Analysis of Performance of ML Algorithms in Detection of Flowers

[1]Prof. Satish J. Manje, [2]Mr. Pawankumar Ravish, [3]Mr. Sujit Yadav, [4]Mr. Deep Marade, [5]Ms. Apeksha Malgaonkar

[1]Asst. Prof. ,[2,3,4,5]UG Student,[1,2,3,4,5]Computer Engg. Dept. Shivajirao S. Jondhle College of Engineering & Technology, Asangaon, Maharashtra, India.

[1]satishmanje93@gmail.com, [2]ravishpawan10@gmail.com, [3]sujity925@gmail.com, [4]deepmarade@gmail.com, [5]apeksham47@gmail.com.

Abstract - In many areas of agriculture and medical research, it is highly helpful to identify different kinds of flowers and leaves based on their properties. In this article, machine learning techniques are used to identify flowers based on their features. A data set of flowers is used to apply the machine learning algorithms K-nearest neighbor, Random Forest, and Decision Tree, and their precision is calculated. Python is used to implement algorithms on a data set. It has been discovered that the KNN machine learning algorithm performs best in the detection of flowers.[1]

Keywords- Machine Learning, K-Nearest Neighbor, Flower Identification, Random Forest, Decision Tree.

I. INTRODUCTION

Machine learning is a technique used in artificial intelligence (AI) to discover patterns in data. Machine learning is a method of artificial intelligence that enables machines to learn just like humans. It makes use of some computational algorithms to provide the machine the being able to learn. In the extremely dynamic corporate environment, ML enables businesses to accelerate digital innovation and enter the age of automation. Some may also contend that AI/ML is essential for some industries to stay relevant, such as financial fraud detection, digital payments, and product recommendations. Machine learning that is supervised involves, providing the computer with properly labelled data. It is using labelled data to teach the computer. Unsupervised learning involves training an unregistered, unclassified machine. Reinforcement learning is the most widely used form of machine learning. There are no solutions in this kind of learning process instead, the agent determines how to use the data. This agent picks up knowledge via examples. If the action taken was successful, it attempts to maximize it; if it wasn't, it corrects itself.[1]

II. AIMS AND OBJECTIVE

a) Aim

The purpose of this paper is to implement an automated method for identifying various species of flowers. Implementing an automated flower recognition system has several applications, including quick recognition for educational purposes because automated methods accelerate learning. People with little knowledge of flower species could identify a flower's species according to automated flower recognition. It aims to provide accurate resutls with high precisions.[3]

b) Objective

- The main objective of this paper is to offer information about the species of flowers, including their identity, life forms and group which is generated depending on the type of flower species.
- Recognizing the image by using techniques that have been evolved using training and testing data.
- Using textural attributes of the image to generate a histogram to accurately and precisely identify each flower.[3]

III. LITERATURE SURVEY

Paper 1: Identification of flower diseases using artificial neural networks:

Flower disease is an example of potential difficulties that could arise during the cultivation process. The automatic diagnosis of flower diseases presented in this paper is using image processing methods. Images of healthy and diseased flowers are gathered, and the regions of interest are then pre-processed and split to create a knowledge base. With the aim of reducing the dimensionality of the features, they computed seven different measures of dispersion and central tendency from the textural features of the images using Gabor feature extraction. Training is carried out using an artificial neural network using eight output nodes that reflect the eight kinds of diseases taken into consideration in this work and seven input features retrieved from individual photos. Then, unknown samples of flower photos are

 DOI : 10.5281/zenodo.7857411

examined based on the training model, and they were able to identify the floral diseases with a mean rate of accuracy 83.3%.[4]

Paper 2: Flower species identification and coverage estimation based on hyper spectral remote sensing data:

Accurately tracking grass species and coverage contributes significantly to research on species diversity and the long-term health of the grassland ecosystem. Grassland plants typically have distinctive spectral traits in florescence. Species identification is simpler during florescence than it is during the nutrient stage. In this work, flowers such as Hemerocallis citrina Baroni, Galium verum Linn., Clematis hexapetala Pall., Lilium concolor var. pulchellum, Serratula centauroides Linn., Artemisia frigida Willd and Lilium pumilum. were identified using optical spectroscopy analysis and feature extraction techniques. Validation demonstrates that the accuracy of identification methods will exceed 90% when flower coverage is greater than 10 %. The results exhibit that the linear unmixing model, with a mean retrieval error of about 4%, is a useful technique for estimating the coverage of grassland flowers.[5]

Paper 3: Herb flower recognition system (HFRS):

The purpose of this study is to develop a Minimum Distance Method-based automatic method for identifying Thai herb flowers. The digital camera images of the herbs and flowers were captured in their natural surroundings. Our classification algorithms are created using the traits of herb flowers, which include the size of the herb flower, the edge of the petals, and the typical RGB colors. More than 380 images representing 16 different species of flowering herbs are used in the experiments. There are about 220 pictures in the training data set. A training data set of 110 images and a test data set of 50 images are used to evaluate the system.[6]

IV. EXISTING SYSTEM

Flowers are present in a range of shades, making it challenging to differentiate between them. A color-based algorithm is presented by researchers, although it is challenging to separate different flower species only based on the color because numerous flowers of the identical species and variety share shades. SVM is used to identify flowers, however it fails miserably. As the SVM process utilizes the slowest and yields the least accurate results in comparison to the other machine learning algorithm. The major drawbacks of the current system are the data collection issue, SVM algorithm's highest time consumption, and lowest accuracy.[2]

V. COMPARATIVE STUDY

Sr No.	Author	Project Title	Publication	Technology	Purpose
1.	Yaregal Assabie Getahun Tigistu	Automatic identification of flower diseases using artificial neural networks	IEEE, 2015	ANN	Applying image processing techniques to automatically identify diseases in flowers.
2.	Wenjie Fan, Xiru Xu, Yuanzhen Zhang, Yingying Gai.	Flower species identification and coverage estimation based on hyperspectral remote sensing data	IEEE,2011	Hyperspect-ral remote sensing.	Estimating the mean retrieval error to predict the grassland flower coverage
3.	Ponrath Sakunreraratsae, Chomtip Pornpanomchai,	Herb flower recognition system (HFRS)	IEEE,2010	Minimum Distance Method	To generate a Minimal Distance Method-based automated flower recognition system.

Table.1: Comparative Study

VI. PROBLEM STATEMENT

Several websites and cellular applications were developed and after analysis, it was found that some of them offer their services only to customers who purchase their premium editions and the ones which are free had very poor execution. Some of them are mainly focused on identifying a single characteristic of flowers, while others have terrible user interfaces that are extremely difficult for the average user to grasp.[8]

VII. PROPOSED SYSTEM

It provides wea method for automatically classifying flowers based on an image's rgb characteristics, which are constructed using matplotlib in Python and the decision trees of the random forest classifier, utilizing computer vision and machine learning methods. Thus, the approach makes the process of classifying flower species quick, simple, and substantially less expensive, and it may be adopted in the study of botany by scientists or botanist without the burden of learning specifics about the flower species. When a flower is photographed, the system accurately and quickly recognizes the species of flower and provides the user with information about the flower.[1]

VIII. ALGORITHM

1. Import necessary libraries:

import os, np, matplotlib.pyplot, plotly.graph_objs, tensorflow.keras

 DOI : 10.5281/zenodo.7857411

2. Set up paths to the image directories:

Images_path = ‘./flower_images/’ dirs = os.listdie(images_path) labels = sorted(dirs)

3. Get the number of images in each flower directory:

num_images = [ ] for d in dirs: num_images.append(len(os.listdir(images_path + d)))

4. Get the dimensions of all images in all directories:

dims = [ ]

for d in dirs: for f in os.listdir(images_path + d):

dims.append(imread(images_path + d + '/' + f).shape)

5. Define the image shape and get a sample image:

img_shape = (128, 128, 3) sample_img = imread(images_path + dirs[0] + '/' + os.listdir(images_path + dirs[0])[0])

6. Define the image data generator for the training and validation data:

datagen = ImageDataGenerator(rescale=1./255, shear_range=0.2, zoom_range=0.2, horizontal_flip=True, validation_split=0.2)

7. Define the KNN model:

model = Sequential([Conv2D(32, (4, 4), activation='relu', input_shape=img_shape), MaxPool2D((2, 2)), Conv2D(64, (4, 4), activation='relu'),MaxPool2D((2,2)),Flatten(),Dense(64,activation='relu'),Dropout(0.5),Dense(len(labels),activation='softmax')])

8. Evaluate the performance of the KNN model:

y_true = test_dataset.classes

y_pred=np.argmax(model.predict(test_dataset), axis=-1) report = classification_report(y_true, y_pred, target_names=labels) print(report)

IX. MATHEMATICAL MODEL

The mathematical equations for flower detection using machine learning algorithms such as Random Forest, and K-Nearest Neighbors are as follows:

1. Decision Tree:

Decision Tree is a tree-based model that uses a tree-like structure to represent Decisions and the potential effects. The mathematical equation for Decision Tree can be represented as:

f(x) = g(T, x)

where f(x) is the final prediction, g is a function that maps the input features x to a leaf node in the decision tree T, and T is the decision tree.

2. Random Forest:

An ensemble learning technique called Random Forest makes predictions by using numerous decision trees. The predictions of the individual trees are combined to produce a final prediction. The mathematical equation for Random Forest can be represented as:

f(x) = mode (f1(x), f2(x), ..., fN(x))

where f(x) is the final prediction, mode is the statistical mode (i.e., the value that appears most frequently in a data set), fi(x) is the prediction of the itch decision tree, and N is the total number of trees in the forest.

3. K-Nearest Neighbors (KNN):

KNN is a non-parametric and instance-based

learning method that makes predictions based

on the k-nearest neighbors in the training set. The mathematical equation for KNN can be represented as:

f(x) = mode(y_i), where i ∈ {1, 2, ..., k}

where f(x) is the final prediction, mode is the statistical mode, y_i is the label of the ith nearest neighbor, and k is the number of neighbors to consider.

In all three cases, the input features x are the measurements or attributes of the flower, and the output f(x) is the predicted class of the flower (e.g., species, color, etc.).

X. SYSTEM ARCHITECTURE

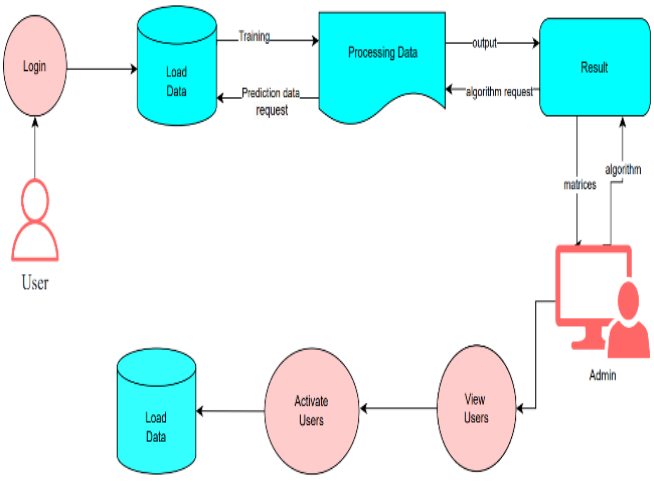

Fig.1: System Architecture

Description:

1)The user will enter the system by logging in. 2)User authentication will be done by adminstrator. 3) The user has the option to submit a flower picture or browse to live prediction. 4)The system uses trained model and three ml algorithms which are Random Forest, Decision tree and KNN algorithm. 5) The uploaded flower image's prediction will be shown as final output.

 DOI : 10.5281/zenodo.7857411

XI. ADVANTAGES

- The ability to recognize a flower kind is helpful in a variety of professions including farming, Ayurveda, gardening, botanical study, and floriculture.[7]
- With simply a flower photograph, this program may be used to distinguish a specific species of flower from millions of others.
- GUI is user-friendly and visually attractive.
- It uses less computing resourses.

XII. DESIGN DETAILS

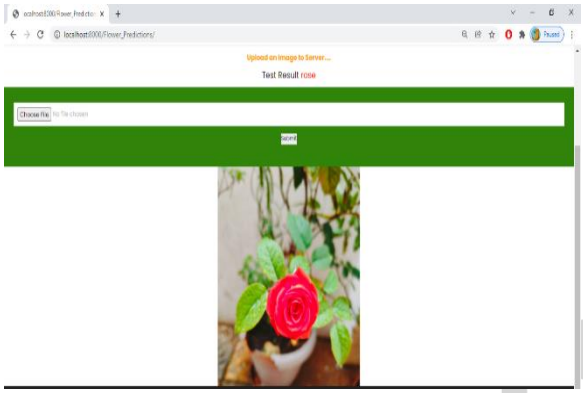

Fig 2: Result

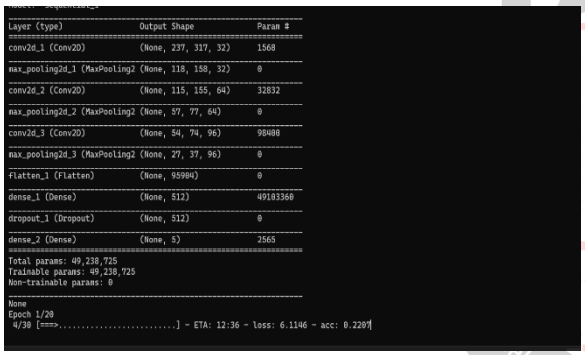

Fig 3: Model summary

The model was trained using 20 epochs and a batch size of 30. The classification report that was received during the training and validation phase is displayed in fig 3. The training loss, estimated time of arrival, and training accuracy are displayed for each period. 90% overall accuracy was reached by the model. An accurate prediction with 98.46% accuracy was made using a real-time photograph of a rose captured on a mobile device as input.

XIII. CONCLUSION

Thus, we have tried to implement the paper "Analysis of Performance of ML Algorithms in Detection of Flowers", "Vinod Jain and Anupam Yadav", IEEE 2021, and the conclusion is that although various ML techniques, including Random Forest, Decision Trees, and KNN, can be used to analyse flowers, not all of them yield great accuracy. According to the comparisons, The Knn method is found to be more accurate than the Random Forest method and the Decision Tree Algorithm based on the accuracy we gained in the detection of flowers. These projects can be updated in the future with richer datasets and cutting-edge technology. Henceforth, the above project was implemented successfully.

REFERENCE

[1] Vinod Jain; Anupam Yadav, "Analysis of Performance of Machine Learning Algorithms in Detection of Flowers", Proceedings of the Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV 2021). IEEE Xplore Part Number: CFP21ONG-ART; 978-0-7381-1183-4 INSPEC Accession Number: 20513090 DOI: 10.1109/ICICV50876.2021.9388599

[2] Prof. Vishal R. Shinde, "Image Processing on Watermark" in IJREAM, ISSN:2494-9150 Special Issue 01, MAR 2016."

[3] Ravi Verma, "Flower Species Classification using Random Forest Classifier", International Journal of Combined Research & Development (IJCRD) eISSN:2321-225X; pISSN:2321-2241 Volume: 5; Issue: 8 August -2016

[4] G. Tigistu and Y. Assabie, "Automatic identification of flower diseases using art- ificial neural networks," *AFRICON 2015*, Addis Ababa, 2015, pp. 1-5. doi: 10.1109/AFRCON.2015.7332020

[5] Y. Gai, W. Fan, X. Xu and Y. Zhang, "Flower species identification and coverage estimation based on hyperspectral remote sensing data", *2011IEEE International Geoscience and Remote Sensing Symposium*, Vancouver, BC, 2011, pp. 1243-1246. doi:10.1109/IGARSS.2011.6049424

[6] C. Pornpanomchai, P. Sakunreraratsame, R. Wongsasirinart and N.Youngtavicha- vhart, "Herb flower recognit ion system (HFRS)," *2010 International Conference on Electronics and Information Engineering*, Kyoto, 2010, pp. V1-123-V1-127. doi: 10.1109/ICEIE.2010.5559906

[7] Deepti.Aralikatti, Vinuta.Bangarshettar, Ashwini.Koti , Shree.Halbhavi, "Automated Flower Species Detection and Recognition using Neural Network", International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Published NCAIT - 2020 Conference Proceedings

[8] W. Liu, Y. Rao, B. Fan, J. Song and Q. Wang, "Flower classificat ion using fusion descriptor and SVM," 2017 International Smart Cities Conference (ISC2), Wuxi, 2017, pp. 1-4. doi:10.1109/ISC2.2017.8090865

 DOI : 10.5281/zenodo.7857411