# Design and Development of Real-time Object Detection

Anwesa Naskar[1], Bishwajeet Kumar[1] , Shivam Raj[1], Pratik Pradhan[1] ,Ketan Kumar Mahato[1], and Sohham Das[1] [1]NSHM Knowledge Campus, Durgapur

**Abstract** Real-time object detection constitutes an expansive, dynamic, and intricate domain within computer vision. When singling out a specific item in an image, it is referred to as Image Localization; whereas, the identification of multiple objects within an image falls under the category of Object Detection. This process involves discerning semantic objects of a particular class in both digital images and videos. The practical applications of real-time object detection span various fields such as object tracking, video surveillance, pedestrian detection, people counting, autonomous vehicles, facial recognition, sports ball tracking, and more. Convolutional Neural Networks stand out as a prominent tool in deep learning, utilised for object detection through Tensor-flow—a programming library primarily designed for real-time computer vision.

**Keywords** Computer vision, Deep Learning, Convolution Neural Networks Tensorflow.js.

## 1. Introduction

Object Detection involves identifying and recognizing real-world objects, such as cars, bikes, TVs,

flowers, and humans, within images or videos. This technique enables a detailed understanding of the content by facilitating the recognition, localization, and detection of multiple objects in an image. Commonly applied in image retrieval, security, surveillance, and advanced driver assistance systems (ADAS), Object Detection utilises various approaches, including Feature-Based Object Detection, Viola-Jones Object Detection, SVM Classifications with HOG Features, and Deep Learning.

These days, a crucial task in video surveillance applications is detecting objects within a video. The technique employed for object detection helps identify the desired objects in video sequences and organises pixels related to these objects. Detecting objects in video sequences is particularly significant in applications like video surveillance. The process of object detection in a video stream involves steps such as pre-processing, segmentation, foreground and background extraction, and feature extraction. Unlike computers, humans have the innate ability to effortlessly spot and recognize objects in an image. The human visual system is both swift and precise, capable of undertaking complex tasks, such as identifying multiple objects, with minimal conscious effort. Thanks to the abundance of data, faster GPUs, and improved algorithms, computers can now be trained effectively to detect and classify multiple objects within an image with remarkable accuracy..

## 2. Background studies

### 2.1. TensorFlow.js

Tensorflow.js, created by Google, is a JavaScript framework that empowers developers to create and train machine learning models for use in web browsers. Additionally, developers can choose to utilise pre-trained models instead of training their own.

### 2.2. Machine Learning

Machine learning, a branch of artificial intelligence (AI), allows computers to learn and improve their task execution without explicit instructions. Machine learning programs can access and utilise data to autonomously learn.

A machine learning program operates by creating a model, similar to a mathematical model, with various inputs and outputs. The model is trained with data, and through iterations, the program adjusts the model's parameters to minimise the model's error over time. In essence, the program modifies the model to make its output align with the desired output.

Tensorflow.js, leveraging WebGL (a JavaScript graphics API), achieves excellent performance, making it a fast framework. Additionally, there is a Node.js version of TensorFlow called tfjs-node, which further enhances performance compared to the browser version.

TensorFlow.js enables the utilisation of pre-trained models, making it possible for developers who are new to machine learning to seamlessly integrate it into their papers. This means that individuals keen on incorporating machine learning into their applications, like creating an audio recognition app, can begin immediately without the need to invest time in training their own models. This accessibility aspect marks a significant improvement, making machine learning more approachable.

TensorFlow.js provides the capability to load models that have been trained using the Python version of TensorFlow. This allows developers to construct and train a model using Python and subsequently integrate it into their JavaScript applications. The advantage of this approach lies in the substantial enhancement of performance, as there's no need to conduct model training within the browser.

Moreover, TensorFlow.js facilitates the retraining and customization of pre-existing machine learning models using client-side data in the browser, such as image or video data from a webcam. The possibilities with this functionality are truly limitless.

## 2.3. Getting started with TensorFlow.js

You can find a wealth of online tutorials to get started with building applications using TensorFlow.js. Valuable tutorials can be accessed on the official TensorFlow.js website, available at tensorflow.org/js/tutorials.. Here, you can gain insights into constructing various applications, including an image classifier and an app capable of recognizing hand-drawn digits. Another excellent resource for beginners is a tutorial series hosted by The Coding Train, covering fundamental concepts and guiding the audience through building regression models like linear regression and polynomial regression.

Projects for Beginners to Explore

1. In this paper, you will learn how to build an image classifier from scratch, enabling you to train your own custom classifier and utilize a pre-trained model for accurate image classification.
2. Artistic creations: Discover the artistry of machine learning as you learn to craft your own distinctive and beautiful artworks.
3. Real-time object detection: Dive into a tutorial where you'll build a web app capable of detecting objects in real-time through the webcam.
4. Game playing: With TensorFlow.js, even beginners can delve into creating their initial AI video game, marking an accessible entry point into the world of artificial intelligence in gaming.

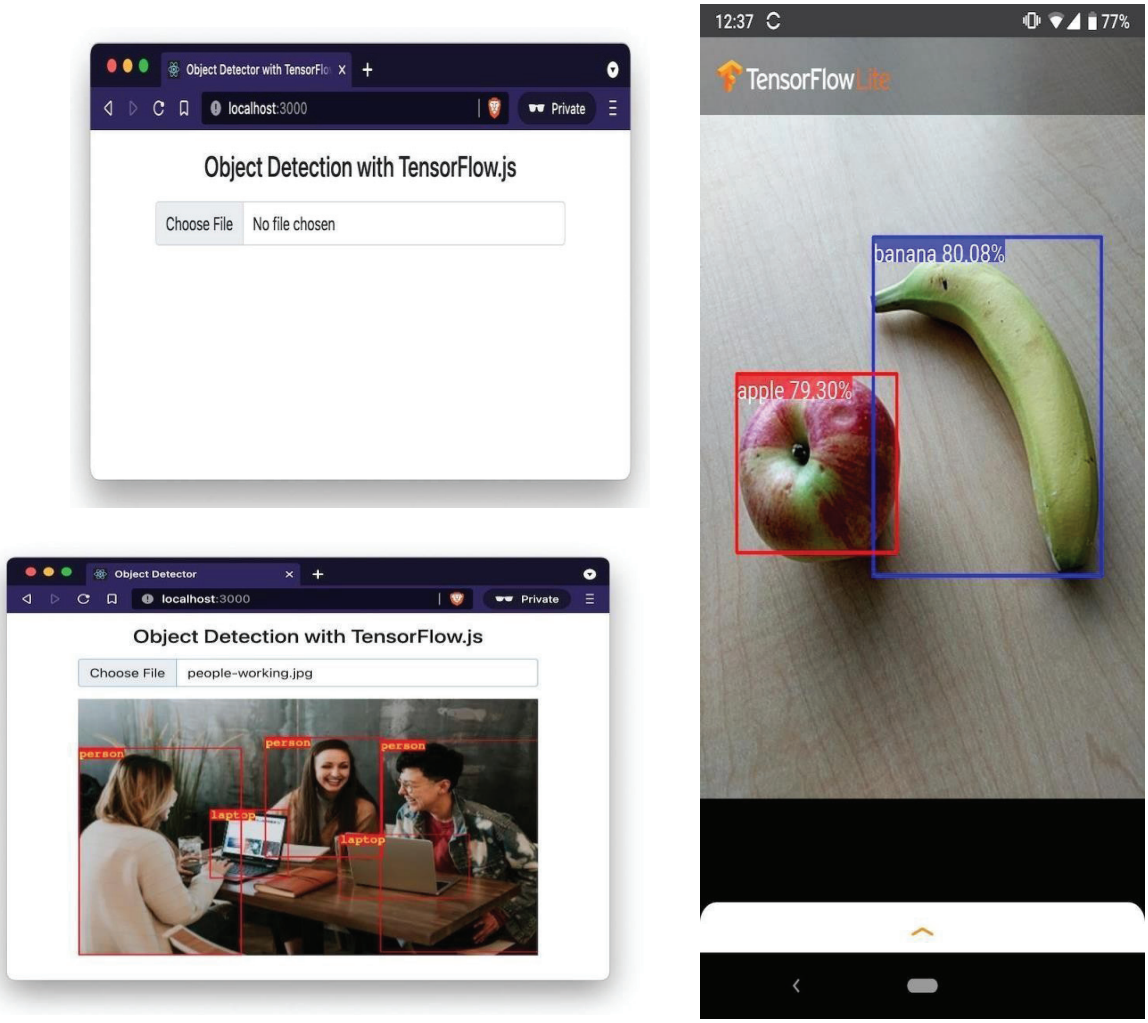


Figure 1: Tensor Flow process

## 2.4 Applications

The major applications of Object Detection are:

## FACIAL RECOGNITION:

Deep Face, a deep learning facial recognition system, has been developed by a team of researchers at Facebook. Google also utilizes its own facial recognition system in Google Photos, which automatically organizes photos based on individuals. Facial recognition involves the analysis of various facial components, such as the eyes, nose, mouth, and eyebrows, to identify a person's face.

## PEOPLE COUNTING:

People counting, a subset of object detection, serves multiple purposes, including identifying individuals or potential criminals, analysing store performance, or gathering crowd statistics during festivals. This process can be challenging as people often move out of the frame quickly.

## INDUSTRIAL QUALITY CHECK:

In industrial processes, particularly in product identification, object detection assumes a crucial role. It streamlines various tasks including sorting, inventory management, machining, quality control, and packaging. The automatic counting and localization of objects contribute significantly to enhancing inventory accuracy.

## SELF-DRIVING CARS:

Self-driving technology holds great promise for the future and involves complex mechanisms that combine radar, laser light, GPS, odometer, and computer vision to perceive the surrounding environment. Advanced control systems interpret sensory information, enabling navigation systems to recognize obstacles even at high speeds.

## SECURITY:

Object Detection is pivotal in security, used in applications like Apple's face ID and retina scans from sci-fi movies. Governments employ this technology to access security feeds, matching them with databases to identify criminals or detect objects related to criminal activities.

## OBJECT DETECTION WORKFLOW AND FEATURE EXTRACTION:

All Object Detection Algorithms operate on the same principle, focusing on extracting features from input images to determine image class. The variations lie in the specific workings of each algorithm.

Prior to the emergence of TensorFlow.js, Google had developed TensorFlow, an open-source machine learning framework written in C++. TensorFlow integrates various machine learning and deep learning models and algorithms. It utilises Python as a user-friendly front-end API for constructing applications, while the execution of these applications occurs in C++.

Before TensorFlow.js, executing machine learning in JavaScript typically involved utilising an API where the server would deploy the model upon the client's request using JavaScript. In 2017, Google introduced DeepLearn.js, a paper that enabled machine learning in JavaScript without the need for an API. Subsequently, in 2018, the DeepLearn.js team merged with the TensorFlow team, resulting in the creation of TensorFlow.js.

It is worth noting that there is no requirement to include page numbers or running heads in your paper. We will take care of those details on our end. If the title of your paper is too long for a running head, we will appropriately shorten it. We appreciate any suggestions you may have for a concise version.

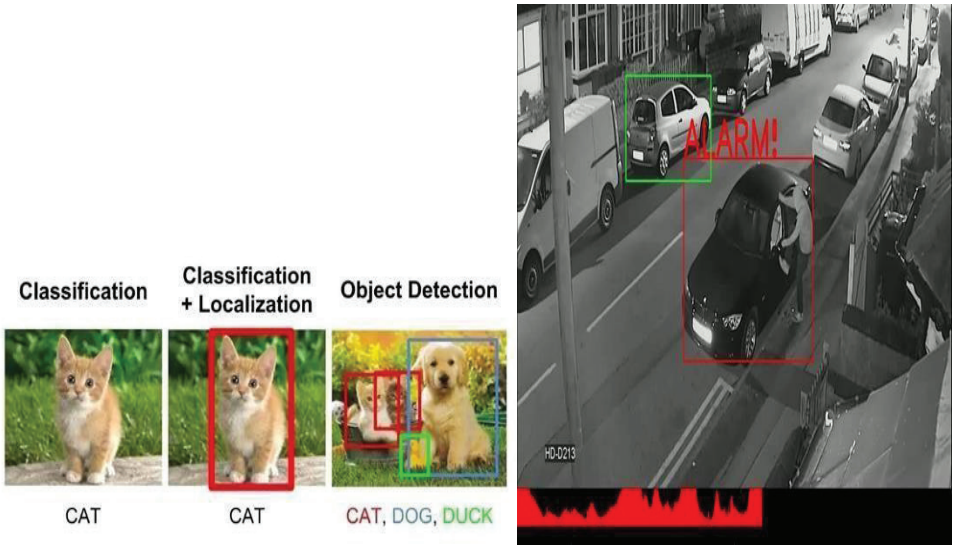


Figure 2: Object detection

## 3. Result and analysis:

For this paper, we have incorporated around 15 to 20 objects to be detected during the training phase. Examples of these objects include 'person', 'car', 'cell phone', 'laptop', 'cup', 'bottle', and more. The paper output showcases the detected objects with a rectangular box surrounding each object. Additionally, a label indicating the object's name is displayed on top of it, providing precise detection accuracy. This system is capable of identifying multiple objects within a single image with confidence.

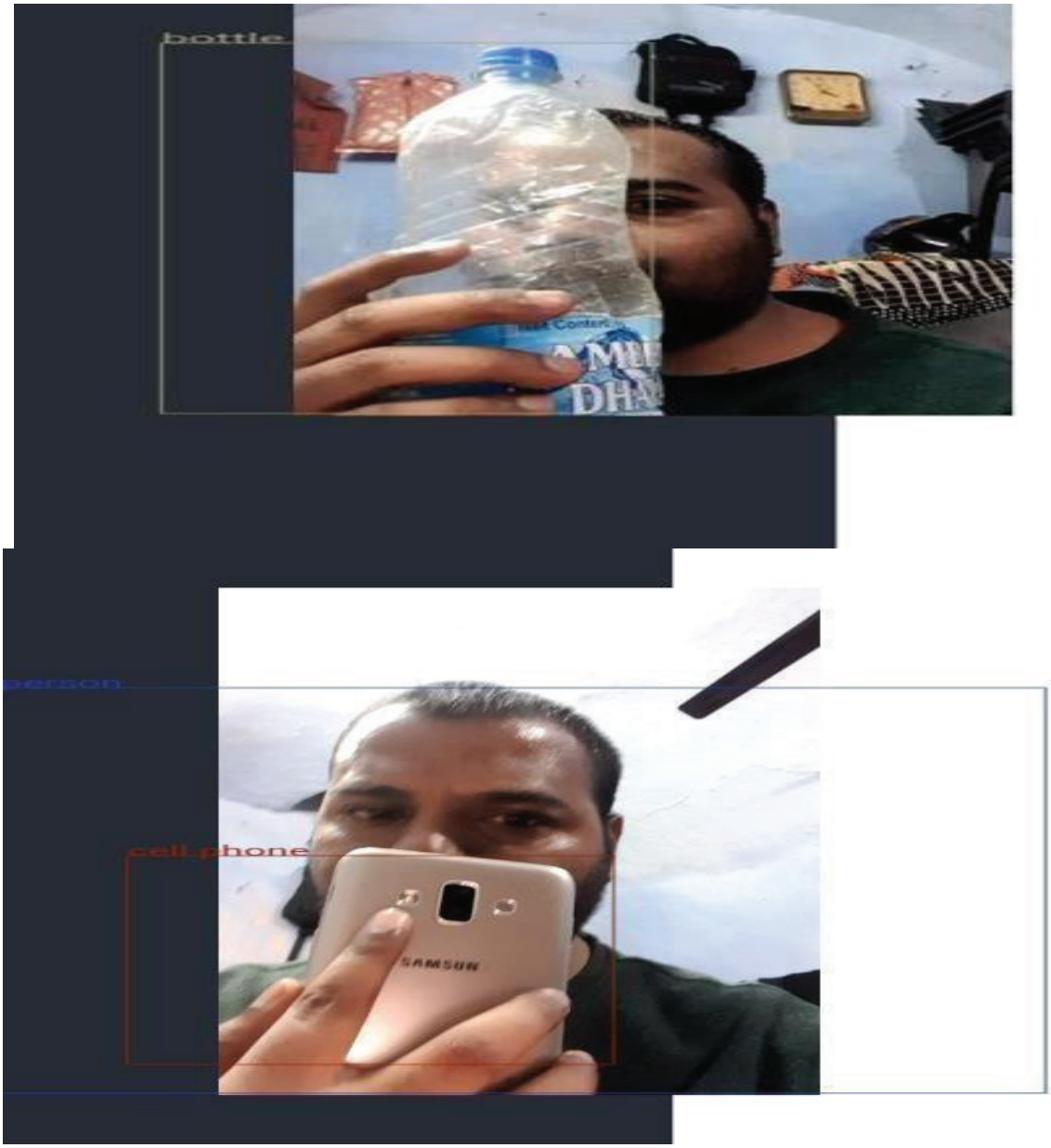


Figure 3: Result Analysis

## 4. Conclusion

TensorFlow stands out as an excellent library for both numerical and graphical computation, particularly in the development of deep learning networks. It holds the position as the most widely utilized library across various applications, including real-time object detection in Google Search, Google Translate, Google Photos, and more. Machine learning has paved the way for incredible achievements, such as healthcare applications, recommendation engines for movies and music, personalised ads, and social media

sentiment mining. Amidst these awe-inspiring advancements in machine learning and artificial intelligence, TensorFlow emerges as a crucial tool, contributing significantly to the realisation of these ambitious goals.

# REFERENCES

1. Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7464-7475. 2023.

2. Cai, Yuxuan, Hongjia Li, Geng Yuan, Wei Niu, Yanyu Li, Xulong Tang, Bin Ren, and Yanzhi Wang. "Yolobile: Real-time object detection on mobile devices via compression-compilation co-design." In Proceedings of the AAAI conference on artificial intelligence, vol. 35, no. 2, pp. 955-963. 2021.

3. Liu, Jiang-Jiang, Qibin Hou, Ming-Ming Cheng, Jiashi Feng, and Jianmin Jiang. "A simple pooling-based design for real-time salient object detection." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 3917- 3926. 2019.

4. Lyu, Chengqi, Wenwei Zhang, Haian Huang, Yue Zhou, Yudong Wang, Yanyi Liu, Shilong Zhang, and Kai Chen. "Rtmdet: An empirical study of designing real-time object detectors." arXiv preprint arXiv:2212.07784 (2022).

5. Shi, Guangsheng, Ruifeng Li, and Chao Ma. "Pillarnet: Real-time and high- performance pillar-based 3d object detection." In European Conference on Computer Vision, pp. 35-52. Cham: Springer Nature Switzerland, 2022.