Optical Character Recognition Using Modified Direction Feature and Nested Multi Layer Perceptrons Network

Mahmud Dwi Sulistiyo, Deni Saepudin, Adiwijaya
Telkom Institute of Technology
Bandung, Indonesia
mahmuddwis@gmail.com, dns@ittelkom.ac.id, adw@ittelkom.ac.id

***Abstract*—The studies of Optical Character Recognition (OCR) are being developed since it still needs a performance improvement. The previous study of alphanumeric character recognition had been conducted by Blumenstein and Liu using Modified Direction Feature (MDF) and Multi Layer Perceptrons (MLP) network. The study reaches the accuracy rate of 70.22% for lowercase characters and 80.83% for uppercase characters.**

In this study the OCR system is proposed to improve the existing performance and have a capability to recognize all case-sensitive alphanumeric characters simultaneously. One of the problems is that there are several characters having similarities in gesture and shape, so that the classifier of the OCR system encounters many ambiguities when classifying some particular characters, especially when recognizing all case-sensitive alphanumeric characters.

To overcome those problems, this study proposes a technique of grouping. All character classes are clustered into some groups using Fuzzy C-Means (FCM) clustering method. The Nested MLP is the novelty of the previous method that is implemented in this study. This is a kind of multi-level MLP network that classifies the problem domain hierarchically. The first level classifies the character into the designated group and the second level continues the classification into the recognized character class. The OCR system using the methods to recognize all case-sensitive alphanumeric characters yields an accuracy rate of 84.38% for the uppercases, 76.43% for the lowercases, and 78.92% for the digits respectively. Any misclassified characters are mostly happened in distinguishing several uppercase and lowercase characters having similarities in gestures and shapes.

Keywords—recognition; OCR; MDF; Nested MLP; case-sensitive alphanumeric characters

I. INTRODUCTION

Technology nowadays produces a system enabling computer to read handwriting as an input, and processes it in order to provide useful information. This kind of system is called Optical Character Reader or Optical Character Recognition (OCR). An OCR system is able to read handwriting patterns and recognize them as characters corresponding with the intention of writing.

One of the utilizations of OCR system in the real world is the implementation of migrating data from the paper-based data form to the digital text database because of some reasons. It is known that paper-based data is very risky, easily damaged, and hard to manage. It is also not environment friendly, wasting money, and space consuming. Moreover, in the future, all kind of data will be integrated as digital data. With the OCR system, hopefully, the paper-based data migration will be easier and faster.

In the past research, it has been investigated that the combination of Modified Direction Feature (MDF) and Multi Layer Perceptrons (MLP) network is used for character recognition. The recognition results toward training and testing data set combination are 70.22% for the lowercase characters and 80.83% for the uppercase characters respectively [1].

The difficulty of human handwriting recognition is that there are several characters having similar gestures and shapes, so that the classifier of the OCR system encounters many ambiguities while classifying some particular characters. Besides, human handwriting coming from various types of individual is complicated and reading it is not easy because it has many variations and is not consistent. Moreover, combining the lowercase and the uppercase makes more complexities and ambiguities.

However, the performance still needs an accuracy improvement if it is going to be implemented in the real system which needs high accuracy, such as reading demographic data for ID card making. The system also needs a capability to recognize all case-sensitive alphanumeric characters simultaneously. Therefore, recent days, researchers are still conducting several studies to develop an optimal system to produce a better performance so that it can be implemented in the real world problems.

Based on the rationale above the objective of this research is to improve the OCR system performance by building an optimal system construction, so it has a capability to recognize all case-sensitive alphanumeric characters and handle various human handwritings. Aside from that, this research is aimed to observe and analyze the optimal classifier training parameters in order to obtain the optimal solution of OCR system based on the real handwritings.

II. DATA COLLECTION

In this study, 62 case sensitive alphanumeric characters were recognized by OCR system, including alphabet letters and Arabic digits: ‘A’ – ‘Z’, ‘a’ – ‘z’, and ‘0’ – ‘9’. In the real condition, people handwritings had many variations and inconsistencies in gestures and shapes.

The respondents collected for this study came from various fields. They represented handwriting type of people who used to write in their daily life and also people who rarely wrote.

The respondents for collecting the data were students, lecturers, office boys, housewives, security guards, villagers, and surrounding people.

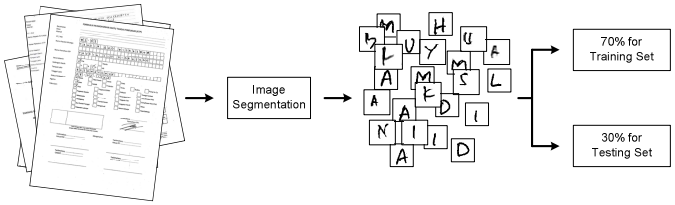

Figure 1. Data preparation scheme

In collecting the data, 3 kinds of forms were distributed to respondents to fill out with their handwriting.

1. ID card application form. This form type obtained from a village in Tasikmalaya area, West Java. This data was the original handwriting of local people who wanted to make an ID card.
2. Artificial data form. This form type contained some general academic data.
3. Alphanumeric characters form. This form contained all alphanumeric characters recognized by the system. This form was for completing alphanumeric characters data and used as training data sets.

Table 1. Data distribution of the collected characters

Character	A	B	C	D	E	F	G	H
Training (70%)	126	16	11	34	32	14	20	27
Testing (30%)	54	7	5	15	14	6	9	12
Character	I	J	K	L	M	N	O	P
Training (70%)	65	10	21	25	32	42	23	19
Testing (30%)	28	5	10	12	14	19	11	9
Character	Q	R	S	T	U	V	W	X
Training (70%)	7	34	37	30	28	7	12	7
Testing (30%)	4	15	17	13	13	4	6	3
Character	Y	Z	A	a	b	c	d	e
Training (70%)	22	10	8	18	7	14	10	11
Testing (30%)	10	5	4	8	3	6	5	5
Character	F	G	H	i	j	k	l	m
Training (70%)	10	13	11	23	7	9	12	27
Testing (30%)	5	6	5	10	3	5	6	12
Character	N	O	P	q	r	s	t	u
Training (70%)	8	23	7	7	13	11	12	11
Testing (30%)	4	11	3	3	6	5	6	5
Character	V	W	X	y	z	1	2	3
Training (70%)	7	7	7	11	7	72	44	26
Testing (30%)	3	3	3	6	4	31	20	12
Character	4	5	6	7	8	9	0	
Training (70%)	18	21	25	29	32	35	94	
Testing (30%)	9	10	12	13	15	15	41	

III. ALGORITHM

This study used Fuzzy C-Means (FCM) for data grouping mechanism, Modified Direction Feature (MDF) for feature extraction, and Nested MLP for the classification.

A. Data Grouping

A proposed technique of grouping the data into some clusters was used in this study to reach the research objective, namely to obtain an accuracy improvement of the OCR system. There were at least two reasons why grouping was used for the OCR classification problem. First, there were many ambiguities while distinguishing some characters in the 62 classes of the alphanumeric set because of their shape and gesture similarities. It made the classifier getting harder to classify all 62 classes of the alphanumeric set directly in one level. Hence, the 62 classes of alphanumeric characters needed to be divided into some groups, which the difference between each group was high.

Next, by grouping mechanism, any misclassified character would not be far from the target class. When a character was already placed in the right group, the output class of the system was around the designated character because each group was designed to be almost the same. So, it was devised to have high difference among the groups and high similarity for each character class in each group.

For this study, the character grouping was done by using Fuzzy C-Means Clustering (FCM). FCM was chosen as the clustering method in this study because it implemented fuzzy logic on its clustering process which was trained quickly trained. The clustering process was getting optimal and higher accuracy on many cases [11].

Table 2. Alphanumeric character grouping by FCM

Group	Character						
1	J	U	V	u	V		
2	B	Q	a	5	8	9	
3	N	W	w				
4	E	G	e	z			
5	D	O	α	o	0		
6	b	d	y	6			
7	A	K	X	k	X		
8	T	r					
9	S	g	s	3			
10	R	Z	q	2			
11	F	P	f	p			
12	H	M	m	n			
13	C	L	c	h	T		
14	I	i	j	l	1		
15	Y	4	7				

B. Modified Direction Feature (MDF)

MDF is a feature extraction technique for image data that is developed by combining Direction Feature (DF) technique and Transition Feature (TF) technique. It is mostly used in several OCR problems because it provides gesture and direction features of scratches in the image extracted. By using this method, it is expected that any invariant problems of character handwritings can be helped.

The algorithm of this method is simply explained as follow.

1. Changing matrix into direction values based on a certain direction guide [14].

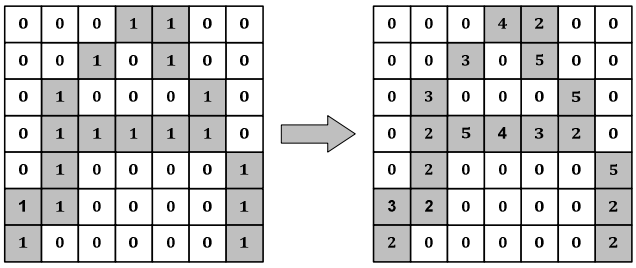

Figure 2. Matrix changing from binary values into direction values

2. Obtaining LT and DT matrices from 4 directions by scanning each row and column from left to right, right to left, top to bottom, and bottom to top [4, 2, 1].
3. LT and DT matrices normalization to ensure that the LT and DT dimensions were normalized in size by re-sampling those matrices [1].

C. The Nested MLP

Since the grouping method was implemented in this study, the satisfied classifier model used to perform the character recognition was a kind of Nested Multi Layer Perceptrons (MLP). Basically, it is a development of general MLP, which the general classifier model is shown in the figure below.

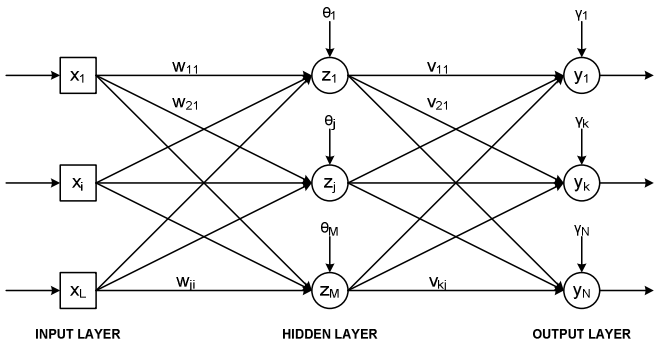

Figure 3. General classifier model of MLP network

The architecture of MLP above has one layer of hidden neurons. It has L number of input neurons (x), M number of hidden neurons (z), and N number of output neurons (y). L is determined by the dimension of the input data and N is determined by the number of designated classes. Each neuron in a layer toward the next layer has a connection with particular weight (w and v). There are also bias weights for hidden and output neurons with corresponding numbers (θ and γ). Meanwhile, M belongs to adjustable parameter that needs to be observed because it used to be determined by trial and error. At this time, there is no certain formula to find the optimal number of hidden neurons exactly. However, a formula below can be applied to approximate the number of hidden neuron [8].

$$M \approx \sqrt{LN} \quad (1)$$

The classifier model is trained in the training mechanism using a certain algorithm. The most popular algorithm to train an MLP network is Back Propagation (BP). It has 2 phases of the training process: the forward propagation to obtain the network outputs and the error between actual outputs and the target outputs, and the backward propagation to back-propagate the error factor for each neuron connection so that the weights are updated and used for next iterations [8].

In the Nested MLP for this case, there will be the grouping network, which is the first network that will be processed, and the sub-grouping network, which is placed inside the each group and will be processed after obtaining the designated group resulted by the grouping network. The sub-grouping network finally determines the character class of the input character image.

In this study, to classify 62 classes of alphanumeric characters correctly, a grouping network with 15 output classes and 15 sub-grouping networks with particular number of output classes in each sub-group were built. Figure 4 shows the architecture of the Nested MLP classifier model used in this study.

D. Training and Testing The Classifier Model

The OCR system in this study used Nested MLP classifier model that needed 2 phases of the training process. The first phase trained sub-grouping networks using training set of particular characters in the corresponding group. The accuracy rate in each group was calculated directly, so that this first phase yielded 15 accuracy rates. After training phase of all sub-grouping networks had been done, it was continued by the second phase to train the grouping network using training set of all alphanumeric characters. This second phase calculated the accuracy rate in placing the correct group for each character in the training set.

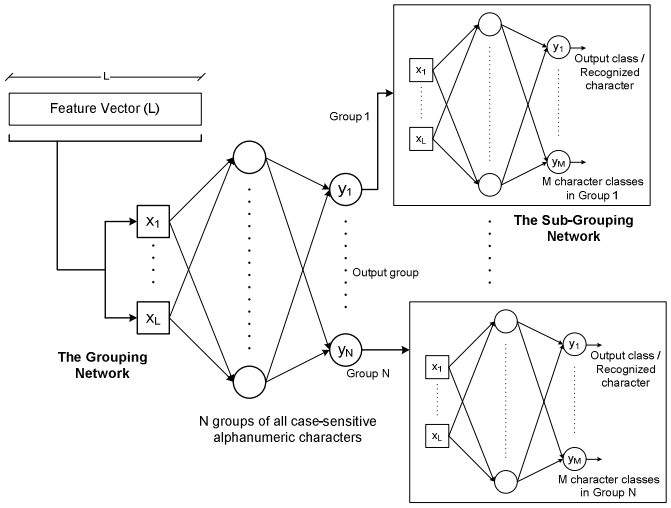

Figure 4. Architecture of The Nested MLP classifier model

After passing the training process of the Nested MLP classifier model, the testing process was continuing. It was done by using the testing set consisting of all alphanumeric characters. Each character was classified by the grouping network first to determine which group the character was classified. Once the character had been placed in a group, it was continued by the corresponding sub-grouping network to classify the class of the character. From this mechanism, two performance parameters of the testing process were obtained including the group accuracy rate and the class accuracy rate.

IV. Result and Evaluation

The all case-sensitive alphanumeric characters had been collected in this study were identified to have some characteristics as follow.

1. Some characters having different gestures and shapes, such as A, B, C, I, etc.
2. Some characters having near gestures and shapes similarities, such as Q, O, K, k, etc.
3. Some characters having exact similarities in gestures and shapes, such as O, o, X, x, U, u, I, l, 1, etc.

The testing results were divided into two phases. The first phase analyzed the results of the sub-grouping networks for independent evaluation, and the second one analyzed the results of the Nested MLP network for overall evaluation. In addition, the results were shown separately between the uppercase, the lowercase, and the digit, for all training and testing data set.

A. The sub-grouping networks testing results using training and testing set

The following table shows individual performance of each sub-grouping network while being tested toward training and testing data set. The performance indicates the capability of each sub-grouping network in classifying the data entered into the corresponding group, if those data is already classified in the correct group. For example, the testing by using training data set in the group 1, since the accuracy rate shown is 100%, the sub-grouping network can correctly classify all data entered into group 1, if those data are already correctly placed in group 1. Meanwhile, for the testing by using testing data set, the network has accuracy rate of 56.67% in classifying all data

entered into group 1, when those data are already placed correctly in group 1. The table also shows the characters contained in each corresponding group.

Table 3. Testing result of the sub-grouping networks

Network	Characters	Using Training Data Set	Using Testing Data Set
Sub-Grouping Network 1	**JUVuv**	**100.00**	**56.67**
Sub-Grouping Network 2	BQa589	100.00	89.09
Sub-Grouping Network 3	NWw	100.00	78.57
Sub-Grouping Network 4	EGez	100.00	81.25
Sub-Grouping Network 5	**DOao0**	**77.08**	**58.14**
Sub-Grouping Network 6	bdy6	100.00	100.00
Sub-Grouping Network 7	AKXkx	100.00	86.84
Sub-Grouping Network 8	Tr	100.00	100.00
Sub-Grouping Network 9	Sgs3	98.86	75.00
Sub-Grouping Network 10	RZq2	100.00	86.05
Sub-Grouping Network 11	FPfp	100.00	69.57
Sub-Grouping Network 12	HMmn	100.00	92.86
Sub-Grouping Network 13	CLcht	98.65	82.35
Sub-Grouping Network 14	**Iijl1**	**76.40**	**53.85**
Sub-Grouping Network 15	Y47	100.00	96.88

Table 3 also shows that 3 lowest network performances are given by the sub-grouping network in group 14, 1, and 5. It is reasonable since the similarity levels of the characters in the groups are very high among other groups. In group 14, there are several characters having near similarities and exact similarities, such as I, i, l, and 1. The sub-grouping networks in the group 1 and 5 yields a poor performance too. In group 1, there are characters of U, u, V, and V, while in group 5, there are characters of O, o, and 0. All of those characters belong to the characters having near similarities and exact similarities.

From the results, the most missing classes yielded by the system are caused by the system that cannot handle ambiguities in differentiating uppercase and lowercase letters having exact similarities in gestures and shapes. The problems can be taken in hand at least by using these two ways.

1. Improve the image preprocessing by dividing the image character into some quadrants.
2. Combine some characters having the exact similarities between the uppercase and the lowercase into one character class.

B. *The Nested MLP network testing results using training and testing set*

The following table shows performances of the Nested MLP network while it was being tested by using training and testing data set. The total data number in the training set was 1417 characters, while the total data number in the testing set was 644 characters, including all case-sensitive alphanumeric characters. By the testing process using training data set, it is seen that the Nested MLP network classifier is able to classify the recognized characters as the correct groups with accuracy rate of 94.78%. Meanwhile for overall evaluation, the classifier could classify the recognized characters correctly with the accuracy rate of 83.39%. It means that most of characters in the training set are able to be classified into correct groups and correct character classes.

Table 4. Testing results of the Nested MLP network using training and testing data set

Accuracy Rate	Using Training Data Set	Using Testing Data Set
Grouping Accuracy	94.78	83.39
Character Accuracy	88.57	64.60

Whereas for the testing process using testing data set, the performance is getting worse because it only yields the accuracy rate of 88.57% for classifying the data into correct groups and the accuracy rate of 64.60% for classifying the data into correct character classes. It is reasonable to be happened while seeing that the most of misclassified characters have near similarities and exact similarities in gestures and shapes for the uppercase and the lowercase letters. For the reasons, the system is said to be having sufficient capabilities to learn and recognize the all case-sensitive alphanumeric characters.

V. Conclusion and Future Work

A. *Conclusions*

The implementation of MDF and Nested MLP network was able to improve the OCR system performance based on the accuracy rates, which were 84.38% of the uppercase, 76.43% of the lowercase, and 78.92% of the digit. It also improved the capability in recognizing all case-sensitive alphanumeric characters simultaneously.

Most of misclassified characters when distinguishing several uppercase and lowercase characters having near and exact similarities in gestures and shapes are reasonable because those characters were still ambiguous to be distinguished by human perception. The system could be improved toward those problems by optimizing the data selection, the grouping process, and the character categorization.

B. *Future Works*

The recognition performance depends on the result of grouping mechanism. Any other clustering methods can be used for the grouping mechanism. But, it is strongly recommended that the grouping mechanism still involves the human perceptions. The better the characters being grouped, the better recognition performance made.

In recognizing all case-sensitive alphanumeric characters, the most misclassified characters are caused by several characters having near and exact similarities in gestures and shapes. It might be solved by improving the image preprocessing by dividing the handwriting character image into some quadrants. The alternative way is to combine some characters having the exact similarities between the uppercase and the lowercase into one character class. In addition, the study might be developed by using various level of the Nested MLP network to reach the optimal solution. The system capability might be improved also by doing the image acquisition process automatically. However, these still need further study to be conducted.

Acknowledgment

I would like to acknowledge the sincere guidance and help of Mr. Deni Saepudin, Dr. and Mr. Adiwijaya, Dr. I am a very lucky person to have you as my supervisors. Thank you so much for accepting me with the various drawbacks. I hope I can still learn many things together with you.

Then, I would like to acknowledge the inspirational instruction of Mr. Suyanto, S.T., M.Sc. Thank you very much for your sincerity of being my 'third supervisor' and guiding me to solve many problems I have. I want to be a man whom many people love to like you.

References

[1] Blumenstein M., X.Y. Liu, and B. Verma. A Modified Direction Feature for Cursive Character Recognition. International Joint Conference on Neural Networks, Budapest, Hungary, pp. 2983-2987, 2004.

[2] Blumenstein M., X.Y. Liu, and B. Verma. An Investigation of the Modified Direction Feature for Cursive Character Recognition.

[3] ESCAP. Guidelines on the Application of New Technology to Population Data Collection and Capture, Bangkok, pp. 52-62, 2001.

[4] Gader, P.D., M. Mohamed, and J.H. Chiang. Handwritten Word Recognition with Character and Inter-Character Neural Network. IEEE Transaction on Systems, Man, and Cybernetics-Part B: Cybernetics, vol. 27, pp. 158-164, 1997.

[5] Mamedov, Fakhraddin and Jamal Fathi Abu Hasna. Character Recognition Using Neural Networks. Near East University, North Cyprus, Turkey via Mersin-10, KKTC.

[6] Pradeep J., E. Srinivasan, and S. Himavathi. Diagonal Feature Extraction Based Handwritten Character System Using Neural Network. International Journal of Computer Application, vol. 8, no. 9, pp. 17-22, 2010.

[7] Sarkar, Deepayan. 2003. Optical Character Recognition using Neural Networks. Department of Statistics University of Wisconsin, Madison.

[8] Suyanto. 2007. Artificial Intelligence: Searching, Reasoning, Planning, and Learning. Bandung: Penerbit Informatika.

[9] Suyanto. 2008. Soft Computing, Membangung Mesin Ber-IQ Tinggi. Bandung: Penerbit Informatika.

[10] Trier, Øivind Due, Anil K. Jain, and Torfinn Taxt. Feature Extraction Methods For Character Recognition – A Survey. Pattern Recognition, vol. 29, no. 4, pp. 641-662, 1996.

[11] Unknown. "A Tutorial on Clustering Algorithms: Fuzzy C-Means Clustering". <http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/cmeans.html>

[12] Umami, Muhammad Amrul. 2010. Analysis and Implementation of Arabic Letter Recognition by Using Modified Direction Feature Extraction and Learning Vector Quantization. Bandung: IT Telkom.

[13] Widara, Primalita, Adiwijaya, and Tjokorda Agung B.W. 2009. Optical Character Recognition (OCR) Using Modified Direction Feature Extraction and Kohonen Self Organizing Maps. Bandung: IT Telkom.

[14] Wirayuda, Tjokorda Agung B.W., I Gede Rudy H., and Retno Novi D. Balinese Character Recognition Using MDF and LVQ. Konferensi Nasional Sistem Informatika, pp. 7-12, 2009.