International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 9, pp: 372-380 www.ijaem.net ISSN: 2395-5252

Scene-Text Grit and Localization by LDN Descriptor and Linked Map

Dr.Beula Bell. T [1] and Dr. Jeya Kumar M. K [2]
[1] *Associate professor,N.M.C.C,Marthandam, 627165, Tamil Nadu, India.*
[2] *Professor, Department of Computer Applications, Noorul Islam University, Kumaracoil 629180, Tamil Nadu, India.*

Date of Submission: 10-11-2020 Date of Acceptance: 25-11-2020

ABSTRACT: The scene-text from the onscene images can be detected with the aid of energetic edge detector by using the Sobel and Gaussian novel pattern masks. The LDN descriptors of Sobel and Gaussian helps to obtain the edge information. By using the edge information a linked map is generated to link the textual information contains in the onscene. Finally the scene text area is marked by rectangle shape markings as output.
Keywords: Scene_text,Sobel,Gaussian,LDN,Grit

I. INTRODUCTION:

The scene-text detection is the process of segmenting text from the scene-images / onscene-images. The scene-text detection process is called by another name as scene-text localization which marks the scene-text area of the input scene image. The Text-segmentation process of document-image is not the same one with scene-text segmentation, because of multicolour texts area, different size text area, and complex background information by images with texts, illusions and inadequate contrast [1][2]. This work proposes a novel method for scene-text detection namely 'Scene text detection using Energetic edge detectors and LDN descriptor' to accomplish the task with higher accuracy. The main contribution in this work is the Energetic Edge Detectors and the modified LDN. The architecture diagram of this method is depicted in Fig.1.1.

II. ENERGETIC SOBEL EDGE DETECTION

The gra**y**scale noise-free image is given as input to this module.

A new Sobel based edge detector is designed to detectthe edges in a potential way. Edge detection is performed in the surrounded eight directions to receive energetic edges.The traditional Sobel detection is working based on a 3x3 dimension overlapped window which produces less-power edges.

To increase the power of Sobel edge detector (with energetic edges), a new 5x5 size mask is designed with potential values which pattern is illustrated in Fig.1.2.

In the traditional Sobel method the entire elements of the window is used, and that system increases the complexity of the edge detection. To reduce the complexity, the proposed scene-text detection method STD-EED-LDN uses only 17 elements among the 25 elements of the 5x5 size window. A pattern which indicates the specific elements of 5x5 size window is illustrated in Fig.1.3, whereas the mnemonic 1 means needed location and 0 means discarded location. This pattern is used to the convolution process which is the basis of edge detection. The dark gray locations in this pattern are the concerned location for convolution and the light gray locations indicate the discarded locations in Fig.1.3.

For LDN computation, there is a need of eight values, so that this work is progressed by multi-direction oriented Sobel edge detection. Herein, eight directions are considered to construct the eight-outputs of edge detection[3][4][5]. The eight direction Energetic Sobel masks can be constructed via the rotation process, and this phenomenon is presented by equation (1.1).are alike with Sobel mask values. The newer edition of Energetic Sobel

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 9, pp: 372-380 www.ijaem.net ISSN: 2395-5252

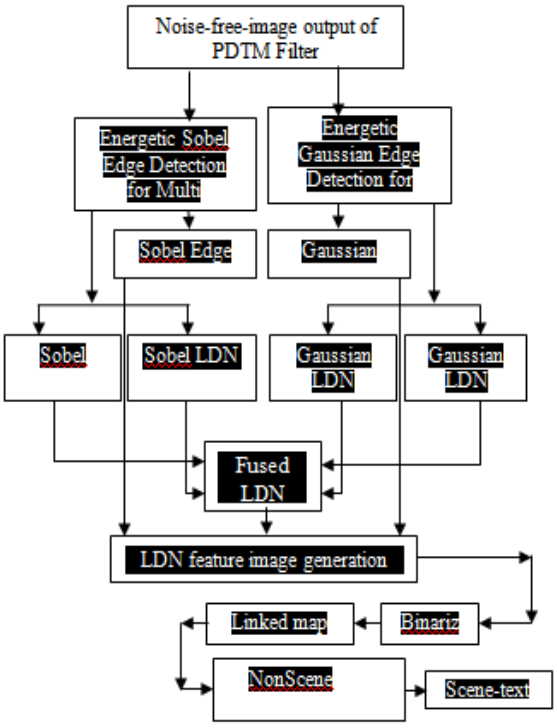


Fig.1.1. Architecture diagram of the proposed method.

$$M_{SOBEL} = \begin{Bmatrix} \{1,0,2,0,1\},\{1,1,0,0,0\},\{0,0,0,0,0\}, \\ \{0,0,0,-1,-1\}\{-1,0,-2,0,-1 \end{Bmatrix} \quad (1.1)$$

+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

Fig.**1.2.** Illustration of energetic Sobel 5x5 size mask.

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

Fig.1.3. Pattern window used for convolution.

The 0^0degree rotation mask $M_{ROT}^{K=0}$ is obtained using equation (1.2).

$M_{ROT}^{K=0} = M_{SOBEL}$(1.2)

$X_{ORIG} = \{0,0,0,2,4,4,4,2,1,1,1,2,3,3,3,2,2\}$(1.3)

$Y_{ORIG} = \{0,2,4,4,4,2,0,0,1,2,3,3,3,2,1,1,2\}$(1.4)

$X_{REPLACE} = \{2,0,0,0,2,4,4,4,2,1,1,1,2,3,3,3,2\}$(1.5)

$Y_{REPLACE} = \{0,0,2,4,4,4,2,0,1,1,2,3,3,3,2,1,2\}$(1.6)

The Energetic masks for the degrees such as 45°, 90°, 135°, 180°, 225°, 270°, and 315° are constructed based on equation (1.7), corresponding with $K \in [1, R-1]$ where R is the total rotation concerned.

$M_{ROT}^{K}(Y_{ORIG}^{p}, X_{ORIG}^{p}) = M_{SOBEL}^{K-1}(Y_{REPLACE}^{p}, X_{REPLACE}^{p})$(1.7)

$p \in [0, q-1]$

Where,

$M_{ROT}^{K} - Rotated\ mask\ for\ k^{th}\ direction$

$q - Total\ position\ of\ considered - locations(herein, it\ is\ 17)$

The convolution process is happened by making the mean process on sum of product of the corresponding elements of the Noise-free-image and the Rotated-Sobel- mask, within the valid locations of pattern mask. This phenomenon is described in equation (1.8).

$$I_S^{k,i,j} = \frac{\sum_{m=-2}^{+2}\sum_{n=-2}^{2}(I_{NF}^{i+m,j+n} * M_{ROT}^{K}) * M_p}{c} \quad (1.8)$$

i ϵ [2, IH-2]

j ϵ [2, IW-2]

Where

$I_S^{K,I,J}$ - Energetic Sobel edge image in the k^{th} direction

M_p - Pattern mask

c– Number of elements which are settled with 1_S, in pattern mask

In this way, the entire 8 directional Sobel Eenergetic edge images are generated[6]. The Energetic Sobel fussed image is computed by composing the eight edge images into a single edge image based on equation (1.9).

$I_{SF}^{i,j} = \sum_{k=0}^{8-1} I_S^{k,i,j} * \frac{1}{R}$(1.9)

k ϵ [0, 8-1]

i ϵ [0, IH-1]

j ϵ [0, IW-1]

Where

I_{SF} – Energetic Sobel fused edge image

R - Number of rotations

III. ENERGETIC GAUSSIAN EDGE DETECTION

The grayscale noise-free image is the input to detect the edges by Gaussian masks. Gaussian edge detector is one of the powerful edge detection method in image processing tasks. In this work, the traditional Gaussian mask is modified to an Energetic version which is named as Energetic Gaussian mask [7][8][10]. This new energetic mask formulates a better edge image than the traditional Gaussian edges. The same pattern mask that is described in Fig.1.3 is used here to form the edge

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 9, pp: 372-380 www.ijaem.net ISSN: 2395-5252

detection output. Herein the new Energetic Gaussian edge mask is defined based on Fig.1.1.

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

Fig.1.1. Energetic Gaussian Edge Mask.

In this work, the eight directional edge detector is performed by the eight masks which are derived by the rotation process. The Energetic Gaussian mask is presented by equation (1.10).

$$M_{Gauss} = \begin{Bmatrix} \{0,0,-1,0,0\},\{0,0,-1,0,0\},\{-1,-1,8,-1,-1\}, \\ \{0,0,-1,0,0\}\{0,0,-1,0,0\} \end{Bmatrix} (1.10)$$

The rotation supportive vectors are designed based on equations from Equation (1.3) to Equation (1.6). The rotation masks for other degrees such as 45^0, 90^0, 135^0, 1800, 225^0, 270^0, and 315^0 are constructed based on equation (1.7) corresponding with k ϵ [1, R-1], where R is the total rotation concerned.

$$M_{ROT}^{K}\left(Y_{ORIG}^{p}, X_{ORIG}^{p}\right) = M_{Gauss}^{K-1}\left(Y_{REPLACE}^{p}, X_{REPLACE}^{p}\right)(1.11)$$

p ϵ [0, q-1]

The convolution process is proceeded by equation (1.8). The resultant Gaussian edge images are stored as $I_G^{k,i,j}$, where k ϵ [0, 8-1], i ϵ [0, IH-1] and j ϵ [0, IW-1].

The Energetic Gaussian fused image is computed by composing the eight- edge-images into a single-edge-image based on equation (1.12).

$$I_{GF}^{i,j} = \sum_{k=0}^{8-1} I_G^{k,i,j} * (\frac{1}{R})(1.12)$$

Where,

I_{GF} – Energetic Gaussian Fused Edge Image

1V. LDN DESCRIPTOR FOR ENERGETIC SOBEL EDGES

LDN descriptor is expanded as Local Directional Number pattern. Generally, LDN descriptor generates an image representation from the edge representation of an input image. This is a powerful image descriptor. But this LDN descriptor is modified to produce two image descriptors from the set of eight directional edge images. The input edge images are taken from the Energetic Sobel edge images. These two image descriptors are formed based on maximum-value and minimum-value given indices through the newly designed two formulae [11][12]. A 3x3 size block is virtually constructed based on Fig.1.5.In Fig.1.5, the eight directional edge information are positioned in the surrounded blocks of the centre pixel $I_{NF}^{i,j}$. Shortly speaking, the [i,j]th pixel of Noise-free-image and the corresponding eight edges data are projected in Fig.1.5. The maximum positive value provider's index is found based on equations (1.13) to (1.15).

$$Z_N = \{I_S^{0,i,j}, I_S^{1,i,j}, I_S^{2,i,j}, I_S^{3,i,j}, I_S^{4,i,j}, I_S^{5,i,j}, I_S^{6,i,j}, I_S^{7,i,j}\}(1.13)$$

$$V_{PMAX} = FuncPositiveMax(Z_N)(1.14)$$

index=7 $I_S^{7,i,j}$ 315^0 rotation	index=0 $I_S^{0,i,j}$ 0^0 rotation	index=1 $I_S^{1,i,j}$ 45^0 rotation
index=6 $I_S^{6,i,j}$ 270^0 rotation	$I_S^{i,j}$	index=2 $I_S^{2,i,j}$ 90^0 rotation
index=5 $I_S^{5,i,j}$ 225^0 rotation	index=4 $I_S^{4,i,j}$ 180^0 rotation	index=3 $I_S^{3,i,j}$ 135^0 rotation

Fig.1.5. Illustration of eight edges of Sobel for LDN computation

$$\alpha = \begin{cases} 0, & if\ Z_N^0 = V_{PMAX} \\ 1, & else\ if\ Z_N^1 = V_{PMAX} \\ 2, & else\ if\ Z_N^2 = V_{PMAX} \\ 3, & else\ if\ Z_N^3 = V_{PMAX} \\ 4, & else\ if\ Z_N^4 = V_{PMAX} \\ 5, & else\ if\ Z_N^5 = V_{PMAX} \\ 6, & else\ if\ Z_N^6 = V_{PMAX} \\ 7, & otherwise \end{cases} (1.15)$$

Where

$Z_N - A\ set\ containing\ eight\ Sobel\ edges\ from\ Fig.1.5.$

$FuncPositiveMax - Function\ to\ compute\ Max + ive\ value$

$V_{PMAX} - Maximum\ Positive\ value$

$\alpha - Maximum\ Positive\ value\ provider's\ index$

The minimum negative value provider's index is formed based on equations from equation (1.16) to equation (1.17)

$$V_{NMIN} = FuncNegativeMax(Z_N)(1.16)$$

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 9, pp: 372-380 www.ijaem.net ISSN: 2395-5252

$$\beta = \begin{cases} 0, & else\ if \quad Z_N^0 = V_{NMIN} \\ 1, & else\ if \quad Z_N^1 = V_{NMIN} \\ 2, & else\ if \quad Z_N^2 = V_{NMIN} \\ 3, & else\ if \quad Z_N^3 = V_{NMIN} \\ 4, & else\ if \quad Z_N^4 = V_{NMIN} \\ 5, & else\ if \quad Z_N^5 = V_{NMIN} \\ 6, & else\ if \quad Z_N^6 = V_{NMIN} \\ 7, & otherwise \end{cases} \quad (1.17)$$

Where

$FuncNegativeMin$

$-\ Function\ to\ compute\ Minimum - ive$

$V_{NMIN} - Minimum\ Negative\ value$

β

$-\ Minimum\ Negative\ value\ provider's\ index$

The First Energetic Sobel edge based LDN is derived by a new approach through equation (1.18).

$$I_{LDN\_S\_1}^{i,j} = 8 * \alpha + \beta \quad (1.18)$$

Where

$I_{LDN\_S\_1} - LDN\ descriptor_1\ by\ Sobel\ edges$

The Second Energetic Sobel edge based LDN image is obtained by a new approach through equation (1.19).

$$I_{LDN\_S\_2}^{i,j} = 8 * \beta + \alpha \quad (1.19)$$

Where

$I_{LDN\_S\_2}$

$-\ LDN\ descriptor_2 by\ Sobel\ edges$

The data ranges of these two descriptors are 0 to 63. In this way the entire pixels of the edge images are processed and these two descriptors are formed.

V. LDN DESCRIPTOR FOR ENERGETIC GAUSSIAN EDGES

LDN descriptor can also be derived from the Gaussian edge information. This section generates two image descriptor related to Gaussian edges through a novel way. These two derivations are generated via the maximum-value and minimum-value given indices of the eight directional edges. Herein, a 3x3 size block is virtually constructed similar with Fig.1.6.

The [i, j]th pixel of noise-free image and the corresponding eight edge data are projected in Fig.1.6. The maximum-Positive-Value provider's index is found based on equation (1.20).

$$Z_N = \{I_G^{0,i,j}, I_G^{1,i,j}, I_G^{2,i,j}, I_G^{3,i,j}, I_G^{4,i,j}, I_G^{5,i,j}, I_G^{6,i,j}, I_G^{7,i,j}\} \quad (1.20)$$

Where

$Z_N - A\ set\ containing\ eight\ Gaussian\ edges\ from\ Fig.1.6$

index=7 $I_G^{7,i,j}$ 315^0 rotation	index=0 $I_G^{0,i,j}$ 0^0 rotation	index=1 $I_G^{1,i,j}$ 45^0 rotation
index=6 $I_G^{6,i,j}$ 270^0 rotation	$I_{NF}^{i,j}$	index=2 $I_G^{2,i,j}$ 90^0 rotation
index=5 $I_G^{5,i,j}$ 225^0 rotation	index=4 $I_G^{4,i,j}$ 180^0 rotation	index=3 $I_G^{3,i,j}$ 135^0 rotation

Fig.1.6. Illustration of eight edges of Gaussian for LDN computation

ThePositive-Maximum-Valuecomputation is performed based on equation (1.14). The Maximum-Positive-Value-Provider's index is computed by using equation(1.15), which is notified by the term α.

The Minimum-Negative-Value-Provider's index β is found based on equation (1.16) and equation (1.17). The First LDN descriptor based on Gaussian edges is a derived by a new approach through equation (1.21).

$$I_{LDN\_G\_1}^{i,j} = 8 * \alpha + \beta \quad (1.21)$$

Where

$I_{LDN\_G\_1}^{i,j} - LDN\ descriptor - 2\ by\ Gaussian\ edges$

The Second LDN descriptor for energetic Gaussian edge is accomplished through equation (1.22).

$$I_{LDN\_G\_2}^{i,j} = 8*\beta+\alpha \quad (1.22)$$

$I_{LDN\_G\_2}^{i,j}$

$-\ LDN\ descriptor\_2\ by\ Gaussian\ edges$

The data ranges of these two descriptors are also 0 to 63, similar with Sobel based LDN descriptors[15][16]. In this way, the entire pixels of the Gaussian edge images are processed and these two descriptors are formed.

Finally, the Fused-LDN descriptor is computed, by integrating the four LDN descriptors. The added-information of the $I_{LDN\_S\_1}, I_{LDN\_S\_2}, I_{LDN\_G\_1}\ and\ I_{LDN\_G\_2}$ descriptors is the resultant LDN descriptor, and it is described in equation (1.23).

$$I_{LDN}^{i,j} = I_{LDN\_S\_1}^{i,j} + I_{LDN\_S\_2}^{i,j} + I_{LDN\_G\_1}^{i,j} + I_{LDN\_G\_2}^{i,j} \quad (1.23)$$

Where

$I_{LDN} - Fused\ LDN\ image\ descriptor$

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 9, pp: 372-380 www.ijaem.net ISSN: 2395-5252

The data range of I_{LDN} descriptor is 0 to 252 because, the maximum range of individual descriptor is 63.

The LDN based feature is computed to represent the text information clearly, which may help to the upcoming-process like linked-map generation. The LDN feature image is formulated by using three parameters and they are: Sobel Fused image I_{SF}, Gaussian Fused image I_{GF} and I_{LDN}. This feature image is computed using equation (1.24).

$$I_{LDN\_F}^{i,j} = \begin{cases} I_{LDN}^{I,J}, if\ I_{SF}^{i,j} > T1 I_{GF}^{i,j} > T2\ \ (I_{LDN} < T3\ I_{LDN} \geq T4) \\ 0 \qquad\qquad , else \end{cases} \quad (1.24)$$

Where

$I_{LDN-F} - LDN\ feature\ image$

$T1 - Threshold\ 1\ (Herein, it\ is\ 10)$

$T2 - Threshold\ 2\ (Herein, it\ is\ 10)$

$T3 - Threshold\ 3\ (Herein, it\ is\ 234)$

$T4 - Threshold\ 4\ (Herein, it\ is\ 80)$

In the equation (1.24), the weak edge data of Sobel data and Gaussian data are neglected. A range of information from the LDN-image is approved and others are neglected. The neglected intensities are filled by zeros.

The LDN feature image I_{LDN_F} is binarized by applying the equation (1.25).

$$I_{LDN-FB}^{I,J} = \begin{cases} 1, & if\ I_{LDN\_F}^{i,j} > 0 \\ 0, & else \end{cases} \quad (1.25)$$

i ϵ [0, IH-1]

j ϵ [0, IW-1]

Where

I_{LDN_FB}

$- LDN\ feature\ image\ in\ binary\ format$

VI. LINKED MAP GENERATION

Linked-Map is a procedure of linking the text information of a scene text. It can be used to discard the false scene-texts from the I_{LDN_FB} image. If the gap of consecutive pixels between two non-zero points in the same row is shorter than 5% of the image width, they are filled with 1s. For example consider a pixel [i,j], if it is 1 then the column k of next-coming 0 in the horizontal direction of the same i^{th} row is found. Suppose the gap between these two points ([i,j] and [i,k-1]) are shorter than 5% of the image width, then a line of pixels from [i,j] and [i,k-1] are filled by 1s. If it is not true, that line of pixels are filled by 0s. If these connected components are smaller than the threshold value, then they are removed [17]18]. The threshold value is empirically selected by observing the minimum size of scene-text region. Afterwards each and every connected components is reshaped to possess smooth boundaries. Normally scene-texts are shaped in rectangular format, so a rectangular bounding box is created by linking four points. The four linking points can be observed as (minimum-y, minimum-x), (minimum-y, maximum-x), (maximum-y, minimum-x) and (maximum-y, maximum-x). The text areas which are not set with rectangular contents are also removed.

VII. SCENE-TEXT DETERMINATION AND LOCALIZATION

The determination of whether the selected block is a scene-text or not, is progressed by measuring the texture property of the block, through the usage of LDN-feature output, i.e. I_{LDN_F} (already computed by equation (1.24)). Normally in scene-text regions the texture property is higher than the scene-areas. The texture property γ is found based on equation (1.26).

$$\gamma = D_{LDN\_F} * d_{LDN\_F} \quad (1.26)$$

Where

$\gamma - Texture\ property\ value$

$D_{LDN\_F} - Density\ of\ LDN - features$

$d_{LDN\_F} -$

$Count\ of\ differnent\ LDN\ values$

The density of LDN_features D_{LDN_F} is computed through the division of 'number_of_LDN_features of each candidate region 'by the 'size of each candidate region'. The d_{LDN_F} is computed by counting the unique LDNs corresponding to the candidate region. If the texture property γ is larger than the threshold 24, the specific region is decided as the scene-text region, otherwise it is concluded as scene area[19][20]. The threshold value is empirically set to 24 based on various experiments on scene-text detection.

The multiline text regions can be separated into single-line texts by processing the histogram analysis for the specified regions. For example, the horizontal-direction oriented texts can be separated via horizontal projection histogram on noise-free image I_{NF} which can be easily understand via Fig.1.7.

Fig.1.7. Illustration of horizontal-orientation-text splitting by horizontal projection of histogram.

The vertical orientation scene-text regions are also processed by vertical projection and the new regions are undergone updated labelling

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 9, pp: 372-380 www.ijaem.net ISSN: 2395-5252

process. In this way, the entire scene-text regions are processed to handle the multi-line scene-texts.

The extracted scene-text regions are marked by rectangle shaped markings as the output of scene-text detection or localization, and this binary output is labelled as scene-text marked imageI_{ST}. Thus the scene-texts are detected from the scene-text images using the proposed STD_EED_LDN method.

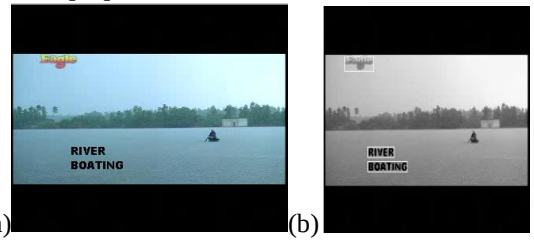
VIII.EXPERIMENTAL RESULTS AND ANALYSIS

This research implements the proposed onscene text detection and recognition method and the experimental results are tabulated in tables and drawn as chart to get results with clarity. This paper analyzes the proposed method against the following three existing methods related with onscene text detection and recognition.

1. Tong He et al. method [9]
2. Youbao Tang et al. method [13]
3. Sezer Karaoglu et al. method [14]

This analysis part is analysed with the state-of-the-art analytic methods to construct a better report on the performance behaviour of the proposed method compared with the existing methods.

This research uses two state-of-the-art onscene image databases namely KAIST [17] and ICDAR [18].



(a) (b)

Fig.2. Scene text detection by the proposed method, a) Input onscene image, b) scene text marked output.

Table1: Peak Signal to Noise Ratio (PSNR) for Text segmentation

Database Name	Image Name	PSNR(in db)			
		Tong. He method	Youbaou method	Sezerk method	Proposed method
KAIST	KAIST Img 1	58.34	59.18	61.27	**64.23**
	KAIST Img 2	58.86	60.48	62.21	**65.14**
	KAIST Img 3	57.08	58.87	60.46	**63.17**
	KAIST Img 4	57.52	59.44	61.92	**64.57**
ICDAR	ICDAR Img 1	56.82	58.43	60.73	**63.51**
	ICDAR Img 2	58.91	60.27	62.03	**64.31**
	ICDAR Img 3	56.28	57.10	59.54	**62.84**
	ICDAR Img 4	57.13	58.12	59.94	**62.52**

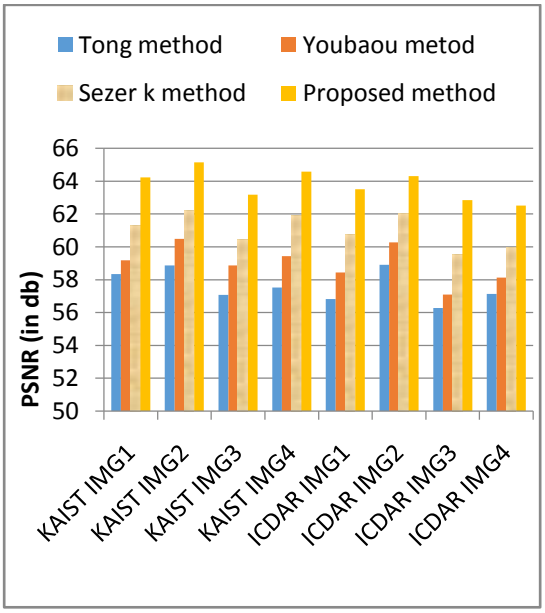


Fig.3. PSNR analysis chart for Scene text segmentation.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 9, pp: 372-380 www.ijaem.net ISSN: 2395-5252

The Figure 2 illustrates the output of the onscene text segmentation of the proposed method. The onscene texts are separated and marked in the Figure 2.b.

The PSNR analysis is performed using the binarized version of segmented image and the ground-truth image using the equation1.27

$$\text{PSNR} = 10\log_{10}\left(\frac{255^2}{\text{MSE}}\right) \quad (1.27)$$

Where

MSE –mean square error

The proposed method obtains the higher PSNR than the existing methods and it can be proven by the Table 1 and Figure 3. The Figure 3 represents the PSNR values for the KAIST database and ICDAR database [21][22]. This PSNR analysis clearly indicates the better segmentation of the proposed method than the existing methods because the proposed method reaches high PSNR than others for both databases. The highest value of PSNR related with proposed method for KAIST database is 65.14 and for ICDAR database it is 64.31.

Table2: Recall analysis for Scene text recognition

Database Name	Image Name	Recall			
		Tong. He method	Youbaou method	Sezer k method	Proposed method
KAIST	KAIST Img 1	72	74	77	**82**
	KAIST Img 2	72	73	76	**81**
	KAIST Img 3	71	74	78	**84**
	KAIST Img 4	70	72	76	**82**
ICDAR	ICDAR Img 1	72	73	75	**81**
	ICDAR Img 2	71	74	78	**83**
	ICDAR Img 3	72	73	77	**82**
	ICDAR Img 4	70	73	76	**83**

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (1.28)$$

Where

TP -True Positive

FN -False Negative

The positive class is referred by all the text in the recognized output file (whether it may be correctly or incorrectly recognized). The term true positive refers the correctly recognized items in the positive class [23].

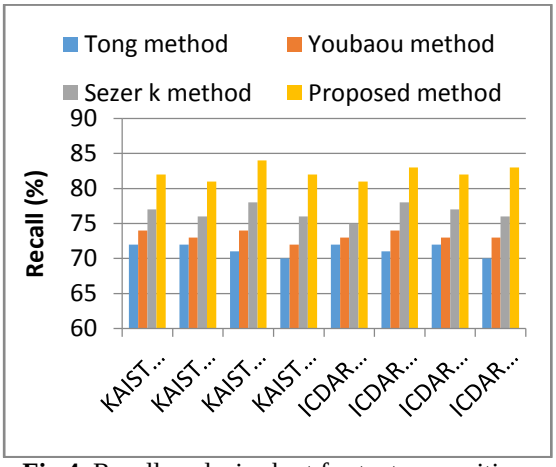


Fig.4. Recall analysis chart for text recognition.

The false negative term indicates the items that are not labelled as belonging to the positive class, but should have been recognized. The highest recall percentage of KAIST database for the proposed method is 84 and the highest recall of ICDAR database is 83. The proposed method achieves better recall percentage than the existing methodology which shows the extended character recognition power of the proposed method.

IX. CONCLUSION

The proposed method efficiently detects the texts positioned in onscene images and recognized them with higher accuracy. The main contribution of this manuscript is set with the scene text character recognition. The proposed method can be calibrated with set of n fonts. The proposed method produces higher recall for text recognition than the existing methods. The average recall obtained by the proposed method by considering the two databases is 83.5 which are much higher than the existing methods. This novel research can compete with the current onscene recognition schemes with better reliability and accuracy. In future this approach can be improved by optimized training with more fonts related with additional languages like Tamil and Hindi.

REFERENCES

[1]. Xilin Chen, Jie Yang, Jing Zhang, and Alex Waibel,"Automatic Detection and Recognition of Signs From Natural Scenes",

IEEE transactions on image processing, vol. 13, no. 1, pp. 87-99, 2004.

[2]. 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, pp. 401-411, 2009.

[3]. 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, no. 3, pp. 800-813, 2011.

[4]. Hyung Il Koo, and Duck Hoon Kim, "Scene Text Detection via Connected Component Clustering and Nontext Filtering",IEEE transactions on image processing , vol. 22, no. 6, pp. 2296-2305, 2013.

[5]. Cun-Zhao Shi, Chun-Heng Wang, Bai-Hua Xiao, Song Gao, and Jin-Long Hu, "Scene Text Recognition Using Structure-Guided Character Detection and Linguistic Knowledge" IEEE transactions on circuits and systems for video technology, vol. 24, no. 7, pp. 1235-1250, 2014.

[6]. Chucai Yi and Yingli Tian,"Scene Text Recognition in Mobile Applications by Character Descriptor and Structure Configuration", IEEE transactions on image processing, vol. 23, no. 7, pp. 2972- 2982, 2014.

[7]. Yao Li, Wenjing Jia, Chunhua Shen, and Anton van den Hengel,"Characterness: An Indicator of Text in the Wild", IEEE transactions on image processing, vol. 23, no. 4, pp. 1666-1677, 2014.

[8]. Huan Yang, Shiqian Wu, Chenwei Deng, and Weisi Lin,"Scale and Orientation Invariant Text Segmentation for Born-Digital Compound Images", IEEE transactions on cybernetics, vol. 45, no. 3, pp. 533-547, 2015.

[9]. Tong He, Weilin Huang, Yu Qiao, and Jian Yao,"Text-Attentional Convolutional Neural Network for Scene Text Detection", IEEE transactions on image processing, vol. 25, no. 6, pp.2529-2541, 2016.

[10]. Beula Bell T and M. K. Jeya Kumar, "Scene Text Segmentation and Recognition by Applying Trimmed Median Filter Using Energetic Edge Detection Schemes and OCR",Journal of Network Communications and Emerging Technologies, vol. 7, issue 12, pp. 57-66, 2017.

[11]. Sezer Karaoglu, Ran Tao, Jan C. van Gemert, and Theo Gevers,"Con-Text: Text Detection for Fine-Grained Object Classification", IEEE transactions on image processing, vol. 26, no. 8, pp. 3965 -3980, 2017.

[12]. Chun Yang, Xu-Cheng Yin, Wei-Yi Pei, Shu Tian, Ze-Yu Zuo, Chao Zhu, and Junchi Yan,"Tracking Based Multi-Orientation Scene Text Detection: A Unified Framework With Dynamic Programming",IEEE transactions on image processing, vol. 26, no. 7, pp. 3235-3248, 2017.

[13]. Youbao Tang and Xiangqian Wu,"Scene Text Detection and Segmentation Based on Cascaded Convolution Neural Networks",IEEE transactions on image processing, vol. 26, no. 3, pp. 1509- 1520, 2017.

[14]. Sezer Karaoglu, Ran Tao, Theo Gevers, and Arnold W. M. Smeulders,"Words Matter: Scene Text for Image Classification and Retrieval", IEEE transactions on multimedia, vol. 19, no. 5, pp.1063-1076, 2017.

[15]. Minghui Liao, Baoguang Shi, and Xiang Bai ,"TextBoxes++: A Single-Shot Oriented Scene Text Detector", IEEE transactions on image processing, vol. 27, no. 8, pp. 3676-3690, 2018.

[16]. Yingying Zhu, Minghui Liao, Mingkun Yang, and Wenyu Liu,"Cascaded Segmentation-Detection Networks for Text-Based Traffic Sign", IEEE transactions on intelligent transportation systems, vol. 19, no. 1, pp. 209-219, 2018.

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

[18]. ICDAR-TEXT-DATASET, https://github.com>Total-Text-Dataset

[19]. 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.

[20]. Boran Yu and Hongjie Wan,"Chinese Text Detection and Recognition in Natural Scene Using HOG and SVM", International Conference on Information Technology for Manufacturing Systems, pp.148-152, 2016.

[21]. P. Shiva Reddy and M.N.Giri Prasad, "Extracting text from natural scene images by HOG Character Descriptor", International Journal of Advanced Research in Electronics and Communication Engineering, vol. 5, pp. 2518-2523, 2016.

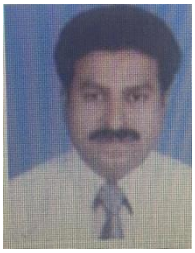
[22]. Ankush Gautam,"Segmentation of Text from Image Document", International Journal of Computer Science and Information Technologies, vol. 4 (3), pp. 538-540, 2013.
[23]. Everton B. Lacerda and Carlos A.B Mello,"Segmentation of Touching Handwritten Digits Using Self-Organizing Maps", IEEE International Conference on Tools with Artificial Intelligence, 2011.

Authors

T. Beula Bell is an Associate Professor, Department of Computer Applications in

Nesamony Memorial Christian College, Marthandam, India. She has 21 years of teaching experience. She has completed research under the guidance of Dr. M. K. JeyaKumar he is working as a Professor in the Department of Computer Applications, Noorul Islam University, Kumaracoil, Tamilnadu, India . She received B.Sc, M.C.A., M.Phil degree from M. S. University, Thriunelveli, India. Her area of interest is image processing, Computer Graphics.

M. K. JeyaKumar 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.