

Plant Disease Detection Using Deep Learning: A Focus on Pathogen-Based Classification

S.N.Tirumala Rao¹, Sireesha Moturi², M.Mounika Naga Bhavani³, Mata Ajay Kiran⁴, Kasthala Huggayi⁵, and D.Venkatareddy⁶

¹Professor, Department of CSE, Narasaraopeta Engineering College,
Narasaraopet-522601, Palnadu, Andhra Pradesh, India.

² Associate Professor, Department of CSE, Narasaraopeta Engineering College,
Narasaraopet-522601, Palnadu, Andhra Pradesh, India.

^{3,6} Assistant Professor, Department of CSE, Narasaraopeta Engineering College,
Narasaraopet-522601, Palnadu, Andhra Pradesh, India.

^{4,5} Student, Department of CSE, Narasaraopeta Engineering College,
Narasaraopet-522601, Palnadu, Andhra Pradesh, India.

sireeshamoturi@gmail.com

Abstract. Agriculture forms a very crucial section in every country's economy. Its sustainable practice requires monitoring plant health; it is essential for the disease to be detected before any visible symptoms appear so that damage to the crop is minimized and the negative effects of chemicals are reduced. This paper introduces an automated approach to the detection and classification of plant diseases using deep transfer learning models. For this work, it focused on the use of Keras-based EfficientNetV2. The system not only detects diseases but also indicates the pathogens responsible, focusing upon images from the Agri-ImageNet dataset along with real-world photographs of cauliflower and sunflower plants. In contrast to such methods, this one does not face the limitations that traditional methods do by making controlled environments through uniform background since it utilized deep models trained in images taken in natural settings. It exhibits a better accuracy with the proposed methodology, as the series EfficientNetV2 reached an accuracy of 96% in the testing procedure. This contribution also advances real-time accurate detection of plant diseases, thereby enabling early intervention and more sustainable agricultural practice.

Keywords: VGG-16, VGG-19, ResNet50, ResNet152V2, ConvNeXt models, NasNet models, EfficientNet models, EfficientNetV2 models, MobileNet models.

1 Introduction

Most of the disease of plant detection rely on visual appearance of the plant, which is not usually possible because by the time the effect of infection starts to show, the plant might already be damaged in an irreparable way. The area of automated detection of diseases in plants holds much promise. The major causes

of plant diseases include poor nutrition, invasion by microbes, rats, and adverse conditions of the environment[1]. Plants are more prone to diseases because of the high number of pathogen organisms surrounding them. Plant disease can be defined as any morphological or physiological disorder arising due to a living organism. Some type of plant pathogens or environmental factors is the cause of plant diseases. Globally, the low yield in agriculture is mainly brought about by plant infection caused by a pathogen. The various families of pathogens can infect the plant singly or together, which may be the cause of the severity of the disease[2]. Plant infection is a threat for food and result in food scarcity and enhanced food costs.

1.1 Pathogen-Plant Interactions

Pathology of plant is a scientific discipline that deals with diseases in plants. Most plant pathogens are microscopic organisms including bacteria, viruses, nematodes, fungus, and protozoa as well as several parasitic plants and algae[3]. They consume nutrients, kill plants, and spread disease by secreting enzymes, toxins, and other substances. Plant diseases create starvation, loss of crops, financial loss, and even extinction of whole species of plants because of poisonous food. Plant pathogens enter plant tissues and cause disease by taking cell growth, multiplication, and spread. The infected plant releases various toxins and enzymes. Dead patch tissues in leaves, shoots, fruits, and roots are manifestations of a pathogen. Blights are characterised by sudden death of leaves and shoots [4].

1.2 AI, ML, DL, TL, AND DTL IN PLANT DISEASE IDENTIFICATION

Advanced computer approaches have revolutionized the identification of plant diseases in agricultural sectors. Traditional methods of identifying plant diseases included visual inspection, mainly resulting in late diagnosis and massive losses to crops. However, AI, ML, DL, TL, and DTL combined have really hastened and improved the accuracy and speed of illness diagnosis.

- **Agriculture and Artificial Intelligence** Artificial intelligence was beginning to appear as a revolutionary tool in agriculture, which had answers to most problems, such as disease control, detection of pests and their infestation pattern, crop monitoring. AI-based systems, comprising of specific ML/DL algorithms, may process vast volumes of data and sound predictions for plant health.
- **Machine Learning for Detection of Disease** Being a subset of AI, ML gives it the ability to learn from data and take decisions without being explicitly programmed. In the classification of diseased and healthy plant samples, some of the techniques used in ML include decision trees, SVM, and KNN; however, these traditional ML approaches often require hand-crafted feature extraction, which severely limits its applicability to very complex patterns in plant disease data[5].

- **Deep Learning for Performance Improvement** DL has elevated the art of plant disease identification to unprecedented levels. CNNs have been used and have gained widespread acceptance in various classification tasks related to images, including disease identification in plants. These features are automatically learned by the CNNs from images, without requiring manual feature engineering.

2 RELATED WORK

The review offered a sound overview of the detection methods that are used to identify plant diseases, especially those that feature AI, ML, and DL. So that it sounds even more interesting, the literature review could also lead further into the issues associated with traditional approaches, like the problem of large datasets and extractions made manually for the characteristics. For example, the earlier techniques with the underlying platforms of SVM and decision trees were handicapped by their reliance on handcrafted features and, therefore, had limited ability to succeed in the more intricate agricultural environments.

There are reports that lately, advancements in deep learning, especially in convolutional neural networks (CNNs), have improved the detection accuracy in plant diseases. However, the images are usually collected based on controlled environments that are mostly not similar to natural conditions. Therefore, the advent of transfer learning, especially deep transfer learning (DTL), is a rich solution to these challenges. DTL, with the assistance of pre-trained models, requires no large data sets for disease detection and is thus suitable for agricultural applications where data scarcity is a significant problem.

Malik et al. developed a hybrid CNN-based approach for the disease detection of sunflower leaves. The four highlighted diseases of their research were Verticillium wilt, Phoma blight, downy mildew, and Alternaria leaf blight. In the authors' approach, the stacking methodology that falls under the ensemble learning method was used with two deep models, VGG-16 and MobileNet, to produce an effective classifier. In fact, the proposed model outperformed other methods using the same dataset at 89.2%.

Liu et al. examined an alternative strategy focusing on feature extraction with 19 colour and texture features. They classified diseased regions by a random forest method at a recognition rate of 95%. Similarly, Sara et al. shared the dataset of sunflower flowers and leaves to aid in developing algorithms for the plant disease detection system. They used a multi-step approach, including image augmentations, scaling, and developed models in a selection range so that they could perform reliably in scenarios[6].

While applying multi-scale recovery methods derived from the GoogLeNet model, Huang et al achieved 92.0%. They enhanced disease identification in different crops using the region proposal networks and also using a Softmax classifier. Liu et al attached a sparse self-encoder with CNN to improve disease diagnosis even more to an 87.2% detection rate[7].

Xue et al. utilized machine vision technologies to classify cauliflower samples according to their grade of rotting. They achieved 95% and 90.9% accuracy using PLS-DA and ELM, respectively. For crop monitoring research in agriculture, the clear and diverse set of images offered by Deng et al.'s dataset, ImageNet, allowed for the easy performance of tasks related to classification[8].

Krizhevsky et al.'s neural network architecture changed the face of deep learning, and the important milestones in object recognition give a platform for models such as ResNet and EfficientNet, which are now used widely for the purpose of detecting plant diseases[9].

3 METHODOLOGY

3.1 DATASET DESCRIPTION

The study relies on three major datasets: Agri-ImageNet, Sunflower, and Cauliflower, using images with natural background hence overcoming one of the major shortcomings of most past works that apply a more uniform background. The dataset Sunflower concentrates upon diseases which are leaf scars, grey mould, and downy mildew.

- **Sunflower Dataset** This dataset consists of images of sunflower leaves infected with leaf scars, grey mould, and downy mildew among many other diseases. To have uniform resolution in training, all the images are of 512×512 pixels. The dataset is mainly centered on these three types of fungal diseases that most often come with sunflower crops. Unlike other datasets such as PlantVillage, whose images are frequently captured against homogeneous backgrounds, pictures will be taken in real environments in order to avoid the constraints of controlled environments[10].
- **Cauliflower dataset** This archive contains images of healthy and diseased leaves and bulbs of cauliflower. Among the prevalent diseases, the cauliflower mosaic virus, sclerotinia stem rot, clubroot, downy mildew, powdery mildew, blackleg, black rot, and bacterial spot rot are included in this dataset. Information also covers issues that include the common dominance of black rot, bacterial spot rot, and downy mildew in the cauliflower crop.
- **Agri-ImageNet** Agri-ImageNet is a crop, fruit, and vegetable specialized subset of the popular ImageNet dataset, which is designed for agricultural applications in image recognition. The collection comprises a number of images of lemon, zucchini, strawberries, spaghetti squash, pineapple, mushrooms, and many others. Target applications from this dataset include crop monitoring, disease detection, and yield estimation for agribusinesses[11].
- **Preprocessing and data augmentation** Images preprocessing The datasets will become uniform through proper image preprocessing. All images will be resized to 512×512 pixels, and a variety of data augmentation techniques including rotation, zoom, and flipping are applied. This will further increase the diversity in the training set and prevent overfitting on some specific orientations or backgrounds.

- **Transfer Learning Models** It utilises deep transfer learning with pre-trained models, wherein the applied models, EfficientNetV2, have been demonstrated to balance accuracy and computational efficiency. The fine-tuned final layers classify plant diseases while the initial layers of the pre-trained model extract basic visual features.
- **Feature Extraction and Fine-Tuning** This technique is feature extraction with fine-tuning: it is not trained from scratch but instead extracts universal features such as edges, textures from pre-trained models like EfficientNet. It improves accuracy on a specific task-the task of detecting diseases in plants-whose training time remains reduced.





Sunflower Dataset				
Disease Name	Gray Mold	Leaf Scars	Downy Mildew	Disease-Free
Number of Original Images	72	120	141	134
Number of Augmented Images	398	470	509	491

Fig. 1. Sunflower(flora) dataset






Cauliflower Dataset					
Disease Name	Downy Mildew	Black Rot	Bacterial Spot Rot	Healthy Leaf	Healthy Bulb
Number of Original images	177	100	173	18	188
Number of Augmented Images	500	500	500	1000	1000

Fig. 2. Cauliflower(flora) dataset

3.2 Transfer Learning in Deep Learning

DTL, in fact, is a very powerful although advanced method that surpasses what DL has covered in its discussions, with a simple idea of allowing models to use information from one task or domain to another related but unrelated activity. DTL has evidently achieved outstanding success in the automation of recognition

and classification of plant diseases resulting from a broad spectrum of pathogens, including viruses and fungi[12]. In this work, deep learning that is used for the goal of plant disease detection. It works in the following way, in detail:

- **Pre-trained Models** Most large deep learning projects start from one of the extremely popular architectures that have been pre-trained on extensively large datasets such as ImageNet. They are VGG, ResNet, Inception, and EfficientNet. They have learned to recognize an incredibly broad variety of features, including edges, textures, and patterns, from millions of photos. They make very good places to start when you're taking on new activities[13].
- **Feature extraction** Deep transfer learning does not require training the entire model from scratch. The bottom layers of these pre-trained models which can learn low-level, fundamental features can be used directly as feature extractors. These layers represent universal properties of vision, such as edges, forms, or textures, that are informative across different contexts.
- **Fine-Tuning** The model needs to be fine-tuned to a specific task of disease identification in plants. It is achieved by stacking the last layers of the pre-trained model with new layers newly designed, solely for this purpose.
- **Good use when the dataset size is small:** Deep transfer learning comes out particularly effective when there size of target task dataset is relatively smaller . Even though the plant disease dataset in this study is not as large as that of ImageNet, the usage of pre-trained learning from ImageNet feeds the model on vast knowledge emanating from it, which improves its performance even with the tiny number of pictures of plant diseases.
- **Shorter Training Time and Lower Possibility of Overfitting with Smaller Datasets** The biggest advantages attributed to transfer Learning in Deep Learning is significantly lower training time as well as overfitting, especially with smaller datasets.
- **Real World Application** Deep transfer learning is very useful for plant disease diagnosis applications. The models are efficient and cost-effective because they do not need an initial training.

3.3 BLOCK DIAGRAM

- **Input Image** The image of the plant usually shows the damaged zones as illustrated by the impact on leaves or blossoms. Fields or datasets for training offer such images.
- **Preprocessing** At this point, input image preprocessing is done. This includes the following step: Ensure all images have uniformity in size by resizing it to the desired size, say, 512×512 pixels.
- **Normalisation** The pixel size will be decreased and input will be taken for deep learning. The process of transforming the set of data and enhancing the robustness of the model against transformation like rotation, zoom, flip, etc is called Data augmentation.
- **Pre-trained Model (Feature Extraction)** To extract features from an image. ResNet, Inception, or EfficientNet other pre-trained models are used.

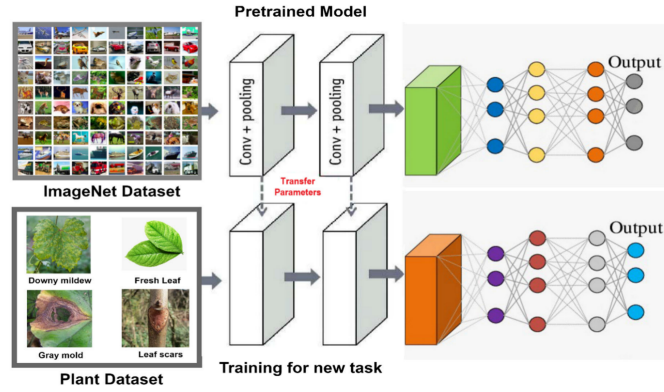


Fig. 3. Transfer learning for plant disease detection

Early layers of the models detect the general visual features like the edges, textures and shapes that might be useful for most tasks.

- **Fine-Tuning Layers** After extraction of feature, new trainable layers are added to the pre-trained models. These new layers are trained to identify the different patterns of diseases in plants for the specific task of classifying plant diseases.
- **Prediction Layer (Classification):** Output layer of the model gives probabilities for every possible class of plant disease and provides classification result. The class which is predicted by the model will be the class with maximum probability.
- **Output Prediction** Whether the plant is diseased or not is the final output predicted. If it has any disease, then which particular sickness is figured out (like grey mould downy mildew). Data helps for advising what next actions would be taken. Like for precautionary or treatment measures[14].

3.4 MODELS

- **VGGNet** Definition: VGGNet is a deep learning model known for its simplicity and use of small (3x3) convolutional filters. It is widely used for image classification tasks, including plant disease recognition, due to its effective feature extraction.
- **ResNet (Residual Networks)** Definition: ResNet is an architecture which employs skip connections so that very deep networks are viable and the problem of vanishing gradients is avoided. It has been successful in plant disease classification especially in the high-resolution images cases.
- **MobileNet Model** Definition: MobileNet is one of those deep learning models, which is lightweight and, therefore, best suited for embedded and mobile devices.
- **Inception Networks** Definition: Inception networks use convolutions of different sizes simultaneously in one layer to capture various feature scales.

Inception-based models are beneficial in plant disease detection due to their ability to extract multi-scale features.

- **NASNET** Definition: The NasNet, or Neural Architecture Search Network, was developed through the utilization of the process of neural architecture search for optimizing neural network architectures.
- **ConNeXT** Definition: ConvNeXt is a deep-learning model updating the traditional CNN architecture based on the Visual Transformer Design Cues. These designs streamline and simplify the structure of CNNs but generate competitive performance in tasks like object detection and picture categorization.
- **EfficientNet** Definition: A family of deep learning models called EfficientNet, using systematic scaling of breadth, depth, and resolution, tries to strike a balance between the size of the model, precision, and efficiency. It is highly useful for tasks like image classification because compound scaling reduces the computational cost without efficiency trade-offs. Unlike previous models, the EfficientNet models reported an advanced level of accuracy with significantly fewer parameters and computing power.
- **EfficientNetV2** EfficientNetV2 models, an evolution of the original EfficientNet, enhance both efficiency and performance through several key improvements. Introduced in the paper "EfficientNetV2: Smaller Models and Faster Training," these models refine the scaling strategy with a focus on optimizing both network architecture and training procedures. This results in models that offer faster training times and higher accuracy with even fewer parameters and computational resources compared to their predecessors, making EfficientNetV2 highly effective for a range of image classification and computer vision tasks.

4 PERFORMANCE EVALUATION AND RESULTS

4.1 PERFORMANCE EVALUATION METRICS

Deep transfer learning approach used for the different models of plant illnesses categorization success is largely dependent on performance evaluation and results. The various pre-trained transfer learning models have been evaluated and tested by using dataset consisting of Agri-ImageNet, sunflower, and cauliflower images. The models were tested with regard to many types of metrics, such as accuracy, F1-score, and also the complexity of the model, to determine which was the most efficient in the categorization of plant diseases.

- **Evaluation Metrics** The performance of the deep transfer learning models was measured using several key evaluation metrics:
- **Accuracy** It indicates the overall correctness of the model in predicting plant diseases.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

- **F1-Score** This metric combines precision and recall to provide a balanced evaluation of the model’s classification performance.

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (2)$$

- **Model Complexity:** The number of parameters (in millions) trained in each model was considered as a measure of complexity. Lower complexity models are preferred as they are computationally efficient.
- **Precision** Precision is defined as the ratio of correctly predicted positive instances to all expected positives. It gives an idea of the percentage of optimistic projections that came to pass.

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

- **Confusion Matrix** The ROC curve plots the true positive rate (recall) against the false positive rate (1 - specificity) at different threshold settings. The AUC measures the entire two-dimensional area beneath the ROC curve and is used to evaluate the model’s ability to distinguish between classes.

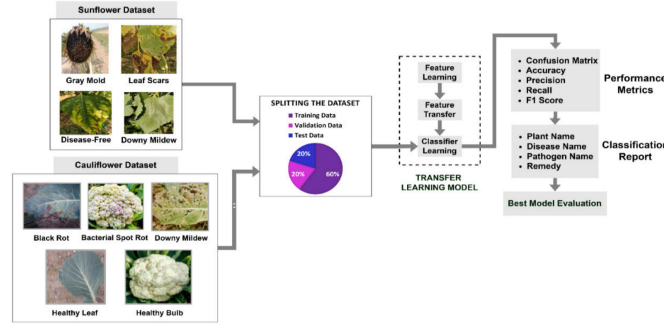


Fig. 4. detection of plant disease dataflow diagram.

4.2 Best Performing Models

The highest accuracy in consistent cases was obtained using the sunflower, cauliflower, and Agri-ImageNet datasets for the EfficientNetV2 series - namely, EfficientNetV2B2 and EfficientNetV2B3-from among the 38 tested transfer learning models, found better than those others that had established much complexity or resulted in worse performance, such as VGG-16, ResNet, and InceptionV3[15]. In terms of performance, EfficientNetV2B2 appeared to supersede deeper architectures such as InceptionResNetV2, with surprising accuracy when reduced

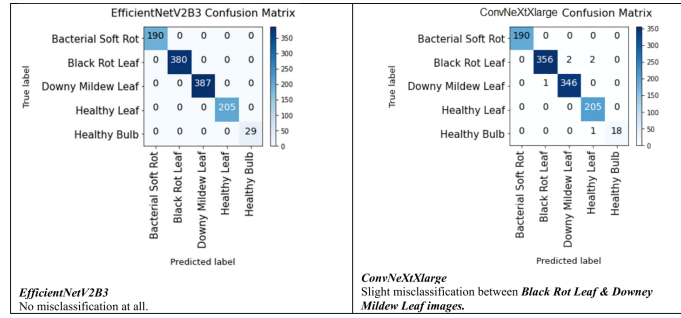


Fig. 5. Confusion matrix report of cauliflower(flora) dataset

by parameters. In addition, it showed a method of balancing the tradeoff between computing efficiency and accuracy, thus optimal for real-world applications where resources might be limited. Although InceptionResNetV2 is a standard benchmark for the evaluation, it was actually proven that the accuracy of this model was lower than those of EfficientNet models on quite complex datasets such as Agri-ImageNet.

Optimized scaling, and lower computational complexity with EfficientNetV2 allow it to be more versatile as well as robust to tackle several types of agricultural images with quite variable lighting, angles and backgrounds. Long benchmark has been InceptionResNetV2-type deep networks for the task of image classification; interesting to observe that such a much-deeper network than EfficientNetV2B2 scored better in performance but only with reduced numbers of parameters without a drop in precision.

SL.NO	MODELS	Accuracy (sunflower)	Accuracy (cauliflower)	Accuracy (ImageNet)
1	VGG_16	90.6	75.1	3.2
2	VGG_19	86.7	76.3	3.1
3	ResNet50	84.3	52.9	70.2
4	ResNet152V2	85	71.2	22.6
5	MobileNet	87.3	76.9	25.2
6	NasNetMobile	87.7	85.4	16.6
7	EfficientNetB0	93.8	88.3	80.9
8	EfficientNetB1	91.6	89.2	85.3
9	EfficientNetV2B1	92.3	81.3	90
10	EfficientNetV2B2	96	94	90.3
11	EfficientNetV2B3	96.3	94.7	91.4
12	ConvNextsmall	78	85	84
13	ConvNextLarge	69.3	89.4	81

Fig. 6. Accuracy of sunflower, cauliflower datasets.

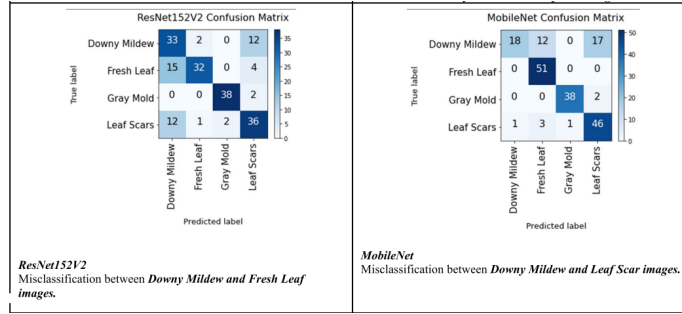


Fig. 7. Confusion matrix report of sunflower(flora) dataset

5 Conclusion

This paper has delivered very deep in analyzing the effectiveness of a set of transfer learning models for the detection task of plant diseases while taken into consideration three datasets, namely Sunflower, Cauliflower, and Agri-ImageNet. As we can see through the results, the InceptionResNetV2 being a benchmark model is still showing a wide accuracy range in identifying plant diseases, though the EfficientNetV2 series is taking over the game. These models are good balance scores for accuracy and computational efficiency, which makes them perfect for real-world applications where the constraints of resources are very usual.

Generalizing, the models did reasonably well, but there were particular problems with the Agri-ImageNet dataset: several of the models performed poorly in comparison with the other data sets, and this is also the case for EfficientNetV2. More work would be needed to explain such differences, perhaps involving factors such as overall image quality and a greater diversity of crops within Agri-ImageNet.

This work further reveals that the choice of transfer learning models must be context-dependent according to the need of individual tasks, such as the available computational resources or the size of the dataset, as well as the target accuracy. The consideration of EfficientNetV2 models, in particular, is more desirable due to the lightweight characteristics of the models, but it has good performance. ConvNeXt models also have encouraging results under proper constraints.

In conclusion, this research gives valid insight into the application of transfer learning models for plant disease detection and promotes more utilization of precision agriculture. Future work will be necessary in exploring diversely inclusive datasets, more sophisticated model tuning, and additional sources of data, be it environmental, to possibly add stronger filtering capability to the disease detection systems. Improving models for mobile-based application development will empower farmers to take remedial action in a timely manner, thereby making farming sustainable.

References

1. Godfray, H. C. J., Beddington, J. R., Crute, I. R., et al. (2010). Food Security: The Challenge of Feeding 9 Billion People. *Science*, 327(5967), 812-818. DOI: 10.1126/science.1185383.
2. Agrios, G. N. (2005). *Plant Pathology* (5th ed.). Elsevier Academic Press. ISBN: 978-0120445653.
3. Strange, R. N., Scott, P. R. (2005). Plant Disease: A Threat to Global Food Security. *Annual Review of Phytopathology*, 43, 83-116. DOI: 10.1146/annurev.phyto.43.113004.133839.
4. Dodds, P. N., Rathjen, J. P. (2010). Plant Immunity: Towards an Integrated View of Plant-Pathogen Interactions. *Nature Reviews Genetics*, 11(8), 539-548. DOI: 10.1038/nrg2812.
5. Sireesha Moturi, S.N.Tirumala Rao, Srikanth Vemuru, Grey wolf assisted dragonfly-based weighted rule generation for predicting heart disease and breast cancer, *Computerized Medical Imaging and Graphics*, Volume 91, 2021, 101936, ISSN 0895-6111 <https://doi.org/10.1016/j.compmedimag.2021.101936>.
6. J. Liu, F. Lv, and P. Di, "Identification of sunflower leaf diseases based on random forest algorithm," in *Proc. Int. Conf. Intell. Comput., Autom. Syst. (ICICAS)*, Dec. 2019, pp. 459-463, doi: 10.1109/icicas48597.2019.00102.
7. L. Tianyu and F. Quan, "Detecting grape leaves based on convolutional neural network," *J. Northwest Univ. Natural Sci. Ed.*, vol. 47, no. 4, pp.505-512, Apr. 2015.
8. J. Xue, L. Huang, B. Mu, K. Wang, Z. Li, H. Sun, H. Zhao, and Z. Li, "Detection of rotten fresh-cut cauliflowers based on machine vision technology and watershed segmentation method," *Amer. J. Biochem. Biotechnol.*, vol. 18, no. 2, pp. 155-167, Feb. 2022.
9. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no.6, pp. 84-90, May 2017, doi: 10.1145/3065386.
10. M. Sireesha, Srikanth Vemuru and S. N. Tirumala Rao, "Coalesce based binary table: an enhanced algorithm for mining frequent patterns", *International Journal of Engineering and Technology*, vol. 7, no. 1.5, pp. 51-55, 2018.
11. Chen, J., Shi, Y., Zhang, Y., Wu, Y. (2020). Applications of Deep Transfer Learning in Agriculture: A Survey. *IEEE Access*, 8, 151393-151413. DOI: 10.1109/ACCESS.2020.3016716.
12. Mohanty, S. P., Hughes, D. P., Salathe, M. (2016). Using Deep Learning for Image-Based Plant Disease Detection. *Frontiers in Plant Science*, 7, 1419. DOI: 10.3389/fpls.2016.01419.
13. M. Sireesha, Srikanth Vemuru and S. N. Tirumala Rao, "Coalesce based binary table: an enhanced algorithm for mining frequent patterns", *International Journal of Engineering and Technology*, vol. 7, no. 1.5, pp. 51-55, 2018.
14. K. P. Asha Rani and S. Gowrishankar, "Pathogen-Based Classification of Plant Diseases: A Deep Transfer Learning Approach for Intelligent Support Systems," *IEEE Access*, vol. 11, pp. 64476-64493, June 2023.
15. Sunayna, S.S., Rao, S.N.T., Sireesha, M. (2022). Performance Evaluation of Machine Learning Algorithms to Predict Breast Cancer. In: Nayak, J., Behera, H., Naik, B., Vimal, S., Pelusi, D. (eds) *Computational Intelligence in Data Mining. Smart Innovation, Systems and Technologies*, vol 281. Springer, Singapore. <https://doi.org/10.1007/978-981-16-9447-9-25>