



Text investigation for fine-grained object using Context technique

N. Bakya, M. Gayathri, R. Krithika

Department of Computer Science and Engineering,
GKM College of Engineering and Technology, Chennai, Tamil Nadu, India

ABSTRACT

Fine-grained classification using recognized scene text in natural images. In this we extract the text from the image and the extracted text is translated to user known language by using language translator. We apply this method in military services. In this the users create their account by giving their details. Now, the user have their user name and password for their further process. The user sends the image to the end user in encrypted type and they can send document also. Encryption is performed by using RSA algorithm. Now, the end user receive the image and they view the image in decrypted type. The end user extract the text from image. The extraction is performed by using OCR algorithm. We subtract the background by background filtering. Once text regions are detected, it perform text recognition. We used two methods for extraction i.e., character extractor and line extractor. The character extractor generates the bounding boxes of words. Each character is compared with ASCII code for translation. In line extractor, it extracts line by line in the image. The extracted text is translated to user known language by using language translator. The accuracy obtained was 85 to 90 percent.

Keywords: *fine-grained classification, text detection, text recognition, text saliency, language translation*

I. INTRODUCTION

Fine-grained object classification refers to distinguishing among object categories at subordinate levels. Fine-grained classification using recognized scene text in natural images. While the state-of-the-art relies on visual cues only, to combine textual and

visual cues. Another novelty is the textual cue extraction. Unlike the state-of-the-art text detection methods, we focus more on the background instead of text regions. Regions are detected they are further processed by two methods to perform text recognition. Then, to perform textual cue encoding bi and tri grams are formed between the recognized characters by considering proposed spatial pair wise constraints. Finally, extracted visual and textual cues are combined for fine grained classification. The text is detected from the image and the detected text is extracted by the user using optical character recognition. And the extracted text is translated to user known language by language translator.

II. RELATED WORKS

S. Karaoglu, J. C. van Gemert, and T. Gevers, "Object reading: Text recognition for object recognition," in *Proc. ECCV Workshops*, 2012, We propose to use text in natural images to aid visual classification. To detect the text in a natural image, we propose a new saliency method that is based on low-level cues and novel contextual information integration. We show that this saliency method outperforms the state-of-the-art end-to-end scene text recognition. B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform," in *Proc. CVPR*, Jun. 2010, There are several possible extensions for this work. The grouping of letters can be improved by considering the directions of the recovered strokes. This may allow the detection of curved text lines as well. We intend to explore these directions in the future. B. Erol and J. J. Hull, "Semantic classification of business images,"

Proc. SPIE, vol. 6073, pp. 139–146, Jan. 2006. In this paper we presented a novel method for classifying digital camera images captured in a business environment that yields a good performance. Our method is based purely on image analysis. It is possible to use other metadata about an image, such as time and location information of the picture, and the user's calendar to improve the classification results. For example, if a picture is taken during the time the user is scheduled to be attending a conference session, the picture is likely to be a slide or a regular image.

III. PROPOSED SYSTEM

We propose a generic and computationally efficient character detection algorithm without any training involved. Unlike the state-of-the-art text detection methods which try to detect scene text directly, the proposed method detects the background to infer the location of the text. We experimentally show that removing background reduces cluster and subsequently improves the character recognition performance of standard OCR systems. A fine-grained classification approach which combines textual and visual cues to distinguish objects.

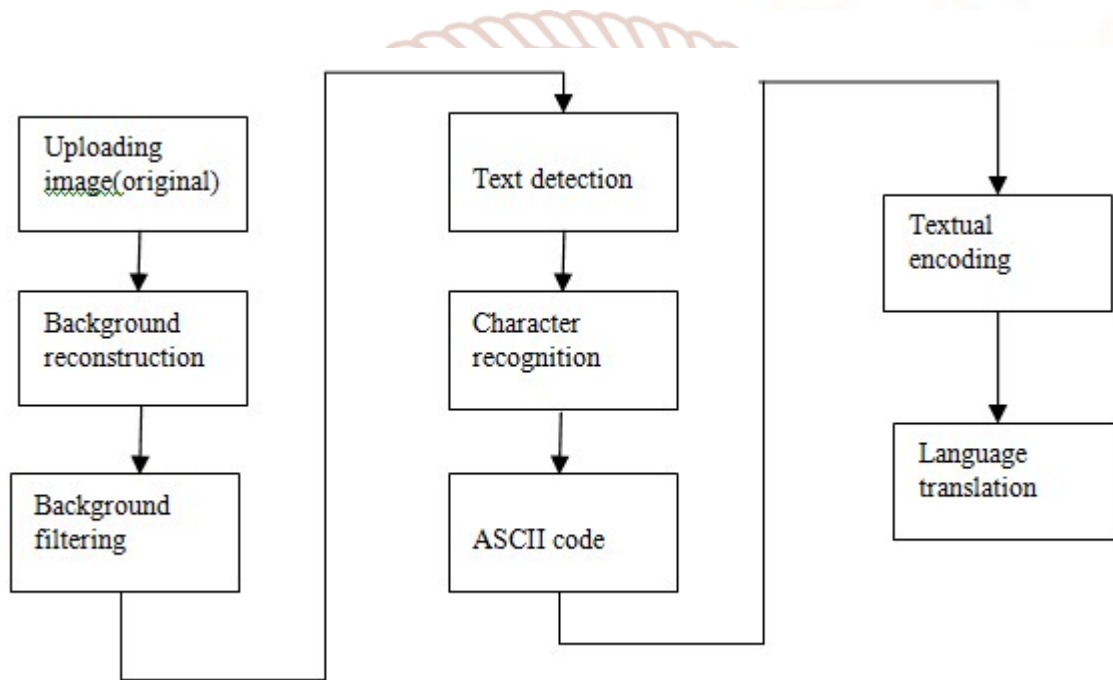


Fig 1: Architecture

A. Authentication

The user register their personal details to create user name and password for their account. The user can also view their registered details. Now, the user is provided with user name and password.

B. Uploading image

The user upload the image and send to the end user. The uploaded image is saved in encrypted type. The user can also upload file or document, this also saved in encrypted type. The encryption is performed by RSA algorithm.

C. Background filtering

Text can appear on unknown background with unknown text size, style and orientation in natural

scene image. Background filtering has two methods i.e., 1. background seed selection and 2. text saliency.

In background seed selection, color boosting approach is used to enhance the saliency of colorful text/background transition and to suppress the background region. Curvature saliency is due to contrast between text and its background, text regions result in high response to curvature saliency even for colorless edge transitions. Spatial context is described by the likelihood of finding an object in certain position to this end the text location priors are used to obtain background location prior. In text saliency, text saliency map is obtained by subtracting the background from the input image. The proposed method outputs a text saliency map which provides information about how likely a region contains text.

This saliency map is further processed to extract textual cues.

D. Text detection

Text detection methods aim at automatically detecting and generating bounding boxes of words in natural scene images. Text detection in images or videos is an important step to achieve multimedia content retrieval. In text detection, it detects the each character in character extractor. Also it detects each line in line extractor.

E. Character recognition

In character recognition method, the detected characters are compared with ASCII code. We use OCR engine to perform character recognition on text saliency. The recognized characters are directly used for textual cue encoding. The output of recognized characters are used to form bi and tri grams without considering their spatial relations.

F. Language translation

A translator or programming language preprocessor is a computer program that performs the translation of a program written in a given programming language into a functionally equivalent program in another computer language (the target language), without losing the functional or logical structure of the original code.

IV. CONCLUSION

A method has been introduced to combine textual with visual cues for fine grained classification .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 for fine-grained classification. 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 region but instead aims to detect background to infer text location. Remaining region after eliminating background are considered as text region. Then, text candidates have been processed by two methods o perform text recognition i.e., OCR engine an state-of-the-art character recognition algorithm. Bi and trigrams have been formed between

the recognized characters by using proposed spatial encoding. Finally the extracted text is translated to user known language.

REFERENCES

1. Wang, K., Babenko, B., Belongie, S.: End-to-end scene text recognition. In: ICCV. (2011)
2. Epshtein, B., Ofek, E., Wexler, Y.: Detecting text in natural scenes with stroke width transform. In: CVPR. (2010) 2963–2970
3. A. Bissacco, M. Cummins, Y. Netzer, and H. Neven, “PhotoOCR: reading text in uncontrolled conditions.” ICCV, 2013.
4. M. Iwamura, M. Tsukada, and K. Kise, “Automatic labeling for scene text database,” in 12th International Conference on Document Analysis and Recognition (ICDAR 2013), Aug. 2013, pp. 1397–1401.
5. Berg and P. N. Belhumeur. POOF: Part-Based One-vs-One Features for Fine-Grained Categorization, Face Verification, and Attribute Estimation. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2013.
6. Y. Movshovitz-Attias, Q. Yu, M. C. Stumpe, V. Shet, S. Arnoud, and L. Yatziv, “Ontological supervision for fine grained classification of street view storefronts,” in *Proc. CVPR*, 2015, pp. 1693–1702.
7. G. Tolias, R. Sircé, and H. Jégou. (2015). “Particular object retrieval with integral max-pooling of CNN activations.” [Online]. Available: <http://arxiv.org/abs/1511.05879>
8. J. Almazán, A. Gordo, A. Fornés, and E. Valveny, “Word spotting and recognition with embedded attributes,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 12, pp. 2552–2566, Dec. 2014.
9. S. K. Ghosh and E. Valveny, “Query by String word spotting based on character bi-Gram indexing,” in *Proc. ICDAR*, Aug. 2015, pp. 881–885.
10. T. Wang, D. J. Wu, A. Coates, and A. Y. Ng, “End-to-end text recognition with convolutional neural networks,” in *Proc. ICPR*, Nov. 2012, pp. 3304–3308.