ISSN 2348 – 9928
IJAICT Volume 4, Issue 11, March 2018 Doi:01.0401/ijaict.2018.03.20 Published on 05 (04) 2018

Robust System for Uprooting Unconstrained Scene Text from Natural Image

Mrs. K. N. Jayapriya
Assistant Professor

Ms. S. Archana
UG Scholar

Ms. B. Greeshma
UG Scholar

Ms. M. Rathika
UG Scholar

Department of Computer Science and Engineering,
VSB College of Engineering Technical Campus,
Coimbatore, Tamilnadu, India

Abstract *– Scene text localization and recognition is still an open problem. Factors contributing to the complexity of the problem include non uniform background. The Scene text recognition is a sub filed of computer vision whose goal is to develop techniques of automatic acquisition of text depicted on photographs taken in the natural environment. It is sometimes called scene text understanding. The objective is to develop a robust scene text detection system to extract the unconstrained text from the natural images. The extracted text can be any form such as plain text or document. It is possible to uproot the text to some other extension. This paper is the first work which proposes to combine both textual and visual cues. The textual cues are extracted by more focusing on background instead of text regions. The regions are processed by Character recognition engine and background detection algorithm. The proposed method is validated on four publicly data sets: ICDAR03, ICDAR13, con-text and Flickr logo.*

Keywords *– Scene Text, Visual Cues, Textual Cues, Data Set*

I. INTRODUCTION

Many existing object recognition methods are focused on distinguishing objects such as horses, bicycles and cars. Object recognition results obtained for different benchmarks show that there has been a significant progress to recognize these distinct object categories. However, the performance of these methods may deteriorate to distinguish categories of objects that only slightly differ in appearance such tasks include fine-grained classification. Fine-grained classification is the problem of assigning images to sub ordinate classes in which objects differ only in subtle details. For example, flower types, bird species. Although, visual cues such as color, texture and shape can be used to distinguish visually distinct objects, the same visual cues may lack discriminative power to differentiate object of similar appearance. Therefore, existing fine grained classification approaches increase the discriminative power of these visual cues by exploiting part information or, implying geometrical constraints. In this paper, we address the problem of fine grained object classification by combining textual and visual cues. Further, we propose a method for textual cue extraction. The success of the proposed fine grained object classification method highly depends on the completeness of the extracted textual image cues. Therefore, robust character localization and a text detection system is proposed.

II. LITERATURE SURVEY

In 2014, Avinash N Bhute and B.B. Meshram et.al proposed text data present in multimedia which contain useful information for automatic annotation, indexing. Extracted information used for recognition of the overlay or scene text from a given video or image. The Extracted text can be used for retrieving the videos and images. This paper reviewed the different techniques for text extraction from images and videos and indexing and retrieval of image and videos by using extracted text.

In 2014, Luka´s Neumann et.al proposed an efficient method for text localization and recognition in real-world images. Thanks to effective pruning, it is able to exhaustively search the space of all character sequences in real time (200ms on a 640×480 image). The method exploits higher-order properties of text such as word text lines. This paper demonstrates that the grouping stage plays a key role in the text localization performance and that a robust and precise grouping stage is able to compensate errors of the character detector. The method includes a novel selector of Maximally Stable Extremal Regions (MSER) which exploits region topology. Experimental validation shows that 95.7% characters in the ICDAR dataset are detected using the novel selector of MSERs with a low sensitivity threshold. The proposed method was evaluated on the standard ICDAR 2003 dataset where it achieved state-of-the-art results in both text localization and recognition.

In 2014, Sezer Karaoglu and Jan C. van Gemert et.al proposed to use text recognition to aid in visual object class recognition. To this end propose a new algorithm for text detection in natural images. The proposed text detection is based on saliency cues and a context fusion step. The algorithm does not need any parameter tuning and can deal with varying imaging conditions. There are three different tasks: 1. Scene text recognition, which increase the state-of-the-art by 0.17 on the ICDAR 2003 dataset. 2. Saliency based object recognition, which outperforms other state of the art saliency methods for object recognition on the PASCAL VOC 2011 dataset. 3. Object recognition with the aid of recognized text, where the first to report multi-modal results on the IMET set. Results show that text helps for object class recognition if the text is not uniquely coupled to individual object instances.

In 2010 Galleguillos and Carolina et.al proposed goal of object categorization is to locate and identify instances of an object category within an image. Recognizing an object in an image is difficult when images include occlusion, poor quality, noise or background clutter, and this task becomes even more challenging when many objects are present in the same scene. Several models for object categorization use appearance and context information from objects to improve recognition accuracy. Appearance information, based on visual cues can successfully identify object classes up to a certain extent. Context information, based on the interaction among objects in the scene or global scene statistics, can help successfully disambiguate appearance inputs in recognition tasks. In this work the problem of incorporating different types of contextual information for robust object categorization in computer vision is addressed. The paper reviews different ways of using contextual information in the field of object categorization, considering the most common levels of extraction of context and the different levels of contextual interactions. It also examine common machine learning models that integrate context information into object recognition frameworks and discuss scalability, optimizations and possible future approaches

In 2013 Rodrigo Minetto and Nicolas Thome et.al proposed snooper text, an original detector for textual information, publicly available, embedded in photos of building facades such as names of stores, products and services that are developed for the iTowns urban geographic information project. Snooper text locates candidate characters by using toggle-mapping image segmentation and character or non-character classification based on shape descriptors. The candidate characters are then grouped to form either candidate words or candidate text lines. These candidate regions are then validated by a text or non text classifier using a HOG based descriptor specifically tuned to single line text regions. These operations are applied at multiple image scales in order to suppress irrelevant detail in character shapes and to avoid the use of overly large kernels in the segmentation. This paper describes two metrics to evaluate the end to end performance of text extraction systems, and show that the use of snooper text as a pre filter significantly improves the performance of a general purpose Optical Character Recognition algorithm when applied to photos of urban scenes.

III. SYSTEM ARCHITECTURE

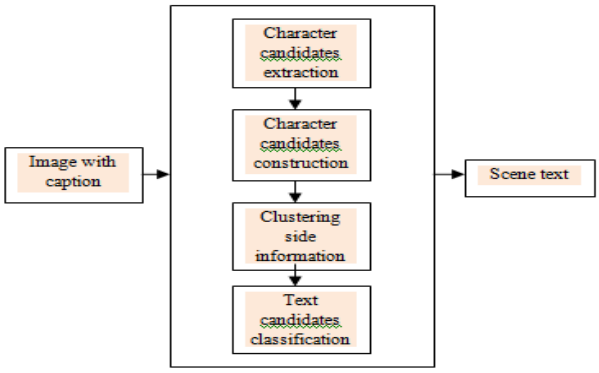

Fig 1 : System Architecture for scene text extraction system

IV. PROPOSED WORK

The proposed work is a robust and accurate Optical Character Recognition based scene text detection method. The OCR and Background detection algorithm relies on a set of learned characters. It compares the characters in the scanned image file to the characters in this learned set. Generating the learned set is quite simple. Learned set requires an image file with the desired characters in the desired font be created, and a text file representing the characters in this image file. A novel self-training distance metric learning algorithm is used to learn distance weights and clustering threshold automatically. Character candidates are clustered into text candidates by the single-link clustering algorithm using the learned parameters. A Character classifier is used to estimate the posterior probabilities of text candidates corresponding to non-text and remove text candidates with high non-text probabilities. Such elimination helps to train a more powerful text classifier for identifying text. Finally, by integrating the above ideas, we build an accurate and robust scene text detection system.

V. CONCLUSION

A method has been introduced to uproot the scene text from the natural language. Scene text localization and recognition

is an open problem since the complex background, poor quality of images. While the state-of-the-art relies on visual cues only, this paper is the first work which proposes to combine recognized scene text and visual cues .To extract text cues, we have proposed a generic, efficient and fully unsupervised algorithm for text detection. The proposed text detection method does not directly detect text regions but instead aims to detect background to infer text location. Remaining regions after eliminating background are considered as text regions. The extracted text will be any form such as plain text or document. This method is also suitable for those input images in which the pixel quality is worse. One of the major applications of text retrieval from images is to extract the information and then recognize its character. This is helpful for indexing the images within storage media.

This method extracts the text in the global language only. In future, we expand the service to some native languages.

References

[1]. R. Arulmurugan and H. Anandakumar, "Early Detection of Lung Cancer Using Wavelet Feature Descriptor and Feed Forward Back Propagation Neural Networks Classifier," Lecture Notes in Computational Vision and Biomechanics, pp. 103–110, 2018. doi:10.1007/978-3-319-71767-8_9

[2]. Haldorai, A. Ramu, and S. Murugan, "Social Aware Cognitive Radio Networks," Social Network Analytics for Contemporary Business Organizations, pp. 188–202. doi:10.4018/978-1-5225-5097-6.ch010

[3]. Haldorai and A. Ramu, "The Impact of Big Data Analytics and Challenges to Cyber Security," Advances in Information Security, Privacy, and Ethics, pp. 300–314. doi:10.4018/978-1-5225-4100-4.ch016

[4]. H. Anandakumar and K. Umamaheswari, "A bio-inspired swarm intelligence technique for social aware cognitive radio handovers," Computers & Electrical Engineering, Sep. 2017. doi:10.1016/j.compeleceng.2017.09.016

[5]. R. Arulmurugan, K. R. Sabarmathi, and H. Anandakumar, "Classification of sentence level sentiment analysis using cloud machine learning techniques," Cluster Computing, Sep. 2017. doi:10.1007/s10586-017-1200-1

[6]. H. Anandakumar and K. Umamaheswari, "An Efficient Optimized Handover in Cognitive Radio Networks using Cooperative Spectrum Sensing," Intelligent Automation & Soft Computing, pp. 1–8, Sep. 2017. doi:10.1080/10798587.2017.1364931

[7]. S. Nandni, R. Subashree, T. Tamilselvan, E. Vinodhini, and H. Anandakumar, "A study on cognitive social data fusion," 2017 International Conference on Innovations in Green Energy and Healthcare Technologies (IGEHT), Mar. 2017. doi:10.1109/igeht.2017.8094075

[8]. H. Anandakumar and K. Umamaheswari, "Supervised machine learning techniques in cognitive radio networks during cooperative spectrum handovers," Cluster Computing, vol. 20, no. 2, pp. 1505–1515, Mar. 2017. doi:10.1007/s10586-017-0798-3

[9]. M. Suganya and H. Anandakumar, "Handover based spectrum allocation in cognitive radio networks," 2013 International Conference on Green Computing, Communication and Conservation of Energy (ICGCE), Dec. 2013. doi:10.1109/icgce.2013.6823431

[10]. Roshini and H. Anandakumar, "Hierarchical cost effective leach for heterogeneous wireless sensor networks," 2015 International Conference on Advanced Computing and Communication Systems, Jan. 2015. doi:10.1109/icaccs.2015.7324082

[11]. S. Divya, H. A. Kumar, and A. Vishalakshi, "An improved spectral efficiency of WiMAX using 802.16G based technology," 2015 International Conference on Advanced Computing and Communication Systems, Jan. 2015. doi:10.1109/icaccs.2015.7324098

[12]. K. Mythili and H. Anandakumar, "Trust management approach for secure and privacy data access in cloud computing," 2013 International Conference on Green Computing, Communication and Conservation of Energy (ICGCE), Dec. 2013. doi:10.1109/icgce.2013.6823567

[13]. Anandakumar, "Energy Efficient Network Selection Using 802.16g Based GSM Technology," Journal of Computer Science, vol. 10, no. 5, pp. 745–754, May 2014. doi:10.3844/jcssp.2014.745.754