**Geez number optical character recognition project**

**Scope and limitation**

The scope for this type of project can be classified like the following:

**Input image scope:**

To facilitate scanning of input hand written geez number character, the paper which holds the characters should have the following characteristics:

* The paper should be strong, white and clean to get clear scanned input characters.
* The paper size should preferably be A4, viz. 210 mm x 297 mm or 8 1/2 by 11 inches (which is the de-facto North-American standard).
* Sheets should be free of creases, holes and should not be rolled.
* The paper should not be absorbent in order to avoid smearing of the ink
* Double-sided writing of the paper should be avoided.
* The characters should be solid black on a white background.
* A minimum margin of 2 centimeters should be present at the top, bottom and sides of each sheet, and a minimum margin of 2.5 centimeters on the left side of each sheet.
* Each number character must write clearly and neatly with enough spacing between them.
* The written character must not be underlined for best clarity.
* The scanned image data must be converted to bmp format for the following reasons
* Bitmap graphics should be created or scanned either at, or very close to, the size required for the finished product.
* You can’t successfully change the size of the image in your word processor. You need to load the image into a suitable program and re-size it by re-sampling
* If you do need to re-size an image, avoid drastic changes in size.
* When re-sizing an image, always use a program that has some form of smart sizing feature and make sure that it is turned on.
* If you are creating a large image, you will need plenty of memory available
* The geez number which is written on the paper sheet must at list include the following numbers for training purpose:

፩ ,፪, ፫, ፬ ,፭, ፮ ,፯, ፰ ,፱,፲, ፳,፵, ፶,፷,፸,፹, ፺, ፻

The scanned image data must be pre processed before feed to the neural network system. This scanned image data must pre process by the following specification:

* The pre process is done by dividing the image using its pixel; each number character is divided horizontally in 50 pixels and vertically 70 pixels.
* Based on the divided pixel we must identify the shadow part form light shadow and empty parts.
* It will not accept values which are different from the above specified characters, if there is it will discard and work on other.

The input layer neurons are determined by the number input parameter which is sample for our condition as listed above like(፩ ,፪, ፫, ፬ ,፭, ፮ ,፯, ፰ ,፱,፲, ፳,፵, ፶,፷,፸,፹, ፺, ፻)

**Output text scope:**

* The output text must be understandable by the abanet, power geez, visual geez Unicode or other font type.
* The size must be 12 or 14 point as a default value.
* Then we can modify, update and edit the file