lr = 10^-4

normalizzazione

no augmentations

optimizer = AdamW