Handwritten Marathi Consonants Recognition using Multilevel Classification

C. H. Patil Research Scholar Bharati Vidyapeeth Deemed University Pune

ABSTRACT

This paper presents approach for the recognition of handwritten Marathi consonants. In order to recognize handwritten Marathi consonants, a database of handwritten Marathi consonants is developed to carry recognition experiments. Problem of handwritten Marathi consonant recognition is simplified using multilevel classificationwhich improves recognition rate. Total 36 Marathi consonants are transformed using instance simplification technique into six sub classesdepending on special property of consonants. Suitable features are extracted from different sub classes and further classification is carried out using SVM and k-NN classifiers.We have used database of 7920 characters for testing and found recognition accuracy 78.27% using SVM classifier and 73.29% using k-NN classifier.

Keywords

Marathi consonents, multilevel classification, svm, knn, pattern recognition.

1. INTRODUCTION

Devanagari alphabet developed from eastern variants of the Gupta script called Nagari, which first emerged during the 8th century. Devanagari is written from left to right direction in horizontal lines.Devanagari script is used to writeAwadhi, Bhojpuri, Hindi. Kashmiri. Konkani,Kurukh,Maithili,Marathi,Marwari,Mundari,Nepal BhasaNewari, Nepali, Pali, Rajasthani, Sanskrit, Saraiki, Sindhi, S unuwar,Sylhetiand many other languages.All of the Indo-Aryan languages originated from Sanskrit. Three Prakrit languages, simpler in structure, emerged fromSanskrit.These

S. M. Mali Department of Computer Science, MAEER's Arts, Com and Sci college,

Pune

were Saurseni, Magadhi and Maharashtri. Marathi is said to be a descendent of Maharashtri which was the Prakrit spoken by people residing in the region of Maharashtra. Marathi is one of the 23 official languages of India and is the official language of Maharashtra. There are around 100million speakers of Marathi language which is fourth largest number of native speakers in India.Marathi language consist of 49 characters out of which 12 are vowelsas shown in Fig. 1(a)and 37 are consonantsas shown in Fig.1(b), one consonant sis rarely used in Marathi, so in this part of work only 36 consonants are considered. Handwriting recognition is the ability of a computer to receive and interpret intelligible handwritten input from source like paper documents. Recognition of handwritten Marathi character is a challenging task. Different writers have different writing styles, pens, papers. Every character has a horizontal line at the top known as the header line. The header line joins all the characters in a word.Marathi also has a complex system of compound characters in which two or more consonants are combined forming a new special symbol. There are few reasons that create problems in Marathi handwritten character recognition. (i)Some characters are similar in shape do and(ii)Total number of characters is more with different strokes as compare to English language. (iii) Different users or same user can write same characters differently at different times, depending on the pen or pencil, the width of the line, the rotation of the paper, the type of paper and the mood and stress level of the person. (iv)The character can be written at different location on paper. (v)Sometimes characters are overlapped and slanted. (vi) Handwritten Marathi characters could be of various shapes andsize. Because of the above mentioned reasons the recognition of handwritten Marathi character process needs to be more efficient and accurate. This has

Г	ள	FE.	

ओ

3

t

а

अ र

b		

	छ	පා	đ	ফ	91	E	च	स्त	ज	ञ
¥	अं	ঞা:	¥	5	3	ß	he	TUT	त	হ্য
-11			ਵ	ET	ብ	দ	4	ब	H	म
			য	হ	त्य	a	2 T	च	स	3
			00	ধ্ব	석	슉	છં			

Fig 1: Handwritten Marathi Vowels and Consonents

motivated us to undertake problem of handwritten Marathi character recognition.Handwriting recognition is popular research area in academicians and researcher since 1870 [4]. Researchers have reported some work on Devanagri [1-3, 5,7,8,9,11,12,15,20,23-26,28-31,33-35], Marathi [4,6,10,13-14,18,22,27,32], Hindi [19] languages. Ajmire P.E. et. al [4] achieved 62% recognition for Marathi vowels. Archana Jain et. al. [6] used smoothing and feature set to improve recognition rate of similar shaped Marathi characters. K.V. Kale et. al. [18] used Zernike moment features set to reported 98.37% for compound Marathi characters. Vinaya Tapkir et. al. [22] used pixel density feature and Euclidean minimum distance classifier to recognize Marathi characters and reported 92.77% recognition rate. Sushma Shelke et. al. [32] reported multistage handwritten Marathi compound character recognition using neural networks and achieved 97.95%

result. Handwritten Marathi character recognition has wide range of applications like bank cheque processing, postal automation, handwritten form processing and historical document preservation. Due to wide range of applications and complexity in recognition of Marathi character is point of attraction for researchers and academicians from long time. We have used analytical approach in which word is segmented into indivisible characters and these characters aretreated as input for recognition system.



Fig 2:Phases in handwritten Marathi character recognition

Phases for handwritten Marathi character recognition are shown in the Fig. 2. Different phases are discussed as follows: Section II explains data collection phase, section III elaborates pre-processing, section IV focuses on multilevel classification, section V explains feature extraction techniques, section VI focusing on classification phase, in section VII results are discussed and section VIII concludes the research paper.

2. DATA COLLECTION

Literature review indicates that benchmark database for handwritten Marathi consonant is not available to carryexperiments. Since a benchmark database is not available our first attempt was to developsufficient database for handwritten Marathi consonants. Two A4 size sheets were designed for data collection and distributed to 20 writers which includestudents, teachers, and clerks. Every writer has to write consonants in square provided and no other constraint was imposed on writers. Specially designed sample handwritten datasheets are shown in Fig. 3.

ъ	ক্র	চ	ক্ত	æ	ক	ক্ত	ন্দ	Т	ক	ক্ত	रू
ख	হুৱ	ত্ত	ত্ত	ত্ত	ত্ত	হর	হ্য	হ্য	ত্ত্ৰ	হ্য	A
ग	ন	91	ন	ল	ল	76	16	21	σT	aL	эT
घ	ত্র	হ	হা	ET	ସ	ত	য	য	ত	হা	ত
च	च	-च	-य	-च	च	न्य	च	-च	-रा	न्च	-ग
ন্ত	8	8	8	2	89	छ	ন্থ	63	E	80	ça
জ	ज	ਤ	ज	স	ज	म	F	ञ	ज	ज	F
झ	ज्ञ	द्भ	ञ	ক	अ	ञ	Ħ	হন	झ	<u>इ</u> न	झ
त्र	X	F	F	त्र	স	স	त्र	F	¥	ਜ	म्
5	ਣ	T	2	2	S	ন্থ	S	2	2	3	2
ठ	5	2	ক	5	3	2	5	ত	5	ত	8
ड	3	3	3	3	6	ड	Z	3	ङ	उ	3
ਫ	5	2	2	2	2	2	5	2	5	2	2
ण	চ্য	তা	70	U	10	ण	OT	U	σT	T	σ
त	ਜ	ਸ	न	ਜ	Ъ.	75	Ъ	F	ਜ	F	त
थ	হ	হ	থ	য	গ্র	T	T	T	T	T	গ্র
द	द	द	र	द	द	द	द	द	Z	2	35
ध	ত	ฮ	ন্ম	ত্য	গ্র	ন্থ	ध	T	T	T	গ
न	ਸ	न	4	7	4	70	40	न	10	F	न
प	म	দ	Q	Ŧ	प	T	শ	А	प	प	Ţ
দ্দ	P	똬	ক	দ্দ	F	দ্দ	P	ক	Th	দ	দ
-	1								-		
ब	đ	Ø	ৰ	a	न	ৰ	đ	19	à	ति	a,
ब भ	র্ব ঙ্গ	ର ୨୩	ৰ স্থ	ाब जि	न अ	র ঙ্গ	র স	10 A	ते अ	ते अ	8
ब भ म	ব ক স	জ জ স	ৰ ঙ্গ স	। জি সি সি	न ज्र	র ঙ্গ ক্ষ	র স স	10 87 87	वे अ म	জ উদ স	100 No 100
ब भ म य	त्र इ म य	ন্থ প স ম	ৰ ক্স স ম	ৰ জি স	न क्र म य	র ঙ্গ শ ম	ৰ প্ল প্ল	ने स्र य	वे अ म म रा	बे इस म य	র প্র ম
ब भ म य र	র্ব ন্স স্ব ম্ব	ब इन म म म म	ৰ ক্স দ ম ম	। ब जि म म य भ भ	ब क्र म य र	র ঙ্গ ফ ম ম	র প্র স য ম	6 87 87 21 21 21	ने भ्र म य र	ब इस इस इस्	N R R G
ब भ म य र ल	त्र म राष्ट्र म राष्ट्र	ब म म म म म म म	18 18 F 12 18 18	। ব কি দ ম ম ম ম ম	ন জ্ব দ য र ত	দ দ স্থ স স স স স স স স স স স স স স স স স স	A A A A A A A A A A A A A A A A A A A	10 87 87 87 87 87 87 87 87 87 87 87 87 87	ते अम् म म म म र ल	बि इस सं स स स	21 24 R R R G
ब भ म य र ल	में हो हो से स्व में रहे हैं।	व कि म य र र र र	ৰ স্পি স স ম দ দ ব	। ব জি দ মি প । কি ব	न इन्ह म य र लि	অ (S) ২ দ দ দ দি	ब भ भ भ भ भ भ भ भ भ भ भ भ भ	ति स स स स स स स	ते अम म म र र ल व	बि इस् में में में में में में में में में में	A 21 A A A B
ब भ म य र ल व श	ह का स स स स स स	बि मि म म म म म म म म म	ब क्र म म म म म म म म म म म म म म म म म म म	ब जि म य र म य र लि न म	ब क्रा म य र ल व का	10 3 7 R 7 8 8	गि भ भ भ भ भ भ भ भ भ भ भ भ भ	10 57 57 72 7 10 70 10 10 10 10 10 10 10 10 10 10 10 10 10	में म म म म म म म म म म म म म म म म म म म	बि इस स स स स स	10 10 H H H H 10 10
ब भ म य र ल ब श	ब म म म म म म म म म म म म म म म म म म म	बि कि म य तर क कि म य तर ब	ब के म य क ल म य क ल म	ब कि म म म म म म म म म म	न म् म् य र ल व न्	म म म म म ज ज ज भ म	न भ भ भ भ भ भ भ भ भ भ भ भ भ भ भ भ भ	10 की क रूप क रिंग ज	ति म म या रू लि व रिंग म	बि इस् से य स स् य स सि वि ब	A RE AN EN H RE RO BI
ब म य र ल व श ष स	ब इस स स स स स स	बि मि म य र म छ प छ य म	ब क्रम म म म म म म म म म म म म म म म म म म	ब क्र म य २ ति व म भ य २ ति व म	न जा जा म य र ल व न् न् न्	ह जिम् म म म म	ा में	10 x x x x x x x	वि म म म म म म म म	बि इस् में य स स सि वि म स	12 B B A A A B B B B
ब भ म य र ल व श ष स ह	র ন্দ ন ন ন ন ন ন ন ন ন ন ন ন ন ন ন ন ন ন ন	ब कि म म य र छ य य म र र	ब के म म य क त त ब ब क न व क न व क न व	TA TA TA TA TA TA TA TA	न इन इन इन इन इन इन इन्ह इन इन्ह इन इन्ह इन इन्ह इन इन्ह इन इन्ह इन इन्ह इन इन्ह इन इन्ह इन्ह	अ म म र छ ज ज ज म	जि अस्त स्त स्त स्ति व स्त स्ति	5 F F F F F F F F F F F F F F F F F F F	LA H K E & N R H K K	बि इस इस इस हर हर हर ब ब ब ब स हर	SA B B A A A B B B B
ब भ म य र ल व श ष स ह ळ	ब म म म न न न न न न न न न न न न न न न न	B 5 H B 3 7 K R R	ब अप्ति म म म म स स स स स स स स स स	ब कि म य र कि व म व मि टि छ	न ज म य र ल व श म र छ	क जिस्स जिस्स जिस्स जि	ज भ भ भ भ भ भ भ भ भ भ भ भ भ भ भ भ भ भ भ	8 ~ 1	B S # R P P N N R R R R	बि इस् मेर य र सि विगे वि	BI FI B B F F F F B B B B
ब भ म य र ल व श ष स ह ळ क्ष	ब क म य र ल न न न न न न न न न न न न न न न न न न	ब कि म य र छ य य म ठ छ छि	ब इ म म म म म म म म म म म म म म म म म म	ब कि म य र लि न म य रि छ मि	न का म य र ल व का व म र छ छ जि	म म म म म म म म म म	ा भाषा का साम भाषा का साम का	10 BR R Z 2 10 R A A A C B R	THE COLOR THE	बि इस् स् स् स् स् स् स् स् स् स् स्	THE BE S TH TH THE THE THE THE
ब भ म य र ल व श ष स ह ळ क्ष ज्ञ	র ন্দ ন ন ন ন ন ন ন ন ন ন ন ন ন ন ন ন ন ন ন	मा की की की मा की	ब अप्ति म य प स त व श व श व श व श व श व श व श त व श त श्र	ब क म य र लि व म य र ल कि म	न ज म य र ल व ज म हि छ जि म	म म म र ए क का लग स र क मि मि	ब भ म य म ल व झ य स द छ छ झ	10 B B D A D A D A D A B B A A	THE	के आ में ये रे लि के से रे लि सि स	TO TO EN TO

Fig 3:Sample data sheets for handwritten Marathi consonants

The data sheets were scanned using a flat bed scanner at a resolution of 1200 dpi and stored as gray scale images. From the scanned gray scale image, the character images were cropped manually and stored in respective class folders. For each consonant approximately 220 samples were stored in respective class folders. Finally database of 7920 consonants images are ready to carry experiments. Fig. 4 shows some gray scale characters cropped from the scanned image of a datasheet.



Fig 4: Sample handwritten Marathi consonants

3. PRE PROCESSING

Image preprocessing techniques enhancesquality of data images prior to extracting features for recognition. Also preprocessing is the technique used to reduce noise introduced in the image by various reasons like scanner, ink, paper etc.Firstly,threshold of input image is computed by using Ostu's method, which chooses the threshold to minimize the intra-class variance of the threshold black and white pixels. Using computed threshold value intensity imageconverted to black and white. Morphological thinning operation is performed to remove pixels so that an object without holes shrinks to a minimally connected stroke, and an object with holes shrinks to a ring halfway between the hold and outer boundary. The binarized character image is mapped onto a standard plane so as to give a representation of fixed dimensionality for classification. Character normalization is used to reduce the inter-class variation of the shapes of the characters in order to facilitate feature extraction process and improve their recognition accuracy. Linear normalization method used to standardize the character images. The standard plane is considered as a square of size 60 x 60. The aspect ratio of the character image is not disturbed due to normalization. Thinned image of each character contains header cap known as 'shirorekha' placed above every character, this cap is identified and removed from the character images. To remove header cap we have computed horizontal projection for each character. In this methodwe have identified a row which contains maximum black pixels in top 33% rows of the character. All pixels values in the identified row are set to zero which will remove the *shirorekha* from the character.

4. MULTILEVEL CLASSIFICATION

Total numbers of Marathi consonants are 36 and in these consonants interclass similarities are more. By experiment it has been observed that single feature is not sufficientfor classification because recognition rate using single feature is very low. Also classifying and recognizing 36 consonants is time consuming, hence to solve the problem of classification well known approach of problem solving based on transformation is used. Instance simplification is type of transformation which solves a problem by first transforming into simple instance of the same problem with some special property that makes the problem easier to solve. In phase I to phase IV problem is simplified by classifying 36 consonants into six sub classes according to their characteristics as shown in the Fig. 5. Following is the discussion on phase I to phase IV sub classification where 36 consonants are classified into six classes:

4.1.Phase I Sub classification:

This sub classification is based on bar characters and no bar characters. Marathi consonants are broadly classified into two major categories bar characters and no bar characters. Bar character are those characters having presence of vertical bar. In order to verify whether a bar is present in the character, vertical projection of imagewas taken and if any column contains more than 70% of black pixels, then label it as bar character otherwise label as no bar character. As shown in the Fig. 6(a) column number 57 contains 86% black pixels hence labeled as bar character and as shown in the Fig.6(b) all columns contains less than 70% black pixels hence labeled as no bar character. Thus 36 consonants are classified into two subclasses, first contains 27 bar characters and second contains nine no bar characters as shown in phase I of fig 5.



Fig 6(a):Bar character



Fig 6(b): No Bar character



Fig. 5:Phases in Multilevel classification

4.2.Phase II sub classification:

This sub classification is based on enclosed region. In phase II, 27 bar characters are broadly classified into two major categories having enclosed region or not. To verify whether enclosed region is present, the numbers of holes are counted in the character using eight connectivity as shown in the Fig. 7(a). If one or more than one enclosed region exists in the character then label it as enclosed region character otherwise not enclosed region character. Now 27 bar characters are sub classified into 18 enclosed region bar characters and nine not enclosed region bar characters as shown in phase II of Fig. 5. Similarly nine no bar characters and four not enclosed region no bar characters as shown in phase II of Fig. 5.



Fig 7(a): Enclosed Region character



Fig 7(b): Not enclosed region character **4.3.** Phase III sub classification:

This sub classification is based on number of components. In phase III 21 bar enclosed region characters are classified into two subclasses, depending on whether number of component is one or more. Presence of a component can be verified using region properties for each labeled region. In all 18 bar enclosed Region characters are sub classified into 14 characters having a component and four characters having two component.





Fig. 8(b): One component character 4.4.Phase IV sub classification:

This sub classification is based on number of rows containing at least one black pixel. We have 14 bar enclosed region characters having one component and further classified into two subclasses, depending on whether character's 80% rows contains at least one black pixel in first 75% columns. Out of 18 bar enclosed region characters having one component we get 10 characters satisfying above condition and four characters does not satisfy the condition.



Fig. 9 (a): 80% row contains at least one black pixels character



Fig 9 (b): less than 80% row contains at least one black pixels character.

Using above sub classification method problem of handwritten Marathi consonant recognition is simplified into six sub classes as follows:

Sub class I. Barnot enclosed region (9 characters)



10(a): Consonants having bar and enclosed region.

Sub class II: Barenclosed region with two components (4

characters).

Fig. 10(b): Consonants having bar, enclosed region and having two components.

Sub class III. Sub class III: Barenclosed region with one componentand having 80% rows contains at least one black pixel in first 75% columns (**10 characters**)



Fig. 10(c): Consonants having bar, enclosed region, one component and black pixels

Sub class IV: Barenclosed region with one component and less than 80% rows contains at least one black pixel in first 75% columns (**4 characters**).



Fig. 10(d): Consonants having bar, enclosed region, one component and not black pixels.

Sub class IV: No bar enclosed region (5 characters)



Fig. 10(e): Consonants does not have bar and having enclosed region.

Sub class V: No bar not enclosed region (4 characters).



Fig. 10(f): Consonants does not have bar and enclosed region.

Now we have 36 consonants divided into 6 different sub classes as discuss above. Feature extraction methods suitable for different sub classes are discussed in the Section VI.

5. FEATURE EXTRACTION

Feature for any character is the distinguishable property of the character and is extracted such that it quantifies some significant characteristics of the character.

6.5.Symmetric Density

By experiment it is observed that symmetric density feature is more suitable for sub class III and sub class VI. In the proposed method, the binary image representing the handwritten character is pre-processed and is normalized to a size of 60 x 60 pixels. The size-normalized image is divided into n equal zones. As shown in Fig. 11(a-e) consonant image is divided into equal size 4 zones, 9 zones, 16 zones, 25 zones and 36 zones and density feature for every zone is calculated and is stored into feature vector. The density of each zone is computed by taking the ratio of total number of object pixels to total number of pixels in that zone. This is carried out for every zone in the image. Finally, 90 features are extracted from the image and feature vector stores there 90 features. Density features calculated by using the following equation.



6.6.Horizontal Features

By experiment it is observed that horizontal feature is more suitable for sub class I, subclass IV and sub class V.To extract horizontal feature of the binary image representing the handwritten character is first preprocessed and is normalized to size of 60 x 60 pixels. The size-normalized image is divided into 36 equal zones as shown in Fig. 12(a); each zone is of size 10 x 10 as shown in Fig. 12(b). Each zone has 10 horizontal lines, each horizontal line is summed to get a single sub feature and thus 10 sub-features are obtained from the each zone.These 10 sub-features values are averaged to form a single feature value and assigned as horizontal feature to the corresponding zone. This procedure is sequentially repeated for the all the zones. Finally, 36 features are extracted for 36 zones for each character.

6.7.Diagonal Features

By experiment it is observed that horizontal feature is more suitable for sub class 2.In the proposed method, the binary image representing the handwritten character is preprocessed and is normalized to a size of 60 x 60 pixels. The sizenormalized image is divided into 36 equal zones each of size is 10 x 10. Each zone has 19 diagonal lines, each diagonal line is summed to get a single sub feature and thus 19 sub-features are obtained from the each zone. These 19 sub-features values are averaged to form a single feature value and placed in the corresponding zone. This procedure is sequentially repeated for the all the zones. Finally, 36 features are extracted for each character. In addition, 12 features are obtained by averaging the values placed in zones row wise and columnwise, respectively. As a result, for every character is 36 + 12 features total 48 features are stored in feature vector.

	<u>p</u>															
	0	0	1	0	0	0	0	0	0	•	 		_			
→	•	0	0	1	0	0	0	þ	U	♪	с		*			
a	-0-	0	0	1	0	0	0	0	0	•	0	0	(1)	0	0	0
	Ð	0	0	-1-	0	0	0	•	-0-	•	0.8	0.1	1.2	0	0	0
U	•	0	0	1	0	0	-0	0	0	•	1	0	0.9	0.2	0.6	0.8
	۵	0	0	1	0	0	-0-	•	0	-0	0.4	0	1	0	0	0.9
	•	0	-1	0	-0	0	0	0	0	•	1.3	0.7	0.8	0	0	0.8
≫−₹	-0-	0	1	0	0	0	-0	0	0	->	0.5	0.4	0.4	0.2	0.2	0.4
	┛	0	1	•	•	0	0	_0_	0	•						
	۰	0	1	-0-	0	0	0	-0-	0	- 🕪						

Fig. 12 (a): Character divided into 36 zonesFig. 12 (b): Zone z3 of size 10x10. Fig. 12 (c): Total 36 zones average values



Fig. 12 (a): Character divided into 36 zonesFig. 12 (b): Zone z3 of size 10x10. Fig. 12 (c): Total 36 zones average values

6. CLASSFICATION

Pattern recognition from labeled 'training' data is called supervised learning, and when no labeled data is available it is called as unsupervised learning. Here we have used supervised learning approach in which 80% samples are used for training and 20% samples for testing. As discussed in section V feature vector is created for all images and k-NN as well as SVM classifiers are used to carry experiments. k-NN and SVM classifiers are discussed below:

6.1.K-NN

The k-Nearest Neighbor (k-NN) classifies an unknown sample based on the known classification of its neighbors [9, 11, 13,

16, 18, 24, and 25]. Suppose that a set of samples with known classification is available, the so-called training set. Intuitively, each sample should be classified similarly to its surrounding samples. Therefore, if the classification of asample is unknown, then it could be predicted by considering the classification of its nearest neighbor samples. Given an unknown sample and a training set, all the distances between the unknown sample and all the samples in the training set can be computed. The distance with the smallest value corresponds to the sample in the training set closest to the unknown sample. Therefore, the unknown sample may be classified based on the classification of this nearest neighbor. k- NN is an instance-based learning type classifier, or lazy learning where the

function is only approximated locally and all computation is deferred until classification. Euclidean distance is used.

6.2.SVM

Support vector machines (SVM) are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification [5, 7, 9, 13-15, 18, 20, 24-26]. SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier.

7. RESULTS

Database of 7920 samples are used to carry out the experiments out of which 80% samples are used for training and 20% samples are used for testing. Multilevel

classification approach is introduced here to simplify the handwritten Marathi consonant recognition problem. Once consonants are classified into six sub classes horizontal, diagonal and symmetric density features are used. Recognition accuracy for subclass I is 74.17%, subclass II is 76.77%, subclass III is 60.63%, subclass IV is 67.92%, subclass V is 74.00% and subclass VI is 86.25% using k-NN classifier as shown in Table I. Recognition accuracy for subclass I is 79.58%, subclass II is 85.83%, subclass III is 64.17%, subclass IV is 77.92%, subclass V is 81.50% and subclass VI is 80.62% using SVM classifier as shown in Table I. Average recognition rate for all subclasses achieved is 73.29% and 78.27% using k-NN and SVM classifier as shown in Table I. Table II shows recognition accuracy for all 36 consonants separately by using k-NN and SVM classifier.

Table I: Recognition Rate for sub	class I to VI using k-NN and SVM Classifier
-----------------------------------	---

Sr. No.	Class	Character set	Feature used	No. of features	k-NN Classifier	SVM Classifier
1	Subclass I	6 characters	Horizontal	36	74.17	79.58
2	Subclass II	3 characters	Diagonal	48	76.77	85.83
3	Subclass III	12 characters	Symmetric Density	90	60.63	64.17
4	Subclass IV	6 characters	Horizontal	36	67.92	77.92
5	Subclass V	5 characters	Horizontal	36	74.00	81.50
6	Subclass VI	4 characters	Symmetric Density	90	86.25	80.62
			Average Rec	ognition Rate	73.29	78.27

Table II: Recognition accuracy for each consonant using k-NN and SVM Classifier

SR. NO.	SUB CLASS	CONSONANT	k-NN	SVM	SR. NO.	SUB CLASS	CONSONANT	k-NN	SVM
1		फ	75	65	19		H	47.5	47.5
2		च	97.5	100	20		হা	55	47.5
3		দ	90	85	21	ш	स	72.5	57.5
4		य	52.5	70	22		F	47.5	45
5	I	ज	90	82.5	23		\$	65	47.5
6		F	77.5	85	24		9	72.5	90
7		त	72.5	75	25	IV	ब	67.5	90
8		CI	52.5	62.5	26		d	75	70
9		*	55	72.5	27		D	50	62.5
10		ফ	47.5	92.5	28		E	57.5	67.5
11	п	TE	72.5	87.5	29	v	8	77.5	82.5
12		TU	85	82.5	30		ne	85	90

International Journal of Computer Applications (0975 – 8887) National Conference on Digital Image and Signal Processing 2016

13		হা	72.5	87.5	31		र	82.5	97.5
14		ab .	95	82.5	32		00	67.5	70
15		ভ	85	62.5	33		5	97.5	87.5
16	III	ञ	60	57.5	34	VI	5	100	97.5
17		E	47.5	35	35	VI	ĥu	50	57.5
18		H	30	40	36		उ	97.5	80

8. CONCULSION

Single level classification is not sufficient for recognition of Marathi consonants. Hence, in this paper multilevel classification approach is introduced. Multilevel classification is depends on the structural feature like bar, region, number of component, black pixels. After multilevel classification suitable features are extracted from different sub classes and recognition is carried out using k-NN and SVM classifiers. Overall recognition rate for all subclasses achieved is 73.29% and 78.27% using k-NN and SVM classifier. SVM classifier is more suitable for consonants classification. In future recognition accuracy can be improved by extracting more suitable features and combining classifiers.

9. ACKNOWLEDGMENTS

The author is grateful to Dr. M. S. Prasad for their guidance, helpful discussion and encouragement duringthis work.

10. REFERENCES

- R. Karwankar, A.S. Bhalchandra, "Stroke Pattern Identification In Devanagari Character Recognition", Published In International Journal Of Advanced Engineering & Application, Jan. 2010.
- [2] Aarti Desai, Latesh Malik, "A Novel Approach To Thinning Of Devnagari Characters", International Journal Of Engineering Science And Technology (IJEST)ISSN: 0975-5462February 2011.
- [3] Aarti Desai, Latesh Malik, Rashmi Welekar, "A New Methodology For Devnagari CharacterRecognition", Jmacademy Of It & Management Volume -1 Issue 1issn: 2229-6115january, 2011
- [4] Ajmire P.E., Warkhede S.E., "Handwritten Marathi Character (Vowel) Recognition", Advances In Information Mining, ISSN: 0975–3265, Volume Issue 2, 2010, Pp-11-132010
- [5] Anilkumar N. Holambe, Ravinder C.Thool, Sushilkumar N. Holambe, "Printed And Handwritten Character& Number Recognition OfDevanagariScript Using Gradient Features", Advances In Multimedia - An International Journal (AMIJ)
- [6] Archana P.Jane, Mukesh. A.Pund, "Recognition Of Similar Shaped HandwrittenMarathi Characters Using Artificial NeuralNetwork", Global Journal Of Computer Science And TechnologyNeural & Artificial IntelligenceVolume 12 Issue 11 Version 1.0 Year 2012Online ISSN: 0975-4172 & Print ISSN: 0975-43502012.
- [7] Ashutosh Aggarwal, Rajneesh Rani, Renudhir, "Handwritten Devanagari Character Recognition Using Gradient Features", International Journal Of Advanced

Research In Computer Science And Software Engineering Volume 2, Issue 5ISSN: 2277 128X41030

- [8] Bikash Shaw, Swapan Kr. Parui, Malayappan Shridhar, "Offline Handwritten Devanagariword Recognition: A Holistic Approach BasedOn Directional Chain Code Feature And HMM", International Conference On Information Technologyieee ICIT 978-0-7695-3513-5/08 20082008
- [9] Brijmohan Singh, Ankush Mittal, Debashis Ghosh, "An Evaluation Of Different Feature Extractors And ClassifiersFor Offline Handwritten Devnagari Character Recognition", Journal Of Pattern Recognition Research 2 (2011) 269-2772011
- [10] C.Namrata Mahender, K.V.Kale, "Structured Based Feature Extraction Of Handwritten Marathi Word", International Journal Of Computer Applications (0975 – 8887)Volume 16– No.6, February 20112011
- [11] Chavan S.V., Kale K.V., Kazi M.M., Rode Y.S., "Recognition Of Handwritten Devanagari Compound Character A MomentFeature Based Approach", International Journal Of Machine IntelligenceIssn: 0975-2927 & E-Issn: 0975-9166, Volume 5, Issue 1, 2013, Pp.-421-425.2013
- [12] Dipak V. Koshti, Sharvari Govilkar, "Segmentation Of Touching Characters In HandwrittenDevanagari Script", UACEE International Journal Of Computer Science And Its Applications - Volume 2: Issue 2 [ISSN 2250 - 3765]
- [13] G. G. Rajput S. M. Mali, "Fourier Descriptor Based Isolated Marathi HandwrittenNumeral Recognition", International Journal Of Computer Applications (0975 – 8887)Volume 3 – No.4,40330
- [14] G. G. Rajput, S. M. Mali, "Isolated Handwritten Marathi Numerals RecognitionBased Upon Fourier Descriptors AndFreeman Chain Code", International Journal Of Computational Intelligence Research ISSN 0973-1873 Volume 6, Number 2 (2010), Pp. 289– 2982010
- [15] Gaurav Agrawal, Kshitij, Amitabha Mukerjee, Nimit Kumar, "Handwritten Devanagari ScriptSegmentation Using Support Vector Machines"
- [16] Holambe A.N., Thool R.C., Shinde U.B. And Holambe S.N., "Brief Review Of Research On Devanagari Script", International Journal Of Computational Intelligence Techniques, ISSN: 0976–0466, Volume 1, Issue 2, 2010, Pp-06-092010
- [17] J.Pradeep, E.Srinivasan, S.Himavathi, "Diagonal Feature Extraction Based Handwritten Character

System Using Neural Network", International Journal Of Computer Applications (0975 – 8887) Volume 8– No.9, October 2010

- [18] Karbhari V. Kale, , Prapti D. Deshmukh Shriniwas V. Chavan, Majharoddin M. Kazi, , Yogesh S. Rode, "Zernike Moment Feature Extraction For HandwrittenDevanagari (Marathi) Compound Character Recognition", (IJARAI) International Journal Of Advanced Research In Artificial Intelligence, Vol. 3, No.1, 2014
- [19] M. Hanmandlu, O.V. Ramana Murthy, Vamsi Krishna Madasu, "Fuzzy Model Based Recognition Of Handwritten Hindi Characters", Digital Image Computing Techniques And Applications 0-7695-3067-2/07 2007 IEEEDOI 10.1109/DICTA.2007.82IEEE2007
- [20] Mahesh Jangid, Devanagari "Isolated Character RecognitionBy Using Statistical Features", International Journal On Computer Science And Engineering (IJCSE) Vol. 3 No. 6 ISSN: 0975-3397 40695
- [21] Mansi Shah, Gordhan B Jethava, "A Literature Review On Hand Written Character Recognition", Indian Streams Research JournalVol -3, Issue –2, March.2013 Issn:-2230-78502013
- [22] Vinaya. S. Tapkir, Sushma.D.Shelke, "OCR For Handwritten Marathi Script", International Journal Of Scientific & Engineering Research Volume 3, Issue 8, August-2012 ISSN 2229-55182012
- [23] N. B. Mapari, A. L. Telang, R. K. Rajbhure, "A Study Of Devnagri Handwritten Character Recognition System", International Journal Of Computer Science And Applications Vol. 4, No. 2, June2011 ISSN: 0974-10032011
- [24] Neeraj Pratap, Shwetank Arya, "A Review Of Devnagari Character Recognition From Past To Future", International Journal Of Computer Science And Telecommunications Volume 3, Issue 6, June 2012
- [25] Neha Sahu, R. K. Rathy, Indu Kashyap, "Survey And Analysis Of Devnagari CharacterRecognition Techniques Using Neural Networks", International Journal Of Computer Applications (0975 – 888) Volume 47– No.15, June 2012
- [26] O. V. Ramana Murthy, M. Hanmandlu, "Zoning Based Devanagari Character Recognition", International Journal Of Computer Applications (0975 – 8887)Volume 27–No.4, August 2011
- [27] P.E.Ajmire, RV Dharaskar, V M Thakare, "A Comparative Study Of Handwritten Marathi Character

Recognition", National Conference On Innovative Paradigms In Engineering & Technology (NCIPET-2012)Proceedings Published By International Journal Of Computer Applications (IJCA)2012

- [28] Ratnashil N Khobragade, Nitin A. Koli, Mahendra S Makesar, "A Survey On Recognition Of Devnagari Script", International Journal Of Computer Applications & Information TechnologyVol. II, Issue I, January 2013 (ISSN: 2278-7720)41275
- [29] S. Arora, D. Bhattacharjee, M. Nasipuri , D.K. Basu, M.Kundu, L.Malik, "Study Of Different Features On Handwritten Devnagari Character", Second International Conference On Emerging Trends In Engineering And Technology, ICETET-092009
- [30] S. Arora, D. Bhattacharjee, M. Nasipuri, D. K. Basu, M. Kundu, "Recognition Of Non-Compound Handwritten Devnagari Characters Using A Combination Of MLP And MInimum Edit Distance", International Journal Of Computer Science And Security (IJCSS), Volume (4): Issue (1)
- [31] Sandhya Arora, Debotosh Bhattacharjee, Mita Nasipuri,Dipak Kumar Basu, Mahantapas Kundu, "Combining Multiple Feature Extraction TechniquesFor Handwritten Devnagari Character Recognition", 2008 IEEE Region 10 Colloquium And The Third ICIIS, Kharagpur, INDIA December 8-10.2008
- [32] Sushama Shelke, Shaila Apte, "A Multistage Handwritten Marathi Compound CharacterRecognition Scheme Using Neural Networks And Wavelet Features, International Journal Of Signal Processing", Image Processing And Pattern RecognitionVol. 4, No. 1, March 2011
- [33] Veena Bansal, R. M. K. Sinha, "Integrating Knowledge Sources In Devanagari Text Recognition System", IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART A: SYSTEMS AND HUMANS, VOL. 30, NO. 4, 1083–442736708
- [34] Veena Bansal And R. M. K. Sinha, "Segmentation Of Touching Characters In Devanagari"
- [35] Veena Bansal And R. M. K. Sinha, "SEGMENTATION OF TOUCHING AND FUSEDDEVANAGARI CHARACTERS".
- [36] Vikas J Dongre, Vijay H Mankar, "A Review Of Research On Devnagari Character Recognition",International Journal Of Computer Applications (0975 – 8887)Volume 12– No.2, November 2010.