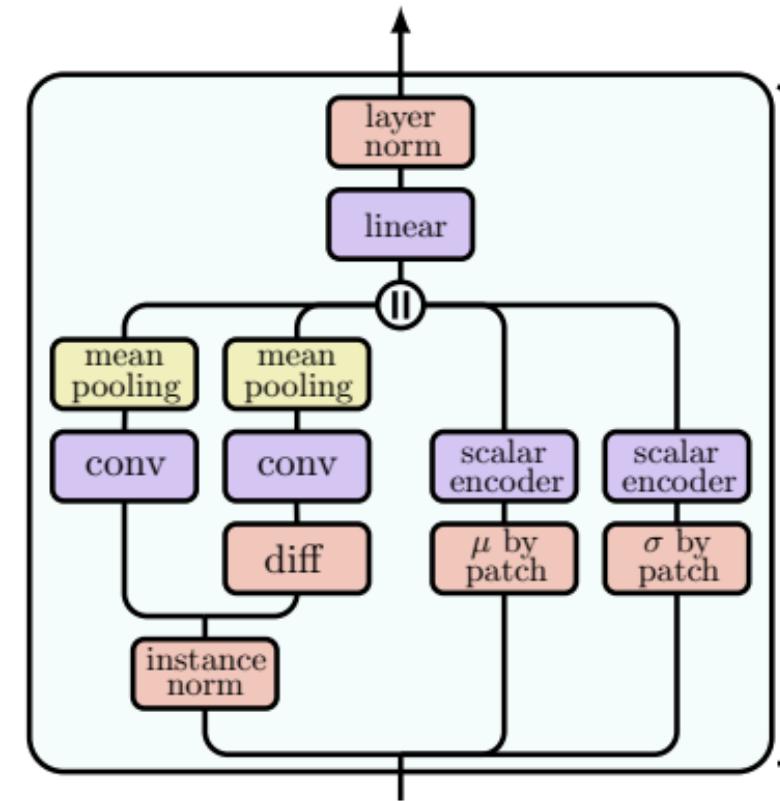


Token Generator Unitoutput for
pre-training

projector

prediction

prediction
head

vit unit

token
generator unitinput **ViT Unit**