3d hand is predicted and mapped on the image and the background is removed. The dataset The image with black background just containing the lines of hand is sent to CNN for spatial prediction then the predicted softmax value in saved in deque data type in python. The queue length is 15. When the stack of 15 is full then softmax predicted list of 15 values is sent to LSTM for temporal feature prediction. Then after the temporal based prediction, each of predicted temporal value is saved in the Counter data type. It is maintained in the Counter data type limited to 10 counts in terms of for example softmax predicted values of ((0:2), (1:1), (2:6), (3:1), (4:4), (5:9), .... , (10:3)). The first value of each in the counter is ordered count value while the other value is predicted value. The Counter is limited to 10 lengths each. The counter is type of special type of List in Python. In the real time situation as the predicted spatial features is increasing, it means it has the highest probability in the maintained list of predicted temporal values. Its means it is the predicted value. The value is then mapped to the word it is corresponding to. The word is the predicted value. This word is then further sent to Sentence Post Processing for grammatical correction. The words sequence probability is predicted only on last 3 words