IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2022
Impact Factor: 6.252

Real Time Text Detection and Recognition using Pytesseract

Rekha Ahirwar[1], Rashi Saxena[2], Harshal Sunil Sangle[3]
Department of Computer Engineering, S. R. Government Polytechnic College, Sagar, MP, India[1]
Department of Computer Science and Engineering, Ujjain Polytechnic College, Ujjain, MP, India[2]
Department of Computer Technology, Sanjivani K.B.P. Polytechnic, Kopargaon, Maharashtra, India[3]
rahirwar1983@rediffmail.com, rashi8808@gmail.com, harshalsangle12@gmail.com

Abstract: *In this Research paper we are discussing the implementation of Real time Text Detection and Recognition using Google's Pytesseract. The Text detection and recognition in natural as well as in artificial scene images has variety of applications in computer vision systems like license plate detection, automatic street sign translation, image retrieval and help for visually impaired people. Here the important consideration like the Scene text may have complex background, blur image, variations in font-styles, image noise and varying illumination. So Detecting and Recognizing Real- time text is a stimulating computer vision problem. Here we are describing various steps required to extract text from any image file and store the extracted data in new text file. In this we are using CV2 OpenCV library with Python language is used for image processing and Tessaract is used for extracting text from the processed image..*

Keywords: Text Detection, Text Recognition, image processing, Text Extraction, Computer Vision

I. INTRODUCTION

The demand for textual information extraction from many sources has increased dramatically in this era of digitization. Fortunately, recent improvements in computer vision have enabled us to make significant progress in simplifying the challenge of text detection as well as other document analysis and comprehension[1],[2],[3]. Optical Character Recognition is a technique used in Computer Vision to translate text found in photographs or scanned documents into a machine-readable format that can be altered, searched, and utilized for further processing (OCR). Text is something we read virtually every minute of our lives as humans[10].

Wouldn't it be wonderful if our robots or systems could read text in the same manner that we do? But the most pressing concern is, "How can we get our machines to read?" This is when OCR (Optical Character Recognition) comes in handy.[8] In today's commercial world, OCR has a wide range of uses. The following are a handful of them:

- At airports, passport recognition is possible.
- Data Entry Automation
- Recognition of Vehicles number plates
- Adding information from business cards to contacts
- Creating electronic pictures from handwritten documents
- Creating PDFs That Can Be Searched
- Make files that can be heard (text to audio)
- Recognition of Vehicle Number Plates
- Car That Drives Itself

II. AIM AND OBJECTIVE

The goal is to discover the textual region by plotting the bounding box for a given Natural Scene/Image, and then to recognize the identified text.

Text can be in a variety of languages, colors, fonts, sizes, orientations, and shapes in natural scene photographs. We must deal with these texts in images of nature scenes that are more diverse and variable.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2022
Impact Factor: 6.252

Backgrounds with patterns or items with a shape that is highly similar to any text can cause difficulty when detecting texts in Natural Scenes.

Poor latency is necessary to identify, recognize, and translate text in real-time in disrupted images (low quality/resolution/multi-orientation).

III. Available For Text Detection And Recognition

There are several publicly accessible datasets that may be utilized for this work; the different datasets are mentioned below, along with the year of release, image number, text orientation, language, and essential properties.

3.1 Dataset Overview & Description

The Street View Text Dataset is provided by Kai Wang EBU3B, Room 4148 Department of Comp. Sci. and Engr., University of California, San Diego[7].

3.2 Description

The name of dataset is Street View Text (SVT). In this data, image text is highly variable and frequently has low resolution. We notice two features while working with outside street level photos. Business signage frequently contains image text, and business names are easily found using geographic business searches. These features make the SVT set ideal for word spotting in the wild: the aim is to recognize words from surrounding businesses using a street view picture. Our publication, Word Spotting in the Wild, contains further information about the data set[7].

3.3 Conclusions Drawn From Dataset Analysis

In Street View Text (SVT) dataset all the images are of similar sizes(1366x768) and extensions(.jpg). There are 350 photos in the Street View Text (SVT) collection. Because all photos have the same height and width, we don't need to calculate mean height and mean width. All of the words are included in little areas in the majority of the photographs. All of the materials are written in English. The majority of the writings consist of single words rather than characters and phrases, and the words are multi-oriented. We need to create a model that can anticipate the text[7].

IV. Implementation

To obtain the text information or printed text or handwritten text encoded in the scanned copy of the picture, we are utilizing Windows 10 (64 bit) OS with Development Environment containing Python, Pycharm IDE, Tesseract, OpenCV, matplotlib[4],[6],[12].

This python tesseract program is used to detect text in images from a dataset. There are OCR programs available, however the tesseract software written in Python is more efficient. It has a better ability to detect text included in a scanned copy of a picture.

Many programming languages may be used to access Tessaract software. The text encoded in the picture is recognized using this python tesseract program.

This python tesseract program can also identify text that is contained in very huge documents or files. The term "OpenCV" stands for "open source computer vision." A library for image processing that is open source is known as open source computer vision.

This open source library is a machine learning software library that is based on machine learning. Open source computer vision was created to offer the framework for computer vision-based applications. The code of the open source computer vision software may be altered or modified to meet our needs[5],[8],[10]..

This open source computer vision software includes a variety of methods and examples for doing various tasks. In the open source computer vision community, there are over 2,500 algorithms. These algorithms can be used for a variety of tasks, including detecting faces, recognizing faces, identifying objects, differentiating human actions in videos, tracking camera movements, tracking object movements, extracting objects from 3D models, stitching images together to provide a single image containing both images as an entire scene with the highest resolution, and finding images in a database.to remove the excess light, color, flash from the image, to identify the eye movements in a video etc. there are several

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2022
Impact Factor: 6.252

thousands of people who are using this open source computer vision[3][11]. This open source computer vision is hugely used in many companies, government, research people etc. some of the companies that are using this open source computer vision are Microsoft, Google, Honda, IBM, Yahoo, Sony, Intel etc. not only these well established companies, there are various number of startups that are using this open source computer vision library. The startup companies that are using this open source computer vision library are VideoSurf, Applied Minds etc

Thousands of individuals use this open source computer vision to eliminate excess brightness, colour, and flash from images, and to recognise eye movements in videos, among other things.

Many enterprises, governments, and academic organizations utilize this open source computer vision software. Microsoft, Google, Honda, IBM, Yahoo, Sony, Intel, and others are among the firms that use open source computer vision. This open source computer vision library is being used by a variety of startups in addition to these well- established corporations. Video Surf, Applied Minds, and other startups are leveraging this open source machine vision library[14]. As illustrated below, there are three easy stages to follow:

1. Loading a stored image from a computer or downloading one from a browser and then loading it. (Any image including text)
2. Image Binarization (Converting Image to Binary).
3. After that, we'll run the image via an OCR system.

V. Results, Discussion And Conclusions

We tested following images with pytesseact from the mentioned dataset.

Original Text Detection

Image Source Street View Text (SVT) dataset

Original Text Detection

Image Source Street View Text (SVT) dataset

Original Text Detection

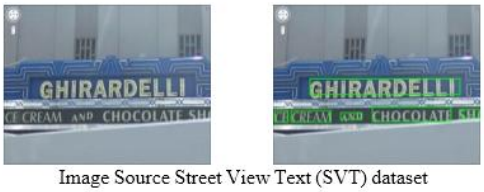

Image Source Street View Text (SVT) dataset

Original Text Detection

Image Source Street View Text (SVT) dataset

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2022
Impact Factor: 6.252

To recognize text in less time, process the images. The accuracy of the acquired data is increased, and mistakes are decreased, thanks to the use of python-tesseract software in text recognition and extraction. Tesseract works best when the foreground text and background text are well separated. In fact, ensuring these sorts of setup may be exceedingly difficult. Tesseract may not provide excellent quality output for a variety of reasons, such as if the picture contains noise in the background. The higher the quality of the image (size, contrast, and lighting), the better the identification result.
To improve OCR performance, some preprocessing is required; photos must be sized suitably, have as much image contrast as feasible, and the text must be horizontally aligned. Tesseract OCR is a strong tool, however it does have certain drawbacks. The following is a list of Tesseract restrictions.[7]

1. The OCR isn't as accurate as some of the commercial solutions we have.
2. Doesn't work well with artefacts such as partial occlusion, warped perspective, or a complicated backdrop.
3. It does not have the ability to recognise handwriting.
4. It's possible that it'll detect nonsense and report it as OCR output.
5. The results may be poor if a document contains languages other than those listed in the -l LANG options.
6. It isn't always successful in determining a document's natural reading order. It may, for example, fail to identify that a document has two columns and attempt to connect text across them.
7. Poor scan quality can result in poor OCR.
8. It does not reveal the font family to which the text belongs.

Our significant objective in this challenge was to find out with regards to Text Recognition and Detection utilizing the Python Tesseract Library. We might utilize a colossal dataset to prepare our models to further develop their presentation. We may likewise utilize an alternate acknowledgment model to further develop text acknowledgment.

ACKNOWLEDGMENT

We would like to express our special thanks of gratitude to Dr. K.V.Arya, Associate Professor, ABV IITM Gwalior.

REFERENCES

[1]. Online:https://searchcontentmanagement.techtarget.com/definition/O CR-optical-character-recognition.
[2]. Online:https://www.anaconda.com/.
[3]. Online:https://sourceforge.net/projects/opencvlibrary.
[4]. Online:https://docs.opencv.org/master/d9/df8/tutorial_root.Html.
[5]. Online:https://matplotlib.org/3.2.1/tutorials/index.html.
[6]. Online:https://www.edureka.co/blog/pythonmatplotlibtutorial/
[7]. Online: http://vision.ucsd.edu/~kai/svt/
[8]. Prasanna Lakshmi, K., Reddy, C.R.K.A survey on different trends in Data Streams ICNIT 2010 -2010 International Conference on Networking and Information Technology, art. no. 5508473, pp. 451- 455.
[9]. Kumar, P., Singhal, A., Mehta, S., Mittal, A.Real-time moving object detection algorithm on high-resolution videos using GPUs (2016) Journal of Real-Time Image Processing, 11 (1), pp. 93-109.
[10]. Padmavathi, K., Sri Ramakrishna, K.Classification of ECG signal during Atrial Fibrillation using Autoregressive modeling(2015) Procedia Computer Science, 46, pp. 53-59.
[11]. Divya, P., Varma, M., & Mouli, U. R. (2021). Web based optical character recognition application using flask and tesseract. Materials Today: Proceedings.
[12]. Thakare, S., Kamble, A., Thengne, V., & Kamble, U. R. (2018, December). Document Segmentation and Language Translation Using Tesseract-OCR. In 2018 IEEE 13th International Conference on Industrial and Information Systems (ICIIS) (pp. 148-151). IEEE.
[13]. Palekar, R. R., Parab, S. U., Parikh, D. P., & Kamble, V. N. (2017, April). Real time license plate detection using openCV and tesseract. In 2017 international conference on communication and signal processing (ICCSP) (pp. 2111-2115). IEEE.
[14]. Rekha, M. (2021). Educational Training For Processing Invoice Of Vendor Identification And Payments Using Python-Tesseract. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(11), 224-228.