

1 Experiment Set Difference: Separable Convolutions (256 filters) v. Separable Convolutions (512 filters)

1.1 Average Treatment Effect (Per Experiment in Experiment Set)

| | |
|-----------------------------------|--------------------------------------------------------------------|
| cpu_hours | -1.00595 +/- 0.00564456 (stderr) (p=1.37914e-06, Welch's t-test) |
| gpu_hours | -0.917247 +/- 0.000934335 (stderr) (p=3.73604e-13, Welch's t-test) |
| estimated_carbon_impact_kg | -0.0830775 +/- 0.00908498 (stderr) (p=0.00259775, Welch's t-test) |
| total_power | -0.238989 +/- 0.0184335 (stderr) (p=4.10683e-06, Welch's t-test) |
| kw_hr_gpu | -0.139161 +/- 0.000459415 (stderr) (p=1.15838e-07, Welch's t-test) |
| kw_hr_cpu | -0.0120973 +/- 0.011862 (stderr) (p=0.15137, Welch's t-test) |
| exp_len_hours | -0.928019 +/- 0.000631014 (stderr) (p=7.02618e-15, Welch's t-test) |
| average_realtime_carbon_intensity | -31.5587 +/- 12.7815 (stderr) (p=0.0187973, Welch's t-test) |

1.2 Cumulative Difference (Per Experiment Set)

| | |
|-----------------------------------|-----------------------|
| cpu_hours | -2.3826 |
| gpu_hours | -2.176298294842155 |
| estimated_carbon_impact_kg | -0.19753628926363603 |
| total_power | -0.5454715963709869 |
| kw_hr_gpu | -0.32644445908594344 |
| kw_hr_cpu | -0.018790728490630654 |
| exp_len_hours | -2.20137431078487 |
| average_realtime_carbon_intensity | 207.88596474763267 |