

Training APE on avg InfoNCE loss
($N = 6$, $M = 1$), evaluation done on ViT-S/16

