OPTICAL CHARACTER RECOGNITION
(HINDI/ENGLISH CHARACTERS)

2018115049 - Kevin Paul
2018115109 - Srijeyarankesh
2018115134 - Vishwa Kumar

DETECTION : -

- The Detection mechanism is already implemented by the Craft algorithm.

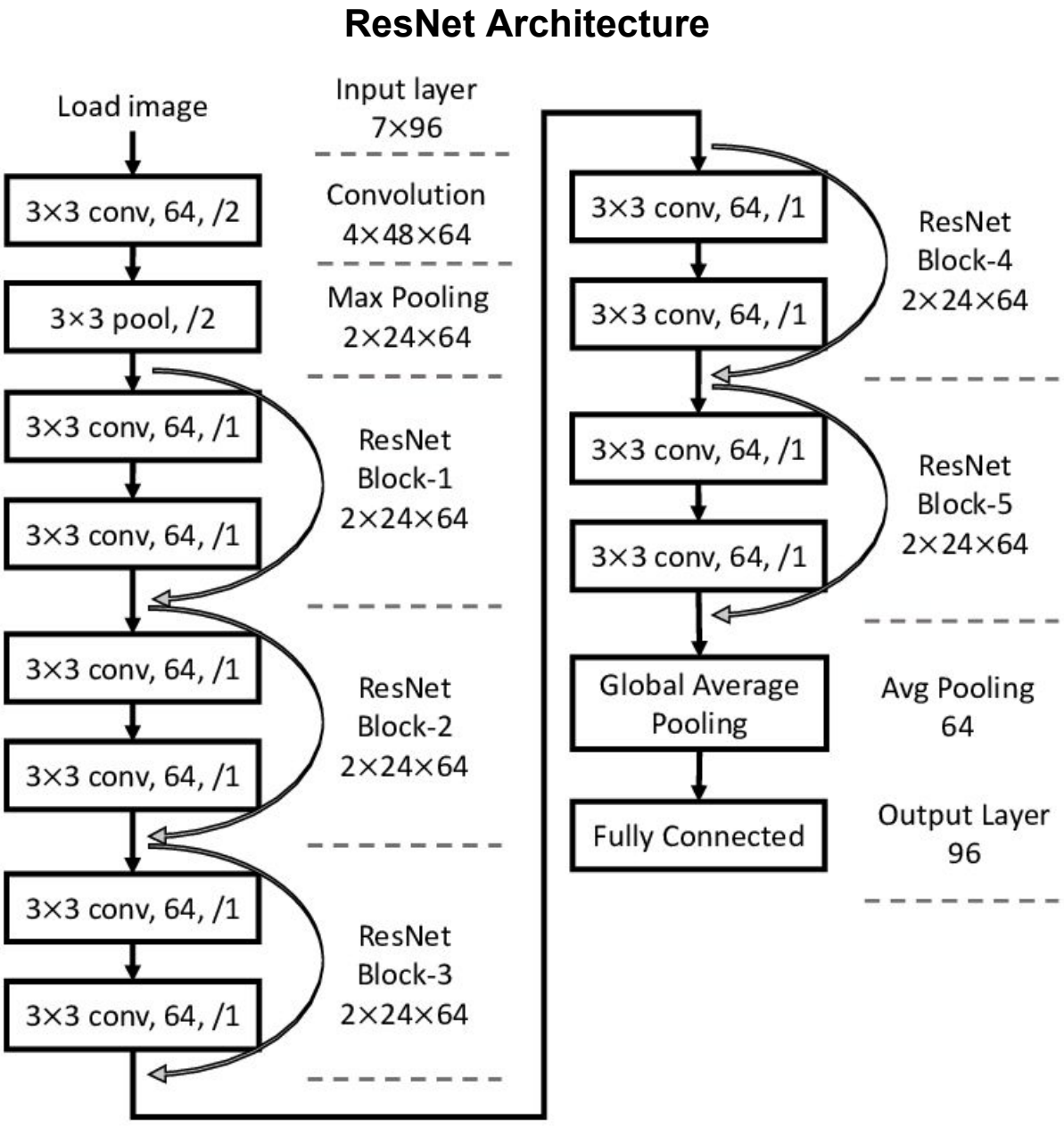
RECOGNITION : -

- After the bounding box has been drawn over the respective texts, we then have to predict what text is present inside the image.
- We thereby use CRNN architecture to identify the text and print it in plain text format from a jpg/png file format.
- The CRNN technique used is a combination of ResNet, LSTM and CTC loss function.
- Thus we have successfully created a deep learning model capable of identifying Hindi texts.

ResNet : -

The ResNet is a neural network used extensively in the field of image detection. ResNet network is implemented by using repeating sequences of double or triple layers that contain non-linear function(ReLU) and batch normalization. This model was proposed by Microsoft and emerged as the most successful model succeeding GoogleNet which was it's predecessor.

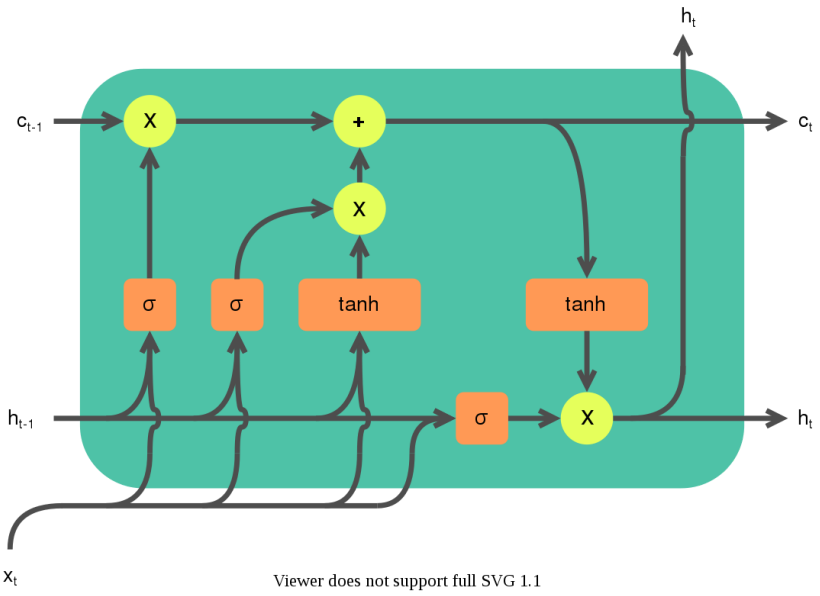
In our project we use ResNet so as to learn the parameters of the text bound by the bounding box in the image.



LSTM : -

Long short term memory cells are an enhanced form of RNNs capable of remembering long/lengthy sequences without trouble. Used in the field of sequence prediction they form the basis of encoders and decoders.

In our project, we use LSTMs to predict the textual data from the parameters learned by ResNet architecture. Thus every successive character is predicted based on the previous characters.



CTC loss function : -

Whenever there is sequence prediction involved the characters predicted don’t always produce the required result. Thus we need some form of equation which can identify the variation in

the actual output and predicted output so that the variation can be minimized. CTC comes into picture.

$$p(Y \mid X) = \sum_{A \in \mathcal{A}_{X,Y}} \prod_{t=1}^{T} p_t(a_t \mid X)$$

The CTC conditional **probability** **marginalizes** over the set of valid alignments computing the **probability** for a single alignment step-by-step.

Though the above equation may seem overwhelming, in our project we are going to use a library associated with the ctc loss function which is present in Tensorflow.