

TEXT DETECTION IN AN IMAGE USING PYTHON

A Project Report Submitted By,

Vivek D Naik	4NM19EC199
Manoj S Bhat	4NM20EC409
P Vignesh Prabhu	4NM20EC411
Sampath Kumar	4NM20EC413

Under the Guidance of

Ms. Shubha B.

Assistant Professor

Bachelor of Engineering

in

Electronics and Communication Engineering

from

Visvesvaraya Technological University, Belagavi



Nitte-574110, Karnataka, India June-2022

Abstract

Text detection or recognition in an image is a very useful step to get the recovery of multimedia content. The text detection in an image using python system used to detect the text in an image automatically and remove horizontally associated text with difficult backgrounds. This project is based on colour decrease technique, a technique for edge detection, as well as the localization of text areas and geometrical belongings. Here OpenCV and OCR are used to detect and extract the text in the image. Tesseract engine is used to perform the character recognition system and the pytesseract python package to interact with tesseract in python.

Acknowledgement

Our project would not have been successful without the encouragement, guidance and support by various personalities. First and foremost, we would like to express our sincere gratitude towards our project guide **Ms. Shubha B**, Assistant Professor, Department of Electronics and Communication Engineering, N. M. A. M. Institute of Technology, Nitte, for their guidance, encouragement and inspiration throughout the project work.

We extend our gratitude to **Dr. K. V. S. S. S. Sairam**, Professor and Head, Department of Electronics and Communication Engineering, N. M. A. M. Institute of Technology, Nitte, for his encouragement and providing necessary facilities in the department.

Finally we would like to thank all our friends and well-wishers who helped us whenever needed and supported us.

Vivek D Naik

Manoj S Bhat

P Vignesh Prabhu

Sampath Kumar

Table of Contents

Abstract	ii
Acknowledgement	iii
Table of Contents	iv
1 Introduction.....	1
1.1 General Introduction.....	1
1.2 Aim	1
1.3 Objectives	1
2 Software Description	2
2.1 Software	2
2.2 Visual Studio Code.....	3
2.3 Python.....	3
2.4 OCR.....	4
2.4 OpenCV.....	4
2.4 Tesseract.....	5
2.4.1 pytesseract.....	5
3 Implementation.....	6
4 Results.....	7-8
5 Conclusion	9

1 Introduction

Text detection in an image using python is a project which will help the user to detect and extract the texts automatically in an image which might also include difficult background.

1.1 General Introduction

Text detection and extraction in an image using python is a software based project which will help the user to detect and extract the text automatically in an image even with the difficult backgrounds. This project will help the user to copy the text detected from the software and paste it or note it anywhere possible which will help to save a lot of time as well as energy.

1.2 Aim

To build a code to detect and extract text in an image using python.

1.3 Objectives

- Capturing or saving the image given by the user.
- Recognising and detecting the text and text boundaries in an image
- Extracting the character from the text using OCR
- Obtaining the output

2 Software Description

The libraries that are used in the projects are as follows:

2.1 Visual Studio Code

Visual Studio Code is a lightweight but powerful free source-code editor made by Microsoft for Windows, Linux and macOS.

Edit, build, and debug with ease

Visual Studio Code features a lightning fast source code editor, perfect for day-to-day use. With support for hundreds of languages, VS Code helps you be instantly productive with syntax highlighting, bracket-matching, auto-indentation, box-selection, snippets, and more. Intuitive keyboard shortcuts, easy customization and community-contributed keyboard shortcut mappings let you navigate your code with ease.

Visual Studio Code includes an interactive debugger, so you can step through source code, inspect variables, view call stacks, and execute commands in the console.

VS Code also integrates with build and scripting tools to perform common tasks making everyday workflows faster. VS Code has support for Git so one can work with source control without leaving the editor including viewing pending changes diffs.

VS Code includes enriched built-in support for Node.js development with JavaScript and TypeScript, powered by the same underlying technologies that drive Visual Studio. VS Code also includes great tooling for web technologies such as JSX/React, HTML, CSS, SCSS, Less, and JSON.

Robust and extensible architecture

Architecturally, Visual Studio Code combines the best of web, native, and language-specific technologies. Using Electron, VS Code combines web technologies such as JavaScript and Node.js with the speed and flexibility of native apps. VS Code uses a newer, faster version of the same industrial-strength HTML-based editor that has powered the "Monaco" cloud editor, Internet Explorer's F12 Tools, and other projects. Additionally, VS Code uses a tools service architecture that enables it to integrate with many of the same technologies that power

Visual Studio, including Roslyn for .NET, TypeScript, the Visual Studio debugging engine, and more.

Visual Studio Code includes a public extensibility model that lets developers build and use extensions, and richly customize their edit-build-debug experience.

2.2 Python

Python is an interpreted, object-oriented, high-level programming language. It emphasizes readability and therefore reduces the cost of program maintenance.

Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently where as other languages use punctuation, and it has fewer syntactical constructions than other languages.

- **Python is Interpreted** – Python is processed at runtime by the interpreter. You do not need to compile your program before executing it. This is similar to PERL and PHP.
- **Python is Interactive** – You can actually sit at a Python prompt and interact with the interpreter directly to write your programs.
- **Python is Object-Oriented** – Python supports Object-Oriented style or technique of programming that encapsulates code within objects.
- **Python is a Beginner's Language** – Python is a great language for the beginner-level programmers and supports the development of a wide range of applications from simple text processing to WWW browsers to games.

Benefits that make Python the best fit for machine learning include simplicity and consistency, access to great libraries and frameworks for machine learning (ML), flexibility, platform independence, and a wide community. Hence python 3.8 is preferred for coding machine learning algorithms.

2.3 OCR

OCR (Optical Character Recognition) is the use of technology to distinguish printed or handwritten text characters inside digital images of physical documents, such as a scanned paper document. The basic process of OCR involves examining the text of a document and translating the characters into code that can be used for data processing. OCR is sometimes also referred to as text recognition.

OCR systems are made up of a combination of hardware and software that is used to convert physical documents into machine-readable text. Hardware, such as an optical scanner or specialized circuit board is used to copy or read text while software typically handles the advanced processing. Software can also take advantage of Artificial Intelligence (AI) to implement more advanced methods of Intelligent Character Recognition (ICR), like identifying languages or styles of handwriting.

The process of OCR is most commonly used to turn hard copy legal or historic documents into PDFs. Once placed in this soft copy, users can edit, format and search the document as if it was created with a word processor.

2.3 OpenCV

OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today's systems. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features.

The first OpenCV version was 1.0. OpenCV is released under a BSD license and hence it's free for both academic and commercial use. It has C++, C, Python and Java interfaces and supports Windows, Linux, Mac OS, iOS and Android. When OpenCV was designed the main focus was real-time applications for computational efficiency. All things are written in optimized C/C++ to take advantage of multi-core processing.

OpenCV Functionality

- Image/video I/O, processing, display
- Object/feature detection
- Geometry-based monocular or stereo computer vision
- Computational photography
- Machine learning & clustering
- CUDA acceleration

2.4 Tesseract

Tesseract is an optical character recognition engine with open-source code, this is the most popular and qualitative OCR-library.

It is available for Linux, Windows and Mac OS X. However, due to limited resources it is only rigorously tested by developers under Windows and Ubuntu.

Tesseract is finding templates in pixels, letters, words and sentences. It uses two-step approach that calls adaptive recognition. It requires one data stage for character recognition, then the second stage to fulfil any letters, it wasn't insured in, by letters that can match the word or sentence context.

2.4.1 Pytesseract

Pytesseract or Python-tesseract is an OCR tool for python that also serves as a wrapper for the Tesseract-OCR Engine. It can read and recognize text in images and is commonly used in python OCR image to text use cases.

It is also useful as a stand-alone invocation script to tesseract, as it can read all image types supported by the Pillow and Leptonica imaging libraries, including jpeg, png, gif, bmp, tiff, and others.

3 Implementation

- Import necessary packages and configure pytesseract with the Tesseract engine.
- Reading images using OpenCV.
- Detect text and numbers from the image.
- Detect each character individually.
- Create Bounding boxes over each character in the image.
- Create a Bounding box over each detected text in the image.

4 Results

Input

Input is taken through the camera and there by capturing the image or video in front of it using a button which will take a photo and save it for further process. Figure 4.1. shows the input taken from the image saved

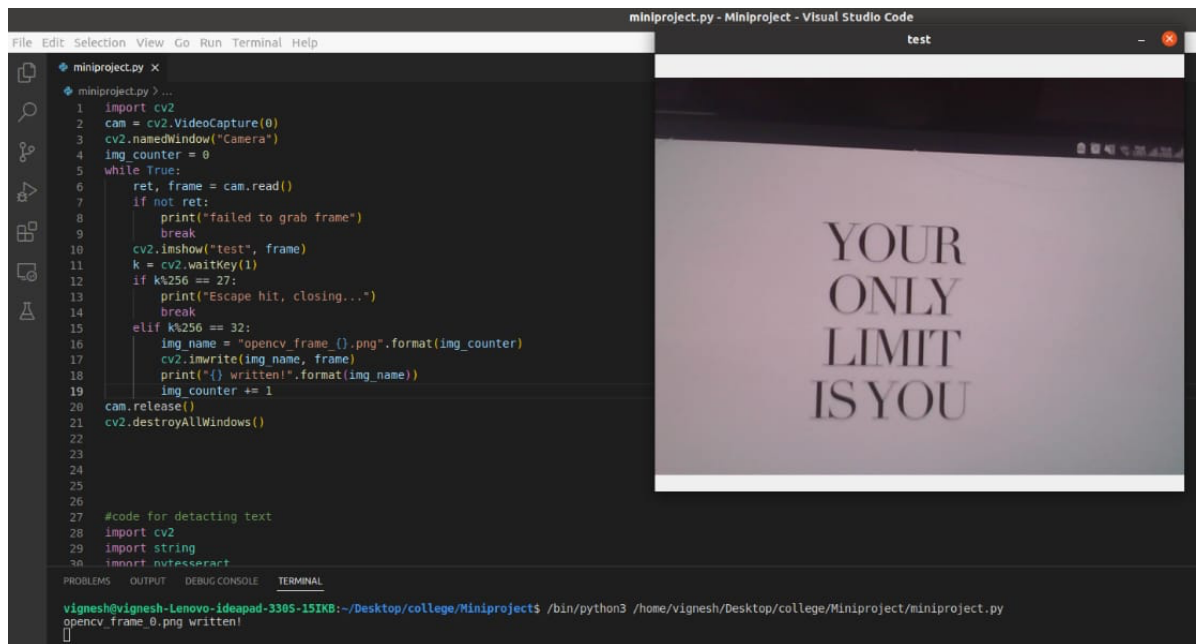


Figure 4.1. Input

Output

The output is the detection and extraction of the text from the saved image. Figure 4.2. shows the output of the project

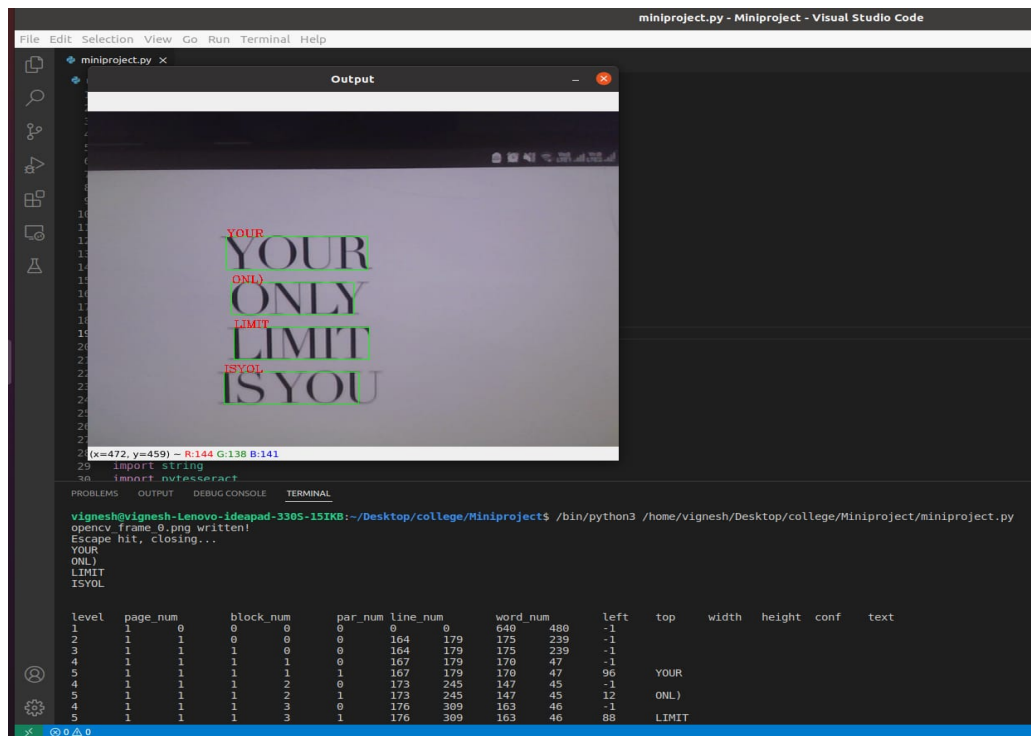


Figure 4.2. Output

5 Conclusion

In this advancing world python is best suited for text detection in the images and also for many major activities. This project is implemented using various packages and libraries of scripting which made it work for low end to end recognition performance. However this approach leaves some problems while processing which can be resolved in the future approach with newer technologies.

Future Scope

This project can also be implemented using various controllers or development boards like raspberry Pi or microcontroller/microprocessors with the real time image as well as video capture and text, object and face detection in it using the camera module, which can be made more accurate with the help of Deep Learning and AI.

References

- [1] Q. Ye and D. Doermann, "Text Detection and Recognition in Imagery: A Survey," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 7, pp. 1480-1500, 1 July 2015, doi: 10.1109/TPAMI.2014.2366765.
- [2] L. Cao, H. Li, R. Xie and J. Zhu, "A Text Detection Algorithm for Image of Student Exercises Based on CTPN and Enhanced YOLOv3," in *IEEE Access*, vol. 8, pp. 176924-176934, 2020, doi: 10.1109/ACCESS.2020.3025221.