Volume 5, Issue 11, November 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Multilingual Text Detection in Natural Scene Images using Wavelet based Edge Features and SVM Classifier

Shivananda V. Seeri
Department of MCA, KLE Technological University, BVBCET Campus, Hubballi, Karnataka, India

J. D. Pujari
Department of I S E, S.D.M. College of Engg. & Tech., Dharwad, Karnataka, India

P. S. Hiremath
Department of MCA, KLE Technological University, BVBCET Campus, Hubballi, Karnataka, India

Abstract— The aim of this study is to propose a new methodology for text region localization and non- text region removal in natural scene images with complex background. In this paper, a new hybrid approach is proposed which locates multilingual text with different complex backgrounds. The proposed approach involves two steps. First, potential text regions in an image are extracted based on edge features in wavelet transformed image. In the second step, potential text regions are tested for text or non-text content using GLCM features and SVM classifier. The text localization algorithm is designed to locate text in natural scene images. The performance of the proposed approach is demonstrated by presenting promising experimental results in terms of precision and recall rates, for a set of natural scene images taken from ICDAR dataset and own data set. The results demonstrate effectiveness of the proposed method, which can be used as an efficient pre-processing technique for text recognition in natural scene images.

Keywords— Natural scene image, Text localization, Haar wavelet, Sobel edge detector, SVM classifier, GLCM

I. INTRODUCTION

Artificially or naturally contained texts in natural scene images have significant and detailed information about the scenes. If one develops a method that extracts and recognizes texts accurately in real-time, the method can be applied to many important applications such as visually impaired people's assistance, helping foreigners with language barrier, automatic detection of vehicle number plate, identification of parts in industrial automation, street signs, robotics and intelligent transport systems. The extraction of text information involves the detection and recognition of text from the input natural scene image captured by a digital camera. Automatic text localization in natural scene images is a primary computer vision task in such applications. Text appearing in images can provide very useful semantic information to describe the image content.

(a)

(b)

Figure 1: Sample digital images containing (a)) Scene text (b) Caption text

In general, text appearing in images can be classified into two groups: scene text and caption text. A scene text is a part of an image, whereas a caption text is inserted on an image in a later stage. Caption text is often used for successful indexing and retrieval of images (or videos). The text extraction from real images faces numerous challenges due to lower resolution, unknown text colour, size and position, or complex backgrounds. The Figure 1 shows these two types of text present in the sample images.

In this paper, a new hybrid approach is proposed which locates multilingual text with different complex backgrounds. The proposed approach involves two steps. First, potential text regions in an image are extracted based on edge features in wavelet transformed image. In the second step, potential text regions are tested for text or non-text content using Gray Level Co-occurrence Matrix (GLCM) features and Support Vector Machine (SVM) classifier.

II. BACKGROUND LITERATURE

Text extraction in an image comprises four stages, namely, (i) Detection, (ii) Localization, (iii) Extraction and (iv) Recognition (OCR). Text detection and text localization are closely related and more challenging stages which has attracted the attention of many researchers. The goal of the two stages is to mark exact bounding boxes for all text objects in a natural scene image and provide a unique identity to each text. In this section, the recent literature focused on text detection and localization is reviewed.

A. Region based methods

Region-based methods use the properties of the colour or gray-scale in a text region or their differences with the corresponding properties of the background. These methods use a bottom-up approach by grouping small components into successively larger components until all regions are identified in the image. A geometrical analysis is needed to merge the text components using the spatial arrangement of the components so as to filter out non-text components and mark the boundaries of the text regions.

A region based method is basically classified in to two sub categories, namely, edge based and connected component (CC) based methods. An edge based method is mainly focused on the high contrast between text and background region. In this method, firstly text edges are identified in an image and are merged. Finally, some heuristic rules are applied to discard non-text regions. Edges are a reliable feature of text regardless of colour/intensity, layout, orientations, etc. Edge strength, density and the orientation variance are three distinguishing characteristics of text embedded in images, which can be used as main features for detecting text. The edges of the text boundary are identified and merged, and then several techniques are used to filter out the non-text regions. Edge-based text extraction algorithm is a general-purpose method, which can quickly and effectively localize and extract the text from both document and natural images.

A connected component based method considers text as a set of separate connected components, each having distinct intensity and colour distributions. It is based on the analysis of the geometrical arrangement of edges or homogeneous colour and gray-scale components that belong to characters.

The edge based methods are robust to low contrast and different text size, whereas CC based methods are somewhat simpler to implement, but fail to localize text in images with complex backgrounds.

Vijayakumar et al. [1] proposed a method for detecting video text regions containing player information and score in sports videos. Key frames from the video were extracted using colour histogram technique to minimize the number of video frames and then converted to gray images. Text image regions were cropped. Canny edge detection algorithm was applied to detect edges on the cropped image. From this edge detected images, text region was identified and fed to an Optical Character Recognition (OCR) system which produces index-able keywords.

Chitrakala et al. [2] proposed an unified method to extract a text region from heterogeneous images such as scene text images, caption text images and document images by using contourlet transform. Contourlets possess the main features of wavelets (namely, multiscale and time-frequency localization), and also offer a high degree of directionality and anisotropy.

Xiaopei Liu et al. [3] proposed new scheme for character detection and segmentation from natural scene images. In the detection stage, stroke edge is employed to detect possible text regions, and some geometrical features are used to filter out obvious non-text regions. A graph model of candidate text regions is set up, and the graph cut algorithm is utilized to classify candidate text regions as text or non-text.

Ankita Sikdar et al. [16] proposed a method based on histogram thresholding, entropy filtering and connected components to extract Bengali text and Bengali characters from a certain level of complex multimedia images with a desired accuracy level.

Xiaoqing et al. [4] has developed a method consisting of three stages: candidate text region detection, text region localization and character extraction. In the first stage, the magnitude of the second derivative of intensity as a measurement of edge strength is used, since it allows better detection of intensity peaks that normally characterize text in images. The edge density is calculated based on the average edge strength within a window. Considering effectiveness and efficiency, four orientations (0^0, 45^0, 90^0, 135^0) are used to evaluate the variance of orientations, where 0^0 denotes horizontal direction, 90^0 denotes vertical direction, and 45^0 and 135^0 are the two diagonal directions. Edge detection is carried out by using a multi-scale strategy, where the multi-scale images are produced by Gaussian pyramids after successively applying low-pass filter and down-sample the original image reducing the image in both vertical and horizontal directions. In the second stage, characteristics of clustering can be used to localize text regions. In the third stage, existing OCR engine was used for character recognition.

Xin Zhang et al. [5] has used the colour and edge features to extract the text from a video frame. In this work, two methods are combined, called colour-edge combined algorithm, to remove text background. One method is based on the exponential changes of text colour, called transition map model, while the other one uses gray level image. After removing complex background, text location is determined using the vertical and horizontal projection method. This algorithm is robust to the images with multilingual text. To improve the efficiency of this method, the edge feature is added to remove background, and then edge detection is performed on each colour component image using Canny operator and morphology operation. Finally, the background of text is removed with the help of transition map model.

Parthasarathi Giri et al. [14] have compared two basic approaches for extracting text region in images, which are edge-based and connected-component based. The algorithms are implemented and evaluated using a set of images that vary in terms of lighting, scale and orientation. Accuracy, precision and recall rates for each approach are analyzed to determine the success and limitations of these approaches.

B. Morphology based methods

Mathematical morphology is a topological and geometrical based approach for image analysis. It provides powerful tools for extracting geometrical structures and representing shapes in many applications. Morphological feature extraction techniques have been efficiently applied for character recognition and document image analysis. It is used to extract dominant text contrast features from the processed images. The feature is invariant under geometrical image transformation, namely, translation, rotation, and scaling. Even after the lighting condition or text colour is changed, the feature can still be maintained.

Jui- Chen Wu et al. [6] presented a morphology based text line extraction algorithm for extracting text regions from cluttered images. The method defines a novel set of morphological operations for extracting important contrast regions as possible text line candidates. In order to detect skewed text lines, a moment-based method is then used for estimating their orientation. According to the orientation, an x-projection technique can be applied to extract various text geometries from the text-analogue segments for text verification. However, due to noise, a text line region is often fragmented into different pieces of segments. Therefore, after the projection, a novel recovery algorithm is then proposed for recovering a complete text line from its pieces of segments. After that, a verification scheme is then proposed for verifying all extracted potential text lines according to their text geometries.

Rama Mohan et al. [7] performed the edge detection operation using the basic operators of mathematical morphology. Using the edge image, the algorithm found out candidate connected components with text. These components have been labeled to identify different components of the image. Once the components have been identified, the variance is found for each connected component considering the gray levels of those components. Then the text is extracted by selecting those connected components whose variance is less than some threshold value.

C. Texture based methods

Texture based methods consider texts as regions with distinct textural properties. Methods of texture analysis are used to analyze text regions. The main drawback of such methods is their complexity, while these are more robust than the CC based methods in dealing with complex background.

Chu Duc et al. [8] presented a novel texture descriptor based on line-segment features for text detection in images and video sequences, which is applied to build a robust car license plate localization system. Unlike most of the existing approaches which use low level features (colour, edge) for text/non-text discrimination, the aim is to exploit more accurate perceptual information. Scale and rotation invariant texture descriptors which describe the directionality, regularity, similarity, alignment and connectivity of group of segments are proposed. An improved algorithm for feature extraction based on local connective Hough transform has also been investigated.

Kwang et al. [9] used a novel texture based method for detecting texts in images. A support vector machine (SVM) is used to analyze the textural properties of texts. The intensities of the raw pixels that make up the textural pattern are fed directly to the SVM, which works well even in high-dimensional spaces. Next, text regions are identified by applying a continuously adaptive mean shift algorithm (CAMSHIFT) to the results of the texture analysis. The combination of CAMSHIFT and SVMs produces both robust and efficient text detection, as time-consuming texture analyses for less relevant pixels are restricted, leaving only a small part of the input image to be texture-analyzed. The performance criterion was the classification accuracy of the SVMs for text and non-text patterns rather than the overall text detection results. For this purpose, 100 training images were divided into two different classes of 70 training images and 30 validation images from which training patterns and validation patterns were collected, respectively. The SVMs were then trained using the training patterns and then tested using the validation patterns.

Neha et al. [10] proposed a new hybrid approach to locate text in different backgrounds. The text localization algorithm system is designed to locate text in different kinds of images and eliminates the need to devise separate method for various kinds of images. The approach uses various concepts, namely, Haar discrete wavelet transform (DWT), Sobel edge detector, morphological dilation and connected components. Finally, using some specific condition, the text is extracted in a bounding box. This algorithm is not sensitive to image colour or intensity, uneven illumination and reflection effects. It can be used in large variety of application fields, e.g. vehicle license plate detection to detect number plate of vehicle, mobile robot navigation to detect text based land marks, object identification, identification of various parts in industrial automation, analysis of technical papers with the help of charts, maps, and electric circuits. The algorithm only analyses a text box but not a character. Therefore, it requires less processing time which is essential for real time applications.

Goel et al. [11] proposed the algorithm for text extraction in colour images using 2D Haar wavelet transform (2D-DWT) along with mathematical morphological operators. The algorithm is tested for vehicle plate text extraction and also the document images.

Sumathi et al. [12] proposed a new methodology for text region extraction and non-text region removal from complex background coloured images. This method is based on Gamma correction by determining a gamma value for enhancing the foreground details in an image. The approach also uses gray level co-occurrence matrices, texture measures and threshold concepts. The proposed method is a useful pre-processing technique to remove non text region and to show the text region in an image. Experimental results show good performance of the proposed method for extracting text regions in an image.

Adesh Kumar et al. [13] proposed an approach based on Haar DWT, Sobel edge detector, and morphological operator. These mathematical tools are integrated to detect the text regions from the complex images. The proposed method is robust against language scripts and font size of the texts. The proposed method is also used to decompose the blocks including multi-line texts into single line text.

Deepa et al. [15] proposed text extraction algorithm for Tamil text extraction from images. The method uses dual tree complex wavelet transform and morphological dilation. It is robust against various conditions such as shadows, degradations, non-uniform illuminations, highlights, different font style and size and low contract images. The experimental results show that the proposed method reasonably extracts text regions with elimination of most non -text regions. This can be further binarized and used by visually challenged persons through text-to-audio conversion software.

Satish Kumar et al. [17] proposed a method to localize potential text regions using connected component based heuristics from colour components (RGB). The system takes coloured images as input; it detects text on the basis of certain text features such as frequency, orientation and spatial cohesion.

Keshava Prasanna, et al. [18] proposed an efficient algorithm to detect, localize and extract Kannada text from images with complex backgrounds. The proposed approach is based on the application of a colour reduction technique, a standard deviation based method for edge detection, and the localization of text regions using new connected component properties.

Murthy et al. [19] proposed a novel approach of extracting a text from image using Haar wavelet transformation and k-means clustering. The proposed system has been experimented with images with single / multiple text, multiple text of different sizes / style / languages, images with uniform and non-uniform background.

Angadi et al. [20] proposed a methodology to detect and extract text regions from low resolution natural scene images. The proposed work is texture based and uses DCT based high pass filter to remove constant background. The texture features are then obtained on every 50x50 block of the processed image and potential text blocks are identified using newly defined discriminant functions. Further, the detected text blocks are merged and refined to extract text regions. The proposed method is robust and achieves a detection rate of 96.6% on a variety of 100 low resolution natural scene images each of size 240x320.

Bhattacharya et al. [21] proposed a novel and effective scheme based on analysis of connected components for extraction of Devanagari and Bangla texts from camera captured scene images. The proposed scheme uses mathematical morphology operations for the extraction. The proposed algorithm works well even on slanted or curved text components of Devanagari and Bangla.

Seeri et al. [22] proposed a hybrid approach, which locates text in natural scene images with different backgrounds. The proposed approach is based on edge feature extraction in wavelet domain and fuzzy classification, and is tested for multilingual text in natural scene images with complex backgrounds.

Leon et al. [23] presented a new robust technique for caption text detection. The technique combines texture information using Haar wavelet decomposition and geometric features extract caption text objects in the scene. The region-based image model is adopted to extract caption text from the images.

Syal et al. [24] proposed a hybrid method to extract the text in images, based on fusion of Daubechies DWT, radiant difference and SVM along with some pre-processing and post-processing steps. This algorithm is tested with images taken from various newspapers, magazines, commercial products and images collected from ICDAR 2003 dataset.

Ezaki et al. [25] proposed a system that reads the text encountered in natural scenes with the aim to provide assistance to the visually impaired persons. The method was based on Sobel edge detection, Otsu binarization, connected component extraction and rule-based connected component filtering.

III. PROPOSED METHODLOGY

The block diagram of the proposed method for text localization in a natural scene image is shown in the Figure 2. The input image may be a colour or gray scale image. If the image is colour image, then it is converted into grayscale image and pre-processing operation is applied on the resultant image. In the algorithm, a colour image is an input to the system and the localized text with black background is the output. The proposed method consists of four stages, namely, pre-processing, feature extraction, potential text region detection, and classification which are explained below.

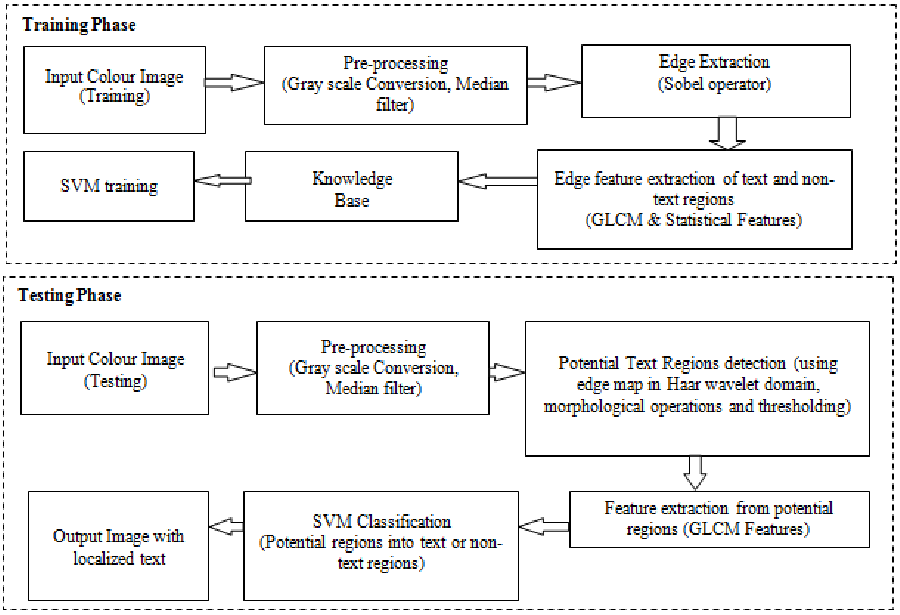


Figure 2: Block diagram of the proposed method.

A. Preprocessing

The input image is pre-processed to facilitate easier detection of text regions and to reduce the processing overload. The input image is RGB colour image. The input image is converted to grayscale image. It will reduce the computational complexity as well as memory requirements significantly. The grayscale image is then filtered using a median filter which removes any noises and preserves sharp edges in the image.

B. Feature extraction

Textural features are used to train Support Vector Machine (SVM) to classify the extracted component from natural scene image as a text or non-text. Texture analysis refers to the characterization of regions in an image by their texture content. Texture is one of the significant characteristics used in identifying objects or regions of interest in an image. Texture can be defined as a regular repetition of an element or pattern on a surface. Texture is an important approach to describe the regions of image. The Gray Level Co-occurrence Matrix (GLCM) features are used to train the SVM for the classification.

Gray Level Co-occurrence Matrix (GLCM) has proved to be a popular statistical method of extracting textural feature from images. From the co-occurrence matrix, Haralick defines fourteen textural features to extract the characteristics of texture statistics of remote sensing images. The four important features, namely, Contrast, Homogeneity, Energy and Entropy that are considered for training SVM are described below.

GLCM is defined as a two dimensional histogram of gray levels for a pair of pixels, which are separated by a fixed spatial relationship. GLCM of an image is computed using a displacement vector, defined by its radius δ and orientation θ. In the present study, the chosen values are δ=1 and θ=0, 45, 90 and 135 degrees. The GLCM C(i, j) represents joint probability of occurrence of pixels with intensities i and j.

Contrast: Contrast evaluates local gray level variation. This statistic measures the spatial frequency of an image and is a difference moment of GLCM.

$$F_1 = ij(i-j)^2 \sum\sum C(i,j) \quad (1)$$

Homogeneity/Inverse Difference Moment: It is a measure of the amount of local uniformity present in the image. It is high when local gray level is uniform and inverse GLCM is high.

$$F_2 = \sum\sum \frac{C(i,j)}{[1+(i-j)*2]} \quad (2)$$

Energy: This statistic, also called Uniformity or Angular Second Moment (ASM), describes image smoothness. It measures the textural uniformity, that is, pixel pair repetitions. It is the sum of squares of entries in the GLCM. ASM measures the image homogeneity. It is high when image has very good homogeneity or when pixels are very identical.

$$F_3 = \sum\sum C(i,j) \quad (3)$$

Entropy: This statistic measures the disorder or complexity of an image. The entropy is large when the image is not texturally uniform or has complex textures.

$$F_4 = \sum\sum ijC(i,j) \log C(i,j) \quad (4)$$

In the training phase, Sobel edge detector is used to locate strong edges that are essential features in this proposed method. The GLCM features F1 to F4 and statistical features, namely, standard deviation and mean absolute deviation from the training text and non-text images are computed, which are stored as knowledge base and then used to train SVM classifier.

C. Potential text region detection and SVM classification

In the testing phase, Haar wavelet transform is applied on edge map of the test image as obtained in the pre-processing stage. The morphological operations are applied on H, V and D components (Figure 3) of wavelet transformed image, which yield segmented images. The connected components in the segmented images are filtered based on thresholding the geometric properties of the connected components. The filtered connected components are the detected potential text regions in the test image, which will be classified as text or non-text region using the trained SVM classifier.

A	H
V	D

Figure 3: Components of 2-D Haar wavelet transform

Training Algorithm:

Input: A digital natural scene image containing text.

Output: Feature vectors of the text and non-text regions in the input image.

Step 1: Input RGB Image I1.

Step 2: Convert input image I1 to grayscale image I2 and apply median filter to obtain image I3.

Step 3: Apply Sobel edge operator on I3 to obtain edge map I4.

Step 4: Compute GLCM features F1 to F4 and statistical features, namely, standard deviation and mean absolute deviation from edge map I4 and store in knowledge base.

Step 5: Repeat Step 1 to Step 4 for all the natural scene images in the training set.

Step 6: Train the SVM classifier using knowledge base obtained in the Step 4.

Step 7: Stop

Testing Algorithm:
Input: A digital natural scene image I1.
Output: The natural scene image I1 with localized text contained in bounding boxes.
Step 1: Input Image I1
Step 1: Convert input RGB image I1 to grayscale image I2 and apply median filter to obtain image I3.
Step 2: Apply Sobel edge operator on I3 to obtain edge map I4. Then apply Haar wavelet transform to I4 to yield H, V, D components.
Step 3: Apply morphological operation (dilation with structuring element disk of radius 3 fixed empirically) on H, V and D components obtained in Step 4.
Step 4: Combine H, V and D components to obtain image CI = (H|V)&D. Then multiply CI to A, H, V and D components. The resulting components are used for reconstruction of segmented image I5 using inverse Haar wavelet transform.
Step 5: After labeling connected components in I5, geometric properties, namely, Area, Eccentricity, Extent and Aspect Ratio are computed for each labeled component.
Step 6: The components are filtered using the following empirical criteria: (Area >=300 AND Eccentricity > 0.4) AND (Extent > 0.5 OR Aspect Ratio > 1.5)
The filtered connected components are the detected potential text regions in I5 and the corresponding region in the input test image I1 are cropped.
Step 7: Each cropped image is subjected to feature extraction process and then classified as text or non-text using the trained SVM classifier.
Step 8: Draw the bounding box for each detected text region in I1 and remove the background, thus yielding localization of text in the input image I1.
Step 9: Stop

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

The proposed algorithm is implemented using Intel Core i5 processor @ 2.5 GHz 4GB RAM and MATLAB R2009a. The experimental data set contains 324 natural scene images (taken from ICDAR data set and own data set), out of which 224 are used as training images and 100 as testing images. The training set contains 112 text images and 112 non-text images, which are used for training the SVM classifier. Each image is a RGB colour image in JPEG format. These images contain multilingual texts with complex background, which vary with respect to font style, font size, scale, lighting and orientation of text in the image. The scripts of the text in the training images are English, Kannada and Hindi, whereas text in test images are in different languages, namely, English, Kannada, Hindi, Tamil, Bengali, Telugu, Urdu, Punjabi, Oriya, Malayalam, Japanese, Russian and Chinese. The sample results of the proposed algorithm are shown in the Figure 4. The performance of the proposed algorithm has been evaluated in terms of precision, recall, F-measure and accuracy, which are defined below.

True Positives (TP) are those regions in the image that are actually text and have been detected by the algorithm as text.

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.

True Negatives (TN) are those regions in the image which are actually not characters of a text and have been detected by the algorithm as non-text.

False Negatives (FN)/ Misses are those regions in the image which are actually text characters, but have not been detected by the algorithm.

$$\text{Precision} = \frac{TP}{TP+FP} \text{ X } 100 \quad (5)$$

$$\text{Recall} = \frac{TP}{TP+FN} \text{ X } 100 \quad (6)$$

$$F-\text{measure} = \frac{2*\text{Precision}*\text{Recall}}{(\text{Precision}+\text{Recall})*100} \quad (7)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \text{ X } 100 \quad (8)$$

The proposed method has yielded the precision and recall rates of 97.14% and 89.81%, respectively. The F-measure is obtained as 0.94 which is close to 1. The accuracy of the proposed method is 88.99%. The experimental results demonstrate the effectiveness of the proposed method, which are better than the results obtained by the method in [22]. Further, it is observed that the method is able to localize text regions of different scripts in natural scene images even though the algorithm is trained with only three scripts, namely, English, Kannada and Hindi, which demonstrates the robustness of the proposed method. It is also noticed that the localized text regions may contain multiple lines of text of different scripts, which need to be segmented for character extraction. The performance comparison of the proposed method with the other methods in literature is given in the Table 1.

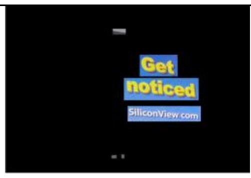
TABLE 1. THE PERFORMANCE COMPARISON OF THE PROPOSED METHOD WITH THE OTHER METHODS

Methods	Techniques Used	Precision (P) % Recall (R)% F-measure (F) Accuracy (A)%	Benefits

1. Proposed method (2015)	Wavelet Transform GLCM, statistical features, geometric properties, Mathematical Morphology, SVM classifier	P=97.14 R=89.81 F=0.94 A= 88.99	Multilingual, Robust to style, font and complex backgrounds
2. Seeri, et al. [22] (2015)	Wavelet Transform, Statistical features, fuzzy classification, K-Means clustering	P=95.54 R=89.14 F=0.92	Multilingual, Robust to style, font and complex backgrounds
3. Leon, et al. [23] (2013)	Wavelet Transform, Hierarchical image model	P=91.0 R=79.55 F=0.85 A=85.78	Insensitive to Different size, colour, complex background for caption text
4. Sumathi, et al. [12] (2012)	Gamma correction, GLCM and Thresholding concepts	P=78.0 R=91.0 F=0.84	Robust to size, style, font and complex background for English text
5. Pavithra, et al. [24] (2014)	Wavelet Transform, Gabor filter, K-Means clustering	P=98.0 R=80.20 F=0.88	Multilingual text detection, Robust to horizontal texts.

V. CONCLUSION

Text localization in a natural scene image with complex background is a difficult, challenging and important problem. In the present study, a novel method is proposed for text detection and localization that involves two steps. First, potential text regions in an image are extracted based on edge features in wavelet transformed image. In the second step, the potential text regions are tested for text content or non-text using GLCM features and SVM classifier. The proposed method has yielded the F-measure and Accuracy of 0.94 and 88.99%, respectively. The experimental results demonstrate the effectiveness and robustness of the proposed method in detection and localization of multilingual text regions with variations in font style, font size, scale, lighting and orientation of text in an image. The localized text regions may contain multiple lines of text of different scripts that need to be segmented for character extraction, which will be considered in future work.

Sl. No	Input Image	Text localized Image	Sl. No	Input Image	Text localized Image
1			11		
2			12		
3			13		
4			14		
5			15		

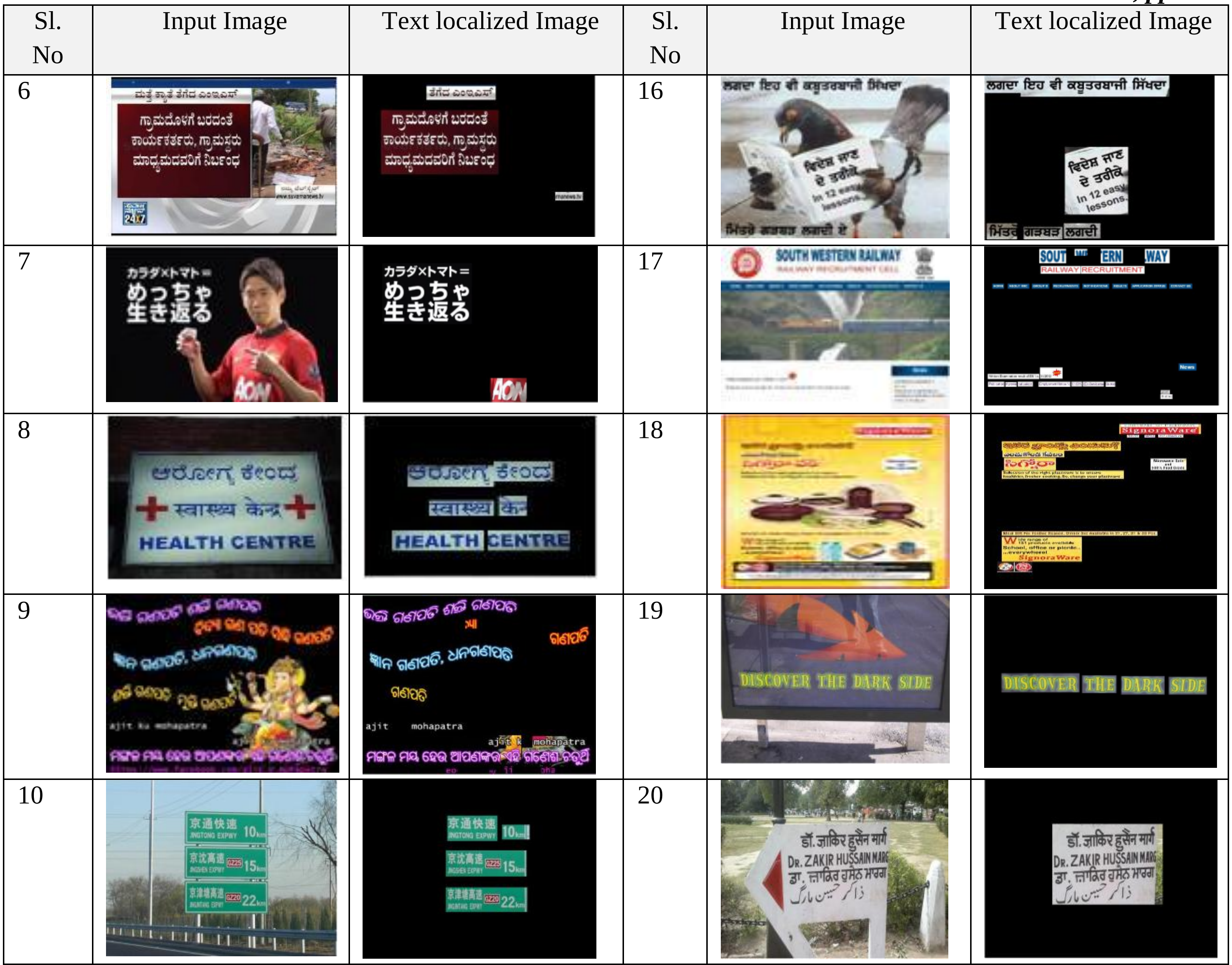


Figure 4: Sample results of the proposed method showing input images and corresponding text localized output images.

ACKNOWLEDGMENT

The authors acknowledge the management of K.L.E Technological University, Hubballi, for the financial assistance sanctioned under Capacity Building Projects scheme.

REFERENCES

[1] V.Vijayakumar, R.Nedunchezhian, A novel method for super imposed text extraction in a sports video, International Journal Of Computer Applications,Volume 15– No. 1, 0975 – 8887, February 2011, pp. 1-6.

[2] Chitrakala Gopalan and Manjula, Text region segmentation from heterogeneous images, IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.10, October 2008, pp. 108-113.

[3] Xiaopei Liu, Zhaoyang Lu, Jing Li, Wei Jiang, Detection and segmentation text from natural scene images based on graph model, WSEAS Transactions on Signal Processing, E-ISSN: 2224-3488, Volume 10, 2014, pp. 124-135.

[4] Xiaoqing Liu and Jagath Samarabandu, Multiscale edge-based text extraction from complex images , multimedia and expo, 2006 IEEE International Conference, ISBN 1-4244-0366-7, 2006, pp. 1721-1724.

[5] Xin Zhang, Fuchun Sun, Lei Gu, “A combinedalgorithm for video text extraction”, 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery,ISBN 978-1-4244-5931-5,2010, pp. 2294 – 2298.

[6] Jui-Chen Wu, Jun-Wei Hsieh, Yung-Sheng Chen, Morphology-based text line extraction, International Journal of Machine Vision and Applications, P-ISSN: 0932-8092, Volume 19, Issue 3, 2008, pp. 195-207.

[7] Rama Mohan Babu, G. Srimaiyee, P. Srikrishna, Text extraction from heterogeneous images using mathematical morphology, Journal Of Theoretical And Applied Information Technology, Vol. 16 Issue 1/2, 2010, pp. 39-47.

[8] Chu Duc Nguyen, Mohsen Ardabilian and Liming Chen, Robust car license plate localization using a novel texture descriptor, Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, E-ISBN : 978-0-7695-3718-4, 2009, pp. 523 – 528.

[9] Kwang In Kim, Keechul Jung, And Jin Hyung Kim, Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm, IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 25, No. 12, ISSN: 0162-8828, 2003, pp. 1631 – 1639.

[10] Neha Gupta, V.K. Banga, Localization of text in complex images using haar wavelet transform, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-1, Issue-6, November 2012, pp. 111-115.
[11] Akash Goel, Yogesh Kumar Sharma, Text extraction of vehicle number plate and document images using discrete wavelet transform in Matlab, IOSR Journal of Computer Engineering (IOSR-JCE), e-ISSN: 2278-0661, p- ISSN: 2278-8727, Volume 16, Issue 2, Ver. V (Mar-Apr. 2014), pp. 117-126.
[12] Sumathi C. P. and G. Gayathri Devi, Automatic text extraction from complex coloured images using gamma correction method, Journal of Computer Science, (Online): 10 (4) 2014, ISSN: 1549-3636, 2014, pp. 705-715.
[13] Adesh Kumar, Pankil Ahuja, Rohit Seth, Text extraction and recognition from an image using image processing in Matlab, Conference on Advances in Communication and Control Systems 2013 (CAC2S 2013), pp. 429-435.
[14] Partha Sarathi Giri, Text information extraction and analysis from images using digital image processing techniques, Special Issue of International Journal on Advanced Computer Theory and Engineering (IJACTE), ISSN (Print) : 2319 – 2526, Volume-2, Issue-1, 2013, pp. 66-71.
[15] S.T. Deepa, S.P.Victor, Tamil text extraction, International Journal of Engineering Science and Technology (IJEST), ISSN: 0975-5462, Vol. 4 No.05 May 2012, pp. 2176-2179.
[16] Ankita Sikdar, Payal Roy, Somdeep Mukherjee, Moumita Das and Sreeparna Banerjee, A Two Stage Method For Bengali Text Extraction From Still Images Containing Text, NatarajanMeghanathan, et al. (Eds): SIPM, FCST, ITCA, WSE, ACSIT, CS & IT 062012. © CS & IT-CSCP 2012, DOI : 10.5121/csit.2012.2306, 2012, pp. 47–55,
[17] Satish Kumar, Sunil Kumar, S. Gopinath, Text extraction from images, International Journal of Advanced Research in Computer Engineering & Technology, ISSN: 2278 – 1323, ISSN: 2278 – 1323, 2012, pp. 34-36.
[18] Keshava Prasanna, Ramakanth Kumar P, Thungamani.M, Manohar Koli, Kannada text extraction from images and videos for vision impaired persons, International Journal of Advances in Engineering &Technolog, ISSN: 2231-1963, Nov 2011, pp. 189-196.
[19] Narasimha Murthy K N, Y S Kumaraswamy, A novel method for efficient text extraction from real time images with diversified background using Haar discrete wavelet transform and k-means clustering, IJCSI International Journal of Computer Science Issues, Vol. 8, Issue , No 3, ISSN (Online): 1694-0814,September 2011, pp. 235-245
[20] S. A. Angadi, 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, November 2009, pp. 229-245.
[21] U. Bhattacharya, S. K. Parui and S. Mondal, Devanagari and Bangla text extraction from natural scene images, 10th International Conference on Document Analysis and Recognition, 978-0-7695-3725-2/09 $25.00 © 2009 IEEE, DOI 10.1109/ICDAR.2009, pp. 171-178.
[22] S. V. Seeri, J. D. Pujari, P. S. Hiremath, Multilingual text localization in natural scene images using wavelet based edge features and fuzzy classification, International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 4, Issue 1, ISSN 2278-6856, February 2015, pp. 210-218.
[23] Miriam Leon, Veronica Vilaplana, Antoni Gasull and Ferran Marques, Region-based caption text extraction, Lecture Notes in Electrical Engineering, Springer, vol. 158, ISSN1876-1100, 2013, pp. 21-36
[24] NitiSyal, Naresh Kumar Garg, Text extraction in images using DWT, gradient method and SVM classifier, International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 4, Issue 6, June 2014, pp. 477-481.
[25] Nobuo Ezaki, Marius Bulacu, Lambert Schomaker, Text detection from natural scene images: towards a system for visually impaired persons, Proc. of 17th Int. Conf. on Pattern Recognition (ICPR 2004), IEEE Computer Society, vol. II, 23-26 August, Cambridge, UK.,2004, pp. 683-686.

