

Handwritten *Bangla* Digit Recognition Using Classifier Combination Through DS Technique

Subhadip Basu¹, Ram Sarkar², Nibaran Das², Mahantapas Kundu²,
Mita Nasipuri^{2,*}, and Dipak Kumar Basu²

¹ Computer Sc. & Engg. Dept., MCKV Institute of Engineering,
Liluah, Howrah-711204, India

² Computer Sc. & Engg. Dept., Jadavpur University,
Kolkata-700032, India

Abstract. The work presents an application of Dempster-Shafer (DS) technique for combination of classification decisions obtained from two Multi Layer Perceptron (MLP) based *classifiers* for optical character recognition (OCR) of handwritten *Bangla* digits using two different feature sets. *Bangla* is the second most popular script in the Indian subcontinent and the fifth most popular language in the world. The two *feature sets* used for the work are so designed that they can supply complementary information, at least to some extent, about the classes of digit patterns to the MLP classifiers. On experimentation with a database of 6000 samples, the technique is found to improve recognition performances by a minimum of 1.2% and a maximum of 2.32% compared to the average recognition rate of the individual MLP classifiers after 3-fold cross validation of results. The overall recognition rate as observed for the same is 95.1% on average.

1 Introduction

The work presented here deals with an OCR technique of handwritten *Bangla* digits. In doing so, classification decisions from two complementary sources of information are combined using DS technique of belief combination [1]. It is in line with the recent trend for improving classification decisions through the *classifier combination approach*.

Bangla is the second most popular script in the Indian subcontinent. As a script, it is used for *Bangla*, *Ahamia* and *Manipuri* languages. *Bangla* is also the national language of Bangladesh and is the fifth most popular language in the world. Evidences of research on OCR of handwritten *Bangla* digits are a few in numbers, as found in the literature [2,3].

Techniques developed for OCR of handwritten digits are in general required for many potential *applications* related to reading amounts from bank cheques, extracting numeric data from filled in forms, interpreting handwritten pin codes from mail pieces and so on. Considering the need of a wide cross section of world population using

* Corresponding Author. *Email address:* nasipuri@vsnl.com

Bangla either as a language or a script, OCR of handwritten *Bangla* digits has a high value of *commercial importance* in the present time. Fig. 1 shows the typical digit patterns of the decimal digits of *Bangla* script.

০	১	২	৩	৪	৫	৬	৭	৮	৯
0	1	2	3	4	5	6	7	8	9

Fig. 1. The decimal digit set of Bengali script

Previously Bhattacharya et al. [2] employed group specific MLP classifiers for recognition of handwritten *Bangla* digits after they are hierarchically classified into some groups of digit classes on the basis of features like loops, junctions etc., extracted from *skeletal shapes* of the digit patterns. The system so developed is trained with a data set of 1880 samples. On being tested with a separate data set of 3440 samples, it shows a recognition performance of 90.56%.

The authors previously developed a *two pass approach* to pattern classification [3] and applied the same for OCR of handwritten *Bangla* digits. In this approach, an unknown digit pattern is first coarsely classified into a group of digit classes on the basis of some preselected *global features* and then classified into some digit class within the group on the basis of some group specific *local features*. Performances of the system developed under the work are tested with a *training set* and a *test set* of 300 and 200 randomly selected sample patterns respectively. It is experimentally observed that application of the two-pass approach improves the classification rate of digit patterns to 93.5% in the second pass from 90.5% observed in the first pass. Due to lack of a common database used for the above two techniques, performances of the same cannot be compared straightway. A major *limitation* of the two-pass approach is that it cannot refine a classification decision on an unknown digit pattern, which is wrongly classified into a group of candidate patterns in the first pass. The *present technique*, aims to overcome this limitation by combining classification decisions of the two MLP classifiers through DS techniques.

The work presented here not only targets the development of suitable *feature sets* for representation of handwritten *Bangla* digits in the feature space but also the development of techniques for dealing with *uncertainty* and *imprecision* in recognizing the same.

2 The Present Work

The two feature sets considered for the present work are so designed that they contain *complementary information* on *Bangla* digit features for improving recognition performances through combination of classification decisions of two MLP classifiers by DS technique. One of the two feature sets selected for the present work consists of 40 features and the other 36 features. Out of these two, the first one consists of octant based 24 *shadow feature* and octant based 16 *centroid features* as described in [3]. The other feature set consists of 36 *longest run features* [4] computed from 9 overlapping windows on the image frame.

To deal with imprecision present in handwritten digit patterns due to variation in writing styles of different individuals, MLPs, a kind of *feed forward* Artificial Neural Networks (ANNs), are used as classifiers for the present work. ANNs are in general famous for their *learning* and *generalization abilities*, necessary for dealing with imprecision in input patterns. The two MLPs employed here to work as two pattern classifiers with the two above mentioned feature sets are of the sizes (40-40-10) and (36-50-10).

Uncertainty of classification decisions occurs because of the fact that unknown handwritten digit patterns may appear close to some other digit pattern(s), which does not belong to its actual class. So, no classification decision taken by any of the two classifiers designed for the present work is considered to be true with the highest degree of probability. One of the major points of *criticism* of probability theory is that it represents *ignorance* by *indifference* or by *uniform probabilities*. Another heavily debated point about the probability theory involves the way it follows in fixing the probability of a hypothesis negation ($p(\sim H)$) once the probability of its occurrence ($p(H)$) is known. In DS theory, it is claimed that, in many situations, evidence that only partially favours a hypothesis should not be construed as also partially supporting its negation.

2.1 DS Technique and Its Application for Combination of Classification Decisions

To deal with the above problems, DS technique involves the following. *Firstly*, it assigns separate *probability masses* to all subsets of a universe of discourse (U) containing all possible hypotheses. For handwritten digit recognition, the value of |U| is 10. The basic probability assignment can be represented as follows:

$$m: 2^U \rightarrow [0,1]$$

The *basic probability assignment* (m) satisfies the relations given as $m(\phi) = 0$ where ϕ denotes a null hypothesis and $\sum_{A \subseteq U} m(A) = 1$.

Secondly, DS technique introduces the concept of *belief* (B) associated with each subset of U as $\mathbf{Bel}(A) = \sum_{B \subseteq A} m(B)$ and $\mathbf{Bel}(\phi) = 0$

The belief function (B) when applied to a set (A) gives a measure of the *total support* or belief committed to all hypotheses constituting A. It sets a *minimum value* for its *likelihood*. The belief function (B) satisfies the relation given as $\mathbf{Bel}(A) + \mathbf{Bel}(\sim A) \leq 1$

In DS technique, only those subsets of U, which have non zero basic probability assignments, are of interest. Each such subset is called a *focal element* of the belief function (Bel) over 2^U .

Thirdly, DS technique introduces a concept of *plausibility* (Pl) of a set of hypotheses (A) giving the maximum amount of belief that can possibly be assigned to A as $\mathbf{Pl}(A) = 1 - \mathbf{Bel}(\sim A)$. Introduction of the concept of plausibility makes it possible to define the *confidence* in A as the belief subinterval $[\mathbf{Bel}(A), \mathbf{Pl}(A)]$ of [0,1]. The

quantity ($\mathbf{Pl}(\mathbf{A}) - \mathbf{Bel}(\mathbf{A})$) is considered as the *uncertainty* in \mathbf{A} . This quantity can never be negative. That is, $\mathbf{Pl}(\mathbf{A}) - \mathbf{Bel}(\mathbf{A}) \geq 0$.

Fourthly, and finally, DS technique also defines rules for combination of belief functions when evidence is available from two or more independent knowledge sources. If m_1 and m_2 be two probability mass assignment functions designed for two such knowledge sources then a function ' $m_1 \oplus m_2$ ' for combining these two can be given as

$$m_1 \oplus m_2 (\mathbf{A}) = \mathbf{K} \cdot \sum_{X \cap Y = \mathbf{A}} m_1(X) \cdot m_2(Y), \text{ where } \mathbf{K}^{-1} = \sum_{X \cap Y \neq \emptyset} m_1(X) \cdot m_2(Y).$$

For totally contradictory evidences $\mathbf{K}^{-1} = 0$.

An Example. The technique described above can be illustrated with the following examples formed out of the experimentation conducted for the present work.

The universe of discourse for the handwritten digit recognition problem consists of ten hypotheses given as $\mathbf{U} = \{ '0', '1', '2', '3', '4', '5', '6', '7', '8', '9' \}$. Each hypothesis in \mathbf{U} denotes the assignment of an unknown digit pattern to a specific digit class represented by the class number placed within a pair of quotation marks.

From the responses of two MLPs employed here to recognize independently an unknown digit pattern on the basis of two different feature sets, basic probability assignments, denoted by m_1 and m_2 , for the focal elements of \mathbf{Bel}_1 and \mathbf{Bel}_2 are determined. Instead of considering all elements of the power set $2^{\mathbf{U}}$ for forming the focal elements of the two belief functions, only the following *three subsets* are considered from the responses of each of the two classifiers.

{ The first highest-ranking class }, { The first highest-ranking class, the second highest-ranking class }, { The first highest ranking class, the second highest-ranking class, the third highest-ranking class }.

The probability assignment for a singleton subset is computed by dividing the *maximum output* of the MLP classifier by the *sum* of its all outputs. For computing the same for the other subsets, a subset is first expressed as the union of singletons, each of which contains an element of the subset. Then the probability assignments for all these singletons are to be computed. And, finally by taking the *minimum* of the probability assignments of the constituent singletons of a subset, the probability assignment of the entire subset is determined.

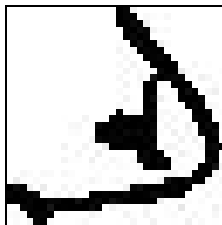


Fig. 2. A sample digit image of '9'

Following the above procedure, the basic probability assignments are made from the responses of the two MLP classifiers to a common *input digit patterns* of '9', shown in Fig. 2. All these basic probability assignment are given below.

$$m_1(\{'6'\})=0.578644, m_1(\{'6','9'\})=0.356895, m_1(\{'6','9','1'\})=0.051056, m_1(U) = 0.013405 \text{ and } m_2(\{'3'\})=0.375894, m_2(\{'3','9'\})=0.271633, \\ m_2(\{'3','9','1'\})=0.205506 \text{ } m_2(U) = 0.146967.$$

To combine basic probability assignments for the non null intersections of the focal elements of the two belief functions formed with the responses of the two MLP classifiers, Table 1 is formed. In Table 1, probability mass assignments of the all possible intersections are shown by computing the numerators of previously given formula for computation of ' $m_1 \oplus m_2$ '. Out of these probability mass assignments, the assignments for singleton cross sections $\{3\}$, $\{6\}$ and $\{9\}$ are only of interest for the present work. By considering the maximum of the probability mass assignments of singleton intersections, the input pattern is finally classified in to the digit class '9', which is its *true class*. It is worth noting that both the classifiers have misclassified the input pattern by their top choices. But by considering other choices through DS technique, it has been possible to obtain the correct one. Moreover, it is due to power of DS technique of belief combination that it has been possible to reach the correct classification decision by enhancing the beliefs for the second top choices of the two MLP classifiers. With a simpler classifier function based on the policy of reaching a consensus through majority *voting*, it could have not been possible.

Table 1. Combined probability mass assignments ($m_1 \oplus m_2$) of all possible intersections of focal elements

$m_1 \backslash m_2$	$\{'3'\}$ (0.375894)	$\{'3','9'\}$ (0.271633)	$\{'3','9','1'\}$ (0.205506)	U (0.146967)
$\{'6'\}$ (0.578644)	Φ (0.217509)	Φ (0.157179)	Φ (0.118915)	$\{'6'\}$ (0.085042)
$\{'6','9'\}$ (0.356895)	Φ (0.134155)	$\{'9'\}$ (0.096944)	$\{'9'\}$ (0.073344)	$\{'6','9'\}$ (0.052452)
$\{'6','9','1'\}$ (0.051056)	Φ (0.019192)	$\{'9'\}$ (0.013868)	$\{'9','1'\}$ (0.010492)	$\{'6','9','1'\}$ (0.007504)
U (0.013405)	$\{'3'\}$ (0.005039)	$\{'3','9'\}$ (0.003641)	$\{'3','9','1'\}$ (0.002755)	U (0.375894)

3 Results and Discussion

For preparation of the *training* and the *test sets* of samples, a database of 6,000 *Bangla* digit samples is formed by randomly selecting 600 samples for each of 10 digit classes from a larger database of 10,000 samples. The larger database is prepared in CVPR unit, ISI, Kolkata. Samples constituting the database are collected from pin codes used on postal mail pieces. A *training set* of 4000 samples and a *test set* of 2000 samples are finally formed. Three such pairs of the training and the test sets are

formed in all with the entire database of 6000 samples for *3-fold cross validation* of results at the time of experimentation. All these samples are first enclosed within a *minimum bounding box* each and then scaled to 32x32 pixel images. These images are finally converted to *binary images* through thresholding.

Table 2. Recognition Performances on Test sets

Classifier	% Recognition Rate		
	Test Set # 1	Test Set # 2	Test Set # 3
MLP # 1	93.1	93.2	94.8
MLP # 2	91.35	92.8	93.7
DS Technique	94.55	95.3	95.45

Recognition performances of the individual MLP classifiers and their combined classification decisions, as recorded experimentally on the three sets of test data are shown in Table 2. The overall recognition rate the technique finally achieves is **95.1%** on average.

Compared to the average recognition performances of the two classifiers, classifier decision combination through DS technique shows a minimum of **1.2%** and a maximum of **2.32%** increase in recognition rates of the classifier ensemble after *3-fold cross validation* of results. Improvement thus achieved is significant especially in the context of high precision applications relating to OCR of handwritten digits.

Acknowledgements

Authors are thankful to the “Center for Microprocessor Application for Training Education and Research”, “Project on Storage Retrieval and Understanding of Video for Multimedia” of Computer Science & Engineering Department, Jadavpur University, for providing infrastructure facilities during progress of the work. Authors are also thankful to the CVPR Unit, ISI Kolkata, for providing the necessary dataset of handwritten Bengali Numerals. One of the authors, Mr. S. Basu, is thankful to MCKV Institute of Engineering for kindly permitting him to carry on the research work.

References

1. Kenng-Chi Ng and Bruce Abramson, “Uncertainty management in expert systems”, IEEE Expert, p.p. 29-48, April, 1990.
2. U. Bhattacharya et. al. “Recognition of hand printed Bangla numerals using neural network models”, AFSS 2002, p.p. 228-235.
3. S. Basu et al., “A two pass approach to pattern classification”, Proc 11th ICONIP, Nov-2004.
4. S. Basu et al., “Handwritten ‘Bangla’ alphabet recognition using an MLP based classifier”, Proc. 2nd NCCPB, Dhaka, Feb-2005.