Cuneiform symbols recognition by support vector machine (SVM)

Ali Adel Saeid   Abdul Monem S. Rahma   Abdul Mohssen .j. Abdul Hossen

Department of Computer Sciences , University of Technology
zeenali736@gmail.com   monem.rahma@yahoo.com   Abdulmoohsen53@yahoo.com

Abstract

Cuneiform character recognition represents a complex problem in pattern recognition as result of problems that related to style of this type of writing and the diversity of its features according to distortion and shadows problems. This research proves that polygon approximation method is an optimal feature extra action method , which has been adopted for recognition task compeer with elliptic Fourier descriptor, according to the achieved high accuracy recognition results after applying multiple classes of support vector machine classifier along with depending on its discriminate functions .This work is applied by using two Data set , the first one contains 320 images of cuneiform symbols patterns for evaluate the optimal feature extraction method. The second contains 240 images of cuneiform characters to evaluate the recognition system, agents training dataset consists of 2D four triangular patterns.

Ali .A/ Abdul Monem .S/ Abdul Mohssen .j

1. Introduction

Cuneiform writing is one of the oldest language systems which emerged in the third millennium BC, where the first character of writing was invented in Urk city in south of Iraq. Writing system is subjected to many stages of development to facilitate its characteristics about the shape of symbols and numbers that represent a development state of old Sumerian scrip language to Babylonian and Assyrian cuneiform language . At the beginning of the 19th century thousands of cuneiform tablets were discovered In Iraq and Iran, Which represent various Assyrian Babylonian and Persian civilizations. Today these tablets reside in many museums and the process of interpretation requires experience and time .However, the need for information technology was required to solve this problem about recognition task, therefore this research 2 presents a recognizing way for the cuneiform symbols by applying OCR principle, through applying its chine sequential principals steps (preprocessing , segmentation , feature extraction and classification). The aim of research is to review a comparison state on features extraction methods, particularly between elliptic Fourier descriptor and polygon approximation methods to evaluate which one of them is adopted to design a recognition system. Where the adopted classifier is a support vector machine (SVM) according its discriminant functions. Finally, it must be noted in this research that the cuneiform character will be treated via collecting of cuneiform symbols, each of them will be classified according to two factors: the first one is to determine the cuneiform patterns and the second is related to its direction.

2-Literature Review

In 2017 Rahma.A. proposed a new method for recognized cuneiform symbols by adopting a polygon approximation as features extraction method. Where K-nearest neighbor classifier is adopted and the recognition raita that was achieved is 91% [1] .In 2014 Mostofi f. proposed intelligent recognition system for Ancient Persian Cuneiform Characters that is based on supervised back propagation neural network model, (classification model). The training data set is created by subjecting the original training set to Gaussian Filter with different values of stander devotion. The otsu's binirized model was adopted for computing global threshold value. The recognition achieved   rate was 89-100 %,[2].In 2013 Naktal M. proposed a method for recognizing the cuneiform symbols depending on statistical and structure features derived by projection histogram, center of gravity

and connected component features. However to separate each distinguish feature according to each class of symbols, the k-mean clustering was used. Multilayer Neural Network (MLP) was applied for classifying a task where the recognition rate of accuracy level was different according to each class from 83.3% to 95.1%[3].In 2001 Al-Aany proposed recognition approach for cuneiform symbols depending on extract recognition features that generated from binary cuneiform image symbol by depending on suggested seven transform forms applied on each pixel's with their neighbors. However, each cuneiform symbol will have distinguish features related to directions used for recognition task .The classification process is implemented by indexing process that was distributed on tree structure [4].

3-Cuneiform writing

The Assyrian cuneiform language represents one of the stages of the development of cuneiform writing in Mesopotamia, which continued from the beginning of the first millennium to 600 BC. It relies on drilling cuneiform symbols on clay tablets or tablets of stone from left to right to form cuneiform groups which reflect the basic language meanings. This language consists of a set of letters. Each one consists of one or set of cuneiform symbols, these symbols or wedges are organized in different directions either horizontal, vertical, Oblique or diagonal, therefore, these letters with their symbols vary from one character to another according to the (number of symbols, their direction and their location)[5] figure (1)
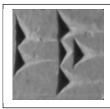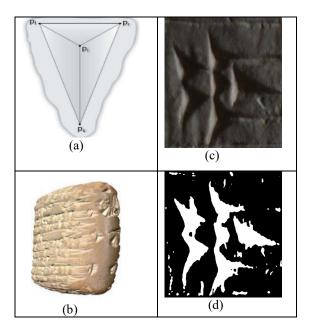

Figure (1) cuneiforms   character images from Iraqi museum

The problems related with recognition tasks have more than one face. The First problem relates to the nature of the writing medium, whether it is stone or clay(2.c), with a three-dimensional writing form. The second relates to the nature of geometry of the cuneiform symbol, which takes the three-dimensional form (three surfaces) figure(2.a),The third problem is the cuneiform writing style, it does not depend on writing on one face but may take writing on all the surfaces of the tablet [5], figure(2.b) .

Ali .A/ Abdul Monem .S/ Abdul Mohssen .j

The last problem that interfaces the cuneiform recognition relates with undesirable results like spots after being subjected to segmentation process,   figure (2.d), where their features differ from one image to another which affect negatively on the recognition task [1].

(a)

(c)

(b)

(d)

Figure(2)   cuneiform writing style . a) the   3-D geometry shape of cuneiform symbol , b) the style of cuneiform   writing take more than one surface. ,c) cuneiform image character ,d) spots problem
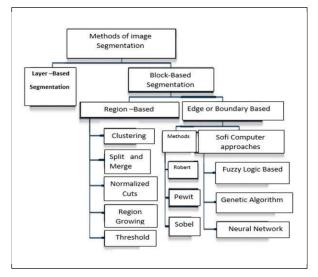
4. Image enhancement:

Image enhancement is one of the most important image processing techniques, which leads to reconstruct the image's features to suit the nature of application's requirements . The primary objective is to treat all the associated problems related with blurring ,contrast and noise. The process of enhancement task takes two directions: the first one submits to human vision as criterion for the evaluation, and the second is moving towards supporting and improving image qualities used to support the identity process by machine vision. Enhancement techniques can be classified into two categories[6]:

1- Frequency Domain.

2- Spatial Domain.

In this research the frequency domain is adopted for applying the enhancement process as the ideal low pass filter.

5.Image segmentation .

It is an image processing technique that leads to segment the image's pixel to segments of regions where each of them has distinguished labels. This simplification process is used to simplify images features to easier or meaningful feature form to be used to support the advanced analysis's or recognition stages. Image segmentation techniques are categorized into two branches: its block and layer based segmentation as seen in following figure (3) [7].

Figure(3):image segmentation techniques

Thresholding is a popular image segmentation technique that is adopted by large number of binriztion methods. It leads to separate the image into two sets group of regions based on selected threshold value (T). If pixel intensity color value is larger than the threshold, it will represent foreground region in the opposite case. It is considered as background, as mathematical formula below.

$$G(x,y)=\begin{cases}1 \, , if \ f(x,y) > T \\ 0 \, , if \ f(x,y) < T\end{cases} \quad …(1)$$

Therefore,to apply image segmentation by thresholding, two formals are adopted to apply this task, these are Niblack and Sauvola's method and choosing one of them depends on statistical Skewness metric [1].

6-Image labeling by Extraction of Connected Components

To reach to labeled image, (Multi-scan strategy) will be applied (which is represented by **Extraction of Connected Components**) on binary image, *A* as it contains foreground pixels with labeled value equals (1) and background, their pixel labeled value is (0). This process is implemented iteratively with restricted condition depending on dilation concepts.

Ali .A/ Abdul Monem .S/ Abdul Mohssen .j

Initial step starts by locating the first foreground pixel (p) which it represents seed point for reconstructed matrix X_K with the structure element **B,** scan the image **A**   for computing the following form[8],[9] .

$$X_k = (X_{k-1} \oplus B) \cap A, \quad k = 1, 2, 3, \ldots \quad \ldots(2)$$

let X_0=p where K=0,1,2…n

This iterative process would be terminated after the terminated condition was satisfied   as $X_K=X_{K-1}$. For applying image labeling concept, each regenerated connected   component   will   be   assigned   by **distinguished label** for each connected component element

7. Elliptical Fourier Descriptor:

Elliptical Fourier Descriptor (EFD) represents boundary shape descriptor where the   classification feature   is   generated   from   Fourier coefficients [ai,bi,ci,di]   about each **N** harmonic to recognize closed contour of **k** elements   , Where harmonic coefficients are defined as follows[10]:

The closed coordinate points on N harmonic can be calculated as follows :

$$a_n = \frac{T}{2n^2\pi^2} \sum_{i=1}^{k} \frac{dx_i}{dt_i} \left[\cos\frac{2n\pi t_i}{T} - \cos\frac{2n\pi t_{i-1}}{T}\right] \quad \ldots(3)$$

$$b_n = \frac{T}{2n^2\pi^2} \sum_{i=1}^{k} \frac{dx_i}{dt_i} \left[\sin\frac{2n\pi t_i}{T} - \sin\frac{2n\pi t_{i-1}}{T}\right] \quad \ldots(4)$$

$$c_n = \frac{T}{2n^2\pi^2} \sum_{i=1}^{k} \frac{dy_i}{dt_i} \left[\cos\frac{2n\pi t_i}{T} - \cos\frac{2n\pi t_{i-1}}{T}\right] \quad \ldots(5)$$

$$d_n = \frac{T}{2n^2\pi^2} \sum_{i=1}^{k} \frac{dy_i}{dt_i} \left[\sin\frac{2n\pi t_i}{T} - \sin\frac{2n\pi t_{i-1}}{T}\right] \quad \ldots(6)$$

$$X_i = X_c + \sum_{n=1}^{N} a_n \cos\frac{2n\pi t_i}{T} + b_n \sin\frac{2n\pi t_i}{T} \quad \ldots(7)$$

$$Y_i = Y_c + \sum_{n=1}^{N} a_n \cos\frac{2n\pi t_i}{T} + b_n \sin\frac{2n\pi t_i}{T} \quad \ldots(8)$$

Xc and Yc is centroid's coordinate. Defined by below equations.

$$X_c = \frac{1}{T} \sum_{I=1}^{K} \frac{dx_i}{2d_i}(t_i^2 - t_{i-1}^2) + \beta_i(t_i^2 - t_{i-1}^2) \quad \ldots(9)$$

$$y_c = \frac{1}{T} \sum_{I=1}^{K} \frac{dy_i}{2d_i}(t_i^2 - t_{i-1}^2) + \alpha_i(t_i^2 - t_{i-1}^2) \quad \ldots(10)$$

T= The length of chain code.

$$dt_i = 1 + \left(\frac{\sqrt{2}-1}{2}\right)(1-(-1)^{ui}) \quad \ldots(11)$$

$$t_n = \sum_{i=1}^{n} dt_i \quad \ldots(12)$$

8. Polygon approximation with Dominant point approaches

In 2007 Asif Masood [11] proposed a new approach for polygon approximation, it is defined as **revers polygonization** principles which provide   a good representation about boundary of 2D shapes with high

accuracy for   data reduction , shape matching and pattern recognition [11]. Basically, this   method implementation depends on dominant points (DP) concept , assigned for endpoints of approximated line segments ,where the initial set of (DP) starting with detected break points(BP) is extracted from shape boundary. However, (BP) is derived after applying **freeman's chain code** on shape boundary points. Therefore, any boundary point is defined as   break point (BP) that its chain code (Ci) value is not equal the chain code of prior point [12][11], However to apply polygon approximation technique, there are two optimization approaches being used to solve this approximation problem[14][13] .

1-Min-ϵ **problems:** let polygon **P** vertices to be approximated   by   another   polygon   **Q**   with predefined number of segments **M** along with minimizing approximation error.

2-Min-# **problems**: let polygon **P** vertices to be approximated   by   another   polygon   **Q**   with minimum numbers of   segments M,   where the error dose not skip the predefined tolerance value .
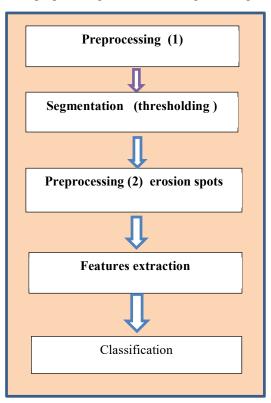
9.Support Vector Machine (SVM)

In 1995, Vapnik proposed a binary classification model for the supervised learning, it is a linear classifier in a feature space. Generally, it is a classification task with two classes once the training parameter is determined, however this model is based on custom functions for classification .It is   defined   as   kernel   functions. Therefore the SVM algorithm with initial form is like a decision boundary that separates between two classes, but this model can be improved to increase the separation   process   in new space ,Which results in creating non liner decision boundary (Kernel function) . The Kernel functions(K) in table(1) was used by SVM to map input data feature space to high feature space when   the   training data   are   not   linearly separable**[15],[16]**.

table 1: kernel functions models

Core	Formula
Linear	K(x,y) = x.y
Sigmoid	K(x,y) = tanh(ax.y +b)
Radial Basis Function (RBF)	K(x,y) = exp(-‖x-y‖2/σ2)
Polynomial	K(x,y) = (ax.y +b)^d

Ali .A/ Abdul Monem .S/ Abdul Mohssen .j

10.Proposed recognition system

This section presents the main diagram figure(4) for cuneiform recognition system starting with preprocessing stage (1)to image segmentation , preprocessing stage (2),features extra action and classification , in addition to review the proposed algorithms according each stage.
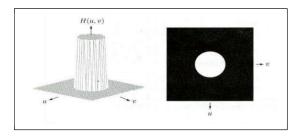

Figure(4):proposed recognition system

10.1Preprocessing stage (1)

Image enhancement. In this research the frequency domain is adopted for applying the enhancement process. The ideal low pass filter is dependent ,where as ideal low pass filters can be defined as follows:-
:-

$$H(u,v) = \begin{cases} 1 & if\ D(u,v) \leq D0 \\ 0 & if\ D(u,v) > D0 \end{cases} \quad ....\ (13)$$

Where D0 is non negative value which represents the radios of cutoff frequency and $D(u,v)$ is distance value starting from point to center of frequency figure (5).

Figure(5): ideal low pass filter.

10.2. Image segmentation (Thresholding)

Therefore for applying image segmentation by thresholding, two formals that are adopted to apply this task are Niblack and Sauvola methods, these are defined respectively in following forms, choosing one of them depends on statistical Skewness metric [1].

$$T=M+k\sigma\ldots(14)$$

$$T=m(1-k(1-\frac{\sigma}{r}))\ldots(15)$$

Where k is constant, (m, σ) represents the mean and standard deviation respectively.

10.3 Preprocessing stage (2)Spot removing

For eliminate the unwanted elements segments like (spots) ,figure(6.b) that resulted from subjecting the gray cuneiform image ,figure (6.a) to thrsholding process. However the target of this process is to create uniform features about cuneiform image symbols to be clear from their elements figure (6.c). Therefore to satisfy this principle, this research adopts the Image connected-component labeling (CCL) concepts to erase the spots according to bellow algorithm .This process get satisfied after applying bellow Algorithm several times to extra each connected elements and assign distinguish   label for each one and apply the erosion process for each segments which has small ratio to all image's pixels.

Ali .A/ Abdul Monem .S/ Abdul Mohssen .j

Algorithm (1) : Connected Components Extraction. **Input :Binary image.** **Output : Connected Components image**
Step1: read input binary image IB

Step2: locate the first foreground pixel p and it's location p(x,y).

Step3:initlize the stricture element B .

Where B={1…9}=1;

Step4:k=0;

Step4:intilize the connection component matrix $X_K(0,0)$.

Step5 :set $X_K(x,y) = p(x,y)$.

Step6: **repeat**

Y=Xk

Applied the dilation process on X_k and interest the result with original

Image *IB* as following formula.

X_{k+1}= dilation(B,X_K)∩ *IB* .

K=K+1.

Untiled (Y==X_{K+1})

Step7: set CC_MATRIX=Y.

Step8:return (CC_MATRIX).
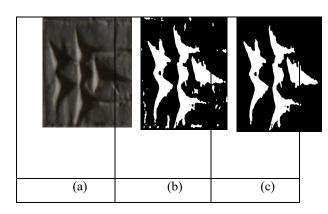

(a)	(b)	(c)

Figure (6) spots erosion a) cuneiform image ,b) binary image with spots ,c) spots off image

10.3.Feature extraction

This section presents two features extraction methods depending on boundary pixels to apply the comparison state between them

10.3.1Feature extraction by Elliptical Fourier Descriptor (EFD).

The features vector generated by EFD is based on quadruple Fourier coefficient (*ai,bi.ci,di*) values that are defined by equations (3-6) ,where the predefined number of harmonic *N*, determines the size of features vector. That means that each one of these four parts corresponds to the quadruple coefficient sequentially and its length agrees with degree of harmonic. However, for extracting the boundary coefficient to generate quadruple Fourier coefficient, each cuneiform image's symbol is subjected to extract its boundary by (canny operators) as closed contour to generate freeman chain code that leads to determine the length of chain (*T*)as seen in figure (7). The Elliptical Fourier algorithm works follows:

Algorithm 2: quadruple Fourier oefficient **Input**: cuneiform binary image symbol **Output:**. Fourier coefficient (*ai,bi.ci,di*)
Step1: read binary image symbol **I$_s$.**

step2: extract boundary pixels of cuneiform symbol by edge detection method   and save the results which represent a boundary image I$_B$.

Strp3: apply the thinning operation on boundary pixel .

Step4: compute edge encoding process to generate freeman direction code **U$_i$**. where

$U_{i=\{0,1,2,3,4,5,6,7\}}$.

Step5: compute the length of previous direction code dt$_i$

Where $dt_i = 1 + \left(\frac{\sqrt{2}-1}{2}\right) * (1-(-1)^{ui}))$.

Step6: compute   the   harmonic   coefficient (*ai,bi.ci,di*.)

(a)	(c)
(b)	(d)

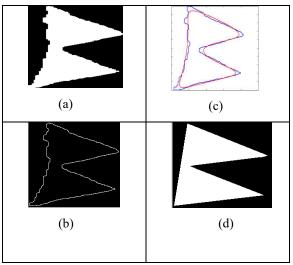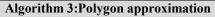Figure (7) Elliptical Fourier Descriptors ,a) binary cuneiform   symbols,   b)   boundary   image . c)approximated boundary figure ,d) matching class .

Ali .A/ Abdul Monem .S/ Abdul Mohssen .j

10.3.2 Feature extraction by polygon approximation with dominate point

The proposed algorithm to generate features vectors about each cuneiform symbols depends on combining the approximation principals (**Min-ϵ problems, Min-# problems**) in approximation approaches as illustrated in (7). That means briefly with first approximation principles (**Min-ϵ**) the number of segments determined previously is compatible with same number of cuneiform classis segments as seen in figure(11) and the tolerance value will be pre-defined where it decreases gradually in each iteration until the termination condition gets satisfied,[1]. The termination condition is satisfied if the number of segments equals the number of one the cuneiform patterns segments   (3,5,7,9). the proposed algorithm is as follows:-

---

Algorithm 3:Polygon approximation
  Input: binary cuneiform image.

  Output: approximated points

Step1: read binary image symbol I_s
Step2: apply edge detection method with suitable filter,
Step3: apply thinning technique ;
Step4: compute Freeman's chain code for boundary .
Step5: find break points DP_b.
Repeat
Step6: compute AVE for all DP
Step7: repeat

Step8: determine DP that minimum value DP_{min}
Step9: remove DP_{min} from dominat table
Step10: recalculate AVE for DP_{min}'s adjacent neighbor.
Step11: compute max $_{error}$
step12: until (max $_{error}$<th$_)$
step13.z= Remaining points about DP's is approximate polygon points
step14: deleted all DP'S which construct with its neighbors a straight angle
step15: th=th-eps; < eps epsilon value ex=0.009 >
step16   : until ((z==3)or(z==5)or(z==7)or(z==9))
step 17:   end   <where z= number of head >
step 18:return approximate points vector .

---

the recognition process depends on generating a **features vector**   for the training   set and testing symbol, as illustrated in figure (8), where each feature vector consists of Cartesian coordinate   (approximate points) for each class .
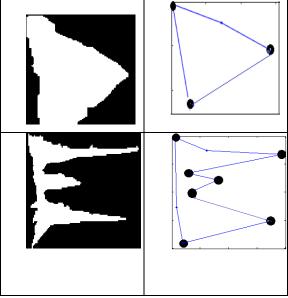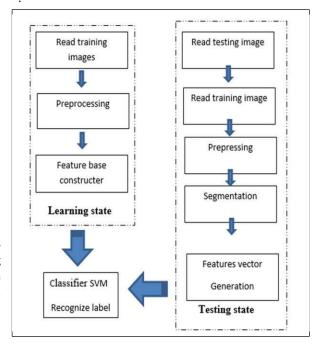

Figure (8): approximated points.

10.4 Classification

The classification task is applied with each cuneiform symbols by SVM classifier as a generated feature vector is adopted **for each features extraction method**,   to apply the comparison state. However, the classification stage concerns with two steps. The First one is a training state to generate features base with assigning a distinguished labile for each class (triangle class) .The other step applies the classification task with test feature vector. Below is the classification diagram figure (9) .
.

Figure(9): classification diagram steps.
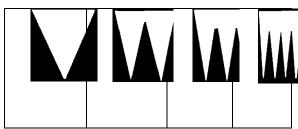
Ali .A/ Abdul Monem .S/ Abdul Mohssen .j

However, after applying the multiple classification task for each symbol the constructed recognition code consists of three parts , which represent number of cuneiform symbol distributed to numbers of   vertical, horizontal and diagonal. For example as seen in figure (10) below the cuneiform character, the recognition code (6#2#1#3) which means the total cuneiform symbols equals 6 , which are distributed to 2 vertical ,1 horizontal and 3 diagonal.

Figure(10):cuneiform character .

11.Proposed training dataset

The design of the dataset is one of the most important aspects that play an important role in the classification process. Therefore this proposed training dataset is a virtual dataset consists of 2D four triangular patterns with forms compatible with three-dimensional geometric shape form of the cuneiform symbols, figure (11)   . Covering all the possibilities and situations is taken by the cuneiform character's symbols with different directions (horizontal, vertical or diagonal), and compatible with patterns of cuneiform symbols to solve the problem of shadows. These patterns are distributed in 16 classes compatible with cuneiform directions.

Figure(11) :cuneiform patterns

11.Results and discussion:

This section presents the evaluation state according to features extraction metrics (EFD and polygon approximation) .The first testing state consists of 320 symbols related to probabilities of conform symbols to determine the recognition accuracy. Therefore, each binary cuneiform symbol is subjected to extra

boundary pixels and thinning process to generate features vectors according to each method of comparison state. Therefore, **the second experiment state** which deals with cuneiform character dataset consists of 240 of cuneiform image characters, after evaluating the comparison state. However after applying the first testing process the recognition accuracy with average **processing time** is illustrated in the following tables (table 2) according to each feature extraction method.

Table 2: Comparison of recognition results about each features extraction method with processing time.

Experiment/n	Classification approaches	Accuracy results with (EFD)	Processing time	Accuracy results with (Polygon approximation)	Processing time
1	SVM with RBF kernel	0.6093	0.332	0.940	0.452
2	SVM with **polyn-omial** kernel	0.6845	0.322	0.925	0.421
3	SVM with linear kernel	0.6594	0.320	0.613	0.441

As seen in previous table, the higher recognition result is achieved when polygon approximation method is adopted to construct features vector especially with RBF discriminate function (Experiment 1). Therefore, to evaluate the reliability about the proposed features extraction method the previous testing set is subjected to resizing process according to the different sizes. Bellow the accuracy results about each size is illustrated in table (3)

Table 3: Comparison of recognition results according to different image sizes, where the SVM classifier is adopted with RBF function.

Image size	Accuracy
256x256	0.940
128x128	0.865
64x64	0.853

Along with applying Gaussian filter on testing set many times according each standard deviation values **δ** as seen in figure (12) to evaluate the recognition reliability.

Ali .A/ Abdul Monem .S/ Abdul Mohssen .j

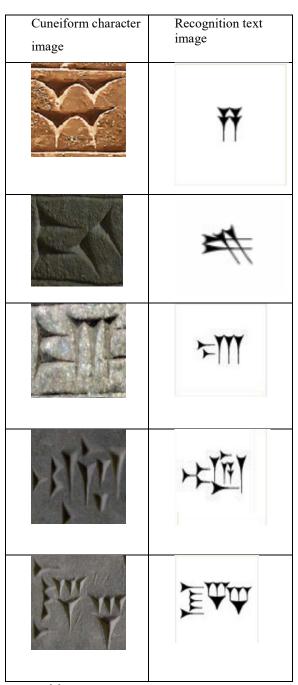| Binary cuneiform symbols	δ =1	δ =3	δ=7

Figure (12): sample of conform symbol deformation. applied Gaussian filter on each cuneiform symbols (in first dataset)   with different values of each standard deviation .

Therefore, the first dataset is subjected to deformation process by Gaussian filter several times   according to each experiment depending on different standard deviation value, the accuracy results are illustrated in table (4) .

Table4 : Comparison of recognition accuracy results according to   different standard deviation values **δ.**

standard deviation value	Accuracy
1	0.860
3	0.867
4	0.870
7	0.860

However,   now   after   adopting   the   polygon approximation with SVM classifier, according to its discriminant function (RBF)   the recognition result about the **second dataset** is satisfied against each cuneiform character image which can be seen in the following figure(13),**where the recognition state for each character results from a cumulative**

Cuneiform character image	Recognition text image

recognition process.

Figure(13): cuneiform character image sample recognized by proposed system.

When adopting spatial low pass filter domain for enhancement process (according to the second dataset) the output of accuracy character recognition results are illustrated in following table (Table 4):-

Table 4:Comparison of results of recognition accuracy ratio after applying different LPF with different sizes

Ali .A/ Abdul Monem .S/ Abdul Mohssen .j

Filter size	3X3	5X5	7X7
Medina	0.496	0.512	0.496
Gaussian	0.536	0.440	0.512
Average	0.472	0.544	0.504

Table 5: Comparison of recognition accuracy results

According to cut of frequency values

Experiment/no	Cutoff frequency values	accuracy
1	0.2	0.51
2	0.4	0.61
3	0.6	0.94
4	0.8	0.62
5	0.9	0.67

As seen in the previous table (5), the higher accuracy character recognition is archived in experiment (3) ,where the value of cut off frequency equals (0.4).

Now the evaluated recognition state is applied among the discriminant functions about SVM with (**second dataset**), where the value of cutoff frequency equals (0.4) .bellow are accuracy character recognition values and **processing time** in table (6).

Experiment/no	Discriminant function	Accuracy	Average Processing Time(s)
1	RBF	0.92	4.028
2	liner	0.77	4.136
3	polynomial	0.75	4.418

As it can be seen in the previous table, the highest accuracy character recognition with low processing time when adopting the RBF is compared with other discriminant functions LPF..

Conclusion

This research presents a comparison state between two feature extracting methods ,the elliptic Fourier descriptor (EFD) and polygon approximation methods, where the support vector machine is adopted as a classifier model with its discriminant kernel functions .Consequently, the evaluated recognition results indicate that the polygon approximation by dominate points is more accurate to be adopted in classification model as the achieved accuracy result is 94% with RBF kernel function compared with (EFD) method . To achieve more reliable decision about the polygon approximation method, the testing set is subjected to deformation state by Gaussian filter with different values of stander divisions, accordingly, the accuracy result is maintained to high quality .The testing state is applied by two testing data set. The frequency domain with low pass enhancement filter is more accurate than the spatial domain according to the accuracy of the achieved results that were compared between them..

REFERENCES

[1] Rahma A," Recognize Assyrian cuneiform characters by virtual dataset " 6th International Conference on Information and Communication Technology and Accessibility (ICTA) Muscat, Oman, 2017.

[2] FahimehMostofi , Adnan Khashman . "Intelligent Recognition of Ancient Persian Cuneiform Characters" . Intelligent Systems Research Center, 1Department of Computer Engineering. August 2014.

[3] Naktal M., "Cuneiform Symbols Recognition Based on K-Means and Neural Network " ,Fifth Scientific   Conference   Information Technology,Raf. J. of Comp. & Math's. , Vol. 10, No. 1, 2013.

[4]AL-Ani s.," Image   enhancement and recognition of cuneiform writing   ",instate of higher   students   for   computer   and information .,2001.

[5] Al-Ani H., "Data Extraction of Cuneiform Tablets Digital Images ",Al Rasheed College University of Technology.,September 2006.

[6]Anil K.,Robert P., and Mao J.," Statistical Pattern Recognition: A Review"   , IEEE Transactions ON Pattern Analysis and Machine intelligence , VOL. 22,   NO. 1, JANUARY 2000

Ali .A/ Abdul Monem .S/ Abdul Mohssen .j

[7]Zaitoun N.,Aqel M,"Survey on Image Segmentation Techniques",International Conference on Communication, Management and Information Technology (ICCMIT) ,Procedia Computer Science 65 (2015) 797 – 806), 2015.

[8]Rafael C. Gonzalez , Richard E. Woods "Digital Image Processing", Pearson Education International , (2002).

[9] Sonka M ., Hlavac ., V and Boyle R., "Imaga processing , Analysis and Machine Vision ", Thomson ,Third edition , 2008.

[10] Ballaru B. ,Reas P and Tegolo D.,"Elliptical Fourier Descriptors for   shape   retrieval in biological images ", Conf. on Electronics, Control & Signal , 2002.

[11]Masood A .,"Dominant point detection by reverse polygonization of digital curves ",Image and Vision Computing 26 702–715, 2008.

[12] Abbasi H.,Olyaee M., and Ghafari H.,"Rectifying Reverse Polygonization of Digital Curves for Dominant Point Detection" IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 3, No 2, May 2013.

[13] Grigore O et al. "On the Implementation of Polygonal Approximation Algorithms". Department of Information and Computing Sciences, Utrecht University. technical report UUCS-2003-005, 2003..

[14] Kolesnikov A., andFränti P. "Polygonal Approximation of Closed Contours " .,13th Scandinavian Conference, SCIA Halmstad, Sweden, June 29 – July 2, 2003.

[15] Sammut C., and Webb G., "Encyclopedia of Machine Learning", US: Springer ,2010

[16]N. Cristianiniand J. Shawe-Taylor. " A n Introduction to Support Vector Ma chines. Cambridge University Press, 2000.

تميز الحروف المسمارية باستخدام مصنف الدعم الأتجاهي

علي عادل سعيد   عبد المنعم صالح رحمه   عبد المحسن جابر عبد الحسين

قسم علوم الحاسوب ، الجامعة التكنولوجية ، بغداد العراق

المستخلص :

يهدف هذا البحث الى اجراء مقارنه بين طريقتي بناء متجه الصفات بين طريقة تقارب المنحنى وطريقة الواصف الأهليجي الدوري مع بيان اي الطريقتين يمكن اعتمادها في انشاء متجه الصفات ،وذلك اعنمادا على النتائج المتحققه المتمثله بعاملي الوقت والدقة. اظهرت النتائج المتحققه الى تفوق الطريقه الأولى في بناء متجه الصفات على الطريق الثانية . تم ومن خلال البحث اجراء الأختبارات المطلوبه على مجموعتي اختبار ،الأولى مكونه من ٣٢٠ ثنائية اللون لأختبار اي الطريقتين افضل في بناء متجه الصفات . اما مجموعة البيانات الأختبارية الثانيه فهي مكونه ٢٤٠ صوره للحروف المسماريه الأشوريه يتم اختبارها بعد استكمال المرحله الأولى .تم اعتماد المصنف الدعم الأتجاهي بدواله المميزه الخطيه ،متعدده الحدود.وداله توزيع كاوس. حيث اثبتت النتائج ا لأختباره على افضليه اعتماد طريقة تقارب المنحنى في بناء متجه الصفات حينما تكون الداله المميزه كاوس معتمده.