

Table 1: The five best runs according to accuracy.

	Run Number	Mean GPU Power Draw (in W) per epoch	Accuracy (in %)	Number of Parameters	Efficiency (acc/gpu)
B	1	117.845	1.2	23792612	0.01
1.	78	183.957	95.954	23792612	0.522
2.	52	121.277	94.792	23792612	0.782
3.	47	182.02	92.838	23846055	0.51
4.	14	119.071	91.066	23792612	0.765
5.	93	144.019	91.026	23846055	0.632

Table 2: The five best runs according to GPU.

	Run Number	Mean GPU Power Draw (in W) per epoch	Accuracy (in %)	Number of Parameters	Efficiency (acc/gpu)
B	1	117.845	1.2	23792612	0.01
1.	35	111.874	35.016	23792612	0.313
2.	44	111.905	19.14	23792612	0.171
3.	92	112.538	1.04	23792612	0.009
4.	98	112.627	1.883	23792612	0.017
5.	45	113.435	1.16	23792612	0.01

Table 3: The five best runs according to efficiency (acc/gpu).

	Run Number	Mean GPU Power Draw (in W) per epoch	Accuracy (in %)	Number of Parameters	Efficiency (acc/gpu)
B	1	117.845	1.2	23792612	0.01
1.	52	121.277	94.792	23792612	0.782
2.	14	119.071	91.066	23792612	0.765
3.	80	125.453	86.005	23792612	0.686
4.	2	139.575	88.515	23792612	0.634
5.	93	144.019	91.026	23846055	0.632

Table 4: Parameter values for the winning run in accuracy (run number 78).

Parameter	Value
model	resnet50
preprocessing	standardization
augmentation	None
precision	float16
batch_size	64
partitioning	80-10-10
lr	0.0008
lr_schedule	exponential
optimizer_momentum	0.5
optimizer	RMSProp
internal	jit_compilation
seed	22
n_parameters	23792612

Table 5: Parameter values for the winning run in GPU (run number 35).

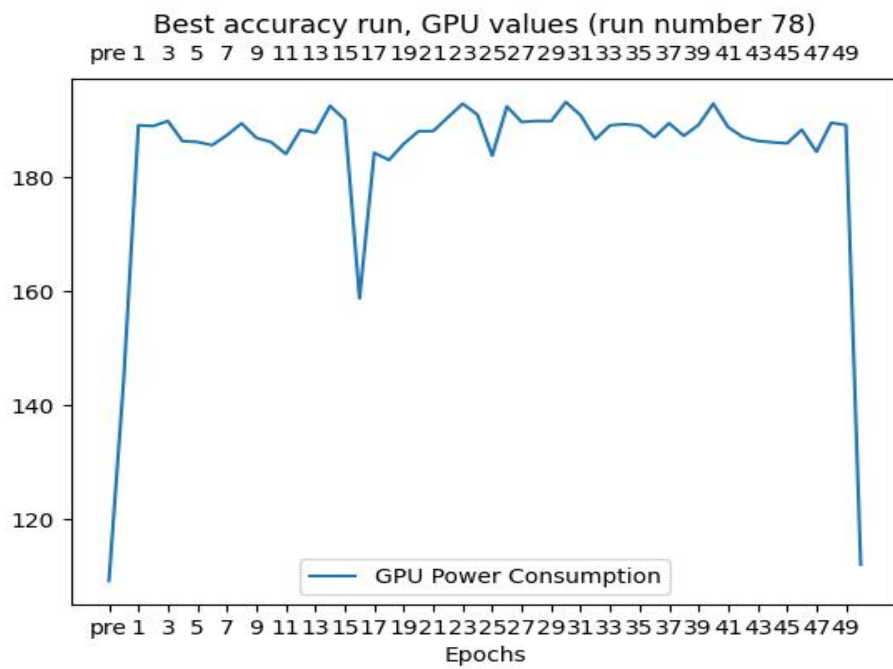
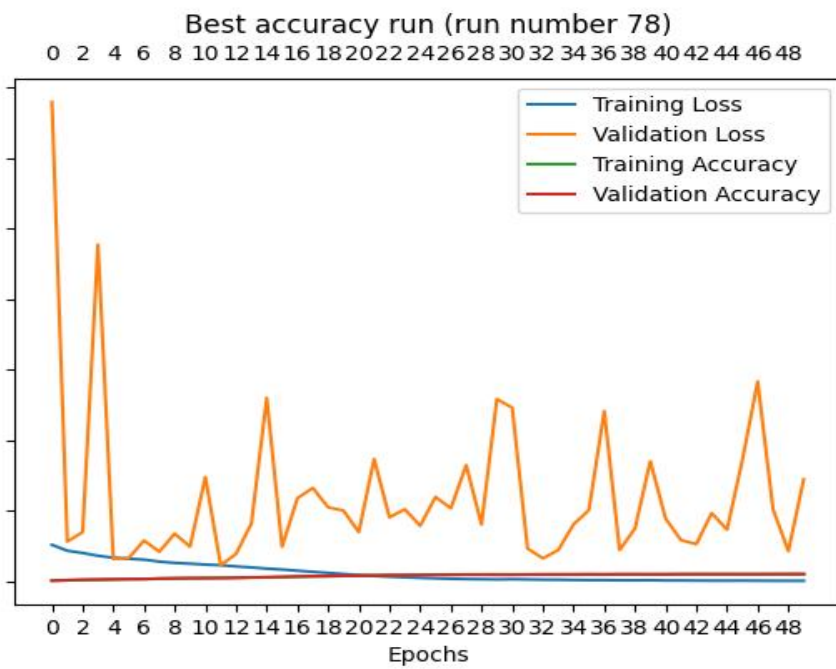
Parameter	Value
model	resnet50
preprocessing	None
augmentation	cutmix
precision	global_policy_float16
batch_size	64
partitioning	90-5-5
lr	0.01
lr_schedule	exponential
optimizer_momentum	0.5
optimizer	Adam
internal	post_quantization
seed	22
n_parameters	23792612

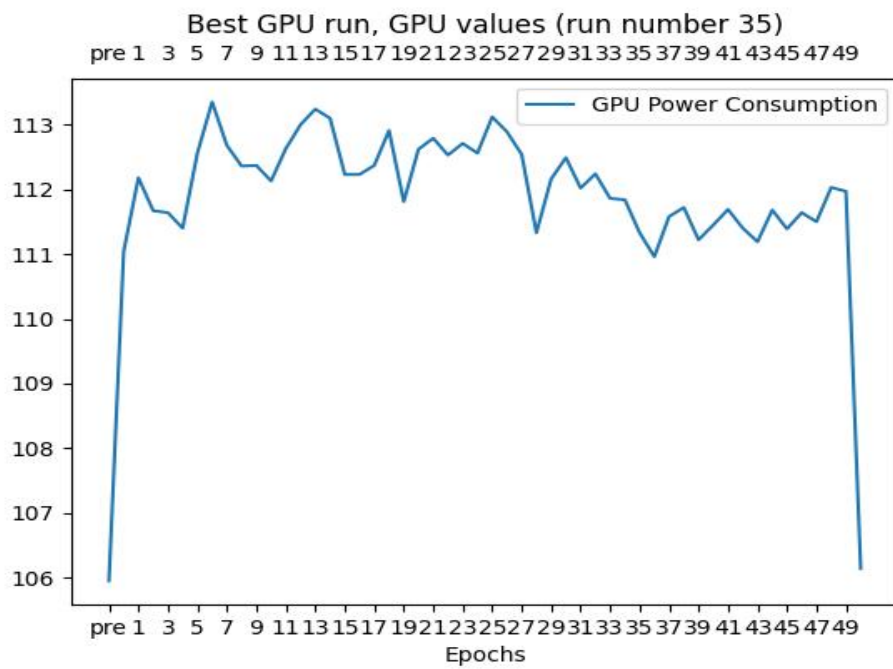
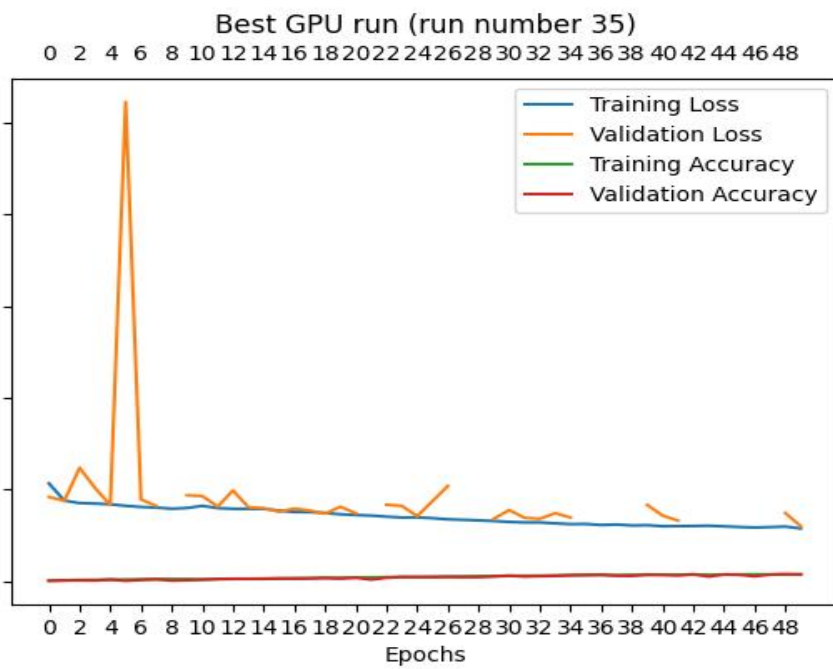
Table 6: Parameter values for the winning run in efficiency (run number 52).

Parameter	Value
model	resnet50
preprocessing	robust_scaling
augmentation	None
precision	global_policy_float16
batch_size	128
partitioning	80-10-10
lr	0.00015
lr_schedule	constant
optimizer_momentum	0.5
optimizer	RMSProp
internal	pre_quantization
seed	22
n_parameters	23792612

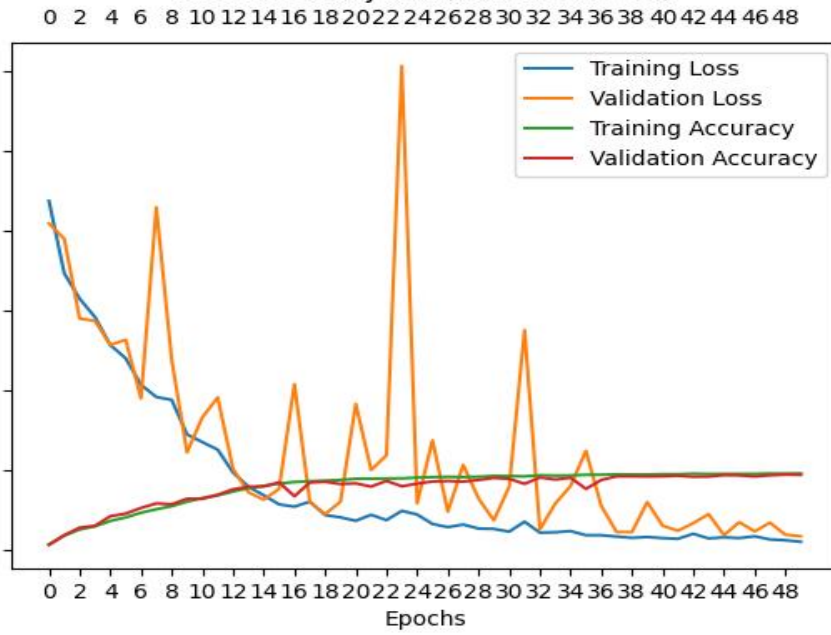
Table 7: Parameter values for the baseline run.

Parameter	Value
model	resnet50
preprocessing	None
augmentation	None
precision	float16
batch_size	1
partitioning	60-20-20
lr	0.01
lr_schedule	constant
optimizer_momentum	0.0
optimizer	RMSProp
internal	None
seed	22
n_parameters	23792612





Best efficiency run (run number 52)



Best efficiency run, GPU values (run number 52)

