MJIT 2017 Malaysian Journal of Industrial Technology

RECOGNITION OF HANDWRITTEN ARABIC DIGITS IN WORST-CASE SCENARIOS

R. F. Raja Suleiman[1, a], and N. Md Said [1,b]

*Corresponding author :
fazliza@unikl.edu.my

[1] Industrial Automation Section, Universiti Kuala Lumpur Malaysia France Institute, Section 14, Jalan Teras Jernang, 43650 Bandar Baru Bangi, Selangor Darul Ehsan

[a]fazliza@unikl.edu.my, [b]norzanah@unikl.edu.my

Abstract

In general, character recognition focuses on average recognition performances. In this study, we aim to maximize the probability of correct classification of handwritten Arabic digits in worst-case scenarios. Here, a worst-case scenario refers to a digit that is poorly written compared to the typical form of its category. Besides focusing on the worst-case rate, this paper also highlights the recognition of Arabic digits which are less explored in the literature in contrast to the Latin or Chinese digits. For these experiments: first, we will build minimax dictionaries of Arabic digits from the training dataset obtained from the MADBase (Modified Arabic Digits dataBase). For comparison purposes, we also train SVD (singular value decomposition) dictionaries from the same database. Each digit represents a class, thus we have 10 classes of training examples and test examples. Then, using these learned dictionaries (minimax and SVD), we evaluate the recognition rate in worst-case scenarios. The experiments are executed 100 times to allow random permutation between the samples. Results show that in most cases, the minimax approach performs better in recognizing the poorly handwritten Arabic digits.

Keywords: Pattern recognition; Worst-case classification; Minimax; Arabic characters;

1.0 INTRODUCTION

Pattern recognition relies on training the available samples (i.e., learning the known database) to construct a model (also called a *dictionary*) based on specified criterion, and then "make sense" of the inputs in order to generate the output(s). The training process can be divided in two categories: supervised learning and unsupervised learning. The first type of learning is where the known data are *labeled* with the corresponding output. The latter category is the case when the data are *unlabeled,* thus the learning algorithm needs to recognize the input's pattern to generate the corresponding output(s). This paper deals with unsupervised learning where we train dictionaries of Arabic handwritten digits w.r.t. *minimax* criterion. The performance is then evaluated in term of worst-case classification rate.

For the readers who are unfamiliar with some of the machine learning terms, a known database is called a library of training samples, say $\boldsymbol{L} \in \mathbb{R}^{MxN}$. From this training library, we learn or construct a dictionary $\boldsymbol{D} \in \mathbb{R}^{MxK}$ where $\boldsymbol{D} = [\boldsymbol{d}_1, \ldots, \boldsymbol{d}_K]$ (usually $K < M$). Each column vector $\boldsymbol{d}_i$ in the learned dictionary is called an atom. Thus, for a single column dictionary, we have one learned atom $\boldsymbol{d}$. Another set of samples is called the test samples. This is the input data that we have to recognize. The learned dictionary will be used to classify the input data into *K* classes w.r.t. a specified criterion.

The next Section 2 discusses some backgrounds on character and digits recognition (subsection 2(a)), and highlights two different classification and learning criterion: the usual average criterion (minimum error) versus the *minimax* criterion (minimize the maximum error)(subsection 2(b)). In the third part of this paper (Section 3), we will explain the conducted experiments and present the simulation results obtained for the minimax approach and the SVD (Singular Value Decomposition) method. Results show that in most cases, the minimax approach performs better in recognizing the poorly handwritten Arabic digits. The final section (Section 4) sums up the finding of this research work.

2.0 RESEARCH BACKGROUND

Before moving into details of the conducted experiments, we present some research backgrounds related to our works.

Character and Digits Recognition

Often in literature, we found a large number of research works done on the recognition of Latin and Chinese characters and digits. These types of recognitions are applicable in finance domain to read cheques, in mail service to sort letters and in smartphones to read the user's handwriting.

For instance, a well-known database of handwritten Latin digits is the MNIST database introduced by LeCun *et al.* in 1998 [1]. This database consists of 60000 training samples and 10000 test samples of handwritten digits from 0 to 9, collected from nearly 250 writers. Each image of the digit has dimension of 28x28 pixels, organized row-wise and the pixel values ranging from 0 (white background) to 255 (black foreground). Various learning and classification algorithms have been tested using the MNIST database. One of the classification method that focused on deep convolutional neural networks, has successfully obtained a very low error rate: 0.23% [2].

An example of widely used handwritten characters and texts is the CASIA (Institute of Automation of Chinese Academy of Sciences) database [3]. There are a large number of Chinese characters, thus the corresponding database is huge: CASIA comprises of 3.9 million samples that represent 7185 Chinese characters and 171 symbols recorded from 1020 writers. During the 2011's Chinese handwriting recognition competition, it was reported [4] that the highest correct classification rate was 94.77% for offline character recognition presented by Fujitsu.

In our work, we choose to test two recognition algorithms on the Arabic handwritten digits database: MADBase (Modified Arabic Digits database) [5], which has similar format to the MNIST database. In general, the recognition of Arabic digits and characters is more challenging compared to Latin's because they are more cursive. They are also less studied in machine learning compared to other characters such as Latin and Chinese. The MADBase is constructed from 700 writers from different institutes. Each writer wrote the same digit 10 times, for a total of 100 digits: ranging from zero (٠ : *sifr*) to nine (٩: *tiss'a*). For our work, we choose smaller samples: 2012 training samples from 212 writers and 1000 test samples from 100 writers. This allows a faster computations since we are working on a standard laptop (64 bits, Intel Core i7-4510U CPU @ 2 GHz). The list of Arabic digits is shown in Table 1.

Average versus Minimax (Worst-case) Criterion

Minimaxity here refers to minimizing the maximum classification error rate, or equivalently the *maximin*: maximizing the probability of correct classification in worst-case scenarios (i.e., worst-case is when a digit is poorly handwritten compared to the usual form). This is an application of previous work presented in [6], where for a single atom dictionary $\boldsymbol{d}$, (K=1) the problem can be written as

$$\begin{aligned} \boldsymbol{d}^* &= \arg \max_{\boldsymbol{d}:||\boldsymbol{d}||_2=1} \min_{i=1,\dots,N} (\boldsymbol{d}^T \boldsymbol{l}_i)^2 \\ &= \arg \max_{\boldsymbol{d}:||\boldsymbol{d}||_2=1} \min_{i=1,\dots,N} |\boldsymbol{d}^T \boldsymbol{l}_i| \end{aligned} \quad (1)$$

where (^{T}) is the transpose of the vector (or matrix) and $\boldsymbol{l}_i$, i ={1, …,N} are the known samples from library $\boldsymbol{L} \in \mathbb{R}^{MxN}$. Under some conditions, the exact solution of $\boldsymbol{d}^*$ can be obtained by solving a convex optimization problem in the form of quadratic programming:

$$\begin{aligned} \boldsymbol{d}^* = &\ \textit{minimize } -t \\ &\ \textit{subject to } t - \boldsymbol{d}^T \boldsymbol{l}_i \le 0,\ i = \{1, \dots, N\} \\ &\ ||\boldsymbol{d}||_2 \le 1 \end{aligned} \quad (2)$$

where we assume that all samples $\boldsymbol{l}_i$, i={1, …,N} are in the positive orthant ($\boldsymbol{l}_i \in \mathbb{R}_+^N$), hence $\boldsymbol{d}^* \in \mathbb{R}_+^N$. Fig. 1 shows the illustration of this method.

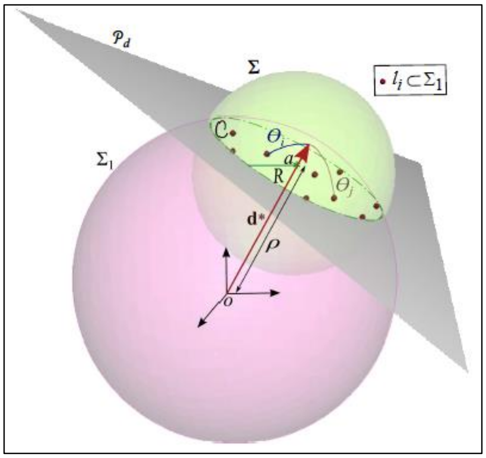


Fig. 1. Ilustration of the minimax critierion. ***d*** * is the minimax vector, *C* is the smallest circle containing all samples $\boldsymbol{l}_i$, Σ_1 is the unit-norm spherical boundary, Σ is a boundary sphere and P_d is an optimal plane.

On the other hand, average criterion refers to minimizing the overall classification error rate (Mean Square Error criterion):

$$\min_{\hat{L}} ||\boldsymbol{L} - \hat{\boldsymbol{L}}||_F^2 \qquad (3)$$

$\hat{\boldsymbol{L}}$ is the approximation of $\boldsymbol{L} \in \mathbb{R}^{MxN}$, where in our case: $\hat{\boldsymbol{L}} = \boldsymbol{DY}$, and ***D*** is the dictionary that we have to learn, ***Y*** is an unknown representation matrix. $||\,.\,||_F$ denotes the Frobenius norm. This type of criterion tends to represent the "average behaviour" or patterns of the samples ***L***. According to the Eckart-Young theorem [7], problem in (Eq. 3) can be solved analytically via the truncated Singular Value Decomposition (SVD) of ***L***, subject to rank($\hat{\boldsymbol{L}}$) ≤ *K* :

$$\hat{\boldsymbol{L}} = \boldsymbol{U}_K\ \boldsymbol{\Sigma}_K\ \boldsymbol{V}_K^T \qquad (4)$$

where $\boldsymbol{U}_K \in \mathbb{R}^{MxK}$ is the truncated (in *K*) left singular matrix, which is often used to represent the library ***L*** in lower dimension. $\boldsymbol{\Sigma}_K \in \mathbb{R}^{KxK}$ is a diagonal matrix containing the ***K*** largest singular value of ***L***, and $\boldsymbol{V}_K \in \mathbb{R}^{NxK}$ is the truncated right singular matrix. $\boldsymbol{\Sigma}_K\boldsymbol{V}^T$ is a row-wise sparse matrix. If we identify (Eq. 4) with the approximation $\hat{\boldsymbol{L}} = \boldsymbol{DY}$, the learned dictionary ***D*** = $\boldsymbol{U}_K$, and the unknown representation matrix ***Y*** = $\boldsymbol{\Sigma}_K\boldsymbol{V}_K^T$.

Comparing both of these criterion, minimax is useful when we do not want to miss recognizing or detecting an item, but at the cost of affecting the average performance. An example is for the detection of gas leakage at a power plant. In this case, we want to reassure that we are able to detect the minimum leakage where the value may not reach the average leakage threshold.

While the average criterion has been shown in the literature to be effective where the objective is to minimize the overall error rate, but at the cost of negliging the item that has lower value than the average threshold. An example is when we want to detect spectral lines. The lines that have a "common" shapes will be detected, but the line that is a bit distorted in shape, but nevertheless important will not be detected using this criterion.

3.0 RESULT AND DISCUSSION

In the first part of this Section, we explain the executed simulations. In the second part, we show the results and discuss the findings.

Worst-case Recognition

The objective of this paper is to learn dictionaries (from the training samples of the database) that maximize the probability of correct classification in the worst-case of the test samples. For example: a dictionary is constructed from training samples of class 1 in order to be robust for all instances of handwritten digits ١ of the testing samples.

Say we name the library (database) of training samples as ***L***. For each digit *c*, the corresponding library is: $\boldsymbol{L}^c$ which consists of 212 samples of handwritten digit *c*. Thus, for all digits: *c* = ٠, ١, …, ٩ ; we have ten libraries: $\boldsymbol{L}^{٠}$, $\boldsymbol{L}^{١}$, …, $\boldsymbol{L}^{٩}$. For each library $\boldsymbol{L}^c$, we learned the corresponding minimax atom $\boldsymbol{d}^{*c}$ by using (Eq. 2) [6]. For all ten Arabic digits, we then have ten minimax atoms, and we concatenate them into a matrix ***M**** = [$\boldsymbol{d}^{*٠}$, …, $\boldsymbol{d}^{*٩}$]. We performed the same procedure for the SVD method ($\boldsymbol{U}_K$ from Eq. 4), resulting in a matrix $\boldsymbol{M}^{SVD}$. Figure 1 illustrates the learned atoms (digits) for both of these approaches.

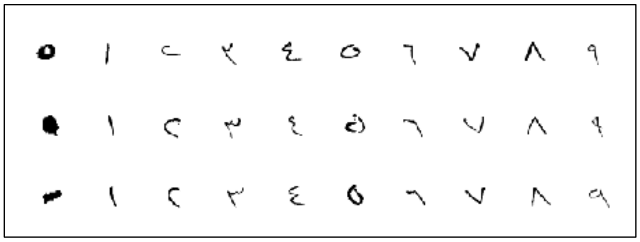
To evaluate the worst-case recognition rate, say for the minimax approach, we identify for each class *c* which of the 100 test samples has the *minimum correlation* with ***M**** = [$\boldsymbol{d}^{*٠}$, …, $\boldsymbol{d}^{*٩}$]. This identified test sample is denoted as $\boldsymbol{l}_i^{*c}$. Correct classification for this worst-case (poorly handwritten) digit sample if :

$$\arg\ \max_{j=٠,\dots,٩} \boldsymbol{d}^{*c}\ \ \boldsymbol{l}_i^{*c} = c \qquad (5)$$

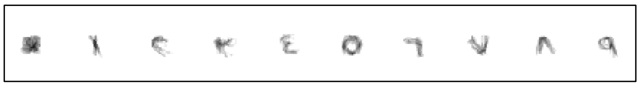
The same evaluation (Eq. 5) goes for the SVD approach, by using the SVD learned matrix $\boldsymbol{M}^{SVD}$.

Simulation Results

For both approach, the experiments were executed 100 times to allow random permutation between the test samples and the training samples. Fig. 2 depicts the samples of the database and the learned digits. Table 1 summarizes the results.



*(a) Some samples of the handwritten digits in the library **L** (from MADBase)*



(b) The learned minimax digits, $\mathbf{M}^* = [\mathbf{d}^{*\text{٠}}, \ldots, \mathbf{d}^{*\text{٩}}]$

(c) The learned SVD digits, $\mathbf{M}^{SVD} = [\mathbf{u}^{\text{٠}}, \ldots, \mathbf{u}^{\text{٩}}]$

Fig. 2. (a) Some digits samples from the database (library) are shown. (b) and (c) depict the learned atoms (digits) for K=1 (single column dictionary representing each digit), for minimax and SVD methods. We can see that the minimax approach (b) take into accounts the dissimilar forms of the digit (e.g., ٠,١), while the SVD approach represents "average, usual and smooth" forms.

Table 1. The list of Arabic digits and the simulation results of worst-case recognition rate obtained from 100 executions for Minimax and SVD methods.

Arabic Digits	Worst-case recognition rate (%)	
	Minimax	**SVD**
٠	*61*	48
١	*100*	6
٢	19	23
٣	30	39
٤	64	94
٥	89	24
٦	69	62
٧	66	62
٨	*84*	58
٩	83	96

We can see from Table 1 that Minimax approach performs better than SVD in 6 cases. These 6 cases are when the worst handwritten digits have dissimilar form from the usual form, e.g.,: for digit ١, minimax recognizes the worst handwriting 100% of the time, while SVD miss most of it, and only recognizes 6% over the 100 experiments.

For other cases, SVD performs better because the poorly handwritten digits (٢ , ٣ , ٤ , and ٩) have forms that are similar to the usual pattern (average Arabic digit form).

These experiments show that by using the minimax criterion, the worst handwriting can be classified correctly compared to the SVD method.

4.0 CONCLUSION

Throughout this paper, we have discussed on some handwritten characters and digits databases, the highlight on Arabic digits, and the recognition in worst-case scenarios compared to the average recognition. In usual case, the performance of a classification algorithm is evaluated based on an "average" criterion where the classification error rate is minimized. In this paper, we focused on "minimax" criterion where the objective is to classify the handwritten digits by minimizing the maximum error rate. In other words: we classify the digits in a worst-case scenario where the active input digit is written poorly w.r.t. the standard form. Results show that the minimax approach performs better in most cases compared to the SVD approach.

For future works, we can improve the algorithm to suits properly the fundamental of machine learning, and achieved best results for all cases (digits ٠ to ٩).

References

[1] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), November 1998, pp. 2278-2324.
The MNIST database can be downloaded from: http://yann.lecun.com/exdb/mnist/

[2] J. Schmidhuber. Multi-column deep neural networks for image classification. *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (*CVPR). IEEE Computer Society, Washington, DC, USA, 2012, pp. 3642-3649.

[3] C. L. Liu, F. Yin, D. H. Wang and Q. F. Wang. CASIA Online and Offline Chinese Handwriting Databases. In *International Conference on Document Analysis and Recognition (ICDAR)*, Beijing, 2011, pp. 37-41.
The CASIA database can be downloaded from: http://www.nlpr.ia.ac.cn/databases/handwriting/Home.html

[4] C. L. Liu, F. Yin, Q. F. Wang and D. H. Wang. Chinese Handwriting Recognition Competition. In *International Conference on Document Analysis and Recognition (ICDAR)*, Beijing, 2011, pp. 1464-1469.

[5] E. EL-Sherif and S. Abdleazeem. A Two-Stage System for Arabic Handwritten Digit Recognition Tested on A New Large Database. In International Conference on Artificial Intelligence and Pattern Recognition, 2007, pp. 237–242.
The MADBase can be downloaded from http://datacenter.aucegypt.edu/shazeem/

[6] Raja Fazliza Raja Suleiman, David Mary and André Ferrari. Subspace learning in minimax detection. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, 2014, pp. 3062–3066.

[7] Carl Eckart and Gale Young. The approximation of one matrix by another of lower rank. Psychometrika, vol. 1, no. 3, 1936, pp. 211–218.