International Journal of Computing Science and Information Technology, **2015, Vol. 3, Iss. 3, 09-17**
ISSN: 2278-9669, July 2015 (http://ijcsit.org)

A Framework for Text Extraction from Document Images

[1]Vikas K. Yeotikar, [2]Manish T. Wanjari, [3]Dr. Mahendra P. Dhore
[1,2]Department of Computer Science, SSESA's, Science College Congress Nagar, Nagpur (MH), India
[3]Department of Electronics & Computer Science, RTM Nagpur University, Nagpur (MH), India
[1]vkyeotikar@gmail.com, [2]mwanjari9@gmail.com, [3]mpdhore@rediffmail.com

Abstract

We propose a Framework by using MATLAB with three text extraction technique i.e. Gabor, Wavelet and Hough to detect text from document images. Extraction of the information in the form of text involves detection, localization, tracking, extraction, enhancement, and recognition of the text from a given document image. A large number of techniques have been proposed to address this problem and the purpose of this paper is to review and implementation of our method to evaluate the performance of the proposed method on ICDAR_UK and MSRA_TD500 dataset.

Keywords: Document Image Analysis (DIA), Text Extraction, Text Detection, Text Localization, Text Enhancement, Gabor, Wavelet, Hough and Canny.

1. Introduction

Text Extraction from images is concerned with extracting the relevant text data from a collection of images. Rapid development of digital technology has resulted in digitization of all categories of materials. Lot of resources is available in electronic medium. Many existing paper-based collections , historical manuscripts , records, books, journals, scanned document , book covers , video images, maps, manuscripts, pamphlets, posters, broadsides, newspapers,, micro facsimile, microfilms, university archives, slides and films, book plates, pictures, painting, graphic materials, coins and currency, stamps, magazines, clipping files, educational , TV programs , business card, magazines, advertisements, web pages , mixed text-picture-graphics regions etc are converted to images. These images present many Challenging research issues in text extraction and recognition. Text extraction from images have many useful applications in document analysis , detection of vehicle license plate, analysis of article with tables, maps, charts, diagrams etc., keyword based image search, identification of parts in industrial automation , content based retrieval, name plates, object identification, street signs, text based video indexing, video content analysis, page segmentation, document retrieving, address block location etc.[1]

The features can be divided into two main groups: the first is related to local features, according to which one feature is extracted for each point in the input domain, in the second group global features are evaluated on sets of pixels (e.g. a word), on a region or even on the whole document. [2]

1.1. Pixel level

When features are computed at a local level some values are obtained for each pixel. In Leydier et al. propose a word-spotting method to access the textual data of medieval manuscripts. This approach does not require image binarization and layout segmentation and is tolerant to low resolution and image degradations. The informative parts of the images are represented through a set of features provided by gradient orientation.

1.2. Column level

Some approaches require a segmentation phase, such as the segmentation of words and characters. In this case a method based on the analysis of column pixels in segmented objects can be exploited. The segmentation of words and characters is done by finding the connected components and, for each pixel column of the character, a set of six features is calculated: vertical projection profile on the gray level image, upper character profile position, lower character profile position, vertical histogram, number of ink/non-ink transitions and middle row transition state.

1.3. Sliding window

One technique related to column level representation adopts a sliding window. In this case a fixed size window is moved across the word image and some features are evaluated for each position. This strategy is frequently used to obtain the input descriptors for supervised classifiers such as the Multilayer Perception (MLP) neural network.

1.4. Stroke and primitive level

When the objects in document images are complex and important spatial relationships among primitives are possible, such as in sketches and in trademarks, one structural representation, which is able to represent the variety of connections, is essential propose the combination of structural and global features: the structural part describes the interconnections among primitives and the global features reflect the object as a whole.

1.5. Connected-component level

In the processing of handwritten or ancient printed documents, it is not always easy to segment a document and to identify text lines, words, and characters. Especially the character segmentation is a difficult task because of the variability of handwriting and the presence of touching characters.

1.6. Word level

In most word spotting applications it is possible to assume that the word segmentation in indexed documents is not problematic. In this case the retrieval is carried out considering the word as a whole.

1.7. Line and Page level

Deal with the script identification among three different on-line handwritten scripts: Arabic, Roman and Tamil. After the detection of text lines, they extract a set of features at line-level such as the horizontal and vertical inter-stroke direction, horizontal and vertical stroke direction, average stroke length, stroke density and the reverse direction. Some features can be extracted at page level by means of geometric transformations develop a system for paleographers and literary experts, to support their work on manuscripts dating and authentication through different historical periods. The approach is based on the Curve let transform to compose a unique signature for each handwritten page.

1.8. Shape Descriptor

Shape descriptors are frequently used in image analysis to compare 2D object silhouettes. Recently they have been adopted also in document image analysis to compare symbol images in a recognition-free approach. According to the object representation the shape descriptors are evaluated on, three main categories can be identified. In the first category contour based descriptors are evaluated on the object contours; in the second, image based descriptors include the shape descriptors based on the overall image pixel values; in the last category, skeleton based descriptors are evaluated on the image skeletons. [2]

2. Related Work

Angadi et.al [9] proposed a methodology to detect and extract text regions from low resolution natural scene images. Their proposed work used Discrete Cosine Transform (DCT) based high pass filter to remove and suppress the constant background. The texture feature matrix was computed on every 50x50 block of the processed image. A newly defined discriminant function was used to classify text blocks. The detected text blocks were merged to obtain new text regions. Finally, the refinement phase was a post processing step used to improve the detection accuracy. This phase used to cover small portions of missed text present in adjacent undetected blocks and unprocessed regions. The proposed methodology had been conducted on 100 indoor and outdoor low resolution natural scene images containing text of different size, font, and alignment with complex backgrounds containing Kannada text and English text. The approach also detected nonlinear text regions and can be extended for text extraction from the images of other languages with little modifications.

Pan et.al[8] proposed a novel hybrid method where in a text region detector was designed to generate a text confidence map. A Local binarization approach was used to segment the text components using text confidence map. A Conditional Random Field (CRF) model was used to label components as text or non-text which was solved by minimum classification error (MCE) learning and graph cuts inference algorithm. A learning based method by building neighbouring components into minimum spanning tree (MST) and cutting off interline edge with an energy minimization model to group the text components into text lines.

Fabrizio et.al[7] offered a region based approach that starts by isolating letters, then groups them to restore words. The process was based on a new segmentation method based on morphological operator called Toggle Mapping Morphological Segmentation (TMMS) and a classification step based on a combination of multiple SVM classifiers. The training data base composed of 32400 examples extracted from various urban images and different configurations of classifiers have been tested to get the highest classification accuracy.

Kohei et.al [3] introduced a new approach to detect and extract text from commercial screenshot images. Their approach implemented edge-based method and connected component labeling method known as blob extraction method. Combination of homogeneity edge detection filter and appropriate threshold number separated the text from the image.

A method for localizing text regions within scene images was introduced by Luz et.al [5]. A set of potential text regions was extracted from the input image using morphological filters. Connected Components (CC) were identified using ultimate attribute openings and closings, and selected a subset of text region after combining some of the CCs. Decision tree classifier were used to distinguish text or non-text regions.

Shivakumara et.al[4] proposed a new method based on Maximum Color Difference (MCD) and Boundary Growing Method (BGM) for detection of multioriented handwritten scene text from video. Average of RGB channel was calculated of the original frame to sharpen the text edges and increase the contrast of text pixels. .Maximum Color Difference was computed to increase the gap between text and non-text pixel. Text clusters were obtained by K-means clustering algorithm .These clusters were used to obtain the text candidates and also help in eliminating false positives. To fix boundary for handwritten text, Boundary Growing Method (BGM) based on the nearest neighbor concept was used. The method made to appear the characters and words in regular spacing in one direction and it can grow based on orientation of text. The concept of intrinsic and extrinsic edges was used to eliminate false positives.

The unique approach of Shyamaet.al [6] projected a text segmentation technique to extract text from any type of camera grabbed frame image or video. Color based segmentation methodology was used to link consecutive pixels in the same direction by exploiting the general text properties. Light Edge Enhancement (LEE) was used to find a set of consecutive candidate points and enhance the edge between them. Next, heavy edge enhancement (HEE) was applied to remove or reduce motion blur from camera image sequences. This helped to treat camera images and video frames in the same manner.

3. Methodology

Document images are acquired by scanning journal, printed document, degraded document images, handwritten historical document, and book cover etc. The text may appear in a virtually unlimited number of fonts, style, alignment, size, shapes, colors, etc. Extraction of text from text document images and from complex color background is difficult due to complexity of the background and mix up of colors of fore-ground text with colors of background. In this section, we present the main ideas and details of the proposed algorithm.

The task of Document Analysis may encounter various types of input sources such as Printed character, Text, Drawing images, Magazine objects, Newspaper objects, Color objects, etc. Thus, the aim is to improve the range and volume of publications such as newspaper, magazines, various types of manuals, other documents needs to be processed through various computer aided text processing techniques. Large amount of documentation needs to be converted into a computer readable format to avoid data entry. Document Analysis meets this need. The methodology followed is in following sequence.

• The Document Analysis
• Document Understanding
• Character /Word recognition

Implementation of any system needs the study of features, it may be symbolic, numerical or both. An example of a symbolic feature is color; an example of numerical feature is weight. Features may also result from applying a text extraction algorithm or operator to the input data. The related problems of feature selection and feature extraction must be addressed at the outset of any text recognition system design. The key is to choose and to extract features that are computationally feasible and reduce the problem data into a manageable amount of information without discarding valuable information.

Different methods used for text extraction from document images (as shown in Proposed Algorithm) include:

Algorithm of Proposed Method:

Input: Image Recognition f(x,y) $1\leq x\leq M$, $i\leq y\leq N$
Output: Recognize Characters Array $C_R(i)$ $0\leq i\leq NC$
Method:
Step 1: Read an Input Image f(x,y)
img = imread (FileName) // image file name
Step 2: if isrgb (img)
img1 = rgb2gray (img) // If img is RGB then convert to Gray
else
img1 = img ;
end
Step 3: if isrgb (img)
// Extract color features
R = colorfeature (img);
G = colorfeature (img);
B = colorfeature (img);
else
end
Step 4: Texture = TextureFeature (img1); // Texture Feature Extraction

Step 5: NOC = Number of classes;
var (i) = variance of class i
p (i) = probability of class i
IC = ExpMax (img1, NOC, var(i), p(i));
Step 6: Create Gabor filter bank
// Gabor & Wavelet Algorithm
Calculate Gabor-Wavelet Features.
Step 7: Apply Hough Transform for Line Detection.
Step 8: Combine Image Classification, Gabor-Wavelet features, Hough Transform and Canny to Extract Characters.
Step 9: Extract Characters from the image and display each one by one.
Step 10: Calculate Tp, fp, fn, Precision, Recall and F-measure and display results.
Step 11: STOP

3.1. Feature Extraction

Feature extraction involves the extracting the meaningful information from the document image. The features are classified in to Global features and Local features. Features that are extracted from whole image are known as the global features and the features that are extracted from blocks identified during segmentation or from subdivision of the document are known s local features. They can be divided into several categories: textural, geometric, component, structural and content based. The extractions of global and local features provide input to classification algorithm/techniques.

3.2 Color Extraction (RGB)

There are three types of images which include Binary image, gray scale and RGB(color) image. Binary image consists of Logical array containing only 0s and 1s, interpreted as black and white, respectively. Grayscale image is also known as an intensity, gray scale, or gray level image. Array of class uint8, uint16, int16, single, or double whose pixel values specify intensity values. For single or double arrays, values range from [0, 1]. For uint8, values range from [0,255]. For uint16, values range from [0, 65535]. For int16, values range from [-32768, 32767]. True color image is also known as an RGB image. A true color image is an image in which each pixel is specified by three values one each for the red, blue, and green components.

3.3. Classification (EM-Algorithm)

The EM algorithm is an iterative algorithm for calculating the maximum-likelihood or maximum-a-posterior estimates when the observations can be viewed as incomplete data. Each iteration of the algorithm consists of an expectation step followed by a maximization step.

We now define the EM algorithm, starting with cases that have strong restrictions on the complete-data specification f(x | Φ), then presenting more general definitions applicable when these restrictions are partially removed in two stages. The simplicity of description and computational procedure, and thus the appeal and usefulness, of the EM algorithm are greater at the more restricted levels. [21]
Suppose first that f(x | Φ) has the regular exponential-family form:
$f(x \mid \Phi) = b(x) \exp (\Phi t(x)^T) / a(\Phi)$,

3.4. Gabor Filter Method

The Gabor Transform is also referred to as the Short Time Fourier Transform (STFT). Filtering the time-frequency content of a signal is indeed one of the main applications of Gabor multipliers. Gabor analysis is an essential part of time-frequency analysis, initiated by the seminar paper of Denis Gabor in 1946.

Some properties of Gabor filters:

- A tunable bandpass filter
- Similar to a STFT or windowed Fourier transform
- Satisfies the lower-most bound of the time-spectrum resolution (uncertainty principle)
- It's a multi-scale, multi-resolution filter
- Has selectivity for orientation, spectral bandwidth and spatial extent.
- Has response similar to that of the Human visual cortex (first few layers of brain cells)
- Computational cost often high, due to the necessity of using a large bank of filters in most applications [10-11]

3.5. Wavelet Transform Method

Wavelets are functions that satisfy certain mathematical requirements and are used in presenting data or other functions, similar to sines and cosines in the Fourier transform. However, it represents data at different scales or resolutions, which distinguishes it from the Fourier transform. Wavelet transform is an increasingly popular tool in computer vision and image processing. Many applications, such as compression, detection, recognition, image retrieval have been investigated. Wavelet transform has nice features of space-frequency localization and multi-resolutions. The wavelet transform of a 1-D signal *f(x)* is defined as:

$$(\mathbf{W_a f})(\mathbf{b}) = \int \mathbf{f}\,(\mathbf{x})\,\boldsymbol{\Psi a, b}\,(\boldsymbol{x})\boldsymbol{dx}$$

With

$$\boldsymbol{\Psi}\ \boldsymbol{a}, \boldsymbol{b}\ (\boldsymbol{x}) = \frac{1}{\sqrt{a}}\ \boldsymbol{\Psi}\left(\frac{x-b}{a}\right)$$

International Journal of Computing Science and Information Technology, **2015, Vol. 3, Iss. 3, 09-17**
ISSN: 2278-9669, July 2015 (http://ijcsit.org)

The mother wavelet Ψ has to satisfy the admissibility criterion to ensure that it is a localized zero-mean function.[12-13]

3.6. Hough Transform Method

Hough Transform technique is used to extract text from document images. Hough Transform (HT) is recognized as a powerful tool for graphic element extraction from images due to its global vision and robustness in noisy or degraded environment. The method herein proposed detects text lines on document images which may include either lines oriented in several directions, erasures, or annotations between main lines. At each stage of the process, the best text-line hypothesis is generated in the Hough Transform domain.

The Hough transform is a feature extraction technique used in image analysis, computer vision, and digital image processing. [14-15]

3.7. Edge Detection Method (Canny)

The edge representation of a document image significantly reduces the quantity of data to be processed; it retains necessary information regarding the shape of character in document image. There are many edge detection methods in the literature for document images. Most of the used discontinuity based edge detection methods are reviewed. Those methods are Prewitt, Sobel, Canny, Roberts, Zerocross and Laplacian of Gaussian. [16] The Canny edge detector is regarded as one of the best and standard edge detectors recently in use; Canny's edge detector ensures good noise immunity and at the same time detects true edge points with minimum error. The Canny edge detector is an edge detection operator that uses a multi-stage algorithm to detect a wide range of edges in document images. It is developed by John Canny considered the mathematical problem of deriving an optimal smoothing filter given the criteria of detection, localization and minimizing multiple responses to a single edge. [17]

Framework of Proposed Method:

This research work presents new approach to text extraction from Document images. For this purpose the research work is organized in following way.

a. For Experimentation, Document images are taken from ICDAR_UK and MSRA_TD500 databases.
b. Hierarchical structure of document images in logical way.
c. Convert it into Gray Image if image is in color format.
d. Extracting RGB values if document image is in color format.
e. Extraction of feature of document image.
f. Apply Expectation Maximum Algorithm for Classification.
g. Apply our methodology for text extraction.
h. Experimental results are calculated from following Framework.

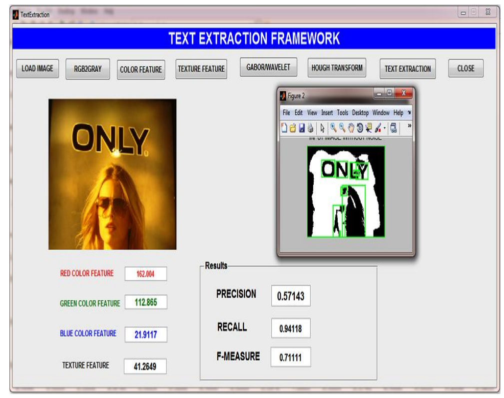


Fig 1: Framework of our proposed method

4. Experiments & Results

The goal of our proposed method is text extraction that achieves the highest recognition accuracy and fast performance. In this section, we demonstrate the efficiency of the proposed method on 20 Document Images of ICDAR_UK and MSRA_TD500 databases, evaluation and success criteria, experiments, comparison and results. For that we used following Mathematical and Statistical methods:

1) **F_N (False Negative)**
False Negatives (FN) / Misses are those regions in the image which are actually text characters, but have not been detected by the algorithm.
2) **F_P (False Positive)**
False Positives (FP) / False alarms are those regions in the image which are actually not characters of a text, but have been detected by the algorithm as text.
3) **T_P (Total Positive)**
Total Positive (T_P) is the correctly detected characters.
4) **Precision**
Precision rate (P) is defined as the ratio of correctly detected characters to the sum of correctly detected characters plus false positives.
Eq(1)…
$$P = \frac{TP}{TP+FP}$$
5) **Recall**

Recall rate (R) is defined as the ratio of the correctly detected characters to sum of correctly detected characters plus false negatives.

Eq(2)…

$$P = \frac{TP}{TP+FN}$$

6) **F-Measure**

F-Measure is the harmonic mean of recall and precision rates.

Eq(3)…

$$F = \frac{2*PR}{P+R}$$

ICDAR-UK Database

ICDAR 2003/2005 text locating competition dataset is the most widely used benchmark for scene text detection. The dataset contains 258 training and 251 test images with various sizes from 307 x93 to 1280 x 960.[18-19]

MSRA-TD500 Database (Multilingual Image Dataset**)**

MSRA Text Detection 500 Database (MSRA-TD500) is collected and released publicly as a benchmark to evaluate text detection algorithms, for the purpose of tracking the recent progresses in the field of text detection in natural images, especially the advances in detecting texts of arbitrary orientations.

MSRA Text Detection 500 Database (MSRA-TD500) contains 500 natural images, which are taken from indoor (office and mall) and outdoor (street) scenes using a packet camera. The indoor images are mainly signs, doorplates and caution plates while the outdoor images are mostly guide boards and billboards in complex background. The resolutions of the images vary from 1296x864 to 1920x1280. [20]

4.1 Sample Results on ICADAR-UK Database

Fig 2a : Original Image

Fig 2b : RGB to Gray Image

Red = 123.473
Green = 123.6
Blue = 58.733
Texture = 35.618
2c: Features Extracted of 2a Image

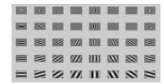


Fig 2d : Real Part of Gabor Filters

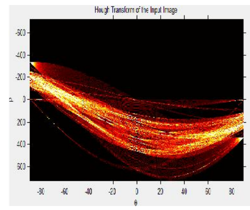


Fig 2e: Hough Transform of 2a Image

Fig 2f: Detected Text of of 2a Image

Fig 2g: Resulting Text Extracted

4.2 Sample Results on MSRA-TD500 Database

Fig 3a : Original Image

Fig 3b : RGB to Gray Image

Red = 162.004
Green = 112.865
Blue = 21.911
Texture = 41.264
3c: Features Extracted of 3a Image

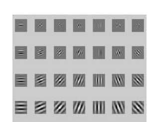


Fig 3d : Real Part of Gabor Filters

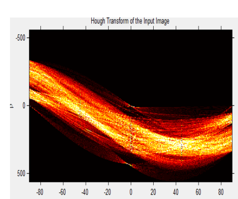


Fig 3e: Hough Transform of 3a Image

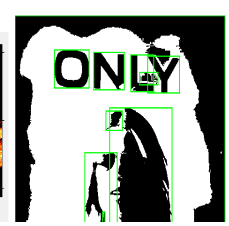


Fig 3f: Detected Text of 3a Image

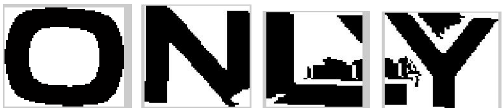


Fig 3g: Resulting Text Extracted

4.3 Some Sample Results on both Databases

Fig 4a: Original Image

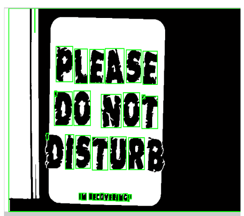


Fig 4b: Detected Text Image

Fig 5a: Original Image

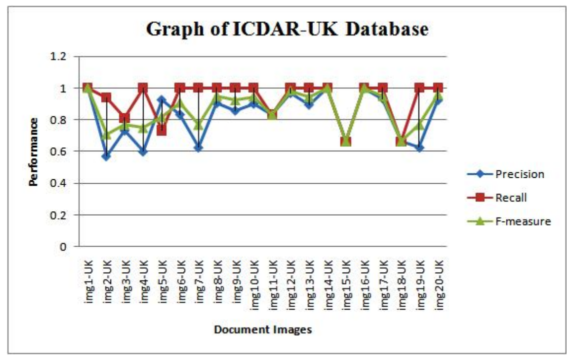
Fig 5b: Detected Text Image

Table 1: Results Calculated on ICDAR-UK Database

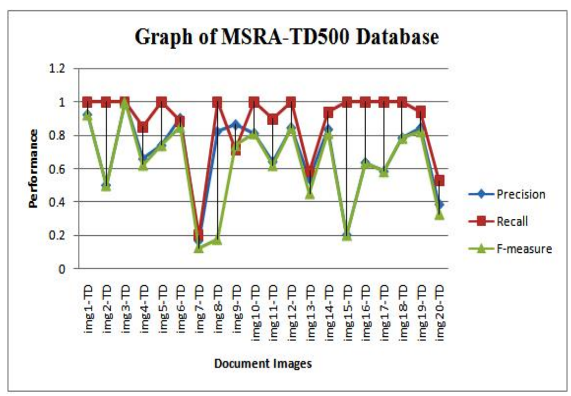
F-Name	Precision	Recall	F-measure
img1-UK	1	1	1
img2-UK	0.571	0.941	0.711
img3-UK	0.732	0.812	0.770
img4-UK	0.6	1	0.75
img5-UK	0.925	0.735	0.818
img6-UK	0.833	1	0.909
img7-UK	0.625	1	0.769
img8-UK	0.904	1	0.95
img9-UK	0.857	1	0.923
img10-UK	0.9	1	0.947
img11-UK	0.833	0.833	0.832
img12-UK	0.967	1	0.983
img13-UK	0.894	1	0.944
img14-UK	1	1	1
img15-UK	0.666	0.666	0.665
img16-UK	1	1	1
img17-UK	0.933	1	0.949
img18-UK	0.666	0.666	0.665
img19-UK	0.625	1	0.769
img20-UK	0.923	1	0.959

Table 2: Results Calculated on MSRA-TD500 Database

F-Name	Precision	Recall	F-measure
img1-TD	0.926	1	0.926
img2-TD	0.5	1	0.5
img3-TD	1	1	1
img4-TD	0.659	0.852	0.623
img5-TD	0.740	1	0.740
img6-TD	0.903	0.882	0.852
img7-TD	0.166	0.2	0.125
img8-TD	0.823	1	0.176
img9-TD	0.865	0.717	0.740
img10-TD	0.812	1	0.812
img11-TD	0.642	0.9	0.620
img12-TD	0.846	1	0.846
img13-TD	0.538	0.583	0.451
img14-TD	0.837	0.939	0.815
img15-TD	0.2	1	0.2
img16-TD	0.636	1	0.636
img17-TD	0.583	1	0.583
img18-TD	0.785	1	0.785
img19-TD	0.846	0.942	0.825
img20-TD	0.384	0.526	0.327



Graph 1: Performance on ICDAR-UK Database



Graph 2: Performance on MSRA-TD500 Database

Table3: Performance Analysis with different Methods

S.N	Author	Year	Accuracy
1	Kohei et.al [14]	2011	94.66%
2	Shivakumaraet.al[16]	2010	89.67%
3	Luz et.al [15]	2010	85.93%
4	Shyama et.al [17]	2009	94%
5	Fabrizio et.al[13]	2009	88.83%
6	Pan et. Al [12]	2009	83.44%

7	Angadi et.al [11]	2009	86.60%
8	OUR METHOD on ICDAR-UK dataset	2015	90.56%
9	OUR METHOD on MSRA-TD500 dataset	2015	84.33%

The performance evaluation of our method on ICDAR-UK and MSRA-TD500 dataset is listed in Table 2 along with the best algorithms reported and the algorithms proposed. We can see that the proposed text detection method achieved better performance than most of the listed algorithms/techniques.

5. Conclusion & Future Work

In this paper, we have presented a unified framework for detection and recognition of text from document images. The proposed method is evaluated on the dataset (ICDAR_UK and MSRA_TD500). Text detection and recognition are accomplished concurrently with exactly the same features and classification scheme. The proposed system is capable of detecting and recognizing texts of different scales, colors, fonts and orientations, in diverse real-world. Extensive experiments demonstrate that compared to existing methods in the literature the proposed algorithm achieves state-of-threat or very competitive performance on various challenging benchmarks.

Future Work

The proposed algorithm can assist numerous applications that require text information extraction from images. In the future, we will devote ourselves to the development of such practical systems for text information extraction from videos, such as video search, target relocation, and automatic navigation, based on the proposed algorithm.

There are several directions where this work can be extended. Another avenue for research would be to implement other text extraction technique on the same data set. In future, two or more Classification technique can be combined to achieve better results. In future this research work will be extended to improve the results of the algorithms which have been used for further analysis of document images.

REFERENCES

[1] C.P. Sumathi1, T. Santhanam and G.Gayathri Devi "A Survey on Various Approaches Of Text Extraction In Images" *International Journal of Computer Science & Engineering Survey (IJCSES) Vol.3, No.4, August 2012.*

[2] Liang, J., Rogers, R. and Chanda, B. 1999. The prototype of a complete Document Image Understanding System., *Intelligent system Laboratory, University of Washington.*

[3] Kohei Arai1 and Herman Tolle. "Text Extraction From Tv Commercial Using Blob Extraction Method", *International Journal Of Research And Reviews In Computer Science Vol. 2, No. 3, 2011.*

[4] Shivakumara P, A Dutta, U Pal And C L Tan. "A New Method For Handwritten Scene Text Detection In Video", *International Conference On Frontiers In Handwriting Recognition, pp: 16-18, 2010.*

[5] Wonder Alexandre Luz Alves And Ronaldo Fumio Hashimoto. "Text Regions Extracted From Scene Images By Ultimate Attribute Opening And Decision Tree Classification", *Proceedings of the 23rd Sibgrapi Conference On Graphics, Patterns And Images 2010.*

[6] Shyama Prosad Chowdhury, Soumyadeep Dhar, Karen Rafferty, Amit Kumar Das and Bhabatosh Chanda. "Robust Extraction Of Text From Camera Images Using Colour And Spatial Information Simultaneously",*Journal Of Universal Computer Science,Vol. 15, No.18 , pp:3325-3342, 2009.*

[7] J. Fabrizio, M. Cord and B. Marcotegui. "Text Extraction From Street Level Images,", *CMRT, Vol. Xxxviii, Part 3/W4 , pp. 199–204, 2009.*

[8] Yi-Feng Pan, Xinwen Hou and Cheng-Lin Liu. "Text Localization in Natural Scene Images Based On Conditional Random Field," *ICDAR, pp 6-10, 2009.*

[9] S. A. Angadi and M. M. Kodabagi. "A Texture Based Methodology For Text Region Extraction From Low Resolution Natural Scene Images", *International Journal Of Image Processing (Ijip) Vol 3, Issue 5, 2009.*

[10] Simona E. Grigorescu, Nicolai Petkov, and Peter Kruizinga; Comparison of Texture Features Based on Gabor Filters; *IEEE Transactions on Image Processing, Vol. 11, No. 10, OCT. 2002; pp 1160-1167.*

[11] V. K. Yeotikar, M. T. Wanjari & Dr. M. P. Dhore, "A Novel Approach for Text Extraction from Document Images", International Journal of Research in Computer Engineering and Electronics, Vol. 3, Issue 6 (Nov.-Dec.-14), ISSN 2319-376X.

[12] E. J. Stollnitz, T. D. DeRose and D. H. Salesin" Wavelets for computer graphics: a primer, part I,"*IEEE Computer Graphics and Applications, vol.15, No. 3, pp. 76-84, May 1995.*

[13] V. K. Yeotikar, M. T. Wanjari & Dr. M. P. Dhore,"Text Extraction of Document Images using Wavelet Transform", Proceeding of International Conference on Science, Technology and Mathematics Education-Trends and Alignment with Industry Needs-TechEd-2015.

[14] Duda R. O. and P. E. Hart, "Use of the Hough Transformation to detect Lines and Curves in

Pictures," *Comm. ACM, Vol. 15, pp. 11-15, Jan-1972.*

[15] V. K. Yeotikar, M. T. Wanjari, J. K. Keche & Dr. M. P. Dhore,"Text Extraction from Document Images Based on Hough Transform Technique", Proceeding of National Conference on ATCON-2015.

[16] Muthukrishnan.R and M.Radha (Dec. 2011). Edge Detection Techniques For Image Segmentation. *International Journal of Computer Science & Information Technology (IJCSIT) Vol 3, No 6.*

[17] M. T. Wanjari and Dr. M. P. Dhore, "Segmenting Document Images using Edge Detection Technique", *Proceeding of NCRAECS-2014, pp 249-253, ISBN 978-81-906457-5-1.*

[18] Datasets-Algoval-University of Essex. *[Online]. Available:http://algoval.essex.ac.uUicdar/Datasets.html, accessed Jul. 16, 2014.*

[19] ICDAR 2N5 Competitions: Text Locating-Algoval.[Online].Available: *h@://algoval.essex.ac.uk:8080/icdar2005/ indexjsp?page=textlocate.html, accessed Jul. 16, 2014.*

[20] MSRA-TD500, C. Yao, X. Bai, W. Liu, Y. Ma and Z. Tu. Detecting Texts of Arbitrary Orientations in Natural Images. *CVPR , pp. 1083-1090, 2012.*

[21] P. DEMPSTERN,. M. LAIRDa nd D. B. RubIN, "Maximum Likelihood from Incomplete Data via the EM Algorithm" *Journal of the Royal Statistical Society. Series B (Methodological), Vol. 39, No. 1. (1977), pp. 1-38.*