Recognizing Arabic Handwriting Using Statistical Hierarchical Architecture

Ramadhan A. M. Alsaidi[1†], Ayed R.A. Alanzi [2††], Saleh R. A. Alenazi[3†††] and Taufiq H. Ghilan [4††††]

Department of Mathematics, Jouf University, Gurayat 77454, Saudi Arabia
Department of Mathematics, Majmaah University, Majmaah 11952, Saudi Arabia
Computer Technology department, Tabuk College of Technical, Tabuk, Saudi Arabia
Researcher at University Utara Malaysia, UUM, towfeeq2k5@yahoo.com

Summary
Among Artificial Intelligence research, feature extraction of Arabic handwriting is still an important topic and interesting challenge. This work is based on Smale's framework which is proposed in his paper "Mathematics of the Neural Response". According to his proposal, hierarchical architectures can be seen as two sequential and different steps or stages: creating templates and measurements of similarities respectively. The main task of this work is to enhance the accuracy of recognizing Arabic handwriting using Smale's framework. To achieve this goal, this paper presents a statistical developed hierarchical method to improve the extraction of features which have been proved to be more effective in the stated research of Arabic hand-writing recognition. This method introduces one criterion for the selection of informative template based on the arithmetic mean. On the other hand, it considers the Square Pearson Correlation Coefficient (SPCC) technique as a similarity measurement.
Key words:
arithmetic mean – derived kernel – neural response – template - Squared Pearson correlation Coefficient - IfN/ENIT databases.

1. Introduction

When we look for numbers of layers which are related to computation , there are two kinds of single layer such as; splines[1], support vector machines [2] and regularization networks[3]. On the other hand, multi-layer architecture where each layer of the architecture is performing similar computations consisting of the creating template and measuring similarity operations. Smale et al. introduced a theoretical of hierarchical architecture framework in [4]. There are many researches proposed to improve the achievement of it such that Jake et al.[5] suggested a generalized representation of derived kernel and neural response. In [6] Ramadhan et al. combined Sparse representation with hierarchical architectures to build hierarchical sparse method (HSM). For image classification a new study a sparse-based neural response feature extraction technique was proposed[7]. In other studies, to enhance the effectiveness of a choosing template, there are many attempts appeared such that , by using k-mean[8, 9] and entropy[10]. Recently, new work was suggested in [11] which combined an entropy method which uses as a criteria to select informative template with Squared Pearson correlation Coefficient (SPCC) technique which adapts to similarity measurement in order to improve the performance of the hierarchical method. On the other hand, there are many researches that have been presented for recognizing Arabic handwriting. A few of these works are for isolated character recognition, while most for whole word recognition. For instance, an off-line character recognition was improved by Ab-uhabiba. et al. [12]. In that work an algorithm was introduced this based on skeletons that reflect the structural relationships of the letter components. In other works[13], the researchers take out moment-based features in order to recognize handwritten Arabic character. Using structural systems to identify off-line Arabic handwriting was presented by Mohammad et al.[14]. Author's in [15] presented a novel and simple system to recognize off-line Arabic handwriting characters through three steps creating database, segmentation and similarity. With the same goal but different data and method, Ramadhan et al. [8] presented new way based on hierarchical procedures. For same data, a new method that used shape-based alphabet for handwriting Arabic recognition showed in[16]. The [17] paper suggested a system that reached high accuracy using effective segmentation, feature extraction, and recurrent neural network (RNN) respectively.
A new approach which dealt with offline recognition cursive Arabic handwritten text was presented by [18] which built on Hidden Markov Models (HMMs) .
Our approach is based on Smale's hierarchical framework with some modifications that he has obtained templates by randomly extracting image patches from an image set segmentation but this work suggests a new criterion based on arithmetic mean to select an effective template; however, Smale's utilized a dot product as similarity measure. Consequently, here Squared Pearson correlation Coefficient (SPCC) technique was considered as similarity measures for more accuracy.
The rest of this paper is structured as follows. In Section 2, some basic terminologies and concepts will be introduced and reviewed to be used throughout the paper. Choosing the effective templates technique based on

Manuscript received August 5, 2020
Manuscript revised August 20, 2020

arithmetic mean is presented in section 3. Section 4 presents Squared Pearson correlation Coefficient (SPCC) techniques as similarity measures. The proposed method to recognize Arabic handwriting data and discussion the results is presented in Section 5. Finally, in Section 6 the conclusions and future work are discussed.

2. Basic Concepts and Terminology

2.1 Arithmetic mean (AM) of images

The arithmetic mean criterion is defined as the average of all pixels within a local region of an image.
The definition can be elaborated as follows:

$$\bar{x} = \frac{1}{n}(x_1 + \cdots + x_n)$$

Since this work deals with statistical concepts, the full image is imagined as a population and the templates as samples. According to the statistical hypothesis testing approach, there is a relationship between AM of samples and AM of population. We can benefit from such an approach by experimentally designing a threshold that allows us to choose some informative templates. We encourage the reader for a more detailed study to consult [19, 20].

2.2 Squared Pearson correlation Coefficient (SPCC)

Pearson correlation is a commonly used for the similarity between different brain states [21-27]. For more efficient authors in[11]used the Squared Pearson Correlation Coefficient (SPCC) as a similarity measures like us in this research.
The mathematical expression for (PCC) is

$$r = \frac{\sum(x_i - x_m)(y_i - y_m)}{\sqrt{(x_i - x_m)^2}\sqrt{(y_i - y_m)^2}}$$

Where x_i is the strength of the pixel in image 1,

And y_i is the strength of the pixel in image 2,

While x_m is the mean intensity of image 1,

and y_m is the mean intensity of image 2,

As we mention above ,this paper utilizes squared Pearson correlation coefficient (SPCC) which equals $\sqrt{r^2}$ [11]

2.3 Framework of Hierarchical Architectures (FHA)

Since our study is based on the FHA which was introduced by Smale et al, the main elements of this method can be mentioned as follows:

1) Ground works.

The components needed to state the FHA are given:

a) Limit architecture defined by nested layers i.e. the case of architecture in this work formed by three layers u, v and sq such that

$u \subset v \subset sq$.

b) Between adjacent layer there were transformation function as:

$$u \xrightarrow{H_u} v$$

$$v \xrightarrow{H_v} sq$$

c) On each layer we defined Function spaces.

$$F_u = \{f : u \longrightarrow [0,1]$$

$$F_v = \{f : v \longrightarrow [0,1]\}$$

$$F_{sq} = \{f : sq \longrightarrow [0,1]\}$$

d) Connecting a real world setting to the mathematical model by templates

$$T_u \subset T_v \subset T_{sq}$$

.

2) Similarity measure

When the previous objects are ready, we turn to the next step which is sometimes called derived kernel.
We will choose the SPCC on u as, i.e.

$$R_u(x,y) = \frac{\sum(x_i - x_m)(y_i - y_m)}{\sqrt{(x_i - x_m)^2}\sqrt{(y_i - y_m)^2}}$$

The normalized kernel $R_u(x,y)$ can be found by

$$\tilde{R}_u(x,y) = \sqrt{R^2}$$

3) Neural response

After that we compute the neural response of first layer as

$$N_v(x)(y) = \max_{h \in H_u} \tilde{R}_u(x \circ h, y))$$

$$where\ x \in I_v,\ \ y \in I_u, and\ h \in H_u\ i.e.\ (x \circ h) \in I_u$$

e) According to the hierarchical architecture view, this work uses a recursive definition of the derived kernel and associated neural response at every layer. For more detail see[4-6, 9].

3. Contribution of this work

The key semantic component in the SHA method is a template because it is always extracted from data. On the other side, the ability of recognizing patterns relies on a hypothesis that classes occupy different regions in the feature space. According to that, we need to select an effective template to satisfy discriminative information and small numbers to enhance the efficiency and recognition accuracy. For this goal, this paper introduces a template selection method based on AM.

It is worth noting that they are two kinds of templates:

The first layer templates: which has size u (i.e. It is reasonably small since it covers only the basic elements of the images.) $\mathrm{T_u} \subset \mathrm{I_u}$.

The second layer templates: which has size v (i.e. It is large enough to include approximately full (almost) image where more discriminative structure.) $\mathrm{T_v} \subset \mathrm{I_v}$.

It is worth noting that the best size of the template is determined experimentally.

The suggested template selection technique is offered as follows

First of all, let us denote by $\mathrm{T_u}, \mathrm{T_v}$ the initial template set of layers u and v respectively which arbitrarily take out sub image patches from an image set of size u and v. The target of SHA method was to choose the informative template sets S_u and S_v through eliminate the redundancy of initial template sets Tu and Tv such that

$S_u \subset \mathrm{T_u}$ and $S_v \subset \mathrm{T_v}$

For the threshold, it designs experimentally based on idea that the mean of sub-image (i.e. sample) is closer to the mean of an image patch (i.e. population).

a) Compute the AM of an image patch, i.e., $\mu = mean(I)$
b) Randomly create sub-image (i.e. initial template (T_1^u)) from image patch
c) Compute the AM of an initial template, i.e., $\bar{x} = mean(T_1^u)$
d) Test the threshold. If it is violated return to step b)
   Else, add T_1^u to S_u and then return to step b.
e) Continue do these iterations until reach a certain tolerance.
f) The iteration is terminated.

Algorithm 1 choosing of S_u

1. Input: Full image (I) and number of template d
2. Output: S_u
3. count=1
4. $S_u = \emptyset, m = \emptyset$
5. μ=mean (I)
6. While count < d
7. create t_u^1 (randomly)
8. x=mean (t_u^1)
9. if x~ μ
10. S_u=S_u∪t_u^1, count=count+1.
11. m=m ∪ x
12. Else go to 4
13. End if
14. if mean(m) ≈ μ
15. goto 16
16. end if.
17. end while
18. S_u =S_u∪∅

4. Experimental and results in depth

In this section, we demonstrate the accuracy and effectivity of the presented system on the special case of IfN/ENIT databases. The testing was performed on a computer with 4 GB RAM and 3.70 GHz Intel Core i3 processor and the code was applied in MATLAB.

4.1 Description of Database

Arabic script appears the great rank due to several facets, for example, every Muslims (around ¼ of (the) people on the earth) should be learned Arabic language this is because it is the language of Al-Quran, the holy book of Muslims, in addition to Arabic characters adopted to use in many languages like Malay, Jawi, Urdu, Persian and Pishtu, (more than a half of a billion people). Furthermore, the study area of recognizing hand writing Arabic is still challenge due to limited studies and therefore, more research is required. On the other side, IfN/ENIT databases are usually used as the main database accessible for Arabic handwriting recognition task. It consists of 26,459 images of 937 cities and names of Tunisian towns and it was written by 411 different writers.
The performance of the suggested technique is evaluated on the special case of IfN/ENIT databases which was described and used in[8]. It characterizes as extracted from IFN/ENIT Database, Number of categories is 37 and consists of 209201 samples in total. Note that, letters were distributed on the basis of shape regardless of its position in the word in one group, the total classes in this test was 37. For a more detailed discussion the reader is encouraged to refer[8]. Fig1 shows the sample of images from this database.

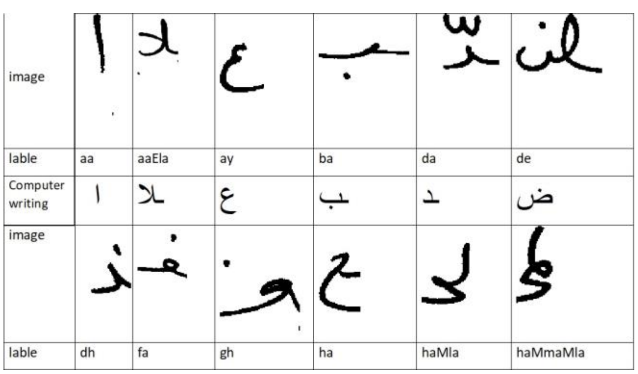

Fig. 1 Samples of database

4.2 Experiment procedures

First of all, this study concerning three-layer architecture. It is worth pointing out that, we scaled images to 50x50 using imresize function in Matlab because the original images have different size. To construct template set we randomly selected five images from each subject. For each class, we randomly divided the reference data into two equal groups one for training and other for testing. However, most of the previous research on IfN/ENIT database had tested on the set-e and trained their systems on the sets a–d.
This is due to the need for large training data for learning by the statistical classifiers. The testing of this study defines the dimension of the first layer template as 21*21 and that of the second layer template was defined as 29*29. To the best of our knowledge, the performance of the developed method is significantly affected by the following parameters: size of the available training set (Tr), patch size (u,v), selections of template size (Te). After carefully studying the above experiments, we acheive the ideal parameters values utilized to reach the best accuracy are: u = 20; v = 28; and Te = 6

4.3 Discussion of the results:

As for accuracy of the recognition itself, the developed method indicates that the performance comparable with the state-of-the-art results on the same data. As it can be showed in Table 1, the proposed system has more accuracy comparable to the[8]. In this test, the proposed technique was adopted as a feature extraction step while Support Vector Machines (SVM) was used for a classification step. Note that, the effectiveness of the developed technique was demonstrated by comparing it with previouse work in [8] which dealt with the same data.
Consequently, figure2 indicates that the suggested system has better performance than other previous approaches on the off-line Arabic handwriting recognition task, especially when the writer of object abides the rules of the Arabic writing.

Authors	***Accuracy***
Ramadhan et al.[8]	63.75
Developed	74.43

Fig. 2 Comparative results

Moreover, figure 3 displays that the developed method passed to recognize every object at least 18% from total of object which means SHA is acceptable method to recognize off-line Arabic handwriting.

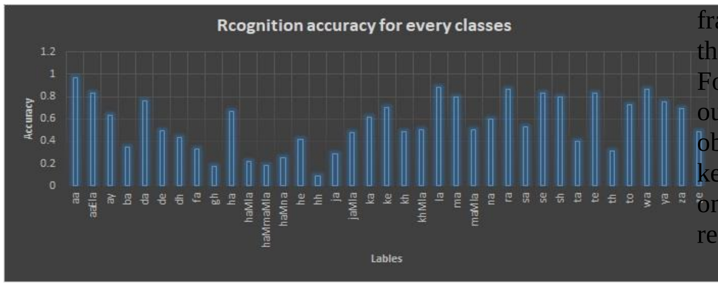

Fig. 3 Recognition accuracy for every classes

To reach the adequate analysis of ability recognizer, the most important of the main source of errors were highlighted as follows:

a) Here, testing dealt with a lot of database, so the errors increased.
b) Most of those who wrote this data did not follow the Arabic writing rules.
c) In special case many letters in Arabic only differ by dot so the differentiate between letters is a very hard task specially when writer not care about it.
d) Some errors comes from segmentation, so the segmentation step requires more tuning. Figure 4 displays side of challenges in this research which shows how same objects can be appeared as different way according to position and writer.

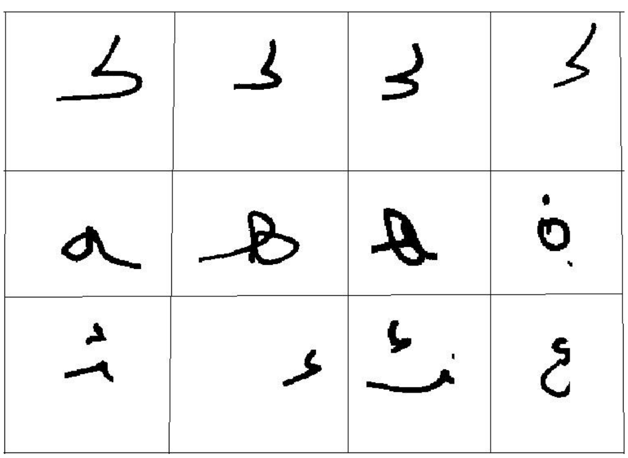

Fig. 4 Same objects has a different shape

5. Conclusion and future work:

This paper presented experiment results on off-line Arabic handwriting text recognition using statistical hierarchical architecture. The developing method is a hybrid Smale's framework and statistical measurement in order to enhance the accuracy of recognition off-line Arabic handwriting.
For future work it is required to use more database to test our algorithm and also the word segmentation should be observed into in the first place, because it is presently the key source of recognition errors. On the other hand, focus on design of the relation between letter-body and a dot recognizer should increase the recognition rate.

Acknowledgments

The authors would like to thank Jouf University for supporting this research and providing the computing facilities. The authors also thank the anonymous reviewers for their comments that improved this work.

References

[1] Wahba, G., Spline models for observational data. 1990: SIAM.
[2] Cortes, C. and V. Vapnik, Support vector machine. Machine learning, 1995. 20(3): p. 273-297.
[3] Evgeniou, T., M. Pontil, and T. Poggio, Regularization networks and support vector machines .Advances in computational mathematics, 2000. 13(1): p. 1.
[4] Smale, S., et al., Mathematics of the neural response. Foundations of Computational Mathematics, 2010. 10(1): p. 67-91.
[5] Bouvrie, J., et al., Generalization and properties of the neural response. 2010.
[6] Alsaidi, R.A.M., et al., Hierarchical sparse method with applications In vision and speech recognition. International Journal of Wavelets, Multiresolution and Information Processing, 2013. 11(02): p. 1350016.
[7] Li, H., et al., Sparse-based neural response for image classification. Neurocomputing, 2014. 144: p. 198-207.
[8] Alsaidi, R., H. Li, and Y. Wei, Improved Hierarchical Sparse Method with application to Offline handwritten Arabic character recognition. International Journal of Scientific, Engineering Research, 2014. 5.(03)
[9] Li, H., et al., Similarity learning for object recognition based on derived kernel. Neurocomputing, 2012. 83: p. 110-120.
[10] ALSAIDI, R.A., et al., SELECTION OF INFORMATIVE TEMPLATE IN HIERARCHICAL SPARSE METHOD . .2018
[11] Ramadhan A. M. Alsaidi, A.R.A.A., Saleh R. A. Alenazi, Madallah Alruwaili, Using Derived kernel as a new Method for Recognition a Similarity Learning. International Journal of Engineering and Advanced Technology (IJEAT), 2020. 9(3): p. 7.
[12] Abuhaiba, I.S., S.A. Mahmoud, and R.J. Green, Recognition of handwritten cursive Arabic characters. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994. 16(6): p. 664-672.
[13] Abandah, G. and N. Anssari, Novel moment features extraction for recognizing handwritten Arabic letters. Journal of Computer Science, 2009. 5(3): p. 226.

[14] Parvez, M.T. and S.A. Mahmoud, Arabic handwriting recognition using structural and syntactic pattern attributes. Pattern Recognition, 2013. 46(1): p. 141-154.
[15] Abdullah, M.A., L.M. Al-Harigy, and H.H. Al-Fraidi, Off-line Arabic handwriting character recognition using word segmentation. arXiv preprint arXiv:1206.1518, 2012.
[16] Menasri, F., et al. Shape-based alphabet for off-line Arabic handwriting recognition. in Ninth international conference on document analysis and recognition (ICDAR 2007). 2007. IEEE.
[17] Abandah, G.A., F.T. Jamour, and E.A. Qaralleh, Recognizing handwritten Arabic words using grapheme segmentation and recurrent neural networks. International Journal on Document Analysis and Recognition (IJDAR), 2014. 17(3): p. 275-291.
[18] Mouhcine, R., A. Mustapha, and M. Zouhir, Recognition of cursive Arabic handwritten text using embedded training based on HMMs. Journal of Electrical Systems and Information Technology, 2018. 5(2): p. 245-251.
[19] Quin, G., B. Pinel-Puysségur, and J.-M. Nicolas. Comparison of harmonic, geometric and arithmetic means for change detection in SAR time series. in EUSAR 2012; 9th European Conference on Synthetic Aperture Radar. 2012. VDE.
[20] Xudong, J.J.I.T.I.P., Iterative truncated arithmetic mean filter and its properties. 2012. 21(4): p. 1537-1547.
[21] Davis, T. and R.A.J.C.C. Poldrack, Quantifying the internal structure of categories using a neural typicality measure . :(7)24 .2014p. 1720-1737.
[22] Davis, T., et al., Global neural pattern similarity as a common basis for categorization and recognition memory. 2014. 34(22): p. 7472-7484.
[23] Kriegeskorte, N., M. Mur, and P.A.J.F.i.s.n. Bandettini, Representational similarity analysis-connecting the branches of systems neuroscience. 2008. 2: p. 4.
[24] Kriegeskorte, N., et al., Matching categorical object representations in inferior temporal cortex of man and monkey. 2008. 60(6): p. 1126-1141.
[25] Nili, H., et al., A toolbox for representational similarity analysis. 2014. 10(4): p. e1003553.
[26] Gardella, C., O. Marre, and T.J.P.o.t.N.A.o.S. Mora, Blindfold learning of an accurate neural metric. 2018. 115(13): p. 3267-3272.
[27] Bobadilla-Suarez, S., et al., Measures of neural similarity. 2019: p. 1-15.

Dr. Ramadhan A. M. Alsaidi, was born in Ibb, Yemen in 1974. He received the B.S. in mathematics & computer from the Sana'a University, in 1996, M.S. degrees in computation mathematics from USTC, China, in2010 and the Ph.D. degree in applied mathematics (computer software and technology) from HUST, China, in 2014. From 2007 to 2014, he was a Research Assistant with the computation and applied mathematics Laboratory. Since 2017, he has been an Assistant Professor with the Mathematics Department, Jouf University, science & art College in Gurayat. He is the author of more than 7 articles. His research interests include applied mathematics and image processing,

Dr. Ayed Rheal A. Alanzi, received the B. E. degree in Mathematics from Aljouf University, Kingdom of Saudi Arabia (KSA), in 2006, the M.S. degree in Statistics from Malaya University, Malaysia, in 2009, the M.S. degree in Mathematics from Southern Illinois University, Carbondale, Illinois, USA, in 2014, and the PhD degree in Statistics from The University of New Mexico, Albuquerque, New Mexico, USA, in 2017. He is currently an Assistant Professor of Statistics at Majmaah University, Majmaah, KSA. His research interests include Bayesian statistics, data analysis, regression analysis, statistics in phylogenetics, biostatistics, and the analysis of stock market time series data.

Dr. Saleh R. A. Alenazi, received the B.S. degree in Computer Science (Third Hons) from Multimedia University (MMU), Malaysia, in 2012, the M.S. degree in Computer Science from Northern University of Malaysia, Malaysia, in 2014, and the Ph.D degree in Computer Science from The National University of Malaysia, Malaysia, in 2019. He is currently an First Trainer A of Computer Sciences at Technical and Vocational Training Corporation. His research interests include Data Analysis, Big Data, Programing.

Dr. Taufiq H. Ghilan, was born in Yemen 1971. He received his B.Sc. in electronic engineering from Aleppo University, Aleppo, Syria, the M.Sc. degree in information and communication technology in 2012 and the PhD degree in Information Technology (Management) in 2018 from University Utara Malaysia, Kedah, Malaysia.

He is an author of a number of articles. His research interests include cloud computing and Green IT, Technology and innovation, social impact of information technology, information behavior, virtual/e-community, e-government, teleworking, ICT, e-banking, e-government, e-education, e-library, Smart Cities, Blockchain studies, human and computer interaction, statistical analysis using PLS-SEM, and open source studies.

From 2007 to 2010, he was the Director of Network Services Department of Yemen Center for Information and Communication Technology at Higher Education (YCIT-HE). Since 2008 to present, he has been appointed as the IT Consultant at the Central Organization for Control and Auditing (C.O.C.A), The Presidency of The Republic, Sana'a, Yemen. In 2007, he was assigned the task to be the Head of Network and Security Project in (C.O.C.A), Sana'a, Yemen. Since 2004 till 2006, he was the deputy manager of Network and Technical Support department in (C.O.C.A). From 2006 to 2007, he was also working as a System Engineer at Safer Oil Company, the National Oil Company, Sana'a, Yemen.

He is a volunteer reviewer with different peer-review journals including IEEE Access. He is a member of different system and

network entities such as Cisco microsystems and Microsoft corporation. He has achieved many certificates in network, security, and systems from Microsoft and Cisco systems.
Contact him at towfeeq2k5@yahoo.com or LinkedIn: https://www.linkedin.com/in/dr-taufiq-hail-ghilan-al-madhagy-07881339/