Compositional Generalization: Distributed vs Winner-Takes-All

Pattern Similarity

gap=0.350

gap=0.023

+29.5%

- - - Chance level
- Train
- Held-out

Winner-Takes-All (Baseline)

Distributed ATL (Ours)