

“© 2018 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

# Cognitive Analysis for Reading and Writing of Bengali Conjuncts

Chandranath Adak<sup>\*†</sup>, Bidyut B. Chaudhuri<sup>‡</sup>, Michael Blumenstein<sup>\*†</sup>

<sup>\*</sup>CAI, School of Software, University of Technology Sydney, Australia-2007

<sup>†</sup>IIS, School of ICT, Griffith University, Gold Coast, Australia-4222

<sup>‡</sup>CVPR Unit, Indian Statistical Institute, Kolkata, India-700108

adak32@gmail.com

**Abstract**—In this paper, we study the difficulties arising in reading and writing of Bengali conjunct characters by human-beings. Such difficulties appear when the human cognitive system faces certain obstructions in effortlessly reading/writing. In our computer-based investigation, we consider the reading/writing difficulty analysis task as a machine learning problem supervised by human perception. To this end, we employ two distinct models: (a) an auto-derived feature-based Inception network and (b) a hand-crafted feature-based SVM (Support Vector Machine). Two commonly used Bengali printed fonts and three contemporary handwritten databases are used for collecting subjective opinion scores from human readers/writers. On this corpus, which contains the perceptive ground-truth opinion of reading/writing complications, we have undertaken to conduct the experiments. The experimental results obtained on various types of conjunct characters are promising.

**Index Terms**—CNN; Human cognition; Inception network; Reading difficulty; SVM; Writing difficulty.

## I. INTRODUCTION

“Writing” is a process of making patterns/graphical symbols to represent the phonology of a spoken language and “Reading” is a process to comprehend the meaning of the writing. For such comprehension, human vision begins at the sensory organ eyes and is processed in the visual cortex of the brain. In the brain, *cognition* plays a vital role in this comprehension. “Cognition” refers to “the mental action or process of acquiring knowledge and understanding through thought, experience, and the senses”<sup>1</sup>.

It is noted that some human beings find it difficult to properly read or write some characters of a certain script [1]. This problem may be interesting to understand the cognitive perspective of the reading/writing difficulty in comprehension/generation.

In this paper, we try to investigate the difficulties which may arise in reading and writing. We analyze the subjective opinion score [2] of difficulty collected from several humans who assess the degree of difficulty using their own cognitive system. Providing such subjective rating scores for video quality assessment is quite common in the video broadcasting domain [2]. Here, we aim to impart to a machine a human-like insight to make an artificial cognitive system that understands the complications appearing in reading/writing. Moreover, such an analysis can be used in the pre-processing module of

an OCR (*Optical Character Recognition*) engine for sorting the texts with respect to the reading/writing difficulties [3]. Then the complex one may be sent to a different branch of the OCR system, which boosts the accuracy.

In the *Educational research*<sup>2</sup> domain, the philosophical aspects of reading and writing have been studied for more than 100 years [4]. Some *education*-based logical strategies for English learning can be found in [5]. It is also popular in *psychological* studies [6], where the main concern is the mental development process of reading and writing skills. Some small case-based statistical analysis can also be found, e.g., work on Arabic script learning by Taouk and Coltheart [7]. However, a computer-assisted automated approach for such reading/writing analysis is rare. A very preliminary computational approach, mainly discussing the rational aspects of remembering Bengali basic characters, is reported in [1].

This paper is an early attempt at the automated analysis of cognitive perspectives related to reading/writing complications. We perform our task on an Abugida Indic script, Bengali [8], [9], owing to its complex patterns. Here, we formulate our reading/writing difficulties analysis task as a feature-based classification problem. For our task, we use two separate models: (i) an auto-derived feature-based Inception network, and (ii) a hand-crafted feature-based SVM (*Support Vector Machine*). We also employ several human individuals to put some cognitive opinion score on Bengali conjunct characters as groundtruth. We note that a large population faces problems in reading/writing of some Bengali conjunct characters rather than basic characters [1]. Therefore, we analyze our method on Bengali conjuncts only. In *Section II*, we discuss Bengali conjunct characters.

The rest of the paper is arranged as follows. *Section II* presents some characteristics of Bengali script. *Section III* describes the proposed methodology. After that, the experiments and results are presented in *Section IV*. Finally, *Section V* concludes the paper.

## II. BRIEFINGS ON BENGALI SCRIPT

The Bengali (endonym: *Bangla*) language is used by more than 265 million native speakers located mostly in south-east Asia. It is the national and foremost official language of Bangladesh, and one of the official languages of India.

<sup>1</sup>Oxford Dictionary, Online: <https://en.oxforddictionaries.com> .

<sup>2</sup>Journal of Education, Online: <http://www.bu.edu/journalofeducation> .

Bengali script, evolved from Brahmi script, is about 1000 years old and its present form came into being in 17<sup>th</sup> century AD [8]. In modern Bengali script, more than 300 characters are present, among which 50 characters are basic [9]. More than one basic character (up to four characters) can be joined together to create a conjunct character. The number of such conjunct characters is more than 250. Besides, some conjunct characters are written in several shapes. This makes the conjunct character set even larger and more complex in structure. Basically, two types of conjunct characters are formed by typologists. They are called transparent and non-transparent [1]. In transparent conjuncts, the basic character shape remains nearly the same. However, in the non-transparent conjuncts, the basic character shape may change drastically. Some examples of printed and handwritten Bengali conjuncts are shown in Fig. 1 and Fig. 2, respectively.

হ্ + ম	ঞ + চ	ক্ + ত্ + ব
হ্ম	ঞ্	ক্‌ত্ব
ক্ষ	ঞ্চ	ক্ব

Fig. 1. Printed Bengali conjunct characters: 1<sup>st</sup> row shows conjoining basic characters, 2<sup>nd</sup> row depicts transparent conjuncts (*Bangla Akademi* font), 3<sup>rd</sup> row represents non-transparent conjuncts (*AdorshoLipi* font). Each column is comprised of the same conjunct.

Moreover, handwritten Bengali conjuncts introduce additional patterns due to the intensive variation of writing styles and cursiveness [10], [11].



Fig. 2. Handwritten Bengali conjunct characters: Each box contains a pair of the same conjunct, written in two different ways.

More detailed descriptions of Bengali language and script characteristics can be found in [1], [8]–[10].

### III. PROPOSED METHOD

In this section, at first, we formulate the problem and then discuss our proposed methodology.

#### A. Problem Formulation

The sense of reading and writing difficulties depends on human perception and here we attempt to teach the machine this cognitive perceptivity in a quantitative way.

Some conjunct characters both in isolated form and word-within form are shown to human volunteers and asked to read/write them. The volunteers are requested to put a subjective score of reading-difficulty within a continuous range of  $[r_1, r_2]$ ;  $r_1, r_2 \in \mathbb{R}$ . Similarly, for writing-complication analysis, the human writer is asked to provide a score in

a continuous range  $[w_1, w_2]$ ;  $w_1, w_2 \in \mathbb{R}$ . During data collection, the subjective scores are marked in a continuous range, since primarily we avoid restricting the scores to a fixed discrete opinion scale, e.g. *Likert scale* [12]. For each conjunct character, several individuals provide their opinion scores. The arithmetic mean of all these scores is calculated and labeled with each character as its *mean opinion score*.

After data collection, the reading-complication score range  $[r_1, r_2]$  is partitioned into  $n_r$  bins. Therefore, all the *reading mean opinion scores* of the entire character set fall into these  $n_r$  bins. In other words, all the conjunct characters are categorized into  $n_r$  classes with respect to the reading-complication score. Similarly, the conjunct characters are tagged with *writing mean opinion scores* distributed into  $n_w$  bins of score range  $[w_1, w_2]$ .

Thus, we map the cognitive complication analysis of reading and writing into classification problems. Now, the task is to label a given conjunct into  $n_r$  classes of reading and  $n_w$  classes of writing opinion score. Such a categorization of opinion scores is quite a well-known topic in the video quality assessment domain [2] and relatively new in the field of document image analysis [13].

The reading/writing difficulty analysis task is performed using two separate methods: an auto-derived feature-based Inception network and a hand-crafted feature-based SVM. These methods are discussed as follows.

#### B. Auto-derived Feature-based Inception Network

In this subsection, we discuss the auto-derived features employed for classification of conjunct characters with respect to the reading/writing effort opinion score.

A standard CNN (*Convolutional Neural Network*) architecture [14] can be considered as a conjoint of two parts: *front* and *rear*. The *front* part is the convolutional part, considered as an automated feature extractor, which takes an image input and produces the feature maps. The *rear* part is usually an MLP (*Multi-Layer Perceptron*) that is used as a classifier. The auto-extracted features are fed to the MLP for obtaining the classified output through the stacked fully connected layers.

For our reading/writing analysis task, we adopt the deep neural architecture of [15] where the concept of an “*inception*

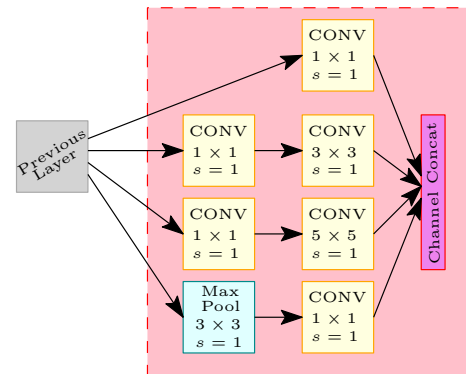


Fig. 3. Inception Module.

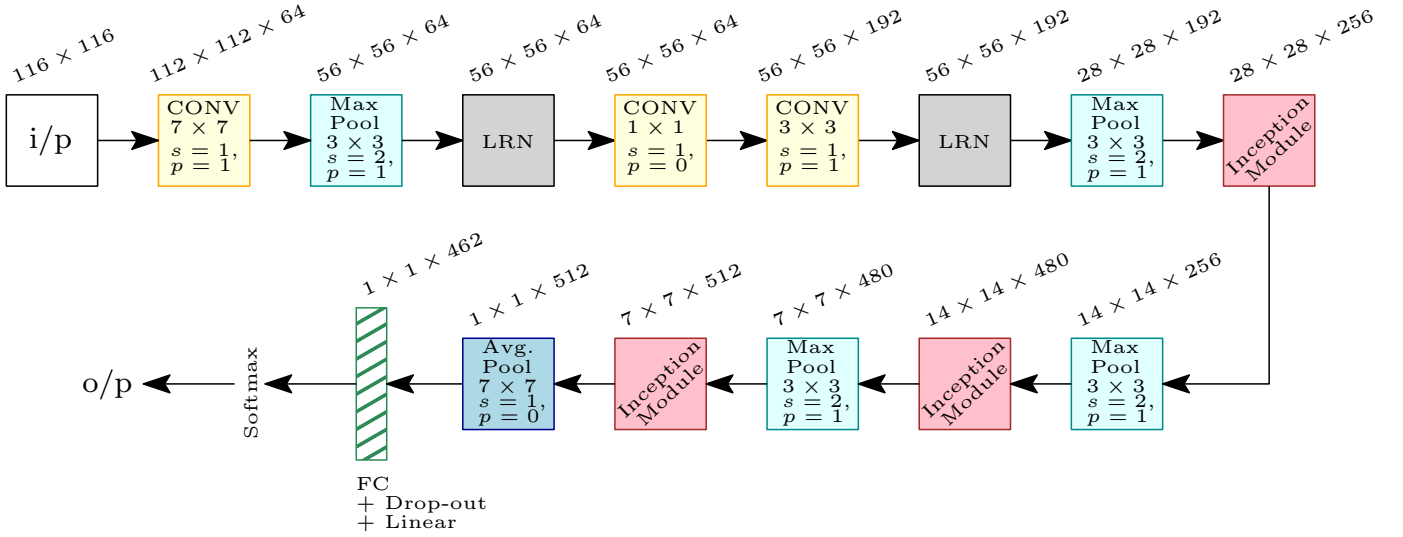


Fig. 4. Our adopted Inception Network model.

module” is reported. In Fig. 3, we show the inception module used in our work. Here, “CONV” denotes the convolutional layer, “Max Pool” represents the max-pooling layer,  $s$  denotes the stride value. The filter size and stride value used in our experiment are provided in Fig. 3.

We adopt the idea of an inception module and work with a relatively smaller version of [15] due to fewer classes involved in our problem and less memory complexity requirement. Our adopted version of the Inception Network is presented in Fig. 4. This architecture takes a fixed-sized character image input ( $i/p$ ). We feed the normalized character of size  $116 \times 116$  to the network. In the front part of the Inception network, the first convolutional layer (CONV) contains 64 feature maps of size  $112 \times 112$ . Each of these feature maps is linked with a  $7 \times 7$  neighborhood of the  $i/p$  image. Here, stride  $s = 1$  and padding  $p = 1$  are used. The following max-pooling (Max Pool) layer, LRN (*Local Response Normalization*), CONV layers are shown in Fig. 4 with the corresponding employed filter size, the number of feature maps, map size, stride, padding etc. The actual inner-view of inception module of Fig. 4 is shown in Fig. 3. In total, we use three inception modules here. We use ReLU (*Rectified Linear Unit*) [16] as the activation function for all the convolutions including those inside the inception module.

After the last inception module, the rear part of the Inception network contains an average pooling (“Avg. Pool”) layer which obeys the technique of [17]. The following layer is fully connected (FC) with a *dropout* of 40% [18]. The linear layer with *softmax activation* is used for predicting the class. The output size of each layer is shown in Fig. 4.

### C. Hand-crafted Feature-based SVM

In this subsection, we discuss the hand-crafted features employed for categorization of characters with respect to the reading/writing effort opinion score. We use an SVM for this categorization.

1) *Feature extraction*: Here, we design some hand-crafted features of a character to analyze the reading/writing difficulties. The extracted features are discussed below.

a) *Structural Keypoint*: A character stroke contains multiple structural keypoints [10], such as *end*, *junction*, *curvature* and *cusp* points [19]. The *end* point denotes the start/stop point of a stroke. The *junction* point is actually the intersection position of multiple strokes or the furcation point of the same stroke. The *curvature* point represents the changing orientation spot of a stroke. The *cusp* point on a stroke is a point where the direction of the stroke changes sharply with the non-existence of a derivative. These keypoints may influence the reading/writing of a character.

Here we count the end points, junction points, curvature points and cusp points on a character stroke, and use them as feature  $f_1$ ,  $f_2$ ,  $f_3$ , and  $f_4$ , respectively.

b) *Loop*: Some Bengali characters contain loop/hole/delta-like structures [20]. Such structures may be useful for analyzing complications in reading/writing. The count of holes is used as feature  $f_5$ .

Let,  $LP$  be the set of object pixels ( $p \in LP$ ) in a loop region, whereas we assume  $LC$  to be the set of pixels on the loop region boundary. The numbers of object pixels in  $LP$  and  $LC$  are denoted by  $n(LP)$  and  $n(LC)$ , respectively. The center of gravity ( $CG$ ) inside the loop region is ( $CG.x$ ,  $CG.y$ ) given by:

$$CG.x = \frac{1}{n(LP)} \sum_{p \in LP} p.x ; \quad CG.y = \frac{1}{n(LP)} \sum_{p \in LP} p.y .$$

We calculate the standard deviation ( $\sigma_L$ ) of the loop from not being an ideal circle [20] and use it as feature  $f_6$  to fit our task.

$$f_6 \equiv \sigma_L = \sqrt{\frac{1}{n(LC)} \sum_{p \in LC} (X_p - \mu_L)^2} \quad (1)$$

where,  $X_p = \sqrt{(p.x - CG.x)^2 + (p.y - CG.y)^2}$  and  $\mu_L = \frac{1}{n(LC)} \sum_{p \in LC} X_p$ .

We measure another loop-feature ( $f_7$ ) reliant on the *roundness* of an object [20], as follows.

$$f_7 \equiv \text{roundness} = \frac{4\pi \cdot n(LP)}{(\sum_{p \in LC} X_p)^2} \quad (2)$$

c) *Isolated component*: A character may be combination of some isolated components. Such components increment pen up/down movements which may obstruct effortless writing/reading. We count the number of such components and use that as feature  $f_8$ .

d) *Concavity*: The character cavity region can be perceived as a “water reservoir” in 2-Dimensions, where “water” can be “poured” from one side of the character shape [21]. We consider four sides, i.e., top, bottom, left, and right, from where “water” can fill the character’s concavity or “water reservoir” until a spill-over occurs (refer to Fig. 5(a)). Such a character concavity may effect the process of graceful reading/writing. Therefore, we count the number of top, bottom, left, and right reservoirs and use these as features  $f_9, f_{10}, f_{11}$ , and  $f_{12}$ , respectively.

We also calculate the *area* of a reservoir that signifies its capacity or concavity. The total area of top reservoirs is computed as  $A_T$ . Similarly, the total area of the bottom, left, and right reservoirs are  $A_B, A_L$ , and  $A_R$ , respectively. The top-bottom reservoir area difference ( $A_{TB}$ ) is calculated and employed as feature  $f_{13} \equiv A_{TB} = \frac{|A_T - A_B|}{(A_T + A_B)}$ . Likewise, the left-right reservoir area difference ( $A_{LR}$ ) is computed to use as a feature  $f_{14} \equiv A_{LR} = \frac{|A_L - A_R|}{(A_L + A_R)}$ .

e) *Straight line*: A straight line in a character shape is formed by combining collinear object pixels. The presence of straight lines in a character may influence how easy it is to read/write. We count the horizontal (—), vertical (|), front-diagonal (/), and back-diagonal (\) straight lines, and use these counts as features  $f_{15}, f_{16}, f_{17}$ , and  $f_{18}$ , respectively.

f) *Loci*: We use a pseudo-structural descriptor called *Loci* [22], which generate an 8-length ternary code, named *Locus Number (LN)* [23] by the crossing counts on strokes in 8-directions from the character’s *CG (Center of Gravity)*. We convert the 8-bit ternary number (base = 3) to its corresponding decimal (base = 10).  $LN(CG) = \sum_{k=0}^{K-1} n \cdot 3^k$ ; where  $n$  is the number of intersections and  $n \in 0, 1, 2$ . Here,  $K = 8$ , since we inspect 8 directions from the *CG*. This  $LN(CG)$  is used as feature  $f_{19}$  for our task. We study this feature due to its obstruction capturing ability, traveling from the character’s *CG* towards the periphery.

g) *Effort-measure*: The *effort* of scribbling a character stroke pattern may be considered the summation of efforts from one pixel to the next. Let,  $e$  be the effort-measure if the stroke is drawn from the one left pixel to the next right pixel through the horizontal axis. From a central pixel, the 8-directional effort measure, adjusted empirically, is shown in Fig. 5(b).

Adding effort-measures ( $e$ ) of all object pixels, let the sum be  $E \times e$ , where  $E$  is a natural number. The *effort-measure* of a character is  $(E/N)e$ , which undertakes normalization by  $N$ . This  $N$  is a scale insensitive quantity, calculated as  $N =$

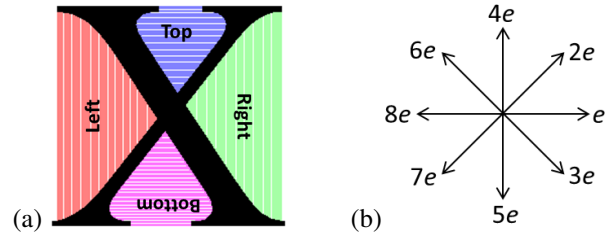


Fig. 5. (a) Concave water reservoirs: Top, Bottom, Left, and Right reservoirs are shown by the color blue, magenta, red, and green, respectively. (b) Effort measure in 8-directions from a central pixel.

$N_{BB} \times E_{BB}$ , where  $N_{BB}$  is the total number of black pixels in the character bounding box ( $BB$ ) and  $E_{BB}$  is the elongation of the  $BB$ ;  $E_{BB} = \max(H_{BB}, W_{BB}) / \min(H_{BB}, W_{BB})$ , where,  $H_{BB}$  and  $W_{BB}$  represent the height and width of the  $BB$ . The effort measure  $(E/N)e$  can be used as a feature having an impact on reading/writing analysis. Let us assume  $e$  to be 1, and compute the feature  $f_{20} = (E/N)$ .

In this manner, we obtain 20 single-valued features from a Bengali conjunct character.

2) *SVM classifier*: The feature vector of size 20 is fed to an SVM classifier to categorize a character corresponding to its reading/writing difficulty opinion score. The SVM usually works well for multi-class classification in a wide range of pattern recognition applications [24]. Concerning SVM multiclass classification for handwriting-related tasks, the *one-against-all* strategy works better than the *one-against-one* [25]. It can also be noted from [26], [27] that the SVM with an RBF (*Radial Basis Function*) kernel [28] works well compared to some other classifiers such as k-NN (*k-Nearest Neighbors*), MLP (*Multi-Layer Perceptron*), MQDF (*Modified Quadratic Discriminant Function*), and SVM-linear while dealing with Bengali characters. Hence, we use one-against-all SVM-RBF for our task.

The SVM-RBF hyper-parameters ( $\gamma$  and  $C$ ) are essential to be tuned to regulate the decision boundary and to avoid overfitting [29]. For the optimal performance of the classifier, the hyper-parameters are selected from a tuning set, called the *model selection* step. We use the traditional *grid-searching* technique for this model selection [28]. Here, *k-fold cross-validation* is used.

The best performance is obtained for  $\gamma = 2^4$  within the range  $[2^{-3}, 2^{-2}, \dots, 2^6]$  and  $C = 2^3$  within the range  $[2^{-3}, 2^{-2}, \dots, 2^7]$ . Here, we use 5-fold cross-validation.

#### IV. EXPERIMENTS AND DISCUSSION

In this section, at first, we discuss the employed database for experimentation.

##### A. Employed Database

For the experimentation, we required a database having a subjective score of difficulties arising in character reading/writing. We did not find any standard and publicly available databases with such ground-truth. Therefore, we generated our own corpus containing subjective opinion ground-

truth scores of complications arising in Bengali conjunct reading/writing. For this, we gathered Bengali conjunct characters, on which we asked volunteers to put some subjective scores while reading and writing separately, as discussed in Section III-A.

For both the reading and writing score, we selected the range [0, 10], i.e.  $r_1 = 0$ ,  $w_1 = 0$ ,  $r_2 = 10$ ,  $w_2 = 10$  (refer to Section III-A). This range was partitioned into 20 bins, i.e.  $n_r = 20$ ,  $n_w = 20$ . Therefore, both the reading and writing difficulty analysis tasks can be perceived as 20-class classification problems. The idea behind such fixing of the class number is adopted from [2]. The class-1 is represented by the lowest score, i.e. most easy in reading/writing, and class-20 as the highest score, i.e. most difficult in reading/writing.

For each Bengali conjunct character, at least 100 human volunteers provided subjective reading and writing difficulty opinion scores, and the arithmetic mean opinion score was selected as the gold standard for each character.

To obtain the reading opinion score, a volunteer was shown the cropped Bengali conjunct characters arbitrarily, since sequential characters shown from a text-line/word may hint at the actual content due to the semantic knowledge. For delivering the reading score, the human readers were invited to see the character image at a distance of 12-18 inches approximately.

The writing opinion score was also collected at the same time while asking to copy (write) the conjunct character in an offline mode. All volunteers were requested to use a general 0.5 mm ball-point smooth black-inked pen of a particular brand/model to write on a 75 GSM ( $g/m^2$ ) blank white page placed on a common good writing surface.

A continuous subjective scoring may exhaust human eyes and may deteriorate the sensitivity of the human cognitive system. To alleviate this, the volunteer was given a 5 to 10 minutes break after providing reading and writing scores of 50 characters.

The reading/writing scores were collected from native Bengali persons who were able to read and comprehend Bengali writing. The volunteers were of both genders in the age group of 14 to 66 years, having various academic backgrounds and passed at least school grade VIII. The volunteers were generally healthy human-beings without having any known serious illnesses (e.g., Dyslexia [6], Parkinson’s disease etc.) which may obstruct reading/writing skill. All the scores were collected during the healthy, awakened state of the volunteers.

For subjective opinion score collection, we employed two commonly used Bengali printed fonts and three databases of Bengali contemporary handwriting; those are discussed as follows.

(i) *Bangla Akademi* (say,  $D_{BA}$ ): This printed font has been developed by *Paschimbanga Bangla Akademi*, India. In this font, most of the Bengali conjunct characters are made simple for being in the transparent form [1] (refer to Fig. 1). We typed 220 usual conjunct characters using this font to generate database  $D_{BA}$ .

(ii) *AdorshoLipi* (say,  $D_{AL}$ ): This printed Bengali font contains traditional non-transparent conjunct characters [1] (refer

to Fig. 1). A total of 220 conjunct characters were typed in this font to generate database  $D_{AL}$ .

(iii) *NewISIdb: HwC* [10] (say,  $D_C$ ): A total of 133 distinct types of Bengali conjunct characters written by 100 writers are present in this database.

(iv) *NewISIdb: HwW* [11] (say,  $D_W$ ): In this database, 111 words containing Bengali conjuncts, written by 100 writers are present.

(v) *NewISIdb: HwP\_Conjunct* [11] (say,  $D_{PC}$ ): This database contains 50 handwritten copies of a text, contributed by 50 writers. Each copy contains 587 words containing Bengali conjunct characters.

The conjunct characters of the databases  $D_C$ ,  $D_W$ ,  $D_{PC}$  are separated if they belong to a word. This separation is performed using a semi-automatic approach, since such conjuncts are the focus of this paper. A volunteer was shown at least 30 individual handwritten samples of each handwritten conjunct.

## B. Results and Evaluation

In this subsection, the experimental results are presented to analyze the performance of our system.

1) *Auto-derived Feature-based Analysis*: We conducted our experiment separately on the previously mentioned five databases  $D_{BA}$ ,  $D_{AL}$ ,  $D_C$ ,  $D_W$ , and  $D_{PC}$  (refer to Section IV-A). For all the databases, the training, validation and test sets were partitioned in the ratio of 2:1:1. These training, validation and test sets contained unique conjunct characters with no overlap.

The performance of auto-derived feature-based analysis employing an Inception network on the test set, with respect to the F-Measure (FM), is presented in TABLE I.

TABLE I  
AUTO-DERIVED FEATURE-BASED ANALYSIS

Database	F-Measure (FM) %	
	Reading	Writing
$D_{BA}$	92.85	87.47
$D_{AL}$	84.38	80.92
$D_C$	86.67	83.33
$D_W$	88.18	85.94
$D_{PC}$	76.02	73.39
$D_{BA} + D_{AL}$	86.58	81.65
$D_C + D_W + D_{PC}$	77.23	75.84

We obtained the highest performance on databases  $D_{BA}$  and the lowest performance on  $D_{PC}$ . We also tested on the overall printed database and overall handwritten database, separately. The performance on the total printed databases was better than the total handwritten database. On the total printed (handwritten) database, we obtained 86.58% (77.23%) and 81.65% (75.84%) FM in reading and writing analysis, respectively. Overall, the reading analysis performance was better than the writing analysis.

2) *Hand-crafted Features with SVM-based Analysis*: In this experiment also, we used the training, validation and test set setup of the auto-derived feature-based analysis.

The performance analysis employing a hand-crafted feature-based SVM on the test set is shown in TABLE II.

TABLE II  
HAND-CRAFTED FEATURE-BASED ANALYSIS

Database	F-Measure (FM) %	
	Reading	Writing
$D_{BA}$	86.28	82.94
$D_{AL}$	78.65	75.54
$D_C$	79.78	77.83
$D_W$	82.40	80.28
$D_{PC}$	69.27	64.65
$D_{BA} + D_{AL}$	81.19	77.04
$D_C + D_W + D_{PC}$	72.62	69.86

Here also, the highest performance was obtained on  $D_{BA}$ , and the lowest was on  $D_{PC}$ . The writing analysis performance was worse than the performance for the reading analysis case. On the total printed database, we obtained 81.19% and 77.04% FM in reading and writing analysis, respectively. For the total handwritten database, these reading and writing performances were 72.62% and 69.86% FM, respectively.

In Fig. 6, we show the overall barcode representation of reading and writing difficulty analysis.

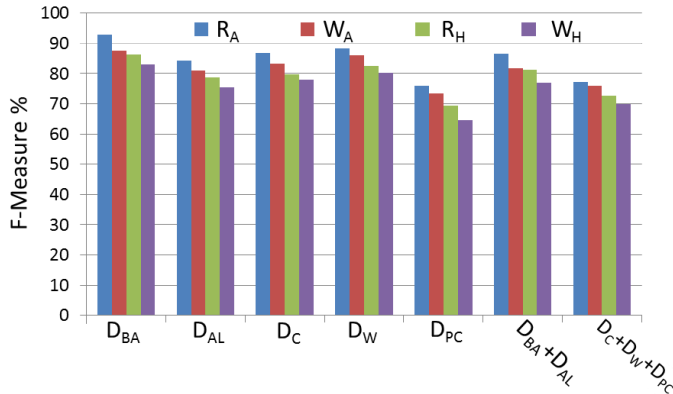


Fig. 6. Barcode representation of reading/writing difficulty analysis.  $R_A$  and  $W_A$  denote reading and writing analysis using Auto-derived features,  $R_H$  and  $W_H$  represent reading and writing analysis using Hand-crafted features, respectively.

Overall, the auto-derived features worked better than the hand-crafted features. The performance on  $D_{BA}$  was better, since it contained mostly the transparent printed conjuncts. The  $D_{PC}$  showed poor performance owing to the fact that it contained most of the handwritten complex conjunct characters.

### C. Correlation Testing of Conjunct Character Reading and Writing Difficulties Analysis

Sometimes, the tasks of reading and writing analysis can be thought of as a single problem. Therefore, it is essential to know whether reading and writing complications of Bengali conjuncts correlate. Based on the reading/writing subjective opinion score ranking on conjunct characters, we perform the “*Spearman’s rank correlation*” [30] testing and obtain the coefficient value as 0.43. This smaller value signifies lower correlation between reading and writing effort for our task.

However, the positive coefficient value denotes that the reading and writing difficulty analysis tasks are not fully opposed.

### D. Comparison

To the best of our searching capacity in both online and offline media, we have not found any papers related to this work for comparison purposes. The work of [1] is very preliminary and mainly a philosophical analysis of remembering Bengali characters, with a lack of proper ground-truthing and experimental setup to perform repeatedly.

## V. CONCLUSION

In this paper, we worked on Bengali conjunct character reading and writing complication analysis with respect to the subjective cognitive score. The reading and writing analysis tasks were formulated as feature-based classification problems. The auto-derived and hand-crafted features were used for both reading and writing analysis. For experimental analysis, we collected human cognitive scores of difficulties arising in printed and handwritten Bengali conjunct reading/writing and generated a corpus. By employing auto-derived features on the overall printed (handwritten) corpus, we obtained 86.58% (77.23%) and 81.65% (75.84%) FM in reading and writing analysis, respectively. We also obtained 81.19% (72.62%) and 77.04% (69.86%) FM in reading and writing analysis, respectively, while using the hand-crafted features on the overall printed (handwritten) corpus. Although we worked on Bengali conjuncts, our method can be extended to the similar problem of some other scripts. In future, we shall try to exploit the reading and writing cognitive complexities in multilingual environments.

## ACKNOWLEDGMENT

We heartily thank all the volunteers for their help in database ground-truth generation. Partial support to one of the co-authors (B. B. Chaudhuri) by INAE and SERB, New Delhi, India, is gratefully acknowledged.

## REFERENCES

- [1] B.B. Chaudhuri, “Learning an Indian Abugida Script : Bangla”, Proc. International Graphonomics Society (IGS) Conference, pp. 22-25, 2011.
- [2] Q.H.-Thu, M.N. Garcia, F. Speranza, P. Corveau, A. Raake, “Study of Rating Scales for Subjective Quality Assessment of High-Definition Video”, IEEE Trans. on Broadcasting, vol. 57, no. 1, pp. 1-14, 2011.
- [3] S.L. Chou, S.S. Yu, “Sorting Qualities of Handwritten Chinese Characters for Setting Up a Research Database”, Proc. Int. Conf. on Document Analysis and Recognition (ICDAR), pp. 474-477, 1993.
- [4] B.V.B. Donsky, “Trends in Elementary Writing Instruction, 1900-1959”, Language Arts, vol. 61, no. 8, pp. 795-803, 1984.
- [5] C.B. Olson, R. Land, “A Cognitive Strategies Approach to Reading and Writing Instruction for English Language Learners in Secondary School”, Research in the Teaching of English, vol. 41, no. 3, pp. 269-303, 2007.
- [6] A.W. Ellis, “Reading, Writing and Dyslexia: A Cognitive Analysis”, Psychology Press, New York, 1993.
- [7] M. Taouk, M. Coltheart, “The Cognitive Processes Involved in Learning to Read in Arabic”, Reading and Writing: An Interdisciplinary Journal, vol. 17, pp. 27-57, 2004.
- [8] R.D. Banerji, “The Origin of the Bengali Script”, University of Calcutta, 1919.
- [9] B.B. Chaudhuri, U. Pal, “A Complete Printed Bangla OCR System”, Pattern Recognition, vol. 31, no. 5, pp. 531-549, 1998.

- [10] C. Adak, B.B. Chaudhuri, "Writer Identification from Offline Isolated Bangla Characters and Numerals", Proc. Int. Conf. on Document Analysis and Recognition (ICDAR), pp. 486-490, 2015.
- [11] C. Adak, B.B. Chaudhuri, M. Blumenstein, "Offline Cursive Bengali Word Recognition using CNNs with a Recurrent Model", Int. Conf. on Frontiers in Handwriting Recognition (ICFHR), pp. 429-434, 2016.
- [12] R. Likert, "A Technique for the Measurement of Attitudes", Archives of Psychology, New York, no. 140, pp. 1-55, 1932.
- [13] C. Adak, B.B. Chaudhuri, M. Blumenstein, "Legibility and Aesthetic Analysis of Handwriting", Proc. Int. Conf. on Document Analysis and Recognition (ICDAR), pp. 175-182, 2017.
- [14] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based Learning Applied to Document Recognition", Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, 1998.
- [15] C. Szegedy et al., "Going Deeper with Convolutions", Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 1-9, 2015.
- [16] V. Nair, G.E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines", Proc. Int. Conference on Machine Learning (ICML), pp. 807-814, 2010.
- [17] M. Lin, Q. Chen, S. Yan, "Network in Network", CoRR, abs/1312.4400, 2013.
- [18] N. Srivastava, G.E. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting", Journal of Machine Learning Research (JMLR), vol. 15, no. 1, pp. 1929-1958, 2014.
- [19] H. Hilton, "Plane Algebraic Curves", Chap.II, pp.18-36, Oxford, 1920.
- [20] C. Adak, B. B. Chaudhuri, M. Blumenstein, "Writer Identification by Training on One Script but Testing on Another", Proc. Int. Conference on Pattern Recognition (ICPR), pp. 1148-1153, 2016.
- [21] U. Pal, S. Sinha and B. B. Chaudhuri, "Multi-Script Line Identification from Indian Documents", Proc. Int. Conf. on Document Analysis and Recognition (ICDAR), pp. 880-884, 2003.
- [22] H.A. Glucksman, "Classification of Mixed-font Alphabets by Characteristic Loci", Annual IEEE Computer Conference, pp.138-141, 1967.
- [23] D.F. Mota, P. Riba, A. Forns, J. Llads, "On the Influence of Key Point Encoding for Handwritten Word Spotting", Proc. Int. Conf. on Frontiers in Handwriting Recognition (ICFHR), pp. 476-481, 2014.
- [24] Y. Ma, G. Guo, "Support Vector Machines Applications", Springer International Publishing Switzerland, 2014.
- [25] J. Milgram, M. Cheriet, R. Sabourin, " "One Against One" or "One Against All": Which One is Better for Handwriting Recognition with SVMs?", Proc. Int. Workshop on Frontiers in Handwriting Recognition (IWFHR), 2006.
- [26] C.L. Liu, C.Y. Suen, "A New Benchmark on the Recognition of Handwritten Bangla and Farsi Numeral Characters", Pattern Recognition, vol. 42, no. 12, pp. 3287-3295, 2009.
- [27] B.B. Chaudhuri, C. Adak, "An Approach for Detecting and Cleaning of Struck-out Handwritten Text", Pattern Recognition, vol. 61, pp. 282-294, 2017.
- [28] V.N. Vapnik, "The Nature of Statistical Learning Theory", Springer-Verlag, New York, 2000.
- [29] K. Duan, S.S. Keerthi, A.N. Poo, "Evaluation of Simple Performance Measures for Tuning SVM Hyperparameters", Neurocomputing, vol. 51, pp. 41-59, 2003.
- [30] J.L. Myers, A.D. Well, R.F. Lorch Jr., "Research Design and Statistical Analysis" Mahwah, N.J. : Lawrence Erlbaum Associates, 2003.