HuSc3D: Human Sculpture dataset for 3D object reconstruction

Weronika Smolak-Dyżewska[1,2][0009−0009−1454−3157], Dawid Malarz[1,3][0009−0002−3501−6021], Grzegorz Wilczyński[1,2,3][0009−0002−6053−4410], Rafał Tobiasz[1,2,3][0009−0002−9265−6148], Joanna Waczyńska[1,2,3][0009−0003−8593−9307], Piotr Borycki[1,2,3][0000−0002−5715−4428], and Przemysław Spurek[1,3][0000−0003−0097−5521]

[1] Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland
[2] Doctoral School of Exact and Natural Sciences, Jagiellonian University, Kraków, Poland
[3] IDEAS Research Institute

Abstract. 3D scene reconstruction from 2D images is important tasks in computer graphics. Unfortunately, existing datasets and benchmarks concentrate on idealized synthetic or meticulously captured realistic data. Such benchmarks fail to convey the inherent complexities encountered in newly acquired real-world scenes. In such scenes the background is often dynamic, and by popular usage of cell phone cameras, there might be discrepancies in, e.g., white balance. To address this gap, we present HuSc3D, a novel dataset specifically designed for rigorous benchmarking of 3D reconstruction models under realistic acquisition challenges. Our dataset features six highly detailed, fully white sculptures characterized by intricate perforations and minimal textural variation. Furthermore, the number of images per scene varies significantly, introducing the additional challenge of limited training data for some instances alongside scenes with a standard number of views. By evaluating 3D reconstruction methods on this diverse dataset, we demonstrate the distinctiveness of HuSc3D in effectively differentiating model performance, particularly highlighting the sensitivity of methods to fine geometric details, color ambiguity, and varying data availability – limitations often masked by more conventional datasets. The project page is available here.

Keywords: Gaussian Splatting · NeRF · Benchmark · Novel View Synthesis.

1 Introduction

The burgeoning field of 3D scene reconstruction has witnessed remarkable advancements, largely propelled by novel neural representations such as Neural Radiance Fields (NeRF) [1] and, more recently, 3D Gaussian Splatting (3DGS) [2]. These methods have demonstrated impressive capabilities in generating photorealistic novel views and detailed 3D geometries from collections of 2D images.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29924-6_54

Consequently, a significant body of research now focuses on improving the fidelity, speed, and robustness of these techniques. Central to this progress is the availability of diverse and challenging datasets for training and, critically, for benchmarking the performance of these algorithms. However, majority of existing datasets, while valuable, present a somewhat sanitized view of the data acquisition process. Synthetic datasets, by their nature, offer perfect camera poses and controlled environments. Even datasets captured in the real world often undergo meticulous curation, extensive calibration, or are captured under studio-like conditions, resulting in data that is cleaner and more consistent than what an average user might produce. This discrepancy is particularly relevant as methods like 3DGS become more accessible, inviting users to capture scenes spontaneously with readily available devices like smartphones. Such "in-the-wild" captures, often intended for casual 3D reconstruction, inherently possess a different set of challenges not fully represented by current benchmarks. To address this gap, we introduce HuSc3D dataset which consists of 6 scenes of white sculptures. The dataset was made by three different methods and showcases several challenges that the average user encounters when starting with the reconstruction task. Our dataset introduces several real-world challenges, including a low number of COLMAP-matched images [3,4], resulting in limited training data, variations in automatic white balance (AWB) typical of smartphone video capture, dynamic background elements (e.g., passing people in public spaces), and objects with intricate geometry but low contrast and minimal texture relative to their surroundings.

2 Related Works

Reconstructing 3D scenes from 2D images is a central problem in computer vision and graphics. Recent advances have been driven by learning-based approaches, particularly Neural Radiance Fields (NeRF) [1] and 3D Gaussian Splatting (3DGS) [2], along with their many extensions [5,6,7,8,9,10,11,12,13,14]. Progress in this area is closely tied to the availability of benchmarking datasets.

While numerous datasets exist supporting progress in 3D reconstruction, most are either synthetic or rely on professionally acquired real-world data. Here, we highlight widely recognized benchmarks that represent key data types and challenges, providing context for the gap addressed by HuSc3D. Synthetic datasets such as NeRF-Synthetic [1] offer fully controlled environments with perfect camera parameters and ground truth geometry. While useful for development, they omit real-world issues like calibration errors, noisy sensors, lighting variation, and dynamic elements—factors emphasized by HuSc3D. Among real-world datasets, Mip-NeRF 360 [5] focuses on large-scale, unbounded scenes with high-fidelity 360° captures. In contrast, HuSc3D targets object-centric challenges involving difficult materials (e.g., low-texture or white surfaces) and consumer-grade capture artifacts such as white balance inconsistencies. The LLFF dataset [18] addresses sparse, forward-facing views from handheld cameras. While HuSc3D also includes sparse inputs, it extends this to 360° settings with additional

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29924-6_54

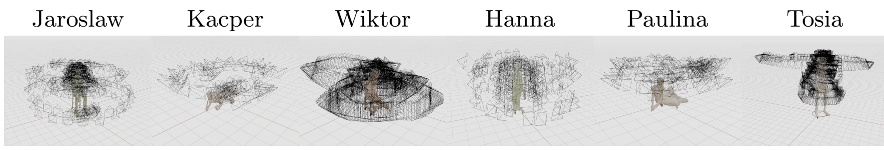

Fig. 1. Visualisation of cameras in scenes in HuSc3D. The visualisation was made in Blender [15] using addons Photogrammetry-based Framework [16] and Kiri [17].

photometric and texture-related challenges. CO3D [19] represents large-scale, category-centric datasets with diverse multi-view object sequences. Rather than focusing on category generalization, HuSc3D emphasizes reconstruction of particularly challenging objects under everyday capture conditions. Finally, datasets like DeepBlending [20] focus on controlled indoor scenes with artificial lighting and room-scale reconstruction, differing from the object-centric focus of HuSc3D.

3 HuSc3D Dataset

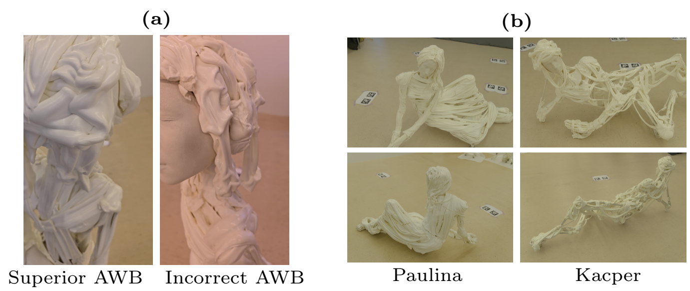

Fig. 2. **(a)** Example of variety in AWB in Tosia scene. **(b)** Examples of Paulina and Kacper scenes' images.

Data acquisition and challenges The HuSc3D dataset consists of six scenes, each featuring a different sculpture made by the Polish sculptor Paweł Althamer. The dataset was acquired in the Foksal Gallery Foundation headquarters. Captured sculptures were made in 2012 and are a part of a Almech project. All sculptures are white with multiple perforations. Among the dataset there are three ways of capturing the data with camera positions presented in Fig. 1.

Kacper and Paulina scenes To simulate sparse-view challenges, the Kacper and Paulina scenes use images captured with a Nikon D7000 and fixed camera settings. After COLMAP-based preprocessing, the final sets contain only 66 (Kacper) and 104 (Paulina) images (see Fig. 2b). This limited data makes high-fidelity reconstruction difficult, especially for detailed objects. Methods like NeRF are particularly sensitive to this sparsity, yielding the weakest results compared to approaches such as Splatfacto-MCMC (see Tab. 1 and Fig. 3a).

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29924-6_54

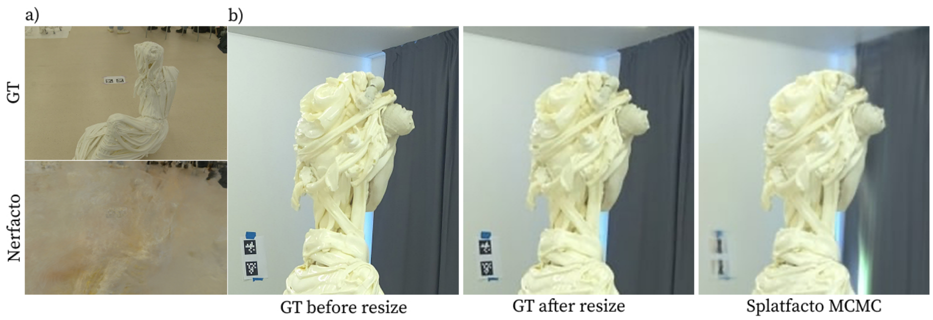

Fig. 3. (a) In scenes with low training samples (Paulina) methods based on NeRF struggle with reconstruction. **(b)** The necessary 8x downsizing of input images (Hanna) causes substantial detail loss, hindering satisfactory reconstruction.

Tosia and Wiktor scenes To simulate common mobile capture issues, two scenes were recorded using a Samsung Galaxy S24 Ultra, yielding 680 and 596 images by extracting every 20th frame. A key challenge is photometric inconsistency caused by automatic white balance (AWB), which introduces varying color casts (e.g., reddish tint) across frames. This reflects real-world conditions where on-device processing affects data quality. The Tosia scene contains 13.2% affected images, while Wiktor has 28.4%. Although these inconsistencies are easily noticeable to humans (see Fig. 2a), they typically require preprocessing for reconstruction algorithms.

Jarosław and Hanna scenes Two scenes in HuSc3D highlight a key limitation in modern 3D reconstruction: handling ultra-high-resolution images. Captured with a Samsung Galaxy S24 Ultra in PRO mode (fixed ISO, aperture, shutter), the datasets include 348 (Jarosław) and 190 (Hanna) images at 6000×8000 px resolution. While rich in detail, their use is constrained by computational cost—requiring 8× downscaling, which removes fine details. These scenes therefore act as a stress test for methods to either process high-resolution data efficiently or remain robust to information loss. Even top-performing methods like Splatfacto-MCMC show noticeable detail degradation (see Fig. 3b).

4 Experiments

To benchmark HuSc3D, six widely used 3D reconstruction methods were evaluated: 2DGS [13], Mip-Splatting [21], NeRF (Nerfacto, Nerfstudio [22]), 3DGS (Splatfacto, Nerfstudio), 3DGS-MCMC (Splatfacto-MCMC, Nerfstudio) [23], and Instant-NGP [24] (Nerfstudio). These represent key approaches in the field, spanning explicit point-based and implicit volumetric methods.

Experimental Setup Reconstruction quality and efficiency were evaluated using PSNR, SSIM, LPIPS, training time, and FPS.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29924-6_54

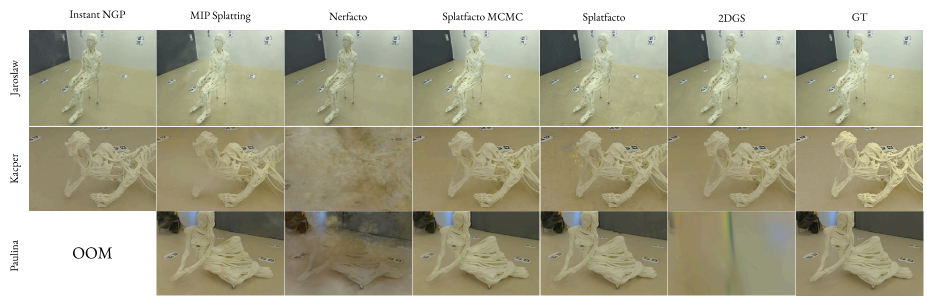

Fig. 4. Qualitative comparison. The majority of methods struggled on scenes with sparse number of training views (Kacper and Paulina).

For all methods we used publicly available implementations with minimal changes. 2DGS and Mip-Splatting follow their original repositories and default settings, while Nerfacto, Splatfacto, Splatfacto-MCMC, and In-stant-NGP are run via Nerfstudio with default configurations. Images were downsampled by 4× (8× for Hanna and Jarosław). Experiments run on an NVIDIA RTX 4090 (256 GB RAM, Ubuntu 24.04). Following 3DGS protocol, 12.5% of images are held out for testing, and all metrics are computed on this subset.

Quantitative results The quantitative performance on the HuSc3D dataset is presented in Tab. 1. The superior performance of Gaussian Splatting (GS) methods (2DGS, Mip-Splatting, Splatfacto, Splatfacto-MCMC) in both reconstruction quality and rendering speed suggests that explicit, point-based representations are more robust to the challenges in HuSc3D. In particular, Splatfacto-MCMC achieved the best metrics. In contrast, NeRF-based methods (Nerfacto, Instant-NGP) struggled more significantly. Nerfacto showed consistently lower fidelity, while Instant-NGP encountered memory limitations on multiple scenes. These issues likely stem from sensitivity to dynamic backgrounds, limited training views, and photometric inconsistencies such as white balance shifts, which are further exacerbated by the texture-poor nature of the objects.

Table 1. Quantitative results across HuSc3D. Best scores per metric are bolded. Time (mm:sec). *Instant-NGP training failed with OOM for Hanna and Paulina scenes.

Method	**SSIM ↑**	**PSNR ↑**	**LPIPS ↓**	**Time ↓**	**FPS ↑**
2DGS	0.920	29.96	0.222	10:21	187.17
Mip-Splatting	0.930	30.62	0.192	11:29	**191.74**
Nerfacto	0.807	21.98	0.359	12:12	1.39
Splatfacto	0.944	31.38	0.136	**7:38**	165.71
Splatfacto-MCMC	**0.953**	**32.45**	**0.109**	16:06	82.34
Instant-NGP	0.908*	27.435*	0.231*	9:55*	1.23*

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29924-6_54

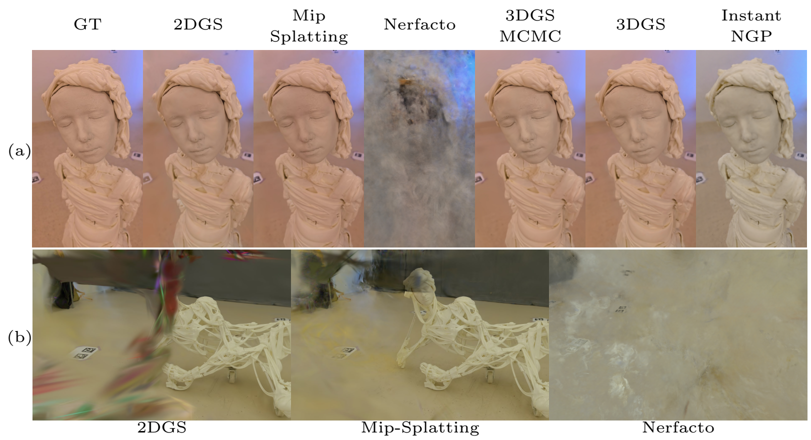

Fig. 5. (a) Qualitative results. Instant-NGP and Nerfacto favor the dataset's average white balance (WB), causing errors or blur, while 3DGS methods successfully reconstruct sudden WB shifts. **(b)** Significant artifacts in sparse views. 3DGS - Splatfacto

Qualitative results Qualitative results are presented in Fig. 4 and Fig. 5(a). It represents renders made on test set with different methods. Even though, Splatfacto-MCMC achieved the best results both qualitatively and quantitatively, it still could misses intricate details as previously mentioned in Fig. 3. Multiple methods failed at reconstructing correctly Paulina and Kacper scenes with sparse views. For fair comparison, we used the same frame throughout every method, however, multiple methods failed to correctly reconstruct every point of view in testing sets creating a significant amount of artifacts located on the object of interest. The methods that had at least one frame failed were 2DGS, Mip-Splatting, Nerfacto (which failed on almost every image). Additionally, Instant-NGP OOM on Paulina scene and on multiple renders in Kacper. We showcased these exemplary problematic frames in Fig. 5(b).

5 Conclusion

HuSc3D is a novel dataset addressing the limitations of existing 3D reconstruction benchmarks. While previous datasets often rely on idealized synthetic or meticulously captured data, HuSc3D focuses on realistic challenges faced by non-expert users. Given recent advancements in reconstruction methodologies, we believe the field is now equipped to tackle these real-world complexities.

Limitations and Future Work The dataset size is the biggest limitation of HuSc3D. Future Work could include more sculptures, presumably from different artists. Additionally, every scene could contain captures showcasing every combination of the mentioned challenges.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29924-6_54

6 Acknoledgments

We are deeply grateful to the amazing artist Paweł Althamer for its wonderful art and his curtosy in enabling us to scan and share his work. We also want to thank Foksal Gallery Foundation for providing us with space for scanning, a warm welcome, and helping in contacting the artist. The work of W. Smolak-Dyżewska, D. Malarz, G. Wilczyński and P. Spurek was supported by the project *Effective Rendering of 3D Objects Using Gaussian Splatting in an Augmented Reality Environment* (FENG.02.02-IP.05-0114/23), carried out under the First Team programme of the Foundation for Polish Science and co-financed by the European Union through the European Funds for Smart Economy 2021–2027 (FENG). The work of P. Borycki was supported by the National Centre of Science (Poland) Grant No. 2025/57/N/ST6/04389. The work of J. Waczyńska was supported by the National Centre of Science (Poland) Grant No. 2025/57/N/ST6/04390.

References

1. Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In *ECCV*, 2020.
2. Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4), July 2023.
3. Johannes Lutz Schönberger, Enliang Zheng, Marc Pollefeys, and Jan-Michael Frahm. Pixelwise view selection for unstructured multi-view stereo. In *European Conference on Computer Vision (ECCV)*, 2016.
4. Johannes Lutz Schönberger and Jan-Michael Frahm. Structure-from-motion revisited. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
5. Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, Peter Hedman, Ricardo Martin-Brualla, and Jonathan T. Barron. MultiNeRF: A Code Release for Mip-NeRF 360, Ref-NeRF, and RawNeRF, 2022.
6. Dor Verbin, Peter Hedman, Ben Mildenhall, Todd Zickler, Jonathan T. Barron, and Pratul P. Srinivasan. Ref-NeRF: Structured view-dependent appearance for neural radiance fields. *CVPR*, 2022.
7. Albert Pumarola, Enric Corona, Gerard Pons-Moll, and Francesc Moreno-Noguer. D-NeRF: Neural Radiance Fields for Dynamic Scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020.
8. Yang Hong, Bo Peng, Haiyao Xiao, Ligang Liu, and Juyong Zhang. Headnerf: A real-time nerf-based parametric head model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 20374–20384, June 2022.
9. Kangle Deng, Andrew Liu, Jun-Yan Zhu, and Deva Ramanan. Depth-supervised nerf: Fewer views and faster training for free. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12882–12891, June 2022.
10. Guanjun Wu, Taoran Yi, Jiemin Fang, Lingxi Xie, Xiaopeng Zhang, Wei Wei, Wenyu Liu, Qi Tian, and Wang Xinggang. 4d gaussian splatting for real-time dynamic scene rendering. *arXiv preprint arXiv:2310.08528*, 2023.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29924-6_54

11. Yi-Hua Huang, Yang-Tian Sun, Ziyi Yang, Xiaoyang Lyu, Yan-Pei Cao, and Xiaojuan Qi. Sc-gs: Sparse-controlled gaussian splatting for editable dynamic scenes. *arXiv preprint arXiv:2312.14937*, 2023.
12. Joanna Waczyńska, Piotr Borycki, Joanna Kaleta, Sławomir Tadeja, and Przemysław Spurek. D-miso: Editing dynamic 3d scenes using multi-gaussians soup. *arXiv preprint arXiv:2405.14276*, 2024.
13. Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2d gaussian splatting for geometrically accurate radiance fields. In *SIGGRAPH 2024 Conference Papers*. Association for Computing Machinery, 2024.
14. Dawid Malarz, Weronika Smolak-Dyżewska, Jacek Tabor, Sławomir Tadeja, and Przemysław Spurek. Gaussian splatting with nerf-based color and opacity. *Computer Vision and Image Understanding*, 251:104273, 2025.
15. Blender Online Community. *Blender - a 3D modelling and rendering package*. Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018.
16. Sebastian Bullinger, Christoph Bodensteiner, and Michael Arens. A photogrammetry-based framework to facilitate image-based modeling and automatic camera tracking. *International Conference on Computer Graphics Theory and Applications*, 2021.
17. Kiri-Innovation. 3dgs render blender addon by kiri engine. https://github.com/Kiri-Innovation/3dgs-render-blender-addon, 2025.
18. Ben Mildenhall, Pratul P. Srinivasan, Rodrigo Ortiz-Cayon, Nima Khademi Kalantari, Ravi Ramamoorthi, Ren Ng, and Abhishek Kar. Local light field fusion: Practical view synthesis with prescriptive sampling guidelines. *ACM Transactions on Graphics (TOG)*, 2019.
19. Jeremy Reizenstein, Roman Shapovalov, Philipp Henzler, Luca Sbordone, Patrick Labatut, and David Novotny. Common objects in 3d: Large-scale learning and evaluation of real-life 3d category reconstruction. In *International Conference on Computer Vision*, 2021.
20. Peter Hedman, Julien Philip, True Price, Jan-Michael Frahm, George Drettakis, and Gabriel Brostow. Deep blending for free-viewpoint image-based rendering. *ACM Transactions on Graphics (SIGGRAPH Asia Conference Proceedings)*, 37(6), November 2018.
21. Zehao Yu, Torsten Sattler, and Andreas Geiger. Gaussian opacity fields: Efficient high-quality compact surface reconstruction in unbounded scenes. *arXiv:2404.10772*, 2024.
22. Matthew Tancik, Ethan Weber, Evonne Ng, Ruilong Li, Brent Yi, Justin Kerr, Terrance Wang, Alexander Kristoffersen, Jake Austin, Kamyar Salahi, Abhik Ahuja, David McAllister, and Angjoo Kanazawa. Nerfstudio: A modular framework for neural radiance field development. In *ACM SIGGRAPH 2023 Conference Proceedings*, SIGGRAPH '23, 2023.
23. Shakiba Kheradmand, Daniel Rebain, Gopal Sharma, Weiwei Sun, Yang-Che Tseng, Hossam Isack, Abhishek Kar, Andrea Tagliasacchi, and Kwang Moo Yi. 3d gaussian splatting as markov chain monte carlo. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2024. Spotlight Presentation.
24. Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph.*, 41(4):102:1–102:15, July 2022.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29924-6_54