	Aishell-1			HKUST		LibriSpeech100					Hub5' 00		
	#Params(M)	dev	test	#Params(M)	dev	#Params(M)	dev clean	dev other	test clean	test other	#Params(M)	swbd	callhm
Transformer	12.47	5.7	6.3	12.47	23.5	12.47	10.3	23.0	11.1	24.1	12.47	8.4	17.0
Conformer	13.06	5.3	5.9	13.06	21.7	13.06	9.3	21.7	9.5	22.1	13.06	8.1	16.2
Lite Transformer	11.31	5.6	6.2	11.31	21.6	11.31	9.1	21.6	9.3	22.4	11.31	8.2	16.5
Random Search	12.20	5.2	5.9	12.23	22.4	12.57	10.3	22.8	10.7	23.7	12.35	8.3	16.5
SCA(η=0.1)	10.41	**4.8**	**5.2**	10.88	**21.0**	10.55	**9.0**	**21.3**	**9.2**	**21.5**	12.35	**7.9**	**15.7**

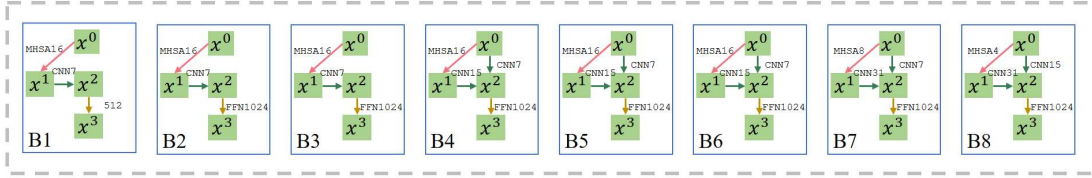In our paper, we show the results(CER/WER) of the human-designed baselines and SCA over four benchmarks.

For Conformer and Transformer, we set the attention dim as 256, head number as 4, and hidden dimension as 2048.

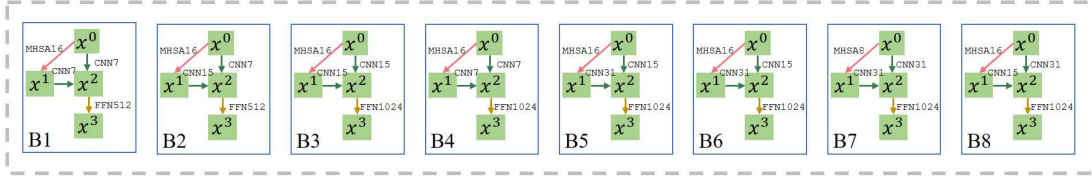For Lite Transformer, we set the attention dim as 128, head number as 4, and hidden dimension as 512.

For the general decoder, we set the attention dim as 256, head number as 4, and hidden dimension as 2048.

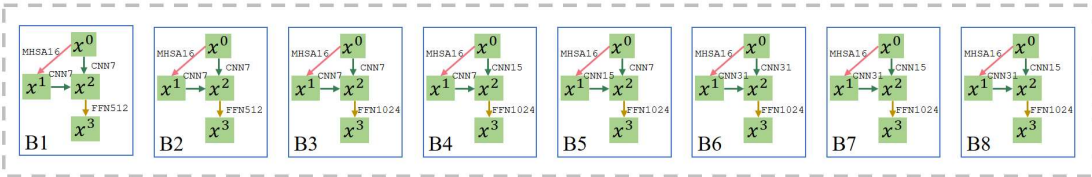For SCA, we set η=0.1 by default and show the searched architectures on four datasets as follow:
Aishell-1:

HKUST:

LibriSpeech100:

SWBD: