

apl6\_olr continued

7\_olr, ct-cn, id-id, ja-jf, ko-kr, ru-ru, vi-vn, zh-cn  
 lang identification raw label: LSTM, 1024 cell  
 tdnn\_baseline: tdnn, (-4, -3, -2, -1, 0, 1, 2, 3, 4 0 -2, 2 0 -4, 4 0)  
 wsj-asr-lstm: from WSJ, LSTM, 1024 cell  
 wsj-asr-tdnn: from WSJ, tdnn, (-4, -3, -2, -1, 0, 1, 2, 3, 4 0 -2, 2 0 -4, 4 0)  
 lw-asr-tdnn: 16k, 7\*1200 relu  
 cnn: input window (-10, 10), filter(64), 2 layers

model	Cavg frame/utt	EER% frame/utt	log-likelihood /accuracy	conclusion about how to get long term memory
i-vector	svm: - 0.0383	svm: - 3.49		
tdnn_baseline	0.2599 0.1900	26.14 18.94	-0.1277 0.9758	
lstm_cell_1024	0.2467 0.1983	25.70 20.33	-0.1451 0.9548	
lstm-3_cell_1024	0.2295 0.1908	23.93 20.35	-0.0879 0.9703	multi-layer lstm and bilstm help a little
Baseline bilstm_cell_1025	0.2197 0.1787	22.73 18.13	-0.0635 0.9788	
gru_cell-1024				
cnn_lstm		<b>divergent</b>		cnn before lstm may not help
tdnn-lstm	0.2203 0.1685	23.19 17.75		
wsj-r3_delta0				
wsj-r4_delta0				
wsj-r5_delta0				
Neural i-vector wsj-r6_delta0 (best:lda+rbf)	null 0.0784	null 7.19		1. although not surpass the baseline, it shows that phonetic feat is informal. 2. a best configure for neural i-vector waits to be found, maybe bottleneck or with delta order 3.
wsj-r6_delta3				
lw-asr-tdnn-svd_as_lstm_input	0.1126 0.0522	11.74 6.32	-0.0604 0.9798	
asr-tdnn_r1-as-lstm-in	0.2008 0.1412	20.40 14.54	-0.0353 0.9885	
asr-tdnn_r2-as-lstm-in	0.1858 0.1245	19.50 13.42	-0.0537 0.9835	
PTN asr-tdnn_r3-as-lstm-in	0.1757 0.1039	18.25 11.12	-0.0926 0.9691	
asr-tdnn_r4-as-lstm-in	0.1712 0.0986	17.64 10.68	-0.1269 0.9563	
<b>asr-tdnn_r5-as-lstm-in</b>	<b>0.1598 0.0648</b>	<b>16.28 7.16</b>	<b>-0.1876 0.9383</b>	1. not the last layer gives the best performance, maybe nearer to the output loses much more info 2. the log-likelihood/accuracy monotonically decreasing, does phonetic feat avoid overfitting?
asr-tdnn_r6-as-lstm-in	0.1683 0.0697	17.08 7.57	-0.2283 0.9190	
long term PTN asr-tdnn_r6-as-lstm-in multi-label utt level label as target				