IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A STUDY ON TEXT DETECTION AND CLASSIFICATION IN NATURAL IMAGES

[1] Lakshmi Aravind,[2] Shabin P

[1] M. tech Student,[2] Assistant Professor

[1] Electronics and Communication Engineering,

[1] College of Engineering Thalassery, Kannur, Kerala, India

Abstract: Text recognition is a major area of experimentation under image processing domain. It is a process by which the system locates any kind of text is present and extract them from an image. The extracted text must be converted to human readable form after several processing and if required is classified them into meaningful classes based on the content present. The platform used here in discussion is MATLAB. This paper provides a detailed study on the evolution of text detection in natural images. It analyzes and also discusses the different methods to overcome existing challenges in text detection. This paper presents the different types of datasets which are used to identify text from natural images and comparative study of different text detection methods. The paper is concluded by a method to recognize and classify the multi-oriented text present in an image based on MSER and CNN.

Index Terms **– MSER (Maximally Stable Extremal Regions), CNN (Convolution Neural Network)**

I. INTRODUCTION

Text present in the natural images comprises of different contents. The need of extraction and retrieval of this content, indexing them is having vast demand. The main aim of content-based classification is to group the images into various meaningful classes based on the content such as Restaurants, Airports, Hospitals etc. Even though there are many works existing related to the text localization,recognition and extraction of text from different kind of images, they lack classification of texts into meaningful classes. The proposed method is used for many applications such as assisting visually impaired peoples, assisting tourists by providing standardized instructions of notice boards and sign boards, Automation in supermarkets and bus terminals, content-based indexing and retrieval in large databases etc.

Challenges in Text Detection The recognition and extraction processes are facing manydifficulties such as multiple orientations, highly illuminated textual content, similar and complex backgrounds, unclear fonts and font style etc. The classification process faces issues like unclear or unknown content and the images which cannot be included to any of the predefined meaningful classes or does not support any of the keyword in certain class attribute.

II. TRADITIONAL METHODS FOR TEXT DETECTION

There are various techniques developed by the researchers for text detection. They are roughly explained below and classified in Fig 2.1

1) Texture Based Method

This method uses the texture-based properties such as Fourier transform, local intensity, filter response and wavelet coefficients for distinguishing the text part and non-text part from the natural images.

2) Region Based Method

This method uses the properties like color, intensity and edge similarity for distinguish the text and non-text part in natural images. It is categorized into three types:

2.1) Edge Based:

This method used the edge detector operator to detect the edges of the images. Usually, two types of edge detection methods are applied such as canny and Sobel edge operator.

2.2) Connected Component Based:

In this method, character components are identified using clustering and edge detectionmethods. Maximum Stable Extremal Region is one of the major techniques of this method.

2.3) Stroke Width:

In this method, text features can be identified through stroke of the components. Character components having constant strokes are treated as text and remaining are treated as non-text. Stroke width transform operator is used for this operation in text detection.

3) Hybrid Method

To overcome the limitations of all the previously mentioned techniques, the combination of two or more techniques is used known as hybrid technique.

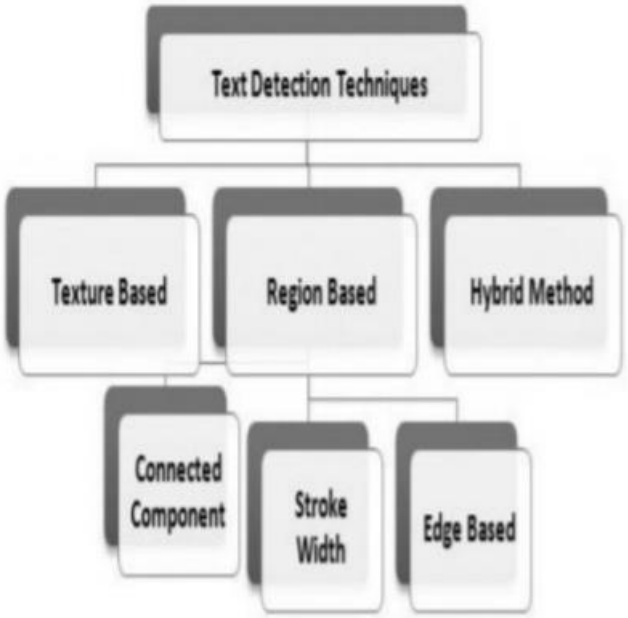

Figure-2.1 Techniques for Text Detection

III. TYPES OF DATASETS

SAMPLE ICDAR 2011 DATASET SAMPLE OSTD DATASET

Figure-3.1 ICDAR 2011 and OSTD Dataset (sample images)

Figure-3.2 ICDAR 2013 Dataset (sample images)

Figure-3.3 MSRA-TD500 Dataset (sample images)

ICDAR 2011 and 2013: It is a subset of QUWI. There are totally 475 images, 4 hand writing images. This is used for recent text detection technique. This dataset is inherited from the previous dataset QUWI. Fig.3.1 and Fig.3.2 shows the sample dataset images.

OSTD: The Oriented Scene Text Database dataset contains 89 images. This can be used for multi oriented text in natural scene images. Fig.3.1 shows the sample OSTD dataset images.

MSRA-TD500 dataset is a text detection dataset that contains 300 training images and 200 test images. Text regions are arbitrarily orientated and annotated at sentence level. Different from the other datasets, it contains both English and Chinese text.

IV. COMPARATIVE STUDY

Earlier numerous methods have been developed for text detection. Text detection was a natural extension of document analysis moving from scanned pages images into camera captured image, preprocessing, detection. Scene text has been regarded as presenting a more difficult challenge yet very little work had been done with it. Recently, researchers have explored approaches that prove effective for text captured in various configurations. The different approaches are SWT (Stroke 2), Convolutional neural network etc.

The Table 4.1 describes a detailed study about the text detection in past and current development techniques. Based on the literaturesurvey the latest technique convolutional neural network provides 93% of accuracy result. This is a better method to apply for producing good result in future.

Table-4.1 Text Detection Techniques Implemented

SL. NO	TITLE	AUTHORS	YEAR OF PUBLICATION	METHOD	CLASSIFIER	DATA SET
1	CRF based text recognition for natural images using CNN and context information	Yanna Wang and Cunzhao Shi	2018	Condition-al Random Field	Convolutional Neural Network	ICDAR 2005,2011, 2013 and SVT
2	Automated text recognition and character recognition in natural scenes based on local image features and contour processing techniques.	Remigiusz Baran and Pavol Partila	2018	MSER and Connected Componen t based	OCR	ICDAR 2017
3	Multi- oriented text recognition from natural scene images based on a CNN and pruning non-adjacent graph edges	Yuanwang and Shen	2018	Exhaustive Segmentati on	CNN	ICDAR 2013 and OSTD
4	Detecting oriented text in natural images by linking segments	B. Shi and Xiang B	2017	SegLink algorithm	CNN	ICDAR 2015
5	Robust text recognition in natural images with edge-enhanced maximally stable extremal regions	H Chen and S Sam	2011	MSER	SWT	ICDAR

V. PROPOSED METHOD

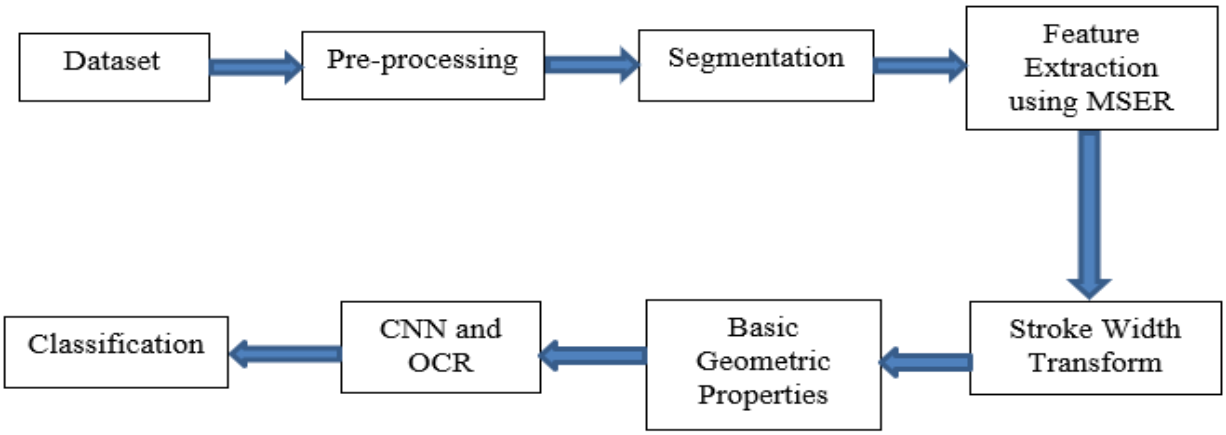

Figure-5.1 Architecture of Text Recognition

Fig.5.1 shows proposed method of text detection, primarily some preprocessing technique is applied on dataset to reduce the noise content. Image converted from RBG to gray scale. The final step of segmentation is to threshold the gray scale images. Thresholding is a process of separating the image into two parts; the background and the foreground. In this work, we have implemented a bounding box method to segment the text content from the whole preprocessed image. This technique is almost similar to Seg Link algorithm. A bounding box is created on the text similar patterns but this may alsocontain non text contents. For that further feature extraction is done on the same. Feature extraction is the process of obtaining interested components from an image as a vector. Applying MSER to those segmented images considerablyremoves less likely textual contents. It works well because most of the text regions will be having uniform and stable intensity profiles. The purpose of applying SWT is to eliminate non-text regions. This is a common matric utilized to distinguish text and non-textual regions by computing the span of the curves and lines. Text regions normally have a compact SW background will have larger disparities.

Although most of the textual area are detected by above mentioned processes, the system has to make sure that the detected component is exactly the text. Thus, we compare those components with the basic geometric properties of the common text. Typically, it is a rule-based method to filter thecomponents of text using uncomplicated threshold values. The different geometric properties that are fine to use in this case are Aspect ratio, Solidity, Eccentricity, Euler number and extend. These properties are calculated by also considering MSER properties. OCR (Optical character recognition) is utilized to gain the individual characters into the actual words

by finding smallest neighboring regions. Classification in this scenario means, the grouping of the detected text and grouping them into meaningful classes based on the content. The advantage of this method grouping the contents for various applications that are useful for the society; that is categorizing the images based on Restaurants, Hospitals, Airports etc. The future scope is that these grouped contents can be utilized and converted to audio to help the visually impaired people. And also, the text detected from messages, conversations, social media comments etc. can be used to detect the emotion of the conveyor.

VI. CONCLUSION

This survey is based on text detection using image processing technique. Different types of difficulties are faced during the time of text detection in from natural images. To overcome the difficulties, different techniques has been introduced: OCR, SVM, CNN etc. These techniques are different types of classifiers to detect text from natural images.

And, also a method is proposed to detect an classify the text from an image. Text recognition from natural images is a vast field of research under image processing. Many datasets like ICDAR, MSRA are available for researches and experimental purposes which is having complexities. The proposed method is one the easy method to extract the content of the images of multiple orientations and to classify them. The system is implanted in MATLAB R2021a.

REFERENCES

[1] Wang, Y., Shi, C., Xiao, B., Wang, C., & Qi, C. (2018). CRF based text recognition for natural scene images using convolutional neural network and context information. *Neurocomputing*, *295*, 46-58.

[2] Baran, R., Partila, P., & Wilk, R. (2018, January). Automated text recognition and character recognition in natural scenes based on local image features and contour processing techniques. In *International Conference on Intelligent Human Systems Integration* (pp. 42-48). Springer, Cham.

[3] Shi, B., Bai, X., & Belongie, S. (2017). Detecting oriented text in natural images by linking segments. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2550- 2558).

[4] Zhang, X., Gao, X., & Tian, C. (2018). Text recognition in natural scene images based on color prior guided MSER. *Neurocomputing*, *307*, 61-71.

[5] Chen, H., Tsai, S. S., Schroth, G., Chen, D. M., Grzeszczuk, R., & Girod, B. (2011, September). Robust text recognition in natural images with edge-enhanced maximally stable extremal regions. In *2011 18th IEEE International Conference on Image Processing* (pp. 2609-2612). IEEE.

[6] P, R., Shamjiith, & K, R. (2020). *Multi-Oriented Text Recognition and Classification in Natural Images using MSER. 2020 International Conference for Emerging Technology (INCET).*

[7] Namrata Choudhary, Kirti Jain. Video Text Detection and Recognition Based on a Transferred DeepConvolutional Neural Network, *International Journal of Computer Sciences and Engineering*.