

DLHLP HW3 Report

組長 Github ID: lou-tun-chieh

學號：b05902111 系級：資工四 姓名：婁敦傑

學號：b05902010 系級：資工四 姓名：張頌平

1. (5%)請記錄 evaluate.log 裡面的SiSNR 數值，和當時所用的hyperparameter(這一題請3-1不用PIT, 3-2用PIT)

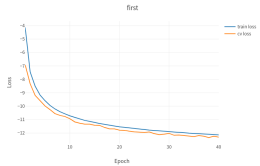
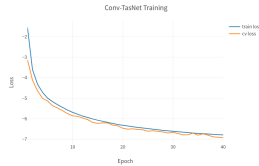
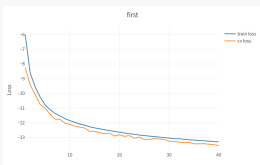
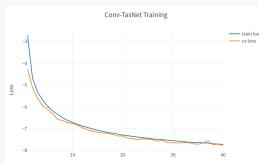
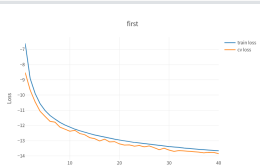
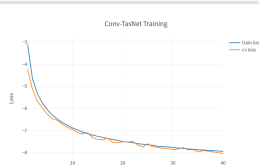
•		3-1	3-2
	N	64	256
	L	16	16
	B	64	128
	H	128	128
	P	3	3
	X	8	8
	R	3	3
	norm_type	gLN	gLN
	causal	0	0
	mask_nonlinear	relu	relu
	C	2	2
	epoch	60	60
	Average SISNR	Average SDR improvement: 19.02 Average SISNR improvement: 18.84	Average SDR improvement: 10.66 Average SISNR improvement: 10.15

2. (5%)嘗試調整不同的hyperparameter，比較其差異，並試著分析結果(至少針對2種不同的hyperparameter進行實驗)

- 3-1: L = 16, B = 32, H = 64, P = 3, X = 7, R = 1, epoch = 40, pit = 0

3-2: L = 16, B = 32, H = 64, P = 3, X = 7, R = 1, epoch = 40, pit = 1

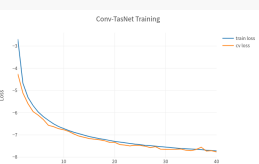

隨著N的變大average SISNR是有變好的，雖然有些變好的幅度不是很大，但loss都是明顯變低的，N的增加，代表encoder和decoder的basis signals也增加了， 有更多的basis signals訓練有使結果變的更好，paper提到當N增加同時增加了 overcompleteness of the basis signals因而使結果變好，我所作的實驗也是符合paper所述。

N	3-1 loss learning curve	3-1 Average SISNR	3-2 loss learning curve	3-2 Average SISNR
32		Average SDR Improvement: 12.58 Average SISNR Improvement: 12.16		Average SDR Improvement: 5.79 Average SISNR Improvement: 5.46
64		Average SDR Improvement: 13.68 Average SISNR Improvement: 13.47		Average SDR Improvement: 6.52 Average SISNR Improvement: 6.13
128		Average SDR Improvement: 14.02 Average SISNR Improvement: 13.67		Average SDR Improvement: 6.97 Average SISNR Improvement: 6.32

- 3-1: N = 64, L = 16, B = 32, H = 64, P = 3, R = 1, epoch = 40, pit = 0

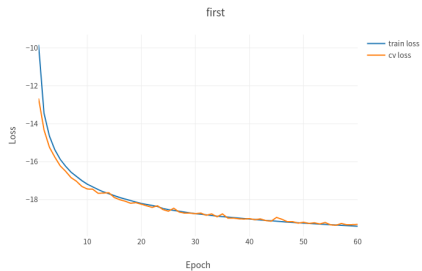
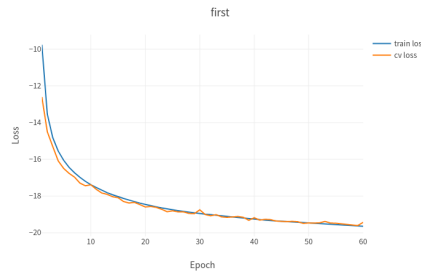
3-2: N = 64, L = 16, B = 32, H = 64, P = 3, R = 1, epoch = 40, pit = 1

當X變大，不論是loss還是average SISNR都是明顯的變好，增加convolution block的數量，使模型參數變多模型變得更複雜，這使結果變得更好，paper寫到可能是因為增加了model capacity而改善performance，實驗結果也是符合paper所述。

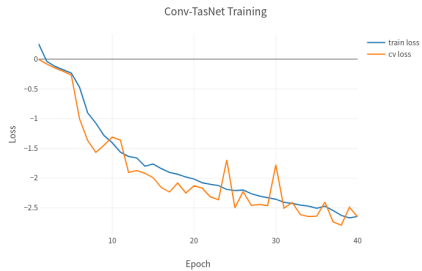
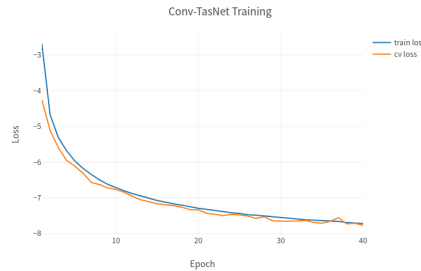
X	3-1 loss learning curve	3-1 Average SISNR	3-2 loss learning curve	3-2 Average SISNR
4		Average SDR Improvement: 12.44 Average SISNR Improvement: 12.26		Average SDR Improvement: 5.85 Average SISNR Improvement: 5.53
7		Average SDR Improvement: 13.68 Average SISNR Improvement: 13.47		Average SDR Improvement: 6.52 Average SISNR Improvement: 6.13
10		Average SDR Improvement: 14.00 Average SISNR Improvement: 13.67		Average SDR Improvement: 7.02 Average SISNR Improvement: 6.63

3. (3%) 3-1, 3-2請分別試看看有無PIT的差異並記錄結果(loss learning curve, Si-SNR).

- 3-1: N = 64, L = 16, B = 64, H = 128, P = 3, X = 8, R = 3, epoch = 60

	pit = 0	pit = 1
loss learning curve		
Average SISNR	<p>Average SDR improvement: 19.02 Average SISNR improvement: 18.84</p>	<p>Average SDR improvement: 19.91 Average SISNR improvement: 19.72</p>

- 3-2: N = 64, L = 16, B = 32, H = 64, P = 3, X = 7, R = 1, epoch = 40

	pit = 0	pit = 1
loss learning curve		
Average SISNR	<p>Average SDR improvement: 3.97 Average SISNR improvement: 2.10</p>	<p>Average SDR improvement: 6.52 Average SISNR improvement: 6.13</p>

4. (2%) 思考一下為何有無PIT會影響3-1, 3-2的結果並寫下你的看法

- 3-1 是 speaker dependent 因此在沒有使用 pit 的情況下，SISNR也是很好的可以成功分離聲音，在有 pit 的情況下，loss 和 SISNR 是有進步的，我們認為可能是因為 data 有 noise 或語者的聲音特徵很相像，造成對應 groundtruth 計算出的 loss 不是最小的，因此用 pit 可以使模型得到最小的 loss，使模型得到比較好的結果，我們認為 speaker dependent 使用 pit 會稍微比沒有 pit 再進步一點。
- 3-2 是 speaker independent 如果不使用 pit，那在對應 groundtruth 計算 loss 時會因為多個語者不同特性，造成 permutation issue，沒有辦法訓練出一個好的模型，實驗結果也很明顯，如果沒有使用 pit，SISNR 是非常差的無法分離聲音。

Bonus: (2%)

Nothing