"COLOR DETECTION USING OPEN-CV"

In partial fulfillment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY
IN
INFORMATION TECHNOLOGY

Submitted by

G PHANINDAR	G ABHINAYA SRI	M AKILA
17911A1215	17911A1214	17911A1232

Under the Esteemed Guidance of
Mr. Devakishan Adla
Associate professor
Department of Information Technology

DEPARTMENT OF INFORMATION TECHNOLOGY

VIDYA JYOTHI INSTITUTE OF TECHNOLOGY

(Accredited by NBA, Approved by AICTE, Affiliated to JNTU Hyderabad)
Aziz Nagar Gate, C.B.Post, Chilkur Road, Hyderabad – 500075
2020- 2021

VIDYA JYOTHI INSTITUTE OF TECHNOLOGY

(Accredited by NBA, Approved by AICTE, Affiliated to JNTU Hyderabad)
Aziz Nagar Gate, C.B.Post, Chilkur Road, Hyderabad – 500075

DEPARTMENT OF INFORMATION TECHNOLOGY

CERTIFICATE

This is to certify that the project report on **"COLOR DETECTION USING OPEN-CV"** is a bonafide work done by **G PHANINDAR (17911A1215)** , **G ABHINAYA SRI (17911A1214) and M AKILA (17911A1232)** in partial fulfillment of the requirement for the award of the degree of Bachelor of Technology in **"INFORMATION TECHNOLOGY"** JNTU Hyderabad during the year **2020- 2021.**

Project Guide   **HOD**

Mr. Devakishan Adla   Mr. B. Srinivasulu
Associate Professor, IT Dept   Professor, IT Dept

External Examiner

VIDYA JYOTHI INSTITUTE OF TECHNOLOGY

(Accredited by NBA, Approved by AICTE, Affiliated to JNTU Hyderabad)
Aziz Nagar Gate, C.B.Post, Chilkur Road, Hyderabad – 500075

DEPARTMENT OF INFORMATION TECHNOLOGY

DECLARATION

We **G PHANINDAR (17911A1215)**, **G ABHINAYA SRI (17911A1214) and M AKILA (17911A1232)** hereby declare that the project report entitled "COLOR DETECTION USING OPEN-CV" is submitted in the partial fulfillment of the requirement for the award of **Bachelor of Technology** in Information Technology at **VIDYA JYOTHI INSTITUTE OF TECHNOLOGY**, affiliated to JNTUH is an authentic work and has not been submitted to any other university or institute for the degree.

G PHANINDAR (17911A1215)
G ABHINAYA SRI (17911A1214)
M AKILA (17911A1232)

ACKNOWLEDGEMENT

CONTENTS

ABSTRACT

In today's world color also plays an important role in various areas. The concept of delivering the color name to the user found as a helpful task to the user and for the people who deal with colors as a major part. So, our project is built upon this idea of easily responding to the user with the specified color name. This project has a lot of scope in future and it is user-friendly, easily understandable to the user.

Color detection is the process of detecting the name of any color. For humans this is an extremely easy task but for computers, it is not straightforward, because it has to register the colors with in it to find out the color name. Human eyes and brain's work together to translate light into color. Human eyes can register 17 million colors and Light receptors that are present in our eyes transmit the signal to the brain. Our brain then recognizes the color. Since childhood, we have mapped certain lights with their color names. We will be using the somewhat same strategy to detect color names.

In this color detection project, we are going to build an application through which we can automatically get the name of the color by clicking on the Image. So, for this, we will have a data file that contains the color name and its RGB and Hex values. We are using a dataset

colors.csv from kaggle.com which contains 2000+ records and then we will calculate the distance from each color and find the shortest one. The dependencies used in this project are Open-CV, pandas, NumPy and etc.

In this OpenCV color detection system there are major modules like scan image, match colors and system results. If the color matched with the color of RGB color model then the system results with the correct output. This project has various applications in the field of image processing and other fields like medical diagnosis, skin color detection, animation, image editing apps and many of the drawing apps. It plays a very important role in the sign language converter where it uses the algorithm in which it takes the color name of the user and if that matches then only the hand is visible. The outcome of the Color Detection project is to get the colors of an image and it is useful in numerous images editing and drawing apps.

LIST OF FIGURES

LIST OF ABBREVATIONS

Open- CV	Open Source Computer Vision
NumPy	Library Numerical Python
CSV	Comma-Separated value
RGB	Red-Green-Blue
LED	Light Emitting Diode
PIP	Package Installer for Python
PyPI	Python Package Index
SRS	Software Requirements Specification
SyRS	System Requirements Specification
SDLC	Software Design Life Cycle
UML	Unified Modelling Language

1.INTRODUCTION

The main idea of the project is to detect the respective color of the image when the cursor is placed at a particular point on the image. It allows the user to find out the name of the color at any tiny pixel point. It is very effective in the field of Image Processing. The development of this project is done through some of the packages in python like OpenCV, Pandas, NumPy and etc. It provides an environment where user can easily deal with colors. Human eyes can see nearly 17million of colors, but it is not that easy to remember all those names of colors. And there exist some colors which human eye can't register. So, our project objective is to detect the color and display the name of the color to the user. The project is initially trained with a data set which has different colors names registered in it. Since childhood, we have mapped certain lights with their color names. We will be using the somewhat same strategy to detect color names in this project. As the program is trained with the data set, it can detect any of the image color name effectively. If user wants to know the color of an image pixel point, then the user can simply provide the image and place the cursor on the image and double click on it to get the color name. The input is taken as an image and there some RGB calculations of the pixel will be taken place and then the color name is displayed as output. The output also represents the RGB values. Working with the dataset is the major task in the data analysis and this can be done using the Pandas package where it is known as the backbone of the data projects.

Pandas is used to extract the data from the CSV file and load it into the dataset. Its major task is to calculate the statistics, clean the data, visualize the data and the store back the data into the CSV file. NumPy is used to work with the Arrays as it is 50 times faster than working with the lists. This project has applications in various fields like Numerous Image Editing Apps, Drawing Apps, In the field of Image Processing, RGB LED consistency control, Animations, Industrial process control and Medical diagnosis Control.

This project can be operated through windows, Linux, Mac and the coding language we used to build this project is python. Tools used as Visual studio code to write the code and some of the packages like OpenCV, Pandas, NumPy and etc. It performs all the operations in finding out the name of the color and if it is successfully executed then the result is displayed as the output as the color name and the respective RGB values. The hardware components used to build the project are system with i3 core or higher, hard disk of 25GB (SSD preferred), 2GB ram, Internet connection of 1Mbps or higher. The Unified Modelling diagrams used in the project are Class Diagram and Sequence Diagram.

1.1 MOTIVATION:

The motivation behind making this project is to make the life of people easier as much as possible, to solve the complex problems of this world through which we can make this globe a better place for living. This world is full of colors which drives us to make a setup which can tell someone about the color of a particular thing. So this is the driving force which make this imagination turning into reality. The main motto of this project is to provide the user and easy and an efficient way to know the color name. A lot of efforts are made in implementing this project. The project responsiveness also plays a major role. In the last decade, there has been vast increase in the development of color recognition techniques for the analysis of images. This increase stems from the fact that there are many important real world applications in which color detection can be applied like vehicle color detection, skin color detection, image segmentation, clothing color recognition. The process of creating an application which will be helpful to the users and which will grab the attention of them with its advantages is a major task in today's world. It allows the user to find out the name of the color at any tiny pixel point. It is very effective in the field of Image Processing. It provides an environment where user can easily deal with colors. This idea of developing a useful project on the basis color turned into motivation in implementing this project.

1.2 PROBLEM DEFINATION:

There exist some colors which human eye can't register. So, our project objective is to detect the color and display the name of the color to the user. The project is initially trained with a data set which has different colors names registered in it. Since childhood, we have mapped certain lights with their color names. We will be using the somewhat same strategy to detect color names in this project. As the program is trained with the data set, it can detect any of the image color name effectively.

Our project aims to provide

- Better Accuracy in segmentation under various illuminations

- Less time-consuming process

1.3 OBJECTIVE OF PROJECT:

1. Human eyes can see nearly 17 million of colors, but it is not that easy to remember all those names of colors. And there exist some colors which human eye can't register. So, our project objective is to detect the color and display the name of the color to the user.

2. The project is initially trained with a data set which has different colors names registered in it. So that when the user gives the input and request for the output it results in the accurate answer.

3. To provide an efficient way in retrieving the color to the user.

4. Since childhood, we have mapped certain lights with their color names. We will use the somewhat same strategy like mapping the color names to the dataset to detect color names in this project. This provides an efficient way of detecting color names.

5. To provide the easy access to the user in detecting the color is a major objective.

1.4 LIMITATIONS OF THE PROJECT:

1. Although the Project has benefits on its way, it also has some of the limitations. In which one of the factors is the major one which is the project works only with Images.

2. The project only accepts images as the input and gives the color name, Hex value and RGB values as output. It cannot work with the videos.

3. Even though the project is built on the packages like Open-CV which is a major source of computer vision, there is no scope of dealing with the live camera actions as this project is exclusively built for images.

1.5 ORGANIZATION OF DOCUMENTATION:

In this project documentation, we have described the motivation towards the project, problem definition and objective of the project initially, and followed by Literature survey, Analysis, Design, implementation & Results and Testing & validations phases.

After the brief description, it will be followed by the conclusion. The conclusion part consists of the project conclusion and the future enhancement of the project.

Later it is followed by the references.

2.LITERATURE SURVEY

2.1 INTRODUCTION:

As there is a lot of applications in the field of image editing, drawing apps, animations and etc., this project create a lot of use in those areas. As we know that a color is most important factor in image editing, the color is made as the major role. This project is built on different packages in python like OpenCV, Pandas, NumPy and etc. A lot of study has undergone to know about the packages most importantly. Firstly, about OpenCV there are lots of advantages of this package while working with the computer vision tasks. It gives an easy understanding to the user to implement the application. As it is a major source for the computer vision tasks, it helped to build the project successfully.

Later coming to pandas , the basic thing we know about pandas is that it is very fast and flexible tool to work in data analytics field but while working on the project we understood how actually does the pandas play a very important role in loading the dataset and calculate the Statistics and answer questions about the data. Clean the data by doing things like removing missing values and filtering rows by some criteria. Visualize the data with the help from Matplotlib, Plot bars, Lines, histograms. Store the cleaned, transformed, data back into a csv. NumPy is used for working with the arrays. We are using this because it is 50 times faster than the python lists. NumPy arrays are stored at one continuous place in memory unlike lists, so processes can access and manipulate them very efficiently. It also contains N-Dimensional array object, broadcasting functions, Tools for Integrating and it also uses linear algebra, Fourier transformations. Argparse is used to load the image, which is the recommended command line parsing module in python library giving the user the ability to enter command line arguments which provides flexibility.

 A lot of study is made on collecting the data set. We have gone through some websites like dataworld, AnalyticalVidhya, UCI machine learning Repository, Kaggle.com. Finally, we collected the data set from Kaggle.com. They have more than 350 datasets in total – with more than 200 as Featured datasets. While some of the initial datasets were usually present at other places, we have seen a few interesting datasets on the platform, not present at other places. Along with new datasets, another benefit of the interface is that we can see scripts and questions from community members on the same interface. In the project when the mouse double click event occurs it is important for one to study how the function retrieves the pixels. And we made an interesting effort in doing that. Image basics: Images consist of a set of pixels. These pixels are the fundamental building blocks of an image. A Pixel can be thought of the "color" or the "intensity" of light that appears at any given place in our image. Each pixel can be in two forms, the first one is Grayscale where each pixel has a value between 0 and 255 and This value represents each darker color shift from white through to black. And the second one is Color where Color is made up of three RGB components (red, green and blue). We store three sets of values between 0 and 255. Each of these values is normally an 8-bit unsigned integer. These three 8bit values are combined into an RGB tuple in the form of (red, green, blue). The image coordinates start from the top left as (0,0). NumPy provides a technique called array slicing.

2.1.1 OPEN-CV:

OpenCV is a cross-platform library using which we can develop real-time computer vision applications. It mainly focuses on image processing, video capture and analysis including features like face detection and object detection.

Let's start the chapter by defining the term "Computer Vision". Computer Vision:

Computer Vision can be defined as a discipline that explains how to reconstruct, interrupt, and understand a 3D scene from its 2D images, in terms of the properties of the structure present in the scene. It deals with modelling and replicating human vision using computer software and hardware.

Computer Vision overlaps significantly with the following fields −

    ☐ Image Processing − It focuses on image manipulation.

    ☐ Pattern Recognition − It explains various techniques to classify patterns.

    ☐ Photogrammetry − It is concerned with obtaining accurate measurements from images.

Computer Vision Vs Image Processing:

Image processing deals with image-to-image transformation. The input and output of image processing are both images.

Computer vision is the construction of explicit, meaningful descriptions of physical objects from their image. The output of computer vision is a description or an interpretation of structures in 3D scene.
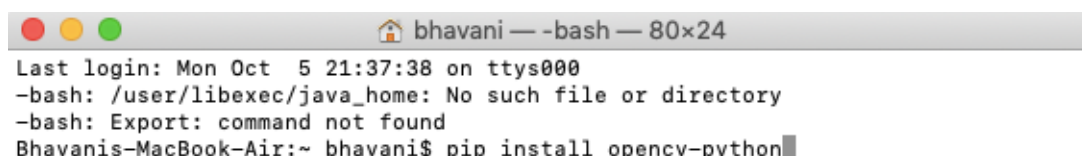
Installation of open-cv:

In-order to install Open-cv we need PIP.PIP is a package management system used to install and manage software packages/libraries written in Python. These files are stored in a large "on- line repository" termed as Python Package Index (PyPI). pip uses PyPI as the default source for packages and their dependencies. pip uses PyPI as the default source for packages and their dependencies.

>>pip -v

OpenCV can be directly downloaded and installed with the use of pip (package manager). To install OpenCV, just go to the command-line and type the following command:

>pip install opencv-python

```
● ● ●                    🏠 bhavani — -bash — 80×24
Last login: Mon Oct  5 21:37:38 on ttys000
-bash: /user/libexec/java_home: No such file or directory
-bash: Export: command not found
Bhavanis-MacBook-Air:~ bhavani$ pip install opencv-python ▉
```

Fig 1. Installation of open-cv

To check if OpenCV is correctly installed, just run the following commands to perform a version check:

Python3
>>>import cv2
>>>print(cv2._version_)

Fig 2. Verifying of opencv installation

How Does Computer Vision Work?

As a human, processing images is a natural process that we tend to give little to no thought to in our daily lives. Computers, on the other hand, have to abide by their own unique processes in order to analyse massive amounts of media. In order for deep learning computer vision to take root, thousands on thousands of photos, videos, and other images need to be compiled for an effective AI to become useful.
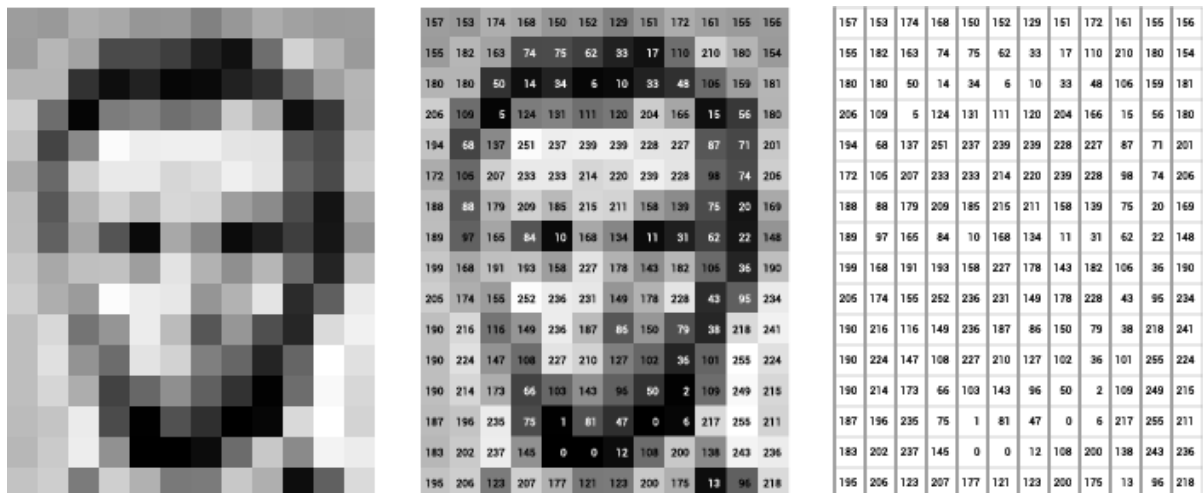

Fig 3. RGB values of an Image

6

By breaking each image down to base components labelled by numeric value, the computer is then able to interpret and store each image based on the resulting array of numbers. The image on the left is what is provided, while the image on the right is how the computer actually processes the image.

The numerical value for each pixel can range from 0-255 in a basic image like the one provided, but that's only because it's in grayscale. One you throw in color, things get a little more complicated.

When accounting for color, the numerical value per pixel jumps from one to three. Every color can be generated through a mix of the three primary colors: red, green, and blue. In order for the computer vision software to properly interpret a color image, it needs to apply a numerical value to each of the three primary colors in the image - hence the three numerical values.

ADVANTAGES OF OPEN-CV:

1. Fast prototyping
   Since we can focus more on the design, we can now experiment with more design ideas. Python is well-suited for implementing new features. Libraries like OpenCV are written in C++ and make Python have slower runtime as it will still call C/C++ libraries. This means we will have the development advantage from Python while we can have performance optimization from C++.

2. Ease of coding
   "Code as Plain English" is Python's primary goal. This allows programmers to focus on the design and not on coding. This is perfect for those who are just getting started with machine learning or basic programming. This advantage is very beneficial, especially when faced with complex scenarios.

3. Vast libraries for machine learning
   Python is commonly used for machine learning. Data scientists invest their time contributing since it's easy to code, and it's free. CV developers don't have to worry much about projects that they're working on since most of their cases are already covered by Python libraries.

4. Low RAM usage (approximately 60–70 mb)

5. It is portable as OpenCV can run on any device that can run C

2.1.2 PANDAS:

Python is a great language for doing data analysis, primarily because of the fantastic ecosystem of data-centric python packages. Pandas is one of those packages and makes importing and analysing data much easier.

Pandas is a Python package that provides fast, flexible, and expressive data structures designed to make working with structured (tabular, multidimensional, potentially heterogeneous) and time series data both easy and intuitive. It aims to be the fundamental high-level building block for doing practical, real world data analysis in Python. Additionally, it has the broader goal of becoming the most powerful and flexible open source data analysis / manipulation tool available in any language. It is already well on its way toward this goal.

pandas is well suited for many different kinds of data:

- Tabular data with heterogeneously-typed columns, as in an SQL table or Excel spreadsheet
- Ordered and unordered (not necessarily fixed-frequency) time series data.
- Arbitrary matrix data (homogeneously typed or heterogeneous) with row and column labels
- Any other form of observational / statistical data sets. The data actually need not be labeled at all to be placed into a pandas data structure

It provides highly optimized performance with back-end source code is purely written in c or python. Pandas is the backbone of most data projects.
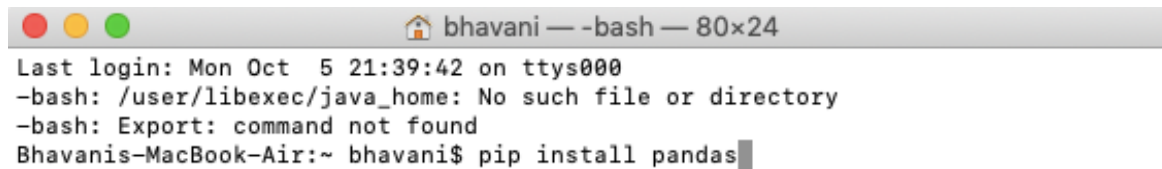
Pandas will extract the data from that CSV into data frame-a table, basically then we can perform the things we like:

1. Calculate statistics and answer the questions about the data

2. Clean the data by doing things like removing missing values and filtering rows by some criteria

3. Visualize the data with the help from Matplotlib, Plotbars, lines, Histograms and etc.

4. Store the cleaned, transformed data back into the CSV , other files or database.

Installation of Pandas:

A lightweight alternative is to install NumPy using popular Python package installer, pip.

Command as follows pip install pandas

Fig 4. Installation of Pandas

To check if Pandas is correctly installed, just run the following commands to perform a version check:

Python3
>>>import pandas
>>>print(pandas._version_)

Fig 5. Verifying Pandas Installation

2.1.3 CSV (comma-separated value)

CSV (comma-separated value) files are a common file format for transferring and storing data. The ability to read, manipulate, and write data to and from CSV files using Python is a key skill to master for any data scientist or business analysis.

A "CSV" file, that is, a file with a "csv" filetype, is a basic text file. Any text editor such as Notepad on windows or Text edit on Mac, can open a CSV file and show the contents. Sublime text is a wonderful and multi-functional text editor option for any platform.
CSV is a standard for storing tabular data in text format, where commas are used to separate the different columns, and newlines (carriage return / press enter) used to separate rows. Typically, the first row in a CSV file contains the names of the columns for the data.

Example table data set and the corresponding CSV-format data is shown in the diagram below.
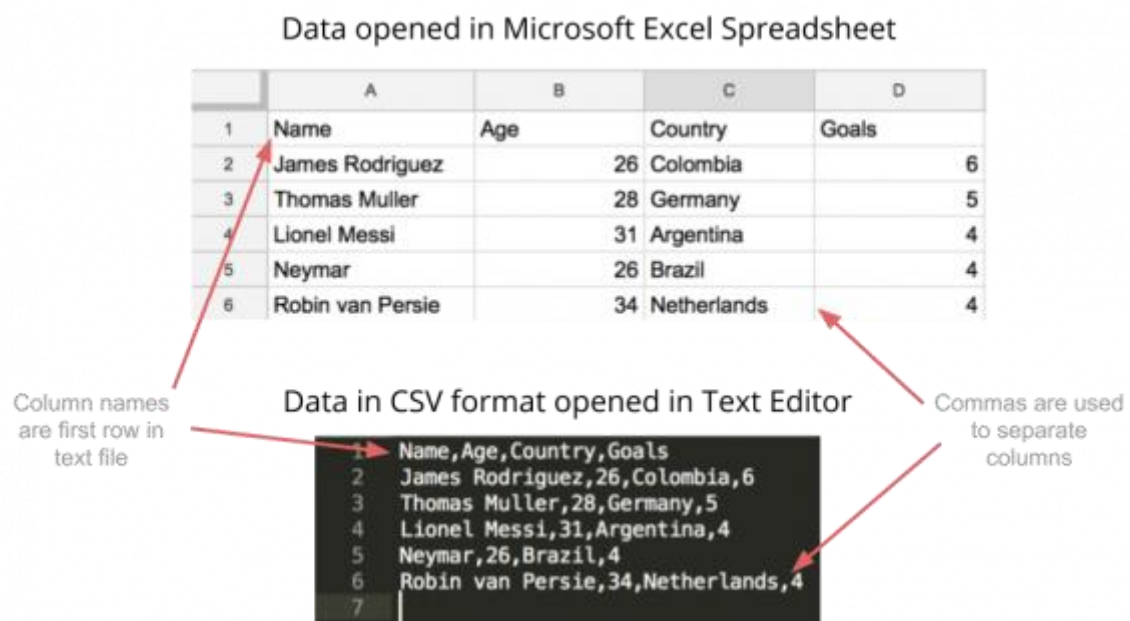
Fig 6. CSV File-format data

Comma-separated value files, or CSV files, are simple text files where commas and newlines are used to define tabular data in a structured way.

Any tabular data can be stored in CSV format – the format is popular because of its simplicity and flexibility. We can create a text file in a text editor, save it with a .csv extension, and open that file in Excel or Google Sheets to see the table form.

Loading a CSV File to Python pandas:

Import pandas import pandas as pd

reading csv file pd.read_csv("filename.csv")

Opening a CSV file through this is easy. But there are many others thing one can do through this function only to change the returned object completely. For instance, one can read a csv file not only locally, but from a URL through read_csv or one can choose what columns needed to export so that we don't have to edit the array later.

Key Features of Pandas

- ☐ Fast and efficient Data Frame object with default and customized indexing.
- ☐ Tools for loading data into in-memory data objects from different file formats.
- ☐ Data alignment and integrated handling of missing data.
- ☐ Reshaping and pivoting of date sets.
- ☐ Label-based slicing, indexing and sub setting of large data sets.
- ☐ Columns from a data structure can be deleted or inserted.
- ☐ Group by data for aggregation and transformations.
- ☐ High performance merging and joining of data.
- ☐ Time Series functionality.

Areas where python does well:

- ☐ Easy handling of missing data (represented as NaN) in floating point as well as non- floating point data
- ☐ Size mutability: columns can be inserted and deleted from Data Frame and higher dimensional objects
- ☐ Automatic and explicit data alignment: objects can be explicitly aligned to a set of labels, or the user can simply ignore the labels and let Series, Data Frame, etc. automatically align the data for you in computations
- ☐ Powerful, flexible group by functionality to perform split-apply-combine operations on data sets, for both aggregating and transforming data
- ☐ Make it easy to convert ragged, differently-indexed data in other Python and NumPy data structures into Data Frame objects
- ☐ Intelligent label-based slicing, fancy indexing, and subletting of large data sets
- ☐ Intuitive merging and joining data sets
- ☐ Flexible reshaping and pivoting of data sets

2.1.4 NUMPY:
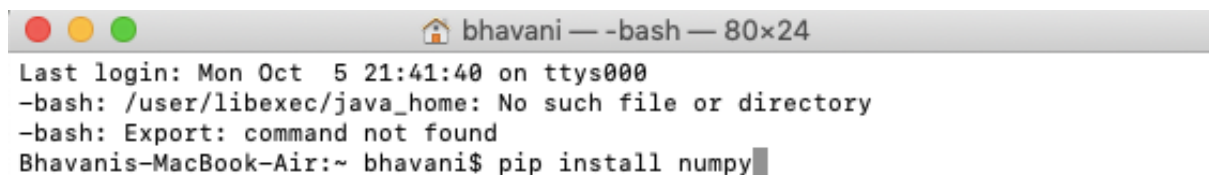
NumPy, which stands for Numerical Python, is a library consisting of multidimensional array objects and a collection of routines for processing those arrays. Using NumPy, mathematical and logical operations on arrays can be performed.

NumPy is a python library used for working with arrays. In Python we have lists that serve the purpose of arrays, but they are slow to process. NumPy aims to provide an array object that is up to 50x faster than traditional Python lists. The array object in NumPy is called ndarray, it provides a lot of supporting functions that make working with ndarray very easy. Arrays are very frequently used in data science, where speed and resources are very important.

NumPy arrays are stored at one continuous place in memory unlike lists, so processes can access and manipulate them very efficiently. This behaviour is called locality of reference in computer science. This is the main reason why NumPy is faster than lists. Also it is optimized to work with latest CPU architectures. NumPy is a Python library and is written partially in Python, but most of the parts that require fast computation are written in C or C++.

Installation of NumPy:

pip install numpy

Fig 7. Installation of NumPy

To check if Numpy is correctly installed, just run the following commands to perform a version check:

Python3
>>>import numpy
>>>print(numpy._version_)

Fig 8. Verifying NumPy Installation

12

Advantages of NumPy:

POWERFUL N-DIMENSIONAL ARRAYS:

Fast and versatile, the NumPy vectorization, indexing, and broadcasting concepts are the de- facto standards of array computing today.

NUMERICAL COMPUTING TOOLS

NumPy offers comprehensive mathematical functions, random number generators, linear algebra routines, Fourier transforms, and more.

INTEROPERABLE

NumPy supports a wide range of hardware and computing platforms, and plays well with distributed, GPU, and sparse array libraries.

PERFORMANT

The core of NumPy is well-optimized C code. Enjoy the flexibility of Python with the speed of compiled code.

EASY TO USE

NumPy's high level syntax makes it accessible and productive for programmers from any background or experience level.

It provides:

- ☐ 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

- ☐ and much more

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

2.2 EXISTING SYSTEM:

1. Existing system is based on the color joint probability function, in this it will look for the centroid of the colors and color edge co-occurrence histogram, so that the accuracy of this system will less.

2. It will give the RGB values as the output.

3. Centroid of colors and histogram plays an important role.

2.2 DISADVANTAGES OF EXISTING SYSTEM:

1. Based on the cursor movement the RGB values dynamically change where it creates a slight confusion to the user.

2. It is moderately sensitive to the background noise.

3. It is unclear to the user about the color name on the cursor movement.

2.3 PROPOSED SYSTEM:

1. Here we are going to use the RGB calculations for the color detection, and then calculate the distance from RGB values and then map the data set color names to detect the color name.

2. Here , this project gives us the name of the color along with the RGB values.

3. In comparison to the existing system it has an additional feature of giving the color name as the output.

4. This system clearly displays the color name along with the RGB values, so it will be a bit clear to the user to understand.

5. It is very less sensitive to the background noise.

6. Calculating the distance and mapping to the color plays an important role.

3.ANALYSIS

3.1 INTRODUCTION:

The Analysis is where we break down the deliverables in the high-level Project into the more detailed business requirements. The Analysis Phase is also the part of the project where we identify the overall direction that the project will take through the creation of the project strategy documents.

Gathering requirements is the main attraction of the Analysis Phase. The process of gathering requirements is usually more than simply getting to know what the users need and writing their answers down. Depending on the complexity of the application, the process for gathering requirements has a clearly defined process of its own. This process consists of a group of repeatable processes that utilize certain techniques to capture, document, communicate, and manage requirements. once we have requirements, we can start to set the overall direction for training in a Training Strategy document. The same strategy is implemented in our project such that firstly we collect the requirements in the form of the dataset. A lot of study is made on collecting the data set. We have gone through some websites like dataworld, AnalyticalVidhya, UCI machine learning Repository, Kaggle.com. Finally, we collected the data set from Kaggle.com. They have more than 350 datasets in total – with more than 200 as Featured datasets. While some of the initial datasets were usually present at other places, we have seen a few interesting datasets on the platform, not present at other places. Along with new datasets, another benefit of the interface is that we can see scripts and questions from community members on the same interface. We have a dataset with more than 2000+ records.

3.2 SOFTWARE REQUIREMENT SPECIFICATION:

Clear requirements will help the development teams create the right product. And a software requirements specification (SRS) helps us lay the groundwork for product development. A software requirements specification (SRS) is a document that describes what the software will do and how it will be expected to perform. It also describes the functionality the product needs to fulfil all stakeholders (business, users) needs. A Typical SRS includes the specific requirements. The best SRS document define how the software will interact when embedded in hardware — or when connected to other software. Good SRS documents also account for real-life users.

A software requirements specification (SRS) includes in-depth descriptions of the software that will be developed.

A system requirements specification (SyRS) collects information on the requirements for a system

"Software" and "system" are sometimes used interchangeably as SRS. But, a software requirement specification provides greater detail than a system requirements specification.

Why Use an SRS Document?

A software requirements specification is the basis for our entire project. It lays the framework that every team involved in development will follow.

It's used to provide critical information to multiple teams — development, quality assurance, operations, and maintenance. This keeps everyone on the same page.

Using the SRS helps to ensure requirements are fulfilled. And it can also help us to make decisions about your product's lifecycle — for instance, when to retire a feature.

Writing an SRS can also minimize overall development time and costs. Embedded development teams especially benefit from using an SRS.

3.2.1 USER REQUIREMENT SPECIFICATION:

1. As the project is to find out the color name of the specified color in an image, User task is to provide the image as an input.
2. In-order to provide the input as an image, The image format should be as .jpg, .jpeg, .png .
3. The User requires to install the python package Open-CV in his/her system which is the Open source computer vision to perform computer vision tasks.
4. The User requires to install the python package pandas in his/her system which is the fast, flexible and powerful tool to work with data projects. Where it is used to load the csv file.
5. The user requires to install the python package Numpy to work with the project
6. There should be a online programming environment.

3.2.2 SOFTWARE REQUIREMENT SPECIFICATION:

- Operating System :- Windows , Mac , Linux

- Coding Language   :-   Python

- Tools   :-   open-CV ,NumPy, Pandas

3.2.3 HARDWARE REQIUREMENT SPECIFICATION:

- System   :- core i3 or higher

- Hard Disk   :- 25GB (SSD Preferred)

- Ram   :- 2GB

- Internet connection :- 1Mbps or higher

3.3 CONTENT DIAGRAM OR ARCHITECTURAL DIAGRAM:

The Architecture diagram can help system designers and developers visualize the high-level, overall structure of their system or application to ensure the system meets their users' needs. You can also use architecture diagrams to describe patterns that are used throughout the design. It's somewhat like a blueprint that can be used as a guide for the convenience of discussing, improving, and following among your team.

WHY DO WE NEED ARCHITECTURAL DIAGRAMS:

The architectural diagrams in particular and the documentation, in general, will be primarily used for collaboration, communication, vision and guidance inside the team and across teams. It also include the significant design decisions in the project (taken at a certain moment of time), but nothing more.

Architectural diagrams will help everybody to see the big picture and to understand the surroundings. This will be the fundamental reason behind creating and maintaining the architectural diagrams.

Sometimes it's helpful to get a big-picture view of something for context — and an architectural diagram does just that. In a nutshell, it shows how elements within a system interact with each other in a wider process.

There are many different kinds of architectural diagrams, all of which vary depending on what it is you're trying to accomplish. They're used in construction, engineering, security, IT, sales — basically any process that involves stages and stakeholders.

DESIGN Vs ARCHITECTURAL DIAGRAMS:

There is a confusion between the two, but they're completely different things. An architecture diagram describes what we are building, how stakeholders interact with it, and where constraints lie. A design diagram explains how to build it.

Two ways architectural diagrams can help in :

1. They help with comprehension:

A picture is worth a thousand words, or so the saying goes. Similarly, architectural diagrams help convey complex information in a single image.

Architectural diagrams show systems. Displaying information visually allows the viewer to see everything at a glance, including how things interact. This is especially useful when making changes: we will be able to see the downstream effects of a given change more clearly.

Architectural diagrams also break down complex systems and processes into layers. So rather than trying to comprehend everything at once, we can zoom in and focus on smaller sub- processes or systems.

2. They improve communication and collaboration

One of the main issues software engineers face is consistency. When we are working on anything that involves multiple people, there's always a risk of miscommunication and discrepancies between project teams and developers. It is crucial to standardize information, which is where an architectural diagram comes in handy.
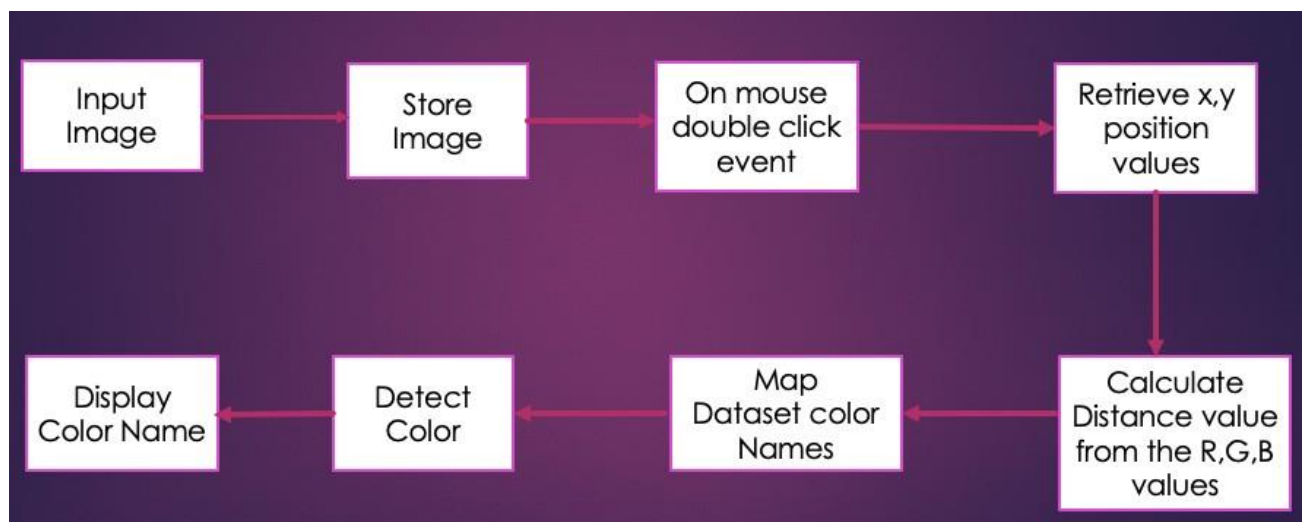
Fig 9. Architectural diagram

The above is the architectural diagram concerned with our project. Firstly input an image and then store that image, when the mouse double click event happens, it retrieves the X,Y values, then we calculate the distance from the R,G,B values where we map to the data set and then detect the color. Later it will display the color. It gives a clear way to understand how the project flows.

3.4 ALGORITHMS AND FLOWCHART:

FLOWCHART:

A flowchart is simply a graphical representation of steps. It shows steps in sequential order and is widely used in presenting the flow of algorithms, workflow or processes. Typically, a flowchart shows the steps as boxes of various kinds, and their order by connecting them with arrows.

A flowchart is a graphical representations of steps. It was originated from computer science as a tool for representing algorithms and programming logic but had extended to use in all other kinds of processes. Nowadays, flowcharts play an extremely important role in displaying information and assisting reasoning.

They help us visualize complex processes, or make explicit the structure of problems and tasks. A flowchart can also be used to define a process or project to be implemented.

Flowchart Symbols

Different flowchart shapes have different conventional meanings. The meanings of some of the more common shapes are as Decision, Flow, data, terminator, On page reference, off page reference and etc.

Terminator

The terminator symbol represents the starting or ending point of the system.

Process

A box indicates some particular operation.

Document

This represents a printout, such as a document or a report.

Decision

A diamond represents a decision or branching point. Lines coming out from the diamond indicates different possible situations, leading to different sub-processes.

Data

It represents information entering or leaving the system. An input might be an order from a customer. Output can be a product to be delivered.

On-Page Reference

This symbol would contain a letter inside. It indicates that the flow continues on a matching symbol containing the same letter somewhere else on the same page.

Off-Page Reference

This symbol would contain a letter inside. It indicates that the flow continues on a matching symbol containing the same letter somewhere else on a different page.

Delay or Bottleneck

Identifies a delay or a bottleneck.

Flow

Lines represent the flow of the sequence and direction of a process.
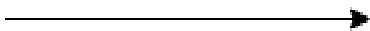
Comment/Annotation

Annotation Symbol indicates additional information regarding a step in a process.

Comments

Multi-document

Multiple documents

FLOWCHART

```
                    ┌─────────────────┐
                    │     START       │
                    └─────────────────┘
                             │
                             ▼
                    ╱─────────────────╲
                   ╱  User Input an    ╲
                  ╱      Image          ╲
                  ╲                     ╱
                   ╲───────────────────╱
                             │
                             ▼
                    ┌─────────────────┐
                    │   Stores an     │
                    │                 │
                    │    Image        │
                    └─────────────────┘
                             │
                             ▼
                    ┌─────────────────┐
                    │ On mouse double │
                    │ click event the │
                    │ flag is updated │
                    └─────────────────┘
                             │
                             ▼
                         ╱───────╲
                        ╱ If Flag ╲
                        ╲ =="TRUE" ╱──── YES ────┐
                         ╲───────╱               │
                             │                    ▼
                            NO            ╱─────────────────╲
                             │           ╱  Display the color╲
                             │           ╲    to the User    ╱
                             │            ╲─────────────────╱
                             │                    │
                             ▼                    │
                    ┌─────────────────┐           │
                    │     STOP        │◄──────────┘
                    └─────────────────┘
```
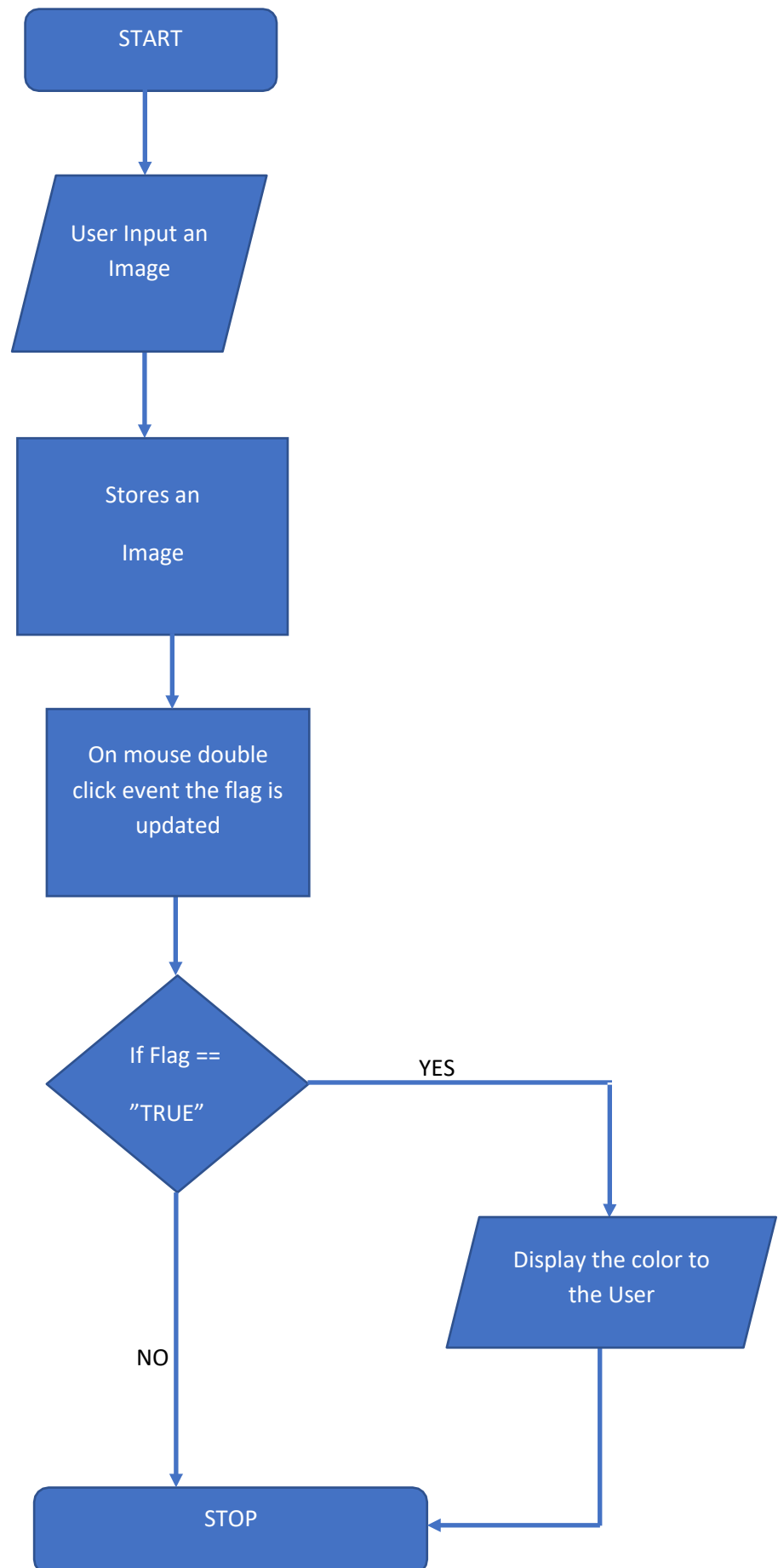
The above is the flowchart concerned to our project. Firstly, the flowchart starts with the terminator symbol i.e., start then the user is allowed to input an image in order to get the output. Then the system initially loads the image and stores it. If there is an action like mouse double click then the flag value in that respective function gets updated to "true"

If the flag value is true then the flow of control is YES then it move to Display the color name to the user. If the flow of control is false then it directly moves to the stop state.

Advantages of Flowcharts:

1. It helps to clarify complex processes.
2. It identifies steps that do not add value to the internal or external customer, including delays; needless storage and transportation; unnecessary work, duplication, and added expense; breakdowns in communication.
3. It identifies steps that do not add value to the internal or external customer, including delays; needless storage and transportation; unnecessary work, duplication, and added expense; breakdowns in communication.
4. It helps team members gain a shared understanding of the process and use this knowledge to collect data, identify problems, focus discussions, and identify resources.
5. It serves as a basis for designing new processes.
6. It allows the user easily understand what the project is all about.
7. One can easily understand the conditions and the output of that conditions base on the flow control.
8. The Way of expressing the whole project idea in a simple way is possible with the flowcharts.
9. Problem Solving: Flowcharts break a problem up into easily definable parts. The defined process displayed by the flowchart demonstrates the method of solving a complex problem. A flowchart reduces the chance that a necessary step for solving a problem will be left out because it appears obvious.
10. Effective Analysis: With the help of flowchart, problem can be analyzed in more effective way. It specifically shows what type of action each step in a process requires. Generally, a rectangle with rounded edges defines the beginning or end of the process, a diamond shape shows the point at which a decision is required, and a square block shows an action taken during the process.

4.DESIGN

4.1 INTRODUCTION:

During the Design Phase, the system is designed to satisfy the requirements identified in the previous phases. The requirements identified in the Requirements Analysis Phase are transformed into a System Design Document that accurately describes the design of the system and that can be used as an input to system development in the next phase.

The Objectives of the design phase Include:

• Transformation of all requirements into detailed specifications covering all aspects of the system

• Assessment and planning for security risks

• Approval to progress to the Development Phase

The purpose of the Design Phase is to transform the requirements into complete and detailed system design specifications. Once the design is approved, the Development Team begins the Development Phase.

Software design is a process to transform user requirements into some suitable form, which helps the programmer in software coding and implementation.

For assessing user requirements, an SRS (Software Requirement Specification) document is created whereas for coding and implementation, there is a need of more specific and detailed requirements in software terms. The output of this process can directly be used for implementation in programming languages.

Software design is the first step in SDLC (Software Design Life Cycle), which moves the concentration from problem domain to solution domain.

4.2 UML DIAGRAMS:

UML, which stands for Unified Modeling Language is a way to visually represent the architecture, design, and implementation of complex software systems. When we are writing code, there are thousands of lines in an application, and it's difficult to keep track of the relationships and hierarchies within a software system. UML diagrams divide that software system into components and subcomponents.

The Unified Modeling Language (UML) was created to forge a common, semantically and syntactically rich visual modeling language for the architecture, design, and implementation of complex software systems both structurally and behaviourally. UML has applications beyond software development, such as process flow in manufacturing.

It is analogous to the blueprints used in other fields, and consists of different types of diagrams. In the aggregate, UML diagrams describe the boundary, structure, and the behaviour of the system and the objects within it.

UML is not a programming language but there are tools that can be used to generate code in various languages using UML diagrams. UML has a direct relation with object-oriented analysis and design.

Object-oriented languages dominate the programming world because they model real-world objects. UML is a combination of several object-oriented notations: Object-Oriented Design, Object Modeling Technique, and Object-Oriented Software Engineering.

UML uses the strengths of these three approaches to present a more consistent methodology that's easier to use. UML represents best practices for building and documenting different aspects of software and business system modeling.

UML diagrams are classified into two types:

1.Structural Diagrams
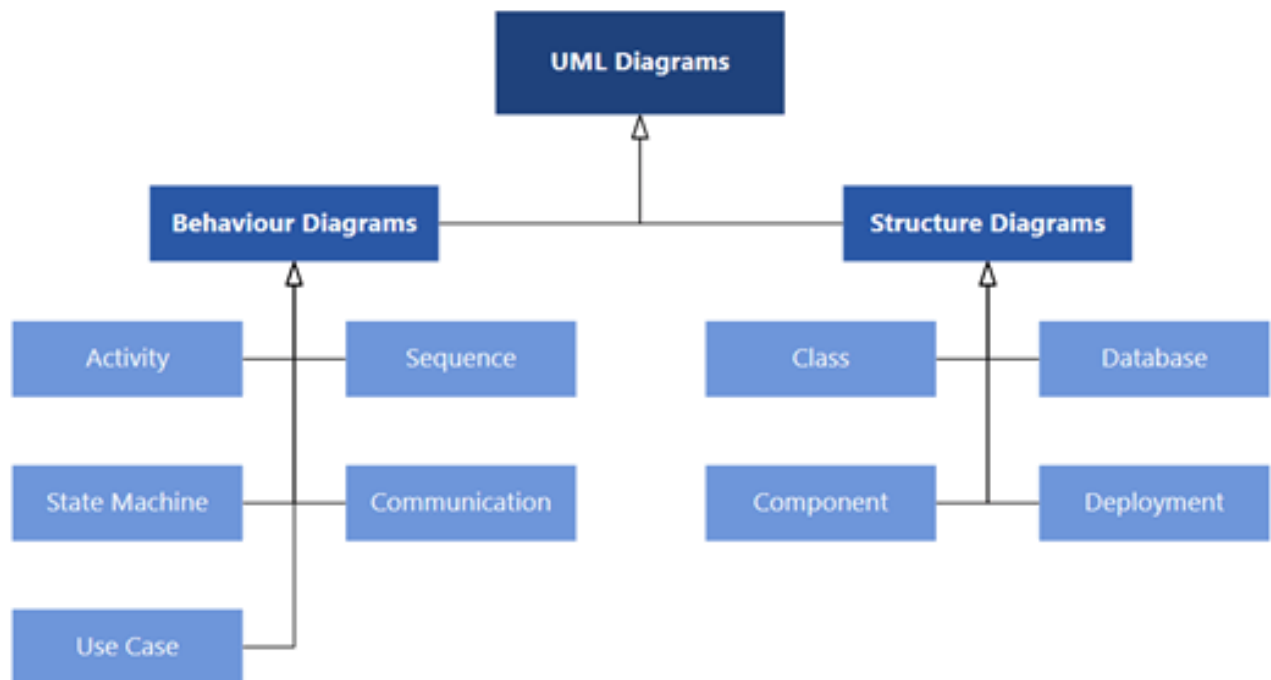
2.Behavioural Diagrams

Fig 10. Structure of UML Diagrams

In our project we implemented the Class Diagram from the structural diagrams and sequence diagram from behavioural diagrams.

Structural Diagrams:

Structure diagrams depict the static structure of the elements in your system. It shows the things in the system - classes, objects, packages or modules, physical nodes, components and interfaces. They also show the relationships between these things - classes that inherit from other classes, objects that own other objects, what classes belong to what packages, what nodes are connected to each other.

Behavioural Diagrams:

Behaviour diagrams depict the dynamic behaviour of the elements in your system. It shows how the system behaves and interacts with itself and other entities (users, other systems). They show how data moves through the system, how objects communicate with each other, how the passage of time affects the system, or what events cause the system to change internal states.

4.2.1 Class Diagram:

Class diagrams are the most common diagrams used in UML. Class diagram consists of classes, interfaces, associations, and collaboration. Class diagrams basically represent the object- oriented view of a system, which is static in nature.

Active class is used in a class diagram to represent the concurrency of the system.

Class diagram represents the object orientation of a system. Hence, it is generally used for development purpose. This is the most widely used diagram at the time of system construction.

Class diagrams are an aspect of UML that describe a static design of the objects, and their relationships to each other, in an application. During analysis, class diagrams may just be the names of objects and how they interact, but as the design develops the details of each class, including attributes and methods.

□ Shows the static structure of the model.

□ Collection of static modeling elements such as classes and their relationships connected as a graph.

□ Provides visual representation of objects, relationships and their structures.

27
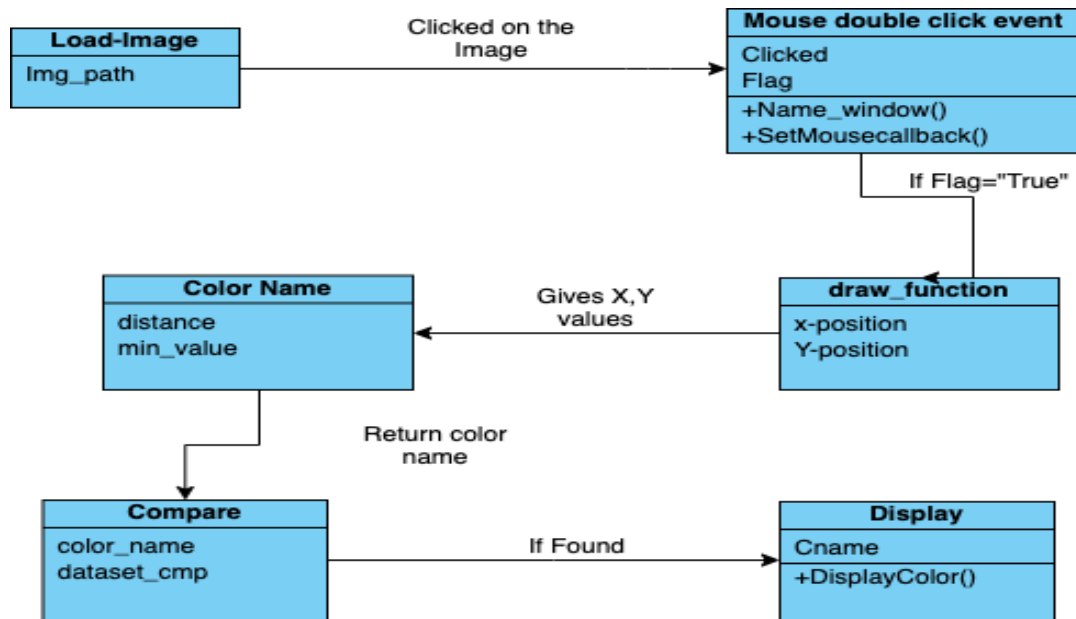
Fig 11. Class Diagram

The above diagram is the class diagram to the project. We have classes like load_image, mouse double click event, draw_func, colorname and etc. load_image has an attribute where the path of the image is stored then after viewing the image, if clicked on the image then their takes place a mouse double click event, initially the flag attribute is false, on clicking it becomes true and if it is true we move to the draw_func where the X and Y position values are retrieved and then those values are submitted to color name class to find out the color name if the value mapping to the data set is successful then the color name is displayed to the user.

The purpose of implementing Class Diagram to the project is:

1. Class diagrams give you a sense of orientation.

2. They provide detailed insight into the structure of your systems. At the same time they offer a quick overview of the synergy happening among the different system elements as well as their properties and relationships.

3. Class diagrams are at the heart of UML. They are based on the principles of object orientation and can be implemented in various phases of a project. Class diagrams are simple and fast to read. With the right software they are also easy to create. They are the foundation for creating systems.

4.2.2 Sequence Diagram:

A sequence diagram is an interaction diagram. From the name, it is clear that the diagram deals with some sequences, which are the sequence of messages flowing from one object to another.

Interaction among the components of a system is very important from implementation and execution perspective. Sequence diagram is used to visualize the sequence of calls in a system to perform a specific functionality.

Sequence diagrams, also known as event diagrams or event scenarios, illustrate how processes interact with each other by showing calls between different objects in a sequence. These diagrams have two dimensions: vertical and horizontal. The vertical lines show the sequence of messages and calls in chronological order, and the horizontal elements show object instances where the messages are relayed.

1. To create a sequence diagram, write the class instance name and class name in a rectangular box.
2. Draw lines between class instances to represent the sender and receiver of messages.
3. Use solid arrowheads to symbolize synchronous messages, open arrowheads for asynchronous messages, and dashed lines for reply messages.

Sequence Diagrams captures:

☐ the interaction that takes place in a collaboration that either realizes a use case or an operation
☐ high-level interactions between user of the system and the system, between the system and other systems, or between subsystems

Purpose of Sequence Diagram

☐ Model high-level interaction between active objects in a system

☐ Model the interaction between object instances within a collaboration that realizes a use case

☐ Model the interaction between objects within a collaboration that realizes an operation

☐ Either model generic interactions (showing all possible paths through the interaction) or specific instances of an interaction (showing just one path through the interaction)
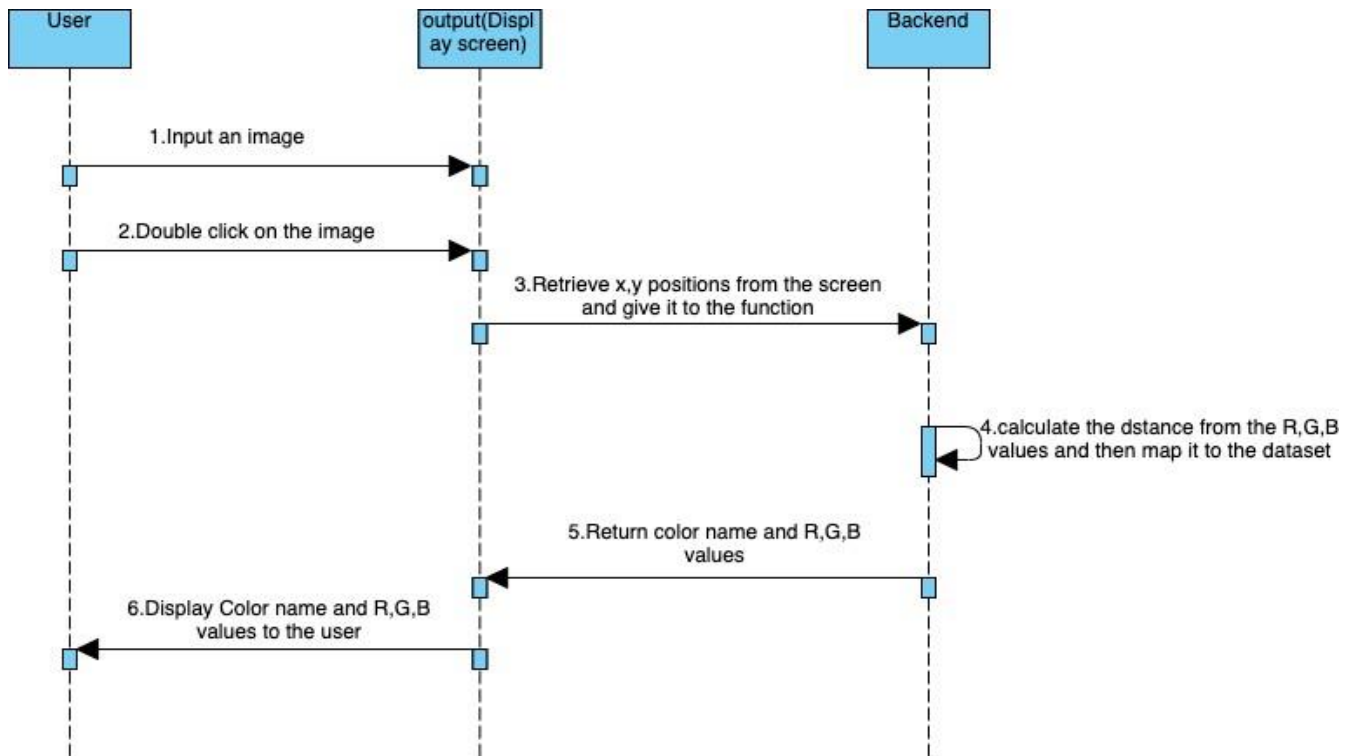
Fig 12. Sequence Diagram

The above diagram is the sequence diagram of the project. Firstly, the user display and backend act as the actors to this diagram. User inputs an Image then the user Double click on the image . After this the X and Y values are submitted to the backend ,there the calculation of the distance takes place, then it return color name and RGB values which it displays to the user.

The purpose of implementing Sequence Diagram to the project is:

- Sequence diagrams are easier to maintain.
- Sequence diagrams are easier to generate.
- Sequence diagrams can be easily updated according to the changes within a system.
- Sequence diagram allows reverse as well as forward engineering.

4.3 Module Design and Organization:

The following are the modules related to our project:

1.User Module:

This module is to perform the necessary actions on the image to find out the color name.

2.Image Module:

This module is for managing the image loading and storing of the image tasks. This module explains how the image is loaded in order to perform the task

3.Mouse Click event module:

This module is for managing the mouse double click event and perform necessary tasks based on the actions. This module orientates how the user module and image module are inter connected and shows how the user module action plays an important role in obtaining the output.

5. IMPLEMENTATION AND RESULTS

5.1 INTRODUCTION:

Implementation is the carrying out, execution, or practice of a plan, a method, or any design, idea, model, specification, standard or policy for doing something. As such, implementation is the action that must follow any preliminary thinking in order for something to actually happen.

In the project the color name, draw function, and mouse double click event plays an important role. Firstly, when the user clicks on the particular point the Mouse double click event occurs. There happens some statistics and finally the color name is displayed to the user.
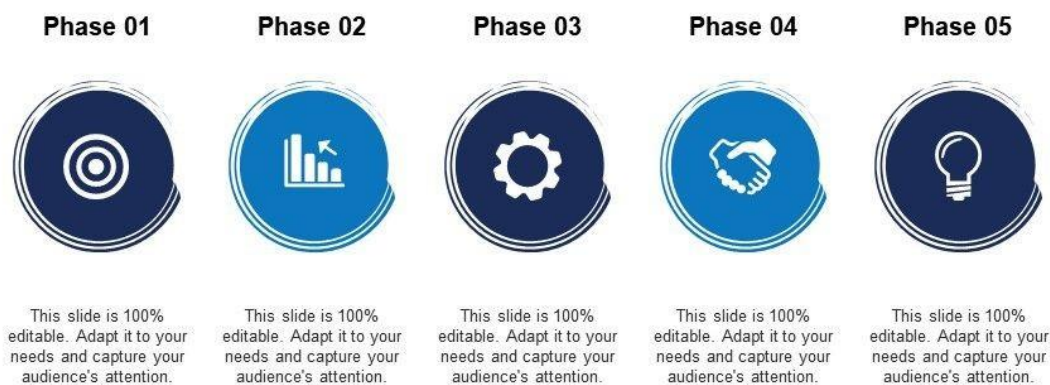

Fig 13. Implementation phases

The above picture describes the phases of the implementation process. The stages in this process are Marketing, Business, Management, Planning, Strategy. All the phases are editable and we can adapt it to our needs and can capture the audience attention.

5.2 Explanation of Key Functions:

1. User:

On running the program , the user is requested to given an image in order to find the color names of the respective image. Once the user gives the image through the command line then the image will be displayed to the user, here clicking on the image gives the RGB values and color name to the user. If there are any errors in the CSV file then that error message will be displayed in the terminal.

2. Input an image:

We load the image using argparse. Argparse is the "recommended command-line parsing module in the Python standard library." Giving the user the ability to enter command line arguments provides flexibility to the user.

3. Mapping the data set:

After the loading of the Image there takes a double click event where the x and y position values are retrieved from the draw function and submitted to the color name function to calculate the distance and later compare that value to the dataset to get the color name.

On required statistics and successful mapping a record is selected and the displayed to the user. The following is the data set:

Acid green	Acid green	#B0BF1A	176	191	26
Aero	Aero	#7CB9E8	124	185	232
Aero blue	Aero blue	#C9FFE5	201	255	229
African violet	African violet	#B284BE	178	132	190
Air Force blue	Air Force blue	#5d8aa8	93	138	168
Air Force blue (USAF)	Air Force blue (USAF)	#00308F	0	48	143
air_superiority_blue	Air Superiority Blue	#72a0c1	114	160	193
Alabama crimson	Alabama crimson	#AF002A	175	0	42

Fig 14. Mapping to a data set

5.3 METHOD OF IMPLEMENTATION:

1. Loading the image:

We load the image using argparse. Argparse is the "recommended command-line parsing module in the Python standard library." Giving the user the ability to enter command line arguments provides flexibility to the user. The argparse module makes it easy to write user- friendly command-line interfaces. The program defines what arguments it requires, and argparse will figure out how to parse those out of sys.argv. The argparse module also automatically generates help and usage messages and issues errors when users give the program invalid arguments. Firstly, we need to create an Argument Parser object which contains all the information necessary to parse the command line into Python data types. The next task is to parsing arguments and this can be done using parse_args() method. Argument Parser parses arguments through the parse_args() method. This will inspect the command line, convert each argument to the appropriate type and then invoke the appropriate action. Here in this project we give the image path through the command line.

```python
#Creating argument parser to take image path from command line
apr = argparse.ArgumentParser()
apr.add_argument('-i', '--image', required=True, help="Image Path")
args = vars(apr.parse_args())
img_path = args['image']
```

Fig 15. Source code for Loading Image

2. Retrieving X and Y values from the Image:

In the project when the mouse double click event occurs it is important for one to study how the function retrieves the pixels. And we made an interesting effort in doing that. Image basics : Images consist of a set of pixels. These pixels are the fundamental building blocks of an image. A Pixel can be thought of the "colour" or the "intensity" of light that appears at any given place in our image. Each pixels can be in two forms, The first one is Grayscale where each pixel has a value between 0 and 255 and This value represents each darker color shift from white through to black. And the second one is Color where Color is made up of three RGB components (red, green and blue).

We store three sets of values between 0 and 255. Each of these values is normally an 8-bit unsigned integer. These three 8bit values are combined into an RGB tuple in the form of (red, green, blue). The image coordinates start from the top left as (0,0). NumPy provides a technique called array slicing. Something very important to note about the structure of the RGB tuple is that OpenCV stores the RGB channels in reverse order. So the normal [RED, GREEN, BLUE] is actually [BLUE, GREEN, RED]. A study on the unified model diagrams is also executed.

```
def draw_func(event, x,y,flags,param):
    if event == cv2.EVENT_LBUTTONDBLCLK:
        global b,g,r,xposition,yposition, clicked
        clicked = True
        xposition = x
        yposition = y
        b,g,r = img[y,x]
        b = int(b)
        g = int(g)
        r = int(r)
```

Fig 16. Source code for Retrieving position values

It will calculate the RGB values of the pixel which we double click. The function parameters have the event name, (x,y) coordinates of the mouse position, etc. In the function, we check if the event is double-clicked then we calculate and set the r,g,b values along with x,y positions of the mouse.

3. Set a mouse callback event on a window:

First, we created a window in which the input image will display. Then, we set a call-back function which will be called when a mouse event happens.

With these lines, we named our window as 'image' and set a call-back function which will call the draw_function() whenever a mouse event occurs.

```
cv2.namedWindow('image')
cv2.setMouseCallback('image',draw_func)
```

Fig 17. Source code for mouse call back event
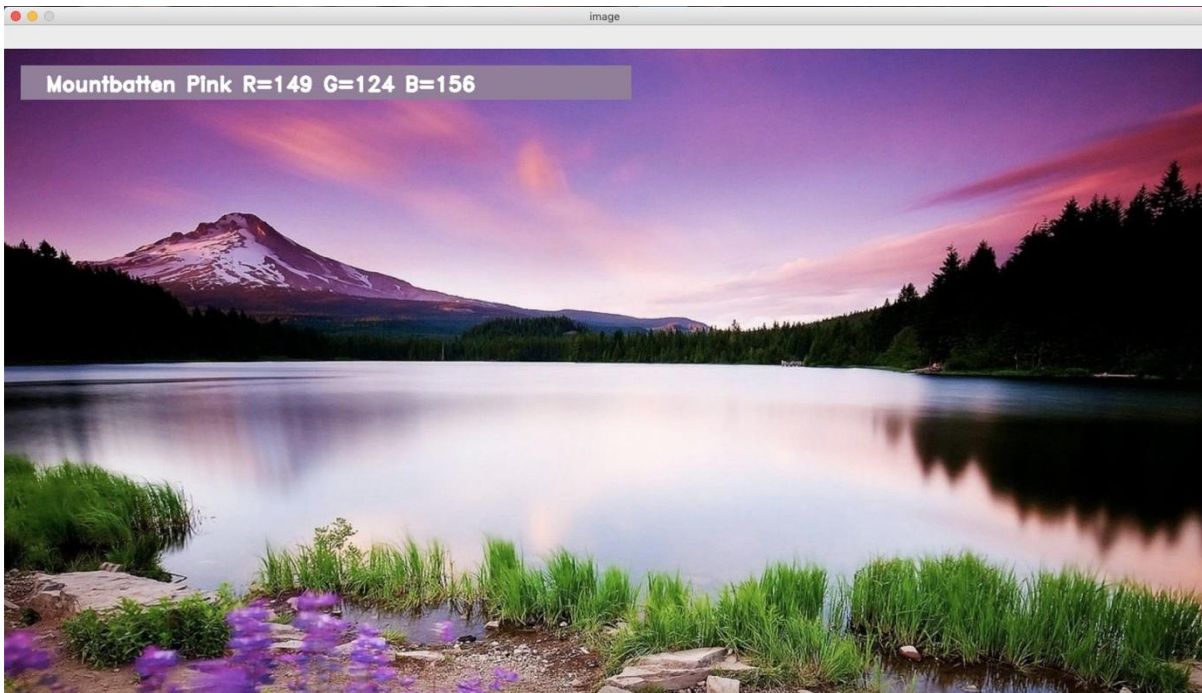
5.3.1 FORMS AND OUTPUT SCREENS: OUTPUT-1:

Fig 18. Output Screen – 1

OUTPUT-2:

Fig 19. Output Screen – 2

OUTPUT-3:

Fig 19. Output Screen - 3

5.4 RESULT ANALYSIS:

We made a detailed analysis of the project by using various factors like results and their expected outputs.

☐ COLLECT DATA SET: Data is collected from a variety of resources. We collected data set with more than 2000 records with required and necessary attributes like color name, Hex value, R, G, B values from Kaggle.com.

☐ CLEAN DATA SET: After collecting the data set the major task is to clean the data set. Cleaning the data set includes the removing the duplicate values, removing null values. The advantage of cleaning the data set is to filter unwanted outliers, handle missing data and remove irrelevant observations.

☐ VISUALIZATION OF DATA: After cleaning the data we need to visualize the data where it provides an accessible way to see and understand trends, outliers, and patterns in data. It helps in dealing the data very easily. It is a graphical representation of information and data. A good visualization helps in removing the noise from data and highlight the useful information.
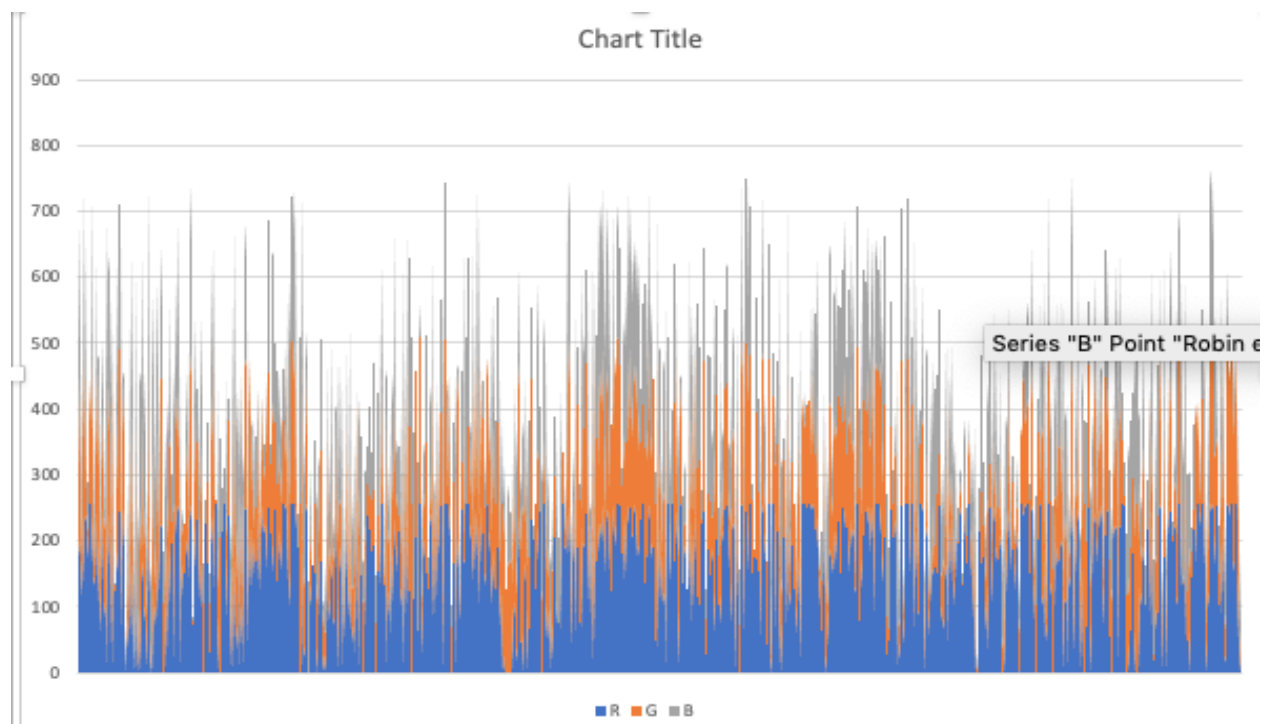

Fig 21. Visualization of Data

DATA VISUALIZATION:

Data visualization is the practice of translating information into a visual context, such as a map or graph, to make data easier for the human brain to understand and pull insights from. The main goal of data visualization is to make it easier to identify patterns, trends and outliers in large data sets. The term is often used interchangeably with others, including information graphics, information visualization and statistical graphics.

Data visualization is one of the steps of the data science process, which states that after data has been collected, processed and modelled, it must be visualized for conclusions to be made. Data visualization is also an element of the broader data presentation architecture (DPA) discipline, which aims to identify, locate, manipulate, format and deliver data in the most efficient way possible.

Importance of Data visualization:

Data visualization provides a quick and effective way to communicate information in a universal manner using visual information. The practice can also help businesses identify which factors affect customer behaviour; pinpoint areas that need to be improved or need more attention; make data more memorable for stakeholders. understand when and where to place specific products; and predict sales volumes.

Other benefits of data visualization include:

- the ability to absorb information quickly, improve insights and make faster decisions;

- an increased understanding of the next steps that must be taken to improve the organization;

- an improved ability to maintain the audience's interest with information they can understand.

- an easy distribution of information that increases the opportunity to share insights with everyone involved;

- eliminate the need for data scientists since data is more accessible and understandable; and

- an increased ability to act on findings quickly and, therefore, achieve success with greater speed and less mistakes

☐ ANALYZE DATA: Analysis of data is to convert the data into information useful for decision making by users. Data analyzed to answer the questions and test hypothesis.

☐ COLLECT THE X, Y COORDINATE VALUES OF THE IMAGE ON A PARTICULAR PIXEL POINT: Whenever the mouse double click event occurs it collects the X and Y coordinate values from the image. This plays a major role in finding out the color name at that particular point.

☐ MAKE REQUIRED STATISTICS AND DISPLAY THE COLOR NAME: After the X and Y coordinate values of a particular pixel point is obtained then required statistics are made to find out the color name. On successful attempt of doing this step leads to the display of the result to the user as the color name and the RGB values.

Expected	Actual
Collect Data Set	YES
Clean Data Set	YES
Visualize Data	YES
Analyze Data	YES
Collect the X,Y coordinate values of the Image on a particular pixel point	YES
Make Required Statistics and display the color name	YES

Fig 22. Result Analysis

6.TESTING AND VALIDATION

6.1 INTRODUCTION:

Software testing can be stated as the process of verifying and validating that a software or application is bug free, meets the technical requirements as guided by it's design and development and meets the user requirements effectively and efficiently with handling all the exceptional and boundary cases.

The process of software testing aims not only at finding faults in the existing software but also at finding measures to improve the software in terms of efficiency, accuracy and usability. It mainly aims at measuring specification, functionality and performance of a software program or application.

Software testing can be divided into two parts:

1. Verification: it refers to the set of tasks that ensure that software correctly implements a specific function.

2. Validation: it refers to a different set of tasks that ensure that the software that has been built is traceable to customer requirements.

Verification: "Are we building the product right?"

Validation: "Are we building the right product?"

6.2 DESIGN OF TESTCASES AND SCENARIOS:

Data Cleaning:

Data cleaning is the process of preparing data for analysis by removing or modifying data that is incorrect, incomplete, irrelevant, duplicated, or improperly formatted. This data is usually not necessary or helpful when it comes to analysing data because it may hinder the process or provide inaccurate results. There are several methods for cleaning data depending on how it is stored along with the answers being sought. Data cleaning is not simply about erasing information to make space for new data, but rather finding a way to maximize a data set's accuracy without necessarily deleting information.

For one, data cleaning includes more actions than removing data, such as fixing spelling and syntax errors, standardizing data sets, and correcting mistakes such as empty fields, missing codes, and identifying duplicate data points. Data cleaning is considered a foundational element of the data science basics, as it plays an important role in the analytical process and uncovering reliable answers. Most importantly, the goal of data cleaning is to create data sets that are standardized and uniform to allow business intelligence and data Analytics tools to easily access and find the right data for each query.

Test Case	Operations	Input	Expected Output	Actual Output	Result
1	Eliminate Null Values	Data set	Null Values Eliminated	Replaced with its actual value	PASS
2	Eliminate Duplicate values	Cleaned Data set	Duplicate Values are Eliminated	Dataset with Distinct values	PASS

Fig 23. Test cases of Data Cleaning

Data Analysing:

Data Analysis is a process of collecting, transforming, cleaning, and modelling data with the goal of discovering the required information. The results so obtained are communicated, suggesting conclusions, and supporting decision-making. Data visualization is at times used to portray the data for the ease of discovering the useful patterns in the data.

Data that is processed, organized and cleaned would be ready for the analysis. Various data analysis techniques are available to understand, interpret, and derive conclusions based on the requirements. Data Visualization may also be used to examine the data in graphical format, to obtain additional insight regarding the messages within the data.

Statistical Data Models such as Correlation, Regression Analysis can be used to identify the relations among the data variables. These models that are descriptive of the data are helpful in simplifying analysis and communicate results.

The process might require additional Data Cleaning or additional Data Collection, and hence these activities are iterative in nature.

Test Case	Operations	Input	Expected Output	Actual Output	Result
1	Analyze the Image	Image	Display the color name and RGB Values	Displays the color name and RGB Values	PASS

Fig 24. Test cases of Data Analysis

TEST CASES OF THE PROJECT:

Test Case ID	Description	Test Steps	Expected Output	Actual Output	Result
1	Reading the input from command prompt and display it using a frame	▪ Run the source code. ▪ Specify the input image path ▪ Resizing the Image	Input image should be displayed in a window Resized Image should be displayed	Input mage is displayed in a window Resized Image Should be displayed	PASS
2	Reading the CSV file and perform operations	Read the CSV file Using pandas	Dataset should be loaded	Dataset is loaded	PASS

Fig 25. Test Cases of the project - 1

Test Case ID	Description	Test Steps	Expected Output	Actual Output	Result
3	Mouse Event Happens	A call-back function is set which will be called when a mouse event happens.	It should call the call-back function when the mouse event occurs	It calls the call-back function when the mouse event occurs	PASS
4	When the mouse double click event occurs	Draws the text on the window using cv2.rectangle and cv2.putText () functions	When the user double clicks the window, it should draw a rectangle and get the color name to draw text on the window	Whenever a double click event occurs, it will update the color name and RGB values on the window.	PASS

Fig 26. Test Cases of the Project - 2

6.3 VALIDATION:

Validation is the documented process of demonstrating that a system or process meets a defined set of requirements. There are a common set of validation documents used to provide this evidence.

Design Validation is a process of evaluating the software product for the exact requirements of end-users or stakeholders. The purpose of design validation is to test the software product after development to ensure that it meets the requirements in terms of applications in the user's environment.
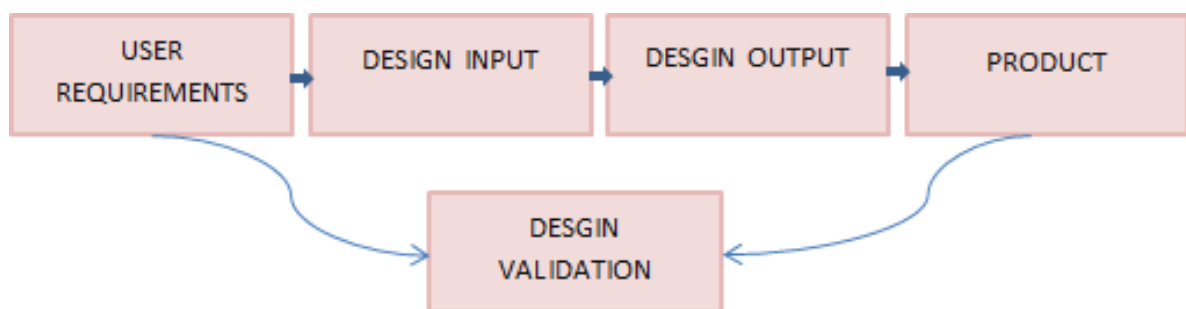

Fig 27. Validation

Validation is concerned with demonstrating the consistency and completeness of design with respect to the user needs. This is the stage where we actually build a version of the product and validate against the user requirements.

In our project the validation process gets clear understanding with the test cases given in the area of the test cases and scenarios. It is related to the complete design and the final product that met the requirements correctly. The requirements, then the input gives the final output which is same as the actual expected output then the final product is made.

7. CONCLUSION

7.1 Project Conclusion:

In this project, using open-cv we detect colors of images using RGB Model with minimum distance to detect red, blue, and green which will be helpful in various fields. In addition to this we also display the name of the selected color in an image. The results seem to be improved and it is tested on a set of varied colored images with various intensities. The Results are found to be efficient and proximity. This system is user-friendly and has a simple interface. It is used in a variety of the image editing apps, drawing apps, animations and in medical diagnosis. It gives us better accuracy in segmentation under various illumination. The best role of the project is lesser time-consuming process and which is very responsive to the users. From this project we conclude that color plays a vitally important role in the world we live. This tend to be most important part in object detection.

7.2 Future Enhancement:

This project has various applications. It can be used in skin color detection where it finds out the human faces and colors. One best example for this is sign language converter. In color name models for image editing uses the image selection prototype, where it displays the most probable color names for those selected pixels on the panel. The other major use of this occurs in endoscopy and laparoscopy.

Applications:

- Numerous Image Editing Apps

- Drawing Apps

- In the field of Image Processing

- RGB LED consistency control

- Animations

- Industrial process control and Medical diagnosis Control

8. REFERENCES

1. Open-cv: https://opencv.org/intel-acquires-itseez.html

2. Adrian Kaehler; Gary Bradski (14 December 2016). <u>Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library</u>.

3. <u>API Reference" in pandas documentation</u>

4. OpenCV- Image: https://handmap.github.io/exploring-opencv-with-python/ 5.argparseg -

<u>https://docs.python.org/3/library/argparse.html</u>