IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

Soybean Leaf Disease Classification using MobileNetV2 and EfficientNetB0

Ankita Dnyanoba Joldapke[1] and Ms. Snigdha Shukla[2]
Student, Master of Computer Applications[1]
Assistant Professor, Master of Computer Applications[2]
Sadhu Vaswani Institute of Management Studies for Girls, Pune
Savitribai Phule Pune University, Pune, Maharashtra, India.

Abstract: *Soyabean is an important crop used for food, animal feed, and industrial application. Disease symptoms on leaves can reduce crop health and productivity, making timely identification valuable. This paper presents a deep-learning workflow for classifying soybean leaf images into six dataset categories: Bacterial Blight, Ferrugun, Healthy_pic, Powdery Mildew, Sudden Death Syndrome and Yellow Mosaic. Two transfer-learning architectures; MobieNetV2 and EfficientNetB0, were tested on a prepared data set of 957 images. The data were organized into training, validation, and tests. They consisted of 654, 157, and 146 images in the subsets, respectively. MobileNetV2 test accuracy (recorded) was 99.32%. The new EfficientNetB0 evaluation claimed a 100% accuracy; however, this result is dependent on the following reproduction due to different previous assessments. The paper addresses the issue of preprocessing, model design, model analysis, and model implementation, evaluation of metrics, limitations, and a proposed bilingual prediction interface. The findings are supporting evidence and do not set equivalent performance on unstimulated field conditions only does so for the selected test split..*

Key Words: Soybean leaf disease, Convolutional neural network, Transfer Learning, MobileNetV2, EfficientNetB0, Image Classification

1. INTRODUCTION

Agriculture is involved in food production, livelihoods, and economic development in rural communities. Plant diseases can impact crop growth, quality, and yield. Soybean is grown for its protein dense seeds and vegetable oil to monitor the health of the oil is important for the farmer and agriculture stakeholders. Disease symptoms may manifest in changes in color, texture, shape, or surface appearance.

However, the identification of disease is often based on traditional practice and agricultural knowledge. While inspection is useful, it can be time consuming and may need access to trained staff. Symptoms of diseases, nutrient deficiencies, environmental stress, and insect damage may be indistinguishable on the plant. Image capture conditions, including lighting, camera quality, leaf orientation, and background complexity, may also influence interpretation.

Deep learning offers a method to learn from labelled images of visual patterns. Convolutional neural transfer learning can make use of pretrained networks (CNNs) for image classification to a desired dataset of features. This study aimed to compare the performance of MobileNetV2 and EfficientNetB0 in the classification of classifies soybean leaf images into one of six categories. The contribution was supposed to be a reproducible not autonomous agricultural diagnosis, but rather a workflow and a baseline evaluation.

In addition to disease classification, the system is intended to provide some symptom-based preliminary guidance. The deep learning model will produce the top 3 possible soybean leaf disease predictions from an uploaded leaf image. The visible symptoms associated with those predicated categories can then be used as a guide to possible diseases, preventative practices, and management options. This is intended to be a guide to help with initial understanding and is

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

not a guarantee of a confirmed disease diagnosis. If is any doubt as to the actual disease present, then it is advised to seek the advice of an agronomic expert.

Problem statement and Objectives

This problem is classification of soyabean leaf photographs into six categories of data sets under limited data settings. The model should be able to recognize that two classes may be very similar visually but are subtly different in terms of the appearance of the image. Also, has to be transparent reporting of performance and limitations of the classification system, particularly when the test set is limited in size.

The Objectives are:

1. Labelled images need to be split into a Training, Validation, and Test set.
2. Use appropriate image processing techniques for architecture.
3. Carry out transfer learning for training and evaluation of MobileNetV2 and EfficientNetBo.
4. Interpret accuracy, precision, recall, F1 score and loss.
5. Compare model results for the same tests.
6. Provide a vision of an interface for top three predictions with accurate educational material in English and Marathi.

II. LITERATURE REVIEW

The areas of plant disease classification research include efficient CNN design, transfer learning, segmentation, attention mechanisms, and severity assessment. The seven papers included below have been selected to explain the architectural basis and agricultural relevance of this study. The reported metrics are from different experimental protocols and datasets and should not be interpreted to be a direct comparison.

Sandler et al proposed an efficient CNN architecture called MobileNetV2, which is based on inverted residuals between the two types of bottlenecks, blocks, and linear bottlenecks [1]. It has depth-wise separable convolutions to reduce useful feature representations while computational. This is useful for applications that have size and inference speed constraints. The current project is based on the MobileNetV2 trained as a feature extractor, and a six-class head is added. This original work does not guarantee the performance of soybeans.

Tan and Le introduced a new family of CNNs called EfficientNet, which scales up network depth-width and input resolving in a coordinated manner [2]. EfficientNetB0 is one of the smaller variants that can be learned to perform a target image-classification task using transfer learning. This gives the architecture element for the second model assessed in this study. To make a fair comparison, the same data set must be used, the same order of classes must be used, the same test protocol must be used, and the different architectures must have the same processing requirements.

Karlekar and Seal introduced a computer vision method for soyabean leaf disease, called SoyNet classification [3]. They are able to integrate leaf region processing with CNN classification in their pipeline. Detaching a leaf from the background can minimize the amount of noise due to irrelevant parts of the image. The work is relevant in that it is possible that background clutter can cause a problem with the soyban leaf images. But this has been performed on a different set of data and cannot be compared directly to the present test results.

Research into the identification of soybean diseases with original images of the fields and transfer-learning techniques was carried out [4]. Multiple categories of disease and deficiency were included, and the effect of model selection and data preparation was explored. It is important to field images, as they will bring in natural lighting, variable backgrounds, and different orientations of leaves. The study encourages independent field testing, not just images that are similar to the training images.

The authors of this article suggest an enhanced method for soybean leaf disease classification using ConvNeXt, which is proposed by Wu et al [5]. The approach involves both feature extraction and attention mechanisms designed to focus on salient areas of the image and minimize the impact of intricate background textures. Data augmentation is also employed to boost robustness to image variation. This work highlights the importance of feature representation and background variation, which can be investigated in future improvements to current models.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

In this study, Goshika et al. Studied the deep learning approach to classify soybean leaf damage and assess severity of the damage [6]. Severity assessment is different from the classification of the disease categories as it tries to estimate how much damage has been done to the plants that are visible. This is a helpful future avenue for the current project. If a good severity model were required, it would have to be expert labelled, have a strict assessment protocol, and be validated separately. Severity estimation is not stated to be a finished part of the present classifier.

A study on plant leaf disease classification using EfficientNet has been performed on plant disease image using EfficientNet and evaluated the transfer-learning based CNNs [7]. The work facilitates learning visual representation from leaf images using models trained with EfficientNet-family models. But the accuracy predicted by a model on one plant data set does not indicate its accuracy on soybean photographs. The differences in class labelling, image sources, and evaluation methods need to be taken into account.

Research Gap

The reviewed work describes efficient CNN, soybean disease classification, field-image, variation, background processing, and severity assessment. In the present study, MobileNetV2 and EfficientNetB0 are compared on the same data set grouped by six classes for soybeans and is proposed to use a bilingual interface for ranking. Reproducibility needs to be confirmed, and the generalization assessed in independent field images is the key point. Disease specific advice should not be added to the data set until the data set label "Ferrugun" has been validated with its source.

III. MATERIALS AND METHODS

A. Dataset and Split

The prepared dataset comprises 957 images over 6 classes. There are 654 images in the training subset, 157 images in the validation subset, and 146 images in the test subset. To minimize the like hood of creating a near-duplicate image being delivered to subsets, a grouped split was employed. This will decrease one leakage source but not ensure Plant or Field Independent images.

Subset	Image count	Percentage
Training	654	68.34%
Validation	157	16.41%
Testing	146	15.26%
Total	957	100.00%

Table 1. Distribution of the prepared soybean leaf image dataset

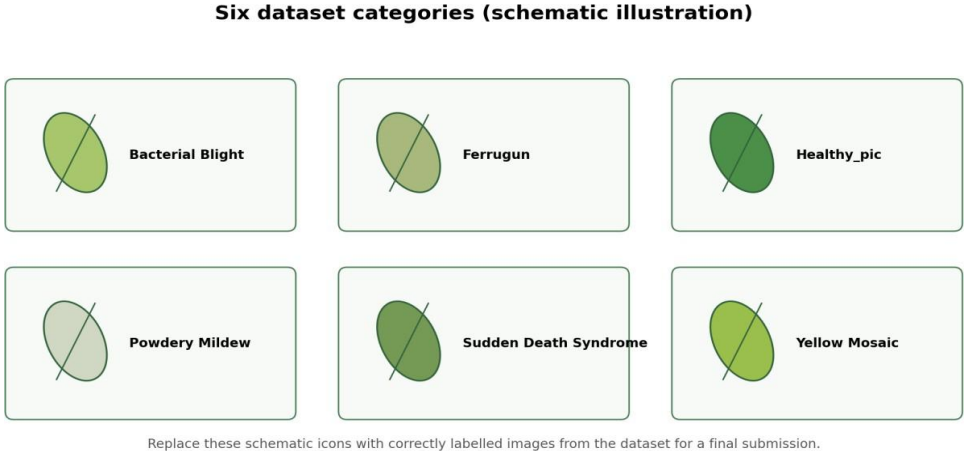


Figure 1. Schematic representation of the six dataset categories

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

B. Image Preprocessing

Images are scaled to 224 × 224 size and converted to RGB. The image values are scaled to the expected input range by the corresponding preprocessing function that is used by MobileNetV2. Preprocessing needs to be used with caution for EfficientNetB0 because rescaling is a part of the standard EfficientNet Keras implementation. In the evaluation, do not randomly augment the test images, and the class index should be consistent with the order in which the images were trained.

C. MobileNetV2 Model

The model MobileNetV2 is a feature extractor trained on a subset of the ImageNet data. The convolutional feature maps are summarized by a global average pooling layer, regularized by a dropout layer and then scored by a dense SoftMax layer yielding six class scores. The recorded configuration features a 224 × 224 RGB input/ dropout of 0.25. The base network was first pretrained and then frozen to keep the pre-trained network representations unchanged during the training of the task specific classification head.

D. EfficientNetB0 Model

A second pretrained convolutional backbone is used: EfficientNetB0. The feature maps are then passed through a global average pooling, dropout, and a 6-unit SoftMax output layer. The recorded configuration is based on 224 × 224 RGB images and dropout of 0.30. Adam, with a learning rate of 0.001 and sparse categorical cross-entropy, was used to train the model. The ultimate training when setting up the configuration for the actual submission should be based on the saved notebook or model configuration.

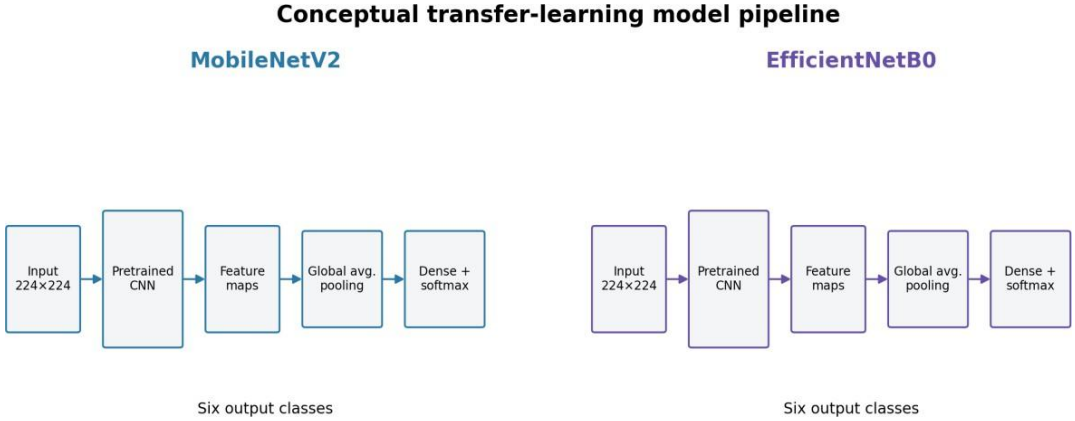


Figure 2. Conceptual transfer-learning pipeline for MobileNetV2 and EfficientNetB0; not a layer-by-layer model graph.

E. System Workflow

The Proposed system takes leaf images, converts the image and resizes it, runs the object detection algorithm, and then outputs the detected object information. Corrects the model used and rates the six output scores. It is possible to confidently show the three categories with the highest scores if there is a user-facing interface. Only disease labels that have been verified and agricultural sources that are reliable may be used to link any symptom or prevention of information. The interface should make it clear that the model output is indicative and not definitive.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

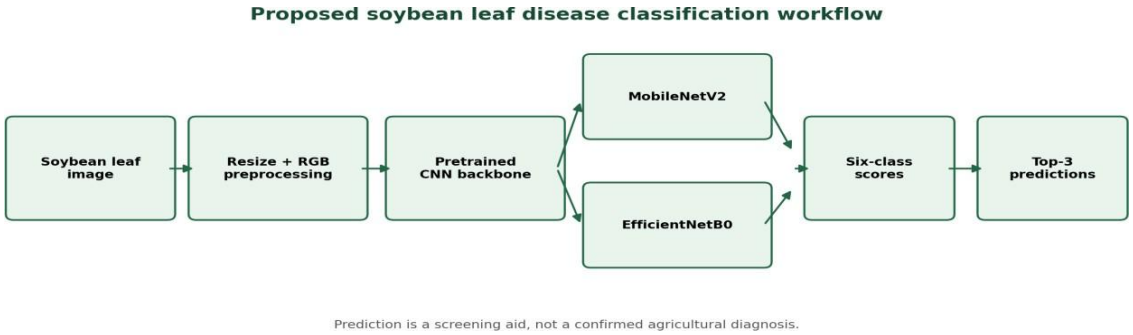
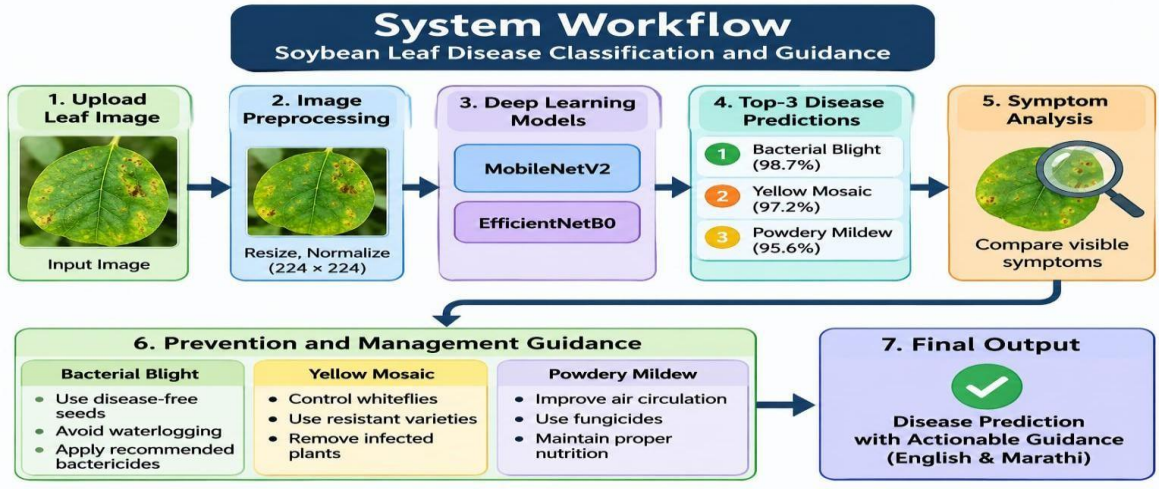


Figure 3. Proposed soybean leaf disease classification workflow

F. Research Methodology

Top 3 Prediction and Symptom- Based Guidance

The System proposed makes use of a two- stage method. In the first stage, the deep-learning model that has been trained looks at the image of the soybean leaf that has been uploaded and produces the top three predicted categories together with the relevant model scores. In the second stage, consideration is given to the visible symptoms linked to these possible disease types in order to offer appropriate advice on prevention and management. This advice is based on a match between the symptoms and reliable agricultural information. When the symptoms available do not clearly match any of the predicated categories, the system must show that it is uncertain instead of giving a definite diagnosis.



G. Evaluation Metrices

The accuracy is the percentage of correctly identified images out of the total number of test images. Precision is the percentage of predictions that are correct, and recall is the percentage of the actual images in the class that are predicted correctly. The F1 score is the harmonic mean of precision and recall. Test loss is the sum of prediction error using the loss function used. It's important to have class-wise metrics because overall accuracy could lead to poor performance in specific classes.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

IV. RESULTS

A. Overall Test Performance

The recorded MobileNetV2 evaluation recorded a test accuracy of 99.32% or 145. It made 146 correct predictions in 146 test images, and the test loss was 0.0372. For the new EfficientNetB0, the accuracy was 100% with a loss of 0.0289. The result is different from the earlier EfficientNetB0 evaluation, so it should be reproduced prior to considering it final.

Metric	MobileNetV2	EfficientB0	Status
Test images	146	146	Same split
Correct	145	146*	Verify*
Incorrect	1	0*	Verify*
Accuracy	99.32%	100.00%*	Verify*
Test loss	0.0372	0.0289*	Verify*

Table 2. Recorded model performance. EfficientNetB0 values marked * require a consistent rerun

B. Class-Wise Observations

High precision and recall on the class-wise classification report were recorded for MobileNetV2. Two images of Powdery Mildew were forecasted as Healthy_pic. This error suggests that the model might have misclassified a disease image as a healthy class when the visible features are not strong enough, or the image is not like the training images. The latest EfficientNetB0 report indicated perfect class-wise scores on the test set. However, these results need to be confirmed against the same saved model, class ordering, and test subset.

C. Interpretation

The results recorded suggest good performance with the test split that was prepared; however, this small test size does not allow for generalization conclusion. The 100% does not imply that all future images will be correctly classified by the model. External images may show new disease stages and soybean varieties, backgrounds, or camera characteristics. Unless calibrated, the confidence score should not be used for interpretations as probabilities.

V. DISCUSSION

This could be applied as transfer learning to transfer features extracted from a limited agricultural dataset to a large one dataset. MobileNetV2 is optimized for computation efficiency while EfficientNetB0 uses compounding to balance network dimensions. Their comparison is limited to cases where the preprocessing and evaluation is correct for each architecture. The preprocessing techniques may vary should be materially changed from predictions and should be documented.

Prior to use of the model, test it on a field dataset that is collected separately; agricultural practice monitoring was discussed. Confusion matrices class-wise should also be incorporated in the evaluation, matrices, and repeated splits (when possible). If the app provides information about the diseases in English, the mapping and translated educational materials should be checked by a competent and Marathi agricultural source.

Limitations

A set of 957 images, and a test set of 146 images. Classes are not evenly distributed, and some classes do not have many test examples. The nature of the grouped split results in a lower risk of near duplicate leakage, but is not necessarily independent from plant to plant, farm to farm, or collection session to collection session. Furthermore, the Ferrugun label has been left unchanged as per dataset and needs to be verified from the source.

The present system carries only image classification. It cannot confirm the diagnosis of disease, provide estimates of area and severity of disease, or validate recommendations for chemical treatment. The EfficientNetB0 performance is also not consistent across evaluations and should be reproduced prior to publication.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

Future Work

A larger and more diverse set of images and data should be gathered in the future to include various soybean varieties, seasons, disease stages, and image- capture conditions, as well as different farms. It is recommended to prioritize evaluation of an independent field dataset. A documented validation protocol could be used to consider fine tuning, data augmentation, class balancing, and class calibration.

Explainability methods like Grad-CAM might be tested to visualize areas of the image which affect a prediction. Severity assessment can be developed based on the severity categories labelled by experts or lesion segmentation. The three predictions might be displayed in a bilingual interface, along with validated symptoms descriptions and recommended preventive measures. These components should be separately implemented and tested before using them for agriculture decision making.

Future work will be directed towards implementing and verifying a guidance module which links the model's top- three predictions with visible symptoms of the disease and with information on preventive measures and management. For this module, it will be necessary to use verified agricultural references and have the recommendations assessed by agricultural experts in order to evaluate their accuracy and usefulness.

VI. CONCLUSION

The authors developed a six-class soybean leaf image-classification workflow based on MobileNetV2 and EfficientNetB0 transfer learning. The prepared dataset includes 957 images split into training, validation, and test sets. The test accuracy of MobileNetV2 was recorded as 99.32%. According to the most recent announcement in EfficientNetB0 report, it was able to achieve 100% accuracy, with numbers being reflected, but it can be reproduced as a previous evaluation was different. The findings are for the selected data set and do not reflect optimal field performance. Independent testing verified class labels and careful communication with users are required steps for actual deployment.

ACKNOWLEDGMENT

The authors thank all the authors of the dataset, all the open-source machine- learning libraries and pretrained neural-network architectures employed in this study. Update and change this as needed to credit the provider, supervisor, and institution of the actual data set.

REFERENCES

[1] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4510-4520, 2018. doi:10.1109/CVPR.2018.00474.

[2] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," Proceedings of Machine Learning Research, vol. 97, pp.6105-6114,2019.

[3] A. Karlekar and A. Seal, "SoyNet: Soybean leaf diseases classification," Computers and Electronics in Agriculture, vol.172, article 105342,2020. doi: 10.1016/j.compag.2020.105342.

[4] "Soybean disease identification using original field images and transfer learning with convolutional neural networks," Computers and Electronics in Agriculture, 2022. Verify full author and publication details from the publisher's record before submission.

[5] Q. Wu et l., "A classification method for soybean leaf diseases based on an improved ConvNeXt model," Scientific Reports, vol.13, articles 19141, 2023.doi:10.1038/s41598-023-46492-3.

[6] S. Goshika, K. Meksem, K.R. Ahmed, and N. Lakhssassi, "Deep Learning Model for Classifying and Evaluating Soybean Leaf Disease Damage," International Journal of Molecular Sciences, vol. 25, no.1, article 106, 2024. doi:10.3390/ijms25010106.

[7] "Plant leaf disease classification using EfficientNet deep learning model," Ecological Informatics, vol. 61, article 101182, 2021. Verify the complete author list from the publisher to record before submission.