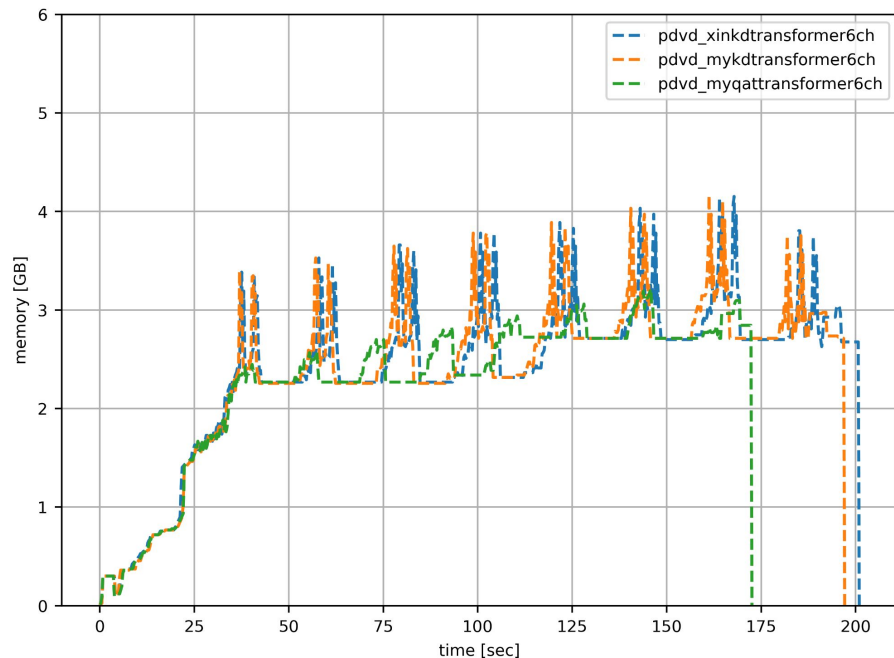




Status report on **DNNROI sigproc**

Hokyeong Nam
Chung-Ang University

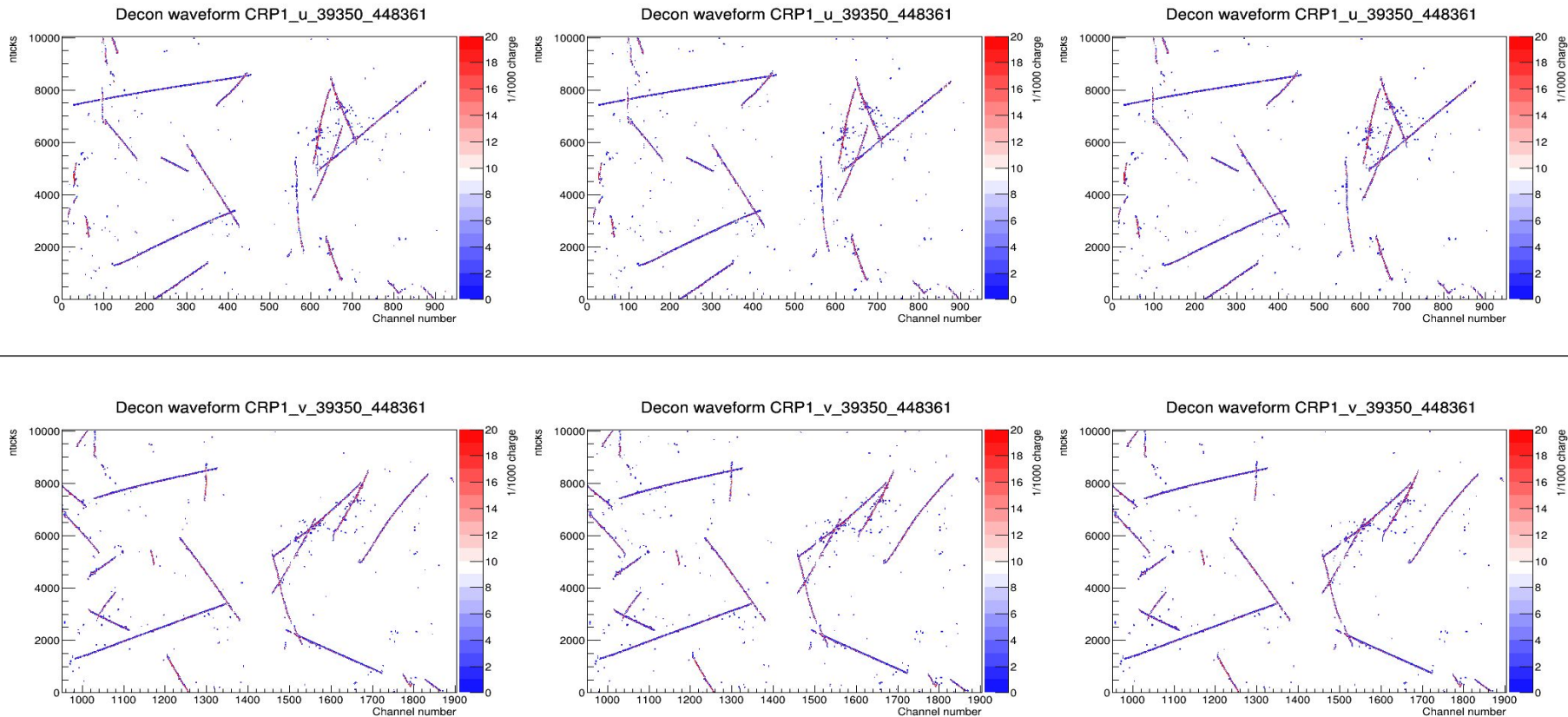
PDVD DNN-ROI + L1SP Memory Usage



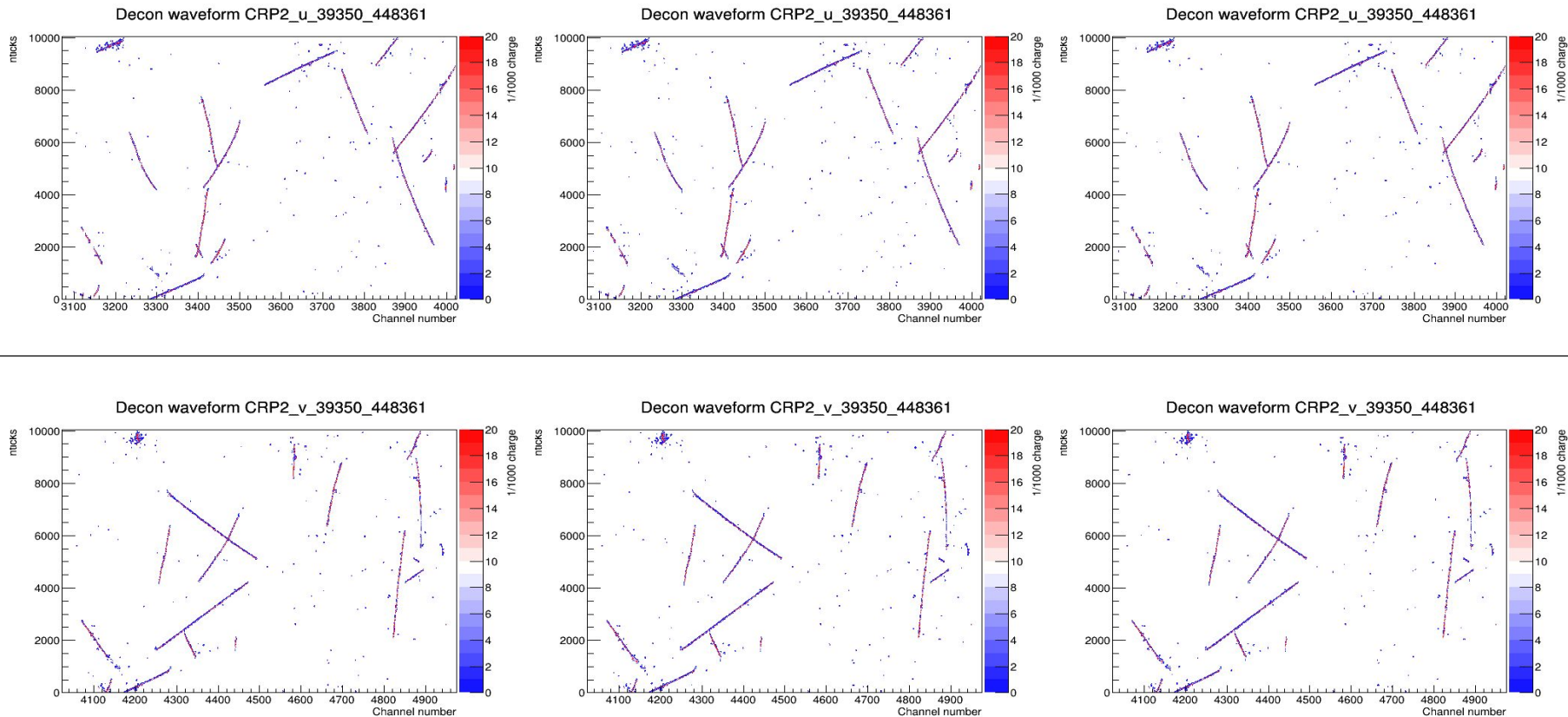
Model	Dice	eff_pix	pur_pix	eff_roi	pur_roi
Xin's KD	0.7680			0.7437	0.8506
My KD	0.7709	0.7239	0.8281	0.7487	0.8494
My QAT	0.7649	0.7089	0.8338	0.7450	0.8520

- Compared Xin's KD, my KD, and my QAT Transformer model
- All the tests are done on dunegpvm machines using same event (run39350 event 448361)
- Peak memory usage for each model: 4240 MB, 4253 MB, 3303 MB

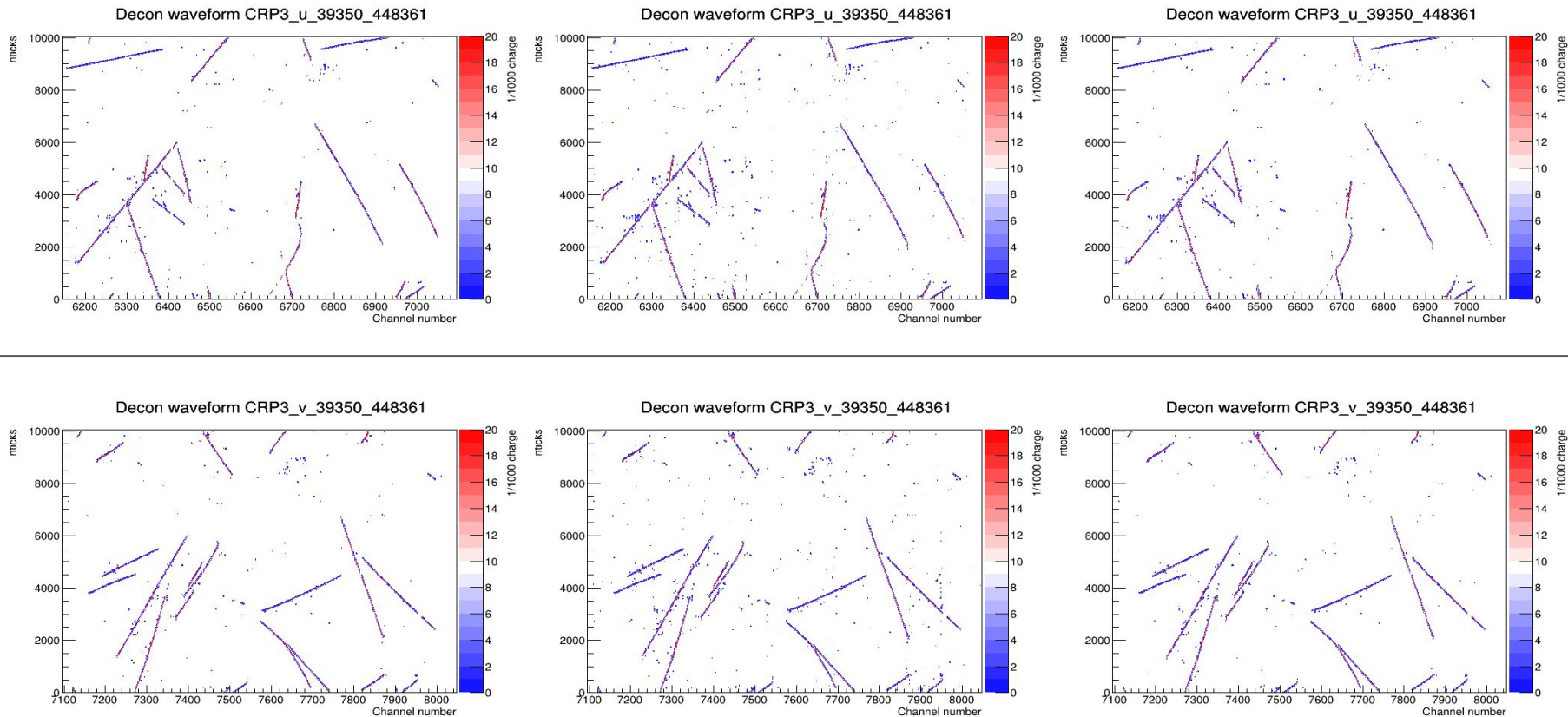
CRP1: Xin's KD vs. My KD vs. My QAT Transformer



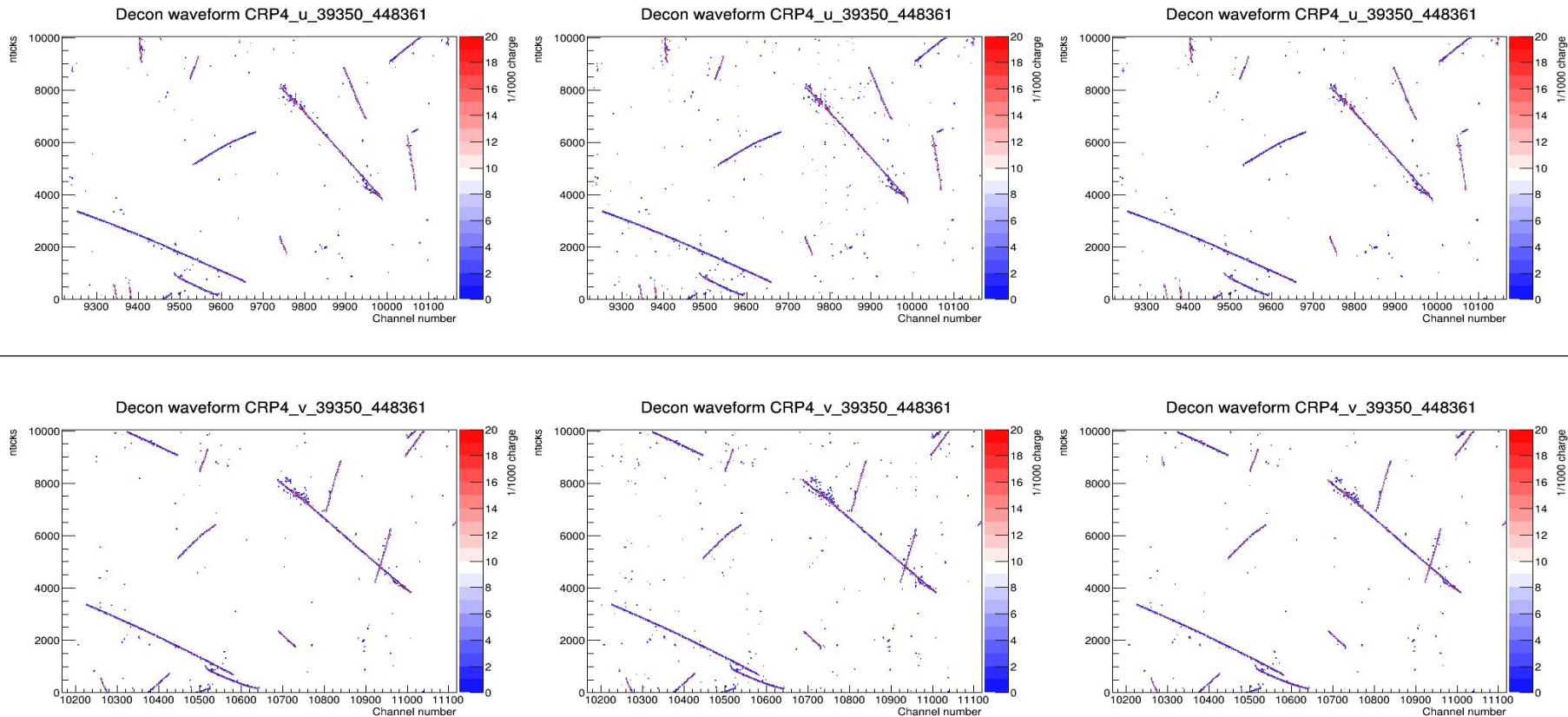
CRP2: Xin's KD vs. My KD vs. My QAT Transformer



CRP3: Xin's KD vs. My KD vs. My QAT Transformer



CRP4: Xin's KD vs. My KD vs. My QAT Transformer



Back Up

PDVD - Full Dataset (Top)vs. Small Dataset (Bottom)

Category	Precision	Dice	eff_pix	pur_pix	eff_roi	pur_roi
MobileNetV3 baseline	FP32	0.7762	0.7301	0.8298	0.7505	0.8521
NestedUNet baseline	FP32	0.7953	0.7642	0.8321	0.7634	0.8537
Transformer baseline	FP32	0.7858	0.7430	0.8432	0.7551	0.8512
KD NestedUNet	FP32	0.7662	0.7135	0.8310	0.7462	0.8504
KD Transformer	FP32	0.7709	0.7239	0.8281	0.7487	0.8494
QAT NestedUNet	INT8	0.7657	0.7190	0.8308	0.7470	0.8477
QAT Transformer	INT8	0.7649	0.7089	0.8338	0.7450	0.8520

Category	Precision	Dice	eff_pix	pur_pix	eff_roi	pur_roi
MobileNetV3 baseline	FP32	0.7898	0.7612	0.8213	0.7586	0.8464
NestedUNet baseline	FP32	0.8071	0.7875	0.8285	0.7728	0.8525
Transformer baseline	FP32	0.8089	0.8009	0.8176	0.7760	0.8465
KD NestedUNet	FP32	0.7754	0.7368	0.8193	0.7579	0.8379
KD Transformer	FP32	0.7801	0.7378	0.8288	0.7535	0.8527
QAT NestedUNet	INT8	0.7813	0.7542	0.8113	0.7571	0.8369
QAT Transformer	INT8	0.7871	0.7661	0.8102	0.7602	0.8384

- Same test set is used for this evaluation: pdvd_test_400ev_6ch.h5
- After QAT, the performance is slightly dropped
- Full dataset is better in terms of purity and small dataset has better efficiency

Comparison with previous result

Run	precision	Dice	eff_roi	pur_roi
pdvd_mobilenetv3_all_6ch	FP32	0.7538	0.7135	0.8594
pdvd_teacher_nestedunet_6ch	FP32	0.7818	0.7429	0.8570
pdvd_teacher_transformer_6ch	FP32	0.7733	0.7385	0.8485
pdvd_distill_nestedunet_6ch	FP32	0.7816	0.7520	0.8537
pdvd_distill_transformer_6ch	FP32	0.7680	0.7437	0.8506
pdvd_qat_nestedunet_6ch	INT8 (ep19)	0.7459	0.7153	0.8567
pdvd_qat_transformer_6ch	INT8 (ep19)	0.7550	0.7303	0.8448

- Xin's results (Top) as a reference
- QAT degraded eff_roi about 3% in his result
- QAT_nestedunet
 - roi efficiency is improved about 4%
 - There is a trade-off in pur-roi
- QAT_transformer
 - roi efficiency in improved about 3%
 - purity degradation is minor in this model

category	Δ Dice	Δ eff_roi	Δ pur_roi
mobilenetv3_baseline	+0.0360	+0.0451	-0.0130
nestedunet_baseline	+0.0253	+0.0299	-0.0045
transformer_baseline	+0.0356	+0.0375	-0.0020
kd_nestedunet	-0.0062	+0.0059	-0.0158
kd_transformer	+0.0121	+0.0098	+0.0021
qat_nestedunet	+0.0354	+0.0418	-0.0198
qat_transformer	+0.0321	+0.0299	-0.0064