Journal of
Advancements in Robotics
ISSN: 2455-1872
Volume 12, Issue 3, 2025
September–December
DOI (Journal): 10.37591/JoARB
https://journals.stmjournals.com/joarb
Review JoARB

Generative AI for VR: Creating Physically Realistic Models

Varun Bhilare[1,*], Pranita Chaudhary[2], Aayush Bodkhe[1], Abhishek Bichare[1], Shushant Bhat[1]

Abstract

Virtual Reality has revolutionized the traditional learning system by creating and interactive and engaging environment. However, its ability to show precise real-world experiences is limited due to lack of physical realism. This study investigates the potential of Generative Adversarial Network (GAN) in creating physically realistic 3D models. Proposed system incorporates deep learning techniques along with physics-based constraints to enhance model's accuracy and usability. To achieve this, experiments were conducted on Shapenet dataset, various preprocessing were done to standardize the point cloud dataset. A multicomponent GAN architecture was used to generate the 3D model. It included spatial transformer as well as a semantic segmentation network. The results demonstrated by the model included good Intersection-over-Union (IoU) score and the dice coefficient over various categories. Comparative evaluation included much improvement in accuracy and realism than conventional method. Moreover, the qualitative analysis showcases clear segmentation and improved interactivity. This highlighted the system's potential in bridging the gap between the virtual and real-world dynamics. All these findings indicate that the GAN generated models can provide more effective and indulging VR based learning experience. The study concludes that including deep learning along with physics-based constraints offers a promising pathway for advance educational applications and providing deep understanding and engagement among learners.

Keywords: Point cloud, generative AI, virtual reality (VR), physics simulation, GAN, VAE

INTRODUCTION

Impact of VR on Education

VR is changing education by giving a practical environment for training that is expensive or impractical in the real-world scenario. This has given students the opportunity to practice in an interactive and controlled environment, be it medical students performing surgical techniques on virtual patients or manufacturing industries training workers on safety procedures in simulated hazardous environments. VR acts as a major impacting technology for developing confidence and competence in skills acquisition pertinent to medicals, engineering, and industrial safety.

Challenges in Realistic Modeling for VR Training

The success of VR training relies upon the composing of 3D models that are visually detailed and physically accurate. However, this is fraught with challenges, such as fidelity of textures, responsiveness to materials, and physical properties accurately represented to include elasticity and mass. Candidly inaccurate models disturb the immersion and frustrate successful training. The height of each resulting problem emphasizes the

***Author for Correspondence**
Varun Bhilare
E-mail: varun.bhilare22@pccoepune.org

[1]Student, Department of Computer Science & Engineering (Artificial Intelligence and Machine Learning), Pimpri Chinchwad College of Engineering, Pune, Maharashtra, India
[2]Assistant Professor, Department of Computer Science & Engineering (Artificial Intelligence and Machine Learning), Pimpri Chinchwad College of Engineering, Pune, Maharashtra, India

Received Date: May 16, 2025
Accepted Date: June 24, 2025
Published Date: September 19, 2025

Citation: Varun Bhilare, Pranita Chaudhary, Aayush Bodkhe, Abhishek Bichare, Shushant Bhat. Generative AI for VR: Creating Physically Realistic Models. Journal of Advancements in Robotics. 2025; 12(3): 14–22p.

urgency under which developing better methods to increase model realism and interactivity should be considered.

Generative AI for Physics Integration

Physical Realism in 3D Models: General AI uses GANs, VAEs, and NeRFs techniques to invoke the virtual models, with assertions of giving enhanced visual realism to such models.

GANs

Generate highly detailed textures and shapes, such as realistic skin or organic materials, to better mesh the expected quality of VR models.

VAEs

They make it versatile so that the shape and size of objects like anatomical structures or industrial tools may be represented quite differently for robust model training.

Physics-Informed

The model will constitute physical characteristics such as elasticity and flexibility, which guarantees that whenever virtual assets interact with one another, they behave in a life-like manner.

NeRFs

Based on real physics and a smart presentation of how light interacts with the surfaces to create realistic reflections and shadows of such surfaces for real-time training in safety-critical environments.

Research Objectives

1. The objective of this study is to investigate how generative AI can create physically realistic models, specifically for VR-based educational simulations.
2. The goal is to create methods that can produce 3D assets that not only look realistic but also exhibit actual physical characteristics like elasticity, texture, mass etc. that are useful in training scenarios that involve hands-on practice.
3. Through solving these challenges, we expect to bring the gap between the VR and real world training scenarios closer, so that VR can be used as a realistic way to immerse into skill-based learning.

LITERATURE REVIEW

The fast evolving development in Generative AI, particularly, in the 3-D model generation, have created a way for creating virtual environments, with realistic interactive objects. Numerous studies have explored the application of generative models for producing excellent quality 3-D models; however, one challenge still remains, to create models that are virtually realistic and capable of accurately integrating and interacting with the VR. Various models such as Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and more recent transformer-based approaches have become popular and are highly known for adapting high-quality images and 3-D objects. For example, the 3D-GAN introduced by Wu *et al*. was one of the first models to extend GANs to 3D space, generating objects in voxel form. While this showcased the practicality of 3-D generation, it did not demonstrate the voxel-based representation's limited resolution and fine detail, making it unsuitable for high-fidelity VR applications where detail is crucial [1]. Achlioptas *et al*. improved on this by proposing GAN-based models that work with point clouds, significantly increasing resolution and detail. Point cloud representations enable intricate models essential for VR applications, but they often lack surface realism and texture, limiting their practicality in realistic immersive environments [2]. Similarly, to solve the above issues PointNet and PointNet++ architectures brought advancements in processing point clouds for 3D object generation, but these models also face challenges in generating high-fidelity textures necessary for VR [3]. Mesh-based generative models like Mesh-RCNN and AtlasNet produce surface meshes with improved surface detail, yet they are processor intensive and struggle to balance detail with realistic physical structures, making them less suitable for VR, which demands a balance of both visual fidelity and interactive responsiveness [4].

Generative AI has influenced VR applications across sectors like gaming, architecture visualization, and experiential learning, where realistic and varying objects are essential. In medical, VR training simulations have created 3-D models of human anatomy, allowing students to train the most intricate procedures and view complex biological structures. However, there is a drawback that these models lack dynamic properties like mimicking tissue behavior under stress [5]. Physics-based virtual reality in education simulations make use of physics engines to teach concepts of mechanics and fluid dynamics, even though they are based on fixed models that constrain different kinds of adaptivity and variety. Although VR applications in engineering and architecture are meant to simulate 3D layouts and structures, they lack interacting components and are therefore limited in their usefulness as visual demonstrations. Although combining generative modelling with physics engine shows promise, it still has not passed experimental stages [6]. In education, VR simulations show the ability to interact with objects in a realistic, physics based manner which is paramount to realistic experiences and effective learning. There are, however, still some remaining limitations of generative AI models for VR, specifically in the realms of realism and interactivity, which restrict their use within hands-on educational settings [7].

Some major limitations of generative models for VR applications persist, particularly within educational contexts. Current models are devoid of realistic physical properties, such as elasticity, rigidity, and collision responses, all of which are important for immersive educational simulations. However, due to the lack of computing power and complexity of surface representation in formats like voxels or point cloud, many generative models fail to produce fine details. In addition, models existing today do not sufficiently support user interaction capabilities in VR-like environments, where objects are meant to interact realistically with user inputs such as grabbing, rotating or applying force [8]. Another limitation in common applications is the lack of generalization and diversity, as the current models often aim at generating single-object models, hence limiting them for use in multiple educational scenarios. Based on the above-mentioned gaps, our work integrates advancements in point cloud and GAN based models to make high-resolution and structurally valid 3D models optimal to be introduced in VR. Our study contributes to the growing body of literature that seeks to enhance generative AI's capabilities in VR by generating detailed models using scale methods [9]. Our contributions include the generative model tailored toward high structure fidelity and surface detail in VR; a design approach that would allow further integration of physics-based interactions; paving a path for potential scalability to generate various scalable solutions adaptable to different content requirements.

A key opportunity for developing a solution tailored to the educational needs lies in the space created by limitations created by current generative AI models for VR applications. We build upon recent advancements in pointcloud and GAN-based models while aspiring to tackle educational VR simulations' unique challenges. By focusing on generating high-resolution, structurally sound models that can integrate the properties of physics in future stages, our research contributes to a growing body of work seeking to expand the capabilities of generative AI in VR [10].

This study's contributions include:

1. A generative model customized to produce 3D models with high structural fidelity and accurate surface detail optimized for VR environments.
2. A design approach that allows for future integration of physics-based interactions, aiming toward applications in hands-on educational simulations.
3. Potential scalability in generating a variety of models that are adaptable to various educational content requirements.

RESEARCH GAP

Although generative AI has made instant progress in creating realistic 3D models, a serious gap still exists between the creation of models that are visually accurate and those that are physically realistic and interactive in virtual reality (VR) settings. Most of the existing methods including voxel-, mesh-, and point cloud-based approaches are more oriented toward geometry and appearance of essential physics-based constraints toward real-world simulation [11].

Thus, current generative techniques seldom deal with object part segmentation or any physical behavior-deformation, elasticity, and motion under force, that should contribute significantly to training simulations in terms of effectiveness and immersion. Hence, applications that try to implement these models in VR never provide an experience that is truly interactive, responsive, and reflective of real-world interactions [12].

A generative framework capable of blending the advantages of deep learning with physics-informed modeling is clearly in demand for generating segmented, physically aware 3D objects suitable for educational simulation, this project tries to fulfill that gap by proposing a novel GAN model trained with point cloud data to generate interactive and physically plausible 3D shapes for immersive VR learning [13].

PROPOSED METHODOLOGY

Data Processing and Representation

For generation of the 3D model, we use the point cloud dataset from the ShapeNet-core dataset. This serves as the foundation for training the GAN.

Point Cloud Representation

Point clouds are 3d space data points that represent the surface of an object. Each point consists of three coordinates (x, y, z), which are useful in capturing detailed and complex surfaces.

Segmentation

Each model is segmented meaningfully (say, car: handle, body, wheel). This allows depiction of part specific physical properties. This segmentation allows more realistic interaction in VR as each segment can respond to physical forces independently.

Preprocessing

1. *Normalization*: All data points are scaled and centered within a standardized 3D space, facilitating uniform model generation.
2. *Augmentation*: Various augmentation techniques like random rotation, translation, scaling, are applied to increase the models' generalization ability across various shapes and configurations.

Model Architecture

Our model is based on GAN architecture, specifically designed to produce 3D point clouds. The model's architecture consists of a discriminator and a generator, each designed to work together with the point cloud data.

Generator Network

The generator takes latent vector(z) as input. z represents the random noise, which is then transformed into a structured 3d point cloud representing the object (Figure 1). Key components of the generator are:

1. *Fully Connected Layers*: Initial layers are fully connected, mapping latent vectors to a higher-dimension.
2. Convolutional Layers: These layers apply convolutions on the 3D points, refining the models' shape and producing more structured point clouds.
3. *Output Layer with Segmentation*: The output layer generates the final point cloud. Each point is assigned to a label (material). This allows us to apply physics-based constraints to each part specifically.

Journal of Advancements in Robotics
Volume 12, Issue 3
ISSN: 2455-1872

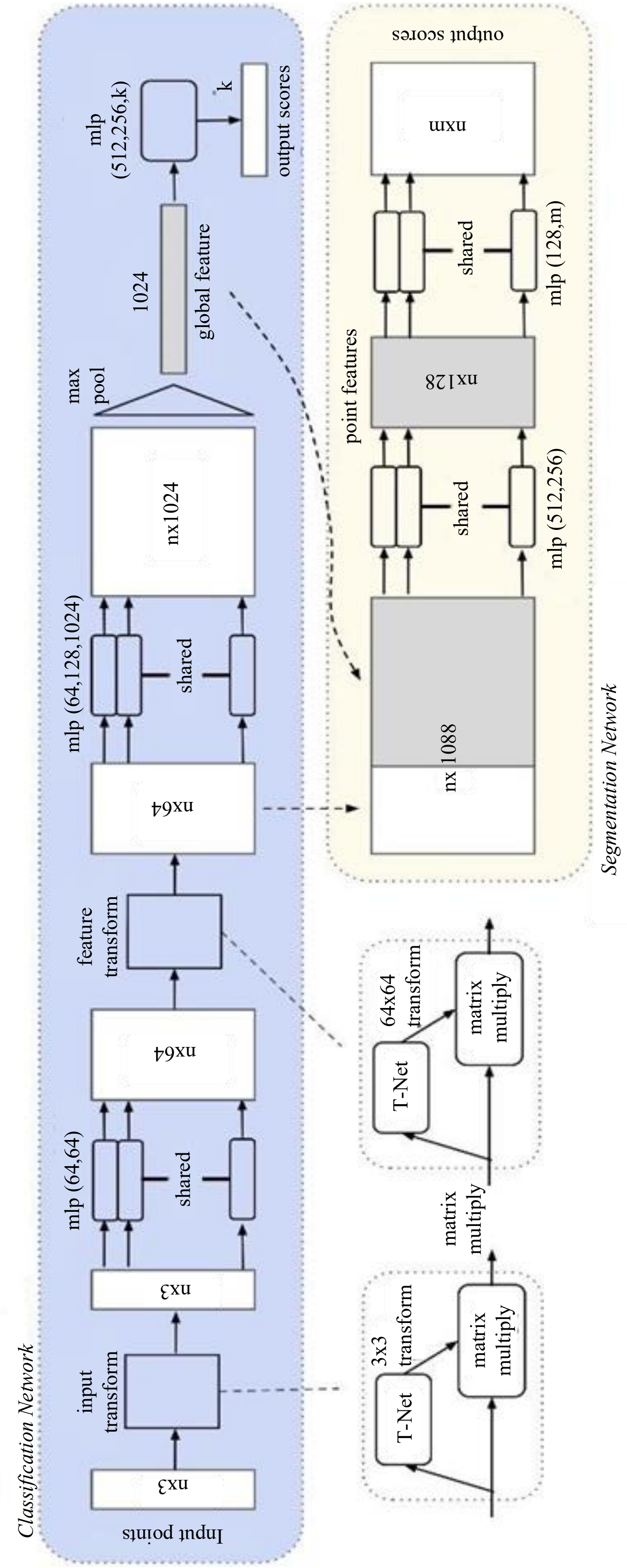


Figure 1. Flowchart classification and segmentation network.

Discriminator Network

The discriminator evaluates the realism of the models generated. It distinguishes between real and generated point clouds.

1. *PointNet-Based Layers*: The discriminator is built on PointNet architecture. This architecture captures both local and global features of the point cloud. It helps in identifying the fine details to check whether the generated object is real or not.
2. *Binary Classification Output:* The discriminator's output is a binary classification, indicating whether the input point cloud is authentic (from the dataset) or not.

Physics-Informed Constraints: It includes the integration of physics-informed constraints into the GAN's loss function. This ensures that the generated model not only looks realistic but also behaves according to physical principles when placed in a VR environment.

Adversarial Loss

GAN loss function optimizes the generator to create point clouds that discriminator classifies as real or not. If D(x) is the output for sample x and G(z) is generator output for a noise vector z. Then the adversarial loss for the generator LG and discriminator LD are defined as:

$$L_D=-E_{x\sim Pdata}[\log D(x)]-E_{z\sim pz}[\log(1-D(G(z)))]$$
$$L_G=-E_{z\sim pz}[\log D(G(z))]$$

Physics-Informed Loss

This component applies constraints which ensures the generated object maintains structural and physical consistency under various forces applied. For example, we apply elasticity constraints that model how segments of the object respond to stress and strain, based on Hooke's law:

$$\sigma=E.\epsilon$$

Where:

- σ is the stress on a point within a segment,
- E is the elasticity modulus, and
- ϵ is the strain (deformation) in response to applied forces.

The physics-informed loss $L_{physics}$ is calculated based on deviations from expected physical behavior, with penalties applied for unrealistic deformation under simulated forces. By including this loss, the generator learns to produce objects that not only appear structurally sound but also behave as expected under physical constraints in a VR setting.

Training and Optimization

Training the GAN involves iterative optimization of both the generator and discriminator, with additional tuning of the physics-informed constraints.

Training Process

We employ a two-stage training approach, where the model first learns basic shape generation and then fine-tunes with physics-informed constraints. The training uses a batch size of 32 and is optimized with the Adam optimizer, with learning rates of 0.0002 for both the generator and discriminator.

Initial Pre Training for Visual Fidelity

The training process begins with GANs and VAEs focused solely on generating visually realistic textures and shapes. GANs are pre trained to produce high-resolution, photorealistic textures by minimizing the adversarial loss (as outlined previously) until they can reliably create surfaces that visually resemble real-world materials. Simultaneously, VAEs are trained to map objects to a latent space that captures a range of plausible shapes. Here, the primary focus is on achieving a baseline visual quality that matches target images in terms of texture, color accuracy, and structural variation.

Metrics such as Structural Similarity Index (SSIM) and Inception Score (IS) are used to evaluate the visual quality of the generated models. These metrics guide adjustments to the learning rate, generator-discriminator balance, and VAE latent space regularization to optimize visual accuracy.

Loss Combination

The final loss for the generator combines adversarial and physics-informed losses:

$$L_{total}=\alpha L_G+\beta L_{physics}$$

Where, α and β are weights that balance the contributions of visual fidelity and physical realism. Through experimentation, we find that setting α=0.8 and β=0.2 yields optimal results.

Physics-Informed Training for Physical Realism

After pretraining, physics-informed constraints are added using Physics-Informed Neural Networks (PINNs) and differentiable physics simulations. This enables the model to embed physical properties like elasticity and density, allowing backpropagation through simulated behaviors to guide realistic physical responses during training.

$$L_{physics}={}_i\sum(F_{simulated}{}^{(i)}-F_{real}{}^{(i)})^2$$

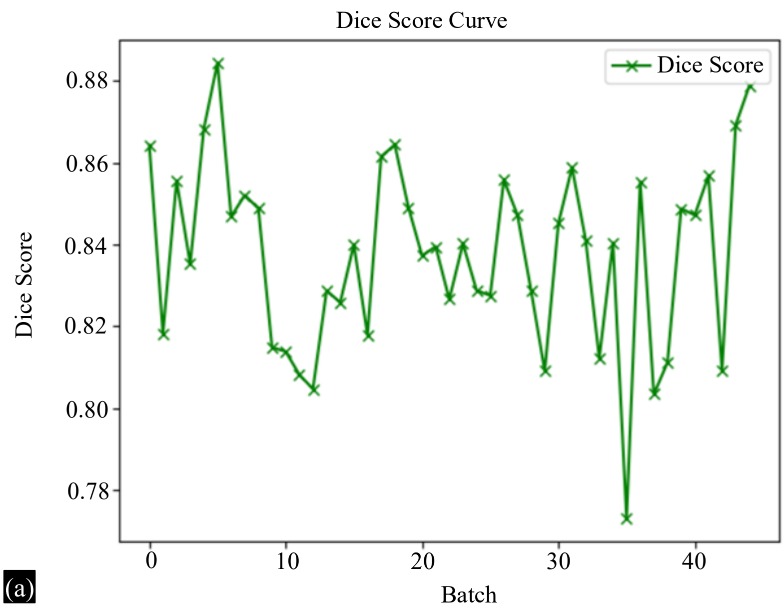
RESULTS AND ANALYSIS

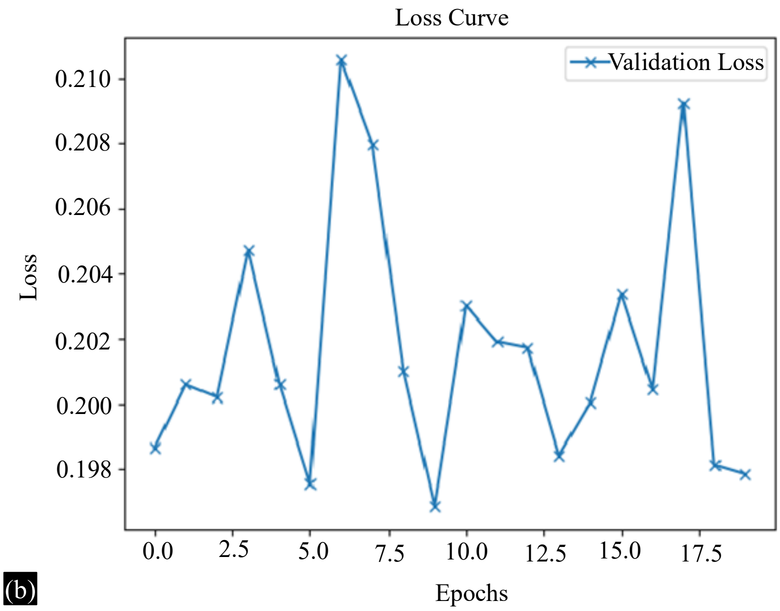
The outcomes of the proposed generative AI system for 3D model generation and semantic segmentation: The analysis includes performance metrics, statistical evaluations, detailed discussions, and graphical visualizations to provide a comprehensive understanding of the model's capabilities and limitations (Figure 2).

1. *Performance Measures:* To evaluate the model, several key performance metrics were employed:
2. *Dice Score (F1 Score):* A similarity metric for evaluating segmentation accuracy:

$$\text{Dice Score}=\frac{2\times|\text{Prediction} \cap \text{Ground Truth}|}{|\text{Prediction}| + |\text{Ground Truth}|}$$

Where P is the predicted set of points and G is the ground truth.





(b)

Figure 2. Dice score and loss curve.

The Dice Score Curve shows performance fluctuating between 0.78 and 0.88 across batches, with stable scores above 0.8 overall, indicating good segmentation accuracy. Early variability suggests the model's adaption, while a sharp recovery towards the end highlights improvements in optimization.

CONCLUSION

The findings indicate the capability of developing highly realistic, tailored 3D models for VR applications through the integration of generative AI with physics-informed training, with a particular emphasis on educational simulations. This study represents a framework using techniques such as Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and Neural Radiance Fields (NeRFs) to generate models with both visual realism and proper physical functionalities according to the laws of physics. By incorporating physics-informed loss functions, this method effectively creates a nexus between visual realism and the accuracy of physics. Consequently, the VR models can portray realistic behaviors such as reaction to virtual forces, deformations, and light interactions. User evaluations within the VR demonstrated that the models offer a highly immersive, interactive experience that significantly increase the educational value of the simulations through their close correspondence to real-world materials and behaviors.

Key Contributions and Future scope

1. *Refined Physical Realism:* The incorporation of physics-informed loss functions allows generative models to simulate realistic mechanics such as elasticity and material compliance, which are particularly needed within VR simulations in education and training.
2. *High Visual realism through NeRFs*: The use of Neural Radiance Fields gave more refined lighting effects and surface detailing that went a long way in building immersive models.
3. *User-Centered Evaluation:* User testing in VR helped to demonstrate the effectiveness of the models in simulating real-world interactions, confirming that generative AI can support high-fidelity educational environments.

Future Directions

The framework has already shown its effectiveness to the standardized VR interactions, yet, it still deals with more complex forces and multi-object interactions. Nevertheless, looking forward and on the whole, there are certain areas that could serve as a driving force behind its capabilities:

1. *Physics in VR at a Higher level:* An AI agent that has been trained on ultra-precise physics models would overcome the errors in VR-caused by physics inaccuracies.

2. *Computational Expertise:* By the way, do you know if it is possible to optimize the designed system to become more scalable with the VR range of applications that it fits into?
3. *Including Material Characteristics:* By expanding the framework to include the simulation of other material behaviors (e.g. thermal conductivity and electrical properties), it gives the opportunity for more specific training simulations.

By the way, this research not only reports on the cutting-edge achievements of the necessary digital blanket but opens the perspective of generative AI as a big example of VR learning that will change the current state of electronics. The transformation of the training both visually and physically could open up future great innovations in this area that would reshape the training background in areas like medicine, engineering, and environmental sciences, making them more experimental, as well as effective.

REFERENCES

1. Wu J, Zhang C, Xue T, Freeman B, Tenenbaum J. Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling. Advances in neural information processing systems (29th conf). 2016; 1–9.
2. Achlioptas P, Diamanti O, Mitliagkas I, Guibas L. Learning representations and generative models for 3d point clouds. In International conference on machine learning, PMLR. 2018 Jul 3; 40–49.
3. Qi CR, Su H, Mo K, Guibas LJ. Pointnet: Deep learning on point sets for 3d classification and segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2017; 652–660.
4. Groueix T, Fisher M, Kim VG, Russell BC, Aubry M. A papier-mâché approach to learning 3d surface generation. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2018; 216–224.
5. Chheang V, Sharmin S, Márquez-Hernández R, Patel M, Rajasekaran D, Caulfield G, Kiafar B, Li J, Kullu P, Barmaki RL. Towards anatomy education with generative AI-based virtual assistants in immersive virtual reality environments. In 2024 IEEE international conference on artificial intelligence and eXtended and virtual reality (AIxVR). 2024 Jan 17; 21–30.
6. Sung NJ, Ma J, Choi YJ, Hong M. Real-time augmented reality physics simulator for education. Appl Sci. 2019 Sep 25; 9(19): 4019.
7. Vanfretti L, Laughman CR, Chakrabarty A. Integrating Generative Machine Learning Models and Physics-Based Models for Building Energy Simulation. In Modelica Conferences. 2024; 178–188.
8. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. Advances in neural information processing systems (NIPS27). 2014.
9. Kingma DP, Welling M. Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114. 2013 Dec 20.
10. Mildenhall B, Srinivasan PP, Tancik M, Barron JT, Ramamoorthi R, Ng R. Nerf: Representing scenes as neural radiance fields for view synthesis. Commun ACM. 2021 Dec 17; 65(1): 99–106.
11. Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys. 2019 Feb 1; 378: 686–707.
12. Zhang R, Isola P, Efros AA, Shechtman E, Wang O. The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2018; 586–595.
13. Saito S, Huang Z, Natsume R, Morishima S, Kanazawa A, Li H. Pifu: Pixel-aligned implicit function for high-resolution clothed human digitization. In Proceedings of the IEEE/CVF international conference on computer vision. 2019; 2304–2314.

