iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** V **Month of publication:** May 2020

DOI: http://doi.org/10.22214/ijraset.2020.5155

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

Text Localization and Recognition

Anarghya U [1], Sagar U S [2], Arpitha Rodrigues [3]
[1, 3]*B. Tech. Student,* [2]*Assistant Professor, Dept. of Computer Science and Engineering, Srinivas Institute of Technology-Mangaluru,*

Abstract: ***The text localization and recognition in real time scene text images has become a significant issue in the current scenario. Mobile application and digitization has given a significant and broad impact on real time scene text images. However, the accuracy of recognition rate is based upon the text localization, i.e., higher the clarity of text background segmentation and decomposition, higher the speed of accuracy for the image recognition. In this project, we present a brand new scene text detection algorithm which supports Stroke detection and Hog Transform method. The tactic introduces an approach for character detection and recognition which integrates the benefits of Feature extraction techniques and distribution of Connected Components. Characters are detected and recognized on the image regions and then the image is segmented, then each segment is transformed into a set of connected strokes. The tactic was evaluated on a standard dataset consisting mostly real time images where it achieves advanced results in both text localization and recognition. The results clearly depict the higher accuracy in terms of localization and recognition for real time images.***
Keywords*: *decomposition, localization, recognition, segmentation.

I. INTRODUCTION

OCR is the acronym for Optical Character Recognition. This technology allows to automatically identify characters through an optical mechanism. Eyes are optical mechanism for humans. The input for brain is the image seen by Eyes, the ability to know these inputs differs in all, in line with many factors. Although OCR isn't able to compete with human reading capabilities, OCR could also be a technology that functions like human ability of reading. OCR can recognize both handwritten and printed text. But the performance of OCR is directly dependent on the quality of input documents. OCR is designed to process images that contain texts, with a little non-textual character which is obtained from the picture captured by camera. Microsoft Office Document Imaging (MODI) enables editing and annotating the texts from the documents. Text generated from the OCR can be saved using MODI. However, MODI produces TIFF files that violate some standard specifications. The OCR engine will de-skew and re-orient the page wherever required, in its default mode. Using its built-in OCR engine, it can convert scanned real time images to text under program control.

A. Problem Statement

The major issues in this area are follows: i) the text intensity affected by lighting conditions such as shadows, low and high light which affects the resolution of natural scene text, ii) a variety of text fonts and colors, iii) language of the text, whether it is English or local language. Text characters and strings in natural scene can provide valuable information, which may be useful in data entry. Because of diverse text patterns and variant background interferences, extracting text directly from natural scene images or videos has become difficult. Reading text from photos is a challenging problem which has received a significant amount of attention and efforts have been made in this area. Two key components of most systems are (i) text detection from real time images and (ii) character localization and recognition, and many recent methods and technologies have been proposed to design models for both.

B. Existing System

Converting images of machine-printed characters into machine-readable characters is the function of OCR engines. Bitmap is utilized to extract images of machine-printed characters. An imaging scanner, faxed, or computer generated is also accustomed to scan the forms to produce the bitmap. There are two separate methods in previous OCR system: First extraction of texts from documents and then text to speech translation. Input characters are converted by an optical scanner, in a very typical OCR system. Each character is then located and segmented. Once the segmentation is completed the resulting character image is obtained. The character image is then fed into a pre-processor for noise reduction and normalization. Certain characteristics are the obtained from the character image for classification.

©IJRASET: All Rights are Reserved

The feature extraction is a critical process and many of various techniques exist for feature extraction but each has its own strengths and weaknesses. After classification the identified characters are grouped accordingly. The characters are then accustomed to reconstruct the primary symbol strings, and context may then be applied to detect and rectify the errors.

C. *Proposed System*

Solving this problem has a potential for significant impact. As an example, reading scene text can play a very important role in navigation for automobiles equipped with International Journal of Signal Processing, street-facing cameras in outdoor environments and in assisting a blind man to navigate in certain indoor environments (e.g., a shopping store).
Text Localization is of fundamental importance in image processing and text based retrieval. For example the localization should be achieved before Optical Character Recognition (OCR). We propose effective algorithms of text localization from detected text regions in scene image.

II. LITERATURE SURVEY

1) *End-to-End Scene Text Recognition:* It proposed a method for End-to-End Scene Text Recognition. They focused on the matter of word detection and recognition in natural images. The matter of word detection was significantly tougher than reading text in scanned documents, and had only recently gained attention from the computer vision community. They occupied this gap by constructing and evaluating two systems. The first, representing the actual state-of-the-art, was a two stage pipeline consisting of text detection followed by variety one OCR engine. The second was a system rooted in generic perception. They showed that the last-mentioned approach achieved superior performance. The proposed research started with the target of processing and refining image dataset that we are employing within the proposed framework and algorithm and through this process, following steps and processes evolved [1].
2) *A Robust and Real Time Approach for Scene Text Localization and Recognition in Image Processing:* The method introduces an approach for character detection and recognition which integrates the advantages of Feature Extraction and distribution of Connected Component methods. Characters are detected and recognized on the image regions and then the image is segmented, then each segment is transformed into a set of connected strokes and these strokes are efficiently detected by rotating the image gradient field with a set of oriented bar filters. The method was evaluated on a standard dataset consisting mostly real time images where it achieves better results in both text localization and recognition. The results clearly depict the higher bit of accuracy in terms of localization and recognition for a collected dataset [2].
3) *Text Images with Unsupervised Feature Learning:* They worked over Text Detection and Character Recognition in Scene Images with Unsupervised Feature Learning. They found that reading text from photographs was a challenging problem that had received a big amount of attention. Two key components of most systems were (i) text extraction from images and (ii) text recognition, and lots of recent methods had been proposed to style better feature representations and models for both [3].
4) *Real-Time Scene Text Localization and Recognition:* They worked over Real-Time Scene Text Recognition System for the sequential selection from the set of External Regions (ERs). In the first classification stage, the likelihood of each External region being a character was estimated using novel features calculated with complexity per region tested. Only External regions with locally maximal likelihood were selected for the second stage, where the classification was upgraded using more computationally high-priced features. For each External region a highly efficient exhaustive search with feedback loops was applied to group External regions into words and to select the most likely character segmentation. Finally, text was recognized in an OCR stage trained using synthetic fonts [4].

III. SYSTEM DESIGN

In this paper there are four modules. They are Login, Registration, Scan image and Extract Text.

A. *Login*

Login module is required for authentication purpose. User need to login by giving his valid username and password. If he fails to provide valid details then he/she cannot get access into system.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

1) The Pseudo code for Login Module is

```
Begin
        Input username, password
        If username exists, then
                Check whether password matches
    If password matches, then
      Successful Login
    Else
      Display error message
    End if
  End if
End
```

B. Registration

The Admin creates the account for Staff to check user Transaction. In the Registration form, the Users and Admin can register by selecting the user type accordingly. When submit button is clicked it redirects to the login page.

C. Scan Image

In this module, the image is captured. The captured image is stored as .jpg file. Image is captured using the user selected camera application. The printed paper is held upfront the camera and the image is captured automatically. The image of the printed paper is saved in the system.

1) The Pseudo code for Scan Image module is

```
Begin
  Get video input device
  Set video capture device
  For each video capture device
    Add name to list
  Next
  Set filename to today's date and time
  If Picture Box is not null then
    Set Bitmap
    Pass PictureBox1.Image as parenthesis
    Set file path
    Set file format to jpeg
    Save image (file path, file format)
    Dispose object
  End if
End
```

D. Extract Text

In this step characters from uploaded image is extracted. Extracted characters are temporarily saved in word document which is overwritten every time new characters are extracted from a new image. English language is given along with image as input to the OCR. The OCR technology of Microsoft office document imaging is used in this process.

1) The Pseudo code for Extract Text Module is

```
Begin
  Read file path (.file)
  Get file name (file path)
```

©IJRASET: All Rights are Reserved

```
    Save file to temporary folder
    Set extract text to extract text from image (file path)
  Create object of document
  Create document (file path)
    modiDocument.OCR(Languages=English)
    Read modiDocument.Images[0]
    Set extract text to modiImage.Layout.Text
  Close Document
End
```

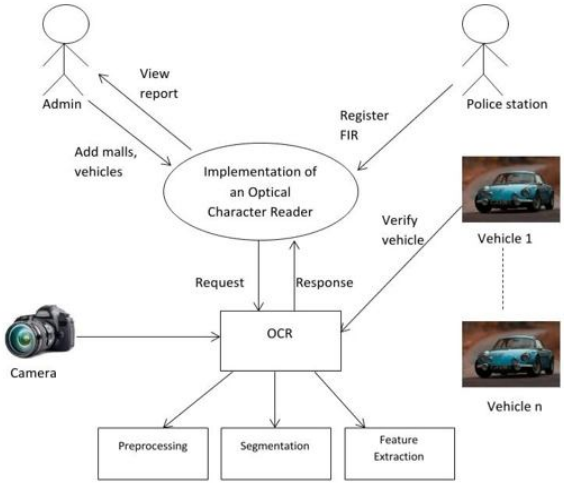


Figure 1: System Architecture of RTSLR

Figure 1 explains that the admin can add malls, the vehicle entering the mall is captured and the vehicle number id detected and captured in the camera then with the help of OCR camera movement detection takes place. The OCR consists of three sections preprocessing, segmentation and feature extraction.

IV. RESULTS

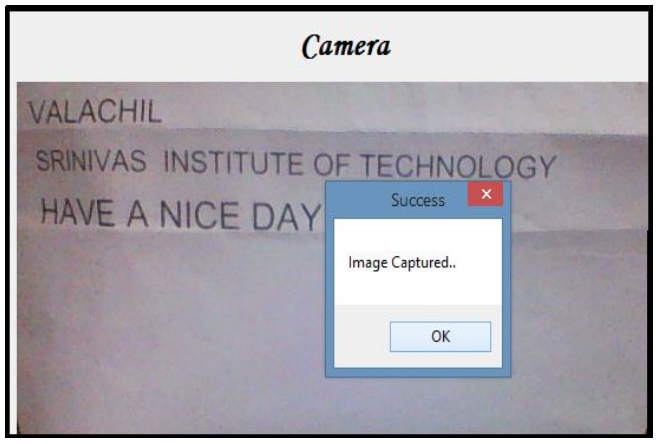


Figure 2: Capturing the Image

Figure 2 displays the capturing of the image. Whenever an image is displayed front of the camera it will be captured automatically.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

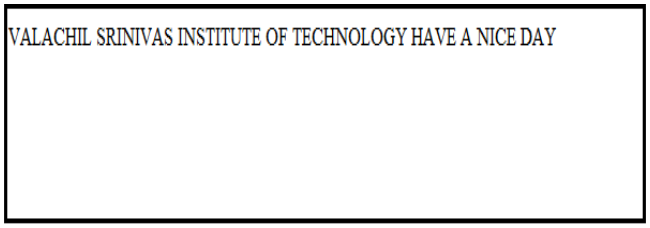


Figure 3: Displaying the Text

Figure 3 shows the extraction of the Text from the Image. First the image is captured using a camera and then from that image the characters are extracted.

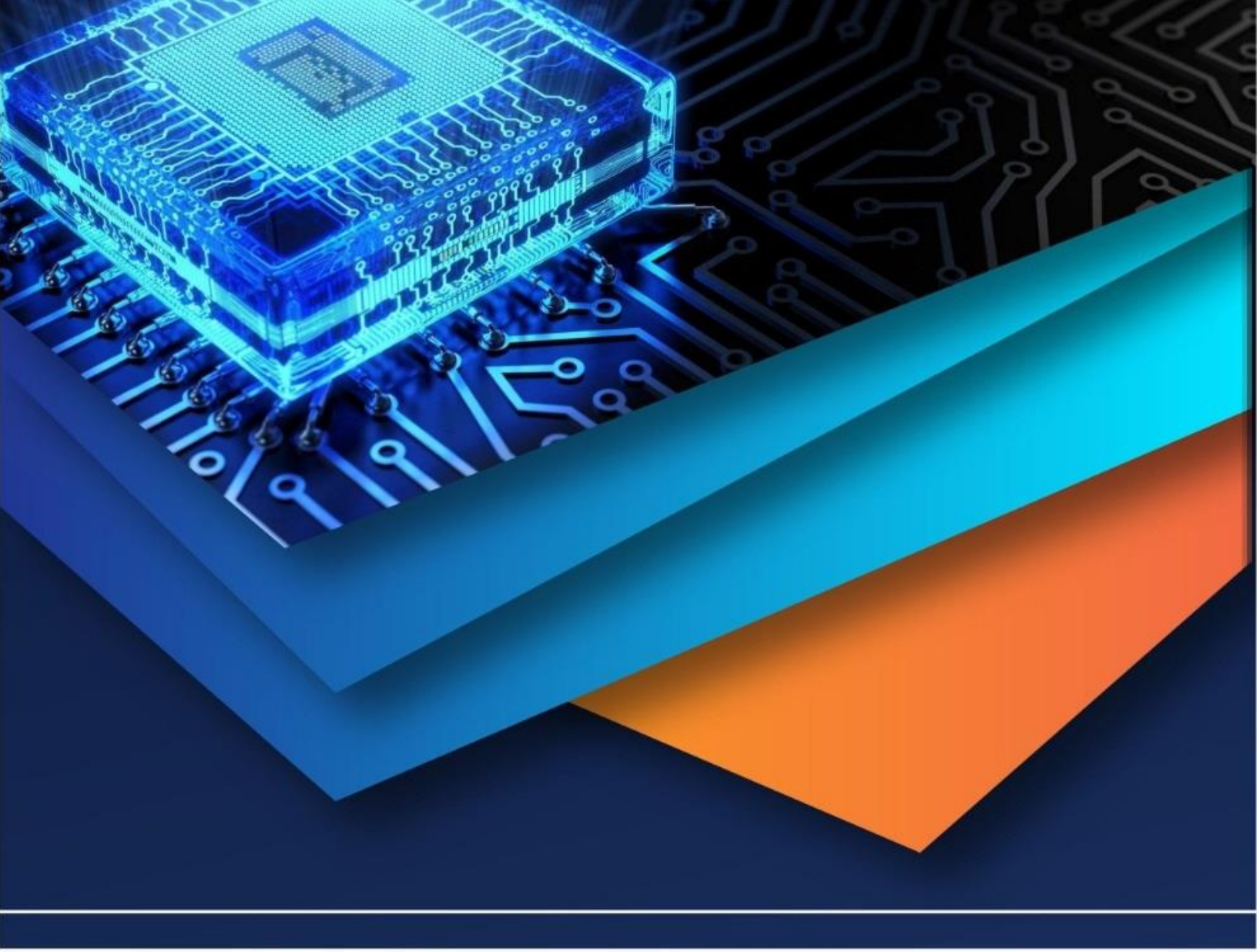
V. CONCLUSION

"Localization and Recognition" is aimed at reliably interpreting text from real world photos which is a challenging problem. The input pattern of a character is read pre-processed feature extracted and recognized, recognized character is displayed. The proposed system is implemented using MODI. Maintaining and getting the contents from and to the books is very difficult. The proposed system can provide paperless environment. It is a user friendly application, it can be accessed by people of varying categories with ease and comfort.

REFERENCES

[1] K. Wang, B. Babenko and S. Belongie, "End-to-End Scene Text Recognition", IEEE, (2011).
[2] Er.Ananta Singh1 and Er. Dishant Khosla. "A Robust and Real Time Approach for Scene Text Localization and Recognition in Image Processing". International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.9, No.1 (2016).
[3] Coates, B. Carpenter, C. Case, S. Satheesh, B. Suresh, T. Wang, D. J. Wu, A. Y. Ng, "Text Detection and Character Recognition in Scene Images with Unsupervised Feature Learning", IEEE, (2011).
[4] L. Neumann and J. Matas, "Real-Time Scene Text Localization and Recognition", International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.9, No.1 (2016).

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)