

COHESIVE MULTI-ORIENTED TEXT DETECTION AND RECOGNITION STRUCTURE IN NATURAL SCENE IMAGES REGIONS HAS EXPOSED

Imran Siddiqui¹ and Dr. Varsha Namdeo²

¹Department of Computer Engineering, RKDF IST, Bhopal
, Bhopal, INDIA

¹Department of Computer Engineering, RKDF IST, Bhopal

ABSTRACT

Scene text recognition brings various new challenges occurs in recent years. Detecting and recognizing text in scenes entails some of the equivalent problems as document processing, but there are also numerous novel problems to face for recognizing text in natural scene images. Recent research in these regions has exposed several promise but present is motionless much effort to be entire in these regions. Most existing techniques have focused on detecting horizontal or near-horizontal texts. In this paper, we propose a new scheme which detects texts of arbitrary directions in natural scene images. Our algorithm is equipped with two sets of characteristics specially designed for capturing both the natural characteristics of texts using MSER regions using Otsu method. To better estimate our algorithm and compare it with other existing algorithms, we are using existing MSRA Dataset, ICDAR Dataset, and our new dataset, which includes various texts in various real-world situations. Experiments results on these standard datasets and the proposed dataset shows that our algorithm compares positively with the modern algorithms when using horizontal texts and accomplishes significantly improved performance on texts of random orientations in composite natural scenes images.

KEYWORDS

Text extraction, scene text, Stroke Feature Transform (SFT), Text Region, Stroke Width Transform (SWT).

1. INTRODUCTION

During the recent years handwriting recognition has become an exhaustive research area, very precise domains. Only very a small amount of schemes are known which take in hand the domain of universal printed text recognition. Among these schemes, two essentially unusual methods have can be examined. Here in this paper present [1] a new propose a method that follows the first approach, i.e. it segments complete lines of printed text into single words. Segmenting a on paper text into lines and essentially to extra segment it into particular declarations is easier said than done and error prone for the reason that it is a pure bottom-up procedure and no high stage information about the text is reachable. The word segmentation algorithm used in our method includes some designs from [1]. Word recognition form scanned or printed image has been also became a significant job in image processing. That's why segmentation based recognition algorithm is explained in [2][3]. Both produce the idea for off-line handwritten text.

The methods of Chen et al. [4] and Epshtein et al. [5] are modern efforts in the direction of end-to-end text recognition schemes. These methods essentially address the problem of text detection and accomplish subsequent recognition using off-the-shelf OCR engines. In [4], Chen et al. trained a flow classifier to choose text regions and apply commercial OCR software to the

detected areas. Epshtein et al. [5] proposed a novel image operator called Stroke Width Transform (SWT) to confine texts in natural scene images. For recognition, off-the-shelf OCR engine was accepted to recognize the masks generated by SWT. These techniques have accomplished exceptional piece on natural images and motivated various researchers in this area [6-8]. On the other hand, these techniques take care of text detection and recognition as separated phases and only focus on horizontal texts. The recognition correctness of these techniques is moderately incomplete as off-the-shelf OCR engines are exclusively planned for texts in document images not natural scene images. Several research works has been done on text detection and extraction technique in natural scene image in the earlier years. Various researchers have been proposed or studied many methods of text extraction and give precise results on the basis of performance metrics. In this paper, we propose a new scheme which detects texts of arbitrary directions in natural scene images. Our algorithm is equipped with two sets of characteristics specially designed for capturing both the natural characteristics of texts using MSER regions using Otsu method. To better estimate our algorithm and compare it with other existing algorithms, we are using existing MSRA Dataset, ICDAR Dataset, and our new dataset, which includes various texts in various real-world situations.

2. SCENE TEXT

There is a long record of research and development of automatic readers. Text extraction from environment is concurrently connection to successive steps of character segmentation and recognition. This combined chain essentially plans at combining color, intensity and spatial information of pixels for strength and accurateness. Additionally, text almost always become visible in canonical horizontal or vertical orientations and various methods have been expanded to approximation the comprehensive rotation essential to horizontally “level” text lines in the image plane. As confirmed in [9], in predictable end-to-end scene text recognition schemes [5], [10], decisions of unusual components in the pipeline are chronological and separated which constructs it unfeasible to develop criticism information between different elements. Modern expansions have also been made in the investigation of pages using camera-based achievement methods. More difficult world transforms must be measured in these cases, but frequently a precise model of the page can be utilized to correct the text and revisit it to a planar exterior. Uneven lighting can be difficult in these circumstances, but the general binary nature of text on a page still creates local processing possible for on condition that good binarizations.



Figure-1: Images for scene text reading

The problem of finding text in an arbitrary image of a scene can be radically more complex. First, the contents of the input image are generally more varied. From urban structures to more natural

subjects like trees, the variety of potential image contents is vast, and it occupies the entire input image. Text regions in scene images need not be well-bounded the way they usually are in documents. Furthermore, text in scenes is often only a few words in one place. There are no large paragraphs or long lines to analyze. Although text is generally designed to be readable, there are often adverse effects of the imaging conditions that can make it easier said than done to identify. Distance from the camera can make text small and low resolution without much element. Specularities can mix text regions with a replicated image. Perspective distortion can produce text with a contrasting font size or orientation. Additionally, unlike document processing there is no comprehensive page model whose transform constraints can be approximation. Because text may be printed on bricks, wood or complex conditions, the easy binarization and text zoning algorithms of document processing will be inadequate. Small amounts of text can become visible anywhere, at any amount, with any world orientation and on any surface. All of these problems have a propensity to make text location in scene images usually more challenging than document text detection and pre-processing.

3. LITERATURE SURVEY

This section consist different earlier works that had been previously proposed by more than a few researchers. Some widespread methods are also talking about here that effort professionally in area of text recognition.

G. Seni and E. Cohen [11] utilized External Word Segmentation of Off-Line Handwritten Text Lines. The initial step in on condition that a recognition / retrieval implement is to repeatedly segment handwritten pages into words. There has been small effort on full handwritten pages and this effort has frequently involved testing on uncontaminated digital image documents generated for the reason of examine. Historical manuscript images, contains noise and much more challenges for retrieval. A novel scale space algorithm for automatically segmenting printed or handwritten (historical) documents into words is described. In initial, the page is cleaned to eliminate boundaries and tagged by a gray-level projection outline algorithm for discovering lines in images. Each line image is then filtered with an anisotropic Laplacian at numerous scales. This process generates blobs which match to sections of characters at undersized scales and to expressions at superior levels. Algorithm is amount collection specifically discovering the optimum level at which blobs match up to expressions. This is completed by discovering the greatest over level of the amount or region of the blobs. This level maximum is approximation using three different methods. The blobs improved at the most favorable scale are then restricted with a rectangular box.

Nallapareddy Priyanka in [12], make available a novel idea for Line and Word Segmentation method for Printed Documents, this comprises printed or scan English text image binarized to make segmentation. They are discussing some original approach for line, word and character segmentation of printed English document. A small number of efforts has been done for optical character recognition on other Indian script on the other hand in case of English language it is approximately insignificant. So this work is an attempt to produce a report on segmentation of documents containing English script forms. Here initial they are conferring about the arrangement of English language and then they talk about some plan for segmentation of line, word and character from English document.

Azadboni, M.K., Shahed Univ., Tehran, Iran in [13] proposed an approach for text detection and localization is proposed. For this reason, they initial localize text location and then establish characters' pixels. The suggested text detection method is a two-stage algorithm that in initial period they be appropriate low pass filter on image in FFT domain to eliminate noisy element and then they are appropriate Laplacian operator to the ensuing image to emphasize high contrast

regions in the image. Then the item for consumption of corner dilated points and Laplacian enhanced image is estimated and text blocks are removed using image vertical and horizontal projection. In the subsequent period of the algorithm the extracted text blocks are confirmed using an SVM classifier. Text textures for ex. text angles and variance, momentum, entropy in co-occurrence matrix of text block are utilized for SVM training. They presupposed that the characters of each text block have the similar color. Consequently, they initial approximation environment color using image pixels in borders of detected text regions. Then the text color is approximated using the color clusters of pixels in text block and background color. They employ color segmentation to remove character pixels.

W. Huang in [14] used Stroke Feature Transform (SFT) followed by a text constituent classifier and a text-line classifier consecutively to remove text areas. As a final point, text areas are positioned by the text-line self-assured map. They present a novel method for text localization in natural images by discriminating text and non-text areas at three levels: pixel, component and text line stages. Initially, a powerful low-level filter called the Stroke Feature Transform (SFT) is recommended which expands the widely utilized Stroke Width Transform (SWT) by incorporating color prompts of text pixels most important to considerably improved show on inter-component partition and intra-component association. Secondly, supported on the output of SFT they are appropriate two classifiers a text part classifier and a text-line classifier successively to remove text areas.

S. A. Angadi & M. M. Kodabagi in [15], proposed a method for ‘Text Region Extraction from Low Resolution Natural Scene Images using Texture Features’. It is texture based and functions on low resolution natural scene images incarcerated by cameras implanted in mobile phones to detect and segment text areas. The method utilizes high pass filter in the DCT domain to restrain the environment and texture characteristics such as homogeneity and distinction to notice and segment text areas. The dealing out is carried out on 8x8 sized image blocks during conditions suppression stage and the left over phases use 50x50 sized image blocks. There are numerous advantages of with huge dimension image blocks for removing texture characteristics. One such advantage is the superior size image blocks cover additional features and for this reason extracted characteristics give adequate information for accurate categorization of blocks into text and non-text group. The other benefits include; robustness and insensitiveness to variation in size, text and alignment.

Kwang in Kim, Keechul Jung, and Jin Hyung Kim in [16] texture based approach is discussed in which text is considered as a specific form of texture. A support vector machine (SVM) is utilized to examine the textural properties of texts. No outside texture characteristic removal component is utilized but to a certain extent the intensities of the unprocessed pixels that make up the textural prototype are fed in a straight line to the SVM which efforts well still in high-dimensional spaces. Subsequently, text areas are recognized by be appropriating an always adaptive mean shift algorithm (CAMSHIFT) to the consequences of the texture analysis. The arrangement of CAMSHIFT and SVMs manufactures both full-bodied and well-organized text detection, as time-consuming consistency analyses for smaller amount appropriate pixels are limited departure only a small element of the input image to be texture examined.

Andrej Ikica, Peter Peer in [17] Text detection in natural images has increased much notices in the last existence as it is a most important step in the direction of fully independent text recognition. They primarily hold various composite texts of different dimensions styles and colors with difficult conditions. It utilizes a position of heuristic policy to eradicate detection of non-text regions. The technique is estimated on CVL OCR DB, an annotated image database of text in natural scenes.

4. PROPOSED ARCHITECTURE SYSTEM

A. PROBLEM DEFINITION

In this paper various issues related with topic is mainly segmentation using different texts related. Here we discussed various segmentation issues and challenges to maintain a good segmentation method, means the challenges involved in the segmentation of the text-lines, words and characters can be proposed or studied many methods of text extraction and give precise results on the basis of performance metrics. In this paper, we propose a new scheme which detects texts of arbitrary directions in natural scene images. When dealing with printed or handwritten text, line segmentation has to solve some difficulties that are exceptional in modern printed text. Among the most identification are:

- **Input text** contains noisy documents image.
- **Image documents** include non-constant spaces between text lines, words and also with characters.
- **Input image** consist insignificant text.
- **The text** with various text sizes coexists in the printed image.
- **Text image** with graphical designs and ornamental characters.
- **Image document** whose color text is skewed and/or wrapped.
- **Multi column** text document image.
- **Skewed Lines:** Lines of text in wide-ranging are not straight. These lines are not similar to each other. Here the difficulty happens that the printed or handwritten lines are not properly aligned.
- **Fluctuating lines:** Lines of text are moderately or fully connected to other text-lines. There is not any unique line to be directly well means it can be varied.

B. PROPOSED METHOD

In this part unusual technique for segmentation of printed English documents into lines, words and characters are recommended. This discusses segmentation approach for text detection. In text, letters should have uniform color and text's family within a text string. This property is used in color based text detection and texts based time analysis approaches. The general algorithm for color based and text based line, word and character text detection approach is as follows. Initially Text image is segmented. As we can see the proposed method or model used for segmentation (line, word and character). Text information Extraction system consist four main phases which will used in our proposed work. The input image is to be processed and passed between all these phases for our purpose and achieving output. Input RGB image with printed or handwritten or scan text, noise can be raised during capturing or passes through electronic medium. This could be blurring, break pixels etc. Image with containing noise passes through out proposed method, removes noise and produces output.

In TIE system we discuss four phase for according to the segmentation need. By this model, text image of English script generates by scanners, cameras and machine can passes through every phase for producing the accurate output. RBG text image includes different types of texts namely- Arial rounded, Times new roman, Calibri etc of segmentation of English script is totally text based segmentation process. For this process some basic routine operation is as follow:

Pre-Processing: In pre processing the printed or scan text image is transformed into prepared to employ format or the noise in the image is decreased by using these techniques in the scanned image the noise would be like character shapes are not precisely scanned document was not

related appropriately the document is skewed to uncommon degrees, edges are not flat, multiple colored characters, So, the eliminate such noise, they have to be appropriate some logics as computer doesn't be acquainted with the variation between the noise and the precise character. To remove the noise the following algorithms can be applied to make it ready to use for character recognition.

- **De-skew:** This method is utilized if the document was not aligned appropriately when scanned it may require to be twisted a small number of degrees clockwise or anti-clockwise to facilitate make lines of text completely horizontal or vertical.
- **Binarization:** This method is utilized to change an image from RGB color to grayscale to black-and-white to make it suitable for character recognition and it is also known as "binary image" for the reason that there are only two colors left in the representation after the binarization i.e.. "BLACK" and "WHITE" Algorithm: Binarization [18]

Step-1: Convert the color English text image into grayscale text image.

Step-2: Find the gray level text image of input image $I(x, y)$.

Step-3: Find the black pixel point and select that point as threshold (T).

Step-4: Output binary image $B(x, y)$ is characterized as

$$B(x, y) = \{1, 0\}: \text{for } I(x, y) > T, \text{ Otherwise.}$$

- **Thersholding:** In order to increase the segmentation processing speed and to reduce storage space thresholding is done on color text image or gray scale text image. The main purpose of threshold (in image) is to separate areas of foreground (ink or on pixels) from the background (off pixels) by generating a histogram of an image.

Feature Detection: The features are known as some interesting parts of an image. An image contains any face, text, scene and texture etc so in such a way this includes the most interesting part of information. Feature detection is the second phase of TIE system with binary image. Detecting the features in an image is a low-level image processing operation because it is often performed as the first operation on the images. It examines each pixel to check, if there any feature on that pixel is present. When a feature detection algorithm is performed and there is a time constraint then a higher level algorithm may be used to supervise the feature detection phase.

- **Text Edge:** The text edegs are the points which includes boundary between two image regions of texts. The text edges are basically as sets of pixel points in the text image which have a high gradient magnitude.
- **Region of Interest / Blobs:** Region of interest (ROI) provide a detail discription of text image structures in terms of regions. It can detect text areas in an image which smooths in previous phase of TIE system.

Horizontal Projection: In order to extract individual text line of English script, this technique based on projection is used. A horizontal projection is a histogram which is giving the several number of ON (value 1) pixels accumulated along parallel lines of text. Thus a horizontal projection is a one-dimensional [1D] array, where every the elements denote the number of ON (value 1) pixels along a row in the text image. Correspondingly a vertical projection outline provides the all column sums. It is simple to distinguish that separating lines by give the impression of bringing for minima in horizontal projection of the input text image and then it can divide words by give the impression of being at minima in vertical projection of one line. Such projection based methods are used for line. To part the handwritten text document image into number of text lines the valleys of the horizontal projection calculated by a row-wise sum of black pixels are utilized.

Morphology Based Projection: For extracting text word from text image of English script with different texts, then Mathematical morphology can be used as a tool. Image components are useful in the representation and description of region of text. Dilation is a primitive morphological operation that grows the objects in a binarized image and the disconnected components are connected using dilation. A shape which controls the extent of this thickness in a specific manner referred to as a structuring element for the segmentation of the lines into words.

Bounding Box: With the purpose of extracting individual texts words and characters this methods supported on bounding box is utilized. First the input RGB text (printed or and written) image is converted to gray scale and histogram corresponding image is plotted. Next find the white spaces and identify the ON pixel measurement with the region props property which calculates centroid for connected or non-connected components in the image. An area prop calculates different properties of the character objects in the binary image. In conclusion with facilitate of dimensions of centroids individual words and characters are removed from the text image.

Proposed Algorithm for Text recognition is following steps is performed:

- **Step 1:** Load image
- **Step 2:** Determine Bounding Boxes Enclosing Text Regions
- **Step 3:** Complete Optical Character Recognition on Text Region
- **Step 4:** Apply the Text Detection Process to Other Images
- **Step 5:** As basically ICDAR and MSAR dataset images are procedures and make changing it into rgb2gray for removing the color.
- **Step 6:** Detecting input noise and Remove noise using noise filter and also detecting the low density background edges and high density foreground edges.
- **Step 7:** Binarization the processed image and bounding rectangular region on to the text into number of connected components.
- **Step 8:** Performing Connected component analysis to find recognized text form image dataset.
- **Step 9:** Calculating precision, recall and F-measure for both ICDAR and MSAR dataset and compare with existing for better result.
- **Step 10:** Calculating CPU time, taken by the overall recognition process on different image dataset and find difference of time in proposed work for ICDAR and MSAR dataset.

Otsu's process named after Nobuyuki Otsu is utilized to repeatedly complete clustering-based image thresholding, or, the decrease of a graylevel image to a binary image. The algorithm presumes that the image restrains two classes of pixels subsequent bi-modal histogram (foreground pixels and background pixels), it then computes the optimum threshold unscrambling the two classes so that their joined spread i.e. intra-class variance is negligible or consistently i.e. for the reason that the sum of pair wise squared expanses is steady so that their inter-class discrepancy is maximal. As a result, Otsu's technique is just about a one-dimensional discrete analog of Fisher's Discriminant Analysis.

Method: This type of thresholding is global thresholding. It amasss the intensities of the pixels in an array. The threshold is premeditated by using entirety mean and variance. Based on this threshold value each pixel is set to either 0 or 1. i.e. background or foreground. Thus at this point the change of image gets position only formerly.

- The following formulas are used to calculate the total mean and variance.

- The pixels are divided into 2 classes, C_1 with gray levels $[1, \dots, t]$ and C_2 with gray levels $[t+1, \dots, L]$.
- The probability distribution for the two classes is:

$$C_1: p_1/w_1(t), \dots, p_t/w_1(t) \text{ and}$$

$$C_2: p_{t+1}/w_2(t), \dots, p_L/w_2(t)$$

$$\text{Where } w_1(t) = \sum_{i=1}^t p_i$$

$$\text{and } w_2(t) = \sum_{i=t+1}^L p_i$$

- Also, the means for the two classes are

$$\mu_1 = \sum_{i=1}^t ip_i/w_1(t)$$

and

$$\mu_2 = \sum_{i=t+1}^L ip_i/w_2(t)$$

- Using Discriminant Analysis, Otsu described the between-class discrepancy of the thresholded image as

$$\sigma_B^2 = w_1 (\mu_1 - \mu_T)^2 + w_2 (\mu_2 - \mu_T)^2.$$

- For bi-level thresholding, Otsu confirmed that the optimal threshold t^* is selected so that the between-class variance σ_B is exploited; specifically,

$$t^* = \underset{1 \leq t \leq L}{\text{Arg Max}} \{ \sigma_B^2(t) \}$$

Performance (regarding our experiments): Otsu efforts well with some images and executes defectively with some. The majority of the consequences from Otsu have excessively much of noise in the outline of the environment being detected as foreground. Otsu can be utilized for thresholding if the noise elimination and character recognition completions are actually high-quality. The most important benefit is the effortlessness of estimate of the threshold. In view of the fact that it is a global algorithm it is compatible only for the images with equal intensities. This strength not gives a good quality effect for the images with assortments of variation in the intensities of pixels.

C. PROPOSED ARCHITECTURE

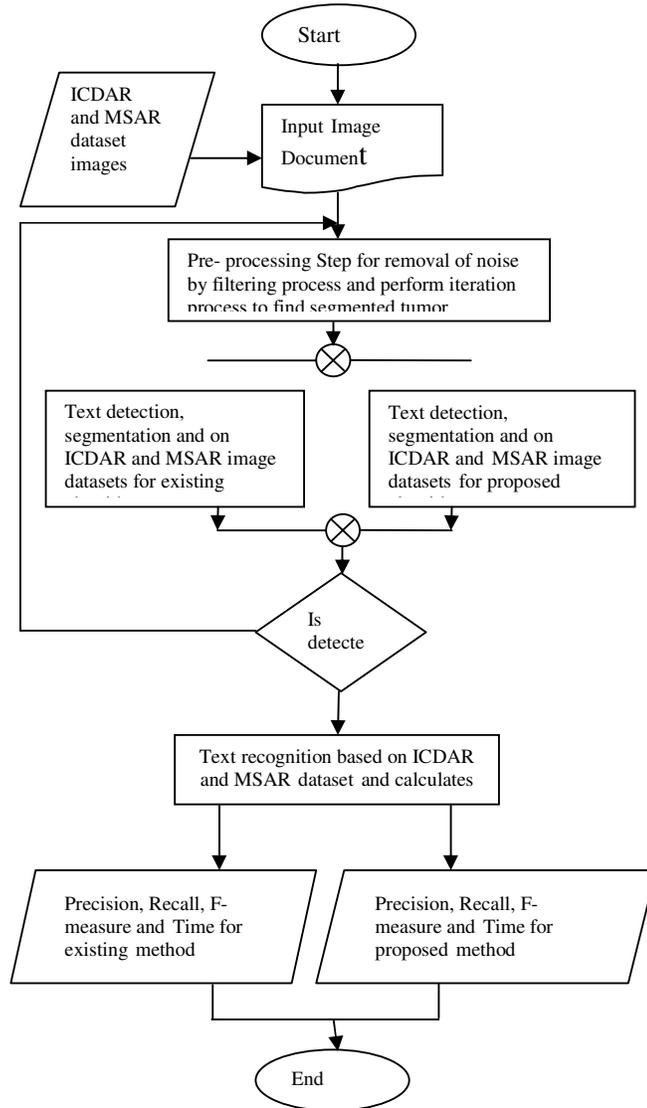


Figure 2: Proposed Architecture of our method

5. EXPERIMENTAL RESULTS

An experimental result is shown below is the analysis and comparison of existing text detection and the proposed text detection methodologies. Its performance based on ICDAR Dataset which consists of training and testing images. The existing and proposed methodology when applied on ICDAR Dataset provides precision and Recall and F-measure. Towards proposing new performance measures for comparing the outcomes of IR experiments (two such recent attempts can be found in ([18, 19])). The proposed methodology implemented here for the text detection provides more Precision & Recall as compared to the existing Text Detection Methodology.

Table 4.1 Analysis & Comparison on ICDAR Dataset

Algorithm	ICDAR Dataset		
	Precision	Recall	F-Measure
Proposed Work	0.5089	0.036	0.0675919
Existing Work	0.4911	0.036	0.0674296

The table exposed below is the examination and contrast of offered text detection and the suggested text detection methodologies. The Experimental results are performed on ICDAR Dataset which consists of training and testing images. The Existing and proposed methodology when applied on ICDAR Dataset provides Computational time in seconds. The proposed methodology implemented here for the text detection provides less Computational time as compared to the existing Text Detection Methodology.

Table 4.2 Analysis & Comparison of Time on ICDAR Dataset Images

No. of ICDAR Images	Time in sec	
	Existing Work	Proposed Work
10	8.4865	2.5272
20	17.6749	3.2916
30	20.5609	5.2884
40	26.7854	6.1776
50	28.0022	7.41
60	32.5886	8.3149
70	40.5915	8.9077
80	44.4603	7.7844
90	52.3071	11.2165
100	54.2103	14.7421

The table shown below is the analysis and comparison of existing text detection and the proposed text detection methodologies. The Experimental results are performed on MSRA Dataset which consists of training and testing images. The Existing and proposed methodology when applied on MSRA Dataset provides Computational time in seconds. The proposed methodology implemented here for the text detection provides less Computational time as compared to the existing Text Detection Methodology.

Table 4.3 Analysis & Comparison of Time on MSRA Dataset Images

No. of MSRA Images	Time in sec	
	Existing Work	Proposed Work
10	13.5565	4.5708
20	32.417	8.5801
30	40.1079	11.1853
40	66.4252	11.4349
50	66.5812	16.2397
60	82.1033	21.1693
70	104.4427	26.1614
80	94.0062	31.5122
90	128.264	32.8694
100	135.0345	32.027

The table shown below is the analysis and comparison of existing text detection and the proposed text detection methodologies. The Experimental results are performed on MSRA Dataset which consists of training and testing images. The Existing and proposed methodology when applied on MSRA Dataset provides precision and Recall and F-measure. The proposed methodology implemented here for the text detection provides more Precision & Recall as compared to the existing Text Detection Methodology.

Table 4.4 Analysis & Comparison on MSRA Dataset

Algorithm	MSRA Dataset		
	Precision	Recall	F-Measure
Proposed Work	0.6	0.5	0.5455
Existing Work	0.4	0.5	0.44

The table shown below is the analysis and comparison of existing text detection and the proposed text detection methodologies. The Experimental results are performed on MSRA & ICDAR Dataset which consists of training and testing images. The Existing and proposed methodology when applied on ICDAR & MSRA Dataset provides Computational time in seconds. The proposed methodology implemented here for the text detection provides less Computational time as compared to the existing Text Detection Methodology.

Table 4.5 Comparison of Computational Time on various Datasets

Algorithm	ICDAR Dataset	MSAR Dataset
Proposed Work	22.0429	39.8895
Existing Work	110.8855	277.3074

The figure shown below is the analysis and comparison of existing text detection and the proposed text detection methodologies. The Experimental results are performed on ICDAR Dataset which consists of training and testing images. The Existing and proposed methodology when applied on ICDAR Dataset provides precision and Recall and F-measure. The proposed methodology implemented here for the text detection provides more Precision & Recall as compared to the existing Text Detection Methodology.

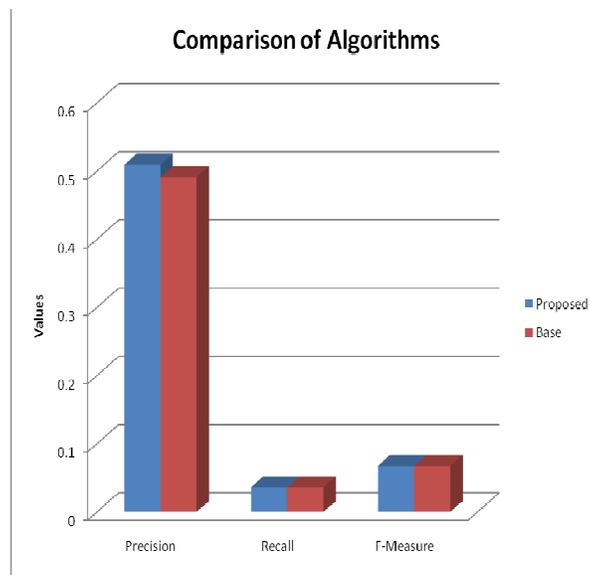


Figure 3: Analysis & Comparison on ICDAR Dataset

The figure shown below is the analysis and comparison of existing text detection and the proposed text detection methodologies. The Experimental results are performed on ICDAR Dataset which consists of training and testing images. The Existing and proposed methodology when applied on ICDAR Dataset provides Computational time in seconds. The proposed methodology implemented here for the text detection provides less Computational time as compared to the existing Text Detection Methodology.

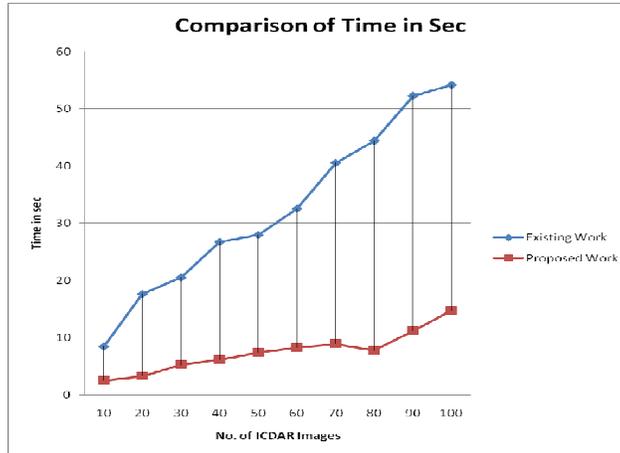


Figure 4: Analysis & Comparison of Time on ICDAR Dataset Images

The figure shown below is the analysis and comparison of existing text detection and the proposed text detection methodologies. The Experimental results are performed on MSRA Dataset which consists of training and testing images. The Existing and proposed methodology when applied on MSRA Dataset provides Computational time in seconds. The proposed methodology implemented here for the text detection provides less Computational time as compared to the existing Text Detection Methodology.

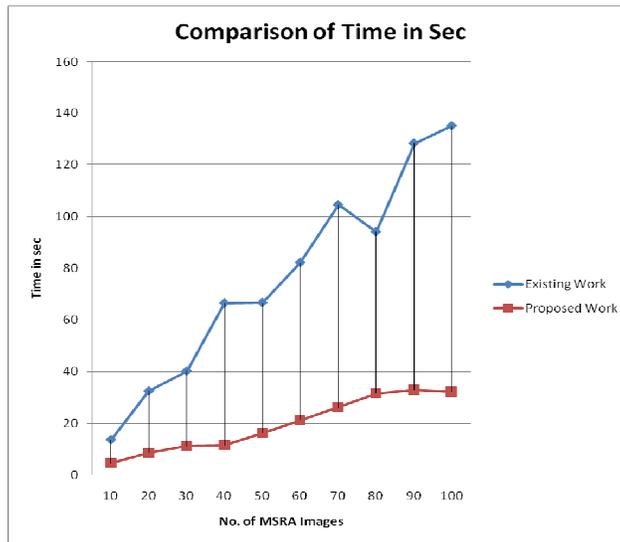


Figure 5: Analysis & Comparison of Time on MSRA Dataset Images

The figure shown below is the analysis and comparison of existing text detection and the proposed text detection methodologies. The Experimental results are performed on MSRA Dataset which consists of training and testing images. The Existing and proposed methodology when applied on MSRA Dataset provides precision and Recall and F-measure. The proposed methodology implemented here for the text detection provides more Precision & Recall as compared to the existing Text Detection Methodology.

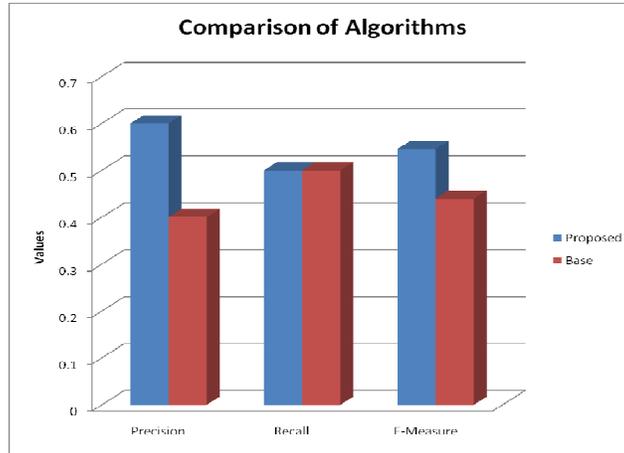


Figure 6: Analysis & Comparison on MSRA Dataset

The figure shown below is the analysis and comparison of existing text detection and the proposed text detection methodologies. The Experimental results are performed on MSRA & ICDAR Dataset which consists of training and testing images. The Existing and proposed methodology when applied on ICDAR & MSRA Dataset provides Computational time in seconds. The proposed methodology implemented here for the text detection provides less Computational time as compared to the existing Text Detection Methodology.

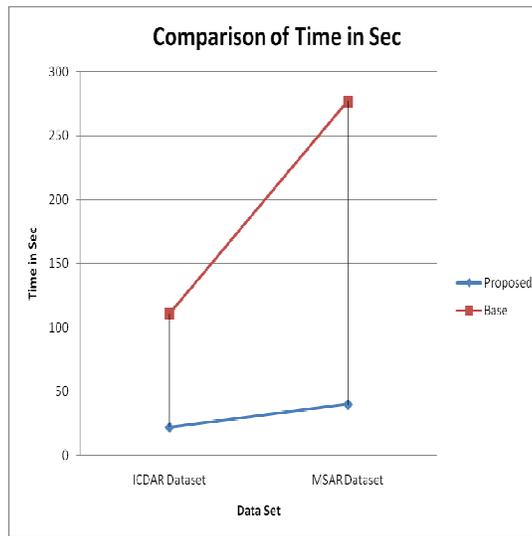


Figure 7: Overall Time Comparison of both dataset in (sec)

6. CONCLUSIONS

In this paper we have exhibited a interconnected outline for detection and recognition of multi-oriented view texts. Experimental analysis here is revealed that the proposed technique efforts well for low resolution text images also when it efforts for high resolution images. Sample results of the recommended and existing techniques are shown here that the proposed technique gives enhanced results than the existing techniques. The quantitative results of the proposed method and existing methods at image, line, word and character level given by both MSRA & ICDAR

Dataset where it is noticed that the proposed method at word gives better results compared to existing methods in terms of character recognition rate at all levels.

REFERENCES

- [1] R. K. Yadav, B. D. Mazumdar, "Detection of Bold and Italic Character in Devanagari Script", In: International Journal of Computer Applications, vol. 39, no. 2, pp. 19-22, 2012.
- [2] G. Seni and E. Cohen, "External Word Segmentation of Off-Line Handwritten Text Lines," Pattern Recognition, vol. 27, no. 1, pp. 41- 52, 1994.
- [3] Nallapareddy Priyanka, Srikanta Pal, Ranju Mandal, "Line and Word Segmentation Approach for Printed Documents", IJCA Special Issue on "Recent Trends in Image Processing and Pattern Recognition" RTIPPR, pp.31-33, 2010.
- [4] X. Chen and A. Yuille, "Detecting and reading text in natural scenes," in Proc. of CVPR, 2004.
- [5] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform," in Proc. of CVPR, 2010.
- [6] Y. Pan, X. Hou, and C. Liu, "A hybrid approach to detect and localize texts in natural scene images," IEEE Trans. Image Processing, vol. 20, no. 3, pp. 800–813, 2011.
- [7] H. Chen, S. S. Tsai, G. Schroth, D. M. Chen, R. Grzeszczuk, and B. Girod, "Robust text detection in natural images with edge-enhanced maximally stable extremal regions," in Proc. of ICIP, 2011.
- [8] A. Mosleh, N. Bouguila, and A. B. Hamza, "Image text detection using a bandlet-based edge detector and stroke width transform," in Proc. Of BMVC, 2012.
- [9] J. J. Weinman, "Unified detection and recognition for reading text in scene images," Ph.D. dissertation, University of Massachusetts Amherst, MA, 2008
- [10] K. Wang, B. Babenko, and S. Belongie, "End-to-end scene text recognition," in Proc. of ICCV, 2011.
- [11] G. Seni and E. Cohen, "External Word Segmentation of Off-Line Handwritten Text Lines," Pattern Recognition, vol. 27, no. 1, pp. 41- 52, 1994.
- [12] Nallapareddy Priyanka, Srikanta Pal, Ranju Mandal, "Line and Word Segmentation Approach for Printed Documents", IJCA Special Issue on "Recent Trends in Image Processing and Pattern Recognition" RTIPPR, pp.31-33, 2010.
- [13] Azadboni, M.K. Behrad, A. "Text detection and character extraction in color color images using FFT domain filtering and SVM classification", 2012 Sixth International Symposium on Telecommunications (IST), 2012 IEEE ,pp 794-799. DOI: 10.1109/ISTEL.2012.6483094.
- [14] Weilin Huang, Zhe Lin , Jianchao Yang ,Jue Wang, "Text Localization in Natural Color images Using Stroke Feature Transform and Text Covariance Descriptors", International Conference on Computer vision (ICCV), IEEE ,pp 1241-1248, 2013.
- [15] S. A. Angadi & M. M. Kodabagi, "Text Region Extraction from Different texts Natural Scene Color images using Texture Features", International Journal of Color image Processing, vol. 3, issue 5, pp. 229-245, 2009.
- [16] Kwang In Kim, Keechul Jung, And Jin Hyung Kim, "Texture-Based Approach For Text Detection In Color images Using Support Vector Machines And Continuously Adaptive Mean Shift Algorithm" , IEEE Transactions On Pattern Analysis And Machine Intelligence , Vol. 25, No. 12, 2003.
- [17] Andrej Ilic, Peter Peer, "An improved edge profile based method for text detection in color images of natural scenes", EUROCON-International Conference on Computer as a tool (EUROCON) IEEE: 1-4, 2011.
- [18] Mizzaro, S.: A new measure of retrieval effectiveness (or: What's wrong with precision and recall). In: In T. Ojala editor, International Workshop on Information Retrieval (IR'2001). (2001)
- [19] Jarvelin, K., Kekalainen, J.: Cumulated gain-based evaluation of IR techniques. ACM Transactions on Information Systems 20 (2002).