Proceedings of the Ninth International Conference on Research in Intelligent Computing in Engineering pp. 79–87
DOI: 10.15439/2024R15
ISSN 2300-5963 ACSIS, Vol. 42

Enhancing Plant Disease Detection Through Image Analysis Using SSD Mobilenet V2 and ResNet-50

Paribartan Timalsina
Department of Computer Science and Engineering
Kathmandu University
Dhulikhel, Nepal
timalsinapari015@gmail.com

Subarna Bhattarai
Department of Computer Science and Engineering
Kathmandu University
Dhulikhel, Nepal
sbrnbhttr@gmail.com

Shaswot Paudel
Department of Computer Science and Engineering
Kathmandu University
Dhulikhel, Nepal
shaswotpaudel58@gmail.com

Sudan Jha*
0000-0003-0074-2584
Department of Computer Science and Engineering
Kathmandu University
Dhulikhel, Nepal
jhasudan@ieee.org

***Abstract*—Being an agrarian nation, Nepal possesses huge economic value from the agricultural sectors with more than half of its population contribution in farming, thus bringing a great potential share in the nation's Gross Domestic Product (GDP). Despite this substantial contribution, there are still issues with advancement that hinders growth in this field. This study, therefore presents a model for the detection of plants and diagnosis of diseases associated with the plants using real-time camera feeds or by analyzing captured images. Existing models often suffer from unrelated irrelevant images; confusing one plant for another which results in failure of proper diagnosis of diseases. Addressing this issue, our model first detects the presence of a plant within the frame or image before disease classification. Plant detection is performed using the Single-Shot Detector (SSD) Mobilenet V2 model. Disease classification process is initiated only when the plant is detected. The Residual Network (ResNet)- 50 model performs the plant disease classification taking the clipped image from detected plant. By focusing only on the detected plant, we reduce background complexity and improve classification accuracy. The model has demonstrated a high level of accuracy in both detecting plants and classifying diseases based on the prevalent diseases associated with specific plants.Our model achieved a mean Average Precision (mAP) score of 0.61 for the object detection model which includes 9 classes and achieved average accuracy of 98% for the classification of 9 plant species having a total of 33 classes. This innovative model is hence targeted at providing reliable and accurate plant detection and disease diagnosis to address some of the key challenges in agricultural technology in Nepal. On near future it has the potential for automated plant diseases detection and real time monitoring of plant diseases. Additionally, we can deploy it in mobile applications, allowing farmers to identify plant diseases using smartphone cameras, facilitating timely detection and reduction of crop losses.**

***Index Terms*—Deep Learning, SSD Mobilenet V2, ResNet-50**

This research was supported by Kathmandu University.
*Corresponding author: jhasudan@ieee.org

I. Introduction

Agriculture being one of the most critical macroeconomic sectors is making a substantial economic contribution in the nation's economy. Approximately 60.4% of the population is engaged in agriculture, which accounts for 27% of Nepal's Gross Domestic Product (GDP) [1]. Despite this fact, small-scale farmers often face challenges related to ignorance and lack of access to cutting-edge, affordable technologies such as precision agricultural tools and technology for crops diseases prediction. This technological gap shortfall affects the crop production and overall productivity. Likewise, the existing bio-molecular technologies like antigen-antibody reaction, DNA sequence amplification are very sensitive for small-scale use and are very costly.

To address the issue, we propose a model suitable for identification of plant species and classification of diseases using object detection and Deep Learning (DL) methodologies. The model utilizes either live camera feeding or image uploading for plant detection and disease classification. The proposed model through commonly available smartphones aims to make the solution accessible and user-friendly. For plant species detection, we employ an object detection model based on Convolutional Neural Networks (CNNs). Specifically, we utilize the Single-Shot Multibox Detector (SSD) with a MobileNet V2 architecture, allowing us to build a lightweight model that is computationally efficient and suitable for real-time applications. SSD MobileNet V2 is different as it uses the depthwise separable convolution thus reducing the parameters of the model. SSD MobileNet V2 is mostly used beacuse of it's lightweight characterstics and maintaining the balance between speed and accuracy. The SSD generates bounding boxes with confidence scores which allows the system to correctly locate the position of a plant within an image.

For disease classification, we make use of the ResNet-50 architecture among the best CNN architectures in the strong extraction of deep features from images. This architecture enables the model to effectively find out the disease classification depending on the features in the detected region of the plant. The ResNet-50 architecture can effectively manage network depth, increase performance and generalization capacity. The key factor for using ResNet-50 lies in it's ability to learn set of residual functions to map the input to desired output. Balancing lightweight object detection for plant identification with a more complex classification model enables us to maintain good accuracy while minimizing resource consumption for fast and efficient system. The proposed method has demonstrated the potential for drastic improvement in agriculture. Productivity is conveyed by providing smallholder farmers with a reliable means and accessible tool for plant and disease identification. Using less complex SSD MobileNet V2 for detection followed by more complex ResNet-50 for classification, we achieve an optimal trade-off between performance and computational efficiency.

II. Literature Review

Numerous bio-molecular research studies have been conducted to increase the early and accurate detection of plant diseases like Enzyme Immunoassay/Enzyme-Linked Immunosorbent Assay(ELISA) and Real-Time Polymerase Chain Reaction (RT-PCR). ELISA, based on the antigen-antibody reaction to detect the diseases in plants by quantifying various kinds of molecules like proteins, hormones and toxins can give high false positive or negative results if the experiment isn't done with great care and also depending on various factors like freshness, concentration of solution and types of organism [2], [3]. Another method, Real-Time Polymerase Chain Reaction(RT-PCR) used to raise and quantify the targeted DNA molecule at the same time is effective because of it's rapid testing with greater sensitivity but is still very expensive for daily applications because of high costs of machines and reagents [4].

As the biological molecular methods are very sensitive and can be very costly researchers have come up with an idea of detecting the plants and it's related diseases using deep learning techniques. One early study implemented a multi-step process, including color transformation, masking of green pixels and removal with specific thresholds, and segmentation through equal-sized patches, followed by classification using a database of 500 plants [5]. This approach used Support Vector Machine (SVM) classifier with an accuracy of 94.74%.

Remote Sensing technologies including airborne multispectral/hyperspectral imagery and high-resolution satellite sensors enable image acquisition using airborne or satellite sensors followed by pre-processing and feature extraction with mapping of crop diseases at last [6]. [7] allow for the analysis of spectral properties, revealing changes in reflectance patterns that indicate stress or disease in crops where healthy vegetation reflects significant light in the near-infrared region, while stressed or diseased plants show altered reflectance.

Several research studies were focused on isolating lesions for plant disease detection by removing the background from leaf images. An effective technique for image segmentation was presented based on a Chan-Vese model and Sobel operator [8]. This method includes three steps: extracting leaf contours via the Chan–Vese model and detecting edges using an enhanced Sobel operator, removing the background by identifying high green-level pixels, and isolating the target leaf in complex backgrounds by combining the Chan-Vese and Sobel results. A proposed model, combined traditional CNN with squeeze-and-excitation (SE) module and global pooling layer to identify the plant diseases and achieved an overall accuracy of 91.7% [9].

Another study involved CNNs for the classification and detection of diseases affecting the plant species potatoes, tomatoes, and peppers. Using the Plant Village dataset consisting of 20,636 images from fifteen classes and reported 98.29% training accuracy and 98.029% testing accuracy [10]. A study proposed a system to detect the varieties of plants like Apple, Grape, Potato, Corn, Sugarcane and Tomato and the diseases associated with them comprising of 35000 dataset with an overall average accuracy of 96.25% and the system being able to give 100% accuracy confidence in classification [11]. Lately, object detection methods have been widely used with aim to minimize loss on a given dataset with enhanced accuracy [12], [13]. Models such as Faster R-CNN and You Only Look Once(YOLO) were employed for object detection [12], [13] while ResNet [14] was utilized for image classification.

Object detection models have also been used in detecting plant diseases. [15] explored various deep learning object detection methods like Regions with Convolutional Neural Networks(R-CNN), Region-based Fully Convolutional Network(R-FCN), and SSD, which were employed in the identification of diseases in plant leaves. These models learn complex scenarios from the plants area and hence can have good accuracy. [16] proposed a multistage method used to detect and classify the leaves of the given plant using YOLOV3 as plant detection model and ResNet-18 as classification model. [17] proposed a model that leverages a Conditional Generative Adversarial Network for generation of synthetic data, a Convolutional Neural Network for feature extraction, and a Logistic Regression classifier for quick classification of the plant species. The model was trained for plants like apple, corn, grapes, potato, sugarcane, and tomato, and gave an accuracy of 96.5% on the combined dataset and 99% to 100% on the dataset of individual species.

The use of CNNs has had a significant impact on computer vision tasks, but the accuracy of CNNs can be significantly impacted if the images in the dataset are diverse [18]. Images and frames with unwanted objects besides the area of interest, called background noise, can greatly impact the efficacy of CNNs. The in-situ plant images have unwanted parts such as soil, rocks, and/or human body parts that result in cluttered backgrounds and hence reduced accuracy [18].

[19], [20] found that average accuracy of CNN models can worsen with a higher number of classes, while performance may increase with fewer categories. Similarly, class imbalance or the over representation of some classes relative to others, is a common issue in large-scale image classification datasets. This lopsided class distribution can negatively affect the model's performance [21].

ResNet-50 model was used for image classification due to its' superior performance in image classification task with respect to other models [22]. The architecture of ResNet-50 is of a deep residual in nature, thus preventing vanishing gradient descent. This ultimately helps the model learn effectively the complex features that make it accurate for classification tasks.Likewise, SSD MobileNet V2 architecture has been used for leaf detection since it can quickly process images [23]. Although this model is relatively lower as compared to other models regarding accuracy, such as inception and fasterRCNN, our object detection models do not require extreme complexity or exceptionally high accuracy.

Recent advancement has focused on developing and training classification models based on Vision Transformer(ViT) and CNN and checking if the recent advancements outperforms the previously given models. A research conducted by [24] demonstrated that the ViT2 model outperformed other models and accquired accuracy upto 99.7%. for identifying the diseases in tomato plants. Another research by [25] developed smartphone-based application using a ViT-based model that used self-attention mechanism and achieved accuracy of 90.99% during experimentaion. So, vision based transformers are being widely used in detection of diseases of plants in present time.

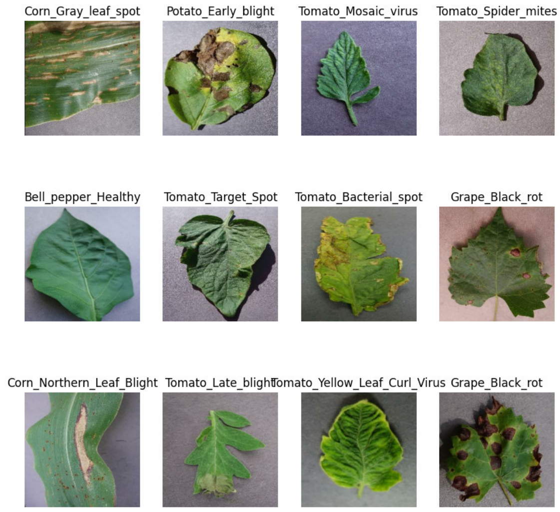


Fig. 1: Sample images from the dataset showing different plant species

III. Methodology

A. Data Collection

The images of the different types of plant species required for research were collected from various sources like GitHub and Kaggle. Images were collected such that there is the inclusion of variety of plants and associated diseases. These images were used for plant leaf detection and image classification.

B. Dataset Preprocessing

All the images were checked for the acceptable files like (jpg, jpeg, png and bmp). This was done to ensure data integrity and to make sure that only acceptable files are processes as valid input. During training of plant detection model, the dataset was checked for distribution of data like class imbalance problem, and oversampling and undersampling was done to ensure that there is uniform distribution of data and making it suitable for training. For the plant leaf detection task the target number for each types of plant was set to 2000. Plants with fewer data samples were oversampled by generating additional samples and for large number of samples the number of samples were reduced. Our original file containing the annotated image information was then changed into csv format which is further changed into TensorFlow Record(TFRecord) format making it compatible with TensorFlow to read and write the large datasets.

During the training of image classification model, images were correctly labeled. The images size were resized to 256*256 to establish the uniformity during training. The images were then normalized in the range of [0,1] for getting stable gradient descent updates during back-propagation and getting faster convergence. The dataset were then splitted in the ratio of 80%, 10% and 10% for train, test and validation data. The sample of images used during training is shown in Figure 1.

C. Data Augmentation and Pipelining

For ensuring the diversity in data and to improve generalization data augmentation is performed. The augmentation was done by flipping the images with horizontal and vertical rotation upto 20%. This process was done during the training phase to ensure that our model learns from the broader dataset. In addition to it the dataset was cached, shuffled, and prefetched for optimization of data pipeline. Caching helps to reduce the data loading time by storing the dataset in memory after first epoch. Shuffling was done having a buffer size of 1000 for presenting the data in random order during training thus preventing overfitting. Prefetching was done to assure that the next batch of data is loaded during the model training and thus making the smooth and efficient training process. To improve the generalization ability of model plants images were collected from different lighting conditions and also the different stages of diseases in a particular plant was taken into consideration. In addition to it the adjustment of contrast and brightness were applied to ensure model in robust among varying environmental condition.

D. Data Analysis and Description

After completing data pre-processing and augmentation, the total number of images required for both models were finalized. The plant leaf detection model is trained on nine different plants, and a separate image classification model is trained for disease classification in each of these plants. The number of plants used in the dataset for the image classification model is enlisted in Table I along with the class names in classification are provided in Table II respectively.

TABLE I: Object Detection and Image Classification Dataset

Plant Species	Object Detection (Number of Images)	Image Classification (Number of Images)
Grape	500	5544
Potato	547	4400
Tomato	703	12523
Apple	562	5511
Corn	545	5643
Bell Pepper	560	3131
Strawberry	462	3012
Peach	604	3112
Cherry	220	3122
Total	**4693**	**40998**

TABLE II: Disease Classification for Different Plants

Plant	Diseases	Number of Classes
Apple	Apple_Black_rot, Apple_Healthy, Apple_Scab,Cedar_Apple_rust	4
Bell Pepper	Bell_pepper_Bacterial_spot, Bell_pepper_Healthy	2
Cherry	Cherry_Healthy, Cherry_Powdery_mildew	2
Corn	Corn_Common_rust, Corn_Gray_leaf_spot, Corn_Healthy, Corn_Northern_Leaf_Blight	4
Grape	Grape_Black_Measles, Grape_Black_rot, Grape_Healthy, Grape_Isariopsis_Leaf_Spot	4
Peach	Peach_Bacterial_spot, Peach_Healthy	2
Potato	Potato_Early_blight, Potato_Healthy, Potato_Late_blight	3
Strawberry	Strawberry_Healthy, Strawberry_Leaf_scorch	2
Tomato	Tomato_Bacterial_spot, Tomato_Early_blight, Tomato_Healthy, Tomato_Late_blight, Tomato_Leaf_Mold, Tomato_Mosaic_virus, Tomato_Septoria_leaf_spot, Tomato_Spider_mites, Tomato_Target_Spot, Tomato_Yellow_Leaf_Curl_Virus	10

IV. Proposed Model Design

The proposed model employs a two-tiered architecture for plant detection and disease classification. At first detection of the species of plant is done using SSD MobileNet V2 and then diseases associated with that plant is identified using ResNet-50 model architecture. The workflow of the proposed model is shown in Figure 2 with the description in the following subsections.

A. Plant Detection

Object detection works by classifying and detecting objects by creating bounding boxes around them. In the proposed plant detection model, SSD Mobilenet V2 is used to identify the plant species, mentioned in Table I. SSD Mobilenet V2 is selected due to its lightweight nature and ability to balance computational efficiency with accuracy, making it suitable for mobile devices. This step confirms that only the relevant part

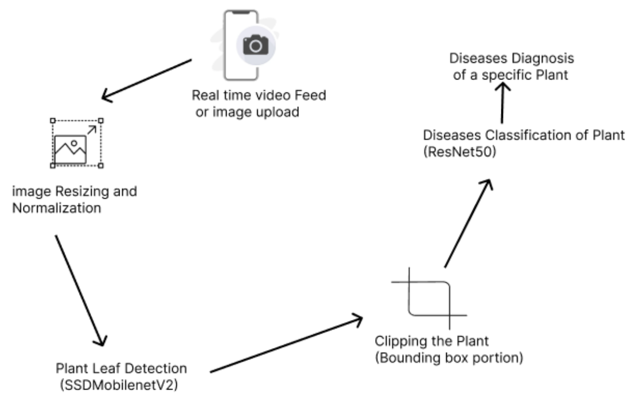


Fig. 2: Workflow of Proposed Model

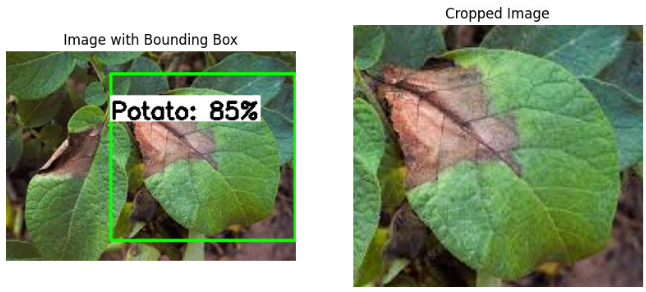


Fig. 3: Cropped region used for Potato disease classification

of the image is further processed for image classification and the clustered or irrelevant background is clipped out. This will make the image classification model only to focus on the foreground provided by clipping the bounding box portion given by first step. The process begins by detecting leaves in the image followed by selection of the identified leaf region.This is illustrated in Figure 3, Figure 4, and Figure 5, where each step in the detection and selection pipeline is visualized.

B. Diseases Classification

After the completion of above step the clipped portion of image is passed to ResNet-50 model to classify the diseases of a detected plant. ResNet-50 model architecture incorporates the concept of residual learning thus allowing the model to accurately classify complex pattern making it ideal to classify various plant diseases. It ensures the precise and reliable diseases diagnosis as it has to focus only on the relevant part of the image having plant leaf.

V. Results and Discussion

A. Plant Detection Model

The performance of plant detection model was evaluated using a method called mean average precision(mAP). This

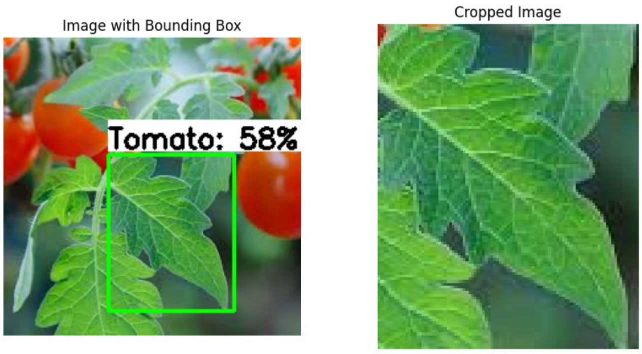


Fig. 4: Cropped region used for Tomato disease classification

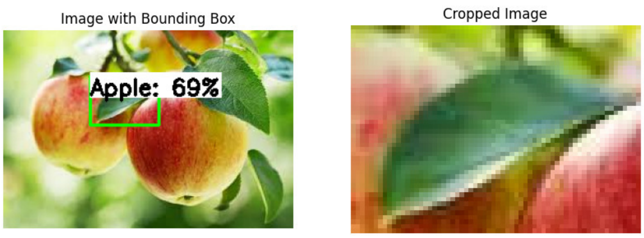


Fig. 5: Cropped region used for Apple disease classification

metric is commonly used in the object detection model's performance evaluation. The mAP is combination of precision and recall for each class giving the average value of the model performance. The mAP @ 0.5:0.95 for different plants is presented in Table III.

TABLE III: mAP @ 0.5:0.95 for Plant Detection Model

Plant	**mAP (%)**
Tomato	40.57
Peach	54.74
Corn	78.64
Apple	69.45
Potato	53.43
Bell Pepper	38.50
Cherry	81.88
Grape	75.27
Strawberry	64.65
Overall	**60.91**

The overall mAP was obtained to be 0.62 resulting good accuracy for plant detection model. The plant detection model has to be lightweight making it run effectively on mobile devices in real-time. Hence, this model is made less complex and maintaining reasonable accuracy.

B. Diseases Classification model

The performance of classification model was evaluated using accuracy, f1-score and confusion matrix. The f1 score is the combination of precision and recall and tells us how well is our model performing. Additionally, the confusion matrix gives us the description of True Positive(TP), True Negative(TN), False Positive(FP) and False Negative(FN) allowing to compare the true labels with the predicted ones.

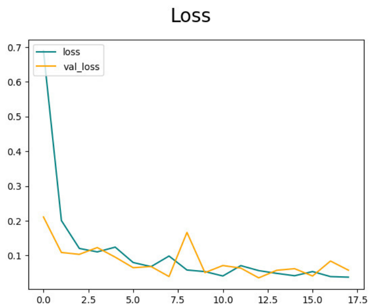
The Figure 6 illustrates the loss plots for the classification models of various plant species, including Apple, Corn, Cherry, Peach, Potato, Tomato, Bell Pepper, Strawberry, Grape, as well as the Combined Model. All the above mentioned loss plots are used as the performance evalutation metric of the ResNet-50 model. In each figures we can observe that the both training loss and validation loss are decreasing which assures that the model is performing well and is converging after each epochs. The fluctuations in the loss plots represent that the model is attempting to adjust it's weight parameters to minimize the error during training. But as the epoch are increasing the loss is decreasing indicating that model learns and improves. The fluctuations are common and indicate that during certain epochs the model is encountering challenges in certain epochs leading certain increase in loss during some epochs. To confirm there is no overfitting the early stopping techniques are applied with a patience of five. Likewise the best weights are taken as the final weights for classification model.

Figure 6j represents the combined model having 33 classes of all plants diseases. The loss value is decreasing after each epochs showing that the model is learning and converging. In comparison, the average test loss value for the individual plant models is 0.06937 with a standard deviation of 0.074, whereas the test loss value for the combined model is 0.182. This indicates that the study that we have done to isolate the foreground from the cluttered background and making the image classification model focus only on the relevant part of plant has the increased performance than the combined model having all thirty three classes. Additionally it suggests that the model with trained with less number of classes tends to perform more accurately than will higher number of classes in the plant detection task.

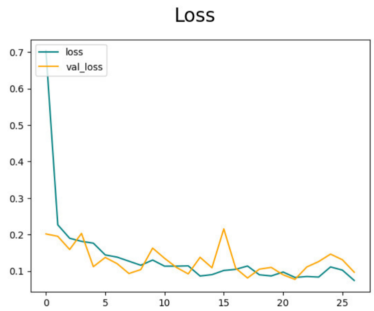
The Figure 7 presents the confusion matrices for various plant species, including Apple, Corn, Cherry, Peach, Potato, Tomato, Bell Pepper, Strawberry, and Grape. Additionally, the confusion matrix for the combined model is also included in Figure 7j. Each matrix provides insights into the model's performance in accurately classifying the respective plant species. The confusion matrix of different plants under study reveals that all the plant models are performing very well with high accuracy in testing datasets and effectively classifying the diseases related to a specific plant species. The high diagonal values indicate that models are very effective and the values outside the diagonal are lesser for most of the plants indicating a high accuracy is maintained. Likewise, after viewing the combined plant model having 33 classes it also has high diagonal values indicating the model performs well in the combined model as well but on average comparison the individual plant model outperforms the combined model.

The performance of the image classification model for different plants is summarized in Table IV. The table presents the test loss and accuracy along with the F1-score for each plant.

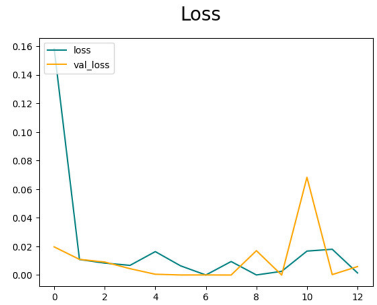
The proposed two-tier architecture for the plant leaf detection and it's diseases classification demonstrated high accuracy



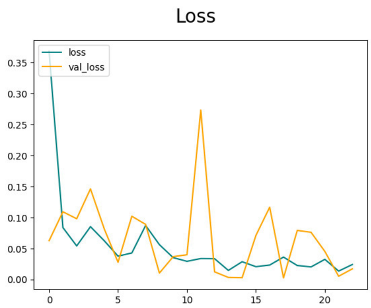
(a) Apple



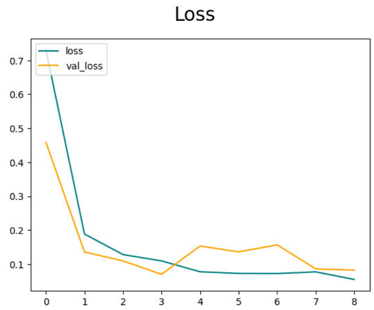
(b) Corn



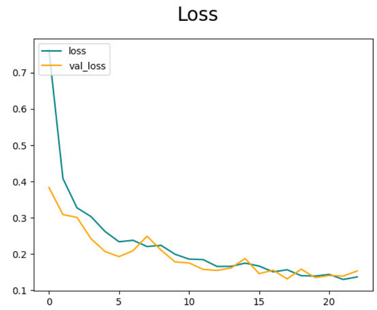
(c) Cherry



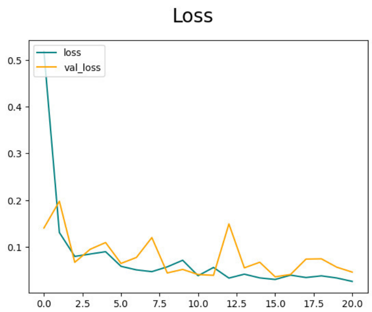
(d) Peach



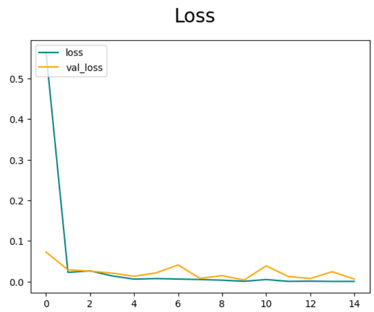
(e) Potato



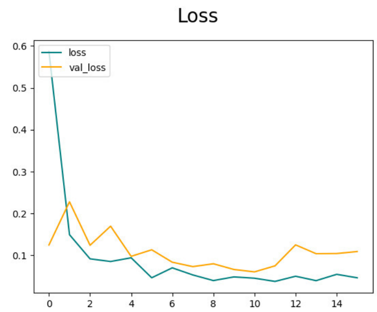
(f) Tomato



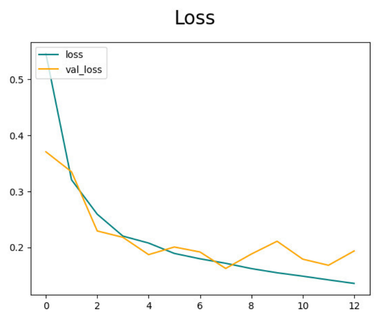
(g) Bell Pepper



(h) Strawberry



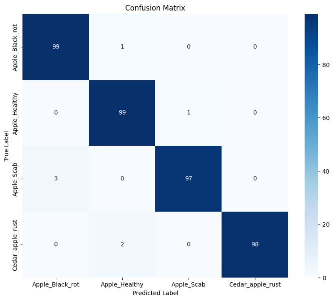
(i) Grape



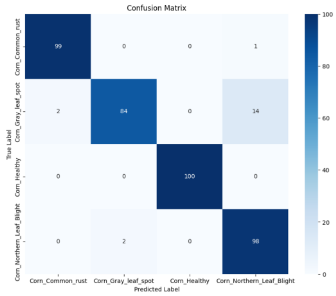
(j) Combined Model

Fig. 6: Loss plots for different plants and the combined model.

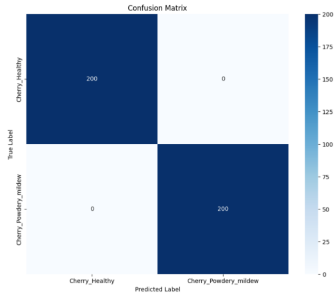
of 98.3% as compared to the combined model. The suggested approach narrowed down the classification task yielding high accuracy. In contrast, the plant species with combined 33 classes has an accuracy decreased to 94.70%. The reduction in accuracy in the combined model highlights the benefits of isolating plant species before disease classification, which simplifies the model's task and enhances overall accuracy. Similarly, the input images' cluttered and rough backgrounds



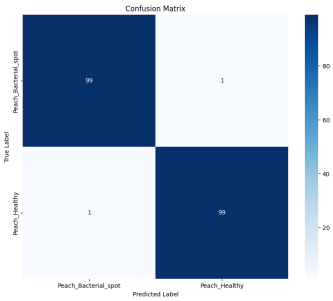
(a) Apple



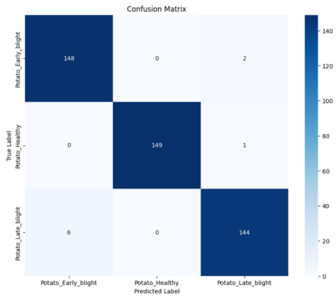
(b) Corn



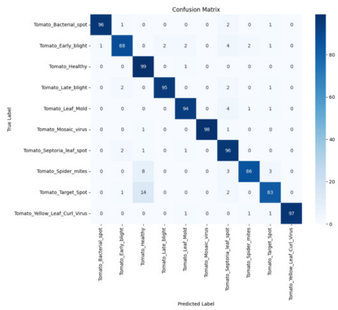
(c) Cherry



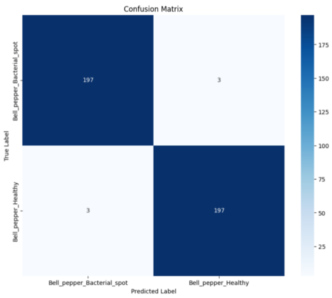
(d) Peach



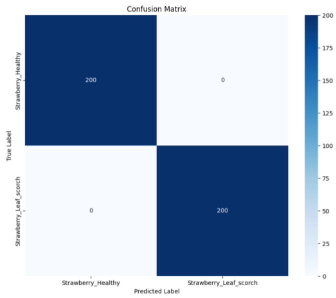
(e) Potato



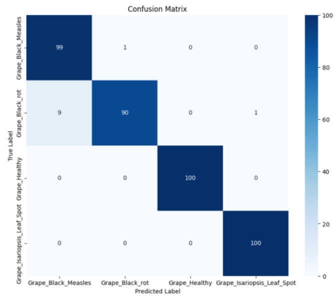
(f) Tomato



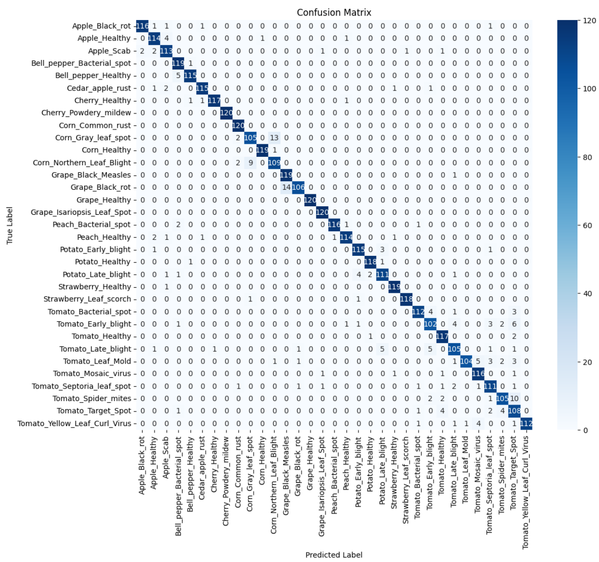
(g) Bell Pepper



(h) Strawberry



(i) Grape



(j) Combined Model

Fig. 7: Confusion matrices for different plants and the combined model.

TABLE IV: Image Classification Results

Plant	Test Loss	Test Accuracy (%)	F1 Score
Apple	0.0802	98.65	0.98
Corn	0.1123	97.5	0.96
Peach	0.0388	99.50	0.99
Cherry	0.000335	100.00	1.00
Grape	0.0605	97.25	0.97
Strawberry	0.0020	99.84	0.99
Bell Pepper	0.0282	98.50	0.98
Tomato	0.152837	95.70	0.95
Potato	0.0567	98.00	0.98
Average	–	**98.30**	**0.98**

made it difficult for the combined model to extract pertinent information, which decreased the accuracy of the classification. On the other hand, the two-tiered model successfully identified the plant in the picture and then carried out the disease classification. Using this method there is a less chance of misidentification of one plant's diseases as another thus increasing prediction accuracy.The ability to first isolate the plant before classification allowed the model to handle complex backgrounds more effectively, contributing to the overall improvement in performance.

Similarly it can also be observed that the individual classification models where there are higher number of classes tends to have lower accuracy when compared with the models with fewer classes. In addition to it there was plant species like Corn, Grape, Tomato where there was impact of background noise in the dataset and the spots on the leaves for diseases detection seems to be similar which made it challenging for the model to accurately classify the diseases. Because of these factors, the accuracy of certain plant diseases detection model has declined in comparision to others as seen in the confusion matrix.

C. Comparative Analysis

In this section we compare our model with the similar and existing systems. The comparison is based on the accuracy, the number of data samples, dataset augmentation used and the model architecture.The comparison is given in the following table.

The table V gives the comparative analysis of our proposed model with some of the existing models on the basis of the testing accuracy, number of samples used, number of classes and the model architecture. From the table it is clearly observable that our proposed model with an of SSD Mobilenet V2 and ResNet-50 outperforms other existing models either with the less dataset used or the high accuracy with the less number of training samples as compared to others. Our proposed model addresses the significant improvement in the image classification task specially in the complex and cluttered backgrounds.

Similarly in the comparative anaysis of the dataset augmentation [16] and [10] applied the advanced data augmentation technique like brightness variations and image shearing of images. Other models applied simple augmentation techniques like flipping and rotation and the absence of techniques like contrast variation, brightness variation can negatively impact the performance of model on new and unseen datatsets. Since our model incorporated advanced augmentation techniques like varying light conditions images and adjustment in contrast and brightness, it has demonstrated high accuracy and greater robustness as compared to other models shown in Table V.

Having the entire focus on relevant section of plant our model enhances the performance in terms of accuracy for the real-time applications in agriculture. The successful implementation of this can address the technological gap between the farmers and improved agricultural methods resulting in decreased diseases related loss and thereby improving plants management and increasing productivity.

VI. Conclusion

The proposed model effectively detects plants and classifies plant diseases using SSD Mobilenet V2 for object detection and ResNet-50 for disease classification. The performance and accuracy results achieved through the proposed model prove very promising in further enhancing agricultural practices. The promising results indicate the approach to be significantly benefit in plant health monitoring.

Despite the encouraging outcomes, there is still a room for further improvement that can be achieved by increasing the number of available training data on the plant species. This would further lead to a much-improved performance with regard to the lower detection and classification accuracy in plant species. Improvement can still be attained by trying other more advanced object detection algorithms for model's robustness under varying environmental conditions. While our model is typically deployed as mobile application it can work well in both the controlled environment and in various weather conditions. Only the factor that can impact the performance of model is because of the background noised in the real-time image frames and other factors such as motion blur, poor lighting conditions and occlusions.

The proposed model will also be further enhanced to support the extended range of plant species within the application and adaptation of the model for diverse geographical locations and farming practices. With these in consideration, our model aims to create a holistic agricultural management tool that strives for better health and productivity in crops, considering the factors of the above mentioned.

References

[1] B. R. Neupane, "Contribution of expenditure to agriculture growth in nepal," pp. 119–131, 2023. [Online]. Available: https://doi.org/10.3126/qjmss.v5i1.56502

[2] S. Sakamoto, W. Putalun, S. Vimolmangkang, W. Phoolcharoen, Y. Shoyama, H. Tanaka, and S. Morimoto, "Enzyme-linked immunosorbent assay for the quantitative/qualitative analysis of plant secondary metabolites," *Journal of natural medicines*, vol. 72, pp. 32–42, 2018.

[3] F. Martinelli, R. Scalenghe, S. Davino, S. Panno, G. Scuderi, P. Ruisi, P. Villa, D. Stroppiana, M. Boschetti, L. R. Goulart *et al.*, "Advanced methods of plant disease detection. a review," *Agronomy for sustainable development*, vol. 35, pp. 1–25, 2015.

[4] K. Alemu, "Real-time pcr and its application in plant disease diagnostics," *Adv. Life Sci. Technol*, vol. 27, pp. 39–49, 2014.

TABLE V: Comparison of Proposed Model with Existing Systems

Study	Model Architecture	Number of Sample Images	Dataset Used	Number of Classes	Accuracy
[10]	CNN	20,636	PlantVillage	12	98.02%
[16]	YOLOV3+ResNet-18	36,000	PlantVillage	29	96%
[11]	CNN	35,000	Not Mentioned	21	96.5%
[9]	CNN with squeeze-and-excitation(SE)	-	Dataset from challenger.ai	10	91.7%
Proposed Model	**SSD MobileNet V2 + ResNet-50**	**40,998**	**Custom + Augmented Data**	**33**	**98.3%**

[5] S. Arivazhagan, R. N. Shebiah, S. Ananthi, and S. V. Varthini, "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features," *Agricultural Engineering International: CIGR Journal*, vol. 15, no. 1, pp. 211–217, 2013.

[6] C. Yang, "Remote sensing and precision agriculture technologies for crop disease detection and management with a practical application example," *Engineering*, vol. 6, no. 5, pp. 528–532, 2020.

[7] N. Gogoi, B. Deka, and L. Bora, "Remote sensing and its use in detection and monitoring plant diseases: A review," *Agricultural Reviews*, vol. 39, no. 4, pp. 307–313, 2018.

[8] Z. Wang, K. Wang, F. Yang, S. Pan, and Y. Han, "Image segmentation of overlapping leaves based on chan–vese model and sobel operator," *Computers and Electronics in Agriculture*, vol. 148, pp. 51–58, 2018.

[9] J. Hang, D. Zhang, P. Chen, J. Zhang, and B. Wang, "Classification of plant leaf diseases based on improved convolutional neural network," *Sensors*, vol. 19, no. 19, 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/19/4161

[10] M. A. Jasim and J. M. AL-Tuwaijari, "Plant leaf diseases detection and classification using image processing and deep learning techniques," pp. 259–265, 2020.

[11] S. V. Militante, B. D. Gerardo, and N. V. Dionisio, "Plant leaf detection and disease recognition using deep learning," in *2019 IEEE Eurasia conference on IOT, communication and engineering (ECICE)*. IEEE, 2019, pp. 579–582.

[12] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in Neural Information Processing Systems*, vol. 28, 2015. [Online]. Available: https://arxiv.org/abs/1506.01497

[13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, 2016. [Online]. Available: https://doi.org/10.1109/CVPR.2016.91

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016. [Online]. Available: https://doi.org/10.1109/CVPR.2016.90

[15] M. Akila and P. Deepan, "Detection and classification of plant leaf diseases by using deep learning algorithm," *International Journal of Engineering Research & Technology (IJERT)*, vol. 6, no. 7, pp. 1–5, 2018.

[16] A. Venkataramanan and P. Agarwal, "Plant disease detection and classification using deep neural networks," 08 2019.

[17] P. S. Kanda, K. Xia, and O. H. Sanusi, "A deep learning-based recognition technique for plant leaf classification," *IEEE Access*, vol. 9, pp. 162 590–162 613, 2021.

[18] K. KC, Z. Yin, D. Li, and Z. Wu, "Impacts of background removal on convolutional neural networks for plant disease classification in-situ," *Agriculture*, vol. 11, no. 9, p. 827, 2021, this article belongs to the Special Issue Latest Advances for Smart and Sustainable Agriculture. [Online]. Available: https://www.mdpi.com/2077-0472/11/9/827

[19] Y. Zhang, J. Zhang, Y. Wang, and X. Liu, "How does the data set and the number of categories affect cnn-based image classification performance?" *Journal of Software*, vol. 14, no. 4, pp. 168–181, 2019.

[20] A. K. Ali, A. M. Abdullah, and S. F. Raheem, "Impact of the classes' number on the convolutional neural networks performance for image classification," *International Journal of Advanced Science Computing and Engineering*, vol. 5, no. 2, pp. 119–128, 2023.

[21] P. P. Wagle and M. Manoj Kumar, "A comprehensive review on the issue of class imbalance in predictive modelling," *Emerging Research in Computing, Information, Communication and Applications: Proceedings of ERCICA 2022*, pp. 557–576, 2022.

[22] S. Mascarenhas and M. Agarwal, "A comparison between vgg16, vgg19 and resnet50 architecture frameworks for image classification," in *2021 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON)*, vol. 1, 2021, pp. 96–99.

[23] N. N. F. Giron, R. K. C. Billones, A. M. Fillone, J. R. Del Rosario, A. A. Bandala, and E. P. Dadios, "Classification between pedestrians and motorcycles using faster r-cnn inception and ssd mobilenetv2," pp. 1–6, December 2020.

[24] R. A. Boukabouya, A. Moussaoui, and M. Berrimi, "Vision transformer based models for plant disease detection and diagnosis," in *2022 5th International Symposium on Informatics and its Applications (ISIA)*. IEEE, 2022, pp. 1–6.

[25] U. Barman, P. Sarma, M. Rahman, V. Deka, S. Lahkar, V. Sharma, and M. J. Saikia, "Vit-smartagri: vision transformer and smartphone-based plant disease detection for smart agriculture," *Agronomy*, vol. 14, no. 2, p. 327, 2024.