3D GAN Object Generation and Reconstruction

Gendong Zhang, Zixuan Zhou, Liuming Zhao
Department of Electrical Engineering, Stanford University

Introduction

Generative Adversarial Networks (GANs) is one of promising techniques used in Computer Vision and unsupervised machine learning, and it can generate superficially authentic images that is hard to distinguished by human observers. Combined with GANs, 3D object generation and reconstruction can be practical use when dealing with robot grasping and obstacle avoidance. In this project, we investigated the training results of GANs with varied GAN architectures and also reconstructed 3D object from one surface of the object.

Dataset

- ModelNet10 [1] contains approximately 57000 objects with dimensions of 32x32x32.
- 3D GAN generation model was trained on "Chair" dataset (900).
- 15 surfaces of each object is generated randomly as input for this model.
- randomly chose 10% samples as validation set, and 10% of the dataset is hidden from the model to evaluate the final system.

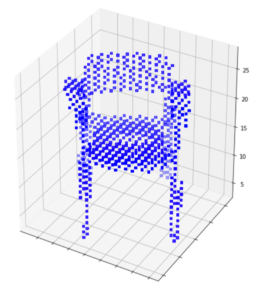

Figure 1. Example of Ground Truth of Chair

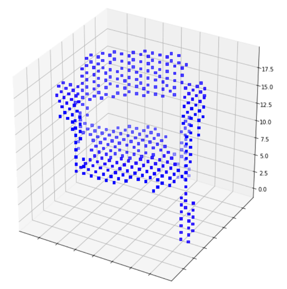

Figure 2. Synthetic Surface of Figure 1

Reference

[1]. Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao. 3d shapenets: A deep representation for volumet- ric shapes proceedings of 28th ieee conference on computer vision and pattern recognition (cvpr2015). 2015. Oral Pre- sentation 3D Deep Learning Project Webpage.

[2]. E. J. Smith and D. Meger. Improved adversarial sys- tems for 3d object generation and reconstruction. *CoRR*, abs/1707.09557, 2017

Model Structures

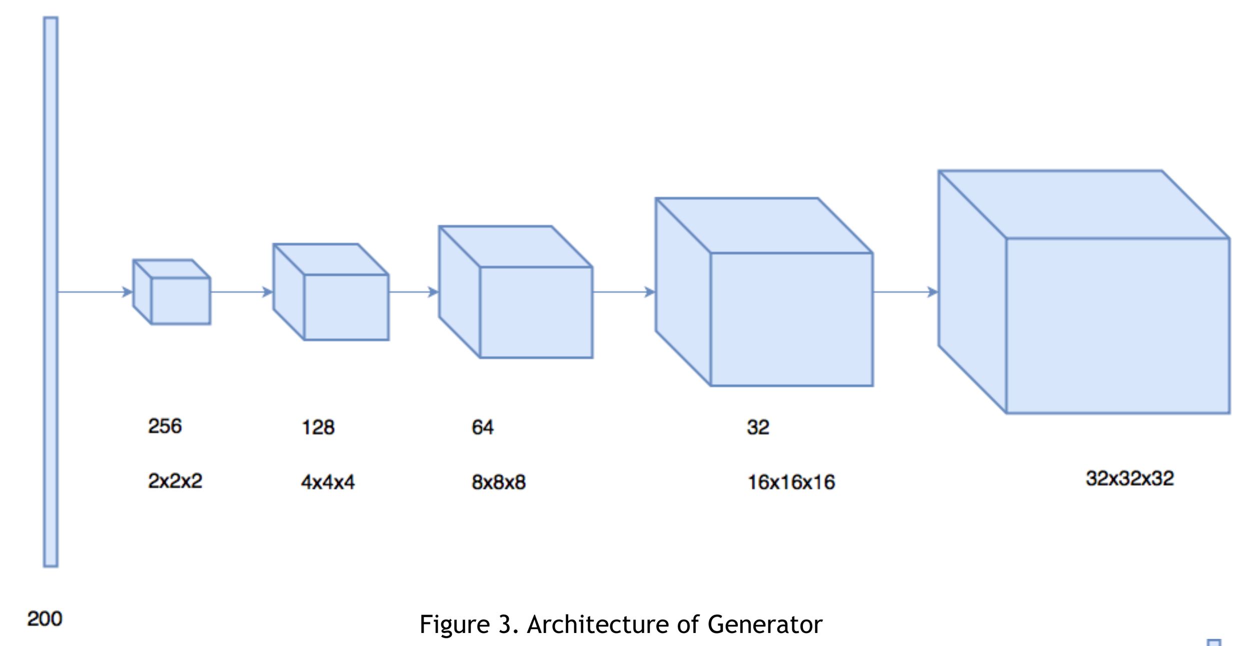

Figure 3. Architecture of Generator

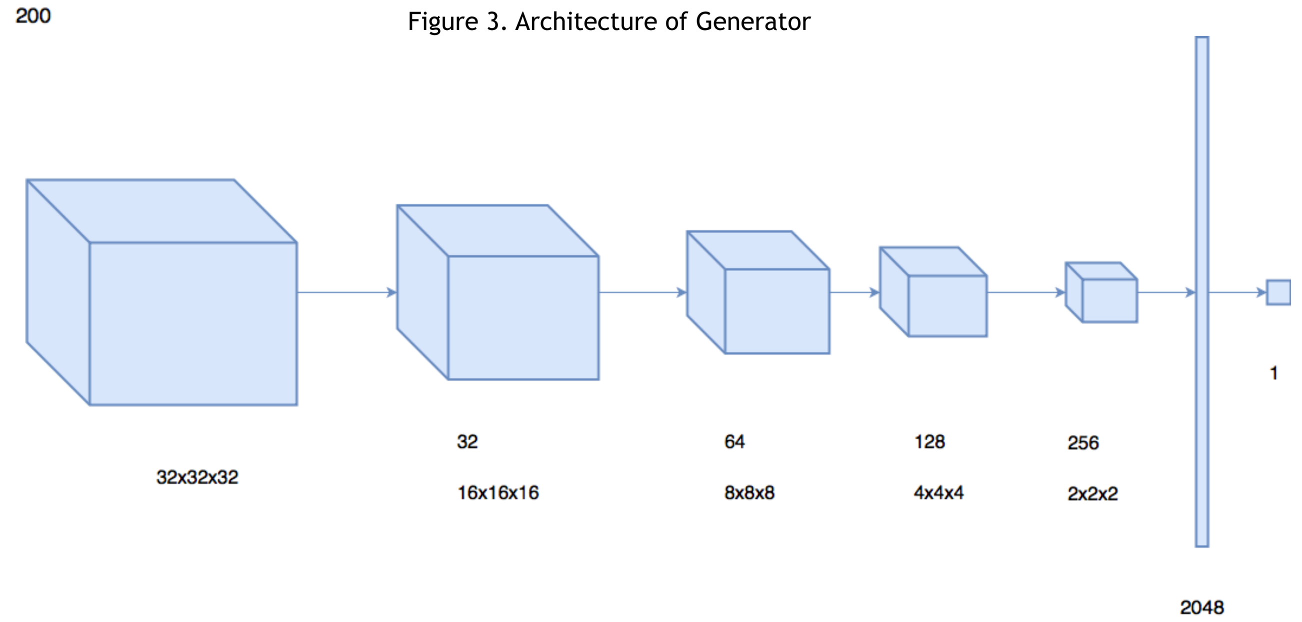

Figure 4. Architecture of Discriminator

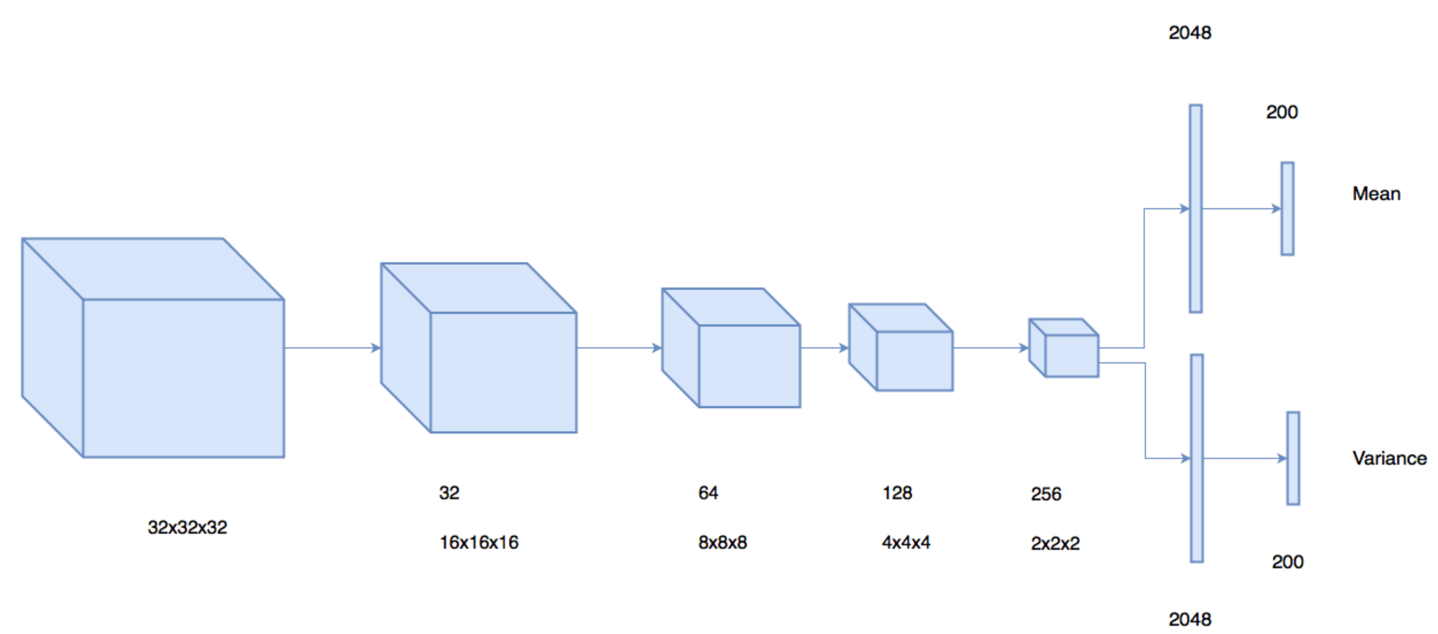

Figure 5. Architecture of Encoder of VAE

Equations

$$L = L_{3D-GAN} + \alpha_1 L_{KL} + \alpha_2 L_{recon},$$

where,

$$L_{3D-GAN} = logD(x) + log(1 - D(G(z))),$$
$$L_{KL} = D_{KL}(q(z|y)||p(z)),$$
$$L_{recon} = ||G(E(y)) - x||_2$$

Equation 1. Overall Loss function of Reconstruction

$$W(P_r, P_g) = \underset{\gamma \sim \prod(P_r, P_g)}{inf} E_{(x,y)\sim\gamma}[||x - y||]$$

Equation 2. Wasserstein distance

$$L_D = -E_{x\sim P_r}[D(x)] + E_{x\sim P_g}[D(x)] + \lambda E_{x\sim P_{\hat{x}}}[||\Delta_x D(x)||_p - 1]^2$$

Equation 3. Improved Loss Function with W-distance and Gradient Penalty

Results

3D-GAN

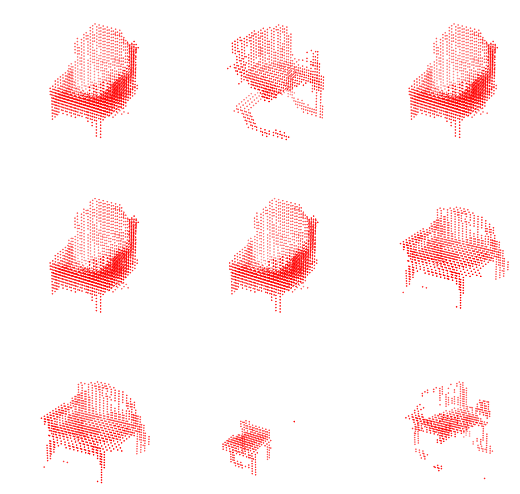

Figure 6. Results of 3D-GAN at 995 epoch

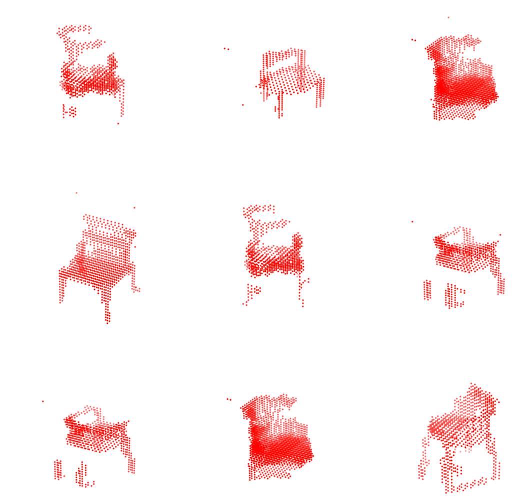

Figure 7. Results of 3D-GAN at 1495 epoch

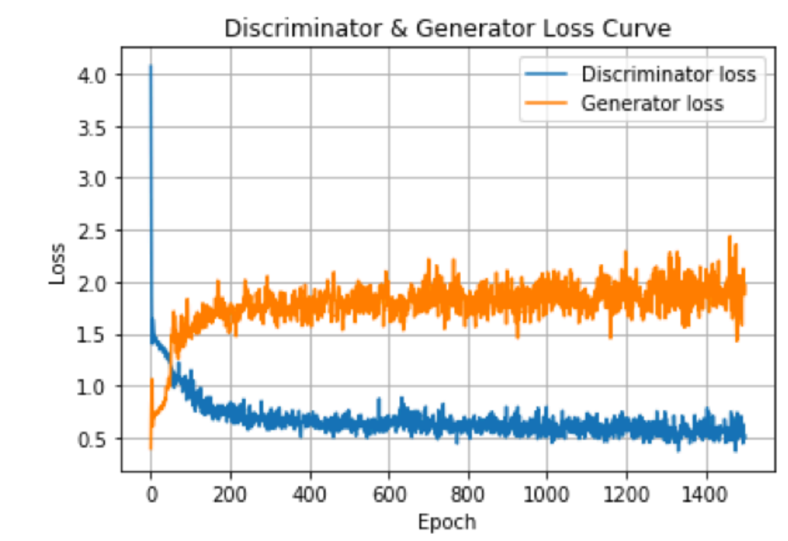

Figure 8. Loss curve of 3D-GAN

3D-WGAN-GP (Diversity Increasing)

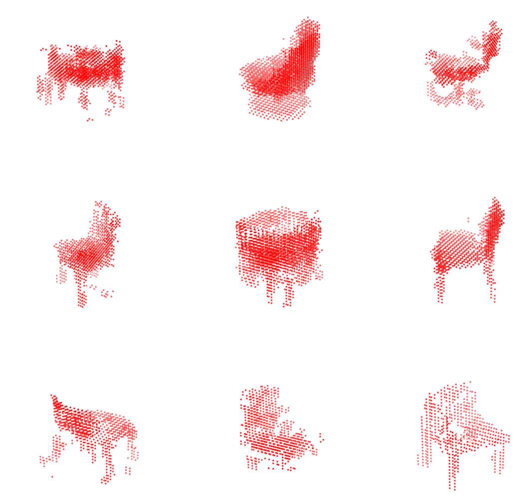

Figure 9. Results of 3D-WGAN-GP at 995 epoch

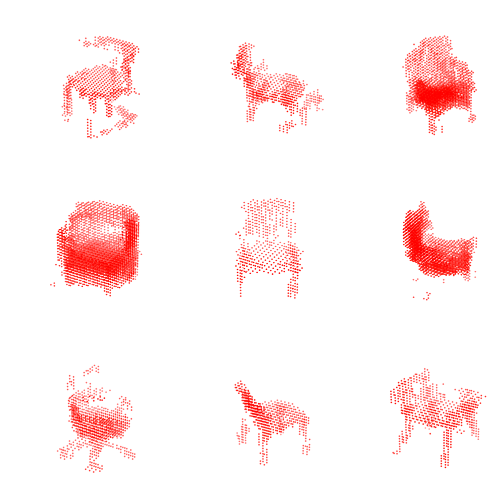

Figure 10. Results of 3D-WGAN-GP at 1495 epoch

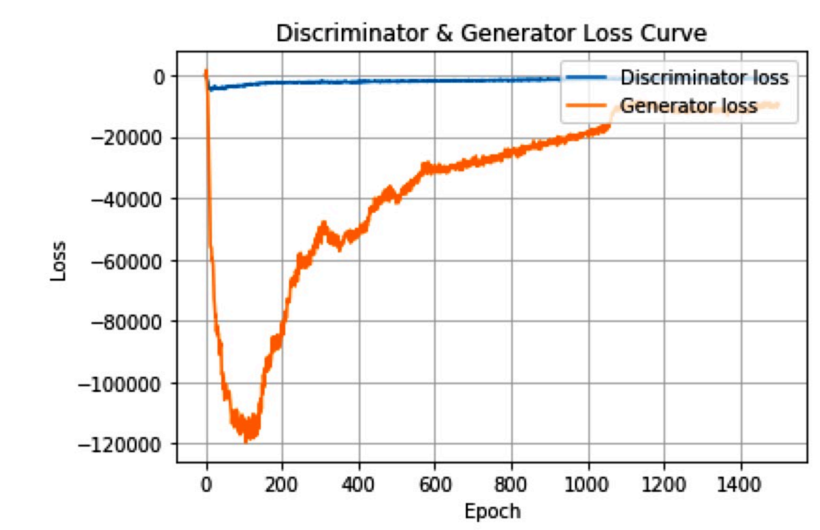

Figure 11. Loss curve of 3D-WGAN-GP

3D-Reconstruction (From Single View)

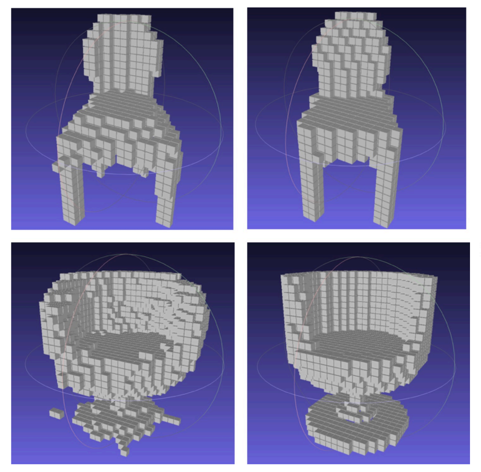

Figure 12. Examples of Reconstruction

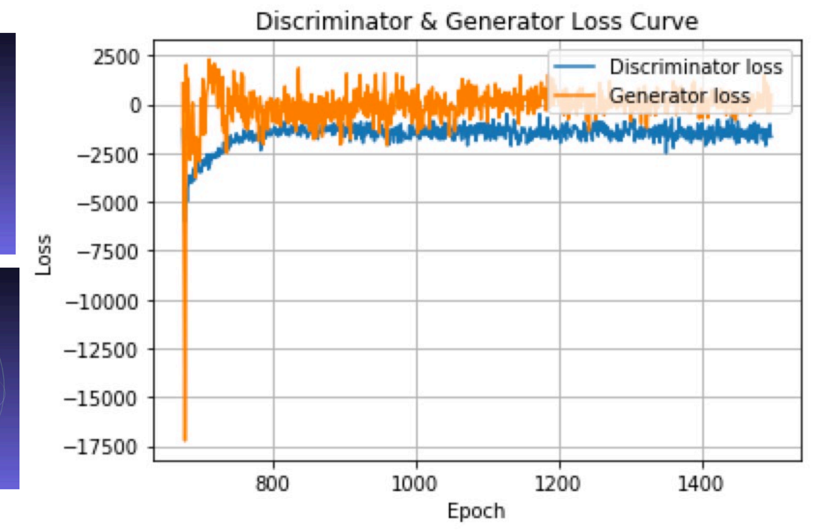

Figure 13. Loss curve of Discriminator & Generator

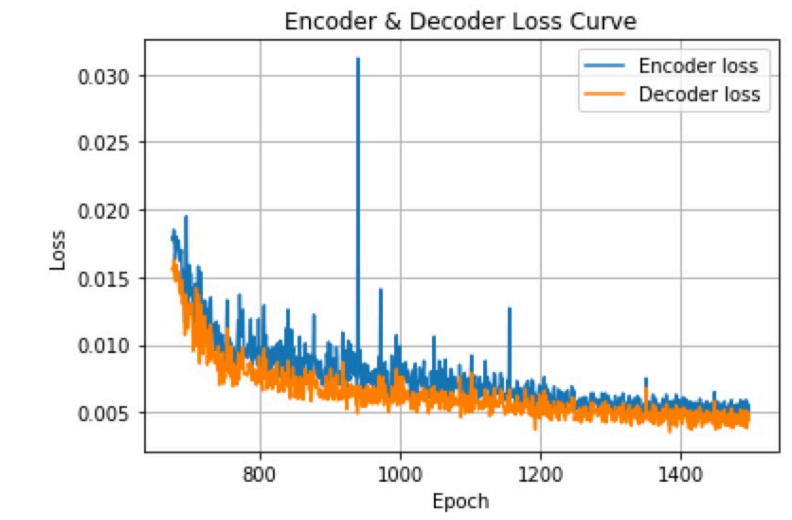

Figure 14. Loss curve of Encoder & Decoder

Conclusion

- Compared WGAN-GP with 3D-GAN object generation model, both of them perform well.
- WGAN-GP outperforms the basic 3D-GAN in diversity of generated objects
- Experiment demonstrated promising results for the reconstruction 3D object from one single view