Test Loss curves for modular addition for different amounts of training data

Legend:
- train_0.95
- test_0.95
- train_0.9
- test_0.9
- train_0.8
- test_0.8
- train_0.7
- test_0.7
- train_0.6
- test_0.6
- train_0.5
- test_0.5
- train_0.4
- test_0.4
- train_0.3
- test_0.3
- train_0.28

Y-axis: Loss (1, 0.01, 100μ, 1μ)
X-axis: Epoch (0, 5k, 10k, 15k, 20k)