**Object detection**

**A MINI PROJECT REPORT**

**18CSC305J - ARTIFICIAL INTELLIGENCE**

### Submitted by

**KSHITIJ KUMAR (RA2011003010724)**

*Under the guidance of*

**Mr. N GHUNTUPALLI MANOJ KUMAR**

### (Assistant Professor, Department of Computer Science and Engineering)

### in partial fulfillment for the award of the degree of

**BACHELOR OF TECHNOLOGY**

in

**COMPUTER SCIENCE & ENGINEERING**

of

**FACULTY OF ENGINEERING AND TECHNOLOGY**



#### S.R.M. Nagar, Kattankulathur, Chengalpattu District

**MAY 2023**

**SRM INSTITUTE OF SCIENCE AND TECHNOLOGY**

(Under Section 3 of UGC Act, 1956)

**BONAFIDE CERTIFICATE**

Certified that Mini project report titled **“OBJECT DETECTION”** is the bona fide work of  **KSHITIJ KUMAR(RA2011003010724)** who carried out the minor project under my supervision. Certified further, that to the best of my knowledge, the work reported herein does not form any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

**SIGNATURE SIGNATURE**

|  |  |
| --- | --- |
| Ghuntupalli Manoj Kumar N  **GUIDE**  Assistant Professor  Department of Computing Technologies | Dr. M. Pushpalatha  **HEAD OF THE DEPARTMENT**  Professor & Head  Department of Computing Technologies |

**ABSTRACT**

Object detection is widely used in the field of computer vision and crucial for variety of applications, e.g., self-driving car. During the development of half a century, object detection methods have been continuously developed, and generated numerous approaches which obtained promising achievements. At present, the approach of object detection has been largely evolved into two categories which are traditional machine learning methods utilizing varied computer vision techniques and deep learning method. This article presents a review of object detection techniques. Firstly, the existing methods based on traditional machine learning are summarized and introduced. Then, two main schools of deep learning methods, R-CNN and YOLO, are selected for analysis and introduction. At the end of the article, the methods mentioned are briefly compared and discussed.

**TABLE OF CONTENTS**

[**ABSTRACT**](#_bookmark1) **iii**

**TABLE OF CONTENTS iv**

[**LIST OF FIGURES**](#_bookmark2) **v**

[**ABBREVIATIONS**](#_bookmark3) **vi**

[**1 INTRODUCTION**](#_bookmark4) **7**

[**2 LITERATURE SURVEY**](#_bookmark20) **8**

**3** [**SYSTEM A**](#_bookmark30)**RCHITECTURE AND DESIGN 9**

3.1 Architecture diagram of proposed IoT based smart agriculture project 9

3.2 Description of Module and components10

**4 METHODOLOGY 14**

4.1 Methodological Steps 14

**5 CODING AND TESTING 15**

**6 SREENSHOTS AND RESULTS**

6.1 Buzzer and PIR 19

6.2 Ultrasonic and Scarecrow 19

6.3 Soil moisture sensor with water pump 20

6.4 LCD Screen 20

6.5 Whole Circuit 21

6.6 Thingspeak Server 22

**7 CONCLUSION AND FUTURE ENHANCEMENT 23**

7.1 Conclusion

7.2 Future Enhancement

**REFERENCES 24**

**LIST OF FIGURES**

##### [3.1.1](#_bookmark12) Architecture block 9

[3.2.1](#_bookmark16) **Soil moisture sensor** 10

3.2.3  [**Temperature Sensor (DHT 11)**](#_bookmark22)  10

[3.2.4](#_bookmark29) **PIR Motion Sensor**  10

[3.2.5 **Motor**](#_bookmark32) 11

* + 1. **Ultrasonic Sensor**  11

3.2.7 [**Buzzer**](#_bookmark36)  12

3.2.8 **WIFI MODULE ESP 8266**  13

[6.1.1](#_bookmark43) **Buzzer and PIR simulation on tinkercad** 19

[6.2.1](#_bookmark48) **Ultrasonic sensor and Scarecrow simulation on tinkercad** 19

[6.3.1](#_bookmark52) **Soilmoisture sensor and motor simulation on tinkercad** 20

[6.4.1](#_bookmark54) **LCD Screen with output** 20

* + 1. **Whole Circuit** 21

6.6.1 **Screenshot of thingspeak server** 22

v

**ABBREVIATIONS**

**IOT** Internet of Things

**PIR** Passive Infrared

**LCD** Liquid Crystal Diode

**DHT** Distributed hash table

**IR** Infra red

**UART** Universal Asynchronous Receiver/Transmitter

**IDE** Integrated Development Environment

**vi**

**CHAPTER 1**

**INTRODUCTION**

A few years ago, the creation of the software and hardware image processing systems was mainly limited to the development of the user interface, which most of the programmers of each firm were engaged in. The situation has been significantly changed with the advent of the Windows operating system when the majority of the developers switched to solving the problems of image processing itself. However, this has not yet led to the cardinal progress in solving typical tasks of recognizing faces, car numbers, road signs, analyzing remote and medical images, etc. Each of these "eternal" problems is solved by trial and error by the efforts of numerous groups of the engineers and scientists. As modern technical solutions are turn out to be excessively expensive, the task of automating the creation of the software tools for solving intellectual problems is formulated and intensively solved abroad. In the field of image processing, the required tool kit should be supporting the analysis and recognition of images of previously unknown content and ensure the effective development of applications by ordinary programmers. Just as the Windows toolkit supports the creation of interfaces for solving various applied problems.

Object recognition is to describe a collection of related computer vision tasks that involve activities like identifying objects in digital photographs. Image classification involves activities such as predicting the class of one object in an image. Object localization is refers to identifying the location of one or more objects in an image and drawing an abounding box around their extent. Object detection does the work of combines these two tasks and localizes and classifies one or more objects in an image. When a user or practitioner refers to the term “object recognition“, they often mean “object detection“. It may be challenging for beginners to distinguish between different related computer vision tasks.

So, we can distinguish between these three computer vision tasks with this example: Image Classification: This is done by Predict the type or class of an object in an image. Input: An image which consists of a single object, such as a photograph.

Output: A class label (e.g. one or more integers that are mapped to class labels).

Object Localization: This is done through, Locate the presence of objects in an image and indicate their location with a bounding box.

Input: An image which consists of one or more objects, such as a photograph. Output: One or more bounding boxes (e.g. defined by a point, width, and height).

Object Detection: This is done through, Locate the presence of objects with a bounding box and types or classes of the located objects in an image.

**CHAPTER 2**

**LITERATURE SURVEY**

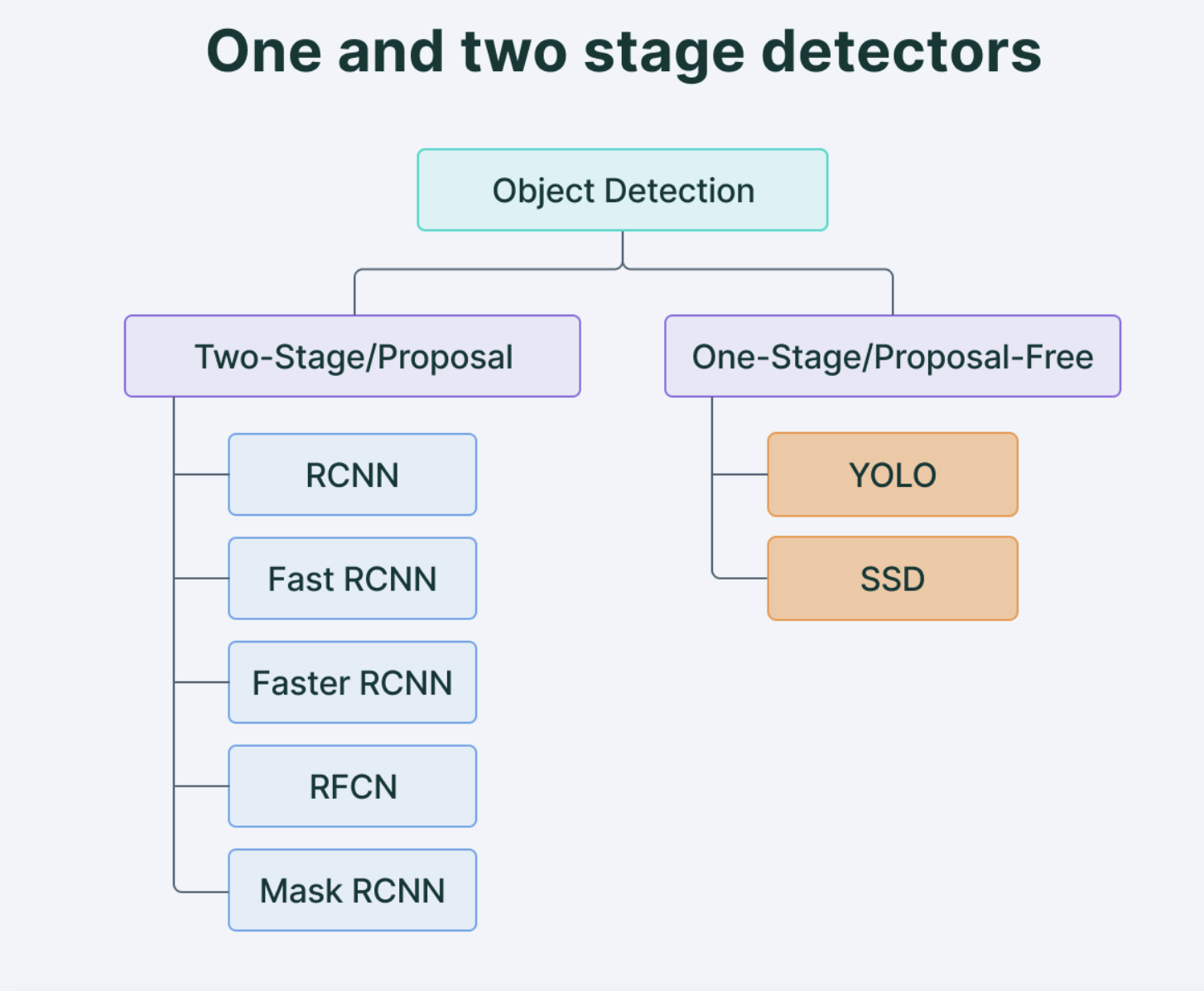
Among the many uses for object detection, security is one of the most crucial ones. Of course, there are many applications of object detection utilizing machine learning in the realm of security. The topic that we opted to conduct a literature review on has not yet been covered by a publication, as far as we can tell. And to our knowledge, no one has conducted a thorough literature study on the subject we have chosen. The two studies that are the closest to being a survey that we have discovered are either specific or cover particular methodologies rather than analyze other papers.

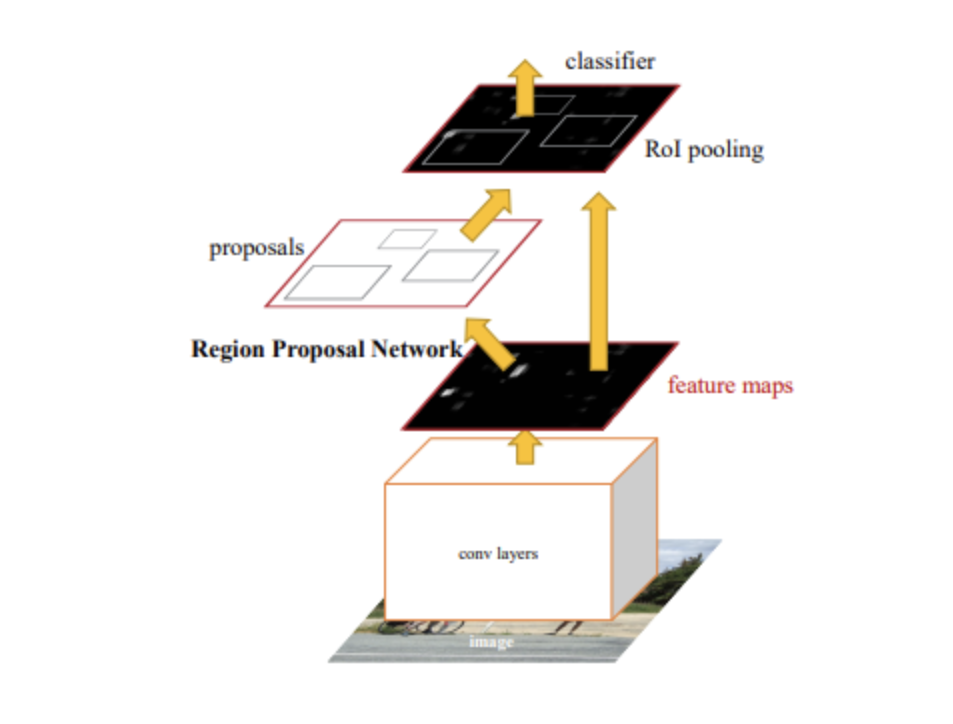
A. A study on several techniques for object detection and tracking in video surveillance footage was done in one work by Murugan et al. [1]. The paper highlights how video surveillance has existed as a technology since the 1950s and reinforces the point stated in the Introduction section about how watching security cameras can be tiresome and have an adverse effect on a person's mental health. It also illustrates how automating the procedure was the answer to that tedious chore by introducing a sophisticated surveillance system. This paper's main goal is to describe the various techniques for object detection and tracking in videos. The first technique described in the study is backdrop subtraction. The report claims that one of the most popular techniques for spotting moving items is backdrop removal. To show only the moving object's pixels, background subtraction includes identifying the background and removing it. The issue with it is that the results of the procedure are impacted if the background is not static and changes as a result of illumination or specific weather conditions. There are other algorithms for background subtraction as well.

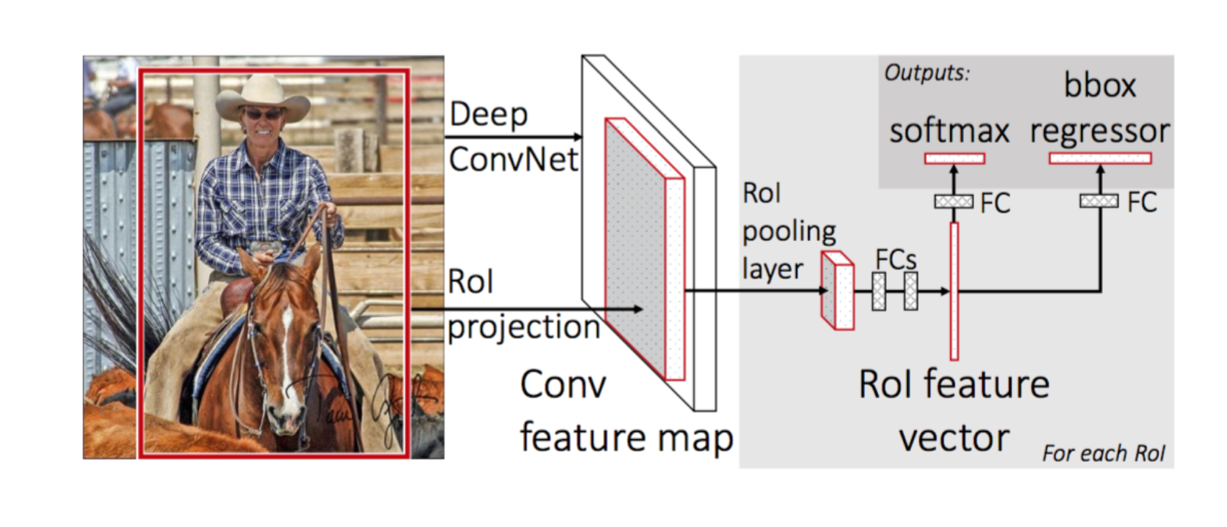
B. A study by Flitton et al. [2] compares various 3D interest point descriptors for CT images of baggage at airports. Finding interesting stuff during baggage x-ray inspections is the core concept here. Five distinct methods were compared in the paper: density, density histogram, density gradient histogram, rotation invariant feature transform, and scale invariant feature transform. The report does a decent job of assessing each while providing crucial indicators, although the research was only done on a relatively narrow issue. Both of the aforementioned studies do a great job of describing the various strategies, but in comparison to our paper, they are too narrow in scope. Instead of separating the focus on machine learning and object detection, both publications provide deeper explanation of the object detection techniques. The two articles cover fewer papers than our Systematic Literature Review does because of their narrower scopes.

**CHAPTER 3**

**SYSTEM ARCHITECTURE AND DESIGN**

****

****

****

**CHAPTER 4**

**METHODOLOGY**

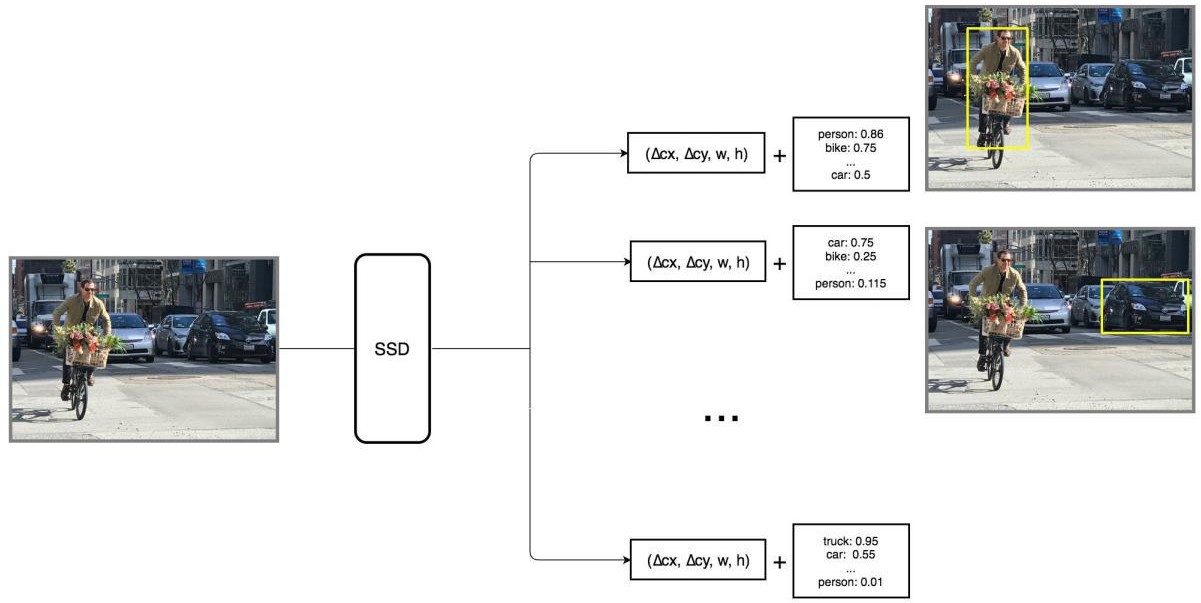
In more traditional ML-based approaches, computer vision techniques are used to look at various features of an image, such as the color histogram or edges, to identify groups of pixels that may belong to an object. These features are then fed into a regression model that predicts the location of the object along with its label.

On the other hand, deep learning-based approaches employ convolutional neural networks (CNNs) to perform end-to-end, unsupervised object detection, in which features don’t need to be defined and extracted separately. For a gentle introduction to CNNs, [check out this overview](https://heartbeat.fritz.ai/a-beginners-guide-to-convolutional-neural-networks-cnn-cf26c5ee17ed).

Because deep learning methods have become the state-of-the-art approaches to object detection, these are the techniques we’ll be focusing on for the purposes of this guide.

**CHAPTER 5**

**CODING AND TESTING**



**Code:**

document.getElementById("ai").addEventListener("change", toggleAi)

document.getElementById("fps").addEventListener("input", changeFps)

const video = document.getElementById("video");

const c1 = document.getElementById('c1');

const ctx1 = c1.getContext('2d');

var cameraAvailable = false;

var aiEnabled = false;

var fps = 16;

/\* Setting up the constraint \*/

var facingMode = "environment"; // Can be 'user' or 'environment' to access back or front camera (NEAT!)

var constraints = {

audio: false,

video: {

facingMode: facingMode

}

};

/\* Stream it to video element \*/

camera();

function camera() {

if (!cameraAvailable) {

console.log("camera")

navigator.mediaDevices.getUserMedia(constraints).then(function (stream) {

cameraAvailable = true;

video.srcObject = stream;

}).catch(function (err) {

cameraAvailable = false;

if (modelIsLoaded) {

if (err.name === "NotAllowedError") {

document.getElementById("loadingText").innerText = "Waiting for camera permission";

}

}

setTimeout(camera, 1000);

});

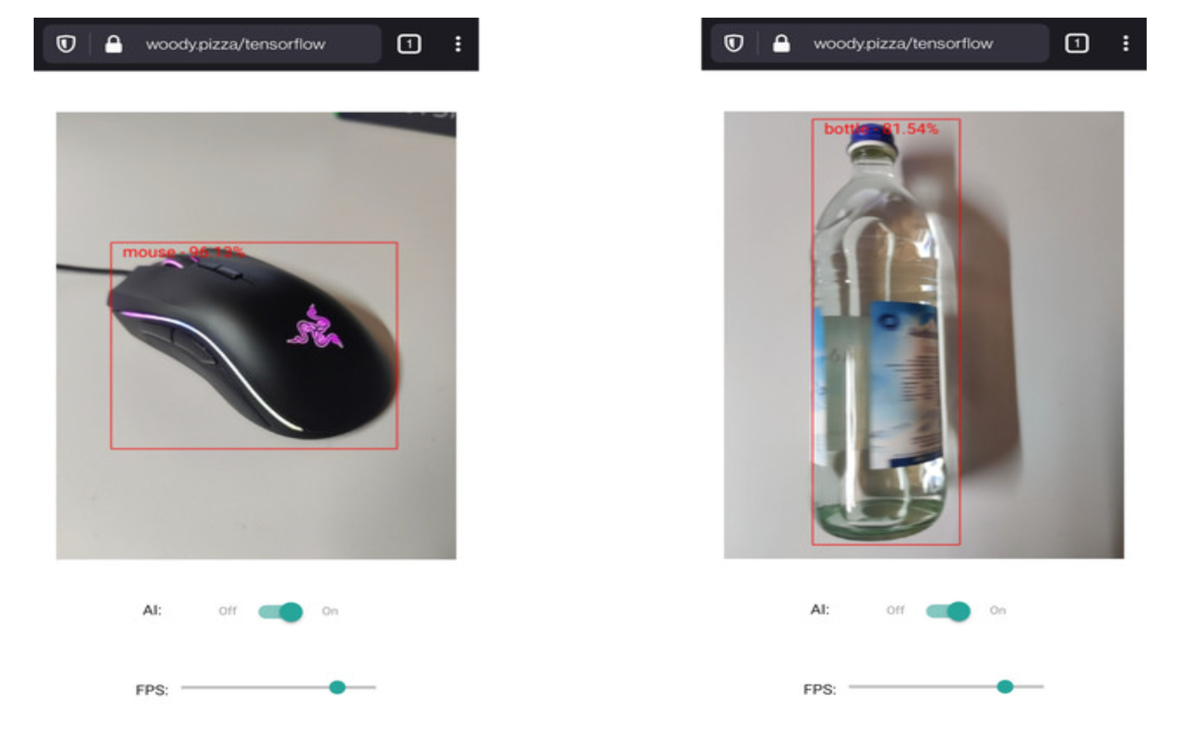
}

To evaluate object detection models like R-CNN and [YOLO](https://blog.paperspace.com/how-to-implement-a-yolo-object-detector-in-pytorch/), the **mean average precision (mAP)** is used. The mAP compares the ground-truth bounding box to the detected box and returns a score. The higher the score, the more accurate the model is in its detections.

In my last article we looked in detail at the [confusion matrix, model accuracy, precision, and recall](https://blog.paperspace.com/deep-learning-metrics-precision-recall-accuracy). We used the Scikit-learn library to calculate these metrics as well. Now we'll extend our discussion to see how precision and recall are used to calculate the map.

**CHAPTER 6**

**SCREENSHOTS AND RESULTS**



**CHAPTER 7**

**CONCLUSION AND FUTURE ENHANCEMENTS**

#### Object detection is a key ability required by most computer and robot vision systems. The latest research on this area has been making great progress in many directions. In the current manuscript, we give an overview of past research on object detection, outline the current main research directions, and discuss open problems and possible future directions.

#### 

#### Object detection methods can be grouped in five categories, each with merits and demerits: while some are more robust, others can be used in real-time systems, and others can be handle more classes, etc.

#### This approach considers object and part models and their relative positions. In general, it is more robust that other approaches, but it is rather time consuming and cannot detect objects appearing at small scales. It can be traced back to the deformable models ([Fischler and Elschlager, 1973](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full" \l "B29)), but successful methods are recent ([Felzenszwalb et al., 2010b](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full" \l "B28)). Relevant works include [Felzenszwalb et al. (2010a)](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full" \l "B27) and [Yan et al. (2014)](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full#B86), where efficient evaluation of deformable part-based model is implemented using a coarse-to-fine cascade model for faster evaluation, [Divvala et al. (2012)](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full" \l "B19), where the relevance of the part-models is analyzed, among others [e.g., [Azizpour and Laptev (2012)](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full" \l "B5), [Zhu and Ramanan (2012)](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full#B93), and [Girshick et al. (2014)](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full" \l "B33)].

#### n such architectures, the parameters of predefined operators and the combination of the operators are learned, sometimes considering an abstract notion of fitness. These are general-purpose architectures, and thus they can be used to build several modules of a larger system (e.g., object recognition, key point detectors and object detection modules of a robot vision system). Examples include trainable COSFIRE filters ([Azzopardi and Petkov, 2013](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full#B6), [2014](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full#B7)), and Cartesian Genetic Programming (CGP) ([Harding et al., 2013](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full#B34); [Leitner et al., 2013](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full#B41)).

#### Most methods used in practice have been designed to detect a single object class under a single view, thus these methods cannot handle multiple views, or large pose variations; with the exception of deformable part-based models which can deal with some pose variations. Some works have tried to detect objects by learning subclasses ([Wu and Nevatia, 2007](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full#B83)) or by considering views/poses as different classes ([Verschae and Ruiz-del-Solar, 2012](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full" \l "B75)); in both cases improving the efficiency and robustness. Also, multi-pose models [e.g., [Erol et al. (2007)](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full#B24)] and multi-resolution models [e.g., [Park et al. (2010)](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full#B56)] have been developed.

**REFERENCES**

[1] Zhang, X., Davidson, E. A,"Improving Nitrogen and Water Management in Crop Production on a National Scale", American Geophysical Union, December, 2018.How to Feed the World in 2050 by FAO.

[2] Abhishek D. et al., "Estimates for World Population and Global Food Availability for Global Health", Book chapter, The Role of Functional Food Security in Global Health, 2019, Pages 3-24.Elder M., Hayashi S., "A Regional Perspective on Biofuels in Asia", in Biofuels and Sustainability, Science for Sustainable Societies, Springer, 2018.

[3] Zhang, L., Dabipi, I. K. And Brown, W. L, “Internet of Things Applications for Agriculture". In, Internet of Things A to Z: Technologies and Applications, Q. Hassan (Ed.), 2018.

[4] S. Navulur, A.S.C.S. Sastry, M.N. Giri Prasad,"Agricultural Management through Wireless Sensors and Internet of Things" International Journal of Electrical and Computer Engineering (IJECE), 2017; 7(6) :3492-3499.

[5] E. Sisinni, A. Saifullah, S.Han, U. Jennehag and M.Gidlund, "Industrial Internet ofThings: Challenges,Opportunities, and Directions," in IEEE Transactions on Industrial Informatics, vol. 14, no. 11, pp. 4724-4734, Nov. 2018.

[6] M. Ayaz, M. Ammad-uddin, I. Baig and e. M. Aggoune, "Wireless Possibilities: A Review," in IEEE Sensors Journal, vol. 18, no. 1, pp. 4-30, 1 Jan.1, 2018.