

Table 1: The five best runs according to accuracy.

| | Run Number | Mean GPU Power Draw (in W/hr) | Accuracy (in %) | Number of Parameters | Efficiency (acc/gpu) |
|----|------------|-------------------------------|-----------------|----------------------|----------------------|
| 1. | 78 | 1.003 | 0.96 | 23792612 | 0.957 |
| 2. | 52 | 0.485 | 0.948 | 23792612 | 1.953 |
| 3. | 47 | 0.729 | 0.928 | 23846055 | 1.273 |
| 4. | 14 | 2.494 | 0.911 | 23792612 | 0.365 |
| 5. | 93 | 1.457 | 0.91 | 23846055 | 0.625 |

Table 2: The five best runs according to GPU.

| | Run Number | Mean GPU Power Draw (in W/hr) | Accuracy (in %) | Number of Parameters | Efficiency (acc/gpu) |
|----|------------|-------------------------------|-----------------|----------------------|----------------------|
| 1. | 60 | 0.324 | 0.867 | 23792612 | 2.674 |
| 2. | 58 | 0.331 | 0.391 | 23792612 | 1.181 |
| 3. | 54 | 0.412 | 0.215 | 23792612 | 0.522 |
| 4. | 61 | 0.437 | 0.008 | 23792612 | 0.019 |
| 5. | 88 | 0.442 | 0.08 | 23792612 | 0.181 |

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

| | Run Number | Mean GPU Power Draw (in W/hr) | Accuracy (in %) | Number of Parameters | Efficiency (acc/gpu) |
|----|------------|-------------------------------|-----------------|----------------------|----------------------|
| 1. | 60 | 0.324 | 0.867 | 23792612 | 2.674 |
| 2. | 52 | 0.485 | 0.948 | 23792612 | 1.953 |
| 3. | 43 | 0.512 | 0.823 | 23792612 | 1.608 |
| 4. | 57 | 0.47 | 0.75 | 23792612 | 1.597 |
| 5. | 66 | 0.489 | 0.775 | 23792612 | 1.586 |

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 60).

| Parameter | Value |
|--------------------|-----------------------|
| model | resnet50 |
| preprocessing | standardization |
| augmentation | mixup |
| precision | global_policy_float16 |
| batch_size | 128 |
| partitioning | 80-10-10 |
| lr | 0.00625 |
| lr_schedule | exponential |
| optimizer_momentum | 0.9 |
| optimizer | SGD |
| internal | jit_compilation |
| seed | 22 |
| n_parameters | 23792612 |

Table 6: Parameter values for the winning run in efficiency (run number 60).

| Parameter | Value |
|--------------------|-----------------------|
| model | resnet50 |
| preprocessing | standardization |
| augmentation | mixup |
| precision | global_policy_float16 |
| batch_size | 128 |
| partitioning | 80-10-10 |
| lr | 0.00625 |
| lr_schedule | exponential |
| optimizer_momentum | 0.9 |
| optimizer | SGD |
| internal | jit_compilation |
| 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 |

