

## **IMAGE SHAPE PREDICTION AND DETECTION**

Addanki Abhaya Vikas<sup>a</sup>, Ammanamanchi Ravisankar Bharadwaj<sup>b</sup>, Jonnadula Narasimharao<sup>c</sup>

<sup>a</sup> *B.Tech Student, Department of Computer Science and Engineering, CMR Technical Campus, Medchal, Hyderabad, Telangana, India, vikasaddanki14962@gmail.com*

<sup>b</sup> *B.Tech Student, Department of Computer Science and Engineering, CMR Technical Campus, Medchal, Hyderabad, Telangana, India, arbharadwaj120@gmail.com*

<sup>c</sup> *Associate Professor, Department of Computer Science and Engineering, CMR Technical Campus, Medchal, Hyderabad Telangana, India, jonnadula.narasimharao@gmail.com*

### **ABSTRACT**

Human Vision is an incredible process of identifying different objects. We can differentiate the objects with the help of the vision Object Dimensions Measurement program used to measure the dimensions of objects. This helps in the automation, capturing, and identification of a particular instance of the object. These days, real-time object detection and dimensioning of objects are essential issues in many areas of industry. This is a necessary topic for computer vision problems. There is a need for an enhanced technique for detecting objects and computing their measurements in real time from pictures. For this, we will take the reference as the a4 sheet. After that, the dimensions of the object are measured. In real-time two different objects or states could help fill in and differentiate the different objects in actuality. We'll refer to this part of the architecture as the "backbone" network, which is usually pre-trained as an image classifier to learn how to extract features from an image more cheaply. This is a result of the fact that data for image classification is easier (and thus cheaper) to label as it only requires a single label as opposed to defining bounding box annotations for each image. Thus, we can train on an extensive labeled dataset (such as ImageNet) in order to learn good feature representations.

**Keywords:** Open cv, COCO dataset, dimensions, GTTS, python, Tkinter

### **Introduction:**

This paper is titled "Image Shape Prediction and detection". The traditional way of finding the object's dimension and finding its shape is a time taking process and two different processes, it is not accurate and the output will not be saved. This will be automated using python and its libraries like OpenCV and Numpy and return the output to the user in the audio or textformat.

This paper has been developed to identify the image and its dimension in the image and also helps to return the desired output that the user requested like audio and text. Previously the process of predicting the image and its dimension was a time taking process and not accurate and it doesn't get saved. This paper

helps to develop a solution in an accurate manner and also gives the output in the desired format. Our paper's main feature is to predict an object's shape and detect its dimension. Our paper gives accurate results of object dimension detection and also will predict objects, These two tasks are done in a single process and save the output and also give the output in audio or text format according to the user requirement Image data can generally be described through two dimensions (rows and columns), with a possible additional shaping. However, sometimes further dimensions are required for more accurate and detailed image analysis in specific applications and domains. For example, you may want to study a three-dimensional (3D) volume, measuring the distance between two parts or modeling how that 3D volume changes over time (the fourth dimension). In these instances, you need more than two dimensions to make sense of what you are seeing. This also includes multiple identifications of the images in the taken surface of the area. By this, we can differentiate objects into primary and secondary objects which give importance to the main object for the primary analysis of the data taken from the COCO dataset, but all the images in the input are to be included in the taken data set for the purpose of identification. The output is also formatted into different types for the sake of different users and their point of the outcome at the time of usage, mainly the output is formatted into text, audio, and more for the point of usage and the end user can have his outcome of data in well-designed page to avoid confusion among the data and to store the particular data. Multidimensional image processing gives you the flexibility to perform functions for traditional two-dimensional filtering in scientific applications.

### **LITERATURE SURVEY**

Dr. Piyush Pratap Singh et al. have researched OpenCV which provides a common infrastructure for computer vision applications and accelerates the use of machine perception in commercial products. It is a Python library that allows you to perform image processing and computer vision tasks. It provides a wide range of features, including object detection, face recognition, and tracking. It helps us in comprising integers that represent the pixels and intensity- hence, by indexing and slicing portions of the NumPy array, we are essentially isolating specific pixels thereby isolating specific portions of the image itself, thus allowing us to effectively

Charles R. Harris et al. have researched on NumPY and Array programming provides a powerful, compact, expressive syntax for accessing, manipulating, and operating on data in vectors, matrices, and higher-dimensional arrays. NumPy is the primary array programming library for the Python language. It plays an essential role in research analysis pipelines in fields as diverse as physics, chemistry, astronomy, geoscience, biology, psychology, material science, engineering, finance, and economics. Here we show how a few fundamental array concepts lead to a simple and powerful programming paradigm for organizing and analyzing scientific data. NumPy is the foundation upon which the entire scientific Python universe is constructed. It is so pervasive that several papers, targeting audiences with specialized needs, have

developed their own NumPy-like interfaces and array objects. Because of its central position in the ecosystem, NumPy increasingly plays the role of an interoperability layer between these new array computation libraries.

Rohith Dwived et.al researched Pycharm an open source and very helpful for accessing all the libraries is a very popular python IDE which is a cross-platform IDE that JetBrains developed to use for python development. These days I believe that python is the best language for building software applications by writing clean and readable code. Python is the favorite language for many, especially people working in Data Science and Machine Learning. The benefit of IDE is to permit a user to work with a different database without integration with other tools. Pycharm also allows you to create HTML, CSS, and Javascript files. Also, users can personalize their interface of the IDE.

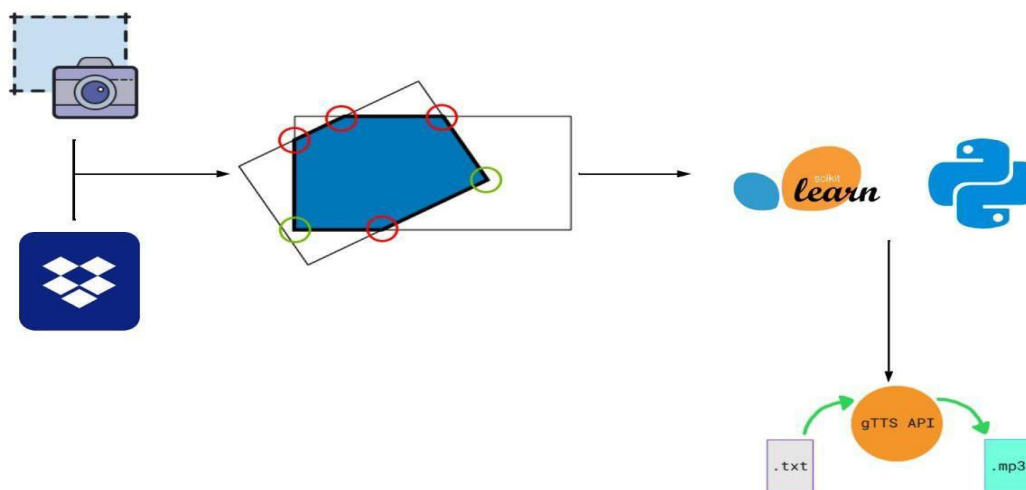
Xingkui Zhu et.al has researched the YOLOv5 which had been used by us for the compound-scaled object detection models trained on the COCO dataset and includes simple functionality for Test Time Augmentation(TTA),, and export to ONNX, and TFLite. With the help of the paper, we detect the unknown object from the data set. Object detection on scenarios is a recent popular task. As they always navigate in different altitudes, the object scale varies violently, which burdens the optimization of networks. Moreover, high-speed and low-altitude flight bring in the motion blur on the densely packed objects, which leads to a great challenge of object distinction. To solve the two issues mentioned above, he proposes TPH-YOLOv5. Based on YOLOv5, they add one more prediction head to detect different-scale objects to be in a particular manner. Then we replace the original prediction heads with Transformer Prediction Heads (TPH) to explore the prediction potential with the self-attention mechanism. We also integrate the convolutional block attention model (CBAM) to find attention regions in scenarios with dense objects. To achieve more improvement of our proposed TPH-YOLOv5, we provide bags of useful strategies such as data augmentation, multiscale testing, multi-model integration, and utilizing extra classifier

A. Karmel et.al researched the GTTs used for the conversion of the final output in our paper from text to images or audio or vice-versa. He addresses the problems faced by differently-abled people such as visually, audibly, and vocally challenged, through a single device is a tough job. Much research has been done on each problem and solutions have been proposed separately. All of them are not proposed together in a singular way. The aim of the paper is to create a single dev

### **PROPOSED METHODOLOGY**

We are proposing object identification software from an image. In an image, there are many objects. This software measures the different areas and dimensions of different objects in the image, It finds the points at the edge and gives its shape as an output. It also gives the details with reference to the given data source.

After taking the input, the image gets blurred, and later it gets converted into the BGR colors to grey so all the objects in the image are converted into black and white. After converting the real task is to convert the border of an object by giving values using a track bar that is adjustable to get or to convert the image into a completely black and white thus it can only visible the edges and borders or any directions, then later it gets to filters and finally, we get to see the modified image, form than counters get calculated that will be used to find the area and the length and width and all the other functions including the identification take place, form then it will give us the output in the desired format.



## REQUIRED COMPONENTS

The following are the required components:

### 1.1 Python(>3.2)

Python is required for running the python script in the system it helps to compile and do the specific task in the specific pattern

### 1.2 Object Detection

These packages are used to detect the objects in an image used in python

### 1.3 Opencv-python

This package is used to find the objects in images that are having different areas, It uses the counters and borderlines to specify the objects and helps to show the location and detection of an object

### 1.4 gTTS

This package is used to convert the text format to the audition format as an output.

### **1.5 OS**

The os module in this python model uses the system operating system-dependent functionality such as opening an audio file, running a program without any user interaction, and many more system-dependent operations.

### **1.6 COCO dataset**

This is data collection used to detect objects on a large scale which helps in our identification process. It consists of several features like Segregation, Recognition of the objects in a huge set of data.

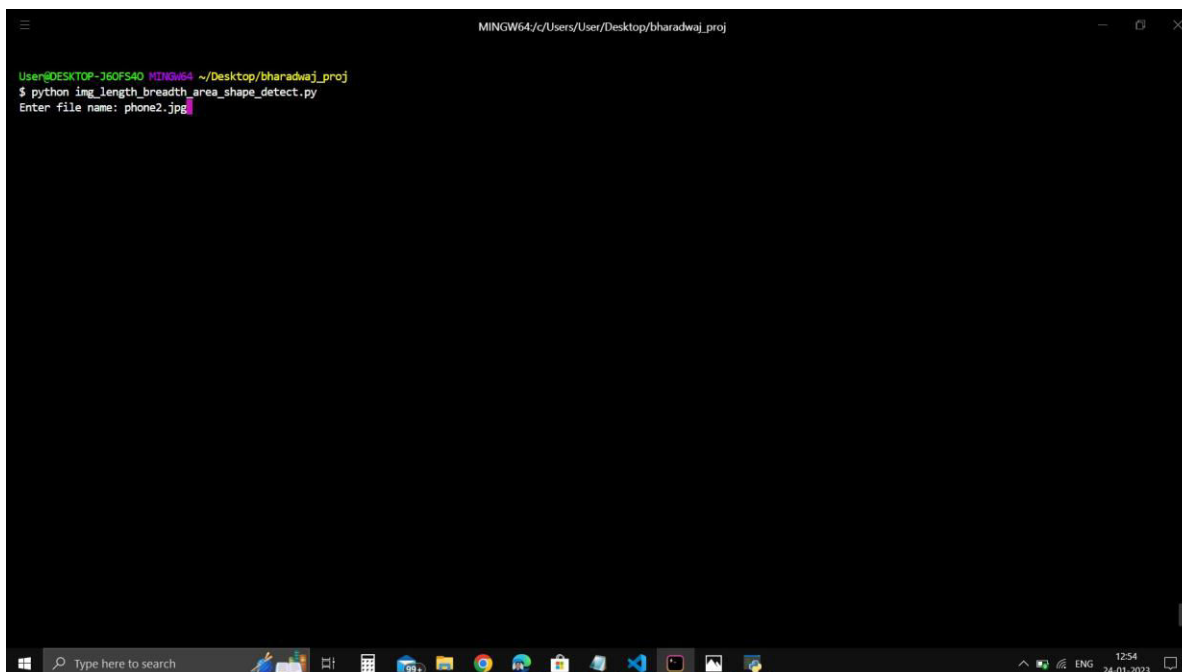
## **RESULTS**

The images of the developed model are shown below:

### **SCREENSHOTS**

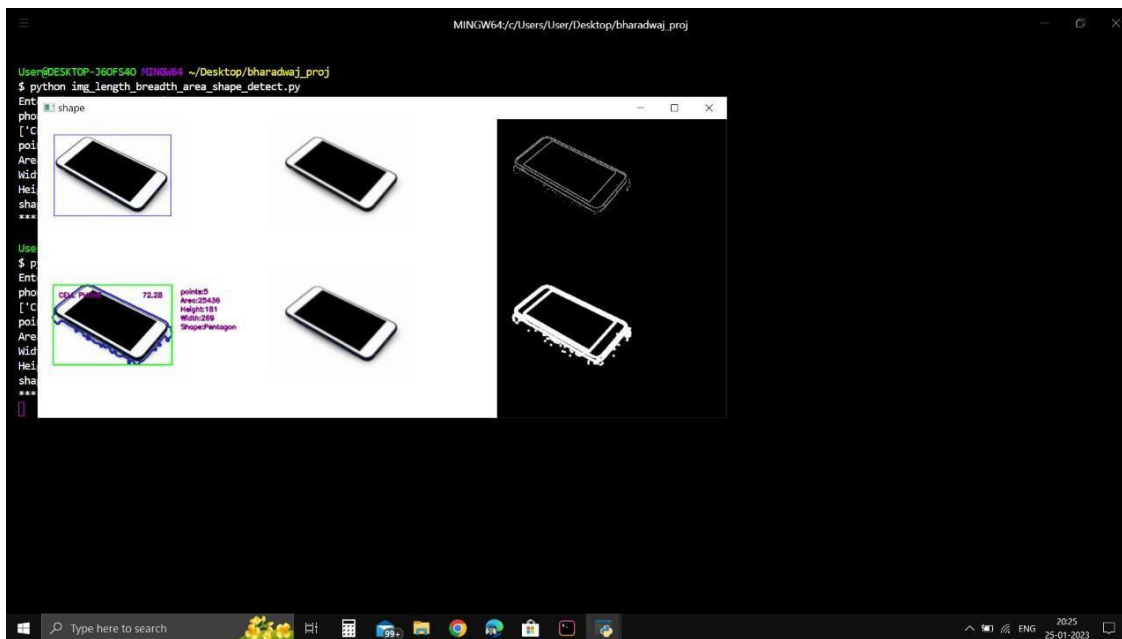
#### **1. IMAGE SELECTION:**

This image shows the way image uploading and selection by the user for Finding its dimension, and shape, by selecting the image from the file saved in the user's machine or dropbox. So the system can identify its information.



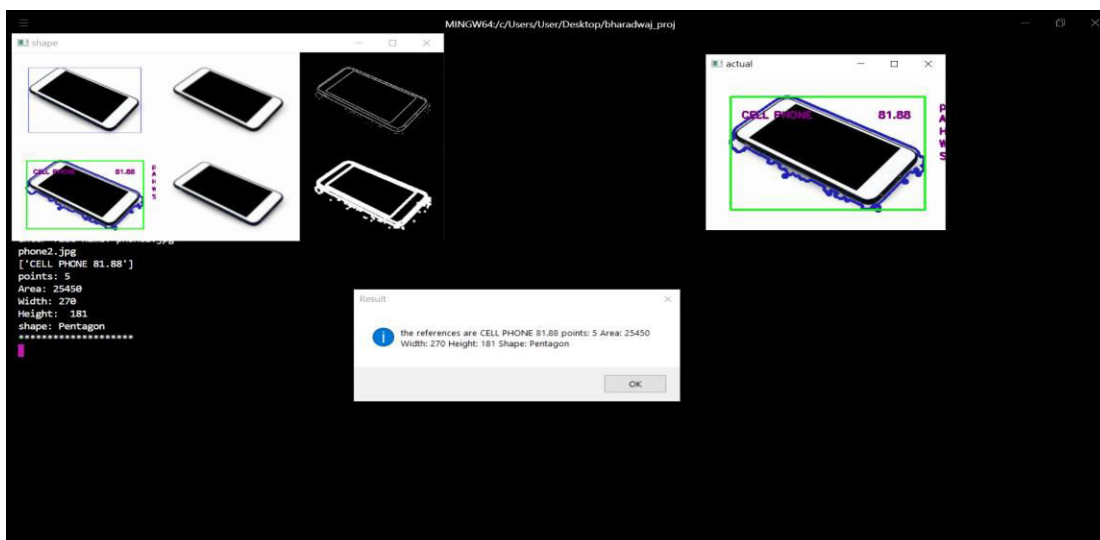
### IMAGE DIFFERENTIATION:

This image identifies the surfaces present in the image uploaded by the user and reads the image for dimension detection and identifying the corners and edges of the image which help to find the shape, and area of the image taken by the user and removes the unwanted surfaces in the uploaded image.



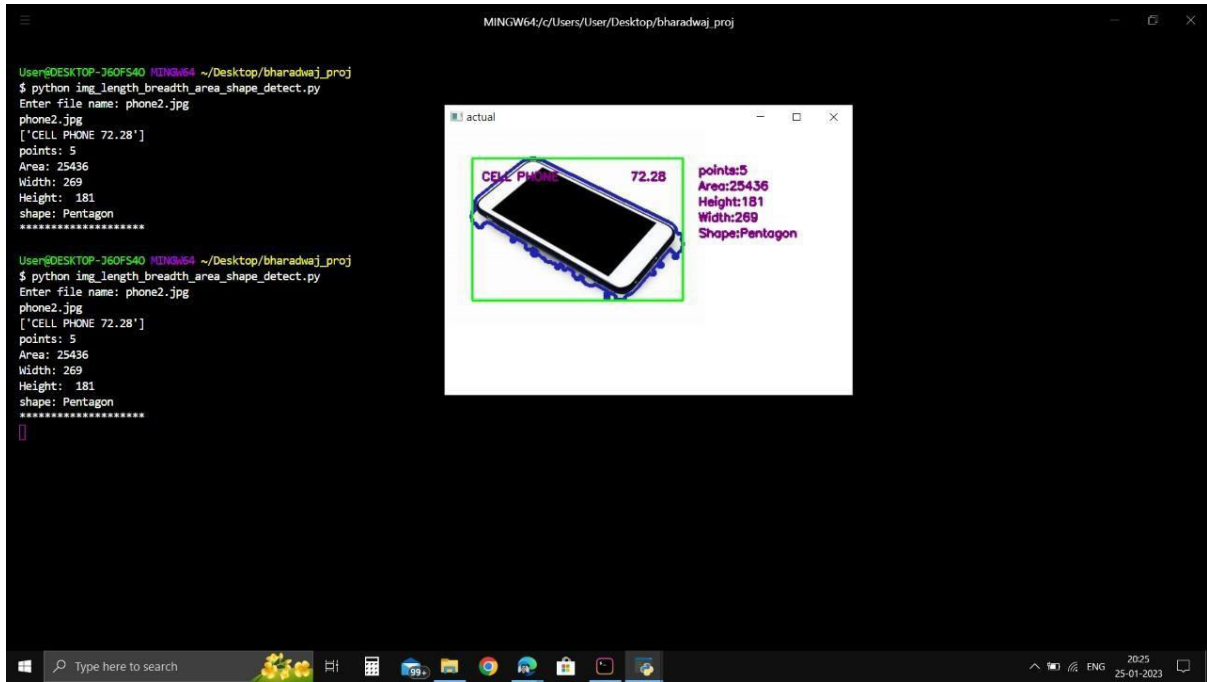
### IMAGE ANALYSIS

This image detects the corners of the taken object and makes its outline compute its dimension and predicts the shape of the object which we compare in our data set to identify the accurate object by analyzing both the data set and image processing



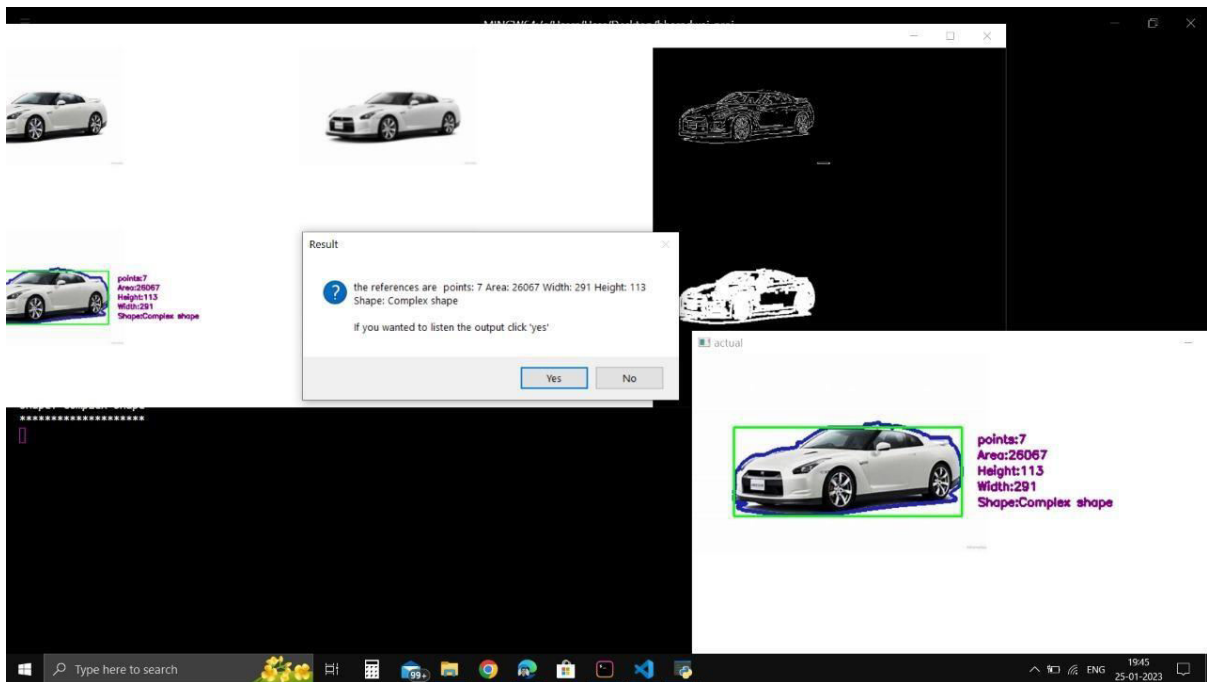
**OUTPUT:**

This screenshot shows the dimensions, shape, area, and corners present in the taken object which gives complete information about the image uploaded by the user.



**OUTPUT FORMAT:**

This paper gives the scope of formatting the output which helps the user to identify the object and its characteristics in the audio and text format based on the user's desire.



## **Conclusion**

This paper focuses on the size detection of an unspecified object, predicts its dimensions, and evaluates its schema accurately. This also helps in the shape identification of unknown objects which might be hazardous or virulent objects. The accurate size evolution helps in finding the perfect shape and helps in its algebra which helps in predicting the object. The COCO data set taken in this paper helps to analyze the data from a large collection of data to improve the accuracy of the output. gTTS gives us the audio output so disabled people can predict the object's shape prediction and detect that image successfully.

## **Future enhancements:**

This paper can be used to detect an object and its dimension, so we can develop a device that can continuously take the reference from the real world and analyze it, which can generate the output in a format that consumes less time.

## **REFERENCES:**

- [1] Haiyang Zhang, Cheng Lia .., ‘ Uncovering the optimal structural characteristics offlocs for microalgae flotation using Python-OpenCV’ journal of cleaner production 2022
- [2] DonghaiZhai, XiaoboZhang, XunLi, XichenXing, YuxinZhou, ChangyouMa ‘Object detection methods on compressed domain videos: An overview, comparative analysis, and new directions’ Measurement 2022
- [3] LukasMalburg, Manfred-PeterRieder, Ronny Seiger, Patrick KleinRalphBergmann ‘Object Detection for smart factory process by machine learning’ Procedia Computer Science 2021
- [4] Li Liu, Wanli Ouyang, Xiaogang Wang, Paul Fieguth, Jie Chen, Xinwang Liu & Matti Pietikäinen ‘Deep Learning for Generic Object Detection: A Survey Springer 2019.
- [5] Qiang Fu ‘Application and Analysis of RGB-D Salient Object Detection in Photographic Camera Vision Processing’ Hindawi 2022
- [6] Gopi Krishna Erabati ‘Object Detection in Traffic Scenarios - A Comparison of Traditional and Deep Learning Approaches’ ResearchGates 2020
- [7] Akhil Kumar, Akashdeep Sharma & Arvind Kalia ‘A Review of Research of Object Detection Area: Current and Future Trends’ Springer 2019
- [8] Felzenszwalb, P., Girshick, R., McAllester, D., Ramanan, D.: ‘Object detection with discriminatively trained part-based models.’ IEEE Transact. Pattern Anal. Mach. Intell. 32, 1627–1645 (2010)
- [9] Ivan Culjak; David Abram; Tomislav Pribanic; Hrvoje Dzapov; Mario Cifrek ‘A brief introduction to OpenCV’ IEEE 2012
- [10] A. Ambikapathy, Jyotiraditya Sandilya, Ankit Tiwari, Gajendra Singh & Lochan



Varshney ‘Analysis of Object Following Robot Module Using Android, Arduino, and Open CV, Raspberry Pi with OpenCV and Color Based Vision Recognition’ springer 2021

[11] Aakanksha Mishra, Kartikey Bhardwaj, Vibhor Sharma ‘Smart Attendance System Using OpenCV’ ResearchGates 2023

[12] AASaber, E. and Tekalp, A., ‘Face Detection and Facial Feature Extraction Using Color, Shape, and Symmetry Based Cost Functions, International Conference on Pattern Recognition.’ Department of Electrical Engineering and Center for Electronic Imaging Systems University of Rochester, Rochester, NY 14627-0126 ,1996

[13] Yuehang Song, Borun Chen, Xiaobin Liu, Hu Weijun ‘Audio and video editing system design based on OpenCV’ ResearchGates 2022

[14] Sarita Gautam & Anuj Kumar ‘Image-based automatic traffic lights detection system for autonomous cars’ Springer 2023

[15] Subodh Kumar, Shristi Kiran, Nitin Mishra ‘Face mask detection using OpenCV’ ResearchGates 2022.