Training APE on avg InfoNCE loss
6 combinations, evaluation done on ViT-S/16