Object and Character Recognition

A Synopsis Submitted

in Partial Fulfillment of the Requirements

for the Degree of

Bachelor of Technology

In

Computer Science and Engineering

by
Aryan Porwal (1816410072)
Aryan Gupta (1816410074)
Ashutosh Choudhary (1816410078)
Ashutosh Singh (1816410080)

Department of Computer Science & Engineering
Pranveer Singh Institute of Technology
Dr. A.P.J.Abdul Kalam University, Lucknow

Synopsis

1.Introduction

Object detection is a technology that falls under the broader domain of Computer Vision. It deals with identifying and tracking objects present in images and videos. Object detection has multiple applications such as face detection, vehicle detection, pedestrian counting, self-driving cars, security systems, etc.

On the other hand , Optical Character Recognition involves the detection of text content on images and translation of the images to encoded text that the computer can easily understand. An image containing text is scanned and analyzed in order to identify the characters in it. Upon identification, the character is converted to machine-encoded text.

2. Objective

This project is mainly aimed towards helping blind people . With the help of this project we'll be developing an application that would be capable of scanning objects and text from static images or real time input and then read out the scanned text or speak the name of the objects present in the image .

The user would only require to give a voice command to the application and then the application would give a desired output . For instance if the user asks the application to read out the text written infront of him then the application would search the image for any text and if present then speak the output.

The two major objectives of object detection include:

1. To identify all objects present in an image.
2. Filter out the object of attention.

3. Feasibility Study:

Object Detection is the process of finding real-world object instances like car, bike, TV, flowers, and humans in still images or Videos. It allows for the recognition, localization, and detection of multiple objects within an image which provides us with a much better understanding of an image as a whole. It is commonly used in applications such as image retrieval, security, surveillance, and advanced driver assistance systems (ADAS).

Since , the people who are blind are unable to recognize everyday objects and have a difficulty in reading text and sign boards . Therefore , this project will help them to recognize faces of individuals and help them in improving their optical vision .

4. Methodology/ Planning of work

To build our deep learning-based real-time object detector with OpenCV we'll need to (1) access our webcam/video stream in an efficient manner and (2) apply object detection to each frame.
The image is first scanned and the text and graphics elements are converted into a bitmap, which is essentially a matrix of black and white dots. The image is then pre-processed where the brightness and contrast are adjusted to enhance the accuracy of the process.

The image is now split into zones identifying the areas of interest such as where the images or text are and this helps kickoff the extraction process. The areas containing text can now be broken down further into lines and words and characters and now the software is able to match the characters through comparison and various detection algorithms. The final result is the text in the image that we're given.

The process may not be 100% accurate and might need human intervention to correct some elements that were not scanned correctly. Error correction can also be achieved using a dictionary or even Natural Language Processing (NLP).
The output can now be converted to other mediums such as word documents, PDFs, or even audio content through text-to-speech technologies.

5. Facilities required for proposed work

5.1 Hardware Requirements

1. Processors : Intel® Core™ i5 processor 4300M at 2.60 GHz or 2.59 GHz
2. RAM : 8 GB of DRAM
3. Disk space : 1 TB
4. GPU : GTX 1050ti (Mobile)
5. Webcam
6. Laptop

5.2 Software Requirements

1. Operating systems: Windows 10 Pro for Workstation
2. Python 3.6.X
3. Jetbrains PyCharm

6. References

1. https://software.intel.com
2. https://www.geeksforgeeks.org
3. https://medium.com
4. https://stackabuse.com
5. https://towardsdatascience.com
6. https://www.python.org
7. https://pypi.org