International Journal of Electrical Electronics & Computer Science Engineering
Volume 6, Issue 2 (April, 2019) | E-ISSN : 2348-2273 | P-ISSN : 2454-1222
Available Online at www.ijeecse.com

Recognition of Myanmar Lottery Ticket Number

Thi Thi Soe[1], Zarni Sann[2]
[1]Professor, Faculty of Computer Science, [2]Professor, Faculty of Computer Systems and Technologies
[1-2]University of Computer Studies (Mandalay), Myanmar
[1]hlaing.nn@gmail.com, [2]zarnisann@gmail.com

Abstract: One of the most popular research in the fields of computer vision is optical character recognition (OCR).OCR is the mechanical or electronic conversion of typed, handwritten or printed text images into editable form; notepad, excel, word. Methods widely used for OCR of Asian languages are mainly involved pattern matching techniques. In this paper we are applying OCR techniques to Myanmar lottery ticket, and trying to extract each alphanumeric present in the lottery ticket number region. The recognition operations consist of following: image acquisition, pre-processing, interest region extraction, element segmentation, and recognition. Myanmar lottery ticket number recognition system is developed by computing the correlation coefficient to compare the similarity for alphanumeric patterns via image processing techniques with MATLAB programing language. To demonstrate the effectiveness of the proposed system the numbers of lottery ticket images having various alphabets and numerals are experimented. Experimental results show that the system achieves with high degree of recognition accuracy.

Keywords: Myanmar Lottery Ticket Number, OCR, Template Matching.

I. INTRODUCTION

Every language in the world has different character and digit recognition algorithm. Many researchers concentrated on character and digit recognition to reduce the processing time with high accuracy. This can be accomplished by using computer technology [1,2]. OCR is a field of research in pattern recognition, artificial intelligence and computer vision. As a result of vigorous efforts on OCR, systems are available with respect to any language; English, Bangla and Japanese languages [3,4,5]. Until present time the idea of OCR technology is leading to a fruitful further research and applications in many places such as postal codes, bank cheque, cars plates, and passport ID. Banerjee et al. [6] developed an automated system to recognize the characters present in the payee namea nd handwritten cheque amount from a bank cheque image. In their attempt used region growing for image segmentation, and Centre of Gravity (COG) of the character for feature points extraction. Kumar et al. [7] achieved very high accuracy in extracting connected components from vehicle license plate using tree data structure and forward scanning. The most likely connected components was determined by applying the SAGAP algorithm.Passport has a special structure called MRZ (machine Readable Zone) which contains a fixed data format. Various recognition methods on MRZ recognition are proposed.MRZ recognition method using enhanced RBF network based on smearing and contour tracking is studied in [8]. Most of proposed methods are OCR based recognition of passport information. A number of research work on character and digit extraction from various areas are exploded as described above, but not many focus on Myanmar lottery tickets. Due to these facts, new methods are currently needed for Myanmar alphanumeric extraction from lottery ticket. This work is to design and develop an architecture that can recognize printed alphanumeric present in the Myanmar lottery ticket number. Lottery number is formally printed with the same font, size, font style, and font color for every month. Not variation in font attributes of the printed numbers makes the recognition task simply. Therefore, pattern matching approach used in our system is adequate for similarity score of recognized number. The organization of remaining sections is as follows: Section II describes the structure of Myanmar lottery ticket. Section III is devoted to creation of template database used for correlation. The proposed approach for Myanmar lottery ticket number recognition is discussed in section IV. Experimental results are demonstrated in section V. In Section VI conclusion of the work is given.

II. LAYOUT OF MYANMAR LOTTERY TICKET

(a)

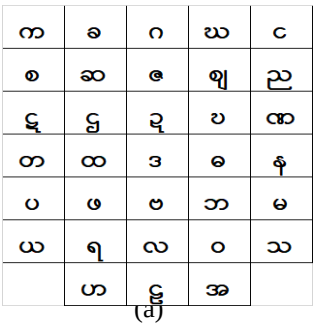
(b)

Fig. 1. Myanmar Lottery Ticket Showing (a) Front (b) Back Sides

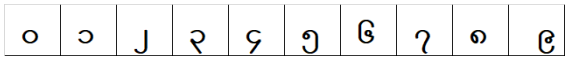
Myanmar lottery (also known as Aung Bar Lay, "may you win") is Myanmar's official state lottery. The monthly lottery is administered by the Ministry of Finance's State Lottery Department [9]. In every

International Journal of Electrical Electronics & Computer Science Engineering
Volume 6, Issue 2 (April, 2019) | E-ISSN : 2348-2273 | P-ISSN : 2454-1222
Available Online at www.ijeecse.com

month 500 kyats lottery ticket and 1000 kyats lottery ticket are selling by department. This paper only concentrated on 500 kyats lottery ticket. The front and back sides of 500 kyats lottery ticket is shown in fig. 1. The size of the lottery ticket is about 5.65 x 2.65 inch in width and high. The ticket number is red color and printed over the security thread at the top right corner. The lottery ticket number is composed of alphabet and six digits. Two number types are issued every month. In one number type, single alphabet followed by six digits. In this type, total element is seven. On the other hand, combined two alphabets append with six digits. In this type, total element is eight. Fig. 2 describes the 33 consonants and 10 digits in Myanmar language which are used in Myanmar lottery ticket number.



(a)

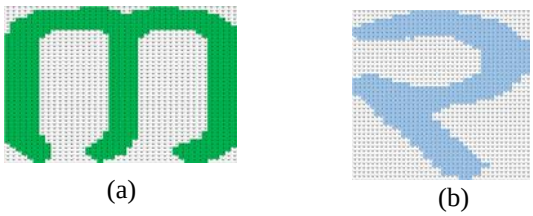


(b)

Fig. 2. Myanmar Language Showing (a) 33 Consonants and (b) 10 Digits

III. TEMPLATE LIBRARY

We initially collect various lottery tickets to obtain 33 consonants and 10 digits. Each alphabet of 33 consonants and each of 10 digits extracted from lottery ticket number regions are regarded as templates. All reference templates used in this system are two-level format and 40x40 pixels size. 'mat2cell' function transforms the binary images to a set of reference alphabet and digit arrays. In alphabet array contains 33 cells and in digit array consists of 10 cells along a single row and each cell having 40 rows and 40 columns. Then these two arrays are saved separately in template library and it is called by the lottery ticket number recognition module for correlation analysis. Fig. 3 is an example template for alphabet က and digit ၃. After collecting the necessary information from lottery ticket number we designed the proposed system.



(a) (b)

Fig. 3. Templates (a) alphabet က, and (b) digit ၃

IV. SYSTEM ORGANIZATION

The process flow of the system is illustrated in fig. 4. Some of the lottery tickets are taken by HUAWEI G520-0000 camera and some are taken with MiMax3 camera model. To obtain the desired image part, unnecessary sections are removed by paint tool and saved as JPEG (Joint Photographic Experts Group) format. These images are inputted to our recognition system. Input images are full color images. Pre-processing operations are needed. Pre-processing is the process to prepare and transform raw original data for speeding up and easier the successive stages.

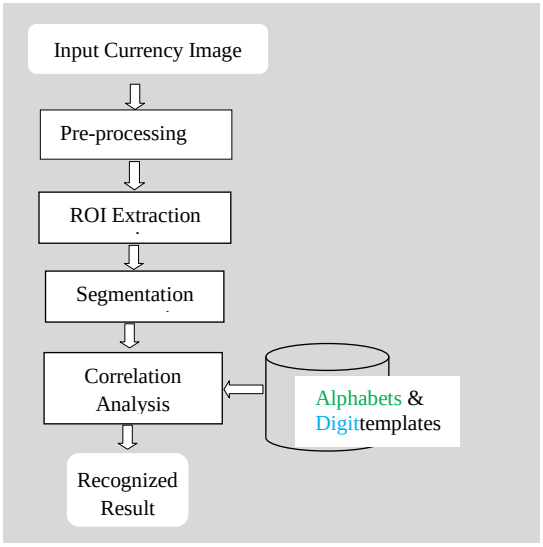


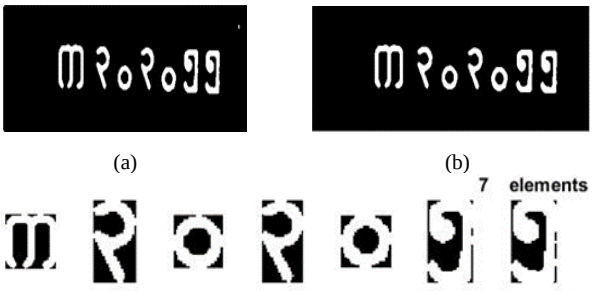
Fig. 4. The Process Flow of Proposed System

Image Binarization: After converting the gray image, the region of interest (ROI), lottery ticket number region, is extracted. In binarization, Otsu's method [10] finds the threshold level of image and then assigned value 0 for black pixel and assigned value 1 for white pixel by minimizing calculated threshold. Fig. 5(a) is binary image.

Image Enhancement: Noise may affect segmentation and pattern matching results. In this system, 5x5 non-linear median filter mask eliminates salt and pepper noise. It preserves edges while removing unwanted noise [11]. In some filtered results, fig. 5(b), there are some outliers at the border and inner region of the image. 'bwareaopen' clean the pixels fewer than 40 pixels and 'imclearborder' function remove any outliers that are connected to the border of the image.

Element Segmentation: In order to separate each alphabet and digit present in the ROI, we have used the connected component labelling algorithm [11]. The individual elements are cropped into different sub images as in fig. 5(c). Sub images are fixed to be ensured that the templates have the same 40x40 size of pixels for pattern matching. Numbers of connected components are recorded that are useful for recognition routine.

International Journal of Electrical Electronics & Computer Science Engineering
Volume 6, Issue 2 (April, 2019) | E-ISSN : 2348-2273 | P-ISSN : 2454-1222
Available Online at www.ijeecse.com



(a) (b) (c)

Fig. 5. (a) Binary Image (b) Filtered Image (c) Segmented Image

Correlation Analysis: There are two possible cases in the segmentation result. These cases are checked for correlation analysis.

Case 1: if the number of connected component =7,

- first element is compared with 33 templates and
- each of other six elements is compared with 10 templates

Case 2: if the number of connected component =8,

- each of first two elements is compared with 33 templates and
- each of other six elements is compared with 10 templates

by calling the alphabet and digit templates in the library respectively. For this comparison process 'corr2' computes the correlation coefficient between two matrices X and Y using the following equation:

$$c = \frac{\sum_m \sum_n (X_{mn} - \bar{X})(Y_{mn} - \bar{Y})}{\sqrt{(\sum_m \sum_n (X_{mn} - \bar{X})^2)(\sum_m \sum_n (Y_{mn} - \bar{Y})^2)}} \quad (1)$$

where, $\bar{X}$ and $\bar{Y}$ are mean2(X) and mean2(Y) respectively.

After computing the correlation coefficients, the operation detects the maximum value of it by 'find' function. In this way, recognition routine produces the recognized lottery ticket number.

During the training stage, output of some alphabets produced misrecognition result due to Myanmar characters are circular structures in nature and have similar forms. Similarity score for below alphabets are closely matched. To distinguish this misclassification, we applied hole filling operation on right side character.

Fig. 6. Closely Matched Two Myanmar Character

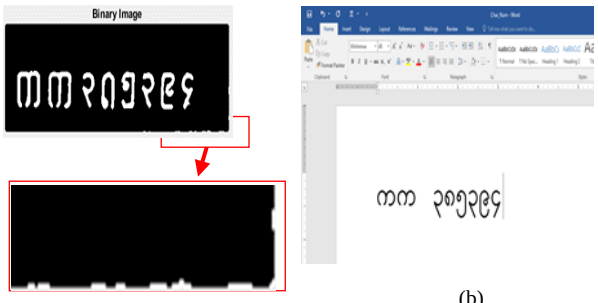
V. EXPERIMENTAL RESULTS

In the proposed lottery ticket number recognition system, 113lottery ticket images for thirteen months are used as testing input. In the first experiment, we test the image in which single alphabet followed by six digits as given in fig.7(a). Cropped number region is shown in left side of fig. 7 (b) with enhanced result. Transformed two-level image is the right side of the figure. Figure 7 (c), the results of correctly recognized lottery ticket number are outputted with figure window and printed form in MS word.

(a) (b) (c)

Fig. 7. (a) Original Image, (b) Pre-process Result, and (c) Recognized Number

Next experiment of lottery ticket number is the combined alphabet with six digits. In this experiment, the number of elements is equal to eight. The resultant binary image has some outliers at the border and zooming these affects are shown in fig. 8 (a). Clear border operation removed it. Finally recognized ticket number is obtained as in fig. 8 (b).



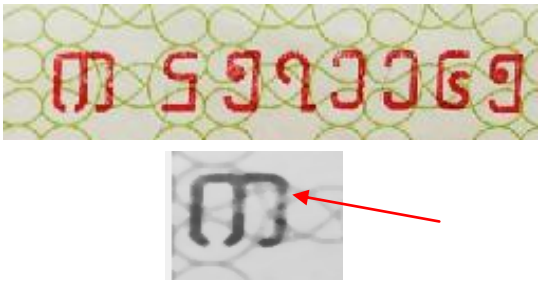
(a) (b)

Fig. 8. (a) Defected Binary Image (b) Recognized Result

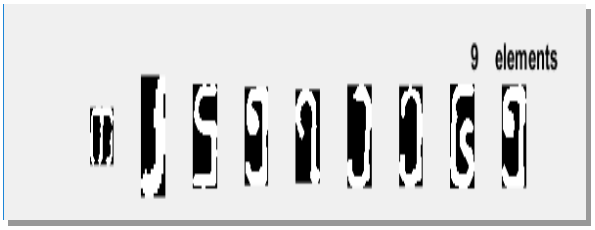
In our tested input image, if the printed tone in the number region tend to very soft, and pastel and pale

International Journal of Electrical Electronics & Computer Science Engineering
Volume 6, Issue 2 (April, 2019) | E-ISSN : 2348-2273 | P-ISSN : 2454-1222
Available Online at www.ijeecse.com

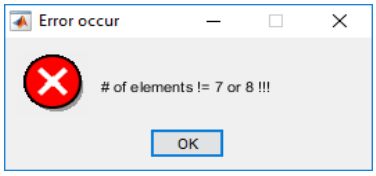
color, some of the elements are blended with the background security thread as in fig. 9(a). This may lead to the incorrect segmentation result. In fig. 9(b), number of connected components is 9 instead of 8. The system unable to recognize the ticket number for this type of input image and return the error message to user as in fig. 9(c).



(a)



(b)



(c)

Fig. 9. Unrecognized Ticket Image with Error

Table 1 shows the number of tested lottery tickets in this experiment. The recognition rate is defined as the ratio of the number of correctly recognized lottery ticket numbers to the total number of tested lottery tickets. Experimental results give the acceptable recognition rate.

Table I. Lottery Tickets for the Experiment

No.	Issued Times	Number of tickets		Total	Error
		Type I	Type II		
1.	9	5	-	5	-
2.	11	8	-	8	-
3.	13	6	-	6	-
4.	14	8	-	8	-
5.	16	15	1	16	-
6.	17	8	-	8	-
7.	18	5	3	8	-
8.	19	8	6	14	-
9.	20	9	5	14	2
10.	21	8	5	13	-
11.	22	3	8	11	1
12.	23	5	5	10	-
13.	24	5	-	5	-

VI. CONCLUSIONS

Myanmar Lottery ticket number recognition system is presented in this paper. The recognition process is based on simple correlation analysis. Different lottery ticket images with different ticket numbers are extracted and recognized. The output of the system is machine printed modifiable form in MS word. The proposed system worked adequately throughout the experiments and recognition accuracy is 88.5%. This approach can be regarded as a part of operations to check for winning lottery ticket numbers. To validate the winning lottery ticket, contents of four regions must be matched; lottery ticket number, opening date, issued time, and monthly ticket logo. In the future works, approaches for three necessary requirements will attempt in developing the checking tool for lucky lottery tickets.

VII. REFERENCES

[1] A. Sharma, K. K.Paliwal and G. C. Onwubolu , "Class-Dependent PCA, MDC and LDA: A Combined Classifier for Pattern Classification, " Pattern Recognition", Vol. 39, Issue 7, July 2006, PP. 1215-1229.

[2] Sergios Theodoridis and Konstantinos Koutroumbas, Pattern Recognition,4th Academic Press, 2008.

[3] O. Samanta, A. Roy, S. K. Parui, U. Bhattacharya, "An HMM framework based on spherical-linear features for online cursive handwriting recognition", Article in Information Sciences, · 441, February 2018, pp. 133-151.

[4] Debi Prasad Bhattacharya, S. Koner, "English Alphabet Recognition using Chain Code and LCS", Indian Journal of Computer Science and Engineering (IJCSE), ISSN: 0976-5166 Vol. 3 No. 2 Apr-May 2012, pp. -335.

[5] H. C. Fernando, N. D. Kodikara, and S. Hewavitharana, "A Database for Handwriting Recognition Research in Sinhala Language", Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003).

[6] S. Banerjee, B. R. Ghosh and A. Kundu, "Handwritten Character Recognition from Bank Cheque", International Journal of Computer Sciences and Engineering, Vol. 4, Special Issue 1, Feb 2016, pp. 99-104, E-ISSN: 2347-2693.

[7] P. Kumar, S. R. Giri, G. R. Hegde and K. Verma, "A Novel Algorithm to Extract Connected Components in a Binary Image of Vehicle License Plates", IJECCT, Vol. 2 (2), 2012, pp. 27-32.

[8] K.-B. Kim, Y.-J. Kim, A.-S. Oh, "An Intelligent System for Passport Recognition Using Enhanced RBF Network," LNCS, Vol. 3314, pp. 762-767, Springer 2004.

[9] https://en.m.wilipedia.org/wiki/Burmese_lottery.

[10] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on System, Man, and Cybernetics," 1979, Vol. 9, pp. 62-66.

[11] R. C. Gonzalez, R.E. Woods, "Digital Image Processing," Addison Wesley, 1998.