Chapter 4

Recognition of Online Handwritten Assamese Characters: Initial Results

This chapter presents a discussion on character recognition experiments performed for online handwritten Assamese characters. For the experiments, the characters are taken from TU-OHAC dataset and standard features reported in the literature (for other scripts) are explored. The discussion on character recognition experiment includes preprocessing steps performed on the characters followed by the feature computation stage. Three feature vectors FS:1, FS:2 and FS:3 are formed. The chapter concludes with a description on the classification results for the above feature vectors.

4.1 Preprocessing

The preprocessing steps are performed before the feature computation stage. Preprocessing is a general first step in a recognition system. Preprocessing plays a very important role which can directly affect the recognition performance. Three types of preprocessing tasks performed on the dataset of online handwritten Assamese characters are namely, size normalization, smoothening and resampling.

4.1.1 Size Normalization

Since the characters are written in different sizes at different locations of the text input box, all of the characters have different ranges of values for the X and Y coordinates. Therefore, the different ranges of values for the coordinate points are normalized to a uniform range of values. This step is performed by normalizing the values of X and Y coordinates of the characters to the range (0, n) as per the size normalization algorithm presented in Algorithm 4.1. An original character is shown in Figure 4.1(a) and the character after size normalization is shown in Figure 4.1(b).

Algorithm 4.1

INPUT: Online handwritten Assamese character

OUTPUT: Normalized character

Read x values of all the coordinate points (x,y) of the character

Read y values of all the coordinate points (x,y) of the character

$X \leftarrow$ x value of the coordinate point (x,y)

$Y \leftarrow$ y value of the coordinate point (x,y)

$X_{min} \leftarrow$ Smallest value of x

$Y_{min} \leftarrow$ Smallest value of y

$X_{max} \leftarrow$ Largest value of x

$Y_{max} \leftarrow$ Largest value of y

Compute: $\frac{1}{(X_{max}-X_{min})}$

Compute: $\frac{1}{(Y_{max}-Y_{min})}$

$S_x \leftarrow \frac{1}{(X_{max}-X_{min})}$

$S_y \leftarrow \frac{1}{(Y_{max}-Y_{min})}$

Compute: $S_x*(X - X_{min})$

Compute: $S_y*(Y - Y_{min})$

$X_{normalized} \leftarrow S_x*(X - X_{min})$

//Comment: $0 \leq X_{normalized} \leq 1$

$Y_{normalized} \leftarrow S_y*(Y - Y_{min})$

//Comment: $0 \leq Y_{normalized} \leq 1$

//Comment: To scale the range [0, 1] of values of $X_{normalized}$ and $Y_{normalized}$ to the range [0, n], multiply the values of //$X_{normalized}$ and $Y_{normalized}$ by the scaling constant n.

$X_{normalized} \leftarrow (X_{normalized})*n$

//$0 \leq X_{normalized} \leq n$

$Y_{normalized} \leftarrow (Y_{normalized})*n$

//$0 \leq Y_{normalized} \leq n$

(a) (b)

Figure 4.1. (a) Original character and (b) character after size normalization

4.1.2 Smoothening

The smoothening of characters is required to reduce the sharp changes or roughness (jitters) from the stroke. It removes jitters from the data while preserving underlying patterns. Here, the smoothing of the characters was performed using a 3-point moving average filter in order to remove jitters from the characters. The mathematical expression s of the 3-point moving average filter are given in the equations 4.1 and 4.2.

$$x'(n) = (1/3)[x(n-1)+x(n)+x(n+1)] \quad (4.1)$$

$$y'(n) = (1/3)[y(n-1)+y(n)+y(n+1)] \quad (4.2)$$

The coordinate point $(x'(n), y'(n))$ belongs to the smoothened trajectory of the character. Figure 4.2(b) represents the smoothing result of the character shown in the Figure 4.2(a).

(a) (b)

Figure 4.2 (a) Original character and (b) character after smoothing

4.1.3 Resampling

Since an online handwritten character is a string of coordinate points, different characters have a different number of points along the trajectory from the start to the end. Therefore, different characters need to be resampled to fixed number points. Here, all of the online handwritten Assamese characters have been resampled to a fixed number of 100 points. The resampling is performed with the `resample` function which changes the sampling rate for a sequence to any rate that is a ratio of two integers. The mathematical expressions of the resampling are given in the equations 5.1 and 5.2.

$$x'(n) = \text{resample}(x(k),p,q) \quad (5.1)$$

$$y'(n) = \text{resample}(y(k),p,q) \quad (5.2)$$

Here the function "resample ( )" in equation 5.1 resamples the sequence $x(k)$ at p/q times the original sampling rate. The length of the result $x'(n)$ is p/q times

the length of x(k). Similarly, the function "resample ( )" in equation 5.2 resamples the sequence y(k) at p/q times the original sampling rate. The length of the result $y'(n)$ is p/q times the length of y(k). Figure 4.3(b) represents the resampling of the character shown in the Figure 4.3(a) to 100 points.

(a) (b)

Figure 4.3. (a) Original character and (b) character after re-sampling

4.2 Feature Set

A general description of various types of features of online handwritings is presented in Chapter 2. In this chapter, standard Feature Vector for online handwritten Assamese characters consisting of geometrical features, structural features and statistical features is discussed. Considering various combinations of features computed from online handwritten Assamese characters three feature vectors are formed for the character recognition experiments.

4.2.1 Feature Computation

The features of online handwritten Assamese characters are computed from the preprocessed characters. The (X, Y) coordinate points of the resampled characters are inputs to the feature computation module and subsequently the feature vectors are generated. Three feature vectors namely, FS:1, FS:2 and FS:3 are generated all for online handwritten Assamese numerals, basic alphabetic characters and conjunct consonants (Juktakkhors).

4.2.1.1 Geometrical Features

The geometrical features computed from online handwritten Assamese characters are namely, X and Y coordinate values of the resampled characters, coordinates of the start and end points of each character, position of the stylus (up or down), direction angle with reference to the horizontal axis (cosine of the angle), direction angle with reference to the vertical axis (sine of the angle), curvature with reference to the horizontal axis and curvature with reference to the vertical axis.

4.2.1.2 Structural Feature

The structural feature computed from the online handwritten Assamese characters is the headline feature.

4.2.1.3 Statistical Features

The statistical features computed from the online handwritten Assamese characters are namely, mean of the X and Y coordinate values, normalized horizontal coordinates, normalized vertical coordinates, number of strokes in the character, distance between the start and the end points of the character. The zone feature is another statistical feature which was computed from online handwritten Assamese characters. The values of the X and Y coordinates of the characters are divided into sixty-four zones or sub-intervals of the normalized range of values (0, 200) of X and Y. The number of points in a character falling into each sub-interval of (0, 200) is considered as a feature of the character (zone feature).

4.2.2 Initial Feature Vectors

The feature set FS:1 is an initial trial [10,104]. The feature set FS:2 is based on [38,39] and the feature set FS:3 is an enhancement of the feature set FS:2. The feature set FS:3 includes all the features of FS:2 along with five other features.

4.2.2.1 Feature Vector FS:1

Feature set FS:1 consists of geometrical features, structural features and statistical features. This feature set is attempted as an initial effort to classify online handwritten Assamese characters. The number of attributes (including the class attribute) in feature set FS:1 is 278. The following is the listing of features in the feature set FS:1.

- X and Y coordinate values of the resampled characters
- Zone feature
- Coordinates of the start and end points of each character
- Distance between the start and the end points of the character
- Number of strokes in the character
- Mean of the X and Y coordinate values
- Standard deviation of X and Y coordinates values
- Headline feature

4.2.2.2 Feature Vector FS:2

The feature set FS:2 was suggested in [38,39] and the same was used in UNIPEN dataset as a script independent feature vector [39]. Here, an effort is made to use this feature set in online handwritten Assamese characters dataset. The number of attributes (including the class attribute) in feature set FS:2 is 591. The following is the listing of features in the feature set FS:2.

- Normalized horizontal coordinates
- Normalized vertical coordinates
- Direction angle with reference to the horizontal axis (cosine of the angle)
- Direction angle with reference to the vertical axis (sine of the angle)
- Curvature with reference to the horizontal axis

- Curvature with reference to the vertical axis
- Position of the stylus (up or down)

4.2.2.3 Feature Vector FS:3

The feature set FS:3 was formed in this dissertation in an attempt to improve to the recognition results of online handwritten Assamese characters. By adding five features taken from the feature set FS:1 to the feature set FS:2 the feature set FS:3 is formed. The number of attributes (including the class attribute) in feature set FS:3 is 663.

- Normalized horizontal coordinates
- Normalized vertical coordinates
- Direction angle with reference to the horizontal axis (cosine of the angle)
- Direction angle with reference to the vertical axis (sine of the angle)
- Curvature with reference to the horizontal axis
- Curvature with reference to the vertical axis
- Position of the stylus (up or down)
- Zone feature
- Distance between the start and the end points of the character
- Number of strokes in the character
- Mean of the X and Y coordinate values
- Standard deviation of the X and Y coordinate values

4.3 Classification

Here we present the classification results of online handwritten Assamese characters, which include numerals, basic alphabetic characters and conjunct consonants (Juktakkhors). Character recognition experiments were performed on the 10 numeral classes (total of 45×10=450 numerals), 52 classes of basic

alphabetic characters (total of 45×52=2340 basic alphabetic characters) and 121 classes of conjunct consonants (total of 45×121=5445 conjunct consonants) available in the dataset. In order to conduct the classification, we explored the support vector machine (SVM) [85]. The architecture for the classification is shown in Figure 4.4.

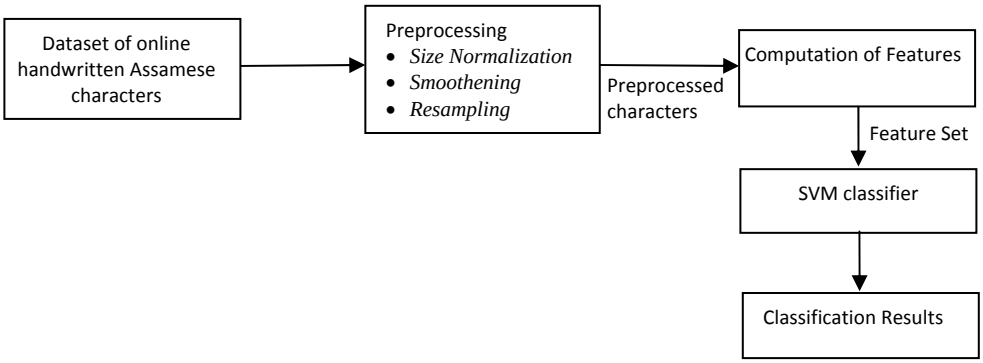

Figure 4.4 Architecture for the classification of characters

4.3.1 SVM based Classifier

SVM based classifications were performed for the classification of online handwritten Assamese numerals, basic alphabetic characters and *Juktakkhors*. A *k*-fold cross validation procedure was used for the training and testing of the SVM classifier with various parametric settings. In *k*-fold cross-validation, the original sample was randomly partitioned into *k* equal size subsamples. Of the *k* subsamples, a single subsample was retained as the validation data for testing the model, and the remaining $k-1$ subsamples was used as training data. The cross-validation process was then repeated *k* times, in which each of the *k* subsamples were used exactly once as the validation data. The *k* results from the folds were then combined to produce a single estimation. All of the observations were used for both training and validation, and each observation was used for validation exactly one time. 10-fold cross-validation was done. Classification experiments are performed using linear kernel, polynomial kernel and Gaussian radial basis

functions (RBF) which are the common choices for kernel functions in SVM. For the classification using linear kernel, the default setting of SVM parameters of linear kernel was attempted. Grid search operation was used to obtain the optimal values of the SVM parameters pair (C, E) in case of polynomial kernel, where C is the penalty parameter of the error term and E is the exponent of polynomial kernel. Similarly, grid search operation was used to obtain the optimal values of the SVM parameters pair (C, gamma) in case of RBF kernel, where C is the penalty parameter of the error term and gamma is the value of γ in the equation $K(x_i, x_j) = \exp(-\gamma \| x_i - x_j \|^2), \gamma > 0$ of RBF kernel.

4.3.2 Classification Results

We present the classification results of online handwritten Assamese numerals, basic alphabetic characters and *Juktakkhors*. Classification results are based on the feature vectors FS:1, FS:2 and FS:3. We present individual recognition rates of each character belonging to a specific class followed by the average recognition results for the class. We also present the overall recognition results considering all the 8235 online handwritten Assamese characters (consisting of numerals, basic alphabetic characters and Juktakkhors) based on feature set FS:3. The individual recognition rates in case of all 183 characters are also determined.

4.3.2.1 Classification Results: Numerals

The classification results for online handwritten Assamese numerals are presented below. Individual recognition rates of numerals are presented in Table 4.1, average recognition rates of numerals are presented in the Table 4.2 and parameters settings for kernels including plots of Grid search for RBF kernel are presented after the tables.

Table 4.1 Individual recognition rates of online handwritten Assamese numerals

Sl. No.	Class	Recognition rate (%)								
		Feature set FS:1			Feature set FS:2			Feature set FS:3		
		Types of SVM Kernel			Types of SVM Kernel			Types of SVM Kernel		
		Linear kernel	Polynomial kernel (C=3 & E=4)	RBF kernel C=9 & gamma=0.07)	Linear kernel	Polynomial kernel (C=1 & E=4)	RBF kernel C=8 & gamma=0.03125)	Linear kernel	Polynomial kernel (C=1 & E=4)	RBF kernel C=8 & gamma=0.03125)
1	SUNYA	97.80	97.80	97.80	95.60	100	100	95.60	97.80	100
2	EK	100	100	100	100	100	95.60	100	100	95.60
3	DUI	97.80	100	97.80	95.60	97.80	97.80	95.60	97.80	97.80
4	TINI	100	97.80	97.80	97.80	97.80	97.80	97.80	97.80	97.80
5	CARI	93.30	97.80	97.80	82.20	93.30	95.60	88.90	91.10	95.60
6	PAC	95.60	100	100	100	97.80	97.80	97.80	97.80	97.80
7	CAY	100	100	100	100	100	100	100	100	100
8	XAT	97.80	100	100	97.80	97.80	97.80	97.80	97.80	97.80
9	ATH	100	100	100	100	100	100	100	100	100
10	NAA	97.80	97.80	97.80	97.80	95.60	95.60	97.80	97.80	95.60

Table 4.2 Average recognition rates of online handwritten Assamese numerals

Online handwritten Assamese numerals					
Type of SVM kernel	Total Number of Instances	Correctly Classified Instances	Average recognition rate (Across Classes)	Standard Deviation (SD) of recognition rate (Across Classes)	Average $\pm$ SD
Feature Set FS:1					
Linear	450	441	98.00	2.21	98.00 $\pm$ 2.21
Polynomial (C=1, E=4)	450	446	99.11	1.54	99.11 $\pm$ 1.54
RBF (C=3,Gamma=0.05)	450	445	98.89	1.16	98.89 $\pm$ 1.16
Feature Set FS:2					
Linear	450	435	96.67	5.37	96.67 $\pm$ 5.37
Polynomial (C=1, E=4)	450	441	98.00	2.21	98.00 $\pm$ 2.21
RBF (C=8,Gamma=0.03125)	450	440	97.78	1.79	97.78 $\pm$ 1.79
Feature Set FS:3					
Linear	450	437	97.11	3.32	97.11 $\pm$ 3.32
Polynomial (C=1, E=4)	450	440	97.78	2.56	97.78 $\pm$ 2.56
RBF (C=8,Gamma=0.03125)	450	440	97.78	1.79	97.78 $\pm$ 1.79

Linear Kernel

The average recognition rates achieved for the online handwritten Assamese numerals using linear kernel were 98.00% (based on FS:1), 96.67% (based on FS:2), and 97.11% (based on FS:3) with a 10 fold cross validation process. A total of 450 characters were used as samples in the numeral recognition experiment (refer to Table 4.2). The kernel parameter settings C = 1 and E = 1 are associated with the linear kernel.

Polynomial Kernel

The polynomial kernel parameter setting C = 1 and E = 4 for the feature sets FS:1, FS:2 and FS:3 was obtained by grid search operation. The average recognition rates achieved for the online handwritten assamese numerals using polynomial kernel were 99.11% (based on FS:1), 98.00% (based on FS:2), and 97.78% (based on FS:3) with a 10 fold cross validation process. A total of 450 characters were used as samples in the numeral recognition experiment (refer to Table 4.2).

RBF Kernel

- *Parameter Settings: Grid Search*

The kernel parameter settings (C=3,Gamma=0.05) for the feature set FS:1, (C=8,Gamma=0.03125) for the feature set FS:2 and (C=8,Gamma=0.03125) for the feature set FS:3 were obtained by grid search operation. The plots of grid search are shown in the Figure 4.5 through Figure 4.7.

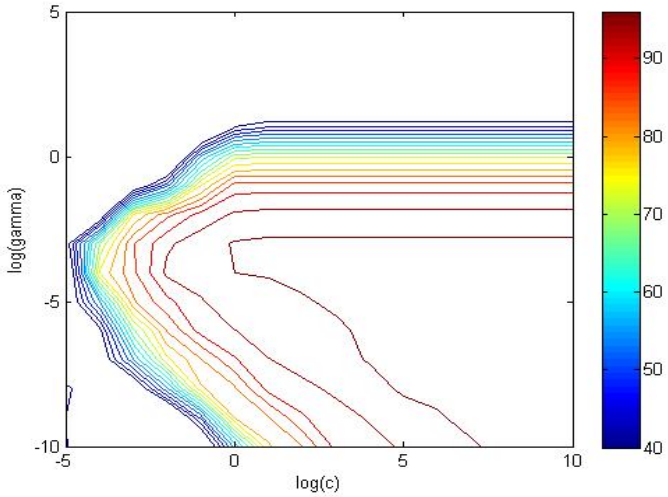

Figure 4.5 Grid search plot of C and gamma for RBF kernel (using FS:1 for numerals)

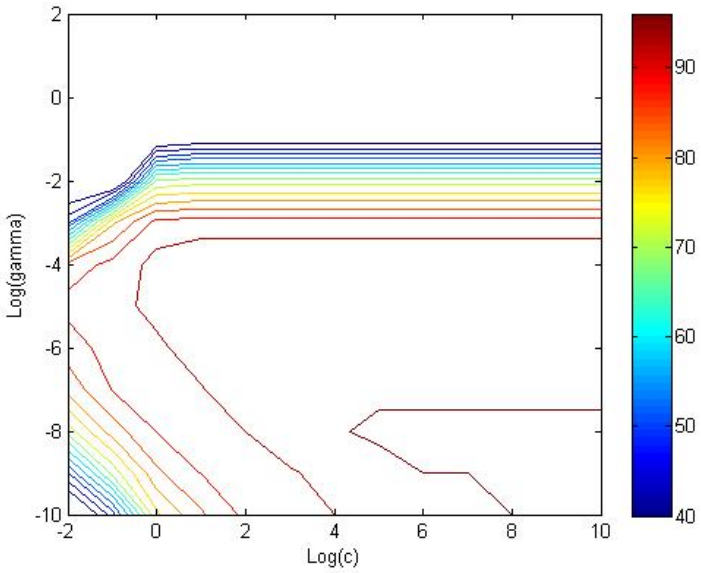

Figure 4.6 Grid search plot of C and gamma for RBF kernel (using FS:2 for numerals)

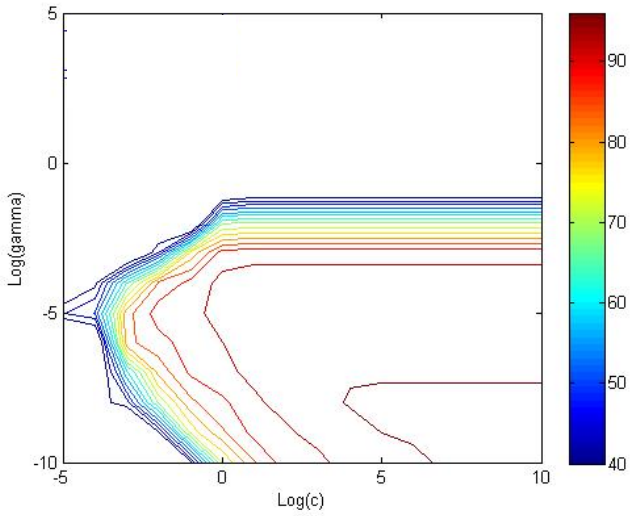

Figure 4.7 Grid search plot of C and gamma for RBF kernel (using FS:3 for numerals)

- *Experimental Results*

The average recognition rates achieved for the online handwritten Assamese numerals using RBF kernel were 98.89% (based on FS:1), 97.78% (based on FS:2), and 97.78% (based on FS:3) with a 10 fold cross validation process. A total of 450 characters were used as samples in the numeral recognition experiment (refer to Table 4.2).

4.3.2.2 Classification Results: Basic Alphabetic Characters

The classification results for online handwritten Assamese basic alphabetic characters are presented below. Individual recognition rates of alphabetic characters are presented in Table 4.3 (a), Table 4.3 (b) and Table 4.3 (c) for feature set FS:1, FS:2 and FS:3 respectively. The average recognition rates of alphabetic characters are presented in the Table 4.4 and parameter settings for kernels including plots of Grid search for RBF kernel are presented after the tables.

Table 4.3 (a) Individual recognition rates of the online handwritten Assamese basic alphabetic characters for feature set FS:1

Sl. no.	Class	Recognition rate (%) Feature set FS:1 Types of SVM Kernel			Sl. no.	Class	Recognition rate (%) Feature set FS:1 Types of SVM Kernel		
		Linear kernel	Polynomial kernel (C=3 & E=4)	RBF kernel C=9 & Gamma=0.07)			Linear kernel	Polynomial kernel (C=3 & E=4)	RBF kernel C=9 & Gamma=0.07)
1	A	57.80	64.40	68.90	27	TA	75.60	75.60	73.30
2	AA	80.00	73.30	80.00	28	THA	84.40	91.10	88.90
3	E	62.20	71.10	66.70	29	DA	68.90	77.80	86.70
4	EE	48.90	57.80	66.70	30	DHA	66.70	64.40	68.90
5	U	62.20	68.90	80.00	31	NA	46.70	53.30	55.60
6	UU	71.10	75.60	80.00	32	PA	71.10	73.30	77.80
7	REE	80.00	80.00	82.20	33	PHA	48.90	60.00	68.90
8	AE	100	97.80	100	34	BA	68.90	71.10	77.80
9	OI	84.40	86.70	91.10	35	BHA	57.80	64.40	71.10
10	O	91.10	91.10	91.10	36	MA	57.80	48.90	48.90
11	OU	88.90	82.20	84.40	37	AJA	46.70	57.80	60.00
12	KA	75.60	77.80	77.80	38	RA	55.60	60.00	62.20
13	KHA	66.70	75.60	80.00	39	LA	44.40	55.60	64.40
14	GA	84.40	80.00	84.40	40	WA	71.10	66.70	71.10
15	GHA	55.60	66.70	66.70	41	TXA	84.40	86.70	88.90
16	NG	93.30	93.30	93.30	42	MXA	46.70	60.00	57.80
17	CA	80.00	86.70	86.70	43	DXA	31.10	37.80	44.40
18	CCA	80.00	80.00	86.70	44	HA	73.30	82.20	84.40
19	JA	80.00	88.90	91.10	45	KHYA	66.70	77.80	77.80
20	JHA	88.90	93.30	91.10	46	AYA	60.00	53.30	60.00
21	NIYA	100	100	100	47	DRA	80.00	91.10	91.10
22	MTA	80.00	88.90	88.90	48	DHRA	57.80	64.40	71.10
23	MTHA	71.10	86.70	80.00	49	KTA	91.10	95.60	100.00
24	MDA	73.30	66.70	68.90	50	ANSR	84.40	91.10	97.80
25	MDHA	73.30	73.30	77.80	51	BXG	97.80	97.80	97.80
26	MNA	66.70	73.30	75.60	52	CBN	86.70	91.10	88.90

Table 4.3 (b) Individual recognition rates of the online handwritten Assamese basic alphabetic characters for feature set FS:2

Sl. no.	Class	Recognition rate (%) Feature set FS:2 Types of SVM Kernel			Sl. no.	Class	Recognition rate (%) Feature set FS:2 Types of SVM Kernel		
		Linear kernel	Polynomial kernel (C=3 & E=4)	RBF kernel C=8 & Gamma=0.03125)			Linear kernel	Polynomial kernel (C=3 & E=4)	RBF kernel C=8 & Gamma=0.03125)
1	A	82.20	86.70	82.20	27	TA	80.00	86.70	88.90
2	AA	77.80	82.20	77.80	28	THA	93.30	93.30	93.30
3	E	80.00	77.80	77.80	29	DA	82.20	91.10	88.90
4	EE	64.40	68.90	66.70	30	DHA	68.90	62.20	68.90
5	U	82.20	77.80	77.80	31	NA	51.10	66.70	66.70
6	UU	73.30	73.30	82.20	32	PA	75.60	77.80	82.20
7	REE	75.60	84.40	84.40	33	PHA	60.00	68.90	64.40
8	AE	97.80	97.80	100	34	BA	60.00	64.40	68.90
9	OI	91.10	93.30	93.30	35	BHA	82.20	75.60	88.90
10	O	95.60	95.60	93.30	36	MA	53.30	62.20	64.40
11	OU	91.10	88.90	91.10	37	AJA	55.60	62.20	62.20
12	KA	75.60	77.80	80.00	38	RA	55.60	66.70	60.00
13	KHA	82.20	80.00	82.20	39	LA	53.30	60.00	68.90
14	GA	71.10	75.60	77.80	40	WA	71.10	73.30	75.60
15	GHA	64.40	68.90	66.70	41	TXA	88.90	86.70	86.70
16	NG	95.60	97.80	95.60	42	MXA	55.60	62.20	64.40
17	CA	86.70	88.90	91.10	43	DXA	37.80	44.40	48.90
18	CCA	86.70	84.40	86.70	44	HA	88.90	93.30	88.90
19	JA	77.80	82.20	84.40	45	KHYA	86.70	91.10	91.10
20	JHA	88.90	86.70	88.90	46	AYA	51.10	55.60	64.40
21	NIYA	100	100	100	47	DRA	82.20	86.70	88.90
22	MTA	77.80	80.00	73.30	48	DHRA	77.80	80.00	73.30
23	MTHA	75.60	77.80	75.60	49	KTA	93.30	91.10	91.10
24	MDA	80.00	82.20	80.00	50	ANSR	82.20	88.90	91.10
25	MDHA	75.60	77.80	77.80	51	BXG	97.80	95.60	93.30
26	MNA	73.30	77.80	77.80	52	CBN	82.20	86.70	86.70

Table 4.3 (c) Individual recognition rates of the online handwritten Assamese basic alphabetic characters for feature set FS:3

Sl. no.	Class	Recognition rate (%)			Sl. no.	Class	Recognition rate (%)		
		Feature set FS:3					Feature set FS:3		
		Types of SVM Kernel					Types of SVM Kernel		
		Linear kernel	Polynomial kernel (C=3 & E=4)	RBF kernel C=8 & Gamma=0.03125)			Linear kernel	Polynomial kernel (C=3 & E=4)	RBF kernel C=8 & Gamma=0.03125)
1	A	82.20	82.20	82.20	27	TA	82.20	86.70	88.90
2	AA	80.00	84.40	80.00	28	THA	93.30	93.30	95.60
3	E	80.00	73.30	75.60	29	DA	82.20	91.10	88.90
4	EE	64.40	68.90	71.10	30	DHA	71.10	68.90	71.10
5	U	80.00	82.20	80.00	31	NA	64.40	68.90	64.40
6	UU	75.60	77.80	80.00	32	PA	75.60	82.20	82.20
7	REE	77.80	82.20	86.70	33	PHA	62.20	66.70	66.70
8	AE	97.80	97.80	100	34	BA	62.20	66.70	71.10
9	OI	93.30	95.60	93.30	35	BHA	80.00	86.70	88.90
10	O	95.60	97.80	93.30	36	MA	53.30	60.00	64.40
11	OU	91.10	91.10	91.10	37	AJA	53.30	66.70	64.40
12	KA	77.80	80.00	80.00	38	RA	57.80	64.40	60.00
13	KHA	82.20	82.20	84.40	39	LA	57.80	60.00	68.90
14	GA	77.80	71.10	77.80	40	WA	75.60	73.30	75.60
15	GHA	66.70	71.10	68.90	41	TXA	91.10	88.90	88.90
16	NG	95.60	97.80	95.60	42	MXA	53.30	64.40	64.40
17	CA	88.90	93.30	91.10	43	DXA	44.40	44.40	51.10
18	CCA	80.00	82.20	86.70	44	HA	91.10	95.60	88.90
19	JA	80.00	82.20	84.40	45	KHYA	91.10	91.10	91.10
20	JHA	91.10	91.10	91.10	46	AYA	55.60	60.00	66.70
21	NIYA	100	100	100	47	DRA	84.40	84.40	88.90
22	MTA	77.80	86.70	82.20	48	DHRA	80.00	84.40	77.80
23	MTHA	75.60	82.20	77.80	49	KTA	95.60	91.10	91.10
24	MDA	80.00	84.40	80.00	50	ANSR	86.70	93.30	91.10
25	MDHA	80.00	82.20	77.80	51	BXG	97.80	95.60	93.30
26	MNA	80.00	80.00	77.80	52	CBN	82.20	84.40	86.70

Table 4.4 Average recognition rates of basic alphabetic characters

Online handwritten Assamese basic alphabetic characters					
Type of SVM kernel	Total Number of Instances	Correctly Classified Instances	Average recognition rate (Across Classes)	Standard Deviation (SD) of recognition rate (Across Classes)	Average ± SD
Feature Set FS:1					
Linear	2340	1674	71.54	15.77	71.54 ± 15.77
Polynomial (C=3, E=4)	2340	1764	75.39	14.39	75.39 ± 14.39
RBF (C=9,Gamma=0.07)	2340	1834	78.38	13.31	78.38 ± 13.31
Feature Set FS:2					
Linear	2340	1795	76.71	14.29	76.71 ± 14.29
Polynomial (C=3, E=4)	2340	1862	79.57	12.23	79.57 ± 12.23
RBF (C=8, Gamma=0.03125)	2340	1879	80.29	11.50	80.29 ± 11.50
Feature Set FS:3					
Linear	2340	1834	78.38	13.45	78.38 ± 13.45
Polynomial (C=3, E=4)	2340	1896	81.03	12.24	81.03 ± 12.24
RBF (C=8,Gamma=0.03125)	2340	1899	81.15	11.08	81.15 ± 11.08

Linear Kernel

The average recognition rates achieved for the online handwritten Assamese basic alphabetic characters using linear kernel were 71.54% (based on FS:1), 76.71% (based on FS:2), and 78.38% (based on FS:3) with a 10 fold cross validation process. A total of 2340 characters were used as samples in the numeral recognition experiment (refer to Table 4.4). The kernel parameter settings C = 1 and E = 1 are associated with the linear kernel.

Polynomial Kernel

The polynomial kernel parameter setting C = 3 and E = 4 for the feature sets FS:1, FS:2 and FS:3 was obtained by grid search operation. The average recognition

rates achieved for the online handwritten Assamese basic alphabetic characters using polynomial kernel were 75.39% (based on FS:1), 79.57% (based on FS:2), and 81.03% (based on FS:3) with a 10 fold cross validation process. A total of 2340 characters were used as samples in the numeral recognition experiment (refer to Table 4.4).

RBF Kernel

- *Parameter Settings: Grid Search*

The kernel parameter settings (C=9,Gamma=0.07) for the feature set FS:1, (C=8,Gamma=0.03125) for the feature set FS:2 and (C=8,Gamma=0.03125) for the feature set FS:3 were obtained by grid search operation. The plots of grid search are shown in the Figure 4.8 through Figure 4.10.

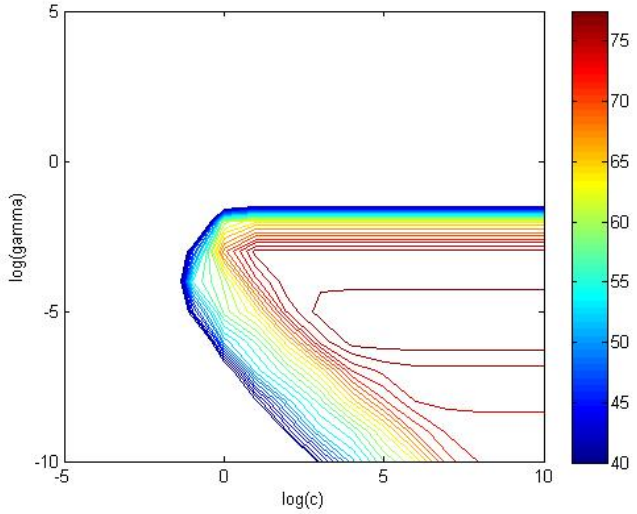

Figure 4.8 Grid search plot of C and gamma for RBF kernel (using FS:1 for basic alphabetic characters)

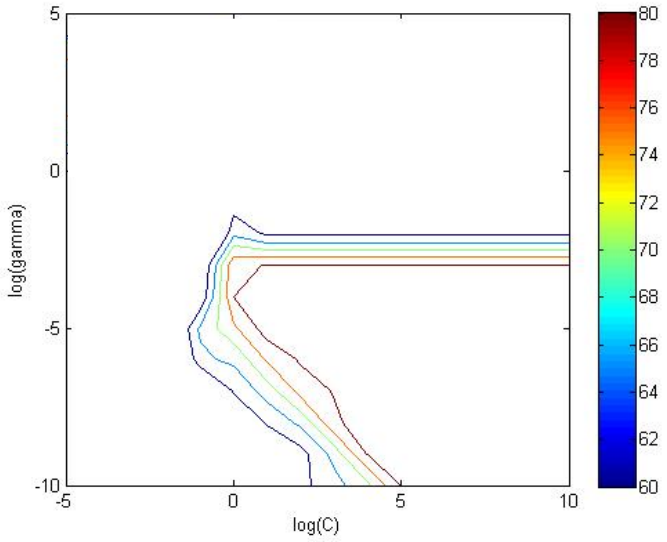

Figure 4.9 Grid search plot of C and gamma for RBF kernel (using FS:2 for basic alphabetic characters)

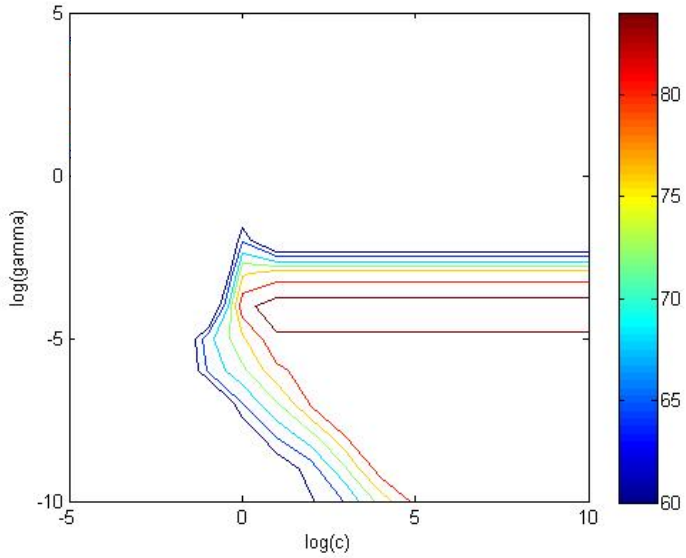

Figure 4.10 Grid search plot of C and gamma for RBF kernel (using FS:3 for basic alphabetic characters)

- *Experimental Results*

The average recognition rates achieved for the online handwritten Assamese basic alphabetic characters using RBF kernel were 78.38% (based on FS:1), 80.29% (based on FS:2), and 81.15% (based on FS:3) with a 10 fold cross validation process. A total of 2340 characters were used as samples in the basic alphabetic character recognition experiment (refer to Table 4.4).

4.3.2.3 Classification Results: Juktakkhors

The classification results for online handwritten Assamese conjunct consonants (*Juktakkhors*) are presented below. Individual recognition rates of *Juktakkhors* are presented in Table 4.5 (a) through Table 4.3 (i) for feature sets FS:1, FS:2 and FS:3. Table 4.5 (a), Table 4.5 (b) and Table 4.5 (c) present individual recognition rates of *Juktakkhors* based on feature set FS:1 for linear, polynomial and RBF kernels respectively. Table 4.5 (d), Table 4.5 (e) and Table 4.5 (f) present individual recognition rates of *Juktakkhors* based on feature set FS:2 for linear, polynomial and RBF kernels respectively. Table 4.5 (g), Table 4.5 (h) and Table 4.5 (i) present individual recognition rates of *Juktakkhors* based on feature set FS:3 for linear, polynomial and RBF kernels respectively. The average recognition rates of *Juktakkhors* are presented in the Table 4.6 and parameters settings for kernels including plots of Grid search for RBF kernel are presented after the tables.

Table 4.5(a) Individual recognition rates of online handwritten Assamese conjunct consonants (*Juktakkhors*) for feature set FS:1 using Linear Kernel of SVM

Individual recognition rates of *Juktakkhors* (%)											
Feature set FS:1											
Type of SVM Kernel: Linear Kernel											
Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%
1	KK	62.20	32	NTH	42.20	63	FN	80.00	94	NKH	88.90
2	KT	64.40	33	NDH	48.90	64	FT	28.90	95	NGH	82.20
3	KTT	68.90	34	NNN	15.60	65	SK	33.30	96	NGKH	80.00
4	KS	60.00	35	NB	26.70	66	SSTH	51.10	97	TTH	91.10
5	KL	57.80	36	NS	31.10	67	SSM	24.40	98	PN	40.00
6	KM	55.60	37	NM	20.00	68	SSN	24.40	99	HN	57.80
7	GL	40.00	38	DB	62.20	69	SSB	31.10	100	XN	77.80
8	CC	73.30	39	QJ	100.00	70	ST	35.60	101	MF	53.30
9	CCC	60.00	40	PTT	75.60	71	SP	33.30	102	BB	62.20
10	JJ	77.80	41	PL	44.40	72	SPH	48.90	103	LB	24.40
11	JB	48.90	42	DV	77.80	73	STH	40.00	104	LM	15.60
12	BJ	55.60	43	BL	53.30	74	SKH	51.10	105	BHM	68.90
13	GN	44.40	44	BD	60.00	75	NGG	48.90	106	ML	15.60
14	TN	42.20	45	TB	57.80	76	NGC	93.30	107	SL	57.80
15	JJB	80.00	46	MM	40.00	77	FP	42.20	108	PS	22.20
16	LG	26.70	47	MV	46.70	78	NGN	77.80	109	KHR	84.40
17	TT	55.60	48	MP	31.10	79	XM	64.40	110	GR	60.00
18	GDH	75.60	49	MN	6.70	80	NGJ	80.00	111	GHR	44.40
19	GM	64.40	50	NTR	53.30	81	MNTH	73.30	112	JR	64.40
20	GHN	35.60	51	MB	46.70	82	NGK	53.30	113	TRR	64.40
21	MDD	60.00	52	LK	44.40	83	KR	71.10	114	DRR	62.20
22	NT	66.70	53	MND	95.60	84	TRU	64.40	115	DHRR	66.70
23	NN	62.20	54	FK	48.90	85	BHR	51.10	116	PRR	42.20
24	NMM	60.00	55	LD	26.70	86	THB	86.70	117	BRR	66.70
25	TTT	66.70	56	LL	20.00	87	DG	46.70	118	MRR	28.90
26	TTB	75.60	57	LP	28.90	88	DGH	51.10	119	TSR	84.40
27	TM	57.80	58	LT	55.60	89	DD	68.90	120	DSR	24.40
28	TR	68.90	59	SN	51.10	90	DDH	64.40	121	HRR	71.10
29	NTT	42.20	60	SC	88.90	91	HR	71.10			
30	RRG	77.80	61	SM	68.90	92	GGU	91.10			
31	NDD	26.70	62	SB	57.80	93	GGN	73.30			

Table 4.5(b) Individual recognition rates of online handwritten Assamese conjunct consonants (*Juktakkhors*) for feature set FS:1 using Polynomial Kernel of SVM

Individual recognition rates of *Juktakkhors* (%)											
Feature set FS:1											
Type of SVM Kernel: Polynomial Kernel with C=1 & E=4											
Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%
1	KK	64.40	32	NTH	48.90	63	FN	86.70	94	NKH	84.40
2	KT	77.80	33	NDH	57.80	64	FT	31.10	95	NGH	84.40
3	KTT	64.40	34	NNN	20.00	65	SK	42.20	96	NGKH	77.80
4	KS	53.30	35	NB	31.10	66	SSTH	53.30	97	TTH	91.10
5	KL	64.40	36	NS	31.10	67	SSM	15.60	98	PN	40.00
6	KM	55.60	37	NM	31.10	68	SSN	31.10	99	HN	55.60
7	GL	46.70	38	DB	62.20	69	SSB	31.10	100	XN	77.80
8	CC	75.60	39	QJ	100.00	70	ST	37.80	101	MF	57.80
9	CCC	71.10	40	PTT	82.20	71	SP	31.10	102	BB	57.80
10	JJ	84.40	41	PL	46.70	72	SPH	44.40	103	LB	33.30
11	JB	64.40	42	DV	80.00	73	STH	40.00	104	LM	22.20
12	BJ	62.20	43	BL	51.10	74	SKH	62.20	105	BHM	80.00
13	GN	46.70	44	BD	64.40	75	NGG	66.70	106	ML	20.00
14	TN	42.20	45	TB	64.40	76	NGC	91.10	107	SL	55.60
15	JJB	77.80	46	MM	53.30	77	FP	55.60	108	PS	42.20
16	LG	26.70	47	MV	57.80	78	NGN	71.10	109	KHR	84.40
17	TT	64.40	48	MP	37.80	79	XM	68.90	110	GR	60.00
18	GDH	73.30	49	MN	13.30	80	NGJ	82.20	111	GHR	51.10
19	GM	68.90	50	NTR	57.80	81	MNTH	73.30	112	JR	68.90
20	GHN	37.80	51	MB	46.70	82	NGK	60.00	113	TRR	71.10
21	MDD	68.90	52	LK	51.10	83	KR	73.30	114	DRR	57.80
22	NT	64.40	53	MND	93.30	84	TRU	77.80	115	DHRR	66.70
23	NN	64.40	54	FK	51.10	85	BHR	51.10	116	PRR	48.90
24	NMM	55.60	55	LD	33.30	86	THB	86.70	117	BRR	71.10
25	TTT	73.30	56	LL	31.10	87	DG	57.80	118	MRR	31.10
26	TTB	77.80	57	LP	24.40	88	DGH	60.00	119	TSR	88.90
27	TM	60.00	58	LT	51.10	89	DD	75.60	120	DSR	40.00
28	TR	73.30	59	SN	53.30	90	DDH	64.40	121	HRR	77.80
29	NTT	48.90	60	SC	91.10	91	HR	73.30			
30	RRG	77.80	61	SM	68.90	92	GGU	93.30			
31	NDD	48.90	62	SB	51.10	93	GGN	77.80			

Table 4.5(c) Individual recognition rates of online handwritten Assamese conjunct consonants (*Juktakkhors*) for feature set FS:1 using RBF Kernel of SVM

Individual recognition rates of *Juktakkhors* (%)											
Feature set FS:1											
Type of SVM Kernel: RBF Kernel with C=1 & gamma=0.07											
Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%
1	KK	60.00	32	NTH	44.40	63	FN	88.90	94	NKH	91.10
2	KT	71.10	33	NDH	57.80	64	FT	22.20	95	NGH	77.80
3	KTT	64.40	34	NNN	17.80	65	SK	35.60	96	NGKH	75.60
4	KS	51.10	35	NB	33.30	66	SSTH	51.10	97	TTH	88.90
5	KL	64.40	36	NS	42.20	67	SSM	24.40	98	PN	24.40
6	KM	55.60	37	NM	11.10	68	SSN	22.20	99	HN	60.00
7	GL	33.30	38	DB	57.80	69	SSB	28.90	100	XN	71.10
8	CC	73.30	39	QJ	91.10	70	ST	42.20	101	MF	48.90
9	CCC	64.40	40	PTT	73.30	71	SP	37.80	102	BB	57.80
10	JJ	77.80	41	PL	55.60	72	SPH	53.30	103	LB	33.30
11	JB	62.20	42	DV	71.10	73	STH	44.40	104	LM	11.10
12	BJ	62.20	43	BL	44.40	74	SKH	60.00	105	BHM	68.90
13	GN	40.00	44	BD	46.70	75	NGG	53.30	106	ML	17.80
14	TN	22.20	45	TB	62.20	76	NGC	93.30	107	SL	53.30
15	JJB	75.60	46	MM	40.00	77	FP	51.10	108	PS	28.90
16	LG	26.70	47	MV	55.60	78	NGN	73.30	109	KHR	91.10
17	TT	48.90	48	MP	24.40	79	XM	66.70	110	GR	46.70
18	GDH	80.00	49	MN	8.90	80	NGJ	84.40	111	GHR	44.40
19	GM	51.10	50	NTR	53.30	81	MNTH	66.70	112	JR	71.10
20	GHN	17.80	51	MB	42.20	82	NGK	66.70	113	TRR	75.60
21	MDD	64.40	52	LK	35.60	83	KR	71.10	114	DRR	71.10
22	NT	73.30	53	MND	93.30	84	TRU	71.10	115	DHRR	66.70
23	NN	57.80	54	FK	31.10	85	BHR	48.90	116	PRR	51.10
24	NMM	60.00	55	LD	22.20	86	THB	84.40	117	BRR	68.90
25	TTT	53.30	56	LL	15.60	87	DG	57.80	118	MRR	6.70
26	TTB	80.00	57	LP	22.20	88	DGH	53.30	119	TSR	86.70
27	TM	64.40	58	LT	35.60	89	DD	71.10	120	DSR	26.70
28	TR	73.30	59	SN	48.90	90	DDH	60.00	121	HRR	84.40
29	NTT	35.60	60	SC	86.70	91	HR	64.40			
30	RRG	66.70	61	SM	62.20	92	GGU	91.10			
31	NDD	31.10	62	SB	55.60	93	GGN	77.80			

Table 4.5(d) Individual recognition rates of online handwritten Assamese conjunct consonants (*Juktakkhors*) for feature set FS:2 using Linear Kernel of SVM

Individual recognition rates of *Juktakkhors* (%)											
Feature set FS:2											
Types of SVM Kernel: Linear kernel											
Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%
1	KK	66.70	32	NTH	68.90	63	FN	93.30	94	NKH	84.40
2	KT	86.70	33	NDH	60.00	64	FT	46.70	95	NGH	88.90
3	KTT	66.70	34	NNN	28.90	65	SK	37.80	96	NGKH	93.30
4	KS	60.00	35	NB	37.80	66	SSTH	48.90	97	TTH	86.70
5	KL	57.80	36	NS	22.20	67	SSM	42.20	98	PN	37.80
6	KM	55.60	37	NM	24.40	68	SSN	37.80	99	HN	71.10
7	GL	42.20	38	DB	60.00	69	SSB	33.30	100	XN	64.40
8	CC	88.90	39	QJ	100.00	70	ST	57.80	101	MF	60.00
9	CCC	77.80	40	PTT	84.40	71	SP	42.20	102	BB	64.40
10	JJ	88.90	41	PL	55.60	72	SPH	48.90	103	LB	28.90
11	JB	66.70	42	DV	82.20	73	STH	57.80	104	LM	24.40
12	BJ	66.70	43	BL	55.60	74	SKH	64.40	105	BHM	82.20
13	GN	40.00	44	BD	64.40	75	NGG	77.80	106	ML	40.00
14	TN	57.80	45	TB	68.90	76	NGC	86.70	107	SL	53.30
15	JJB	75.60	46	MM	42.20	77	FP	53.30	108	PS	26.70
16	LG	37.80	47	MV	73.30	78	NGN	82.20	109	KHR	84.40
17	TT	55.60	48	MP	53.30	79	XM	73.30	110	GR	80.00
18	GDH	73.30	49	MN	26.70	80	NGJ	84.40	111	GHR	60.00
19	GM	71.10	50	NTR	66.70	81	MNTH	82.20	112	JR	73.30
20	GHN	60.00	51	MB	51.10	82	NGK	64.40	113	TRR	62.20
21	MDD	80.00	52	LK	66.70	83	KR	93.30	114	DRR	75.60
22	NT	66.70	53	MND	91.10	84	TRU	82.20	115	DHRR	77.80
23	NN	57.80	54	FK	55.60	85	BHR	73.30	116	PRR	57.80
24	NMM	60.00	55	LD	26.70	86	THB	84.40	117	BRR	75.60
25	TTT	84.40	56	LL	24.40	87	DG	64.40	118	MRR	55.60
26	TTB	73.30	57	LP	42.20	88	DGH	73.30	119	TSR	93.30
27	TM	73.30	58	LT	44.40	89	DD	71.10	120	DSR	26.70
28	TR	93.30	59	SN	57.80	90	DDH	55.60	121	HRR	82.20
29	NTT	48.90	60	SC	93.30	91	HR	77.80			
30	RRG	75.60	61	SM	88.90	92	GGU	91.10			
31	NDD	40.00	62	SB	60.00	93	GGN	82.20			

Table 4.5(e) Individual recognition rates of online handwritten Assamese conjunct consonants (*Juktakkhors*) for feature set FS:2 using Polynomial Kernel of SVM

Individual recognition rates of *Juktakkhors* (%)											
Feature set FS:2											
Types of SVM Kernel: Polynomial Kernel with C=3 & E=4											
Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%
1	KK	66.70	32	NTH	68.90	63	FN	88.90	94	NKH	86.70
2	KT	84.40	33	NDH	66.70	64	FT	48.90	95	NGH	91.10
3	KTT	66.70	34	NNN	33.30	65	SK	35.60	96	NGKH	91.10
4	KS	62.20	35	NB	48.90	66	SSTH	57.80	97	TTH	88.90
5	KL	68.90	36	NS	31.10	67	SSM	46.70	98	PN	42.20
6	KM	60.00	37	NM	33.30	68	SSN	37.80	99	HN	77.80
7	GL	46.70	38	DB	68.90	69	SSB	37.80	100	XN	71.10
8	CC	91.10	39	QJ	97.80	70	ST	66.70	101	MF	62.20
9	CCC	80.00	40	PTT	86.70	71	SP	44.40	102	BB	66.70
10	JJ	91.10	41	PL	60.00	72	SPH	55.60	103	LB	37.80
11	JB	71.10	42	DV	91.10	73	STH	66.70	104	LM	33.30
12	BJ	68.90	43	BL	57.80	74	SKH	68.90	105	BHM	80.00
13	GN	44.40	44	BD	68.90	75	NGG	77.80	106	ML	44.40
14	TN	60.00	45	TB	75.60	76	NGC	84.40	107	SL	62.20
15	JJB	80.00	46	MM	51.10	77	FP	60.00	108	PS	42.20
16	LG	42.20	47	MV	71.10	78	NGN	84.40	109	KHR	88.90
17	TT	60.00	48	MP	55.60	79	XM	77.80	110	GR	77.80
18	GDH	73.30	49	MN	37.80	80	NGJ	82.20	111	GHR	66.70
19	GM	68.90	50	NTR	75.60	81	MNTH	80.00	112	JR	75.60
20	GHN	66.70	51	MB	51.10	82	NGK	73.30	113	TRR	62.20
21	MDD	82.20	52	LK	62.20	83	KR	91.10	114	DRR	84.40
22	NT	71.10	53	MND	88.90	84	TRU	86.70	115	DHRR	68.90
23	NN	66.70	54	FK	55.60	85	BHR	86.70	116	PRR	64.40
24	NMM	60.00	55	LD	37.80	86	THB	84.40	117	BRR	75.60
25	TTT	86.70	56	LL	35.60	87	DG	68.90	118	MRR	53.30
26	TTB	80.00	57	LP	51.10	88	DGH	82.20	119	TSR	93.30
27	TM	71.10	58	LT	55.60	89	DD	82.20	120	DSR	37.80
28	TR	91.10	59	SN	55.60	90	DDH	66.70	121	HRR	86.70
29	NTT	60.00	60	SC	93.30	91	HR	77.80			
30	RRG	77.80	61	SM	86.70	92	GGU	91.10			
31	NDD	51.10	62	SB	66.70	93	GGN	80.00			

Table 4.5(f) Individual recognition rates of online handwritten Assamese conjunct consonants (*Juktakkhors*) for feature set FS:2 using RBF Kernel of SVM

Individual recognition rates of *Juktakkhors* (%)											
Feature set FS:2											
Types of SVM Kernel: RBF Kernel with C=2 & gamma=0.04											
Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%
1	KK	66.70	32	NTH	60.00	63	FN	77.80	94	NKH	77.80
2	KT	82.20	33	NDH	62.20	64	FT	51.10	95	NGH	77.80
3	KTT	60.00	34	NNN	42.20	65	SK	31.10	96	NGKH	84.40
4	KS	57.80	35	NB	46.70	66	SSTH	64.40	97	TTH	88.90
5	KL	48.90	36	NS	28.90	67	SSM	35.60	98	PN	51.10
6	KM	55.60	37	NM	37.80	68	SSN	46.70	99	HN	73.30
7	GL	48.90	38	DB	71.10	69	SSB	40.00	100	XN	71.10
8	CC	86.70	39	QJ	91.10	70	ST	60.00	101	MF	57.80
9	CCC	84.40	40	PTT	84.40	71	SP	46.70	102	BB	64.40
10	JJ	91.10	41	PL	55.60	72	SPH	48.90	103	LB	37.80
11	JB	80.00	42	DV	80.00	73	STH	66.70	104	LM	28.90
12	BJ	60.00	43	BL	55.60	74	SKH	64.40	105	BHM	82.20
13	GN	48.90	44	BD	66.70	75	NGG	80.00	106	ML	33.30
14	TN	51.10	45	TB	66.70	76	NGC	84.40	107	SL	57.80
15	JJB	75.60	46	MM	37.80	77	FP	62.20	108	PS	46.70
16	LG	35.60	47	MV	75.60	78	NGN	84.40	109	KHR	80.00
17	TT	60.00	48	MP	64.40	79	XM	71.10	110	GR	80.00
18	GDH	75.60	49	MN	31.10	80	NGJ	84.40	111	GHR	53.30
19	GM	62.20	50	NTR	75.60	81	MNTH	73.30	112	JR	82.20
20	GHN	55.60	51	MB	51.10	82	NGK	71.10	113	TRR	64.40
21	MDD	73.30	52	LK	46.70	83	KR	82.20	114	DRR	77.80
22	NT	68.90	53	MND	93.30	84	TRU	84.40	115	DHRR	73.30
23	NN	55.60	54	FK	62.20	85	BHR	77.80	116	PRR	64.40
24	NMM	57.80	55	LD	44.40	86	THB	80.00	117	BRR	75.60
25	TTT	80.00	56	LL	26.70	87	DG	68.90	118	MRR	42.20
26	TTB	73.30	57	LP	44.40	88	DGH	73.30	119	TSR	86.70
27	TM	68.90	58	LT	35.60	89	DD	77.80	120	DSR	31.10
28	TR	88.90	59	SN	44.40	90	DDH	62.20	121	HRR	82.20
29	NTT	53.30	60	SC	86.70	91	HR	77.80			
30	RRG	66.70	61	SM	86.70	92	GGU	93.30			
31	NDD	57.80	62	SB	62.20	93	GGN	68.90			

Table 4.5(g) Individual recognition rates of online handwritten Assamese conjunct consonants (*Juktakkhors*) for feature set FS:3 using Linear Kernel of SVM

Individual recognition rates of *Juktakkhors* (%)											
Feature set FS:3											
Types of SVM Kernel: Linear Kernel											
Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%
1	KK	68.90	32	NTH	66.70	63	FN	93.30	94	NKH	88.90
2	KT	88.90	33	NDH	57.80	64	FT	51.10	95	NGH	91.10
3	KTT	66.70	34	NNN	28.90	65	SK	40.00	96	NGKH	95.60
4	KS	57.80	35	NB	37.80	66	SSTH	55.60	97	TTH	86.70
5	KL	60.00	36	NS	26.70	67	SSM	40.00	98	PN	51.10
6	KM	60.00	37	NM	22.20	68	SSN	42.20	99	HN	68.90
7	GL	44.40	38	DB	66.70	69	SSB	40.00	100	XN	71.10
8	CC	91.10	39	QJ	100.00	70	ST	60.00	101	MF	60.00
9	CCC	82.20	40	PTT	86.70	71	SP	42.20	102	BB	66.70
10	JJ	88.90	41	PL	55.60	72	SPH	53.30	103	LB	28.90
11	JB	71.10	42	DV	84.40	73	STH	57.80	104	LM	24.40
12	BJ	71.10	43	BL	57.80	74	SKH	60.00	105	BHM	82.20
13	GN	42.20	44	BD	64.40	75	NGG	77.80	106	ML	37.80
14	TN	62.20	45	TB	71.10	76	NGC	88.90	107	SL	57.80
15	JJB	80.00	46	MM	48.90	77	FP	60.00	108	PS	26.70
16	LG	46.70	47	MV	71.10	78	NGN	82.20	109	KHR	84.40
17	TT	66.70	48	MP	60.00	79	XM	75.60	110	GR	80.00
18	GDH	75.60	49	MN	28.90	80	NGJ	84.40	111	GHR	62.20
19	GM	73.30	50	NTR	68.90	81	MNTH	84.40	112	JR	75.60
20	GHN	60.00	51	MB	51.10	82	NGK	66.70	113	TRR	64.40
21	MDD	84.40	52	LK	62.20	83	KR	93.30	114	DRR	77.80
22	NT	66.70	53	MND	91.10	84	TRU	82.20	115	DHRR	80.00
23	NN	60.00	54	FK	64.40	85	BHR	75.60	116	PRR	60.00
24	NMM	62.20	55	LD	37.80	86	THB	84.40	117	BRR	77.80
25	TTT	88.90	56	LL	24.40	87	DG	62.20	118	MRR	55.60
26	TTB	77.80	57	LP	40.00	88	DGH	77.80	119	TSR	93.30
27	TM	75.60	58	LT	57.80	89	DD	71.10	120	DSR	28.90
28	TR	95.60	59	SN	51.10	90	DDH	60.00	121	HRR	82.20
29	NTT	51.10	60	SC	93.30	91	HR	75.60			
30	RRG	82.20	61	SM	88.90	92	GGU	91.10			
31	NDD	46.70	62	SB	60.00	93	GGN	82.20			

Table 4.5(h) Individual recognition rates of online handwritten Assamese conjunct consonants (*Juktakkhors*) for feature set FS:3 using Polynomial Kernel of SVM

Individual recognition rates of *Juktakkhors* (%)											
Feature set FS:3											
Types of SVM Kernel: Polynomial Kernel with C=4 & E=6											
Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%
1	KK	66.70	32	NTH	68.90	63	FN	91.10	94	NKH	84.40
2	KT	88.90	33	NDH	66.70	64	FT	60.00	95	NGH	93.30
3	KTT	68.90	34	NNN	37.80	65	SK	48.90	96	NGKH	93.30
4	KS	60.00	35	NB	46.70	66	SSTH	62.20	97	TTH	91.10
5	KL	71.10	36	NS	31.10	67	SSM	48.90	98	PN	60.00
6	KM	57.80	37	NM	31.10	68	SSN	42.20	99	HN	80.00
7	GL	46.70	38	DB	73.30	69	SSB	44.40	100	XN	80.00
8	CC	91.10	39	QJ	97.80	70	ST	73.30	101	MF	64.40
9	CCC	88.90	40	PTT	88.90	71	SP	60.00	102	BB	68.90
10	JJ	93.30	41	PL	62.20	72	SPH	60.00	103	LB	40.00
11	JB	77.80	42	DV	91.10	73	STH	68.90	104	LM	35.60
12	BJ	73.30	43	BL	57.80	74	SKH	73.30	105	BHM	82.20
13	GN	48.90	44	BD	66.70	75	NGG	77.80	106	ML	40.00
14	TN	55.60	45	TB	77.80	76	NGC	88.90	107	SL	57.80
15	JJB	86.70	46	MM	46.70	77	FP	64.40	108	PS	53.30
16	LG	55.60	47	MV	68.90	78	NGN	80.00	109	KHR	86.70
17	TT	66.70	48	MP	66.70	79	XM	82.20	110	GR	77.80
18	GDH	77.80	49	MN	44.40	80	NGJ	84.40	111	GHR	64.40
19	GM	66.70	50	NTR	84.40	81	MNTH	77.80	112	JR	75.60
20	GHN	68.90	51	MB	57.80	82	NGK	75.60	113	TRR	71.10
21	MDD	80.00	52	LK	62.20	83	KR	86.70	114	DRR	84.40
22	NT	68.90	53	MND	88.90	84	TRU	88.90	115	DHRR	66.70
23	NN	64.40	54	FK	60.00	85	BHR	82.20	116	PRR	66.70
24	NMM	60.00	55	LD	44.40	86	THB	80.00	117	BRR	73.30
25	TTT	88.90	56	LL	37.80	87	DG	68.90	118	MRR	53.30
26	TTB	82.20	57	LP	55.60	88	DGH	84.40	119	TSR	93.30
27	TM	71.10	58	LT	62.20	89	DD	86.70	120	DSR	35.60
28	TR	93.30	59	SN	53.30	90	DDH	66.70	121	HRR	84.40
29	NTT	60.00	60	SC	93.30	91	HR	86.70			
30	RRG	80.00	61	SM	86.70	92	GGU	91.10			
31	NDD	60.00	62	SB	66.70	93	GGN	82.20			

Table 4.5(i) Individual recognition rates of online handwritten Assamese conjunct consonants (*Juktakkhors*) for feature set FS:3 using RBF Kernel of SVM

Individual recognition rates of *Juktakkhors* (%)											
Feature set FS:3											
Types of SVM Kernel: RBF Kernel with C=3 & gamma=0.02											
Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%	Sl. No	Class	%
1	KK	66.70	32	NTH	66.70	63	FN	86.70	94	NKH	77.80
2	KT	88.90	33	NDH	66.70	64	FT	44.40	95	NGH	86.70
3	KTT	64.40	34	NNN	35.60	65	SK	44.40	96	NGKH	91.10
4	KS	64.40	35	NB	48.90	66	SSTH	60.00	97	TTH	91.10
5	KL	66.70	36	NS	26.70	67	SSM	48.90	98	PN	55.60
6	KM	60.00	37	NM	31.10	68	SSN	51.10	99	HN	75.60
7	GL	42.20	38	DB	75.60	69	SSB	46.70	100	XN	75.60
8	CC	91.10	39	QJ	97.80	70	ST	66.70	101	MF	57.80
9	CCC	88.90	40	PTT	86.70	71	SP	64.40	102	BB	66.70
10	JJ	91.10	41	PL	60.00	72	SPH	55.60	103	LB	42.20
11	JB	75.60	42	DV	80.00	73	STH	64.40	104	LM	37.80
12	BJ	73.30	43	BL	60.00	74	SKH	73.30	105	BHM	80.00
13	GN	53.30	44	BD	71.10	75	NGG	82.20	106	ML	48.90
14	TN	57.80	45	TB	73.30	76	NGC	84.40	107	SL	62.20
15	JJB	86.70	46	MM	51.10	77	FP	66.70	108	PS	46.70
16	LG	55.60	47	MV	75.60	78	NGN	80.00	109	KHR	84.40
17	TT	64.40	48	MP	57.80	79	XM	80.00	110	GR	77.80
18	GDH	75.60	49	MN	35.60	80	NGJ	84.40	111	GHR	64.40
19	GM	66.70	50	NTR	80.00	81	MNTH	75.60	112	JR	82.20
20	GHN	62.20	51	MB	60.00	82	NGK	73.30	113	TRR	66.70
21	MDD	86.70	52	LK	64.40	83	KR	86.70	114	DRR	80.00
22	NT	71.10	53	MND	91.10	84	TRU	93.30	115	DHRR	66.70
23	NN	64.40	54	FK	57.80	85	BHR	80.00	116	PRR	64.40
24	NMM	64.40	55	LD	44.40	86	THB	82.20	117	BRR	77.80
25	TTT	82.20	56	LL	31.10	87	DG	75.60	118	MRR	48.90
26	TTB	84.40	57	LP	48.90	88	DGH	75.60	119	TSR	88.90
27	TM	68.90	58	LT	51.10	89	DD	82.20	120	DSR	28.90
28	TR	95.60	59	SN	53.30	90	DDH	64.40	121	HRR	84.40
29	NTT	62.20	60	SC	88.90	91	HR	80.00			
30	RRG	75.60	61	SM	91.10	92	GGU	95.60			
31	NDD	60.00	62	SB	64.40	93	GGN	82.20			

Table 4.6 Average recognition rates of *Juktakkhors*

Online handwritten Assamese Conjunct Characters(Juktakkhors)					
Type of SVM kernel	Total Number of Instances	Correctly Classified Instances	Average recognition rate (Across Classes)	Standard Deviation (SD) of recognition rate (Across Classes)	Average ± SD
Feature Set FS:1					
Linear	5445	3014	55.35	20.37	55.35 ± 20.37
Polynomial (C=3, E=4)	5445	3212	58.99	19.29	58.99 ± 19.29
RBF (C=1,Gamma=0.07)	5445	2980	54.73	21.63	54.73 ± 21.63
Feature Set FS:2					
Linear	5445	3466	63.65	19.41	63.65 ± 19.41
Polynomial (C=3, E=4)	5445	3671	67.42	17.17	67.42 ± 17.17
RBF (C=2,Gamma=0.04)	5445	3495	64.19	17.04	64.19 ± 17.04
Feature Set FS:3					
Linear	5445	3576	65.68	18.96	65.68 ± 18.96
Polynomial (C=4, E=6)	5445	3791	69.62	16.30	69.62 ± 16.30
RBF (C=8,Gamma=0.03125)	5445	3726	68.43	16.41	68.43 ± 16.41

Linear Kernel

The average recognition rates achieved for the online handwritten Assamese conjunct consonants (*Juktakkhors*) using linear kernel were 55.35% (based on FS:1), 63.65% (based on FS:2) and 65.68% (based on FS:3) with a 10 fold cross validation process. A total of 5445 characters were used as samples in the *Juktakkhors* recognition experiment (refer to Table 4.6). The kernel parameter settings C = 1 and E = 1 are default associated with the linear kernel.

Polynomial Kernel

The polynomial kernel parameter settings (C = 3 and E = 4), (C = 3 and E = 4) and (C = 4 and E = 6) for the feature sets FS:1, FS:2 and FS:3 respectively were

obtained by grid search operation. The average recognition rates achieved for the online handwritten Assamese *Juktakkhors* using polynomial kernel were 58.99% (based on FS:1), 67.42% (based on FS:2), and 69.62% (based on FS:3) with a 10 fold cross validation process. A total of 5445 characters were used as samples in the *Juktakkhors* recognition experiment (refer to Table 4.6).

RBF Kernel

- *Parameter Settings: Grid Search*

The kernel parameter settings (C=1,Gamma=0.07) for the feature set FS:1, (C=2,Gamma=0.04) for the feature set FS:2 and (C=8,Gamma=0.03125) for the feature set FS:3 were obtained by grid search operation. The plots of grid search are shown in the Figure 4.11 through Figure 4.13.

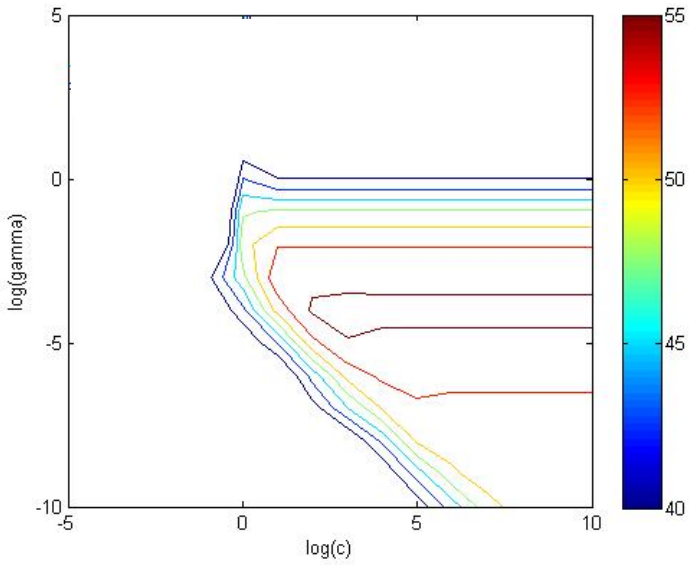

Figure 4.11 Grid search plot of C and gamma for RBF kernel (using FS:1 for Juktakkhors)

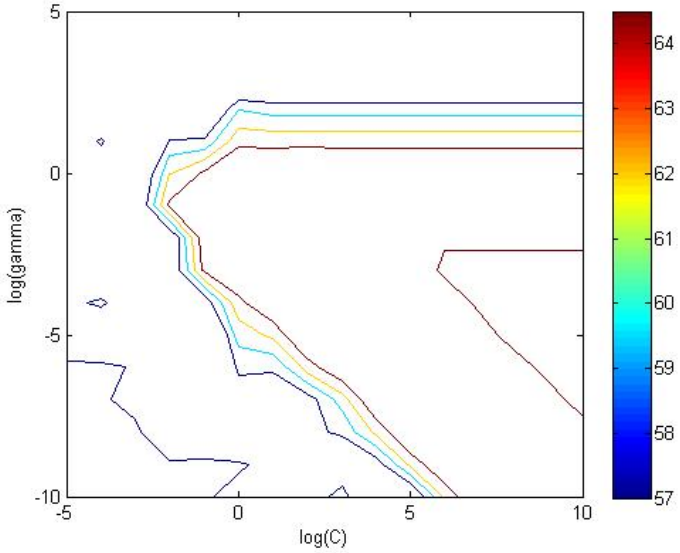

Figure 4.12 Grid search plot of C and gamma for RBF kernel (using FS:2 for Juktakkhors)

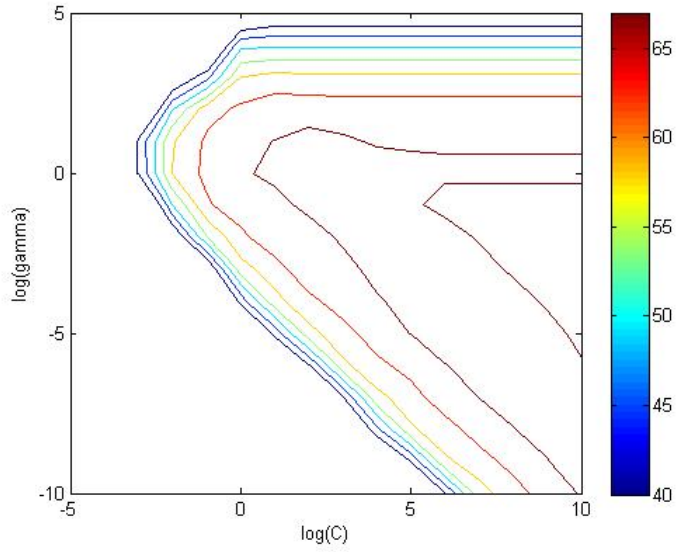

Figure 4.13 Grid search plot of C and gamma for RBF kernel (using FS:3 for Juktakkhors)

- *Experimental Results*

The average recognition rates achieved for the online handwritten Assamese Juktakkhors using RBF kernel were 54.73% (based on FS:1), 64.19% (based on

FS:2), and 68.43% (based on FS:3) with a 10 fold cross validation process. A total of 5445 characters were used as samples in the *Juktakkhors* recognition experiment (refer to Table 4.6).

4.3.2.4 Overall recognition Rates

In the previous sections we presented the recognition rates of online handwritten Assamese numerals, basic alphabetic characters and *Juktakkhors* for three SVM kernels based on feature vectors FS:1, FS:2 and FS:3. Here we present the overall recognition rates of all the 183 characters (a total of 8235 samples which include numerals, basic alphabetic characters and *Juktakkhors*) based on the best performing feature vector FS:3 using linear, polynomial and RBF kernels of SVM. Individual recognition rates of all the 183 characters are presented in Table 4. 7, average recognition rates of all the 183 characters are presented in Table 4.8 and parameters settings for kernels including plots of Grid search for RBF kernel are presented after the tables.

Table 4.7 Individual recognition rates of all the 183 online handwritten Assamese characters

Sl. No.	Class	Recognition rates (%)		
		Feature Set FS:3		
		Types of SVM Kernel		
		Linear Kernel	Polynomial Kernel (C=2 & E=3)	RBF Kernel (C=1 & gamma=0.01)
1	SUNYA	95.60	95.60	82.20
2	EK	97.80	100.00	91.10
3	DUI	93.30	95.60	93.30
4	TINI	93.30	95.60	88.90
5	CARI	86.70	86.70	68.90
6	PAC	97.80	97.80	88.90
7	CAY	100.00	100.00	95.60
8	SATH	97.80	95.60	93.30
9	ATH	97.80	97.80	97.80
10	NAA	93.30	95.60	88.90
11	A	68.90	80.00	64.40

Sl. No.	Class	Recognition rates (%)		
		Feature Set FS:3		
		Types of SVM Kernel		
		Linear Kernel	Polynomial Kernel (C=2 & E=3)	RBF Kernel (C=1 & gamma=0.01)
12	AA	75.60	80.00	68.90
13	E	75.60	77.80	62.20
14	EE	62.20	68.90	53.30
15	U	73.30	77.80	60.00
16	UU	82.20	77.80	68.90
17	REE	73.30	80.00	73.30
18	AE	95.60	97.80	97.80
19	OI	93.30	93.30	88.90
20	O	91.10	91.10	86.70
21	OU	88.90	91.10	86.70
22	KA	75.60	80.00	68.90
23	KHA	82.20	84.40	68.90
24	GA	73.30	73.30	66.70
25	GHA	60.00	64.40	46.70
26	NG	95.60	95.60	93.30
27	CA	86.70	88.90	77.80
28	CCA	77.80	82.20	73.30
29	JA	77.80	82.20	80.00
30	JHA	88.90	88.90	84.40
31	NIYA	100.00	100.00	100.00
32	MTA	80.00	82.20	64.40
33	MTHA	80.00	77.80	55.60
34	MDA	71.10	71.10	64.40
35	MDHA	82.20	84.40	60.00
36	MNA	71.10	80.00	57.80
37	TA	75.60	84.40	68.90
38	THA	91.10	95.60	82.20
39	DA	77.80	93.30	84.40
40	DHA	66.70	71.10	60.00
41	NA	64.40	68.90	55.60
42	PA	77.80	80.00	71.10
43	PHA	62.20	66.70	44.40
44	BA	62.20	66.70	57.80
45	BHA	77.80	80.00	57.80
46	MA	62.20	73.30	44.40

Sl. No.	Class	Recognition rates (%)		
		Feature Set FS:3		
		Types of SVM Kernel		
		Linear Kernel	Polynomial Kernel (C=2 & E=3)	RBF Kernel (C=1 & gamma=0.01)
47	AJA	60.00	62.20	55.60
48	RA	62.20	64.40	46.70
49	LA	48.90	60.00	24.40
50	WA	77.80	77.80	66.70
51	TXA	86.70	88.90	80.00
52	MXA	53.30	60.00	57.80
53	DXA	37.80	42.20	26.70
54	HA	84.40	93.30	66.70
55	KHYA	77.80	95.60	60.00
56	AYA	55.60	62.20	55.60
57	DRA	71.10	82.20	66.70
58	DHRA	71.10	77.80	53.30
59	KTA	91.10	91.10	91.10
60	ANSR	80.00	86.70	75.60
61	BXG	95.60	95.60	75.60
62	CBN	82.20	84.40	82.20
63	KK	66.70	68.90	68.90
64	KT	86.70	86.70	66.70
65	KTT	66.70	66.70	66.70
66	KS	62.20	62.20	55.60
67	KL	64.40	75.60	51.10
68	KM	55.60	60.00	53.30
69	GL	46.70	44.40	31.10
70	CC	86.70	91.10	86.70
71	CCC	77.80	88.90	77.80
72	JJ	91.10	93.30	82.20
73	JB	68.90	75.60	66.70
74	BJ	62.20	64.40	48.90
75	GN	42.20	46.70	48.90
76	TN	64.40	62.20	42.20
77	JJB	82.20	86.70	77.80
78	LG	40.00	51.10	44.40
79	TT	62.20	66.70	51.10
80	GDH	66.70	77.80	66.70
81	GM	64.40	68.90	66.70

Sl. No.	Class	Recognition rates (%)		
		Feature Set FS:3		
		Types of SVM Kernel		
		Linear Kernel	Polynomial Kernel (C=2 & E=3)	RBF Kernel (C=1 & gamma=0.01)
82	GHN	64.40	66.70	55.60
83	MDD	82.20	84.40	66.70
84	NT	62.20	68.90	71.10
85	NN	51.10	62.20	57.80
86	NMM	66.70	62.20	62.20
87	TTT	73.30	82.20	64.40
88	TTB	77.80	86.70	71.10
89	TM	60.00	68.90	55.60
90	TR	82.20	93.30	71.10
91	NTT	51.10	51.10	42.20
92	RRG	77.80	84.40	71.10
93	NDD	46.70	55.60	24.40
94	NTH	62.20	64.40	51.10
95	NDH	62.20	64.40	64.40
96	NNN	31.10	42.20	20.00
97	NB	35.60	51.10	31.10
98	NS	24.40	31.10	31.10
99	NM	17.80	24.40	22.20
100	DB	60.00	73.30	57.80
101	QJ	95.60	97.80	93.30
102	PTT	84.40	88.90	80.00
103	PL	46.70	55.60	53.30
104	DV	86.70	91.10	82.20
105	BL	57.80	57.80	55.60
106	BD	64.40	68.90	62.20
107	TB	64.40	68.90	55.60
108	MM	48.90	48.90	44.40
109	MV	71.10	64.40	62.20
110	MP	51.10	60.00	57.80
111	MN	35.60	44.40	26.70
112	NTR	75.60	77.80	75.60
113	MB	51.10	57.80	57.80
114	LK	64.40	68.90	40.00
115	MND	91.10	91.10	88.90
116	FK	48.90	62.20	42.20

Sl. No.	Class	Recognition rates (%)		
		Feature Set FS:3		
		Types of SVM Kernel		
		Linear Kernel	Polynomial Kernel (C=2 & E=3)	RBF Kernel (C=1 & gamma=0.01)
117	LD	31.10	46.70	20.00
118	LL	24.40	40.00	26.70
119	LP	37.80	51.10	28.90
120	LT	48.90	60.00	20.00
121	SN	44.40	46.70	44.40
122	SC	91.10	91.10	82.20
123	SM	84.40	86.70	84.40
124	SB	57.80	64.40	55.60
125	FN	91.10	88.90	77.80
126	FT	48.90	55.60	31.10
127	SK	40.00	42.20	35.60
128	SSTH	51.10	60.00	40.00
129	SSM	46.70	51.10	37.80
130	SSN	35.60	44.40	48.90
131	SSB	42.20	48.90	42.20
132	ST	57.80	68.90	44.40
133	SP	46.70	46.70	44.40
134	SPH	44.40	46.70	37.80
135	STH	60.00	66.70	57.80
136	SKH	71.10	68.90	64.40
137	NGG	71.10	80.00	71.10
138	NGC	93.30	91.10	86.70
139	FP	51.10	55.60	53.30
140	NGN	75.60	82.20	75.60
141	XM	75.60	77.80	73.30
142	NGJ	86.70	84.40	80.00
143	MNTH	71.10	75.60	57.80
144	NGK	66.70	66.70	60.00
145	KR	88.90	91.10	82.20
146	TRU	86.70	86.70	77.80
147	BHR	77.80	82.20	60.00
148	THB	82.20	84.40	80.00
149	DG	55.60	64.40	66.70
150	DGH	73.30	80.00	64.40
151	DD	62.20	77.80	68.90

Sl. No.	Class	Recognition rates (%)		
		Feature Set FS:3		
		Types of SVM Kernel		
		Linear Kernel	Polynomial Kernel (C=2 & E=3)	RBF Kernel (C=1 & gamma=0.01)
152	DDH	55.60	62.20	60.00
153	HR	73.30	86.70	64.40
154	GGU	93.30	91.10	88.90
155	GGN	84.40	80.00	71.10
156	NKH	84.40	86.70	75.60
157	NGH	93.30	93.30	80.00
158	NGKH	97.80	95.60	95.60
159	TTH	86.70	88.90	86.70
160	PN	44.40	57.80	42.20
161	HN	75.60	80.00	64.40
162	XN	71.10	77.80	68.90
163	MF	62.20	62.20	57.80
164	BB	64.40	71.10	57.80
165	LB	31.10	35.60	33.30
166	LM	26.70	28.90	15.60
167	BHM	73.30	80.00	71.10
168	ML	28.90	35.60	22.20
169	SL	57.80	62.20	53.30
170	PS	40.00	51.10	37.80
171	KHR	86.70	86.70	84.40
172	GR	75.60	77.80	68.90
173	GHR	62.20	64.40	55.60
174	JR	80.00	82.20	64.40
175	TRR	64.40	64.40	68.90
176	DRR	75.60	80.00	64.40
177	DHRR	80.00	75.60	62.20
178	PRR	66.70	68.90	64.40
179	BRR	66.70	73.30	64.40
180	MRR	46.70	51.10	22.20
181	TSR	93.30	93.30	91.10
182	DSR	33.30	35.60	28.90
183	HRR	88.90	86.70	75.60

Table 4.8 Average recognition rates of all the 183 online handwritten Assamese characters based on feature set FS:3

Online handwritten Assamese characters (numerals, basic alphabetic characters & Juktakkhors)					
Type of SVM kernel	Total Number of Instances	Correctly Classified Instances	Average recognition rate (Across Classes)	Standard Deviation (SD) of recognition rate (Across Classes)	Average ± SD
Feature Set FS:3					
Linear	8235	5675	68.91	18.67	68.91 ± 18.67
Polynomial (C=2, E=3)	8235	6029	73.21	16.96	73.21 ± 16.96
RBF (C=1,Gamma=0.01)	8235	5148	62.51	19.15	62.51 ± 19.15

Linear Kernel

The overall recognition rate achieved for the online handwritten Assamese combined characters using linear kernel is 68.91% (based on FS:3), with a 10 fold cross validation process. All the 8235 characters of the dataset were used as samples in the character recognition experiment (refer to Table 4.8). The kernel parameter settings C = 1 and E = 1 are default associated with the linear kernel.

Polynomial Kernel

The polynomial kernel parameter setting C = 2 and E = 3 for the feature set FS:3 was obtained by grid search operation. The overall recognition rate achieved for the online handwritten Assamese combined characters using polynomial kernel is 73.21% (based on FS:3) with a 10 fold cross validation process. All the 8235 characters of the dataset were used as samples in the character recognition experiment (refer to Table 4.8).

RBF Kernel

- *Parameter Settings: Grid Search*

The kernel parameter settings C=1 and gamma=0.01 for the feature set FS:3 was obtained by grid search operation. The plot of grid search operation is shown in the Figure 4.14.

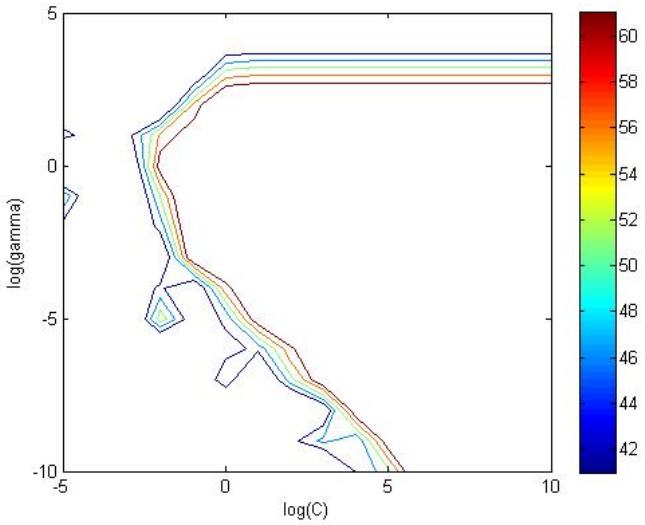

Figure 4.14 Grid search plot of C and gamma for RBF kernel (using FS:3 for the combined set of characters)

- *Experimental Results*

The overall recognition rate achieved for the online handwritten Assamese combined characters using RBF kernel was 62.51% (based on FS:3) with a 10 fold cross validation process. All the 8235 characters of the dataset were used as samples in the character recognition experiment (refer to Table 4.8).

4.4 Conclusion

For recognition of online handwritten Assamese characters the effectiveness of standard features reported in literature (for other scripts) is explored. We have performed character recognition experiments based on standard features. We have

explored feature sets consisting of combinations of geometrical, structural and statistical features for the classification of online handwritten Assamese characters. It is observed that there are scopes to improve the recognition rates of online handwritten Assamese basic alphabetic characters and *Juktakkhors*. However, improvement in the recognition rates has been noticed from feature set FS:1 through feature set FS:3 in case of basic alphabetic characters and *Juktakkhors*. Furthermore, in the context of classification, three different kernels of SVM are tried. We have started our experiment using linear kernel. Linear kernel gave us baseline results. Nonlinear kernels, namely polynomial kernel and RBF kernel were used subsequently to improve the recognition results. In majority of the experiments, nonlinear kernels gave us improved recognition results over linear kernel. Overall recognition of 73.21% is observed for FS:3. *Juktakkhors* and certain complex shaped basic alphabetic characters exhibited very low recognition rates. In order to distinguish *Juktakkhors* and basic alphabetic characters involving complex shaped strokes, the next two Chapters viz. Chapter 5 and Chapter 6 explore qualitative representation and qualitative feature sets.