OBJECT DETECTION WITH DEEP LEARNING

A Project Work Synopsis

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

BACHELOR OF ENGINEERING

IN

BE CSE - AIML

Submitted by:

VUPPUTURI BHARATH

20BCS6586

Under the Supervision of:

Ms. Sukhman Kaur

CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413, PUNJAB

MONTH & YEAR
OCTOBER - 2022

Table of Contents

ABSTRACT

Due to object detection's close relationship with video analysis and image understanding, it has attracted much research attention in recent years. Traditional object detection methods are built on handcrafted features and shallow trainable architectures. Their performance easily stagnates by constructing complex ensembles that combine multiple low-level image features with high-level context from object detectors and scene classifiers. With the rapid development in deep learning, more powerful tools, which are able to learn semantic, high-level, deeper features, are introduced to address the problems existing in traditional architectures. These models behave differently in network architecture, training strategy, and optimization function. In this paper, we provide a review of deep learning-based object detection frameworks. Our review begins with a brief introduction on the history of deep learning and its representative tool, namely, the convolutional neural network. Then, we focus on typical generic object detection architectures along with some modifications and useful tricks to improve detection performance further. As distinct specific detection tasks exhibit different characteristics, we also briefly survey several specific tasks, including salient object detection, face detection, and pedestrian detection.

Timeline/Gantt Chart

S.N	Strategies	1st week	2nd week	3rd week	4th week	5th week	6th week
1)	Problem Identification						
2)	Research & Analysis						
3)	Design						
4)	Coding						
5)	Implementation & testing						
6)	Project finalisation						
7)	Documentation						

1. INTRODUCTION

Humans can easily detect and identify objects present in an image. The human visual system is fast and accurate and can perform complex tasks like identifying multiple objects and detect obstacles with little conscious thought. With the availability of large amounts of data, faster GPUs, and better algorithms, we can now easily train computers to detect and classify multiple objects within an image with high accuracy. This explores terms such as object detection, object localization, loss function for object detection and localization, and finally explore an object detection algorithm. The ultimate purpose of object detection is to locate important items, draw rectangular bounding boxes around them, and determine the class of each item discovered. Applications of object detection arise in many different fields including detecting pedestrians for self-driving cars, monitoring agricultural crops, and even real-time ball tracking for sports. Researchers have dedicated a substantial amount of work towards this goal over the years.

1.1 Problem Definition:

- The first major complication of object detection is its added goal: not only do we want to classify image objects but also to determine the objects' positions, generally referred to as the *object localization* task.
- Object detection algorithms need to not only accurately classify and localize important objects, they also need to be incredibly fast at prediction time to meet the real-time demands of video processing.
- The limited amount of annotated data currently available for object detection proves to be another substantial hurdle. Object detection datasets typically contain ground truth examples for about a dozen to a hundred classes of objects, while image classification datasets can include upwards of 100,000 classes.

1.2 Project overview:

- In this project, a dataset is taken for reference of particular type of objects for instance Fashion MNIST dataset. After training and testing, the model is ready to detect the object accordingly. The only drawback/ limitation is objects outside the dataset can't be detected by the model.

1.3 Hardware Specifications:

- Processor – 64-bit eight-core, 2.5GHz per core.
- RAM – Minimum 4GB required.
- Hard Disk – SSD or HDD minimum 40GB free space required.

1.4 Software Specifications:

- Edition - Windows 10/11
- OS build 19043.1526
- Experience Windows Feature Experience Pack 120.2212.4170.0
- Python installed – version 3.7 to 3.10
- OpenCV installed on windows
- Mediapipe installed
- Any python compiler or system terminal

Media-pipe:

Media-Pipe is a framework which is used for applying in a machine learning pipeline, and it is an open-source framework of Google. The Media-Pipe framework is useful for cross platform development since the framework is built using the time series data.

Open-CV:

Open-CV is a computer vision library which contains image-processing algorithms for object detection. Open-CV is a library of python programming language, and real-time computer vision applications can be developed by using the computer vision library. The Open-CV library is used in image and video processing and also analysis such as face detection and object detection.

2. LITERATURE REVIEW

Literature review has been selected as the research method. The literature review is selected in order to gain knowledge and deep understanding about various deep learning models and their efficiency so that the most suitable and efficient method can be selected from the identified models.

Object detection is the procedure of determining the instance of the class to which the object belongs and estimating the location of the object by outputting the bounding box around the object. Detecting single instance of class from image is called as single class object detection, whereas detecting the classes of all objects present in the image is known as multi class object detection. Different challenges such as partial/full occlusion, varying illumination conditions, poses, scale, etc are needed to be handled while performing the object detection.

2.1 Exsisting System:

The first object detector came out in 2001 and was called the Viola Jones Object Detector. Although, it was technically classified as an object detector, it's primary use case was for facial detection. It provided a real time solution and was adopted by many computer vision libraries at the time. The field was substantially accelerated with the advent of Deep Learning. It classified each part of the image as an object/non object and subsequently combined the results to generate the final set of predictions.

2.2 Proposed System:

In this project, the model is created by machine learning algorithms. Model has access for a dataset called MNIST that includes images of some various types of clothing. With the help of the dataset, the model is able to differentiate clothing types. This model is included with input functions and output functions after calculating its accuracy. The model precise that input image is given then detects it and shows the type of clothing.

PROBLEM FORMULATION

Problem Statement for Object Detection

There are always two components to constructing a deep learning model. The first component is responsible for dividing the training data into input and targets. The second component is deciding upon the neural network architecture and training regime. The input for these models is an image. The targets are a list of object classes relaying what class the object belongs to and their corresponding coordinates. These coordinates signify where in the image the object exist. There are 4 types of coordinates- the center x and y coordinates and the height and width of the bounding box. We shall use the term bounding box to denote the box formed by applying these 4 coordinates on the image. The network is trained to predict a list of objects with their corresponding locations in the form of bounding box coordinates.

OBJECTIVES

The proposed work is aimed to carry out work leading to the development of an approach for Object Detection with Deep learning The proposed aim will be achieved by dividing the work into followingobjectives:

1. create a model that detects any object based on dataset.
2. Dataset MNIST is taken for model.
3. model is trained and tested with the classes in the dataset.
4. After training and testing, find its accuracy accordingly
5. Accuracy gives you whether the model is working fine or not
6. Make sure the accuracy is above 80%
7. Now give a image input
8. It detects the image and find which type of cloth it is.

METHODOLOGY

The following methodology will be followed to achieve the objectives defined for proposed research work:

1. Detailed study of Object Detection with Deep learning will be done.
2. Installation and hand on experience on existing approaches of object will be done. Relative pros and cons will be identified.
3. Various parameters will be identified to evaluate the proposed system.
4. Comparison of new implemented approach with exiting approaches will be done.
5. Image of object taken by the camera or any object taken from web.
6. The model checks the image on the basis of all classes in dataset
7. If majority of the classes fits for a image it detects and shows its type.
8. If object isn’t matched with types of objects with in dataset it doesn’t detect it.

REFERENCES

[1] Manjula S, Lakshmi Tamilselvan (2016) A study on object detection.

[2] Jun Deng, Xiaojing Xuan, Weifeng Wang, Zhao Li, Hanwen Yao, Zhiqiang Wang (2020) A review of research on object detection based on deep learning.

[3] Ajeet Ram Pathak, Manjusha Pandey, Siddharth Rautaray (2018) Applications of Deep Learning for object detection.

[4] Sunil, Gagandeep (2019) Study of object detection methods and applications on digital images.

[5] Zhong-Qiu Zhao, Shou-tao Xu, Xindong Wu (2017) Object detection with deep learning: A review.

[6] Jun Wang Tingjuan Zhang, Yong Cheng, Najla Al-Nabhan (2021) Deep Learning for Object detection: A survey.

[7] C.Stauffer and W.Grimson.Adaptive background mixture models for real-time tracking. Proceedings of IEEE Conf.Computer Vision and pattern Recognition,vol.2,1999, 246-252.

[8] Girshick, R., Donahue, J., Darrel, T.,Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In: Computer Vision and Pattern Recognition. Columbus.2014, pp. 580-587.

[9] Ren, S.Q., He, K.M., Girshick, R., Sun, J. Faster R-CNN: towards real-time object detection with region proposal networks. In: Advances in neural information processing systems. Montreal.2016, pp. 91-99.

[10] Liu, W., Anguelov, D., Erhan, D., et al. SSD: Single Shot MultiBox Detector. European Conference on Computer Vision, 2016, pp. 21-37.

[11] Dr. Sabeenian R.S, N Deivanai (2020) Wild animals intrusion detection using deep learning techniques.

[12] Khyati Zalawadia ,Helly Shah ,Radhika Shringarure (2015) Traffic Detection and Diversion System: Case Study.