An efficient multiple-classifier system for Arabic calligraphy style recognition

1st Zineb Kaoudja
LINATI Laboratory
university of Kasdi Merbah
Ouargla, Algeria
zinebkaoudja@gmail.com

2nd Mohammed Lamine Kherfi
LAMIA laboratory
université du Québec
à Trois-Rivières, Canada and
LINATI laboratory
university Kasdi Merbah
Ouargla, Algeria
Kherfi@uqtr.ca

3rd Belal Khaldi
LINATI Laboratory
university of Kasdi Merbah
Ouargla, Algeria
bilel.khaldi@gmail.com

Abstract—**This work endeavors to present an automatic system for Arabic-calligraphy style classification. The main issue one faces in order to fulfill such a task is the unavailability of data due to the lack of related works. Therefore, we put forward, and make publically available, a new Arabic calligraphy style dataset. Additionally, we introduce a system for Arabic calligraphy style recognition based on the combination of multi-classifier decisions.**

Keywords— Arabic calligraphy handwriting, optical font recognition (OFR), Local Phase Quantization, Classifier combination.

I. Introduction

Arabic calligraphy handwriting is an artistic way for text writing, it contains several styles each one has its own features. These styles were born over the epochs, e.g., Kufic, Diwani, Naskh, Thuluth, Roqa'a, Maghribi, etc. The font type and style give important information about the corresponding document such as content, parts, history, origins, etc. Some of these styles are illustrated in Fig. 1. In order to differentiate Arabic artistic styles, one must holds a certain level of expertise. However, due to the huge amount of available digital documents in different fonts, manually classification became an extremely hard and tedious task. Therefore, computer vision and machine learning techniques should be involved in order to alleviate such a burden. Optical Font Recognition (OFR) is an approach that is widely used for automatic digital document processing and font recognition.

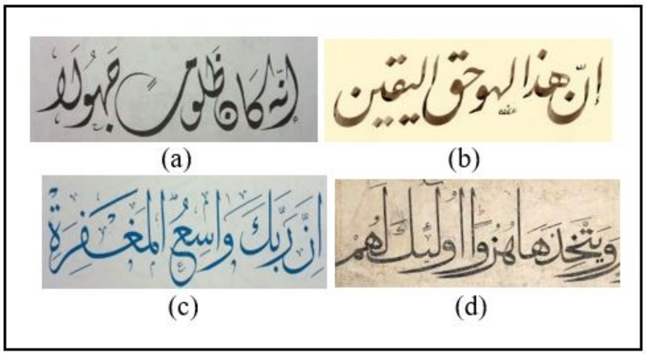

Fig.1: illustrative examples of some Arabic calligraphy styles (a) Diwani (b) Parsi (c) Thuluth (d) Mohakek.

Owing to the lack of data and works on Arabic calligraphy handwriting, difficulties confront researchers working on such a field, which raises a serious challenge. Another problem when dealing with Arabic text is the complex shape of the Arabic letters, especially in artistic texts (Fig. 1). The letters in Arabic have different shapes based on the appearance position (at first, middle or end of the word) as it is illustrated in Fig.2.

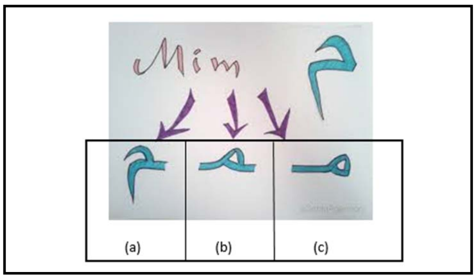

Fig.2: the arabic letter "Mim" shape's depends on its position (a) at the end (b) in the middle(c) at the beginning of the word.

Arabic handwriting style has a cursive shape, and its characteristic could be affected by the writer's style, or furthermore, the writer's emotions. All the aforementioned difficulties raises a serious challenge to propose a system for recognizing Arabic words or defining text styles. In Contrast to Arabic language, a great amount of work has been done for other languages such as Latin [1], Hebrew [2], and Chinese [3, 4]. Therefore and to the best of our knowledge, there is no publically available benchmark for styles of Arabic handwritten text. In all the previous works [5, 6, 7, 8], the researchers have used their small, auto-collected, very simple and personal image dataset. Some works in literature have used the digital printed text [9, 10], as an alternative of handwritten text [5, 6, 7, 8], for different tasks such as text extraction [11], segmentation [12], font recognition [13], word recognition [14], or writer identification [15]. However, opting for digital printed text, as an alternative of handwritten text, does not reflect the real word problems.

Indeed, there exists some works in literature where authors have tested their methods on real handwritten arabic artistic text. In [5], a method based on statistical features, named Edges Direction Matrix (EDMS), for Arabic calligraphy recognition has been proposed where the images is represented by 22 statistical moments. Nevertheless, this method has been evaluated using a small dataset that contains 700 images categorized into seven styles. Another work in [7] has been proposed aiming to classify the JAWY (Malizian language) text styles. In this work, 20 letters must be manually segmented before been analyzed. The main weakness of this study is the difficulty (sometimes impossible in artistic texts) of manually

978-1-7281-2642-5/19/$31.00 ©2019 IEEE

isolating the different letters in a word. Adam & al [8] have adopted a similar technic based on letter isolation. Instead, they have used Local Binary Pattern (LBP) and Gabor Filters (GF) as image features and SVM as a classifier.

For Arabic text classification systems based on classifiers combination technics, several works have been done. In [16] a proposed system for handwriting words recognition based on the combination of three classifiers, namely the multilayer perceptron (MLP), the support vector machine (SVM) and the Extreme Learning Machine (ELM). Where, they use several combination methods at the decision level between MLP, SVM and ELM. The classifiers were trained with Chebyshev moments (CM) and Contour-based Features (SCF). Another work have been done in [17] for writer identification from handwritten documents, using SVM classifiers combination. The combination model is based on Dempster–Shafer Theory (DST).

The aforementioned methods have achieved good results but in very simple conditions : a small dataset, a limited number of styles, styles that are easy to distinguish, Additionally, the used datasets lack complexity and hold images of words with easy-separable letters, which is not the case in most artistic calligraphy styles.

In this study, we put forward the richest handwritten Arabic calligraphy dataset that comprises 1685 images categorized into 9 categories. The introduced dataset offers a more serious and real words problems in Arabic calligraphy styles. To meet real word cases, the data has been collected from heterogeneous sources including books, manuscripts and web. In addition, we have proposed a new system for automatic style classification of handwritten Arabic calligraphy using a combination of multiple classifiers. The proposed system has shown promising results as it will be shown in the experimental section.

The rest of the paper is organized as follow: In section 2, we give a general overview of the proposed system and the different methods of classifier combination. Section 3 is devoted to present our collect dataset. In Section 4, experimentation will be conducted and results will reported. Finally, we draw some conclusions.

II. The Proposed System

Our main aim is to recognize Arabic calligraphy text style in an automatic manner. The proposed solution is based on multi-classifier technique that have recently been widely utilized in different tasks such as object classification [18] and character recognition [19]. For images representation we choose LPQ texture descriptor. texture-based methods has been established successfully in various applications of digital documents analyses , this is one of the recent works that investigate the effect of using different texture discriptors with several datasets of digital text images[20]. A general scheme of the proposed system is given in Fig. 3.

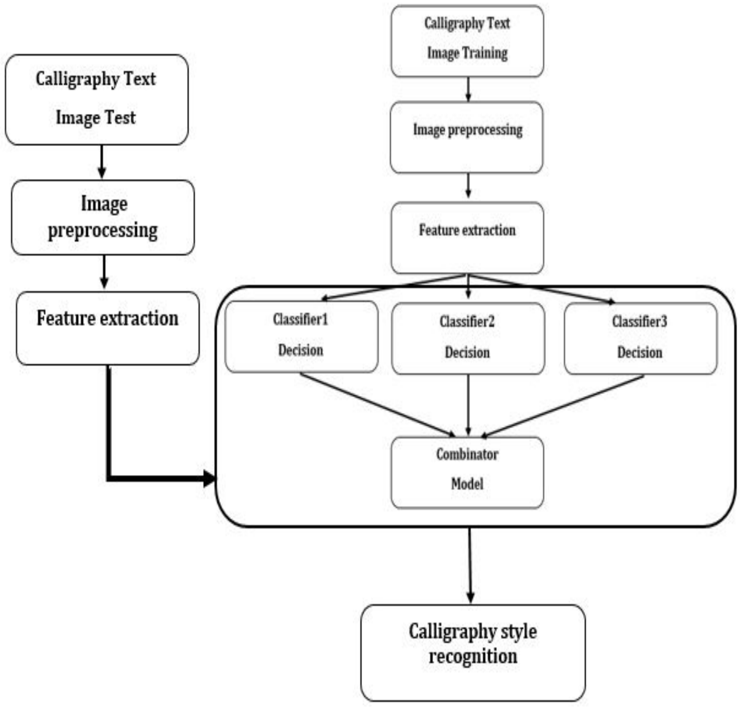

Fig.3. illustrative scheme of the proposed system.

Firstly, the image is subjected to a preprocessing step in order to clean noise, binarize and crop the text. From each image in the training set, a Local Phase Quantization (LPQ) descriptor is extracted then fed to a multi-classifier module for training purposes. Likewise, test images are passed through the same process for identification purposes. The main reason behind using multiple classifiers instead of one is the intuition stating that multiple decision is better than one.

A. *Preprocessing*

At this stage, we binarize all the images to get clean text without extra information that could affect the recognition process. The adopted method is Otsu's [21], which establishes a threshold that minimizes the intra-class variance (Fig.4).

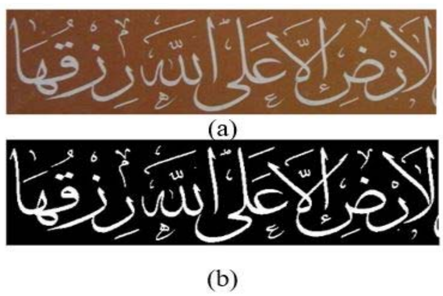

Fig.4. image binarization using Otsu's method (a) color image (b) binary image.

B. *local phase quantization LPQ(feature extraction)*

The local phase quantization is a blur insensitive texture descriptor, mainly invented to deal with image blurring. LPQ method founded on the blur invariance property of the Fourier phase spectrum [22]. At each pixel position x of the image, the local phase information computed over M-by-M neighborhood using the 2-D DFT (Fig.5) or, more precisely, a short-term Fourier transform (STFT) using the following equation.

$$(u, x) = \sum y \in Nx\ (x - y)\ {-j2\pi u}\ Ty = wuTf$$

Where wu is the basis vector of the 2-D DFT at the frequency u, and f is another vector containing all M2 image samples from Nx.

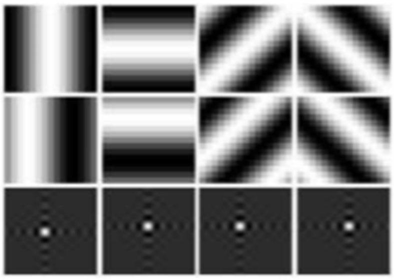

Fig.5. the short-term Fourier transform (STFT “uniform window”).

C. Classification & combination

This step aims to assign each image to its corresponding class by employing, in parallel manner, the decision of three different classifiers namely: Multilayer Perceptron (MLP), Support Vector Machine (SVM), and K-Nearest Neighbor (KNN). These three classifiers have shown sufficient results on our dataset. It should be mentioned that there are several methods for combining multiple classifiers decisions’ each with its pros and cons. In order to improve the classification results, we have tried multiple combination techniques namely majority vote, maximum, minimum, Sum. We use this simple classifiers combination technics for their effectiveness, ease of use and because they do not require a complex calculation. As shown in this paper [23] the best classifier combinations technic is depends on your data.

- Majority/Plurality vote means the correct class is the one chosen (voted) by the most of the classifiers. Even the sum of those votes did not exceeds 50% of the all votes obtained from the ensemble of classifiers. If all the classifiers indicate different classes, then the final class is the class with the overall maximum output score value.
- Maximum, the final score is the maximum between the classifiers output scores. That means if one classifier insists on a specific class for a given test sample, final decision assigns it to that class, even if all other classifiers disagree.
- Minimum, the final assigned score is the maximum between the minimum classifiers output scores. That means for each classifier we take the minimum score for a given sample, then we choose the maximum between them.
- Sum, summing the output scores of each base classifiers and assigns the final class label with the maximum score to given input sample.

III. THE NEW DATABASE PRESENTATION

As it has been mentioned before, the literature lacks a public benchmark for handwritten Arabic calligraphy style classification. Although researches have used some personal datasets, these datasets suffer from the smallness, the homogeneity and the simplicity, which dispose them from being challenging. To tackle this issue, we propose and make publically available a new dataset that holds the heights number of images and categories among others. The proposed dataset holds the following advantages:

- Richness: Different categories represent the basis calligraphy styles.
- Size: Sufficient number of images in each category.
- Heterogeneity (variation): Scale and sentence length variation for the same category.

The proposed dataset contains **9** artistic styles of calligraphy which are Naskh, Diwani, Parsi, Rekaa, Thuluth, Maghribi, Kufic, Mohakek, and Square-Kufic where each style represents a category /class as it is presented in Fig.6. The number of images in each class is between 180 and 195 images as given in Table 1. For the capturing process, a camera with a resolution of 18 megapixels has been used to get images with a sufficient quality. To grant variation, the pictures have been taken from different sources such as books, manuscripts and from the web. The texts within the images have been carefully and manually segmented to formulate phrases with different lengths. Finally, the obtained images (i.e., 1685 image) have been categorized in their appropriate font styles.

Fig.6. the database Arabic calligraphy styles.

TABLE I. THE DIFFERENT STYLES WITH THEIR RESPECTIVE NUMBER OF IMAGES IN OUR DATASET.

Category style	N^0. Of images
Diwani	190
Naskh	190
Parsi	180
Rekaa	185
Thuluth	195
Maghribi	180
Kufi	185
Mohakek	190
Squar-kufic	190
Total images n^0	**1685**

III. THE EXPERIMENTAL RESULTS:

In this section, we accomplish two different experiments. The first one is devoted to separately evaluate each classifier whereas the second one is for evaluating the effect of decision combination. In order to avoid over-fitting, K-Fold (k=3) cross-validation has been used. The motivation to use cross validation techniques is that when we fit a model, we are fitting it to a training dataset. Without cross validation we only have information on how does our model perform to our in-sample data. Ideally, we would like to see how the model performs when we have a new data in terms of accuracy of its predictions [24] [25]. The dataset is firstly divided into three equal sub-sets, and then, per each test sub-set, we use the other two for training. The average accuracy of the three folds, for each classifier, have been presented in TABLE II and Fig 6.

TABLE II. AVERAGE ACCURACY FOR EACH SINGLE CLASSIFIER

Classifier	Average accuracy
MLP	94.50%
KNN	93.79%
SVM	94.60%

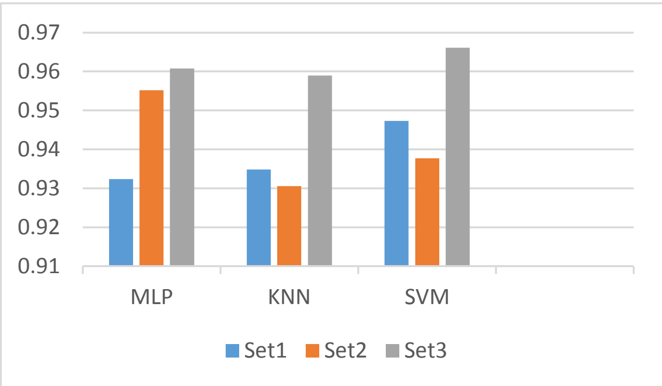

Fig.6. the results of the three folds cross-validation for the individual classifiers.

We can see that the classifiers have reported similar results with an error rate of 5.4%-7%. In order to reduce this error rate, the three classifiers have been combined and evaluated using the same configuration. We evaluated the performance of the proposed system by comparing it with the recent Edges Direction Matrix (EDMs) proposed by Bathaina & al [5]. The EDMs statistically analyses the relationship between pixels of the boundary edges of a binary image. The final feature vector is calculated from the edges EDM1 and EDM2. The proposed method extracts 22 statistical moments (correlation, homogeneity, etc.). TABLE III and Fig 7 presents the results obtained with our system using different combination methods.

TABLE III. THE AVERAGE ACCURACY OF MULTIPLE-CLASSIFIERS SYSTEM WITH DIFFERENT COMBINATION METHODS

Combination methods	Avg accuracy
Majority voting	96.44%
Maximum	93.83%
Sum	**96.91%**
Minimum	96.56%
EDMs+ Decision-Tree [5]	70.09%

As it is shown in Table III, the proposed system has yielded the best result among all by reducing the error rate to 3.01% instead of 5.4% reported by SVM separately. This confirms the intuition stating that opting for multiple decision instead of one improves results. Additionally, we can see that Bathaina & al [5] has yielded poor results compared to the others. This could be attributed to the sensitivity of the later method to the complex and heterogeneous artistic calligraphy texts. It, also, seems that Sum combination methods has reported better results than the others have. This is because Sum method involves all the classifiers in the decision-making in contrast to the others where the decision is made by the most (resp. least) dominant(s).

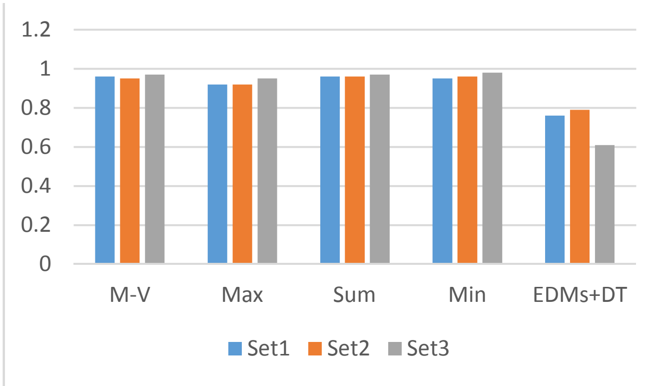

Fig.7. the results of the three folds cross-validation for the classifier combination technics and the EDMs[5].

The correction rate results for each type of calligraphy are shown in Table IV. With Sum combination method, the highest accuracy was achieved for Square-Kufi, which achieve 99.08%, at the same time, the lowest accuracy was obtained with the Diwani calligraphy at 94.30%.

Likewise, the highest accuracy for EDMs method was Square-Kufi with 83.73%, and the lowest accuracy was for Mohakek style with 58.00%. Considering these results, the proposed System obtained an accuracy rate for each type of calligraphy higher than the EDMs method.

TABLE IV. THE CORRECTION RATE FOR EACH CLASS BY PROPOSED SYSTEM AND EDMS.

	Our System	EDMs+DT[5]
Diwani	94.30%	67.27%
Naskh	98.60%	76.40%
Rekaa	98.27%	65.36%
Parisi	96.44%	69.71%
Thuluth	96.87%	59.99%
Maghribi	96.56%	71.23%
Kufic	98.09%	73.74%
Mohakek	95.48%	58.00%
Square-Kufic	99.08%	83.73%

IV. CONCLUSION

In spite of the importance of Arabic calligraphy handwriting, an intensive work should be carried out trying to alleviate the problems concern with text/style recognition. This paper put forward a major twofold contribution to Arabic calligraphy font styles. In one hand, it provides a rich and diverse image dataset that might be used for different tasks of text and style recognition. In the other hand, we provide an automatic system based on the LPQ texture descriptor and multiple classifiers combination for Arabic calligraphy styles recognition. Overall, the obtained results demonstrate the effectiveness of combining multiple decisions by yielding the highest performance compared to single decision or to the work of Bataineh & al [5].

REFERENCES

[1] Ramanathan, R., et al. "A novel technique for english font recognition using support vector machines." 2009 International Conference on Advances in Recent Technologies in Communication and Computing. IEEE, 2009.

[2] Yosef, Itay Bar, et al. "Classification of Hebrew calligraphic handwriting styles: preliminary results." First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.. IEEE, 2004.

[3] Zhu, Yong, Tieniu Tan, and Yunhong Wang. "Font recognition based on global texture analysis." IEEE Transactions on pattern analysis and machine intelligence 23.10 (2001): 1192-1200.

[4] Zhang, Xiafen, and George Nagy. "Computational method for calligraphic style representation and classification." Journal of Electronic Imaging 24.5 (2015): 053003.

[5] Bataineh, Bilal, Siti Norul Huda Sheikh Abdullah, and Khairuddin Omar. "A novel statistical feature extraction method for textual images: Optical font recognition." Expert Systems with Applications 39.5 (2012): 5470-5477.

[6] Allaf, S. R., and R. Al-Hmouz. "Automatic Recognition of Artistic Arabic Calligraphy Types." Journal of King Abdulaziz University: Engineering Sciences 27.1 (2016).

[7] Azmi, Mohd Sanusi, et al. "Arabic calligraphy classification using triangle model for Digital Jawi Paleography analysis." 2011 11th International Conference on Hybrid Intelligent Systems (HIS). IEEE, 2011.

[8] Adam, Kalthoum, Somaya Al-Maadeed, and Ahmed Bouridane. "based classification of Arabic scripts style in ancient Arabic manuscripts: Preliminary results." 2017 1st International Workshop on Arabic Script Analysis and Recognition (ASAR). IEEE, 2017.

[9] Slimane, Fouad, and A. I. Smpai. "Multi-font and multi-size low-resolution arabic text recognition using apti-database." Document Analysis and Text Recognition: Benchmarking State-of-the-Art Systems. World Scientific Publishing Co, 2016.

[10] Luqman, Hamzah, Sabri A. Mahmoud, and Sameh Awaida. "KAFD Arabic font database." Pattern Recognition 47.6 (2014): 2231-2240.

[11] Zayene, Oussama, et al. "A dataset for Arabic text detection, tracking and recognition in news videos-AcTiV." 2015 13th International Conference on Document Analysis and Recognition (ICDAR). IEEE, 2015.

[12] Al-Thubaity, Abdulmohsen, and Abdullah Al-Subaie. "Effect of word segmentation on Arabic text classification." 2015 International Conference on Asian Language Processing (IALP). IEEE, 2015.

[13] Lutf, Mohammed, et al. "Arabic font recognition based on diacritics features." Pattern Recognition 47.2 (2014): 672-684.

[14] Alghamdi, Mansoor, and William Teahan. "Printed Arabic Script Recognition: A Survey." INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS 9.9 (2018): 415-428.

[15] Hannad, Yaacoub, Imran Siddiqi, and Mohamed El Youssfi El Kettani. "Writer identification using texture descriptors of handwritten fragments." Expert Systems with Applications 47 (2016): 14-22.

[16] Tamen, Zahia, Habiba Drias, and Dalila Boughaci. "An efficient multiple classifier system for Arabic handwritten words recognition." Pattern Recognition Letters 93 (2017): 123-132.

[17] Kessentini, Yousri, Sana BenAbderrahim, and Chawki Djeddi. "Evidential combination of SVM classifiers for writer recognition." Neurocomputing 313 (2018): 1-13.

[18] Tulyakov, Sergey, et al. "Review of classifier combination methods." Machine learning in document analysis and recognition. Springer, Berlin, Heidelberg, 2008. 361-386.

[19] Rahman, Ahmad Fuad Rezaur, and Michael C. Fairhurst. "Multiple classifier decision combination strategies for character recognition: A review." Document Analysis and Recognition 5.4 (2003): 166-194.

[20] Alaei, Fahimeh, et al. "A comparative study of different texture features for document image retrieval." Expert Systems with Applications 121 (2019): 97-114.

[21] Otsu, Nobuyuki. "A threshold selection method from gray-level histograms." IEEE transactions on systems, man, and cybernetics 9.1 (1979): 62-66.

[22] Ojansivu, Ville, and Janne Heikkilä. "Blur insensitive texture classification using local phase quantization." International conference on image and signal processing. Springer, Berlin, Heidelberg, 2008.

[23] Mohandes, Mohamed, Mohamed Deriche, and Salihu O. Aliyu. "Classifiers combination techniques: A comprehensive review." IEEE Access 6 (2018): 19626-19639.

[24] Andonie, Răzvan. "Extreme data mining: Inference from small datasets." International Journal of Computers Communications & Control 5.3 (2010): 280-291.

[25] Smelser, Neil J., and Paul B. Baltes, eds. International encyclopedia of the social & behavioral sciences. Vol. 11. Amsterdam: Elsevier, 2001.

NOTE :

The web link to our Arabic Calligraphy Database:

https://drive.google.com/open?id=1dC7pwzT_RHL9B42H8-Nzf5Sant-86NV6