Model

cvt_io_v1_1 :

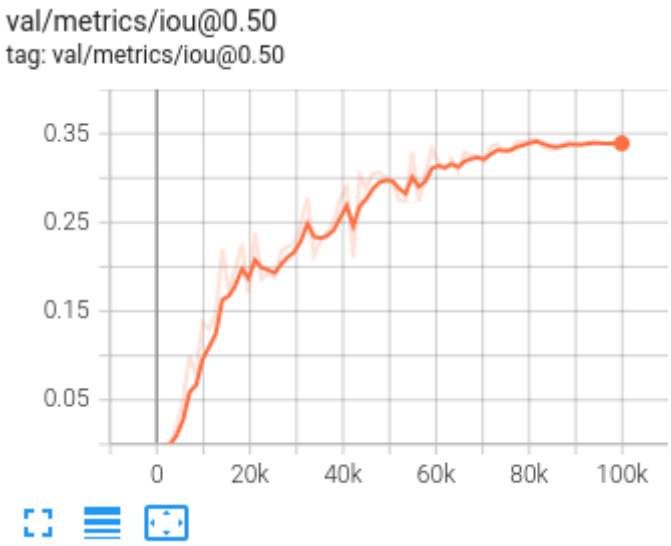
latent array： N = 10， L = 6， self attention shape：[b, n, d, H, W]

100K steps, lr: 4e-3, wd: 1e-7

cvt_io_v2_1 :

latent array： N = 10， L = 6， self attention shape：[b, d, H, W]

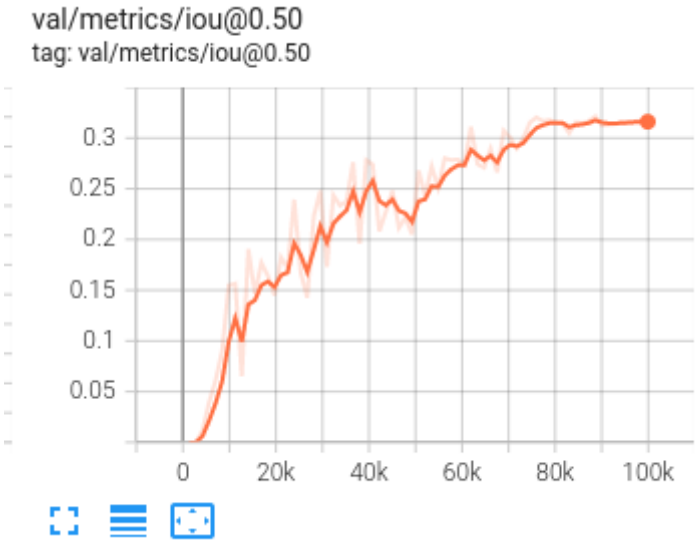
100K steps, lr: 4e-3, wd: 1e-7



cvt_io_v2_2 :

latent array： N = 5， L = 6， self attention shape：[b, d, H, W]

100K steps, lr: 4e-3, wd: 1e-7

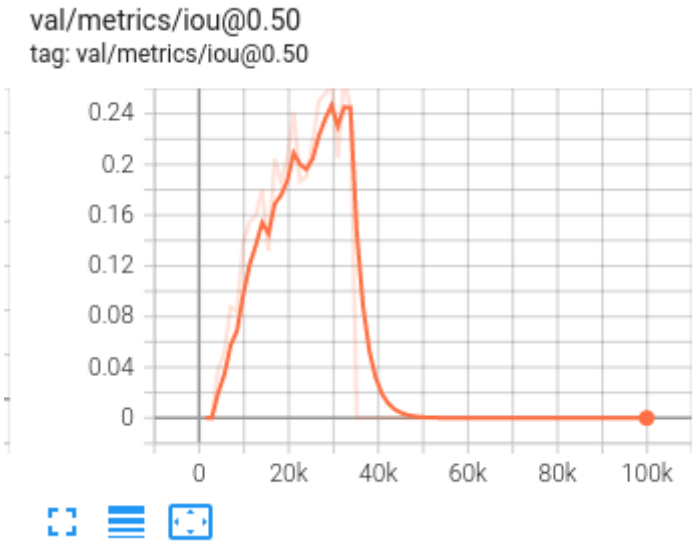


cvt_io_v2_3:

latent array： N = 10，L = 6，self attention shape：[b, d, H, W]，self attention不共享权值

100K steps, lr: 4e-3, wd: 1e-7

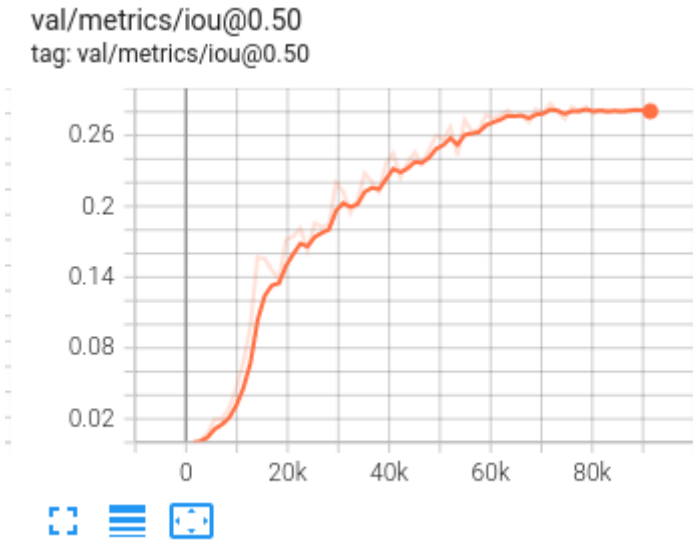
梯度爆炸，loss突然暴涨，不收敛



cvt_io_v2_4:

latent array： N = 10，L = 6，self attention shape：[b, d, H, W]，self attention不共享权值

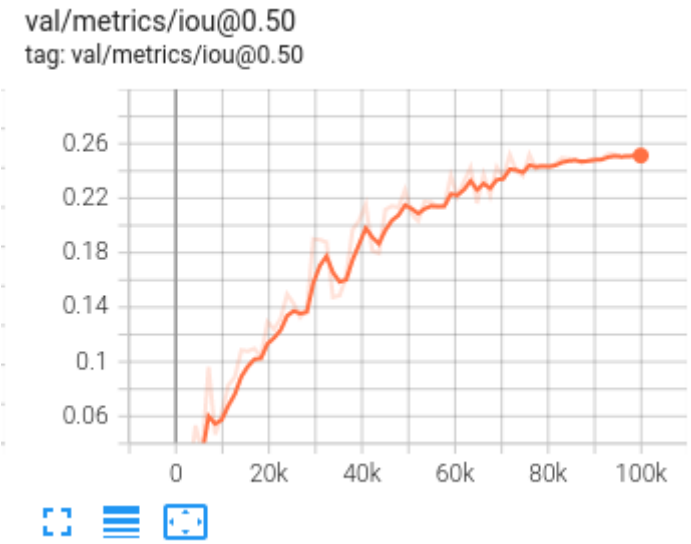
100K steps, lr: 1e-3, wd: 1e-7



cvt_io_v2_5:

latent array： N = 3，L = 1（实验设置出错），self attention shape：[b, d, H, W]，self attention共享权值

100K steps, lr: 4e-3, wd: 1e-7



cvt_io_v2_6:

latent array： N = 5，L = 12，self attention shape：[b, d, H, W]，self attention共享权值

100K steps, lr: 4e-3, wd: 1e-7

30k steps之后nan了，cost-time34.12ms，可以看出self-attention是真不怎么耗时

cvt_io_v2_7:

latent array： N = 5，L = 10，self attention shape：[b, d, H, W]，self attention共享权值

100K steps, lr: 2e-3, wd: 1e-7

cvt_io_v2_8:

latent array： N = 3，L = 6，self attention shape：[b, d, H, W]，self attention共享权值

100K steps, lr: 4e-3, wd: 1e-7

cvt_io_v2_9:

latent array： N = 5，L = 4，self attention shape：[b, d, H, W]，self attention共享权值

100K steps, lr: 4e-3, wd: 1e-7

	IOU@0.50	Params	FLOPs	Cost-Time
cvt	0.36	670.32K	6.934G	46.67ms
cvt_io_v1_1	0.34	936.94K	7.438G	36.91ms
cvt_io_v2_1	0.34	936.94K	6.918G	35.17ms
cvt_io_v2_2	0.328	936.94K	6.81G	33.08ms
cvt_io_v2_5	0.25	936.94K	6.78G	29.8ms
cvt_io_v2_7	0.28			
cvt_io_v2_8	0.26			