FLOWER IDENTEFICATION

Abdelrahman Ahmed Ramzy, Ahmed Fawzi Hosni & Moaz Khairy Hussien

Abstract

There are hundreds of thousands of species of flowering plants around the world, even the keenest botanist or flower enthusiast in the world would never be able to recognize and differentiate between all these species.

To be able to help the world correctly recognize and differentiate between these flowers without the need to look it up in a book and in a fast and orderly manner, science has intervened in this task at hand and that is where our project comes in.

We tried to make a prototype software application based on computer vision and machine learning algorithms to correctly identify flower images based on a set of pre-identified images used to train the used classifier.

We were able to produce an accuracy of 67.5% which is acceptable given the dataset capacity at hand.

I. Introduction

a. Problem Definition

There are approximately 369,000 species of flowering plants around the world [1], which poses a very troublesome task for a single person to be able to correctly identify and memorise all of them.

To make use of the technologies at hand, various engineers and computer scientists have been publishing papers and research solutions for classifiers that can correctly recognize these flowers and identify any new image given to that said classifier. To be able to correctly identify or differentiate between any number of flowers, a huge dataset is required as there are various flowers that have the same shape, colour, texture and even all of these together but differ in the number of petals.

b. Others work

There is a team that collected 2.6 million flower images from various contributions around the world that is accessed by anyone to contribute to, and with that, they produced an accuracy of 90% [2].

Another team (*Amira Ben Mabrouk et al.*), used the Oxford university dataset with 13 classes each holding 45 images to a total of 585 image and were able to reach an accuracy of 84.01% in some cases using the CIELAB colour space and Speeded-Up Robust Features (SURF) as features [3].

II. Materials and Methods

a. Materials

To implement our coding at hand, we used the Python programming language and a GUI by PyQt5. Our algorithm was implemented using the “Flower Color Images” dataset on Kaggle, which held 210 images for 10 different species, and added 40 images to make each specie have an equal number of images of 25 [4].

b. Methods

In order to correctly classify flowers, four main features were used [5-9]

- Haralick texture features.
- Image moments using Hu Moments.
- Colour histogram in the HSV colour domain.
- Colour histogram in the LAB colour domain.

We divided the dataset into 84% training and 16% testing, then applied the classifiers on the datasets to see which gave the best accuracy. The overall steps were followed by a repository given in the references [10].

The workflow of the overall software can be seen in Figure 1.

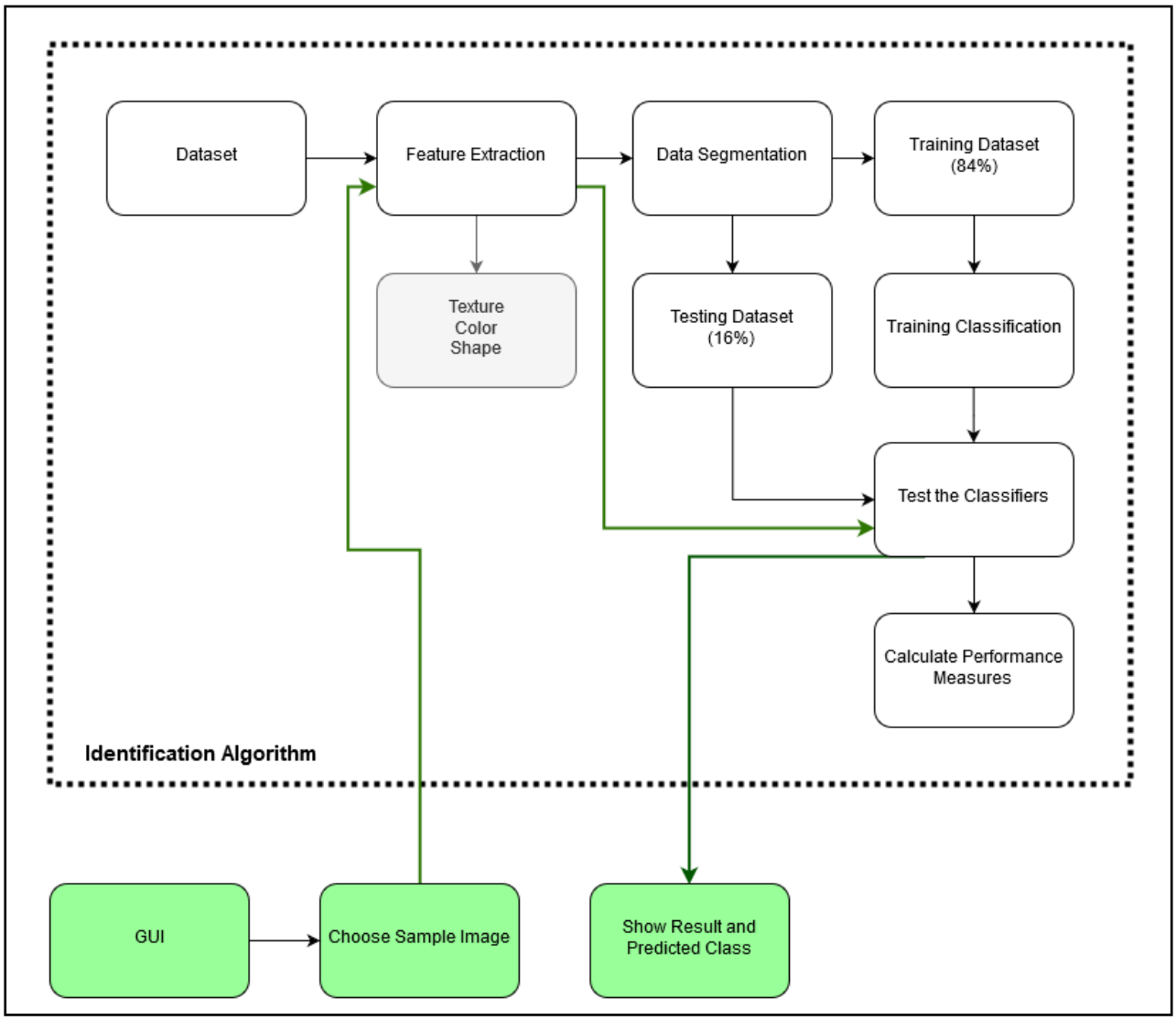

Figure 1 Block Diagram of the overall System

c. Classifiers

We applied the following classifiers on the dataset to find the one that yields the highest performance.

Logistic Regression (LR), Linear Discriminant Analysis (LDR), K-Nearest-Neighbours (KNN), Decision Tree (DTC), Random Forest (RF), Gaussian-NB (NB) and Support Vector Machine (SVM).

III. Results

The performance results can be seen in Figure 2 as a comparison among the seven used classifiers.

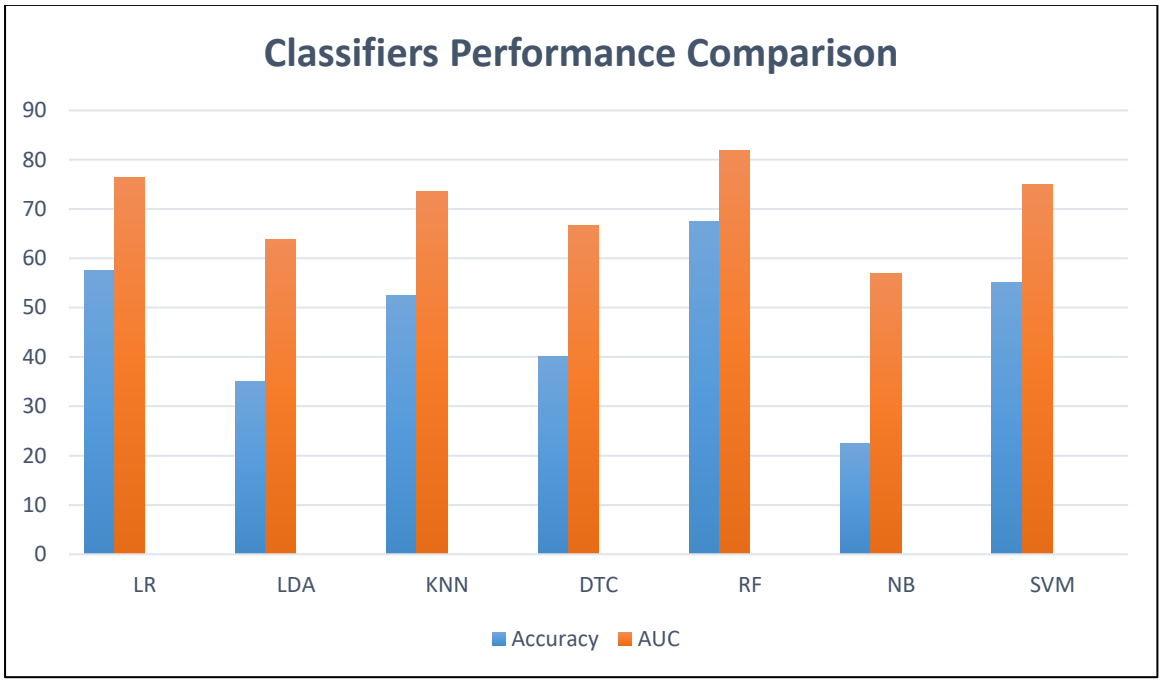

Figure 2 Classifiers Performance Comparison

IV. Discussion

From the above results we can see that the highest performing classifier was the RF classifier with an accuracy of 67.5%, while the NB was the poorest performing one with an overall accuracy of 22.5%. The highest AUC was 81.9% for the RF classifier while the lowest was 56% for the NB classifier. So, the overall performance was highest for the Random Forest while the Gaussian produced the lowest performance.

V. References

[1] https://news.mongabay.com/2016/05/many-plants-world-scientists-may-now-answer/

[2] https://www.i-programmer.info/news/105-artificial-intelligence/9951-machine-learning-for-flower-recognition.html

[3] Mabrouk, Amira & Najjar, Asma & Zagrouba, Ezzeddine. (2014). Image Flower Recognition based on a New Method for Color Feature Extraction. VISAPP 2014 - Proceedings of the 9th International Conference on Computer Vision Theory and Applications. 2. 10.5220/0004636302010206.

[4] https://www.kaggle.com/olgabelitskaya/flower-color-images

[5] https://en.wikipedia.org/wiki/Image_moment

[6] https://www.youtube.com/watch?v=O-hCEXi3ymU

[7] https://www.youtube.com/watch?v=QanGMHlFn5o

[8] http://murphylab.web.cmu.edu/publications/boland/boland_node26.html

[9] https://www.learnopencv.com/color-spaces-in-opencv-cpp-python/

[10] https://gogul09.github.io/software/image-classification-python#feature-extraction