A Review of Feature Extraction Techniques for Handwritten Arabic Text Recognition

Bouchra El qacimy*, Ahmed Hammouch
Laboratory LRGE, ENSET of Rabat,
Mohamed V University,
Rabat, Morocco
*elq.bouchra@gmail.com, hammouch_a@yahoo.com

Mounir Ait kerroum
Laboratory LARIT, Ibn Tufail University,
Faculty of Science, ENCG of Kenitra,
Kenitra, Morocco
maitkerroum@gmail.com

***Abstract*—Research in Arabic handwritten recognition has been of growing interest in the last few decades. This is mainly due to its broad spectrum of applications in different fields such as bank check processing, form data entry, postal mail sorting, automatic processing of old manuscripts, etc. In the literature, numerous techniques have been proposed for feature extraction and applied to various types of images. This work provides a comprehensive review of these methods for off-line handwritten Arabic text recognition. It also presents recognition rates and descriptions of the databases used for the discussed approaches. This paper includes background on the field, discussion of feature extraction methods, and future research directions.**

***Keywords*—Features Extraction; Arabic handwriting; Text recognition; Off-line recognition.**

I. INTRODUCTION

Handwritten Arabic script recognition made object of intense researches justified by the importance of its numerous potential applications. Indeed, several entities of various domains showed interest in this field, in particular the banks for automatic checks processing, post offices for the automation of post mail sorting, insurances for automatic form processing and many other industries [1]. Moreover, Arabic is the 5^{th} spoken language world wide after Chinese, Spanish, English and Hindi. With a population exceeding 237 million, Arabic is the first language in 60 countries [2]. However, automatic recognition of Arabic handwritten script remains a challenging task due, in part, to its inherent characteristics such as cursiveness, overlapping characters, presence of diacritic points, etc. On the other hand, unlike research in Latin script, for which several standard databases were developed to serve as benchmark for researchers [3] [4] [5] [6], there is a lack of standard databases for handwritten Arabic text recognition. Any recognition system needs a large database for training and testing. Most data from insurance companies or banks are confidential for non-lucrative research. Moreover, many works in the Arabic handwritten text recognition field have used databases that are not available to the public [7] [8] . However, over the past decade, several large standard databases have been developed. Margner and ElAbed [9] presented a survey of databases used for Arabic text recognition and their availability. Alma'adeed et al. [10] presented an Arabic handwritten database (AHDB) from 100 writers. This database was developed for reading check amounts research. Al-Ohali et al. [11] developed an Arabic check database for the Center of Pattern Recognition and Machine Intelligence (CENPARMI) in Canada. This database has 1547 legal amounts for checks written in words, 23325 sub-words, and 9865 Indian digits which are used in Arabic handwriting normally. The most used database for handwritten text recognition is the IFN/ENIT Arabic database of handwritten Arabic words. It was released by Pechwitz et al [12] and it is available for free download at the IFN/ENIT home page. The database consists of 937 Tunisian town/villages names along with their postcodes. All the forms were scanned at 300 dpi and converted to binary images. The images are divided into five sets so that researchers can use some of them for training and some for testing.

According to Arica et al. [13], recognition systems are classified based on the data acquisition into two types: on-line based or off-line based systems. In on-line recognition systems, the computer recognizes the text as the user is writing it. This is done using a special digitizer tablet and pen. A digitizer is an electromagnetic tablet which transfers the coordinates of the pen position to the computer at a constant rate [14]. On the other hand, in off-line recognition systems, the image of the text is converted into a computerized format by an optical scanner or camera, then the text is recognized after the printing or writing is completed. Our work focuses on Off-line handwritten Arabic text recognition, especially on feature extraction techniques. In this paper we identify and discuss different techniques used in Arabic handwriting text recognition. We provide descriptions of databases used and recognition rates for the discussed approaches.

The rest of this paper is organized as follows: in section II we give a brief overview of the Arabic script characteristics. Next, in section III we describe the components of a typical off-line handwritten Arabic text recognition system. Then, we present a comprehensive review of state of the art feature extraction techniques. Finally, the conclusion gives a brief summary and addresses some future research directions in Arabic handwriting text recognition.

II. ARABIC SCRIPT CHARACTERISTICS

The Arabic script is distinguished from other writings with various points. There are few basic principles of Arabic script which apply whichever style of writing is used [15]:

- There are 28 letters in the alphabet ;
- The script is written right to left ;
- There are no capital letters ;

978-1-4799-7479-5/15/$31.00 ©2015 IEEE

- Words are written in cursive, or 'joined-up', writing. All the letters join to the letter before in a word and all but six join to the letter after also.

Unlike English, each Arabic letter has either two or four shapes depending on the letter's position in the text. The shape of a letter changes with its position, which may be at the beginning, middle or end of a word, or in isolated form [16]. Although generally, Arabic is cursive, there are some non-cursive letters. There are 22 cursive letters with four different shapes and 6 non-cursive letters with only two shapes corresponding to the isolated and end position. Figure 1 shows the different forms for each letter. For example, letter Ayn (ع) has the following shapes: عـ at start, ـعـ at middle, ـع at end, and (ع) when isolated. In addition, Arabic language uses various diacritical marking to distinguish some words from each other, including 'fattha', 'dumma', 'kasra', 'hamza', 'shadda', or 'madda'. The main objective of the diacritical marking is to control the pronunciation and meaning of words. Indeed, the presence or absence of vowel diacritical indicates different meaning [14]. For example, كتب could mean books or the action of writing depending on the diacritics used. However, diacritical marking are not normally written in handwriting. On the other hand, dots are essential in Arabic script in order to distinguish some characters from each other. One letter can have up to three dots which may be placed either above or below the main body of the letter.

EF	MF	BF	IF	EF	MF	BF	IF
ـض	ـضـ	ضـ	ض	ـا			أ
ـط	ـطـ	طـ	ط	ـب	ـبـ	بـ	ب
ـظ	ـظـ	ظـ	ظ	ـت	ـتـ	تـ	ت
ـع	ـعـ	عـ	ع	ـث	ـثـ	ثـ	ث
ـغ	ـغـ	غـ	غ	ـج	ـجـ	جـ	ج
ـف	ـفـ	فـ	ف	ـح	ـحـ	حـ	ح
ـق	ـقـ	قـ	ق	ـخ	ـخـ	خـ	خ
ـك	ـكـ	كـ	ك	ـد			د
ـل	ـلـ	لـ	ل	ـذ			ذ
ـم	ـمـ	مـ	م	ـر			ر
ـن	ـنـ	نـ	ن	ـز			ز
ـه	ـهـ	هـ	ه	ـس	ـسـ	سـ	س
ـو			و	ـش	ـشـ	شـ	ش
ـي	ـيـ	يـ	ي	ـص	ـصـ	صـ	ص

Figure 1: Arabic letters forms.

The non cursive letters are not connected from the left resulting in the separation of a word into sub-words or pieces of Arabic words (PAW) [17]. Figure 2 shows three Arabic words with one, two and three sub-words respectively.

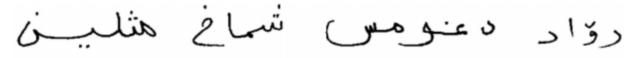

Figure 2: Words containing respectively one, two, three, and four sub-words.

III. Overview of the off-line handwriting Arabic text recognition process

A typical text recognition system depends on five main components. These components are : preprocessing, representation, segmentation, feature extraction, and classification. Some of these components are omitted or merged by some researchers [13].

Figure 3 illustrates a general model for Arabic off-line handwritten text recognition system.

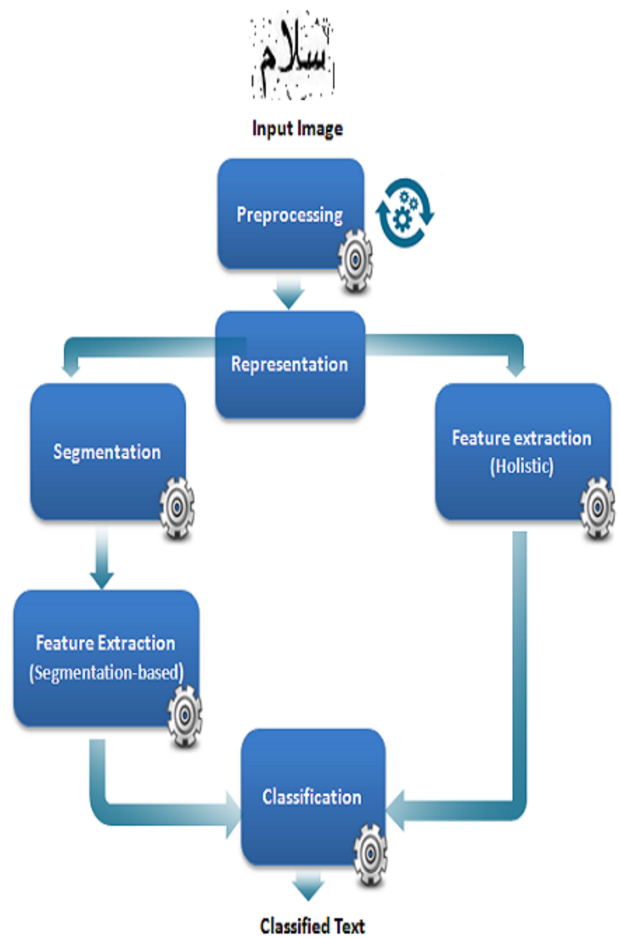

Figure 3: A general model of Arabic off-line handwritten text recognition system.

The input to the system is a scanned text image. The scanned image may need to go through some preprocessing steps, where the image is enhanced before recognition. Common preprocessing tasks include noise removal, skew detection and correction, etc. After preprocessing, the text image may need to be segmented into lines and/or words/sub-words/characters/sub-characters. Then features are extracted.

Features are used to train the classifier to build the models (in the training phase) or to classify based on the previously generated models (testing phase). We can add another step for validation which is post-processing. Post-processing improves recognition accuracy by refining the decisions taken by the previous stage and possibly recognizing words by using the context [18].

Next, we'll describe the different phases outlined previously.

A. *Preprocessing*

The preprocessing phase is considered one of the most important phases in any recognition systems. The main objective of preprocessing is to remove irrelevant information that has a negative effect on the recognition [19]. Once the input image is acquired, preprocessing enhances the image for better recognition. Preprocessing usually includes techniques such as noise removal, thresholding, thinning, skew/slant correction, baseline estimation and normalization. The choice of preprocessing steps utilized depends on the desired recognition algorithm [20].

B. *Representation*

The image is often converted to a more concise representation prior to recognition as shown in figure 4. A skeleton is a one-pixel thick representation showing the centerlines of the text. Skeletonization, or thinning, facilitates shape classification and feature detection. An alternative is the Freeman chain code of the border (contour) of the text [21], [22]. Chain code stores the absolute position of the first pixel and the relative positions of successive pixels along the contour. Difficulties with thinning include possible mis-localization of features and ambiguities particular to each thinning algorithm. The contour approach avoids these difficulties since no shape information is lost.

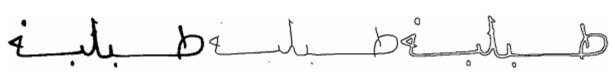

Figure 4: Different image representations [17].

C. *Segmentation*

In Arabic text recognition, the segmentation phase is a crucial phase, because any error in segmentation will impact the recognition stage. Besides, the literature distinguishes between two distinct approaches for Arabic text recognition, namely: the segmentation-based systems and segmentation free systems, or holistic systems [23]. In the latter, recognition is performed on the whole representation of a word without segmentation. While in the segmentation-based systems, the segmentation is done according to three different approaches: Segmenting word into characters, segmenting words into its primitives and segmentation after recognition.

D. *Feature extraction*

The purpose of feature extraction is the measurement of the attributes that are most pertinent to a given classification task. Feature extraction methods can be categorized into structural features, statistical features and feature space transformations methods.

1) Structural Features: The structural features involve the geometrical and topological characteristics of an input image [1]. The structural features includes but are not limited to: strokes, width of the stroke, stroke directions, chain codes, end points, number of maxima and/or minima, numbers of cusps, the number of dots and their position relative to the baseline, loops, etc [13] [23].

2) Statistical Features: The statistical features are extracted from the statistical distribution of pixels which describe the characteristic measurements of the input image pattern [23]. The statistical features provide low complexity and high speed. The major statistical features can be summarized as: zonings, characteristic loci, crossings and distances [13].

3) Feature space transformations: The image can be represented by a linear combination of a series of simpler well defined functions. The series expansion provides a compact encoding by the coefficients of the linear combination. The common transforms and series expansions include [20]: Fourier Transform, Walsh Hadamard Transform, Rapid Transform, Hough Transform, Gabor Transform, Wavelets, Karhunen Loeve Expression and moments.

E. *Classification*

The aim of this phase is to assign each observation with a class label or membership scores to the defined classes. A number of classifiers have been used for recognition of Arabic handwritten words. These include Hidden Markov Model (HMM), Support Vector Machines (SVM), Artificial Neural Networks (ANN),k-Nearest Neighbors (kNN), and others.

IV. BENCHMARK OF FEATURE EXTRACTION TECHNIQUES

In the literature, numerous types of features were proposed by researchers.

As shown in Table I, there is an intense inclination for using structural features. Indeed, Almuallim [24] and Goraine [25] used a set of structural features as loops, dots and stroke directions along with a classifier based on rules. They achieved respectively 90 and 91% as recognition rate on their particular datasets. We remark that most systems are based on Hidden Markov Models (HMM) such as [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36] and [37]. Those systems are more likely to use structural features. Fewer systems used Support Vector machine. Khalifa et al. [38] implemented a holistic recognition system using DCT features and SVM Classifier and achieved a recognition rate of 91,70% on the IFN/ENIT Arabic Standard Database which is very competitive. Neural networks are also broadly used for handwritten Arabic text recognition [39], [40], [41]. Haboubi et al. [42] proposed a segmentation-based recognition system using pixel and structural description, Gabor Filter and Fourier descriptors. He attained 89% on the IFN/ENIT database. Finally, Parvez et al. 2013 [18] proposed a holistic recognition system based on structural and syntactic pattern attributes that were fed to a fuzzy polygon matching recognizer. They achieved a recognition rate of 92,20% on the well known IFN/ENIT database.

TABLE I: Summary of Results for Arabic Handwritten Text/Word Recognition

Publication	Database	Recognition approach	Features	Recognition engine	Recognition rate
Almuallim 1987	400 images of two writers	Segmentation-based	Strokes endpoints, lengths, frames, connection points	Set of recognition rules	91%
Goraine 1992	180 images of 3 writers	Segmentation-based	Stroke directions, sub-word positions, loops, secondary strokes, dots	Rules for recognition	90%
El-Hajj 2005	IFN/ENIT	Holistic	Pixel densities, density transitions, concavity configuration along frames and with respect of baselines, frame derivatives, centers of gravity with respect to baselines	HMM	87,20%
Fahmy 2001	600 words written by one writer	Segmentation-based	Locating endpoints, junctions, turning points, loops, generating frames and detecting strokes	ANN	69,70%
Dehghan 2001	17,000 images of the 198 names of cities in Iran	Holistic	Histograms of Freeman chain code directions in regions of vertical frames	Discrete HMM	65%
Snoussi Madouri 2002	2,070 images with a lexicon of 70 words	Holistic	Ascenders, descenders, loops, dots, sub-word position, normalized Fourier descriptors	Transparent four-layer NN	97%
Pechwitz 2003	IFN/ENIT	Holistic	Pixel intensities	Semi-continuous HMMs	89%
Khorsheed 2003	Manuscript gamhart alnasab libni lkalbi	Holistic	Dots, endpoints, branch, cross, turning points, loops, lines	HMM	97%
Souici-Meslati 2004	480 for training and 1,200 words for testing	Holistic	loops, dots, connected components, ascenders, and descenders	Rule based ANNS	93%
Farah 2004	1200 words by 100 writers for training and 3,600 words for testing	Holistic	Loops, dots, connected components, ascenders, and descenders	Parallels neural network, k-nearest-neighbor and fuzzy k-nearest-neighbor	96%
Almaadeed 2004	AHDB	Holistic	Ascenders, descenders, and other structural features, frame-based features	HMM and rule based system	60%
Pechwitz and Margner 2006	IFN/ENIT	Holistic	Skeleton directions, Pixel values	Semicontinuous 1-D HMM	89,10%
Margner et al. 2006	IFN/ENIT	Segmentation-based	Pixel values	HMM	74,69%
Kundu et al. 2007	IFN/ENIT	Segmentation-based	Geometrical and topological features	Variable Duration HMM	60%
Menasri et al. 2007	IFN/ENIT	Segmentation-based	Shape-based Graphemes	Hybrid HMM/NN	87,40%
Touj et al. 2007	IFN/ENIT	Segmentation-based	Directional values, loops, connection of diacritics Planar graphemes, presence of diacritics	Planner HMM	86,10%
Khorsheed 2007	32,000 words from manuscripts	Holistic	Spectral features	HMM (one for each word)	85%
Haboubi et al. 2009	IFN/ENIT	Segmentation-based	Pixel description, structural description, Gabor Filter, Fourier Descriptors	ANN	87,10%
Mohamad et al. 2009	IFN/ENIT	Holistic	Distribution and concavity features	Multiple HMM	90,26%
Kessentini et al. 2010	IFN/ENIT	Holistic	Directional, contour and density features	Multi-stream HMM	79,60%
Khalifa et al. 2011	IFN/ENIT	Holistic	DCT features	SVM	91,70%
Parvez et al. 2013	IFN/ENIT	Holistic	Structural and syntactic pattern attributes	Fuzzy polygon matching algorithms	92,20%

V. Future work and conclusions

In this work we provided a comprehensive review of feature extraction methods for off-line handwritten Arabic text recognition. We presented a description of the databases used and the recognition rates of the discussed approaches. Most of the recent work on off-line Arabic text recognition has focused on isolated characters, digits, and words recognition. There are very few attempts on Arabic page of text recognition [18]. Therefore, more research effort is needed for unconstrained Arabic handwritten text recognition. Moreover, features used for Arabic text recognition are mostly imported from other languages or modifications of features of other languages. More novel features, taking into consideration the characteristics of Arabic text is highly needed [18].

References

[1] M. Cheriet, N. Kharma, C.-L. Liu, and C. Suen, *Character recognition systems: a guide for students and practitioners*. John Wiley & Sons, 2007.

[2] M. P. Lewis *et al.*, *Ethnologue: Languages of the world*. SIL international Dallas, TX, 2009, vol. 9.

[3] J. J. Hull, "A database for handwritten text recognition research," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 16, no. 5, pp. 550–554, 1994.

[4] P. J. Grother, "Nist special database 19 handprinted forms and characters database," *National Institute of Standards and Technology*, 1995.

[5] L. Deng, "The mnist database of handwritten digit images for machine learning research," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 141–142, 2012.

[6] U.-V. Marti and H. Bunke, "The iam-database: an english sentence database for offline handwriting recognition," *International Journal on Document Analysis and Recognition*, vol. 5, no. 1, pp. 39–46, 2002.

[7] S. N. Srihari, G. R. Ball, and H. Srinivasan, "Versatile search of scanned arabic handwriting," in *Arabic and Chinese Handwriting Recognition*. Springer, 2008, pp. 57–69.

[8] S. Abdulla, A. Al-Nassiri, and R. A. Salam, "Off-line arabic handwritten word segmentation using rotational invariant segments features." *Int. Arab J. Inf. Technol.*, vol. 5, no. 2, pp. 200–208, 2008.

[9] H. El Abed and V. Margner, "Arabic text recognition systems-state of the art and future trends," in *Innovations in Information Technology, 2008. IIT 2008. International Conference on*. IEEE, 2008, pp. 692–696.

[10] S. Al-Ma'adeed, D. Elliman, and C. A. Higgins, "A data base for arabic handwritten text recognition research," in *Frontiers in Handwriting Recognition, 2002. Proceedings. Eighth International Workshop on*. IEEE, 2002, pp. 485–489.

[11] Y. Al-Ohali, M. Cheriet, and C. Suen, "Databases for recognition of handwritten arabic cheques," *Pattern Recognition*, vol. 36, no. 1, pp. 111–121, 2003.

[12] M. Pechwitz, S. S. Maddouri, V. Märgner, N. Ellouze, H. Amiri *et al.*, "Ifn/enit-database of handwritten arabic words," in *Proc. of CIFED*, vol. 2. Citeseer, 2002, pp. 127–136.

[13] N. Arica and F. T. Yarman-Vural, "An overview of character recognition focused on off-line handwriting," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 31, no. 2, pp. 216–233, 2001.

[14] A. Amin, "Off-line arabic character recognition: the state of the art," *Pattern recognition*, vol. 31, no. 5, pp. 517–530, 1998.

[15] J. Wightwick and M. Gaafar, *Mastering Arabic script: a guide to handwriting*. Palgrave Macmillan, 2005.

[16] A. Amin, H. Al-Sadoun, and S. Fischer, "Hand-printed arabic character recognition system using an artificial network," *Pattern recognition*, vol. 29, no. 4, pp. 663–675, 1996.

[17] L. M. Lorigo and V. Govindaraju, "Offline arabic handwriting recognition: a survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 5, pp. 712–724, 2006.

[18] M. T. Parvez and S. A. Mahmoud, "Offline arabic handwritten text recognition: a survey," *ACM Computing Surveys (CSUR)*, vol. 45, no. 2, p. 23, 2013.

[19] B. Q. Huang, Y. Zhang, and M. T. Kechadi, "Preprocessing techniques for online handwriting recognition," in *Proceedings of the Seventh International Conference on Intelligent Systems Design and Applications*. IEEE Computer Society, 2007, pp. 793–800.

[20] J. H. Y. AlKhateeb, "Word based off-line handwritten arabic classification and recognition: design of automatic recognition system for large vocabulary offline handwritten arabic words using machine learning approaches," Ph.D. dissertation, University of Bradford, 2010.

[21] H. Freeman, "On the encoding of arbitrary geometric configurations," *Electronic Computers, IRE Transactions on*, no. 2, pp. 260–268, 1961.

[22] S. Madhvanath, G. Kim, and V. Govindaraju, "Chaincode contour processing for handwritten word recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 21, no. 9, pp. 928–932, 1999.

[23] M. S. Khorsheed, "Off-line arabic character recognition–a review," *Pattern analysis & applications*, vol. 5, no. 1, pp. 31–45, 2002.

[24] H. Almuallim and S. Yamaguchi, "A method of recognition of arabic cursive handwriting," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 5, pp. 715–722, 1987.

[25] H. Goraine, M. Usher, and S. Al-Emami, "Off-line arabic character recognition," *Computer*, vol. 25, no. 7, pp. 71–74, 1992.

[26] R. El-Hajj, L. Likforman-Sulem, and C. Mokbel, "Arabic handwriting recognition using baseline dependant features and hidden markov modeling," in *Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on*. IEEE, 2005, pp. 893–897.

[27] M. Dehghan, K. Faez, M. Ahmadi, and M. Shridhar, "Handwritten farsi (arabic) word recognition: a holistic approach using discrete hmm," *Pattern Recognition*, vol. 34, no. 5, pp. 1057–1065, 2001.

[28] M. Pechwitz and V. Maergner, "Hmm based approach for handwritten arabic word recognition using the ifn/enit-database," in *2013 12th International Conference on Document Analysis and Recognition*, vol. 2. IEEE Computer Society, 2003, pp. 890–890.

[29] M. S. Khorsheed, "Recognising handwritten arabic manuscripts using a single hidden markov model," *Pattern Recognition Letters*, vol. 24, no. 14, pp. 2235–2242, 2003.

[30] S. Alma'adeed, C. Higgins, and D. Elliman, "Off-line recognition of handwritten arabic words using multiple hidden markov models," *Knowledge-Based Systems*, vol. 17, no. 2, pp. 75–79, 2004.

[31] M. Pechwitz, V. Maergner, H. El Abed *et al.*, "Comparison of two different feature sets for offline recognition of handwritten arabic words," in *Tenth International Workshop on Frontiers in Handwriting Recognition*, 2006.

[32] V. Märgner, H. El Abed, M. Pechwitz *et al.*, "Offline handwritten arabic word recognition using hmm-a character based approach without explicit segmentation," in *Actes du 9ème Colloque International Francophone sur l'Ecrit et le Document*, 2006, pp. 259–264.

[33] A. Kundu, T. Hines, J. Phillips, B. D. Huyck, and L. C. Van Guilder, "Arabic handwriting recognition using variable duration hmm," in *Document Analysis and Recognition, 2007. ICDAR 2007. Ninth International Conference on*, vol. 2. IEEE, 2007, pp. 644–648.

[34] S. M. Touj, N. Ben Amara, and H. Amiri, "A hybrid approach for off-line arabic handwriting recognition based on a planar hidden markov modeling," in *Document Analysis and Recognition, 2007. ICDAR 2007. Ninth International Conference on*, vol. 2. IEEE, 2007, pp. 964–968.

[35] M. S. Khorsheed, "Hmm-based system for recognizing words in historical arabic manuscript," *International Journal of Robotics and Automation*, vol. 22, no. 4, pp. 294–303, 2007.

[36] R. Al-Hajj Mohamad, L. Likforman-Sulem, and C. Mokbel, "Combining slanted-frame classifiers for improved hmm-based arabic handwriting recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 7, pp. 1165–1177, 2009.

[37] Y. Kessentini, T. Paquet, and A. Ben Hamadou, "Off-line handwritten word recognition using multi-stream hidden markov models," *Pattern Recognition Letters*, vol. 31, no. 1, pp. 60–70, 2010.

[38] M. Khalifa and Y. BingRu, "A novel word based arabic handwritten recognition system using svm classifier," in *Advanced Research on Electronic Commerce, Web Application, and Communication*. Springer, 2011, pp. 163–171.

[39] L. Souici-Meslati and M. Sellami, "A hybrid approach for arabic literal amounts recognition," *Arabian Journal for Science and Engineering*, vol. 29, no. 2, pp. 177–194, 2004.

[40] N. Farah, L. Souici, L. Farah, and M. Sellami, "Arabic words recognition with classifiers combination: An application to literal amounts," in *Artificial Intelligence: Methodology, Systems, and Applications*. Springer, 2004, pp. 420–429.

[41] M. M. M. Fahmy and S. A. Ali, "Automatic recognition of handwritten arabic characters using their geometrical features," *Studies in Informatics and Control*, vol. 10, no. 2, pp. 81–98, 2001.

[42] S. Haboubi, S. Maddouri, N. Ellouze, and H. El-Abed, "Invariant primitives for handwritten arabic script: A contrastive study of four feature sets," in *Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on*. IEEE, 2009, pp. 691–697.