Synthesis of Dimensionality: a Distinctive Approach for High-Fidelity 3D Surface Prediction

Susmita Ghosh and Abhiroop Chatterjee

January 14, 2024

Synthesis of Dimensionality: A Distinctive Approach for High-Fidelity 3D Surface Prediction

Susmita Ghosh, *Member IEEE*
Computer Science and Engineering Department
Jadavpur University
Kolkata, India
susmitaghoshju@gmail.com

Abhiroop Chatterjee, *Member IEEE*
Computer Science and Engineering Department
Jadavpur University
Kolkata, India
abhiroopchat1998@gmail.com

Abstract— This paper introduces an implementation of Neural Radiance Fields (NeRF) for 3D surface prediction. NeRF is a powerful approach for synthesizing complex 3D scenes from sparse 2D observations. In this work, we present a concise neural network architecture for NeRF and utilize synthetic 3D data to train the model. The training process involves optimizing the model parameters to minimize the mean squared error loss between predicted and ground truth surfaces. The results showcase the model's ability to accurately predict 3D surfaces, as demonstrated through visualizations of both ground truth and predicted surfaces. The simplicity of our implementation serves as an accessible entry point for researchers and practitioners interested in exploring NeRF and its applications in 3D surface prediction.

Keywords: Neural Radiance Fields, 3D Surface Prediction, Machine Learning, Synthetic Data, Neural Networks

Introduction

In recent years, the intersection of computer vision, machine learning [1-8], and deep learning [9-14] has witnessed remarkable advancements, especially in the realm of three-dimensional (3D) scene reconstruction. Neural Radiance Fields (NeRF) stand out as a cutting-edge approach within this landscape, providing a potent methodology for synthesizing detailed 3D scenes from sparse 2D observations.

This paper presents a straightforward implementation of NeRF, a neural network-based model designed for predicting intricate 3D surfaces. The fundamental premise of NeRF involves training a neural network to learn the radiance field of a scene, enabling the synthesis of novel views and rendering high-fidelity images from arbitrary viewpoints.

The primary objective of our implementation is to provide a clear and accessible entry point for researchers and practitioners interested in experimenting with NeRF. We adopt a minimalist neural network architecture and leverage synthetic 3D data for training, facilitating ease of understanding and application.

Throughout this paper, we delve into the specifics of our NeRF model, detailing the architectural components and the training process. The results, presented through visualizations of both ground truth and predicted surfaces, demonstrate the efficacy of our approach in accurately capturing 3D structures.

By simplifying the NeRF implementation, we aim to contribute to the broader exploration of 3D surface prediction methodologies, fostering a more inclusive understanding of this innovative field.

Related Research

PixelNeRF [15] is a novel learning framework designed to address limitations in existing methods for constructing neural radiance fields (NeRFs). Unlike traditional approaches that require extensive calibration and compute time for each scene, PixelNeRF employs a fully convolutional architecture that conditions a NeRF on input images. This enables training across multiple scenes, learning a scene prior and facilitating efficient novel view synthesis from sparse views (even just one). Leveraging NeRF's volume rendering approach, PixelNeRF can be trained directly from images without explicit 3D supervision. Extensive experiments on ShapeNet benchmarks, including single image novel view synthesis tasks and multi-object scenes, demonstrate PixelNeRF's superior performance over current state-of-the-art baselines in both category-specific and category-agnostic settings. The model also excels in real scenes from the DTU dataset, showcasing its flexibility and efficacy in various 3D reconstruction scenarios.

BungeeNeRF [16] addresses the challenge of modeling multi-scale scenes, such as cityscapes, landscapes, and intricate 3D models, where imagery exhibits significant variations in scale. Unlike conventional Neural Radiance Fields (NeRF) that struggle with diverse scales, BungeeNeRF introduces a progressive approach. It starts by fitting distant views with a shallow base block and, as training advances, dynamically appends new blocks to handle emerging details in closer views. This strategy activates high-frequency channels in NeRF's positional encoding inputs progressively, unfolding complex details over training. Demonstrating superior performance on various data sources, including city models, synthetic, and drone-captured data, BungeeNeRF excels in rendering high-quality, detailed scenes across a wide range of scales.

Methodology and Visualization

Our NeRF implementation comprises a three-layered neural network. The network takes a three-dimensional point as input, undergoes two ReLU-activated fully connected layers with 256 units each, and outputs a single unit representing the predicted radiance value.

Data Preparation:

Synthetic 3D Data: Generate synthetic 3D data by sampling points uniformly within the range [-1, 1]. Compute the ground truth for each point using a predefined function that combines sine and cosine terms to create a complex surface.

Training Set: Split the generated data into training and testing sets. Use the training set for updating model parameters and the testing set for evaluating model performance.

Training Loop:

Epoch Iteration (5000 epochs in this case):

Forward Pass: Feed the training points through the NeRF model to obtain predicted radiance values.

Loss Computation: Calculate the mean squared error loss between the predicted radiance values and the ground truth.

Backward Pass: Propagate the loss backward through the network to compute gradients.

Optimization Step: Use the Adam optimizer to update the model parameters based on the computed gradients.

Monitoring and Evaluation

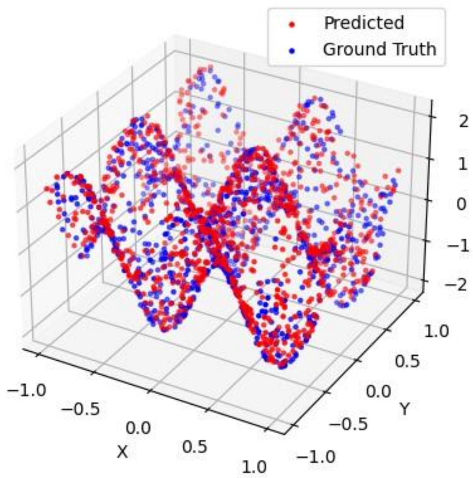

Fig. 1: Model Performance: Ground Truth (Blue) vs. Predicted (Red) Surfaces.

The image (Fig.1) illustrates a comparison between ground truth and predicted 3D surfaces. In the visualization, the ground truth surfaces are represented in blue, while the predicted surfaces are depicted in red. The juxtaposition of these surfaces allows for a visual assessment of the model's performance in capturing the intricacies and details of the underlying 3D structures. This comparison provides valuable insights into the accuracy and fidelity of the neural network's predictions in the context of 3D surface reconstruction.

Conclusion

This paper presented a simplified implementation of Neural Radiance Fields (NeRF) for 3D surface prediction. The straightforward neural network architecture, coupled with synthetic 3D data, aimed to provide an accessible entry point for researchers and practitioners interested in exploring NeRF.

Through 5000 epochs of training, the model demonstrated the capacity to predict intricate 3D surfaces, as evidenced by the visualizations comparing ground truth and predicted surfaces. The optimization process, facilitated by the Adam optimizer, allowed the model to learn the underlying radiance field and generalize to unseen data.

While the results showcase the potential of NeRF in 3D surface prediction, there exist opportunities for further exploration and improvement. Future work could involve experimenting with different neural network architectures, incorporating real-world data for training, and exploring advanced optimization techniques to enhance model performance.

This implementation serves as a foundational step in understanding and applying NeRF in the context of 3D scene reconstruction. As the field continues to evolve, this work contributes to the broader conversation surrounding the intersection of neural networks and computer vision for synthesizing detailed 3D environments.

References

[1] Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., & Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer genomics. *Cancer genomics & proteomics*, *15*(1), 41-51.

[2] Myles, A. J., Feudale, R. N., Liu, Y., Woody, N. A., & Brown, S. D. (2004). An introduction to decision tree modeling. *Journal of Chemometrics: A Journal of the Chemometrics Society*, *18*(6), 275-285.

[3] Zheng, A., & Casari, A. (2018). *Feature engineering for machine learning: principles and techniques for data scientists*. " O'Reilly Media, Inc.".

[4] Tseng, M. H., Chen, S. J., Hwang, G. H., & Shen, M. Y. (2008). A genetic algorithm rule-based approach for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *63*(2), 202-212.

[5] Halder, A., Ghosh, A., & Ghosh, S. (2009). Aggregation pheromone density based pattern classification. *Fundamenta Informaticae*, *92*(4), 345-362.

[6] Dehuri, S., Ghosh, S., & Cho, S. B. (Eds.). (2011). Integration of swarm intelligence and artificial neural network (Vol. 78). World Scientific.

[7] Halder, A., Ghosh, S., & Ghosh, A. (2013). Aggregation pheromone metaphor for semi-supervised classification. Pattern Recognition, 46(8), 2239-2248.

[8] Datta, A., Ghosh, S., & Ghosh, A. (2016). Supervised feature extraction of hyperspectral images using partitioned maximum margin criterion. IEEE Geoscience and Remote Sensing Letters, 14(1), 82-86.

[9] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

[10] Niu, Z., Zhong, G., & Yu, H. (2021). A review on the attention mechanism of deep learning. *Neurocomputing*, *452*, 48-62.

[11] Ghosh, S., & Chatterjee, A. (2023). Automated COVID-19 CT Image Classification using Multi-head Channel Attention in Deep CNN. *arXiv preprint arXiv:2308.00715*.

[12] Ghosh, S., & Chatterjee, A. (2023). Introducing Feature Attention Module on Convolutional Neural Network for

Diabetic Retinopathy Detection. *arXiv preprint arXiv:2308.02985*.

[13] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., ... & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data*, *8*, 1-74.

[14] Ghosh, S., & Chatterjee, A. (2023). T-Fusion Net: A Novel Deep Neural Network Augmented with Multiple Localizations based Spatial Attention Mechanisms for Covid-19 Detection. *arXiv preprint arXiv:2308.00053*.

[15] Yu, A., Ye, V., Tancik, M., & Kanazawa, A. (2021). pixelnerf: Neural radiance fields from one or few images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4578-4587).

[16] Xiangli, Y., Xu, L., Pan, X., Zhao, N., Rao, A., Theobalt, C., ... & Lin, D. (2022, October). Bungeenerf: Progressive neural radiance field for extreme multi-scale scene rendering. In European conference on computer vision (pp. 106-122). Cham: Springer Nature Switzerland.