

A Robust Approach to Detect the Text in Images and Video Frames

Dr. Puttegowda D¹, Bhomika P², Chandana A S³, Krithika G⁴, Srivathsa S Raghavan⁵

¹Department of Computer Science and Engineering ATME College of Engineering, Mysuru, India

²⁻⁵8th semester, Department of Computer Science and Engineering ATME College of Engineering, Mysuru, India

Abstract—Text extraction in images and videos is an imperative role in image study and computer vision. The text detection is the process of identifying the text present in the image and extracting the text. The problems that are faced during the identification of text in image are low resolution image, varying background and many other factors. In this paper, a robust approach is implemented to overcoming the problems mentioned above.

Keywords: Text Detection, Computer Vision, Text extraction, Robust

I. INTRODUCTION

Due to the development of network and multimedia technology, there are plenty of unstructured data on the internet. So the retrieval of information has become one of the key area for research. Since texts in video and images are often very important and it gives rich semantic information to understand the video content. So many researchers are engaged in text detection, localization, and segmentation. For text detection and segmentation, there are numerous approaches are developed, they are divided into two methods traditional and deep learning method. The traditional method involves texture based, regional-based and edge-based approaches. In texture based make use of the difference between background and text region. In the region-based method, it uses color features to distinguish the background and text regions. The edge-based approach identifies the abrupt changes to distinguish the text region from the background. In deep learning method, a convolution neural network is employed for segmentation of text.

1.1 Challenges in Text Detection

From the last 2 decades, ample works have been proposed to identify the presence of text in an image or video. Even though, the text detection process is still inspiring and motivating research area due to the following challenges.

Complex Background: The minimum gap between text pixels and background pixels.

Low resolution: Most of the scene texts are blurred in nature due to the quality of the camera.

Illumination: The variation of lightening effect on scene text in the image or video frames.

Orientation: The text present in an image or video may be in any direction.

Different Font styles: The text present in an image or video may consist of different font shapes.

Different Font size: The headings, captions, and descriptions have different font sizes in an image or video.

Different Font color: The text in an image or video may have mono-color or poly-color.

1.2 Applications of Text Detection

Assisting visually impaired persons: It helps visually impaired persons to walk freely on the road by converting the detected text present in signboards and street names into speech.

Assisting Tourists: With the help of text detection process, the google map uses a global position system to search an interested location of a tourist.

Super Markets: It helps customers to search an item faster.

Indexing: Image or video can be indexed on the basis of textual information.

Retrieval: Textual information helps to fetch the precise from the large database.

Identifying the text is a important part in some of the software, for example let us consider autonomous vehicle, where the car or any vehicle will be driving on its own without any assistance from the driver inside, in this scenario it will be important to observer the signs on the sides of the road which will in text format, here comes text detection in video. When it comes to the field of Computer Science, almost everything is filled with text, for example let us consider a book which has to be converted into a digital copy, it takes many days or months to manually type a book, this problem can be resolved by creating a photo copy the book and then giving the image as an input to the software, so that the text in the image will be extracted and the digital copy can be created within minutes and saves lot of time.

II. PROPOSED METHODOLOGY

Image or PDF will be given as a input to the software. The software after taking the image sharpens the image and video, as the image or video might be of low quality or resolution, to improve the rate of text detection, the image or video will be converted into gray scale and will be blurred to remove the noise from the image, then it undergo sharpening to improve the resolution of image or video. Now as the quality of the image and video has been improved, image can be used for further process. Coming to the text detection process, whole document cannot be processed at a time, Instead each character is identified, after identifying all the characters in a word, moves to the next word in the same line, similarly for the rest of the words in the line, in the same way all the line are processed.

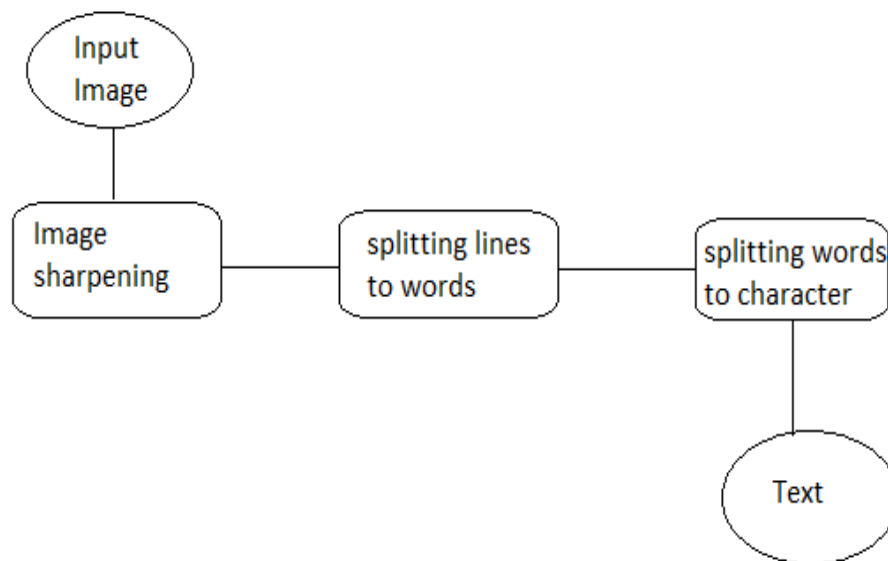


FIG.1. FLOW CHART OF THE PROPOSED METHODOLOG.

III. EXPERIMENTAL RESULTS

Experimentation conducted on a image which has text. The stages are shown in the below figure, fig1 represents the original image on which the experimentation is conducted, fig2 is the stage where the words are identified, after the identification of the word, characters in the words have to be identified, this is represented in fig3, after each text is identified the final result can be seen in fig4.



FIG.1. Original Image.



FIG.2. Image with boundary for word



FIG.3. Image with boundary for single character.

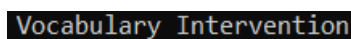


FIG.4. Final Result.

IV. CONCLUSION

In some situations it is necessary to get the text from image or video, the information present in image and video will take lot of time if processed manually, instead if we make use of text detection algorithms, the amount of time will be saved, the above work also represents the same, Image containing text will be given as an input, the image undergoes sharpening, image without noise will be processed and words are identified, in these words, every single character is identified and finally all the characters are combined back to word and will be given as output. Future scope of the project will be based on text detection in videos.

References

- [1] Chen, D., Odobez, J. M., & Bourlard, H. (2004). Text detection and recognition in images and video frames. *Pattern recognition*, 37(3), 595-608.
- [2] Jiang, R. J., Qi, F. H., Xu, L., Wu, G. R., & Zhu, K. H. (2007). A learning-based method to detect and segment text from scene images. *Journal of Zhejiang University-Science A*, 8(4), 568-574.
- [3] Puttegowda, D., & Padma, M. C. (2016, August). Human Motion Detection and Recognising their Actions from the Video Streams. In *Proceedings of the International Conference on Informatics and Analytics* (pp. 12).
- [4] Punith Kumar M B, Dr P S Puttaswamy. (2015 july). Moving Text Line Detection and Extraction in TV video frames.
- [5] Dayananda, K. J., & Puttegowda, D. (2019). A Comprehensive Study on Text Detection in Images and Videos. *Journal of Innovation in Computer Science and Engineering*, 9(1), 17-23.
- [6] Dayananda, K. J., & Puttegowda, D. (2020). Performance Analysis of Color K-Means and Range Filter for Text Detection in Images or Video.
- [7] J. Gllavata, R. Ewerth, B. Freisleben, "Text Detection in Images Based on Unsupervised Classification of High-frequency Wavelet Coefficients", *ICPR 2004, Proceedings of the 17th International Conference on*, Vol. 1.(2004),pp. 425-428.
- [8] Qixiang Ye, Wen Gao, Weiqiang Wang, and Wei Zeng,"A robust text detection algorithm in images and video frames",*information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia, Proceedings of the 2003 Joint Conference of the Fourth International Conference on*, Volume: 2, 15-18 Dec. 2003, pp.802-806.
- [9] Jiang Wu, Shao-Lin Qu, Qing Zhuo, and Wen-Yuan Wang, "Automatic text detection in complex color image", *Machine Learning and Cybernetics, 2002. Proceedings.2002International Conference on*, Volume: 3, 4-5, Nov. 2002, pp.1167 - 1171.
- [10] Xilin Chenjie Yang, Jing Zhang, Waibel, A, "Automatic detection and recognition of signs from natural scenes", *Image Processing, IEEE Transactions*, Vol. 13. Issue.1. (2004), pp. 87-99.