Journal of Theoretical and Applied Information Technology
31st July 2014. Vol. 65 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

HANDWRITTEN DIGIT RECOGNITION FOR MANAGING EXAMINATION SCORE IN PAPER-BASED TEST

[1,*]APICHAT SURATANEE, [1]NANTAPORN LERTSARI, [1]SETHAWAT KAMPHASEE, [1]KRITSADA SRIKET
[1]Department of Mathematics, Faculty of Applied Science,
King Mongkut's University of Technology North Bangkok, Bangkok, Thailand.
E-mail: *apichatsu@kmutnb.ac.th

ABSTRACT

Calculating scores of a paper-based test with a large number of students is a difficult task. Suppose a large number of students, e.g., more than 1200 students, in a class take a paper-based examination with more than 6 questions in the test. Teachers need to write the score of each question on the examination cover sheet, calculate total scores of all students and fill the scores on the student list. These tasks really take a lot of time. We, therefore, develop an automatic system for reducing some procedures of these tasks for time-saving. Our developed system uses handwritten number recognition that has been widely used in several fields. In this study, we employ the recognition with artificial neural network to automatically identify the student identification number and scores of all questions from a scanned image of the cover page. The summation of scores is calculated automatically. Both of total score and the student identification number are exported into excel format. With the neural network classification to recognize the digits, we obtain high performance with overall accuracy of 99.89%. In conclusion, two main processes are improved from our system: (i) automatic total score calculation and (ii) exportation of scores to excel format. This designed system could successfully reduce the time for evaluating scores of students and yields more accurate score calculation.

Keywords: *Examination Cover Page, Handwritten Digit Recognition, OCR, Artificial Neural Network.*

1. INTRODUCTION

Character and number recognition have been successfully applied into several fields, e.g., automatic signature verification in banking [1], postal address recognition [2], clinical data capture [3,4], etc. The aim of these recognitions is to read characters or digits from papers or read text or digits from documents. To reach the aim, images, or scanned images of the documents need to be translated into a format that computer is able to recognize. The information on the documents could be printed character or handwritten characters. In schools or universities, some elementary subjects have participated students in more than 30 sections and each section has more than 40 students. In every midterm and final examination, each examination has around 6 to 9 questions. That means if the test is paper-based test, the teachers or lectures need to fill scores of each student on the examination cover pages more than 7200 times. Moreover, for each student, the teacher needs to calculate the total score. That means they need to do it for 1200 times. Another difficult task is the finding of student names and filling their scores on the list. In addition, with a number of papers, total score calculation by human is possible to do error. These all processes are difficult and waste a lot of time to proceed. Therefore, the processes should be reduced for saving the operated time and increasing accuracy of the score calculation.

An alternative way to perform the test for large number of students is to use optical mark recognition (OMR). OMR is aimed to capture marked data from specific document forms [5, 6, 7]. Traditional OMR works with a scanner device that beams light onto a form paper. Reflectivity at pre-determined positions on the documents is used to detect the marks. One of the most familiar OMR applications for most students is using HB or 2B pencil marks their answer with bubbles on OMR answer sheet for multiple-choice examination. Most of OMR software [8, 9] could detect the marks of student identification number and student's answers, then, automatically calculate the total scores and export the score with the student identification number into a file. OMR is very useful; however, it is just suitable for multiple-

Journal of Theoretical and Applied Information Technology
31st July 2014. Vol. 65 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

choice testing. Hence, OMR could not answer our problem due to the fact that most of subjects do not perform by multiple-choice examination. Mostly, they are short answer test or essay test that teachers need to read student's answer and write the score manually on the cover sheet with handwriting. It would not be problem if the number of students is not much. However, in real situation with a number of students, the total score calculation and searching score of a student on a big list are certainly a difficult task. Therefore, other more powerful pattern recognition engine, e.g., optical character recognition (OCR), should be applied to solve this problem.

In this study, we aim to develop a systematic tool employing the digit handwritten recognition to recognize student identification number and examination scores on the examination cover page. Finally, total scores with student identification numbers are automatically exported to excel format. The main steps of the recognition from scanned image consist of (i) image processing, (ii) feature extraction and (iii) classification. We employed artificial neural network algorithm for the digit classification. Applying the method to the examination cover page is useful for to the time and complexity reduction.

This paper is organized as follows: In section 2, the overview of this study is described. In section 3, we present the image processing involving in this work and the feature extraction. In section 4, the method of classification using artificial neural network is described. In section 5, performance measurement of the classification is presented, and in section 6, classification results and graphic user interface of our system are showed. In section 7, we present conclusion.

2. OVERVIEW OF THIS STUDY

An example of the examination cover page is shown in Figure 1. The main parts of the page consist of the heading part (name of institute, faculty, subject name, semesters, etc.), the area for writing instructions part, the spaces to fill in the student identification number and name parts, the area for filling score of each question and final score parts. The important parts analyzed in this study are the student identification number and scores of all questions.

All cover pages with filled student identification numbers and scores were scanned as JPEG image data. The images were analyzed by image processing (see section 3) and all features representing the numbers were extracted (see section 4). These features were used for the classification processes using artificial neural network. The digits of the student identification number and scores were automatically identified. After the scores of all questions were recognized, the total score was calculated and exported to excel format.

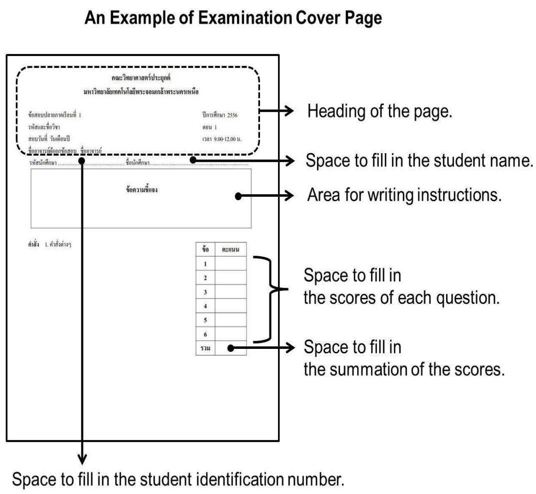

Figure 1: An example of examination cover page.

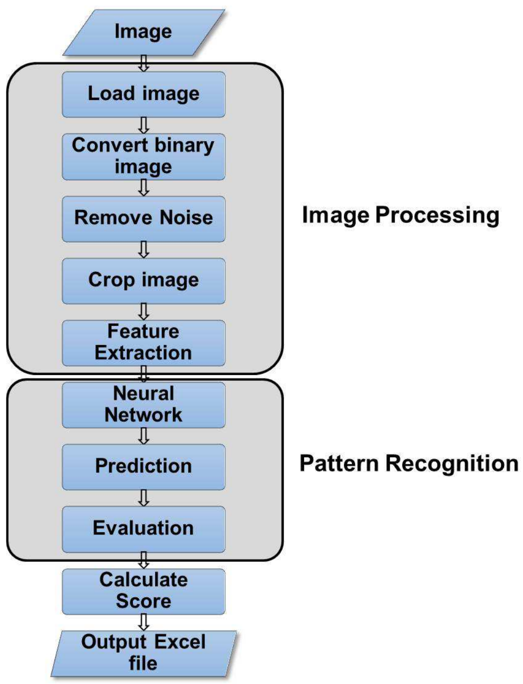

Figure 2: Overview of this study.

Journal of Theoretical and Applied Information Technology
31st July 2014. Vol. 65 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The further main processes could be divided into two parts comprising (i) image processing and feature extraction processes and (ii) pattern recognition using artificial neural network. The overview of this study is presented in Figure 2.

3. IMAGE PROCESSING AND FEATURE EXTRACTION

The main processes of image processing in this study consisted of binarization, noise filtering, image segmentation, and edge detection. Firstly, we adapted the input raw image data to the grayscale image and then transformed to binary image, i.e., 0-1 format. Then, noise filtering technique was applied to erase noised pixels from the image. Foreground/objects and background were identified by thresholding techniques. Let $P(x,y)$ be a pixel at position x and y in the image, $A(x,y)$ be a transformed pixel after the thresholding and T is a threshold value.

The simplest approach of thresholding is defined as follows:

$$A(x,y) = \begin{cases} 0 & ; P(x,y) < T \\ 1 & ; P(x,y) \geq T \end{cases} \quad (1).$$

After the thresholding process, the output image becomes binary image with value of 0 and 1 in the pixels. There are several algorithms [10] for reducing grayscale to binary image. In this study, we used Otsu's algorithm [11]. The algorithm assumed that an image contains two classes of pixels, e.g., foreground/objects and background. Therefore, the algorithm searches for an optimal threshold minimizing intra-class variance or maximizing inter-class variance of these two groups. Intra-class variance is defined as follows:

$$\sigma_{\text{intra}}^2 = W_b \sigma_{\text{b}}^2 + W_f \sigma_{\text{f}}^2 \quad (2).$$

W_b is the background class probability computed from histogram of a threshold T. It is defined as

$$W_b = \sum_{i=1}^{T} p(i) \quad (3).$$

While μ_b is class mean in which defining as

$$\mu_b = \sum_{i=1}^{T} \frac{i \cdot p(i)}{W_b} \quad (4),$$

σ_b which is variance of this background class is defined as

$$\sigma_b^2 = \sum_{i=1}^{T} (i - \mu_b) \cdot \frac{p(i)}{W_b} \quad (5).$$

For the foreground class probability W_f, it is defined by

$$W_f = \sum_{i=T+1}^{I} p(i) \,. \quad (6),$$

where I is the maximum range of the intensity levels. Foreground class mean is defined as follows:

$$\mu_f = \sum_{i=T+1}^{I} \frac{i \cdot p(i)}{W_f} \quad (7),$$

and the variance of the foreground class is defined as

$$\sigma_f^2 = \sum_{i=T+1}^{I} (i - \mu_f) \cdot \frac{p(i)}{W_f} \quad (8).$$

The algorithm tried to find T that minimize the intra-variance as defined in (2) or, likewise, maximize the inter-variance that is defined as

$$\sigma_{\text{inter}}^2 = W_b W_f (\mu_b - \mu_f)^2 \quad (9).$$

A digit segmentation was applied to the image data. Therefore, the isolated digits of the student identification number or scores were obtained with different sizes and ready to extract their features. Due to the different size of each isolated digit, we rescaled the digit to be an image with 70x50 pixels. Hence, the whole rescaled pixel area was divided into 7 and 5 parts according to row and column, respectively. That means the area was divided into 35 small sub-areas. In each sub-area, we calculated the summation of pixel intensity values. Finally, we yielded 35 values taken as features of a digit number. Figure 3 showed an example of using image processing to identify student identification number.

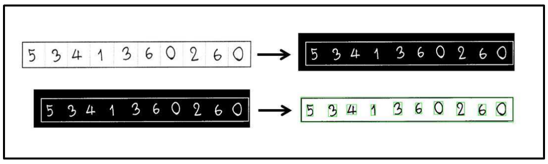

Figure 3: Image processing to identify student identification number

Journal of Theoretical and Applied Information Technology
31st July 2014. Vol. 65 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. ARTIFICIAL NEURAL NETWORK FOR CLASSIFYING DIGITS

In pattern recognition, data is divided into two parts: the training data and testing data. The training data is used for the neural network to learn all patterns and the testing data is used to measure the performance. In this study, all 9530 digit samples were collected from scanned images. Figure 4 shows an example of scanned image using for extracting digit data for training and testing. The figure contains 29 lines of digits. Therefore, we obtained 290 samples from this image.

Each isolated digits was segmented and extracted 35 features for training and testing with neural network. The architecture of the neural network in this study is multi-layered perceptron network with back-propagation learning algorithm. Input layer contained 35 input nodes, the same size as the features. The second layer of the network is called hidden layer. We denoted the number of the nodes in the hidden layer as 35 nodes. The third layer of the network is output layer. In our case the output layer contained 10 output nodes. We used log-sigmoid function as our transfer function for this learning. For the network, we want to predict the isolated digit image which number it is. Therefore, the targets of this training are the number in range of 0 to 9. Here, we used a classical binary encoding to represent each target with 10 output nodes of neural network. Examples of designed target 1, 5, 9 and 0 are shown in Table 1.

For measuring the performance of the algorithm, we used bootstrapping method by randomly selecting 80% of the each group of digits for training and the rest 20% of each group for testing. We repeated overall 5 times and the average of the performance was calculated.

Table 1: An Example of Binary Encoding Designed for Representing Output Targets of Neural Network.

Output Node	Target 1	Target 5	Target 9	Target 0
1	**1**	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	**1**	0	0
6	0	0	0	0
7	0	0	0	0
8	0	0	0	0
9	0	0	1	0
10	0	0	0	1

5. PERFORMANCE MEASUREMENT

Accuracy of all digit predictions was calculated for the performance measuring. The accuracy can be calculated by

$$\text{Accuracy} = \frac{\text{The number of corrected prediction}}{\text{The number of all prediction}} \qquad (10).$$

In addition, we could find how good prediction for each single digit is by calculating accuracy. Similarly with formula in equation (10), the accuracy of each single digit is the number of corrected prediction of that digit divided by the total number of all prediction.

Figure 4: 290 sample digits of scanned image for training data.

6. EXPERIMENTAL RESULTS

6.1 Classification Performance

The training of artificial neural network and the graphic user interface of our system were done by MATLAB software. For training and testing, 7620 samples were randomly selected for training and 1910 samples were used for testing. The best

Journal of Theoretical and Applied Information Technology
31st July 2014. Vol. 65 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

performance met at the mean square error of 1.0E-04 at epoch 475. Figure 5 showed the performance neural network learning. Finally, we yielded 99.92% of accuracy for training process, and 99.89% of accuracy for testing process. We found that the misclassification between 1 and other digits, like 2, 3, 5, and 7, were a bit high. As well, we found the misclassification between 3 and 9 and between 4 and 6. These occurred due to the similarity of characteristic of these numbers. Figure 6 and Figure 7 showed the confusion matrix of training and testing processes.

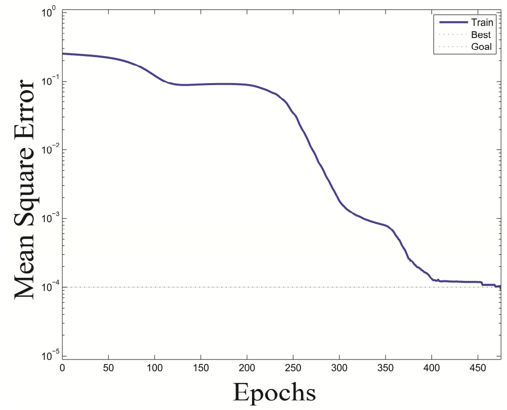

Figure 5: Training Performance with 475 epochs to reach the gold at 1.0E-04.

Targets

Prediction	1	2	3	4	5	6	7	8	9	0
1	756.2	0	0	0	0	0	0	0	0	0
2	2.6	762.0	0	0	0	0	0	0	0	0
3	1.4	0	762.0	0	0	0	0	0	0	0
4	0	0	0	762.0	0	0	0	0	0	0
5	0.2	0	0	0	761.8	0	0	0	0	0
6	1.4	0	0	0	0.2	762.0	0	0	0	0
7	0	0	0	0	0	0	762.0	0	0	0
8	0	0	0	0	0	0	0	762.0	0	0
9	0	0	0	0	0	0	0	0	762.0	0
0	0.2	0	0	0	0	0	0	0	0	762.0

Figure 6: Confusion Matrix of Training Set. The overall accuracy is 99.92%

Targets

Prediction	1	2	3	4	5	6	7	8	9	0
1	189.6	0	0	0	0	0	0	0	0	0
2	0.6	191.0	0	0	0	0	0	0	0	0
3	0.2	0	190.8	0	0	0	0	0	0	0
4	0	0	0	190.6	0	0	0	0	0	0
5	0.4	0	0	0	191.0	0	0	0	0	0
6	0	0	0	0.4	0	191.0	0	0	0	0
7	0.2	0	0	0	0	0	191.0	0	0	0
8	0	0	0	0	0	0	0	191.0	0	0
9	0	0	0.2	0	0	0	0	0	191.0	0
0	0	0	0	0	0	0	0	0	0	191.0

Figure 7: Confusion Matrix of Testing Set. The overall accrucy is 99.89%.

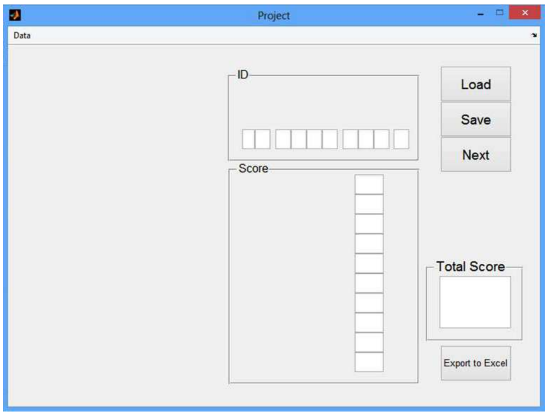

Figure 8:Start page of the program.

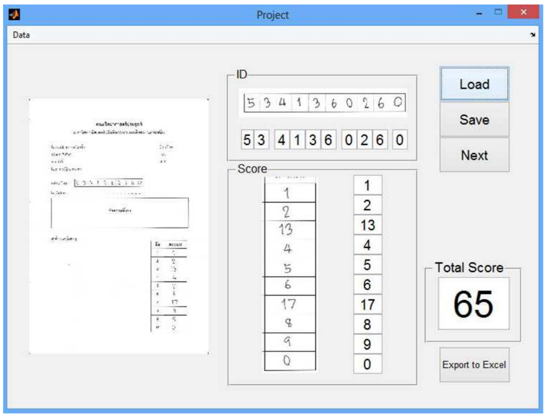

Figure 9: An example of the program with automatic score calculation.

6.2 Graphic user interface of the system

We developed the graphic user interface for easily managing the system. All scanned examination cover pages were stored in a defined input directory. The start page of the program is shown in Figure 8. Mainly, there are 4 buttons comprising "Load", "Save", "Next", and "Export to Excel". By clicking the "Load" button, the program inquires the input directory and then the first image was loaded. For each image, the student identification number and the scores were detected automatically by using image processing and the classification from the artificial neural network algorithm. The digits of student identification number and the scores were appeared on the defined text boxes. If the predicted digits were wrong from the classification algorithm, we can edit them directly from the text box. The summation of the scores was automatically calculated and showed at the bottom right space. If we changed the score of each question in the text box, the total score was automatically recalculated

Journal of Theoretical and Applied Information Technology
31st July 2014. Vol. 65 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

and showed. When we verified that the student identification number and all score including the summation score were corrected, we clicked "Save" button. To go to the next cover page, we can click at "Next". Then, the next cover image will show on the left blank space. Similarity, the same procedures were applied to the new image. We can click "Next" until the last page. Then a message informed the end of image file is presented. Finally, we can click "Export to Excel" to export all data into excel format. An example of the program is shown in Figure 9.

7. CONCLUSION

We developed a tool for improving the score management processes from scanned images of examination cover pages. This tool utilized image processing techniques for detecting and extracting digits from the scanned images of the examination cover pages. Artificial neural network was employed to recognize the numbers from the images. The results yielded that the neural network was successfully recognized the number with high accuracy of 99.92% and 99.89% for training and testing set, respectively. The student identification number and scores were identified from the scanned images. Moreover, the graphic user interface was developed for user-friendly by MATLAB. Users can simply upload all scanned images of examination cover pages located in a directory. Then, the program can easily load them to detect the identification number and scores and, then, calculate the total scores. In addition, user can export all results into excel format. All things considered, there are two main parts are improved from our automatic system, i.e., (i) an automatically total score calculation and (ii) exporting student's scores and their student identification numbers to excel format. This tool is simple, useful and practical for any subjects with a large number of registered students.

Nevertheless, there are some limitations of this tool that could be further improved. In some cases, a number, e.g., of student identification number or score, was misclassified from the neural network algorithm. This could be happened if the user's handwriting of the users is quite different from the data that neural network was learned. This problem could be solved by adding digit handwriting of the users into the system. However, the tool has an option that user can easily edit the number immediately from the textbox in the graphic user interface if they found some mistake from the classification. Our tool is designed to recognize only numbers, e.g., the score and student identification number. It could be further developed to recognize the name of students that the recognition needs character sets to learn.

REFRENCES:

[1] Kishor Mane and Vandana Pujari, "Signature Matching with Automated Cheque System", *2013 International Conference on Intelligent Systems and Signal Processing (ISSP)*, pp. 166-169, 2013.

[2] Sargur N. Srihari and Sangjik Lee, "Automatic Handwriting recognition and Writer Matching on Anthrax-related Handwritten Mail", *Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition (IWFHR'02)*, IEEE Computer Society, 2002.

[3] Paul G. Biondich, J. Marc Overhage, Paul R. Dexter, Stephen M. Downs, Larry Lemmon, and Clement J. McDonald, "A modern optical character recognition system in a real world clinical setting: some accuracy and feasibility observations", *Proc AMIA Symp 2002*, pp. 56-60, 2002.

[4] Bryan Bergeron, "Clinical Data Capture: OMR and OCR and Your Flatbed Scanner", *MedGenMed*, Vol.7, No.2, May 2005.

[5] Krisana Chinnasarn, Yuttapong Rangsanseri, "Image-processing-oriented optical mark reader", Proc. *SPIE 3808, Applications of Digital Image Processing XXII,* October 1999.

[6] Bryan P. Bergeron, "Optical mark recognition", *Postgraduate Medicine*, Vol. 104, No.2, August 1998.

[7] Haag, S., Cummings M., McCubbrey, D., Pinsonnault, A., Donovan, R., "Management Information Systems for the Information Age", 3rd ed., Canada: McGraw-Hill Ryerson, 2006.

[8] Gravic Inc, "Remark Office OMR®", available at http://www.gravic.com/remark/Office-OMR-Scanning-Software/ on May 2014.

[9] Lacuritie Software, "OGAMI OMR", available at http://www.lacuritie.com/en/ogami-omr/ on May 2014.

[10] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation", *Journal of Electronic Imaging*, Vol. 13, No. 1, pp. 146–165, 2004.

[11] Nobuyuki Otsu, "A threshold selection method from gray-level histograms", *IEEE Trans. Sys., Man., Cyber*, Vol. 9, No.1, pp. 62–66, 1979.