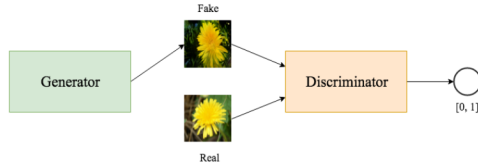# Spherical CNNS

Liangjie Cao

July 14, 2018

1. GAN

Figure 1. High-level view of the GAN framework. The generator produces synthetic images. The discriminator takes images as input and outputs the probability it assigns to the image of being real. A common analogy is that of an art forger (the generator) which tries to forge paintings and an art investigator (the discriminator) which tries to detect imitations

The GAN framework we have learned yesterday is can be seen from Figure 3. The easiest way to define and analyse the game is as a zero-sum game where D_ω and G_θ are the strategies of the two players. Such a game can be described by a value function $V(D, G)$ which in this case represents the payoff of the discriminator. The discriminator wants to maximise V while the generator wishes to minimise it. The payoff described by V must be proportional to the ability of D to distinguish between real and fake samples. The value function from Equation 1 which was originally proposed in [1] comes as a natural choice.

$$V(D, G) = \mathbb{E}_{X\ \mathbb{P}_r}[log(D(x))] + \mathbb{E}_{Z\ P_z}[log(1 - D(G(Z)))] \quad (1)$$

On the one hand, note that $V(D, G)$ becomes higher when the discriminator can distinguish between real and fake samples. On the other hand, V becomes lower when the generator is performing well, and the critic cannot distinguish well between real and fake samples.

The difference from maximum likelihood models is that samples generated by the generator have an implicit distribution $\mathbb{P}_g$ determined by the particular value of θ. $\mathbb{P}_g$ cannot be explicitly evaluated. The discriminator forces the generator to bring $\mathbb{P}_g$ close to $\mathbb{P}_r$.

Zero-sum games are minimax games so the optimal parameters of the generator can be described as in Equation 2.

$$\theta^* = \arg\min_\theta \max_\omega V(D, G) \quad (2)$$
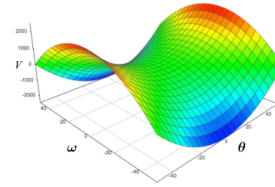
The solution of the minimax optimisation is a Nash Equi-

Figure 2. The Nash Equilibrium corresponds to a saddle point in space where V is at a minimum with respect to θ and at a maximum with respect to ω. Neither of the networks has any interest to change their parameters

librium (Figure 2). Theorem [1] describes the exciting properties of the equilibrium for non-parametric functions. The theorem confirms the intuition that as the two players play this game, the generator will improve at producing realistic samples.

In practice, the generator does not minimise $V(D, G)$. This would be equivalent to minimising $E_{ZpZ}[log(1D_\omega(G_\theta(z)))]$, but the function $log(1D_\omega(G_\theta(z)))$ has saturating gradient when the discriminator is performing well and the function approaches log(1). This makes it difficult for the generator to improve when it is not performing well. Instead, the generator is minimising L_G from 3 whose gradient saturates only when the generator is already performing well. The loss L_D of the critic is included in 4.

$$L_G = -\mathbb{E}_{z\ pz}[log(D_\omega(G_\theta(z)))] \quad (3)$$

$$L_D = -mathbbE_{x\ pr}[log(D_\omega(x))]$$
$$-\mathbb{E}_{z\ pz}[log(1 - D_\omega(G_\theta(z)))] \quad (4)$$

---

[1]

Theorem1 *TheNashequilibriumofthe(non*   —
parametric)GANgameoccurswhen :
 1.Thestrategyofthediscriminatoris $D = \frac{\mathbb{P}_r}{\mathbb{P}_r + \mathbb{P}_g}$
 2.Thestrategy G of the generator makes $\mathbb{P}_g = \mathbb{P}_r$.

2. Conditional Generative Adversarial Networks

The original GAN paper [1] describes how one can trivially turn GANs into a conditional generative model. To generate data conditioned on some condition vector c, c is appended to both the generator and the discriminator in any of their layers. The networks will learn to adapt and adjust their parameters to these additional inputs.
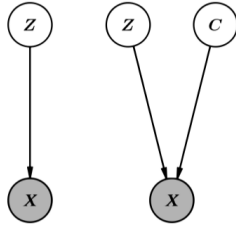

Figure 3. Probabilistic graphical model view of regular GANs (left) and conditional GANs (right)

Conditional GANs can also be seen from the perspective of a probabilistic graphical model (Figure 3). In the case of regular GANs, the noise Z influences the observable X. For conditional GAN s, both Z and C influence X. In the particular case of text to image synthesis, the states c of C are vectors encoding a text description.

References

[1] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *NIPS*, 2014. 1, 2