Heliyon 9 (2023) e18261

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Enhancing sugarcane disease classification with ensemble deep learning: A comparative study with transfer learning techniques

Swapnil Dadabhau Daphal [a,*], Sanjay M. Koli [b]

[a] Department of E&TC Engineering, G. H. Raisoni College of Engineering & Management, Wagholi, Pune, 412207, Maharashtra, India
[b] Department of E&TC Engineering, Ajeenkya DY Patil School of Engineering, Charholi Bk., Pune, 412105, Maharashtra, India

ARTICLE INFO

Keywords:
Agriculture
Deep learning
Disease classification
Sugarcane database

ABSTRACT

Deep learning practices in the agriculture sector can address many challenges faced by the farmers such as disease detection, yield estimation, soil profile estimation, etc. In this paper, disease classification for the sugarcane plant and the experimentation involved thereby is thoroughly discussed. Experimental results include the performances of the well-known existing transfer learning techniques and proposed ensemble deep learning based architecture that incorporates stack ensemble of two networks with one having level-wise spatial attention helping to provide better generalization. A Self-created database of sugarcane leaf diseases is introduced to the research community through this paper. It involves 5 categories with a total of 2569 images. Here, it is observed that best performing transfer learning method, MobileNet-V2 shows an accuracy of around 84% with the lowest number of parameters whereas ensemble model reaching to 86.53% with less epochs and with acceptable number of parameters.

1. Introduction

Most agrarian nation employs a sizable portion of their labor force in agricultural activities. Therefore, a decrease in agricultural productivity has a significant influence on everyday life [26]. Bad weather, inadequate irrigation systems, a poor choice of crops, plant diseases, and a lack of state-of-the-art farming facilities are the main factors that contribute to production loss [16]. Environmental and biological elements that affect agricultural yield are more difficult to manage than technical ones [9]. Modern technological developments have benefited humanity in practically every field. In the healthcare sector, the use of computers to assist in diagnosis has been very important in medical examinations [15]. Medical professionals can make simple diagnoses with the use of transfer learning applied to chest X-ray images [36].

Similarly, the use of technology in agriculture is expanding to unprecedented levels. For instance, an intelligent context-aware irrigation system has been put in place to enable farmers to grow intensively with comparatively less water in order to address the issue of water scarcity. This system's primary components for efficient operation are a data acquisition system, decision support system, and scheduling system [11]. Precision agricultural machinery can perform more effectively, produce better results, have lower production costs, and be simpler for the user if a global positioning system (GPS) and global information system (GIS) are used. The system has demonstrated managerial effectiveness and information management assistance [37]. In [19], the hyperspectral imaging technique was employed to determine the geographic origin of the agricultural product. Based on image processing, there

* Corresponding author.
E-mail address: daphalsd01@gmail.com (S.D. Daphal).

https://doi.org/10.1016/j.heliyon.2023.e18261
Received 26 November 2022; Received in revised form 12 July 2023; Accepted 12 July 2023

Available online 18 July 2023

Table 1
Comparison of deep learning studies taken on various plants.

Reference	Plant	Technology used	Dataset	Data augmentation	application	Accuracy
[20]	Tea	IOT + ML	Self-collected dataset of Environmental conditions monitoring like temperature, humidity, rainfall etc.			
	Yes	blister disease detection	91%			
[38]	Peach, tomato	DL	plant village dataset	Yes	Disease classification	95.48%
[2]	Corn	DL	plant village dataset	Yes	Disease classification	98.56%
[8]	Bean	DL	ibean (AI-Lab-Makerere)	Yes	Disease classification	92.97%
[12]	Corn, tomato, rice, casssva	DL	plant village dataset, rice disease dataset, cassava dataset	Yes	Disease classification	99.39% 99.66% 76.59%
[10]	Rice	DL	Self-collected dataset	Yes	Disease classification	92.46%
[17]	Wheat	DL	CIAGR dataset	Yes	Disease classification	86.50%
[3]	Multi-plant	DL	plant village dataset	No	Disease classification	99%
[1]	pepper, tomato	DL	plant village dataset, Pepper dataset (NIHHS)	Yes	Disease classification	99.69% 99%
[25]	Mango	DL	Self-collected dataset	No	Disease classification	89.41%
[23]	Rice	DL	Self-collected dataset	No	Disease classification	95.31%
[31]	Multi-plant	IOT + DL	plant village database, hyperspectral images	No	Disease classification	95%
[18]	Multi-plant	DL	plant village dataset	Yes	Disease classification	95%

has been an increase in methods used to address agricultural problems. In the agricultural workflow, harvesting is an important stage. It has been difficult for a while to choose ripened fruit that is of excellent quality. The harvesting process can be greatly affected by a lack of skilled labor. In [24], for categorizing different grape types, various semantic segmentation algorithms were compared. Further, these algorithms were used to prepare systems for harvesting purposes.

In several Indian states, sugarcane is an important cash crop. Numerous investigations on the sugarcane crop have been performed in an effort to boost productivity. For calculating the prices of sugarcane derivatives, neural networks were used in view of the sugarcane crop's economic significance and impact on the economy [32]. UAV-acquired images and data from the ground can greatly aid in the estimation of the yield with a high degree of precision. It provides evidence in favour of the claim that increased technological advancements will result in higher sugarcane yields [34]. The integration of IoT and deep learning could be employed effectively to provide precise solution to increase the crop yield and improve product quality [40]. Furthermore, to improve agricultural productivity, it is necessary to manage problems like diseases.

For effective disease control, deep learning techniques are required. Table 1, gives the summary of deep learning practices used for various plants. It highlights the underlying technology in research, dataset used and its accuracy for specific application. The aforementioned techniques employ easily accessible datasets. The availability of the datasets, however, is the main challenge for researchers working on deep learning. Transfer learning techniques where pre-trained models are systematically trained with small databases to overcome overfitting problems. A customized database can be difficult and time-consuming to create. For conducting deep learning experiments, the Plant Village dataset is a great tool. It does not, however, cover every plant species found in every region of the world. Additionally, it only contains images that were taken in a controlled environment. It is very likely that the designed algorithm will function as anticipated in a lab setting but not in real-time [14]. Therefore, addressing the need for a sugarcane disease detection system is the key driver for creating a customized dataset. On the self-created dataset, the well-known transfer learning algorithms are used and compared with proposed ensemble deep learning model. Also, to find the optimum choice sugarcane disease diagnosis system, the trade-off between a number of hyperparameters and accuracy is carefully examined.

The major contributions of the paper are as follows:

1. For the purpose of identifying sugarcane leaf disease, a self-generated sugarcane dataset is proposed. In contrast to laboratory conditions, this collection includes real-time field images of sugarcane leaves.
2. A detailed analysis of several transfer learning methods using the aforementioned dataset.
3. Suggested a transfer learning method that will be the most suitable for a real-time system for diagnosing sugarcane disease.
4. Proposed a ensemble model that trains faster and produces the better classification results than major transfer learning methods, on our sugarcane dataset.

The paper is organized as follows: Section 2, provides a thorough overview of the proposed dataset. The experimentation is discussed in Section 3. Results are discussed in section 4, and conclusions are presented in section 5.

2. Datasets

Dataset are key requirement for successful deployment of deep learning architectures. The task of problem specific data collection is challenging and requires significant efforts. The following subsection focuses on available dataset and highlights the methodology followed in the sugarcane database creation.

2.1. Existing datasets

Plant disease diagnosis and classification is a challenging task. Using deep learning techniques for the classification needs large databases. Plant Village database is commonly used for testing the suitability of any devised deep learning methodology. This database contains over 20000 images for approximately 14 plants. This dataset generally concentrates on the foliar characteristics of major fruit plants including apples, grapes, strawberries, etc. Many researchers have adopted this dataset in the study due to a lack of modern facilities or due to other constraints. However, to address the challenges related to certain plant-specific images associated with it is a key requirement. Since the plant village database does not cover the sugarcane plant, it is not useful for the proposed work. However, the range of plants and different variants of images, it has covered in samples is outstanding. This has inspired the author to create own database for sugarcane plant which best fit for the requirement of farmer community [27]. Researchers have collected 2000 images of wheat leaves for the identification of diseased plant leaves. The specific crop season was chosen in Indian environmental conditions for capturing the images and study of the plant [22]. In another study, for identification of the maydis leaf blight researchers have collected 1547 images of the Maize plant. All images were collected on the field from different areas in West Bengal, India, and New Delhi. Smartphones were one of the major means for collecting images of plant leaves [29]. Few researchers have collected a total of 5235 images for the identification of T. absoluta infected plant leaves of tomato species. The labelling of the dataset was done by experts from the related domain. U-net model was employed for the disease segmentation in the images captured [21]. In an attempt at automatic recognition of strawberry diseases researchers have collected the data using smart phones in cultivated lands. In this work, image annotation was manually done with farmers having rich experience in strawberry cultivation. To minimize the anticipated overfitting issue data augmentation was employed. Total 6608 images were available after the data augmentation. Images were resized to $(128 \times 128 \times 3)$. Total 9 classes were considered for the experimentation including rot, thripes, mildew, etc [7]. An extensive and systematic survey has been carried out regarding application of deep neural network for the detection of plant diseases in [39].

Indian marginal farmers are yet to get the technical expertise and are not well versed with recent digital trends. So, in this database collection i.e. during the image acquisition, keeping the farmers in mind no special attention was provided for the illumination conditions, orientations, and different physical factors.

2.2. Highlights of the dataset

In this section, key highlights of the proposed sugarcane database are given. The suitability of the database for real-time implementation is discussed in this section.

2.2.1. Number of classes

The sugarcane database mainly contains five categories. It includes healthy, mosaic, red rot, rust, and yellow leaf diseases as shown in Fig. 1a, 1b, 1c, 1d, and 1e, respectively. For each disease type, approximately 500 images are taken. Table 2, gives the number of images corresponding to each class.

2.2.2. Number of image capturing devices

General smartphone cameras were used for capturing the images. The ultimate aim of the work is to build a real-time disease detection system. In order to make the system make more universal and reliable to lower-end farmers, mobile phones in the range of INR 8K-13K were chosen.

For the following specifications, the trade-off between cost and imaging resolution is observed satisfactory to the requirement of authors. Detailed specifications of the devices used for the database collection is given in Table 3.

Table 2
Images per class in sugarcane database.

Category name	Number of Images	in percentage
Healthy	520	20.24%
Rust	514	20.07%
Red rot	519	20.20%
Yellow	505	19.65%
Mosaic	511	19.89%
Total	2569	100%

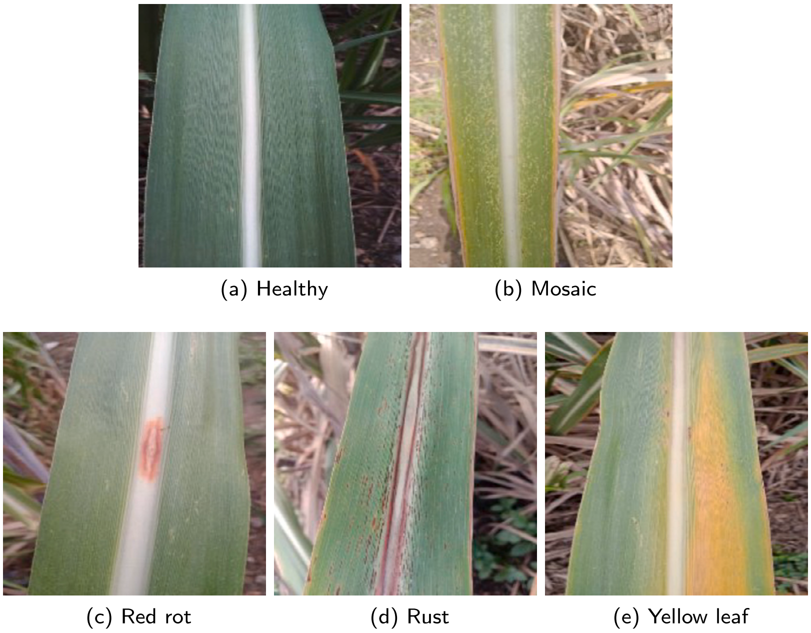


Fig. 1. Different conditions of the sugarcane leaves.

Table 3
Specifications of the smartphone used for image acquisition.

Device	Camera Resolution	RAM	ROM	Processor cores & Frequency
Vivo Y12	13MP + 2MP + 8MP	3GB	64GB	Octa-core 2 GHz
Vivo	13 MP	2GB	32GB	Octa-Core 2.3 GHz
POCO M2 pro	48MP + 8MP + 5MP + 2MP	6GB	64GB	Octa-Core 2.3 GHz

Table 4
Detailed description of the image collection site.

Parameters	Details
Region	Nagargaon, Shirur, MH, India
Climate(Köppen and Geigers)	BSh
Average annual temperature	25.4 degree Celsius
Average rainfall	498 mm
Soil type	Black soil
Annual temperature variation	9.7 degree Celsius
Irrigation facility to site	Surface irrigation
Largest nearby water body	Bhima River-500-1000 m from site

2.2.3. Image collection site

The images were collected in normal environmental conditions during the month of August to September. Samples were taken from different farms from Nagargaon, Pune, India. The region selected is under the local steppe climate category. It is classified as BSh as per Köppen and Geigers scheme of climate classification.

The key features of the climate of the region are mentioned in Table 4.

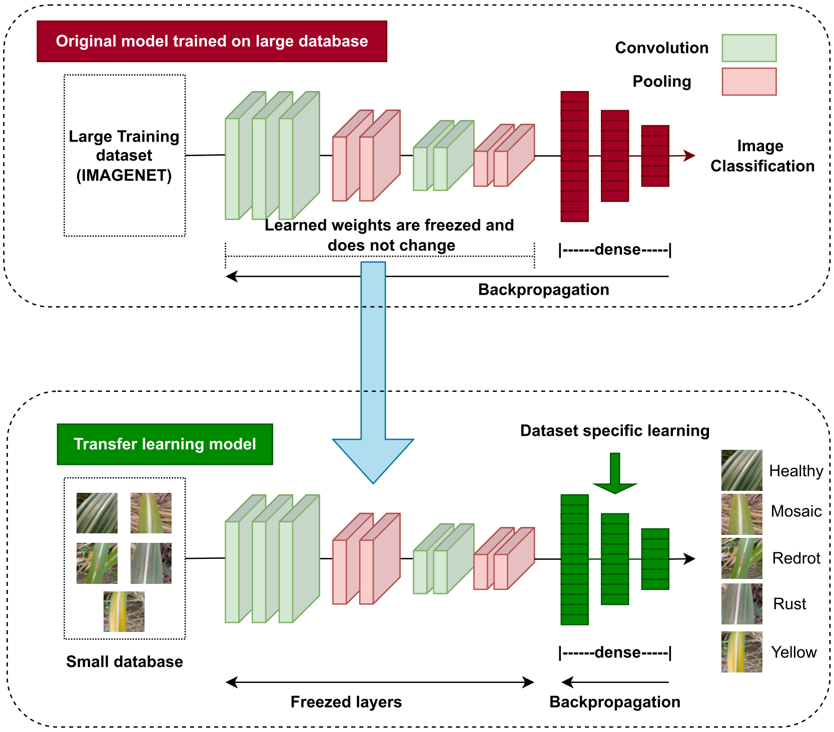


Fig. 2. Transfer learning block diagram.

2.2.4. Annotations

After the collection of dataset images were annotated with the help of agricultural officers and farmers having rich experiences in the cultivation of sugarcane. After the detailed scrutiny images were separated into categories mentioned in Table 2.

2.2.5. Content diversity

Amateur farmers were asked to take the images along with due observations from skilled technical personalities. The reason behind doing so was to anticipate the handling of the system by the farmers only that lack required technical expertise sometimes. The image capturing time was restricted between 9.00 am to 5.00 pm on a normal days (Without haze and dark clouds). Database can be found in [4]

3. Methodology

In this section, various transfer learning approaches are applied to the database collected. The most straightforward way to implement the deep neural network is to use the transfer learning (TL) approach. Work flow of TL is shown in Fig. 2. In classification task, base model is trained with large number of images and possibly more number of classes. The knowledge base (KB) in terms of probable weights and internal graph structure is generated in the subsequent phase. KB can be used effectively to train the model with comparatively less number of images. It is helpful in situations where collecting the dataset is most complicated task. In addition to that the architecture of mobilenetv2 is shown in Fig. 3 MobilenetV2 has characteristics like smaller model size, increased efficiency and faster inference time.

The model is already trained with large number of images of size 60K-90K in numbers. This model then can be used as initial point for task which faces challenge in terms of availability of the sufficient data. VGG-19 [33], ResNet-50 [13], XceptionNet [28], MobileNet-V2 [30] and EfficientNet-B7 [35] are experimented with the database. Performance comparison of all methods is carried out and best suitable method for the sugarcane disease classification is suggested here. Following performance parameters were used for the comparison.

- **Precision:**
  It is the ratio of retrieved and relevant results to the all retrieved results. It is the description of ratio between true positive to the total positive results obtained during the experimentation, as shown in equation (1).

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

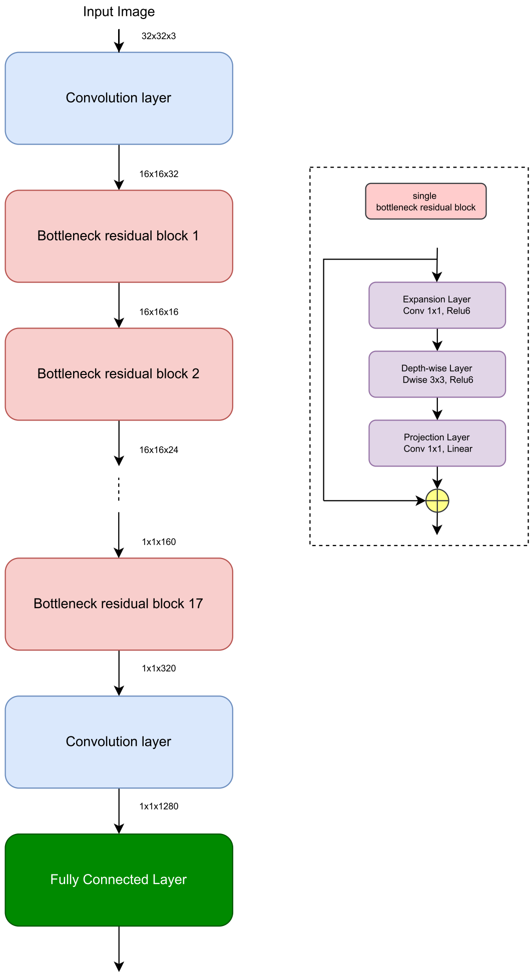


Fig. 3. MobilenetV2 architecture.

- **Recall:**
  Recall sometimes referred as sensitivity is the ratio of retrieved and relevant result to the relevant results. In equation (2), TP is correct detection, FP, is the total number of false detections and FN is number of missed quantities.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

- **F1-Score:**
  The F1 score is also referred as F measure. It signifies the harmony between precision and recall and helps the researcher to trade off between them. Mathematically, it is given by equation (3).

$$\begin{aligned} F1 &= \frac{2 * Precision * Recall}{Precision + Recall} \\ &= \frac{2 * TP}{2 * TP + FP + FN} \end{aligned} \tag{3}$$

- **Accuracy:**
  It is ratio of number of correct identification to the all identifications in the experimentation and is given by equation (4).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

- **Support:**
  It is defined as number of instances of true outcomes that are found in each group of target values.

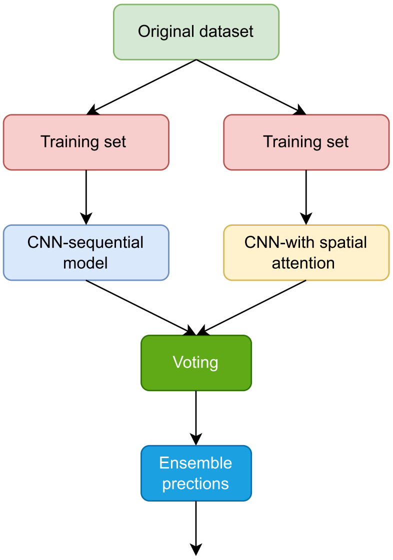


Fig. 4. Mechanism used for model ensemble.

3.1. Proposed model

A machine learning ensemble technique called stacking combines several models to enhance prediction accuracy. Instead of training individual models and combining their predictions based on a simple average or majority vote, stacking uses a ensemble model that learns how to combine the predictions of different models to produce a final prediction. The fundamental concept underlying stacking is to train a variety of base models on training data, then utilise the predictions from those base models as features to train a ensemble model. This ensemble model finds the most effective way to combine the predictions of the basic models to get a final prediction. Any model, including linear models, decision trees, neural networks, or any other machine learning approach, can use stacking.

The stacking method is often carried out in a number of steps:

1. A number of folds (often five or ten) are created from the training data.
2. Predictions are made using the validation set after each base model has been trained using the training data from each fold.
3. The input features to the meta-model are the predictions of each base model on the validation set.
4. The validation set predictions are used to train the ensemble model on how to most effectively integrate them to generate a final prediction.
5. After the ensemble model has been trained, predictions can be made using the test data.

In many machine learning applications, stacking has been proved to be a potent strategy for enhancing prediction performance. Stacking can help to decrease overfitting, enhance generalisation, and increase prediction accuracy by combining the predictions of many models. Here, as part of the stacked ensemble model depicted in Fig. 4, a simple sequential CNN model and another deep CNN model with spatial attention [41] at various levels have been used. Figs. 5 and 6 show the detailed architectures for the simple sequential model and the deep CNN model with spatial attention, respectively. In order to predict leaf diseases that have a lot of morphological similarities, it is imperative to extract stage-specific spatial features to stop the loss of vital spatial information.

4. Results and discussions

We present performance metrics and loss curves for the analysis to show the performance of all models. Table 5 provides the parameter set selected for experimentation. Results from the experimentation of several transfer learning models, which are essential members of the deep learning family, are addressed in this section.

4.1. Performance parameters

Precision, recall, accuracy, F1 score, and support are key metrics to assess any deep neural methodology's effectiveness. Tables 7, 8, 9, 10 and 11 provide a thorough explanation of all parameter values for the various networks VGG19, ResNet50, XceptionNet, MobileNetV2, and efficientNet-B7, respectively. The VGG19 network has a unique architecture with 19 layers of deep convolutional layers. VGG19 produces poor results when compared to other approaches, nevertheless. For any architecture to stand well, it is

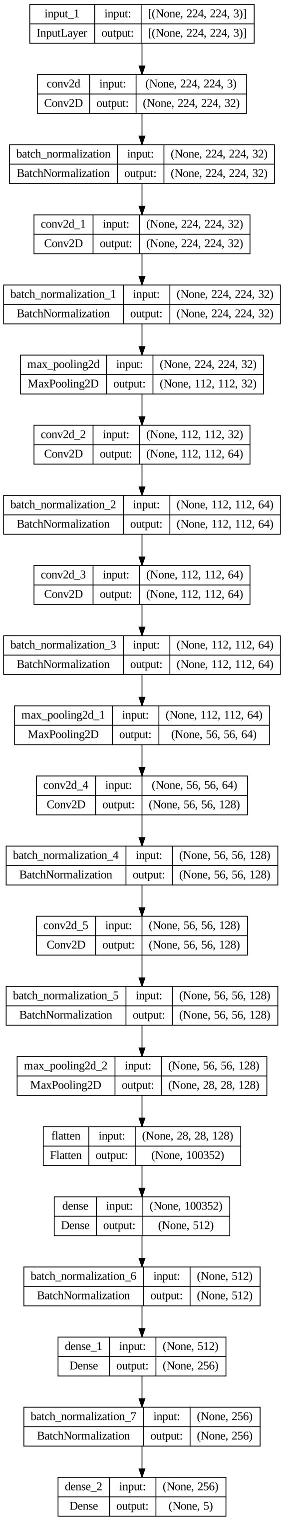


Fig. 5. A simple sequential CNN model.

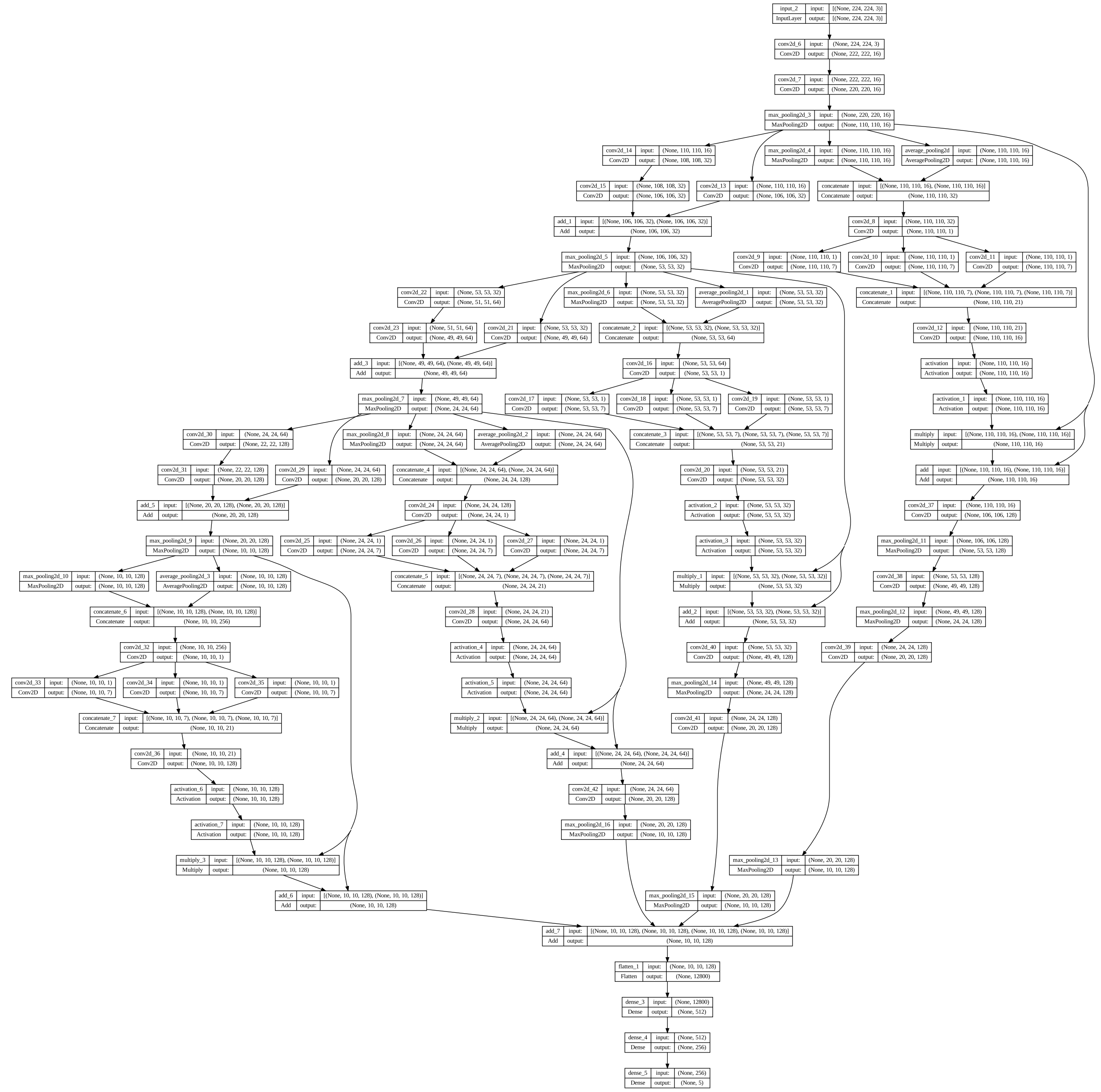


Fig. 6. CNN with spatial attention at different stages of the network.

Table 5
Parameter setting for all models.

Parameter	Details
Batch size	32
Image size	128x128 and 224x224 (both for ensemble)
train:test:val	80:10:10
Classes	5
Epochs	50 and 20 (for ensemble)

Table 6
Impact of change in input size.

Image size	Accuracy
128x128	84.73%
224x224	86.53%

expected to have a higher classification accuracy. By using customised residual networks in the sequential structure, ResNet architectures have proven to be reliable. Even if it performs better than VGG19, an empirical investigation revealed greater room for improvement. XceptionNet contains 71 layers and is deeper than the other two models. Larger parameters are required for improved performance, which adds to the system's computational load. With regard to XceptionNet, the complexity-performance trade-off is not encouraging. The number of parameters employed in the network architecture is shown in Table 12 along with the classification accuracy. The mobilenetV2 model has the best accuracy, at 83.24%. Over the years, there has been a demand for well-known design that is deployment-friendly. The key features of this approach include fewer parameters and better results. The efficientNet class of models is another one. This comparison demonstrates unequivocally that MobileNetV2 delivers superior outcomes to other approaches. MobilenetV2 has demonstrated better accuracy for sugarcane disease classification with practically the same amount of factors. The architecture of the mobilenetV2 is shown in Fig. 2. Depth-wise separable convolutions and linear bottleneck layers used in its architecture, which result in light-filtering, are the main causes of this performance gain. Other important advantages of mobilenetV2 are increased efficiency, smaller size, faster inference time, and improved accuracy which also makes it a better option for real-time implementation. The performance of the model is tested for different parameters and the best possible selected are used for the subsequent experimentation [6] [5]. Along with these findings, the major limitations with the mobilenetV2 are mentioned as follows.

- **Lower accuracy:** For image identification tasks, Mobilenet v2 is not as accurate as previous models.
- **Moderate scalability** It is unsuitable for more difficult tasks since it does not scale well with growing model complexity.
- **Lack of prominent features** It lacks some characteristics found in other models, such as batch normalisation and depthwise separable convolution.
- **Moderate performance with high resolution images:** High-resolution images or pictures with complicated patterns don't work well with it.

The ideal input size for a deep learning model depends on a number of variables, including the task's complexity, the model's architecture, the available computing power, and the size of the dataset. Generally speaking, a higher input size can capture more minute details in the data, but it also uses up more computer resources and can slow down the training process. The training process can be sped up and computational resources saved with a lower input size, but it might not be able to capture all of the important features in the data. The effect of accuracy change on the suggested ensemble model is shown in Table 6. A deep-learning model's performance can be significantly impacted by the quantity of inputs. The model may perform better if there is more information available to it as a result of adding more inputs. The risk of overfitting, in which the model becomes overly specialized to the training data and performs badly on new, unforeseen data, can be increased by adding too many inputs. As a result, it's essential to balance the model's performance and the quantity of inputs. In practice, this frequently means carefully choosing the most relevant characteristics or employing strategies like feature selection or dimensionality reduction to decrease the number of inputs while maintaining the maximum amount of relevant information. Additionally, regularisation methods like dropout can help to avoid overfitting and enhance the model's generalisation capabilities.

Another important factor is coefficient optimization during training of the network. Standard weight initiallization supported in tensorflow & keras is used in the experimentation for e.g. *glorot_uniform*. There are other methods of optimizing the values of coefficients (also known as weights) in training a deep learning model. They are pretrained models, hyperparameter tuning, gradient-based optimization, random initialization etc. Random initialization is opted in this work.

4.2. Visualized analysis

All of the TL models mentioned above have been trained and tested to compare with one another. VGG19 exhibits steady behaviour in Fig. 7a, but it only achieves an accuracy of about 70%. Similarly Fig. 7b, displays the loss curves for ResNet-50 and emphasises its limitations in terms of classification performance. Even though XceptionNet has more parameters, its accuracy has not increased. Fig. 7c displays the loss curves for it. The MobileNet-V2 loss curves are shown in Fig. 7d, where it performs better than all other curves. The model does, however, have a slight tend to overfit as the number of epochs increases. In accordance with epochs, the EfficientNet-B7 loss curves shown in Fig. 7e illustrate how the model's accuracy and loss increase over time. It also fails to surpass the mobilenetV2 in performance. It is also worth noting that if the training and validation curves are far apart from each other model is overfitting or both curves are not improving at all model is underfitting. In this case, the model does not seem to be overfitting or underfitting as well.

The behaviour of all models is compared in the Figs. 8 and 9 in terms of training vs. validation accuracy and training vs. validation loss, respectively. The upper panel of Fig. 8 shows the training accuracy while its lower panel shows the validation accuracy of all

Table 7
Performance metrics for the VGG19.

Class	Precision	Recall	F1-score	Support
0	0.78	0.58	0.67	12
1	0.56	0.71	0.63	7
2	0.00	0.00	0.00	0
3	0.60	1.00	0.75	3
4	1.00	0.60	0.75	10

Table 8
Performance metrics for the ResNet-50.

Class	Precision	Recall	F1-score	Support
0	1.00	1.00	1.00	10
1	0.86	0.86	0.86	7
2	0.80	0.57	0.67	7
3	0.80	1.00	0.89	4
4	0.60	0.75	0.67	4

Table 9
Performance metrics for the XceptionNet.

Class	Precision	Recall	F1-score	Support
0	0.83	0.83	0.83	6
1	0.88	0.78	0.78	10
2	1.00	0.33	0.50	3
3	0.89	1.00	0.94	8
4	0.62	1.00	0.77	5

Table 10
Performance metrics for the MobileNetV2.

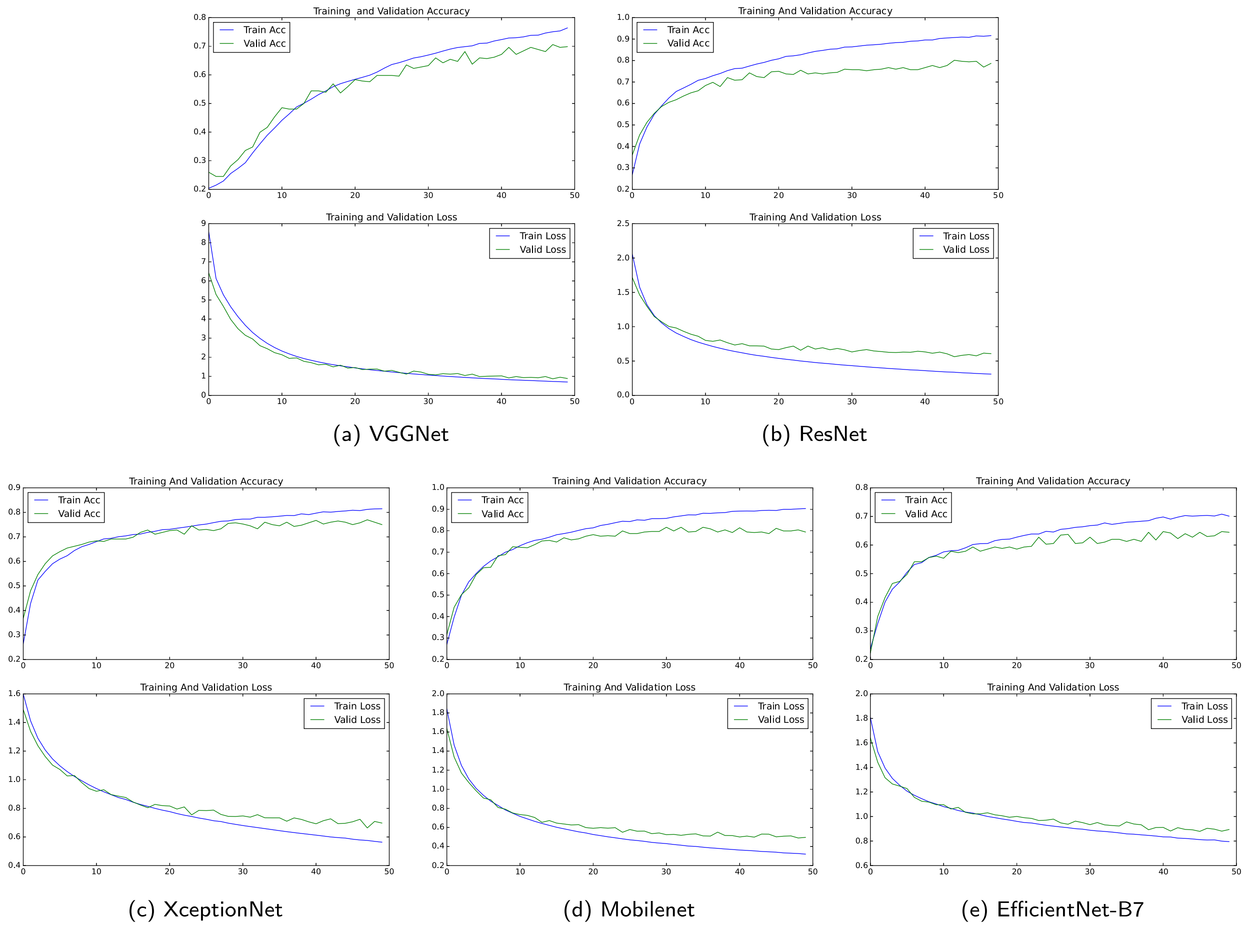
Class	Precision	Recall	F1-score	Support
0	0.88	1.00	0.93	7
1	1.00	0.71	0.83	7
2	1.00	0.5	0.67	8
3	0.62	0.83	0.71	6
4	0.57	1.00	0.73	4

Table 11
Performance metrics for the EfficientNet B7.

Class	Precision	Recall	F1-score	Support
0	0.29	0.50	0.36	4
1	0.83	0.71	0.77	7
2	0.88	0.78	0.82	9
3	0.67	1.00	0.80	2
4	0.88	0.70	0.78	10

Table 12
Performance comparison of different models.

Model	Accuracy	Number of Parameters
VGG19	0.7083	20,026,949
ResNet50	0.8064	23,575,045
XceptionNet	0.7917	20,871,725
MobileNet_V2	0.8324	2,264,389
EfficientNet_B7	0.7272	2,264,389
Proposed model	**0.8653**	8,883,825



(a) VGGNet (b) ResNet (c) XceptionNet (d) Mobilenet (e) EfficientNet-B7

Fig. 7. Loss and accuracy curves for all transfer learning model proposed database.

transfer learning methods that are studied. Similarly, the upper panel of Fig. 9 shows the training loss while its lower panel shows the validation loss of all transfer learning methods.

When the test set's inputs are applied to the MobileNetV2 network that has been trained, the actual prediction results are shown in Fig. 10. Because MobilenetV2 is more accurate, only that model's results are used for predictions. In the aforementioned figure, the MobilenetV2 model's predicted class and actual class labels are displayed. This behaviour was completely anticipated prior to the examination, in contrast to the laboratory setting where these unfavourable factors are manually removed. The whole purpose of doing this is to make the system farmer-friendly and capable of producing results in environments that are close to real-time.

4.3. Ensemble model results

The results for model ensembling are included in this section. A sequential CNN model and deep CNN model with sequential module were stack ensembled and trained with same database for 20 epochs. Comparison of those two models separately is also studied, accuracy and loss curves of both models working separately are given in Fig. 11 and 12, respectively. Fig. 11a and Fig. 11b gives idea about the accuracy of sequential CNN and deep CNN, respectively, whereas Fig. 12a and Fig. 12b are the loss curves of the sequential CNN and deep CNN, respectively. In addition, in order to evaluate the fitness of the model the confusion matrix of those models were also plotted separately. Those performances were again compared with the results of ensemble model that shows the better classification than the former models. The accuracy of 86.53% has been observed with ensemble model that is higher than the other models taken under study. The major advantages that stacked ensembled model offers are as follows,

1. Higher accuracy: The prediction accuracy of stacked ensemble models is frequently higher than that of conventional deep learning models. This is so that complicated patterns in the data can be more effectively captured by stacking, which combines the strengths of various models.
2. Enhanced generalization: More resistant to overfitting than conventional deep learning models are stacked ensemble models. This is due to stacking's use of numerous models with various strengths and weaknesses, which lowers the possibility that one model may become overfit to the training set.

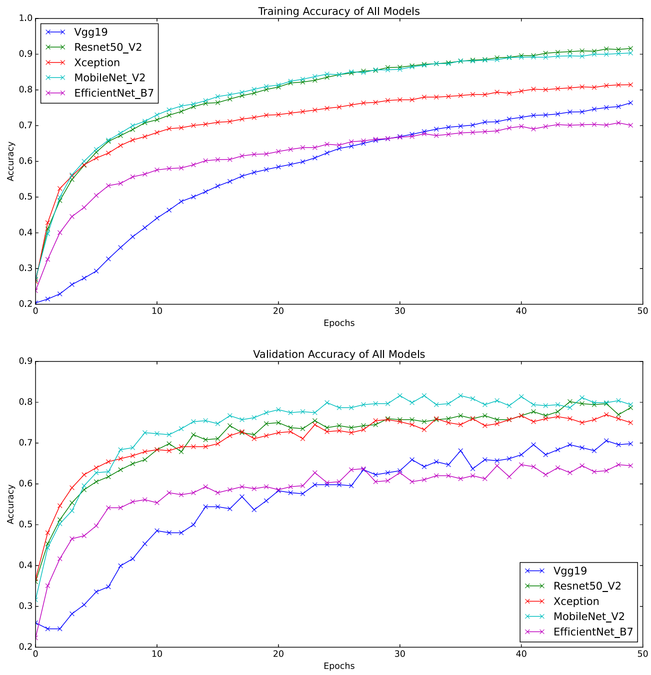


Fig. 8. Accuracy comparison of transfer learning methods: The upper panel shows the training accuracy while the lower panel shows validation accuracy.

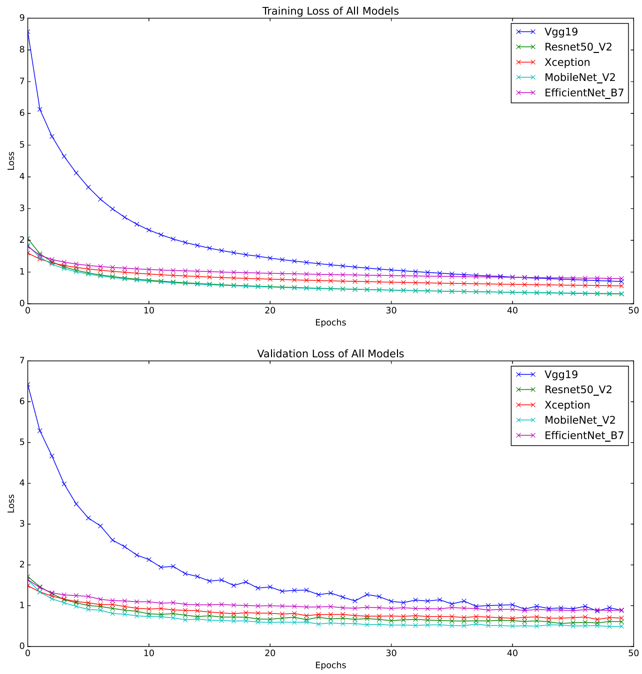


Fig. 9. Loss comparison of transfer learning methods: The upper panel shows the training loss while the lower panel shows validation loss.

3. Robustness: Traditional deep learning models are less resistant to changes in the input data than stacked ensemble models. Stacking makes it more probable that at least one of the models will be able to handle the new data because it employs numerous models with various architectures and hyperparameters.
4. Interpretability: Compared to conventional deep learning models, stacked ensemble models may be easier to understand. The reason for this is that stacking combines the results of various models, making it possible to find significant characteristics and patterns in the data. This may help to shed light on the fundamental mechanisms underpinning the forecasts.

Fig. 10. Predictions with MobileNet V2 for all classes.

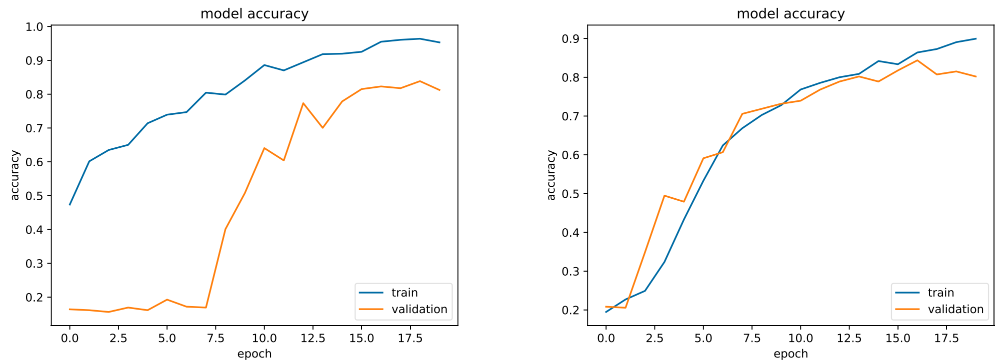
5. Faster training: Stacked ensemble models may train more quickly than conventional deep learning models. This is so that the overall training time can be decreased. The base models can be trained independently, and the ensemble-model can be trained using the predictions of the base models.

Overall, stacked ensemble models provide an effective and scalable method for machine learning that can enhance accuracy in prediction, limit overfitting, and reveal information about the underlying mechanisms influencing the predictions.

Stacked ensemble models have a number of benefits over conventional deep learning models, but they also have certain limitations, such as:

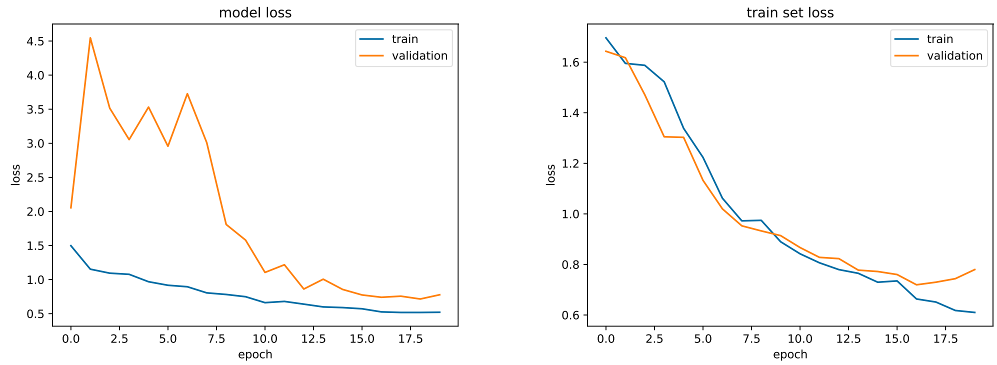
1. Complexity: Compared to conventional deep learning models, stacked ensemble models can be more difficult to develop and maintain. This is due to the fact that they call for the training and integration of numerous models, which can be time-consuming and costly in terms of computing.
2. Overfitting: Stacked ensemble models can help to lessen overfitting, but they are not resistant to it. The stacked ensemble model can still overfit to the training data if the base models are poorly constructed or if the ensemble-model is overfitting to the validation data.
3. Data accessibility: In order to train properly, they need a lot of data. It could be challenging to train numerous models and produce a reliable stacked ensemble model if there is not enough data available.
4. Costly to compute: Especially for large datasets, training many models and a ensemble model can be computationally expensive. Due to this, scaling stacked ensemble models to very large datasets may be challenging.

In general, stacked ensemble models can be an excellent deep learning technique, but they need to be carefully implemented and managed to be reliable. In order to evaluate the fitness of the model confusion matrix for sequential CNN, spatial attention powered deep CNN and ensemble model are given in Fig. 13a, 13b and 13c, respectively. It clearly shows that the ensemble model perform



(a) Sequential model (b) Spatial attention model

Fig. 11. Accuracy curve for sequential and attention model separately.



(a) Sequential model (b) Spatial attention model

Fig. 12. Loss curve for sequential and attention model separately.

well compare to other two models acted separately. Fig. 14 gives the classification report for the ensemble model and that shows the superior precision, recall and f1 score compared to the all previous transfer learning methods that studied so far with the proposed database in the study. Finally, the prediction with the ensemble model shows the good classification results that are totally taken on random basis and are shown in Fig. 15.

4.4. Impact of noise

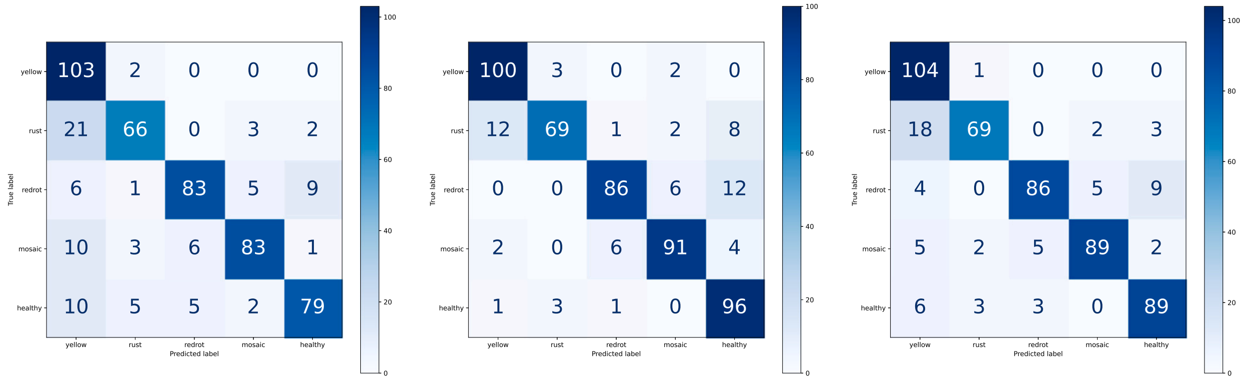
There is always a possibility that noise will affect the classification performance. It is evident that due to their spatial appearance, rust and red rot are simpler to differentiate from the other three categories. However, other classes are very similar to one another. But there are a few factors that influence the classification: They are,

1. **Uneven illumination**: Sometimes the presence of sunlight can make leaves appear yellow, resulting in instances when healthy plants are wrongly categorised as yellow.
2. **Background**: It's possible that the background of a certain image matches the foreground of other classes in a close way, causing the model to predict the latter as the classification output.
3. **Incorrect sampling**: If samples are taken during database collection that are near to the boundary of any two classes, the model might predict incorrect results.

This effect can be seen in Fig. 16, leftmost panel indicates the impact of uneven illumination, next panel highlights the role of background in incorrect classification and the rightmost shows the impact of incorrect sampling. Additionally, because yellow disease and mosaic look so similar to one another, they are commonly misclassified. Appropriate image processing techniques can be advantageous during data input in order to improve results.

5. Conclusion

In this research, a comparative analysis using transfer learning (TL) techniques, significant members of the deep learning family of networks, has been proposed. By adopting a TL-based approach, which also efficiently allows use of trained models for classification



(a) Sequential model (b) Spatial attention model (c) Ensemble model

Fig. 13. Confusion matrix for components of ensemble model.

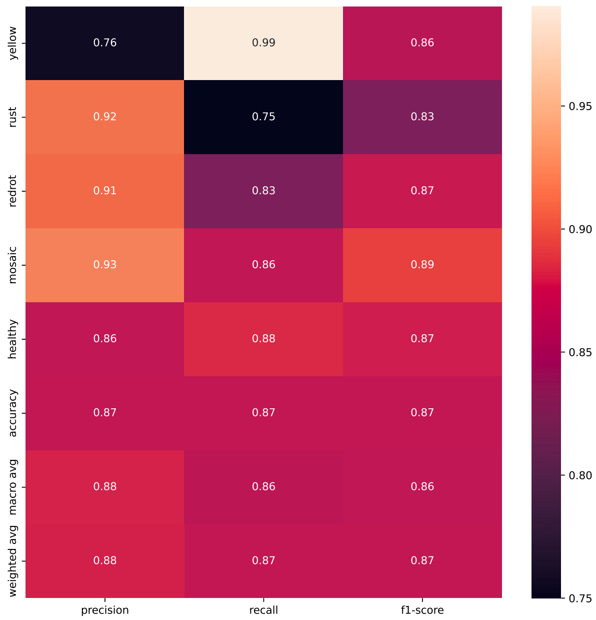


Fig. 14. Classification report and performance metrics for ensemble model.

and detection tasks, the need for large databases can be avoided. Our TL-based technique can automatically extract the unique features of foliar diseases and categorise them into five classes with a maximum accuracy of 84% for MobileNetV2. This empirical study used about 2569 self-collected images, and a pretrained model underwent thorough evaluation. To assist a vast majority of farmers in developing nations like India, the database was created with the need in mind, and all requirements were upheld with the application in mind.

Google Colaboratory, which has GPU support, was used for all experiments. Despite previous studies suggesting that efficientNet performs better, MobileNet beats every model that was employed in the study. It showed 84% accuracy with a parameter that was almost the same as efficientNet. MobileNet can be an amazing Android deployment option with additional parameter fine-tuning. It features fewer parameters and less computing overhead as compared to other variations employed in the experimentation. Loss curves also demonstrate the mobilenetV2's performance with other models on such a small database. However, the length of time required for a model to converge is unquestionably greater. The current study demonstrated an accuracy of 84%.The stack ensemble model proposed here, however, is able to achieve the same results with an improved level of accuracy of 86.53% in just 20 epochs. With such small datasets, this model may be a good choice where the memory restriction is less of an issue.

In short, it can be seen from the comparison of the ensemble model and the transfer learning model that each has its own advantages and disadvantages. The ensemble model, however, performs better than the transfer learning model for small datasets when it comes to higher accuracy and good classification outcomes. The ensemble learning model appears as a more efficient and effective way for enhancing model performance, especially when database is small and you need faster training. Therefore, using

Fig. 15. Predictions with ensemble model.

ensemble models in different applications can improve classification and accuracy, opening up opportunities for future deep learning models that are more advanced.

In the future, the focus of the strategy will progressively be on improving accuracy. Additionally, the size and degree of variation of the current database are relatively small. The database size will be increased to address any potential issues like overfitting in the future. An accurate, swift, and responsive system for diagnosing sugarcane disease will be created in the future.

CRediT authorship contribution statement

SWAPNIL DADABHAU DAPHAL: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Sanjay Koli: Contributed reagents, materials, analysis tools or data.

Fig. 16. (panel from left to right) Impact of uneven lightning condition, incorrect background and incorrect sampling of image, respectively.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Additional information

No additional information is available for this paper.

References

[1] M. Ahmad, M. Abdullah, H. Moon, D. Han, Plant disease detection in imbalanced datasets using efficient convolutional neural networks with stepwise transfer learning, IEEE Access 9 (2021) 140565–140580, https://doi.org/10.1109/ACCESS.2021.3119655.

[2] H. Amin, A. Darwish, A.E. Hassanien, M. Soliman, End-to-end deep learning model for corn leaf disease classification, IEEE Access 10 (2022) 31103–31115, https://doi.org/10.1109/ACCESS.2022.3159678.

[3] B. Chellapandi, M. Vijayalakshmi, S. Chopra, Comparison of pre-trained models using transfer learning for detecting plant disease, in: 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), 2021, pp. 383–387.

[4] Swapnil Daphal, S. Koli, Sugarcane leaf disease dataset. mendeley data, v1, 10.17632/9424skmnrk.1.

[5] S.D. Daphal, S.M. Koli, Transfer learning approach to sugarcane foliar disease classification with state-of-the-art sugarcane database, in: 2021 International Conference on Computational Intelligence and Computing Applications (ICCICA), 2021, pp. 1–4.

[6] S.D. Daphal, S.M. Koli, Efficient use of convolutional neural networks for classification of sugarcane leaf diseases, in: A. Kumar, S. Mozar (Eds.), ICCCE 2021, Springer Nature Singapore, Singapore, 2022, pp. 675–680.

[7] C. Dong, Z. Zhang, J. Yue, L. Zhou, Automatic recognition of strawberry diseases and pests using convolutional neural network, Smart Agric. Technol. 1 (2021) 100009, https://doi.org/10.1016/j.atech.2021.100009, https://www.sciencedirect.com/science/article/pii/S2772375521000095.

[8] E. Elfatimi, R. Eryigit, L. Elfatimi, Beans leaf diseases classification using mobilenet models, IEEE Access 10 (2022) 9471–9482, https://doi.org/10.1109/ACCESS.2022.3142817.

[9] M. Geetha, R. Suganthe, S. Nivetha, R. Anju, R. Anuradha, J. Haripriya, A time-series based yield forecasting model using stacked lstm to predict the yield of paddy in cauvery delta zone in tamilnadu, in: 2022 First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), 2022, pp. 1–6.

[10] S. Ghosal, K. Sarkar, Rice leaf diseases classification using cnn with transfer learning, in: 2020 IEEE Calcutta Conference (CALCON), 2020, pp. 230–236.

[11] W. Haider, A.u. Rehman, M.A. Abbasi, S. Khan, N.M. Durrani, Towards hybrid smart irrigation for mixed-cropping, in: 2022 Global Conference on Wireless and Optical Technologies (GCWOT), 2022, pp. 1–6.

[12] S.M. Hassan, A.K. Maji, Plant disease identification using a novel convolutional neural network, IEEE Access 10 (2022) 5390–5401, https://doi.org/10.1109/ACCESS.2022.3141371.

[13] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[14] D. Hughes, M. Salathé, et al., An open access repository of images on plant health to enable the development of mobile disease diagnostics, preprint, arXiv:1511.08060, 2015.

[15] A. Irfan, A.L. Adivishnu, A. Sze-To, T. Dehkharghanian, S. Rahnamayan, H. Tizhoosh, Classifying pneumonia among chest x-rays using transfer learning, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC), 2020, pp. 2186–2189.

[16] Saranya K, Deena Dhayalan S, Prasanth R, Sathish M, Agriculture based recommendation system with image processing, in: 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT), 2022, pp. 1–6.

[17] D. Kumar, V. Kukreja, N-cnn based transfer learning method for classification of powdery mildew wheat disease, in: 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), 2021, pp. 707–710.

[18] A. Lakshmanarao, M.R. Babu, T.S.R. Kiran, Plant disease prediction and classification using deep learning convnets, in: 2021 International Conference on Artificial Intelligence and Machine Vision (AIMV), 2021, pp. 1–6.

[19] X. Li, H. Jiang, X. Jiang, M. Shi, Identification of geographical origin of chinese chestnuts using hyperspectral imaging with 1d-cnn algorithm, Agriculture 11 (2021), https://www.mdpi.com/2077-0472/11/12/1274.
[20] Z. Liu, R.N. Bashir, S. Iqbal, M.M.A. Shahid, M. Tausif, Q. Umer, Internet of things (iot) and machine learning model of plant disease prediction–blister blight for tea plant, IEEE Access 10 (2022) 44934–44944, https://doi.org/10.1109/ACCESS.2022.3169147.
[21] L. Loyani, K. Bradshaw, D. Machuve, Segmentation of tuta absoluta 's damage on tomato plants: a computer vision approach, Appl. Artif. Intell. 35 (2021) 1–21, https://doi.org/10.1080/08839514.2021.1972254.
[22] S. Nigam, R. Jain, S. Marwaha, A. Arora, Wheat rust disease identification using deep learning, in: Conference Proceedings, 2021, pp. 1104–1107.
[23] R.R. Patil, S. Kumar, Rice-fusion: a multimodality data fusion framework for rice disease diagnosis, IEEE Access 10 (2022) 5207–5222, https://doi.org/10.1109/ACCESS.2022.3140815.
[24] Y. Peng, A. Wang, J. Liu, M. Faheem, A comparative study of semantic segmentation models for identification of grape with different varieties, Agriculture 11 (2021), https://www.mdpi.com/2077-0472/11/10/997.
[25] T.N. Pham, L.V. Tran, S.V.T. Dao, Early disease classification of mango leaves using feed-forward neural network and hybrid metaheuristic feature selection, IEEE Access 8 (2020) 189960–189973, https://doi.org/10.1109/ACCESS.2020.3031914.
[26] Public information bureau, https://pib.gov.in/PressReleasePage.aspx?PRID=1741942.
[27] P. Rubini, P. Kavitha, Deep learning model for early prediction of plant disease, in: 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021, pp. 1104–1107.
[28] A. Rössler, D. Cozzolino, et al., Faceforensics: a large-scale video dataset for forgery detection in human faces, http://arxiv.org/abs/1803.09179.
[29] N. Sabri, N.S. Kassim, S. Ibrahim, R. Roslan, N.N.A. Mangshor, Z. Ibrahim, Nutrient deficiency detection in maize (zea mays l.) leaves using image processing, IAES Int. J. Artif. Intell. (2020).
[30] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.C. Chen, Mobilenetv2: inverted residuals and linear bottlenecks, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510–4520.
[31] K.K. Sarma, K.K. Das, V. Mishra, S. Bhuiya, D. Kaplun, Learning aided system for agriculture monitoring designed using image processing and iot-cnn, IEEE Access 10 (2022) 41525–41536, https://doi.org/10.1109/ACCESS.2022.3167061.
[32] N. Silva, I. Siqueira, S. Okida, S.L. Stevan, H. Siqueira, Neural networks for predicting prices of sugarcane derivatives, Sugar Tech. 21 (2019) 514–523, https://doi.org/10.1007/s12355-018-0648-5.
[33] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, https://arxiv.org/abs/1409.1556.
[34] J. Som-ard, M.D. Hossain, S. Ninsawat, V. Veerachitt, Pre-harvest sugarcane yield estimation using uav-based rgb images and ground observation, Sugar Tech. 20 (2018) 645–657, https://doi.org/10.1007/s12355-018-0601-7.
[35] M. Tan, Q.V. Le, Efficientnet: rethinking model scaling for convolutional neural networks, arXiv:1905.11946 [abs], 2019.
[36] C. Wang, D. Chen, L. Hao, X. Liu, Y. Zeng, J. Chen, G. Zhang, Pulmonary image classification based on inception-v3 transfer learning model, IEEE Access 7 (2019) 146533–146541, https://doi.org/10.1109/ACCESS.2019.2946000.
[37] Z. Weidong, W. Chun, H. Jing, Development of agriculture machinery aided guidance system based on gps and gis, in: 2010 World Automation Congress, 2010, pp. 313–317.
[38] Y. Wu, X. Feng, G. Chen, Plant leaf diseases fine-grained categorization using convolutional neural networks, IEEE Access 10 (2022) 41087–41096, https://doi.org/10.1109/ACCESS.2022.3167513.
[39] V. Dhaka, S. Meena, G. Rani, D. Sinwar, Ijaz M. Kavita, M. Woźniak, A survey of deep convolutional neural networks applied for prediction of plant leaf diseases, Sensors 21 (2021), https://www.mdpi.com/1424-8220/21/14/4749.
[40] N. Kundu, G. Rani, V. Dhaka, K. Gupta, S. Nayak, S. Verma, M. Ijaz, M. Woźniak, IoT and interpretable machine learning based framework for disease prediction in pearl Millet, Sensors 21 (2021), https://www.mdpi.com/1424-8220/21/16/5386.
[41] S. Woo, J. Park, J. Lee, I.C.BA. M Kweon, Convolutional block attention module, CoRR, arXiv:1807.06521 [abs], 2018, http://arxiv.org/abs/1807.06521.