CSE 465 PROJECT REPORT

Ristian Uddin	**ID # 2021188642**
Mohammod Abdullah Bin Hossain	**ID # 2022351642**

To

Dr. Nabeel Mohammed

Associate Professor

Department of Electrical and Computer Engineering

North South University

Individual Contribution Table

Section	Contributing Member Name
IEEE/LaTEX formatting	Mohammod Abdullah Bin Hossain
Grammarly check	Mohammod Abdullah Bin Hossain
Abstract	Mohammod Abdullah Bin Hossain
Keywords	Mohammod Abdullah Bin Hossain
Introduction Motivation	Mohammod Abdullah Bin Hossain
Paper Review	Mohammod Abdullah Bin Hossain
Introduction Second-Last Paragraph (describe your work)	Ristian Uddin
Proposed System (Dataset and Preprocessing)	Ristian Uddin
Proposed System (Model description)	Ristian Uddin
Results	Ristian Uddin
Fig. and Table Title Formatting	Ristian Uddin
Conclusion and Discussion	Ristian Uddin
Equations formatting	Ristian Uddin
References Formatting in IEEE format	Mohammod Abdullah Bin Hossain, Ristian Uddin

Image-Based Plant Disease Classification Based On Deep Learning Techniques

Ristian Uddin
Electrical and Computer Engineering
North South University
Dhaka, Bangladesh
ristian.uddin@northsouth.edu

Mohammod Abdullah Bin Hossain
Electrical and Computer Engineering
North South University
Dhaka, Bangladesh
bin.abdullah@northsouth.edu

***Abstract*: Crop diseases pose a significant threat to food security, yet their prompt identification remains challenging in many regions due to the lack of essential infrastructure. The convergence of rising global smart phone penetration and recent strides in computer vision, facilitated by deep learning, has opened avenues for smart phone-assisted disease diagnosis. Leveraging a public dataset comprising 55,562 images of plant leaves with diseases and in healthy conditions collected under controlled settings, we employed a deep convolutional neural network for the identification of 12 crop species, 26 diseases and 1 background pictures with no leaves. The initially trained model demonstrated noteworthy accuracy, achieving 80.1% accuracy. To enhance efficiency, we opted for knowledge distillation, employing a smaller and less complex student model. Through this process, we achieved a close approximation with 78.1% accuracy, incorporating crucial refinements. This underscores the significance of knowledge distillation in refining models, and despite a slight reduction in accuracy; the added nuances contribute substantially to the effectiveness of smart phone-assisted crop disease diagnosis on a global scale.**

Keywords— Image-based identification, Plant diseases, Model, GoogleNet, Knowledge Distillation

I. INTRODUCTION

The identification and management of plant diseases stand as pivotal components in safeguarding crop health, bolstering agricultural productivity, and addressing the ever-pressing challenges of global food security. Traditional methods of disease identification, reliant on visual inspection by human experts[1], suffer from inherent drawbacks such as subjectivity, time intensiveness, and susceptibility to human error. In contrast, the advent of computer vision and machine learning technologies has ushered in a new era, offering automated and precise solutions for disease identification in plants.

Fig 1: Pictures of healthy and infected leaves

Among these technologies, Convolutional Neural Networks (CNNs) [9] have emerged as a formidable force in image analysis and recognition tasks, with particular prowess in plant disease identification. CNNs exhibit an innate ability to learn intricate features from images, utilizing hierarchical layers of convolutional filters. Through extensive training on large-scale datasets of plant images, CNNs become adept at recognizing nuanced patterns and distinctive characteristics associated with diverse diseases affecting crops. This research paper serves as a comprehensive overview of the application of CNNs in the realm of plant disease identification. Delving into the underlying principles of CNN architecture, the paper elucidates the suitability of this approach for the intricate task of plant disease recognition. Furthermore, it meticulously explores the challenges inherent in dataset acquisition, preprocessing, model training, and evaluation [2]. A key focal point of the paper is the comparative analysis of CNN-based approaches against traditional methods and alternative machine learning algorithms. The study not only sheds light on the superior performance of CNNs but also elucidates the potential implications of this technology for revolutionizing plant disease management. By offering rapid and accurate identification, CNNs empower timely interventions and optimization of crop protection strategies. The automation of disease identification using CNNs has the potential to provide farmers and agricultural practitioners with invaluable insights, facilitating effective disease control and mitigation. This transformative technology contributes to the paradigm shift towards sustainable agriculture practices, thereby addressing global food security concerns with enhanced precision and efficiency[3].

II. LITERATURE REVIEW ON RELATED WORK

Plant pathologists play a crucial role in the identification and classification of plant diseases by meticulously analyzing various plant components, including roots, kernels, stems, and leaves. Their examination involves the careful observation of symptoms, such as discoloration, lesions, or deformities, which can vary across different parts of the plant. In the realm of modern research, Ferentinos [20] has contributed significantly to this field. In his paper, he introduced innovative approaches by implementing various Convolutional Neural Network (CNN) architectures for the classification of plant diseases, achieving remarkable accuracy. This technological integration allows for the automation of disease detection, providing a more efficient and precise means of identifying and categorizing plant ailments. By harnessing the power of machine learning, plant pathologists can enhance their diagnostic capabilities, ultimately contributing to more effective disease management strategies in agriculture. The continuous evolution of these computational methods holds great promise for advancing our understanding of plant pathology and improving overall crop health.

Dheeb Al Bashish et al. introduced an innovative approach for the segmentation and classification of leaf diseases by proposing the utilization of K-means clustering for segmentation and Artificial Neural Networks (ANN) for disease detection. The model they presented achieved an impressive accuracy rate of 93%. Noteworthy strengths of this proposed model include its high effectiveness in disease recognition. However, it is important to acknowledge that finer segmentation and feature extraction are necessary, representing potential challenges associated with its implementation [21]. In a related study, Anand H. Kulkarni et al. put forth a model centered on an ANN classifier for the classification and recognition of diseased leaves. The incorporation of Gabor filter for image filtration and segmentation enhanced the model's capabilities. The accuracy of this proposed model reached 91%. Among its strengths is the development of robust classification and recognition mechanisms. On the downside, there is room for improvement in the form of better classifiers that could potentially enhance recognition rates [22]. These studies collectively underscore the continuous efforts to employ advanced computational techniques for accurate and efficient plant disease diagnosis.

In the pursuit of advancing plant disease classification, S. Arivazhagan and colleagues introduced a Support Vector Machine (SVM) classifier, showcasing a commendable 94% accuracy in their proposed model. The inherent strengths of this model lie in its capacity for automatic detection and classification of leaf diseases. However, it is not without its challenges, as the use of Neural Network (NN) classifiers may be necessary to achieve optimal performance, albeit at the cost of increased complexity [23].On a parallel trajectory, Usama Mokhtar and his team presented a novel approach by employing SVM with diverse kernel functions for the detection of tomato leaf diseases, achieving an impressive annotation accuracy of 99.5%. The merits of their method extend to the delivery of effective and reliable results in disease identification. Nevertheless, it is important to note that the deployment of a large scale of inputs introduces a potential drawback, contributing to a decrease in overall performance [24]. These research endeavors underscore the ongoing exploration of diverse classification techniques, each with its unique advantages and limitations; in the ongoing quest for more robust and accurate plant disease diagnosis and management strategies. In the realm of plant classification research, Atole and collaborators [25] introduced a pioneering study that harnessed the power of a pre-trained AlexNet deep network. Their work focused on classifying rice plants into distinct categories, achieving a noteworthy accuracy of 91.23%. Shifting gears to the realm of grape plant disease identification, A. P. Singh contributed to the literature with a research paper that showcased the utilization of artificial bee colony for feature selection. This innovative approach aimed to determine the optimal feature set crucial for identifying diseases in grape plant images, employing a support vector machine classifier. The algorithm put forth by Singh demonstrated an impressive 92.14% accuracy, emphasizing the effectiveness of feature selection techniques in enhancing the precision of plant disease identification [26]. These research endeavors underscore the diversity of methodologies employed in the pursuit of accurate and efficient plant classification and disease identification.

In the realm of innovative plant disease classification, Akshay Kumar and collaborators introduced CNN-based architectures. Notably, their proposed model demonstrated exceptional performance, with VGGNet leading the way and achieving an impressive 99.25% accuracy. The strengths of this model lie in its ability to reach maximum accuracy with minimal computational loss. However, it is essential to acknowledge the associated challenges, such as the time-consuming nature of VGGNet training and the demand for high-end hardware configurations, which pose potential drawbacks to its widespread application [27].

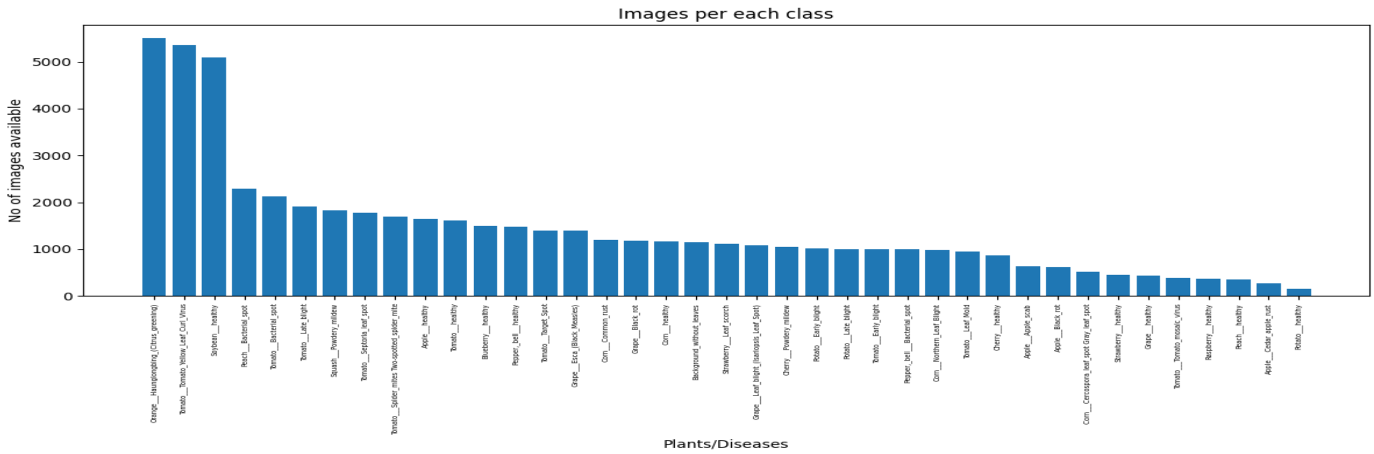
In a separate endeavor, Mehmet Metin Qzguven and team put forth a faster R-CNN architecture, albeit one that demands significant time in disease detection rates. Despite this, the model achieved a noteworthy overall classification accuracy of 99.25%. The advantages of this approach encompass reduced time spent on disease identification and a decrease in human errors during the identification process. Nonetheless, it is crucial to consider the limitations associated with large datasets, as they can result in a decrease in disease detection accuracy, representing a notable drawback to the model's effectiveness [28]. These distinct contributions highlight the multifaceted landscape of plant disease classification methodologies, each with its own set of strengths and challenges, pushing the boundaries of precision and efficiency in this critical domain. In their innovative work, Surampalli Ashok and collaborators introduced a methodology based on Convolutional Neural Networks (CNN) that exhibited an impressive 98% classification accuracy. The strengths of their approach manifest in its high-performance capabilities. However, it is worth noting that the reliance on recent algorithms and classifiers may be essential for optimizing results, presenting a potential limitation [29].

III. DATASET DESCRIPTION

Within this section, we present a summary of the dataset and its characteristics, data preprocessing, and an overview of the models, approaches, and resources used in this study.

We scrutinize a total of 55,562 images depicting various plant leaves, each associated with one of 39 class labels representing distinct crop-disease pairs. The non-leaf class, a distinct category in the dataset, adds an extra layer of complexity to the task. This class might include images of stems, roots, or other plant parts, presenting a unique challenge for CNNs to differentiate and identify the context of the given image. Our objective is to predict the specific crop-disease pair based solely on the image of the plant leaf. Illustrated in Figure 1 are representative examples from each classes sourced from the Plant Village dataset. Throughout the methodologies outlined in this paper, a standard resizing operation is applied to the images, adjusting them to 256 × 256 pixels. Both model optimization and predictions are conducted on these downscaled images to ensure consistency across our analyses.

Fig 2: Histograph diagram of samples in each class



The PlantVillage dataset[10], acknowledged for its prevalence in image classification, stands as a ubiquitous resource offering an extensive array of plant images accompanied by diverse class labels. This dataset significantly facilitates the task of multiclass classification.

IV. APPROACH

In order to gauge the performance of our methodologies on novel and unexplored data, and concurrently monitor the potential occurrence of overfitting in any of our approaches, we execute all experiments across an extensive spectrum of different types of datasets.

In this study, we leverage diverse datasets comprising images of Tomato, Corn, Apple, Grapes, Strawberry, and Potato plants for the identification of plant diseases. We first split the dataset into three sets: train – 0.70, test – 0.15 and valid – 0.15. Then we augmented the dataset into three types, first one is the original one, the second one is randomly augmented i.e. zoomed in to 20% or sometimes rotated to 30 degree or sometimes flipped horizontally and the last one is experimented with a gray-scaled version.

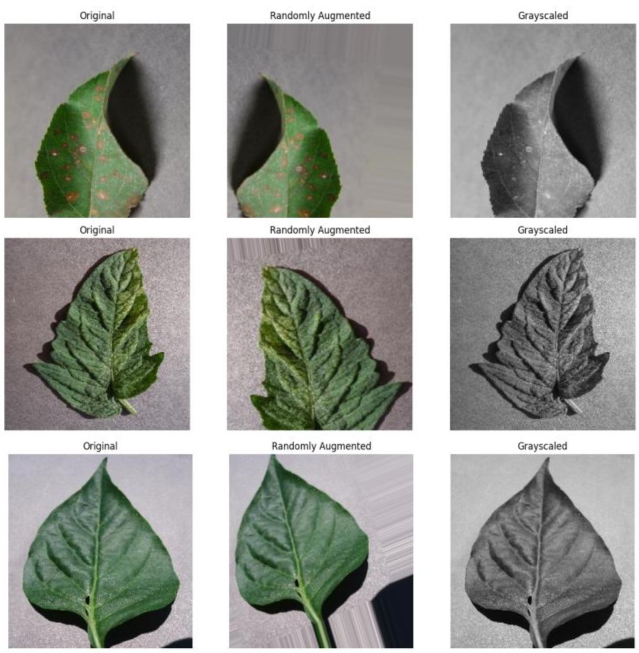


Fig 3: Pictures of samples of the three types of datasets

The intent behind this series of experiments was to ascertain whether the neural network genuinely acquires an understanding of the conceptual framework pertaining to plant diseases or if its learning process merely involves internalizing the inherent biases within the dataset. Figure 3 illustrates diverse renditions of the same leaf, selected randomly for analysis. The aim was to discern the depth of the neural network's comprehension regarding plant diseases, probing whether it encapsulates the essence of these conditions or if its learning is confined to the ingrained biases embedded within the dataset. The visual representation offers a comparative glimpse of various iterations of a leaf, serving as a snapshot of the network's interpretative capabilities.

Since 2015, research in leaf disease identification has predominantly centered on deep learning techniques. Hong et al.[19] characterized deep learning as a transfer learning form, emphasizing its capacity to autonomously determine optimal data[17] representation. The broad applications of deep learning, spanning disease diagnosis, pest identification, quality control, marketing, automation, and big data, stand to significantly impact the agricultural industry[13].

V. METHODOLOGY

As we reached one of the important stages of the process, which is about implementing a Convolutional Neural Network model to train, we decided to do a bit of research about which model to use. As we have seen in the past papers, most of them have implemented either ResNet[4] or AlexNet[5], so we decided to go with another popular architecture, namely GoogleNet Inception v3 model, which was mainly designed in the context of the “Large Scale Visual Recognition Challenge”. This is a pre-trained model and is much wider and deeper architecture with a considerable number of convolutional layers with around 21 million parameters. By applying filters at various layers, this CNN captures spatial and temporal dependencies in images. We also applied transfer learning to retain the knowledge learned by the pre-trained model during training. And towards the end of this complex architecture, we also added a Global Average Pooling layer that averages the values of each feature map and a Dense layer with 39 units and a softmax activation function is added on top of the Global Average Pooling layer. This is the final layer responsible for producing predictions.

The dataset, in three versions (original, augmented, and gray-scale), exhibits distinct performance variations across experiments, while maintaining constant

Table 1: Analysis of plant disease identification using deep learning.

References	CNN Network Used	Dataset Used	Accuracy	No of Classes
[8]	LeNet architecture	PlantVillage (extended)	91%	3
[11]	AlexNet	PlantVillage (extended)	88%	9
[10]	ResNet34	PlantVilllage	96.4%	10
[14]	GoogleNet	Field Images	89.5%	5

experimental configurations. Colored versions yield the best results. Concerned about potential biases related to lighting conditions and data collection methods, we tested adaptability using the gray-scaled dataset. Performance decreased in gray-scaled version (70.1%) compared to colored versions (78.3% and 80.1%), as anticipated.

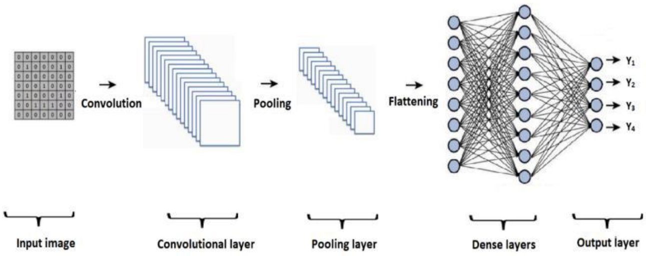


Fig 4: Transition in a CNN

After training the model, we decided to enhance and optimize the training of neural networks, so we used the Knowledge Distillation technique to compress the Inception model into a smaller and more computationally efficient student model. This is crucial for deploying models on resource-constrained devices, where smaller models can achieve similar performance with reduced computational requirements.

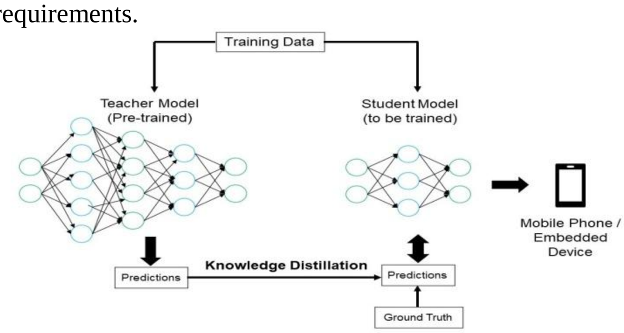


Fig 5: Flowchart of Knowledge Distillation

The student model consists of four convolutional layers with Leaky ReLU activation and max-pooling, progressively reducing spatial dimensions. The final layers include a flatten operation followed by two densely connected layers, the first with 128 neurons and Leaky ReLU activation. The output layer, named 'logits,' has 39 neurons with softmax activation, suitable for multiclass classification.

After generating the soft target labels from the trained teacher model, we applied the distillation method through two Loss functions, adjusting the impact of the soft targets with temperature scaling. This is a typical setup for knowledge distillation, encouraging the student model to not only predict the correct classes but also mimic the teacher's predictions.

$$\text{Hard Loss} = -\sum y(\text{true}_{\text{hard}}) \cdot \log(y(\text{pred}_{\text{hard}}))$$

$$\text{Soft Loss} = \sum y(\text{true}_{\text{soft}}) \cdot \log\left(\frac{y(\text{pred}_{\text{soft}})}{(y(\text{true}_{\text{soft}})}\right)$$

Where, $y(\text{true}_{\text{hard}})$ is the true class labels, $y(\text{true}_{\text{soft}})$ is the soft target prediction score from teacher model, $y(\text{pred}_{\text{hard}})$ is the predicted values from student model and $y(\text{pred}_{\text{soft}})$ is the output of student model. The two loss functions are then added, where the relative importance of the two components is controlled by the temperature, T.

$$\text{Total Loss} = \text{Hard Loss} + \frac{\text{Soft Loss}}{T^2}$$

VI. RESULTS

The performance across all experiments exhibits characteristic variations among the three versions of the dataset (original, augmented, and gray-scale) when keeping the rest of the experimental configuration constant. Notably, the models demonstrate their highest performance levels when applied to the colored version of the dataset. So we decided to proceed with the augmented set for distillation process, where we trained the student model by constantly fine-tuning the hyper-parameter Temperature to get the best possible accuracy to mimic the result of teacher model.

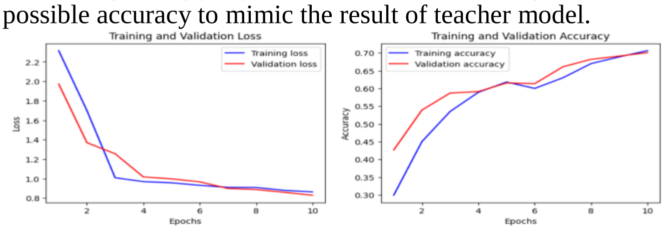


Fig 6: Loss and accuracy for Gray-scale dataset

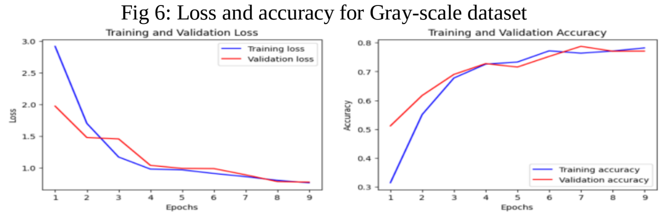


Fig 7: Loss and accuracy for Original dataset

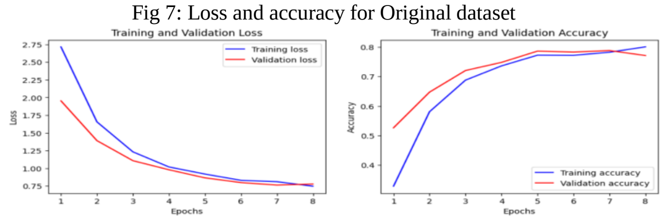


Fig 8: Loss and accuracy for Randomly Augmented dataset

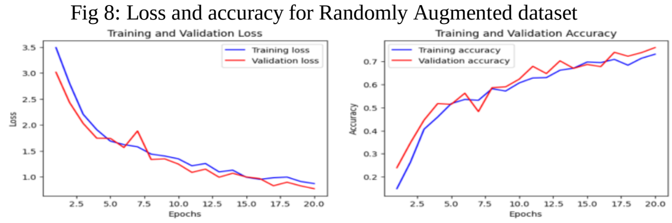


Fig 9: Loss and accuracy from student model when Temperature = 3

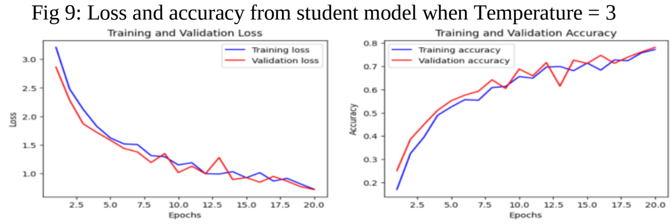


Fig 10: Loss and accuracy from student model when Temperature = 6

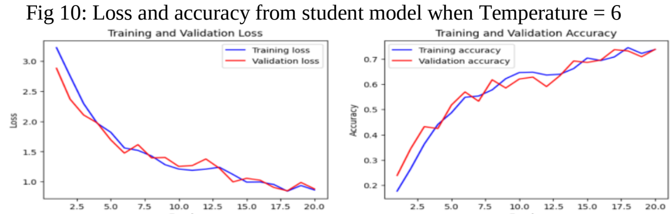


Fig 11: Loss and accuracy from student model when Temperature = 9

We also conducted an experiment on our student model without any distillation or transfer learning method, we just trained the model with raw data to see and compare if the Knowledge Distillation method does many any improvement at all study.

Fig 12: Loss and accuracy without Distillation method

To summarize, we conducted all these different experimental procedures between 10-20 epochs maintaining slight difference between train and valid accuracy, meaning no significant overfitting as the model performed similarly on new and unseen data as it has on the training data.

Types of Dtaset	**CNN Model**	**Accuracy**	**Valid Accuracy**	**Precision**	**F1 Score**
Original	Teacher	78.3%	77.5%	0.71	0.72
Augmented	Teacher	80.1%	78.6%	0.74	0.77
Gray-scale	Teacher	70.7%	70.1%	0.67	0.69
Augmented with soft labels(T=3)	Student	73.1%	75.7%	0.72	0.73
Augmented with soft labels(T=6)	Student	77.1%	78.1%	0.77	0.77
Augmented with soft labels(T=9)	Student	73.7%	74.2%	0.76	0.77
Augmented	Student	68.4%	69.3%	0.70	0.69

Table 2: Experimental scores across various configurations.

VII. CONCLUSION AND DISCUSSION

In recent years, there has been remarkable advancement in the performance of convolutional neural networks (CNNs) for tasks involving object recognition and image classification [16]. Subsequently, learning algorithms were applied in these feature spaces. The efficacy of these traditional approaches was highly contingent upon the predefined features established in the underlying framework. The recent surge in CNNs has significantly transformed this landscape [15], rendering them highly effective and capable in comparison to the previous reliance on hand-engineered features.

Our objective was to mimic the result score of the teacher model or even better, which was around 80.1%, and by fine-tuning the hyper-parameters, we got as close as 78.1% as we can see in Table 2. It was not an ideal score but it was as close as we got, thereby we settled it as our best one. However, we should also consider the nature of the model of both teacher and student. The teacher model was a lengthy and complex one with many possible feature maps for one route, also has around 22 million parameters whereas, the student model has roughly got 2 million with some countable layers, clearly showing that the teacher model is almost 11 times more superior.

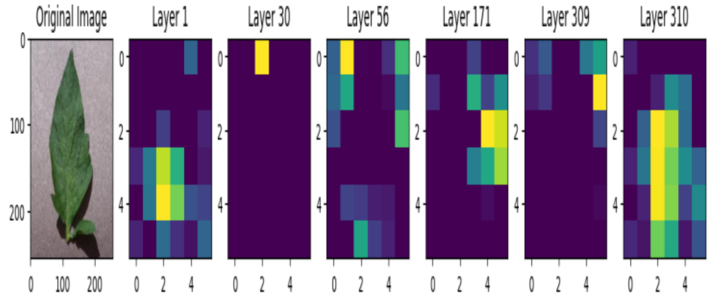


Fig 13: Feature maps extracted at different layers of Teacher Model

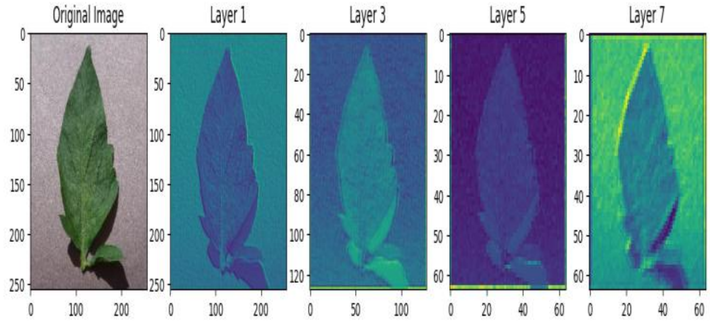


Fig 14: Feature maps extracted at all the layers of Student Model

As we can see from Fig. 13 and 14, how the same image is being transitioned in both the models. The image is being processed and inspected in a more detailed way in teacher model than that of student model. So, it can be safe to say that even after this huge margin of differentiation, the student model did well to learn from the soft target outputs to get as close as the teacher model in terms of accuracy.

While the aforementioned methodologies have demonstrated commendable efficacy on the PlantVillage dataset, meticulously curated within a controlled environment, our evaluation extends to assessing the model's proficiency in handling images extracted from esteemed online sources, notably academic agriculture extension services. The availability of such images is limited, prompting us to employ a strategic approach involving automated downloads from Bing Image Search and IPM Images. A subsequent visual verification step was implemented to ensure data integrity, resulting in the curation of two relatively diminutive yet meticulously verified datasets, comprising 121 images for Dataset 1 and 119 images for Dataset 2 [6][7].

Upon subjecting these datasets to the scrutiny of our best-performing model, we observed an overall accuracy of 31.40% for Dataset 1 and 31.69% for Dataset 2. This performance metric reflects the model's proficiency in correctly predicting the class label from a pool of 38 possible labels, underscoring its adaptability to diverse datasets beyond the confines of a controlled environment.

However, considering the increase of usage of smartphones[12], we believe that our approach can be a practical way to help prevent crop yield loss. Looking ahead, we may enhance the accuracy by adding location and time information to the image data captured by smartphones. It's important to note the rapid progress in mobile technology over the past few years and its expected continuation[18]. With the increasing number and quality of sensors in mobile devices, we anticipate that achieving highly accurate diagnoses through smartphones is just a matter of time.

REFERENCES

[1] S. M. Hassan, A. K. Maji, M. Jasiński, Z. Leonowicz and E. Jasińska, "Identification of Plant-Leaf Diseases Using CNN and Transfer-Learning Approach," MDPI, vol. 10(120, p. 10, 9 June 2021.

[2] J. Chen, J. Chen, Y. S. Defu Zhang and Y. A. Nanehkara, " Using deep transfer learning for image-based plant disease identification, " SEMANTIC SCHOLAR, vol. 173, p. 13, 6 April 2021.

[3] M. Al-Hammadi and A. Bhilare, " PLANT DISEASE IDENTIFICATION USING CNN, " PLANT DISEASES, pp. 4-6, 4 June 2021.

[4] V. K. Vishnoi, K. Kumar, and B. Kumar, *Plant disease detection using computational intelligence and image processing*, no. 0123456789. Springer Berlin Heidelberg, 2020.

[5] J. G. A. Barbedo, "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification," *Comput. Electron. Agric.*, vol. 153, no. July, pp. 46– 53, 2018, doi: 10.1016/j.compag.2018.08.013.

[6] G. Geetharamani and A. P. J., "Identification of plant leaf diseases using a nine-layer deep convolutional neural network," *Comput. Electr. Eng.*, vol. 76, pp. 323–338, 2019, doi: 10.1016/j.compeleceng.2019.04.011.

[7] G. Sambasivam and G. D. Opiyo, "A predictive machine learning application in agriculture: Cassava disease detection and classification with imbalanced dataset using convolutional neural networks," *Egypt. Informatics J.*, vol. 22, no. 1, pp. 27–34, 2020, doi: 10.1016/j.eij.2020.02.007.

[8] R. Sharma, "Overview Of Different Machine Learning TechniquesFor Plant Disease Detection," *J. Gujarat Res. Soc.*, vol. 21, no. 6, pp. 416–425, 2019, [Online]. Available: http://www.gujaratresearchsociety.in/index.php/JGRS/article/view/21 01.

[9] H. B. and R. S. B. M. K. Rafsanjani, "The Classification of Galaxy Images Using Neural Network Algorithm," in *International Conference on Computing, Power and Communication Technologies (GUCON),New Delhi, India*, 2019, pp. 127–131.

[10] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Comput. Electron. Agric.*, vol. 161, no. February, pp. 272–279, 2019, doi: 10.1016/j.compag.2018.03.032.

[11] M. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep Learning for Tomato Diseases: Classification and Symptoms Visualization," *Appl. Artif. Intell.*, vol. 31, no. 4, pp. 299–315, 2017, doi: 10.1080/08839514.2017.1315516.

[12] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Comput. Electron. Agric.*, vol. 145, no. January, pp. 311–318, 2018, doi: 10.1016/j.compag.2018.01.009.

[13] G. Shrestha, Deepsikha, M. Das, and N. Dey, "Plant Disease Detection Using CNN," *Proc. 2020 IEEE Appl. Signal Process. Conf. ASPCON 2020*, no. September, pp. 109–113, 2020, doi: 10.1109/ASPCON49795.2020.9276722.

[14] R. Sujatha, J. M. Chatterjee, N. Z. Jhanjhi, and S. N. Brohi, "Performance of deep learning vs machine learning in plant leaf disease detection," *Microprocess. Microsyst.*, vol. 80, no. December.

[15] P. Jiang, Y. Chen, B. Liu, D. He, and C. Liang, "Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks," *IEEE Access*, vol. 7, pp. 59069–59080, 2019, doi: 10.1109/ACCESS.2019.2914929.

[16] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks," *Neurocomputing*, vol. 267, pp. 378–384, 2017, doi: 10.1016/j.neucom.2017.06.023.

[17] D. Oppenheim and G. Shani, "Potato Disease Classification Using Convolution Neural Networks," *Adv. Anim. Biosci.*, vol. 8, no. 2, pp. 244–249, 2017, doi: 10.1017/s2040470017001376.

[18] A. K. Sandhu and R. S. Batth, "Software reuse analytics using integrated random forest and gradient boosting machine learning algorithm," *Softw. - Pract. Exp.*, vol. 51, no. 4, pp. 735–747, 2021, doi: 10.1002/spe.2921.

[19] A. Cruz *et al.*, "Detection of grapevine yellows symptoms in Vitis vinifera L. with artificial intelligence," *Comput. Electron. Agric.*, vol.157 no. December 10.1016/j.compag.2018.12.028.

[20] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," Computers and Electronics in Agriculture, vol. 145 pp. 311–318, 2018.

[21] Al Bashish, Dheeb, Malik Braik, and Sulieman Bani-Ahmad."Detection and classification of leaf diseases using K-means-based segmentation and." Information technology journal 10.2 (2011): 267-275.

[22] Kulkarni, Anand H., and Ashwin Patil. "Applying image processing technique to detect plant diseases." International Journal of Modern Engineering Research 2.5 (2012): 3661-3664.

[23] Arivazhagan, Sai, et al. "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features." Agricultural Engineering International: CIGR Journal 15.1 (2013): 211- 217.

[24] Mokhtar, Usama, et al. "Tomato leaves diseases detection approach based on support vector machines." 2015 11th International Computer Engineering Conference (ICENCO). IEEE, 2015.

[25] R. R. Atole and D. Park, "A multiclass deep convolutional neural network classifier for detection of common rice plant anomalies,"International Journal of Advanced Computer Science and Applications, vol. 9, no. 1, pp. 67–70, 2018.

[26] A. P. Singh, A. K. Luhach, S. Agnihotri, N. R. Sahu, D. S. Roy et al., "A novel patient-centric architectural framework for blockchain enabled healthcare applications," IEEE Transactions on Industrial Informatics, vol. 17, no. 8, pp. 5779–5789, 2021.

[27] Kumar, Akshay, and M. Vani. "Image Based Tomato Leaf Disease Detection." 2019 10th International Conference on Computing,Communication Networking Technologies (ICCCNT). IEEE,2019.

[28] Ozguven, Mehmet Metin, and Kemal Adem. "Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms." Physica A: Statistical Mechanics and its Applications 535 (2019): 122537.

[29] Ashok, Surampalli, et al. "Tomato Leaf Disease Detection Using Deep Learning Techniques." 2020 5th International Conference on Communication and Electronics Systems (ICCES). IEEE, 2020.