Classification of Immunity Booster Medicinal Plants Using CNN: A Deep Learning Approach

Md. Musa(✉), Md. Shohel Arman(✉), Md. Ekram Hossain(✉), Ashraful Hossen Thusar(✉), Nahid Kawsar Nisat(✉), and Arni Islam(✉)

Daffodil International University, Dhaka 1207, Bangladesh
{musa35-1870,arman.swe,ekram35-1936,ashraful35-1908, nahid35-1889,arni15-8789}@diu.edu.bd

Abstract. Environment has blessed us with various kinds of plants. Some of them uses as resources of medicines as it is called medicinal plant. In Bangladesh medicinal plants are also known as Ayurveda, Homeopathy and Unani. Experts says medicinal plants can be very useful in the fight with recent pandemic which is Covid-19. As we know health of a body depends on its immune system, so it is important to keep immunity stronger. Strong immune system can be influential to any infectious virus, bacteria and pathogens. On the other hand, inactive one can get easily infected with virus and other illness. There are certain medicinal plants which reinforce our immunity. Therefore, classification of these medical plants is very important. For this classification we have collected leaf images for six different classes which's local names are Darchini, Tulshi, Tejpata, Sojne, Neem, Pathorkuchi. In this article we introduced a famous algorithm for classification named CNN (Convolutional neural network). We used CNN (Convolutional neural network) to recognize the plant from leaf images and got 95.58% accuracy. In future infectious virus can appear which can be more threatening than others, our research will help people to know about immune system and medicinal plants which reinforce our immunity, so that they can fight with diseases and viruses.

Keywords: Immunity system · Medicinal plant · Plant classification · Convolutional neural network

1 Introduction

Covid-19 has shown us the importance of our immune system. Immune is our most centric benefactor and its prime duty is to keep us healthy and strong. A strong immune system depends on which foods we are taking and if we take foods which are fresh and contain vitamins and minerals, our immune system will be capable of to do its battle against virus, harmful bacteria, toxins and other pathogens. The immune is not a single entity, it's a system. Balance and harmony are main requirements for function it well. Nature has blessed mankind with enormous medicinal herbs which provide timely and suitable remedies to various health disorders. In the time of health crisis like novel Corona virus, the medicinal herbs are enabling the people to boost their immunity

M. Singh et al. (Eds.): ICACDS 2021, CCIS 1440, pp. 244–254, 2021.
https://doi.org/10.1007/978-3-030-81462-5_23

[1]. Understanding the correlation among immune system, medicinal herbs and Covid-19 in the present time is very necessary. Adequate nutrition is the backbone for the development, maintenance and optional functioning of immune cells in this outbreak of Covid-19.

Based on the principle of plant classification, categorization of plants occurs with plant parts such as roots, flowers, fruits and leaves. However, in various incidents characteristics like roots, flowers and fruits don't show much difference in computationally and mathematically. However, for species identification leaves are reach resources and some of them boost our immunity and also used in medicinal treatment. World Health Organization (WHO) stated that on the purpose of remedies medicinal plants are used by 65% to 80% of the world population and there are around 17810 of different medicinal species in the world [2].

Images are collected from different nurseries and botanical garden and the number of images are 920. This image has six different classes. The local and scientific names of these plants are:

Table 1. Dataset information

No.	Scientific name	Local name	No. of samples
1	Cinnamomumverum	Darchini	150
2	Ocimumtenuiflorum	Tulsi	150
3	Cinnamomumtamala	Tejpatta	150
4	Moringaoleifera	Sojne	150
5	Azadirachtaindica	Neem	150
6	Bryophyllum pinnatum	Pathorkuchi	150
		Total	900

Around the world high percentage of people use medicinal plants leaves for their primary healthcare. Leaves has no side effects and enhancement use of the medicinal plants can reduce medicine side effects. Plants are not only used in medical science but it also uses as resources of cosmetics and other products also. Large numbers of industries are heavily dependent on medicinal plants as plants are their main resources of manufacturing. In this time these plants are in the threat of extinction because of deforestation, urbanization and lack of awareness in medicinal plants. According to the Food and Agriculture organization in 2002 the number of medicinal plants were 50,000 and in 2016 according to the Royal Botanic gardens the number is decreasing in high rate and the estimate number is now 17,810 [3]. So, it is high time to save these plants for our beneficial concern.

2 Literature Review

A lot of work is being done with medicinal plant images. In medical science plants has huge impact and it is considered as great asset. In 2016 D Venkatraman and mangayarkarasi N proposed an automated system based on vision approach which helps a common man on identifying the medicinal plants. Usually, plants are classified on leaves features like – shape, color and texture. Their classification was based on leaves texture. They used GLCM method for classifying the leaves and to find the dissimilarity between the leaves [4].

In year 2017 Adams Begue and VemitheKowlessur collected leaves of 24 medicinal plants species and exterted each leaf on their width, shape, color, length, perimeter, area and vertices number. They used Random Forest algorithm for their classification and gets 90.1% accuracy. Their anticipation for this automated recognition was encourage researchers to develop more techniques on species identification and help common people to know more about medicinal plants [5].

Plants are not only used in medical science it also used in several types of cosmetics and other products as well. In 2016, Prabhakar Poudel and Shyamdew Kumar proposed an automation system for leaf detection. Leaf extraction was based on diverse feature, categorization and pattern identification. Support Vector Machine (SVM) was applied for the classification and for extracting the features they utilized Scale Invariant Feature Transformation (SIFT) [6].

In 2019 C. Amuthalingeswaran had built a Deep Neural Networks model for the identification of medicinal plants. They used 8000 images for four classes and got accuracy of 85%. They collected their images from the open field land areas [7]. Nazish Tunio, For extraction of the region of the interest (ROI) and an algorithm which is image based was proposed by Abdul Latif Memon. It was also used for identify the plant species and to recognize the disease of certain plants as well. Support Vector Machine (SVM) classifier was applied for better outcomes and they got their expected result of 93.5% accuracy with their leaves dataset which had four different classes [8].

A. Aakif proposed an algorithm for identification of plants with three ways i) Preprocessing ii) Feature extraction iii) Classification in 2015. Several leaf feature was extracted and these features became input vector for their build model Artificial Neural Network (ANN). They trained their model with 817 leaf images from 14 different classes and got accuracy of 96% and for checking the effectiveness of the model they trained it with another dataset and also got accuracy of 96% [9].

In 2017 MM Ghazi used Deep Convolutional Network for identification the plants species. They applied popular deep learning architectures like GoogLeNet, AlexNet, and VGGNet and trained it with Life CLEF 2015 dataset. Their combined system had overall accuracy of 80% on validation set [10].

On the other hand, in 2013, R. Janani and A Gopal Artificial Neural Network (ANU) classifier to identify the medicinal plants from leaf images Artificial Neural Network (ANN) was used for less computational complexity and to gain high efficiency. The models give 94.4% of accuracy after testing on 63 leaf images and which was impressive [12].

C. Ananth and Azha. Proposed an automation system of plants identification because manual classification has huge chance of human error.In this research they use MATLAB. They extract leaf images and used models like segmentation, thresholding and applied to neural network [13].

In recent times (2019) digital image Processing for identification of medicinal plants was proposed by P. Chitra and S. JanesPushparani. Speeded Up Robust Feature Transform (SURF) and Scale Invariant Feature Transform (SIFT) both was used for leaf extraction and for distinguishing the structure. They used Support Vector Machine (SVM) classifier to get correct plant identification [14].

RAJANI S and VEENA M.N proposed an automatic identification and classification of medicinal plants. Create awareness and encourage people to know more about medicinal plant is their main goal of this automatic identification. Like many others they only used plant leaves images for the classification but they also used plant flowers, fruits and seeds images for their classification [15].

3 Proposed Methodology

For this study we use convolutional neural network (CNN) for plant detection. Our working procedure is in Fig. 1.

Fig. 1. Work flow

3.1 Dataset Collection

Various sorts of medicinal plants leaves are gathered from several nurseries and botanical garden from Dhaka city. Our dataset contains 900 images with 6 unique classes. For testing we utilize 200 images and other 700 images for training and validation. Now every one of them have 150 images for each. We have named the medicinal plants as their native names for our study. Dataset classes are (Fig. 2):

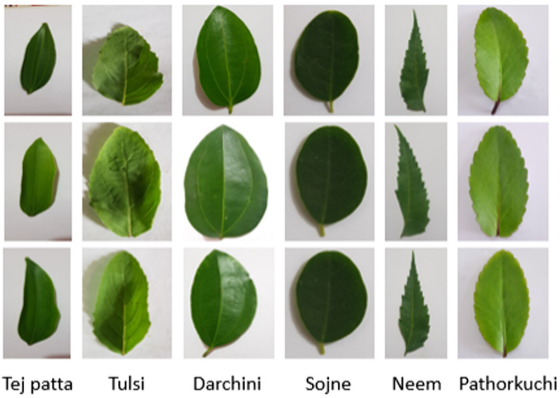

Fig. 2. Dataset sample

3.2 Data Preprocessing

3.2.1 Data Augmentation

In limited dataset overfitting is a typical issue. We may face overfitting problem as well as our dataset is also limited. we implement data augmentation to eliminate overfitting. Augmentation is actually an artificial process which extends the dataset. In our dataset Augmentation Technique modified version of images is absence. Augmentation was done four way and they are height shift range, rotate 90° left, flip horizontally and shear by certain amount.

3.2.2 Data Preparation

Subsequent to completing the increase interaction these data have various shapes. For training and test purpose our proposed model needs fixed shapes. Data was resized into 128 × 128 pixels and we used RGB for getting decent accuracy. For the purpose of testing, we separated 20% images and 80% images for training and per class were individual.

3.3 Convolutional Neural Network

The classification of Artificial Neural Network is an unimaginable mainstream approach to solving pattern recognition problem. Convolutional Neural Network (CNN) is named fully connected neural network and it was one of the vital elements adding to these tests. Out of many advantages of CNN there is a principal advantage which is without human guidance important features are automatically detected. CNN model has two segments starting fragment is for include extraction and the ensuing portion works for classification. The central layer will endeavor to recognize edges and shape a demonstrate to distinguish the edge. At that point or maybe layers may endeavor to solidify them in less complex ways. Include channel to the picture and attempt to extricate the picture edges are the primary layer's work. The second layers are called the pooling layers and just like the first layer, it too includes channels to the image. This layer finds highlights more

profoundly from the figure and the layer has ReLU usefulness that makes a difference to associate the taking after neuron. At that point, the flatten layer changes over 3D picture data to 2D data for classification [16].

3.3.1 Convolution Layer

This is called CNN's heart structure. The convolution layer computes the dab item among width, the tallness of the input layer, and the channels. It also contains filters. If the dimension of image is 128 × 128 × x3 and number of filters is 24 then dimension of output is 128 × 128 × 24 (Fig. 3).

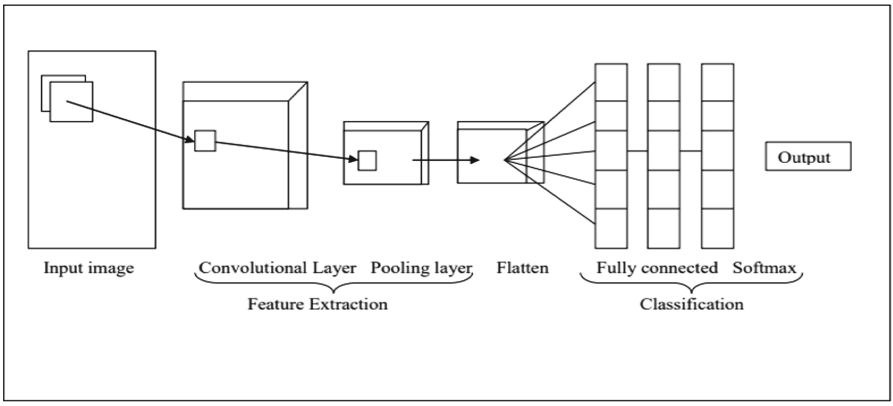

Fig. 3. Architecture of CNN

3.3.2 Activation Function Layer

Activation could be a sort of function and gives the related yield that's bolstered from any input. Functions like Direct and Nonlinear are out. Relu may be a nonlinear enactment work that's utilized in our consideration. ReLU ordinarily changes over negative numbers by setting them to zero.

$$F(x) = max(0, x)$$

3.3.3 Pool Layer

After the activation layer comes the Pool layer. A few sorts of pool layers are there. Max pooling is for the most part utilized pooling layer additionally utilized for our errand. The essential work of max-pooling is lessening the measurement of the pictures. In the event, that pool estimate 2 × 2 and walk 2 utilizing the max-pooling layer, at that point the yield measurement is 64 × 64 × 24. Figure 4 appears the visualization of a max-pooling operation.

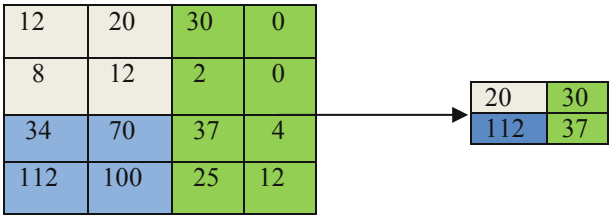

Fig. 4. 2 × 2 Max polling

3.3.4 Fully Connected Layer

This is the final layer of CNN. It composes a complete combination of all former layer and it takes input from the previous layers then makes into id vector.

3.4 Proposed Model

We have created an architecture to recognize the medicinal plants and it is a 12-layered network. The developed layers are –

Convolution Layer: In our model we apply four convolution layer and they are:

i. Convolution (32 – 3 × 3) filter.
ii. Convolution (64 – 3 × 3) filter.
iii. Convolution (128 – 3 × 3) filter.
iv. Convolution (256 – 3 × 3) filter.

Pooling Layer: Pooling layer come after convolutional layer. We use four max pooling layers with a pool size of 2 × 2.

Dropout Layer: In neural network there are dropping out some units and it is called dropout. Dropout is use for avoid overfitting. Our proposed model has two dropout layers with layer rate of 0.25 and 0.5.

Flatten Layer: The 2D vector includes a pool highlight outline and this vector makes a continuous 1D vector where the method is called flatting. Our show contains a flatten layer.

Dense Layer: The dense layer utilized as a classifier.

3.5 Training the Model

We utilized an optimizer known as Adam [17] optimizer to plan our show with a little learning rate. It is likely the foremost utilized optimizer that's quick and dependable to utilize. For 15 epochs of 35 batch, the size we first ought to compile our model at that point begin our model utilizing the fit () strategy. The issue is that multi-class classification is for classification cross-entropy loss (Fig. 5).

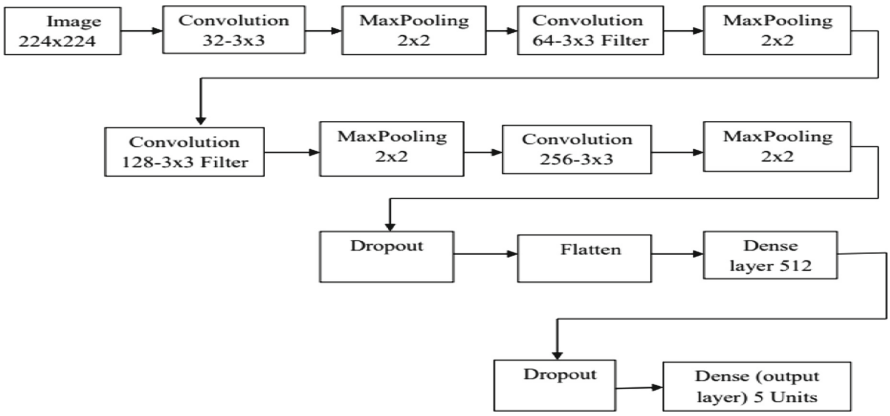

Fig. 5. Proposed model architecture

4 Performance Evaluation

Model assessment is exceptionally imperative in machine learning. This appears how proficient the scoring of the dataset has been by a prepared model. We accomplished our result by completing preparation and approval. This is often called preparing exactness when we apply the show to the preparing data and it is called validity accuracy when we apply the show to the test data with distinctive classes.

Figure 6 appears that the training accuracy of our model is 95.58% and we kept up our approval exactness from 92% to 94.5%.

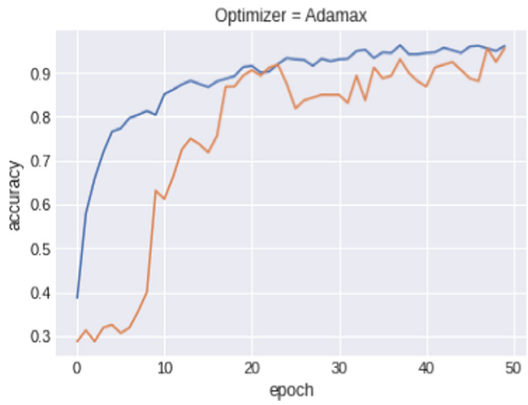

Fig. 6. Training vs validation accuracy

In Fig. 7 we can see the training and validation loss of our model. 0.08 is our validation loss.

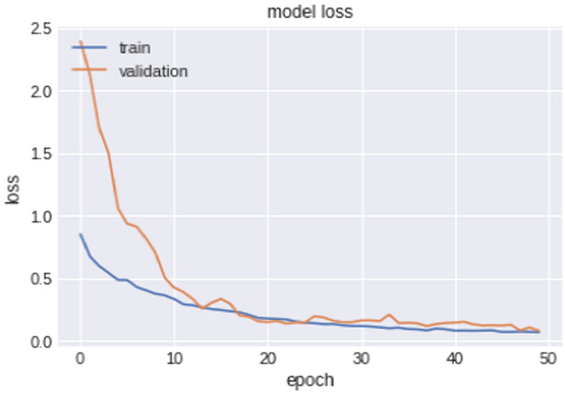

Fig. 7. Training vs validation loss

5 Result and Discussion

The classification report visualizer appears the accuracy, precision, F1-score, and bolster scores for the model. To assist less complex interpretation and issue revelation, the report arranges scientific scores with a shading coded heatmap. All heatmaps are within the reach (0.0, 1.0) to empower a straightforward comparison of classification models over different arrange reports. After completing the training strategy our model has been tested with 200 data. Table 2 appears the classification report which is computed from our model. We assessed precision, recall, and f1-score.

Table 2. Classification report

Class	Precision	Recall	F1-score
Darchini	0.87	1.00	0.93
Tulsi	1.00	0.93	0.96
Tej patta	1.00	1.00	1.00
Neem	0.94	0.97	0.95
Pathorkuchi	1.00	0.87	0.93
Sojne	0.90	0.93	0.92
Average	0.95	0.95	0.95

Confusion matrix of our model is shown in Table 3.

Table 3. Confusion matrix

	Darchini	Tulsi	Tej patta	Neem	Pathorkuchi	Sojne
Darchini	27	0	0	0	0	0
Tulsi	0	27	0	2	0	0
Tej patta	0	0	31	0	0	0
Neem	0	0	0	30	0	1
Pathorkuchi	2	0	0	0	27	2
Sojne	2	0	0	0	0	28

6 Conclusion

In this study, our presented model appears great precision for helpful plant leaf arrangements. We advertised a model by our claim. This method worked well within the field pictures that we attempted. This work makes a difference Botanists get its plants related to their long-term treatment. It can at that point exceptionally well be utilized for reestablishing takes off and bunches in pertinent ranges where leaf arrangement is required.

7 Future Work

For ages, various sorts of therapeutic plants have been found all over the world. Each one of them has its characteristics that remedy an assortment of afflictions. Along these lines, within the future, we have to make this dataset greater with diverse medicinal plant leaf pictures and proposed a (Convolutional Neural Organize) CNN with more upgrades. We make superior treatment solutions utilizing modern propels and movement in computer vision. Within the future individuals will utilize more medicinal herbs clears out as contradicted to medication for the treatment.

References

1. http://www.nytimes.com/2010/07/06/health/06real.html?pagew
2. Arun, C.H., Sam Emmanuel, W.R., Christopher Duraira, D.: Texture feature extraction for identification of medicinal plants and comparison of different classifiers. Int. J. Comput. Appl. **62**(12) (2013). 0975 - 8887
3. O'Donnell, K., Sharrock, S.: Botanic gardens complement agricultural gene bank in collecting and conserving plant genetic diversity. Biopreserv Biobank. **16**(5), 384–390 (2018). https://doi.org/10.1089/bio.2018.0028
4. Venkataraman, D., Mangayarkarasi, N.: Computer vision based feature extraction of leaves for identification of medicinal values of plants. In: IEEE International Conference on Computational Intelligence and Computing Research. IEEE (2016). 978-1-5090-0612-0/16/$31.00 ©2016

5. Begue, A., Kowlessur, V., Mahomoodally, F., Singh, U., Pudaruth, S.: Automatic recognition of medicinal plants using machine learning techniques. Int. J. Adv. Comput. Sci. Appl. **8**(4) (2017)
6. Poudel, P., Kumar, S.: Robust recognition and classification of herbal leaves. **05**(04) (2016). ICESMART-2016
7. Amuthalingeswaran, C., Sivakumar, M., Renuga, P., Alexpandi, S., Elamathi, J., Hari, S.S.: Identification of medicinal plant's and their usage by using deep learning. In: 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, pp. 886–890 (2019). https://doi.org/10.1109/ICOEI.2019.8862765
8. Tunio, N., Memon, A., Khuhawar, F., Abro, G.M.: Detection of infected leaves and botanical diseases using curvelet transform (2019)
9. Aakif, A., Khan, M.: Automatic classification of plants based on their leaves. Biosyst. Eng. **139**, 66–75 (2015). https://doi.org/10.1016/j.biosystemseng.2015.08.003
10. Ghazi, M.M., Yanikoglu, B., Aptoula, E.: Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing **235**(C), 228–235 (2017)
11. Jain, G., Mittal, D.: Prototype designing of computer-aided classification system for leaf images of medicinal plants. J. Biomed. Eng. Med. Imaging **4**(2), 115 (2017). https://doi.org/10.14738/jbemi.42.3053
12. Janani, R., Gopal, A.: Identification of selected medicinal plants leaves using image features and ANN. In: 2013 International Conference on Advanced Electronic System (ICAES) (2013). 978-1-4799-1441-8
13. Ananthi, C., Periasamy, A., Muruganand, S.: Pattern recognition of medicinal leaves using image processing techniques. J. Nano Sci. Nano Technol. **2**(2) (2014). Spring Edition, ISSN 2279 – 0381
14. Chitra, P.L., James Pushparani, S.: SURF points versus SIFT points in identification of medicinal plants. Int. J. Eng. Adv. Technol. (IJEAT) **9**(2) (2019). ISSN : 2249 -8958
15. Rajani, S., Veena, M.N.: Study on identification and classification of medical plants. Int. J. Adv. Sci. Eng. Technol. **6**(2), Spl. Issue-2 (2018). ISSN(p): 2321 –8991, ISSN(e): 2321 –9009
16. Bengio, Y., Lecun, Y.: Convolutional networks for images, speech, and time-series (1997)
17. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. CoRR, abs/1412.6980 (2014)