A Technique for Skew Correction and Text Localization in Natural Scene Images

Navneet Kaur[1], Dr. Rajesh Kumar Bawa[2], Dr. Ganesh Kumar Sethi[3]
[1]Research Scholar, [2]Professor, [3]Assistant Professor,
[12]Punjabi University, Patiala
[3]Multani Mal Modi College, Patiala

Abstract- The image processing is the technique which is used to process the digital information. The text extraction is the major challenge due to robust image information. The images which are natural scene has not the equal contrast and also has the skewness due to which it is difficult to extract the text from the image. In this research work, the PCA algorithm is applied for the extraction and also threshold based technique is applied for the skew removal. The proposed algorithm is implemented in MATLAB and results in terms of certain parameters

Keywords- PCA, MATLAB, Skeness removal

I. INTRODUCTION

The technology through which raw images that are collected from cameras available on various sources can be improved such that the information that is important can be extracted from them is known as image processing. Several techniques have been developed over the time within this technology to provide enhancements such that the complex data can be extracted. Digital image processing is another name of image processing. It however provides both optical and analog image processing in it [1]. Several techniques are applied in this technology, which are studied here. Imaging is defined as the image acquisition. It is feasible to perform optical and analog image processing. Several fields such as computer graphics are available in this approach through which it is possible to generate images. Within various applications, the major issue is to extract text from images or videos. Within surrounding sign boards and objects, there are various text characters and strings found which are categorized into natural scene images and videos. The information related to the region and the objects can easily be known by studying this data. The resolution and quality of the natural scene images is very less and the background is complex as well. Therefore, it is very difficult to recognize, extract or identify the scene text since it is possible that the text is either slant, tilt or occluded partially [2]. Recently, several text detection techniques have been proposed by researcher for natural scene images. It is important to include scene text detection and recognition algorithms that are automatic and efficient such that the text information can be extracted from natural scenes with the help of mobile devices. The text within natural scene images is identified, localized and recognized with the help of this approach. A text region detector is used such that the text regions present within each layer of image can be detected using histogram of oriented gradients (HOG) within the pre-processing stage. For the segmentation of candidate connected components, local binarization known as segmentation is applied [3]. For the filtering of text and non-text components, it is important to consider parameters such as normalized height width ration and compactness within text extraction stage. Further, the zone centroid and image centroid based distance metric feature extraction system is utilized at final stage for performing text recognition. In order to overcome various issues, the text extraction and recognition can be speed up. Achieving a digitized raster image of the document with the help of appropriate scanning system is the initial step of document analysis. Further, page layout analysis and character recognition are performed after this. In order to identify whether the document is skewed or not, a test is performed previous to the structuring of achieved text. In order to correct the skewed document images, several methods have been proposed lately. Through the projection of an image in XOY space within the Hough space, the initiation of Hough transform algorithms is done [4]. Further, the dominance of lines is calculated. The computation time is maximized with the resolution which is a disadvantage of this approach even though is provides very good precision. In order to identify the predominant straight lines present within a document, the Hough transform is used by this class of methods. Thus, the skew angle of image is provided here. The recognition of parallel straight lines plays a very important role within these techniques. On OX and OY axes the pixels of image are projected by program such that the histograms are generated. Various projections are generated with the help of rotating the source image. The angle at which almost all of the lines are perpendicular to projection axes is the one at which most of peaks can be seen. However, the computational intensity of this method is higher in comparison to Hough transform based methods [5]. The fact that there is text available within majority of documents and the characters and aligned and organized collectively is exploited by these algorithms. Initiating from linking the components, grouping them together and then using spatial relationships the document

skew is calculated and this is known to be a bottom-up process. For determining the skew angle, the coherency that is identified within the alignment of blocks of characters is utilized by nearest-neighbor clustering methods. The connected components are discovered by these algorithms though which the characters are represented and the skew angle is calculated through the recognition of dominant slope that exists within them. The centroids or bottom pixels are considered within these calculations. Application of a series of filters and a clustering algorithm such that the lines can be identified is the simplest implementation of this idea. A local skew angle is created by the bottom pixels of each connected component within the cluster. The median value of all the angles identified is known as skew angle. For skew angles, good results are achieved by this method [6]. Further, grouping the components with the help of neighbor is another similar idea for this approach. A histogram achieves contribution from the angles that are present between each connected component. A homogeneous structure of the textural zones is assumed to be present for the textual zones that are present within a skew free image as per these cross correlation approaches. The skew angle of an input image is estimated with the help of utilizing deviation from this benchmark. Initially, the vertical segments are chosen from an input image, which are further used during profiling process. Further, for each vertical segment, the projection profiles are calculated. The skew angle is estimated by this approach by making comparison against the projection profiles of adjacent segments.

II. LITERATURE REVIEW

Veena Rajan, et.al (2017) proposed an enhanced model or new hybrid method in this paper for the detection of text from natural scene images. This method is based on the fractional poisson and also used Laplacian operation that helps in increasing the image quality. In order to obtain the better contrast between the targets and their backgrounds, it is required to have the actual values of pixels that help sin the operation of the image enhancement [7]. There is increase in the brightness of the image and Laplacian images after performing enhancement operations in it. In order to enclose the letters or characters of the words, they adopted a new method for segmenting the word in which bounding box regions is used. Thereby, to combine the character bounding boxes and to obtain the word bounding boxes, there resulting letter spaces are continuously filled.

Rituraj Soni, et.al (2017) proposed a hybrid approach in this paper for the extraction of above mentioned issues on the basis of text confidence score in which three different attributes are used such as dissimilarity, color dissimilarity and occupy rate [8]. With the help of this method text and non-text constituents are distinguished from each other. Therefore, achieving the fast detection and text regions localization in low resolution and blurred images is the main objective of this paper. They used the Bayesian framework for the calculation of the text confidence score on these constituents in which mentioned three attributes are used. Therefore, they performed experiments for the evaluation of the proposed method.

Wenjun Ding, et.al (2017) presented that detection of text from natural images is the major challenge due to presence of different texts and scene complexity [9]. They proposed a robust text detection method in this paper using which textual components are localized at different levels such as pixel, intra-character and inter-character. They adopted a seed growing mechanism at each level in which well-conditioned seed textual components are detected. After which to localize the degraded components they grow all the seeds. In order to aggregate character candidates into text lines, they proposed a random walk in this paper to restart the algorithm. They performed experiments on the proposed method and concluded the effectiveness of the proposed method as compared to others.

Zhandong Liu, et.al (2017) presented in the application of content-based multimedia, the extraction of the text from natural scene images is considered as the essential requirement. Therefore, they proposed a simple and effective text detection method in this paper. They utilized the V-MSER algorithm for the extraction of MSERs using different channels such as G, H, S, O1, and O2, as component candidates [10]. Therefore, to generate the component pairs they adopted the four simple features. They utilized the two public dataset such as ICDAR 2011 and MSRA-TD500 in this paper for the evaluation of the proposed method that achieves the higher accuracy of 82.94 and 75% F-measure. The obtained results show the effectiveness of the proposed method.

Shuping Liu, et.al (2017) presented a novel method in this paper for the detection of scene images from the text. They utilized the morphological component analysis (MCA) in this paper for the decomposition of the scene images [11]. This helps in reducing the adverse effects of complex backgrounds on the detection results. Hence, there is improvement in the discrimination of dictionary with the help of the proposed method in which Laplacian sparse regularization is introduced. They constructed the text component on the basis of text dictionary and the sparse-representation coefficients of text. They implemented certain heuristic rules for the detection of the query image in the text. On the basis of performed experiments, they concluded the effectiveness of the proposed method as compared to others.

Juili P. Bhirud, et.al (2016) presented there is development in the extraction of text automatically from the natural scene images as it provides various effective results in robotic navigation. This technology has the application in the field of image retrieval, number plate detection, content-based web search, intelligent transport system and for the visually

impaired people. They utilized the stroke-width transform method in this paper using which text is extracted from natural images and textual images [12]. In order to find the stroke-width of every pixel, the image operator (SWT) has been utilized. On the basis of performed experiments, shows the better performance of the SWT method which is reliable and powerful in detecting the text independent of scale, direction, background and font. On the basis of the detection of possible texts, the connected component based approaches are categorized by this method.

III. RESEARCH METHODOLOGY

Following are the steps followed in proposed methodology:

1. Pre-Processing :- In this phase, the input image is taken as input and when the input is RGB it will be converted to the gray scale for the further processing

2. Apply PCA algorithm :- The PCA is the algorithm which is applied to high light the headline of the input natural scene images. Linear Discriminant Analysis (LDA), Independent Component Analysis and PCA are a portion of the techniques utilized for feature extraction, among them PCA is intense method in image formation, Data examples, similarities and differences between them are identified effectively. The other principle advantage of PCA is dimension will be reduced by maintaining a strategic distance from redundant information, without much loss. Better comprehension of principal component analysis is through statistics and a portion of the mathematical techniques which are Eigen esteems, Eigen vectors. PCA is a valuable statistical and common method that has discovered application in fields, for example, image recognition and compression. Principal Component Analysis (PCA) is a mathematical methodology that utilizations linear Transformations to map data from high dimensional space to low dimensional space. The low dimensional space can be controlled by Eigen vectors of the covariance matrix.

The steps involved in PCA include:

• The mean valve S of the given data set "S" is found

• Subtract the mean value say from S. from these valves a new matrix is obtained. Let say "A"

• Covariance is obtained from the matrix i.e., C = AAT Eigen values are obtained from the covariance matrixes that are V1V2V3V4…VN,

• Finally Eigen vectors are calculated for covariance matrix C

• Any vector S or $S - \bar{S}$ or can be written as linear combination of eigen vectors shown in Equation below.

• Because covariance matrix is symmetric it form basis $V_1 V_2 V_3 V_4 \dots V_N$,

$$V_N S - \bar{S} = b_1 u_1 + b_2 u_2 + b_3 u_3 + \cdots + b_N u_N$$

• Only Largest eigen values are kept to form lower dimension data set:

$$\hat{S} - \bar{S} = \sum_{I=0}^{1} b_1 u_1 \,; 1 < N$$

The components in lower dimension space are called principal components which are ensured to be independent just if the data set is mutually typically appropriated. PCA is delicate to the relative scaling of the original variables. Contingent upon the field of application, it is additionally named as discrete Karhunen-Loève Transform (KLT), or the Hotelling transform.

3. Apply Chord-to-Distance Algotithm: The chord-to-distance algorithm is applied which will detect the key points from the image which the output of the PCA algorithm. The PCA algorithm will high light the headline of the input image. Radon change is a mapping of Cartesian directions to a separation and point (s,θ) facilitates. The radon work figures projections of a picture grid along indicated headings. A projection of a two-dimensional capacity f(m,n) is an arrangement of line integrals. The radon work processes the line integrals from different sources along parallel ways, or pillars, in a specific bearing. The shafts are divided one pixel unit separated. To speak to a picture, the radon work takes numerous, parallel-pillar projections of the picture from various edges by turning the source around the focal point of the picture. Where the X-beam bars are going through the question f(m,n). At the point when the pillar go through the question, at that point number of photons are ingest by the protest and that number is anticipated in the polar Cartesian framework (s, θ) at edge θ to the even hub

4. Post Processing: - The chord-to-distance algorithm will return the image in which key points are marked and also distance between the key points. The key point angles are also analyzed correspond to the 180 degree. The distances are extracted whose angles are below 5 degree. The image is rotated on the angle difference between the 180 degree and detected angle. This leads to skew removal in the input image and output image will be more clear and visible.

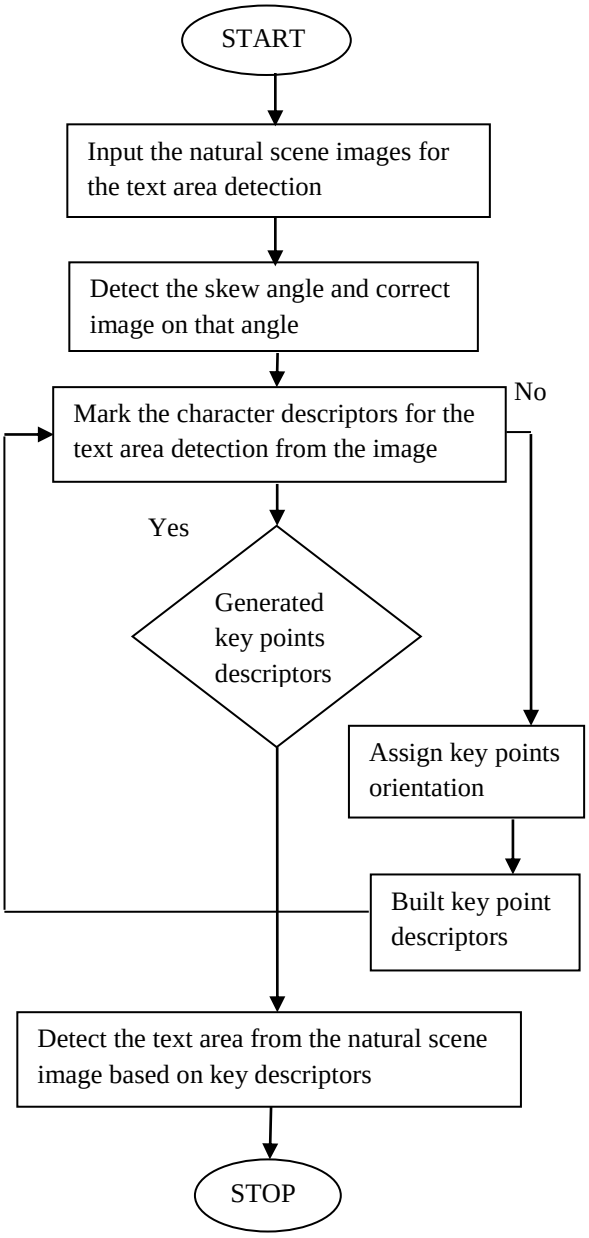

Fig.1: Proposed Flowchart

IV. EXPERIMENTAL RESULTS

The proposed work is implemented in MATLAB and the results are evaluated by making comparisons of proposed and existing techniques in terms of several parameters.

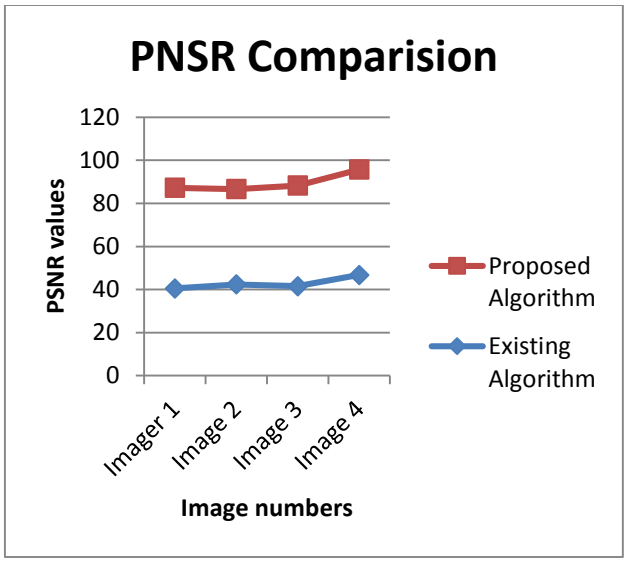

Fig.2: PSNR Comparison

Figure 2 represents the PSNR values of the existing algorithm. The proposed algorithm and the already existing algorithms are compared on the basis of the PSNR and analyzes that the PSNR values of the proposed algorithm are higher than the existing algorithm's values.

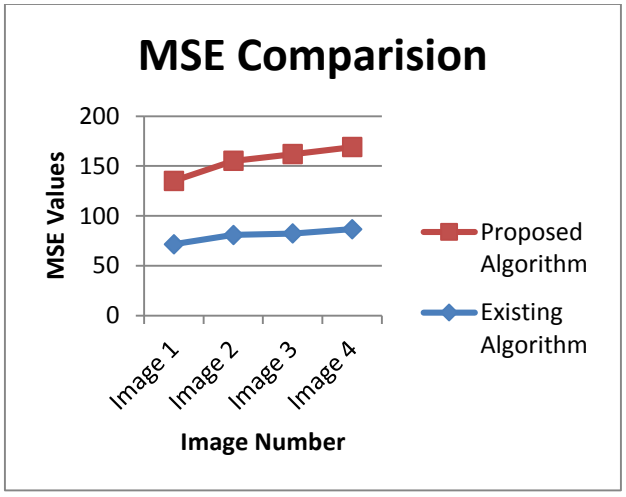

Fig.2: MSE Comparison

Figure 3 shows a comparison between the proposed and the existing algorithm and analyzed that the MSE value of the proposed algorithm is less as compared to the existing ones.

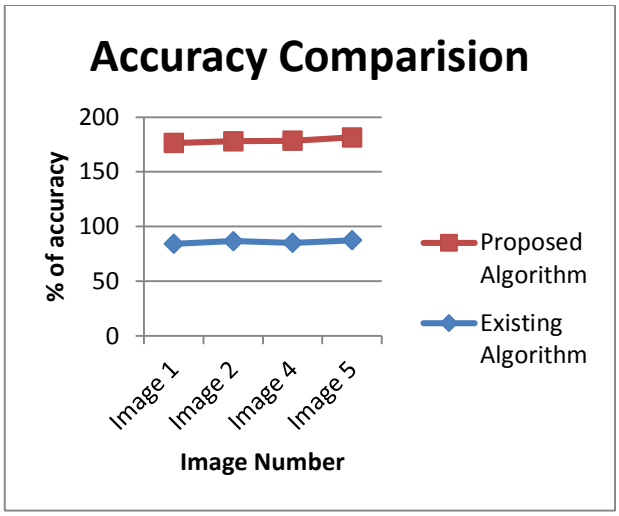

Fig.4: Accuracy Comparison

Figure 4 demonstrates a comparison between the proposed and the existing algorithm on the basis of the accuracy. It has been found that the accuracy of the proposed algorithm is more than the existing algorithm.

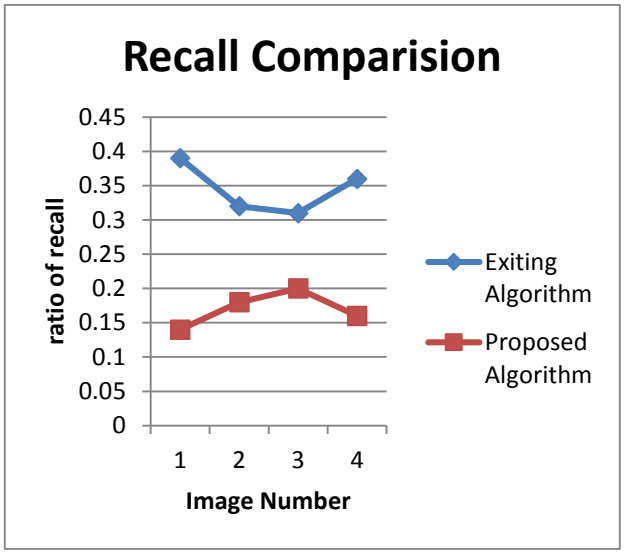

Fig.5: Recall Comparison

Figure 5 compares the performances of the both the existing and the proposed algorithm on the basis of their recall values. It has been found from the figure above that the recall values of the proposed algorithm is less than the existing algorithm.

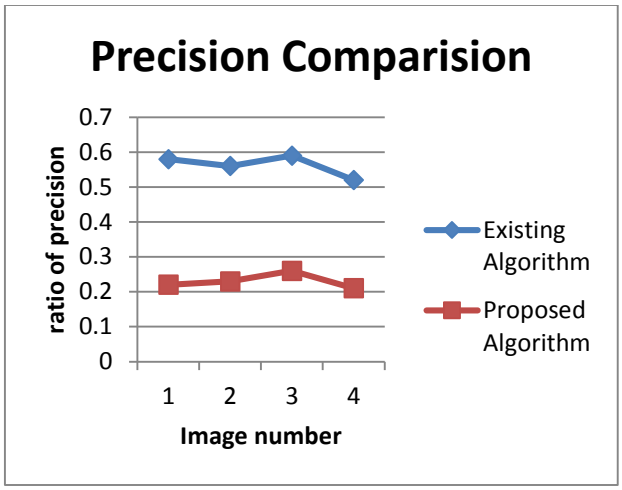

Fig.6: Precision Comparison

Figure 6 shows a comparison between the proposed and the existing algorithm on the basis of their precision values. It has been found that the precision value of the existing algorithm is less than the proposed algorithm.

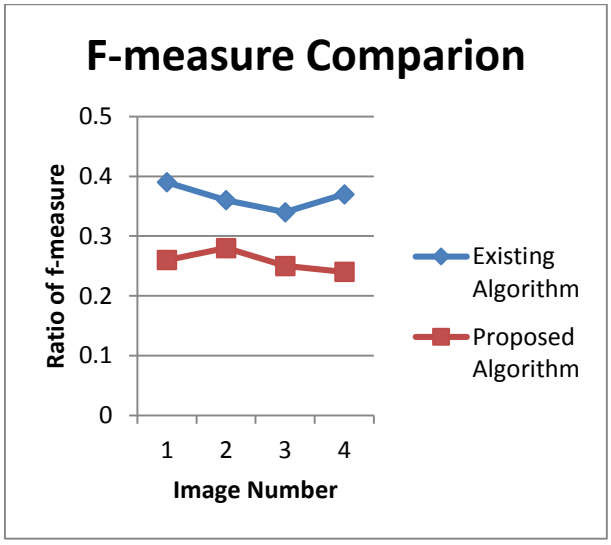

Fig.7: F-measure Comparison

Figure 7 illustrates a comparison between the proposed and the existing algorithm on the basis of the f- measure values. It has been analyzed that the F-measure values of the proposed algorithm is more than the existing algorithms.

V. CONCLUSION

It is concluded that text extraction is the major issue of the image processing. The text extraction is the challenge due to skeness and random color contract of the images. In this research work, the novel approach of skeness removal and feature extraction is proposed. The technique of PCA is applied for the feature extraction and threshold based technique is applied for the skew removal. The proposed algorithm is implemented in MATLAB and results are analyzed in terms of certain parameters.

VI. REFERENCES

[1]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," 2012, in Neural InformationProcessing Systems (NIPS).

[2]. X. C. Yin, X. Yin, K. Huang, and H. W. Hao, "Robust text detection in natural scene images," IEEE Trans. Pattern Analysisand Machine Intelligence (TPAMI), vol. 36, pp. 970–983, 2014

[3]. K. Kim, K. Jung, and J. Kim, "Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm," IEEE Trans. Pattern Analysis and Machine Intelligence (TPAMI), vol. 25, pp. 1631–1639, 2003.

[4]. S. Hanif and L. Prevost, "Text detection and localization in complex scene images using constrained Adaboost algorithm," 2009, in International Conference on Document Analysis and Recognition (ICDAR).

[5]. D. B. Barber, J. D. Redding, T. W. McLain, R. W. Beard, and C. N.Taylor. Vision-based target geo-location using a fixedwing miniature air vehicle. Journal of Intelligent and Robotic Systems, 47(4):361–382,2006.

[6]. Y. K. Ham, M. S. Kang, H. K. Chung, R. H. Park, and G. T. Park. Recognition of raised characters for automatic classification of rubber tires. Optical Engineering, 34(1):102–109, 2005.

[7]. Veena Rajan, Shani Raj, "TEXT DETECTION AND CHARACTER EXTRACTION IN NATURAL SCENE IMAGES USING FRACTIONAL POISSON MODEL", IEEE 2017 International Conference on Computing Methodologies and Communication (ICCMC)

[8]. Rituraj Soni∗, Bijendra Kumar†, Satish Chand, "Text Detection and Localization in Natural Scene Images Using MSER and Fast Guided Filter", 2017 Fourth International Conference on Image Information Processing

[9]. Wenjun Ding, Susu Shan and Feng Su, "TEXT DETECTION IN NATURAL SCENE IMAGES BY HIERARCHICAL LOCALIZATION AND GROWING OF TEXTUAL COMPONENTS", Proceedings of the IEEE International Conference on Multimedia and Expo (ICME) 2017

[10]. Zhandong Liu1,2, Yong Li1 , Xiangwei Qi1 , Yong Yang1 , Mei Nian1 , Haijun Zhang1,3 , Reziwanguli Xiamixiding, "Method for unconstrained text detection in natural scene image", IET Comput. Vis., 2017, Vol. 11 Iss. 7, pp. 596-604

[11]. Shuping Liu, Yantuan Xian, Huafeng Li, and Zhengtao Yu, "Text Detection in Natural Scene Images Using Morphological Component Analysis and Laplacian Dictionary", 2017

[12]. Juili P. Bhirud, Priti P. Rege, "A Modified SWT based Text-image Separation in Natural Scene Images", 2016 Conference on Advances in Signal Processing (CASP) Cummins College of Engineering for Women, Pune. Jun 9-11, 2016