International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 4 Issue: 4 507 - 516

# Text Recognition Past, Present and Future

Amandeep Kaur
CTIEMT, jalandhar
*Amandeepk750@gmail.com*

Manju Bala
CTIEMT, Jalandhar

Pooja
CTIEMT, jalandhar
*poojachoudhary80@gmail.com*

***Abstract*—** Text recognition in various images is a research domain which attempts to develop a computer programs with a feature to read the text from images by the computer. Thus there is a need of character recognition mechanisms which results Document Image Analysis (DIA) which changes different documents in paper format computer generated electronic format. In this paper we have read and analyzed various methods for text recognition from different types of text images like scene images, text images, born digital images and text from videos. Text Recognition is an easy task for people who can read, but to make a computer that does character recognition is highly difficult task. The reasons behind this might be variability, abstraction and absence of various hard-and-fast rules that locate the appearance of a visual character in various text images. Therefore rules that is to be applied need to be very heuristically deduced from samples domain. This paper gives a review for various existing methods. The objective of this paper is to give a summary on well-known methods.

***Keywords***-*Text-Detection,Text-Localization,Text-Recognition,text-segmentation*

*****

## I. INTRODUCTION

In the recent time the amount of capturing digital images and digital videos has increased through internet. Images contain the text also that is very important feature of the images and also the text in the images has sometimes the important information or data. Mostly the data in the images might not be read easily, so if the work is done to detect, extract and recognize the text information, it might be very beneficial for the computer world ,due to this the recognition of image text from various text images has become very prominent topic these days. Basically the images are divided into following types-

Scene images- Scene images include the text of scene images, such as the advertising boards in market, banners in various areas, which is taken naturally somewhere by the camera, therefore scene text is assumed to be embedded in the background as a part of the scene. Scene images are considered very complex because the backgrounds of captured images are complex containing the text of different sizes, styles and alignments, sometimes the resolution of the images is very low.

Born digital images- Born-digital images are generated by computer and are named as well as saved as digital images. As Compared to document images and scene images, it is considered that there are more defects in born digital images, like more complex background, sometimes low resolution, and compression loss.

Document images- Document images are the image-format of the existing hard copy of document. Document images are the area of the text extraction in its early stage.

Text can play an important role in analyzing street view images which might contain text. Large amount of work have been done to recognize text in various types of images. Text detection and Recognition also extraction in videos can also help in video content analysis since text can give concise and direct explanation of the stories contained by the videos. In digital news videos, the superimposed captions mostly show the involved person's name and the summary of the news event. So, the recognized text can be a part of index in a video text retrieval or recognition system.

## II. SOME METHODS FOR TEXT DETECTION AND RECOGNITION

### *A. Method-A*

**Based on born digital images using conditional random field**

The extraction process has been explained in this paper on the basis of CRF (conditional random field)
Steps for extracting the text

IMAGE SEGMENTATION

Binarization is a preprocessing that is done for the text extraction and recognition. N Wavelet filtering is one of the important applications of wavelet that can be applied to the image processing. Binarization can be done through following steps

1. Conversion of color image to binary image
2. Extraction of appropriate background images
3. Extraction of appropriate foreground images
4. Global thresholding
5. Local thresholding
6. Binarization

TEXT EXTRACTION

 ISSN: 2321-8169
507 - 516

Text extraction can be done with the help of conditional random field model. After that the CC analysis is done by applying CC extraction and CC adjacent graph. When we got results from the CC analysis the feature extraction is done with the help of unary component features and binary component features. Finally the designing of CRF classification is done

*B. Method-B*

**Based on complex images**

This method works on the scene text images. Scene text images are those which basically contain the data like advertising boards, banners, posters which are naturally taken by the camera. These images are considered complex because background is complex which mostly contains components except text like buildings, trees, vehicles and also the alignment of the text is with different colors, styles, size somewhere the resolution may also be low. Text extraction and recognition is full of challenges. Method of detection, segmentation and recognition has been described here.

- Some general methods for detecting text are given below
    1. CC based method
    2. Edge based method
    3. Color based method
    4. Texture based method
    5. Corner based method

- Segmentation of text

The separation of pixels from the background in done for the extraction process. MRF markov random field model can be used here, a energy function is taken by which the quality of binarization is determined. Minimization of energy function is done for the optimal binarization with the help of graph cut scheme. A new conditional random field model was designed by combining the features of color, edge, stroke and the relationship between the. Another method was designed on the basis of two assumptions first was color of text pixel which work through Gaussians distribution and second is local portion of the image that has same color distribution as global portion.

- Recognition of characters

Use of multiple-size sliding sub window the automatic recognition of characters using segmentation is done. First the template of the image is done through customized program, after that the matching of the template with the extracted SIFT feature id applied.

*C. Method-C*

**Based on scene text images for large lexicons**

With large amount of work going on the text recognition, here the focus is on the lexicons which reflect the image specific list of words. The basic conditional random field model is defined for initializing the recognition process. It is started by generating a set of candidate words which gives M-best diverse solutions as a result. With potential solutions given by applied algorithm, refinement of the very large lexicon is done by removing words from it and deducing its size which has edit distance of the candidates generated, then re-analyze the M-best diverse solutions of the input text image which is a result of applied algorithm. These steps can be applied number of times. Once previous algorithm has been applied a expected solution can be collected through various sources like with the help of minimum edit distance algorithm.

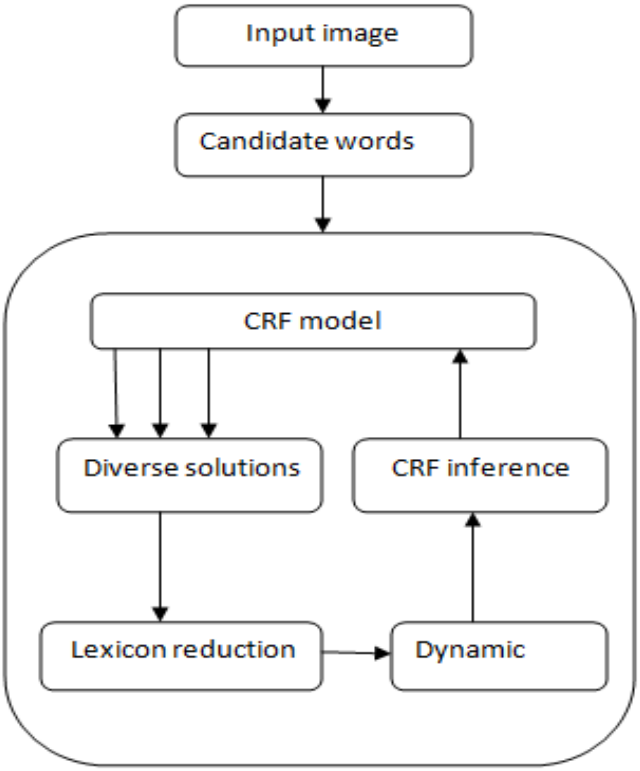


Figure 1. Flow for finding text from lexicons using CRF

The approach shown in figure-1 gives better results for small lexicon settings, and suffer from two main disadvantages: (i) getting a single set of true character windows in a input text image in defined methods is quite difficult, (ii) Pairwise information is less effected when the size of lexicon increases. A similar framework has been adopted for the recognition process. First, generated number of word hypotheses and obtain candidate words set so as to locate the word image from the image that has been inputted. Second, a method has been shown to analyze the large lexicon which depends on edit distances of candidate solutions and lexicon words of input image. This method results in significantly the lexicon size reduction and gives priors more specific to the input. Third, does not match with the existing works which gives a single

solution, this method also gives multiple solutions, and is also a applicable to the text-to-image retrieval process

*D. Method-D*

**Based on video text**

Text detection, recognition and extraction from videos results in analysis of the data and information contained in video and also understanding, as text can give concise and direct explanation of the stories and the information which are present in the videos. In digital news videos, the superimposed captions mostly present the name and the concluded summary of the news event. So, the recognized and extracted text might become a portion of index in a video retrieval and text recognition system. In this method, a new system for text information extraction from news videos has been explained. This system combines text detection, tracking and recognition to retrieve the text information or data from news and also from the commercial videos. Figure 2 shows the diagram of the basic architecture. Different from other approaches, this procedure integrates detecting and tracking on number of frames at one time to give the detection results more accurate and efficient. The quality of the detected or extracted text blocks is then improved by averaging over multiple frames. At last, the resulted averaged text blocks are binarized and finally the blocks are sent to OCR engine for recognition. To enhance the received results, an iterative binarization procedure is used here, to ensure good quality binarization results for recognition of text for the input image. Lot many efforts has been done to deal with the problems of text area detection [10][11][12], text tracking [11][14], and also text recognition [4]. The recent text detection approaches is classified into three categories. The first portion is method based on connected component [10], which can show text area very quickly but have difficulties when text is embedded in complex background like trees, vehicles, buildings or text is in contact with other graphical objects. The next category is texture-based [11][12][13], which is very difficult to find accurate and clear boundaries of text areas in the image and generally results many inaccurate or false alarms in "text-like" background texture areas. The last category is method based on edge [15][16][17]. Mostly, text regions in the image can be divided or decomposed by analyzing and studying the projection profiles of edge intensity maps. This kind of approaches is not that good for handling large-size text.

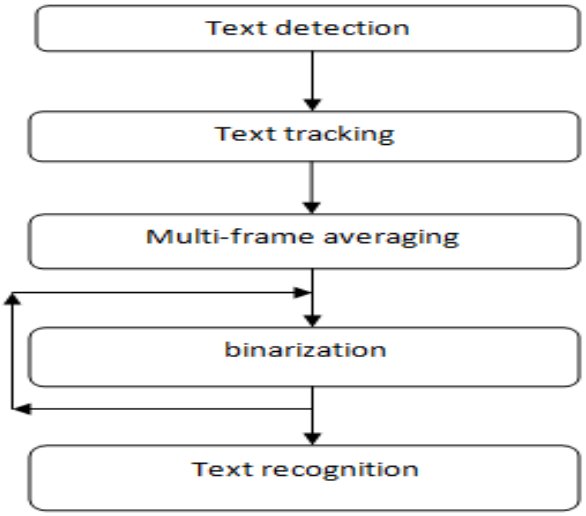


Figure 2. Architecture of recognition process in videos

**Text detection**

The text detection method here basically is texture-based and also the edge maps are used as the texture features of text lines Figure 3 shows the basic flow chart of text detection approach for videos. This procedure is only applicable for detection of horizontal text lines. As vertical text appears very rarely with some very few changes, it could be able to handle vertical text lines appearing in the video. Different from some other methods, the text detection algorithm defined here get more accurate boundaries of the text lines appearing in the sample video.

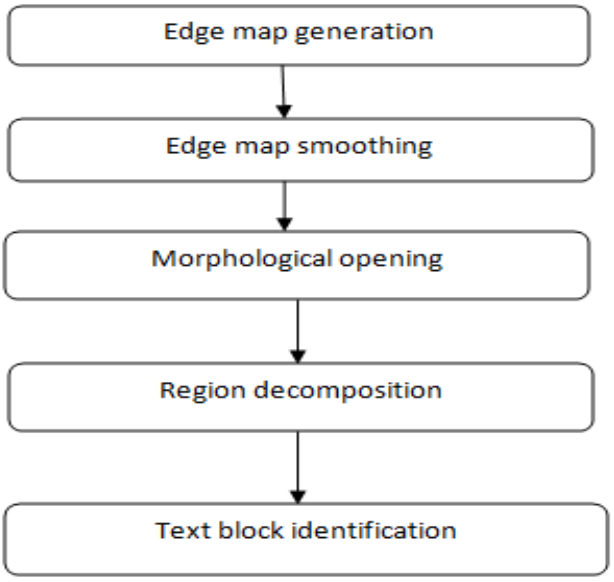


Figure 3. Basic text detection process

While text detection the following processes will be done

- Generating edge map

- Smoothing of edge map and morphological opening
- Decomposition of region
- Identification of text block
- Solving scale problem
- Applying evaluation method for text detection

**Text tracking**

As the text blocks are detected with the help of above defined procedure, the next process which is tracking, it can be applied to enhance detection accuracy and also it defines the use of multi-frames so as to increase the image quality, the blocks are result of above mentioned process.
Following procedure are applied

- Block matching based on SSD
- Multi frame averaging
- Detection and tracking integration

**Binarization and recognition**

From last few years most commercially used OCR software results in good recognition, mostly applicable only on high-resolution images, here enlarging the resolution of the averaged text blocks before converting the images to binary images is done.
For bilinear interpolation following steps can be applied

- Text intensity identification
- Post processing and adaptive thresholding
- Recognition of text

*E. Method-E*

**Based on text detecting and localizing for Scene Images**

To overcome the difficulties that arise in the text recognition, a hybrid approach is described to detect and localize texts robustly in natural scene images by grasping and merging advantages of both region-based and CC-based methods described earlier. From past some years local region detection method can robustly detect scene texts even in images that contain noise, so designed a text region detector to estimate or locate the probabilities of text position and scale, which directly segment candidate text components with an easy and efficient local binarization algorithm. To combine unary component properties that exist in the text image and it can be located with the help of CRF model and binary contextual component relationships, combined with supervised parameter learning. At last, text components that are detected in the image are grouped into text lines/words with an energy minimization method robustly. A basic flow chart that describes the functioning is been described in figure.4.

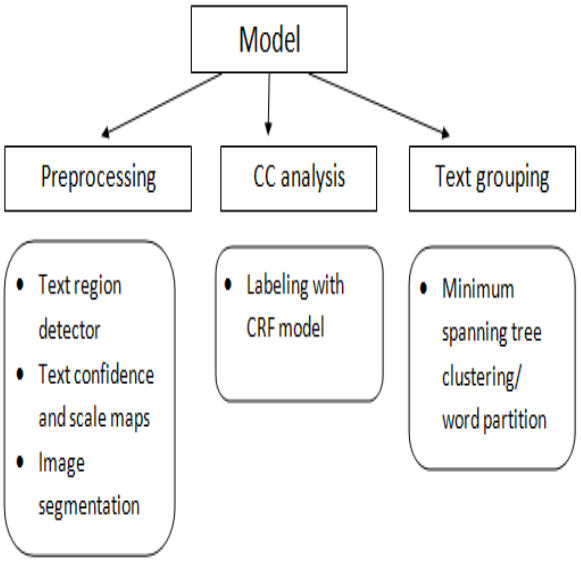


Figure 4. Hybrid method of text detection that combines region based and CC based method

**Preprocessing**

To extract, analyze and utilize local text area information in the image, a text region detector is designed for the estimation of the text confidence and its corresponding scale existing in text image, dependent on which candidate text components can be segmented, analyzed and extracted accurately

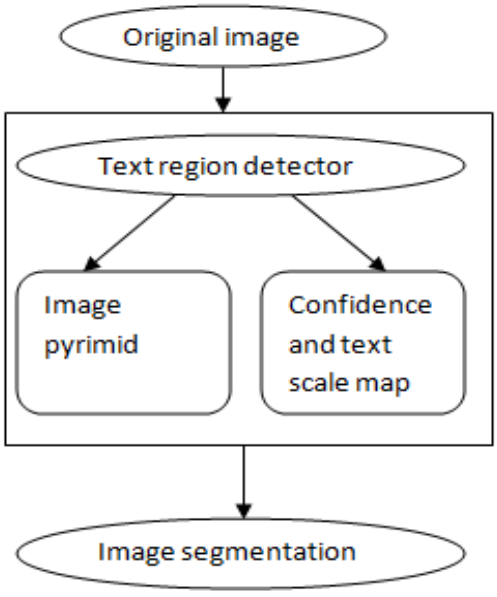


Figure 5. Flow for the preprocessing

**Text region detector**

Within each sliding window in a layer of the input image pyramid, several sub-regions by horizontally as well as vertically partitioning the window has been defined, from where other 4-orientation HOG descriptors are extracted through the input image and then sent for estimating the text confidence of the image.

*F. Method-F*

**Based on Text Detection in Scene text images**

In this technique, a robust and very accurate method based on MSER scene text detection has been explained briefly. initially, by analyzing in detail the hierarchical structure of MSERs and acquiring some simple features, a fast and accurate MSERs pruning algorithm has been developed in this procedure; the number of character candidates those has to be processed is very significantly decreased so as to detect text characters of image with high accuracy. Next thing that has been considered here is, a self-training distance metric learning algorithm, which has an ability to learn distance weights and also clustering threshold automatically when given an image as a input for recognition; candidates characters are clustered or grouped into text candidates through a single-link clustering algorithm with the help of the already learned parameters as computer program. At third step, use of a character classifier to estimate the posterior probabilities of input image text candidates corresponding to the existing non-text in the image are described in this process and also it remove text candidates with high non-text probabilities within the image. These kind of elimination helps to train a more accurate, efficient and reliable text classifier for identifying and recognize text in the image. Lastly, by combining the above described ideas, an accurate and robust scene text detection system has been built.

**ROBUST SCENE TEXT DETECTION**

By applying various key improvements over traditional MSER-based methods, a novel MSER-based scene text detection method has been developed. The proposed scene text detection method includes the following stages:

1. Character candidates extraction from the image
   From the input image Character candidates are extracted by applying the MSERs algorithm; the repeating components are decreased by MSERs pruning algorithm with the help of minimizing regularized variations.
2. Text candidates construction from received results
   Expected distance weights and clustering threshold of the detected text are learned with the help of the metric learning algorithm is used in this method; character candidates are clustered into text candidates through the single-link clustering algorithm by using the already learned parameters.
3. Text candidates elimination.
   The posterior probabilities of the input image text candidates corresponding to non-texts in the image are estimated by the character classifier and text candidates where high non-text probabilities of the image are deleted.
4. Text candidate's classification.
   Text candidates which correspond to real texts of image are identified by using the text classifier. A classifier is trained to finalize whether in an image the text candidate corresponding to the true text or not [28].

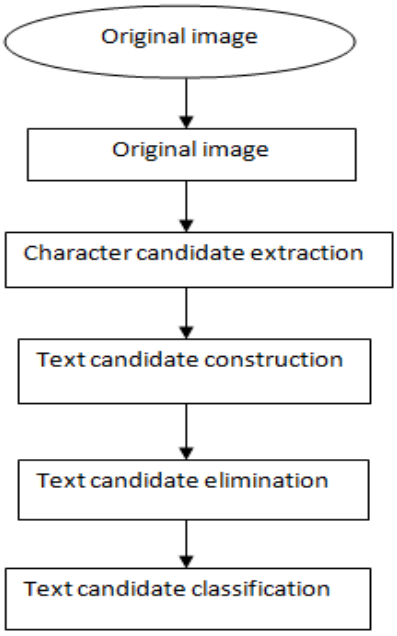


Figure 6. Detection using MSER

*G. Method-G*

**Based on Regression Tree Technique**

Some already defined system gives a hybrid approach(combined approach) to robustly detect , localize and recognize texts in natural scene images by acquiring benefit of regions-based algorithm and CC-based approaches used for detecting candidate regions[3][9]. This system consists of three basic stages preprocessing stage, the connected component analysis stage and text grouping from the input image. At the initial stage that is preprocessing, a text region detector is allocated to detect text regions of input image in each and every layer of the image pyramid and locate the text confidence and also the scale information involved back to the

original image which was inputted, after that scale-adaptive local binarization algorithm is defined to generate various candidate text components of the text involved in the input text image. When it moves to connected component analysis category, [4] [5] [7] a CRF model which means in conditional random field model integration of unary components and also binary contextual component relationships of text is finally used to detect and remove non-text components in the input image. In final stage, the neighboring text components involved in the text image are connected to a learning-based minimum spanning tree (MST) algorithm and also between line / word edges are cut off through an energy minimization model so as to group text components into lines or words of text. There is also description of the binary contextual component relationships defined, also including the unary component properties, all are combined in a CRF model which is a conditional random field model, and after that its parameters are optimized together through supervised learning algorithm. This procedure mostly does not work on some hard-to-segment texts images. If region-based methods are considered, the speed is comparatively quite slow and also the results are dependent on text alignment orientation somewhere. If this approach is considered which is, CC-based approach segmentation is not easy for text components without previous knowledge of the position and scale of text in the input image.

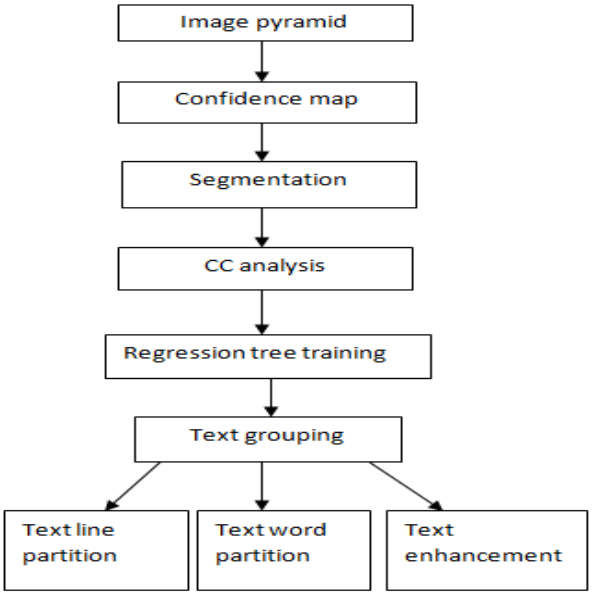


Figure 7. Text recognition using MST algorithm

*H. Method-H*

**Recognition based on fuzzy logic**

Fuzzy sets and fuzzy logic are basically used for presenting fuzzy character as well as for recognition. Algorithms based on fuzzy in which initially segmentation of character is done and then with the help of fuzzy system it results the all possible characters which are similar to the given input and after that using defuzzication system at the end recognizes the character in input image. Initially, from the document the input is taken in the form of scanned copies, these taken images might include large of number of sentences, words or characters. If there are number of sentences, these sentences require division into characters so as to be recognized every character independently. Segmentations are done with the help of algorithms. Therefore the sentences appearing in the input image are divided into number of characters. There is a database involving characteristics features of the characters of some specific language. For the Character recognition of input image it requires matching of features of characters with characteristic features in the existing database. For implementing this, the character recognition, the characters those were detected again should be divided into various numbers of cells. Through a fuzzy system considering as a controller, fuzzy information is transformed into a single value which will help in recognizing the character of input image. The transformation from a fuzzy set to a crisp number is known as defuzzification. Defuzzification defines a procedure of giving a quantifiable result in fuzzy logic. Various methods are available for the process of defuzzification. Such as

1. Method using Centre of gravity (COG)
2. Method that uses Centre of singleton (COS)
3. Method which uses Maximum
4. Margin properties of the centroid methods

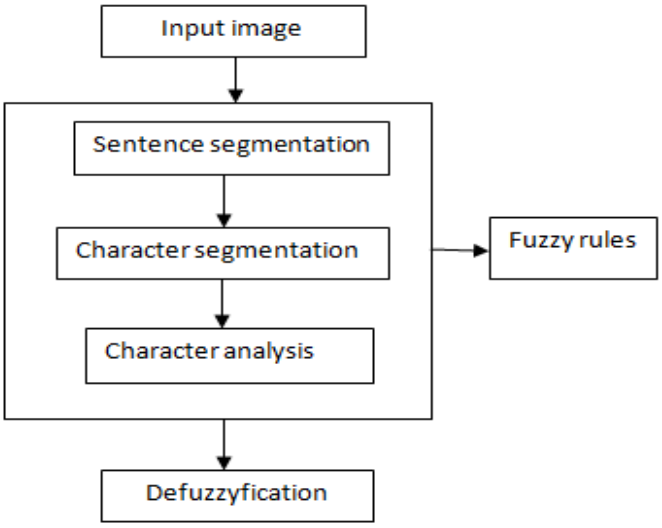


Figure 8. Recognition using fuzzy

*I. Method-I*

**Recognition based on Artificial Neural Network and Genetic Algorithm**

Using artificial neural network and genetic algorithm so as to solve text recognition problem an algorithm has been used here. A hetero-associative neural network is used to train the computer system for recognizing data from PDF or jpeg images which are somewhere not readable. Also, a crossover algorithm which is based on genetic algorithm is designed for getting or recognizing texts from the input image file. Here the main focus is to change the available text data from PDF and converting into digits to recognize characters easily. The defined genetic algorithm periodically performs crossover on sections and also on the portion of text data from an input image file to train the system. The genetic algorithm once trained with text data or information in image will change it into that form which can be recognized. The define algorithm for recognition process wisely handle the issue of deciphering digits and characters from image by parsing image and changing it to a pixel array. The algorithm selects digits, characters and process crossover with trained patterns along with variable heights.

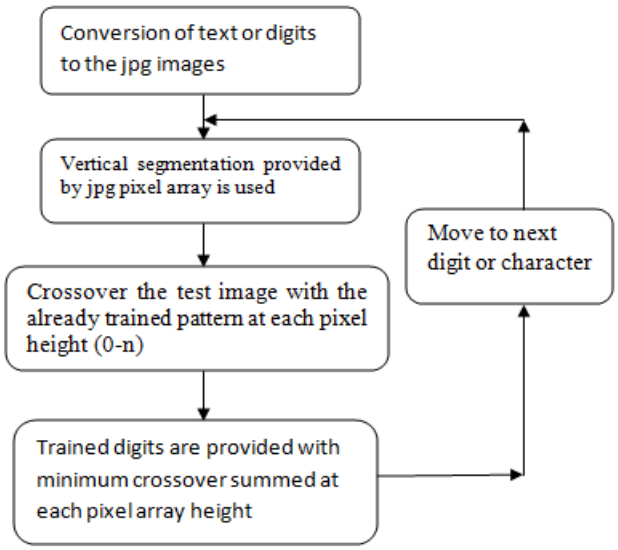


Figure 9. Recognition using artificial neural network and genetic algorithm

*J. Method-J*

**Recognition based on the decorative text**

Sometimes documents found with text imposed on very complex and complicated color background. Reading of text in those documents or images is very difficult due to high complexity of the backgrounds and combinations of color(s) of fore-ground text including all colors of background. In this approach the extraction of foreground text from the image document which is having complex background has been done. The applied flow is a combined approach which relates connected component approach and also texture feature analysis of text region in the image. The method that has been applied uses Canny edge detector to highlight or detect the text region in the input document image.[41]The Connected component approach applied on various edge pixels of the text is to know and identify candidate text regions from the input image. Due to the complex background it is highly possible that a non-text region in the image may be highlighted as a text region in the image [42]. To overcome this issue the texture features of potential text region related to each connected component in the text is used. For the segmentation an unsupervised local thresholding is performed. After that the noisy text region which is detected is reprocessed.

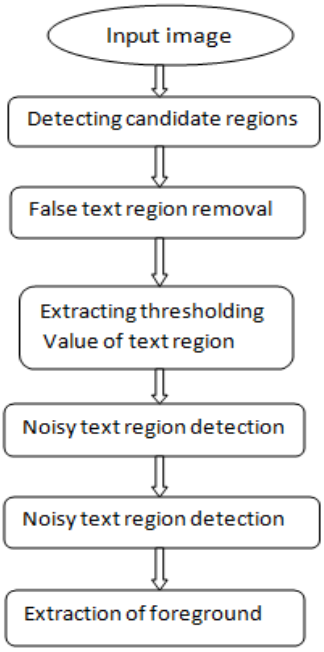


Figure 10. Recognition flow for decorative text images

## III. APPLICATIONS

Assistive technology for blind and visually Since last few years, there are various text related applications for both images and video available in the computer world, which are categorized as multimedia retrieval domain broadly, Visual input of various types of text, as well as industrial automation.

1. **Multimedia recognition domain:** Text in web images is almost related to the data or information of the web pages available on internet. Captions in videos generally reflect information about from where, when also who of the proceeding events [36], [37]. Recognizing, localizing, detecting text and extracting keywords in these types of multimedia resources increase multimedia retrieval.
2. **Visual input of various types of text:** The enhancement in mobile devices including digital cameras has created imaging devices easily available in the market. Through an embedded method, mobile

devices itself input cards name, sometimes slide presentations [38], [39], [40]. Without any pressure to give input through keyboard, users might feel better and can work more efficiently and easily. Signs in various natural scenes have information. Automatic sign recognition and translation systems help users to deal with language barriers.

3. **Automation in Industry:** Recognition of text on various packages, containers, houses, and maps is having wide area of application in industrial automation. Automatic number plate recognition
4. Automatic insurance documents impaired users.
5. Key information extraction from text images uses this domain.
6. Extracting business card information into a contact list.
7. More quickly make textual versions of printed documents, e.g. book scanning.
8. Make electronic images of printed documents searchable, e.g. Google Books.
9. Converting handwriting in real time to control a computer.

## CONCLUSION

The character recognition processes has been developed highly in the last few years. Various types of methods or techniques have grown, and influenced by developments in near domains like image recognition and face recognition. The challenges of resulting recognition are analyzed by the nature of the text to be read and also through text quality. In this paper we have reviewed and analyzed different methods to find text characters from scene images, document images, and born digital images. We have reviewed basic architecture of text recognition from images. In which we have discussed different techniques of image processing in particular sequence for text recognition from scan image .Different techniques on different types of images through table has been discussed which includes name, technique or algorithms used and the challenges. Also, we have discussed some application of text recognition system**.** The idea and method of recognition has been explained with the help of flow charts. Huge future work can be done in the field of text recognition, as preprocessing is done for the process of recognition, for future work it can be replaced by other algorithms to reduce the complexity. As many methods are available for the process of text recognition using all the existing algorithms new methods can be proposed to create new hybrid algorithms. The hybrid approach would be used in such a way that it reduces the complexity of existing algorithms in terms of size and time. If we consider the document images sometimes text is missing, for recognizing that text we can apply artificial intelligence algorithms, for speeding up the procedure fuzzy logic or neural network can be applied. We have included a list of references sufficient to provide a more-detailed understanding of the approaches described.

TABLE I. VARIOUS RECOGNNITION METHODS

| . no | Methods used | Algorithm used | challenges |
|---|---|---|---|
| 1. | Method –A **Based on born digital images using conditional random field** | In this the algorithm is based on the text extraction using conditional random field. In this CRF model uses unary component as well as binary components. | Born digital images contain more defects and have more complex background, also have low resolution, compression loss. |
| 2. | Method-B **Based on complex images** | Text extraction review has been done, existing key methods has been discussed along with the advantages and disadvantages. | What makes difference between text and its background, which method can be directly applied on basis of initial information? |
| 3. | Method-C **Based on scene text images for large lexicons** | This is basically based on the lexicons; a CRF model has been defined on potential character locations and interaction between them. | Lexicon size is the challenging problem, also scene text detection is also challenging. |

| | | | |
|---|---|---|---|
| 4. | Method-D **Based on video text** | Text detecting and tracking integration method has been developed to locate text area in images, multi-frame averaging is also used, and thresholding is used for binarization. | Extracting text from the videos is a difficult task, also it is difficult to detect small characters from the videos. Detection can be done only for horizontal lines. |
| 5. | Method-E **Based on text detecting and localizing for Scene Images** | Text localization and detection is done using hybrid approach robustly, segmentation candidate text components by local binarization, CRF model are used, and learning based energy minimum method is applied. | Complex background, non-uniform illumination, variation of text font, size and line orientation. |
| 6. | Method-F **Based Text Detection in Scene text images** | Pruning algorithm has been designed for MSER as candidate character, minimizing Regularized variations, single link clustering algorithm used, self training distance metric learning algorithm. | Detection of blurred text in images, low resolution text detection is difficult, multilingual text detection is a difficult task. |
| 7. | Method-G **Based on Regression Tree Technique** | Segmentation of candidate text component by local binarization, CRF model is used, learning based energy minimization method. | Difficulty arises due to complex background, non uniform illumination, the variant of text font, size orientation of text. |
| 8. | Method-H **Recognition based on fuzzy logic** | Fuzzy approach is applied for character recognition, segmentation and matching of character through fuzzy rules, defuzzification is also applied at last step. | Variation in different writing styles, when partially Witten text appears it becomes difficult to recognize. |
| 9. | Method- I **Recognition based on Artificial Neural Network and Genetic Algorithm** | Artificial neural network and genetic algorithm used for text recognition, crossover based genetic algorithm used. | Blurred images create problem in recognition. |
| 10. | Method-J **Recognition based on the decorative text** | A hybrid approach which uses connected components and texture feature analysis is used. Also the work has been done with the decorative images. | Extracting the text from complex backgrounds. |

## REFERENCES

[1] Weinman, J., Butler, Z., Knoll, D., Feild, J.: Toward integrated scene text reading. TPAMI **36**, 375–387 (2014)
[2] Bissacco, A., Cummins, M., Netzer, Y., Neven, H.: Photoocr: reading text in uncontrolled conditions. In: ICCV (2013)
[3] Wang, K., Babenko, B., Belongie, S.: End-to-end scene text recognition. In: ICCV (2011)
[4] Mishra, A., Alahari, K., Jawahar, C.V.: Top-down and bottom-up cues for scene text recognition. In: CVPR (2012)
[5] Goel, V., Mishra, A., Alahari, K., Jawahar, C.V.: Whole is greater than sum of parts: recognizing scene text words. In: ICDAR (2013)
[6] Rodriguez, J., Perronnin, F.: Label embedding for text recognition. In: BMVC (2013)
[7] Sheshadri, K., Divvala, S.K.: Exemplar driven character recognition in the wild. In: BMVC (2012)
[8] Tian, S., Lu, S., Su, B., Tan, C.L.: Scene text recognition using co-occurrence of histogram of oriented gradients. In: ICDAR (2013)
[9] Novikova, T., Barinova, O., Kohli, P., Lempitsky, V.: Large-lexicon attributeconsistent text recognition in natural images. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part VI. LNCS, vol. 7577. Springer, Heidelberg (2012)
[10] A. K. Jain, and B. Yu, "Automatic Text Location In Images and Video Frames," Pattern Recognition, Vol. 31, No. 12, 1998.
[11] H. Li, D. Doermann, and O. Kia, "Automatic Text Detection and Tracking in Digital Video," IEEE Trans. Image Processing, Vol. 9, No. 1, Jan 2000.
[12] Y. Zhong, H. Zhang, and A. K. Jain, "Automatic Caption Extraction of Digital Videos," Proc. ICIP'99, Kobe, 1999.

[13] T. Sato and T. Kanade, "Video OCR: Indexing digital news libraries by recognition of superimposed caption," ICCV Workshop on Image and Video retrieval, 1998.
[14] H. Li and D. S. Doermann, "Text Enhancement in Digital Video Using Multiple Frame Integration," Proc. ACM - Multimedia 99, Orlando, Florida, 1999.
[15] W. Qi, et. al. "Integrating Visual, Audio and Text Analysi for News Video". 7th IEEE International Conference on Image Processing (ICIP2000), Vancouver, British Columbia, Canada, 10-13 September 2000.
[16] A. Wernicke, R. Lienhart. "On the Segmentation of Text in Videos". IEEE Int. Conference on Multimedia and Expo (ICME2000), July 2000.
[17] T. Sato, T. Kanade, E. Hughes, and M. Smith. "Video OCR for Digital News Archives". IEEE International Workshop on Content-Based Access of Image and Video Databases, January, 1998.
[18] X. R. Chen and A. L. Yuille, "Detecting and reading text in natural scenes," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR'04), Washington, DC, 2004.
[19] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, 2005.
[20] S. L. Feng, R. Manmatha, and A. McCallum, "Exploring the use of conditional random field models and HMMs for historical handwritten document recognition," in Proc. 2nd Int. Conf. Document Image Analysis for Libraries (DIAL'06), Washington, DC, 2006.
[21] J. Gllavata, R. Ewerth, and B. Freisleben, "Text detection in images based on unsupervised classification of high-frequency wavelet coefficients," in Proc. 17th Int. Conf. Pattern Recognition (ICPR'04), Cambridge, U.K., 2004.
[22] J. M. Hammersley and P. Clifford, Markov Field on Finite Graphs and Lattices, 1971, unpublished.
[23] X.-B. Jin, C.-L. Liu, and X. Hou, "Regularized margin-based conditional log-likelihood loss for prototype learning," Pattern Recogn., vol. 43, no. 7, pp. 2428–2438, 2010.
[24] Y. Jin and S. Geman, "Context and hierarchy in a probabilistic image model," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR'06), New York, NY, 2006, pp. 2145–2152.
[25] B.-H. Juang and S. Katagiri, "Discriminative learning for minimum error classification," IEEE Trans. Signal Process., vol. 40, pp. 3043–3054, 1992.
[26] K. Jung, K. I. Kim, and A. K. Jain, "Text information extraction in images and video: A survey," Pattern Recogn., vol. 37, no. 5, pp. 977–997, 2004.
[27] K. I. Kim, K. Jung, and J. H. Kim, "Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 12, pp. 1631–1639, 2003.
[28] X. B. Liu, H. Fu, and Y. D. Jia, "Gaussian mixture modeling and learning of neighboring characters for multilingual text extraction in images," Pattern Recogn., vol. 41, no. 2, pp. 484–493, 2008.
[29] Y. X. Liu, S. Goto, and T. Ikenaga, "A contour-based robust algorithm for text detection in color images," IEICE Trans. Inf. Syst., vol. E89-D, no. 3, pp. 1221–1230, 2006.
[30] S. M. Lucas, "ICDAR 2005 text locating competition results," in Proc. 8th Int. Conf. Document Analysis and Recognition (ICDAR'05), Seoul, South Korea, 2005, pp. 80–84.
[31] M. Lyu, J. Q. Song, and M. Cai, "A comprehensive method for multilingual video text detection, localization, and extraction," IEEE Trans. Circuits Syst. Video Technol., vol. 15, no. 2, pp. 243–255, 2005.
[32] K.B.M.R. Batuwita, "Meaningful Segmentation of Offline Individual Handwritten Numeric Characters," 2006 IEEE International Conference on Fuzzy Systems Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada July 16-21, 2006.
[33] Md. Shabiul Islam, M.S. Bhuyan, Sawal H. M. Ali, Masuri Othman, Burhanuddin Yeop Majlis, "VHDL Implementation of Fuzzy Based Handwriting Recognition System "ICSE2010 Proc. 2010, Melaka, Malaysia.
[34] Ankit Sharma,Dipti R Chaudhary, "Character Recognition Using Neural Network", International Journal of Engineering Trends and Technology (IJETT) - Volume4lssue4- April 2013
[35] Y. Zhong, H.J. Zhang, and A.K. Jain, "Automatic Caption Localization in Compressed Video," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, pp. 385-392, 2000.
[36] Q. Ye, Q. Huang, W. Gao and D. Zhao, "Fast and Robust Text Detection in Images and Video Frames," Image and Vision Computing, vol. 23, pp. 565-576, 2005.
[37] I. Haritaoglu, "Scene Text Extraction and Translation for Handheld Devices," in Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition, pp. 408-413, 2001.
[38] J. Liang, D. Doermann, H. Li, "Camera-based Analysis of Text and Documents: A Survey," Int'l J. Document Analysis and Recognition, vol.7, pp. 84-104, 2005.
[39] L. Lin, C.L. Tan, "Text Extraction from Name Cards Using Neural Network," in Int'l Joint Conf. Neural Network, pp. 1818-1823, 2005.
[40] E. Kavallieratou and E. Stamatatos, "Improving the quality of degraded document images," Proceedings of 2nd International Conference on Document Image Analysis for Libraries, pp. 340–349, 2006.
[41] S. Nirmala, P. Nagabhushan, "Isolation of foreground-text in document images having known complex background," Proceedings of 2nd International Conference on Cognition and Recognition, pp. 99–106, 2008.
[42] [43] J. Sauvola and M. Pietikäinen, "Adaptive document image binarization," Pattern Recognition, Vol. 33, No. 2, pp. 225–236, 2000.