

Master Experimental Results — Hand Sign CNN Study

Consolidated from result.txt; combined L1+L2 typo corrected to L1 $\lambda=0.05$ as confirmed by the researcher.

1. Preliminary experiments before normalization

Experiment	Setting	Train	Test
Iterations	300	100%	75.8%
Iterations	1000	100%	80.0%
Iterations	2000	100%	75.8%
L2	$\lambda=0$	100%	77.5%
L2	$\lambda=0.1$	100%	77.5%
L2	$\lambda=0.3$	100%	75.0%
L2	$\lambda=0.6$	100%	80.0%

2. Normalized medium-model L2 sweep

Seed	$\lambda=0$	$\lambda=0.1$	$\lambda=0.3$	$\lambda=0.6$	$\lambda=1$	$\lambda=2$
42	91.67%	95.00%	90.00%	90.00%	91.67%	81.60% (train 99.8%)
1137	91.67%	89.17%	93.30%	93.30%	90.00%	—
2024	86.67%	89.16%	88.30%	90.83%	90.00%	—

3. Capacity experiment — no regularization

Small = 16→32→64; Medium = 32→64→128; Large = 64→128→256.

Model	Architecture	Seed	L1	L2	Dropout	Train	Test
Small	16→32→64	42	0	0	0	100%	84.16%
Medium	32→64→128	42	0	0	0	100%	91.67%
Large	64→128→256	42	0	0	0	100%	83.33%
Large	64→128→256	2024	0	0	0	100%	90.80%

4. Dropout × model capacity — seed 42

Model	Architecture	dr=0.1	dr=0.3	dr=0.5
Small	16→32→64	100 / 90.83	100 / 90.83	81.4 / 71.6
Medium	32→64→128	100 / 89.00	100 / 96.66	97.4 / 79.1
Large	64→128→256	100 / 90.83	100 / 92.50	100 / 95.00

Cells show Train / Test accuracy (%).

5. L1 × model capacity — seed 42

Model	Architecture	$\lambda=0.001$	$\lambda=0.01$	$\lambda=0.05$
Small	16→32→64	100 / 87.50	100 / 88.33	100 / 91.66
Medium	32→64→128	100 / 86.67	100 / 85.83	100 / 90.00
Large	64→128→256	100 / 90.00	100 / 88.30	100 / 89.16

Cells show Train / Test accuracy (%).

6. Combined L1 + L2 regularization

Confirmed: L1 $\lambda=0.05$ and L2 $\lambda=0.6$ for all three seeds.

Seed	L1 λ	L2 λ	Train	Test
42	0.05	0.6	100%	88.33%
1137	0.05	0.6	100%	85.00%
2024	0.05	0.6	100%	89.16%
Mean	0.05	0.6	100%	87.50%

7. Selected reproducibility checks

Configuration	Seed 42	Seed 1137	Mean
Small, no regularization	84.16%	84.16%	84.16%
Medium, Dropout 0.3	96.66%	93.34%	95.00%
Large, Dropout 0.5	95.00%	95.00%	95.00%
Medium, L1 0.05	90.00%	89.16%	89.58%

8. Three-seed normalized L2 baseline

Configuration	Seed 42	Seed 1137	Seed 2024	Mean
Medium, $\lambda=0$	91.67%	91.67%	86.67%	90.00%

9. Main observations for later paper analysis

- Normalization improved the baseline relative to the preliminary pre-normalization experiments.
- Moderate L2 regularization sometimes improved generalization; excessive L2 ($\lambda=2$ in the seed-42 sweep) reduced performance substantially.
- Dropout produced the strongest observed accuracies in the capacity sweep: 90.83% (Small), 96.66% (Medium), and 95.00% (Large).
- In the seed-42 capacity sweep, the best tested dropout strength shifted from 0.1/0.3 for Small, to 0.3 for Medium, to 0.5 for Large; this is an observed trend, not a universal claim.
- L1 showed architecture-dependent behavior: $\lambda=0.05$ was best for Small and Medium, while $\lambda=0.001$ was best for Large among the tested values.
- Selected second-seed runs remained reasonably strong, supporting reproducibility of the main configurations.
- Combined L1+L2 (L1=0.05, L2=0.6) did not outperform the strongest individual regularization configurations.

Source: user's result.txt. Experimental values are preserved; only the explicitly confirmed L1+L2 typo was corrected.