

Table 1: The five best runs according to accuracy.

	Run Number	GPU (in kWh)	Accuracy (in %)	Number of Parameters	Efficiency (acc/gpu)
1	78	183.95712	95.954	gpu	0.52161
2.	52	121.27731	94.792	acc	0.78161
3.	47	182.02019	92.838	acc	0.51005
4.	14	119.07077	91.066	acc	0.7648
5.	93	144.01865	91.026	acc	0.63204

Table 2: The five best runs according to GPU.

	Run Number	GPU (in kWh)	Accuracy (in %)	Number of Parameters	Efficiency (acc/gpu)
1	35	111.87385	35.016	acc	0.313
2.	44	111.90538	19.14	acc	0.17104
3.	92	112.5375	1.04	acc	0.00924
4.	98	112.62692	1.883	acc	0.01672
5.	45	113.43481	1.16	acc	0.01023

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

	Run Number	GPU (in kWh)	Accuracy (in %)	Number of Parameters	Efficiency (acc/gpu)
1	52	121.27731	94.792	acc	0.78161
2.	14	119.07077	91.066	acc	0.7648
3.	80	125.45288	86.005	acc	0.68556
4.	2	139.57481	88.515	acc	0.63418
5.	93	144.01865	91.026	acc	0.63204

Table 4: Parameter values for the winning run in accuracy.

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.

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.

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