Multilayer MNIST CNN [Time 726.357s, GPUs 2] (Ir = 0.000100, batch size = 50) Loss Accuracy 10^{-1} 1.000 10^{-2} 0.995 10⁻³ 0.990 10^{-4} 10⁻⁵ 0.985 10^{-6} 0.980 - 10^{-7} train train 10-8 0.975 val val 25 25 50 100 125 150 50 75 100 75 0 125 150