# Parametric 3D Explorations with Adversarial Networks

Y. Veneela Sudha Lakshmi
UG Students,
Malla Reddy University
Hyderabad,500100

Y. Sai Yaswanth
UG Students,
Malla Reddy University
Hyderabad,500100

Y. Manideep Reddy
UG Students,
Malla Reddy University
Hyderabad,500100

Y. Ajay
UG students,
Malla Reddy University
Hyderabad,500100

Thayyaba Khatoon, PhD
Professor & HOD
Malla Reddy University
Hyderabad,500100

## ABSTRACT
Three-dimensional (3D) modeling and exploration are crucial for a wide range of applications.The authors propose a novel approach for sketch-based 3D exploration utilizing Stacked Generative Adversarial Networks (SGANs). The authors highlight the significance of three-dimensional (3D) modeling and exploration across various domains, such as computer-aided design, virtual reality, and gaming. The authors emphasize that traditional techniques for 3D modeling often necessitate expertise in complex software tools and considerable manual effort. The authors note that in recent years, deep learning methods, specifically Generative Adversarial Networks (GANs), have demonstrated impressive capabilities in generating realistic and high-quality 3D models. Building upon this progress, they present their method, which harnesses the power of GANs to produce 3D models from sketch-based inputs. This approach enables users to intuitively and interactively explore 3D scenes.

The research contributes to the field of sketch-based 3D exploration by introducing a unique framework that combines the strengths of GANs with sketch input, resulting in the generation of realistic and interactive 3D models. The authors highlight that their approach holds the potential to transform the way users engage with 3D modeling tools, making the process more intuitive, accessible, and enjoyable.

## Keywords
Sketch-based modeling, 3D shape generation, Stacked Generative Adversarial Networks (SGANs), Interactive 3D exploration, Sketch-based input, Intuitive user interface.

## 1. INTRODUCTION
The authors proposes a novel approach to address the challenges faced by sketch-based 3D modeling in generating realistic and visually pleasing 3D shapes. The approach combines the power of Generative Adversarial Networks (GANs) with sketch-based inputs to enhance the 3D exploration process. The framework presented in the paper is called Stacked GAN (SGAN), which consists of two stages: sketch generation and 3D shape generation.

In the sketch generation stage, a conditional GAN is employed to learn the generation of realistic 2D sketches from user input. This stage captures the relationship between sketches and their corresponding 3D shapes, allowing for accurate representation. By training the GAN on a dataset of sketch-3D shape pairs, it can generate sketches that closely resemble the desired 3D models.

The 3D shape generation stage utilizes a stacked GAN architecture, which progressively refines the generated shapes to achieve high-quality and complex 3D models. The stacked GAN framework facilitates the iterative improvement of the generated shapes, ensuring that they become more realistic and visually pleasing with each refinement step. This approach enables users to explore and create intricate 3D models in an intuitive and natural manner.

By combining the strengths of GANs with sketch-based inputs, the proposed SGAN framework offers a promising solution to the limitations of traditional sketch-based 3D modeling. It opens up new possibilities for users to interact with 3D modeling systems, allowing for the creation of visually appealing and realistic 3D shapes through the seamless integration of sketching and generative modeling techniques.

## 2. LITERATURE SURVEY
"Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images" presents an innovative approach to generate 3D mesh models from 2D RGB images. Published in 2018 by Nanyang Technological University, the paper introduces a deep learning-based method that addresses the challenging task of inferring the underlying 3D geometry of objects from 2D image data.The paper emphasizes the significance of 3D reconstruction from 2D images and highlights its potential applications in computer vision, robotics, and virtual reality. Pixel2Mesh introduces a deep learning architecture that combines convolutional neural networks (CNNs) and graph convolutional networks (GCNs) to generate 3D mesh models. This end-to-end method takes an RGB image as input and produces a 3D mesh as output.

The authors extensively evaluate the performance of Pixel2Mesh through various experiments. They utilized the ShapeNet dataset and provide comprehensive implementation details, including training settings and evaluation metrics. The results showcase the effectiveness of the proposed method in generating high-quality 3D mesh models.To demonstrate the effectiveness of Pixel2Mesh, the authors present visual results and quantitative evaluations. They showcase 3D reconstructions from different object categories and compare the generated meshes against ground truth models. The paper concludes by discussing the limitations of the proposed method, such as handling occlusion and generalizing to unseen object categories. The authors also suggest potential directions for future research, including incorporating semantic information and exploring alternative network architectures.

## 3. DATA COLLECTION
### 3.1 Dataset
The authors utilized existing 3D shape databases, such as ShapeNet, which provide a comprehensive collection of 3D

models across various object categories. This database serves as a valuable source of 3D shapes that can be paired with the collected sketches.

## 3.2 Data Preprocessing

The data preprocessing begins with data cleaning, which involves removing any noisy or irrelevant elements from the collected dataset. This step ensures that the dataset contains high-quality and relevant sketches and associated 3D shape information. By eliminating unwanted artifacts or inconsistencies, the subsequent training process can be more effective.

Normalization is another crucial step in data preprocessing. It involves scaling the data to a consistent range or distribution.This involves scaling the pixel values of sketches to a standardized range, or normalizing the coordinates of the 3D shapes to a specific coordinate system

Alignment is another important aspect of data preprocessing. It aims to ensure spatial alignment between the sketches and the corresponding 3D shapes. This step involves aligning the sketches and 3D shapes in a consistent manner, such as matching the orientations or viewpoints. By aligning the data, the GANs can learn the correspondence between the sketches and the associated 3D shapes more effectively, leading to improved performance during the generation process.

## 4. METHOD

The proposed sketch-based 3D exploration approach employs a Stacked Generative Adversarial Network (SGAN) framework consisting of two stages: sketch generation and 3D shape generation.

In the sketch generation stage, a conditional GAN architecture is utilized. The conditional GAN is trained on a dataset containing paired sketch and 3D shape data, enabling it to learn the correspondence between sketches and their corresponding 3D shapes. Given a user's input sketch, the conditional GAN generates plausible and realistic 2D sketches. The objective of the CGAN can be defined as a minimax game between the generator network (G) and the discriminator network (D). Given an input sketch x and its corresponding 3D shape y, the generator aims to generate a plausible sketch G(x), while the discriminator aims to distinguish between the generated sketches and real sketches.

The loss function for the CGAN can be represented as:

$$CGAN(G,D) = E\big[\log\big(D(x,y)\big)\big] + E\left[\log\left(1 - D\big(x, G(x)\big)\right)\right]$$

Here, E represents the expectation over the training dataset. The first term aims to maximize the probability that the discriminator correctly identifies real sketches, while the second term aims to maximize the probability that the discriminator fails to distinguish between real sketches and generated sketches.

The generator's objective is to minimize the cGAN loss:

$$G = -E\left[\log\left(D\big(x, G(x)\big)\right)\right]$$

In the 3D shape generation stage, a stacked GAN architecture is introduced. The stacked GAN is composed of multiple GAN modules, each responsible for modeling a specific level of detail. The generated 2D sketches from the previous stage are used as inputs to the stacked GAN. Through the stacked architecture, the GAN modules progressively refine the generated shapes, allowing for a gradual increase in complexity and fidelity. The output of the stacked GAN is a high-quality 3D shape that closely resembles the user's intended design. The generator of the stacked GAN can be represented as a composition of individual generators: G = G1 ∘ G2 ∘ ... ∘ GN, where G1 generates low-level details and GN generates high-level details.The discriminator of the stacked GAN evaluates the generated 3D shape's realism at each level of detail. It provides feedback to the corresponding generator, helping to refine the shape generation process.

Similar to the CGAN, the stacked GAN's objective can be formulated as a minimax game between the generator network (G) and the discriminator network (D). The loss function for the stacked GAN can be defined as:

$$SG(G,D) = \sum_i \left[E\big[\log\big(D(i,y)\big)\big] + E\left[\log\left(1 - D\big(i, G(i)\big)\right)\right]\right]$$

Here, SG means Stacked GAN, i represents the level of detail, and D(i, y) and D(i, G(i)) denote the discriminator's outputs for the real shape y and the generated shape at level i, respectively.

The generator's objective is to minimize the stacked GAN loss across all levels of detail:

$$StackedG = -\sum_i \left[E\left[\log\left(D\big(i, G(i)\big)\right)\right]\right]$$

Through this progressive stacking approach, the stacked GAN refines the generated 3D shapes, incorporating details from lower-level to higher-level modules, ultimately producing high-quality 3D models.

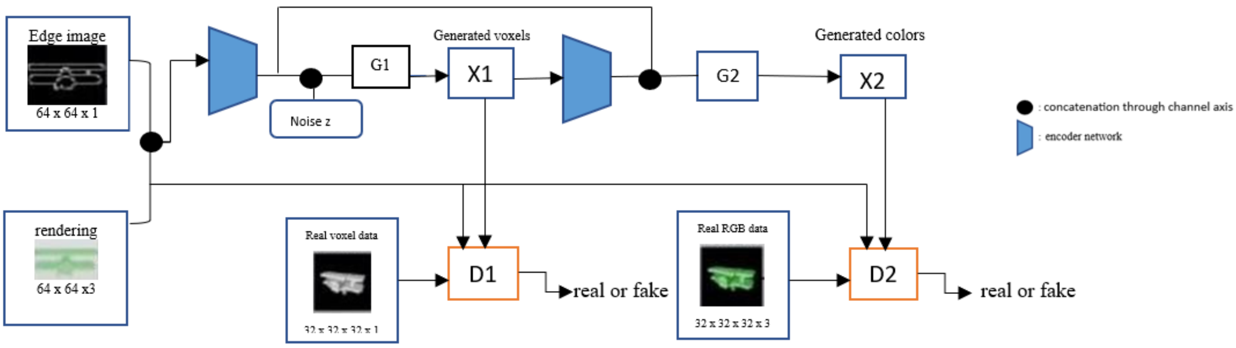


**Fig 1: Architecture**

## 5. RESULTS

The experiments conducted aimed to assess the performance of the proposed approach by analyzing the relationship between the input and output of the system. The input provided to the SGAN framework consisted of sketch-based inputs, which were represented as 2D sketches capturing the desired 3D shapes. These sketches served as the initial point for the generation process. They were fed into the system to initiate the generation of corresponding 3D shapes.

The output of the SGAN framework was the generated 3D shapes that corresponded to the input sketches. The stacked GAN architecture was utilized to progressively refine the generated shapes, incorporating high-level details and enhancing their visual fidelity. Through an iterative process, the system iteratively improved the generated shapes, ensuring the generation of complex and realistic 3D models that closely resembled the designs represented by the input sketches.

The results obtained from the experiments demonstrated the effectiveness of the proposed approach in generating high-quality and visually appealing 3D shapes from sketch-based inputs. The iterative refinement process facilitated by the stacked GAN architecture allowed for the capture of intricate details and the production of realistic and complex 3D models. These generated shapes exhibited a close correspondence to the intended designs represented by the input sketches, showcasing the capability of the SGAN framework in facilitating intuitive and natural exploration of 3D models based on sketch inputs.

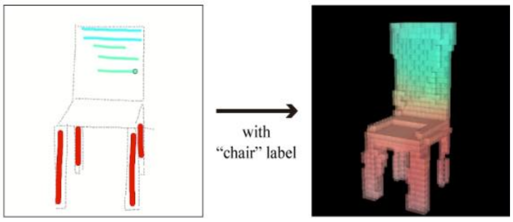


**Fig 2: Image with Chair Label**

Fig 2: The user provides a 2D sketch as input by selecting a label chair, which is fed into stacked GANs. The GANs generate a 3D model from the input sketch. The final output is a 3D model that closely matches the input sketch.

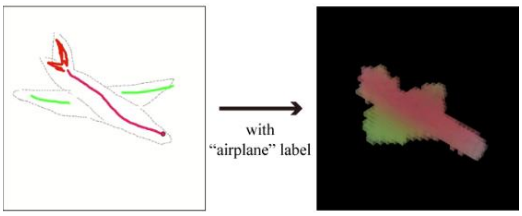


**Fig 3: Image with Airplane Label**

In the figure 4.1.2, the user provides a 2D sketch as input by selecting a label airplane, which is fed into stacked GANs. The GANs generate a 3D model from the input sketch. The final output is a 3D model that closely matches the input sketch.

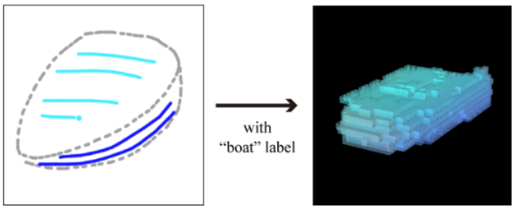


**Fig 4: Image with Boat Label**

In the figure 4.1.3, the user provides a 2D sketch as input by selecting a label boat, which is fed into stacked GANs. The GANs generate a 3D model from the input sketch. The final output is a 3D model that closely matches the input sketch.

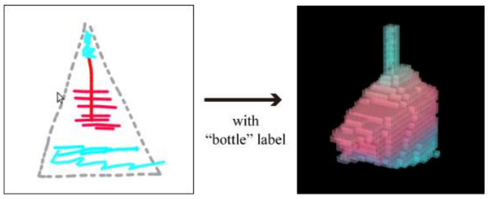


**Fig 5: Image with Bottle Label**

In the figure 4.1.4, the user provides a 2D sketch as input by selecting a label bottle, which is fed into stacked GANs. The GANs generate a 3D model from the input sketch. The final output is a 3D model that closely matches the input sketch.

## 6. CONCLUSION

This model introduces a novel approach for sketch-based 3D exploration using Stacked Generative Adversarial Networks (SGANs). The proposed framework successfully combines the power of GANs with sketch-based inputs, enabling intuitive and natural interaction with 3D modeling systems. The experiments demonstrate the effectiveness of the SGAN framework in generating high-quality and visually pleasing 3D shapes that closely resemble the input sketches.

Future research directions include incorporating semantic information, exploring alternative network architectures, addressing occlusion challenges, and improving generalization to unseen object categories. The proposed approach holds promise for advancing the field of sketch-based 3D exploration and has potential applications in computer vision, virtual reality, and related domains.

## 7. REFERENCES

[1] Jiang, Li, Shaoshuai Shi, Xiaojuan Qi, and Jiaya Jia. "Gal: Geometric adversarial loss for single-view 3d-object reconstruction." In Proceedings of the European conference on computer vision (ECCV), pp. 802-816. 2018.

[2] Liu, Li, Fumin Shen, Yuming Shen, Xianglong Liu, and Ling Shao. "Deep sketch hashing: Fast free-hand sketch-based image retrieval." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2862-2871. 2017.

[3] Xu, Peng, Chaitanya K. Joshi, and Xavier Bresson. "Multigraph transformer for free-hand sketch recognition." IEEE Transactions on Neural Networks and Learning Systems 33, no. 10 (2021): 5150-5161.

[4] Wang, Nanyang, Yinda Zhang, Zhuwen Li, Yanwei Fu, Wei Liu, and Yu-Gang Jiang. "Pixel2mesh: Generating 3d mesh models from single rgb images." *In* Proceedings of

the European conference on computer vision (ECCV), pp. 52-67. 2018.

[5] Wang, Lingjing, Cheng Qian, Jifei Wang, and Yi Fang. "Unsupervised learning of 3D model reconstruction from hand-drawn sketches." In Proceedings of the 26th ACM international conference on Multimedia, pp. 1820-1828. 2018.

[6] Su, Wanchao, Dong Du, Xin Yang, Shizhe Zhou, and Hongbo Fu. "Interactive sketch-based normal map generation with deep neural networks." Proceedings of the ACM on Computer Graphics and Interactive Techniques 1, no. 1 (2018): 1-17.

[7] Lun, Zhaoliang, Matheus Gadelha, Evangelos Kalogerakis, Subhransu Maji, and Rui Wang. "3D shape reconstruction from sketches via multi-view convolutional networks." In 2017 International Conference on 3D Vision (3DV), pp. 67-77. IEEE, 2017.

[8] Delanoy, Johanna, Mathieu Aubry, Phillip Isola, Alexei A. Efros, and Adrien Bousseau. "3d sketching using multi-view deep volumetric prediction." Proceedings of the ACM on Computer Graphics and Interactive Techniques 1, no. 1 (2018): 1-22.

[9] Sun, Xingyuan, Jiajun Wu, Xiuming Zhang, Zhoutong Zhang, Chengkai Zhang, Tianfan Xue, Joshua B. Tenenbaum, and William T. Freeman. "Pix3d: Dataset and methods for single-image 3d shape modeling." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2974-2983. 2018..