

An Investigation on various Text Information Extraction (TIE) Algorithms for Images

K. Kavın Kumar¹, M. E. Paramasivam², P. Anandan³, P. M. Dinesh⁴, R. S. Sabeenian⁵

¹Assistant Professor, Department of Electronics and Communication Engineering, Kongu Engineering College, Tamil Nadu, India.

²Assistant Professor, Department of Electronics and Communication Engineering, Sona College of Technology, Tamil Nadu, India.

³Associate Professor, Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Tamil Nadu, India.

⁴Assistant Professor, Department of Electronics and Communication Engineering, Sona College of Technology, Tamil Nadu, India.

⁵Professor & Head, Department of Electronics and Communication Engineering, Sona College of Technology, Tamil Nadu, India.

ABSTRACT

Automated text recognitions as of innate pictures has received a wide notice with latent application during picture recovery, robotic as well as intellectual transportation systems. Text data into innate scenery pictures are an vital source of several picture-based application like automatic geo-coding, assistive navigation, contents-based picture recovery as well as scene understanding. Conversely, segmenting texts as of a complicated backgrounds through millions of color are still date challenge to scientist. With the availability of digital cameras with higher resolution, image analysis has become a real-time possibility. Numerous approaches have been developed for extracting text information from camera-captured text pictures as well as the presence of commercialized optical character recognition (OCR) systems have supported a lot. This paper deals with an extensive survey on all the literatures that have been presented till date in this regard of extraction of text from real-time scenes.

Key words: Text Extraction Algorithms, Real-Time Scenes, Survey.

1. INTRODUCTION

Textual information in an image is important objects that usually contain vital information about the viewed image. They often carry the essential messages that need to be understood by the image viewer. Machine intelligence for recognizing the textual information from any image would be of high value. The extraction on textual information from videos with high-level semantics would be of great help to end-users [1]. Vital applications like, aids for visually impaired and understanding display boards for guiding

tourists require an automated system for text extraction. Extraction of text data performs a chief function into information recovery depending upon content from images for easier classification. Automated textual localizing into pictures has been a chief job into system visualization. This paper primarily focuses on quite a few algorithms that are described within the literatures towards the location as well as extraction of non handwritten texts from either a real-time scene or any complex image with complex semantic and perpetual content.

The source images for processing are broadly classified into two types in this literature viz., images that are captured in real-time with more complex semantic and perpetual content, which will called as 'Real-time scenes' here on in this literature and images of scanned documents which contain a major portion of text with less amount of perpetual content, which will be called as 'Document Images' here on in this literature. The complexity of the text extraction algorithms would increase with the increased semantic and perpetual content in an image. A sample of a real-time scene and a document image is represented into the Figure 1.



Figure 1: Sample Pictures of Real-time scene and Document Image

Some of the major factor that affects text extraction algorithms from images is listed below, while most of the listed factors are self-explanatory. A set of images with either of the characteristics is also shown in the Figure 2. a) Shaded

else Textured Backgrounds, b) Low Contrast Images, c) Complex Image, d) Variation in Text Orientation and Alignment, e) Variation in Font Size, f) Variation in Font Style g) Variation in Font Colour



Text on Textured Backgrounds



Variation of Font Style



Variation in Text Orientation & Text in Low Contrast Images Alignment



Variation of Font Size and font Complex Images Colour
Figure 2: Images with variation in characteristics

Some other literatures [8] also classify the characteristic variation of text in images.

2. TEXT INFORMATION EXTRACTION (TIE) ALGORITHMS

An algorithm receives an image and segments text available into the images as well as gives this as an output. The input images can be either a gray scale or colour, un-compressed or compressed, or graphic images with texts moving in the image. The challenge of any TIE algorithm can be broadly classified as: (i) image detection, (ii) text extractions and enhancements, (iii) text localization, (iv) text tracking, and (v) character recognition (OCR). Text detections, text localizations, as well as text extractions have been frequently utilized interchangeably into different literatures.

(i) Text Detection

Text detection identifies the presence of texts within each image. In general text detections are utilized systems that are involved in a real-time application. Otherwise, only images with text are fed to the algorithm.

(ii) Text Tracking

Text Tracking performs a vital function in textual identification into videos. The localization of texts in every

adjacent frame is one important perspective in TIE Algorithms.

(iii) Text Localization

Text localization represent towards identifying the region of interest into a specified picture which has texts within. In certain literatures, for human and machine convenience, boxes are generating around the text.

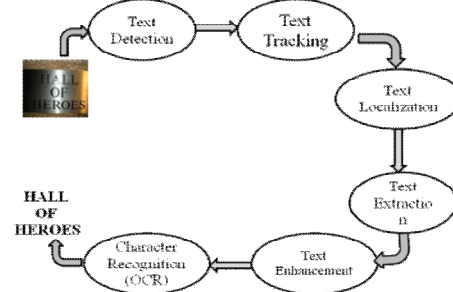


Figure 3: Architecture of a TIE system

(iv) Text Extraction and Enhancement

Despite identification of text in an image, the texts even now need segmentation as of the backgrounds for facilitating their recognitions. In certain cases, the background is more textured or complex and needs to be more processed for converting the text to a binary image.

(v) Optical Character Recognition (OCR)

Optical Character Recognition (OCR) has been a technology where the given input text in the form of binary image is converted into its equivalent printed text. There are a number of algorithms, both commercial and open source algorithms available for OCR and a detailed discussion.

3. TIE ALGORITHMS IN REAL-TIME SCENES

In certain cases, building windows and fences give an impression such that text-like patterns are portrayed. These lead to misclassification as texts. Much different from document images, innate scenic image embeds text into random shape, size as well as orientation in multifaceted backgrounds like depicted into figure 2. This has been highly tedious for recognizing texts into innate scenic image honestly developed TIE algorithms couldn't tackle complicated backgrounds interference as well as non-oriented texting line. An extensive literature research has been carried out in [8], in the year 2003 after which lots of task are carried out within these areas which has been discussed into detail below.

Manmatha and et.al [8] have proposed a four-level system where it could robotically identify as well as dig out texts as of image in [1]. First, the texture segmentation approach is used to focus on regions where the text are present in an image. Strokes present in the image are extracted from the segmented text regions and the backgrounds are cleaned by binarizing the detected text. Lastly, a better bounding box is put up around the text for easier identification. Jain and et.al [2]

proposed text location approach using Connected Component Analysis (CCA), a classical approach. A spatial structure analysis for colour connected components is being used in CCA and can be widely used for book covers, news titles and video captions. Serra [3] introduced a toggle mapping approach, which is a morphological operator. The approach was initially utilized towards dissimilarity development furthermore for noise lessening, was being used into text segmentation in natural scenes by Fabrizio [12]. This approach reduces the number of segmented regions thereby reducing the time of computation.

A text string can be designed through an edge on a rude level signifying a skeleton of qualities. Skeletons edge suits geometric as well as structural limitations like perpendicular else non-parallel towards the edge. Based on these techniques Tran [4] proposed a method, which never depends upon a specific alphabets and also on the orientation of the text string. The proposed method has two stages, wherein first ridges in scale space are identified and then classifying the ridges into text or non-text. In [5], the primary signal of texture region is initially identified as of the detained image and an adaptable colour framing as well as search algorithms are used for discriminating texts and non-texts region. The algorithm proposed within the article is a combination of benefits of numerous earlier methods towards text recognition.

Datong and et.al [6] modeled the textual image like a combination of Gaussian as well as assigned every pixels of a Gaussian Layer. An online estimation of co-efficients using Markov Random Field based on prior contextual information was done. Every layers are after that progressed by a connected element analysis as well as sent towards the Optical Character Recognition systems. Rui and et.al [7] have discussed on picture segmentations depending upon spectral clusterings. They have used a standardized graph cuts like a measures of spectral clusterings as well as weighted matrices for assessing the graphs cut, which in turn gives a lower computational complexity of the algorithm.

An adaptable regioning layout study should position texts into a localized region was proposed by Qixiang and et.al in [10]. A combination of features such as colour, structural layouts, textures as well as statistic upon Optical Character Recognition results have been employed into the area locations, candidates textual locations, texting verifying phases for text detection. The work has used a simple homographic function among 2 planes for restoring the text's place of in-law deformity without affecting the source parameters.

In [11], Kumar and et.al have utilized worldwide linked wavelet filter to extract text regions in a picture. The work was in turn extended to even document images for segmenting text, background and picture components by making use of a multiple, two-class Fisher Classifier. Later on the contextual information in a image was exploited with

the help of a Markov Random Field formulations depending upon pixel's label method to refine the segmented outcomes.

To overcome the drawbacks of the existing binarization methods, an edge-based connected component approach had been projected in Kasar, et.al within [12]. The algorithm projected had the capability of handling documents with multi-colored text with different backgrounds and text of varying sizes which cannot be usually handled by binarization methods. Pan and et.al [13] have made use of a region based method that uses connected component analysis along with Markov Random Field (MRF) Model for text extraction. The algorithm broadly classifies the process of text detection as pre-processing and region analysis, and text localization as text line generation and text extraction. The region analysis technique introduces a new method of feature integral map generation and windows each and every text in the image.

Most of the literatures deal with features such as coloring of the texts, sizes, fonts as well as backgrounds litter for segmentation. An attempt was made in [14] for considering the inherent features of the texts with the help of Stroke filters for local region analysis which is in turn used for determination of texts coloring polarization as well as localized area budding processes. Finally the comment process as of an Optical Character Recognition has been employed for improving the performances of the texts segmentations. A stroke filter [21] generally could eliminate texts candidate that has stronger borders however they don't have texts. In addition on still applying the stroke filters, the texts can be localized in a given image for further processing. Canny Edge operators are used for retrieving the edge information present in an image. The edges retrieved so are grouped into connected components by Wumo in [15]. Every linked components are then classified to texts else non-texts connected component based with making use of picture element leveled label as well as linked components label. Layout analysis has been additionally executed for verifying the text's candidate.

Angadi and Kodabagi [16] have projected a text dependent methods that utilizes DCT dependent higher pass filters for removing the backgrounds component, which is later split into blocks for identifying potential blocks using newly defined functions. The work specifically focuses on identifying regional Indian text viz. Kannada when compared to English texts. Manik and et.al. [17] has specifically focused on cases where in the OCR techniques are not capable of handling the real time scenes. The work focuses on both English and Kannada Text by having two sets of datasets. The performance of the various features is assessed based on nearestneighbour and SVM classification. Bhattacharya and et.al proposed a novel approach within [18], where the concept of connected component analysis is done to identify Devangari and Bangla Text. Also they have carried out the binarization of scene texts with a global thresholding for effective text extraction.

As proposed in [19], text's assurance maps are designed depending upon the text's region detector where the textual component is separated through localized binary method. A conditional random field (CRF) method has the unary and binary neighboring element relationship has been utilized for label "text" as well as "non-text" elements. An energy minimizing method has been utilized for grouping "text" components to text line.

In [20], a Markov Random Field (MRF) and Belief Propagation (BP) concepts has been presented for segmenting handwritten texts, apparatus printed texts as well as noises as of interpreted device printed document. Three level wording qualities have been extorted using a customized K-means cluster algorithms.

Bouman [23] in his literature has made use of the method of uniform luminance and texture to detect the background in the image. For performing it, the images are separated to $k \times k$ non-overlap block, such that k has a power of 2. Once the homogeneous blocks are identified, the text background is segmented separately and text alone is discriminated. Also a process of false background rejection has been discussed in detail to prevent any loss of text from the image. The work has been carried out by making use of Nokia N800 Internet Tablet

Based on the difference on morphological reconstruction and gamma correction a binarization algorithm has been designed for extracting the linked elements of a picture is projected in [24]. The linked component is then categorized into texts and non-texts by means of Random Forest Classifiers. Finally a novel merging algorithm is used for localizing text regions and for further processing.

4. SCOPE OF RESEARCH

An extensive research can be carried out in detecting all the local regional languages. Only certain literatures have dealt with extraction of text such as Bangla and Devanagiri. It is very much evident on an extensive search carried out in this area, that there has been no work carried out in Classical Tamil Text extraction from real-time images. Only a recent work on Handwritten text and with the classical Tamil language using statistical features have been developed by [25, 26], which specifically deals on handwritten text rather than the text from images [27]. On development of better text extraction algorithms for regional text would help to create a cheap PDA device that would in turn enable better tourism in the region [28]. More over even creating a test data set would be an extensive task for evaluating the algorithms developed in extracting the text.

5. CONCLUSION

This paper has provided a comprehensive survey of various text information extractions that can be use for real-time images. For more than a decade, several algorithms have been

presented by scientist. However, each one of them has tried to provide solution for a specific problem. The same approach when implemented on a different problem pattern does not provide satisfactory results due to variations in character size, fonts, texture, colour and background too.

Despite the availability of a large number of public datasets, most of the algorithms proposed in this paper fail occasionally for extracting the texts as of the scenic pictures. Scientist are working to investigate the localization of text on images and also to track text from videos.

REFERENCES

1. Victor Wu, R.Manmatha, Edward M.Riseman, "Text Finder: An Automatic System to Detect and Recognize Text in Images", 1997.
2. Anil K.Jain and B.Yu, "Automatic Text location in images and video frames", Pattern Recognition 31 (1998), pp. 2055-2076.
3. J.Serra, "Toggle Mapping", From Pixels to Features, pp 6172, 1989, J.C. Simon (ed), North Holland, Elsevier.
4. H.Tran, A.Lux, H.L.Nguyen T and A.Boucher, "A novel approach for text detection in images using structural features", in the proceedings of 2010 International Conference on Digital Image Computing: Techniques and Applications (DICTA), pp 635 – 642, DOI: 10.1109/DICTA.2010.115.
5. Jiang Gao and Jie Yang, "An Adaptive Algorithm for Text Detection from Natural Scenes", Proceedings of Computer Vision And Pattern Recognition (CVPR), 2001, pp 84-89.
6. Datong Chen, Jean-Marc Odobez and Herve Bourland, "Text Segmentation and Recognition in Complex Background based on Markov Random Field", 16th International Conference on Pattern Recognition, 2002. Proceedings, 2002, pp: 227 – 230, vol.4, DOI: 10.1109/ICPR.2002.1047438.
7. Rui Wu, Jianhua Huang, Xianglong Tang and Jiafeng Liu "A text image segmentation method based on Spectral Clustering", Journal of Computer and Information Science, Vol.1, No.4, November, 2004, pp 9-15.
8. Text information extraction in images and video: a survey, K. Jung Pattern Recognition, Vol. 37, No. 5. (May 2004), pp. 977-997. doi:10.1016/j.patcog.2003.10.012.
9. Datong Chen, Jean-Marc Odobez and Herve Bourland, "Text detection and recognition in images and video frames", Patter Recognition 37 (2004), pp 595-608.
10. Qixiang Ye, Jianbin Jiao, Jun Huang and Hua Yu, "Text Detection and restoration in natural scene images", Journal of Visual Communication and Image Representation, 2007, DOI: 10.1016/j.jvcir.2007.07.003.
11. Sunil Kumar, Rajat Guptat, Nitin Khanna, Santanu Chaudhury and Shiv Dutt Joshi, "Text Extraction and Document Image Segmentation using Matched Wavelets

- and MRF Model”, in the IEEE Transactions on Image Processing, Vol. 16, No. 8, August, 2007, pp 2117-2128.
12. T.Kasar, J.Kumar and A.G.Ramakrishnan, “Font and Background Color Independent Text Binarization”, Intl. workshop on Camera Based Document Analysis and Recognition (workshop of ICDAR), pp. 3-9, 2007.
 13. Yi F. Pan, Xinwen Hou & Cheng L. Liu, “A Robust System to Detect and Localize Texts in Natural Scene Images”, Proceedings of the 2008 The Eighth IAPR International Workshop on Document Analysis Systems (2008), pp. 35-42.
 14. Cheolkon Jung, Qifeng Liu and Joongkyu Kim, “A new approach for text segmentation using stroke filter”, Elsevier Signal Processing, 2008, pp 1907-1916.
 15. Wumo Pan, T.D.Bui and C.Y.Suen, “Text Detection from scene images using sparse representation”, 2008, 19th IEEE International Conference on Pattern Recognition, ICPR – 2008, pp 1-5.
 16. Angadi, S.A.; Kodabagi, M.M, “A Texture based Methodology for Text Region extraction from low resolution natural scene images”, International Journal of Image Processing, Vol 3, Issue 5, pp 229-245, November 2009.
 17. Teofio E.De Campos, Bodla Rakesh Babu and Manik Varma, “Character Recognition in Natural Images”, in the proceedings of International Conference on Computer Vision Theory and Applications, VISAPP, 2009.
 18. U.Bhattacharya, S.K.Parui and S.Mondal, “Devanagari and Bangla Text Extraction from Natural Scene Images”, in the Proceedings of the 10th International Conference on Document Image Analysis and Recognition, 2009, pp 171-175.
 19. Yi-Feng Pan, Xinwen Hou, Cheng-Lin Liu, “Text Localization in Natural Scene Images based on Conditional Random Field”, in the proceeding of the 10th International Conference on Document Analysis and Recognition ICDAR ‘09, 2009, pp 6-10.
 20. Xujun Peng, Srirangaraj Setlur, Venu Govindaraju, Ramachandrula Sitaram and Kiran Bhuvanagiri, “Markov Random Field Based Text Identification from Annotated Machine Printed Documents”, in the proceedings of 10th International Conference on Document Image Analysis and Recognition, 2009, pp 431-435.
 21. Cheolkon Jung, Qifeng Liu and Joongkyu Kim, “A stroke filter and its applications to text localization”, Pattern Recognition Letters 30, (2009), pp 114-122.
 22. J.Fabrizio, B.Marcotegui and M.Cord, “Text Segmentation in Natural Scenes using Toggle-Mapping”, in the proceedings of ICIP, 2009, pp 2373-2376.
 23. Bouman, K.L., Abdollahian, G., Boutin, M., Delp, E.J., “A low complexity method for detection of text area in natural images”, 2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), 14-19 March 2010, pp 1050 – 1053.
 24. Sezer Karaoglu, Basura Fernando and Alain Tremeau , “A Novel Algorithm for Text Detection and Localization in Natural Scene Images”, 2010 International Conference on Digital Image Computing and Applications (DICTA) – 2010, pp 635 – 642.
 25. Sabeenian.R.S., “Digitalization of Handwritten Classical Tamil Language using First Order and Second Order Statistical Features”, International Journal of Advanced Engineering Informatics 36, (2011), pp 3584 – 3586.
 26. Sabeenian.R.S., and Vidhya.M., “Hand written Text to Digital Text Conversion using Radon Transform and Back Propagation Network (RTBPN)”; Information and Communication Technologies Communications in Computer and Information Science, Springer Berlin Heidelberg, 2010, 498 – 500, 10.1007/978-3-642-15766-0_82.
 27. Yazan Alaya AL-Khassawneh, “The use of Semantic Role Labelling with Triangle-Graph Based Text Summarization”, International Journal of Emerging Trends in Engineering Research, Volume 8. No. 4, April 2020, pp.1162-1169.
 28. M.Govindarajan, “Ensemble of Classifiers in Text Categorization”, International Journal of Emerging Trends in Engineering Research, Volume 8. No. 1, January 2020, pp.41-45.