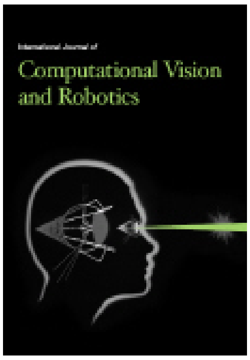


International Journal of Computational Vision and Robotics

ISSN online: 1752-914X - ISSN print: 1752-9131
https://www.inderscience.com/ijcvr

Comparison of convolutional neural networks architectures for mango leaf classification

B. Jayanthi, Lakshmi Sutha Kumar

DOI: 10.1504/IJCVR.2022.10049962

Article History:

Received:	21 February 2022
Accepted:	05 July 2022
Published online:	01 December 2023

Comparison of convolutional neural networks architectures for mango leaf classification

B. Jayanthi* and Lakshmi Sutha Kumar

Department of Electronics and Communication,
National Institute of Technology Puducherry,
Karaikal, India
Email: jayanthib1995@yahoo.com
Email: lakshmi@nitpy.ac.in
*Corresponding author

Abstract: Plant diseases are a threat to the food supply as they reduce the yield, and reduce the quality of fruits and grains. Hence, early identification and classification of plant diseases are essential. This paper aims to classify mango plant leaves into healthy and diseased using convolutional neural networks (CNNs). The performance comparison of CNN architectures, AlexNet, VGG-16 and ResNet-50 for mango plant disease classification is provided. These models are trained using the Mendeley dataset, validation accuracies are found and compared with and without the use of transfer learning models. AlexNet (25 layers, 6.2 million parameters) produces a testing accuracy of 94.54% and consumes less training time. ResNet-50 (117 layers, 23 million parameters) and VGG-16 (16 layers, 138 million parameters) have given testing accuracies of 98.56% and 98.26% respectively. Therefore, based on the accuracies achieved and complexity, this paper recommends AlexNet followed by ResNet-50 and VGG-16 for plant leaf disease classification.

Keywords: convolution neural networks; neural network; image classification; precision agriculture.

Reference to this paper should be made as follows: Jayanthi, B. and Kumar, L.S. (2024) 'Comparison of convolutional neural networks architectures for mango leaf classification', *Int. J. Computational Vision and Robotics*, Vol. 14, No. 1, pp.84–98.

Biographical notes: B. Jayanthi obtained her BE in the Department of Electronics and Communication and ME in VLSI Design and Embedded Systems from Anna University, Chennai, India in 2017 and 2019, respectively. Her area of interest includes deep learning and FPGA hardware accelerator design, etc. She is currently pursuing her PhD from National Institute of Technology, Puducherry, India.

Lakshmi Sutha Kumar received her BEng degree from Bharathidasan University, India in 1994, MTech degree from VIT, India in 2005, and PhD degree from Nanyang Technological University, Singapore in 2012. From 1995 to 1998 and from 1998 to 2002, she was a Lecturer at Bharathidasan University and Pondicherry University, India, respectively. From 2011 to 2012 and 2012 to 2014, she was a Research Engineer and Research Fellow at the School of EEE, Nanyang Technological University, Singapore, respectively. She joined the Department of Electronics and Communication Engineering, NIT Puducherry, Karaikal, Puducherry, India as an Assistant Professor in 2014 and she is currently working as an Associate Professor at the same department. Her research interests include remote sensing, climate and weather-related studies, digital image processing, speech processing, pattern recognition, and machine learning. She has more than 65 publications in reputed journals and conferences.

1 Introduction

Agriculture is the backbone of the any country's economy. It provides the necessary food and fulfils the needs of growing population. Plant diseases affect the food security of the nation. Detection and classification of the plant disease at an early stage avoids problems like quality diminution, decrease in the yield and loss for the farmers. Certain pathogens present in the mango fruits produce toxins which create serious health problems to the consumers. Mango plants grown in Indian mango plants contribute to nearly 40% of the total world's mango production (Sekhar and Prahadeeswaran, 2013). They need three to six years of proper growth and maintenance for giving better yield. Hence, it is essential to design a model that facilitates the early identification and classification of mango leaf diseases.

Artificial Intelligence (AI) is one of the most competent technologies in computer vision and the classification results obtained using AI techniques are more reliable. Deep learning (DL) is a key subset in AI and is often employed in plant disease identification.

There is a considerable amount of research done for the plant disease identification. Plant village dataset is the dataset that is used for plant disease identification in the literature. Kapach et al. (2012), explains the machine learning approaches for automatic image classification. In Bhattacharyya et al. (2019), a facial image is divided into a number of regions and certain features are extracted from the image and then the extracted features are classified using an SVM classifier and the results of the individual regions are combined using a genetic algorithm. In Haider et al. (2021), ML and DL techniques are compared and the results are presented. While comparing all the ML models, DT provides the highest accuracy of 94.70%. The highest accuracy using Dl model is obtained using sequential CNN. It provides a training accuracy of 90.40% and testing accuracy of 97.20%. Thus, DL models provide better results compared to ML results. In Barburiceanu et al. (2021), texture features are extracted from different layers of pre-trained convolutional neural network (CNN) models and it is applied to a machine-learning classifier. Plant village dataset is used for this classification problem. This method is compared with other machine learning and the proposed method provides more efficient result than the other models. And this works well for lesser dataset also. In Pham et al. (2020), ANN and CNN are compared. The CNN architectures used for comparison are AlexNet, VGG-16 and ResNet-50. From the comparison, it has been identified that CNN models provide better accuracy. As mentioned in Al-Barazanchi et al. (2018), the accuracy of the machine learning classifier is based on the quality of the features used. The identification of the best features is a time-consuming process and the set of extracted features are specific for the dataset. If the dataset varies, it is necessary to do the feature identification again. These are the disadvantages of machine learning algorithms. Hence, they conclude that CNN architectures are better than machine learning algorithms.

The type of CNNs architecture used for plant disease identification also vary according to the application. Arya and Singh (2019) have performed classification on tomato and mango leaves. A comparison between a CNN model designed in Arya and Singh (2019) and existing AlexNet model is given. Tomato leaf images are taken from the openly available plant village dataset and for mango leaves the database is captured from real-time environment. The accuracies of their classification schemes are 90.85% and 98% respectively. In Singh et al. (2019), the dataset which consists of 1070 images of the mango leaves is obtained from Shri Mata Vaishno Devi University, Katra, J&K, India. The architecture used in Singh et al. (2019) produced inea train accuracy of 98% and a test accuracy of 96%. In Madiwalar and Wyawahare (2017), comparison of minimum distance classifier and support vector machine (SVM) is done and the accuracy of 79.16% and 83.34% are obtained respectively. In Bhunia et al. (2019), a texture synthesis network (TSN) is pre-trained which takes a texture patch as input and outputs an enlarged view of the texture by injecting newer texture content. By encoding the learnt texture specific information in its intermediate layers. In the second network, the multi-scale feature representations from the TSN's intermediate layers are combined into a dense continuous representation which is converted into a binary hash code with the help of individual and pairwise label information. The new enlarged texture patches from the TSN can be used in data augmentation. This method produces superior results compared to the state of the art networks The model used in Venkatesh et al. (2020) is a combination of VGG Net and InceptionV2 architecture which produces the accuracy of 92% using the Plant village dataset. Generative adversarial networks (GANs) in Gandhi et al. (2018) is used to augment the limited number of images in the dataset. The classification is done by Inception v3 with an accuracy 88.6% and Mobilenet with an accuracy of 92%. In Mukherjee et al. (2017), GoogleNet is chosen for automatic disease classification and this model provides an average accuracy of 85.04%. In Amin et al. (2022), a new CNN model is proposed. The proposed model is a combination of two state of the art architectures. EfficientNetB0 and DenseNet121 are advanced architectures of EfficientNet and DenseNet respectively. After the classification part of both the architecture is completed both the weights are merged and then it is applied to fully connected layer and then to the softmax layer. Thus, it provides an accuracy of 98.56% which is higher than the two baseline models.

In Durmus et al. (2017), AlexNet and SqueezeNet architecture were tested on Plant village dataset and provide an accuracy of 95.65% and 94.3% respectively. AlexNet and VGG architecture were tested in Ferentinos (2018) on a dataset of 87,848 images, containing 25 different plants in a set of 58 distinct classes of [plant, disease] combinations and gave an accuracy of 99.45% and 99.50% respectively. In Kumar et al. (2020), ResNet-50 architecture is tested on a dataset consisting of 15,200 images that covers 14 crops and is divided into 38 different classes and is derived from the original Plant village dataset. The accuracy of their model is 99.40%. It is found from the literature that the successful CNN models for plant disease classification are AlexNet, VGG and ResNet-50 with classification accuracies greater than 90%.

In this paper, three CNN architectures namely, AlexNet, VGG-16 and ResNet-50 are used for the classification of mango leaves into healthy or diseased using Mendeley dataset. The performance measures of the classification are found and the comparison of the three architectures are provided.

The rest of the paper is arranged as follows. Section 2 presents the description of dataset used for the classification. The commonly implemented CNN architectures used for image classification are explained in Section 3. Results and discussions are provided in Section 4 and Section 5 concludes the paper.

2 Image dataset

Plants play a major role in maintaining the ecological balance and also improve the livelihood of human beings by enriching the atmosphere with lot of essential elements. They also serve as the major and only natural source of oxygen and control climate change. In spite of having all these advantages, rise in the population and the exploitation of the cultivation lands and forests have resulted in the destruction and extinction of a lot of plant species. Therefore, in an attempt to preserve the economically and environmentally beneficial plants, Chouhan et al. (2019) have made a contribution by collecting the leaf images in both healthy and diseased condition of 12 plants and his dataset is called Mendeley dataset. These plants include mango, arjun, alstonia scholaris, guava, bael, jamun, jatropha, pongamia pinnata, basil, pomegranate, lemon, and chinar. In this paper, the dataset having the mango leaf images is used for study and analysis. There are 430 images of mango leaves in the Mendeley dataset and Table 1 shows the classes of the mango leaves.

Table 1 Mendeley dataset with classes of mango leaves

Class	*No. of images*
C1 – Healthy leaf	170
C2 – Diseased leaf	260

The images belong to the RGB colour model. The dimension of the images is 6000x4000.The database of Mango leaves is further divided as training set, validation set and testing test before applying them to the selected DL model. The training set, validation set and test set contain 70%, 20% and 10% of the images respectively. Both classes of mango leave images are selected for training, validating and testing at a random fashion. The images are pre-processed and rescaled into a lower resolution based on the input size of the selected DL model. The processing techniques involved before and during the classification methods are explained next.

3 Plant disease classification for mango leaves

3.1 Image pre-processing

The image is first converted into black and white image and the edges of the leaf image is detected. Then, a bounding box is framed to cover the entire leaf including the edges. As the images in the database are of size 6,000 × 4,000, it is necessary to rescale the image based on the CNN model used. Therefore, the image is cropped to the size of the bounding box. Finally, the image is converted back to the RGB colour model and contrast enhancement is performed.

Histogram equalisation is one of the most commonly used technique for contrast enhancement. Even though the contrast enhancement is done, it is important to maintain the mean brightness of the image in order to get the exact information from the image (Rahman et al., 2015). In this paper, dynamic stretching-based brightness preservation (DSBP) method is used to enhance the image by preserving the mean brightness. DSBP method adopts the sub-image separation principle, and the threshold value for separating the sub-image is based on the golden section search approach (Chang, 2009). The image in the RGB space is converted to the hue, saturation and intensity (HSI) space and then the intensity channel alone is separated by using the method in Gonzalez and Woods (2006). The '*I*' channel is then separated into the low (γ_{lo}) and high (γ_{hi}) group based on the threshold obtained by golden search method. Then, histogram equalisation is applied on both the groups to occupy their complete intensity range. The procedure for intensity separation and stretching is summarised below.

$$\gamma_{hi} = \{\gamma(i) \mid i > \gamma_{th}\} \tag{1}$$

$$\gamma_{lo} = \{\gamma(i) \mid i \le \gamma_{th}\} \tag{2}$$

$$\gamma_{ehi} = \gamma_{th-1} + \frac{\left(\gamma_{hi} - \min\{\gamma_{hi}\}\right)}{\left(\max \gamma_{hi} - \min \gamma_{hi}\right)} * \left(I - 1 - \gamma_{th}\right) \tag{3}$$

$$\gamma_{elo} = \frac{\left(\gamma_{lo} - \min\{\gamma_{lo}\}\right)}{\left(\max\{\gamma_{lo}\} - \min\{\gamma_{lo}\}\right)} * \left(\gamma_{th}\right) \tag{4}$$

where I is maximum intensity level, γ_{lo} and γ_{hi} are low and high intensity groups and γ_{elo} and γ_{ehi} are stretched intensity values. The equations (1) and (2) explain the splitting of the intensity levels into low and high groups. The equations (3) and (4) explain the stretching of intensities. The enhanced intensity component is combined with the hue and saturation for forming the HSI image. The enhanced HSI image is then converted to RGB image. Figure 1(a) and Figure 1(b) show the original image and rescaled enhanced image respectively.

Figure 1 (a) Original image (b) Rescaled and enhanced image (see online version for colours)

(a) (b)

3.2 CNN architectures

After pre-processing the original image, the image classification is performed by using the state of the art architectures, AlexNet (Krizhevsky et al., 2012), VGG-16(Simonyan and Zisserman, 2015), Resnet-50 (He et al., 2016). AlexNet is the architecture that won the ImageNet large-scale visual recognition challenge (ILSVRC) in 2012. The inputs to this model are RGB images. It has eight layers with learnable parameters and has five convolution layers with a combination of max-pooling layers followed by three fully connected layers. The activation function used in all the layers except the last layer is Relu. The activation function used in the output layer is Softmax. The architecture of AlexNet is shown in Figure 2. The total number of parameters present in AlexNet are 62.3 million.

Figure 2 AlexNet architecture

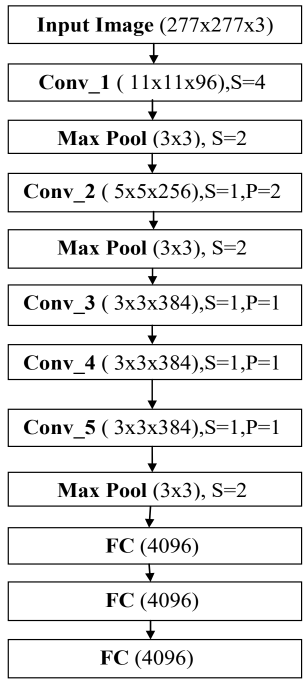
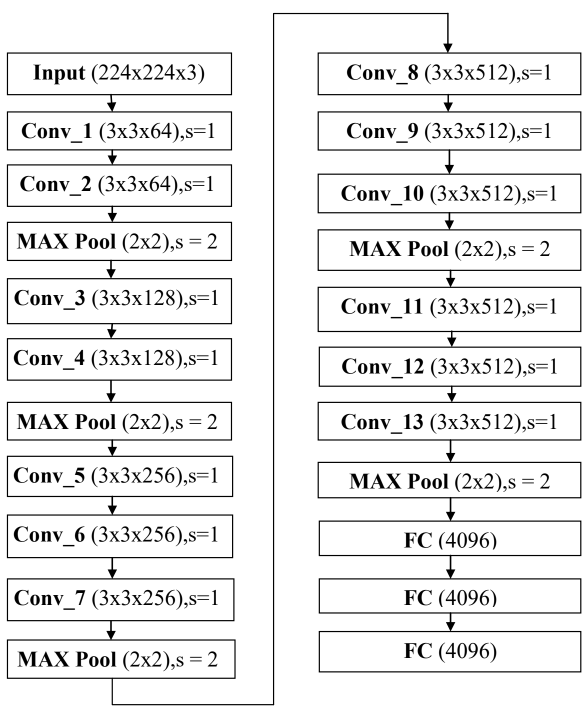
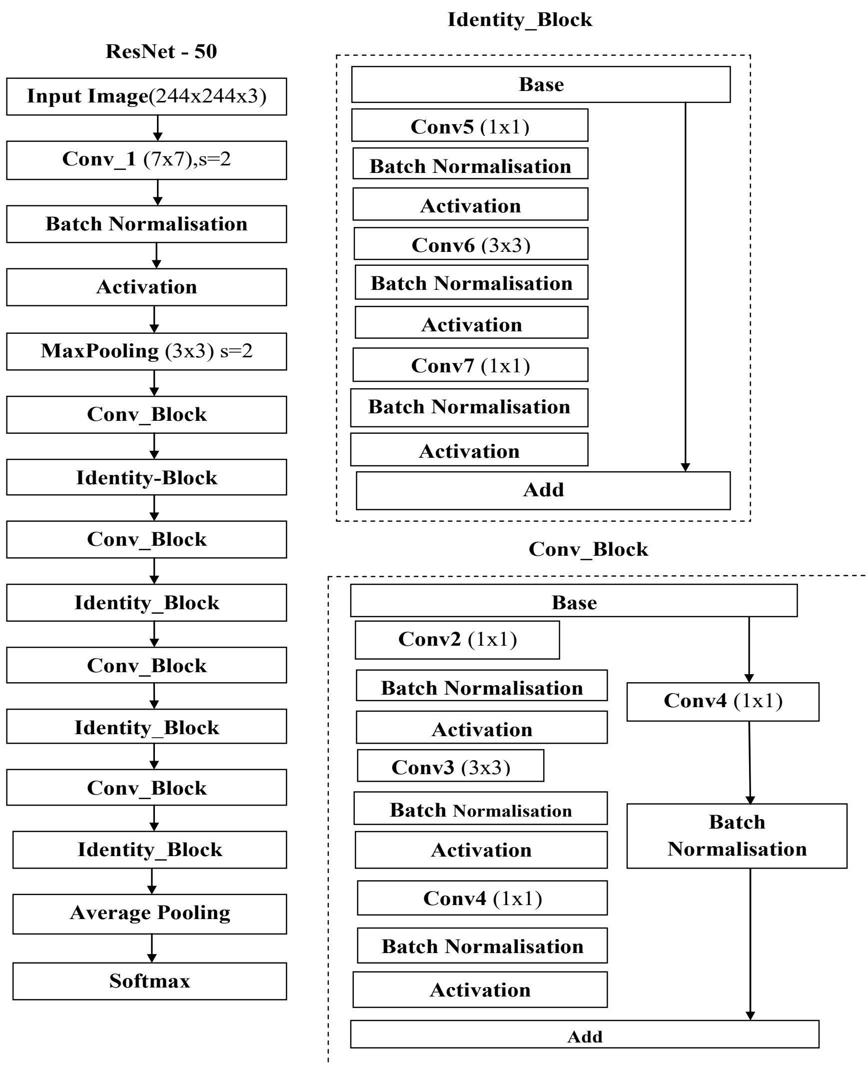


Figure 3 VGG-16 architecture



A combination of max-pooling layers followed by three fully connected layers. The activation function used in all the layers except the last layer is Relu. The activation function used in the output layer is Softmax. The architecture of AlexNet is shown in Figure 2. The total number of parameters present in AlexNet are 62.3 million.

Figure 4 ResNet architecture



VGG-16 is the architecture that won the ILSVRC in 2014.It has 16 layers which have weights. This network is a large network and it has about 138 million parameters approximately. VGG-16 architecture is shown in Figure 3. ResNet-50 is one of the important architectures which is participated in the ILSVRC in the year 2012 along with AlexNet. Many residual Networks are stacked together to form ResNet-50 and is shown in Figure 4. Deep CNNs was trained using the ResNet-50 model. The network has 23 million parameters.

The inputs to the CNN Architectures are the RGB images and their size vary based on the size of the input layer present in the architecture. The CNN model was trained, validated and tested with 70%, 20% and 10% Mendeley dataset images. The no. of images used for training, validating and testing are 273, 117 and 40, respectively. The model was trained with and without transfer learning. Transfer learning is using the model trained for a particular application and applying it for different applications (Chen et al., 2020). If the transfer leaning is used, it is not necessary for the model to learn the weights from the scratch. In this paper, the pre-trained models, trained with ImageNet database are used for plant disease classification. ImageNet database has 1,000 classes and the pre-trained models have already learned the features which are required for plant disease classification. But the use of transfer learning may sometimes not increase the accuracy of classification depending on the dataset used and the type of application. The comparison of classification methods with and without transfer learning model is presented for all the three state-of-the-art-architectures next

4 Results and discussion

All the models are programmed in MATLAB R-2020a and implemented on a desktop PC using a single CPU. The training criteria is fixed for all the CNN models in order to make an effective comparison. The optimiser used is Adam Optimiser. It has a constant learning rate of 0.001. The training cycle has 20 epochs for each training. There are two iteration per epoch, and therefore the maximum number of iterations are 40. The metrics which are used to investigate the model are accuracy, precision, recall and F1-score. They are calculated using the following equations.

$$\text{Accuracy} = \text{Number of correctly classified images}/(\text{Total number of images}) \tag{5}$$

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Recall = \frac{TP}{TP + FP} \tag{7}$$

$$F1Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{8}$$

$$Specificity = \frac{TN}{TN + FP} \tag{9}$$

where TN = true negative, TP = true positive, FN = false negative and FP = false positive.

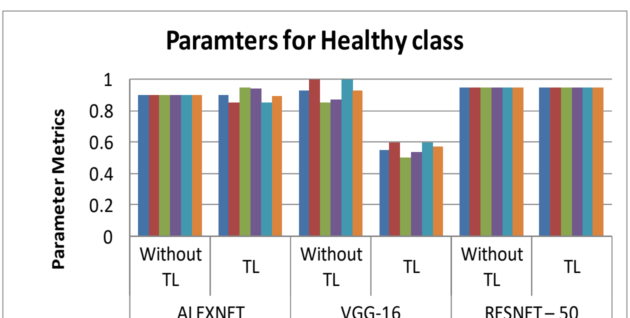
The performance metrics of the three CNN models considered for mango leave classification are calculated from the confusion matrix. The confusion matrix is formed based on the testing data by taking the predicted labels and actual labels as input. True positive is an outcome if the model correctly predicts the healthy class, true negative is an outcome if the model correctly predicts the diseased class, and false positive is produced as an outcome when the model incorrectly predicts the healthy class. And a false negative is an outcome if the model incorrectly predicts the diseased class. Accuracy is a ratio of correctly predicted observation to the total observations Precision is the ratio of correctly predicted healthy observations to the totally predicted healthy observations. Recall is the ratio of correctly predicted images to all the images in that particular class. F1 -Score is the weighted average of precision and recall.

The parameters that are investigated are displayed in Figure 5 and Figure 6. The AlexNet model has a high precision and accuracy. The model is comparatively smaller in terms of the number of layers used for classification. Since the number of layers used are less the training time is also reduced.

For VGG16 architecture, as shown in Figure 5, the precision and accuracy of the model without using the transfer learning is high compared to transfer learning model. This may be due to many factors. One of the reasons may be the model has over fit for the dataset. This problem may be due to the usage of small dataset. The solutions to avoid over fitting may include increasing the size of the dataset by using data augmentation methods. The next technique to overcome over fitting is to use early stopping which means the number of epochs that are used for training may be reduced to avoid over training of the model. Dropout layers may also be added in order to avoid over fitting.

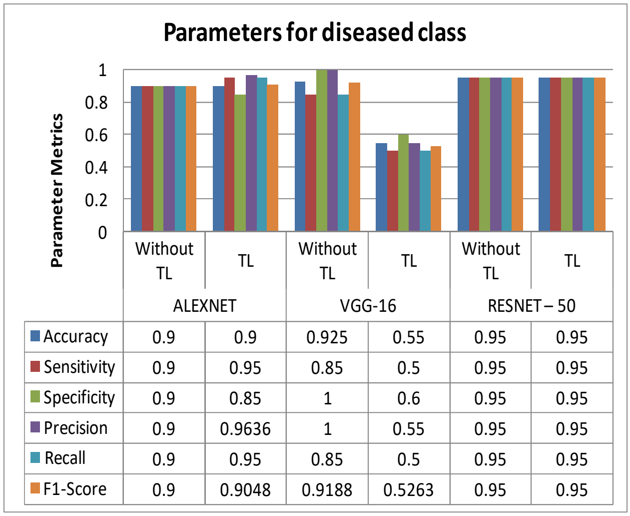
For ResNet-50, the precision for both the classes is the same and also the accuracy of the model is higher. As compared to AlexNet architecture, ResNet-50 provides more accuracy but the number of its layers is higher. With batch normalisation, ResNet-50 has 177 layers whereas AlexNet has only 25 layers. The time taken to train the ResNet-50 model is higher as compared to the AlexNet model. Figure 7 shows the comparison of the time taken to train the CNN architectures. As shown in Figure 7, the training times are different for different architectures. The training time depends on the number of layers present in the CNN architecture. As ResNet-50 and VGG-16 have a greater number of layers compared to AlexNet, their training times are more. Even though VGG-16 has a smaller number of layers as compared to Resnet-50, its training time is higher and the accuracy produced by it is lower. The use of transfer learning trains the models faster as compared to the models without transfer learning.

Figure 5 Performance analysis for healthy class (see online version for colours)



	ALEXNET Without TL	ALEXNET TL	VGG-16 Without TL	VGG-16 TL	RESNET – 50 Without TL	RESNET – 50 TL
Accuracy	0.9	0.9	0.925	0.55	0.95	0.95
Sensitivity	0.9	0.85	1	0.6	0.95	0.95
Specificity	0.9	0.95	0.85	0.5	0.95	0.95
Precision	0.9	0.9444	0.8696	0.54	0.95	0.95
Recall	0.9	0.85	1	0.6	0.95	0.95
F1-Score	0.9	0.8947	0.9302	0.57	0.95	0.95

Figure 6 Performance analysis for diseased class (see online version for colours)



The accuracy of CNN models is compared and shown in Figure 8. As can be seen from Figure 8, AlexNet model has an accuracy of 91.45% and 94.02% without using the transfer learning and with transfer learning respectively. VGG-16 model has an accuracy of 88% without transfer learning and achieved an accuracy of 58% while considering the model with transfer learning. VGG-16 model performed well in Venkatesh et al. (2020) and Ferentinos (2018) with accuracy not less than 90%. However, it need not be the case for all the dataset used (Venkatesh et al., 2020; Ferentinos, 2018). The accuracy comparison in Figure 8 shows that VGG-16 with the transfer learning model using the Mendely dataset has produced very lower accuracy of 58.9%. Therefore, it is concluded that the VGG-16 pre-trained model has not learnt the features well. ResNet-50 model has an accuracy of 97.44% without the transfer learning and with the transfer learning. The basic architecture and pre-trained model of ResNet-50 performs well as compared to VGG-16 architectures.

Figure 7 Comparison of the time taken to train the CNN architectures (see online version for colours)

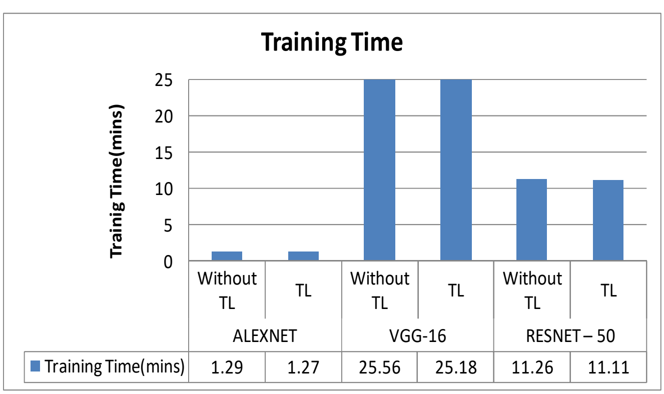
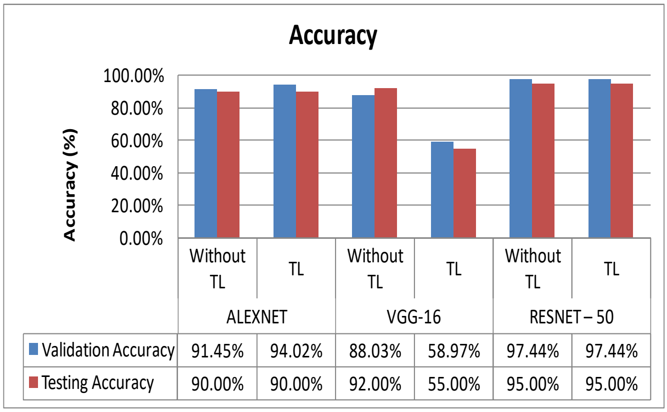


Figure 8 Comparison of accuracy for CNN models (see online version for colours)



It is found from literature that AlexNet model of Ferentinos (2018),VGG-16 model of Ferentinos (2018), and ResNet -50 model of (Kumar et al., 2020) produced an accuracy of 99.45%, 99.5% and 97% without transfer learning respectively. In this paper, AlexNet, VGG-16 and ResNet-50 models (without transfer learning) have provided accuracies of 91.45%, 88.03%, and 97.44% for mango leave classification using Mendely dataset respectively. Even though, AlexNet has given lower accuracy as compared to ResNet-50, the complexity of its architecture is lower and it takes less time to train it. Hence, Alexnet architecture is preferred over ResNet-50. Even though all the three models have provided the accuracy greater than 90%, VGG-16 architecture shows a very large difference in the training and testing accuracies. This may be because of the over-fitting of CNN models due to the smaller size of the dataset. Therefore, data augmentation techniques such as reflection, rotation and scaling were implemented on the original dataset to increase the size of the dataset into a total of 680 diseased mango leaf images and 1060 healthy leaf images. Figure 9 shows the comparison of accuracies for AlexNet, VGG-16 and ResNet models after the use of data augmentation techniques. After using data augmentation, the accuracy of all the three models have been increased.

Table 2 shows the comparison of accuracies of existing CNN models with the results of CNN models of this paper with and without transfer learning. The use of data augmentation techniques improved the training and testing accuracies of VGG-16 and ResNet-50 models and provided comparable performance for AlexNet model. From Table 2 and literature, existing AlexNet models gave accuracy ranging from 88.54 to 98.33%, existing VGG-16 models gave 85.30% accuracy and existing ResNet-50 models gave accuracy ranging from 86.58% to 99.40%. In this paper, AlexNet has given training and testing accuracies of 92.81% to 93.53% and 89.37% to 94.54% after data augmentation. This is a comparable performance with the existing AlexNet systems. For AlexNet model, training accuracy is increased by 2.27% with transfer learning and testing accuracy is increased by 5.04% without transfer learning.

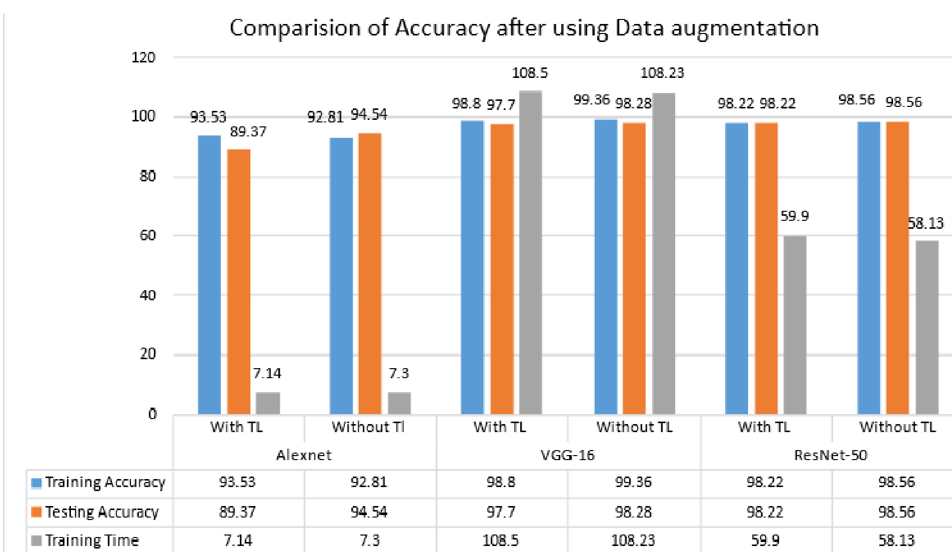
In this paper, for VGG-16 model with transfer learning, data augmentation increases the training and testing accuracies by 12.23% and 6.19% respectively. For VGG-16 without transfer learning, training and testing accuracies are increased by 67.54% and 77.63%. Similarly, for ResNet-50 model with transfer learning, data augmentation increases the training and testing accuracies by 0.80% and 3.38% respectively

Table 2 Performance analysis of CNN models with existing systems

Reference	*Architecture*	*Accuracy*	*Results obtained before data augmentation*		*Results obtained after data augmentation*	
			With TL	*Without TL*	*With TL*	*Without TL*
Arya and Singh (2019)	Alexnet	98.33%	91.45%	94.02%	93.53%	92.81%
Pham et al. (2020)		88.54%	90%	90.00%	89.37%	94.54%
Venkatesh et al. (2020)	VGG-16	85.30%	88.03% 92%	58.97%, 55.00%	98.8% 97.7%	99.36% 98.26%
Kumar et al. (2020)	ResNet-50	99.40%	97.44%	97.44%	98.22%	98.56%
Pham et al. (2020)		86.58%	95%	95%	98.22%	98.56%

For ResNet-50 without transfer learning, training and testing accuracies are increased by 1.15% and 3.75%. All the three CNN models, with and without transfer learning, show significant increase in accuracies for mango leaf disease classification after the use of data augmentation.

Figure 9 Comparison of accuracy for CNN models after data augmentation (see online version for colours)



AlexNet is able to produce testing accuracy of 94.54% with only 25 layers, 6.2 million parameters and consume less training time. ResNet-50 (117 layers, 23 million parameters) and VGG-16 (16 layers, 138 million parameters) have given testing accuracies of 98.56% and 98.26%, respectively. VGG-16 is recommended if higher accuracy is required but at the expense of higher complexity. ResNet-50 has moderate complexity. Therefore, based on the accuracies achieved and complexity, this paper recommends AlexNet followed by ResNet-50 and VGG-16 for plant leaf disease classification.

5 Conclusions

In this paper, mango leaf classification into healthy and diseased is presented using the CNNs. Most commonly used state-of-the-art architectures was identified for plant disease classification. Then, all the three models were trained similarly using the Mendeley dataset to obtain a proper comparison. The AlexNet, VGG-16 and ResNet-50 architectures provide testing accuracies of 90.0%, 92.0% and 95.0% without using the transfer learning model respectively. The testing accuracies of 90%, 55% and 95% with the transfer learning are obtained by AlexNet, VGG-16 and ResNet-50 respectively for original dataset. After data augmentation AlexNet, VGG-16 and ResNet-50 architectures provide testing accuracies of 89.37%, 97.7% and 98.22% with transfer learning and also provide testing accuracies of 94.54%, 98.26% and 98.56% without transfer learning respectively. The use of data augmentation improved the accuracies of VGG-16 and ResNet-50 significantly for mango leaf classification. AlexNet is able to produce testing accuracy of 94.54% with only 25 layers, 6.2 million parameters and consume less training time. ResNet-50 (117 layers, 23 million parameters) and VGG-16 (16 layers, 138 million parameters) have given accuracies of 98.56% and 99.36% respectively. VGG-16 is recommended if higher accuracy is required but at the expense of higher complexity. ResNet-50 has moderate complexity. Therefore, based on the accuracies achieved and complexity, this paper recommends AlexNet followed by ResNet-50 andVGG-16 for plant leaf disease classification.

References

Al-Barazanchi, H., Verma, A. and Wang, S.X. (2018) ‘Intelligent plankton image classification with deep learning’, *International Journal of Computational Vision and Robotics*, Vol. 8, No. 6, pp.561–571, https://doi.org/10.1504/IJCVR.2018.095584.

Amin, H. et al. (2022) ‘End-to-end deep learning model for corn leaf disease classification’, *IEEE Access*, Vol. 10, pp.31103–31115.

Arya, S. and Singh, R. (2019) ‘A comparative study of CNN and AlexNet for detection of disease in potato and mango leaf’, *IEEE International Conference on Issues and Challenges in Intelligent Computing Techniques, ICICT 2019, Dl*, https://doi.org/10.1109/ICICT46931.2019.8977648.

Barburiceanu, S. et al. (2021) ‘Convolutional neural networks for texture feature extraction. Applications to leaf disease classification in precision agriculture’, *IEEE Access*, Vol. 9, pp.160085–160103.

Bhattacharyya, A., Saini, R., Roy, P.P. et al. (2019) ‘Recognizing gender from human facial regions using genetic algorithm’, *Soft Comput*., Vol. 23, pp.8085–8100.

Bhunia, A.K., Perla, S.R.K., Mukherjee, P., Das, A. and Roy, P.P. (2019) ‘Texture synthesis guided deep hashing for texture image retrieval’, *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp.609–618, DOI: 10.1109/WACV.2019.00070.

Chang, Y.C. (2009) 'N-dimension golden section search: its variants and limitations', *Proceedings of the 2009 2nd International Conference on Biomedical Engineering and Informatics, BMEI 2009*, Vol. 3, pp.1–6, https://doi.org/10.1109/BMEI.2009.5304779.

Chen, J., Chen, J., Zhang, D., Sun, Y. and Nanehkaran, Y.A. (2020) 'Using deep transfer learning for image-based plant disease identification', *Computers and Electronics in Agriculture*, March, Vol. 173, p.105393, https://doi.org/10.1016/j.compag.2020.105393.

Chouhan, S.S., Kaul, A. and Singh, U.P. (2019) *A Database of Leaf Images: Practice towards Plant Conservation with Plant Pathology*, DOI:10.17632/hb74ynkjcn.1.

Durmus, H., Gunes, E.O. and Kirci, M. (2017) 'Disease detection on the leaves of the tomato plants by using deep learning', *2017 6th International Conference on Agro-Geoinformatics, Agro-Geoinformatics 2017*, https://doi.org/10.1109/Agro-Geoinformatics.2017.8047016.

Ferentinos, K.P. (2018) 'Deep learning models for plant disease detection and diagnosis', *Computers and Electronics in Agriculture*, January, Vol. 145, pp.311–318, https://doi.org/10.1016/j.compag.2018.01.009.

Gandhi, R., Nimbalkar, S., Yelamanchili, N. and Ponkshe, S. (2018) 'Plant disease detection using CNNs and GANs as an augmentative approach', *2018 IEEE International Conference on Innovative Research and Development, ICIRD 2018*, May, pp.1–5, https://doi.org/10.1109/ICIRD.2018.8376321.

Gonzalez, R.C. and Woods, R.E. (2006) *Digital Image Processing*, 3rd ed., Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Haider, W. et al. (2021) 'A generic approach for wheat disease classification and verification using expert opinion for knowledge-based decisions', *IEEE Access*, Vol. 9, pp.31104–31129.

He, K., Zhang, X., Ren, S. and Sun, J. (2016) 'Deep residual learning for image recognition', in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, June, p.770778, DOI: 10.1109/CVPR.2016.90.

Kapach, K., Barnea, E., Mairon, R., Edan, Y. and Ben-Shahar, O. (2012) 'Computer vision for fruit harvesting robots – state of the art and challenges ahead', *International Journal of Computational Vision and Robotics*, Vol. 3, Nos. 1–2, pp.4–34, https://doi.org/10.1504/IJCVR.2012.046419.

Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) 'ImageNet classification with deep convolutional neural networks', in *Proc. Adv. Neural Inf. Process. Syst.*, p.10971105, DOI: 10.1145/3065386.

Kumar, V., Arora, H., Harsh and Sisodia, J. (2020) 'ResNet-based approach for detection and classification of plant leaf diseases', *Proceedings of the International Conference on Electronics and Sustainable Communication Systems, ICESC 2020, ICESC*, pp.495–502, https://doi.org/10.1109/ICESC48915.2020.9155585.

Madiwalar, S.C. and Wyawahare, M.V. (2017) 'Plant disease identification: a comparative study', *2017 International Conference on Data Management, Analytics and Innovation, ICDMAI 2017*, pp.13–18, https://doi.org/10.1109/ICDMAI.2017.8073478.

Mukherjee, S. et al. (2017) 'Plant disease identification using deep neural networks', *Journal of Multimedia Information System*, Vol. 4, No. 4, pp.233–238.

Pham, T.N., Tran, L.V. and Dao, S.V.T. (2020) 'Early disease classification of mango leaves using feed-forward neural network and hybrid metaheuristic feature selection', in *IEEE Access*, Vol. 8, pp.189960–189973, DOI: 10.1109/ACCESS.2020.3031914.

Rahman, M.A., Liu, S., Lin, S.C.F., Wong, C.Y., Jiang, G. and Kwok, N. (2015) 'Image contrast enhancement for brightness preservation based on dynamic stretching', *International Journal of Image Processing (IJIP)*, Vol. 9, No. 4, pp.241–253.

Sekhar, C. and Prahadeeswaran, M. (2013) 'Production and export of mango in India – a paradigm to the developing nations', *American International Journal of Research in Humanities, Arts and Social Sciences*, Vol. 13, No. 337, pp.78–84.

Simonyan, K. and Zisserman, A. (2015) 'Very deep convolutional networks for large-scale image recognition', in *Proc. 3rd Int. Conf. Learn. Represent.*, p.114 [online] http://arxiv.org/abs/1409.1556 (accessed 5 October 2019).

Singh, U.P., Chouhan, S.S., Jain, S. and Jain, S. (2019) 'Multilayer convolution neural network for the classification of mango leaves infected by anthracnose disease', *IEEE Access*, Vol. 7, pp.43721–43729, https://doi.org/10.1109/ACCESS.2019.2907383.

Venkatesh, N.Y., Sahana, T.S., Swetha, S. and Hegde, S.U. (2020) 'Transfer learning based convolutional neural network model for classification of mango leaves infected by anthracnose', *2020 IEEE International Conference for Innovation in Technology, INOCON 2020*, https://doi.org/10.1109/INOCON50539.2020.9298269.