Camera Captured based Myanmar Character Recognition Using Dynamic Blocking and Chain Code Normalization

Kyi Pyar Zaw[1st], Zin Mar Kyu[2nd]

University of Computer Studies, Mandalay
Mandalay, Myanmar
kyipyarzaw08@gmail.com

DOI: 10.29322/IJSRP.8.8.2018.p8033
http://dx.doi.org/10.29322/IJSRP.8.8.2018.p8033

Abstract- This paper presents a system for Myanmar text extraction and recognition from warning signboard images taken by a mobile phone camera. Camera captured natural images have numerous difficulties compared to the traditional scanned documents. Common problems for camera captured text extraction are variations in font style, size, color orientation, illumination condition as well as the complex background. In this system, color enhancement process is performed to distinguish the foreground text and background color. Color enhanced images are converted into binary images using color threshold range. The detected non-text objects are removed as clearly as possible using width, high, aspect ratio and object region area threshold. In the segmentation process, horizontal projection profile, vertical projection profile and bounding box are used for line segmentation and character segmentation. To recognize the above segmented Myanmar characters, blocking based pixel count and eight-direction chain codes features are proposed. In this system, characters are classified by feature based approach of template matching method by using the proposed features. In this paper, dynamic blocking based pixel count, eight-direction chain codes features and geographic features are used to correctly recognize Myanmar characters.

Index Terms- dynamic blocking, chain codes, Myanmar signboard, text extraction.

I. INTRODUCTION

Character recognition is an important process in the understanding of warning text images by a foreigner or visually impaired person. In the optical character recognition (OCR) system, camera captured images have numerous difficulties compared to the traditional scanned documents. Common problems for camera captured text images are variations in font style, size, color, orientation, illumination condition as well as the complex background. This research presents the Myanmar text recognition system for warning signboard images which are captured by a camera. In this system, color enhancement process is performed to distinguish the foreground text and background color. Color enhanced images are converted into binary images using color threshold range. The detected non-text objects are removed as clearly as possible using width, high, aspect ratio and object region area threshold. Recently, many researchers presented the text extraction and recognition systems for the various languages such as English [2], Chinese, Japanese, Korea, Kannada, Arabic, etc. Myanmar text extraction and recognition in real signboard images remain surprisingly challenging in Computer Vision. The text on the signboard is written in various styles, various fonts, various size and various colors. In our country, warning text papers or signboards may be anywhere and are printed or handwritten. To understand that warning text on the signboard by a foreigner or impaired person it is need to recognize and translate that warning text on the signboard. Myanmar script is considered as a complex script by software developers, as it originated from Indic scripts like Thai or Khmar. And then, Myanmar characters can be divided into two types: basic characters and extended character. The basic characters (consonants) may stand as single character or may be combined with one or more extended characters. The extended characters may be at left or right or top or bottom of the basic character. In Myanmar script language, there exist isolated characters and compound words but, there are no spaces and words as in English language. The remaining Related Works are presented in Section II. Dataset and Nature of Myanmar Script is presented in Section III. Propose Methods is presented in Section IV, Experimental Results are described in Section V. Error Analysis is described in Section VI and Section VII concludes System of the research.

II. RELATED WORKS

Most of the algorithms for detecting text in the literature can be divided into methods based on space-based and connectivity (CC). The regional approach uses a sliding window scheme, which is basically a powerful method. The second way is to localize individual symbols using local parameters of the image (intensity, area, color, gradient, etc.). The selection of a function also plays a decisive role in the image positioning process. The main goal of object extraction is to maximize recognition speed. Many researchers have made research related to this but no technique is almost perfect and they found need to improve the work in more areas at different instants and techniques. Most of the work is found on Machine Printed document Images and handwritten document images of Myanmar script without background noise.

In 2005, T.Swe and P.Tin proposed a Myanmar printed character recognition and translation system using a hoped field neural network. They displayed experimental result using a standard application form. They achieved 97.56% on printed character recognition rate and 94.61% on word recognition rate [9]. In 2008, E.E.Phyu and et.al proposed online handwritten Myanmar compound words recognition system based on Myanmar Intelligent Character Recognition (MICR). They achieved 95.45% and 93.81% recognition rate for typeface and handwritten compound words respectively [2]. In 2010, Y. Thein and S.S. S. Yee contributed an effective Myanmar Handwritten Characters Recognition System using MICR and back propagation neural network. This system only takes 3 seconds average processing time for 1000 word samples and 93% recognition rate for 1000 samples of noise free image [10]. In 2013, Al-Hashemi and Alsharari developed a new system that translates Arabic texts of the signboards into English text by using mobile phone camera. The experimental results of character recognition rate are 81.82% and translation rate is 83.33% [1]. In 2014, Angadi and Kodabagi presented a robust segmentation method for line, word and character extraction from Kannada Text in low resolution display board images. They used projection profile features and pixel distribution statistics for segmentation of text lines. They also used k-means clustering to group inter-character gaps into character and word cluster spaces. This method achieved text line segmentation accuracy of 97.17%, word segmentation accuracy of 97.54% and character extraction accuracy of 99.09% [7].

In 2011, H.P.P.Win proposed a Bilingual OCR System for both Myanmar and English script using multiclass- Support Vector Machine (SVM). They used connected component segmentation method, 25 features of zoning, 60 features of horizontal and vertical profiles methods. This paper achieved 98.89% segmentation rate for 6 Myanmar printed documents [4]. Emmanuel and George, 2013, described a method for the detection and recognition of the Mylayalam text from color natural scene image taken by a mobile phone camera. They used edge profile based method for text detection, bounding box method for text segmentation and template matching method for character recognition. Their algorithm overcomes the problem of challenge like complex background, different font styles, size, and orientations [3]. In 2014, T. Tint and N. Aye proposed Myanmar Text Area Detection and Localization from Video Scenes using connected component labeling approach and geometric properties such as aspect ratio. They used Gaussian filter for eliminating noise from video scenes [12]. In 2016, M. Sayed and S. A. Angadi presented a mobile application for capturing display boards having Kannada text and further giving its English meaning. The character recognition accuracy of this paper is 89.94% on 13 images [6]. In 2017, Ahmed, Saad Bin, et al presented Arabic scene text recognition using Convolutional Neural Networks (ConvNets) as a deep learning classifier. They evaluated on the dataset of 2700 characters of 27 classes and reported the encouraging results on recognition of Arabic characters from segmented Arabic scene images [8]. In 2017, K. P. Zaw and Z. M. Kyu proposed a method of segmentation for Myanmar character recognition using blocked based pixel count and aspect ratio. The segmented characters are classified using feature matching method. In this system, only printed character from the text images can be segmented and recognized by training 98 Myanmar typed-face characters [5]. In 2015, H. S. Mohana created a Template of Kannada characters and each image in a template is in size of 24x42 dimensions. And also they captured Different style, size Kannada Printed and Handwritten images using ordinary mobile camera of 5Mega resolution. This system achieved the recognition accuracy of 92% on handwritten characters and 95% on printed characters [13].

III. DATASET AND NATURE OF MYANMAR SCRIPT

In the proposed system, training dataset of 343 Myanmar connected characters with Zawgyi One 64 points is created using snipping tool. Here, Myanmar characters may be consonants or vowels or compound (i.e. combination of basic consonants and vowel modifiers or consonant modifiers. In this system, 150 warning text signboard images are collected from various environments since there is no available standard dataset for Myanmar text extraction and recognition system, where the signboard images are captured with the resize range of 200x300 to 400x600.

In Myanmar character, there are 33 consonants, 12 vowels, 4 medials, 10 digits, 3 asat, 1 kinzi and 12 independent vowels. And then, characters can be divided into two types such as basic characters and extended characters. In each basic character (consonant), there can have zero or more extended characters to create a compound word. The extended characters may be at left or right or top or bottom of the basic character. Segmentation between basic character and extended characters is the difficult problem in the printed and typed face character recognition system. Since one character is formed from the other by adding different signs and marks, most of the alphabets have some common features. Some Myanmar characters that have similarities with few distinctions are ဟ and တ, စိ and စီ, ွိုး and ြိုး, ဃ and ယ, ပ and ႎ, ထ and ၜ. All Myanmar symbols is shown in Figure 1.

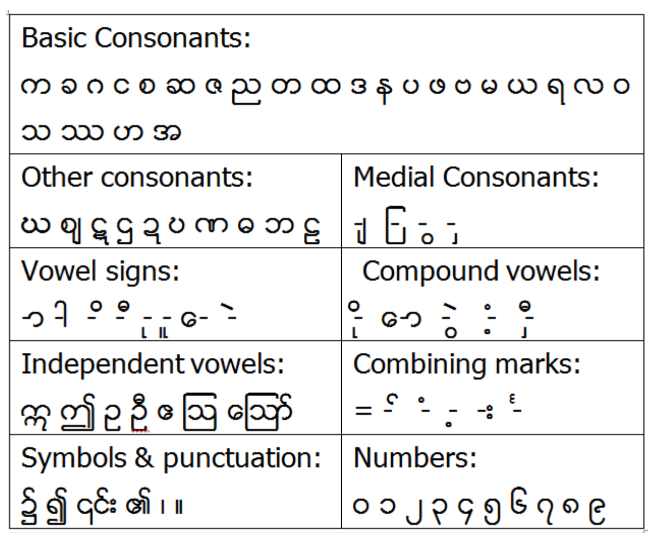

Figure 1: Myanmar basic symbols

IV. Proposed System

In this proposed system, the following four main modules are mainly performed:

A. Text Information Extraction

Text extraction is critical and essential step since the efficiency of the OCR depends upon the accuracy of the text extraction system. In the text extraction step, RGB color enhancement is firstly processed by finding new RGB values using (1, 2 and 3).

$$New\_R = OR + round(Old\_R * 0.8) \quad (1)$$
$$New\_G = Old\_G + round(Old\_G * 0.8) \quad (2)$$
$$New\_B = Old\_B + round(Old\_B * 0.8) \quad (3)$$

Where, NR= new red value, NG=new green value, NB=new blue value, OR = original red value, OG= original green value and OB= original blue value.

After color enhancement, RGB image is converted into binary image by filling ‘1’ in the places where all RGB values are greater than color threshold value 215 and filling ‘0’ in other places. The connected components (CCs) in the binary image may be too large and/or too small connected components. Since these components, obviously, are not text, they are required to be deleted. In this system, the features such as height, width and aspect ratio of detected object are used to filter non-text CCs. And then, the non-text objects around the signboard boundary are removed by finding center location of the signboard image. The remaining too small non-text CCs of this stage are further filtered based on the object region area threshold. The CCs less than object region area threshold values are discard and CCs greater than region area threshold values are dilated for grouping text. After grouping text regions, the remaining non-text regions are removed based on region area threshold of grouping text region. The implementations of various text extraction are shown in Figure 2.

Figure 2 : Step by step text extraction

B. Segmentation

In this system, horizontal projection profile is used for line segmentation and vertical projection profile is used for vertically connected characters. The generated characters from vertical projection profile may have 1 to 5 characters which may be or may not be touching. Therefore, bounding box method is also used to segment non-touching character resulted in previous vertical projection profile until achieving individual character.

The segmentation steps of the horizontal projection profile are as follows:

1) Count the black pixel in each row of the image.
2) Find the rows containing no white pixel.
3) Crop each text-line.
4) Input the cropping text-line image to vertical projection profile step.

The character segmentation steps of vertical projection profile and bounding box method are as follow:

1) Count the black pixel in each column of the image.
2) Find the columns that containing no white pixel.
3) Crop each vertically connected character
4) Extract the individual connected character from the resulted character of step 3 by using connected component labeling with bounding box method.

C. Feature Extraction

The feature extraction is described about the characteristics of an image. It is one of the most important steps for any recognition system, since the classification/recognition accuracy is depending on the features. Before any other extraction of features, the segmented input character is preprocessed by resizing and thinning. In the feature extraction phase, normalization eight direction chain code features and pixel count features are extracted from the blocking image. Eight direction chain codes is one of the shape representations which are used to represent a boundary by a connected sequence of straight line segments.

The direction of each segment is coded by using a numbering scheme as shown in Figure 3. Chain codes based from this scheme are known as Freeman chain codes.

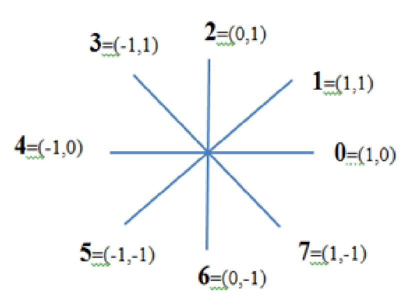

Figure 3 : Eight-direction chain code illustration

The chain code can be generated by following the object boundary clockwise and assigning a direction to the line segment connecting each pair of pixels. The chain code method select the topmost value in the leftmost column as starting pixel of the character object boundary shown in right side of Figure 4. The goal is to find the next pixel on the boundary image. Depending on what it is, a numeric code from 0 to 7 is assigned as chain code. Repeat the process of positioning the next border pixel and assign the code until there is no remaining unvisited pixel and we return to the initial border pixel. Therefore, almost boundary pixels are visited twice to reach initial starting point.

Proposed feature extraction steps are described as follow:

Step 1: Frequency features of chain code directions are extracted on the input character as shown in left side of Figure 4.

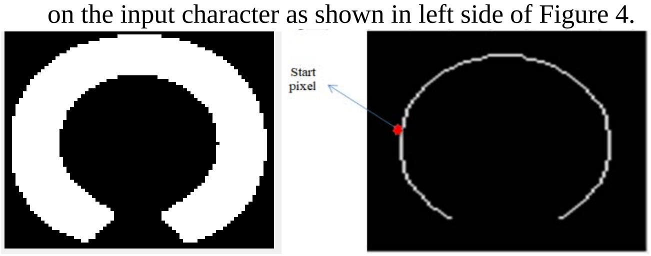

Figure 4 : Original character (left) and preprocessed character (right) of ဂ "gange"

The traversing of chain code directions on the input character is shown in figure 5.

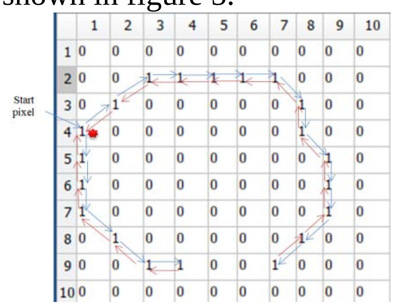

Figure 5 : 10 Traversing example of 10x10 character 'gange' image with binary values

In the above figure 5, blue arrow is first time traversing direction and red arrow is second time traversing direction. According to Figure. 10, row_no and column_no of visited pixels are described as follow:

row_no=[4,3,2,2,2,2,2,3,4,5,6,7,8,9,8,7,6,5,4,3,2,2,2,2,2,3,4,5,6,7,8,9,9,9,8,7,6,5,4]

column_no=[1,2,3,4,5,6,7,8,8,9,9,9,8,7,8,9,9,9,8,8,7,6,5,4,3,2,1,1,1,1,2,3,4,3,2,1,1,1,1]

In this example, the start pixel exist at row_no = 4 and column_no = 1 and second visited pixel is at row_no = 3 and column_no = 2, third visited pixel is row_no = 2 and column_no = 3, and so on.

Based on these row_no and column_no of visited pixels, x value and y value are generated as follow:

x = column_no(next_pixel) – column_no(previous_pixel)

y = row_no(previous_pixel) – row_no(next_pixel)

Chain code direction numbers are assigned based on the following x and y conditions:

```
if x==1&y==0
    Code = 0;
elseif x==1 &y==1
    Code = 1;
elseif x==0&y==1
    Code = 2;
elseif x==-1&y==1
    Code = 3;
elseif x==-1&y==0
    Code = 4;
elseif x==-1&y==-1
    Code = 5;
elseif x==0&y==-1
    Code = 6;
elseif x==1&y==-1
    Code = 7;
end
```

The following chain codes are extracted according to the x values and y values of this example and the frequencies of chain code direction results are shown in TABLE I.

Code = [1 1 0 0 0 0 7 6 7 6 6 5 5 1 1 2 2 3 2 3 4 4 4 4 5 5 6 6 6 7 7 0 4 3 3 2 2 2]

Table I. Frequency of eight chain code directions

Direction	0	1	2	3	4	5	6	7
Frequency	5	4	6	4	5	4	6	4

Step 2: The input character image is divided into N x N blocks (this paper presents 4x4 blocks) shown in Figure 6.

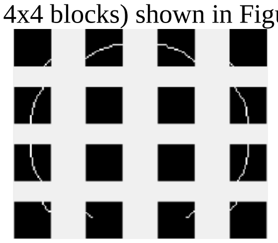

Figure 6 : 4x4 blocks character image

a) Frequency features of 8 chain code directions [3] on each of 4x4 blocks character by performing as the processes of first step and therefore 8xNxN (N=4) features are extracted.
b) The total number of visited directions on each of NxN blocks is extracted by using (4) and therefore 16 features are extracted.

$$B_n = \sum_{i=0}^{7} F(i) \quad (4)$$

Note that, B_n = Total frequency of eight directions on Block n of character, where n= 1, 2, 3, …, N x N (In this paper N is assigned 4).

c) White pixel (binary value 1) density features on each of above 4x4 blocks image are extracted using equation (5). In this step, 16 features are extracted from 16 blocks.

$$W_n = \sum_{r=1}^{100/N} \sum_{c=1}^{100/N} x_{r,c} \quad (5)$$

Where, W_n be sum of pixel values in block n, n=1, 2, 3,…, 16 (if N=4).

$x_{r,c}$ be pixel value in row r, column c of block n.

Step 3: the input character is divided into Nx1 blocks character image and 1xN blocks character image shown in Figure 7(a) and (b) using (6 and 7).

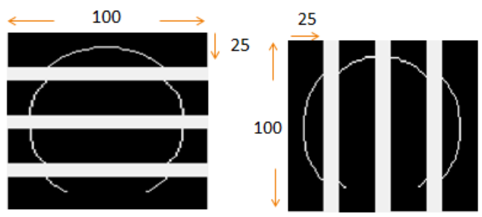

Figure 7 : (a) 4x1 blocks character image and (b) 1x4 blocks character image

$$R_i = \sum_{r=1}^{100/N} \sum_{c=1}^{100} x_{r,c} \quad (6)$$

$$C_i = \sum_{r=1}^{100} \sum_{c=1}^{100/N} x_{r,c} \quad (7)$$

Where, R_i be the sum of pixel values in horizontal block i.
C_i be the sum of pixel values in vertical block i $x_{r,c}$ be the pixel value in row r and column c.

*S*tep 4: Number of segments in upper block and lower block are extracted by dividing the character image into horizontally 2 x1 blocks shown in Figure 8.

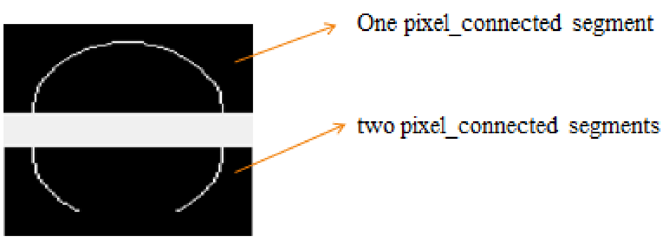

Figure 8 : 2x1 blocks character image

According to the above feature extraction steps, six feature sets are identified and analysis the recognition accuracy resulted from using each feature set for N x N blocks (N = 4).

- Feature_Set_1 includes features from step 1, 2(c), 3, 4 and aspect ratio.
- Feature_Set_2 includes features from step 1, 2(b, c), 3, 4 and aspect ratio.
- Feature_Set_3 includes features from step 1 and 2(a).
- Feature_Set_4 includes features from step 1and 2(a, b).
- Feature_Set_5 includes features from step 1, 2(a, c), 3, 4 and aspect ratio.
- Feature_Set_6 includes features from step 1, 2, 3, 4 and aspect ratio.

Number of features included in the feature sets is different according to the divided number of N x N blocks as shown in TABLE II.

Table II. Different number of features according to different numbers of blocking

Feature_Set Type	No of features on 3x3 blocks	No of features on 4x4 blocks	No of features on 5x5 blocks
Feature_Set_1	26 features	38features	46 features
Feature_Set_2	35features	51features	71features
Feature_Set_3	83features	139features	211features
Feature_Set_4	92features	155features	236features
Feature_Set_5	98features	163features	246features
Feature_Set_6	107features	**179 features**	271features

D. Classification

Characters are recognized by matching the features of input character with the features of all 343 characters that are already stored in the database. The distance values between tested character and the characters trained in the clustered group are calculated using Euclidean distance equation (8). The minimum distance value is selected and returns the position of that value.

$$dist\big((x,y),(a,b)\big) = \sqrt{(x-a)^2 + (y-b)^2} \quad (8)$$

V. EXPERIMENTAL RESULTS

In currently experiment, 343 Myanmar Zawgyi-One font characters that contain isolated characters and 2 to 5 connected component characters are trained. In this system, 3283 objects are extracted from 150 warning text signboards in current time. This system is investigated by KNN classifier, SVM classifier and feature based approach of template matching method using various blocking and various features on 3283 objects. The character classification performances are shown in TABLE III and VI.

Table III. Recognition accuracy of various types of feature sets on various blocking

Number of extracted objects	Feature_Set_Type	Recognition accuracy on 3 x 3 blocks	Recognition accuracy on 4 x 4 blocks	Recognition accuracy on 5 x 5 blocks
3283	**Feature_Set_1**	(66.16%)	(72.55%)	(72.46%)
3283	**Feature_Set_2**	(66.07%)	(74.66%)	(72.74%)
3283	**Feature_Set_3**	(66.28%)	(69.08%)	(64.88%)
3283	**Feature_Set_4**	(67.10%)	(79.89%)	(68.96%)
3283	**Feature_Set_5**	**(73.44%)**	(75.60%)	**(73.8%)**
3283	**Feature_Set_6**	(71.06%)	**(75.63%)**	(72.89%)

Table IV. Character classification performances using various classification methods 100x100 resized image, 4 x 4 blocking and feature_set_6

No of Image	Total Extracted objects	Fine KNN	Medium SVM	Area based matching	Features based matching
150	3283	(67.1%)	67.1%	64.31%	75.63%

VI. ERROR ANALYSIS

There are a few limitations in this system such as:

1. Since Myanmar characters on the natural warning signboard images are various styles and various size, the connected characters or words can not be segmented into individual characters. Therefore, we have to consider the connected word samples in the training

character dataset for classifying the connected words from the warning text images.

2. The Myanmar compound characters on the warning signboard images may be often touching or non-touching and they are handwritten styles. Therefore, if the basic character and extended characters in the compound word were touching as shown in Figure 9, the character segmentation system will not further segment that compound word and if not, the system will further segment that compound word as shown in Figure 10. Therefore, this system identifies the number of final segmented characters as the number of characters.

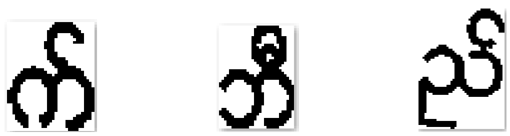

Figure 9 : Touching formation of basic character and extended character (connected compound characters)

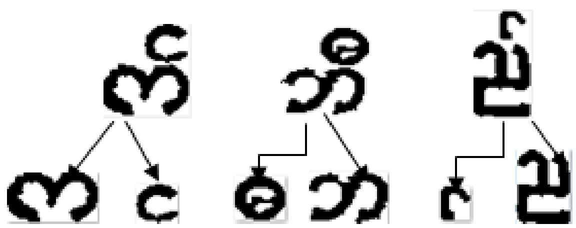

Figure 10 : Non-touching formation of basic character and extended character (non-connected compound characters)

VII. CONCLUSION

The proposed system is investigated with dynamic blocking features, normalization chain code features and various classification methods. Among them, features based approach of template matching method is the most suitable method for this system and achieves 75.63% on 3283 extracted objects of 150 warning signboard images. Since the proposed features are based on pixels, this system can also be used in the character recognition system of any scripts.

REFERENCES

[1] AL-Hashemi and Alsharari, “Instant Arabic Translation System for Signboard Images Based on Printed Character Recognition”, International Journal of Machine Learning and Computing, vol. 3, N0. 4, Augest 2013.

[2] E.E. Phyu, Z.C. Aye, E.P. Khaing and Y. Thein, 2008, “Recognition of Myanmar Handwritten Compound Words based on MICR”, the 29th Asian Conference on Remote Sensing .

[3] Emmanuel, Rosemol, and Jilu George. "Automatic detection and recognition of Malayalam text from natural scene images." IOSR Journal of VLSI and Signal Processing 3.2 (2013): 55-61.

[4] H.P.P.Win, P.T.T. Khine and K.N.N.Tun, 2011 “Bilingual OCR System for Myanmar and English Scripts with Simultaneous Recognition”, International Journal of Scientific & Engineering Research Volume 2, Issue 10, October, ISSN 2229-5518.

[5] K.P. Zaw and Z.M. Kyu, “Segmentation Method for Myanmar Character Recognition using Block based Pixel Count and Aspect Ratio”, 27th International Conference on Computer Theory and Application (ICCTA), October 2017.

[6] M. Sayed and S.A. Angadi, “Mobile Application for Reading Display Boards having Kannada Text”, International Journal of Recent Trends in Engineering and Research, Vol. 02, Iss. 08; August- 2016.

[7] S.A. Angadi, and M. M. Kodabagi. "A robust segmentation technique for line, word and character extraction from Kannada text in low resolution display board images." Signal and Image Processing (ICSIP), 2014 Fifth International Conference on. IEEE, 2014.

[8] S. B. Ahmed, S. Naz, M. I. Razzak, & R. Yousaf. 2017. Deep Learning based Isolated Arabic Scene Character Recognition. arXiv preprint arXiv:1704.06821.

[9] T. Swe and P. Tin, 2006, “Recognition and Translation of Myanmar Printed Text based on Hopfield Neural Network”, IEEE.

[10] Y.Thein and S.S.S. Yee, 2010, “High Accuracy Myanmar Handwritten Character Recognition using Hybrid Approach through the MICR and Neural Network”, International Journal of Computer Science Issues, Vol. 7, Issue-6, November -2010.

[11] S. Sharma and C. Roxanne, 2007, Extraction of Text Region from Natural Images, Master Project Report Book.

[12] T. Tint, and N. Aye, 2014, “Myanmar Text Area Identification from Video Scenes”, International Conference on Advanced in Engineering and Technology (ICAET2014), March, Singapore.

[13] H. S. Mohana, R. Pradeepa, B. Rajithkumar BK and M. Shivakumar, “Printed and Handwritten Mixed Kannada Characters Recognition using Template Matching Method. International Journal of Electronic and Communication Technology, Vol. 6, Iss. 2, 2015.

.