

Two extra layers, 10000 training samples, $L_1 = 0.0001$

