**OBJECT DETECTION**

A

Project Report

For The Award Of the

Degree

**BASIC AND ADVANCED PYTHON**

***Under Guidance Of***

**Mr. HARI SHANKAR**

**Project Carried Out At**



**ASP InfoTech Pvt. Ltd.**

**Submitted By**

**ABHINAB MONDAL**-ASPR 240087

**SNEHA BHAGAT**-ASPR

**MEHULI LAHIRI**-ASPR240057

**DECLARATION**

This is to certify that the dissertation/project proposal entitled **“Object Detection”** is done by **Abhinab Mondal, Sneha Bhagat, Mehuli Lahiri** an authentic work carried out for the partial fulfilment of the requirements for the award of the degree of **Python Programming Language** under the guidance of **Hari Shankar** Sir and **Anand Kumar Prajapati** Sir.

Name of the Students

ABHINAB MONDAL

SNEHA BHAGAT

MEHULI LAHIRI

Date: …………………….

------------------------------------------ -----------------------------------------

|  |  |
| --- | --- |
| Signature of the Guide | Signature of the Head |
| Date: | Date: |

**ACKNOWLEDGEMENT**

Success of any project depends largely on the encouragement and guidelines of many others. We take this sincere opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project work. We extend our heartfelt thanks to everyone who has supported us throughout this minor group project. We are immensely grateful to our professor for providing us with the guidance and support we needed to complete this challenging task. We would also like to acknowledge the contributions made by our fellow students, who provided valuable insights and feedback that helped shape and refine our ideas. Additionally, we would like to thank the university library staff for providing us access to a wide range of resources we used in researching various aspects of this project. Without their help, it would have been nearly impossible for us to create such a comprehensive work. Finally, we want to express gratitude towards all those who have encouraged and motivated us along the way. Your unwavering support was instrumental in helping us set ambitious goals and achieve them successfully.

**TABLE OF CONTENT**

1. Introduction
2. Objective
3. Libraries
4. Snapshot
5. Conclusion
6. Future Enhancement
7. References

**INTRODUCTION**

Object detection is a computer vision technique that identifies and classifies a particular object in a particular setting. The main goal of object detection is to scan digital images or real-life scenarios to locate instances of every object, separate them, and analyse their necessary features for real-time predictions. It forms a ground for other important AI vision techniques like image classification, image retrieval, or object co-segmentation that drives meaningful information out of real-life objects. It can help automatically detect cattle movements, traffic signals, and road lanes for self-driving vehicles to reach their destinations. It works in a similar way as object recognition. The only difference is that object recognition is the process of identifying the correct object category whereas object detection simply detects the object’s presence and location in an image.

Object detection tasks can be performed by two different methods of data analysis techniques:

* **Image processing:** Itis a part of unsupervised learning that doesn’t require historical training data to teach analytics models. The models self-train themselves on the input images and create feature maps to make predictions. Image processing does not require high graphical processing power (GPU) or large datasets for execution.
* **Deep neural network:**A deep neural network is generally a supervised learning algorithm that requires large datasets and high GPU computation power to predict object classes. It’s a more accurate way to classify objects which are partially hidden, complex, or placed in unknown backgrounds in an image.

**OBJECTIVE**

The main objective behind object detection project is:

* **Detect and Localize Objects:** Design and implement a deep learning model to accurately detect and localize objects within images or video frames.
* **Achieve High Accuracy and Precision:** Train the model to achieve a high level of accuracy, precision, and recall in object detection, ensuring reliable identification and localization of objects.
* **Handle Multiple Object Classes**: Extend the model to handle detection and classification of multiple object classes, providing a comprehensive solution for diverse scenarios.
* **Real-Time Object Detection:** Optimize the model for real-time performance, enabling it to process images or video streams efficiently without sacrificing accuracy.

**LIBRARIES**

* **PyTorch:**

PyTorch is an open-source machine learning library based on the Torch library, developed by Facebook’s AI Research lab. It is widely used in deep learning, natural language processing, and computer vision applications. It also provides a dynamic computational graph, which allows for easy and efficient modelling of complex neural networks. It resides inside the torch module. Here the data that has to be processed is input in the form of a tensor. It provides a wide range of pre-built neural network architectures, such as fully connected networks, convolutional neural networks (CNNs), and recurrent neural networks (RNNs). In training the model, it provides a convenient data loading and processing library, torch. utils.data that can be used to load and prepare the data.

* **Matplotlib:**

It is a collection of functions that make matplotlib work like MATLAB. Each pyplot function makes some change to a figure: e.g., creates a figure, creates a plotting area in a figure, plots some lines in a plotting area, decorates the plot with labels, etc. It makes easy things easy and hard things possible. It also used in creating publication quality plots, make interactive figures that can zoom, pan, update, customize visual style and layout. Here various states are preserved across function calls, so that it keeps track of things like the current figure and plotting area, and the plotting functions are directed to the current axes.

In matplotlib generating visualizations is also very quick:

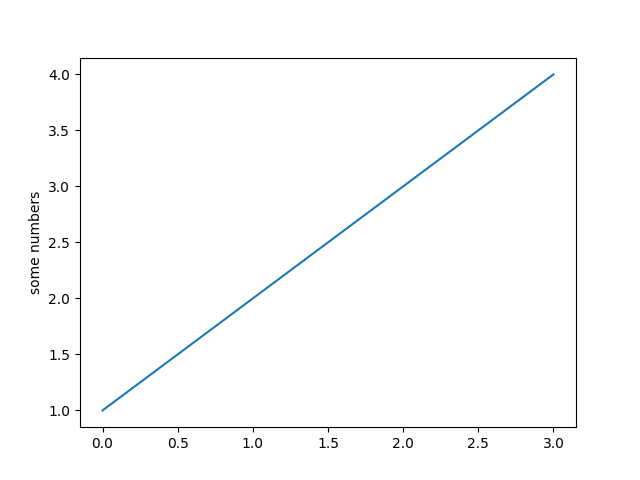
For example,

import matplotlib. pyplot as plt

plt.plot([1, 2, 3, 4])

plt.ylabel('some numbers')

plt. show ()



Here you may be wondering why the x-axis ranges from 0-3 and the y-axis from 1-4. If you provide a single list or array to plot, matplotlib assumes it is a sequence of y values, and automatically generates the x values for you. Since python ranges start with 0, the default x vector has the same length as y but starts with 0; therefore, the x data are [0, 1, 2, 3].

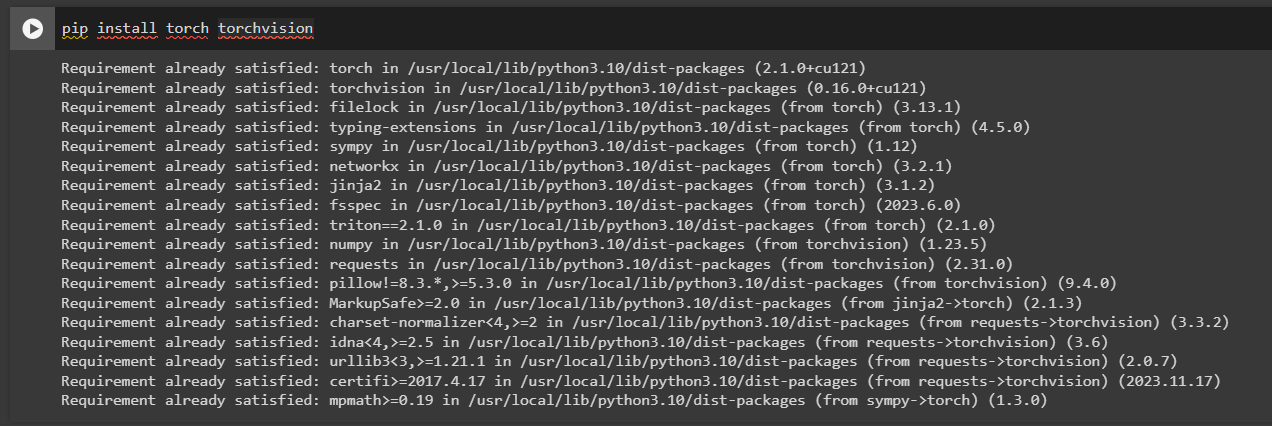
* **Torchvision**

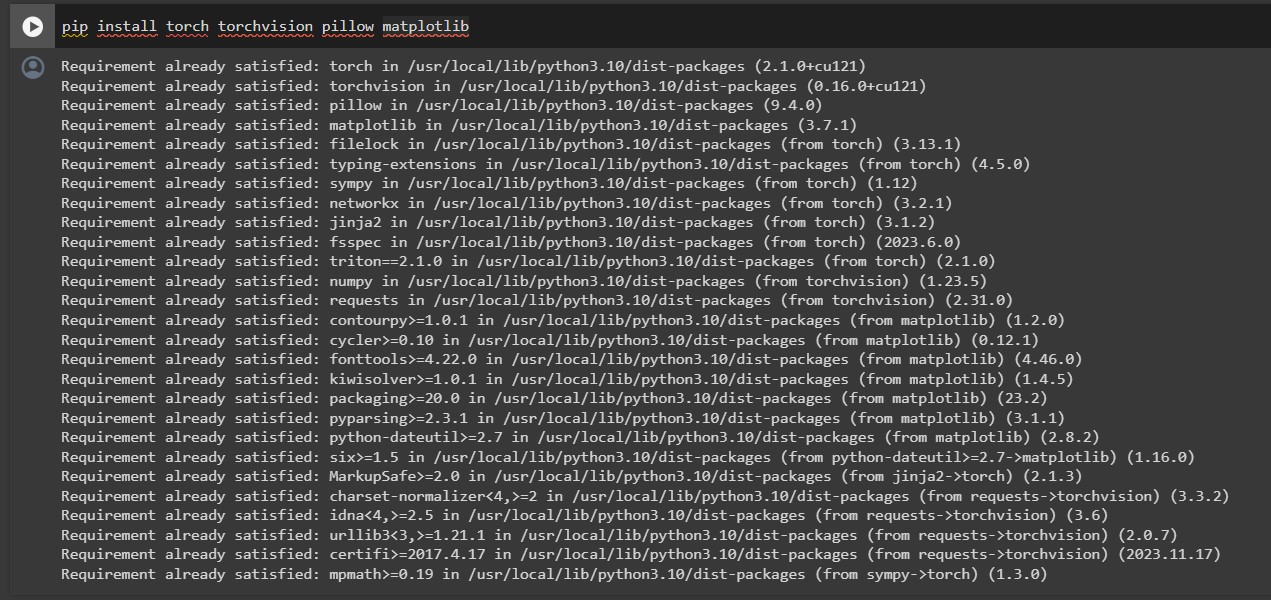
Torchvision is a popular computer vision library in the PyTorch ecosystem. It provides a collection of tools and datasets for working with image and video data. The library is designed to work seamlessly with PyTorch, making it easy for researchers and developers to build and experiment with computer vision models.

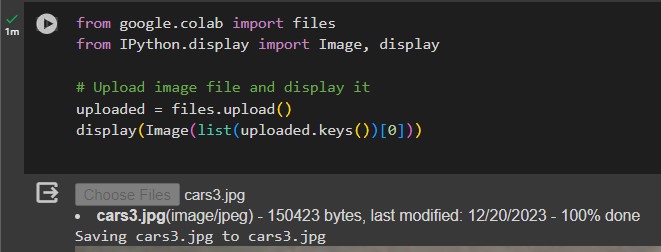
Features:

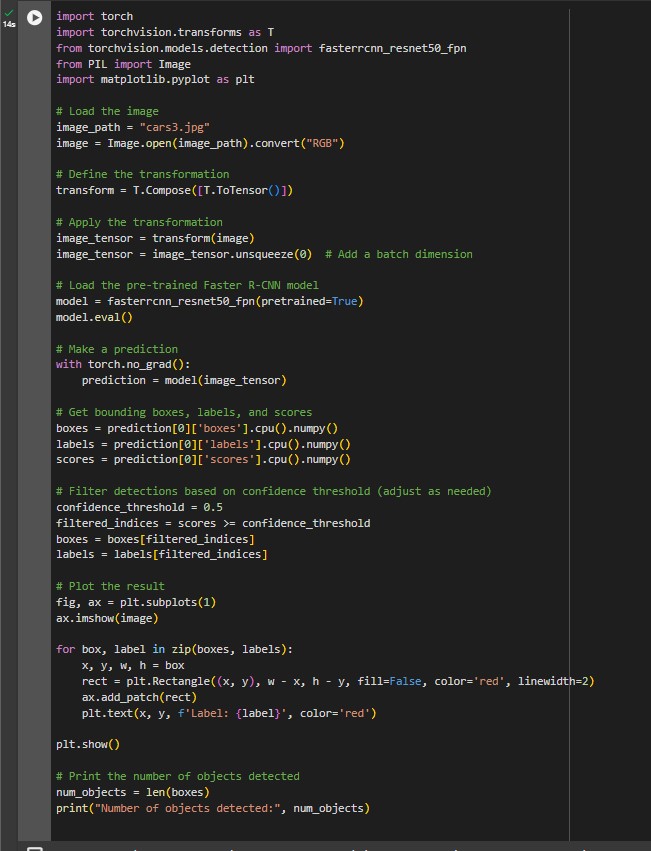
* **Datasets and Transforms:** Torchvision includes a variety of standard datasets for tasks such as image classification, object detection, and segmentation. It also provides data transforms for preprocessing images, including resizing, cropping, and data augmentation.
* **Models:** The library offers pre-trained models for image classification, object detection, and segmentation. These models are based on state-of-the-art architectures and can be easily fine-tuned for specific tasks.
* **Utilities:** Torchvision includes utilities for working with images, such as functions for reading and saving images, drawing bounding boxes, and visualizing model prediction.

**SNAPSHOT**









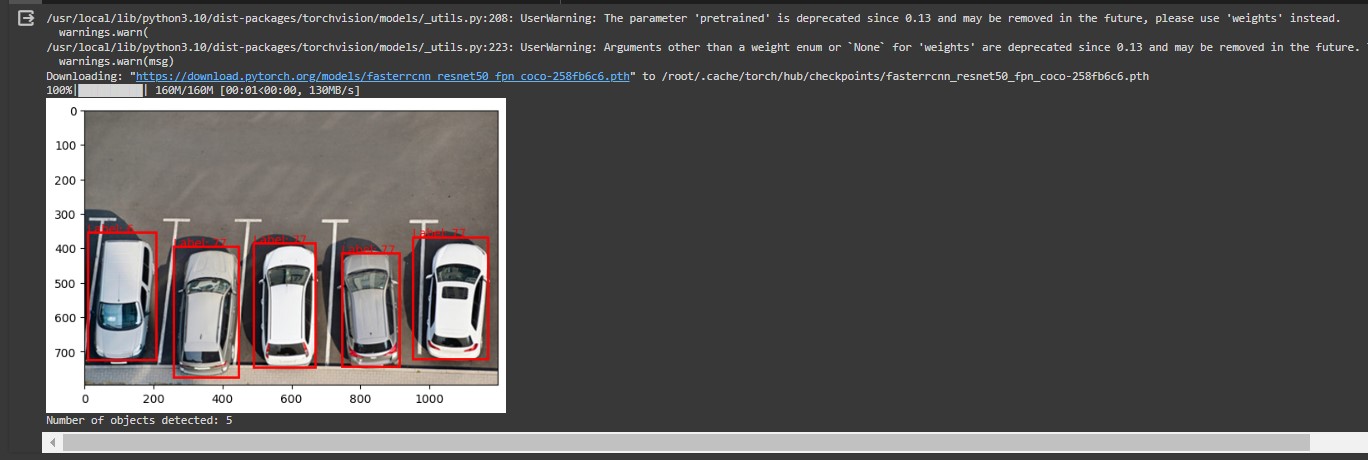


**Figure: Before Detection**

40

10.References

This is a sample image we feed to the algorithm and expect our algorithm to detect and identify objects in the image and label them according to the class assigned to it.



**Figure: After Detection**

**CONCLUSION**

By using this thesis and based on experimental results we are able to detect object more precisely and identify the objects individually with exact location of an object in the picture in x, y axis. This paper also provides experimental results on different methods for object detection and identification and compares each method for their efficiencies. This has also achieved significant milestones. The model demonstrated commendable performance in accurately identifying and localizing objects, thanks to a well-curated and diverse training dataset. While computational efficiency was a consideration, the model's ability to handle real-time applications was noteworthy. Despite challenges faced during the project, including data quality issues and deployment complexities, the transparent acknowledgment of these obstacles contributes to a realistic assessment.

**FUTURE ENHANCEMENTS**

The object recognition system can be applied in the area of surveillance system, face recognition, fault detection, character recognition etc. The objective of this thesis is to develop an object recognition system to recognize the 2D and 3D objects in the image. The performance of the object recognition system depends on the features used and the classifier employed for recognition. This

research work attempts to propose a novel feature extraction method for extracting global features and and obtaining local features from the region of interest. Also, the research work attempts to hybrid the traditional classifiers to recognize the object. The object recognition system developed in this research was tested with the benchmark datasets like COIL100, Caltech 101, ETH80 and MNIST.

As a scope for future enhancement,

* Features either the local or global used for recognition can be increased, to increase the efficiency of the object recognition system.
* Geometric properties of the image can be included in the feature vector for recognition.
* Using unsupervised classifier instead of a supervised classifier for recognition of the object.
* The proposed object system uses grey-scale image and discards the color information. The colour information in the image can be used for recognition of the object. Colour based object recognition plays vital role in Robotics.
* For night time visual tracking, night vision mode should be available as an inbuilt feature in the CCTV camera.
* Splitting and merging cannot be handled very well in all conditions using the single camera due to the loss of information of a 3D object projection in 2D images.

**REFERENCES**

1. Agarwal, S., Awan, A., and Roth, D. (2004). Learning to detect objects in images via a sparse, part-based representation. IEEE Trans. Pattern Anal. Mach. Intel. 26,1475–1490. doi:10.1109/TPAMI.2004.108

2. Alexe, B., Deselaers, T., and Ferrari, V. (2010). “What is an object?” in Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on (San Francisco, CA: IEEE), 73–80.doi:10.1109/CVPR.2010.5540226

3. Aloimonos, J., Weiss, I., and Bandyopadhyay, A. (1988). Active vision. Int. J. Comput. Vis. 1,333–356. doi:10.1007/BF00133571

4. Andreopoulos, A., and Tsotsos, J. K. (2013). 50 years of object recognition: directions forward. Comput. Vis. Image Underset. 117, 827–891. Doi:10.1016/j.cviu.2013.04.0055. Azizpour, H., and Laptev, I. (2012). “Object detection using strongly-supervised deformable part models,” in Computer Vision-ECCV 2012 (Florence: Springer),836–849.

6. Azzopardi, G., and Petkov, N. (2013). Trainable cos fire filters for key point detection and pattern recognition. IEEE Trans. Pattern Anal. Mach. Intel. 35, 490 503.doi:10.1109/TPAMI.2012.106

7. Azzopardi, G., and Petkov, N. (2014). Ventral-stream-like shape representation: from pixel intensity values to trainable object-selective cos fire models. Front. Comput. Neurosci. 8:80.doi:10.3389/fncom.2014.00080

8. Benbouzid, D., Busa-Fekete, R., and Kegl, B. (2012). “Fast classification using sparse decision dags,” in Proceedings of the 29th International Conference on Machine Learning (ICML-12), ICML ‘12, eds J. Langford and J. Pineau (New York, NY: Omni press), 951–958.

9. Bengio, Y. (2012). “Deep learning of representations for unsupervised and transfer learning,” in ICML Unsupervised and Transfer Learning, Volume 27 of JMLRProceedings, eds I. Guyon, G. Dror, V. Lemaire, G. W. Taylor, and D. L. Silver (Bellevue: JMLR.Org), 17–36.

10. Bourdev, L. D., Maji, S., Brox, T., and Malik, J. (2010). “Detecting people using mutually consistent pose let activations,” in Computer Vision – ECCV2010 – 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part VI, Volume 6316 of Lecture Notes in computer Science, eds K. Daniilidis, P. Maragos, and N. Par agios (Heraklion: Springer), 168–181.

The proposed object recognition system uses grey-scale image and discards the color information.The colour information in the image can be used for recognition of the object. Colour based object

40

10.References