lr = 10^-4

optimizer = AdamW

no augmentation

normalizzazione