Computers, Materials & Continua
DOI:10.32604/cmc.2022.021875
Article

Tech Science Press

Deep Learning Based Automated Detection of Diseases from Apple Leaf Images

Swati Singh[1], Isha Gupta[2], Sheifali Gupta[2], Deepika Koundal[3,*], Sultan Aljahdali[4], Shubham Mahajan[5] and Amit Kant Pandit[5]

[1]Department of Electronics and Communication Engineering, University Institute of Technology, Himachal Pradesh University, Shimla, 171005, India
[2]Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
[3]Department of Systemics, School of Computer Science, University of Petroleum and Energy Studies, Dehradun, India
[4]Department of Computer Science, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia
[5]School of Electronics & Communication, Shri Mata Vaishno Devi University, Katra, 185320, India
[*]Corresponding Author: Deepika Koundal. Email: dkoundal@ddn.upes.ac.in
Received: 18 July 2021; Accepted: 30 August 2021

Abstract: In Agriculture Sciences, detection of diseases is one of the most challenging tasks. The mis-interpretations of plant diseases often lead to wrong pesticide selection, resulting in damage of crops. Hence, the automatic recognition of the diseases at earlier stages is important as well as economical for better quality and quantity of fruits. Computer aided detection (CAD) has proven as a supportive tool for disease detection and classification, thus allowing the identification of diseases and reducing the rate of degradation of fruit quality. In this research work, a model based on convolutional neural network with 19 convolutional layers has been proposed for effective and accurate classification of Marsonina Coronaria and Apple Scab diseases from apple leaves. For this, a database of 50,000 images has been acquired by collecting images of leaves from apple farms of Himachal Pradesh (H.P) and Uttarakhand (India). An augmentation technique has been performed on the dataset to increase the number of images for increasing the accuracy. The performance analysis of the proposed model has been compared with the new two Convolutional Neural Network (CNN) models having 8 and 9 layers respectively. The proposed model has also been compared with the standard machine learning classifiers like support vector machine, k-Nearest Neighbour, Random Forest and Logistic Regression models. From experimental results, it has been observed that the proposed model has outperformed the other CNN based models and machine learning models with an accuracy of 99.2%.

Keywords: Deep learning; convolutional neural network; apple leaves; apple scab; support vector machine

1 Introduction

Apple also known as malus domestica is commercially the chief significant moderate fruit in India and its industry is an essential constituent of India's whole agricultural economy. India is the sixth major apple manufacturer in the world. The agriculture of apple plants in farming has become considerable more than merely an earning to feed ever increasing population [1]. Apple is the most substantial energy source, in both human life and the other living creatures that exist in the world. Therefore, identifying the infection at an appropriate time and accurately is the utmost importance [2–4].

Pathology of plants can be perceived in numerous ways. Symptoms that are visible are associated with certain infections that cannot be noticed at all times and several of them seem only when it is too late when any precaution can't be taken care [5–7]. Therefore, there is a need to develop a technique to solve all of these problems [8].

By usage of image-processing in arrangement with recognition of pattern and classification tools, a lot of the difficulties will be determined or reduced. Apple leaves mainly are affected through many diseases like alternaria, black spot, Marsonina, Apple Scab and leaf miner pest. Diseases like alternaria, black spot and leaf miner pest have been detected in the past by the use of image processing but no work has been done on detecting Apple Scab and Marsonina coronaria apple leaf diseases by taking live background of the diseased images. This research work aims in detecting the two apple diseases, Apple Scab and Marsonina coronaria through machine learning and deep learning. Leaf affected with Marsonina coronaria is shown in Fig. 1a and Apple Scab disease is shown in Fig. 1b.

(a) (b)

Figure 1: Diseased apple leaves: (a) Marsonina Coronaria, (b) Apple Scab

This novel saliency detection approach intends to detect the apple leaf diseases by training the proposed CNN algorithm with the infected diseased apple leaf images. Database used is composed of images taken from farms of H.P and Uttarakhand (India). Total of 50000 live background images are collected, from the total database 25000 are of Marsonina coronaria and 25000 of Apple Scab diseases with the help of State Agriculture Universities of Himachal Pradesh and Uttarakhand. The collected data was also verified by the Horticulture department of Himachal Pradesh. The proposed algorithm is compared with already obtainable machine learning techniques and other CNN based architectures. The major contributions of the study are as follows:

- A database of 50000 apple leaf images Marsonina coronaria and Apple Scabs diseases with live background have been collected from apple farms of two different states of India Uttarakhand and Himachal Pradesh.
- The images of the dataset are increased to 2 lakh by using different augmentation techniques and the images were saved separately.
- A Convolutional Neural Network (CNN) based model is proposed having 19 layers with 6 convolutional layers, 3 max-pooling layers, 4 dropout layers, 3 Batch normalization layer, 1 flatten layer and 2 dense layers to detect above said apple diseases.
- The performance of the proposed model is compared with state of art models using various evaluation metrics.

The remaining paper is structured as follows. Section 2 reviews the past literature on machine learning and deep learning techniques. Section 3 gives description of materials and methods and also the proposed methodology. Experiments results and comparison of machine learning algorithms with the designed CNN are given in Section 4 and lastly the Section 5 gives the findings and conclusion.

2 Literature Review

An early detection of diseases apple leaf enhances the possibility of better crop yield, and there is a possible chance of recovery. Efforts are also being done to develop various algorithms that can categorize apple leaf and apple fruit diseases accurately and efficiently.

The existing widespread machine learning classifiers has been technologically advanced for early detection of apple diseases. The commonly used methods are built using support vector machine (SVM), K-nearest neighbor (KNN) method and linear discriminant (LDA) [4]. The studies developed an automated classification system based on initial segmentation and extraction of features to describe the morphology or their arrangement. The study demonstrated a support vector machine based classifier to detect infected potato plants providing the accuracy of 89% [5]. The authors had used Support Vector Machines (SVM) technique for classifying various plant diseases and achieved high accuracies [6]. Chávez et al. proposed a method to detect foliar symptoms of potato with the usage of reflectance in the regions of spectrum of light reflection to be attained with a "trained" human eye [7]. While all mentioned classifiers were designed in previous studies, to classify manually selected regions of interest. Machine learning algorithms are in need of engineering in the field of feature selection or extraction and to organize the data with discriminative information [8]. Sometimes, the power of discriminative is the method which is inadequate to the competitive cost of identification of features for classification. Prior studies introduced a new dataset of diseased apple leaf images as the studies carried down previously were on small dataset. Among different approaches, the convolutional neural network is widely used in the current era of research for the recognition of patterns in different images problems [9]. In [10], proposed a new identification method for disease in rice plants using deep convolutional neural networks (DCNN). The dataset used contains 500 original images from fields of leaves of rice plant, diseased and healthy both. The network was trained for the identification of ten main diseases in rice plants. The average accuracy of 95.84% is attained by this proposed experiment. In [11] a CNN based technique is proposed to identify the chemical stained images of apple leaves and the technique used is a self-adaptive momentum rule for updating the parameters of CNN model. The accuracy result obtained by the experiment done is 96.08%. In [12], the new system is suggested for the identification of disease in leaves of cucumber plant using convolutional neural networks. The accuracy achieved by the proposed CNN system by using the fourfold cross-validation strategy is 94.9% in identification of cucumber leaf between disease leaves and healthy leaves. In [13], a model using deep convolutional neural networks for the detection of disease in plants. The different thirteen common plant type's diseases were identified using the proposed CNN model. The accuracy achieved by the proposed network is 96.3%. In [14], a CNN based model is designed for the detection of twenty six different types of diseases and crop species fourteen in number using the dataset which is publically available of healthy and diseased leaves. Also the study showed that convolutional neural networks are able to exceed outmoded texture descriptors. Moreover, only a little work is done in this field of research which has opened the doors to exploration by using deep learning methods to propose the identification of diseases in apple leaves automatically [15].

The key aim of this research is to propose a computerized model to examine the diseased images of apple leaves. This involves an algorithm which will automatically detect Marsonina Coronaria and Apple Scab from diseased apple images.

3 Proposed Methodology

Here, a CNN based automatic detection model system is developed for automatic classification of diseased apple leaf images into Marsonina Coronaria and Apple Scab. A research methodology has been proposed in the following sections which are divided into various stages as discussed below.

3.1 Dataset Preparation

The two most common diseases in apple leaves are Marsonina Coronaria and Apple Scab. Farmers usually detect these diseases through naked eye, leading to a lot of faulty detection of diseases. The dataset used for the experiment of identification of apple leaf diseases is taken from fields of Himachal Pradesh and Uttarakhand states of India. Live background images have been captured with a camera of 12 megapixels. This dataset consists of a total of 50,000 images of size $50 \times 50 \times 3$, in which 25,000 Marsonina Coronaria leaves and 25,000 Apple Scab leaves. Fig. 2 shows diseased apple leaf images of Marsonina Coronaria and Fig. 3 shows diseased apple leaf images of Apple Scab.

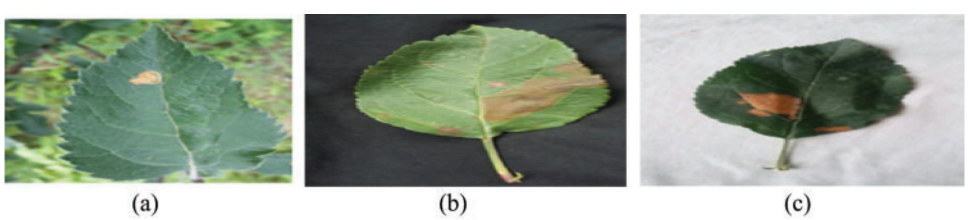

Figure 2: Marsonina diseased apple leaf images: (a) with Live Background, (b) with Black Background, (c) with White Background

3.2 Hardware and Software Used

In order to implement the proposed methodology, a database has been collected from fields of Himachal Pradesh and Uttarakhand states of India. A classification system is developed using Python 3.7 software containing libraries which are tensorflow and opencv that run on system with Intel® Core(TM) i3 processor with specifications configuration given in the Tab. 1.

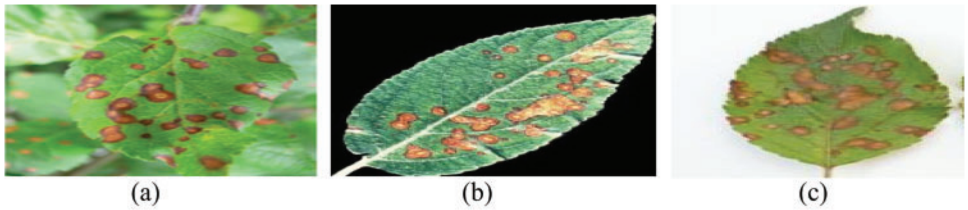

Figure 3: Apple scab diseased images (a) with Live Background, (b) with Black Background, (c) with White Background

Table 1: System configuration

Component	Specifications	
Laptop	Processor	Intel (R) Core (TM) i3-7100U CPU @ 2.40 GHz
	RAM	6 GB
	CPU	Inspiron 15
Software	Python	3.7
Data set collected	From fields of Himachal Pradesh and Uttarakhand	

3.3 Data Normalization

Normalization is performed on the given dataset to sustain value and its stability in the architecture of CNN and is a very important step as preprocessing stem in processing of the images. By using this step, proposed CNN models are likely to learn more rapidly. In this proposed research, the pixel value of the image at input side has been normalized in the range between 0–1. The dataset being used contains the images that are 0–1 scaled images and this rescaling of the images is done by dividing every pixel of image with 255. Each diseased apple image is converted into a single vector form of 7500x1. The machine is trained with the converted single vector form of every single image in the given training database [16–18]. After the training of the machine, testing is done with testing images.

3.4 Data Augmentation

The different CNN modes in deep learning works with larger number of dataset for the training of the architecture designed [19–21]. The images taken from the fields are very less in numbers of 50,000 that why the major concern is to increase the database as it's tough to arrange more number of images from the field which is a tedious task. To come up with this problem the technique data augmentation is performed which helps is increases the number of images by performing some transformation by keeping the class labels and pixel values same as original image [22–25]. The transformations made on the original images are illustrated in Fig. 4. The images were transformed using the various techniques as given: (1) images were scaled by a measure of 15%, (2) Gaussian noise was added with mean zero and 0.25 variance, (3) horizontal flipping was performed, and (4) original image was rotated by the angles of 30 degrees clockwise. All these techniques had been practical done on the given samples of training and some example results for each the proposed technique of data augmentation is shown in Fig. 5 for Apple Scab and in Fig. 6 for Marsonina Coronaria.

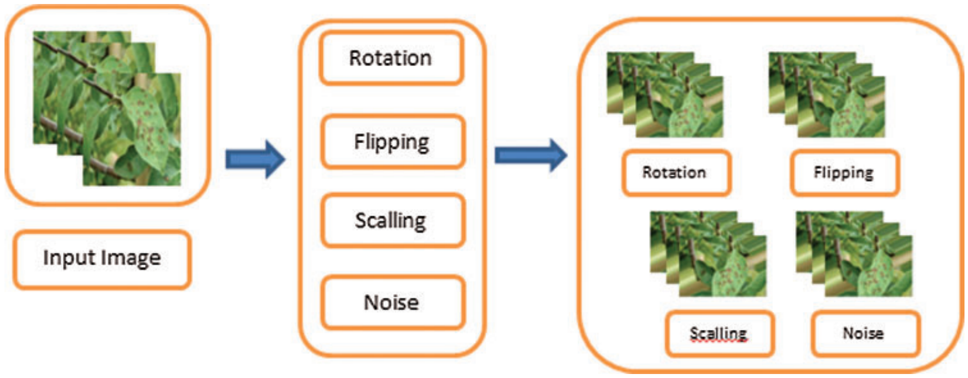

Figure 4: Different augmentation techniques used in this study

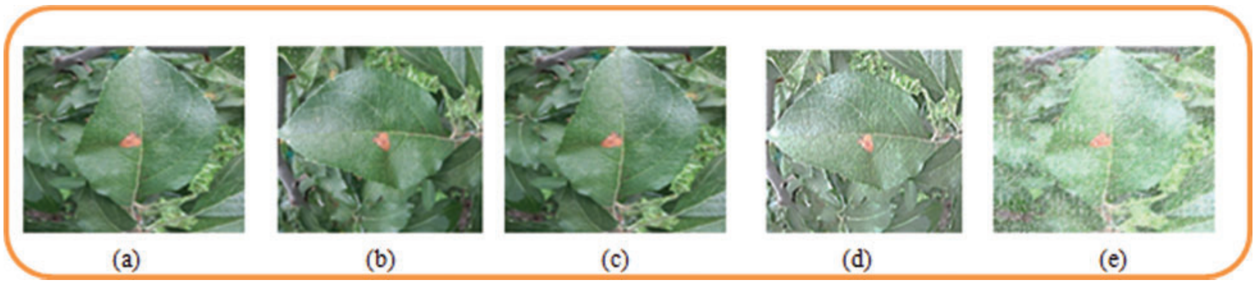

Figure 5: Result of data augmentation of apple scab image: (a) Original image; (b) Rotation clockwise; (c) Horizontal flipping; (d) Scaling by 15%; (e) Gaussian noise

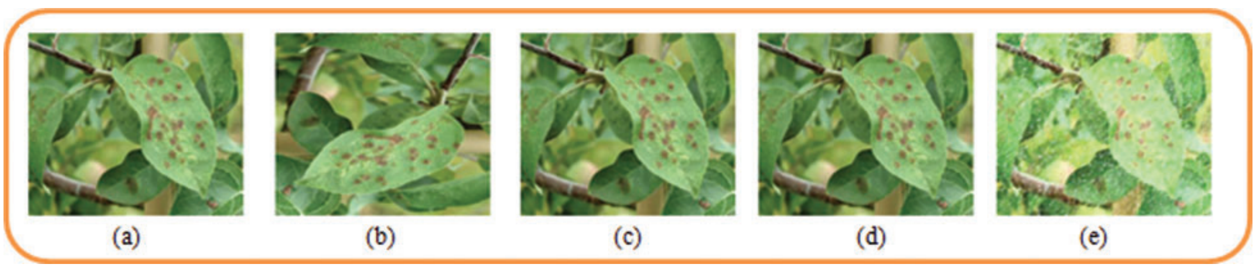

Figure 6: Result of data augmentation of marsonina coronaria image: (a) Original image; (b) Rotation clockwise; (c) Horizontal flipping; (d) Scaling by 15%; (e) Gaussian noise

3.5 Design of CNN Based Models

For the detection of Marsonina coronaria and Apple Scab diseases in apple leaves, a CNN based model is proposed which is divided in three categories; CNN-A Model, CNN-B Model and CNN-C model. CNN-A model consists of 8 layers whereas CNN-B model consists of 9 layers and CNN-C consists of a total of 19 layers. Following subsections describe these models in detail along with their block diagrams and the total number of parameters used for each layer.

3.5.1 Design of CNN-A Model

A block diagram representation of the CNN-A model is shown in Fig. 7 in which a total 8 layers are used. In CNN-A model, a convolutional block is present consisting of two (3 × 3) convolutional layers and a dropout (0.25). A fine-tuned fully connected layer is also present which consists of a flatten layer, dense_1 layer, dropout layer and dense_2 layer. Tab. 2 describes the size of filter, number of filters, input image size, output image size and total number of parameters used at each layer for the designed CNN-A.

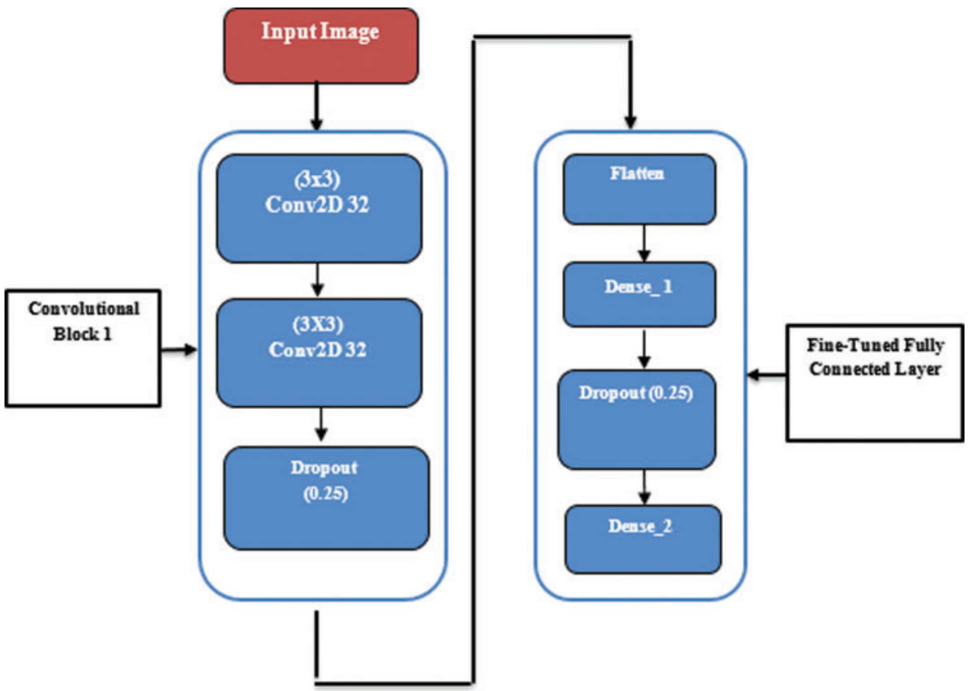

Figure 7: Convolutional neural network CNN-A with 8 layers

Table 2: Parameters of CNN-A with 8 layers

S. No.	Layers	Input image size	Filter size	No. of filter	Activation function	Output	Parameters
1	Input image	50 ∗ 50 ∗ 3	–	–	–	–	–
2	Convolutional	50 ∗ 50 ∗ 3	3 ∗ 3	32	ReLU	48 ∗ 48 ∗ 32	896
3	Convolutional	48 ∗ 48 ∗ 32	3 ∗ 3	64	ReLU	46 ∗ 46 ∗ 64	18496
4	Maxpooling	46 ∗ 46 ∗ 64	Poolsize (2 ∗ 2)	–		23 ∗ 23 ∗ 64	0
5	Dropout	23 ∗ 23 ∗ 64	Dropout (0.25)	–	–	23 ∗ 23 ∗ 64	0
6	Flatten	23 ∗ 23 ∗ 64	–	–	–	33856	0
7	Dense	33856	128		ReLU	128	4333696
8	Dropout	128	Dropout (0.5)	–	–	128	0
9	Dense	128	Num_Clas		Softmax	2	258

3.5.2 *Design of CNN-B Model*

The block diagram representation of CNN-B model is shown in Fig. 8 which consists of a total 9 layers. The detailed description of the model is given in Tab. 3, which describes the size of filter, number of filters, input image size, output image size and total number of parameters used at each layer. In this model, a maxpool layer is added after the first convolution layer and a fine-tuned fully connected layer is also present which consists of a flatten layer, dense_1 layer, dropout layer and dense_2 layer.

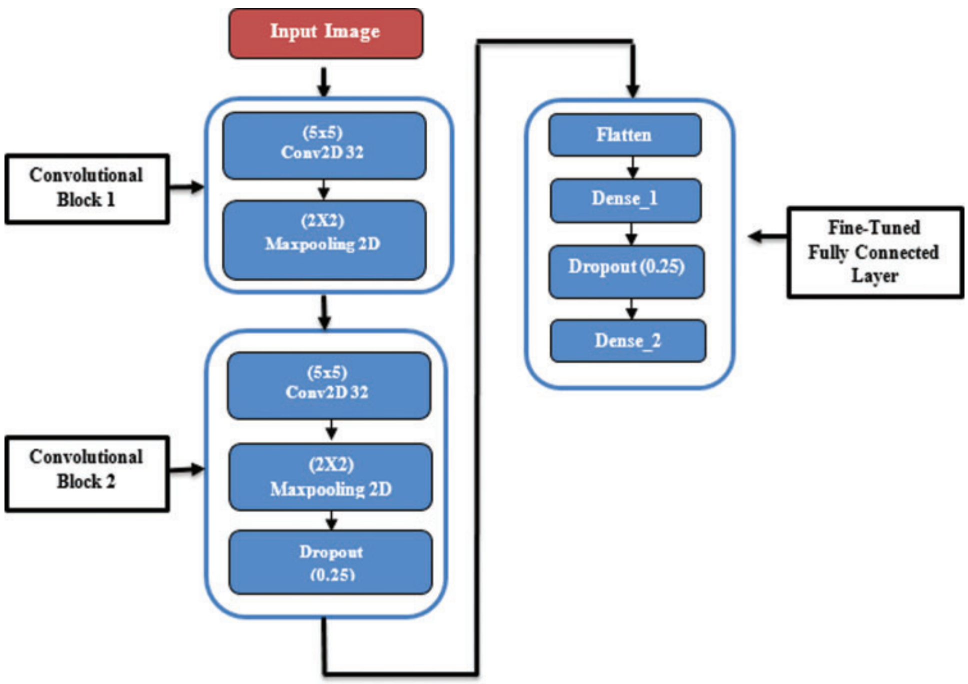

Figure 8: Convolutional neural networks CNN-B with 9 layers

Table 3: Parameters of CNN-B with 9 layers

S. No.	Layers	Input image size	Filter size	No. of filter	Activation function	Output	Parameters
1	Input image	50 ∗ 50 ∗ 3	–	–	–	–	–
2	Convolution	50 ∗ 50 ∗ 3	5 ∗ 5	32	ReLU	46 ∗ 46 ∗ 32	2432
3	Maxpooling	46 ∗ 46 ∗ 32	3 ∗ 3	–	–	15 ∗ 15 ∗ 32	0
4	Convolution	15 ∗ 15 ∗ 32	5 ∗ 5	32	ReLU	11 ∗ 11 ∗ 32	25632
5	Maxpooling	11 ∗ 11 ∗ 32	3 ∗ 3	–	–	3 ∗ 3 ∗ 32	0
6	Dropout	3 ∗ 3 ∗ 32	0.25	–	–	3 ∗ 3 ∗ 32	0
7	Flatten	3 ∗ 3 ∗ 32	–	–	–	288	0
8	Dense	288	64	–	ReLU	64	18496
9	Dropout	64	0.5	–	–	64	0
10	Dense	64	2	–	Sigmoid	2	130

3.5.3 Design of CNN-C Model

The block diagram representation of CNN-C model is shown in Fig. 9 in which a total 19 layers are used whose detailed description is given in Tab. 4. This table describes the size of filter, number of filters, input image size, output image size and total number of parameters used at each layer. CNN-C model consists of three convolution blocks and a fine tuned fully connected layer. The size of the filter is also changed from 32 to 86 in this model.

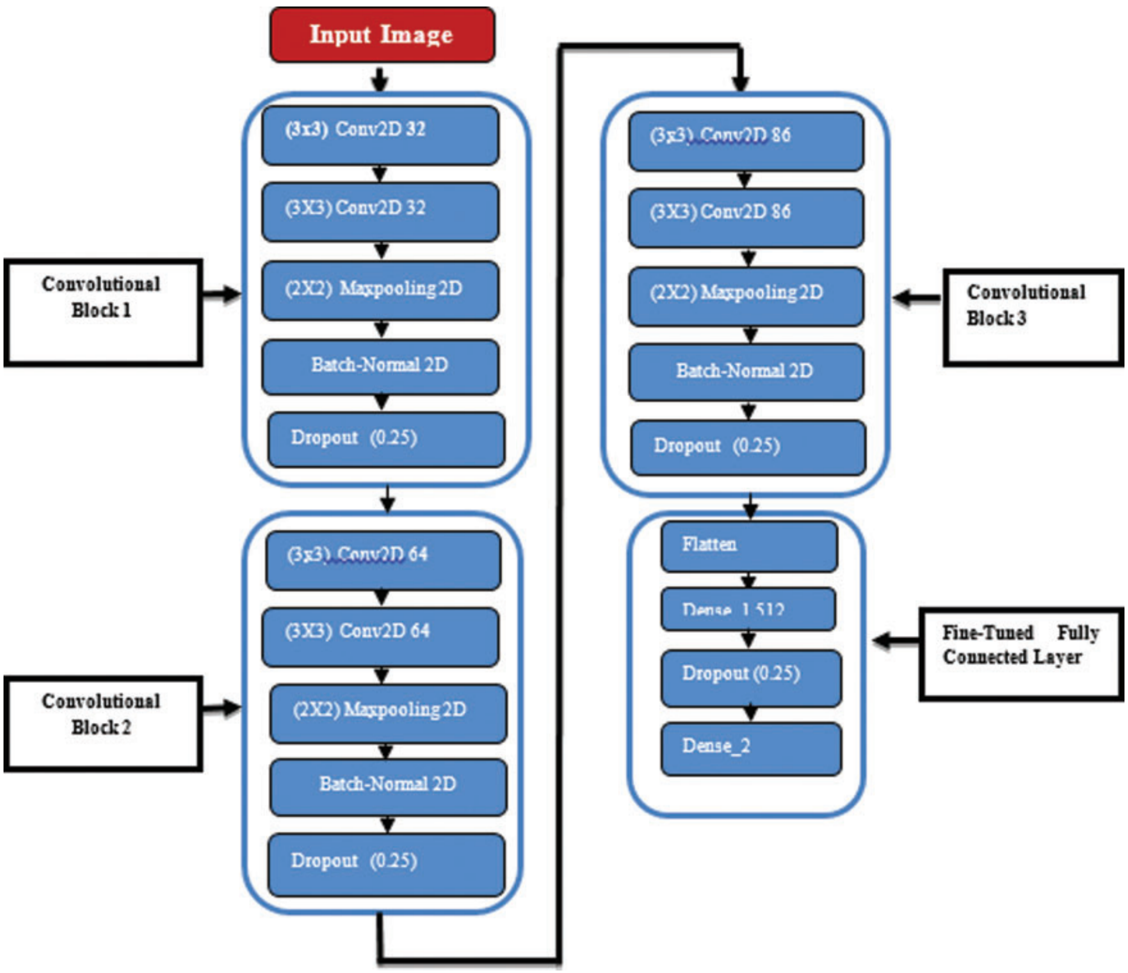

Figure 9: Convolutional neural networks CNN-C with 19 layers

Table 4: Parameters of CNN-C with 19 layers

S. No.	Name of layers	Input image size	Filter size	No. of filter	Activation function	Output with padding	Parameter
1	Input image	50 ∗ 50 ∗ 3	–	–	–	–	–
2	Convolutional	50 ∗ 50 ∗ 3	3 ∗ 3	32	ReLU	50 ∗ 50 ∗ 32	896
3	Convolutional	50 ∗ 50 ∗ 32	3 ∗ 3	32	ReLU	50 ∗ 50 ∗ 32	9248
4	Maxpooling	50 ∗ 50 ∗ 32	2 ∗ 2	–	–	25 ∗ 25 ∗ 32	0
5	Batch normalization	25 ∗ 25 ∗ 32	–	–	–	25 ∗ 25 ∗ 32	128

(Continued)

Table 4: Continued

S. No.	Name of layers	Input image size	Filter size	No. of filter	Activation function	Output with padding	Parameter
6	Dropout	$25 * 25 * 32$	0.25	–	–	$25 * 25 * 32$	0
7	Convolutional	$25 * 25 * 32$	$3 * 3$	64	ReLU	$25 * 25 * 64$	18496
8	Convolutional	$25 * 25 * 64$	$3 * 3$	64	ReLU	$25 * 25 * 64$	36928
9	Maxpooling	$25 * 25 * 64$	$2 * 2$	–	–	$12 * 12 * 64$	0
10	Batch normalization	$12 * 12 * 64$	–	–	–	$12 * 12 * 64$	256
11	Dropout	$12 * 12 * 64$	0.25	–	–	$12 * 12 * 64$	0
12	Convolutional	$12 * 12 * 64$	$3 * 3$	86	ReLU	$12 * 12 * 86$	49622
13	Convolutional	$12 * 12 * 86$	$3 * 3$	86	ReLU	$12 * 12 * 86$	66650
14	Maxpooling	$12 * 12 * 86$	$2 * 2$	–	–	$6 * 6 * 86$	0
15	Batch normalization	$6 * 6 * 86$	–	–	–	$6 * 6 * 86$	344
16	Dropout	$6 * 6 * 86$	0.25	–	–	$6 * 6 * 86$	0
17	Flatten	$6 * 6 * 86$	–	–	–	3096	0
18	Dense	3096	512	–	ReLU	512	1585664
19	Dropout	512	0.25	–	–	512	0
20	Dense	512	2		ReLU	2	1026

4 Results and Discussion

The various results obtained from the experiments are discussed in this section. Firstly, the different proposed CNN models are applied on the dataset, and the experiment results are obtained and then machine learning techniques are applied on the dataset, and the experiment results are obtained. Secondly, proposed CNN network results are compared with the best machine learning classifier for computerized detection of diseased apple leaf images. Lastly, the CNN-algorithm is compared with the previous literature.

4.1 Performance Analysis of Different CNN Models

Investigational results of the designed CNN networks i.e., CNN-A, CNN-B and CNN-C for identification of diseases in apple leaves is shown in the following section. The performance of these networks is examined in terms of accuracy, sensitivity and specificity. The three different models proposed convolution neural networks are trained with different numbers of images. Every experiment is computed with the library of scikit learn. Following subsections shows the parameter analysis of the three designed CNN's model. These parameters are analyzed on four different sets of images i.e., 50 K, 100 K, 150 K and 200 K. 16 epochs are used for training each model with different parameters.

4.1.1 Results for CNN-A Model

Here the CNN-A model is trained on 4,353,346 parameters with 16 epochs. During simulation time taken per epoch is at least 72 seconds for 50 K images whereas 362 seconds for 200 K images with 16 epochs. The performance analysis is done with different sizes of dataset having 50 K, 100 K, 150 K and 200 K respectively shown in Tab. 5. From the table, it is analyzed that highest accuracy and sensitivity of 87.2% and 87% respectively is obtained with 150 K training images whereas the highest specificity is achieved of 91.4% with 200 K images.

Table 5: Parameters analysis of CNN-A with 16 epochs

	No. of images			
Metrics	50 K	100 K	150 K	200 K
Accuracy	0.851	0.856	**0.872**	0.867
Sensitivity	0.840	0.813	**0.870**	0.833
Specificity	0.857	0.868	0.874	**0.914**

4.1.2 Results for CNN-B Model

Here the CNN-B model is trained on 46,690 parameters with 16 epochs. During simulation time taken per epoch is at least 72 seconds for 50 K images whereas 182 seconds for 200 K images with 16 epochs. The performance analysis is done with different sizes of dataset having 50 K, 100 K, 150 K and 200 K respectively shown in Tab. 6. From the table, it is analyzed that highest accuracy and specificity of 88.7% and 81.8% respectively is obtained with 200 K training images whereas the highest sensitivity is achieved of 93.5% with 150 K images.

Table 6: Parameters analysis of CNN-B with 16 epochs

	No. of images			
Metrics	50 K	100 K	150 K	200 K
Accuracy	0.844	0.867	0.871	**0.887**
Sensitivity	0.889	0.911	**0.935**	0.926
Specificity	0.815	0.759	0.775	**0.818**

4.1.3 Results for CNN-C Model

Here the CNN-C model is trained on 1,769,258 parameters with 16 epochs. During simulation time taken per epoch is at least 89 seconds for 50 K images whereas 442 seconds for 200 K images with 16 epochs. The performance analysis is done with different sizes of dataset having 50 K, 100 K, 150 K and 200 K respectively is given in Tab. 7. From the table, it is analyzed that the highest accuracy, sensitivity and specificity of 99.2%, 99.7% and 97.8% respectively is obtained with 200 K training images.

Table 7: Parameter analysis of CNN-C with 16 epoch

	No. of images			
Metrics	50 K	100 K	150 K	200 K
Accuracy	0.889	0.926	0.964	**0.992**
Sensitivity	0.889	0.911	0.986	**0.997**
Specificity	0.815	0.859	0.839	**0.978**

4.1.4 Comparative Analysis of Three Different CNN Models

Tab. 8 shows comparison of the performance parameter values and the architecture of proposed CNN models CNN-A, CNN-B and CNN-C. The graphical representation is shown in Fig. 10. Accuracy is obtained by division of the total correctness in predictions upon the total given predictions. As given in table, the highest accuracy, sensitivity and specificity of 99.2%, 99.7% and 97.8% is obtained with CNN-C model as compared to the CNN-B and CNN-A. It is seen that deeper the convolutional neural networks architecture better are the results when compared on the basis of layers.

Table 8: Comparison result of different CNN

	CNN		
Layers	CNN-A	CNN-B	CNN-C
EPOCH	16	16	16
Convolution	2	2	6
Maxpooling	1	1	3
Dropout	2	2	4
Batch normalization	–	–	3
Flatten	1	1	1
Dense	2	2	2
Total layers	8	9	19
Accuracy	0.872	0.887	**0.992**
Sensitivity	0.87	0.926	**0.997**
Specificity	0.874	0.818	**0.978**

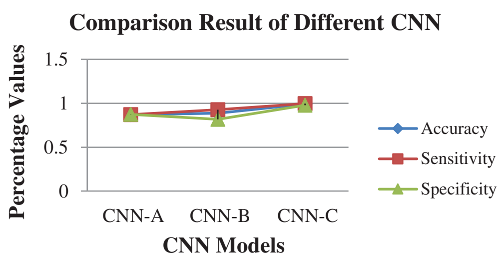

Figure 10: Comparison result of different CNN

4.1.5 Misclassification Result Analysis for the best CNN-C Model

The samples which are misclassified by best CNN model CNN-C is shown in Fig. 11. The misclassification probably happens because of similarity in features images of the Marsonina and Apple Scab.

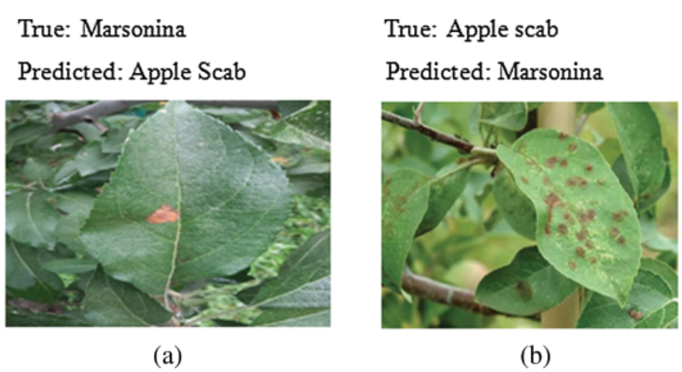

Figure 11: (Continued)

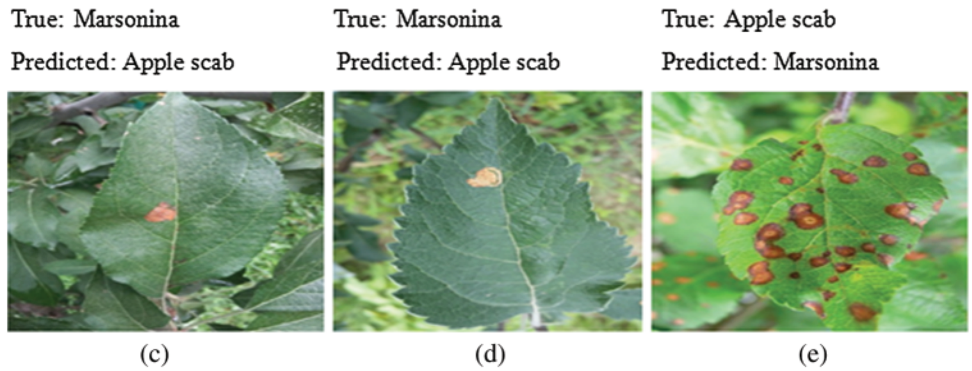

Figure 11: Misclassification results obtained by CNN-C model

4.2 Performance Analysis with Machine Learning (ML) Algorithms

For the result analysis of machine learning algorithms, four different classifiers are implemented on the same dataset. The classifiers considered are SVM, KNN, Logistic Regression and Random Forest. A comparative analysis is done among machine learning and designed CNN models. Finally, the best architecture model is achieved. Initially the different classifiers were analyzed on 24000 training and 6000 testing images.

4.2.1 Results for KNN Classifier

The prepared dataset is analyzed with commonly used classifier KNN. For this, three values of K i.e., 3, 5 and 7 have been taken and their performance has been analyzed in terms of accuracy, classification error, sensitivity, specificity, false positive rate and precision. The result of KNN classifiers with three values of K as 3, 5 and 7 are shown in Tab. 9. From the table it can be analyzed that KNN performs better in terms of sensitivity of 81.7% with K value as 7. The other parameters like accuracy, specificity and precision are almost the same for three values of K. But their results are not satisfactory if compared with the proposed three CNN models.

Table 9: KNN parameter analysis with 6000 testing images

Parameters	K = 3	K = 5	K = 7
Accuracy	0.778	0.78	0.792
Classification error	0.122	0.12	0.108
Sensitivity	0.758	0.782	**0.817**
Specificity	0.791	0.778	0.774
False positive rate	0.108	0.121	0.125
Precision	0.712	0.706	0.711

4.2.2 Results for SVM Classifier

The SVM classifier is analyzed with different kernels of SVM like Sigmoid, Linear and RBF. These kernels are taken for analyzing the performance of a classifier in terms of accuracy, classification error, sensitivity, specificity, false positive rate and precision as shown in Tab. 10.

Table 10: SVM parameter analysis with 6000 testing images

Parameters	Sigmoid	Linear	RBF
Accuracy	**0.820**	0.801	0.713
Classification error	0.22	**0.199**	0.287
Sensitivity	**0.854**	0.725	0.641
Specificity	0.729	0.852	0.761
False positive rate	0.270	0.147	0.238
Precision	0.682	**0.769**	0.646

From the table it is seen that sensitivity of 85.4% is achieved using sigmoid kernel. Also, an accuracy of 82% with sigmoid kernel, classification error of 19.9% with linear kernel and finally precision of 76.9% with linear kernel is obtained. It can be analyzed that the sigmoid kernel performs better in terms of accuracy and sensitivity but when compared to the proposed three CNN models, their results are unsatisfactory.

4.2.3 Results for Random Forest Classifier

Random forest classifier is analyzed with different values of N-estimator i.e., 1000, 1500, 2000 as shown in Tab. 11. From the table it is seen that N-estimator (2000) shows the maximum sensitivity of 80.4% and N-estimator (1500) shows maximum accuracy of 84.3%, whereas classification error decreased to 15.7% with N-estimator (1000). It is also analyzed that precision is highest in case of N-estimator (1500). Again, these results are not satisfactory if compared with the proposed three CNN models.

Table 11: Random forest parameter analysis with 6000 testing images

Parameter	N_estimator (1000)	N_estimator (1500)	N_estimator (2000)
Accuracy	0.827	**0.843**	0.839
Classification error	0.173	**0.157**	0.161
Sensitivity	0.772	0.802	**0.804**
Specificity	0.863	0.870	0.862
False positive rate	0.136	0.129	0.137
Precision	0.794	**0.808**	0.799

4.2.4 Results for Logistic Regression

The prepared dataset of 6000 testing images is analyzed with a logistic regression classifier taking different parameters of logistic regression like binomial, multinomial and ordinal as shown in Tab. 12. From the table it is seen that maximum sensitivity of 77% is obtained with multinomial and ordinal kernels. Also accuracy of 82% and highest precision of 78% is achieved with ordinal. This parameter also gives least classification error of 18% making it the best logistic regression technique as compared with the other two logistic regression parameters but when compared to the proposed three CNN models, their results are unsatisfactory.

Table 12: Logistic regression performance metrics with 6000 testing images

Parameter	Binomial	Multinomial	Ordinal
Accuracy	0.789	0.819	**0.82**
Classification error	0.211	0.181	**0.18**
Sensitivity	0.727	0.770	**0.772**
Specificity	0.834	0.852	0.852
False positive rate	0.167	0.147	0.147
Precision	0.755	0.78	**0.780**

4.2.5 Comparative Analysis Result of four ML Classifiers

A combined performance analysis of the classifier is given in Tab. 12, where parameters have the maximum sensitivity for each classifier. Sensitivity and accuracy plays an important role in disease detection of apple leaves. Sensitivity is the measure that tells what proportion of images that actually had diseases, how sensitive the classifier is in classifying apple diseased images. Wherein, the accuracy gives value of correctness in the estimation made by the designed algorithm. The results from Tabs. 9 to 12 shows that sigmoid kernel is considered for SVM, N-estimator (1000) for RF, ordinal for LR classifier and value of K i.e., 7 for KNN classifier given the good result of accuracy and sensitivity. But, out of all the classifiers; the highest sensitivity of 85% is shown by the SVM classifier and accuracy of 83.9% is shown by RF classifier as shown in Tab. 13.

Table 13: Performance metrics of different classifiers

Performance metrics	SVM	KNN	LR	RF
Accuracy	0.820	0.692	0.82	0.839
Classification error	0.22	0.108	0.18	0.161
Sensitivity	**0.854**	0.717	0.772	0.804
Specificity	0.729	0.674	0.852	0.862
False positive rate	0.270	0.125	0.147	0.137
Precision	0.682	0.611	0.780	0.799

4.3 Comparison of Best ML Classifier with Best Proposed CNN-C Model

Performance analysis is used to examine the efficiency of classification of the designed CNN-A model which outperforms machine learning algorithms like SVM by applying the same dataset to both the techniques. The analysis of both the techniques is done specifically on the basis of accuracy and sensitivity, as these are the most important parameters for classification purposes. Sensitivity shows the proportion of images that actually had diseases was also predicted of having disease or one can say, how sensitive the classifier is in classifying diseased apple leaf images. Wherein, the accuracy gives the accurate predictions made by the predictive model over the rest of the predictions. The results of classification using machine and deep learning techniques is shown in Tab. 14.

Here, the number of dataset taken is varied from 50 K to 200 K and it is seen that with CNN-C model the accuracy of 99.2% is obtained when the number of images taken is 200 K.

Overall, the deep learning technique gives better performance as compared to machine learning techniques.

Table 14: Parameter analysis with varying number of images

Number of images	SVM		CNN-C	
	Accuracy	Sensitivity	Accuracy	Sensitivity
50000	0.842	0.822	0.889	0.889
100000	0.851	0.766	0.926	0.911
150000	0.847	0.797	0.964	0.986
200000	0.839	0.790	0.992	0.997

4.4 Comparison of Best Proposed CNN-C Model with State of Art

When compared with accuracy, the designed 19 layer CNN-C outperforms previous state-of-art. The proposed model has improved the accuracy when compared with other research as a bigger dataset of diseased apple leaves is used and 99.2% of accuracy is being achieved. The table clearly, shows that the proposed model gives the improved output results than the previous work. The comparison of the proposed model accuracy with other research work is depicted in the Tab. 15.

Table 15: Results of previous work using deep learning

Paper ID	No. of images	Classifiers	Classification	Accuracy
[7]	9164	CNN	Apple Scab and Rust Black Rot	88.9%
[8]	2462	DenseNet-121	Six different Apple Diseases	93.71%
[9]	26377	CNN	Alternaria Leaf Spot, Brown Spot and Rust	78.80%
[10]	–	V3 Deep Architecture	Benign and Malignant	97%
[11]	–	CNN	Marsonina Blotch	84.3%
[12]	2561	CNN	Apple Scab and Apple Black Rot	98.54%
[13]	404	LSA-Net	Alternaria and Marsonina Blotch	89.4%
[14]	2446	CNN	Apple Scab and Frog Eye Spot	99.01%
[15]	13689	CNN	Brown Spot, Alternaria and Rust	97.62%
Proposed algorithm	200000	CNN	Marsonina Coronaria and Apple Scab	99.2%

5 Conclusion

Apple is a highly essential product for India's economy, and it has played a significant role in our economy in the past. However, due to numerous diseases harming the crop, it is not as abundant as it once was. The key goal of the research is to detect diseases like Marsonina Coronaria and apple scab in apple leaves using automatic digital system. Firstly, three different convolutional CNN networks are proposed for computerized detection of diseased apple leaf images; secondly, machine learning techniques used for CAD systems, are applied on the dataset, which can help detect the apple diseases. Lastly, the CNN model is compared with the already existing machine learning techniques and also with the previous literature. The focus is on exploring how the developed system performs when various performance parameters are implemented, thus providing a better accuracy and sensitivity than the already existing techniques. The first stage of the classification is done by convolution neural network in which CNN-C gives best result, and then the classification is done by machine learning technique in which SVM classifier gives the best result. Later various performance measures like accuracy, sensitivity, specificity, false positive rate, classification error and precision are evaluated for machine learning. Out of the performance measures, accuracy and sensitivity is chosen as the main parameter to judge the best classifier. Accuracy of 99.2% and sensitivity 99.7% is obtained by taking 200000 histopathology images using proposed CNN model C.

This work successfully exhibits better classification accuracy and sensitivity for the classification of apple leaf diseases images using 19 layer convolutional neural networks to classify Marsonina Coronaria and Apple Scab diseases images of apple leaves. The usage of various machine learning techniques and convolutional neural networks leads to better classification accuracy. Apple diseases like alternaria, fire blight, cork spot, powdery mildew, black rot and phytophthora rot may be considered for future disease detection. For real time solutions, there is a need to develop mobile based applications which will guide the farmers to detect the disease in apple leaves on their own. Further other plant diseases can also be taken for detection and identification of diseases using CNN models and transfer learning models.

Funding Statement: This work was supported by Taif University Researchers Supporting Project (TURSP) under number (TURSP-2020/73), Taif University, Taif, Saudi Arabia.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] Z. Chuanlei, Z. Shanwen, Y. Jucheng, S. Yancui and C. Jia, "Apple leaf disease identification using genetic algorithm and correlation," *International Journal of Agricultural and Biological Engineering*, vol. 10, no. 2, pp. 74–83, 2017.

[2] Z. Peng, L. Gang, L. Minzan and L. Daoliang, "Management information system for apple diseases and insect pests based on GIS," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 22, no. 2, pp. 150–154, 2006.

[3] Y. Atoum, M. J. Afridi, X. Liu, J. M. McGrath and L. E. Hanson, "On developing and enhancing plant-level disease rating systems in real fields," *Pattern Recognition*, vol. 53, no. 5, pp. 287–299, 2016.

[4] P. R. Rothe and R. V. Kshirsagar, "Cotton leaf disease identification using pattern recognition techniques," in *Proc. of the 2015 Int. Conf. on Pervasive Computing (ICPC)*, Pune, India, pp. 1–6, 2015.

[5] P. B. Padol and A. A. Yadav, "SVM classifier based grape leaf disease detection," in *Proc. of the 2016 Conf. on Advances in Signal Processing (CASP)*, Pune, India, pp. 175–179, 2016.

[6] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed *et al.,* "Going deeper with convolutions," in *Proc. of the 2015 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, pp. 1–9, 2015.

[7] S. S. Gaikwad and S. Singh, "Fungi classification using convolutional neural network," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 10, pp. 4563–4569, 2021.

[8] Y. Zhong and M. Zhao, "Research on deep learning in apple leaf disease recognition," *Computers and Electronics in Agriculture*, vol. 168, no. 2, pp. 105146, 2020.

[9] P. Jiang, Y. Chen, B. Liu, D. He, C. L. Jiang *et al.,* "Real-time detection of apple leaf diseases using deep learning approaches based on improved convolutional neural networks," *IEEE Access*, vol. 7, no. 2, pp. 59069–59080, 2019.

[10] M. B. Tahir, M. A. Khan, K. Javed, S. Kadry, Y. D. Zhang *et al.,* "Recognition of apple leaf diseases using deep learning and variances-controlled features reduction," *Microprocessors and Microsystems*, vol. 7, no. 1, pp. 104–127, 2021.

[11] Y. H. Jin, S. C. Hwan and L. D. Hyuk, "Apple leaf disease identification through region-of-interest-aware deep convolutional neural networks," *Journal of Imaging Science and Technology*, vol. 64, no. 2, pp. 20507-1, 2020.

[12] S. Baranwal, S. Khandelwal and A. Arora, "Deep learning convolutional neural network for apple leaves disease detection," in *Proc. of Int. Conf. on Sustainable Computing in Science, Technology and Management (SUSCOM)*, Jaipur-India, Amity University Rajasthan, 2020.

[13] H. J. Yu and C. H. Son, "Leaf spot attention network for apple leaf disease identification," in *Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition Workshops*, China, pp. 52–53, 2020.

[14] Q. Yan, B. Yang, W. Wang, B. Wang, P. Chen *et al.,* "Apple leaf diseases recognition based on an improved convolutional neural network," *Sensors*, vol. 12, no. 2, pp. 3535–3541, 2020.

[15] B. Liu, Y. Zhang, D. J. He and Y. Li, "Identification of apple leaf diseases based on deep convolutional neural networks," *Symmetry*, vol. 10, no. 1, pp. 11–16, 2018.

[16] P. Bansal, R. Kumar and S. Kumar, "Disease detection in apple leaves using deep convolutional neural network," *Agriculture*, vol. 11, no. 7, pp. 617–620, 2021.

[17] I. Gupta, S. Gupta and S. Singh, "Breast cancer detection using deep learning and machine learning techniques," *Journal of Advanced Research in Dynamical & Control Systems*, vol. 11, no. 2, pp. 1487–1498, 2019.

[18] S. Singh and S. Gupta, "Apple scab and marsonina coronaria diseases detection in apple leaves using machine learning," *International Journal of Pure and Applied Mathematics*, vol. 118, no. 5, pp. 1151–1166, 2018.

[19] S. Singh and S. Gupta, "Digital image processing techniques for early detection and classification of different diseased plants," *International Journal of Bio-Science and Bio-Technology*, vol. 8, no. 4, pp. 61–66, 2016.

[20] I. Gupta, S. Gupta and S. Singh, "Invasive ductal carcinoma in histopathology images using machine learning techniques without deploying feature extraction," *Journal of Computational and Theoretical Nanoscience*, vol. 7, no. 16, pp. 2589–2595, 2020.

[21] S. Mahajan, A. Raina, X. Z. Gao and A. K. Pandit, "Plant recognition using morphological feature extraction and transfer learning over SVM and AdaBoost," *Symmetry*, vol. 13, no. 2, pp. 356–363, 2021.

[22] J. Lu, L. Tan and H. Jiang, "Review on convolutional neural network applied to plant leaf disease classification," *Agriculture*, vol. 11, no. 8, pp. 707–710, 2021.

[23] S. Uguz and N. Uysal, "Classification of olive leaf diseases using deep convolutional neural networks," *Neural Computing and Applications*, vol. 33, no. 9, pp. 4133–4149, 2021.

[24] N. Kaur and V. Devendran, "Plant leaf disease detection using ensemble classification and feature extraction," *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 11, pp. 2339–2352, 2021.

[25] S. K. Kim and J. G. Ahn, "Tomato crop diseases classification models using deep CNN-based architectures," *Journal of the Korea Academia-Industrial Cooperation Society*, vol. 22, no. 5, pp. 7–14, 2021.