SN Computer Science (2020) 1:152
https://doi.org/10.1007/s42979-020-00168-1

SURVEY ARTICLE

A Survey on Arabic Handwritten Character Recognition

Amani Ali Ahmed Ali[1,2] · **M. Suresha**[2] · **Heba Ali Mohsin Ahmed**[3]

Received: 27 March 2020 / Accepted: 30 March 2020 / Published online: 6 May 2020

Abstract

There are much heavy studies on handwritten character recognition (HCR) for nearly previous four decades. The research on some of the common script like Arabic, Indian and Chinese has been done. This manuscript presents a survey of character recognition on Arabic script, and most of the popular published paper methods are summarized and also analyzed different methods for building a robust system of HCR and included some future research on recognition direction of handwritten character. The paper analyzed and presented various algorithms with respect to preprocessing methods, segmentation methods, feature extraction methods and various classification approaches of the Arabic character recognition.

Keywords Character recognition of Arabic handwritten · Artificial neural network · Freeman chain code · Fuzzy systems · Genetic algorithms · Hidden Markov model · Neural network · Support vector machine

Introduction

Recognition system of Arabic character endeavor to recognize text characters which converts into machine executable format (digital format) that can be treated thorough characters' processor software either offline or online. Character recognition is active area in pattern recognition and image processing due to its application in various fields and can contribute massively to the automation process development and can develop the interaction among human and machine in many applications, comprising of check verification, office automation and data entry applications.

Arabic character recognition and its nature is a complex problem in numerous applications of artificial intelligence, particularly in the time of characters' recognition of connected cursive. This is because of both the Arabic databases dictionaries and scarcity and the Arabic writing cursive nature in both forms of handwritten and printed. Most recognition methods of characters utilized in Arabic characters' recognition from existing methods are adopted to utilize over handwritten characters of Chinese ad Latin; yet, other methods are improved for segmentation of Arabic character only.

Another dimension of Arabic characters' complexity forms different shapes based on their locations inside the word. The result, unexplored character recognition of handwritten Arabic has.

Character recognition categorizes into two recognitions of Online character and Offline character. Offline recognition means that the handwritten text on a paper optically is captured that then existing like an image. Online recognition means that the text written on any digital device is performed as time function and strokes' order. Recognition of online character is more robust in comparison case with recognition of offline character since extra information about various writing styles, pressure, strokes order etc. is available in case of online character recognition.

Huge studies over handwriting recognition of various languages like Japanese, Chinese and English are found. By contrast, researches on character recognition of Arabic script do not receive enough attention.

Neural network approach considered as arising technique in handwritten character recognition field by utilizing ANN indicated to artificial neural network implementations was that networks employ specific learning rules to change the

This article is part of the topical collection "Advances in Computational Approaches for Artificial Intelligence, Image Processing, IoT and Cloud Applications" guest edited by Bhanu Prakash K N and M. Shivakumar.

✉ Amani Ali Ahmed Ali
dramaniali2@gmail.com

[1] Taiz University, Taiz, Yemen

[2] Kuvempu University, Shimoga, India

[3] Thissufal, Ibb, Yemen

Table 1 Shapes of Arabic characters in different locations

No.	Name	Isolated	Beginning	Middle	End	No.	Name	Isolated	Beginning	Middle	End
1	Alif	ا	ا	ـا	ـا	15	Dhad	ض	ضـ	ـضـ	ـض
2	Baa	ب	بـ	ـبـ	ـب	16	Tah	ط	طـ	ـطـ	ـط
3	Tea	ت	تـ	ـتـ	ـت	17	Dha	ظ	ظـ	ـظـ	ـظ
4	Thea	ث	ثـ	ـثـ	ـث	18	Ain	ع	عـ	ـعـ	ـع
5	Jam	ج	جـ	ـجـ	ـج	19	Ghen	غ	غـ	ـغـ	ـغ
6	Haa	ح	حـ	ـحـ	ـح	20	Fa	ف	فـ	ـفـ	ـف
7	Khaa	خ	خـ	ـخـ	ـخ	21	Qaf	ق	قـ	ـقـ	ـق
8	Daal	د	د	ـد	ـد	22	Kaf	ك	كـ	ـكـ	ـك
9	Thaal	ذ	ذ	ـذ	ـذ	23	Lam	ل	لـ	ـلـ	ـل
10	Raa	ر	ر	ـر	ـر	24	Meem	م	مـ	ـمـ	ـم
11	Zaay	ز	ز	ـز	ـز	25	Noon	ن	نـ	ـنـ	ـن
12	Seen	س	سـ	ـسـ	ـس	26	Ha	ه	هـ	ـهـ	ـه
13	Sheen	ش	شـ	ـشـ	ـش	27	Waw	و	و	ـو	ـو
14	Sad	ص	صـ	ـصـ	ـص	28	Yaa	ي	يـ	ـيـ	ـي

links which are weights among their nodes. Such networks can be fed with data from an input picture and trained to output characters in one or more forms [1]. Many ANN's structures are found as Adaline, back-propagation, Kohonen, Madaline, Perceptron and many others. Back-propagation ANN considered as most common method utilized since it is effective and has very simple implementation [2].

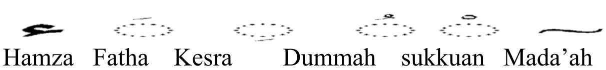

Fig. 1 Diacritic marks

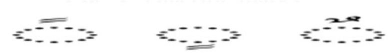

Fig. 2 Various forms of Tanween

Characteristic of Arabic Characters

Arabic utilized as the main language in all Arab countries of East, Northern and Middle Africa. People original language in 26 countries is the Arabic with approximately 300 million from native speakers and 250 million from non-native speakers. It is also central to other languages in the Muslim world such as Farsi (Persian), Pashto, Sindhi and Urdu. Some languages in China like Kirghiz, Kazakh and Uighur are all written utilizing the modified script of Arabic.

Arabic language has 28 characters, and its words from right side to left side are written with no upper case or lower case. Each character has various shapes relying the character location within the word considered as character isolated, middle end and beginning as displayed in Table 1. Arabic characters built with two components which are one fundamental component along with secondary components. Usually the secondary component size is smaller than that of the main component. Both the secondary component and the character shape are used to distinguish between the characters in Arabic language. The number of secondary components varies from one to three points and may be located on either the upper, lower or middle part of the character. The improvement in Arabic recognition systems is a more challenging task.

A character may be formed with mark of diacritic or a vowel written under or over it, as in Fig. 1. Kesra always wrote below letters. Mada'ah, Dummah, Sukkun and Fatha always wrote above letters. Special character or diacritic which can be written with different positions can be Hamza. Diacritics are signs that form short vowels characters or other character sounds, like endings of syllable and Tanween which represent the n sound addition at word end which is usually represented with double Fatha (◌ً), Dummah (◌ٌ) or double Kesra (◌ٍ) as shown in Fig. 2, which are written below the character called double Kesra and which are written over the character called double Fatha and double Dummah.

Another diacritic mark is the Sha'ada which is written above the letter in the Fatha, Kesra and Dummah case, Sha'ada looks like number 3 rotated by 90° clockwise as written in Fig. 3. A Sha'ada utilized to display that there is double of the character which stress the syllable; meaning the character Sukkun after another character of the same along with a various diacritic marks is came; so to avoid the

Fig. 3 Different forms of Sha'ada

ئـ ؤ ئ إ أ ء

Fig. 5 Different forms of Hamza

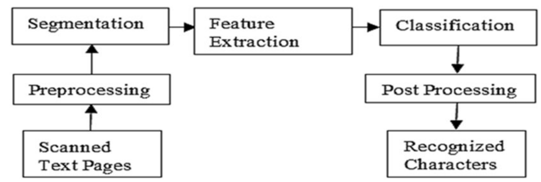

Fig. 6 Stages of Arabic character recognition

case of writing two same characters, a Sha'ada wrote along with short vowel diacritic implicitly forming another character that is not wrote explicitly. For the purpose of recognize among words, the diacritics are very significant in writing, a single diacritic can change completely the meaning of word. For example, "ذهبْ" which is Arabic word with a sukkun above the last letter baa pronounced "thahub" which means gold and with diacritic Fatha above character baa make the word pronounced "thahaba" meaning went which is past tense of word go.

In regular Arabic writing, diacritics not broadly utilized and usually the professional readers understand from the context the meaning of word. But for non-Arabic reader, it may be difficult to capture the meaning and could lead to misunderstanding of the text being read.

Several Arabic alphabet characters share the same shape and are distinguished only in number and position terms of dots on the letters. These dots might be ascenders if put above the Arabic character, or desenders if located below character. Basic shapes are shown in Fig. 4, there are two shapes of which each is used for three letters, six shapes of which each is used for two letters and the remaining ten are used for one letter each.

Arabic writing has also some ligatures those are exceptions to linking character's rules. Most commonly utilized is the laam-alif, "لا", which is the combination of "laam" and "alif" and so many others. These ligatures are not standard and are not presented by all fonts. Various fonts may contain various ligatures depending on the writing style presented by the font.

Additional characters: Ta-Marbuta, "ة ـة", Hamza, ء and Alif-Maqsurah, ى. The Ta-Marbuta can only be added to a character ending and is a special form of the letter. It can be written in two forms: isolated (ة) or final (ـة). Examples of words with Ta-Marbuta are: عائشة, وردة. The Hamza, ء, can be considered as an additional letter, descender or ascender also may be written with various forms; every form is offered a different name. It can be put above the character alif, below character alif and above waaw, by itself with different forms as displayed in Fig. 5.

Hamza is also associated with some characters as displayed in Fig. 5. Also it can be part of the character kaaf, ك. The Alif-Maqsurah, ى, is pronounced Alif (ا) and it is always written at a word ending which is not last letter in the Arabic alphabet ي and is distinguished by the two dots absence.

There are nine Arabic letters; Sad (ص), Dhad (ض), Tah (ط), Dha (ظ), Fa (ف), Qaf (ق), Meem (م), Ha (ه) and Waw (و) that have closed loops. This made the closed loop a significant feature in recognition characters of Arabic script.

Literature Review

Character recognition system involved several stages as displayed in Fig. 6. Individual stages are discussed in detail in the following paragraphs.

Preprocessing

The preprocessing phase goal is used to remove each element in word image which are not helpful in process of recognition. Preprocessing stage involves several techniques utilized to image enhancement to make appropriate segmentation, which are normalization; normalization is the process to converting to standard size for images of the random size. In [3], found a popular defined size, skeletonization and thinning. The technique of thinning will convert the binary image of character into one pixel thick image used by [4]. Thresholding in this technique includes converting gray input images into document images of binary type for getting good features from the image [5], skew detection and correction [6], noise removal; the basic noise objective removal which is to eliminate all bit-patterns which are unwanted those did not have any value in the result. The technique of median filtering utilized in [3] for the purpose of remove noise, horizontal projection profile in [7], filterization in [8] and smoothing in [9–11].

ف ق ك ل م ن ه و ي	ع غ	ط ظ	ص ض	س ش	ر ز	د ذ	ج ح خ	ب ت ث	ا

Fig. 4 Arabic alphabets

Segmentation

The segmentation is considered as the most complex stage in recognition. The goal of segmentation is to dividing the text into its sub-units such as lines, words, characters or strokes. Segmentation is a significant stage because it has an effect on the recognition rate [12, 13] and the basic source of errors in recognition. There are several methods used for this purpose.

Segmenting a line into words; after segmenting the document into lines, the line is segmented into words. Segmenting the lines into words depend on the space between the words. Longer spaces separate words while short spaces split sub-words. Some of the methods analyze the distances between connected components to segment the words [14]. Word segmentation into characters; this part of segmentation includes segmenting a word to individual characters. The Arabic cursive nature made words' segmentation into individual characters a complex work. There are different techniques applied for character segmentation [12, 15]. Technique of segmentation depended on thinning where the character skeleton provides fundamental information about character shape. Number of algorithms suggested for extracting the skeletons [16]. In [17] traced from right side to left side the thinned word for the purpose of the segmented detected points.

Segmentation technique relying on subset approach, in [18] another segmentation method used for text recognition which depends on analysis of subset of documents which are three main subset domains. To avoid detection, failed authors used the skeletonization method, which did the correction in the situation of false alarms to build the segmentation of vertically connected characters.

Segmentation technique relying artificial neural networks denoted by ANNs utilized to achieve the valid segmentation points. In [19], authors proposed a technique for segmenting Arabic handwritten texts based on ANNs. Authors identified points of pre-segmentation through utilizing topographic features of every characters connected block like holes and density of black pixel. The ANNs then used to achieve these segmentation points. The potential manually points are classified to invalid and valid segmentation points also those points pass as input with their features to ANNs to do training [15].

In [20] proposed a new novel algorithm for vertical segmentation. For the purpose to locate the segmentation points used the word image thinning to fetch the width of stroke of one pixel and to detect the ligatures of Arabic characters' geometry and shape are used in the segmentation procedure. The proposed segmentation approach is worked with touching characters, in case of ligatures touching segmentation existing between the characters of consecutive closed and the existence of ligatures within characters in the case of open characters.

Feature Extraction

It considered as the pattern recognition heart. In Arabic documents of handwritten also printed, the features are essential representation of the information which extracted from the text image. This information should have been identical characteristics of the character or the word which make it different from another. Feature extraction is very problem dependent [14]. Features are extracted from the image of a word/character which is expected to represent the shape of the images. This information goes onto the classifier to help in the classification process. Extraction methods of feature are different from one application to another. Techniques that succeed in one application may not be successful for other applications [21]. So, the appropriate selection technique for extract features stills the most significant step for gaining high accuracy rate in recognition. Values which are distinguished for objects in various categories and similar for objects in the same class are considered as good features. Features of Arabic handwritten texts can be categories into the following:

Structural Features

It is like loops, dots, branch points and endpoints. These types of features were used by [22–26]. In [23], structural features used the length and width of the character, if there is a right or left character to connect to it, if it has a loop or not and if there is a character complement such as the shape of zigzag (Hamza), three point, one point or two point. In [24] utilized features like holes, loops, cusps, vertical lines and strokes. In [25, 26] used the features dots and their position, directions, strokes, width and height of the stroke, intersection of line segments and loops.

Statistical Features

Statistical features analyze the spatial pixels' distribution by counting local features at every pixels and leading a statistics set from local features distributions [27]. Statistical features as pixel densities, chain code directions histograms, moments and zoning. Pixel densities features used in [28–30]. In [28, 30] number of pixels and sum of pixels of white and black were used, detecting black and white points were used as statistical features. Moment invariants were used by [31] used scale, orientation and center of gravity. Vertical and horizontal projections used in [7, 23]. In [23], used the longest spike which represented the baseline. In [4], used start-point also end-point of a character features and branch. Wavelet transforms features used by [32]. Determine the body and secondary part, position of the part above or below, loop and Radon transform of the characters were used in [5, 33]. Code of chain feature used in [34]. Zoning where a character separates into no overlapping or overlapping parts and the density character distribution pixels in various regions are analyzed. In [35], authors measured the character contour direction through separating the image character into zones. Then, histograms of chain codes are used to compute contour direction in these regions.

Global Transformation

This technique converts pixel representation into more compact shape. These techniques represent the signal through linear series combination of simpler well-defined functions. The series expansion fits compact encoding by the linear combination coefficients [36]. Common transformation techniques used in character recognition are: Discrete Cosine Transform [37], Hough Transform [38], Fourier Transform [39] and Wavelets [40].

Classification Approaches

Classification is the fundamental character recognition system stage in which the features in feature extraction stage are fed as input to the model for identification and recognition. Common methods in the classification stage are based on: HMM which indicated to Hidden Markov Model, SVM which indicated to support vector machines, artificial neural networks which denoted by ANN, k-NN which indicated to fuzzy logic, k-nearest neighbors, genetic algorithm and others. In [41], methods are applied to achieve good segmentation to get better recognition rate using the Hough transform and skeletonization for text segmentation and Gaussian mixture modeling framework method. That covers the possible false correction to create proficiency vertical connected characters' text segmentation. The segmentation of the Arabic words is pointed as a two class problem.

K-Nearest Neighbor

It is denoted by K-NN that is among the easiest of every machine learning algorithm. It is a method utilized for objects classification relying inside feature space the closest examples of training [42] via the object majority vote being assigned to class most popular between its k closest neighbors with its neighbors the object is classified.

Hidden Markov Models

Hidden Markov Models use a sliding window of text line image to convert a two-dimensional image to a one-dimensional feature vector [43]. In [9], authors used a HMMs algorithm with various explicit distributions for the duration of state and fusion structural and statistical features. In [32], authors used wavelet transform features extraction and presented the edge detection of the character features with Hough transform.

Neural Network

The fundamental purpose of research of neural network is to build a machine which works as human brain works. Neural networks utilized in a widely of various fields to get solution for wide issues' range. Unlike brains of human which can memorize and identify the characters like digits or characters, computers handle them like binary graphics. Thus, algorithms are required to recognize and identify all characters. Back-propagation algorithm was used in [11] for recognition. In [8], authors presented sequenced networks to recognition of the characters with 120 features for each character's image. In [4], the neural network algorithm of back-propagation was also used with histogram features. In [22], they have designed radial-basis neural network classifier, investigated and compared among results of four different artificial neural network models.

Fuzzy Logic Approach (FL)

FL is a soft computing approach relying "degrees of truth" instead of popular logic of "true or false". A lot of researchers have utilized FL for recognition [29–31, 34].

Neuro-Fuzzy

Many researchers used Neuro-Fuzzy approach [24, 44, 45]. In [5], they used Fuzzy logic with ART neural networks. In [45], proposed a method to recognition of the isolated characters of Arabic language using fuzzy logic approach. In [30], described an online Arabic handwritten characters' system. In this system, authors used a fuzzy neural network for classifying Arabic characters. Preprocessing methods used smoothing operation and cursive characters. The extractions of features from all characters are the physiological parameters of neuron of the equation displaying the script curvilinear velocity. The recognition stage is the Beta fuzzy neural network which has 100 neurons inside hidden layer and inside output layer 55 neurons. In [3], proposed a recognition system of Arabic characters using wavelet transform to extract features. Authors used an approach of neuro-fuzzy for character recognition.

Table 2 Recognition system of Arabic handwritten character

Authors	Preprocessing	Features extraction/Classification	Details of result/Description
[46]	Segmentation, enhancement of image and noise elimination	Features of statistical such as statistical pixels' distribution and displaying the characteristic pattern measurements, which contained of moments, density pixels' distribution which counts the zeros and ones, zoning. Also structural that is features of topological and geometrical such as the location regarding to the baseline, dots, loops, endpoints and strokes. HMM this indicated to Hidden Markov Model	The best result achieved thorough utilizing multiple HMMs combination that was 95.15%
[47]	Segmentation	Features of statistic and structural. Neural Network	Fusion independent of context- + MLP of grapheme -HMM displayed highest rate of recognition that was 89.42%
[48]	Skew correction and smoothing	pixel zoning, HMM-based feature and features of statistical such as character array zoning like (crossing distances, the occurrence n-tuples of linked or white or black, calculating the character black pixels' moments, the characteristic position and splitting it into zones of non-overlapping or over-lapping)/HMM the symbol of Hidden Markov Model	Using this new feature extraction algorithm, they obtained 98% of accuracy, a significant improvement compared to the best result utilizing the features of hierarchical achieved 81.45%
[49]	Skeletonizing, smoothing, contouring and thresholding	Statistic like moments of invariants, zoning, descriptors of Fourier and chain code of Freeman. Also structural such as concavities, strokes, relations of stroke, loops, line segments intersections and end points. And features/PNN which indicated to probabilistic neural network, K-NN which indicated to the method of K-nearest neighbor, the algorithm of K-Means and FCM which indicated to the algorithm of Fuzzy C-Means	Fusion of three features achieved accuracy of 70%
[50]	Binarization, segmentation and noise reduction	Compute ratio of aspect from graph of skeleton/HMM the symbol of Hidden Markov Model	The best rate of recognition was around 92% along with segmentation
[51]	Binarization	1) An approach of statistical utilized for forming the pixel values distribution of spatial of image with binary type. 2) Count the intersections number along vertical ray of middle and divide the images to eight parts into four horizontal and four vertical. 3) Feature extraction of flexible meshing directional/MLP which indicated to multi-layer perceptron	Utilizing sixteen hidden classifier layer the rate of recognition achieved 97.62% for testing data
[52]	Normalization	Code information of directional chain and under sampled of bitmaps/SVM which indicated to support vector machines	Utilizing 196 directional features with window-map of overlapping achieved the best rate of recognition which was 96.17%
[53]	Noise reduction, word segmentation and binarization	1) Sliding window along with Feature of Pixel. 2) Direction-based features of skeleton. 3) Sliding window along with features of local/HMM	Utilizing the method of skeleton based for direction features of skeleton and baseline estimation gained rates of recognition of 89% over level of word
[54]	Segmentation and normalization.	Features of chain code with 8-FCC/Used neural network	Achieved average rate of recognition of 99.03%

Table 2 (continued)

Authors	Preprocessing	Features extraction/Classification	Details of result/Description
[55]	Detection of baseline, smoothing, contour and normalization	Densities features of black pixel and directional density/HMM this indicated to Hidden Markov Model	A two-level algorithm of decoding was suggested to increase the speed of the procedure of recognition while preserve the accuracy of recognition and decreased the decoding complexity step
[56]	Not specify	Features of chain code with 16-FCC/SVM, neural network and function of radial-basis	Recognition rates of writer-dependent along with their standard deviation relying on the utilized original characters number compared to rates of reference utilizing 30 or 10 letters per class for classifiers of both SVM and RBFN
[57]	Not specify	Features of chain code with histogram of cyclic FCC and 8-FCC/Pruning/Growing + Sigmoid RBF	Without deformation of skew and slant, feature of CCH created a rate of recognition around 90.1% also 91.4% was for cyclic-CCH. With deformation of skew and slant, the rate of recognition of feature of CCH reduced into 60.3%, also cyclic-CCH reduced into 67.9%
[58]	Thinning	features of chain code with 8FCC/used neural network	Suggested performs of PSO better than the suggested DE which can be appeared through comparing their result of standard deviation, max and average. The time of computation reduced twice

Approaches Summary

In system of recognizer, the most popular approaches are SVM which indicated to support vector machine, ANNs which indicated to artificial neural networks and HMMs which indicated to Hidden Markov Models. But, some of multiple classifier or fusion displayed very excellent and accurate result comparing to others. Different techniques of feature extraction have been discussed and noticed that also scripts type affect the accuracy results. For example, in Arabic script, the dots in among of characters and on below afford researchers many challenges to solve them while characters of Chinese script have many strokes which vary among one individual to another. Handwriting with slant is so complex to recognize which needs appropriate stage of preprocessing to correct them before features extraction. Also continuous writing requires segmentation into isolate characters. In Table 2, discussed handwritten character recognition phase contains classification, feature extraction and preprocessing for Arabic script. In Table 3, discussed some research paper which utilized FCC.

Different Methods Comparison

Recognition of handwriting as mentioned earlier can be done through different methods. From the methods discussed above for Arabic script, often neural network is a method utilized for recognition of writing, containing handwriting. In almost all situations, the neural network accuracy is very high. Other methods are widely reliable with recognition of writing. The simplicity of method did not determine that method will be good for recognition of handwriting. Recognition of handwriting must be accompanied through a procedure of training and also more samples would be better. Every method has their advantages and disadvantages. In methods of HMM, the weakness is that a problem with segmentation. For the method of zoning, the weakness is that in the document image must be many zones due to otherwise the recognition level of accuracy will be minimal. In the neural network, the time needed for long time of training because neural network is containing in the study of deep learning, but neural network has good handwriting recognition accuracy due to more training of neural network will result with more precise recognition of writing. In Table 4, displayed some of other popular and common methods in Arabic script.

Conclusion

The purpose of this paper to supply a detailed survey of published work of research in the sequential stages of Arabic handwriting character recognition along with special concern toward each stage techniques. Different works on

Table 3 Freeman chain code in handwritten character recognition

Authors	Preprocessing	Features extraction/classification	Result details/description
[41]	Not specify	Features of chain code with 16-FCC/SVM, neural network and function of radial basis	Recognition rates of writer-dependent along with their standard deviation relying on the utilized original characters number compared to rates of reference utilizing 30 or 10 letters per class for classifiers of both SVM and RBFN
[50]	Not specify	Features of chain code with Histogram of cyclic FCC and 8-FCC/pruning/growing + sigmoid RBF	Without deformation of skew and slant, feature of CCH created a rate of recognition around 90.1% also 91.4% was for cyclic-CCH. With deformation of skew and slant, the rate of recognition of feature of CCH reduced into 60.3%, also cyclic-CCH reduced into 67.9%
[51]	Thinning	Features of chain code with 8FCC/ used neural network	Suggested performs of PSO better than the suggested DE which can be appeared through comparing their result of standard deviation, max and average. The time of computation reduced twice

Table 4 The disadvantages and advantages among methods

Method	Advantages	Disadvantages
Neural network	It has been proven successful used for many handwriting and computer recognition. When CNN is trained, then the image recognition will be accurate	Long training time. In order for high accuracy in the training process should use many samples
Semi-incremental	Considers the latest strokes and previous segments. Time of waiting won't be so visible	The way it works is more complex than the pure incremental method. It should be accompanied by other methods can not only semi-incremental only
HMM and incremental	The recognition stage is simpler than semi-incremental, also is one of the simplest factors that make it simpler is because it only sees the state	There is a problem with the segmentation process
Line and word segmentation	Line and words segmentation approach for printed documents to project a powerful process	Can't detect pattern
K-Nearest neighbor	Reliable enough for writing recognition because of low accuracy	Low accuracy. In order for high accuracy in the training process should use many samples
Slope and slant correction	Segmentation process simpler and the accuracy of the writing recognition	Not maximal for handwriting recognition
Neuro-fuzzy and ensemble	Accuracy results are quite high	Line segmentation components affect accuracy
Fuzzy logic and zoning	Accuracy results are quite high	The number of zones in an image must be many, because if a little then the accuracy will be smaller

systems of HCR are reported with the trends of research, discussed the method being utilized in the modern systems of HCR, and the hardness happens in the researches. Also in this work discussed the accuracy of recognition and future work. Found that various characteristics such as feature extraction, preprocessing, classification and segmentation method have been utilized with different accuracy levels. The neural network method has the highest accuracy and the method with lowest accuracy. The method of slant and slope correction has the lowest accuracy. Experiments showed that SVM and neural network gained better popularization ability than classifier of both MLP and also K-nearest neighbor. MLP achieves slightly better popularization ability than HMM.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Niu XX, Suen CY. A novel hybrid CNN–SVM classifier for recognizing handwritten digits. Pattern Recogn. 2012;45(4):1318–25.
2. Zeiler MD, Fergus R. Visualizing and understanding convolutional neural networks. arXiv:1311.2901. 2013.
3. Mesleh A, Sharadqh A, Al-Azzeh J, Abu-Zaher M, Al-Zabin N, Jaber T, Odeh A, Hasn M. An optical character recognition. Contemp Eng Sci. 2012;5(11):521–9.

4. Albakoor M, Saeed K, Sukkar F. Intelligent system for Arabic character recognition. In: World congress on nature & biologically inspired computing. New York; IEEE; p. 982–87. 2009.
5. Zaghloul RI, Alrawashdeh EF, Mohammad D, Bader K. Multilevel classifier in recognition of handwritten Arabic characters. J Comput Sci. 2011;7(4):512–8.
6. Ali AAA, Suresha M. A novel approach to correction of a skew at document level using an Arabic script. Int J Comput Sci Inf Technol. 2017;8(5):569–73.
7. Ismail SM, Norul S, Sheikh H. Online Arabic handwritten character recognition based on a rule based approach. J Comput Sci. 2012;8(11):1859–68.
8. Zawaideh FH. Arabic hand written character recognition using modified multi-neural network. J Emerg Trends Comput Inf Sci. 2012;3(7).
9. Benouareth A, Ennaji A, Sellami M. HMMs with explicit state duration applied to handwritten Arabic word recognition. In: 18th international conference on pattern recognition (ICPR'06). New York: IEEE; Vol. 2, p. 897–900, 0-7695-2521-0/06. 2006.
10. Cheung A, Bennamoun M, Bergmann NW. An Arabic optical character recognition system using recognition-based segmentation. Pattern Recogn. 2001;34(2):215–33.
11. Wang X, Huang TL, Liu XY. Handwritten character recognition based on BP neural network. In: Third international conference on genetic and evolutionary computing (WGEC'09). New York: IEEE; p. 520–24. 2009.
12. Alginahi YM. A survey on arabic character segmentation. Int J Docum Anal Recognit. 2002;1–22.
13. Casey RG, Lecolinet E. Strategies in character segmentation: a survey. Int Conf Docum Anal Recognit IEEE. 1995;2:1028–33.
14. Alkhateeb JH, Jiang J, Ren J, Ipson S. Component-based segmentation of words from handwritten Arabic text. Int J Comput Syst Sci Eng. 2009;5(1):54–8.
15. Zeki AM. The segmentation problem in Arabic character recognition the state of the art. In: First international conference on information and communication technologies (ICICT). New York: IEEE; p. 11–26, 0-7803-9421-6/05. 2005.
16. Cowell J, Hussain H. Thinning Arabic characters for feature extraction. In: International conference on information visualization. New York: IEEE; p. 181–185. 2001.
17. Al-Sadoun H, Amin A. A new structural technique for recognizing printed Arabic text. Int J Pattern Recognit Artif Intell. 1995;9(1):101–25.
18. Suresha M, Ali AAA. Segmentation of handwritten text lines with touching of line. Int J Comput Eng Appl. 2018;12(6):1–12.
19. Hamid A, Haraty R. A neuro-heuristic approach for segmenting handwritten Arabic text. In: International conference on computer systems and applications. p. 110–13. 2001.
20. Ali AAA, Suresha M. An efficient character segmentation algorithm for recognition of Arabic handwritten script. In: International conference on data science and communication. New York: IEEE; 978-1-5386-9319-3/19/$31.00 ©2019. 2019.
21. Al-Ma'adeed S. Recognition of off-line handwritten Arabic words. PhD Thesis, University of Nottingham. 2004.
22. Alhamad HA. Use an efficient neural network to improve the Arabic handwriting recognition. In: International conference on signal and image processing applications (ICSIPA). New York: IEEE; p. 269–74. 2013.
23. Aljuaid H, Muhammad Z, Sarfraz M. A tool to develop Arabic handwriting recognition system using genetic approach. J Comput Sci. 2010;6(6):619–24.
24. Almohri H, Gray JS, Alnajjar H. A real-time DSP-based optical character recognition system for isolated Arabic characters using the TI TMS320C6416T. In: Proceedings of the IAJC-IJME international conference ISBN 978-1-60643-379-9. 2008.
25. Amin A, Mari J. Machine recognition and correction of printed Arabic text. IEEE Trans Syst Man Cybern. 1989;19(5):1300–6.
26. Khorsheed MS, Clocksin WF. Structural features of cursive Arabic script. In: Proceedings of tenth British machine vision conference (BMVC). p. 422–31. 1999.
27. Srinivasan GN, Shobha G. Statistical texture analysis. Proc World Acad Sci Eng Technol. 2008;36:1264–9.
28. Abdullah MA, Al-Harigy LM, Al-Fraidi HH. Off-line Arabic handwriting character recognition using word segmentation. J Comput. 2012;4(3):40–4.
29. Abed MA, Abed AL-Asadi HA, Baha Al-Deen ZS, Ismail AN. Fuzzy logic approach to recognition of isolated Arabic characters. Int J Comput Theory Eng. 2010;2(1):119–24.
30. Khedher MZ, Al-Talib G. Recognition of secondary characters in handwritten Arabic using Fuzzy Logic. In: International conference on machine intelligence. 2005.
31. Sarfraz M, Al-Awami ATA. Arabic character recognition using particle swarm optimization with selected and weighted moment invariants. In: 9th international symposium on signal processing and its applications (ISSPA). New York: IEEE; p. 1–4, 1-4244-0779-6/07. 2007.
32. Amor NB, Amara NB. Combining a hybrid approach for features selection and hidden Markov models in multifont Arabic characters recognition. In: IEEE second international conference on document image analysis for libraries 0-7695-2531-8/06. 2006.
33. Abandah GA, Younis KS, Khedher MZ. Handwritten Arabic character recognition using multiple classifiers based on letter form. In: International conference on signal processing, pattern recognition, & applications. p. 128–33. 2008.
34. Hachour O. The combination of fuzzy logic and expert system for Arabic character recognition. In: 3rd international IEEE conference intelligent systems. p. 189–191. 2006.
35. Mohiuddin KM, Mao J. A comparative study of different classifiers for handprinted character recognition. Mach Intell Pattern Recogn. 1994;16:437–48.
36. Alkhateeb JH. Word based off-line handwritten Arabic classification and recognition. PhD Thesis, Computing School, Informatics and Media, Bradford University. 2010.
37. AlKhateeb JH, Ren J, Jiang J, Ipson SS, El-Abed H. Word-based handwritten Arabic scripts recognition using DCT features and neural network classifier. In: 5th international multi-conference on systems, signals and devices. New York: IEEE; p. 1–5. 2008.
38. Ben Amor N, Ben Amara NE. Multifont Arabic characters recognition using hough transform and HMM/ANN classification. J Multi-media. 2006;1(2):50–4.
39. Khorsheed MS, Clocksin WF. Multi-font Arabic word recognition using spectral features. Int Conf Pattern Recognit. 2000;4:543–6.
40. Jannoud IA. Automatic Arabic handwritten text recognition system. Am J Appl Sci. 2007;4:857–64.
41. Ali AAA, Suresha M. Efficient algorithms for text lines and words segmentation for recognition of Arabic handwritten script. Emerging research in computing., Information, Communication and ApplicationsSingapore: Springer; 2019. p. 387–401.
42. Siddharth KS, Jangid M, Dhir R, Rani R. Handwritten Gurmukhi character recognition using statistical and background directional distribution features. Int J Comput Sci Eng. 2011;3(6):2332–45.
43. Parvez MT, Mahmoud SA. Offline Arabic handwritten text recognition: a survey. ACM Comput Surv. 2013;45(2):23.
44. Alimi AM. A neuro-fuzzy approach to recognize Arabic handwritten characters. Int Conf Neural Netw IEEE. 1997;3:1397–400.
45. Alimi AM. An evolutionary neuro-fuzzy approach to recognize on-line Arabic handwriting. Int Conf Docum Anal Recognit IEEE. 1997;1:382–6.

46. AlKhateeb JH, Ren J, Jiang J, Al-Muhtaseb H. Offline handwritten Arabic cursive text recognition using hidden markov models and re-ranking. Pattern Recogn Lett. 2011;32(8):1081–8.
47. Likforman-Sulem L, Mohammad RAH, Mokbel C, Menasri F, Bianne-Bernard AL, Kermorvant C. Features for HMM-based Arabic handwritten word recognition systems. In: Guide to OCR for Arabic scripts. London: Springer; p. 123–43. 2012.
48. Ahmed I, Mahmoud SA, and Parvez MT. Printed Arabic text recognition. Springer, In: Guide to OCR for Arabic scripts. p. 147–68. 2012.
49. Nemouchi S, Meslati S, Farah N. Classifiers combination for Arabic words recognition: application to handwritten Algerian city names. In: International conference on image and signal processing. Berlin: Springer; p. 562–70. 2012.
50. Pechwitz M, El Abed H, Märgner V. Handwritten Arabic word recognition using the IFN/ENIT-Database. In: Guide to OCR for Arabic scripts. Berlin: Springer; p. 169–213. 2012.
51. Moradi M, Poormina MA, Razzazi F. FPGA implementation of feature extraction and MLP neural network classifier for Farsi handwritten digit recognition. In: UKSim European symposium on computer modeling and simulation. New York: IEEE; p. 231–34. 2009.
52. Alaei A, Nagabhushan P, Pal U. A new two-stage scheme for the recognition of Persian handwritten characters. In: International conference on frontiers in handwriting recognition (ICFHR). New York: IEEE; p. 130–35. 2010.
53. El Abed H, Margner V. Comparison of different preprocessing and feature extraction methods for offline recognition of handwritten Arabic words. In: 9th international conference on document analysis and recognition, New York: IEEE; vol. 2, p. 974–78. 2007.
54. Lawal IA, Abdel-Aal RE, Mahmoud SA. Recognition of handwritten Arabic (Indian) numerals using Freeman's chain codes and abductive network classifiers. International conference on pattern recognition. New York: IEEE; p. 1884–87. 2010.
55. Kessentini Y, Paquet T, Hamadou AB. Multi-stream Markov models for Arabic handwriting recognition. In: Guide to OCR for Arabic scripts, Berlin: Springer. p. 335–50. 2012.
56. Bayoudh S, Mouchere H, Miclet L, Anquetil E. Learning a classifier with very few examples: analogy based and knowledge based generation of new examples for character recognition. In: European conference on machine learning. Berlin: Springer; p. 527–34. 2007.
57. Lee CC, Shih CY, Yu CC, Lai WR, Jeng BS. Vision-based fingertip-writing character recognition. Int Signal Process Syst. 2011;64(3):291–303.
58. Hasan H, Haron H, Hashim SZ. Freeman Chain Code extraction using differential evolution (DE) and particle swarm optimization (PSO). In: International conference of soft computing and pattern recognition. New York: IEEE; p. 77–81. 2009.