sgd_pretrained vs adahessian_pretrained