International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-5, January 2020

Assamese Handwritten Character Recognition using Supervised Fuzzy Logic

Kalyanbrat Medhi, Sanjib Kr. Kalita

Abstract*: This paper presents a state of the art supervised fuzzy pattern recognition system for recognition of Assamese handwritten characters. The fuzzy classifier is well suited for applications with ambiguities and handwritten character recognition is such a task. The dataset used in this experiment is taken from ISI Kolkata. After preprocessing images are normalized into uniform size 42x32 and then two features namely distance vector and density vector have been extracted. The experiment has two stages, training and testing. In first stage we extract distance vector and density features from uniform zones of the binary images for training classes and estimate the mean and variance for each class. In second stage we use this mean and variance to calculate the membership values for each unknown character of the testing set of data. An exponential fuzzy membership function is used for this purpose. Finally we recognize an unknown test character as that class for which it gives highest membership value. Finally result is stored in editable document. The highest recognition accuracy achieved in the experiment is 88.29%, 86.55% and 82.74% for numerals, vowels and consonants respectively.

Keywords: Assamese handwritten character, fuzzy logic, fuzzy membership function, pattern recognition, distance vector, density vector, uniform zoning.

I. INTRODUCTION

Handwriting character recognition (HCR) is still a challenging task in the field of pattern recognition. A lot of researches are going on since last four decades but none of them achieved 100% accuracy. The challenges occur due to immense variation in writing. The variation is not only among the writers but also may occur in case of one writer due to different mood of writer or different writing time. It is a crucial task to find out robust features that can uniquely identify a character even after much variation.

Handwriting is till now a primitive method for documenting our thoughts for communicating with others. In this digital age digital document are more convenient for the purpose of storage, communication and retrieval of information. Hence converting handwriting documents into digital document become a trend where people spend lot of time and money in the name of data entry operation. This can be done more conveniently with efficient handwritten character recognition (HCR) system. The outcome of any HCR is a digital document which is editable.

The HCR processes have been classified into two categories. The first one is where recognition performed at

the time of writing; called online recognition. The second is where user has to write on some kind of personal digital assistant (PDA). The HCR system enables these devices to convert the user input into digital document. In another process recognition is performed after writing on paper. The paper needs to be digitized as image by scanner or camera with a good resolution. The image is then put as input to the HCR system and it produces editable digital document. This type of recognition is called off line recognition.

The HCR has very useful application in post office for pin code sorting and in bank for check processing where check number, account holder name and legal amount can be automatically inserted into the database. Now a day form processing is another emerging application of HCR.

Literature review reveals that a few researches on Assamese language have done. Assamese is one of the major northeastern languages of India. It is primarily spoken in the state of Assam. In this work we have focused on features extraction and classifications based on fuzzy with different combination of classifiers and features.

II. LITERATURE REVIEW

Ritesh Sarkhel et al. designed a cost effective isolated handwritten Bangla character and digit recognition model [1]. Their proposed method use a multi-objective region sampling algorithms Non-Dominated Sorting Harmony-Search Algorithm (NSHA) and Non-Dominated Sorting Genetic Algorithm-II (NSGA-II) for region sampling separately together with the Axiomatic Fuzzy Set (AFS) theory. In their experiment the convex haul based global features and CG based quad-tree partitioned longest-run based local features are used with SVM classifier. The proposed method achieved 98.23% recognition accuracy and 12.61% reduction in recognition cost for ISI Kolkata dataset.

Vilém Novák et al. presented a few models to recognize characters printed both on hot as well as on cold metal ingots[2]. They used neural networks, mathematical fuzzy logic and representation of image using fuzzy-valued function. Among these three methods, based on image represented by the fuzzy-valued function achieved the highest character recognition rate 87.7%.

Manuscript published on January 30, 2020.
* Correspondence Author
Kalyanbrat Medhi*, Computer Science, Gauhati University, Guwahati, India. Email: kalyanbrat_medhi@gauhati.ac.in
Sanjib Kr. Kalita, Computer Science, Gauhati University, Guwahati, India. Email: sanjib959@gauhati.ac.in

Sushama Shelke and Shaila Apte proposed a Fuzzy based multi-stage classification model for Handwritten Devanagari script [3]. The first stage of classification is based on fuzzy inference system and second stage of the proposed model is based on structural parameters. Initially, the fuzzy inference systems classify into 24 classes. In the next stage, normalized pixel density features are calculated and fed into feed forward neural network. According to their claim, recognition accuracy is 96.95%.

D. Álvarez et al. present their fuzzy classification model for off-line intelligent word recognition [4]. The words are segmented into single characters followed by labeling each pixel into groups of vertical or horizontal strokes. After that the connected locations between the vertical and horizontal strokes are obtained by using dynamic zoning. They have represented a character in a string of regular grammar with these features and check using a Deterministic Finite Automaton. Finally for recognition purpose they have use Fuzzy Lattice Reasoning classifier on Proprietary IAM dataset and obtained 90.80% accuracy.

Recently Xu-Yao Zhang et al. developed a deep convolution neural network based online and offline handwritten Chinese character recognition model with shape normalization and directional decomposition[9]. They have proposed a new adaption layer to reduce the training and testing mismatch. According to their claim it can also adopt the new writing style of a writer and hence recognition accuracy is improved. The training offline dataset is HWDB1.0-1.2 and online dataset is OLHWDB1.0-1.2. The ICDAR-2013 offline and online competition dataset are considered for testing. They have claimed that without the adaption layer recognition accuracy is 97.55% and when the adaption layer is used recognition accuracy is 97.91%.

A. K. Sampath and N. Gomathi designed a multi kernel support vector machine based on fuzzy triangular membership function[10]. After preprocessing feature extraction is done using histogram oriented gradient descriptor. Then classification is done for chars74K database through the proposed system. According to them the recognition accuracy of their system is 99.4%

Recently K. Nongmeikapam et al. proposed a SVM based model for recognizing Manipuri script[14]. They have extracted histogram oriented gradient based feature. A liner multiclass support vector machine has been used for classification. They have claimed that the classified yielded accuracy of 96.928%

Zhen Xi and George Panoutsos designed a convolutional neural network (CNN) with a fuzzy logic (FL) based rules in one layer and a Radial Basis Function (RBF) layer to maintain the rules of the network[17]. A normalized exponential function is used for defuzzification of the FL. They used MNIST dataset that is consist of 60000 training and 10000 testing samples. It is claimed that highest testing accuracy obtained is around 99% with 64 features.

III. DATASET USED

The dataset used in this experiment is standard dataset taken from ISI Kolkata. This dataset consists of 23392 samples of numerals written by 1106 persons. There are 10 classes of the numerals. These samples had been collected from 465 mail pieces and 268 job application forms and for the rest by using specially designed form. There are 37,858 samples of basic characters in this dataset with 11 classes for vowels and 39 classes for consonants. Some samples from ISI Kolkata dataset are depicted in Fig. 1.

Fig. 1. Sample of isolated Bangla numerals and characters

IV. PREPROCESSING

It is an important part of character recognition. The original grayscale images are converted to binary images using Otsu's method of global thresholding. Fig. 2 shows binary image representation of Assamese digit *ak* (one) and *sunya* (zero)

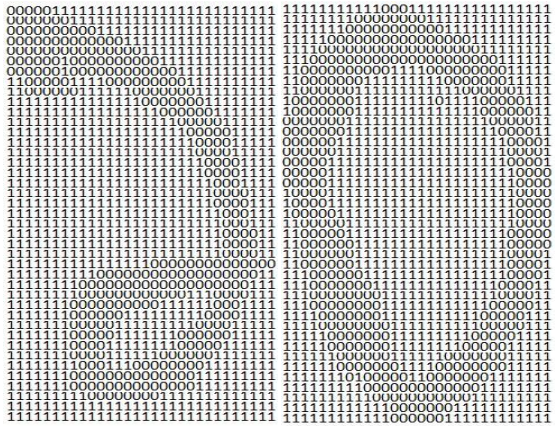


Fig. 2. Binary image of Assamese digit *ak* (one) and *sunya* (zero)

The original data contains a lot of background noises. The noises need to be removed before classification.

In order to remove noises of the images morphological opening and closing operation are performed. The binary images are normalized into 42x32 sizes before feature extraction. Fig. 3 shows the stages of noise removal.

Fig. 3. An original grayscale image (a) binary image (b) with noise, and image (c) after noise removal

V. FEATURE EXTRACTION

The feature extraction is one of the crucial processes in character recognition process. There exists a lot of method for feature extraction including based on projection, histograms, intersections, mathematical transformation, moment, graph, geometrical shape, contour etc. in the present study, zone based feature extraction technique is used to obtain useful local information of patterns image. An outstanding literature about zoning method can be found in an article of D. Impedovo and G.Pirlo [11]. Zoning based feature extraction has two category; static and adaptive. In this experiment we used static zoning method. In static method character image divided into many regions of MxN resulting many sub images of identical shape. The following Fig. 4 shows 7x8 uniform zoning of a binary image digit zero.

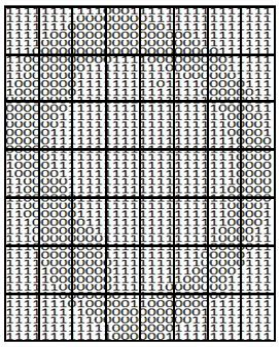


Fig. 4. Uniform zoning 7x8 for digit zero

In the present study we have considered two features for classification namely density and distance vector. In order to extract density feature we consider total number of black pixels in a box. The second feature we consider is vector distance from the bottom left corner. If bottom left corner is absolute origin (0, 0), then the vector distance for pixel in b^{th} box at location (i, j) is calculated as,

$$d_k^b = (i^2 + j^2)^{1/2} \qquad (1)$$

By dividing the sum of distances of all '1's pixels present in that box with the total number, a normalized vector distance (γ_b) for each box is computed as

$$\gamma_b = \frac{1}{n_b}\sum_{k=1}^{n_b} d_k^b \qquad b=1, 2, \ldots\ldots, N \qquad (2)$$

Where n_b is number of pixels in b^{th} box and N is total number of boxes.

The following algorithm is used for distance vector feature extraction from image size 42x32 using MxN zoning.

```
vector_length = N;
i_diff = 42/M;
j_diff = 32/N;
dkb = zeros(vector_length, total_number_of _images);
Image reading loop start
1.  for k = 1 : total_number_of_images
    filename = Image filename
    I= Read the k^th image of filename
    I = Convert the image to binary image
    I = Remove Noise from the image
    I = Normalize the image I to size 42x32
    Initialize boxno = 1;
 Feature extraction loop start
2. for i = 1 : i_diff : 41
       row = i_diff ;
       col = j_diff ;
3.     for j = 1 : j_diff : 31
4.         for x = i : 1 : row+i
5.             for y = j : 1 : col
                   if(I(x,y) == 0)
                   dkb(boxno,k)=dkb(boxno,k)+sqrt((42-x)^2+y^2)
                   end
                   if(y > 31)
                       exit from loop 5
                 else
                       continue loop 5
               if(x > 41)
                   exit from loop 4
               else
                   continue loop 4
               dkb(boxno,k) = dkb(boxno,k)/(i_diff * j_diff);
               col = col + j_diff ;
               boxno = boxno+1;
               j=j+1;
           exit from loop 3
           i=i+1
       exit from loop 2
       k=k+1
    exit from loop 1
```

The algorithm reads all images from file and provides distance vector feature of length N.

The following algorithm is used for density vector feature extraction from image size 42x32 using MxN zoning.

```
vector_length = N;
i_dif f = 42/M;
j_dif f = 32/N;
dkb = zeros(vector_length, total_number_o f _images);
Image reading loop start
1.  for k = 1 : total_number_o f_images
    filename = Image file name
    I= Read the k^th image of filename
    I = Convert the image to binary image
    I = Remove Noise from the image
    I = Normalize the image I to size 42x32
    Initialize boxno = 1;
 Feature extraction loop start
2. for i = 1 : i_diff : 41
       row = i_diff ;
```

```
        col = j_diff ;
3.      for j = 1 : j_diff : 31
          black=0 ;
4.        for x = i : 1 : row+i
5.          for y = j : 1 : col
              if(I(x,y) == 0)

              black=black+1;
              end
              if(y > 31)
                exit from loop 5
            else
                continue loop 5
          if(x > 41)
              exit from loop 4
          else
              continue loop 4
          dkb(boxno,k)=black;
          col = col + j_diff ;
          boxno = boxno+1;
          j=j+1;
        exit from loop 3
        i=i+1
      exit from loop 2
      k=k+1
    exit from loop 1
```

The algorithm reads all images from file and provides distance vector feature of length N.

VI. PROPOSED METHODOLOGY

The performance of the recognition system depends on efficiency of the feature extracted as well as the classifier. The selected features should be able to uniquely differentiate the characters. Literature review reveals that one recognition model that work efficiently for one language may not be efficient for another language. In this work we use two features namely distance vector and density and a recognition model based on fuzzy logic.

A. Fuzzy logic

Fuzzy logic is due to fuzzy set that was introduced by Zadeh[12] in 1965. Its purpose is to provide an approximation to ambiguous or complex pieces of real life problem. Application of fuzzy logic includes numerous artificial intelligence problems such as face recognition, antiskid braking system, air conditioner etc. Fuzzy based supervised pattern recognition was introduced by Bellman et al [13] in 1966. Their proposed methodology consists of two basic operations- abstraction and generalization. In fuzzy set theory abstraction refers to estimate a membership function of fuzzy class from the training samples. In the process of generalization the estimation is use to compute the membership of the unknown samples apart from training set. Chunyan Wang et al. proposed a high resolution remote sensing image classification using interval type-2 fuzzy membership function [8]. In order to analyze the feature data they have used type-1 fuzzy membership function. The total accuracy by using type-1 fuzzy model is reported to be 0.879 with 0.839 kappa. The application of fuzzy logic and membership function can be seen in various works [15, 16, 8]. We will try to implement this method for character recognition. In present study an exponential fuzzy membership function is used for classification. The following Fig. shows fuzzy membership value.

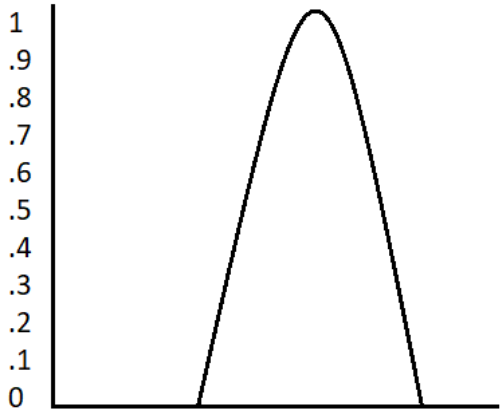
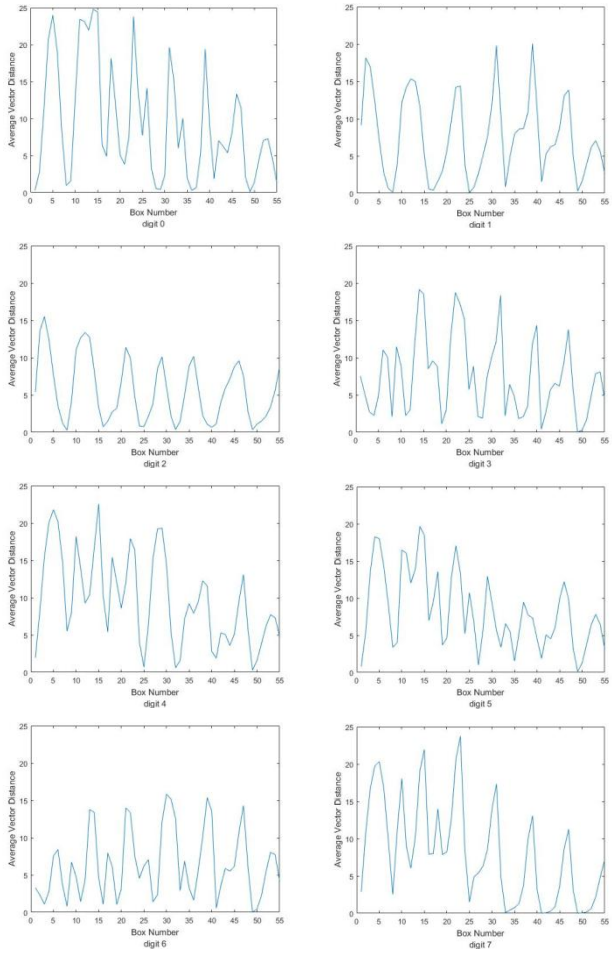


Fig. 5: Fuzzy membership value represented in x- axis

The graph of average vector distances and density vector features for each digit is depicted in Fig. 6 and Fig. 7 respectively.



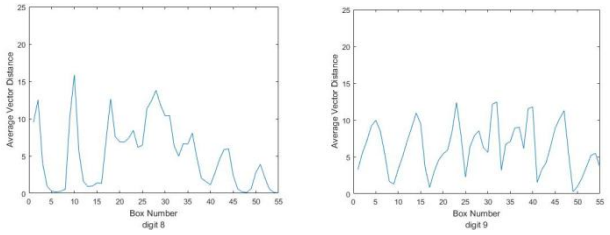


Fig. 6. Average vector distance vs. box numbers

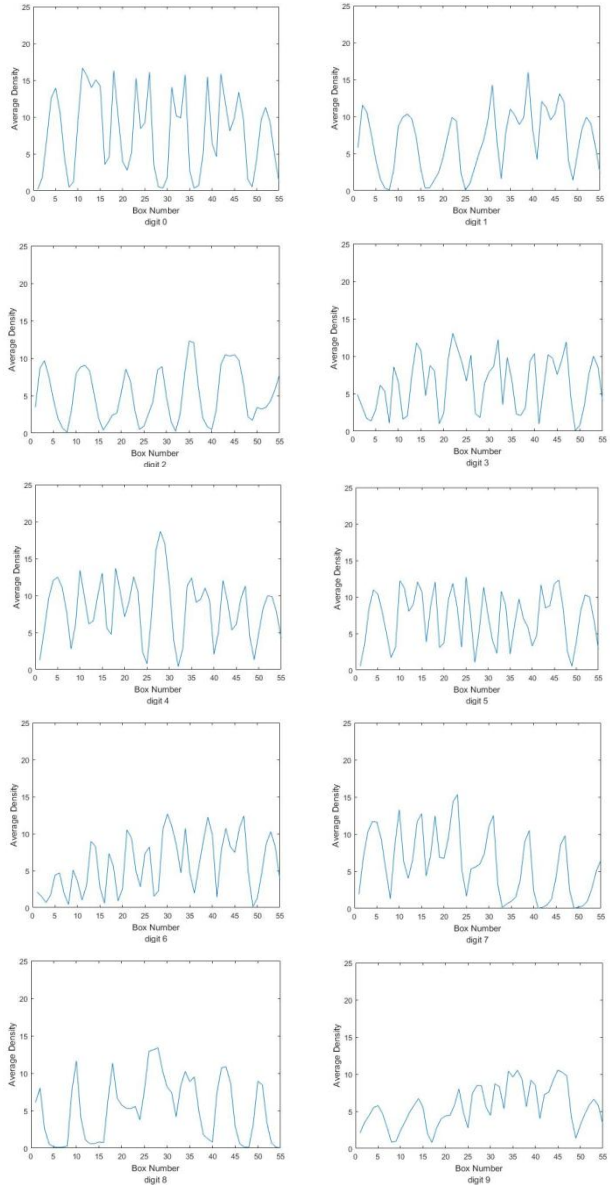


Fig. 7. Average density vector vs. box numbers

The membership values for each unknown characters have been computed against each of the digits. The unknown digit is that digit for which it has highest membership value. The mean m_{r_i} and variance $\sigma_{r_i}^2$ for each fuzzy set of characters are calculated using equation 3 and 4.

$$m_{r_i} = \frac{1}{N_i^r}\sum_{j=1}^{N_i^r} f_{ij} \quad (3)$$

$$\sigma_{r_i}^2 = \frac{1}{N_i^r}\sum_{j=1}^{N_i^r}(f_{ij} - m_{r_i})^2 \quad (4)$$

Where, N_i^r is the number of samples in the i_{th} fuzzy set and f_{ij} stands for the j_{th} sample in the i_{th} feature of r_{th} reference character. The membership values for each unknown characters have been calculated from estimated mean and variance by using the following fuzzification function:

$$\mu_{r_i}(x_i) = exp^{-\left|\frac{(m_{r_i} - x_i)}{\sigma_{r_i}^2}\right|} \quad (5)$$

Where, x_i is i_{th} feature of the unknown character x, m_{r_i} and $\sigma_{r_i}^2$ are the corresponding mean and variance of i_{th} fuzzy set of a r_{th} reference character. The identity of the unknown character x is determined from the reference character r that gives the maximum value for $\mu_r(x)$. Then the average membership value is calculated using Eq. (6).

$$\mu_r(x) = \frac{1}{N^r}\sum_{i=1}^{N^r}\mu_{r_i}(x_i) \quad (6)$$

Where N^r is the number of features.

VII. RESULTS AND DISCUSSION

The dataset used in the experiment consist of isolated grayscale image. After preprocessing the binary images are normalized into size 42x32. The distance vector and density vector features are extracted from each uniform zone of the images. This provides us four different vector lengths 24, 48, 56 and 112 respectively. A distance vector profile of length 24 for digit zero is as follows.

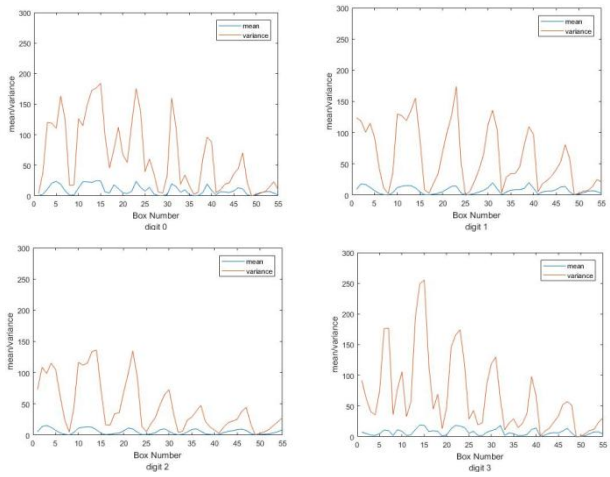
[8.4697 18.79 4.8769 0 13.345 17.918 22.937 2.6804 10.701 3.8884 0.64509 15.268 8.6156 0 0 13.017 2.7053 7.1513 3.6175 12.331 0 3.7686 8.4161 0.45911]

The density vector of length 24 for digit zero is as follows.

[13 27 7 0 24 30 34 4 25 8 1 24 28 0 0 22 11 28 9 24 0 14 23 1]

Then we calculate mean and variance from all the training samples for each vector length

We have considered 90%, 80%, 70%, 60% and 50% samples for training and correspondingly 10%, 20%, 30%, 40% and 50% samples for testing purpose. Then we calculate mean and variance from all training samples for each feature with different vector length.



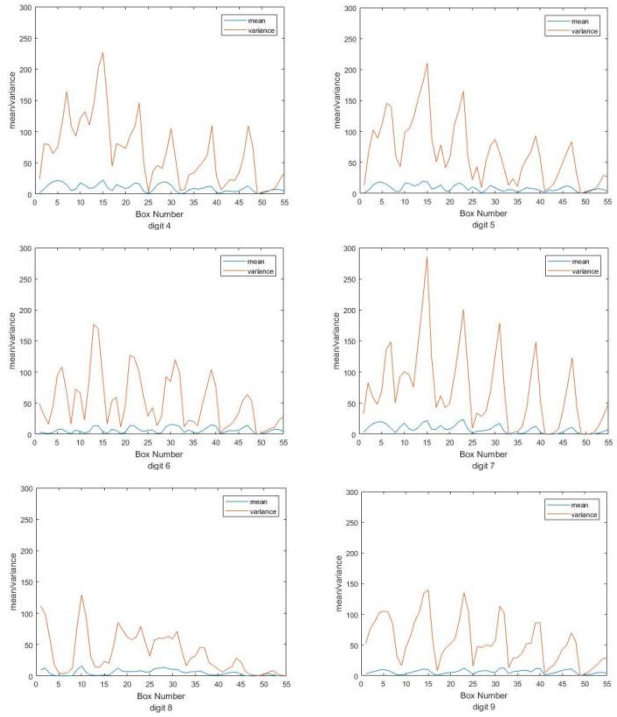


Fig. 8. Mean and variance of distance vector of length 56 when 90% training is considered

The Fig. 8 is graphical representation of mean and variance of distance vector length 56 considering 90% samples for training. This mean and variance are used to calculate membership value. In order to recognize unknown a character the fuzzy membership function is used to calculate membership value. Thus we calculate membership value of an unknown character for all other characters. The unknown charter is that character for which it gives highest membership value. If an unknown character has highest membership value against numeral zero then the character is zero.

It is seen that vector length 56 provides highest recognition accuracy for both features. The following table 1 and table 2 depict recognition accuracy of distance vector feature for numerals and vowels respectively.

Table 1: Recognition accuracy of numerals for vector length 56 with distance vector feature

Digits	90% train 10% test	80% train 20% test	70% train 30% test	60% train 40% test	50% train 50% test
0	90.09	90.11	91.71	93.23	91.75
1	85.34	80	82.29	85.05	86
2	84.31	82.04	76.43	79.89	72.1
3	79.83	80.32	77.71	77.2	67.96
4	94.83	93.98	93.57	91.94	83.93
5	89.74	86.02	87.43	87.1	88.32
6	81.72	78.49	71.14	66.34	63.97
7	87.5	90.32	91.14	92.37	92.44
8	90.95	88.6	88.43	86.45	83.59
9	93.53	90.75	85.29	80.32	74.74
Average	87.78	86.06	84.51	83.98	80.48

The highest individual accuracy is obtained for numeral 4 with 94.83% when 10% testing is considered. The lowest individual accuracy is obtained for numeral 6 with 63.97% when 50% testing is considered.

Table 2: Recognition accuracy of vowel for vector length 56 with distance vector feature

Vowels	90% train 10% test	80% train 20% test	70% train 30% test	60% train 40% test	50% train 50% test
অ	90.2	90	91.5	90.74	85.56
আ	95.49	90.98	91.73	83.23	87.23
ই	82.71	79.7	81.2	81.95	81.95
ঈ	81.2	79.7	79.7	81.95	80.45
উ	78.95	78.95	77.44	76.7	76.69
ঊ	81.2	77.44	77.44	78.2	76.69
ঋ	78.95	76.69	75.19	72.95	76.69
এ	72.18	73.68	69.92	70.44	68.42
ঐ	69.92	74.44	69.17	76.95	70.68
ও	87.97	86.24	82.71	71.43	75.71
ঔ	87.22	85.49	75.94	83.68	73.23
Average	82.36	81.21	79.27	78.93	77.57

The highest individual accuracy is obtained for vowel আ with 95.49% when 10% testing is considered. The lowest individual accuracy is obtained for vowel ঐ with 69.17% when 30% testing is considered. It is seen that as the training sample decreases the accuracy decreases a little bit for both numerals and vowels.

The following table 3 and table 4 depict recognition accuracy of density vector feature for numerals and vowels respectively.

Table 3: Recognition accuracy of numerals for vector length 56 with density vector feature

Digits	90% train 10% test	80% train 20% test	70% train 30% test	60% train 40% test	50% train 50% test
0	93.1	91.81	90.14	94.09	92.7
1	85.34	80.6	79.78	76.24	86.77
2	84.22	83.36	81.86	79.57	60.43
3	83.71	76.72	73.43	73.66	74.48
4	95.69	96.12	94.71	92.47	89.28
5	91.9	93.53	90.14	90.54	92.1
6	83.02	79.66	76.86	70.86	68.44
7	84.05	81.81	89.43	91.18	90.89
8	91.81	87.93	90.57	87.63	89.65
9	90.09	84.48	81.57	77.74	76.68
Average	88.29	85.60	84.84	83.39	82.14

The highest individual accuracy is obtained for numeral 4 with 96.12% when 20% testing is considered. The lowest individual accuracy is obtained for numeral 6 with 68.44% when 50% testing is considered.

Table 4: Recognition accuracy of vowel for vector length 56 with density vector feature

Vowels	90% train 10% test	80% train 20% test	70% train 30% test	60% train 40% test	50% train 50% test
অ	96.09	92.67	90.19	88.71	87.66
আ	85.41	84.62	85.86	86.84	81.2
ই	91.43	86.92	90.15	82.63	89.77

ঈ	81.35	79.32	83.83	79.05	75.79
উ	85.64	84.14	79.62	80.6	82.78
ঊ	87.07	82.03	86.02	84.51	90.53
ঋ	81.95	82.48	80.98	80.23	76.69
এ	91.28	86.84	86.83	87.3	85.05
ঐ	88.87	87.2	84.21	92.48	89.95
ও	82.03	83.31	85.34	80.34	76.84
ঔ	80.94	86.17	81.43	79.93	75.19
Average	86.55	85.06	84.95	83.87	82.86

The highest individual accuracy is obtained for vowel অ with 96.09% when 10% testing is considered. The lowest individual accuracy is obtained for vowel ঔ with 75.19% when 30% testing is considered.

The recognition accuracy for consonant is shown in table 5 and 6 for distance vector and density vector feature respectively.

Table 5: Recognition accuracy of consonant for vector length 56 with distance vector feature

Class	90% train 10% test	80% train 20% test	70% train 30% test	60% train 40% test	50% train 50% test
1	85.52	83.74	81.34	79.62	79.74
2	81.04	80.23	78.14	77.04	76.23
3	77.66	74.73	70.06	73.66	70.73
4	75.04	70.23	74.6	71.04	66.23
5	82.98	83.99	80.68	78.98	79.99
6	76.17	74.75	70.21	72.17	70.75
7	79.91	80.69	82.85	75.91	76.69
8	88.65	90.16	85.35	84.65	86.16
9	68.95	72.49	69.3	64.95	68.49
10	90.04	91.44	88.04	86.04	87.44
11	90.26	88.66	84.94	86.26	84.66
12	88.59	89.93	86.66	84.59	85.93
13	70.62	68.96	66.32	66.62	64.96
14	77.83	77.23	72.59	73.83	73.23
15	75.99	73.99	69.69	71.99	69.99
16	75.51	74.75	77.21	71.51	70.75
17	77.36	68.73	70.36	73.36	64.73
18	80.16	82.72	78.06	76.16	78.72
19	76.8	78.2	77.5	72.8	74.2
20	75.89	77.23	72.59	71.89	73.23
21	78.07	77	73.27	74.07	73
22	82.73	83.47	78.73	78.73	79.47
24	84.36	81.73	83.07	80.36	77.73
25	76.77	74.97	73.47	72.77	70.97
26	64.73	57	57.77	60.73	53
27	68.06	66.7	64	64.06	62.7
28	74.99	73.99	77.68	70.99	69.99
29	74.78	71.21	68.58	70.78	67.21
30	75.91	75.72	70.61	71.91	71.72
31	72.78	71.21	68.48	68.78	67.21
32	74.58	76.4	71.28	70.58	72.4
33	84.05	86.4	85.7	80.05	82.4
34	81.02	85.2	83.42	77.02	81.2
35	85.01	82.9	80.71	81.01	78.9
36	94.28	91.2	90.42	90.28	87.2
37	94.45	89.9	88.65	90.45	85.9
38	92.67	90.73	88.07	86.67	86.73
39	93.38	91.21	86.57	89.38	87.21
Average	80.19	79.21	77.03	76.1	75.21

Table 6: Recognition accuracy of consonant for vector length 56 with density vector feature

Class	90% train 10% test	80% train 20% test	70% train 30% test	60% train 40% test	50% train 50% test
1	87.52	85.24	83.54	81.56	79.66
2	82.94	81.73	80.34	82.31	80.41
3	79.56	76.23	72.26	71.3	69.4
4	76.94	71.73	76.8	74.74	72.84
5	84.88	85.49	82.88	80.73	83.53
6	78.07	76.25	72.41	73.46	76.26
7	81.81	82.19	85.05	80.46	83.26
8	90.55	91.66	87.55	85.42	88.22
9	70.85	73.99	71.5	73.97	76.77
10	91.94	92.94	90.24	88.34	91.14
11	92.16	90.16	87.14	85.76	83.86
12	90.49	91.43	88.86	86.58	84.68
13	72.52	70.46	68.52	65.12	63.22
14	79.73	78.73	74.79	71.81	69.91
15	77.89	75.49	71.89	71.89	69.99
16	77.41	76.25	79.41	79.41	75.51
17	79.26	70.23	72.56	72.56	68.66
18	82.06	84.22	80.26	82.53	78.63
19	78.7	79.7	79.7	79	75.1
20	77.79	78.73	74.79	71.81	67.91
21	79.97	78.5	75.47	73.63	69.73
22	86.63	84.97	80.93	81.25	77.35
24	91.26	90.23	85.27	82.64	78.74
25	78.67	76.47	75.67	73.62	69.72
26	66.63	58.5	59.97	55.63	51.73
27	69.96	68.2	66.2	62.55	58.65
28	76.89	75.49	79.88	74.73	72.83
29	76.68	72.71	70.78	71.67	69.77
30	77.81	77.22	72.81	72.81	70.91
31	74.68	72.71	70.68	70.68	68.78
32	76.48	77.9	73.48	73.48	71.58
33	89.95	87.9	87.9	85.95	84.95
34	88.92	86.7	85.62	81.18	80.18
35	89.91	86.4	82.91	80.1	79.1
36	96.18	96.7	92.62	90.73	89.73
37	96.35	91.4	90.85	91.5	90.5
38	95.57	92.23	90.27	91.64	90.64
39	98.68	92.71	88.77	84.67	83.67
Average	82.74	81.05	79.23	77.82	75.92

The highest recognition accuracy is provided by density feature for numerals, vowels and consonants.

The following table 7 represents highest average recognition accuracy for numerals, vowels and consonants corresponding to various features.

Table 7: Recognition accuracy of numerals, vowel and consonants

	Recognition Rate (%)	
	Distance Vector	Density Vector
Digits	87.78	88.29
Vowels	82.36	86.55
Consonants	80.19	82.74

The average recognition time for one Assamese handwritten character is shown in table 8. The experiment is performed in Intel i3 2.0 GHz processor with 4GB RAM.

Table 8: Average recognition time of a Assamese handwritten character

Vector length		24	48	56	112
Time (ms)	Distance	4.81	5.54	5.89	6.98
	Density	4.14	4.87	5.94	7.69

The average feature extraction time of distance vector feature for one character is 4.5ms, 5.1ms, 5.3ms and 6.1ms for vector length 24, 48, 56 and 112 respectively.

It is found that average feature extraction time of density vector feature for one character is 3.8ms, 4.4ms, 5.4ms and 6.8ms for vector length 24, 48, 56 and 112 respectively. The membership value calculation time for distance vector feature is 0.31ms, 0.44ms, 0.59ms and 0.88ms for vector length 24, 48, 56 and 112 respectively. It is observed that membership value calculation time for density vector feature is 0.34ms, 0.47ms, 0.54ms and 0.89ms for vector length 24, 48, 56 and 112 respectively. The time mentioned in table is detection time excluding the result writing time on document.

It is found that as the vector lengths increase the detection time also increase due to computation time. We have found that the vector length 56 has highest recognition accuracy for both the features. This recognition time is impressive compared to the modified quadratic discriminant function classifier proposed by Xiaohua Wei et al.[18]. From the above experiment we may conclude our worked based fuzzy logic efficient in terms of speed.

VIII. CONCLUSION

In this paper, we present robust supervised fuzzy logic classifier using distance vector and density vector feature. The experiment is performed five various combination of training and testing. As the training sample decrease the accuracy does not decreases much. Hence it is a robust recognition methodology. It is noted that density vector outperforms distance vector feature. The highest recognition accuracy achieved is 88.29%, 86.55% and 82.74% for numerals, vowels and consonants respectively with density vector feature. The contributions of our works are summarized as bellow.

1. The primary objective of the task is the application of statistical feature i.e. static zoning method for handwritten character recognition. Particularly, we adopted two statistical features, distance vector and density vector for creating knowledge base of Assamese handwritten characters. Further they are used with a supervised fuzzy logic classifier.
2. We have analysis various vector lengths for both features and find out that vector length 56 is the most powerful amongst all.
3. Another objective of this work is to check application of supervised fuzzy logic classifier and fuzzy membership function in character recognition.
4. In order to check the robustness of the classifier we have considered 10%, 20%, 30%, 40% and 50% testing sample. It is observed that as the training sample decrease the accuracy does not decrease much more.
5. We observed the speed of supervised fuzzy logic classifier and found that it is feasible for real time applications for handwritten character recognition. It is noted that accuracy of the classifier needs to be improved.

Finally, we conclude that the experiment shows the effective utility of supervised fuzzy logic classifier in handwritten character recognition.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the Department of Computer Science, Gauhati University for all the necessary support.

REFERENCES

1. R. Sarkhel, N. Das, A. K. Saha, M. Nasipuri, “A multi-objective approach towards cost effective isolated handwritten Bangla character and digit recognition”, Pattern Recognition, vol. 58, Oct. 2016 , pp. 172–189.
2. V. Novák, P. Hurtík, H. Habiballa, M. Štepnička, “Recognition of damaged letters based on mathematical fuzzy logic analysis”, Journal of Applied Logic, vol. 13, Jun. 2015, pp. 94–104.
3. S. Shelke and S. Apte, “A Fuzzy based Classification Scheme for Unconstrained Handwritten Devanagari Character Recognition”, 2015 International Conference on Communication, Information & Computing Technology (ICCICT), Mumbai, Jan. 2015, pp. 1–6
4. D. Álvarez, R. A. Fernández and L. Sánchez, “Fuzzy system for intelligent word recognition using a regular grammar”, Journal of Applied Logic, vol. 24, Nov. 2017, pp. 45–53.
5. S. Saha, N. Saha, “A Lightning fast approach to classify Bangla Handwritten Characters and Numerals using newly structured Deep Neural Network”, Procedia Computer Science, vol. 132, Jun. 2018 , pp. 1760–1770.
6. C. W. Hsu, C. J. Lin, “A comparison of methods for multi-class support vector machines”, IEEE Transactions on Neural Networks, vol. 13, Mar. 2002, pp. 415–425.
7. M. Z. Alom, P. Sidike, M. Hasan, T. M. Taha and V. K. Asari, “Handwritten Bangla Character Recognition Using the State-of-the-Art Deep Convolutional Neural Networks”, Hindawi Computational Intelligence and Neuroscience, vol. 2018, Aug. 2018, pp. 1–13.
8. C. Wang, A. Xu and Xi. Li, “Supervised Classification High-Resolution Remote-Sensing Image Based on Interval Type-2 Fuzzy Membership Function”, *Remote Sensing*, vol. 10, May. 2018, pp. 2–22.
9. X. Y. Zhang, Y. Bengio and C. L. Liu, “Online and offline handwritten Chinese character recognition: A comprehensive study and new benchmark”, Pattern Recognition, vol. 61, Jan. 2017, pp. 348–360.
10. A. K. Sampath and N. Gomathi, “Fuzzy-based multi-kernel spherical support vector machine for effective handwritten character recognition”, Sadhana, vol. 42, Sep. 2017, pp. 1513–1525.
11. D. Impedovo and G. Pirlo, “Zoning methods for handwritten character recognition: A survey”, Pattern Recognition, vol. 47, Mar. 2014, pp. 969–981.
12. L. A. Zadeh, “Fuzzy sets”, Information and Control, vol. 8, 1965, pp. 338–353.
13. R. Bellman, R. Kalaba, L. A. Zadeh, “Abstraction and pattern classification”, Journal of Mathematical Analysis and Applications, vol. 13, Jan. 1966, pp. 1–7.
14. K. Nongmeikapam, W. K. Kumar, O. N. Meetei, T. Tuithung, “Increasing the effectiveness of handwritten Manipuri Meetei-Mayek character recognition using multiple-HOG-feature descriptors”, Sādhanā vol. 44, Apr. 2019, pp. 1–14.
15. G. Manogaran, R. Varatharajan, M. K. Priyan, “Hybrid Recommendation System for Heart Disease Diagnosis based on Multiple Kernel Learning with Adaptive Neuro-Fuzzy Inference System”, Multimedia Tools and Applications, vol. 77, Feb. 2018, pp 4379–4399.
16. H. Singh and B. S. Khehra, “Visibility enhancement of color images using Type-II fuzzy membership function”, Modern Physics Letters B, vol. 32, no. 11, pp. 1–12, 2018.

17. Z. Xi and G. Panoutsos, "Interpretable Machine Learning: Convolutional Neural Networks with RBF Fuzzy Logic Classification Rules", 2018 International Conference on Intelligent Systems, 2018, pp.448–454.
18. X. Wei, S. Lu and Y. Lu,"Compact MQDF classifiers using sparse coding for handwritten Chinese character recognition", Pattern Recognition, Vol. 76, pp. 679–690, 2018

AUTHORS PROFILE

Kalyanbrat Medhi, received the B.Sc. degree in computer science from Gauhati University in 2009 and M.C.A. degree from Gauhati University in 2012. He is currently pursuing the PhD in computer science at Gauhati University, Guwahati, Assam, India. His research interests include image processing, pattern recognition and machine learning.

Dr. Sanjib Kr. Kalita, Sanjib Kumar Kalita completed his MCA and Ph.D degree from Gauhati University, Guwahati, Assam, India. Currently he is working as Assistant Professor in the Department of Computer Science, Gauhati University. He has authored two books and published around seventy research articles in different National and International peer reviewed journals. Three research scholars completed Ph. D under his guidance. Presently five research scholars are working under his supervision. His area of interest are speech processing, Image processing and Hypersprectral Image Processing.