**EXTENDED ABSTRACT**

# TEXT IDENTIFICATION TECHNIQUE TO RECOGNIZE NAME BOARDS WITH PARTIALLY DISFIGURED CHARACTERS

Lathagini. S* and Suthaharan. S

University of Vavuniya, Sri Lanka

* lathagini123@gmail.com

(Published 15 October 2021)

**Abstract**

Although text recognition gives rise to many applications, this research aims to identify, locate, and recognize characters in name boards using various font styles and partial letters. English language displayed in name boards is used to predict partial and whole characters in this work. First, the image is preprocessed with grayscale transformation and thresholding, followed by morphological transformations and connected component analysis to localize the text sections. The text is then segmented using the Skeleton analysis method in the next phase. Further, the segmented character image is turned into strings using the Coefficient of Correlation and Structural Similarity Index methods. As a result, characters with disfigurements can be identified using the most similar character image with an accuracy of 81.7%.

**Keywords:** Grayscale transformation, morphological transformations, skeleton analysis, coefficient of correlation, structural similarity index methods

## 1. Introduction

The name boards of places are the most frequent visual aids in roadways. Their use is assisting people who are unfamiliar with the area. Sometimes name boards are disfigured by damages. In case of name boards are partially disfigured, it is complicated to identify by an unfamiliar person. This research aims to find a suitable technique to detect, localize, and recognize the characters in the partially disfigured name boards. Most research works in literature have focused on text recognition in handwritten, printed documents and scene images(Chen et al., 2004),(Epshtein et al., 2010) with clear letters. (Zhang et al., 2013) and (Zhu et al., 2016) presented the survey on scene text localization and recognition. Most of researcher's focus is on different fonts, colors, scales, and orientations in the same scene.

In (Ohya et al., 1994) ,Jun ohya proposed novel to recognize character in scene images. Main steps of this research is image segmentation using Local thresholding, detecting character candidate regions, selecting high similarity patterns for multi font printed Chinese characters by using Mesh feature algorithm and determining and recognizing character patterns by a relaxational operation .In (Kaur et al., 2012), the researchers proposed a novel approach for image recognition of flowers using coefficient of correlation (CoC) and Structural Similarity Index Method(SSIM) in real time environment. The results were obtained that recognition performance of CoC and SSIM gets improved with rotations and discrete wavelet transform. In (Neumann et al., 2013), the authors proposed a method with

Oriented Stroke Detection to localize and recognize text in natural scene images using sliding-window and connected components. They presented a block structure containing Stock detection, Candidate Region detection, Character recognition, word formation, and Word NMS in the literature. The authors proposed a font comparison system by using Multiple Similarity metrics in (Takizawa et al., 2020). They had considered Mean Squared Error ( MSE), Peak Signal to noise ratio(PSNR), Structural Similarity(SSIM), Haar wavelet-based perceptual similarity index(Haar PSI) for assessing image quality, as well as Euclidean distance and cosine similarity in t-distributed stochastic neighbor embedding(t-SNE) to reduce the number of dimensions. They had also ranked the results obtained from five types of fonts. SSIM is the closest among the eight types of methods that they used. This research study has been carried out to consider partially disfigured name boards images as input, segmenting those images, computing the correlation between template and input image. Further, it detects the pattern, identifying the characters, and finally turns into strings.

## 2. Methodology

### *2.1 The proposed system*

In this section, an automated text identification of name boards that have partially disfigured characters shown in Figure 1. It contains different stages, including acquiring name board images with disfigured characters, preprocessing, localization of text regions, Text line segmentation, Character segmentation, Character recognition.

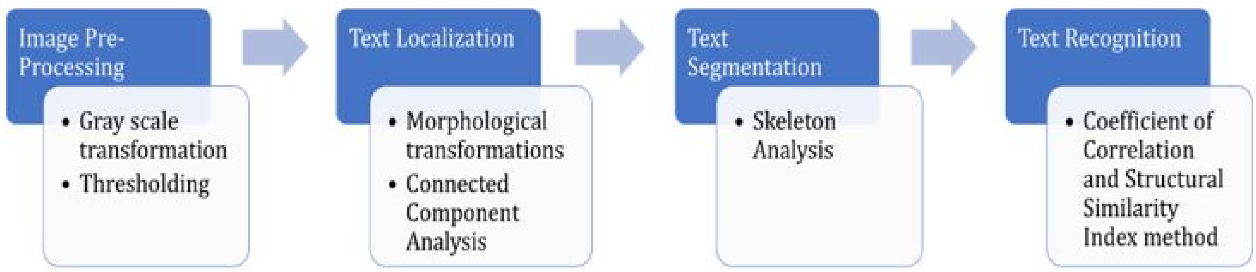


**Figure 1.** Proposed system

First, the image is preprocessed using grayscale transformation and thresholding. In text locaization stage, the binary morphological operations Dilation and Opening are applied to the binary image to connect the characters in a word together and remove noises in the background. The connected component analysis scans an already binarized image and labels its pixels into components based on pixel connectivity. After groups of pixels have been identified, each pixel is labelled with a value based on the component. In the next phase, the Skeleton analysis method is used to segment the text lines. Then segmented character image is converted into strings by using the Coefficient of Correlation and Structural Similarity Index method.

### *2.2 Image preprocessing*

Images acquired from camera or mobile are color images and not suitable for image processing steps of character recognition described below as it takes time to process image processing algorithms. Therefore it has to be performed preprocessing. As this work focuses on predicting characters, grayscale transformation is used to avoid additional color information. As it reduces the amount of data that have to be processed in the image, it helps to reduce computation and also speeds up the processing.

As a result of the grayscale transformation is challenging to extract areas that correspond to significant structures in an image and analyze text areas. Thresholding is used to partition an image


**Figure 2.** Gray scale image

into regions that separate characters from the background. By changing image bit depth to 1 bit consisting of only black and white, recognizing the partially disfigured characters improves results. Thresholding can be considered an essential step in most image processing applications.

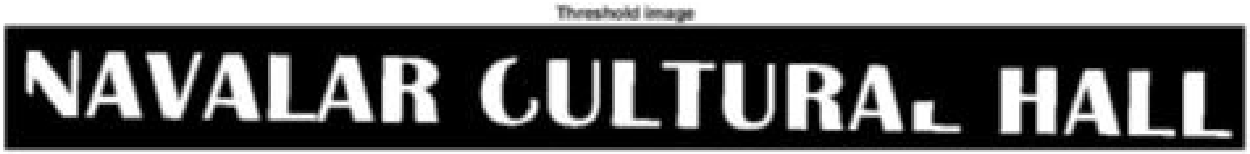


**Figure 3.** Threshold image

### *2.3 Text Localization*

The objective of this method is to localize text regions as well as group them as a word. In this stage, the binary morphological operations Dilation is applied to binary image to connect the characters in a word together, and then Opening is used to remove the small pixel in the background. The

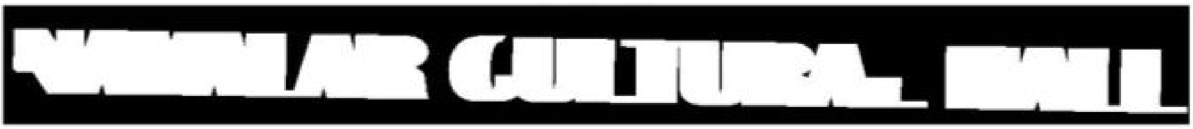

**Figure 4.** Result image of morphological operations.

connected components can be extracted using morphological operations. The connected component analysis scans the binarized image and labels its pixels into components based on pixel connectivity to separate the words. It is used to break the combination of words as single and display it in a new line. Each character is localized and marked with a box for segmentation. The process of extracting and labelling various disjoint and connected components of characters in the name board image makes the recognition result easier and more accurate.

### *2.4 Text segmentation*

Text line segmentation and character segmentation are used to obtain precisely bounded characters. The skeleton analysis (Shivakumara et al., 2010) is proposed text line segmentation method. Text skeletons are extracted from connected components, and a text line is defined as a continuous path on the skeleton from an intersection point to either an endpoint or an intersection point. The minimum and maximum index of the picture that contains a text pixel vertically and horizontally is to be found. Then the cropped image is obtained by the calculated index. Then extract the whole characters line by line. Each row is performed summation to find the first one to be zero. The


**Figure 5.** Word segmentation and Localization of Character

index of the row has to be known to trim. The first line of texts is separated from the remaining if there are several lines. By repeating this procedure, all the lines that contain texts can be separated successfully. Further, each character is trimmed in each line. Function 'bwlabel' is used to achieve this.

The function 'L = bwlabel(BW)' returns the label matrix L that contains labels for the 8-connected objects found in BW. The label matrix, L, is the same size as BW. Finally, every signal character is extracted from each line.


**Figure 6.** Line segmentation and Character segmentation.

## *2.5 Text Recognition*

In this phase first step is to allow the system to learn the patterns and features from the known character and then making predictions accordingly. The basic step is to create character images (bitmap image format), including uppercase and lower case letters (A-Z, a-z). These images are cropped according to the image size of 42 * 24 pixels.

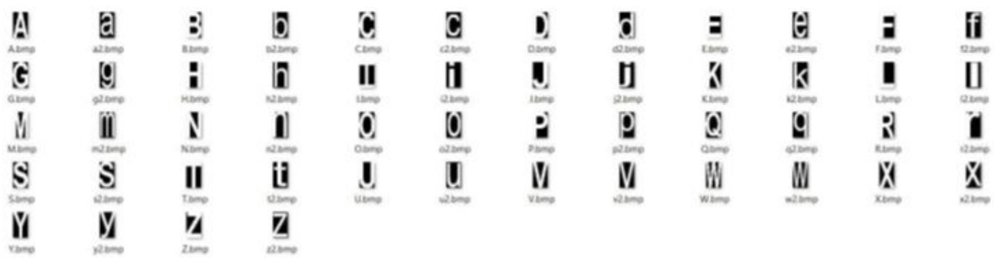

**Figure 7.** Dataset for English Upper case and lower case Characters

Image Recognition using Coefficient of Correlation and Structural Similarity Index (Kaur et al., 2012) method is used to recognize the characters. The goal is to compare the image of each trimmed character with template characters. Due to the fact that each image of characters is made up of numerical pixels, it is just needed to compare pixel by pixel to find the level of similarity. Correlation coefficients are calculated between the unknown character and a set of template characters in the two-dimensional array. The formula in Figure 8 is used to find the value of correlation. Here, A

$$r = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N} (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_{m=1}^{M} \sum_{\sum n=1}^{N} (A_{mn} - \bar{A})^{2})(\sum_{m=1}^{M} \sum_{n=1}^{N} (B_{mn} - \bar{B})^{2})}}$$

**Figure 8.** value of correlation

represents the template image, and B is a trimmed character image. m represents the row of the matrix of images, and n represents the column of the matrix of images. In this particular case, M = 42 and N = 24.Then the maximum value of correlation is found within all the comparisons with template characters. Furthermore, the corresponding uppercase letter or lowercase letter of the max correlation in template data is the recognized symbol for each character. The template which has a higher value means the better match.

## 3. Results and Discussions

Due to the lack of a dataset of name board images with partially visible characters, images of name boards are collected in Jaffna for data. In this study, the basic case template was produced with limited font styles. This research is limited to predicting partial and full visible English characters in name boards. The performance of this proposed method is compared with Automatic text detection and classification in natural images (Baran et al., 2018). Results for the test case images, this algorithm detected text completely with an accuracy of 56%. The proposed method turns name board images with partial characters into strings recorded in a text file to achieve outcomes with an accuracy of 81.7%.

**Table 1.** Performance of the proposed method for text recognition in the partially disfigured name boards

| | Total no of characters | Recognized characters by the proposed method | Recognized characters by the existing method | Accuracy of proposed method | Accuracy of existing method |
|---|---|---|---|---|---|
| Name board 1 | 17 | 15 | 10 | 88.2% | 58.8% |
| Name board 2 | 14 | 12 | 5 | 85.7% | 35% |
| Name board 3 | 10 | 7 | 3 | 70% | 30% |
| Name board 4 | 19 | 15 | 12 | 78.9% | 63.15% |
| Name board 5 | 39 | 30 | 21 | 76.9% | 53.84% |
| Name board 6 | 11 | 9 | 7 | 81.8% | 63.63% |
| Name board 7 | 15 | 13 | 12 | 86.6% | 80% |
| Name board 8 | 25 | 22 | 16 | 88% | 64% |
| Total | 125 | 101 | 85 | 81.7% | 56% |

## 4. Conclusion and future work

Even though some name boards have partially disfigured characters, partial English character prediction is achievable. The findings indicate that this technique performs slightly better in identifying explicit and partially disfigured characters on name boards. The templates achieved an accuracy of 81.7% since they were trained using a limited font style. In order to improve the accuracy rate of character recognition using Machine Learning techniques which can be evaluated with more character templates in future. In Sri Lanka, name boards are usually designed in three languages. This study could be expanded to recognize disfigured text in other languages like Tamil and Sinhala also.

## References

Baran, Remigiusz, Pavol Partila, Wilk, and Rafal. 2018. Automated text detection and character recognition in natural scenesbased on local image features and contour processing techniques. In *International conference on intelligent human systemsintegration* 42–48. Springer.

Chen, Xiangrong, Yuille, and Alan L. 2004. Detecting and reading text in natural scenes. In *Proceedings of the 2004 ieee computer society conference on computer vision and pattern recognition, 2004. cvpr 2004.* 2:II–II. IEEE

Epshtein, Boris, Eyal Ofek, Wexler, and Yonatan. 2010. Detecting text in natural scenes with stroke width transform. In *2010 ieee computer society conference on computer vision and pattern recognition* 2963–2970. IEEE.

Kaur, Avneet, Lakhwinder Kaur, Gupta, and Savita. 2012. Image recognition using coefficient of correlation and structural similarity index in uncontrolled environment. *International Journal of Computer Applications* 59(5).

Neumann, Lukas, Matas, and Jiri. 2013. Scene text localization and recognition with oriented stroke detection. In *Proceedingsof the ieee international conference on computer vision* 97–104.

Ohya, Jun, Akio Shio, Akamatsu, and Shigeru. 1994. Recognizing characters in scene images. *IEEE transactions on patternanalysis and machine intelligence* 16(2): 214–220.

Shivakumara, Palaiahnakote, Trung Quy Phan, Tan, and Chew Lim. 2010. A laplacian approach to multi-oriented text detection in video. *IEEE transactions on pattern analysis and machine intelligence* 33(2): 412–419.

Takizawa, Sho, Taisei HOSHI, CHEN, and Qiu. 2020. Font comparison system based on multiple similarity metrics. *International Journal of Affective Engineering*, IJAE–D.

Zhang, Honggang, Kaili Zhao, Yi-Zhe Song, Guo, and Jun. 2013. Text extraction from natural scene image: a survey. *Neurocomputing* 122: 310–323.

Zhu, Yingying, Cong Yao, Bai, and Xiang. 2016. Scene text detection and recognition: recent advances and future trends. *Frontiers of Computer Science* 10(1): 19–36.