Flower Identification Using Machine Learning

**Submitted by**

Name: Md Hasibur Rahman

Course Code: CSE 316

Course Title: Research and Innovation

Id: 211-15-14616

Sec: 58\_A

Submission date:

**Submitted to**

Dr. Arif Mahmud, Associated Professor, Department of CSE, DIU.

**ABSTRACT**

Beauty is incomplete without flower. Bangladesh is the land of flower. In our everyday life, On the way of walking, beside the rail line or in our garden we used to see a lot of flowers. But in most case, we have no knowledge about that flower. Even we don’t know its name. In that case we choose this idea to research and develop our project. That will introduce people about that unknown flower which they see but don't know about that. Our developed application recognizes the flower in real time by using mobile camera. This project is an attempt at using the concepts of neural networks to create an image classifier by TensorFlow on Android platform. Convolutional neural networks are a popular realm of machine learning and are widely used in image classification. So that we choose this topic to research about image classification by CNN and TensorFlow. Though we developed an application of our research project, but we have a lot of plans to do more research to upgrade our system.

**Keywords:** TensorFlow, CNN

**TABLE OF CONTENTS**

Abstract and Keywords

Table of Contents

1 Introduction.

1.1 Introduction

1.2 Background Problem (with table/chart/graph)

1.3 Motivation

1.4 Problem Statement

1.5 Research Questions

1.6 Research Objectives

2 Literature Review

2.1 Theories and Model

2.1.1 Basics of Theories

2.1.2 Explanation of Variables

2.2 Related Works

2.3 Research Summary

2.4 Scope of The Project

2.5 Challenges

3. Methodology

3.1 Introduction to Machine Learning

3.2 Convolutional Neural Network

3.3 Design of Dataset

3.4 Implementation Requirements

3.5 Design Development

4 Time plan

5 References

APA-7 formatting

**1.1 INTRODUCTION**

Flower is a very important part of nature. Mostly we identify a plant through its flower. Experienced botanists do this identification of flower but a naive person will have to consult flower guidebooks or browse any relevant web pages on the Internet through keywords searching. Our system can recognizes the flower in real time using mobile camera. Currently this Android app can identify around 10 flowers. Most important thing is that this app can fully works in offline. We are continuously working to add more flowers to identify. Everyday we see a huge number of flower species in our house, parks, roadsides, in farms, on our rooftop but we have no knowledge of that flower species or their origin. Even we have no idea about its name. There are several guidebooks for flowers knowledge but it becomes quite difficult to find the name when have the picture. Even the Internet sometimes is not useful. But it is quite difficult for human brain to memorize all the species they see. Even some flower is similar to look at. This application recognizes the flower in real time by using mobile camera. The purpose of this project is to use Tensorflow, an open-source dataflow and machine learning library, to build an image classifying Convolutional Neural Network (CNN) for classifying the flower image. Tensorflow, in addition to providing developers a simple way to build neural network layers, can also be run on mobile platforms such as Android. The ultimate goal of this project is to design and optimize a convolutional neural network for use with flower classification, and eventually build a simple classification app for mobile devices around the trained network. The mobile app will allow users to try and classify flowers while outdoors or offline [1]

**1.2 Background Problem**

Since the invention of computer, it has begun to change our daily life. It improves the quality of our lives. An impressive idea is to let a computer think and learn as like human being. Basically machine learning developed skills to develop the computer itself with the knowledge given. There are many applications of CNNs, one of the most popular of which is image classification, and CNNs have even been used specifically for flower classification. These flower classification neural networks, however generally use datasets with very specific details. Thus, this project is explorative and aimed at learning how to design a neural network using Tensorflow, but ultimately has practical applications for developers, botanists, or nature enthusiasts. The dataset used for training was collected through Google Images, and the images of the plants themselves are very general, rather than of specific characteristics of the plants, and typically include most of the identifying characteristics of a flower [3].

**1.3 Motivation**

There have a special motivation behind this research and project work. Me and my cousin used to walk along the railway line in afternoon. That time we used to observed a lot of unknown flower around the area. We were curious about that flowers but we could not recognize it. That time we discussed about an idea to make a system which can identify our desired flower automatically. Since then our journey has started and dream have come true now. We developed an android application which can introduce people about that unknown flower which they see but don't identify that. Hope our research & developed project will be helpful for the curious people.

**1.4 Problem Statement**

Flowers play a crucial role in maintaining the balance of ecosystems, serving as key indicators of biodiversity and environmental health. However, manual flower identification for ecological studies and conservation efforts is time-consuming and often requires specialized knowledge. As a result, there is a growing need for automated flower recognition systems that leverage advancements in computer vision and machine learning to streamline the identification process.

The existing literature on flower recognition primarily focuses on datasets with limited diversity, making it challenging to develop robust models applicable to real-world ecological settings. Additionally, variations in lighting conditions, background clutter, and the presence of multiple flower species within a single image pose significant challenges for accurate and reliable recognition.

This research aims to address these challenges by developing an innovative flower recognition system capable of identifying diverse flower species in their natural habitats. The system should be resilientto variations in lighting and background, and it should accommodate the simultaneous presence of multiple species within a single image. Furthermore, the research will explore the integration of emerging technologies such as transfer learning, deep neural networks, and edge computing to optimize the recognition process for field applications where computational resources may be limited.

The successful implementation of an automated flower recognition system has the potential to revolutionize ecological monitoring and conservation efforts by providing researchers, conservationists, and citizen scientists with a tool that facilitates rapid and accurate flower species identification. This, in turn, will contribute to a deeper understanding of ecosystem dynamics, aid in the preservation of endangered species, and support the overall goal of sustainable biodiversity management.

**1.5 Research Questions**

1. Can a machine identify flower before the human eye?
2. Can a system recognize different color of same flower using machine learning? Ex: Pink Rose / White Rose
3. Is convolutional neural network (CNN) technique being perfect to identify flower accurately?

**1.6 Research Objectives**

1. Dataset Compilation and Augmentation:

- Curate a comprehensive and diverse dataset of flower images encompassing various species and environmental conditions.

- Augment the dataset to simulate real-world scenarios, accounting for variations in lighting, background clutter, and the simultaneous presence of multiple flower species in a single image.

2. Algorithm Development:

- Investigate and implement state-of-the-art computer vision and machine learning algorithms for flower recognition.

- Explore the effectiveness of transfer learning techniques to leverage pre-trained models and optimize recognition accuracy.

3. Robustness to Environmental Variability:

- Enhance the system's resilience to variations in lighting conditions, diverse backgrounds, and other environmental factors commonly encountered in natural habitats.

4. Multi-Species Recognition:

- Develop and evaluate the system's capability to accurately identify multiple flower species within a single image, reflecting the complexity of real-world ecological settings.

5. Model Interpretability:

- Incorporate mechanisms for model interpretability to provide insights into the decision-making process, enabling users to understand how the algorithm arrives at specific flower identifications.

6. Edge Computing Integration:

- Explore the feasibility of implementing the flower recognition system on edge devices, ensuring its applicability for field research where computational resources may be limited.

7. Performance Evaluation and Benchmarking:

- Conduct rigorous performance evaluations using standard metrics to benchmark the developed system against existing flower recognition approaches.

8. User-Friendly Interface:

- Design and implement an intuitive user interface for the flower recognition system, facilitating easy integration into research workflows and encouraging adoption by a diverse user base, including researchers, conservationists, and citizen scientists.

9. Field Testing and Validation:

- Validate the system's performance in real-world ecological settings by conducting field tests and comparing its results with ground truth data collected by domain experts.

10. Contribution to Conservation and Ecological Understanding:

- Assess the system's impact on biodiversity conservation and ecological research by evaluating its effectiveness in supporting rapid and accurate flower species identification.

11. Ethical Considerations:

- Address ethical considerations related to the use of automated flower recognition technology, including privacy concerns, data security, and potential biases in the algorithm's predictions.

By achieving these research objectives, the study aims to contribute a robust and practical solution to automated flower recognition, advancing the field of ecological monitoring and biodiversity conservation.

**Literature Review**

**2.1 Theories and models**

**2.1.1Basics of theories:**

1. Identify flower in real time which include native Bangladeshi flowers along with some

foreign flowers also.

2. Produce a convolutional neural network which is capable correctly classifying images of

flowers with an average confidence level of 95% or more. Sometimes it provides 100%

accuracy rate for some kind of flowers.

1. Collected a dataset of over 5000 images of flowers using their genus-species classification as the Google Image search term.

The following figure showing the output of the application which identify Rose with highest confidence value by using mobile camera.

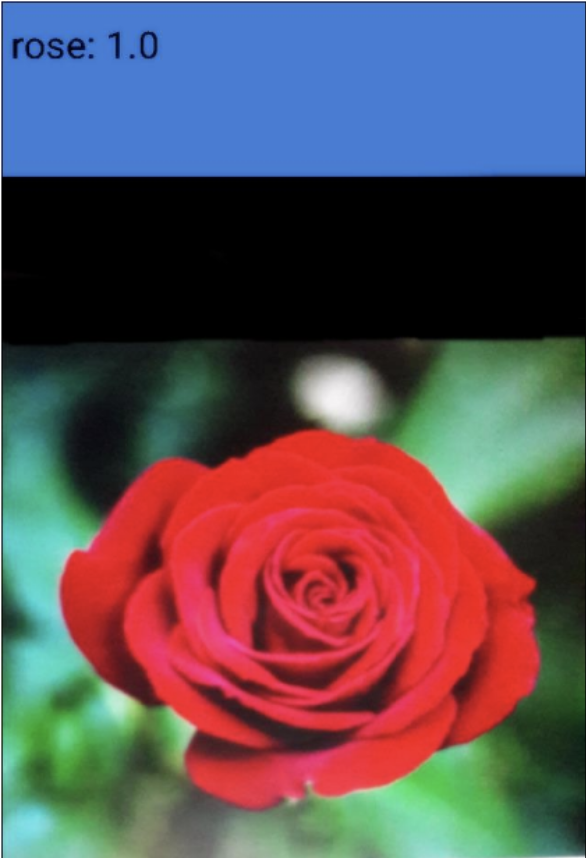


Figure 1.5 : Rose Identification with highest confidence value

**2.1.2 Explanation of variables**

The variables in a flower recognition research paper refer to the factors, elements, or conditions that are studied and manipulated to understand their impact on the performance and outcomes of the flower recognition system. Here's an explanation of key variables that might be considered in such a study:

Independent Variables:

Environmental Conditions:

This variable encompasses factors such as lighting conditions, weather, and background clutter, which may affect the accuracy of flower recognition in natural settings.

Diversity of Flower Species:

The variety of flower species in the dataset serves as an independent variable, influencing the system's ability to generalize across different types of flowers.

Image Quality:

Image resolution, focus, and quality can impact the recognition system's performance. This variable explores how variations in image quality affect the accuracy of flower identification.

Dependent Variables:

Recognition Accuracy:

The primary dependent variable is the accuracy of the flower recognition system, measured as the percentage of correctly identified flower species in comparison to ground truth data.

Robustness to Variability:

This variable assesses the system's ability to maintain accurate recognition under changing environmental conditions, such as variations in lighting and background.

Multi-Species Recognition Accuracy:

For images containing multiple flower species, this variable measures the system's accuracy in correctly identifying each species present.

Controlled Variables (Constants):

Dataset Composition:

Keeping the composition of the dataset consistent across experiments ensures that any observed differences in performance are attributable to changes in independent variables rather thanvariations in the dataset.

Model Architecture:

Maintaining a consistent model architecture and training methodology helps isolate the impact of independent variables, ensuring that changes in performance can be attributed to specific experimental conditions.

Evaluation Metrics:

The metrics used to evaluate the system's performance, such as precision, recall, and F1 score, should remain consistent to facilitate meaningful comparisons between experiments.

Moderating Variables:

Transfer Learning Techniques:

These techniques can moderate the impact of limited data by leveraging knowledge gained from pre-trained models. The study may investigate how different transfer learning approaches influence the system's performance.

Edge Computing Capability:

The system's ability to operate on edge devices may moderate its performance in field settings. Exploring the impact of edge computing on recognition accuracy is a relevant moderating variable.

Extraneous Variables:

Image Artifacts:

Unintended artifacts in images, such as reflections, shadows, or occlusions, may introduce variability. Monitoring and controlling for these extraneous variables enhance the study's internal validity.

Human Bias:

If human annotators are involved in generating ground truth data, their biases can influence the study outcomes. Implementing measures to minimize human bias is essential for the reliability of the results.

By carefully defining and controlling these variables, researchers can conduct a systematic and rigorous investigation into the effectiveness of the flower recognition system, contributing valuable insights to the field of computer vision and ecological monitoring.

**2.2 Related Works**

Convolutional Neural Network (CNN) and Image Classification is the most used technique in machine learning. On the other hand Tensorflow an open source library developed by Google Brain team is the hot topic nowadays in the field of machine learning. There are some good publications of Tensorflow. But It is very little compared to the importance. In the following overview of the related works, we try to cover the works that are mostly related to ours to our best knowledge. Image Classification by Tensorflow is a relatively new topic. But the idea of reliable classification has made it increasingly important. Recent history of modern software technology we can see lots of use of machine learning and image processing in our day to life. Like Police using face detection technology to identify criminal. Facebook using AI technology to match your face ID with any photo of you that has been uploaded. Pinterest use this technology to find similar photos. Traffic police also find out vehicles by using this technology. There is so many big projects on this topic, those are given below,

**» PICTURETHIS**

****

Figure 2.2 : PictureThis application feature graphic image

Among the popular software This is a Plant identification software developed by “Glority Software limited” running on both Android and IOS. This Software let you identify plant by taking a picture with 90% accuracy rate less than a second.

**» PLANTSNAP**

**A green and white logo

Description automatically generated**

Figure 2.3 : PlantSnap application feature graphic image

This is also a popular software like PictureThis, developed by PlantSnap INC. This software let us to identify plants, flowers, cacti, mushrooms etc. by taking a picture of that plant. There are 585,000 species of plants and trees in their database. So their accuracy rate is 90% accurate.

**2.3 Research Summary**

The purpose of this project is to use Tensorflow, an open-source dataflow and machine learning library, to build an image classifying Convolutional Neural Network (CNN) for classifying the flower image. Tensorflow, in addition to providing developers a simple way to build neural network layers, can also be run on mobile platforms such as Android. The ultimate goal of this project is to design and optimize a convolutional neural network for use with flower classification, and eventually build a simple classification app for mobile devices around the trained network. The mobile app will allow users to try and classify flowers while outdoors or offline. After examining, inquiring, researching and comparing all the related works on Flower Identification, we come to have different and many types of conclusion. First of all, none of them are using real time object detection and identification. All of them are just stuck into by taking picture of object and search through their database to identify the object. Secondly is accuracy rate, most of the big project hit the accuracy rate up to 90% so they say. But the community feedback shows the different color. Users complain about the miss detection of plants and less information about it. It also take long time for analyzing and searching their database. Because of our real time flower identification, it is more like instant detection of flower.

**2.4 Scope of the Project**

**Developers**

The project is complete ready for practical use, someone who wishes to learn the basics of CNNs and Tensorflow, substituting their own classes. The project is essential for anyone that wants to improve their project. It have scope to improve both the dataset and the CNN architecture, however, this will require more research and knowledge of CNN design. Additionally, someone who is interested in Tensorflow and Android development or front end development can be benefited from this work.

**Scientist or Botanist**

This research and project will be helpful for botanists or biologists, although any scientist would likely wish to use their unique dataset. The project would most likely be useful to scientists for classifying large amounts of data and generating statistics for the dataset. Using our Android application Botanists or other scientists dealing with flowers to identify easily.

**Traveler or Nature Enthusiasts**

Our developed project will be much helpful while spending time outdoors to identify flowers by a mobile application. This application works completely in offline which use mobile camera to identify flowers. It will be really helpful for Traveler, Hiker and Nature Enthusiasts [1].

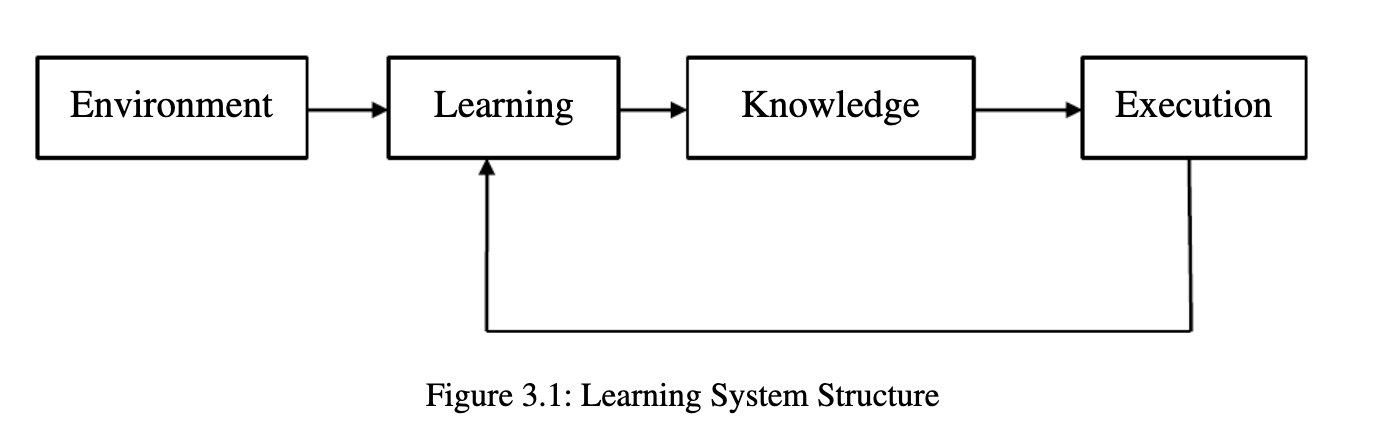
**2.5 Challenges**

Duplicate flower identification is the main challenge of our project. Some flowers are same to look at by size and shape. That time it is so tough to provide the accurate result. We are continue our research to solve this problem. Another challenge is to identify same flower of different specific color. Ex- White Rose, Pink Rose. This is a tough work to identify color using convolutional neural network. But we take it as a challenge. Other challenge is achieving accurate data. In the market of false and duplicate data, it is very hard to get all of the things right. But we are trying are heart and soul to do better.

**Methodology**

**3.1 Introduction to Machine Learning**

In Artificial Intelligence Learning is a very important feature. Many scientists tried to give a proper definition for learning. Many computer scientists, sociologists, logicians and others discussed about this for a long time. Some scientists think that learning is an adaptive skill that can perform the same process better later on (Simon 1987). Others claim that learning is a process of collecting knowledge (Feigenbaum 1977). Although there is no definite definition for learning skills, we still have to define machine learning. In general, machine learning has to be identified on how to improve the computer algorithm automatically through experience (Mitchell 1997). Machine learning is one of the important field of Artificial Intelligence. At the beginning of development of Artificial Intelligence (AI), the system does not have a thorough learning ability so the whole system is not perfect. For instance when the computer faces problems, it can not be self-adjusting. Moreover, the computer cannot automatically collect and discover new knowledge. Therefore, computer only can conducted by already existing truths. It does not have the ability to discover a new logical theory, rules and so on.



**3.2 Convolutional Neural Network**

Convolutional neural networks are a class of machine learning networks which are commonly applied to image visualization problems such as classification. CNNs were inspired by the connections of the neurons and synapses in the brain. The design of these networks is made up of series of convolutional, pooling, and fully connected layers. The convolutional layer does what its name describes, it applies a number of convolutional filters to the input images in order to acquire the learning parameters for the network. Pooling layers are placed in between convolutional layers, and are used to reduce the number of parameters used for learning, and thusreduce the computation required. Finally, fully connected layers are full connections to the previous layer, rather than the small window the convolutional layers are connected to in the input. Convolutional neural networks are commonly used for image classification, however, there are limitations to this application. A human canidentify the contents of certain images much more quickly than a computer, but CNNs have proven to have a 97.6% success rate when applied to facial recognition.

**A diagram of a layer of a layer

Description automatically generated**

**3.3 Design of Dataset**

The diagram of dataset design which shows different step of data collection and dataset handler.

A diagram of a software system

Description automatically generated

Figure 3.3: Design of Dataset

**3.4 Implementation Requirements**

**Software**

* Windows / Linux / Mac OS
* Python pip package / Anaconda
* Tensorflow
* Android Studio

**Hardware**

* Intel i3 Processor
* 4 GB RAM
* 1TB Hard Disk

**3.5 Design Development**

The initial step of this project was to research the available machine learning libraries, convolutional neural network design, and collect datasets. Though there exist some other libraries, but Tensorflow was chosen because there are many tutorials and documentation for the library. After began to have a basic understanding of the Tensorflow library, we retrained the MobileNet model with our own dataset, which proved to be very successful in testing, however, the goal of this project was to learn how to develop and optimize a neural network. The initial designs of the CNN for this project were based on several different tutorials about how to use Tensorflow to design an image classifier. The next step was to begin modifying the initial network to try and find a design that worked for the application of this project. The next step was implementation of android app using Android Studio. Some open source code of Tensorflow collected from github which needed in our implementation work. Final Step was set up the pb file and label file with the application. The design of CNN which we tested is being submit as a deliverable for this project. This model also was trained using the dataset at the genus-species level, which uses approximately 700 images average per class, which is almost enough. But we are working to add more images per class. This model is currently being trained and tested. The parameters of the network such as number of training steps, output directory, and image input directory can all be specified, however, their defaults will place all the output directory in the current working directory. The input image directory must be specified, and the contents of the directory must be folders of images in tf files folder. The Tensorflow trainer will generates two text files: one containing the labels for the classifier, and the other lists which images were selected for training, testing, and validation which called pb file. The classifier uses these to read result for each image classification and show the output result.

**Time Plan**

This project has plenty of room for future work, by myself or a future interested student.

1. Improved CNN design. There much more research and practice is needed to optimize the design.

2. Improve the dataset and add more data.

3. Specific Identification of duplicate flower which is same to look at.

4. We have plan to release this app on Google Playstore.

**Reference**

[1] Kody G. Dangtongdee and Dr. Franz Kurfess, “Plant Identification Using Tensorflow, ”. Computer Engineering Department, California Polytechnic State University, San Luis Obispo, Senior Project Final Report. CPE 462, Winter 2018, pp. 5-15.

[2] Xuanxin Liu, Fu Xu, Yu Sun, Haiyan Zhang, and Zhibo Chen, “Convolutional Recurrent Neural Networks for Observation-Centered Plant Identification,” Journal of Electrical and Computer Engineering, vol. 2018, pp. 1–7, 2018.

[3] Yu Yang and Patric Granholm, “A study of pattern recognition of Iris flower based on Machine Learning, ”. Turku University of Applied Sciences, Turku Finland , Bachelor’s Thesis, 2013.

[4] Wang-Su Jeon and Sang-Yong Rhee, “Plant Leaf Recognition Using a Convolution Neural Network, ”. International Journal of Fuzzy Logic and Intelligent Systems Vol. 17, No. 1, March 2017, pp. 26-34.

[5] Er. Gurleen Kaur Marwa and Er. Salam Din, “Development of Software for Flower Recognition Using Image Processing Technique, ”. Dept. of SEEIT, COAET, PAU, Ludhiana, India, vol. 3, Issue 5, May 2015.

[6] Vijaya Bylaiah, “Leaf Recognition and Matching with MATLAB, ”. A Thesis Presented to the Faculty of San Diego State University, Spring 2014.

[7] Y. Li, Q. Zhu, Y. Cao, and C. Wang, “A leaf vein extraction method based on snakes technique,” in Proceedings of the International Conference on Neural Networks and Brain (ICNN&B '05), pp. 885–888, 2005.

[8] Bache, K.& Lichman, M. 2013. UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

[9] Bishop, C. 2006. Pattern Recognition and Machine Learning. New York:Springer, pp.424-428.

[10] Robin. 2010. Machine-Learning - Artificial Intelligence. Available at: http://intelligence.worldofcomputing.net/category/machine-learning. Consulted 22 Nov 2018.