

Parameter	
Audio resampling freq.	16 KHz
Bands of Mel-spectrogram	80
Hop length	400
Convolution layers, channels, filter, strides	1, 64, 20×5 , 8×2
Recurrent layer size	128
Fully connected size	128
Dropout probability	0.9
Learning Rate	10^{-3}
Max gradient norm	100
Gradient clipping max. value	5