Novel Class Discovery for Representation of Real-World Heritage Data as Neural Radiance Fields (Student Abstract)

Shivanand Kundargi, Tejas Anvekar, Ramesh Tabib, Uma Mudenagudi

Center of Excellence in Visual Intelligence (CEVI), KLE Technological University
Vidyanagar, Hubballi, Karnataka, India.
shivanandkundargi992@gmail.com, anvekartejas@gmail.com, ramesh_t@kletech.ac.in, uma@kletech.ac.in

Abstract

Neural Radiance Fields (NeRF) have been extensively explored as a leading approach for modeling and representing 3D data across various domains. Their ability to capture arbitrary scale point clouds and generate novel views makes them particularly valuable for digitizing cultural heritage sites. However, despite their impressive rendering capabilities, prior methods have often overlooked a significant real-world challenge: handling open-world scenarios characterized by unstructured data containing multiple classes in a single set of unlabeled images. To address this challenge, we propose a novel method NCD-NeRF that leverages Novel-Class Discovery to effectively tackle the complexities inherent in real-world data with unlabeled classes while excelling in producing high-quality NeRF representation. To validate our approach, we conducted a benchmarking analysis using a custom-collected dataset featuring UNESCO World Heritage sites in India. We observe that our proposed NCD-NeRF can parallely discover novel classes and render high-quality 3D volumes.

Introduction

Neural Radiance Fields (NeRFs) (Mildenhall et al. 2021) have risen to prominence as the leading methodology for best representation of 3D data as radiance field. When provided with images of a scene captured from various camera poses, NeRFs (Mildenhall et al. 2021) excel in seamlessly interpolation between these viewpoints. We specifically opt NeRF (Mildenhall et al. 2021) as representation learning framework for Real-World Heritage data due to two major reasons: 1) flexibility to obtain point cloud and mesh data from rendered NeRF (Mildenhall et al. 2021) representation; 2) Arbitrary Scale Sampling of obtained Point clouds.

For collecting real-world data, it is customary to capture images from multiple sites concurrently. However, this collection of unstructured data typically demands a significant level of human intervention for meticulous labeling and organization aligning them with their specific attributes and characteristics.

Our objective is to entirely eradicate the necessity for human intervention in the categorization / segregation of classes derived from real-world data. This aims to make 3D data representation more efficient for real-world data. To accomplish this, we harness the power of emerging Novel Class Discovery techniques (Han, Vedaldi, and Zisserman 2019; Liu and Tuytelaars 2022; Roy et al. 2022; Du et al. 2023), which have demonstrated remarkable efficiency in obtaining the 3D representation of unstructured data.

NCD-NeRF

Problem Setting consists of two distinct stages. The initial stage involves training a Novel Class Discovery model to efficiently identify and categorize unlabeled classes. The second stage entails utilizing the newly discovered classes for neural radiance field representation, following the specifications outlined in (Mildenhall et al. 2021). In the first stage, during training, the dataset is divided into two subsets: a labeled set, D_l, comprising images x_l^i with corresponding class labels y_l^i, and an unlabeled set D_u, consisting of images x_u^i. The primary objective is to leverage both D_l and D_u to unveil C_u novel classes. This is conventionally achieved by partitioning D_u into C_u clusters and assigning labels y_u^i to images within D_u. This setup allows for the categorization of any unlabeled classes into specific categories. Subsequently, in the second stage, the newly discovered novel classes are harnessed for NeRF representation under the assumption that the discovered categories contain multiview images as well discribed in Figure 1.

Results and Discussions

IDH Data nine UNESCO-recognized heritage sites, like Stone Chariot, Kadalekalu Ganesha Temple, Sasvikalu Ganesha Temple, Veerupaksha Temple, Kings Balance, and four pillars at Hampi and Pattadakallu, comprises over 2000 images for each structure, under varied conditions. Employing a pre-trained model trained on a set of five classes, we conduct novel-class discovery on a distinct set of four classes, different from the labeled set.

Effect of NCD on Real-World Data in the context of real-world Heritage data for neural radiance field representation, we evaluate various methods as shown in Table 1. Our benchmark shows that the SMILE approach (Du et al. 2023) surpasses other methods, attributed to its sign-magnitude disentanglement technique effectively managing both inter and intra-class variations. This capability directly addresses

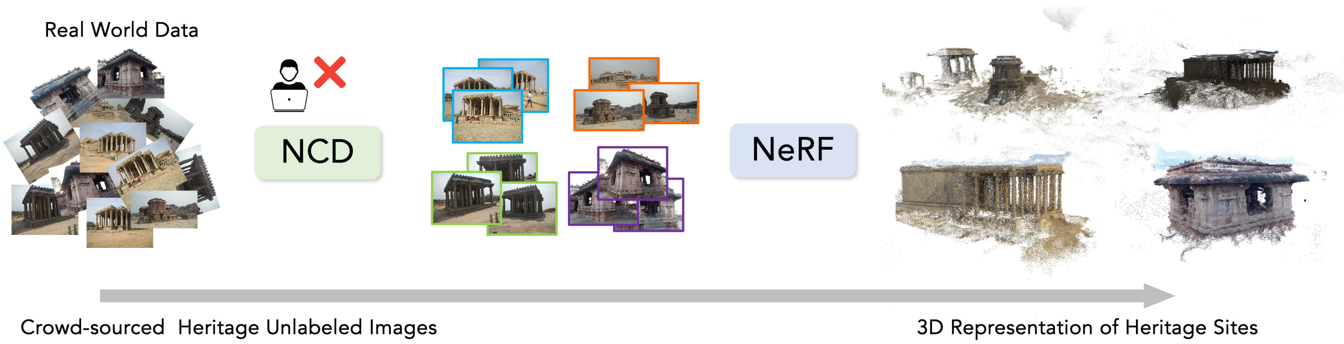

Figure 1: We present an end-to-end pipeline to get the 3D representation of crowdsourced unlabeled data. Proposed NCD-Nerf discovers novel classes from the data and renders all of them parallelly without human intervention. The 3D point clouds from the figure are obtained by sampling the desired number of points from the volume rendered by NeRF.

challenges in real-world data, making SMILE (Du et al. 2023) a suitable solution for NeRF representation.

Methods	IDH		
	Old	New	All
DTC 2019	22.63	52.73	30.7
Restune 2022	43.19	46.97	45.31
Class-INCD 2022	51.63	40.37	55.89
OCD 2023	**63.27**	**65.04**	**64.32**

Table 1: Quantitative Analysis of state-of-the-art Novel Class Discovery methods. **Note:** "*Old*", "*New*", "*All*" depicts Strict-Hungarian accuracy (Du et al. 2023).

Why NCD with NeRF?

From our simple yet effective approach NCD-NeRF, we can obtain high-quality 3D representation from real-world unlabeled crowdsourced heritage data. We specifically employ Novel Class Discovery (NCD) methods to leverage information learned during the supervised stage, enabling the identification of novel classes within unstructured data. These methods outperform trivial unsupervised alternatives like K-Means clustering. Empirically we observe that discovered images from (Du et al. 2023) yield better performance quantitatively. Hence we report their metrics on different sites as specified in the Table 1. NCD-NeRF also achieves parallelism in processing the discovered classes, while maintaining the quality of NeRF (Mildenhall et al. 2021) rendering as depicted in Table 2.

Methods	NeRF 2021		
	SSIM	PSNR	LPIPS
Stone Chariot	22.63	0.69	0.23
Kadalekal Ganesh Temple	22.26	0.59	0.63
Sasvikal Ganesha Temple	21.79	0.58	0.32
Veerupaksha Temple	22.76	0.64	0.43

Table 2: We report the SSIM, PSNR, LPIPS of the respective NeRF 2021 renderings obtained by our methodology.

Conclusions

We proposed a streamlined methodology aimed at mitigating human intervention in the representation of real-world data as Neural Radiance Fields. Our approach involves a two-stage solution: in the initial stage, we harness Novel Class Discovery (NCD) to eliminate the need for human intervention, thereby obtaining multiple categories of unstructured data. Subsequently, in the second stage, we generate Neural Radiance Fields (NeRF) representations for these categorized datasets.

Acknowledgements

This work is partly carried out under the Department of Science and Technology (DST) through the ICPS programme - Indian Heritage in Digital Space for the project "Crowd-Sourcing" (DST/ ICPS/ IHDS/ 2018 (General)).

References

Du, R.; Chang, D.; Liang, K.; Hospedales, T.; Song, Y.-Z.; and Ma, Z. 2023. On-the-Fly Category Discovery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11691–11700.

Han, K.; Vedaldi, A.; and Zisserman, A. 2019. Learning to discover novel visual categories via deep transfer clustering. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 8401–8409.

Liu, Y.; and Tuytelaars, T. 2022. Residual tuning: Toward novel category discovery without labels. *IEEE Transactions on Neural Networks and Learning Systems*.

Mildenhall, B.; Srinivasan, P. P.; Tancik, M.; Barron, J. T.; Ramamoorthi, R.; and Ng, R. 2021. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1): 99–106.

Roy, S.; Liu, M.; Zhong, Z.; Sebe, N.; and Ricci, E. 2022. Class-incremental novel class discovery. In *European Conference on Computer Vision*, 317–333. Springer.