

ass\_geo

asr: combined softmax, LSTM, 1024 cell  
lang: raw label, LSTM, 1024 cell  
asr-lang: LSTM, 1024 cell, multi-task (asr & lid)  
asr-l: asr from ass  
swbd-asr: LSTM, 1024 cell

model	Cavg frame/utt	EER% frame/utt
i-vector		
lstm_lang_cell_1024	0.1545 0.0905	16.20 9.20
asr_as_front_lstm	0.0776 0.028	8.14 3.04
multitask_asr_lang	0.0822 0.0399	8.68 4.10
ASR posterior		
multitask_asr_lang_pst		
asr-lang_r_to_i	0.0762 0.0298	7.99 3.25
asr-lang_r_to_f	0.0765 0.0280	8.01 3.20
asr-lang_r_to_o	0.0752 0.0286	7.91 3.15
asr-lang_r_to_g	0.0726 0.0264	7.57 2.92
asr_r_to_o	0.0801 0.0298	8.38 3.28
asr_r_to_g	0.0795 0.0283	8.38 3.08
lang_r_to_o	0.1310 0.0638	13.67 7.00
lang_r_to_g	0.1279 0.0640	13.45 6.95
asr-l_r_to_o	0.0942 0.0375	9.81 4.10
asr-l_r_to_g	0.0900 0.038	9.49 4.05
swbd-asr_r_to_o		
swbd-asr_r_to_g	0.0970 0.0385	10.18 4.10

ass\_geo

lang: raw label, LSTM, 1024 cell  
asr-lang: tdnn, (2048->256)\*6, multi-task (asr & lid), (-4, -3, -2, -1, 0,

model	Cavg frame/utt	EER% frame/utt
lstm_lang_cell_1024	0.1545 0.0905	16.20 9.20
multitask_asr_lang	0.0822 0.0399	8.68 4.10
tdnn_asr-lang	0.1064 0.0355	11.17 3.75
tdnn ASR posterior		
tdnn_asr_lang_pst		
tdnn_asr-lang_last_to_o		

tdnn asr-lang last to g 0.0435 0.0214 4.45 2.40

---

**ben\_tur**

lang raw label: LSTM, 1024 cell

asr-lang: LSTM from ass\_geo

model	Cavg frame/utt	EER% frame/utt
lstm_lang_cell_1024	0.1740 0.0813	17.47 8.65
asr-lang_r_to_g	0.0869 0.0194	8.70 2.15

---

**ass\_ben\_geo\_tur**

lang raw label: LSTM, 1024 cell

asr-lang: LSTM from ass\_geo

model, output dim 4	Cavg frame/utt	EER% frame/utt
lstm_lang_cell_1024	0.2131 0.1499	21.34 15.75
asr-lang_r_to_g	0.1531 0.0833	15.43 9.27

---

**7\_all, ass\_ben\_can\_geo\_pas\_tag\_tur**

lang raw label: LSTM, 1024 cell

asr-lang: LSTM from ass\_geo

asr-lang-tdnn: **tdnn** from ass\_geo, (-4, -3, -2, -1, 0, 1, 2, 3, 4 0 -2, 2 0

swbd-asr: LSTM, 1024 cell

model	Cavg frame/utt	EER% frame/utt
lstm_lang_cell_1024	0.1987 0.1249	19.11 12.63
asr-lang_r_to_g	0.1454 0.0654	14.12 7.11
asr-lang-tdnn_r_to_g	0.1280 0.062	12.27 6.66
swbd-asr_r_to_g	0.1623 0.0803	15.62 8.34

---

7\_olr, ct-cn, id-id, ja-jf, ko-kr, ru-ru, vi-vn, zh-cn

lang raw label: LSTM, 1024 cell

asr: from WSJ, LSTM, 1024 cell

tdnn\_baseline: tdnn, (-4, -3, -2, -1, 0, 1, 2, 3, 4 0 -2, 2 0 -4, 4 0)

asr-tdnn: from WSJ, tdnn, (-4, -3, -2, -1, 0, 1, 2, 3, 4 0 -2, 2 0 -4, 4 0)

model	Cavg frame/utt	EER% frame/utt
lstm_lang_cell_1024	0.1983 / 0.2467	20.33 / 25.70
lstm-3_baseline		
tdnn_baseline		
asr_r_to_g	0.1961 / 0.1432	20.50 / 14.82
asr-tdnn_r_to_g	0.1857 / 0.1183	18.99 / 11.96
tdnn_to_tdnn-affine6		
tdnn to lstm-3_g123		

datasets	#frames
ass	2694, 8168
ben	3134, 1111
can	6307, 5828
geo	<b>2299, 7268</b>
pas	4000, 8931
tag	4165, 2554
tur	<b>3867, 3022</b>

WER%  
on test1/test2

68.20 / 65.50

68.90 / 66.00

1, 2, 3, 4 0 -2, 2 0 -4, 4 0)

WER%  
on test1/test2

66.4 / 64.4

-4, 4 0)



