README

"BanglaHandwrittenCharSamples.txt"

This file consists of one sample from each of the underlying 209 classes. Thus, it has 209 non-blank lines and one such line is used for one sample. There are 4 entries in each line corresponding to sample ID, column size, row size and the pixel values respectively. Samples are binarized; 0 and 1 have been used to represent foreground and background pixels respectively. The first line of this file is:

0 39 36 111 **...** 1111

Here, '0' represents the sample ID, 39 represents the column size, 36 represents the row size and the sequence '111 **...** 1111' (consisting of 0's and 1's) represents all of its pixels. If this sequence is displayed in an appropriate two-way arrangement, it looks like:

```
111111111111111111111111111111111111111
111111111111111111111111111111111111111
111111111111111111111111111111111111111
111111111111111100000111111111111111111
111111111111000000000000011111111111111
111111111100000000000000000111111111111
111111111000000000000000000001111111111
111111110000000000001111000000111111111
111111110000000000011111100000011111111
111111100000111111111111111000001111111
111111000000111111111111111100000111111
111111000011111111111111111110000011111
111110000011111111111111111111000001111
111110000011111111111111111111100001111
111100000011111111111111111111110001111
111100001111111111111111111111110000111
111100001111111111111111111111111000111
111100001111111111111111111111111000111
111100001111111111111111111111111000011
111100001111111111111111111111111100011
111100000111111111111111111111111100011
111100000111111111111111111111111000011
111110000111111111111111111111111000011
111110000001111111111111111111111000011
111111000011111111111111111111111000011
111111000001111111111111111111110000011
111111100000011111111111111111110000111
111111110000000111111111111111111001111
111111111100000000111111111111100001111
111111111110000000000000000000000011111
111111111111100000000000000000001111111
111111111111111000000000000001111111111
111111111111111111111111000111111111111
111111111111111111111111111111111111111
111111111111111111111111111111111111111
111111111111111111111111111111111111111
```

The original image of this sample is shown below:

"Classes_labels.txt"

Each line of this file provides the ground truth of the sample given in the corresponding non-blank line of the file "BanglaHandwrittenCharSamples.txt". Thus, this file has 209 lines. There are 4 entries in each line of this file, which correspond to sample ID, class ID, Unicode and graphical transliteration of the corresponding entry in the sample file. The first line of this file is:

<div align="center">

0   0   09e6   0

</div>

Here, '0' represents the sample ID of the sample provided in the first line of the file "BanglaHandwrittenCharSamples.txt", next '0' represents its class ID, 09e6 is its Unicode and the last '0' is its graphical transliteration.

Finally, the graphical shapes of each of the underlying 209 numerals / characters are shown in the table of the following page. The class ID's of each character is shown within brackets against respective shapes.

Numerals :

Basic Characters :

Compound Characters :