# Data Augmentation for Recognition of Handwritten Words and Lines using a CNN-LSTM Network

Team Name  - VISION

Archit Kumar 2018201051

Aman Agarwal 2018201006

Siddhesh Savant 2018201031

Abstract : implementation  of two data augmentation and normalization techniques, which, used with a CNN-LSTM, significantly reduce Word Error Rate (WER) and Character Error Rate (CER) beyond best-reported results on handwriting recognition tasks.

- **Profile Normalization** : Images are normalized to compensate for variations in the size of the handwriting.

- **Novel Grid-Based Distortion Augmentation** : Random warp grid distortion (RWGD) is performed as follows. (1) Place control points on a regular grid such that they align to the baseline. For IAM/RIMES we used a 26- pixel interval. (2) Perturb each control point in the x and y direction by randomly sampling from a normal distribution. (3) Warp the image according to the perturbed control points.

- **Test side Augmentation** : . Test-side augmentation is performed by (1) generating N augmented examples for each word/line image in the test set (N=20 in our experiments), (2) performing recognition on the N augmented images, and (3) choosing from the N predictions the one that produces the lowest CTC loss.

- **THE CNN-LSTM NETWORK :**

   Our CNN-LSTM network, uses 6 convolutional layers: 64, 128, 256, 256, 512, and 512 (3x3) filters respectively in the forward direction.

 Batch normalization is applied after the 4th and 5th layers.

Max pooling (2x2 window), stride 2 in both directions are applied after the 1st and 2nd layers. Max pooling (2x2 window) and vertical stride of 2 and horizontal stride of 1 is applied after the 4th and 6th layer.

Two BLSTM layers follow with 512 and 256 hidden nodes respectively with dropout rate of 0.5 before each.

A fully connected layer reduces the output to the character set size and a softmax is applied. It is trained using the CTC loss and the ADADELTA optimizer.

# DataSets :

- **IAM dataset** : The IAM Handwriting Database is a multi-author handwriting recognition database of 115,320 word-level images from 500 authors.

- **IIIT HWR dataset :** synthetically designed dataset by varying intercharacter spacing stroke width , foreground ,background pixel distribution .