IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 27 February 2024, accepted 11 March 2024, date of publication 19 March 2024, date of current version 26 March 2024.

*Digital Object Identifier 10.1109/ACCESS.2024.3378262*

RESEARCH ARTICLE

# Medicinal Plant Classification Using Particle Swarm Optimized Cascaded Network

**MD. TAREQUL ISLAM[1,2], WAHIDUR RAHMAN[2,3], MD. SHAKHAWAT HOSSAIN[4], (Member, IEEE), KANIZ ROKSANA[3], IRMA DOMÍNGUEZ AZPÍROZ[5,6,7], RAQUEL MARTÍNEZ DIAZ[5,8,9], IMRAN ASHRAF[10], AND MD. ABDUS SAMAD[10], (Member, IEEE)**

[1]Department of Computer Science and Engineering, Khwaja Yunus Ali University, Sirajganj 6751, Bangladesh
[2]Department of Computer Science and Engineering, Mawlana Bhashani Science and Technology University, Tangail 1902, Bangladesh
[3]Department of Computer Science and Engineering, Uttara University, Dhaka 1230, Bangladesh
[4]Department of Computer Science and Engineering, Independent University, Bangladesh, Dhaka 1229, Bangladesh
[5]Universidad Europea del Atlántico, 39011 Santander, Spain
[6]Universidad Internacional Iberoamericana, Arecibo, PR 00613, USA
[7]Universidade Internacional do Cuanza, Cuito, Bié, Angola
[8]Universidad Internacional Iberoamericana, Campeche 24560, Mexico
[9]Universidad de La Romana, La Romana, Dominican Republic
[10]Department of Information and Communication Engineering, Yeungnam University, Gyeongsan-si 38541, South Korea

Corresponding authors: Md. Abdus Samad (masamad@yu.ac.kr) and Imran Ashraf (imranashraf@ynu.ac.kr)

This work was supported by European University of the Atlantic.

**ABSTRACT** Medicinal plants are essential to healthcare since ancient times and are integral to developing drugs and other medical treatments. More than 25% of medicines in developed countries are produced from medicinal plants, while in developing countries, approximately 80% of individuals receive primary healthcare from these plants. Traditionally, these plants are identified manually by experts, which is tedious, time-consuming, subjective and dependent on the availability of experts. Furthermore, a wrong detection can result in serious health issues or death. This signifies the need for a more reliable approach to identifying medicinal plants, which is accurate and practical. Several automated methods were proposed previously, utilizing deep learning and traditional machine learning (TML) techniques, but they require singular leaf images and failed to achieve sufficient accuracy when demonstrated in a different setting. Capturing singular leaf images for each plant is also time-consuming and laborious. This paper presents a robust, accurate and practical system to identify medicinal plants from smartphone-captured plant images in the site of plants. The proposed system utilized a cascaded architecture to extract features using a pre-trained ResNet50 model, which were optimized using Particle Swarm Optimization (PSO) to classify the plants using a Support Vector Machine (SVM). The proposed ResNet50-PSO-SVM network classified seven medicinal plants with 99.60% accuracy, outperforming the state-of-the-art (99%). The system was demonstrated for three different smartphones, classifying an image in 0.15 seconds with 97.79% accuracy on average. The system's high accuracy, rapid identification time and robustness ensured its practical use.

**INDEX TERMS** Medicinal plants, particle swarm optimization, feature selection, cascaded network, medicinal plant classification.

## I. INTRODUCTION

Medicinal plants were the primary source of medicines in ancient times and still are globally valuable sources of new drugs. However, the process of using medicinal plant compounds has changed over time. Modern pharmaceutical scientists extract essential components from medicinal plants and use them to produce medicines. Many modern medicines have been produced from bioactive compounds found in medicinal plants, for example, aspirin from willow bark, morphine from opium poppy, and quinine from cinchona bark. Thus, medicinal plants are a valuable resource for

The associate editor coordinating the review of this manuscript and approving it for publication was Alberto Cano.

© 2024 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

modern drug discovery. According to the World Health Organization (WHO), 60% of the population worldwide relies on medicinal plant-based medicines [1]. In the United States, around 80 out of 100 prescribed medicines are produced mainly or partly from medicinal plants [2]. This is even higher in underdeveloped countries. On top of that, the demand for medicinal plants is increasing rapidly throughout the world.

An estimated 350,000 medicinal plants exist worldwide, which is 10% of all vascular plants [3]. Thus, identifying the medicinal plants is a challenging task. In Bangladesh, more than 116 types of medicinal plants are reported by the experts [4]. Traditionally, these plants are identified based on the naked eye observation of experts. Experts mainly observe the shape, color, texture and aroma of leaves and flowers to identify a plant. However, this manual examination is not practical as it takes a lot of time, is prone to fatigue and introduces inter-observer variability. Many researchers have proposed automated methods for medicinal plant identification, motivated by the latest advancements in computer vision and machine learning techniques. These medicinal plant classification methods were proposed for mainly two purposes: detecting leaf disease in medicinal plants and classifying the type of medicinal plants for drug development or treatment. In this study, we mainly focused on classifying medical plants; thus, we have only reviewed the relevant papers proposed to classify medicinal plants [5], [6], [7], [8], [9], [10], [11], [18], [19], [20], [21], [22], [23]. The major limitation of the existing medicinal plant classification methods is their dependency on singular leaf images captured from a close distance, usually less than a meter. Capturing distinct leaf images for each plant is tedious and time-consuming, compromising the primary objective of developing an automated detection system. These methods also fail to classify the plants from plant images. Moreover, they fail when the leaf images are captured using a different device or from a distance. These methods work when the input image contains a singular leaf taken from a close distance. Most of these methods can not provide classification results on the site of examination as they process the captured images in a different location where the method is implemented in a computer. Therefore, the existing methods are not suitable for practical use. Another limitation is their robustness. Most of the methods were tested using homogeneous data. Consequently, they failed when the images were captured using a different device. Additionally, the deep learning-based methods are highly parameterized and computationally heavy, challenging to implement in a low-computing machine such as a smartphone. This paper proposed a cascaded network combining a pre-trained CNN, PSO and SVM to classify the medicinal plants. This method can accurately predict the plant class from plant images taken from a normal distance, regardless of the imaging device. This method is lightweight yet robust.

The major contributions of this paper are listed as follows: 1) the development of a cascaded network for medicinal plant classification in the site of plants from smartphone-captured plant images, 2) an analysis of the pre-trained CNN-PSO-traditional classifier-based cascaded approach for the classification of medicinal plants, 3) an ablation study to evaluate the contribution of PSO in the cascaded network and 4) the assessment of the proposed system in terms of robustness, rapidness and facility to provide results in the site of plants to ensure its practical application.

## II. RELATED WORKS

Several image-based automated methods were proposed to classify medicinal plants. These methods can be broadly classified into three categories: 1) manually selected features with traditional machine learning (TML) or convolutional neural network (CNN) classifiers, 2) CNN-based feature selection and classification and 3) cascade of multiple TML or CNN models. Typically, these methods targeted the local medicinal plants and fine-tuned the methods for classifying them.

Kan et al. proposed a method by manually selecting ten shapes and five texture features to classify 12 medicinal plants using SVM [5]. They collected 240 leaf images from China for the 12 plants. This method achieved an accuracy of 93.3%. Another method based on leaf shape and texture features was proposed by Janani and Gopal [6]. They collected 63 leaf images for six plants and trained a custom-built artificial neural network (ANN) to classify the plants, achieving 94.4% accuracy. Begue et al. extracted 24 features related to shape, color, and texture from 30 leaf images belonging to 20 plants [7]. They then trained a random forest classifier to classify the plant, which obtained 90.1% accuracy. Habiba et al. extracted Modified Local Gradient Pattern-based texture features from 1054 leaf images of 10 medicinal plants and trained an SVM classifier to classify the plants [8]. This study collected images from Bangladesh, similar to ours, but achieved only 96.11% accuracy. Naeem et al. utilized a bank of 65 features, which include texture features, run-length matrix, and multi-spectral features [9]. They captured 600 multi-spectral images of 6 medicinal plants and classified them using Multi-Layer Perceptron (MLP) with 99.13% accuracy. However, multi-spectral imaging is costly, and it is not easy to carry the imaging device to the site of plants; it is not suitable for practical applications. Pacifico et al. also utilized color and texture features of leaves to classify plants using MLP network [10]. They collected 1148 leaf images of 15 plants from Brazil and achieved 97% accuracy. Rohmat et al. proposed another manual feature-based approach. However, they selected the features using principle component analysis and then classified the plans using the CNN method [11]. Azadnia and Kheiralipour utilized 28 color and texture features of leaves to classify six medicinal plants [12]. They used a specially prepared image acquisition system and trained an artificial neural network to classify the quality-enhanced leaf images. This method achieved 100% in training. Anami et al. combined edge features with color and texture to classify three types of plants using SVM [13]. This method achieved 90% accuracy when trained using 278 features from 900 images. Pushpa et al. used the KNN classifier to classify ten plants based on the gray-level texture features with 60% accuracy [14]. Another texture-based SVM classifier was proposed by Puri et al.; however, the accuracy was 91% in their method [15]. Venkataraman and Mangayarkarasi combined histogram-oriented features with texture features and classified using SVM [16]. When few classes and limited training images were available, the TML classifier-based approaches using manually selected features functioned effectively. However, effective feature selection is critical to these models' performance. TML models are lighter and have fewer parameters than deep learning-based techniques like CNN.

On the contrary, deep learning models are highly parameterized but suitable for achieving high accuracy when the number of classes is high, given a large dataset. Duong-Trung et al. utilized lightweight MobileNet-based CNN for classifying ten medicinal plants collected from Vietnam [17]. This method was trained and tested using 2296 leaf images and achieved 98.7% accuracy. Valdez et al. also utilized MobileNet and achieved 97.43% accuracy [18]. Akter and Hosen [19] and Musa et al. [20] proposed a CNN-based method to classify Bangladeshi medicinal plants, similar to our study. Raisa et al. prepared a dataset of 37,693 leaf images of 10 plants and classified them with 71.3% accuracy. In contrast, Musa et al. classified six plants with 95.58% accuracy. Saikia et al. [21] also utilized a neural network for classifying six medicinal plants from 90 RGB leaf images collected from India. Dileep and Pournami combined AlexNet-based feature extraction with SVM classifier to classify 40 medical plants [22]. The model was trained using 2400 images and yielded 96.76% accuracy. A similar approach was proposed by Dewedi et al. utilizing ResNet and SVM model to classify 40 medicinal plants [23]. This model was trained using a public dataset of 6500 images. This method achieved 96.80% accuracy. Uddin et al. [24] proposed an ensemble of CNN models to classify ten medicinal plants. The method was trained using 5000 images for only ten epochs to achieve an accuracy of 99%. Our study, considered this method as the state-of-the-art, as it achieved the highest accuracy for the RGB images. However, this model was highly over-fitted with a loss of over 0.80. Another ensemble of CNN models was proposed by Bahri [25]. This method was trained using 40,000 images for 50 epochs to achieve 97.4% accuracy. This model also suffers from high loss, indicating an over-fitted method. Kyalkond et al. proposed a binary classifier to distinguish medical plants from non-medicinal plants using a custom-built CNN model [26]. This model was trained using 1600 images of 100 medicinal plants for 100 epochs. This method achieved 90% accuracy with a low loss value. However, identifying medicinal plants without knowing their class doesn't serve our study. Similar work was done by Pukhrambam and Sahayadhas in which they trained a DenseNet-based CNN for distinguishing Medicinal Plants from phytochemistry and therapeutics plants [27]. Berihu et al. proposed a GoogLeNet-based method to classify medicinal plants from Ethiopia with 96.7% accuracy [28]. Another neural network-based method was proposed by Kumar et al. to classify 25 different medicinal plant leaves using a Multi-layer Perceptron (MLP) classifier [29]. This method prepared and investigated the performance of 6 different deep learning-based methods and obtained the highest accuracy of 82.51%. Sharma studied the performance of VGG and VGG19 models [30]. The VGG16 model achieved a higher accuracy of 98.52% in classifying 21 plants. Widneh et al. trained a MobileNet model using 15,100 leaf images to classify 52 plants and achieved 92% accuracy [31]. These methods utilized entirely CNN-based models. They obtained excellent accuracy when trained on large datasets; nevertheless, for most of these methods, the loss climbed with the accuracy, particularly when the number of classes was high. This phenomenon suggests an over-fitting model which fails when used in different settings.

Some studies integrated multiple networks to develop a cascaded architecture to utilize different models for feature selection and classification. The cascaded architecture featuring deep learning-based feature selection with traditional machine learning-based classification tends to achieve higher accuracy with low loss when the number of classes is high. Moreover, these systems have less parameters compared to the entirely CNN-based methods, thus less computational complexity. Dileep and Pournami cascaded AlexNet for extracting the features and SVM for classification [22], which achieved 96.76% accuracy. They prepared a dataset of 2400 leaf images of 40 plants, which were captured using a scanner at 1200 or 600 dpi. Another cascaded network was proposed by Diwedi et al. [23]. They also classified 40 medicinal plants using 6500 leaf images from India and achieved 96.8% accuracy. They utilized ResNet50 for feature extraction and SVM for classification.

High accuracy is essential for the medicinal plant classification to be effective because a false positive detection could cause fatalities or major health issues. However, making the system practical and easy to use is also necessary. Recent deep learning techniques such as CNNs [32], vision transformers [33], graph-based neural networks [34] and AI-generated super-resolution images [35] have excelled in various domains, especially where large amounts of labeled data are available and complex patterns need to be captured. Therefore, many researchers developed large datasets and implemented CNN for medicinal plant classification. However, collecting a large dataset of medical plant images takes a lot of work. Some of these datasets contained noisy images and needed to be verified by experts. CNN-based models use local receptive fields to capture small, spatially proximate patterns in the image, which enables them to focus on local features, such as edges and textures, effectively. The convolutional and pooling operations and weight-sharing strategies contribute to CNN's remarkable feature extraction capabilities. Several medical plant classification methods were previously developed utilizing CNN-based models. However, these methods are highly parameterized, computationally intensive, and often produce over-fitted models [24], [25].

This study focused on developing an optimal and practical system for medicinal plant classification. SVM and similar traditional machine learning methods offer computational efficiency compared to complex deep learning models like CNN. SVMs can generalize well with limited data, making them suitable for scenarios where acquiring a large labeled dataset is challenging. It allows explicit feature engineering, where domain knowledge can be leveraged to select relevant features. This is suitable for medical plant classification and motivated many researchers to deploy SVM for classifying plants using different leaf features. However, these methods failed when the complexity of the decision boundary increased with the number of plant classes [5], [7], [8]. Therefore, in this research, we combined CNN's expelling feature extraction capability with SVM's ability to draw a simple decision boundary to develop an optimal method. This method utilized the ResNet50 model for feature extraction. Then, these features were further processed using PSO to select the suitable features to predict the class using SVM. The incorporation of PSO further boosted the accuracy of the system. The features were extracted using pre-trained CNN with no learnable parameters. The learnable parameters of the proposed system are included only in the PSO and SVM module, which makes the system lightweight and computationally effective.

## III. MATERIALS AND METHODS

### A. MEDICINAL PLANT DATASET

In this study, we collaborated with the ethnobotanists to prepare a dataset of 6,427 images of medicinal plants which included 678 aloevera, 1012 hibiscus, 1002 holy basil, 621 lemon grass, 926 neem, 1237 henna and 951 mint leaf images. These images were collected from six distinct places in Bangladesh and captured using an iPhone 12 camera (12 megapixels, f/1.6, 1.4 $\mu m$).

Initially, 7000 images were captured and then the exposure and focus quality of the images were examined to eliminate under-exposed and blurry images. To evaluate the focus and exposure quality of the image, we used a reference-less image quality evaluation method [36], in which we estimated the average width and height of the edges. After that, two ethnobotanists examined the images independently and we took only those images for which the experts' judgement matched. This was done to ensure the confirm the class of each image, which is the ground truth. In the quality and experts evaluations, 573 images were eliminated, resulting in 6,427 images for this dataset, as shown in Table 1.

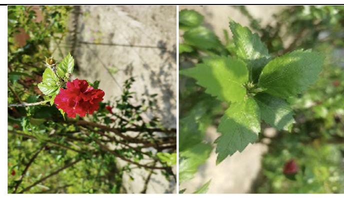
### B. CASCADED ARCHITECTURE FOR THE PROPOSED SYSTEM

In this study, we have considered seven different types of medical plants and proposed a cascaded architecture of AI models to classify them. Multiclass classification is typically challenging as assigning a class label to a new image among the seven classes is more complex than making the same decision where there are fewer classes. The time complexity is also higher for a multiclass classifier. Another challenge is the data imbalance problem. In a multiclass classification problem where some classes are rare and others are common, the model may learn to favor the common classes and neglect the rare ones. Finally, selecting the best model is another critical task. There is no one-size-fits-all solution; the strengths and weaknesses vary among the AI models depending on the data and the task. Therefore, in this study, we undertook an exhaustive search to select the best AI-enabled network to classify medicinal plants from one of the seven classes. We have gone through intensive data preparation, data curation, network tuning, feature selection, model evaluation, network selection and finally, demonstration of the selected network. To deal with data imbalance, we have oversampled the images of the classes with fewer images to balance the class distribution. In order to select the best network, we considered the accuracy, time complexity and interoperability of the networks. We determined the accuracy of the model using cross-validation. Grid-based searching was used to compare and optimize different models and parameters for the best performance.

In this study, we combined the CNN models with the TML models. CNN models are computationally expensive and require a large number of images, especially for a multi-class problem. In comparison, the TML models take less time to train and can achieve good accuracy when trained using a smaller dataset [36]. However, the performance of the TML models depends on the optimally selected features [37]. Therefore, we have utilized the PSO method to select the optimal features for training the TML models. The PSO was used because of its quick convergence speed, minimal memory requirements, and ease of implementation [38]. In this study, we proposed a cascaded architecture where the CNN model works as the feature extractor, which are then processed by the PSO method to select the optical features. Finally, a TML model utilizes these features to predict the output class. We used the pre-trained weights of CNN to reduce the computation and time complexity further. This cascaded architecture allowed for a reduction in training time and resources. Moreover, it achieved sufficient accuracy and high robustness. Fig. 1 shows the architecture of the proposed system.

The proposed system can classify the plants from low-resolution smartphone-captured images. Once the image was captured, the system evaluated the quality of the image first. If the quality of the image was satisfactory, then the image was converted to *sRGB* space to compensate for the color variation and later resized according to the input size of the CNN model. The quality was determined in terms of sharpness and brightness indices. We measured the distances between local minima and local maxima for

**TABLE 1.** Medicinal Plant species dataset information with sample images.

| Local name | Scientific name | Number of images | Sample Images |
|---|---|---|---|
| Aloevera | Aloevera | 678 |  |
| Hibiscus | Hibiscus | 1012 |  |
| Holy Basil | Ocimum tenuiflorum | 1002 |  |
| Lemon grass | Cymbopogon citratus | 621 |  |
| Neem | Azadirachta indica | 926 |  |
| Henna | Lawsonia inermis | 1237 |  |
| Mint Leaf | Mentha piperita | 951 |  |

each edge's gradient, corresponding to the edge's width. The sharpness index was the average width of the edges. A smaller average width value indicates that the image is sharper. Then, we also calculated the absolute difference in intensity between the local minima and maxima, representing the edge's height. The average height of the edges indicates the brightness of the image. A brighter image had a better average height. The sharpness and brightness indices are given by Equation (1) and (2):

$$\text{Sharpness index} = \frac{1}{N}\sum_{i=1}^{N} w(i) \tag{1}$$

$$\text{Brightness index} = \frac{1}{N}\sum_{i=1}^{N} h(i) \tag{2}$$

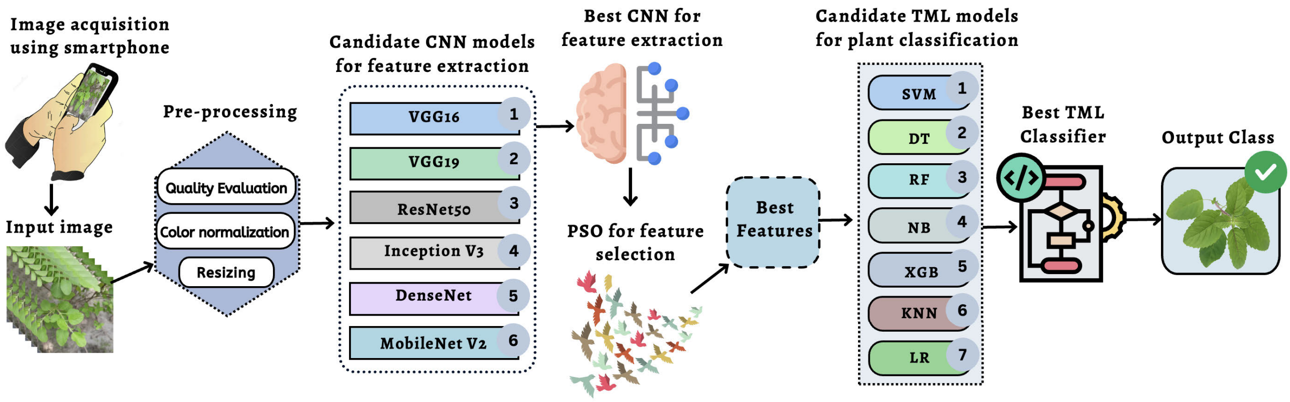



**FIGURE 1.** Architecture of the proposed cascaded network for medicinal plant classification.

In Equation (1) and (2), $N$ is the count of edges, $w(i)$ is the width of edge $i$ and $h(i)$ is the height of edge $i$. If an image has sharpness index higher than 6 or brightness index lower than 6, then it was eliminated for further processing.

Following the pre-assessment, the CNN-PSO-TML cascaded network processed the image to determine its class. The following section contains details of how the appropriate CNN and TML models were selected for the feature extraction and classification. We additionally evaluated the proposed system in this study using three different smartphones to ensure the system is robust.

### C. FEATURE SELECTION FOR CASCADED NETWORK

Particle Swarm Optimization (PSO) is indeed a population-based stochastic optimization algorithm developed based on the social behavior of birds flocking. PSO optimization follows the techniques that a swarm of birds follow to find their desired target in a search space. The swarm's movement is influenced by their current search directions, the best individual results or fitness to this point, the best fitness found by all birds, and a random perturbation. Every bird continues to track its best fitness. Then, it switches its position with other birds in the swarm. The swarm of birds uses this communication system to locate the desired target, something that is not possible for a single bird to do on its own. This method almost always converges on the global optimal.

In our experiment, we relied on the standard PSO algorithm by Kennedy and Eberhart [38] and utilized a similar parameter setup for feature selection as mentioned in [39] to derive the fitness for each particle, using the following the fitness function:

$$F_i = qa_i + \frac{1-q}{N_f^{\text{selected},i}} \tag{3}$$

where $F_i$ is the fitness of the $ith$ particle, $N_f^{\text{selected},i}$ is the number of selected features, $a_i$ is the accuracy for the $ith$ particle and $q$ is a weight. A higher value of $q$ yields higher number of features, maximizing its accuracy, whereas a lower value yields lower number of selected features, compromising accuracy. In our experiment $q$ was 1.

### D. MODEL SELECTION FOR CASCADED NETWORK

In this study, we have performed an exhaustive search to find suitable networks for the proposed cascaded architecture. We have experimented with six different CNN models and seven TML models incorporated with the PSO to get the best network for the proposed system. For the CNN-enabled feature extraction, we investigated the performance of deep convolutional architecture-based networks [40] VGG16 and VGG19, depthwise separable convolutions based network [42] Xception, densely connected layer based network [43] DenseNet121, deep residual learning-based networks [41] ResNet50, and inverted residual structure-based model MobileNet which is optimized for mobile devices [44]. Typically, a CNN model functions in two steps: firstly, the bottom layers of the models, which mainly contain the convectional layers, and pooling layers extract features from the input image. Then, the features extracted by the convolutional base are utilized by the top layers of the model, which mainly contain the dense layers and dropouts to predict the class based on the given features. One of the most significant advantages of using CNN models is their ability to extract useful features suitable for class prediction, which is often difficult to achieve using an independent feature selection method. Therefore, we have utilized the convolutional base extracted features of the CNN models in the cascaded architecture. However, we have utilized the pre-trained weights of the convolutional base, which were derived by using the ImageNet dataset for the CNN models. This allowed us to eliminate the convolutional base's training using our data. Then, these features are processed by PSO to identify the optimal features. PSO identifies the minimum number of best-fitting features. Finally, these features were utilized by seven different TML models, which include Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Naive Bayes (NB), eXtreme Gradient Boosting (XGB), K-Nearest Neighbors (KNN) and Logistic Regression (LR).

For each CNN-based model, we have investigated the performance incorporating PSO and each TML model. This generated 42 cascaded networks, which were then trained and tested using our dataset to select the best one. We performed a hold-out test and 10-fold cross-validation to compare the cascaded networks. In the hold-out test, we used 60% of the images for training, 20% for validation and 20% to test the networks. In the 10-fold cross-validation, the 7000 images were divided into ten groups. The images were randomly distributed to each group to have 700 images, containing 100 images per class.

We also performed an ablation study to determine the role of PSO. The time requirements were also analyzed for the selected network. Finally, the system was implemented and demonstrated using three smartphones to ensure it is robust regardless of the device.

## IV. RESULTS AND ANALYSIS

In this study, we evaluated the performance of 42 distinct networks, which were prepared by cascading six pre-trained CNN and seven TML models with the PSO. Each CNN model was used with PSO and one of the seven TML models, generating seven cascaded networks for each CNN model. Firstly, we selected the best candidate cascaded network of each CNN model based on their accuracy, precision, recall, and F1-Score for the test data. Then, the best candidate networks of all CNN models were compared based on their average accuracy in the 10-fold cross-validation experiment to select the best network for the proposed system. After that, the proposed method was compared with the existing methods including the state-of-art method to ensure its efficacy. We also evaluated the performance of these 42 cascaded networks without incorporating PSO-based feature selection in the ablation study to ensure the role of PSO in the cascaded network. The ablation study revealed that incorporating PSO increases the networks' accuracy. Finally, the performance of the proposed system was investigated for practical use. The results of these experiments are detailed in the following sections.

### A. MODEL SELECTION AND EVALUATION

First, we compared the 42 distinct PSO-incorporated cascaded networks' performance for the unseen test data in terms of average accuracy, precision, recall, and F1-Score, as shown in Fig. 2. This experiment revealed that the best candidate networks for VGG16, VGG19, ResNet50, Xception, DenseNet121, and MobileNet CNN are VGG16-PSO-KNN, VGG19-PSO-SVM, MoblineNetV2-PSO-SVM, and DenseNet121-PSO-SVM, respectively. From Fig. 2, it can be observed that the VGG16, VGG19, and ResNet50-based networks performed better for the other CNN models regardless of the type of TML classifiers. The VGG16 models achieved the highest accuracy of 98.61% when cascaded with the KNN classifier. While VGG19 and ResNet50 achieved their highest accuracy of 98.93% and 99.60%, respectively, with the SVM classifier. The DenseNet121 and MobileNetV2 also obtained their best performance with the SVM classifier. Four out of six CNN models achieved the best performance when SVM was incorporated as the classifier of the cascaded network. Moreover, in our experiment of ablation study, we have also found that the SVM yielded the best performance in the absence of PSO for most CNN models. This explains the suitability of SVM for the cascaded network. The findings of our experiment suggest that the SVM classifier performs significantly better than another classifier when trained with optimally selected features generated by a CNN model. This finding is analogous to the results reported in [36] for artifact classification.

After that, we performed 10-fold cross-validation for the best candidate networks. For this experiment, we used 6,400 images and distributed them randomly to one of the ten groups so that each group contains 67 aloe vera, 101 hibiscuses, 100 holy basil, 62 lemongrass, 92 neem, 123 henna, and 96 mint leaf images. We calculated the average accuracy for the candidate networks in the 10-fold cross-validation experiment. It involves dividing the dataset into ten subsets. The networks were trained and tested ten times, each using a different subset of images as the test set and the remaining data as the training set. This experiment provided a robust and reliable estimate of a network's performance, suitable for selecting the best network for the proposed system. In the cross-validation experiment, the ResNet50-PSO-SVM achieved an average accuracy of $98.39 \pm 0.19$, outperforming the other candidate networks, as shown in Table 2. The second highest accuracy was $96.13 \pm 0.40$ for the VGG19-PSO-SVM network. Therefore, the ResNet50-PSO-SVM was selected for the proposed system. Fig. 2 also showed that Resnet50-PSO-SVM outperformed all the cascaded networks in accuracy, precision, recall, and F1-score for the test data of the holdout experiment.Additionally, we plotted the confusion matrix for the best candidate networks for the test data. Fig. 3 shows the confusion matrices. Further, we plotted the received operating characteristics curves for these cascaded networks, as shown in Fig. 4. The average area under the curve (AUC) value was 1, indicating 100% accuracy of the network. The confusion matrices and the ROC curves also reveal the precedency of the ResNet50-PSO-SVM network over the other cascaded networks.

Finally, we have compared the results of the proposed ResNet50-PSO-SVM network-based method with the existing methods, as shown in Table 3. The proposed method outperformed the existing methods in average accuracy. This method utilized plant images and did not rely on singular leaf images. Moreover, this rapid method can predict the class in 0.17 seconds. Time for other methods was not reported. This Table also shows that the cascaded networks achieved higher accuracy when the number of classes are high compared to the other two approaches. Standalone CNN-based methods also achieved high accuracy; however, they required a large number of images for training. Alternatively, TML models such as SVM and RF performed well when trained using selected features derived from fewer images. Moreover, SVM has few parameters compared to the CNN models, making it lightweight. The proposed method utilized pre-trained

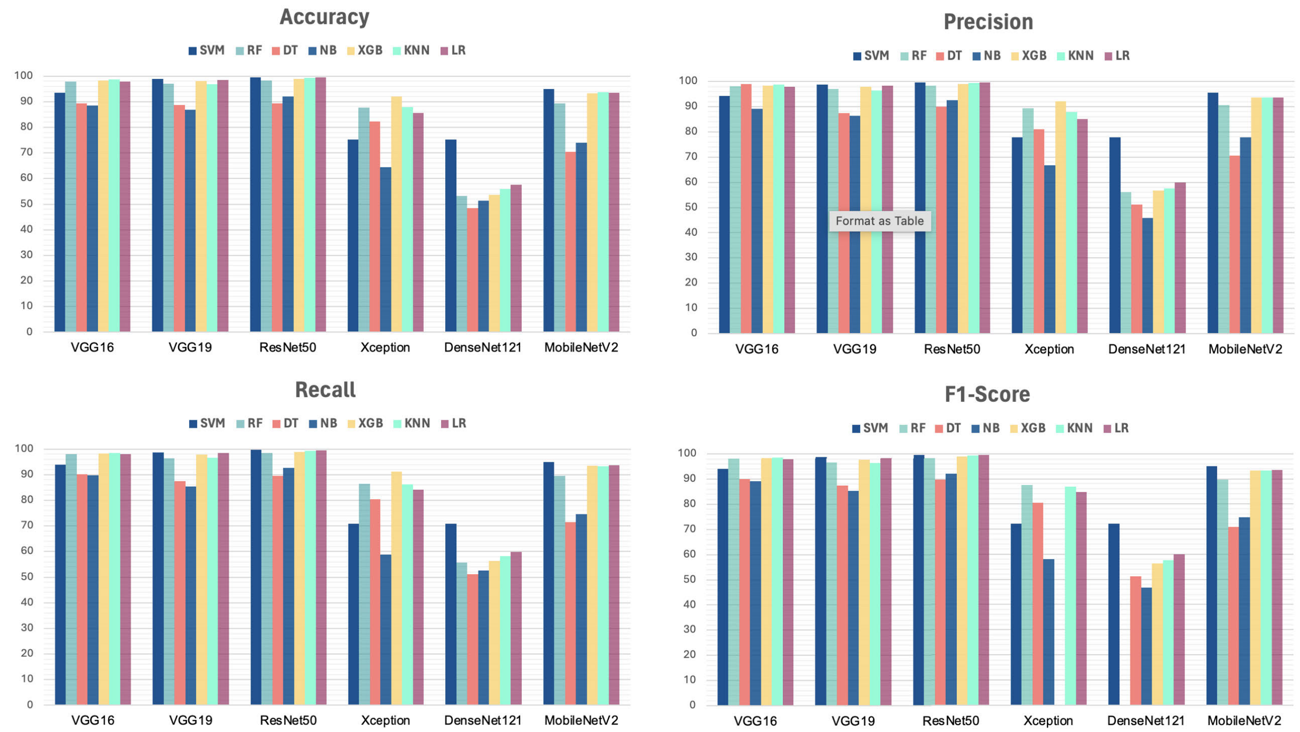



**FIGURE 2.** Evaluation to select the best cascaded network for each CNN.

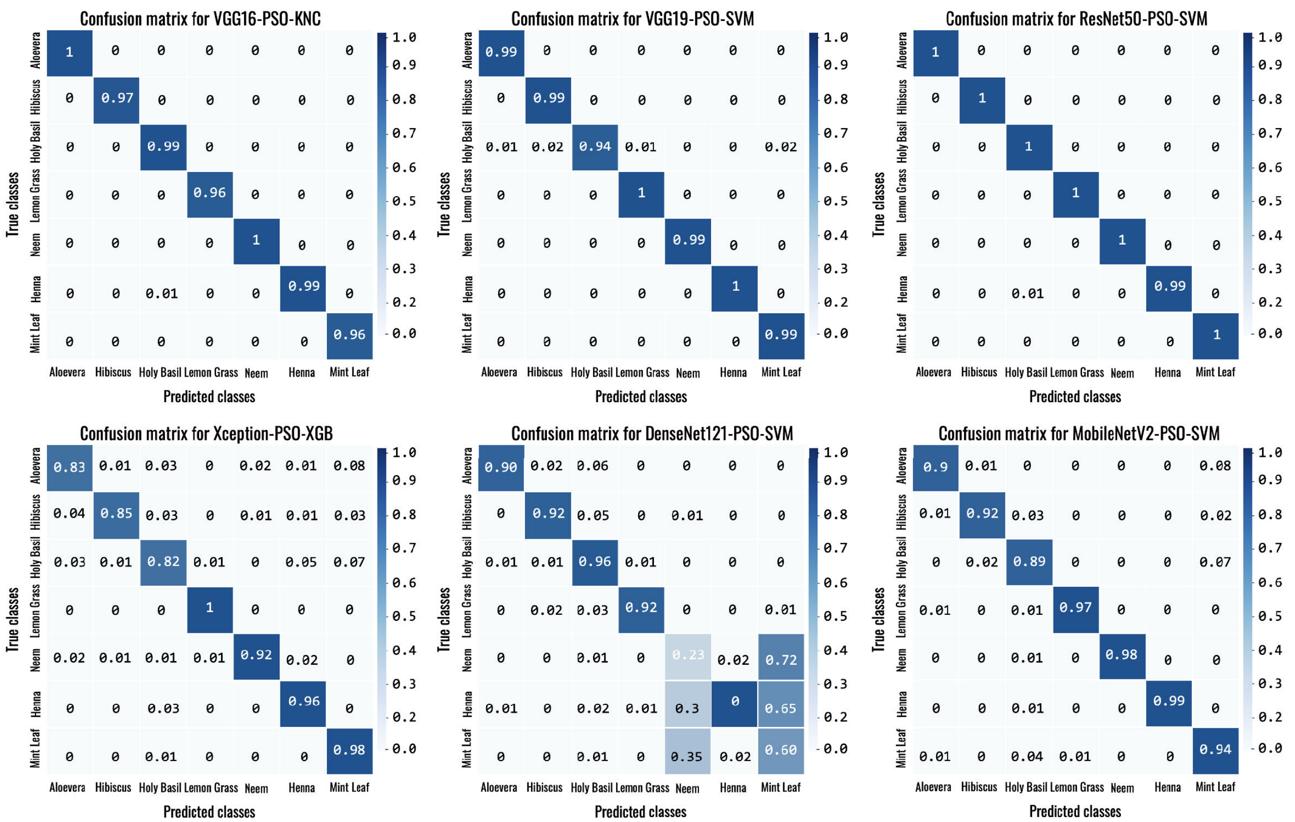



**FIGURE 3.** Confusion matrix for the best cascaded network of each CNN model.

RestNet50 for feature extraction that involves no learnable parameter. The SVM-based classifier and the PSO have a few parameters; thus, this cascaded network is lightweight compared to the standalone CNN-based methods.

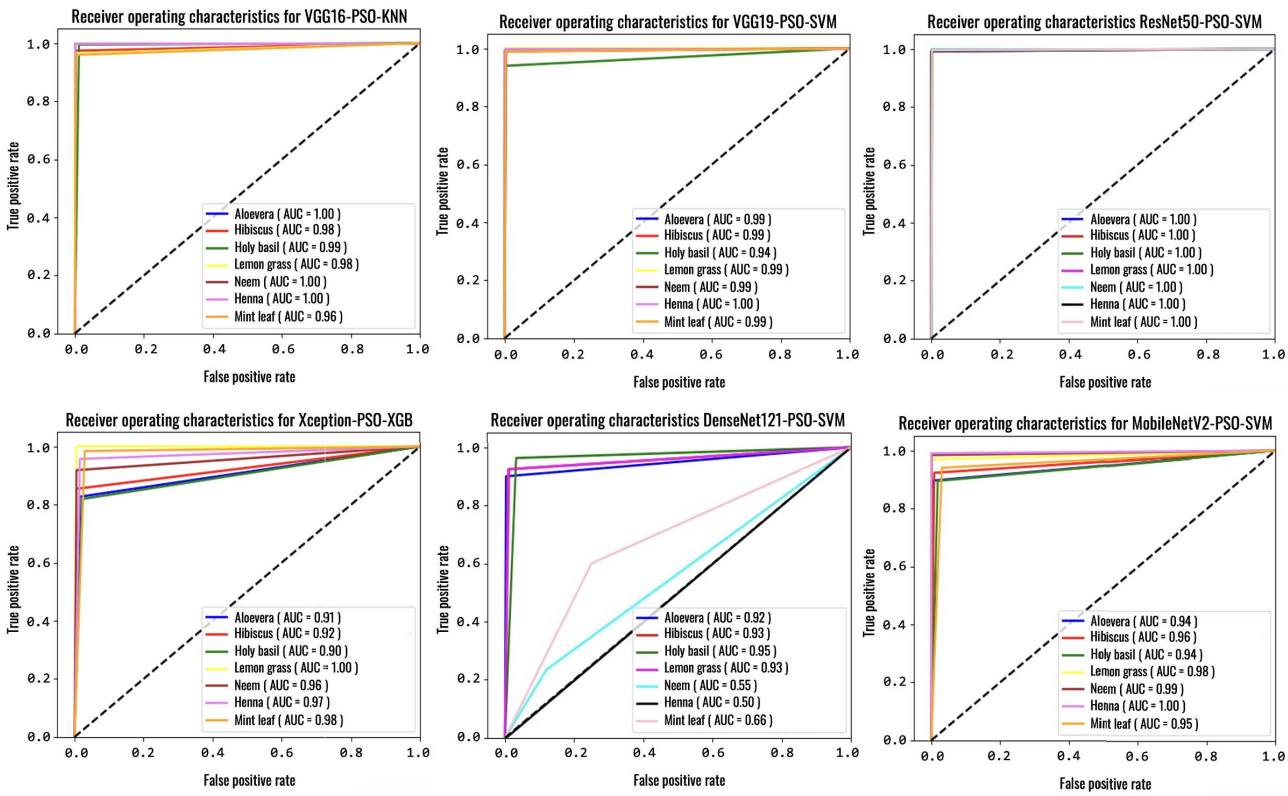



**FIGURE 4.** ROC for the best cascaded network of each CNN model.

**TABLE 2.** The fold-wise accuracy of the best cascaded networks of each CNN model.

| Cascaded Network | Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | Average |
|---|---|---|---|---|---|---|---|---|---|---|---|
| VGG16-PSO-KNN | 85.80 | 87.05 | 87.29 | 88.29 | 88.46 | 88.45 | 88.87 | 88.96 | 89.12 | 88.95 | 88.12 |
| VGG19-PSO-SVM | 95.24 | 95.64 | 96.08 | 96.17 | 96.08 | 96.17 | 96.44 | 96.52 | 96.61 | 96.43 | 96.13 |
| ResNet50-PSO-SVM | 98.93 | 98.17 | 98.33 | 98.25 | 98.25 | 98.33 | 98.33 | 98.41 | 98.41 | 98.40 | 98.39 |
| Xception-PSO-XGB | 81.04 | 82.13 | 83.50 | 84.16 | 84.60 | 85.33 | 85.65 | 85.36 | 85.00 | 86.50 | 84.32 |
| DenseNet121-PSO-SVM | 61.45 | 62.43 | 63.55 | 64.46 | 64.60 | 65.30 | 65.66 | 65.73 | 65.80 | 66.01 | 64.50 |
| MobileNetV2-PSO-SVM | 84.07 | 86.90 | 87.06 | 88.36 | 88.52 | 88.84 | 88.44 | 89.00 | 88.84 | 89.08 | 87.91 |

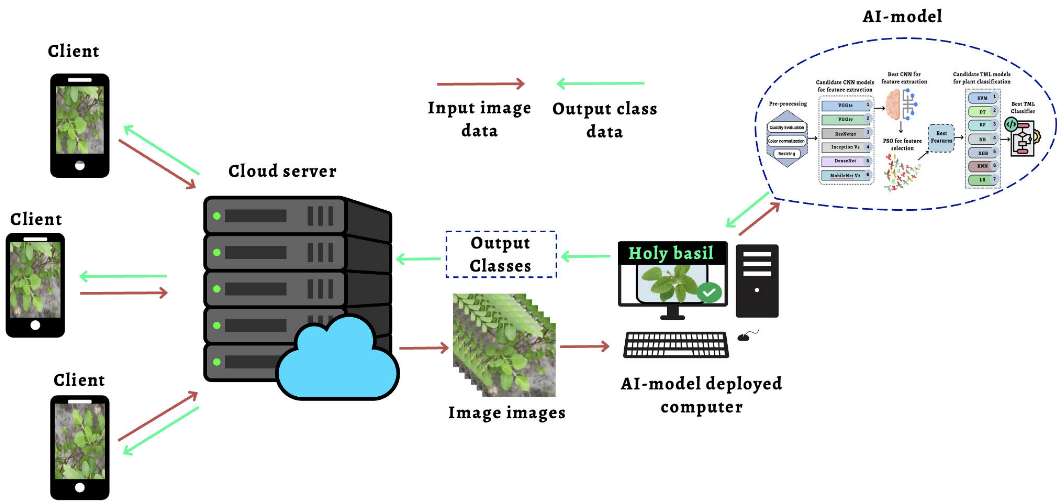



**FIGURE 5.** Implementation of the proposed system for practical use.

**TABLE 3. Comparison between relevant existing methods and proposed method.**

| Methods | Techniques | Data description | Accuracy |
|---|---|---|---|
| Kan et al. [5] | Manually selected features with SVM classifier | 240 leaf images (RGB) of 12 plants captured using digital camera from China | 93.30% |
| Janani et al. [6] | Manually selected features with ANN classifier | 63 leaf images (RGB) of 6 plants captured using digital camera from India | 94.40% |
| Begue et al. [7] | Manually selected features with RF classifier | 30 leaf images (RGB) of 20 plants captured using high resolution digital camera | 90.10% |
| Habiba et al. [8] | Manually selected Local Gradient Pattern-based features with SVM classifier | 1054 leaf images (RGB) of 10 plants captured using smartphone from Bangladesh | 96.11% |
| Naeem et al. [9] | Chi-square feature selection with MLP classifier | 600 leaf images (Multi-spectral) of 6 plants captured using multi-spectral camera from Pakistan | 99.13% |
| Luciano et al. [10] | Manually selected color and texture features with MLP classifier | 1148 leaf images (RGB) of 15 plants captured using digital camera from Brazil | 97.00% |
| Rohmat et al. [11] | PCA-based feature selection with CNN classifier | Not available | 88.67% |
| Azadnia et al. [12] | Manually selected color and texture features with ANN classifier | Specially designed imaging system; images from 6 plants; number not available | 100.0% |
| Anami et al. [13] | Manually selected edge, color and texture features with SVM classifier | 1200 leaf images (RGB) of 3 plants, captured using high resolution digital camera from India | 90.00% |
| Puri et al. [15] | Manually selected texture features with SVM classifier | 514 leaf images (RGB) of 5 plants captured using digital camera from India | 93.54% |
| Trung et al. [17] | CNN based feature extraction with MobileNet classifier | 2296 leaf Images (RGB) of 10 medicinal plants captured using high resolution digital camera from Vietnam | 98.70% |
| Valdez et al. [18] | CNN based feature extraction with MobileNetV3 classifier | RGB Leaf images of 10 plants, caputred using digital camera from Philippines, image number not available | 97.43% |
| Raisa et al. [19] | CNN based feature extraction with Custom CNN classifier | 37,693 leaf images (RGB) of 10 different medicinal plants captured using smartphone from Bangladesh | 71.30% |
| Musa et al. [20] | CNN based feature extraction with Custom CNN classifier | RGB Leaf images of 6 plants captured using digital camera from Bangladesh, image number not available | 95.58% |
| Anurag et al. [21] | Manually selected color and texture features with Back Propagation Neural Network classifier | 90 leaf images (RGB) of 6 plants, captured using digital camera from Assam, India | 98.88% |
| Dileep et al. [22] | AlexNet based feature extraction with SVM classifier | 2400 leaf images (RGB) of 40 medicinal plants captured using high resolution digital camera from India | 96.76% |
| Dewedi et al. [23] | ResNet50 extracted features with SVM classifier | 6500 leaf images (RGB) of 40 medicinal plants captured using digital camera from India | 96.80% |
| **Hasib et al. [24]** | **Ensembled CNN based classifier** | **5000 leaf images (RGB) of 10 medicinal plants captured using smartphone from Bangladesh** | **99.00%** |
| Bahri et al. [25] | Ensembled CNN based classifier | 40,000 leaf images (RGB) of 50 medicinal plants captured using digital camera from Morocco | 97.40% |
| Kyalkond et al. [26] | CNN based medicinal and non-medicinal binary classifier | 1600 leaf images (RGB) of 100 medicinal plants captured using high resolution digital camera from India | 90.00% |
| Sharma et al. [30] | CNN based classifier | RGB leaf images of 21 medicinal plants, captured using smartphone from India, image number not available | 98.52% |
| Widneh et al. [31] | MobileNet CNN classifier | 15,100 leaf images (RGB) of 52 medicinal plants, captured using high-resolution digital camera | 92.00% |
| **Proposed method** | **ResNet50 extracted features were optimized by PSO to classify using SVM** | **6427 leaf images (RGB) of 7 medicinal plants, captured using smartphone from Bangladesh** | **99.60%** |

### B. ABLATION STUDY FOR PSO

Ablation studies help understand the role of a particular component in a complex system, such as the cascaded architecture of multiple machine learning models. This ablation study investigates the performance of cascaded networks by removing PSO-based feature selection to understand its contribution. For this purpose, we trained and validated each of the 42 cascaded networks without the PSO module using the same dataset. Then, the trained networks were tested using the same dataset. Finally, we compared the performances of these networks in terms of average accuracy, precision, recall and F1-Score, as provided in Table 4. This table shows that the performance of most of the networks improves in the presence of the PSO-based feature selection. SVM classifier-based networks yielded the best performance for most cascaded networks for both with and without PSO. Additionally, we have compared the average accuracy and average time requirements for the best-cascaded network of each CNN with and without integrating PSO, shown in Table 5. The time was estimated for 145 unseen test images. This table shows that the PSO improved accuracy without significantly increasing the total computational time of the network. It justifies the contribution of the PSO-based feature selection for the cascaded networks and the proposed system. The SVM-based classifiers tend to cost more time compared to the other classifiers, such as KNN and XGB. XGB classifier showed higher computation speed than KNN and SVM, considering the number of features.

### C. FEASIBILITY OF PROPOSED SYSTEM FOR PRACTICAL USE

A practical system enables capturing the image quickly, such as using smartphones and provides rapid and robust classification results on the site of plants. In this study,

**TABLE 4. Results of the ablation study for the PSO.**

| CNN-based Feature Extractor | Classifiers | Avg. Accuracy (%) | | Avg. Precision (%) | | Avg. Recall (%) | | Avg. F1-Score (%) | |
|---|---|---|---|---|---|---|---|---|---|
| | | **W/O PSO** | **With PSO** | **W/O PSO** | **With PSO** | **W/O PSO** | **With PSO** | **W/O PSO** | **With PSO** |
| VGG16 | SVM | 93.44 | 93.53 | 93.37 | 94.33 | 93.31 | 93.83 | 93.72 | 94.08 |
| | RF | 97.10 | 97.76 | 97.35 | 98.07 | 97.37 | 98.04 | 97.34 | 98.05 |
| | DT | 88.71 | 89.38 | 89.57 | 98.88 | 89.68 | 90.14 | 89.59 | 89.99 |
| | NB | 88.46 | 88.46 | 88.84 | 89.18 | 89.50 | 89.76 | 88.97 | 89.14 |
| | XGB | 98.09 | 98.17 | 98.04 | 98.35 | 98.03 | 98.31 | 98.13 | 98.33 |
| | **KNN** | **98.01** | **98.61** | **98.28** | **98.73** | **98.05** | **98.47** | **98.15** | **98.59** |
| | LR | 97.18 | 97.84 | 97.40 | 97.81 | 97.49 | 98.03 | 97.44 | 97.91 |
| VGG19 | **SVM** | **98.67** | **98.93** | **98.40** | **98.82** | **98.66** | **98.69** | **98.53** | **98.73** |
| | RF | 96.88 | 97.06 | 96.40 | 97.04 | 95.76 | 96.40 | 96.03 | 96.55 |
| | DT | 87.46 | 88.70 | 86.32 | 87.39 | 86.43 | 87.56 | 86.35 | 87.45 |
| | NB | 86.83 | 86.93 | 86.38 | 86.33 | 84.79 | 85.44 | 84.76 | 85.31 |
| | XGB | 97.86 | 98.04 | 97.45 | 97.89 | 97.61 | 97.89 | 97.64 | 97.73 |
| | KNN | 96.35 | 96.89 | 96.12 | 96.44 | 96.35 | 96.55 | 96.13 | 96.50 |
| | LR | 98.40 | 98.49 | 98.14 | 98.26 | 98.11 | 98.48 | 98.18 | 98.29 |
| ResNet50 | **SVM** | **99.13** | **99.60** | **99.15** | **99.60** | **99.19** | **99.66** | **99.17** | **99.63** |
| | RF | 98.01 | 98.25 | 98.08 | 98.35 | 98.16 | 98.38 | 98.10 | 98.35 |
| | DT | 87.92 | 89.27 | 88.23 | 89.98 | 87.78 | 89.58 | 87.94 | 89.73 |
| | NB | 91.81 | 91.97 | 92.22 | 92.49 | 92.44 | 92.74 | 92.20 | 92.25 |
| | XGB | 98.57 | 98.89 | 98.61 | 98.89 | 98.66 | 98.92 | 98.63 | 98.90 |
| | KNN | 99.05 | 99.36 | 99.08 | 99.42 | 99.08 | 99.36 | 99.07 | 99.39 |
| | LR | 99.12 | 99.52 | 99.23 | 99.53 | 99.18 | 99.59 | 99.15 | 99.56 |
| Xception | SVM | 74.21 | 75.26 | 77.78 | 77.85 | 69.75 | 70.85 | 70.45 | 72.07 |
| | RF | 86.12 | 87.74 | 87.75 | 89.41 | 84.65 | 86.38 | 85.65 | 87.47 |
| | DT | 78.35 | 82.27 | 76.24 | 81.02 | 76.00 | 80.48 | 75.85 | 80.46 |
| | NB | 63.70 | 64.47 | 63.80 | 66.73 | 58.22 | 58.82 | 57.83 | 57.92 |
| | **XGB** | **91.73** | **91.94** | **91.54** | **92.05** | **90.86** | **91.18** | **91.11** | **91. 47** |
| | KNN | 86.48 | 87.81 | 86.40 | 87.92 | 84.72 | 86.26 | 85.35 | 86.92 |
| | LR | 85.42 | 85.63 | 84.91 | 85.04 | 84.12 | 84.25 | 84.48 | 84.62 |
| Dense Net121 | **SVM** | **62.95** | **75.26** | **61.94** | **77.78** | **65.02** | **70.85** | **62.02** | **72.07** |
| | RF | 52.97 | 53.14 | 55.68 | 56.09 | 55.56 | 55.63 | 55.61 | 55. 83 |
| | DT | 46.45 | 48.51 | 49.25 | 51.27 | 48.94 | 51.15 | 48.97 | 51.13 |
| | NB | 50.00 | 51.40 | 47.36 | 45.87 | 51.14 | 52.52 | 46.10 | 46.77 |
| | XGB | 53.47 | 53.71 | 56.48 | 56.82 | 55.99 | 56.34 | 56.22 | 56.40 |
| | KNN | 36.35 | 56.02 | 57.48 | 57.59 | 58.08 | 58.16 | 57.57 | 57.68 |
| | LR | 57.67 | 57.67 | 59.95 | 59.98 | 59.80 | 59.91 | 59.85 | 59.94 |
| Mobile NetV2 | **SVM** | **94.34** | **94.99** | **95.16** | **95.48** | **94.28** | **95.04** | **94.36** | **95.05** |
| | RF | 87.31 | 89.41 | 88.70 | 90.56 | 87.28 | 89.58 | 87.49 | 89.86 |
| | DT | 68.39 | 70.49 | 67.65 | 70.56 | 68.75 | 71.46 | 67.93 | 70.90 |
| | NB | 72.76 | 73.97 | 76.20 | 77.90 | 73.13 | 74.53 | 73.26 | 74.68 |
| | XGB | 92.40 | 93.29 | 93.02 | 93.66 | 92.27 | 93.40 | 92.46 | 93.49 |
| | KNN | 92.56 | 93.78 | 92.74 | 93.66 | 92.71 | 93.38 | 92.71 | 93.50 |
| | LR | 93.37 | 93.45 | 93.35 | 93.55 | 93.59 | 93.66 | 93.45 | 93.60 |

**TABLE 5. The impact of PSO in the cascaded network in terms of number of features, accuracy and computation time.**

| | Number of features | | Avg. Accuracy (%) | | Avg. time (seconds) | |
|---|---|---|---|---|---|---|
| Networks | W/O PSO | With PSO | W/O PSO | With PSO | W/O PSO | With PSO |
| VGG16-KNN | 4096 | 2065 | 98.01 | 98.61 | $33.6 \pm 0.97$ | $33.1 \pm 0.73$ |
| VGG19-SVM | 4096 | 2004 | 98.67 | 98.93 | $43.9 \pm 0.12$ | $48.2 \pm 0.21$ |
| ResNet50-SVM | 2048 | 1001 | 99.13 | 99.60 | $13.5 \pm 0.21$ | $14.9 \pm 0.10$ |
| Xception-XGB | 2048 | 1045 | 91.73 | 91.94 | $11.9 \pm 0.22$ | $12.7 \pm 0.21$ |
| DenseNet121-SVM | 1024 | 510 | 62.95 | 75.26 | $15.9 \pm 0.10$ | $18.1 \pm 0.11$ |
| MobileNetV2-SVM | 1024 | 491 | 94.34 | 94.99 | $13.2 \pm 0.40$ | $13.5 \pm 0.63$ |

we assessed the feasibility of the proposed system for practical use. For this purpose, firstly, we tested the proposed system for three different smartphones to determine its robustness. We captured two sets of 145 plant images using two other smartphones: Xiaomi Redmi Note 10 Pro (108 megapixels, f/1.9, 0.7 $\mu m$) and Samsung Galaxy S10 (12 megapixels, f/1.5-2.4, 1.4 $\mu m$). We captured the images for the same trees and then transferred the images to a server using the internet. The proposed method received the images from the server and predicted the output for each set of images captured by the smartphones. Then, the results were sent back to the client's smartphone. Fig. 5 shows how the proposed system is implemented currently for practical use. The accuracy of the system using Redmi Note Pro and Samsung Galaxy S10 was 93.89% and 99.90%, respectively. This ensured the robustness of the method. This system can detect the class of plans from the smartphone-captured plant images taken from a standard distance (1)-2 meters).

Additionally, this method was tested for smartphone-captured leaf images and was found effective.

Secondly, we estimated the time requirements of the proposed system. The proposed method took $35.6 \pm 2.14$ and $22.9 \pm 5.67$ seconds for classifying 145 images captured by Redmi Note Pro and Samsung Galaxy S10 phones, respectively. The accuracy and the evaluation time were 99.60% and $14.9 \pm 0.10$ seconds respectively for the iPhone 12. The proposed method is robust regardless of the device, as demonstrated by the average accuracy and prediction time of $97.79 \pm 2.76\%$ and $0.15 \pm 0.05$ seconds/image for the three devices. The total turnaround time for the proposed system from capturing image to receiving the classification result was less than five minutes for a single plant, regardless the smartphone. This is practical considering the existing system which requires to process image after leaving the site of plants. In the future, an smartphone application will be developed and installed to perform the classification in the client machine and provide results instantly.

## V. DISCUSSION

Traditionally, experts identify medicinal plants meticulously to use them for drug development or treatment. The manual identification is prone to inter-observer variability and is contingent upon the availability of experts. The manual assessment lacks reliability because even one incorrect identification can result in significant health problems, financial losses, and even fatalities. If the image is captured in different settings, existing automated methods fail to detect plants. They rely on singular leaf images taken from a short distance, such as one meter. These systems fail when the image is captured using a different imaging device or smartphone. More importantly, they are required to carry the images and later process them using a computer to get the results. Thus, the existing systems lack practical usability.

In this study, we presented an automated medicinal plant classification method that outperformed previous approaches, achieving 99.60% accuracy. This method is designed to predict the class from plant images. Moreover, this system also works if leaf images are provided. The images were captured on the plant's site and transferred over the internet to get the classification results in the client's phone in less than five minutes. This study demonstrated the system for two different smartphones not used in the training. Regardless of the imaging devices, the system's high accuracy and rapid speed ensured its robustness. This system was effective for different smartphones, for singular leaf and plant images, and for providing the classification results on the site in less than five minutes. This ensured the practical use of the system.

The proposed method incorporated a three-stage image pre-assessment, which includes color normalization to compensate for the color variation due to different imaging devices, sharpness evaluation to ensure that the image is not out of focus or blurry, and brightness evaluation to ensure the images are not underexposed. These pre-assessment steps boosted the robustness of the system. In this study, we incorporated PSO for feature selection, which boosted the system's performance, as found in the experiments. Nonetheless, examining the efficacy of alternative feature selection techniques, such as chi-square, is imperative. This method was designed for only limited medicinal plants, which should be increased in the future. Again, this method was not tested for images captured using other devices, such as digital cameras, or captured by other groups. This is another limitation of this study.

## VI. CONCLUSION

In this paper, we proposed an automatic medicinal plant classification method for Bangladesh. This method is highly accurate and practically sound to identify the class of the plants using the smartphone-captured image on the site of plants within minutes. In the future, more plants will be integrated and a standalone mobile application will be developed for the proposed system.

## REFERENCES

[1] K. Ahmad and I. Ahmad, *Herbal Medicine: Current Trends and Future Prospects*. (New Look to Phytomedicine). New York, NY, USA: Academic, 2019.

[2] M. J. Balunas and A. D. Kinghorn, "Drug discovery from medicinal plants," *Life Sci.*, vol. 78, no. 5, pp. 431–441, Dec. 2005.

[3] L. N. Joppa, D. L. Roberts, N. Myers, and S. L. Pimm, "Biodiversity hotspots house most undiscovered plant species," *Proc. Nat. Acad. Sci. USA*, vol. 108, no. 32, pp. 13171–13176, Aug. 2011.

[4] M. F. Kadir, J. R. Karmoker, M. R. Alam, S. R. Jahan, S. Mahbub, and M. M. K. Mia, "Ethnopharmacological survey of medicinal plants used by traditional healers and indigenous people in chittagong Hill tracts, bangladesh, for the treatment of snakebite," *Evidence-Based Complementary Alternative Med.*, vol. 2015, pp. 1–23, Jan. 2015.

[5] H. X. Kan, L. Jin, and F. L. Zhou, "Classification of medicinal plant leaf image based on multi-feature extraction," *Pattern Recognit. Image Anal.*, vol. 27, no. 3, pp. 581–587, Jul. 2017.

[6] R. Janani and A. Gopal, "Identification of selected medicinal plant leaves using image features and ANN," in *Proc. Int. Conf. Adv. Electron. Syst. (ICAES)*, Sep. 2013, pp. 238–242.

[7] A. Begue, V. Kowlessur, U. Singh, F. Mahomoodally, and S. Pudaruth, "Automatic recognition of medicinal plants using machine learning techniques," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 4, pp. 166–175, 2017.

[8] U. Habiba, M. R. Howlader, M. A. Islam, R. H. Faisal, and M. M. Rahman, "Automatic medicinal plants classification using multi-channel modified local gradient pattern with SVM classifier," in *Proc. Joint 8th Int. Conf. Informat., Electron. Vis. (ICIEV) 3rd Int. Conf. Imag., Vis. Pattern Recognit. (icIVPR)*, May 2019, pp. 6–11.

[9] S. Naeem, A. Ali, C. Chesneau, M. H. Tahir, F. Jamal, R. A. K. Sherwani, and M. Ul Hassan, "The classification of medicinal plant leaves based on multispectral and texture feature using machine learning approach," *Agronomy*, vol. 11, no. 2, p. 263, Jan. 2021.

[10] L. D. S. Pacifico, L. F. S. Britto, E. G. Oliveira, and T. B. Ludermir, "Automatic classification of medicinal plant species based on color and texture features," in *Proc. 8th Brazilian Conf. Intell. Syst. (BRACIS)*, Brazil, Oct. 2019, pp. 741–746.

[11] R. I. Borman, R. Napianto, N. Nugroho, D. Pasha, Y. Rahmanto, and Y. E. P. Yudoutomo, "Implementation of PCA and KNN algorithms in the classification of Indonesian medicinal plants," in *Proc. Int. Conf. Comput. Sci., Inf. Technol., Electr. Eng. (ICOMITEE)*, Oct. 2021, pp. 46–50.

[12] R. Azadnia and K. Kheiralipour, "Recognition of leaves of different medicinal plant species using a robust image processing algorithm and artificial neural networks classifier," *J. Appl. Res. Medicinal Aromatic Plants*, vol. 25, Dec. 2021, Art. no. 100327.

[13] B. S. Anami, S. S. Nandyal, and A. Govardhan, "A combined color, texture and edge features based approach for identification and classification of Indian medicinal plants," *Int. J. Comput. Appl.*, vol. 6, no. 12, pp. 45–51, Sep. 2010.
[14] B. R. Pushpa, N. Megha, and K. B. Amaljith, "Comparision and classification of medicinal plant leaf based on texture feature," in *Proc. Int. Conf. Emerg. Technol. (INCET)*, Jun. 2020, pp. 1–5.
[15] D. Puri, A. Kumar, and J. Virmani, "Classification of leaves of medicinal plants using laws' texture features," *Int. J. Inf. Technol.*, vol. 14, pp. 931–942, Aug. 2019.
[16] D. Venkataraman and N. Mangayarkarasi, "Support vector machine based classification of medicinal plants using leaf features," in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Sep. 2017, pp. 793–798.
[17] N. Duong-Trung, L.-D. Quach, M.-H. Nguyen, and C.-N. Nguyen, "A combination of transfer learning and deep learning for medicinal plant classification," in *Proc. 4th Int. Conf. Intell. Inf. Technol.*, Feb. 2019, pp. 83–90.
[18] D. B. Valdez, C. J. G. Aliac, and L. S. Feliscuzo, "Medicinal plant classification using convolutional neural network and transfer learning," in *Proc. IEEE Int. Conf. Artif. Intell. Eng. Technol. (IICAIET)*, Sep. 2022, pp. 1–6.
[19] R. Akter and M. I. Hosen, "CNN-based leaf image classification for Bangladeshi medicinal plant recognition," in *Proc. Emerg. Technol. Comput., Commun. Electron. (ETCCE)*, Dec. 2020, pp. 1–6.
[20] M. Musa, M. Arman, M. Hossain, A. Thusar, N. Nisat, and A. Islam, "Classification of immunity booster medicinal plants using CNN: A deep learning approach," in *Proc. Int. Conf. Adv. Comput. Data Sci.*, Nashik, India, Apr. 2021, pp. 244–254.
[21] A. P. Saikia, P. V. Hmangaihzuala, S. Datta, S. Gope, S. Deb, and K. R. Singh, "Medicinal plant species classification using neural network classifier," in *Proc. 6th Int. Conf. Commun. Electron. Syst. (ICCES)*, Jul. 2021, pp. 1805–1811.
[22] M. R. Dileep and P. N. Pournami, "AyurLeaf: A deep learning approach for classification of medicinal plants," in *Proc. TENCON IEEE Region 10 Conf. (TENCON)*, Oct. 2019, pp. 321–325.
[23] H. K. Diwedi, A. Misra, and A. K. Tiwari, "CNN-based medicinal plant identification and classification using optimized SVM," *Multimedia Tools Appl.*, vol. 83, no. 11, pp. 33823–33853, Sep. 2023.
[24] A. H. Uddin, Y.-L. Chen, B. Borkatullah, M. S. Khatun, J. Ferdous, P. Mahmud, J. Yang, C. S. Ku, and L. Y. Por, "Deep-learning-based classification of Bangladeshi medicinal plants using neural ensemble models," *Mathematics*, vol. 11, no. 16, p. 3504, Aug. 2023.
[25] A. Bahri, "Dynamic CNN combination for Morocco aromatic and medicinal plant classification," *Int. J. Comput. Digit. Syst.*, vol. 11, no. 1, pp. 239–249, Jan. 2022.
[26] S. Kyalkond, S. Aithal, V. Sanjay, and P. Kumar, "A novel approach to classification of ayurvedic medicinal plants using neural networks," *Int. J. Eng. Res. Technol.*, vol. 11, no. 1, p. 2278, 2022.
[27] B. Pukhrambam and A. Sahayadhas, "Advanced medicinal plant classification and bioactivity identification based on dense net architecture," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 6, pp. 1–7, 2022.
[28] M. Berihu, J. Fang, and S. Lu, "Automatic classification of medicinal plants of leaf images based on convolutional neural network," in *Proc. CCF Conf. Big Data*, 2022, pp. 108–116.
[29] G. Kumar, V. Kumar, and A. Hrithik, "Herbal plants leaf image classification using machine learning approach," in *Intelligent Systems and Smart Infrastructure*. Boca Raton, FL, USA: CRC Press, 2023, pp. 549–558.
[30] P. Sharma, "Automatic classification and diseases identification of Indian medicinal plants using deep learning models," *ECS Trans.*, vol. 107, no. 1, pp. 4631–4639, Apr. 2022.
[31] M. Widneh, A. Workneh, and A. Alemu, "Medicinal plant parts identification and classification using deep learning based on multi label categories," *Ethiopian J. Sci. Sustain. Develop.*, vol. 8, pp. 96–108, Sep. 2021.
[32] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1–7.
[33] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16×16 words: Transformers for image recognition at scale," 2020, *arXiv:2010.11929*.
[34] J. Zhou, "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, pp. 57–81, Jan. 2020.
[35] X. Zhu, K. Guo, H. Fang, L. Chen, S. Ren, and B. Hu, "Cross view capture for stereo image super-resolution," *IEEE Trans. Multimedia*, vol. 24, pp. 3074–3086, 2022.
[36] H. M. Shakhawat, T. Nakamura, F. Kimura, Y. Yagi, and M. Yamaguchi, "[Paper] automatic quality evaluation of whole slide images for the practical use of whole slide imaging scanner," *ITE Trans. Media Technol. Appl.*, vol. 8, no. 4, pp. 252–268, 2020.
[37] M. S. Hossain, M. G. Hanna, N. Uraoka, T. Nakamura, M. Edelweiss, E. Brogi, M. R. Hameed, M. Yamaguchi, D. S. Ross, and Y. Yagi, "Automatic quantification of HER2 gene amplification in invasive breast cancer from chromogenic in situ hybridization whole slide images," *J. Med. Imag.*, vol. 6, no. 4, p. 1, Nov. 2019.
[38] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. ICNN Int. Conf. Neural Netw.*, Nov. 1995, pp. 1942–1948.
[39] M. S. Hossain, N. Hasan, M. A. Samad, H. M. Shakhawat, J. Karmoker, F. Ahmed, K. F. M. N. Fuad, and K. Choi, "Android ransomware detection from traffic analysis using metaheuristic feature selection," *IEEE Access*, vol. 10, pp. 128754–128763, 2022.
[40] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[41] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[42] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1800–1807.
[43] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.
[44] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.

**MD. TAREQUL ISLAM** received the B.Sc. degree from the University of Rajshahi, Bangladesh, and the M.Sc. (Engg.) degree in computer science and engineering from Mawlana Bhashani Science and Technology University (MBSTU), Bangladesh, where he is currently pursuing the Ph.D. degree. He is an Assistant Professor with the Computer Science and Engineering Department, Khwaja Yunus Ali University, Bangladesh. His research interests include computer vision, bioinformatics, the IoT, and blockchain.

**WAHIDUR RAHMAN** received the bachelor's and master's degrees in computer science from the Computer Science and Engineering Department, Mawlana Bhashani Science and Technology University. He is currently a Senior Lecturer with the Department of Computer Science and Engineering. His research interests include machine learning, the Internet of Things, and computer vision in agriculture.

**MD. SHAKHAWAT HOSSAIN** (Member, IEEE) received the Ph.D. degree from Tokyo Institute of Technology, Japan. Later, he was a Research Scientist with the Memorial Sloan Kettering Cancer Center, USA. He also received an offer to join the University of Oxford, U.K., as a Senior Researcher of machine learning in medical imaging. He is currently an Assistant Professor with the Computer Science and Engineering Department, Independent University, Bangladesh. His research interests include using machine learning and whole slide image analysis techniques to gain insight into the treatment of cancer patients. Recently, his work has focused on unraveling the role of the HER2 (human epidermal growth factor receptor 2) gene in the growth of different cancers, such as breast, colon, and gastric. He is a member of the Association of Pathology Informatics, USA.

**KANIZ ROKSANA** received the B.Sc. degree in computer science and engineering from Daffodil International University, Bangladesh, in 2021. She is currently a Lecturer with the Department of Computer Science and Engineering, School of Science and Engineering, Uttara University (UU). Her research interests include machine learning, deep learning, medical image processing, and the Internet of Things (IoT) domain.

**IRMA DOMÍNGUEZ AZPÍROZ** is currently with Universidad Europea del Atlántico, Santander, Spain. She is also affiliated with Universidad Internacional Iberoamericana, Arecibo, PR, USA, and Universidade Internacional do Cuanza, Cuito, Bié, Angola.

**RAQUEL MARTÍNEZ DIAZ** is currently with Universidad Europea del Atlántico, Santander, Spain. She is also affiliated with Universidad Internacional Iberoamericana, Arecibo, PR, USA, and Universidad de La Romana, La Romana, Dominican Republic.

**IMRAN ASHRAF** received the Ph.D. degree in information and communication engineering from Yeungnam University, Gyeongsan-si, South Korea, in 2018, and the M.S. degree (Hons.) in computer science from Blekinge Institute of Technology, Karlskrona, Sweden, in 2010. He was a Postdoctoral Fellow with Yeungnam University. He is currently an Assistant Professor with the Information and Communication Engineering Department, Yeungnam University. His research interests include positioning using next-generation networks, communication in 5G and beyond, location-based services in wireless communication, smart sensors (LIDAR) for smart cars, and data analytics.

**MD. ABDUS SAMAD** (Member, IEEE) received the Ph.D. degree in information and communication engineering from Chosun University, Gwangju, South Korea. He was an Assistant Professor with the Department of Electronics and Telecommunication Engineering, International Islamic University Chittagong, Chattogram, Bangladesh, from 2013 to 2017. He has been a Research Professor with the Department of Information and Communication Engineering, Yeungnam University, South Korea. His research interests include signal processing, antenna design, electromagnetic wave propagation, applications of artificial neural networks, deep learning, and millimeter-wave propagation by interference and/or atmospheric causes for 5G and beyond wireless networks. He won the prestigious Korean Government Scholarship for his doctoral study.