**Available Online at www.ijcsmc.com**

**International Journal of Computer Science and Mobile Computing**

**A Monthly Journal of Computer Science and Information Technology**

**ISSN 2320–088X**
**IMPACT FACTOR: 7.056**

***IJCSMC, Vol. 9, Issue. 12, December 2020, pg.95 – 102***

# Message (Text) Extraction System from Noisy-Image Regions Using Morphological Dialation Technique in Forensics Investigation

**[1]Dr. Adeyemi B.M; [1]Dr. Mosud Y.O; [2]Prof. Odesanya B.O; [3]Dr. Arigbede A.O; [4]Sunmola F.O**

[1]Department of Computer Science and Cyber Security, Caleb University, Lagos, Nigeria
[2]Faculty of Pure and Applied Science, Southwestern University, Okun Owa, Nigeria
[3]Department of Guidance Psychology, Tai-Solarin University of Education, Ijebu Ode, Nigeria
[4]Department of Information Technology, Federal University of Technology, Akure, Nigeria
Corresponding Authors: adejid2000@yahoo.com; myolumoye@yahoo.com; dipoarigbedelasu@yahoo.com; tunde.odes@gmail.com; fosunmola@futa.edu.ng
Tel : +234-1-8089908927

**DOI: 10.47760/ijcsmc.2020.v09i12.013**

***ABSTRACT: In cyber/forensics investigation, it was observed that there is a need to extract text/messages from scene images especially from a noisy-image zones/regions, hence the attention was drawn to study text detection and recognition in natural images using computer vision systems. Text detection and recognition in natural images has applications in computer vision systems like image retrieval, registration plate number detection, automatic traffic sign detection, could be of help for visually impaired people. Scene text, however, has complicated background, blur image, partly occluded text, variations in font-styles, image noise and ranging illumination. Hence scene text recognition could be a difficult computer vision problem. In the previous works, paper connected component method is used to extract the text from background. In this work, horizontal and vertical projection profiles, geometric properties of text, image binirization and gap filling method has been used to extract the text from scene images while the morphological dilation method was introduce to separate text background of the images. precision rate and recall rate of the proposed method and that of the other existing methods were compared and discovered that the performance of our proposed method is excellent overall according to the result of the metrics variables. Therefore, the proposed method is proved to be efficient for extracting the text/messages from the outdoor scene images. Finally text is extracted from images for optimal result.***

***Keywords: Horizontal and vertical projection, Morphological dilation, Geometric properties of text, Image binirization and Gap filling method.***

## I. INTRODUCTION

In the recent years, the world has seen the image and video capturing technologies grow exponentially. Few examples include digital cameras, mobile phones, etc. With the improvisation in image capturing devices and their ease of use, there is being captured and stored, which may contain useful data. The challenge that remains in front of us is that of processing these images to bring out the required detail. One appealing and developing field comprises of extracting the text from images. Digital image processing is used for extracting the text from image. “Image processing is the study of algorithm that takes image as input and returns an image as output”. Images with text have various complexities that arise due to real world backgrounds, font sizes, and text positioning and so on. In some cases, low quality of capturing devices adds another level of difficulty.

Existing methodologies for text extraction include region based, connected component based, texture based. Region based method used gray scale properties to differentiate text from background, whereas texture based method uses textual properties.

There are many hurdles or problems in detection of text, extraction of text and localization of text from images. In the proposed work, in order to overcome all these hurdles and problems that are occur when extracting text from images, texture - based method used in this proposed scheme. In this, image is converted into black and white image and used the filter to remove noises and used morphological operations for the feature extraction from image and for extracting the text from image discrete wavelet transform is used.

The implementation of the following work would prove useful in solving real world problems such as license plate capturing of speeding vehicles, translation of sign/board, capturing city maps, to capture directive of the routes, to capture public notice and to capture advertisement banners etc. All these images contain useful and important information. And this information present in the images is used in the variety of applications.

## II. RELATED WORK

Existing text - detection methods can be divided into region based and texture based methods. Region based methods rely on image segmentation. Pixels are grouped to CCs which are character candidates. These candidates are further grouped to candidate words and text-lines based on geometric features. Texture based methods distinguish text from non-text based on local features and machine learning techniques.

Chen et al. [4] propose a text detection method using MSER. The outlines of MSER are improved by edges detection techniques such as canny edge detection. This makes MSER less responsive to blur images. Based on geometric cues these candidate character regions are then grouped to words and textlines.

Neumann et al.[3] propose ERs for segmenting regions. ERs are extracted on the gradient images, HSI and RGB to recover regions for character candidate. As an alternative of using heuristics as Epshtein et al. [3] for labeling text, an AdaBoost classifier based on geometric features is used. Text- CCs are then grouped to words.

Kim et al. [1] combined a Support Vector Machine (SVM) and continuously adaptive mean shift algorithm (CAMSHIFT) to detect and identify text regions. Gao et al. [9]developed a three layer hierarchical adaptive text detection algorithm for natural scenes. This method was applied in a prototype Chinese sign translation system which mostly has a horizontal and/or vertical alignment.

In Zheng et al.[7] proposed a completely unique image operator is projected to observe and find text in scene images to attain a high recall of character detection, extremely regions are detected as character candidates. 2 classifiers are trained to spot characters, and a algorithmic native search algorithm is projected to extract characters that are incorrectly known by the classifiers.

An efficient pruning technique, which mixes component trees and recognition results, is projected to prune continuation elements. A cascaded technique combines text line entropy with a Convolutional Neural Network model. It's wont to verify text candidates that reduce the quantity of non-text regions. The projected technique is taking a look at on 3 public datasets, i.e. ICDAR2011 dataset, ICDAR2013 dataset and ICDAR2015 dataset.

Wang et al. [6] propose HOG features with a Random Ferns classifier to detect and classify text in an end to end setting. Multiclass detector is trained on letters. Non-maxima of the detector results are concealed. The remaining letters are then combined in a Pictorial Structure framework, where letters are parts of words. For each word in a dictionary, the most plausible character responses are found in the image. Detected words are then rescored based on geometric information and non-maxima suppression is done to remove overlapping word-responses.

In Greenhalgh et al.[2] proposed a unique system for the automated detection and recognition of text in traffic signs. Scene structure is employed to describe search regions at intervals the image, surrounded by traffic sign candidates are then found. Maximally Stable Extremely Regions (MSERs) and hue, saturation, and worth color thresholding are used to locate a large range of candidates, that are then reduced by applying constraints supported temporal and structural data.

A recognition stage interprets the text contained at intervals detected candidate regions. Individual text characters are detected as MSERs and are classified into lines, before being interpreted using optical character cognition(OCR).

 

In Raj et al.[5] proposed connected component approach for extracting devanagiri text. But extraction accuracy is immensely improved through the temporal fusion of text results across consecutive frames.

## III. PROPOSED METHOD

The common OCR systems available require the input image to be such that the characters can be easily parsed and recognized. The text and background should be monochrome and background-to-text contrast should be high. The proposed system briefly shows in Fig.1.

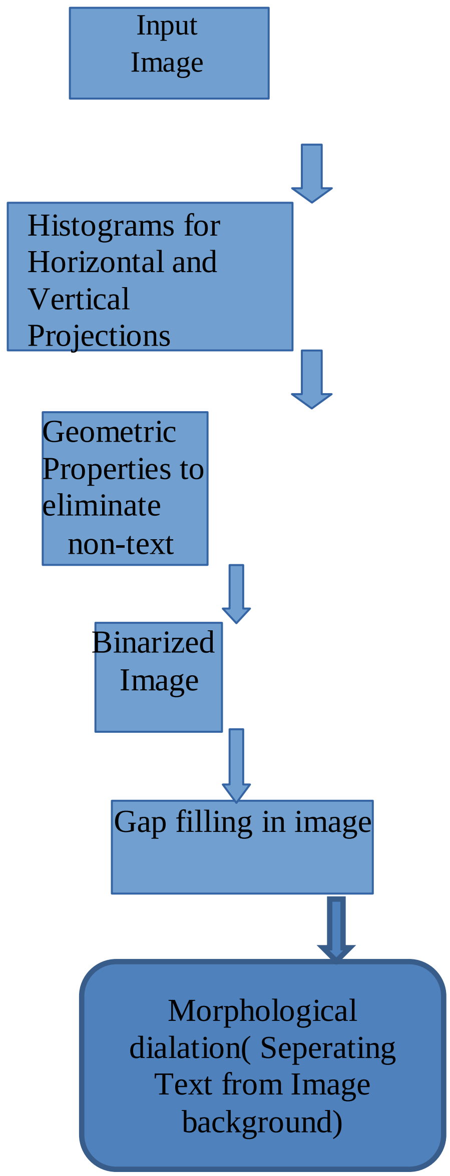


**Fig.1: Flowchart of proposed system**

The steps of the connected-component text extraction algorithm are given below:

1. The horizontal and vertical projection profiles of candidate text regions using a histogram with an appropriate threshold value.
2. Use geometric properties of text such as width to height ratio of characters to eliminate possible non-text regions.
3. Binarize the edge image enhancing only the text regions against a plain black background.

4. The gap-filling process is used to fill the holes where the text are been extracted from the image
5. Morphological dilation is used to separate the text background of the image

### 1. Horizontal and vertical projection using histogram

Projection profile of an image in a particular direction refers to the running sum of the pixels in that direction. In context of text processing, horizontal projection profile is needed to identify or separate out the lines of a text since, the profile exhibits valley points at line boundaries and the location of these minima points mark the line boundaries.

Similarly, vertical projection profile is used to perform word segmentation as the valleys are created corresponding to word gaps. These word boundaries can be identified with the help of these minima points. Threshold selection is made by using the average value of maximum and minimum value of the corresponding projection is taken.

### 2. Eliminate non text regions

To improve the performance of the system, non text regions are eliminated using some rules according with text properties.

The rules are,

1. Text are limited in size.
2. Text must contain edges.
3. Text have special texture properties

### 3. Enhancing non text regions

Binirize the edge image by using Otsu's threshold[4] value.

### 4. Gap filling process

Gap filling process is made by morphological dialation method. The value of the output pixel is the *maximum* value of all pixels in the neighborhood. In a binary image, a pixel is set to 1if any of the neighboring pixels have the value. Morphological dilation makes objects more visible and fills in small holes in objects.

## IV. RESULT AND DISCUSSION

In order to evaluate the performance of proposed method road signal scene text images are used. Here, output of text extraction results when run on are shown below.


**Fig 2.Original image**

**Table 1: MORPHOLOGICAL DIALATION METHOD WITH PRECISION AND CALL RATE**

**PERFORMANCE EVALUATION OF TEXT EXTRACTION USING MORPHOLOGICAL DIALATION METHOD WITH PRECISION AND CALL RATE**

| | | **PRECISION** | | **RECALL** | **REMARKS** |
|---|---|---|---|---|---|
| | | **RATE(%)** | | **RATE(%)** | |
| | | | | | |
| | | | | | |
| **METHOD 1** | | **63.7** | | **82.8** | **Previously** |
| **METHOD 2** | | **89.8** | | **92.1** | **Need to Improve** |
| **PROPOSED METHOD** | | **92.6** | | **94.3** | **Improved** |

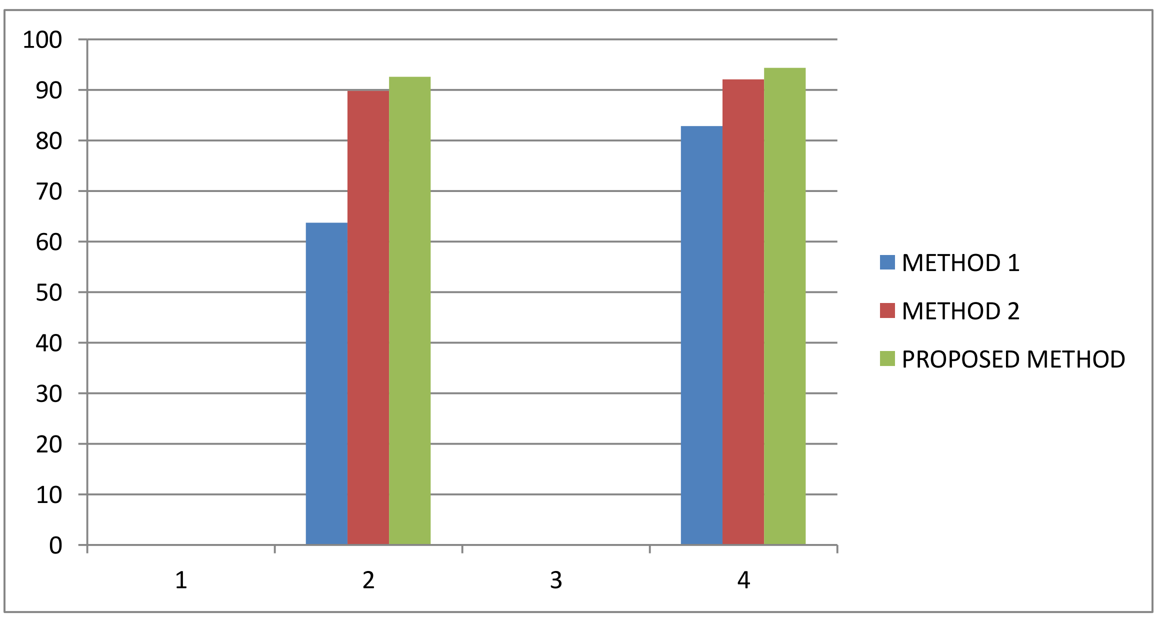


**FIG 3**:Performance Chart

**KEY NOTE: Existing Methods**

METHOD 1 = Kim et al.[1]

METHOD 2 = Wang et al.[6]

PROPOSED Method = Morphological Dialation

Metric variables = Precision and Recall Rate

 


**Fig 4: Output of text extraction**

Precision and recall rate are used as performance measures and the proposed system compared with Wang et al[6] and Kim et al.[1].

Table 2 : Performance results of text extraction

| Method / Measure | Precision rate (%) | Recall rate (%) |
|---|---|---|
| Proposed system | 92.6 | 94.3 |
| Wang et al[6] | 89.8 | 92.1 |
| Kim et al.[1] | 63.7 | 82.8 |

In Table 1,shows ***precision rate and recall rate*** of the proposed method and that of the other existing methods. The performance of our proposed method is excellent overall according to the result of the metrics variables. Therefore the proposed method is proved to be efficient for extracting the text/messages from the outdoor scene images.

## V. CONCLUSION

In this paper, proposed method which is presented very simple and effective algorithm for text extraction( Mophological Dialation). The aim proposed text extraction makes the application easy to deploy and reduce the computation complexity. The aim is satisfied by the proposed work. According to the experimental results, the proposed method is proved to be effective and efficient for extracting the text regions from the complex scene images. There may be so many ideas for future work. But further works have to improve the efficiency of the system. So our intension is to explore the efficacy direction in future.

 

# REFERENCES

**[1].** K.C.Kim, H.R.Byun, Y.J.Song, Y.M. Choi, S.Y.Chi, K.K. Kim and Y.K.Chung, "Scene Text Extraction in natural scene images using hierarchical feature combining and verification", Pattern Recognition, 2004, Aug 2004, vol.2 of ICPR 2004. Proceedings of the 17th International Conference on, pp. 679 – 682.

**[2].** J. Greenhalgh and M. Mirmehdi, "Real-time detection and recognition of road traffic signs," IEEE Transactions on Intelligent Transportation Systems, pp. 1498-1506, Dec. 2012.

**[3].** Neumann et al., "Realtime Scene Text Localization and Recognition", IEEE journal ,2012.

**[4].** [4] Otsu, N., A threshold selection method from gray-level histograms. Automatica, 1975.11(285-296):pp.23-27

**[5].** H. Raj, R. Ghosh, Devanagari Text Extraction from Natural Scene Images, 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE, 2014, pp.513-517.

**[6].** Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (2004) 600–612.

**[7].** J. Zhang, R. Kasturi, A novel text detection system based on character and linkenergies, IEEE Trans. Image Process. 23 (2014) 4187–4198.

 