

1. Describe the difference between WGAN* and GAN**, list at least two differences

WGAN 使用的是 Wasserstein 距離，GAN 使用 KL 散度、JS 散度

WGAN 的 Discriminator 把最後一層 sigmoid 去掉了

Ref: <https://zhuanlan.zhihu.com/p/25071913>

2. Please plot the “Gradient norm” result.

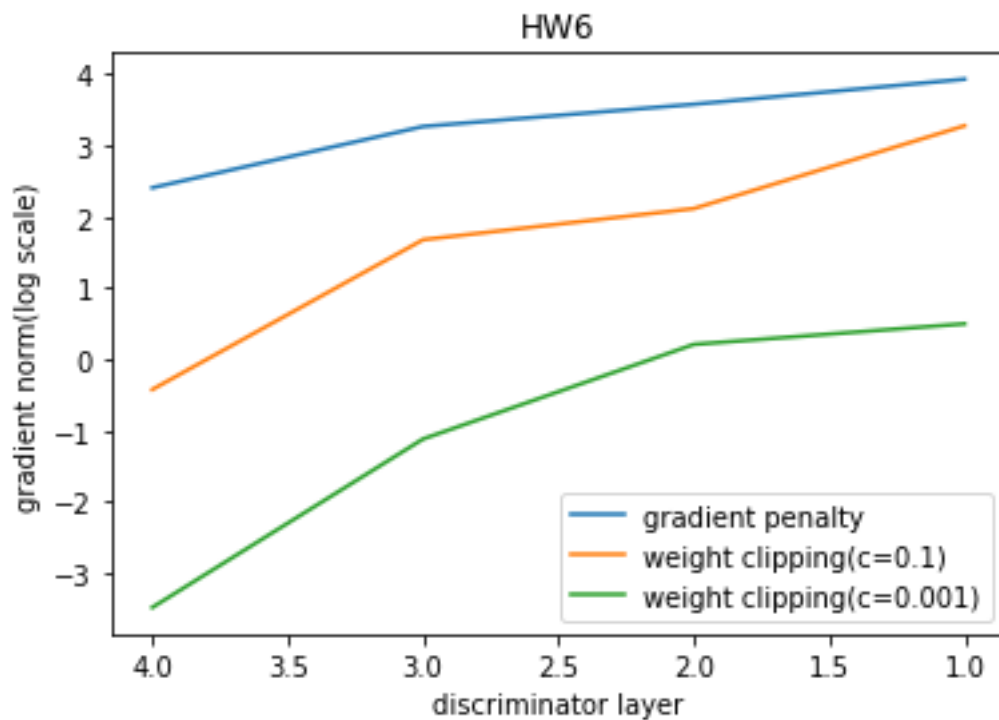
a. Use training dataset, set the number of discriminator layer to 4 (minimum requirement)

b. Plot two setting:

i. weight clipping

ii. gradient penalty

c. Y-axis: gradient norm(log scale), X-axis: discriminator layer number (from low to high)



layer 是取 conv_bn_lrelu