Scene Text Segmentation and Recognition by Applying Trimmed Median Filter Using Energetic Edge Detection Schemes and OCR

Beula Bell . T

Research Scholar, Manomaniam Sundaranar University, Tirunelveli 627012, Tamil Nadu, India.

Dr. M. K. Jeya Kumar

Professor, Dept. Of Computer Applications, Noorul Islam University, Kumaracoil 629180, Tamil Nadu, India.

Abstract – This proposed method is an accurate and strong method for detecting texts in natural scene images. There are many cases that text regions are not clearly notable by background regions due to brightness or illumination variations. The proposed scene text finding process finds out the scene text contents from the natural scene images using the sophisticated edge revealing methods, the local directional number feature and linked map generation process. This proposed method takes natural scene image as input and detects the scene text regions. The detected scene text regions are marked for visual identification for human eyes.Text recognition is done by OCR .

Index Terms – Noise Reduction, Energetic edge detection, Local directional number, Linked map, Non seen text rejection.

1. INTRODUCTION

Natural scene images are camera captured scenes from Nature, which are naturally affected due to bad weather, poor lightings and grimy environment. Natural scene imagery produced massive capacity of Meta information when it was acquired by camera. Content matching in news reading, video set detection, content examination by text matching and extraction of bio medical image text removal are required a tool for programmed text removal from natural scene images. The main challenge in text detection in natural scene image is that the document holds graphics and text information. So the document image has to be segmented to identify text and image regions so that suitable techniques to segment the image and text regions should be processed. The improvement in efficiency and accuracy of natural scene text detection stage can be used as a processing component to increase the text recognition results when applying OCR system.

There are lot of researches are go through on the natural image text segmentation in the recent years. There are various classes of methods for text localization and they are connected component methods, texture methods, threshold methods and line structural methods. The connected component methods [1] [3] [4] employed connected component analysis of the geometrical arrangement of edges, homogeneous color and greyscale components. The implementation methods are not too hard but the robustness of these methods is less when it faces the complicated background.

Threshold methods have been developed to segment characters in document images with relatively simple backgrounds. In this kind of method, Niblack's local threshold method is proved to be the best one [2]. The model and system for separating freely written texts into text line and estimating the line direction and character orientation was reported in [5][6][7].

The papers [16] and [17] express the Scene text detection methods based on transition maps.

Many previous text localization algorithms applied color-based clustering to group the pixels in like colors into own color layers. Thus, text characters and strings could be alienated from the background objects in different colors. Nikolaou *et al*. and Manza proposed an algorithm of color reduction based on color histogram and mean-shift algorithm. It initialized color centers arbitrarily and projected each pixel to the nearest color center. Then mean-shift algorithm was applied to fix the color centers into the mean positions as last color layers [20][21]. Chen *et al*. established Gaussian mixture model (GMM) in five color channels (red, green, blue, hue, and intensity) to analyze the distributions of text pixels and background pixels. The parameters of these characteristic distributions were then used to label candidate text regions [22][23]. Cosine similarity and K-means clustering were, respectively, applied to RGB channels to segment text characters in uniform color [24][25]. In these algorithms, color clustering is a single-variable function that maps each pixel to the adjacent color center. From the high-level viewpoint, it maps each text string in uniform color to the most well-suited color layer. However, these algorithms ignore that text string is attached to the neighboring surface in uniform color in most cases. Based on this feature, the algorithm in [26] is designed for boundary clustering (BC) which is another significant property of text stroke width consistency. As the basic element of text, stroke is defined as a connected region in the form of a band of approximately constant width.

Text features, such as edge distribution, gradient variation, closed component boundary, and edge based filter response, are obtained from boundary maps and gradient maps of scene images to detect and verify text regions. They were connected to the geometrical structure of text. In the above algorithms, multiple pixel-based features were employed to distinguish text characters and strings from background outliers. But most of the processes were based on the individual selection of features and the hard assignment of parameters.

Recently the paper [27] suggests scene text detection through two masks filtering method. The paper [28] describes scene text detection through SWT, MSER and candidate classification. The author Wenjun Ding et al. presents a method for overlay text detection using hierarchical localization and textual components [29].

This paper proposed simple but very effective and pioneering method for document text segmentation. The proposed method is robust against noise and has the capacity of less false text area segmentation. The proposed method reduces the noise using MDBUTMF based noise reduction [10]. An energetic edge detection scheme is proceeded to detect adaptive edges which can be most useful to detect text regions. The novel Local Directional Number [LDN] based feature selection is made for feature extraction. Finally the natural image text detection is carried out and the experimental analysis is gone through the segmentation results.

2. PROPOSED METHOD

The proposed scene text detection method finds out the scene text contents from the natural scene images using the advanced edge detection methods, the local directional number feature and linked map generation process. This proposed method takes natural scene image as input and detects the scene text regions. The detected scene text regions are marked for visual identification for human eyes. This new method can be divided by five sub modules and they are MDBUTMF based noise reduction, Energetic edge detection schemes, Local directional number feature image generation, Linked map generation and non scene text rejection. The figure 1 expresses the architectural diagram of the proposed method. The following sections briefly explain the sub modules of this paper.

2.1 MDBUTMF noise reduction:

The 24 bit input scene image is read out into three color channels. These three color channels are used to generate the greyscale image using equation 1. The input image is converted into greyscale image.

GR=. $G = 0.299R + 0.587G + 0.114B$ (1)

Where　GR is the greyscale image, R is the red channel, G is the green channel, and B is the blue channel.

The color scene images are often affected by impulse noises. The gray image which is converted from the color image also reflects that noises. In order to reduce the noise level the Modified Decision Based Unsymmetrical Trimmed Median Filter (MDBUTMF) is applied on the greyscale image. The MDBUTMF is an existing method [10] and it reduces the noisy level of natural scene images, which may contain scene text also. In this method the impulse noises are found based on the values zeroes and 255.If it is a non noise pixel, then that pixel is left without any change otherwise it undergone the noise reduction process. A 3x3 size 2D window is generated from the noisy pixel. If the surrounding pixels are mixture of noisy pixels then the centre noisy pixel is replaced by mean of the surrounded pixels.

2.2 Energetic edge detection schemes

This section deals with two edge descriptors namely, energetic Sobel and Energetic Gaussian.

If the surrounding pixels are mixture of any gray values then the centre noisy pixel is replaced by the median of the surrounded pixels. In this way whole image is made out of noise free and it can be represented by G'.

The proposed energetic Sobel edge detection is based on a new Sobel derived edge mask and the new pattern mask. The 5x5 size new energetic Sobel Edge Mask [ESM] and the new Pattern mask [PM] is described below.

$$ESM = \begin{bmatrix} +1 & +0 & +2 & +0 & +1 \\ +1 & +1 & +0 & +0 & +0 \\ +0 & +0 & +0 & +0 & +0 \\ +0 & +0 & +0 & -1 & -1 \\ -1 & +0 & -2 & +0 & -1 \end{bmatrix}$$

$$PM = \begin{bmatrix} +1 & +0 & +1 & +0 & +1 \\ +0 & +1 & +1 & +1 & +0 \\ +1 & +1 & +1 & +1 & +1 \\ +0 & +1 & +1 & +1 & +0 \\ +1 & +0 & +1 & +0 & +1 \end{bmatrix}$$

These two masks find out the edges in an energetic way. The 5x5 size overlapping window W which is generated from the pixel (i, j) of G', is undergone the convolution process with the ESM. It can be expressed by equation 2.

ESE(i.j)=W Θ ESM (2)

Where, ESE means Energetic Sobel edge image, (i,j) means position to find edge and Θ means convolution operator. The convolution process takes account the pixels locations which

are having the value 1 in the pattern mask PM. Usually all pixels of the 5x5 ESM mask values are used for edge detection .Here the new pattern mask reduces the time consumption and the thick bulk representation of edges. It can be additional seven ESM's can be derived using the rotation of the mask ESM by seven times ($45^0, 90^0, 135^0, 180^0, 225^0, 270^0, 315^0$). These multi-directional masks are derived and using them another seven Sobel edge images are generated. Over all speaking there are eight multidirectional edge images are available and they can be represented as a 3D array such as ESE (i,j,k) where i represents scene image height ,j represents scene image width and k represents the direction indicator(0 to 7).

This paper also proposes another new edge detection mask which is also energetic and efficient. This new mask is derived from the Gaussian edge descriptor. The Pattern Mask PM is also used for the edge detection. The new 5x5 size new Energetic Gaussian Mask [EGM] descriptor is described below.

$$EGM = \begin{bmatrix} +0 & +0 & -1 & +0 & +0 \\ +0 & +0 & -1 & +0 & +0 \\ -1 & -1 & +8 & -1 & -1 \\ +0 & +0 & -1 & +0 & +0 \\ +0 & +0 & -1 & +0 & +0 \end{bmatrix}$$

Energetic Gaussian edge image is generated by the convolution process of EGM and the 5x5 windows of scene text image (G').The EGM is rotated through 45^0 by seven times. The convolution process is using these multidirectional descriptors. The eight multi directional images can be represented as a 3D array such as EGE(i,j,k).

The Sobel fused image is generated to get the integrated version of eight energetic Sobel edge images. The integration process is performed based on equation 3.

$$FESE(i,j) = \int_{k=0}^{8-1} ESE(i,j,k) * CP \qquad (3)$$

Where FESE is the fused Energetic Sobel Edge image, i is the scene image height, j is the scene image width, k-8 multi directions, k $\epsilon[0,7]$, and CP is the contribution part in fusion process. It can be assigned as 0.125.

In this model, the fused energetic Gaussian edge image is also obtained and it is represented by equation 4.

$$FEGE(i,j) = \int_{k=0}^{8-1} EGE(i,j,k) * CP \qquad (4)$$

Where

FEGE - Fused Energetic Gaussian Edge image.

2.3 Local directional number feature image generation

In this section a new and advanced Local Directional Number (LDN) pattern is proposed. The local directional number pattern is explained in the papers [9][10][11][14][15]. But in that references the authors created only one type of local directional number pattern. In the proposed work, two types of Local Directional Number patterns are emerged and they are called as type I LDN and Type II LDN. The LDN descriptors can be found using figure 2 which contains 3x3 window with the position values.

K=7	K=0	K=1
K=6	(i.j)	K=2
K=5	K=4	K=3

Fig: 2. Demonstrating window with the position values.

The ESE 3D array contains 8 edge image and it can be represented by the term K which range from 0 to 7. The k=1 potion is filled by the zero degree edge image. The K=2 potion is filled by the 45 edge image. Like this all eight edge images are filled. These eight values are related with the (i.j) position of the natural scene image (G').These eight elements vector is considered as V. The maximum positive value given degree among these vectors is computed by equation 5 and equation 6.

$$M = MaxPositive(V) \qquad (5)$$

$$MPI_{I,J} = MaxPositiveValue\ givenIndex(V, M) \qquad (6)$$

Where, M is the maximum positive value from vector V,$MPI_{i,j}$ is the maximum positive value given index. Like this, the maximum negative value given index $MNI_{i,j}$ is also computed. The Type I LDN for energetic Sobel is computed by the equation 7.

$$SLDN\_T1(i.j) = 8*MPI_{i,j} + MNI_{i,j} \qquad (7)$$

Where

SLDN_T1 - Sobel LDN Type1 image

The Type II LDN for energetic Sobel is computed by equation 8.

$$SLDN\_T2(i.j) = 8*MNI_{i,j} + MPI_{i,j} \qquad (8)$$

The maximum positive given index $MPI_{I,J}$ for energetic Gaussian edge image is found using equation 6. The maximum negative value given index $MNI_{I,J}$ for energetic Gaussian edge image is also found. The type1 LDN for energetic Gaussian is computed by equation 9.
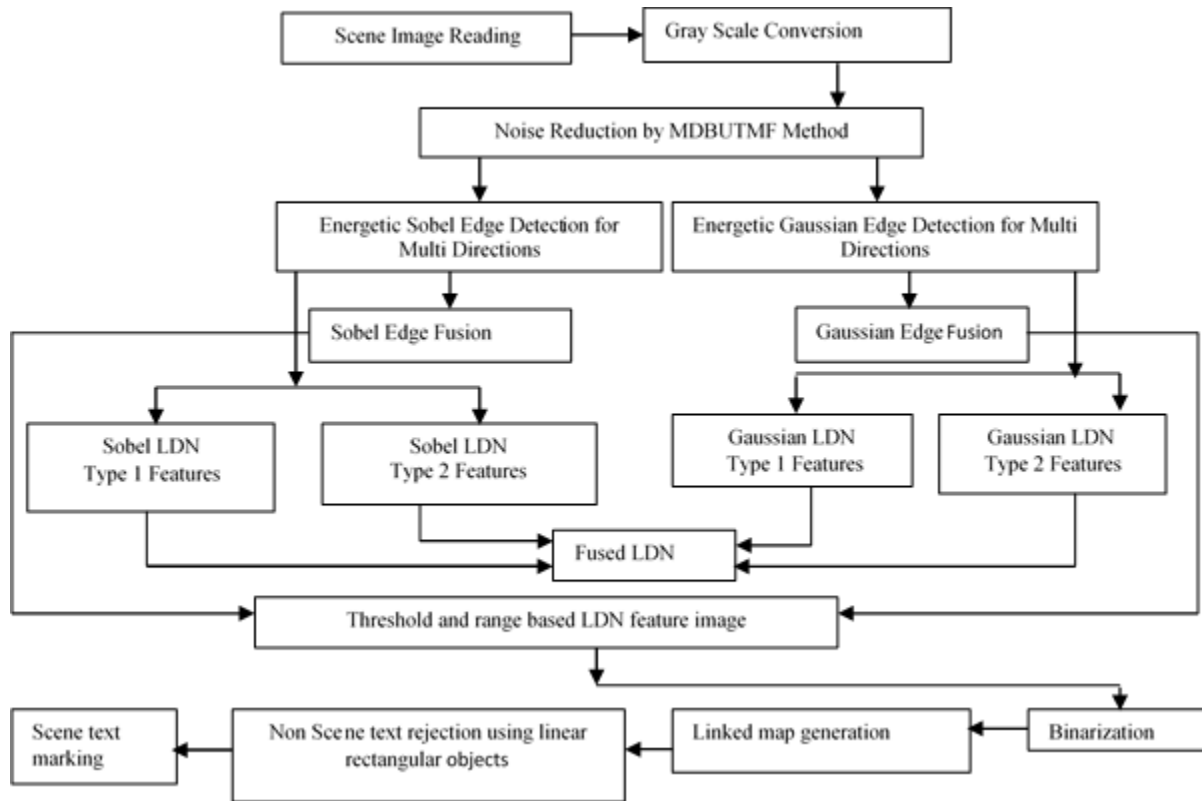
$$GLDN\_T1(i.j) = 8*MPI_{i,j} + MNI_{i,j} \qquad (9)$$

Fig.1. Architectural diagram of the proposed method

Where

GLDN_T1 is the Gaussian LDN Type I image

The Type II LDN for Energetic Gaussian is computed by equation 10.

$$GLDN\_T2(i.j)=8*MNI_{i,j}+MPI_{i,j} \qquad (10)$$

The fused LDN is computed by equation 11.

$$FLDN(i.j)=SLDN\_T1(i.j)+SLDN\_T2(i.j)+$$
$$GLDN\_T1(i.j)+GLDN\_T2(i.j) \qquad (11)$$

The final LDN for scene text supportiveness is achieved by the thresholds. It can be expressed in the equation 12.LDN-

$$
\text{Feature\_img (i.j)} =
\begin{cases}
FLDN(i.j) & \text{if } FESE(i.j)>T1 \\
& \text{and} \\
& FEGE(i.j)>T1 \\
& \text{and} \\
& FLDN(i.j)>=T2 \\
& \text{and} \\
& FLDN(i.j)<=T3 \\
0 & \text{else}
\end{cases}
\qquad (12)
$$

The LDN-Feature_img is the final LDN image and T1, T2 & T3 are thresholds.

2.4 Linked Map Generation

Linked map generation is used to connect or link the two edges whose distances are in between the permitted length in horizontal direction. The entire edges are undergone this linked map generation so that a mapped region by the linking process is obtained [12][13].

The binarization process is the first step to obtain the linked map .The LDN-Feature_img is binarized by a threshold to obtain the binary LDN Feature image. This binary LDN-Feature_img is undergone the edge extension process in horizontal direction, If any pixel is found near to it in horizontal direction within a short range, then a line is drawn from the process originating pixel to the pair end pixel. In this way all nonzero pixels are linked with the neighborhood paired pixel. Finally the individual object lies in the binary image are made connected with its internal pixels.

2.5 Non Scene Text Rejection

In the linked map output each and every individual binary objects which are connected with eight neighbourhoods are extracted. In those binary objects, the objects which satisfy two conditions are extracted. Those two conditions are maintained

to detect the scene text area. Binary shape object can be called as a scene text area if it supports rectangular shape and required height level. The horizontal histogram is computed to detect the rectangular shape. If most of the rows are in approximately same length then it can be considered as the rectangular shape. Binary shape objects which are not coinciding with these conditions are removed and the other objects are extracted for the scene text marking purpose.

The refined binary eight connected objects are undergone the scene text marking process. It is acceptable to guess the scene text regions are generally being in rectangular shapes. So a rectangular bounding box is formulated by connecting the four corner points [11]. These corner points are listed as (start_x,start_y), (end_x,start_y), (start_x,end_y) and (end_x,end_y). These corner points are computed from the linked map. All the scene text objects are marked in this way.

OCR (optical character recognition) is used for the recognition of marked text characters in the scene text. This involves photo scanning of the text character-by-character, analysis of the scanned-in image, and then translation of the character image into character codes, such as ASCII, commonly used in data processing.

3. RESULTS AND DISCUSSIONS

The proposed method is implemented and the analysis is performed using UCID database [18] and KAIST scene text database [19]. For testing purpose hundred UCID images and hundred KAIST images are used. The proposed method performs the scene text detection using MDBUTMF noise reduction, energetic Sobel edge detection and energetic Gaussian edge detection methods. This research work uses LDN Type1 features and LDN Type II features for the detection of scene texts. For analysis purpose, this paper considers the existing scene text detection methodologies by the authors K.C.Kim et al. [7], Xu Chenget et al. [8] and Yi-Fengpan et al. [9].

The experimental scene samples are placed in the figure 3. The figure 3.a indicates the original scene image. In that figure three scene texts are available. They are Eagle, RIVER and BOATING. The figure 3.b shows the results of image conversion from RGB to greyscale .The figure 3.c shows the removal of noises from the greyscale image using MDBUTMF method. The figured indicates the output of edge detection using the new energetic Sobel method. The eight multidirectional energetic Sobel outputs are fused to get the output 3.d. The figure 3.e shows the fused Gaussian edge detection using new energetic Gaussian method. The figure 3.f shows the fused LDN feature image from the Sobel LDN Type I feature, Sobel LDN Type II feature, Gaussian LDN Type I feature and Gaussian LDN Type II feature. The figure 3.g shows the linked map output which creates the links with in the internal fragment of each components .The figure 3.h depicts

the internal screen texts by brighten objects and the background is by darken region. The figure 3.i expresses the marking of the scene texts in the input image. The available three scene texts are surrounded by bounding boxes are indicated by white lines.

The figure 4 depicts New Energetic Sobel multi directional edge detection results

The Table 1 and Fig.5 show the accuracy or the effective true identification of the text content area extraction marking. For UCID database and KAIST the accuracy is analysed in terms of percentage. The proposed and the method identified by us shows greater response over the existing methods for the database images. The proposed method generates an average accuracy of 92.14% for the UCID while the existing K.C.Kim, Xu Chenget and Yi-Fengpan methods generate 84.28, 86.47 and 88.25 respectively and on the other hand the KAIST produced an average accuracy of 91.58% for the proposed method and the existing methods such as K.C.Kim, Xu Chenget and Yi-Fengpan methods has the value of 83.53, 86.12 and 87.94% respectively. The accuracy is calculated by the true positive, false positive, true negative and false negative measures which are computed based on equations 13 to 17.

$$AC= \frac{TP+TN}{TP+TN+FP+FN} * 100 \qquad (13)$$

Where AC is the accuracy and TP is the True Positive.

$$TP = \frac{CTP}{TB} \qquad (14)$$

Where CTP is the number of correctly identified text pixels and TB is the total number of background pixels.

$$FP = \frac{CFP}{TB} \qquad (15)$$

Where CFP is number of incorrectly identified text pixels.

$$TN = \frac{CTN}{TB} \qquad (16)$$

$$FN = \frac{CFN}{TB} \qquad (17)$$

Where CFN is the count of incorrectly identified.

Table 1: Average accuracy analysis

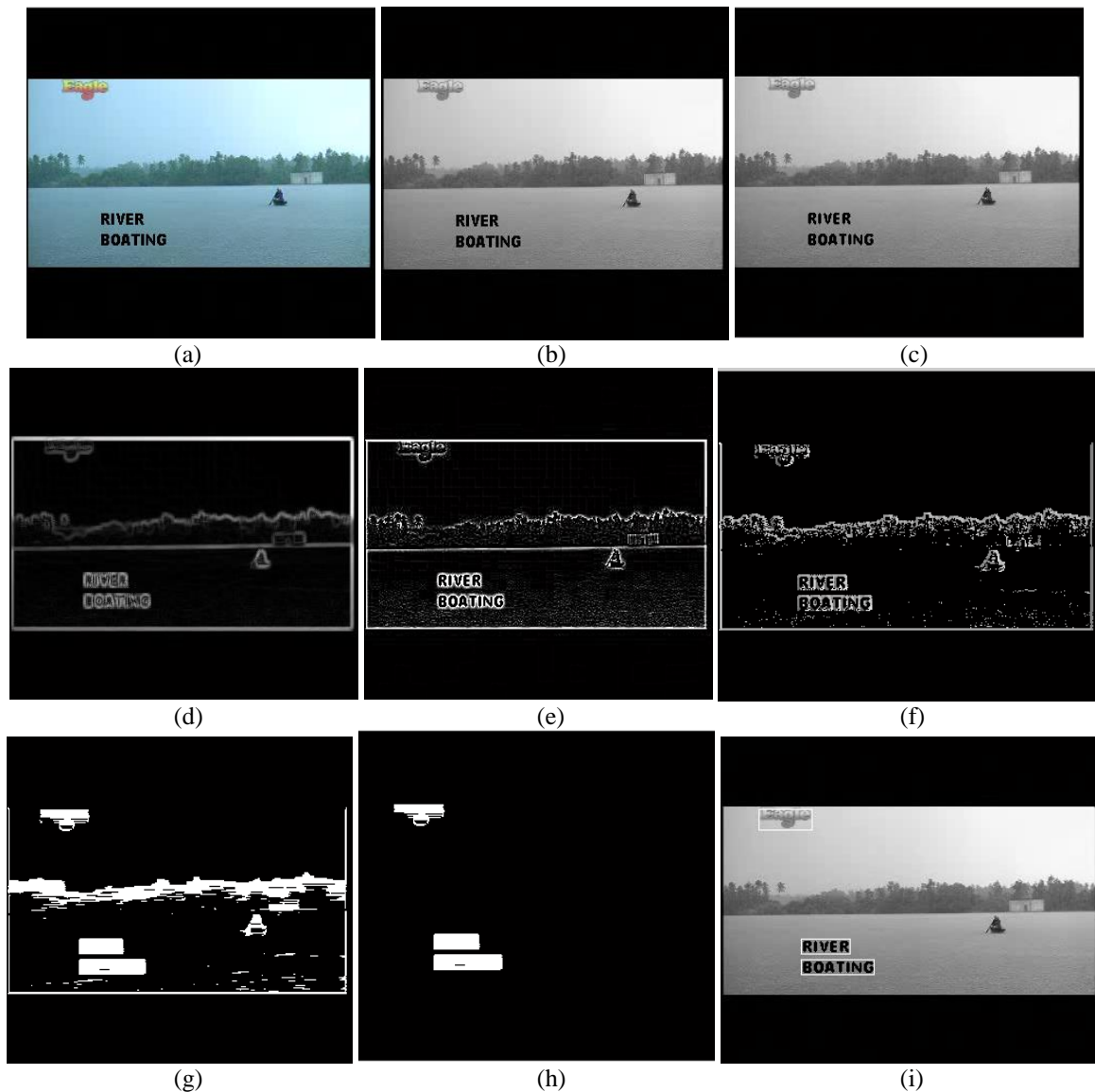| Database Name	Average Accuracy (in percentage)			
	K.C.Kim method	Xu Chenget Method	Yi-Fengpan Method	Proposed Method
UCID	84.28	86.47	88.25	92.14
KAIS	83.53	86.12	87.94	91.58

Fig.3. Proposed Scene detection results: a) Original input scene image b) Gray scale image c) Noise reduction image d) Energetic Sobel fused image e) Energetic Gaussian fused image f) LDN output g) Linked map result h) Scene text detection i)Scene text marking.
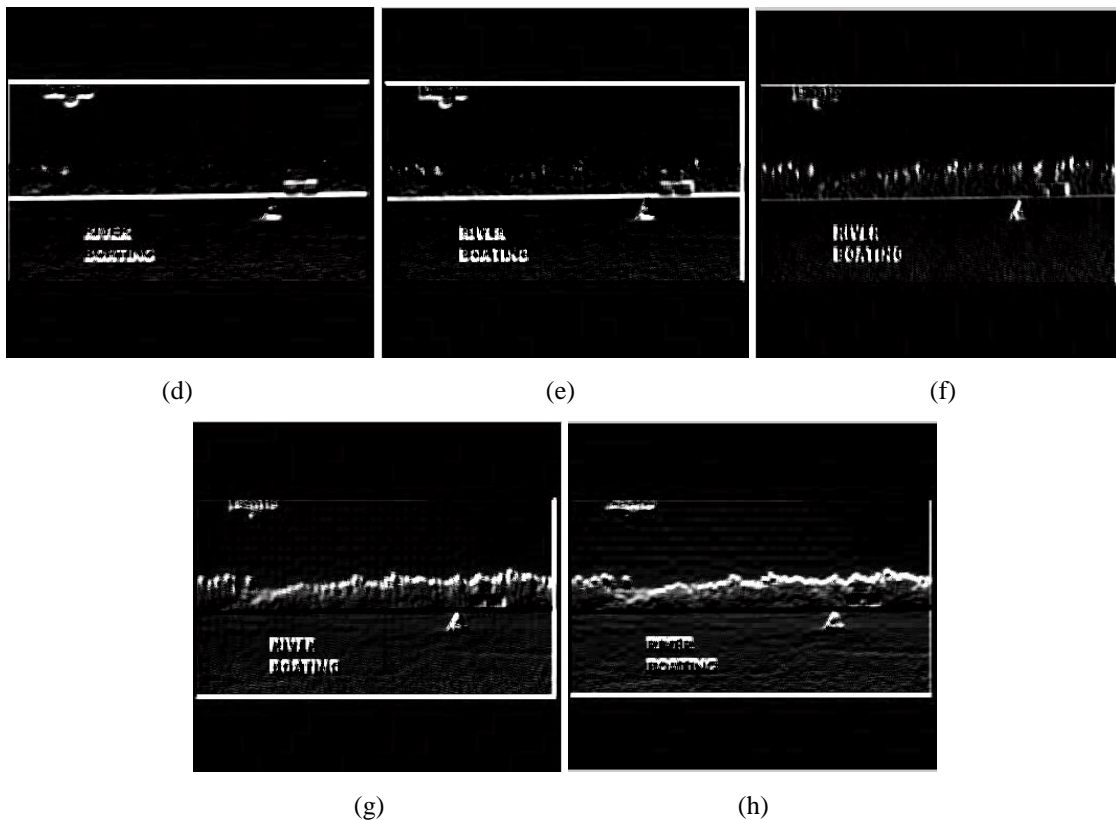

(a)   (b)   (c)

(d)   (e)   (f)

(g)   (h)
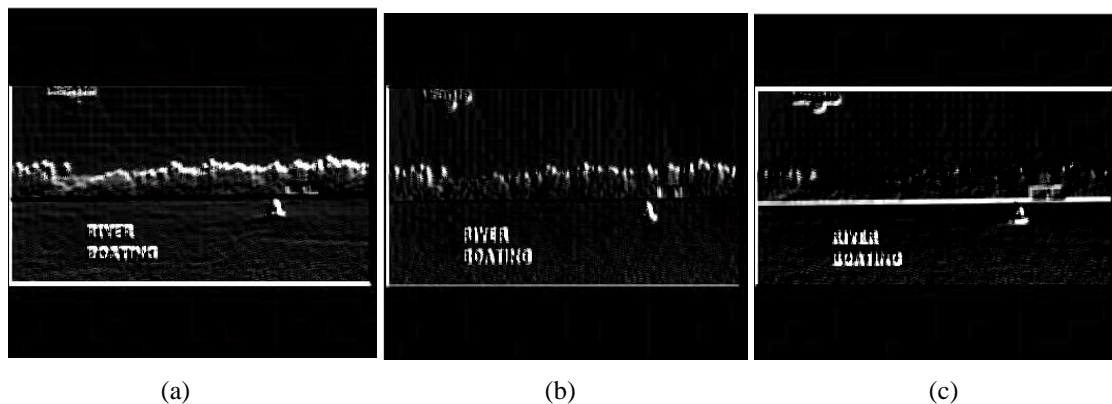
Fig. 4. New Energetic Sobel multi directional edge detection results: a) 0^0 edge detection b) 45^0 edge detection c) 90^0 edge detection d) 135^0 edge detection e)180^0 edge detection f)225^0 edge detection g) 270^0 edge detection h)315^0 edge detection.
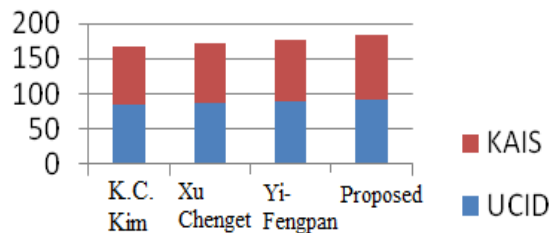

Fig. 5. Average accuracy analysis chart.

Table 2 and Fig.6 give the eye perception based performance evaluation. Eye perception quality index is found from human eye perception based quality measurement.

Table 2: Eye Perception based performance Index analysis

Data base Name	Image Name	Eye Perception based performance Index			
		K.C. Kim method	Xu Chenget Method	Yi-Fengpan Method	Proposed

UCID	UCID Img 1	1	2	3	4
	UCID Img 2	1	2	3	4
	UCID Img 3	1	2	3	4
	UCID Img 4	2	1	3	4
KAIST	KAIST Img1	1	2	3	4
	KAIST Img2	2	1	3	4
	KAIST Img3	2	1	3	4
	KAIST Img4	1	2	3	4

High eye perception performance index means higher segmentation quality. In other words the better segmentation quality providing method is marked by higher index numbers. In the proposed method a high level of index is achieved while compared with the methods.
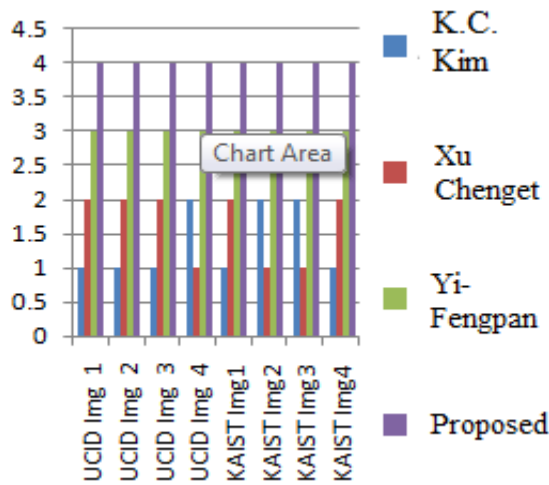
UCID Img 3	0.153	0.119	0.073	0.021
UCID Img 4	0.075	0.110	0.034	0.015

KAIST	KAIST Img1	0.184	0.120	0.066	0.026
	KAIST Img2	0.153	0.27	0.061	0.020
	KAIST Img3	0.295	0.194	0.073	0.024
	KAIST Img4	0.275	0.197	0.124	0.037

Fig.6. Eye Perception based performance Index analysis chart.

In Table 3 and Fig.7, mean square error can be calculated by,

$$\text{MSE} = \frac{1}{H,W} \sum_{i=o}^{H-1} \sum_{j=0}^{w-1} \ (S\_img(i,j) - G\_img\ (i,j))^2$$

$$(18)$$

Where

        H - Height of the image

        W - Width of the image

        S - Segmented image

        G - Ground Truth image

The mean square error (MSE) analysis also has a greater response in the proposed method and it is shown in table 3.

Table 3: MSE analysis

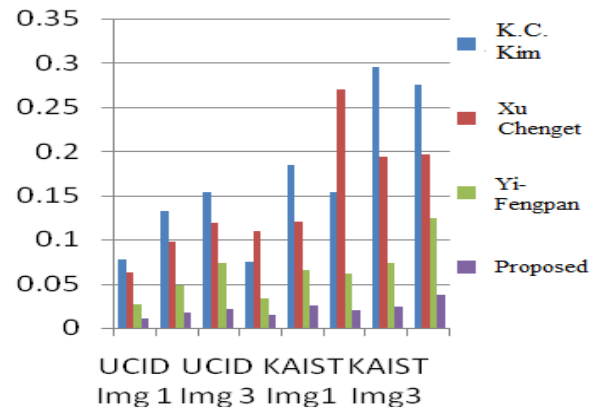| Database Name	Image Name	MSE For Test Data			
		K.C. Kim method	Xu Chenget Method	Yi-Fengpan Method	Proposed
UCID	UCID Img 1	0.077	0.063	0.027	0.011
	UCID Img 2	0.132	0.097	0.048	0.017

Fig.7. MSE analysis chart

In Table 4 and Fig.8, Peak signal to noise ratio (PSNR) is analysed with the set of images using equation 19.

$$\text{PSNR} = 10 \log_{10}\left(\frac{255^2}{MSE}\right) \qquad (19)$$

Table 4: PSNR analysis

Database Name	Image Name	PSNR (in db)			
		K.C. Kim method	Xu Chenget Method	Yi-Fengpan Method	Proposed
UCID	UCID	59.26	60.13	63.72	67.43

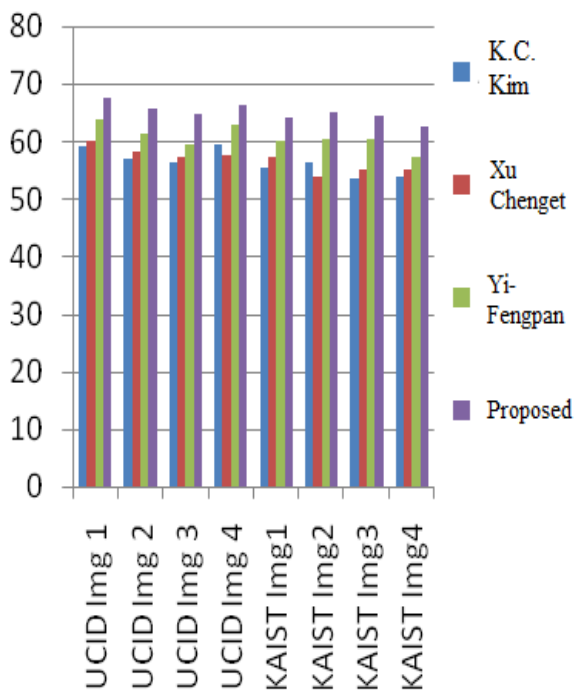|					
	Img 1				
	UCID Img 2	56.92	58.26	61.38	65.77
	UCID Img 3	56.28	57.37	59.49	64.90
	UCID Img 4	59.38	57.71	62.81	66.36
KAIST	KAIST Img1	55.48	57.33	59.93	63.98
	KAIST Img2	56.28	53.81	60.27	65.12
	KAIST Img3	53.43	55.25	60.49	64.31
	KAIST Img4	53.73	55.18	57.19	62.44

Fig.8. PSNR analysis chart

Peak signal to noise ratio for the existing methods such as K.C.Kim, Xu Chenget and Yi-Fengpan with respect to the UCID and KAIST data bases are low while compared with the proposed method. The proposed method produces an output above 60db in all images taken to test purpose.

Speed is a measure of average time taken to detect text in seconds. The speed in the proposed method is increased little bit while compared with the existing methods. The Table 5 and Fig.9 express the Execution speed analysis.

Table 5: Average accuracy

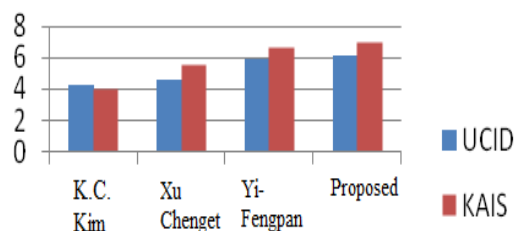| Database Name	Average Accuracy (in seconds)			
	K.C. Kim method	Xu Chenget Method	Yi-Fengpan Method	Proposed Method
UCID	4.24	4.60	5.89	6.08
KAIST	3.98	5.52	6.71	6.97

Fig.9. Average accuracy chart

4. CONCLUSION

An efficient text extraction from complex image background is achieved in this paper. The efficient scene text detection is achieved based on advanced edge detection methods, Local Directional Number feature and linked map generation process. The resultant scene text regions are marked for visual identification of human eyes on natural scene images. The proposed method involves MDBUTMF filter for noise reduction to get robust result. The edge detection process is influenced by novel Energetic edge detection scheme. The proposed method outperforms the existing methods in cases of MSE, PSNR and Accuracy. This work can be extended for deep study about clustering for text segmentation and to identify the characters.

REFERENCES

[1]   K,Jain and B.Yu, "Autamatic text location in   images and video frames", Pattern Recognition, vol. 31, no.12, pp. 2055-2076, 1998.
[2]   Trier and A.K.Jaind, "Goal directed evaluation of Binarization methods", IEEE Trans. on PAMI, vol. 17, no. 12, pp. 1191-1202, Dec 1995.
[3]   V.Wu.R.Manmatha and E.M.Riseman, "Text finder an Automatic system to Detect and Recognize Text in Images", IEEE Trans. on PAMI, vol. 20, pp. 1224-1229, 1999.
[4]   Chucai YI and Yingli Tian, "Localizing Text in scene images by boundary clustering stroke segmentation and string fragment classification", IEEE Transaction on image processing, vol. 21, sep 2012.

[5]   Sunil Kumar,"Text Extraction and Document Image Segmentation Using Matched Wavelets and MRF Model", IEEE Transactions on Image Processing, vol. 16, no. 8, Aug-2007.

[6]   Ranjit Goshal, Anandarup Roy,Swapan and K Parui, "Text Extraction from Scene Images Using Statistical Distributions", IEEE Third International Conference on Emerging Applications Of Information Technology, 2012.

[7]   K.   C.   Kim,   H.   R.   Byun   and   Y.   J.   Song, "Scene Text Extraction in Natural Scene mages using Hierarchical Feature Combining and Verification," IEEE Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04), 2004.

[8]   Xu-cheng Yin, Xuwang Yin, Kaizhu Huang and Hong-wei, "Robust Text Detection in Natural Scene Images", vol. 36, issue 5, pp. 970-983, 2014.

[9]   Yi-Feng Pan, Xinwen Hou and Cheng Lin-Liu, "A Hybrid Approach to Detect and Localize Texts in Natural Scene Images", IEEE Transactions on image processing, vol. 20, issue 3, 2010.

[10]   S. Esakkirajan, T. Veerakumar, Adabala N. Subramanyam, and C. H. PremChand, " Removal of High Density Salt and Pepper Noise Through Modified Decision Based Unsymmetric", IEEE Signal processing letters, vol. 18, no. 5, May 2011.

[11]   O. R. Vincent and O. Folorunso, "A Descriptive Algorithm for Sobel Image Edge Detection", Proceedings of Informing Science & IT Education Conference (InSITE), 2009.

[12]   Manoj K.Vairalkar and S.U.Nimbhorkar, "Edge Detection o f Images Using Sobel Operator", International Journal of Emerging Technology and Advanced Engineering, vol. 2, issue 1, Jan. 2012.

[13]   Ramirez Rivera A., Rojas Castillo and Oksam Chae, "Local Directional Number Pattern for Face Analysis: Face and Expression Recognition", IEEE Transactions on Image Processing, vol. 22 , issue 5, 2013.

[14]   Dong-Ju Kim, Sang-Heon Lee and Myoung-Kyu Sohn, "Face Recognition via Local Directional Pattern", International Journal of Security and Its Applications, vol.7, no. 2, March 2013.

[15]   Manjunatha S.B, Guruprasad A.M and Vineesh P, "Face Analysis By Local   Directional   Number   Pattern",   International Journal   of   Engineering   Research   and General Science, vol. 3, issue 1, pp. 2091-2730, 2015.

[16]   Wonjun Kim and Changick Kim, "A New Approach for Overlay Text Detection and Extraction From Complex Video Scene" IEEE Transactions On Image Processing, vol. 18, no. 2, Feb. 2009.

[17]   Mohamed Shajahan H and Munir M, "An Image based method for Rendering Overlay Text Detection and Extraction Using Transition Map and Inpaint" International Journal of Scientific Research and Innovative Technology, vol. 2, no. 4, April 2015.

[18]   UCID - An Uncompressed Colour Image Database, http://vision.doc.ntu.ac.uk/.

[19]   KAIST Scene Text Database, Prof. Jin Hyung Kim, Email: Jkim @ kaist.ac.

[20]   Nikolaou, N and N. Papamarkos, "Color reduction for complex document images", Int. Journal of Imaging Systems Technology, vol.19, issue 1, pp. 14–26, 2009.

[21]   Manza R.R, B.P. Gaikwad and G.R. Manza,"Use Of Edge Detection Operators For Agriculture Video Scene Feature Ex-Traction From Mango Fruits," Advances in Computational Research, vol. 4, issue 1, pp. 50-53, 2012.

[22]   Chen X, J. Yang, J. Zhang and A. Waibel, "Automatic detection and recognition of signs from natural scenes", IEEE Trans. Image Processing, vol. 13, issue 1, pp. 87–99, 2004.

[23]   Chen, D.T, J.M. Odobez and H. Bourlard, "Text detection and recognition in images and videos frames", Pattern Recognition, vol. 37, pp. 595-608, 2004.

[24]   Mancas-Thillou C. and B. Gosselin, "Spatial and color spaces combination for natural scene text extraction", Proc. IEEE Conf. Image Process., pp. 985–988, 2006.

[25]   Song Y, A. Liu, L. Pang, S. Lin, Y. Zhang, and S. Tang, "A novel image text extraction method based on K-means clustering", Proc. 7th Int. Conf. Comput. Inform. Sci., pp. 185–190, 2008.

[26]   Epshtein B., E. Ofek and Y. Wexler, "Detecting text in nature scenes with Stroke width transform", Proc. IEEE Conf. Comput. Vision Pattern Recogn., pp. 2963–2970, 2010.

[27]   Houssem Turki, Mohamed Ben Halima and Adel M. Alimi, "Text detection in natural scene images using two masks filtering", IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA), 2016.

[28]   Leibin Guan and Jizheng Chu,"Natural scene text detection based on SWT, MSER and candidate classification", 2nd International Conference on Image, Vision and Computing (ICIVC), pp. 26 - 30, 2017.

[29]   Wenjun Ding, Susu Shan and Feng Su,"Text detection in natural scene images by hierarchical localization and growing of textual components", IEEE International Conference on Multimedia and Expo (ICME),

Authors

T.Beula Bell is an Assistant Professor, Department of Computer Applications in Nesamony Memorial Christian College, India. She is doing research under the guidance of Dr. M. K. Jeya Kumar he is working as a Professor in the Department of Computer Applications, Noorul Islam University, Kumaracoil, Tamilnadu, Ind*ia* . She received B.Sc, M.C.A., M.Phil degree from M. S. University, Thriunelveli, India. Her area of interest is image processing.

M. K. Jeya Kumar was born in Nagercoil, Tamilnadu, India on 18th September 1968. He received his Master of Computer Applications degree   from   Bharathidasan   University, Trichirappalli, Tamilnadu, India in 1993. He fetched his M.Tech degree in Computer Science and Engineering   from   Manonmaniam SundarnarUniversity, Tirunelveli, Tamilnadu, India in 2005. He completed his Ph.D degree in Computer Science and Engineering from Dr.M.G.R University, Chennai, Tamilnadu, India in 2010. He is working as a Professor in the Department of Computer Applications, Noorul Islam University, Kumaracoil, Tamilnadu, India since 1994. He has more than twenty three years of teaching experience in reputed Engineering colleges in India in the field of Computer Science and Applications. He has presented and published a number of papers in various national and international journals. His research interests include Mobile Ad Hoc Networks and network security, image processing and soft computing techniques.