# GENERATING 3D MODELS USING 3D GENERATIVE ADVERSARIAL NETWORK

[1] Pranav Gandhi, [2] Adarsh Shaw, [3]Emil Eji, [4] Steffina Muthukumar

[1,2,3]*Student*, [4]*Assistant Professor*

*Computer Science & Engineering,*

*SRM Institute of Science and Technology,Chennai.*

---

***Abstract:*** *In recent years there is an increase in solving the problems which combines Computer vision and Natural Language processing. There has been new algorithm developed to solve these problems. 3D Model Generation is one such problem which fall in the computer vision category. In this paper we are studying the difficulties of generating a 3D object. For this we present a simple framework of 3D Generative-Adversarial-Network(3D-GAN) which produces 3 Dimensional structure from a latent vector by using the modern advancement in convolution networks. Our model is maps the low dimensional space to 3d space. So that it can sample objects without needing to reference a model created on Computer Aided Design Software. Our model learns these features without any supervision and can generates high quality models.*

***IndexTerms*** **– Generative Adversarial Network, 3D Convolution, Unsupervised learning.**

## I. Introduction

3D shape awareness is an old problem in Computer vision which is not yet completely solved by the community. A large amount of work focus on 3D reconstruction. The main goal of the computer vision research is to figure out how the human system accomplishes such task. For example, asking a person to “think of violet cat”, the person will have no problem imagining a violet cat. He will have a clear image of that cat, without having seen it before. Generative AI is Growing Fast will help computers understand the world better.

Creating a 3D model is a complex and a time consuming task. This is a big problem in game industry, interior designing and CAD modeling for engineering. In the past decade, researcher have made astonishing progress in 3D object creation, mostly based on the meshes.

Recently, due to improvement in computing hardware. The sector of deep learn has advanced and the spring up of large dataset like ShapeNet has made it possible to train complicated model which has greatly out performed all other algorithms. The older method of generation based on meshes was challenging due to the high dimensionality of the problem.

Earlier works suggest that it’s possible to generated 3D objects with one image and convolution neural networks. In this paper, we show modeling 3D objects in adversarial manner can be a effective solution as it can have both properties of good generative model variation and realistic. Our model uses volumetric convolutional network. Slightly different from the traditional convolutional network. The training criteria also differs as an adversarial discriminator is introduced to identify whether the generated object is real or fake. The purpose of this paper is to make the voxel based 3d model generation simpler. And to see how significant is the trade between the learning speed and the resolution of the object.

Modeling 3D object in GAN structure offers addition advantage as the variation in the generated model increases without trading off the realism of the object. We demostrate that our model can be used for generating highly detailed and realistic model and the discriminator can be used for 3D object recognition.

## II. LITERATURE SURVEY

**2.1 3D Shape Completion and Isometric view 3D Reconstruction:** In general 3D shape reconstruction is a narrowed down problem and a special case of 3d object generation. Classic 3D shape reconstruction approach can be classified into symmetry based method and data analysis method. The data driven method approaches the shape finalization problem as retrieval and alignment problem. In general, data driven method approaches are will only work by assuming the features about the data category.

**2.2 Modeling and generating 3D shapes**: 3D object generation is a difficult problem in vision and computer graphics. In the past, AI and computer vision researcher have made astonishing attempts to learn 3D objects representation largely based on skeleton and meshes. Most of the algorithms are nonparametric and recombines parts of shapes to create a new shape.

**2.3 Deep learning on 3D data:** The computer vision community has seen a rapid improvement of neural networks in various task. In the task of 3D model recognition Li et al. [2015], Su et al. [2015b], Girdhar et al. [2016] propose to learn a joint inlay of 3d object and generated images. Most of the framework are trained with full or partial supervision compared to ours which is unsupervised.

**2.4 Learning with Adversarial net:** Generative adversarial network proposed to incorporate an adversarial discriminator with the generating model. Recently LAPGAN and DCGAN combined GAN with convolutional neural network for image generation problem and achieve outstanding results. While all the previous approaches were for 2D images we used adversarial component for the 3D objects.

**2.4. Kullback-leibler divergence:** It is also known as relative entropy. It is a method to identify the relation between two probability distribution p(x) and q(x). It measures how one distribution diverges from the other.

$$D_{KL}(p||q) = \int_x p(x) log \frac{p(x)}{q(x)} dx$$

(2.4)

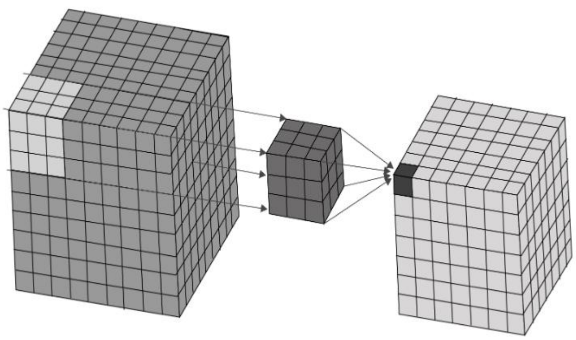
If the KL divergence is zero, it means that p(x) is same as q(x) at every other point.

**2.5. Nash equilibrium:** This concept comes from the game theory. It is a particular state in game theory which can be achieve d in a non-cooperative game where the player chooses the best possible strategy for themselves to get the best possible result for them. The decision is based on strategy to maximize the profit for that player.

## III. PROPOSED SYSTEM

In this section we present our proposed architecture and the approach for 3D Generative adversarial network-based 3D model generator. Our proposed system uses the simplified volumetric pixel instead of the higher dimensional mesh data. By using the recent advancement in GANs the frame work is able to map the 3D space into a low dimensional latent space. Our 3D Generative Adversarial Network takes advantage of 3D convolution network and generative adversarial network.

**3.1. 3D Convolution**: 3D convolution is similar to the traditional convolution except operations apply a 3D filter to input data along all three dimensions. This operation makes a 3 stack of features maps. The shape of the output is a cuboid.



(3.1)

**3.2. 3D Generative Adversarial Network:** As proposed in Goodfellow etal.[2014], the Generative Adversarial Network (GAN) consists of a discriminator and a generator where the prior tries to identify fake object and the real object, and the generator tries of fool the discriminator. In our 3D-Generative Adversarial Network(3D-GAN), the generator maps the two hundred dimensional noise randomly sampled to a 64x64x64 cube, representation of an object in a voxel space which are volumetric pixel. Discriminator gives the probability if the input object is real or the fake.

Same as the Goodfellow etal.[2014], we use the binary cross entropy as the discriminator loss, and present overall model loss as:

$$L_{3D-GAN} = log D(x) + log(1 - D(G(z)))$$

(3.2)

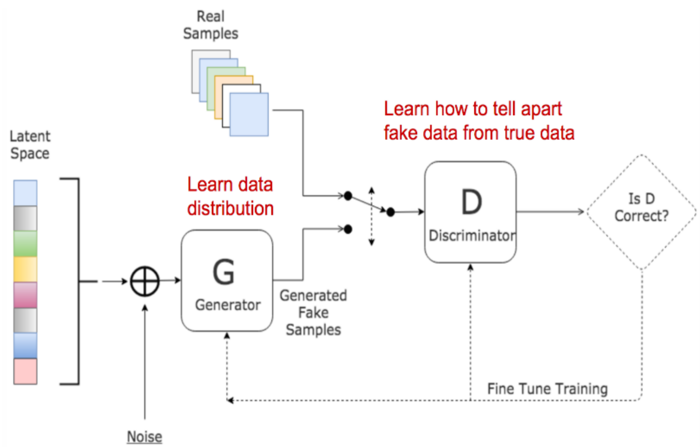
where x is a real object sample and z is the noise sample from random distribution. In each dimension z is distributed over the interval of [0,1].

The advantage of our system is that it is unsupervised there is no assistance in training and all the data is not labeled. The model learns the density distribution of the data. So, it creates the internal representation of the messy and complicated distribution.

## IV. NETWORK Architecture

Motivated by Radford et al. [2016], we designed a deconvolution neural network for 3d object generation. The generator network consists of five volumetric deconvolution layer of filter size 4x4x4 and stride of 2 with batch normalization and leaky ReLU layers and at the end there is a sigmoid layer. The discriminator is essentially an inverted form of the generator except it uses volumetric convolution layers there is no pooling layer in the network.

**4.1. Training Process:** There is a simple training process. It is to update both network in every batch, but like this the discriminator learns much faster than the generator and this make the discriminator much better than the generator much faster and generator is not able to keep up with the pace. So, discriminator is only trained on every other mini-batch and generator is trained on all mini-batch and the learning rate of the generator is much higher than that of the discriminator. We are using the Adam optimizer for both the discriminator and generator.




**(4.1)**

**4.2. Loss Function:** Since we are trying to generate a model that is realistic and it is supervised by the discriminator which judge the generator. So, here we can apply the Nash equilibrium where we want the loss of discriminator to maximize and loss of generator to minimize. And combining the binary cross entropy loss with Nash equilibrium we get the following equation

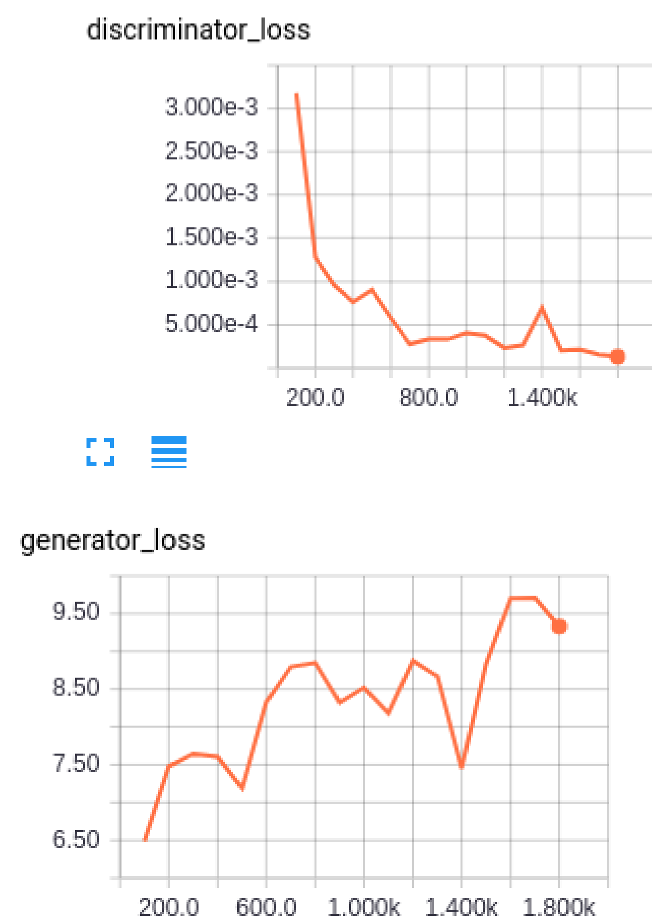
$$\min_{G}\max_{D} V(D,G) = \mathbb{E}_{x\sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_{z}(z)}[\log(1-D(G(z)))].$$

(4.2)

Where D is the discriminator and G is the generator p(x) is real data distribution and p(z) is the fake data distribution.

## V. RESULT

The proposed model was trained on the 3DShapeNet dataset. After training the model we evaluate the result of generated objects. Astonishingly, the model not only can generate new objects but also combine different styles of the object to produce a new realistic object for example combining a Victorian style chair and a modern chair it production new chair with both traits. Since there is no accuracy for the generator network it is evaluated by comparing it against how well it does against the discriminator which has an accuracy matrix as it only classifies between fake and real object. The training is very unstable, but compared to previous works on 3D GANs our model generates both high resolution object with details. Note it is easier to generate low resolution objects compared to higher resolution as the complexity increases exponentially. One of the major concern for the generator is that if it is just recreating the models in the training set. So, we compared the voxel position of generated model with training model using distance formula and the result show a similarity between object but they are the identical. This proves that the generator is not just retrieving a random model from the training set and giving it as the output.




(5.0)

## VI. CONCLUSION

In this paper we have developed a 3D object generation model and a 3d object recognition model. The 3d object recognition models is utilize to improve the generator model. Our model was able to generate novel objects and both of our network learned without any supervision and the discriminator was used to create a low dimensional feature representation of the object which help us explore the latent space of the object representation and object interpolation.

## VII. FUTURE WORKS

In this paper we have shown the proof of concept as the model performed phenomenal well. This is a lot of head space for improvement by further tuning the hyper parameters. We can further improve the generation of object by passing a string in the input which co relates with the training data which can result in generation of specific style of model instead of a random style object. This string could also contain some constraint parameter which will help generator synthesize an object with particular constraints. This may be helpful in designing and engineering where a design needs to have some constraint. As we see there are two use cases to which our model can be adapted.

## VIII. REFERENCE

[1] “Learning 3D Shape Completion under weak Supervision”, International Journal of Computer vision,Springer, David Stutz, Andreas Geiger.

[2] “Learning to generate chairs, tables and cars with convolution networks”.arXiv:1411.5928v4 [cs.CV], Thomas Brox, Jost Tobias Springenberg, Alexey Dosovitskiy, Maxim Tatarchenko.

[3] “Learning a Probablistic Latent space of object space via 3D Generative Adversarial Modeling”. Xue T.,Wu, J., Zhang, C., Freeman, W. T. and Tenenbaum, J. B. [2016].

[4] Despois, J. [2017 (Accessed 2017-10-18)], Autoencoders Deep learning bits, [blogpost], Hackernoon.

[5] “Example based 3d object reconstruction from line drawing”,in CVPR,2012. Xiaoou Tang, Jianzhuang Liu, and Tianfan Xue.

[6] “3D shapenets, A deep representation for volumetric shape”,in CVPR,2015. Xiaoou Tang, Zhirong Wu, Linguang Zhang, Shuran Song, Fisher Yu, Aditya Khosla and Jianxiong Xiao.

[7] “Generative image modeling using style and structure adversarial networks”. In ECCV, 2016.Xiaolong Wang and Abhinav Gupta.

[8] “Unsupervised representation learning with deep convolutional generative adversarial networks”, in ICLR, 2016. Soumith Chintala, Luke Metz and Alec Radford.

[9] Generative-adversarial-networks, in NIPS, 2014. David Warde-Farley, Ian Goodfellow, Yoshua Bengio, Jean Pouget-Abadie, Bing Xu, Sherjil Ozair, Aaron Courville, and Mehdi Mirza.

[10] “Shapenet: An information-rich 3d model repository”. arXiv preprint arXiv:1512.03012, 2015.Angel X Chang, Thomas Funkhouser, Leonidas Guibas, et al.

[11] “Shape completion using 3d-encoder-predictor CNNs and shape synthesis”. In IEEE (CVPR) Dai, A., Qi, C. R., & Nießner, M. (2017).

[12] I. Sutskever,A. Krizhevsky, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in NIPS, 2012, pp. 1106–1114.