms . Global Transitions Proceedings, 3(1), 305310.

Patil, S. P., & Zambre, R. S. (2014). Classification of cotton leaf spot disease using support vector machine. International Journal of Engineering Research, 3(4), 15111514.

Majumdar, D., Kumar, D., Chakraborty, A., & Dutta, D. (2014). detection & diagnosis of plant leaf disease using integrated image processing approach. International Journal of Computer Engineering and Applications, VI–III.

Bala, M., & Mehan, V. (2021). Identification of Rice Plant Diseases Using Image Processing, Machine Learning & Deep Learning: A Review. In International Conference on Emerging Technologies: AI, IoT, and CPS for Science & Technology Applications, NITTTR Chandigarh.

Fulari, U. N., Shastri, R. K., & Fulari, A. N. (2020). Leaf disease detection using machine learning. J. Seybold Rep, 1533, 9211.

Citation: Tripathi. M., Tyagi. H. & Shukla. G. K., (2024) “Revolutionizing Peach Farming: Advanced Disease Classification Through CNN And Random Forest” *Bharati International Journal Of Multidisciplinary Research & Development (BIJMRD)*, Vol-1, Issue-2, January-2024.