A Project Report on

Flowers Detection and Classification.

Submitted by:

Arbaz Khan (roll no - 54)
Amit Potdar (roll no - 58)
Ganesh Jamkar (roll no -59)

Under the guidance of

Prof. Daivashala Deshmukh

In partial fulfillment of the award of Bachelor of Technology in

Computer Science and Engineering

Department of Computer Science and Engineering

Maharashtra Institute of Technology, Aurangabad (M.S)

[2022-2023]

DECLARATION

I declare that this written submission represents my ideas in my own words and where others' ideas words have been included; I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Place: Aurangabad

Date:

Signature and Name of students

Arbaz Khan (54)

Amit Potdar (58)

Ganesh Jamkar (59)

CERTIFICATE

This is to certify that the major project report entitled "**Flowers Detection and Classification.**" submitted by "**Arbaz khan**, **Amit Potdar and Ganesh Jamkar"** is the bonafied work completed under my supervision and guidance in partial fulfillment for the award of Bachelor of Technology in Computer Science and Engineering, Maharashtra Institute of Technology under Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (M.S.).

Place: Aurangabad
Date:

Ms. D.R. Deshmukh
Guide

Dr. S. L. Kasar
Head of Department

Dr.S.P.Bhosle
Director
Maharashtra Institute of
Technology, Aurangabad

APPROVAL CERTIFICATE

This Major Project report entitled "**Flowers Detection and Classification.**" by **Arbaz Khan, Amit Potdar and Ganesh Jamkar** is approved for Bachelor of Technology in Computer Science and Engineering, Maharashtra Institute of Technology under Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (M.S.).

Place: Aurangabad
Date:

Examiner_____
(Signature)

_____
(Name)

ACKNOWLEDGEMENT

There were lots of individual who were certainly very helpful, kind, cooperative and generous along the development of this project. I extend my special thanks to them.

I would like to express my special thanks of gratitude to our project guide Ms.D.R. Deshmukh for their valuable support in completion of my project.

I would also want to extend my gratitude to Dr.S.L.Kasar and our principal Dr.S.P.Bhosle sir for providing such opportunity to enhance our skills.

INDEX/CONTENTS

List of Tables

List of Figures

ABSTRACT

Classification is a supervised machine learning technique which is used to predict group membership for data instances. Neural networks are being introduced to simplify the problem of classification. This model focuses on Iris flower classification using Neural Network. For simplification of classification, we will use scikit learn tool kit. This project mainly focuses on the classification of dataset using scikit learn. The problem concerns that the recognition of Iris flower species (Sentosa, versicolor and verginica) on the basis of the measurements of length and width of sepal and petal of the flower. We can generate classification model by using various machine learning algorithms through training the iris flower dataset and can choose the model with highest accuracy to predict the species of iris flower more precisely. Classification of Iris data set would be detecting patterns from examining sepal and petal size of the Iris flower and how the prediction was made from analyzing the pattern to form the class of Iris flower. By using this pattern and classification, in future upcoming years the unseen data can be predicted more precisely. Artificial neural networks have been successfully applied to problems in pattern classification, function approximations, optimization, and associative memories. The goal here is to model the probabilities of class membership, conditioned on the flower features. In this project we will train our model with data using machine learning to predict the species of iris flower by input of the unseen data using what it has learnt from the trained data.

Chapter 1

INTRODUCTION

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 recognize 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. Every day 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.

It is very important to identify naturally occurring objects and classify its type. It is useful to identify flower type in various fields such as gardening, botany research, Ayurveda, treatment, farming, Floriculture etc. Nature has many kinds of flowers, similarity in some features is found between the flowers. For example, many flowers share the red colour. On the other hand, these red flowers are different from other features. Red flowers do not necessarily share the same shape. These similarities and differences highlight the difficulty of identifying each flower species automatically. Traditional flower classification task is done by a botanist. Many challenges are facing botanist through flower classification task.

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 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

Our project aim is to provide an automated system that detects and classification flower species. The importance of building automated flower classification method stands out in many benefits such as providing fast classification for educational purpose, as automated method accelerates the learning process. Automated flower classification gives the people with limited experience in flower species, the ability to classify the species of a flower, with the advantages. Computerized picture preparing manages control of advanced pictures through a computerized PC. For example, Classification, Feature extraction, Pattern acknowledgment and so forth are helpful to order the pictures.

The main base of the software is a data set containing various images of flowers, which is further split into train sets and test sets. Keeps all the information related to the image of the flower.

1.2 Necessity:

Flowers Detection and Classification observes and analysis flowers, the importance of building automated flower classification method stands out in many benefits such as providing fast classification for educational purpose, as automated method accelerates the learning process. Automated flower classification gives the people with limited experience in flower species, the ability to classification the species of a flower, with the advantages.

1.3 Problem Definition

Given a photograph of a flower, the aim is to identify the type of the flower using machine learning and to display its relevant information. The aim of our project is to provide an automated system that it identifies and recognise flower species. the importance of building automated flower classification method stands out in many benefits such as providing fast classification for educational purpose, as it improves learning process.

Automated flower classification gives the people with limited experience in flower species, the ability to classify the species of a flower, with the advantages.

The main task for our developer finding the proper feature extraction factors relating to plants and flowers. Since there is many flowers of different shape, colour and texture of flowers.

1.4 Objectives:

1) To collect the standard data base.
2) To decide the features of image.
3) To convert every image to features set.
4) To train a neural network.
5) To test the neural network.

1.5 Scope and Limitations

1) It satisfies the user requirement.
2) Be easy to operate.
3) Have a good user interface.
4) Be expandable.
5) To provide more information of flower and their family that might help Botany student for study and research purpose.
6) This helps in predicting the flower present in data set.

Limitation

1) Data security
2) It requires mobile/pc/laptop
3) Chances of technical problems in website
4) Sometimes it cannot help to solve the problem

1.6 Applications

1) Easy to use this application

2) It gives accurate information of Flowers.

3) It can be used on android, Linux, windows, apple.

4) It is very simple.

Chapter 2

LITERATURE SURVEY

Table 2.1 Literature survey Table

Sr. No	Year	Paper Title	Authors	Method used	Result
1	2017	"Flower Detection Species and Recognition from Digital Images".	Aalaa Albadarneh and Ashraf Ahmad	Region growing method is used for segmentation and different features are used combine.	They obtained the recognition accuracy result of 92% by using SGD classifier
2	2018	"A System Flower asked On Image Processing And Neural network".	Huthaifa Almogdady, Dr. Saher Manaseer and Dr.Hazem Hiary	GLCM method is used for feature extraction.RGB and HSV used for colour extractionANN method is used for classification.	They have achieved 81.19% as an accuracy rate for classification.
3	2015	"Literature Review on Flower Classification".	Chaku Gamit, Prof. Prashant B. Swadas and Prof. Nilesh B. Prajapati	Presented a literature survey of various techniques used for flower classification	Gave a conclusion that MLP gives better result as compare to logistic regression, kNN, pNN and SVM.

4	2014	"Flower Recognition System Based on Image Processing".	Tanakorn Tiay, Pipimphorn Benyaphaichit, and Panomkhawn Riyamongkol	They used Hu's seven-moment algorithm for texture and color characteristics such as Red, green, blue, hue, and saturation are derived from histogram after that K-nearest neighbor is used to classify flowers.	The system shows the most nearest 3 flower images with matching flower information of the first nearest flower image. The accuracy of this system is more than 80%.
5	2017	"Flower Species Recognition System using Convolution Neural Networks and Transfer Learning".	I.Gogul, V.Sathiesh Kumar	A Deep learning approach using Convolutional Neural Networks (CNN) is used to recognize flower species and Feature extraction of flower images is performed using a Transfer Learning approach.	Yields impressive Rank-1 accuracies of 73.05%, 93.41% and 90.60% using Over Feat, Inception-v3 and exception architectures, respectively as Feature Extractors on FLOWERS102 dataset.
6	2015	"Flower Detection and Counting Using	Balvant V. Biradar, Santosh P. Shrikhande	Here Gaussian low-pass filter and morphological	The proposed method for automatic detection

		Morphological and Segmentation Technique".		operations for pre-processing the flower images to remove the non-flower region and enhancement of fine details	and counting of marigold flowers is an efficient and robust with accuracy of 92%.
7	2015	"Retrieval of Flower Based on Sketches".	Y H Sharath Kumara, D S Gurub	They have used Kd-tree based indexing approach to index a huge dataset of flowers for given input sketch.	From this model they conclude that the combination of all the features descriptors achieves a good accuracy with indexing approach.
8	2010	"Textural features in flower classification ".	D.S.Guru,Y.H.Sharath Kumar,S.Manjunath	They have proposed a probabilistic neural network-based flower classification method with the use of texture feature such as CTMs, GLCM, and Gabor responses.	They observed that using the proposed textural features one can achieve relatively a good classification accuracy when compared to any other available features.35% for the best single feature to 79% for the combination of all features.

9	2016	"Counting Flowers in Digital Image:A Review".	Kriti Sharma, Anoop Singhal	This paper discuses and well us reviews the algorithms and the technologies which are available for detecting and counting the number of flowers from the flower images.	Result a quick overiew on different flower detecting and counting methodology.
10	2014	"Image Flower Recognition based on a New Method for Color Feature Extraction".	Amira Ben Mabrouk, Asma Najjar and Ezzeddine Zagrouba	SURF method is used for point feature extraction, for the color feature the authors used RGB and HSV method. They used SVM classification to recognize the flower image.	Results have proved that combining features perform better than using a single feature for classification.
11	2012	"Flower image segmentation based on color analysis and a supervised evaluation".	Asma Najjar and Ezzeddine Zagrouba	Here they proposed a flower segmentation schema based on OTSU thresholding technique and Lab colour space.	Segmentation can be done with a small time consuming.

12	2006	"A Visual Vocabulary for Flower Classification".	Maria-Elena Nilsback, Andrew Zisserman	Developed and optimize a nearest neighbor classifier architecture that distinguish categories which have significant visual similarity or to overcome the ambiguities that exist between flower categories.	They conclude that in their larger dataset (OXFORD 17) color feature perform better than shape and texture and texture perform very poorly.
13	2007	"Delving into the Whorl of Flower Segmentation".	Maria-Elena Nilsback, Andrew Zisserman	Described an algorithm for automatically segmenting flowers in colour photographs using a colour model for foreground and background, and a generic shape model for the petal structure.	Reasults a good background substraction model for 13 classes of OXFORD 17 dataset.
14	2008	"Automated Flower Classification over	Maria-Elena Nilsback, Andrew Zisserman	They compute feature like local shape/texture, the	An improvement on performance found, from 55.1% for the

		a Large Number of Classes".		shape of the boundary, the overall spatial distribution of petals, and the colour with SVM classifier on OXFORD 102.	best single feature to 72.8% for the combination of all features.
15	2009	"An automatic visual Flora segmentation and classification of flower images".	Maria-Elena Nilsback	For segmentation grabcut is used.In classification, they combined different features (color,texture and shape).and passed these feature set to SVM classifier for classification.	Found that combination of multiple feature with using a multiple kernel framework with a SVM classifier works good on Oxford 17 category and a 102 category flower database.

Chapter 3

SYSTEM DEVELOPMENT

3.1 Proposed system:

Systems development is the process of defining, designing, testing, and implementing a new software application or program. It could include the internal development of customized systems, the creation of database systems, or the acquisition of third party developed software. Written standards and procedures must guide all information systems processing functions. The organization's management must define and implement standards and adopt an appropriate system development life cycle methodology governing the process of developing, acquiring, implementing, and maintaining computerized information systems and related technology. Systems Development Life Cycle (SDLC) Standards and Procedures:

The aim of our project is to provide an automated system that detects and classify the flower species. The importance of building automated flower classification method stands out in many benefits such as providing fast and an accurate classification of flower for educational purpose, as automated method accelerates the learning process. Automated flower classification gives the people who know limited flowers species, the ability to classify the species of a flower, with the advantages. It is centres around building up a PC framework that can perform preparing on a picture. Computerized picture preparing advancements, for example, Classification, Feature extraction, Pattern acknowledgment and so forth are helpful to order the pictures.

Knowledge of the various kinds of flowers is essential to the protection and control of biodiversity. Also, flowers are the most important part of the food chain. Therefore, adequate recognition of flower species is essential to the protection of biodiversity. There are thousands of flowering plants in various parts of the world. Personal identification of all these flower species is a time consuming and challenging task for even the botanist.

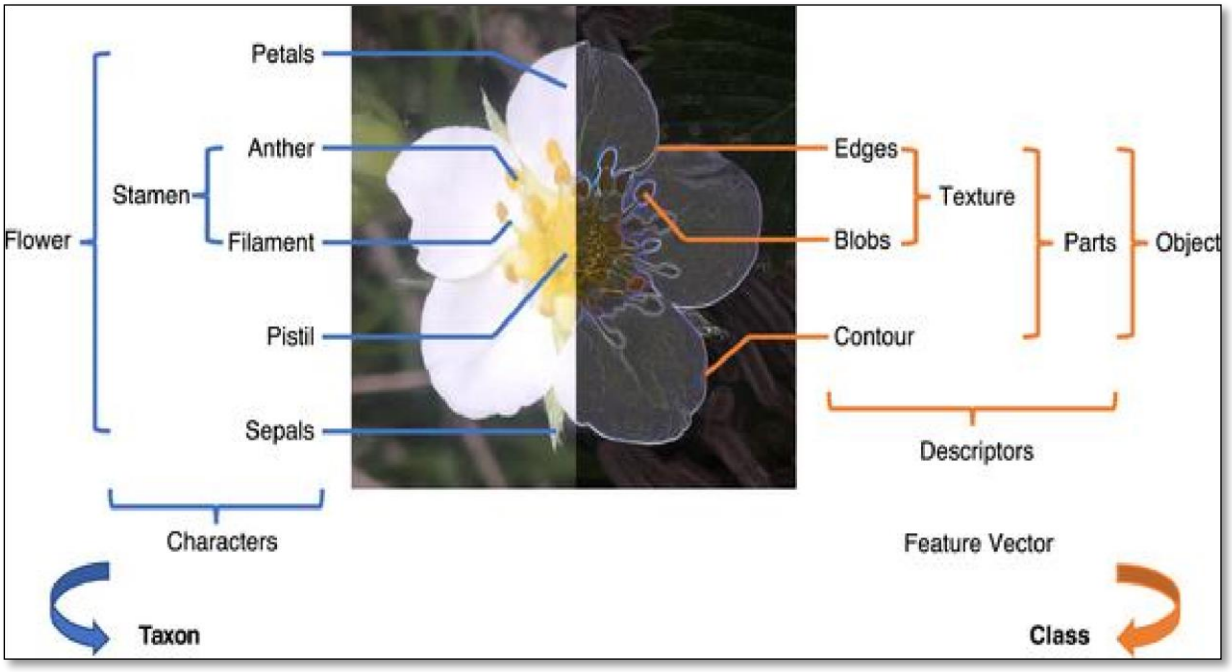


Fig 3.1 Features of Flowers

3.2 Block Diagram:

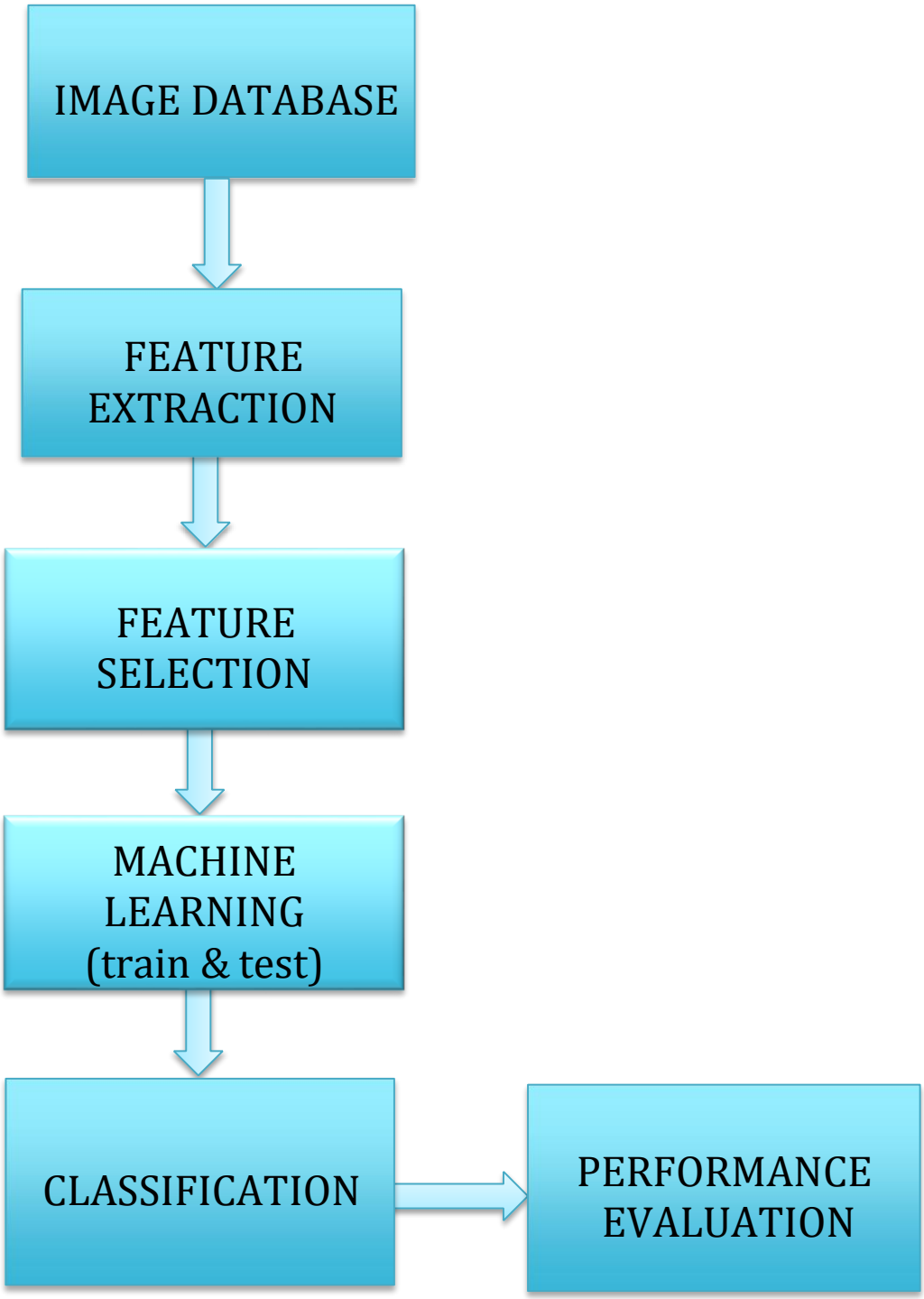


Fig. 3.2 Block Diagram

● **Feature Selection**

Feature selection aims at selecting the most relevant features that allow discriminating among classes. For acute classification, extracted features have been further processed by selecting the most informative ones. There are three main paradigms for feature selection:

1. **Filter-based feature selection** chooses the most relevant features regardless of the learning algorithm. These methods rely on measuring properties of the features, such as correlation, entropy, or intra/inter-class distance to determine how informative they are.

2. **Wrapper-based feature selection** uses the learning algorithm to evaluate the goodness of a subset of features. This approach has the advantage of allowing the detection of likely interaction among characteristics, albeit with a higher computation time than filter-based methods.

3. **Embedded-based feature selection** aims to combine the advantage of both filter and wrapper methods. The feature selection is embedded in the learning process of the algorithm. For instance, decision trees, such as CART, have a built-in mechanism to perform feature selection.

● **Feature extraction**

Feature extraction is a process of dimensionality reduction by which an initial set of raw data is reduced to more manageable groups for processing. A characteristic of these large data sets is many variables that require a lot of computing resources to process. Feature extraction is the name for methods that select and /or combine variables into features, effectively reducing the amount of data that must be processed, while still accurately and completely describing the original data set.

Why is this Useful?

The process of feature extraction is useful when you need to reduce the number of resources needed for processing without losing important or relevant information. Feature extraction can also reduce the amount of redundant data for a given analysis. Also, the reduction of the data and the machine's efforts in building variable combinations (features) facilitate the speed of learning and generalization steps in the machine learning process.

Practical Uses of Feature Extraction

1. **Auto encoders** – The purpose of auto encoders is unsupervised learning of efficient data coding. Feature extraction is used here to identify key features in the data for coding by learning from the coding of the original data set to derive new ones.
2. **Bag-of-Words** – A technique for natural language processing that extracts the words (features) used in a sentence, document, website, etc. and classifies them by frequency of use. This technique can also be applied to image processing.
3. **Image Processing** – Algorithms are used to detect features such as shaped, edges, or motion in a digital image or video.

Feature Extraction techniques can also lead to other types of advantages such as:

- Accuracy improvements.
- Over fitting risk reduction.
- Speed up in training.
- Improved Data Visualization.
- Increase in explain ability of our model.

Flower Recognition using Machine Learning

In this project, I initialize my model architecture and modify its higher level convolutional layer to learn the flower species. I do not handcraft any feature descriptor for the fusion features but introduce convolution layers to learn the filters themselves. To incorporate species of flowers, we first train the layers CNN based on flower classes. After we trained the species layers, I test my model for the classification of the data.

I have used the dataset of 5 different flowers. Even before using Deep learning, Flower Recognition using ML has been made possible, however their accuracies were really low or they had a relatively dataset. Flower Recognition using ML is a classic pattern recognition problem for which researchers have worked since the early days of computer vision.

Implementation of Flower Recognition using ML:

Flower Recognition using ML works in stages as pre-processing, segmentation, feature extraction and recognition using neural networks. Pre-processing includes series of operations to be carried out on document image to.

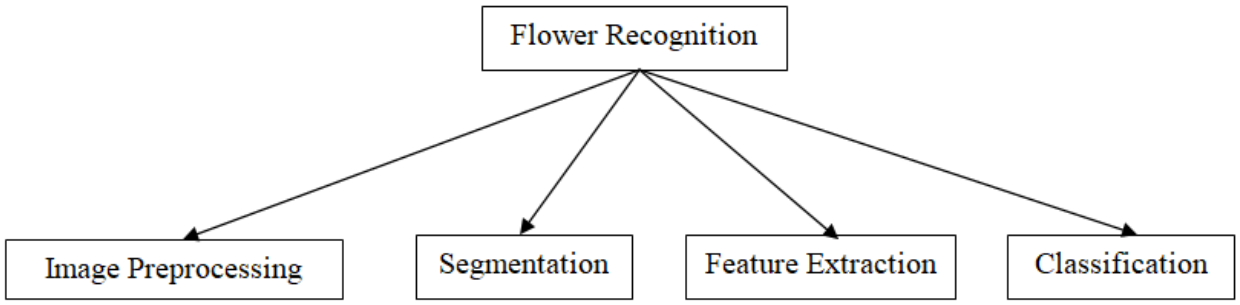


Fig 3.3 Flowers Recognition

BASIC 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 (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 must define machine learning. In general, machine learning must be identified on how to improve the computer algorithm automatically through experience (Mitchell 1997). Machine learning is one of the important fields 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 cannot be self-adjusting. Moreover, the computer cannot automatically collect and discover new knowledge. Therefore, computer only can be conducted by already existing truths. It does not have the ability to discover a new logical theory, rules and so on.

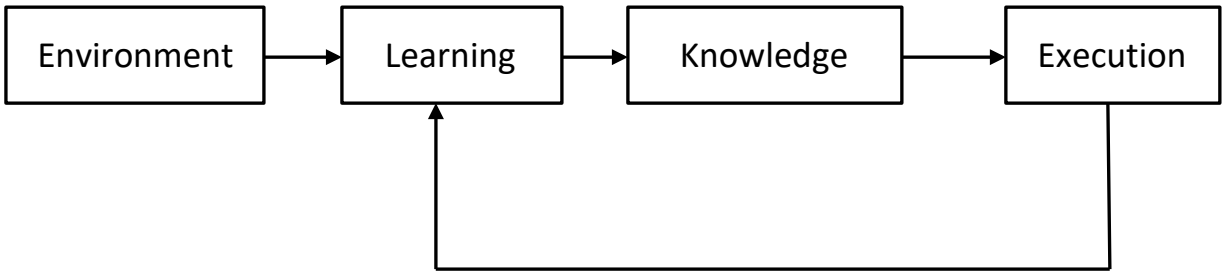


Figure 3.4: Learning System Structure

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 several 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 thus reduce 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 can identify 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.

The following figure showing the typical design of convolutional neural network.

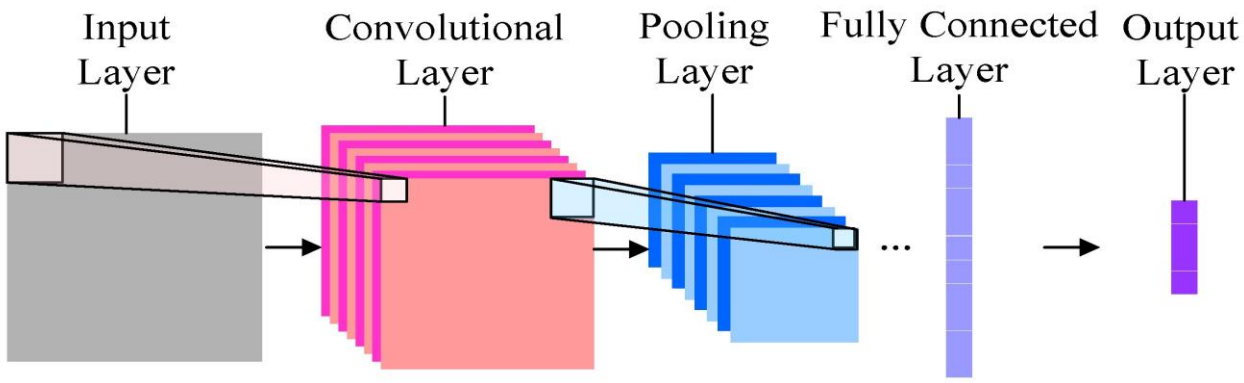


Fig 3.5: Typical CNN design

Convolution — The main building block of CNN is the convolutional layer which is a mathematical operation to merge two sets of information. The convolution is applied on the input data using a convolution filter to produce a feature map. The first layer to extract features from an input image is

convolutions are regularized versions of multilayer perceptron's. Multilayer perceptron's usually mean fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. The "full connectivity" of these networks makes them prone to over fitting data. Typical ways of regularization, or preventing over fitting, include penalizing parameters during training (such as weight decay) or trimming connectivity (skipped connections, dropout, etc.) CNNs take a different approach towards regularization: they take advantage of the hierarchical pattern in data and assemble patterns of increasing complexity using smaller and simpler patterns embossed in their filters. Therefore, on a scale of connectivity and complexity, CNNs are on the lower extreme.

Striding — The number of shifts over the input image is called stride. Example- When stride is 1 then we move the filters to 1 pixel at a time.

Pooling Layer — After a convolution operation we usually perform pooling to reduce the dimensionality. This enables us to reduce the number of parameters, which in turn shortens the training time and reduces chances of over fitting.

Pooling layers down sample each feature map independently, reducing the height and width, keeping the depth intact. The most common type of pooling is max pooling which just takes the maximum value in the pooling space. Contrary to the convolution operation, pooling has no parameters. It slides a window over its input, and simply takes the maximum value in the window. The largest element from the feature map is selected by using this operation.

Why use Neural Network:

Neural network with their remarkable ability to derive meaning from complicated or imprecise data can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyse. This expert can then be used to provide projections given new situations of interest and answer "what if" questions. Other Advantages Include:

Neural Networks find extensive applications in areas where traditional computers don't fare too well. Like, for problem statements where instead of programmed outputs, you'd like the system to learn, adapt, and

change the results in sync with the data you're throwing at it. Neural networks also find rigorous applications whenever we talk about dealing with noisy or incomplete data. And honestly, most of the data present out there is indeed noisy.

With their brain-like ability to learn and adapt, Neural Networks form the entire basis and have applications in Artificial Intelligence, and consequently, Machine Learning algorithms. Before we get to how Neural Networks power Artificial Intelligence, let's first talk a bit about what exactly is Artificial Intelligence.

Adaptive Learning: An ability to learn how to do tasks based on the data given for training or initial experience.

ALGORITHMS FOR TRAINING THE MODEL

LOGISTIC REGRESSION

Logistic Regression is a type of regression that predicts the probability of occurrence of an event by fitting the appropriate or cleaned data to a logistic function. Like several forms of regression analysis, it makes use of many predictor variables that will be either numerical or categorical. For instance, the probability that a email received is spam or not might be predicted from knowledge of the type of data or the history of sender. This regression is quite used in several scenarios such as prediction of customer's propensity of purchasing a product or used in market analysis to improve the business and use to predict the unpredictable scenarios.

What is logistic regression?

Logistic Regression, also known as Logit Model or Logit Regression, is a mathematical model used in statistics to estimate (guess) the probability of an event occurring having been given some previous data. Logistic Regression works with binary data, where either the event happens (True or 1), or the event does not happen (False or 0). So, by giving some feature x it tries to find out whether some event y happens or not. So, y can either be 0 or 1. In the case if the event happens, y is given the value 1. If the event does not happen, then y is given the value of 0. For example, if y represents whether a coin gives head, then y will be 1 if after tossing the coin we get head or y will be 0 if we get tail. This is known as Binomial Logistic Regression.

Logistic Regression can also be used when there is use of multiple values for the variable y. This form of Logistic Regression is known as Multinomial Logistic Regression.

Logistic regression is named for the function used at the core of the method, the logistic function. The logistic function, also called the Sigmoid function was developed by statisticians it's an S-shaped curve that can take any real-valued number and map it into a value between 0 and 1.

K-Nearest Neighbour Algorithm

K-Nearest Neighbour (KNN) is one of the simplest supervised machine learning algorithms mostly used for classification which uses an entire dataset in extreme phase. When prediction is required for unknown data, it searches through the entire training dataset for k more similar instances and the data with the most similar instance is finally returned as the prediction.

What is KNN Algorithm?

K-Nearest Neighbour is an algorithm which stores every accessible case and characterizes the new cases dependent on the component similarity measure or similarity features. We can say that if we are like our

neighbours then we are one of them. For example, if apple looks most like banana, orange or melon rather than a monkey, rat or a cat then most likely apple belongs to the group of fruits. But in-general KNN is using such application where you're looking for similar items i.e., when the task is some form of fine items like this one then you call the search as a KNN search. Let's see the example of scenarios that is used in the industry: So let's see the industrial application of KNN algorithm starting with recommender system, the biggest use case of KNN search is a recommender system, this recommender system is like an automated form of a shop counter person, today when you're asking for a product not only shows you the product but also suggest you or displays you relevance set of products which are related to the item you're already interested in buying this KNN algorithm is applied to recommending products like an Amazon or a recommending media like in case of Netflix or Today's almost all of us must have used Amazon for shopping so just for a knowledge, more than thirty five percent of Amazon dot com's revenue is generated by its recommendation engine and so what's the strategy the Amazon uses recommendation as a targeted marketing tool in both the email campaigns and on most of its website pages. Amazon will recommend many products from different categories based on what we are browsing and it will pull those products in front of us which are likely to buy like the frequently bought together option that comes at the bottom of the product page to tempt us into buying the combo well this recommendation has just one main goal that is increase average order value by providing product suggestions based on items in the shopping cart well based on the product they're currently looking out on site.

What is K in KNN Algorithm?

In K-Nearest Neighbour, K denotes the number of nearest neighbours which are voting for the class of the new data or the testing data. For example, if K=1 then the testing data are given the same label as a close to this example in the training set, similarly if K=3 or more the labels of the three closest classes are checked and most common label are assigned to the testing data. As we can see in figure the value of K increases decision region gets smoother. If K=1 then there will be the case of overfitting as it takes only one neighbour. Most of the value of K lies between 3-10 to generate accurate result.

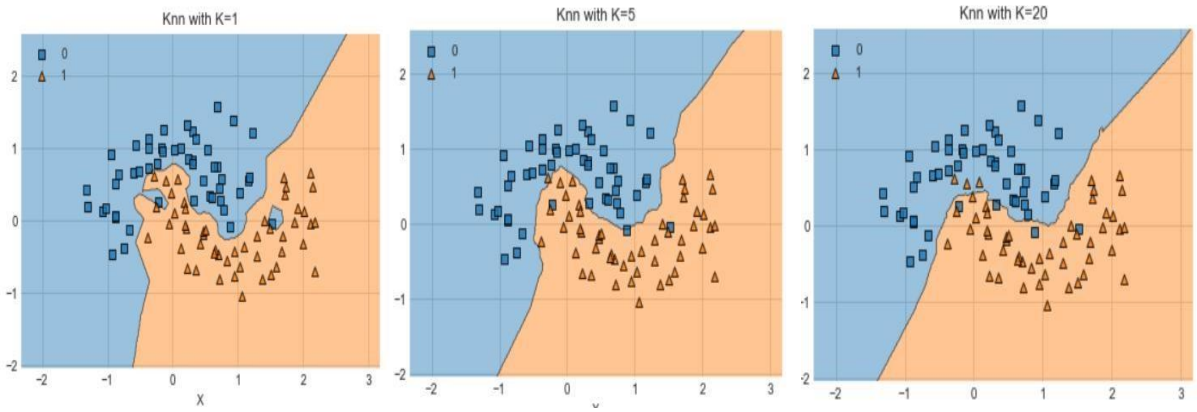


Fig.3.6 Decision Region graph

KNN algorithm is based on feature similarity choosing the right value of K is a process called parameter tuning and is important for better accuracy. There is no method for finding the value of K but for finding k we could keep some points in our mind like K shouldn't be too big, if it's too big then it's going to take forever to process so this going to run into processing issues and resources issues. To prevent from these issues there are some methods to choose a value of K:

1) Sqrt(n), where n is the total number of data points,

2) Odd value of K is selected to avoid confusion between two classes of data,

3) There is another way to choose the value of K i.e., cross-validation, in this way take the small portion of training dataset and call it a validation dataset and then use the different values of K to choose the best performance on the validation set and then choose that value of K to minimizing the validation error.

How does KNN algorithm Work?

In the given below fig. there is blue and orange point on a graph, so these blue points belong to class A and orange ones belong to class B. Now we can see a new point in the form of star and our task is to predict whether this new point belongs to class A, or it belongs to class B. So to start the prediction the very first thing that we have to do is to select the value of K, as we know K in KNN algorithm refers to the number of nearest neighbours that you want to set, for example in this case take K equal to three so

what does that mean it means that select three points which are the least distance to the new point or we can say that select three different points which is closest to the star. KNN uses least distance measures, so when we calculate the distance, we'll get one blue and two orange points which are closest to the star now since in this case as we have a majority of orange points so we can say that for K equal to three the star belongs to class B. all we can see that the star is more similar to the orange points.

Mathematics of KNN algorithm

There are different types of data and similarity measures are dependent on that. The Euclidean distance can be used for real-valued data. Hamming distance can be used for other types of data such as binary or categorical data. The average of the predicted attribute may be returned in the case of regression problems. The most prevalent class may be returned in the case of classification.

Euclidean Distance

Euclidean distance is defined as the square root of the sum of difference between the new point x and existing point y. KNN algorithm use the Euclidean Distance to find the K number of nearest neighbours.

$$\text{Euclidean distance} = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}$$

Manhattan Distance

Manhattan distance has used to calculate the distance between real vector using sum of their absolute difference.

$$\text{Manhattan distance} = \sqrt{\sum_{i=1}^{k} |x_i - y_i|}$$

KNN is called as a lazy learner. KNN as a classification is a very simple algorithm but that's not the reason why it is called lazy learner. KNN is a lazy learner because it doesn't have a discriminative function from the training data but what it does, it memorizes the training data. There is no learning phase of the model, and all of the work happens at the time of prediction is requested that's why KNN is referred as a lazy learner algorithm.

RANDOM FOREST ALGORITHM

Random forest or random decision forest is an ensemble classifier that operates by constructing multiple decision trees models during the training phase. Ensemble models combines the decision from multiple models to choose the result as the final decision.

What is Random Forest?

Random forest is a supervised machine learning algorithm used for classification. In this the trees are trained on subsets which are being selected as random therefore this is called random forest. So, random forest is the collection or an ensemble of decision trees. The whole dataset is used in the decision tress with considering all features but in random forest only the subset of dataset is selected at random, and the number of features are selected at random, that is how the random forest is built upon. Number of decision trees will be grown, and each decision tree will result into certain outcome and random forest just collect results of all the decision trees and the decision of the majority of the trees will be the final result.

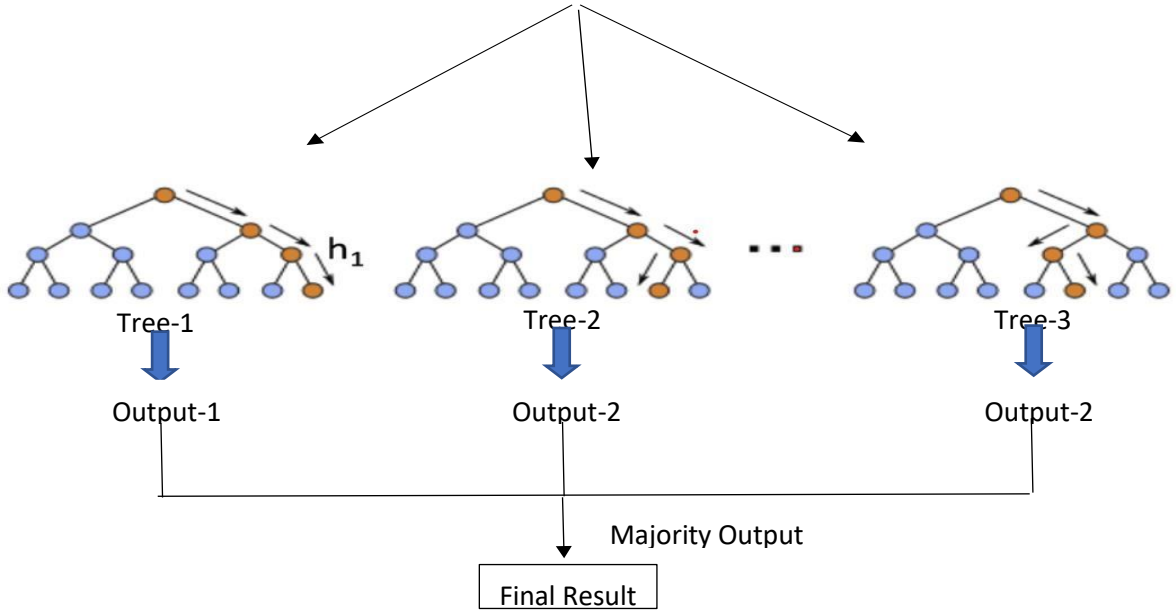


Fig. 3.7 Working Diagram of Random Forest

Random Forest is a versatile algorithm capable of performing both regression and classification. Random Forest has many benefits like

- It uses the multiple trees which reduce the risk of overfitting
- Random Forest has high accuracy that runs efficiently on large database.
- Random Forest algorithm can also estimate missing data. Random Forest can maintain accuracy when a large proportion of data is missing.

SUPPORT VECTOR MACHINE ALGORITHM

A support vector machine abbreviated as SVM was first introduced in the 1960's and late an improvised in the 1990's. SVM is supervised learning machine learning classification algorithm that has become extremely popular nowadays owing to its extremely efficient results, so SVM is implemented in a slightly

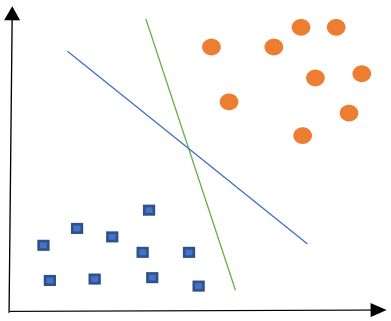
differently than other machine learning algorithms it can perform classification and regression and outlier detection as well.

What is Support Vector Machine?

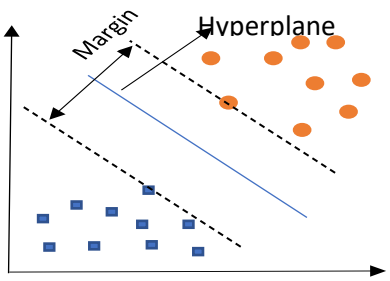
Support vector machine is a discriminative classifier that is formally designed by a separative hyperplane. It is a representation of examples as points in space that are mapped so that the points of different categories are separated by a gap as wide as possible. SVM does not directly provide probability estimates these are calculated using five-fold cross validation. Five-fold cross validation means the dataset will divided randomly into 5 subsets and then take a subset for use as a test set and use remaining subgroups as a training set, then fit a model on the training set and evaluate the score, this will happen until each subset one-by-one considered as a test-set.

The main objective of support vector machine is to separate the given data of different classes by a line (hyperplane) in the best possible way. The nearest points of classes from the hyperplane is known by support vectors. There can be many hyperplanes that will separate the classes, so to choose the appropriate hyperplane SVM algorithm find the nearest points to the hyperplane of both the classes and check the distance between the hyperplane and support vectors. Here, this distance is known by margin. SVM algorithm selects the hyperplane which gives the maximum margin.

In the below figure in Graph- A we can see that there are two lines blue and green. It is clearly seen that blue is placed in the space in which the support vectors give maximum margin from the blue line therefor blue line.



Graph A



Graph B

The distribution of the data can be critical means data will not always be separated in linear form; it can be inseparable in linear plane. So, to separate this type of data there is a trick in SVM i.e., kernel. Kernel is used to transform the input in higher dimensional space so, the separation between the classes can be easy.

Functional Requirements:

The functional requirements for a system describe what system do.

1. The developed system should recognize identify, segment and classify flower into its respective class.
2. System must provide the quality of service to user.

Non-functional requirements:

Performance: Flower in the input image will be recognised with good accuracy.

Flexibility: It provides the users to load the image easily and displays the output.

Learning ability: The software is very easy to use and reduces the learning work.

There are four non-functional requirements?

a) Non-functional requirements or NFRs are a set of specifications that describe the system's operation capabilities and constraints and attempt to improve its functionality.

b) Performance and scalability. ...

c) Portability and compatibility. ...

d) Reliability, maintainability, availability.

The overview of the proposed flower recognition system on image processing is shown in Fig.1. The first step the image acquisitions. The input image file can be a file, which is already downloaded in the computer

memory or stored elsewhere and is readily available. A storage facility ‘in the Cloud’, such as Dropbox enhances the mobility of this system. The image may be a photo downloaded from a smart phone or digital camera. The image data is pre-processed to prepare the image data for analysis. In the image analysis section, Hu’s seven-moment algorithm of shapes together with RGB and HS data are used. After that, two data parts will be combined as a vector and classified by the K-nearest neighbour algorithm. Our system returns the top three most similar flower images. Finally, the output part shows input flower images, a set of candidate flower images and flower information.

Working of machine learning tools:

The use case view models functionality of the system as perceived by outside uses. A use case is a coherent unit of functionality expressed as a transaction among actors and the system.

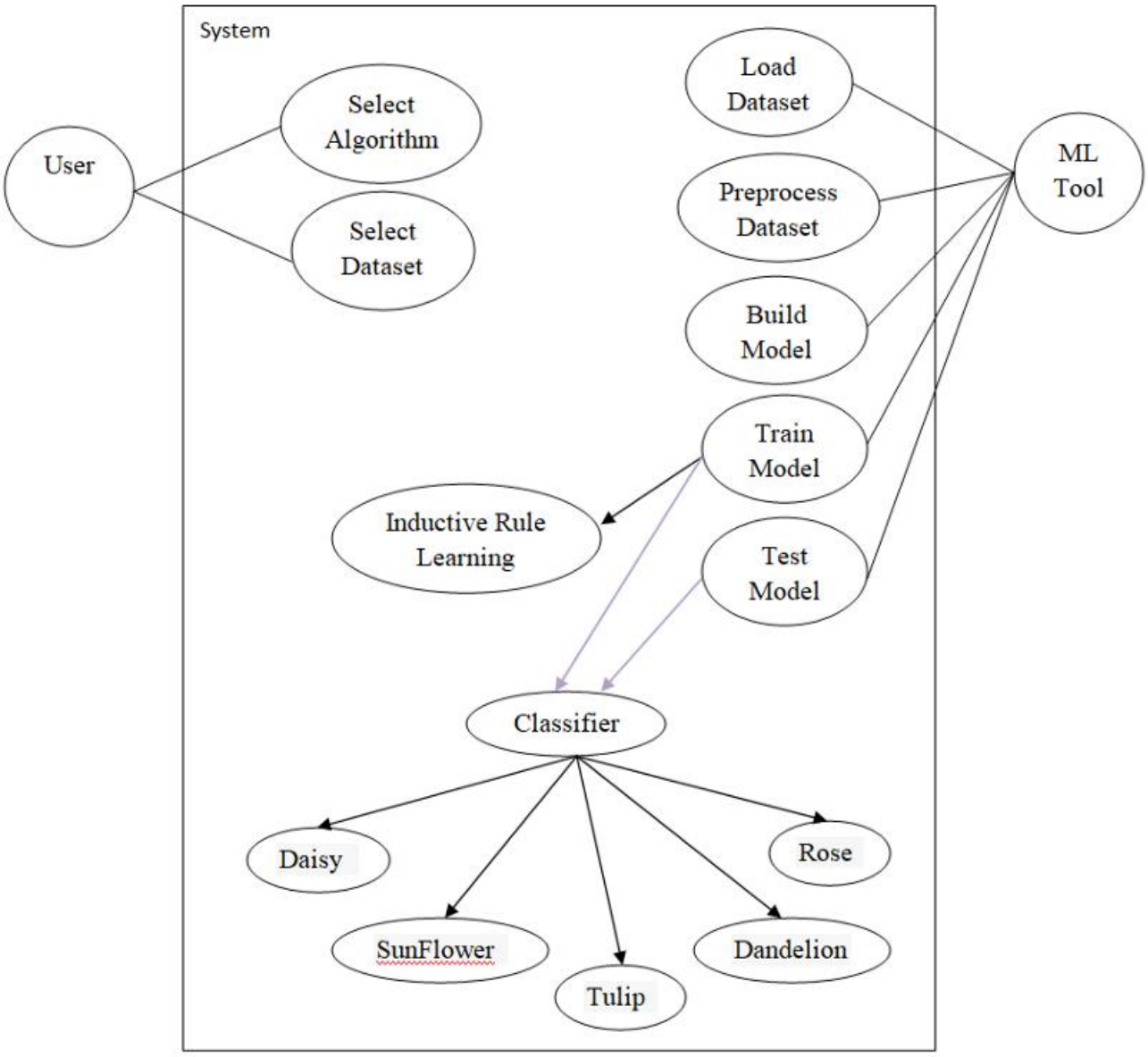


Fig. 3.8 working of machine learning tools

Flow Diagram of flower classification:

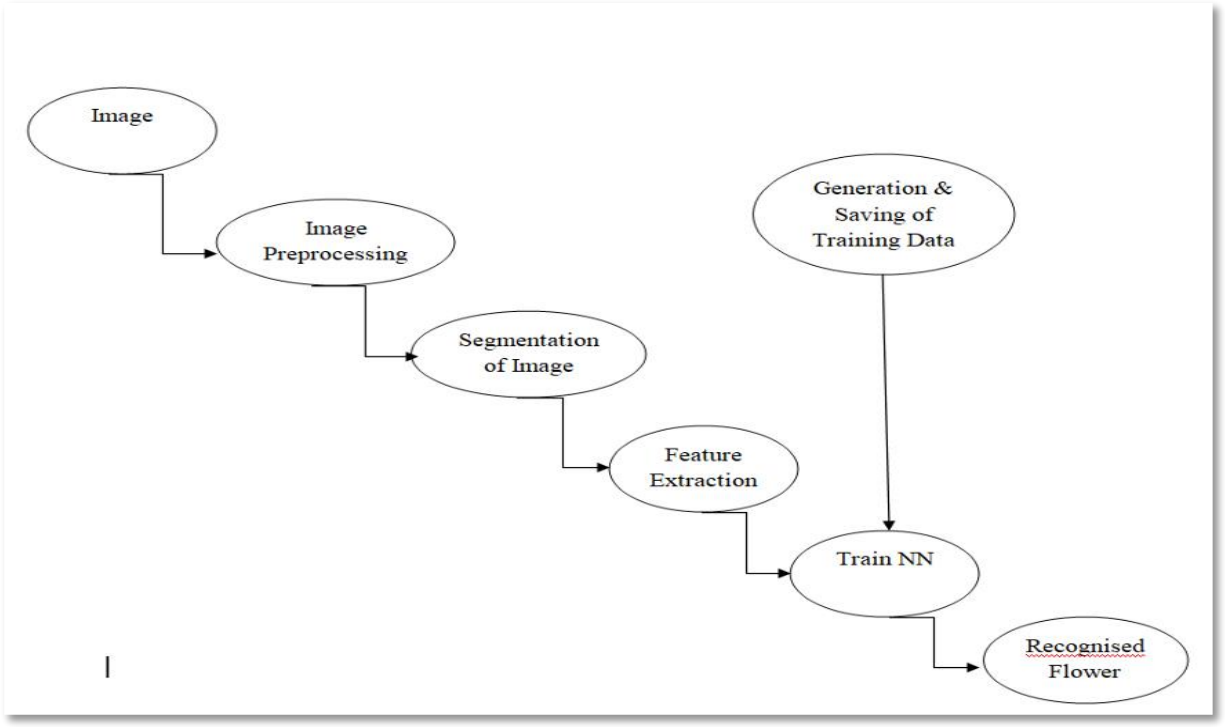


Fig. 3.9 Flow Diagram of flower classification

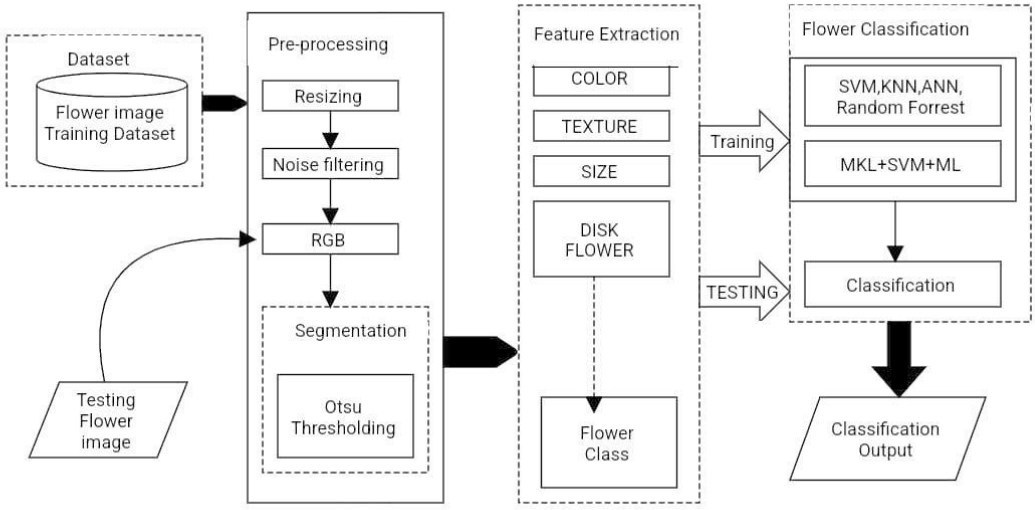


Fig.3.10 Proposed diagram of flower identification and classification

3.3 Database Specification

The Database:

The dataset, used for training the model is from a Kaggle Flowers Detection and Classification.

1. Now a day's it is very important to identify naturally occurring objects and recognize its type.

2. It is useful to identify flower type in various fields such as Gardening, botany research, Ayurveda treatment, Farming, floriculture etc.

3. Our literature survey must find that the different Image processing techniques used to extract features from flower and classify them using computational intelligence algorithms.

4. Digital image processing deals with manipulation of digital images through a digital computer. Digital image processing technologies such as Classification, Feature extraction, Pattern recognition etc. are useful to classify the images.

This dataset contains 7000 images of flowers.

The data collection is based on the data astilbe (flower), Google images, and flowers images.

Our 10 Flowers Dataset:

1. daisy (600+ Images)
2. dandelion (600+ Images)
3. roses (600+ Images)
4. sunflowers (600+ Images)
5. tulips (600+ Images)

1. .daisy (600+ Images)

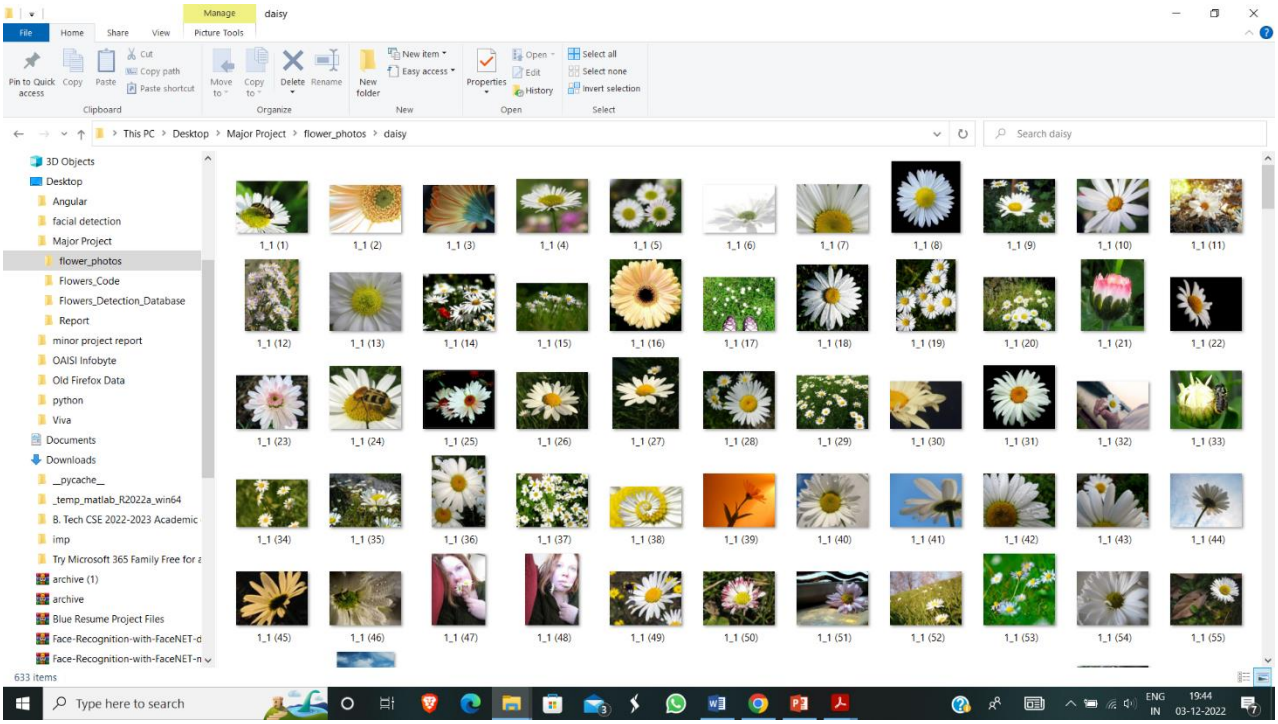
2. dandelion (600+ Images)

3. Roses (600+ Images)

4. Sunflowers (600+ Images)

5. Tulips (600+ Images)

Fig. 3.11 Flowers sample dataset



Flowers Size is same and rename all the flowers in serial no:

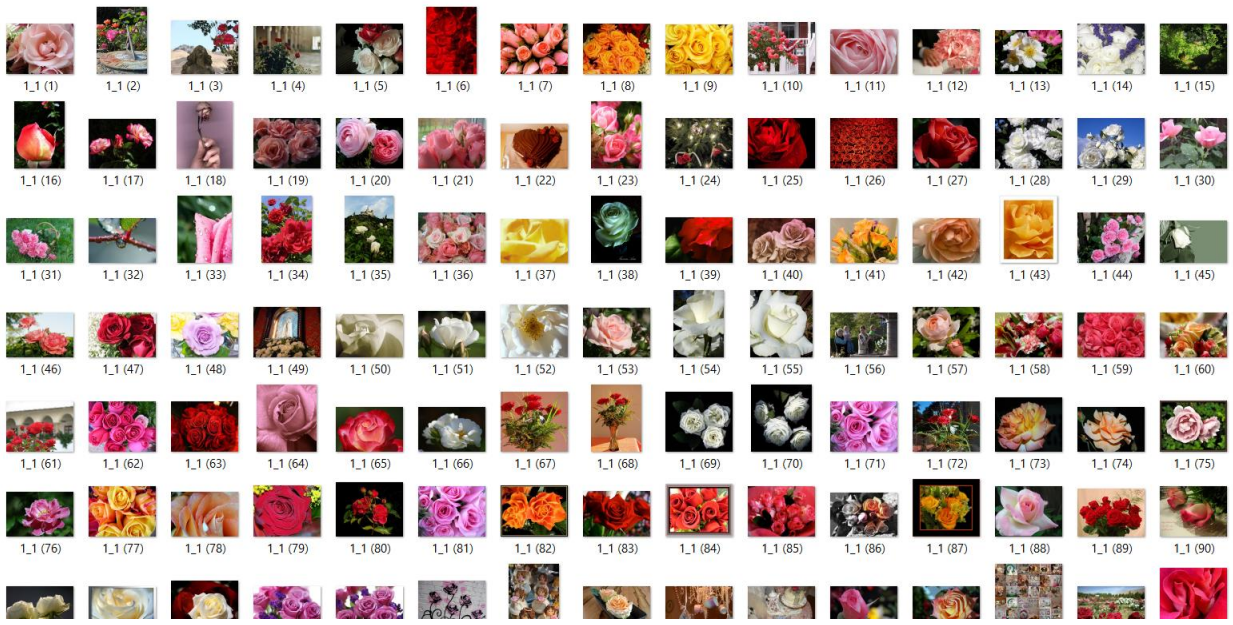


Fig. 3.12 renaming the dataset

EXPECTED OUTCOME

1. 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 flowers.
2. 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.

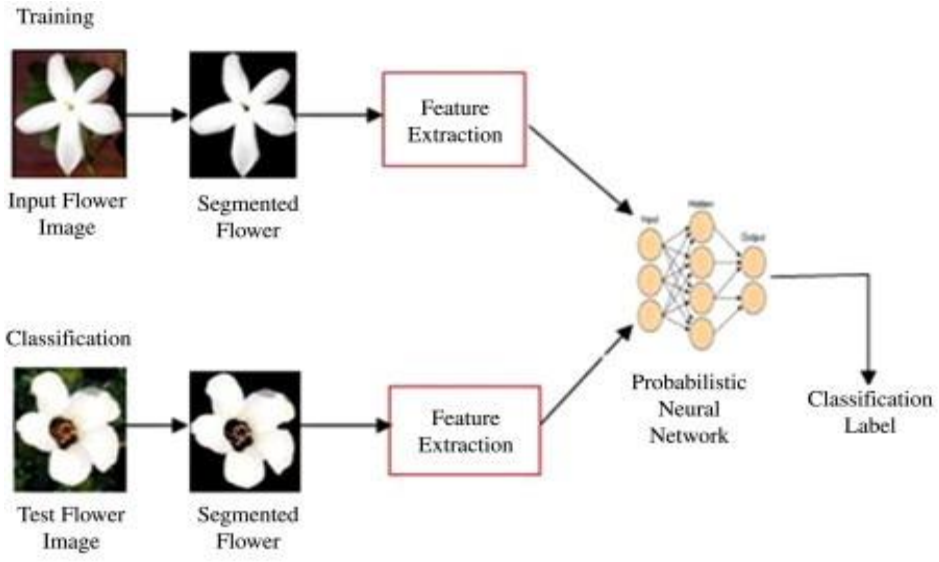
3.4 ALGORITHM:

K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on Supervised Learning technique-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most like the available categories-KNN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm-NN algorithm can be used for Regression as well as for Classification but mostly it is used for the Classification problems-NN is a non-parametric algorithm, which means it does not make any assumption on underlying data. It is also called a lazy learner algorithm because it does not learn from the training set immediately instead it stores the dataset and at the time of classification, it performs an action on the dataset algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much like the new data.

K can be kept as an any odd number so that we can calculate a clear majority in the case where only two groups are possible (e.g., Red/Blue). With increasing K, we get smoother, more defined boundaries across different classifications. Also, the accuracy of the above classifier increases as we increase the number of data points in the training set

Example: Suppose, we have an image of a creature that looks like cat and dog, but we want to know either it is a cat or dog. So, for this identification, we can use the KNN algorithm, as it works on a

similarity measure. Our KNN model will find the similar features of the new data set to the cats and dogs' images and based on the most similar features it will put it in either cat or dog category.



3.13 working of KNN algorithm

Steps for KNN Algorithm

Step 1: Collection of a data set of images, the dataset, used for training the model is from a Kaggle Flowers Detection and Classification, In the Database, there are 10 different flowers and each flowers have 700 images. (10*700) = 7000.

Step 2: Pre-processing of image

Step 3: Feature Selection of all images.

Step 4: Data Preparation.

Step 5: The NumPy array gets passed into the Convolution2D layer.

Step 6: Pooling method called MaxPooling2D that uses (2, 2) windows across the feature map only keeping the maximum pixel value.

Step 7: During training, Neural network Forward propagation and backward propagation performed on the pixel values.

Step 8: The model can show the detail probability to classification the flowers.

3.5 Basic Flower Features

1) **Colour:**

Various colour models including RGB, HSV, Lab, etc. are available to extract colour feature. To extract colour feature from a flower image, two colour models are used i.e., HSV and Lab colour space. HSV colour space normally having better representations about how people relate to colour than RGB colour model and it also generates high quality graphics. In HSV colour space, Hue, Saturation and Value refer to tint, shade and tone respectively. To approximate human vision, normally lab colour model is used. To add more precision for interpreting colour feature value, another prominent model, Lab colour space is also used. Lab colour model represents value for L as “lightness”, value for a as “green to red” and value for b as “blue to yellow”. The figure 4 depicts the values of both colour models.

2) **Size:**

Plants take completely different times to manoeuvre from germination to maturity once flowers develop all their procreative elements and therefore the life cycle begins once more. Width and height ratio is a ratio between the blooming of a flower maturity. Based on the height and width function measures, it is identified that flower is in blooming or bud stage. Size feature is captured by the centroid point and boundary box methods. The flower centroid and boundary tracing algorithm usually start from the beginning point. When it reaches to end, it stops. By this way, it gets the four flower boundaries that are partial and then it will have combined them to form the

whole flower boundary. Using this techniques, shape of flower image can properly have identified and extracted with width and height measurements

3) **Texture**

Texture feature is extracted using grey level co-occurrence matrix, abbreviated as GLCM. GLCM feature extraction is a mixture of grey levels co-occurs in a picture section or image. A quantitative description of a spatial pattern can efficiently provide by applying GLCMs. The four GLCM-derived options extracted from a flower image are homogeneity, energy, contrast, and correlation

4) **Corona**

One of the other prominent morphology features of a flower image is the corona (Crown). It is a characteristic that possesses androecial derived structure and found in Asclepiadaceous flowers with a variable form, colour and dimension, which is incorporated in the gynostemium. The corona typically consists of inner and outer lobes. Corona feature is captured by the k-means clustering method

5) **Petal Type:**

Modified and normally coloured leaves that surround the generative components of flower area unit are known as petals. They are formed to draw an attention of pollinators and they are usually brilliantly coloured. Together, all the leaves of a flower referred to as a Corolla. Petal type can be free or united. The free petal type structure is shown in. Applying the convex hull method and ripple image algorithm, the feature of flower petal-type which is free or united is extracted. Using the convex hull method, as it is used to represent description of region shape of flower image, the petal of flower boundary or edges can be converted in binary image and easily detected. After applying convex hull method, the ripple algorithm is applied for representation of the target in different view angles and sizes.

6) **No. of Petals:**

It is also essential to calculate number of petals remain present in a flower as the quantity of petals in flower has significantly helps to classify a plant. A composed form of all the petals remains in a flower image is represented by a Corolla. Corolla of different species of plants varies greatly in colour and pattern. This feature works well where the roundness of the flower is high. Sample of

several petals remain present in a flower image The Skeleton with DCE (Discrete Curve Evolution) technique is applied for counting the flower petals Discrete Skeleton Evolution (DSE) mainly applied to reduce a morphological or topological skeleton. It is a form of pruning that removes noisy or redundant branches generated by the skeletonization process. Moreover, it preserves information-rich trunk segments. The sample results of how to get number of petals using DSE method.

7) Aestivation of Petal:

The arrangement of petals or sepals within the flower is denoted as Aestivation. Different aestivation of petals is available includes.

I) Valvate: The floral petals are settled side by side.

(ii) Twisted: The petals are settled regular intersecting in one direction.

iii) Imbricate: The flower petals are completely in, one completely out and others are intersecting in one direction.

(iv) Quincuncial and Vexillary: It found only in pentamerous flower, and it is rate one. Aestivation of petal feature is captured by the Laplacian. The Laplacian of an image highlights regions of rapid intensity change and is therefore often used for edge detection. This operation in result produces such images which have greyish edge lines and other discontinuities on a dark background

8) Flower Class:

Monocots or Dicots are types of flower classes which can be easily identified by dividing number of petals. Where Monocots Petals are appeared in combinations of three, example is a Lily flower, Dicots Petals are seemed in combos of five or four. The example of a dicots flower is a Rose. To obtain the flower class, the number of petals is divided by 3 or 4 or 5. It is one of most required morphology features which facilitates the classification of a flower.

3.6 The Library & Packages (module):

1. Guido

Guido van Robot, or GvR for short, is a programming language and free software application designed to introduce beginners to the fundamentals of programming. GvR runs on Windows, Macintosh, and GNU/Linux, in a variety of languages! It's great in both the classroom and the home as a way of introducing people to the basic concepts of programming.

Are you interested in GvR, or do you have a bug report or question? Head on over to the Contact Us' page to find out how to send us an e-mail or how to get the latest version of the GvR source.

What's new in GvR?

2010-19-01

● Many bugfixes and many improvements to the GUI and syntax checking.

2008-11-09

● New redesigned layout to provide more room for the code editor.

2007-05-12

● GvRng has now a 'one window' design

● GvRng is now packed in one window, so you have everything available.

● No longer losing your editor windows or the language reference, everything in one place.

2007-16-06

GvRng is being ported to the olpc laptop

● Olpc stands for One Laptop Per Child, see the olpc website for more info about the laptop project

● Work on getting GvRng suitable for the laptop is just started and involves a redesign of the GUI.

● There's a screenshot available on the screenshots page.

2006-07-14

GvRng supports windows

GvRng is now also available for the windows platform.
There's not yet an installer but you can test the zip package after you've installed Python and GTK+.
A installer will be available soon.

2006-04-17

GvR is dead, long live GvRng !

● GvRng, which stands for GvR Next Generation, is a redesign of the original gvr to make it easier to maintain and more stable then the old one.

● From a users point of view the biggest change is that GvRng uses GTK as it's GUI and GTK Sourceview as the world and program editor.

● The drawback is that GvRng is no longer Windows compatible.

● Window users can use the old gvr which is based on wxPython.

● There are plans to create a Windows frontend for GvRng.

2. Python Programming Language:

Python is a high-level, general-purpose and a very popular programming language. Python programming language (latest Python 3) is being used in web development, Machine Learning applications, along with all cutting-edge technology in Software Industry. Python Programming Language is very well suited for Beginners, also for experienced programmers with other programming languages like C++ and Java.
This specially designed Python tutorial will help you learn Python Programming Language in most efficient way, with the topics from basics to advanced (like Web-scraping, Django, Deep-Learning, etc.) with examples.

Below are some facts about Python Programming Language:

1. Python is currently the most widely used multi-purpose, high-level programming language.
2. Python allows programming in Object-Oriented and Procedural paradigms.

3. Python programs generally are smaller than other programming languages like Java. Programmers must type relatively less and indentation requirement of the language, makes them readable all the time.
4. Python language is being used by almost all tech-giant companies like – Google, Amazon, Facebook, Instagram, Dropbox, Uber… etc.

5. The biggest strength of Python is huge collection of standard libraries which can be used for the following:
    - Machine Learning
    - GUI Applications (like Kivy, Tkinter, PyQt etc.)
    - Web frameworks like Django (used by YouTube, Instagram, Dropbox)
    - Image processing (like OpenCV, Pillow)
    - Web scraping (like Scrapy, Beautiful Soup, Selenium)
    - Test frameworks
    - Multimedia
    - Scientific computing
    - Text processing and many more.

Image classification is a method to classify way images into their respective category classes using some methods like:

- Training a small network from scratch
- Fine-tuning the top layers of the model using VGG16

Let's discuss how to train the model from scratch and classify the data containing cars and planes.

- **Train Data:** Train data contains the 200 images of each car and plane, i.e., in total, there are 400 images in the training dataset
- **Test Data:** Test data contains 50 images of each car and plane i.e., includes a total. There are 100 images in the test dataset.

3. NumPy

1. NumPy is a Python library used for working with arrays.
2. It also has functions for working in domain of linear algebra, Fourier transform, and matrices.
3. NumPy was created in 2005 by Travis Oliphant. It is an open-source project, and you can use it freely.
4. NumPy stands for Numerical Python.

It is open-source software. It contains various features including these important ones:

- A powerful N-dimensional array object
- Sophisticated (broadcasting) functions
- Tools for integrating C/C++ and Fortran code
- Useful linear algebra, Fourier transform, and random number capabilities

Besides its obvious scientific uses, NumPy can also be used as an efficient multi-dimensional container of generic data. Arbitrary datatypes can be defined using NumPy which allows NumPy to seamlessly and speedily integrate with a wide variety of databases.

Installation:

- **Mac** and **Linux** users can install NumPy via pip command:

```
pip install NumPy
```

- **Windows** does not have any package manager analogous to that in Linux or mac. Please download the pre-built windows installer for NumPy from (according to your system configuration and Python version). And then install the packages manually.

What is NumPy for

NumPy is an important library for:

- Data Science
- Machine learning
- Signal and image processing
- Scientific and engineer computing

Summary

- NumPy stands for Numerical Python.
- It's a Python library for numerical calculation.
- Use np as the alias for the NumPy module.

4. Matplotlib

Matplotlib is an amazing visualization library in Python for 2D plots of arrays. Matplotlib is a multi-platform data visualization library built on NumPy arrays and designed to work with the broader SciPy stack. It was introduced by John Hunter in the year 2002.

One of the greatest benefits of visualization is that it allows us visual access to huge amounts of data in easily digestible visuals. Matplotlib consists of several plots like line, bar, scatter, histogram etc.

Installation:

Windows, Linux and macOS distributions have matplotlib and most of its dependencies as wheel packages.

Run the following command to install matplotlib package:

```
Python -mpip install -U matplotlib
```

Importing matplotlib:

```
from matplotlib import pyplot as plt

or

import matplotlib.pyplot as plt
```

5. Google Colab

Google Collaboratory, popularly known as Colab, is a web IDE for python that was released by Google in 2017. Colab is an excellent tool for data scientists to execute Machine Learning and Deep Learning projects with cloud storage capabilities.

Colab is basically a cloud based Jupyter notebook environment that requires no setup. What's more, it provides its users free access to high compute resources such as GPUs and TPUs that are essential to training models quickly and more efficiently!

Google Colab was developed by Google to provide free access to GPU's and TPU's to anyone who needs them to build a machine learning or deep learning model. Google Colab can be defined as an improved version of Jupyter Notebook.

Google Colab Features

Google Colab provides tons of exciting features that any modern IDE offers, and much more. Some of the most exciting features are listed below.

- Interactive tutorials to learn machine learning and neural networks.
- Write and execute Python 3 code without having a local setup.
- Execute terminal commands from the Notebook.
- Import datasets from external sources such as Kaggle.
- Save your Notebooks to Google Drive.
- Import Notebooks from Google Drive.
- Free cloud service, GPUs and TPUs.
- Integrate with PyTorch, Tensor Flow, Open CV.
- Import or publish directly from/to GitHub.

TensorFlow

TensorFlow is an open-source library developed by Google primarily for deep learning applications. It also supports traditional machine learning. TensorFlow was originally developed for large numerical computations without keeping deep learning in mind. However, it proved to be very useful for deep learning development as well, and therefore Google open-sourced it.

TensorFlow accepts data in the form of multi-dimensional arrays of higher dimensions called tensors. Multi-dimensional arrays are very handy in handling large amounts of data.

TensorFlow works based on data flow graphs that have nodes and edges. As the execution mechanism is in the form of graphs, it is much easier to execute TensorFlow code in a distributed manner across a cluster of computers while using GPUs.

TensorFlow allows developers to create dataflow graphs—structures that describe how data moves through a graph, or a series of processing nodes. Each node in the graph represents a mathematical operation, and each connection or edge between nodes is a multidimensional data array, or tensor.

TensorFlow applications can be run on most any target that's convenient: a local machine, a cluster in the cloud, iOS and Android devices, CPUs or GPUs. If you use Google's own cloud, you can run TensorFlow on Google's custom TensorFlow Processing Unit (TPU) silicon for further acceleration. The resulting models created by TensorFlow, though, can be deployed on most any device where they will be used to serve predictions.

The single biggest benefit TensorFlow provides for machine learning development is abstraction. Instead of dealing with the nitty-gritty details of implementing algorithms or figuring out proper ways to hitch the output of one function to the input of another, the developer can focus on the overall application logic. TensorFlow takes care of the details behind the scenes.

TensorFlow offers additional conveniences for developers who need to debug and gain introspection into TensorFlow apps. Each graph operation can be evaluated and modified separately and transparently, instead of constructing the entire graph as a single opaque object and evaluating it all at once. This so-called "eager execution mode," provided as an option in older versions of TensorFlow, is now standard.

How to Install TensorFlow?

Installation of TensorFlow is straightforward if you already have a Python SciPy environment. TensorFlow works with Python 3.3+. You can follow the Download and Setup instructions on the TensorFlow website. Installation is probably simplest via PyPI, and specific instructions of the pip command to use for your Linux or Mac OS X platform are on the Download and Setup webpage. In the simplest case, you just need to enter the following in your command line:

Pip Install TensorFlow

An exception would be on the newer Mac with an Apple Silicon CPU. The package name for this specific architecture is TensorFlow macOS instead:

TensorFlow architecture works in three significant steps:

1. Data pre-processing - structure the data and brings it under one limiting value.
2. Building the model - build the model for the data.
3. Training and estimating the model - use the data to train the model and test it with unknown data.

3.7 HARDWERE / SOFTWARE USED:

1) Software Requirement

i) Python Language
ii) Numpy
iii) Matplotlib
iv) Gradio
v) Google Colab

vi) Tensorflow / Keras

vii) Os

2) Hardware Requirement

Computer System with 8 Gb Ram and I5 Processor.

Chapter 4

PERFORMANCE ANALYSIS

4.1 Performance Analysis

In Machine Learning, the performance evaluation metrics are used to calculate the performance of your trained machine learning models. This helps in finding how better your machine learning model can perform on a dataset that it has never seen before. If you have never used any performance evaluation metrics to evaluate the performance of your machine learning model, then this article is for you. In this article, I will take you through an introduction to some of the best performance evaluation metrics in machine learning.

There are many performance evaluation metrics that you can use to measure the performance of your machine learning models for classification as well as for regression. Below are some of the best performance evaluation metrics that I will recommend you use to evaluate the performance of your machine learning model. All the performance evaluation metrics mentioned below are solved and explained by using the Python programming language.

To measure the performance of the classification algorithms can be obtained through accuracy, precision, recall, and F-measure.

Table 4.1 Description of performance analysis

Sr. No.	**Performance Analysis**	**Formula**
1.	Accuracy	TP+TN / TP+FP+FN+TN
2.	Precision	TP / TP+FP
3.	Recall	TP / TP+FN
4.	F-measure	2*(Recall*Precision) / (Recall Precision)

The performance of KNN, Random Forest, ANN, SVM and MKL+SVM with its accuracy and other parameters. The KNN, Random Forest and ANN have not worked well while SVM achieved good accuracy with 90% accuracy. However, the hybrid MKL + SVM achieved the highest accuracy among the other classifiers with 95% accuracy.

Table 4.2 Performance Analysis with various classifier

Algorithm	Accuracy	Precision	Recall	F-Measure
KNN	0.21	0.195	0.203	0.201
Random Forest	0.69	0.687	0.690	0.688
ANN	0.71	0.705	0.708	0.706
SVM	0.72	0.720	0.721	0.720
MKL-SVM	0.78	0.775	0.770	0.774

4.2 Gantt Chart:

4.3 Output:

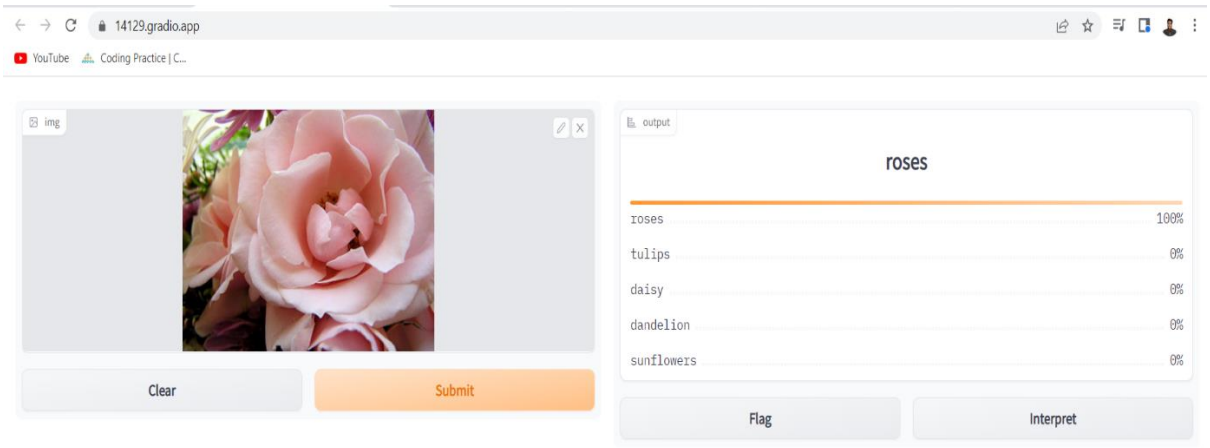


Fig 4.1 Expected Output

Chapter 5

CONCLUSIONS

With the rapid development of technology, AI is being used in various fields. Machine learning is the most basic method to achieve AI. This research describes the work principle of machine learning and an application of machine learning. At the beginning of development of Artificial Intelligence (AI), the AI system does not have a thorough learning ability so the whole system is not perfect. For instance, when the computer faces problems, it cannot be self-adjusting. Moreover, the computer cannot automatically collect and discover new knowledge. Therefore, computer only can be conducted by already existing truths. It does not have the ability to discover a new logical theory, rules and so on. The 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 plants while outdoors or offline. We will continue our research to make the system more efficient.

The question of the total species of flower being known is divided into three parts. First thing is image characteristics are retrieved from the training dataset using Convolution Neural Network and stored to format HDF5 files. Secondly, the network will be trained using various machine learning classifiers, such as Bagging Tress, Linear Classification Analysis, Gaussian Naive Bayes, K-Nearest Neighbour, Logistic Regression, Decision Tress, Random Forests and Stochastic Gradient Boosting. Finally, the random test images are given to the network for label prediction to assess the accuracy of the device. The software correctly identifies dataset machine learning classifier

REFERENCE

[1] Aalaa Albadarneh and Ashraf Ahmad," Automated Flower Species Detection and Recognition from Digital Images", IJCSNS International Journal of Computer Science and Network Security, VOL.17 No.4, April 2017.

[2] Huthaifa Almogdady, Dr. Saher Manaseer and Dr.Hazem Hiary," A Flower Recognition System Based On Image Processing and Neural Networks", INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 7, ISSUE 11, NOVEMBER 2018.

[3] Chaku Gamit, Prof. Prashant B. Swadas and Prof. Nilesh B. Prajapati," Literature Review on Flower Classification", International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181, Vol. 4 Issue 02, February-2015.

[4] Tanakorn Tiay, Pipimphorn Benyaphaichit, and Panomkhawn Riyamongkol," Flower Recognition System Based on Image Processing", 2014 Third ICT International Student Project International Journal of Computer Sciences and Engineering

[5] https://www.neuraldesigner.com/learning/examples/iris-flowers-classification

[6]extension://efaidnbmnnnibpcajpcglclefindmkaj/https://ijariie.com/AdminUploadPdf/Flower_Image_Detection_ijariie16409.pdf

[7] https://www.kaggle.com/datasets/alxmamaev/flowers-recognition

[8] M. Das, R. Magmata, and E. M. Riseman, "Indexing flower patent images using domain knowledge", IEEE Intell. Syst., vol 14, pp. 24-33, 1999.

[9] T. Saitoh and T. Kaneko, "Automatic recognition of wildflowers", System and Computer in Japan, vol. 34, no. 10, 2003.

[10] M. E. Nilsback and A. Zisserman, “A visual vocabulary for flower classification”, In CVPR, 2006.

[11] Y. Chai, “Recognition between a large number of flower species”, Master thesis, University of Oxford 2011