

Wasserstein GAN (WGAN)

ORIGINAL Standard GAN

Distance Metric

Jensen-Shannon divergence

$$JS(p_{data} || p_g)$$

Output Layer

Sigmoid (Discriminator)

Constraint

No explicit constraint

Training

Unstable, oscillating losses

VS

WGAN Wasserstein GAN

Distance Metric

Wasserstein distance (Earth Mover)

$$W(p_{data}, p_g)$$

Output Layer

Linear (Critic, no sigmoid)

Constraint

Lipschitz constraint (weight clipping)

Training

Stable, meaningful loss



Stable Training

More consistent convergence dynamics



Meaningful Loss

Loss correlates with quality



Mode Coverage

Better distribution coverage



WGAN-GP Enhancement

Gradient penalty replaces weight clipping for better Lipschitz constraint enforcement

Original Papers

WGAN

Wasserstein GAN

Arjovsky et al., 2017

WGAN-GP

Improved Training of Wasserstein GANs

Gulrajani et al., 2017