

Training loss: $L(\theta_1, \theta_2) = \frac{1}{2} \|ReLU(\mathbf{A}^{\mathsf{T}} \mathbf{K}_2(\theta_2) \mathbf{F}^{\mathsf{H}} \mathbf{K}_1(\theta_1) \mathbf{F} \mathbf{p}) - \mathbf{x}\|_2^2 + R(\theta_1)$

