International Journal of Computer Applications (0975 – 8887)
Volume 44– No.6, April 2012

# A New Decision Tree for Recognition of Persian Handwritten Characters

Mohammad Rajabi
Department of Computer Engineering, University of Isfahan, Iran

NaserNematbakhsh
Department of Computer Engineering, University of Isfahan, Iran

S.AmirhassanMonadjemi
Department of Computer Engineering, University of Isfahan, Iran

## ABSTRACT
In this paper a binary decision tree, based on Neural Networks, Support Vector Machine and K-Nearest Neighbor is employed and presented for recognition of Persian handwritten isolated digits and characters. In the proposed method, a part of the training data is divided into two clustersusing a clustering algorithm, and this process continues until each subtree reaches clusters with optimum clustering, where the tree leaves are the final obtained clusters. According to the clustering results, classifiers such as ANN and SVM can perform correctly, therefore the decision tree can be built. A part of the test data is selected as validation data and in each node of the tree, a classifier with the highest recognition accuracy on validation data is selected. Recognition accuracy at 8, 20, and 33 clusters have been evaluated and compared with other existing methods. Recognition accuracy of 98.72% and 97.3% on IFHCDB database is obtained respectivelywhen 8-class and 20-class problems is assumed. Again 98.9% accuracy on HODA database is achieved.

## General Terms
Persian handwritten recognition, Feature extraction, Supervised learning, Classification,Binary decision tree, Unsupervised learning.

## Keywords
Support Vector Machine, K nearest neighbor, Decision Tree, Self organized map, Neural Networks, Multi class classification.

## 1. INTRODUCTION
Handwritten Character recognition is one of the still open issues in the pattern recognition and it has diverse applications such as reading the checks, car plate and handwritten postal codes recognition. In recent years the use of SVM and ANN has had good results in recognizing of Persian handwritten digits and characters so that their recognition rate have been reached over 97% and 95%, respectively[1-6].

Ebrahimpur et al. [1]have used Loci Featuresinthefeature extraction phase and Neural Network and combination of experts in the classification phase, They achieved 97.52% accuracy in recognizing Farsi handwritten digits. Abdleazeem and Sherif[7]have used gradientfeaturesandclassifiersofNeural Networks with two hidden layersandSVM in recognition of Arabic handwritten digits. The accuracy of their system was 99.2% and 99.48% using neural networks and SVM. Ziaratban et al. [8] have divided 32 Persian characters into 8 clusters; theycombinedstructuralandstatisticalfeaturesinthefeature extraction phase and achieved 93.15%accuracy using neuralnetworks. In [9] a system is presented for recognizing handwritten courtesy amounts in Persian bank checks by Sadri et al. they have used different algorithms for extracting features such as Outer profile, Chain codes and Zoning, and several algorithms in recognition phase such as Neural Networks and SVM. They used two hidden layers and 74 neurons for each layer in neural networks and RBF kernel for SVM. The best accuracy of their system was 96.5% by neural networks.

Inthispaperwehave useddifferent methods forextracting features,in the recognition phase we used SVM, KNN and neural networks, for SVM we used different methodsto createamulti-class recognitionsystem such as OVA ،OVO andSVM-BDT. In clustering phase and creation of binary decision tree, we used self-organized map. Indifferentstudiesbecauseofsimilaritiesbetween thePersianletters, same characters were tried to be placed into a single cluster. The number of clustersusually is selected 8 or 20. Table1 shows printed and handwritten Persian characters in 32 classes.

The organization of the rest of the paper is as follows: Section 2is devotedto preprocessingandimprovingimages, differentmethods offeature extractionis introducedinSection 3, inSection 4 we introduce different methodsof classification wehave usedin thispaper. We explained details of proposedmethod in section 5, described experimental results and comparisons in section 6, and finally mentioned the conclusions and our future works in Section 7.

## 2. PREPROCESSING AND IMPROVING IMAGES
Atthissection,wetry toremoveimages' noise using,image processingalgorithms and normalize them. Database images are grayscale format, in the first step of pre-processing phase we convert grayscale images into binary images using Otsu algorithm [10].The resultingbinaryimagescontainssomenoiseanddiscontinuity, by thenoiseremovalalgorithms [11]andthemorphological algorithms [12], we fixthem. In the last step, the images are normalized by the normalization method described in [11] and the images' size is converted into 45*45 pixels. Figure 1showsthepre-processing steps inthe paper.

**Table 1. Persian Handwritten and Printed Characters**

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Printed Characters | آ ا | ب | پ | ت | ث | ج | چ | ح | خ | د | ذ | ر | ز | ژ | س | ش |
| Handwritten Characters | | | | | | | | | | | | | | | | |
| Printed Characters | ص | ض | ط | ظ | ع | غ | ف | ق | ک | گ | ل | م | ن | و | ه | ی |
| Handwritten Characters | | | | | | | | | | | | | | | | |

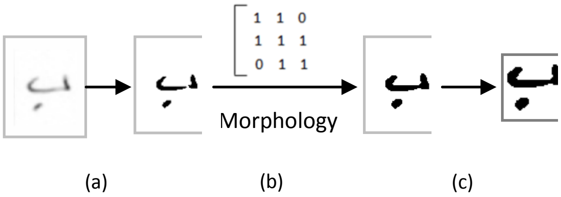


**Figure1. (a) An example of Binarization using Otsu method, (b) Structuring element for morphological operation to filling of discontinued images, (c) Normalization of characters**

# 3. FEATURE EXTRACTION

We used different algorithms for extracting features; in this step some of themost importantFarsicharacterfeature extraction techniquesarebrieflydescribedasfollows: Zoning[13],Outer profiles [14]and Crossing counts[15] .

## 3.1 Zoning

Inthismethod,imagesare divided into3*3zones, the ratio of black pixels number to white pixels numberof each region is calculated, andafeature vectorwith225elements is created.Figure 2 showsan example ofthismethod.

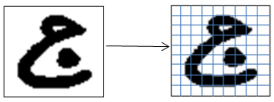

**Figure2. An example of extracting zoning**

## 3.2 Outer Profiles

Each sideofImages(top, bottom, left and right) hasaparticularview, and featurescanbe extractedfrom each ofthem. For example in extractingfeaturesfromtopview, we startfrom thefirst leftpixelandgodowntogetthe firstblackpixel, then we store the obtained row number as the first feature andalsodothis for theothercolumns, 45featuresareextracted fromeachviewandfinallythe featurevectorwith 180elementsmakes up.Figure 3showsan example ofthismethod.

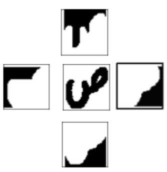

**Figure3. An example of extracting Outer profile features**

## 3.3 Crossing Count

In this method theimageis scannedline by lineandnumber ofchangescolorfromblackto white pixelor vice versa is counted, the number of extracted features in the first phase will be 45. Thisworkis alsodonefor thecolumnsand eventuallyafeaturevectorwith 90elementswill be formed.An exampleofthismethodis showninFigure 4.

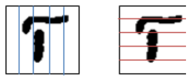

**Figure4. An example of extracting Crossing count**

# 4. CASSIFICATION AND CLUSTRING

The most importantpartof apatternrecognitionsystemislearningand classification. Inthissection, webriefly describe classification methodswehaveusedinthe paper.

## 4.1 K-Nearest Neighbor

Thisclassifier has nolearningprocessand isable to solvemulti-class classification problems. Each Sample of training set has a label that defines its class, whenasamplecomesfrom thetestset for classification; the distancebetween test sample and all training samples is measured. Thisdistanceis obtainedthroughmethodssuch as Euclidean,Hamming, Correlation, and etc.

Let $x_i$be an input sample with $k$ features $(x_{i1},x_{i2},\ldots,x_{ik})$ ,$n$ be the total number of input samples ($i$=1,2,…,$n$) and $k$ the total number of features ($j$=1,2,…,$k$) . Eq. (1) shows the Euclidean distance between sample $\mathbf{x}i$ and $\mathbf{x}l$ ($l$=1,2,…,$n$) :

$$d(x_i,x_l) = \sqrt{(x_{i1}-x_{l1})^2 + (x_{i2}-x_{l2})^2 + \ldots + (x_{ik}-x_{lk})^2} \quad (1)$$

Then k samples (neighbors) that have minimumdistance from the test sample are selected. Using majority voting among K neighbors, class label for the test sample is found. For example with 1-nearest neighbor rule, if $\omega$ be the true class of a training sample, the predicted class of test sample $x$ is set equal to the true class $\omega$ of its nearest neighbor, where $m_i$ is a nearest neighbor to $x$ if the distance $d(m_i,\mathbf{x})=\min_j\{d(m_j,\mathbf{x})\}$.

## 4.2 Support Vector Machine

Support Vector Machine [16] isaset of supervised learningmethods thatare usedforclassification andregression. A data in SVM can be seenasavectorinap-dimensional(or a list ofp numbers). SVM goal is to maximizemargin betweentwoclasses. So it selects a hyperplane that has maximum distance from the nearest data on both sides of this separator.Ifthere isthehyperplane, it is knownasthe maximummarginhyperplane. The decision function to separate the data is determined with a subset of training examples that are called support vectors (closest trainingdatato thehyperplaneseparating). Indeed, theoptimalhyperplaneinSVM is a separatorbetweenthesupportvectors. Iftrainingdataare notseparatedlinearly we mustbuildahyperplane to minimize probability offalseseparation.Theconcept illustrated inFigure 5.

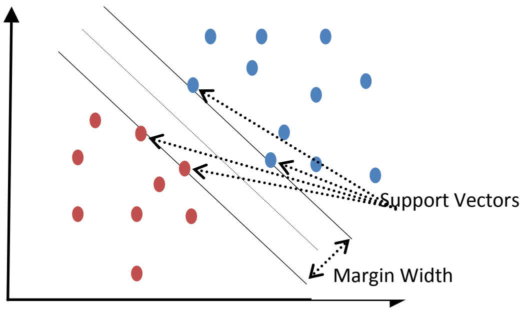


**Figure 5. Linear separating hyperplanes for the separable case**

SVM isinherently abinaryclassifier anditcannotbe used directlyinmulti-class problems. In the following subsections, we review thevarietyofmethodsto createa multi-class classifierbySVM.

Thenumber ofclassesforallmethodsisdenoted byN. During thestudy it wasprovedthatSVM withpolynomial kernel hasbestrecognitionaccuracycompared to the otherkernelssuch asRBFandMLP.

### *4.2.1 OVO (one versus one)*

Amethodforconstructingmulti-class Support Vector Machine is OVO [17]. Inthismethod,each classwith*N*-1otherclassesis taught. Thus $\frac{N(N-1)}{2}$ SVM is required for training. AsshowninFigure 6, for n=4, 6 SVM is required for training. Inputsampleis given $\frac{N(N-1)}{2}$ SVM forrecognition. Disadvantage ofthismethodislow speedinrecognitionphase becauseinthismethod the number ofbinary SVMs istoo much**.**

### *4.2.1 OVA (One versus All)*

There is a binary SVM per-class in OVA training, So that one class has labeled with 1 and *N-1* class have labeled with -1.Figure 7 shows image of training using OVA for a 6-class problem. The number of SVM in training phase will be the same as number of classes. Input sample is given to all SVMs in the test phase. Each SVM may have negative or positive answer. Input image class is selected by SVM that has the most confidence. Disadvantage of this method is complex training for a very large training set, because each class must be taught against all data from all other classes. In the following subsections, we describe methods are better than OVA in terms of speed and accuracy**.**

### *4.2.2 DAGSVM (Directed Acyclic Graph)*

In this method, training phase is the same of OVO. In the test phase, input sample is sent to the first SVM, between class c1 and c2 if c1 is chosen then candidate c2 will be completely removed and *N*-1othercandidates will beremained. Recognitionrateof thismethodis same asOVO butin thetestphase, requirednumber ofSVMis *N*-1anditsrecognitionspeedishigher than OVO.

### *4.2.3 SVM-BDT (Support Vector Machines utilizing Binary Decision Tree)*

In thepresentedmethod describedin[18],datasetare divided intotwo groupsbyaclusteringalgorithm. Then eachsub-groupisdividedinto twoothergroups andthisprocesscontinuesuntildata related to each class is placed into a group.Figure 8 showsthisclusteringfor a 7-class problem.

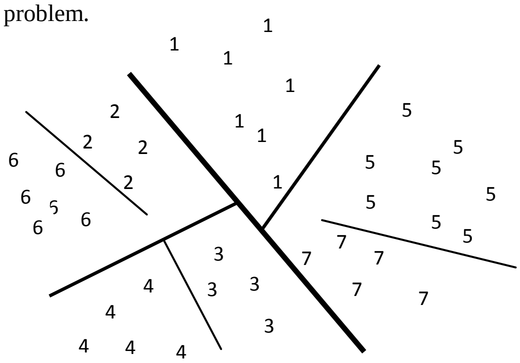

**Figure8.Clustering of the 7-classes problem.**

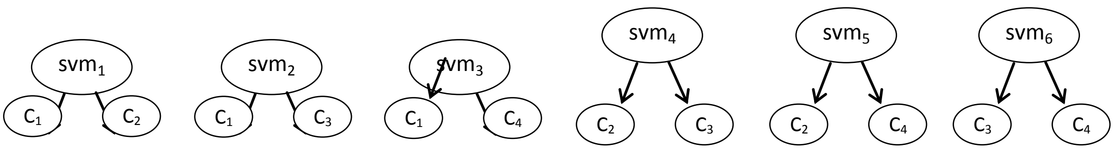


**Figure6. OVO decomposes a k-class classification problem into $\frac{N(N-1)}{2}$ binary classification problems.**

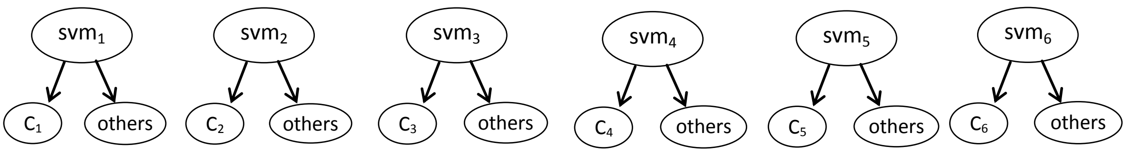


**Figure7.OVA decomposes a k-class classification problem into k binary classification problems. Each binary classifier distinguishes instances of one class form all other remaining classes.**

According totheFigure 8, classes1, 5and 7 have been placedin thefirst groupandclasses2, 3, 4 and 6 have been placedinthe second group.Now the decision tree can be constructed using SVM. Figure 9 showshow to buildthistreeAccording toFigure 8. InFigure9internalnodes with circleandleaveswith trianglesare shown. Number of internal nodes shows the number of needed SVMs for train. There are n-1 internal nodes in a problem with *N*class. More speed and accuracy is the advantage of this method. According to thisalgorithminthetraining phase we need n-1 SVM for training and in the test phase, at most $\lceil \log_2 N \rceil$SVM are required. Lower number of SVMs in test phase,made ittoo fast.

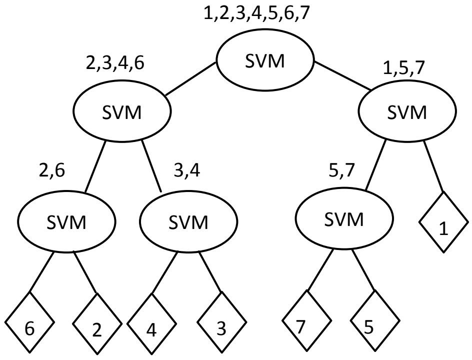


**Figure9.Illustration of SVM-BDT.**

## 4.3 Neural Networks

Artificial neural network (ANN)isacollectionofneuronsthatapproximatethe brain functions. Artificial neural network can be considered a parallel computing system that consists of many computational small units which are connected together. ANN inmachine learningand classification problems is used as a classifier. In this paper, we use a Multi-Layer Perceptron (MLP) network [19] and Scaled conjugate gradient backpropagationalgorithm [20]to learnandmodify thenetworkweights.

## 4.4 SOM (Self Organizes Map)

SOM [21]isanunsupervisedlearning algorithm and isused forclusteringinpattern recognitionandmachine learning problems. It consists of twolayers, Inputlayer and competition layer, thesize ofthe input layer is equalto thefeaturevectorofeach sample. Thenumber ofcompetitiveneurons usuallyis equal tothe number ofclusterswe want tomake. Figure 10 shows architecture SOMwithtwooutputneurons is used in the paper.

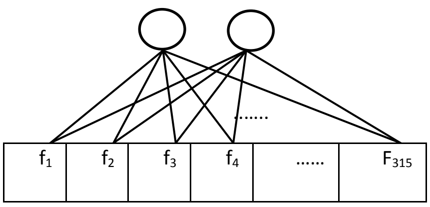


**Figure10. SOM with two neurons in competition layer**

## 5. DECISION TREE BASED ON ANN, SVM AND KNN

In this method we have used SVM-BDT idea. Featuresextractedforeachimageisa combination ofZoning and Crossing counts because they had better performance. The total number of combination features will be 315. Ourresearchissue has 33classes. Initially,33classes are divided intotwoclusters by SOM clustering,and then they are trained by SVM and ANN.

During thestudyit was foundexperimentallythat SVMwithpolynomial kernelhasbetter performancethanother kernels. The number ofhiddenlayersandepochsin eachnode for neural networkis selected separately. Clusteringprocedureis doneforthe nextclusters, so that most of thedataineach clusterbelong to aclass. Figure 11showsdecision treeforourproblem. In this method *N*-1 classifier is required. As describedin section4, decisiontree is a solution for support vector machine to create multi-class classifier,but neural networkcansolve non-linear problemswithlargeclasses. Wehaveofferedto use neural networks ineachnodeofthebinarydecision tree. We have usedoptimized structureforeachnode(such as thenumber ofiterationsandhidden layers) of the networks. Table 2 shows recognition of classifiers for each node.

**Table 2.Resultof ANN, SVM, and KNN classifiersoneachnodeofthe tree (%)**

| | Classifiers | | |
|---|---|---|---|
| **Node** | ANN | KNN | SVM |
| **1** | 99.1 | 99.12 | **99.3** |
| **2** | **99.8** | 99.56 | 99.5 |
| **3** | 99.47 | **99.76** | 99.35 |
| **4** | **99.1** | 98.7 | 98.2 |
| **5** | 99.83 | **100** | 99.83 |
| **6** | **98** | 98 | 98 |
| **7** | **99** | 98.6 | 98.8 |
| **8** | **99.8** | 99.8 | 99.8 |
| **9** | **99.8** | 99.8 | 99.6 |
| **10** | **97** | 95 | 94.5 |
| **11** | **99.25** | 98.5 | 99.25 |
| **12** | **97** | 89.27 | 96 |
| **13** | **94** | 82.5 | 90 |
| **14** | **99.8** | 98.8 | 99.6 |
| **15** | **98.3** | 97 | 97.17 |
| **16** | **99** | 98.89 | 99 |
| **17** | **99** | 95 | 97.16 |
| **18** | **98** | 97.35 | 98 |
| **19** | **99.5** | 98.5 | 98.5 |
| **20** | **98** | 95.45 | 98 |
| **21** | **96.5** | 95 | 96 |
| **22** | **89.5** | 89 | 89.5 |
| **23** | **96** | 95.3 | 95.5 |
| **24** | **98** | 95.65 | 96.5 |
| **25** | **99.3** | 96.6 | 97.67 |
| **26** | **98.75** | 97.25 | 97.28 |
| **27** | 97.5 | 96 | **98.5** |
| **28** | **98.5** | 96 | 98 |
| **29** | **99** | 98.8 | 99 |
| **30** | **98.5** | 95 | 95 |
| **31** | **97.6** | 90 | 92.6 |
| **32** | 96.5 | 94 | **97** |

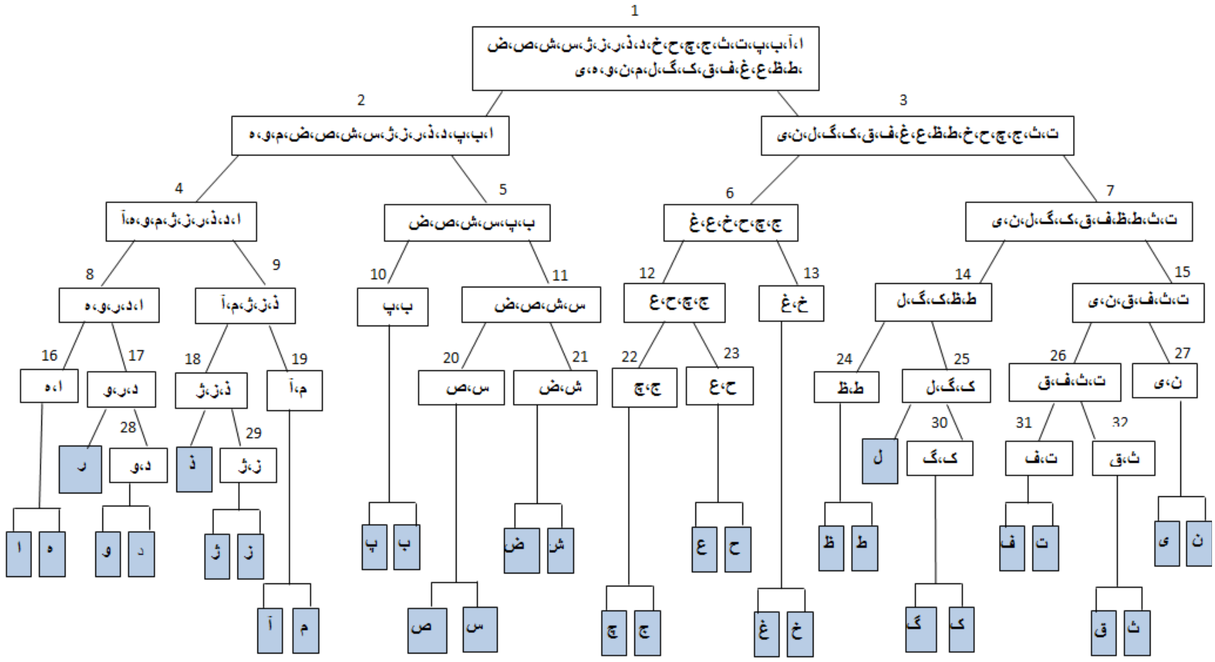

**Figure11.Decision treecreated by SOM**

Aftertrainingclassifiersforeachnode, we select validation data from test data and testrecognitionat eachnodebyKNN, SVM and ANN, and then we choose the classifier which has the best recognition accuracy as the main classifier of the node.According toTable 2inmost of thenodes, neural networks have better recognitionthanKNNandSVM. SVM hasmore accuratelyatnodes1,27and 32than theothertwoclassifiers. KNNalso hasmore accuratelyonly at nodes3 and 5than theothertwoclassifiers. Thus for each node, classifier that has more accuracy is selected.

## 6. EXPERIMENTAL RESULTS

We used two databases to evaluate our method:IFH-CDB [22] and HODA [23] described below, each data set is divided into training, validation, and test set.

**The IFH-CDB database,** thisdatabaseincludes 52320 Persian isolated characters thatresearcherstesttheir methods on it. Images inthe database have beenextracted from inputtestregistrationoftraining centers. The size ofthe images inthis collectionis90*77pixelsin 300 dpi grayscale format. We have used 32400 samples for training, 16620 samples for test and 3300 samples for validation set.

**The HODA database,** this database was introduced by Khosravi and Kabir. They collected handwritten Persian digits using universities’ entrance exam forms. The database consist of 80,000 samples, according [24] we consider it an easy database, so we selected 9000 harder sample which were classified into more than three classes by k-nearest neighbor method with k=6.

Some Persian characters such as'س', 'ش', 'ص', 'ض' or 'ح', 'خ', 'ج', 'چ' are very similar and onlydots,distinguishthemfrom each other. Due tothis,researchersin theirpreviousstudies havetried toputthesecharactersintoacluster and have reducedthenumber ofclasses. Figure 12and 13 showsgroupingproblem into 8and 20 classes, respectively.

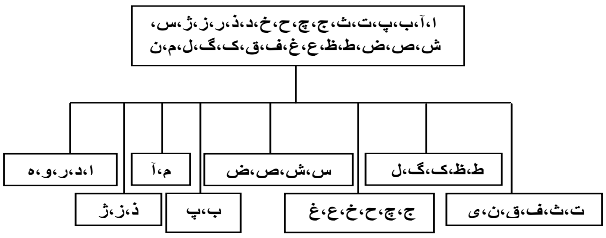

**Figure12. Similar classes of Persian characters grouped into 8 classes**

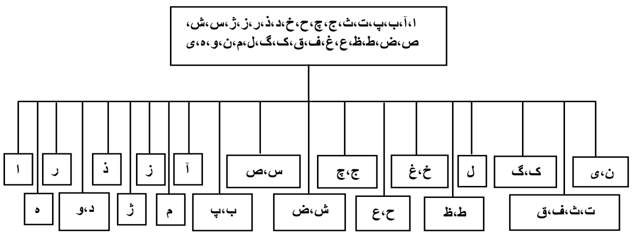

**Figure13. Similar classes of Persian characters grouped into 20 classes**

In the following, recognitionaccuracy oftheproposed method is compared withthe recognition accuracy of other methods and previous researches for8,20 and,33 classes. The number of test data was 16620 and 3300 samples isused for validation. Table 3showsthe accuracyof recognitionwithdifferentmethods.

**Table 3. Different result obtained with different classifiers (%)**

| NO. of Clusters | KNN | | SVM | | | ANN | Our System |
|---|---|---|---|---|---|---|---|
| | 1-NN | 3-NN | OVA | OVO | BDT | | |
| 8 | 97.9 | 97.84 | 96.63 | 97.98 | 97.97 | 98.13 | **98.72** |
| 20 | 96.09 | 96.28 | 93.8 | 96.28 | 96.19 | 96.27 | **97.3** |
| 33 | 91.67 | 92.11 | 88.72 | 92.11 | 93.07 | 93.37 | **94.82** |

As shown in Table 3, recognition accuracy of the proposed method is more than of other methods such as neural networks and support vector machine in all clustering modes. Table 4 shows confusion matrix when the number of classes are considered 8. Also we tested our method on HODA dataset, table5 show confusion matrix for handwritten Persian digits in the dataset.

Table 6 and table 7 compared accuracy of the proposed method with other methods were used in recent researches on IFHCDB and HODA dataset, respectively.

**Table 4. Confusion matrix for the 8-class problem of our method on IFH-CDB dataset (%)**

| Input character | Recognized as | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 |
| C1 | 98.78 | 0 | 0.22 | 0 | 0.11 | 0.33 | 0.11 | 0.44 |
| C2 | 1.1 | 97.96 | 0.37 | 0 | 0.19 | 0 | 0.19 | 0.19 |
| C3 | 0.56 | 0.28 | 99.17 | 0 | 0 | 0 | 0 | 0 |
| C4 | 0 | 0 | 0 | 99.44 | 0 | 0 | 0 | 0.56 |
| C5 | 0 | 0 | 0.28 | 0.14 | 98.19 | 0.14 | 0.69 | 0.56 |
| C6 | 0.09 | 0 | 0 | 0 | 0 | 99.91 | 0 | 0 |
| C7 | 0.22 | 0.22 | 0.1 | 0 | 0.22 | 0 | 98.56 | 0.67 |
| C8 | 0.19 | 0.19 | 0 | 0.09 | 0 | 0.19 | 1.57 | 97.78 |

**Table 5. Confusion matrix of the proposed method on HODA dataset (%)**

| Input Digit | Recognized as | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 0 | 98.8 | 0.53 | 0 | 0 | 0 | 0.67 | 0 | 0 | 0 | 0 |
| 1 | 0.29 | 98.76 | 0.4 | 0 | 0 | 0 | 0.13 | 0 | 0 | 0.42 |
| 2 | 0 | 0.4 | 98.7 | 0.6 | 0.3 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0.22 | 98.68 | 0.1 | 0 | 0.09 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0.28 | 99.31 | 0 | 0.2 | 0 | 0 | 0.31 |
| 5 | 0.58 | 0 | 0 | 0 | 0 | 99.19 | 0 | 0 | 0.23 | 0 |
| 6 | 0 | 0 | 0.19 | 0 | 0.48 | 0 | 99.1 | 0 | 0 | 0.23 |
| 7 | 0.22 | 0 | 0 | 0 | 0 | 0.23 | 0 | 99.55 | 0 | 0 |
| 8 | 0.36 | 0 | 0 | 0 | 0 | 0.85 | 0 | 0 | 98.79 | 0 |
| 9 | 0 | 0.67 | 0.33 | 0 | 0.36 | 0 | 0.39 | 0 | 0 | 98.25 |

**Table 6. Results of different algorithms on IFHCDB dataset**

| Algorithm | Train Size | Test Size | NO. of Clusters | Accuracy (%) |
|---|---|---|---|---|
| Alaei et al. [2] | 36682 | 15338 | 8 | 98.1 |
| Alaei et al. [2] | 36682 | 15338 | 32 | **96.68** |
| Ziaratban et al. [8] | 11471 | 7647 | 8 | 93.15 |
| Mozaffari et al. [25] | 3200 | 2880 | 8 | 87.26 |
| Dehghani et al. [26] | - | - | 8 | 71.82 |
| Mowlai et al. [27] | 3200 | 2880 | 8 | 32.75 |
| Dehghan et al. [28] | 1600 | 1600 | 20 | 96.92 |
| Shanbezade et al. [29] | 1800 | 1200 | 32 | 87 |
| **Proposed method** | 36000 | 13320 | 8 | **98.72** |
| **Proposed method** | 36000 | 13320 | 20 | **97.3** |
| **Proposed method** | 36000 | 13320 | 33 | 94.82 |

**Table 7.Result ofdifferent algorithms on HODA dataset**

| Algorithm | Train size | Test size | Accuracy (%) |
|---|---|---|---|
| Ebrahimpour et al. [1] | 60,000 | 20,000 | 97.52 |
| Ebrahimpour et al. [24] | 6000 | 2000 | 95.3 |
| Javidi and sharifizadeh [30] | 6000 | 2000 | 98.16 |
| Javidi et al. [31] | 6000 | 2000 | 97.73 |
| Moradi et al. [32] | 18000 | 2000 | 96 |
| **Proposed method** | 6000 | 2000 | **98.9** |

## 7. CONCLUSION

In this paper we introduced a binary decision tree for Persian handwritten isolated characters recognition; at each node of the tree, a binary classifier was used. We combined zoning and crossing count methods in feature extraction phase, and created a feature vector with 315 elements. We also employed SOM to create binary decision tree. In the training phase, we trained SVM and ANN for each node of the decision tree. A part of the test data called validation is used for selecting a classifier for each node. The classifier, whose recognition accuracy was the most at each node of the tree, was selected. Recognition accuracy which was obtained for 8 and 20 clusters (98.72% and 97.3%)were higher than those of the previous methods. Most of the misclassified samples were related to the clusters that were very similar to each other. This caused low accuracy when the number of clusters equaledthe number of classes. In future, we plan to use more efficient methods in feature extraction such as connected component to extract main body of the image and its dots information to remove some of the confusions amongst similar classes.

# 8. REFERENCES


[1] R. Ebrahimpur, M. R. Moradian, A. Esmkhani *et al.*, “Recognition of Persian handwritten digits using Characterization Loci and Mixture of Experts,” *International Journal of Digital Content Technology and its Applications,* vol. 3, 2009.

[2] A. Alaei, P. Nagabhushan, and U. Pal, "A New Two-Stage Scheme for the Recognition of Persian Handwritten Characters." pp. 130-135.

[3] D. Ghosh, T. Dube, and A. P. Shivaprasad, “Script Recognition-A Review,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 32, no. 12, pp. 2142-2161, 2010.

[4] S. S. Ahranjany, F. Razzazi, and M. H. Ghassemian, "A very high accuracy handwritten character recognition system for Farsi/Arabic digits using Convolutional Neural Networks." pp. 1585-1592.

[5] C. L. Liu, and C. Y. Suen, “A new benchmark on the recognition of handwritten Bangla and Farsi numeral characters,” *Pattern Recognition,* vol. 42, no. 12, pp. 3287-3295, 2009.

[6] A. Borji, M. Hamidi, and F. Mahmoudi, “Robust Handwritten Character Recognition with Features Inspired by Visual Ventral Stream,” *Neural Processing Letters,* vol. 28, no. 2, pp. 97-111, 2008.

[7] S. Abdleazeem, and E. El-Sherif, “Arabic handwritten digit recognition,” *International Journal of Document Analysis and Recognition (IJDAR),* vol. 11, no. 3, pp. 127-141, 2008.

[8] M. ziaratban, K. Faez, and F. Allahveiradi, "Novel Statistical Description for the Structure of Isolated Farsi/Arabic Handwritten Characters." pp. 332-337.

[9] J. Sadri, Y. Akbari, M. J. Jalili *et al.*, "A New System for Recognition of Handwritten Persian Bank Checks." pp. 925-930.

[10] N. Otsu, “A Threshold Selection Method from Gray-Level Histograms,” *Systems, Man and Cybernetics, IEEE Transactions on,* vol. 9, no. 1, pp. 62-66, 1979.

[11] M. Cheriet, N. Kharma, C. L. Liu *et al.*, *Character Recognition Systems a Guide for Students and Practitioners*: John Wiley & Sons Inc, 2007.

[12] R. C.Gonzalez, R. E.Woods, and S. L.Eddins, *Digital Image Processing Using MATLAB*: Gatesmark Publishing, 2009.

[13] S. V. Rajashekararadhya, and P. V. Ranjan, "Zone based Feature Extraction Algorithm for Handwritten Numeral Recognition of Kannada Script." pp. 525-528.

[14] J. Sadri, C. Y. Suen, and T. D. Bui, “Application of Support Vector Machines for Recognition of Handwritten Arabic/Persian Digits,” in 2nd Conference on Machine Vision and Image Processing & Applications, Iran, 2003, pp. 300-307.

[15] H. Soltanzadeh, and M. Rahmati, “Recognition of Persian handwritten digits using image profiles of multiple orientations,” *Pattern Recognition Letters,* vol. 25, no. 14, pp. 1569-1576, 2004.

[16] V. Vapnik, *Statistical Learning Theory*: Wiley,New York, 1998.

[17] X. Peng, and A. K. Chan, "Support vector machines for multi-class signal classification with unbalanced samples." pp. 1116-1119 vol.2.

[18] G. Madzarov, D. Gjorgjevikj, and I. Ghorbev, "A Multi-Class SVM Classifier Utilizing Binary Decision Tree." pp. 77-81.

[19] P. Y. Simard, D. Steinkraus, and J. C. Platt, "Best practices for convolutional neural networks applied to visual document analysis." pp. 958-963.

[20] M. Martin Fodslette, “A scaled conjugate gradient algorithm for fast supervised learning,” *Neural Networks,* vol. 6, no. 4, pp. 525-533, 1993.

[21] T. Kohonen, “The self-organizing map,” *Proceedings of the IEEE,* vol. 78, no. 9, pp. 1464-1480, 1990.

[22] S. mozaffari, K. faez, F. Faradji *et al.*, "A comprehensive isolated Farsi/Arabic character database for handwritten OCR research." pp. 385-389.

[23] H. Khosravi, and E. Kabir, “Introducing a very large dataset of handwritten Farsi digits and a study on their varieties,” *Pattern Recognition Letters,* vol. 28, no. 10, pp. 1133-1141, 2007.

[24] R. Ebrahimpur, A. Esmkhani, and F. Faradji, “Farsi handwritten digit recognition based on mixture of RBF experts,” *IEICE Electron. Express,* vol. 7, no. 14, pp. 1014-1019, 2010.

[25] S. Mozaffari, K. Faez, and H. R. Kanan, "Recognition of isolated handwritten Farsi/Arabic alphanumeric using fractal codes." pp. 104-108.

[26] A. Dehghani, F. Shabini, and P. Nava, "Off-line recognition of isolated Persian handwritten characters using multiple hidden Markov models." pp. 506-510.

[27] A. Mowlaei, and K. Faez, "Recognition of isolated handwritten Persian/Arabic characters and numerals using support vector machines." pp. 547-554.

[28] M. Dehghan, and K. Faez, "Farsi handwritten character recognition with moment invariants." pp. 507-510 vol.2.

[29] J. Shanbezadeh, H. Pezashki, and A. Sarrafzadeh, "Features Extraction from Farsi Hand Written Letters." pp. 35-40.

[30] M. M. Javidi, and f. Sharifizadeh, “A Modified Decision Templates Method for Persian Handwritten Digit Recognition,” *Journal of American Science,* vol. 8, no. 1, pp. 504-512, 2012.

[31] M. M. Javidi, R. Ebrahimpur, and f. Sharifizadeh, “Persian handwritten digits recognition: A divide and conquer approach based on mixture of MLP experts,” *International Journal of the Physical Sciences,* vol. 6, no. 30, pp. 7007-7015, 2011.

[32] M. Moradi, M. A. Pourmina, and F. Razzazi, “FPGA-Based Farsi Handwritten Digit Recognition System,” *International Journal of Simulation Systems, Science & Technology,* vol. 11, pp. 17-22, 2010.