Figure 1: Fairness evaluation as displayed in Figure 4b in the manuscript.
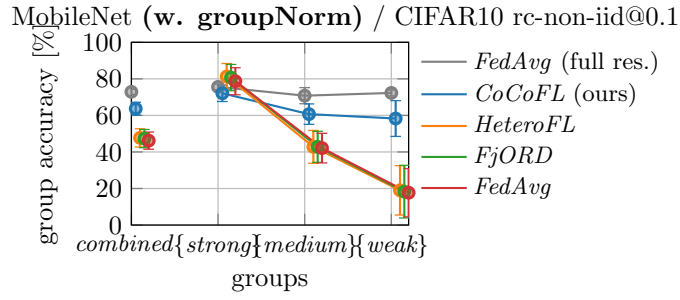

Figure 2: Fairness evaluation in the same setup with groupNorm instead of batchNorm.
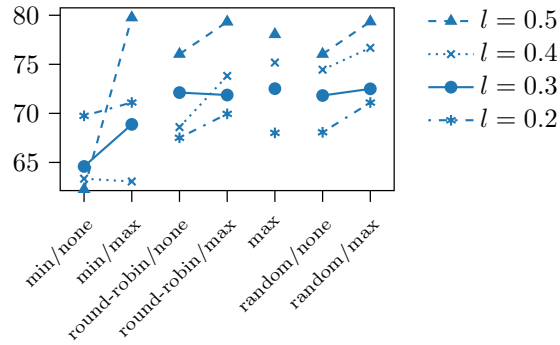

Figure 3: Ablation study for heuristic evaluating with $l \in [0.2, 0.3, 0.4, 0.5]$