Real-Time 3d Scene Generation using Optimized Neural Radiance Fields

Dr.R.Atul Kumar, Manepally Preethika, Kalakunta Srilochan, and Mogulla Hari Charan Reddy

Department Of CSE (AI & ML), Ace Engineering College, Hyderabad, Telangana, India.

ABSTRACT:

High-quality 3D scene reconstruction is challenging, especially when only a limited number of images are available. Neural Radiance Fields (NeRFs) enable the representation of a scene as a continuous volumetric function, allowing novel views to be synthesized from sparse inputs. The scene is encoded in a fully-connected neural network that takes 3D coordinates and viewing directions as input and outputs density and color values. Differentiable volume rendering projects these predictions into 2D images, and techniques such as positional encoding and hierarchical sampling improve detail capture and computational efficiency. This approach allows accurate, high-fidelity visualization of complex real-world scenes while avoiding the memory costs of traditional voxel-based methods.
Keywords: Neural Radiance Fields (NeRF), 3D scene reconstruction, differentiable volume rendering, positional encoding, hierarchical sampling, view synthesis, implicit representation

INTRODUCTION:

Three-dimensional (3D) scene reconstruction is a key task in computer vision and graphics, enabling machines to understand and recreate the physical world from two-dimensional (2D) images. It has applications in augmented reality, robotics, virtual tourism, and film production, where generating realistic novel views from limited inputs is essential. Traditional 3D reconstruction methods, such as voxel grids, meshes, or point clouds, face challenges in scalability, memory efficiency, and handling complex visual effects like reflections or transparency.
Neural Radiance Fields (NeRF) introduce a breakthrough by representing a scene as a continuous volumetric function using deep neural networks. Instead of explicitly modeling geometry, NeRF implicitly encodes scene structure and appearance in the weights of a multilayer perceptron (MLP). By taking 3D coordinates and viewing directions as input and outputting color and density values, NeRF enables photorealistic rendering of novel viewpoints. The use of differentiable volume rendering allows end-to-end optimization directly from a sparse set of RGB images, while positional encoding and hierarchical sampling improve fine detail and computational efficiency. This project focuses on understanding and implementing the NeRF framework to achieve accurate, memory-efficient, and high-fidelity 3D scene visualization.

LITERATURE REVIEW:

The rapid development of deep learning has significantly advanced the field of 3D scene representation and reconstruction. Researchers have explored various neural implicit representation techniques to model complex 3D structures efficiently. These approaches aim to overcome the limitations of traditional voxel, mesh, and point-cloud based methods in terms of memory usage and scalability. Recent studies focus on learning continuous scene representations using neural networks trained directly from image data. Methods such as implicit functions, neural texture spaces, and differentiable volumetric rendering have improved reconstruction accuracy and visual realism. These contributions collectively laid the foundation for modern frameworks like Neural Radiance Fields (NeRF) that enable high-quality novel view synthesis from sparse inputs.

Table 1 : Comparative Analysis of Existing Systems

ISSN: 2583-6129
DOI:10.55041/ISJEM06020

Author(s) & Year	System/Title	Technology Used	Scope	Limitations
Kyle Genova et al., 2020	Local Deep Implicit Functions for 3D Shape	Deep implicit neural functions, surface template anchoring, neural networks	Enables accurate 3D shape representation and reconstruction using localized implicit functions	Limited scalability for very large scenes and requires template alignment
Philipp Henzler et al., 2020	Learning a Neural 3D Texture Space from 2D Exemplars	Neural networks, latent texture space learning, 3D texture synthesis	Generates continuous 3D textures from 2D images for realistic rendering	Focuses mainly on texture synthesis rather than full geometry reconstruction
Chen-Hsuan Jiang et al., 2020	Local Implicit Grid Representations for 3D Scenes	Latent vector grids, autoencoders, implicit neural fields	Efficient representation of large-scale 3D scenes with improved memory efficiency	Training complexity increases with higher scene resolution
Gabriel Rainer et al., 2020	Unified Neural Encoding of Bidirectional Texture Functions (BTFs)	Neural encoding, Bidirectional Texture Functions, appearance modeling	Compact representation of material appearance under varying lighting and viewing directions	Primarily models surface appearance rather than full 3D geometry
Michael Niemeyer et al., 2019	Differentiable Volumetric Rendering: Learning Implicit 3D Representations without 3D Supervision	Differentiable volumetric rendering, implicit neural representations, gradient optimization	Enables learning 3D geometry directly from 2D images without 3D supervision	High computational cost and slower rendering during training

[1] Title: Local Deep Implicit Functions for 3D Shape.

Author: Kyle Genova, Forrester Cole, Daniel Vlasic, Aaron Sarna, William T. Freeman, and Thomas Funkhouser (2020)

Genova and colleagues introduced a deep learning approach that anchors local implicit functions to a surface template to represent complex 3D shapes. Their method effectively captures fine geometric structures and maintains robustness to topological changes, providing detailed and accurate reconstructions. This study demonstrated that localized implicit modeling improves flexibility and scalability in 3D representation, paving the way for efficient neural shape encoding techniques that influence modern neural scene representation methods.

[2] Title: Learning a Neural 3D Texture Space from 2D Exemplars
Author: Philipp Henzler, Niloy J. Mitra, and Tobias Ritschel (2020)

Henzler and co-authors proposed a neural network framework that learns a latent 3D texture space directly from 2D exemplars. The system synthesizes continuous 3D textures for arbitrary geometries, allowing realistic and seamless texture mapping in 3D models. Their approach successfully bridges 2D image data and 3D surface representation, contributing to high-quality visual rendering and advancing the field of neural texture synthesis.

[3] Title: Local Implicit Grid Representations for 3D Scenes

Authors: Chen-Hsuan Jiang, Avneesh Sud, Amit Makadia, Jingwei Huang, Matthias Nießner, and Thomas Funkhouser (2020)

Jiang and his team presented an approach that represents large-scale 3D scenes using a grid of latent vectors encoding local implicit fields. Trained through an autoencoder, this model efficiently captures detailed scene information while maintaining low memory usage. The method demonstrated superior scalability and reconstruction fidelity compared to voxel-based systems, establishing an effective framework for detailed scene representation at reduced computational costs.

[4] Title: Unified Neural Encoding of Bidirectional Texture Functions (BTFs)

Author: Gabriel Rainer, Abhijeet Ghosh, Wenzel Jakob, and Tim Weyrich (2020)

Rainer and collaborators developed a neural network model to encode Bidirectional Texture Functions (BTFs), which describe how surface appearance changes with varying lighting and viewing directions. Their work provided a compact and continuous representation of complex materials, allowing for realistic and memory-efficient rendering. This unified neural encoding framework significantly improved the quality and efficiency of texture-based rendering techniques.

[5] **Title:** Differentiable Volumetric Rendering: Learning Implicit 3D Representations without 3D Supervision

Author: Michael Niemeyer, Lars Mescheder, Michael Oechsle, and Andreas Geiger (2019)

Niemeyer and his team introduced a differentiable volumetric rendering pipeline that connects implicit 3D representations with 2D image data. Their method learns scene geometry directly from 2D images without requiring explicit 3D ground-truth supervision, using gradient-based optimization to align image projections with underlying volumetric representations. This research provided the conceptual foundation for NeRF, demonstrating the power of differentiable rendering in reconstructing accurate 3D geometry from sparse image input.

PROPOSED SYSTEM:

The proposed system implements an optimized **Neural Radiance Field (NeRF)** framework to generate high-quality 3D scenes from a limited set of 2D images. Instead of using traditional 3D representations such as voxel grids or meshes, the system models the scene as a continuous volumetric function using a deep neural network. This approach allows efficient memory usage while maintaining high visual fidelity.

In the proposed method, multiple images of a scene captured from different viewpoints are used as input. Each pixel from these images is projected into a 3D space using camera parameters to obtain spatial coordinates and viewing directions. These coordinates are then encoded using positional encoding to capture high-frequency details and passed into a multilayer perceptron (MLP). The neural network predicts the color (RGB) and density values for every sampled point along a camera ray.

To render the final image, differentiable volume rendering is applied, which integrates the predicted color and density values along each ray to synthesize realistic pixel values. The training process minimizes the difference between the rendered images and the original input images, allowing the network to learn an accurate representation of the scene.

To improve performance and efficiency, the system incorporates hierarchical sampling to allocate more samples to important regions of the scene, enabling better detail reconstruction while reducing computational cost. This optimized NeRF model can generate photorealistic novel views of the scene from viewpoints that were not originally captured.

Overall, the proposed system provides a memory-efficient and scalable solution for real-time 3D scene visualization, making it suitable for applications such as virtual reality, augmented reality, digital twins, and immersive environment reconstruction.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 04 | April – 2026 DOI:10.55041/ISJEM06020
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

SYSTEM ARCHITECTURE:

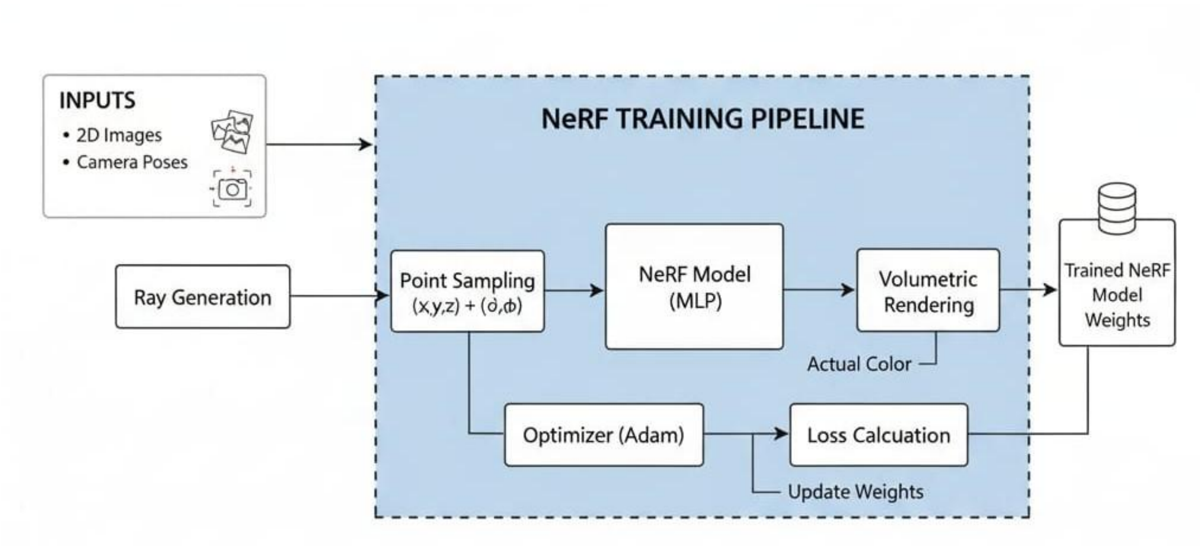

Figure 1 : System Architecture

RESULT AND DISCUSSION:

The implemented Neural Radiance Field (NeRF) model successfully reconstructs high-quality 3D scenes from a limited set of 2D images. During experimentation, the system was able to learn the spatial structure and appearance of the scene by training the neural network on images captured from multiple viewpoints. The rendered outputs demonstrated smooth geometry representation and accurate color reproduction. The use of positional encoding helped the network capture fine details of the scene, while hierarchical sampling improved the efficiency of ray sampling during the rendering process.

The experimental results show that the proposed system generates realistic novel views that closely resemble the original scene. Compared with traditional voxel-based or mesh-based reconstruction methods, the NeRF-based approach provides higher visual fidelity while using less memory for scene representation. The differentiable volume rendering mechanism allowed the model to optimize the scene representation directly from RGB images without requiring explicit 3D supervision. This capability significantly improves flexibility in real-world scenarios where collecting 3D ground-truth data is difficult.

However, the system also has certain limitations. Training the NeRF model requires significant computational resources and time, especially when dealing with high-resolution images or complex scenes. Although the optimized techniques improve rendering efficiency, real-time performance can still be challenging without powerful GPUs. Future improvements may focus on faster training methods, lightweight neural architectures, and integration with real-time rendering techniques to make NeRF-based systems more practical for interactive applications such as augmented reality and real-time simulation.

CONCLUSION AND FUTURE SCOPE:

This project presented a Neural Radiance Field (NeRF)-based approach for 3D scene reconstruction and novel view synthesis. The system successfully implemented the complete pipeline, including dataset preparation, model training, checkpoint management, and rendering of test views. By leveraging volumetric rendering and neural network-based scene representation, the model was able to generate high-quality photorealistic images from previously unseen viewpoints. Overall, the NeRF-based reconstruction framework illustrates the power of neural implicit representations for modeling complex 3D scenes. The project contributes to understanding modern neural rendering techniques and their applications in computer vision and graphics. With further optimization and enhancements, the system can serve as a foundation for advanced research and real-world applications in 3D reconstruction, immersive visualization, and intelligent scene understanding.

ISSN: 2583-6129
DOI:10.55041/ISJEM06020

Compliance with ethical standards

Disclosure of conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES:

1. K. Genova, F. Cole, D. Vlasic, A. Sarna, W. T. Freeman, and T. Funkhouser, "Local Deep Implicit Functions for 3D Shape," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

2. P. Henzler, N. J. Mitra, and T. Ritschel, "Learning a Neural 3D Texture Space from 2D Exemplars," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

3. C.-H. Jiang, A. Sud, A. Makadia, J. Huang, M. Nießner, and T. Funkhouser, "Local Implicit Grid Representations for 3D Scenes," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

4. G. Rainer, A. Ghosh, W. Jakob, and T. Weyrich, "Unified Neural Encoding of Bidirectional Texture Functions (BTFs)," in *Computer Graphics Forum*, vol. 39, no. 2, pp. 535–547, 2020.

5. M. Niemeyer, L. Mescheder, M. Oechsle, and A. Geiger, "Differentiable Volumetric Rendering: Learning Implicit 3D Representations without 3D Supervision," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.