Small data set

Character Wise Accuracy

Model	Number of chars Matched	Total Number of Chars	Ratio
Model 1 - OCR Model	275	510	0.54
Model 2 - Transition Model	338	510	0.66
Model 3 - Combined Model	363	510	0.71

Word Wise Accuracy

Model	Number of Words Matched	Total Number of Words	Ratio
Model 1 - OCR Model	9	104	0.087
Model 2 - Transition Model	27	104	0.26
Model 3 - Combined Model	37	104	0.36

Average Dataset log-likelihood

Model	∑log-likelihood	Total Number of Words	Avg log-likelihood
Model 1 - OCR Model	-812.067	104	-7.808
Model 2 - Transition Model	-738.095	104	-7.097
Model 3 - Combined Model	-653.071	104	-6.279

We can clearly see the increase in accuracy in the three metrics as the model is changed from only OCR to OCR+Transition to OCR+Transition+Skip factor model.
All the words which are corrected by the trans model are listed in 0-1.txt and the words corrected by combined model but are partially corrected by trans and wrongly interpreted by ocr are given in 1-2.txt.

Large Dataset - Combined Model Results

Character-Wise Accuracy

Dataset	Number of chars Matched	Total Number of Chars	Ratio
1	7735	10919	0.708
2	7722	10919	0.707
3	7712	10919	0.706
4	7728	10919	0.7077
5	7760	10919	0.7083

Word-Wise Accuracy

Dataset	Number of Words Matched	Total Number of Words	Ratio
1	689	2188	0.3148
2	696	2188	0.3180
3	699	2188	0.3194
4	697	2188	0.3185
5	729	2188	0.3332

Avg-dataset log-likelihood

Dataset	∑log-likelihood	Total Number of Words	Avg log-likelihood
1	-13722.8	2188	-6.27186
2	-13721.6	2188	-6.2713
3	-13707.9	2188	-6.26506
4	-13714.2	2188	-6.26793
5	-13692.3	2188	-6.25789

I have runned for every large dataset using combined model dataset-5 is having large accuracy in terms of all the three metrics.The character accuracy ratios are higher compared to word accuracy as it is difficult to capture all the relativity between the character combinations that are forming with the model we are having.
To Compare between the three models I have runned it along the dataset-5 and the results are

OCR
Total chars:10919
Matched # of Chars:6391
Total words:2188
Matched # of Words:253
Avg Dataset log-likelihood:-7.857453254399122

OCR+Trans
Total chars:10919
Matched # of Chars:7475
Total words:2188
Matched # of Words:584
Avg Dataset log-likelihood:-7.158425135057614

Combined
Total chars:10919
Matched # of Chars:7760
Total words:2188
Matched # of Words:729
Avg Dataset log-likelihood:-6.257888800554419

we can clearly see the combined model has beat in all three metrics

I have tried to attempt extra credit question by increasing the skip factor what I have found is as if we give more weight to the skip factor the word and character wise have not changed but we can see an increase in average log-likelihood.which is true as expected as in skip factor we are giving weight if img and word are equal at a time.