Application of Neural Graphics Primitives Models for 3D Representation of Devastation Caused by Russian Aggression in Ukraine*

Illia Oholtsov*[1][0000−0003−3807−7647], Yuri Gordienko[1][0000−0003−2682−4668], Mariia Ladonia[1][0009−0009−5874−2406], Sergii Telenyk[2][0000−0001−9202−9406], Grzegorz Nowakowski[2][0000−0002−3086−0947], and Sergii Stirenko[1][0000−0001−5478−0450]

[1] National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
illia.oholtsov@gmail.com
[2] Cracow University of Technology, Cracow, Poland

Abstract. This work investigates the feasibility of applying Neural Radiance Fields (NeRFs) for reconstructing 3D representations of damaged structures caused by the ongoing aggression of Russia against Ukraine. The drone footage depicting the devastation was utilized and three NeRF models, Instant-NGP, Nerfacto, and SplatFacto, were employed. The models were evaluated across various damage levels (0: no damage, 4: high damage) using visual quality metrics like Structural Similarity Index Measure (SSIM), Learned Perceptual Image Patch Similarity (LPIPS), Peak Signal-to-Noise Ratio (PSNR) and rendering speed metrics like frames per second (FPS) and the number of rays per second (NRS). All input data (videos frames) and evaluation results (rendered visualizations) are available as a Kaggle dataset (http://tiny.cc/srasxz). No clear correlation was observed between damage level and reconstruction quality metrics, suggesting these metrics might not be reliable indicators of damage severity. SplatFacto consistently achieved the highest rendering speed (FPS, NRS) and exhibited the best visual quality (SSIM, PSNR, LPIPS) across all damage levels. The findings suggest that NeRFs, particularly SplatFacto, hold promise for rapid reconstruction and visualization of damaged structures, potentially aiding in damage assessment, documentation, and cultural heritage preservation efforts. Moreover, the study sheds light on the potential applications of such advanced modeling techniques in archiving and documenting conflict zones, providing a valuable resource for future investigations, humanitarian efforts, and historical documentation. However, further research is needed to explore the generalizability and robustness of NeRFs in diverse real-world scenarios.

Keywords: Artificial Intelligence · 3D representation · Neural Radiance Fields · Instant Neural Graphics Primitives · Damage Assessment · Ukraine.

* Supported in part by the National Research Foundation of Ukraine (NRFU) grant 2022.01/0199.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23

1 Introduction

Recent advancements in technology demonstrate the feasibility of restoring deteriorated artworks and cultural artifacts to their former glory. Employing advanced 3D reconstruction methodologies, these artifacts can be visually resurrected, even if they have been lost or irreparably damaged. Leveraging photographic documentation or 3D scanning data, it becomes possible to generate accurate three-dimensional models of cultural assets, thereby enabling viewers to engage with them once more. This restoration process extends to various forms of cultural heritage, including rediscovered ruins, damaged historical texts, ancient Egyptian and Greek temples, monumental structures, medieval frescoes, and other significant artifacts, all reconstructed with the aid of state-of-the-art techniques.

The ongoing aggression and war of Russia against Ukraine has left a profound mark on the landscape, particularly evident in the widespread destruction of cities [1,2]. The consequential impact of this aggression underscores the critical need to monitor, visualize, and document the extent of devastation with a depth that surpasses traditional methodologies.

Recently, several studies effectively demonstrated the potential of Neural Radiance Fields (NERFs) for reconstructing 3D representations [3,4]. NERF allows to create 3D models of objects and scenes from images of some scene from different angles. NERF "learns" the scene from these images ("remembering" the object's shape and details) and use this information to generate a realistic 3D model from any viewpoint, even one you didn't take images from before. This can be used for reconstructing damaged structures, capturing historical sites, or even creating immersive virtual experiences.

In this work, the feasibility of application of NERF models for reconstruction of 3D representations for damaged structures caused by Russian aggression in Ukraine is considered. Section 2 contains description of the state of the art, section 3 describes dataset, models, experiments, and the whole workflow, section 4 gives the results obtained during the experiments, section 5 contains discussions of results and resumes them in section 6.

2 Background and Related Work

Neural Radiance Fields (NeRFs) have revolutionized the field of 3D scene representation by offering a novel deep learning approach. NeRFs operate under a paradigm of learning a continuous volumetric representation of a scene from a collection of 2D images captured from various viewpoints [3]. This representation allows for the synthesis of novel views, enabling visualization of the scene from any desired perspective, even those not present in the original capture set.

NeRFs have experienced remarkable progress since their inception [3], effectively addressing initial limitations and paving the way for promising future applications. The Mip-NeRF approach employs a multiscale representation, significantly improving image detail and rendering speed while mitigating aliasing

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23

artifacts [5]. Mip-NeRF 360 tackles the challenge of unbounded scenes by enabling the reconstruction of large and intricate environments with 360-degree camera views [6]. Instant Neural Graphics Primitives utilizes multiresolution hash encoding for efficient training and real-time rendering, significantly improving the speed and efficiency of NeRF models [4]. High-Fidelity Reconstruction combines multi-resolution 3D hash grids with rendering, achieving superior surface reconstruction detail, leading to high-fidelity scene representations [7]. K-Planes presents a "white-box" model using k-planes to represent dynamic scenes like videos [8]. This approach enables efficient optimization and incorporation of specific priors, making it suitable for dynamic scene representation. 3D Gaussian Splatting leverages pre-computed 3D Gaussians and visibility-aware rendering, achieving real-time rendering at high quality, opening doors for interactive applications [9].

These advancements collectively demonstrate the rapid development of NeRF technology. By addressing initial limitations, these innovations push the capabilities of NeRFs towards real-time applications, detailed scene reconstruction, and handling complex scenes with diverse material properties and dynamics.

While NeRFs have seen significant advancements, several key challenges remain. For example, training NeRFs on large datasets with high resolution can be computationally expensive and time-consuming. Ongoing research explores techniques for improving training efficiency and reducing memory footprints. NeRFs can struggle with scenes containing complex materials with non-diffuse BRDF (Bidirectional Reflectance Distribution Function) properties or challenging lighting conditions that deviate from the typical assumptions made during training. Real-world data often contains noise, occlusions, or missing information. Improving the robustness of NeRFs to handle such data remains an ongoing area of research that can potentially improve generalization and reconstruction quality for integrating NeRFs into real-world applications. Addressing these challenges is crucial for further advancing the practical applications and capabilities of NeRF technology. That is why the main of the work is to investigate the feasibility of application of NERF models for reconstruction of 3D representations for damaged structures caused by Russian aggression in Ukraine.

3 Methodology

3.1 Data

The data for this project consisted of drone footage depicting the aftermath of Russian aggression in Ukraine. Due to the ongoing conflict, the initial step involved collecting existing videos from various sources, primarily YouTube. While this approach provided a starting point, the preferred method for future data acquisition would be to utilize drone cameras directly. This would enable high-resolution video capture and ensure comprehensive coverage of all key details and angles. All input data (videos frames) and evaluation results (rendered visualizations) are available as a Kaggle dataset [3].

[3] http://tiny.cc/srasxz

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23

To effectively analyze and categorize the scenes within the dataset, a method of differentiation based on "damage levels" was employed. This categorization was crucial for understanding the varying degrees of destruction observed in each scene. Annotators were instructed to assign damage level scores ranging from 0 to 5 to each scene, representing the extent of destruction. This scoring system enabled a quantitative representation of the severity of damage, ensuring consistency and accuracy in the evaluation process. Initially, the damage level classification was based on the existing dataset, which provided a range of damage levels. However, it is important to note that this classification system can be extended to encompass a broader range of destruction scenarios with the inclusion of new data. For instance, level 5 classifications could represent scenes where structures are completely obliterated, leaving only ruins with no discernible remnants of their original form. Incorporating such detailed classifications aims to provide a more nuanced understanding of the extent of devastation caused by the conflict. These efforts contribute to the comprehensive analysis and visualization of the aftermath captured in the drone footage.

Borodyanka residential area, damage level 4. This dataset focuses on the aftermath of conflict in Borodyanka, characterized by a lack of varied perspectives (Fig.1). The footage reveals a high level of destruction, providing intricate details and figures. The camera path is complex, offering a nuanced view of the affected area. However, the presence of many background objects introduces a challenge, as limited information is available about these elements, potentially impacting the overall interpretability of the dataset.

Fig. 1. Examples of drone imagery for Borodyanka, damage level 4.

Chernihiv Hotel “Ukraine”, damage level 3. This dataset features drone footages capturing the aftermath of Russian aggression on the Chernihiv hotel “Ukraine” (Fig.2). The camera follows a 360-degree path, capturing the scene from all angles and perspectives. The medium to high level of destruction is emphasized, with minimum background objects, focusing primarily on the hotel.

Chernihiv residential area, damage level 2. This dataset captures drone footages of a residential area in Chernihiv (Fig.3). The footage follows 360-degree

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23

Fig. 2. Examples of drone imagery for Chernihiv Hotel “Ukraine”, damage level 3.

camera path but with several jumps. The level of destruction is medium to low, and the main scene contains numerous objects. There are many background objects, which may impact the clarity of the footage.

Fig. 3. Examples of drone imagery for Chernihiv residential area, damage level 2.

Chernihiv Bridge, damage level 1,3. The Chernihiv Bridge dataset showcases drone footage with a forward-flying camera path, providing a singular perspective of the scene (Fig.4). The scene exhibits a low level of destruction on one half and a high level on the other. There are minimum background objects, with the main focus on the Bridge.

Fig. 4. Examples of drone imagery for Chernihiv Bridge, damage level 'mixed'.

Chernihiv cathedral, damage level 0. This dataset showcases drone footage captured using a 360-degree camera path, featuring the Chernihiv Cathedral (Fig.5) with occasional zoom adjustments. The scene focuses primarily on the cathedral and does not depict any destruction.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23

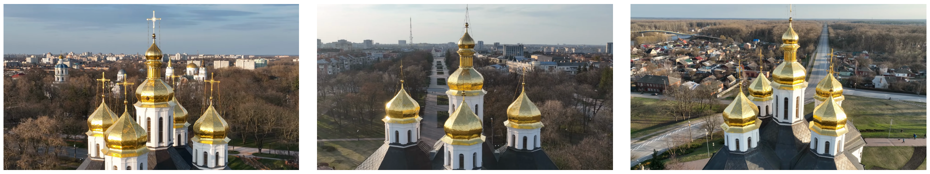

Fig. 5. Examples of drone imagery for Chernihiv cathedral, damage level 0.

3.2 Models

This section outlines the NeRF models employed in our research, trained and evaluated using the user-friendly NerfStudio framework [10]. NerfStudio provides various implementations of popular NeRF-based methods, including those directly relevant to this work[4].

SplatFacto. The SplatFacto implementation within NerfStudio was used to explore the 3D Gaussian Splatting (3D GS) approach [9]. The model uses default parameters to achieve a balance between speed, rendering quality, and the size of the generated splat file. One noteworthy parameter is cull-alpha-thresh (0.1) that defines the opacity threshold used for culling Gaussians, effectively removing those with minimal contribution to the final rendering.

Instant-NGP. For comparison purposes, NerfStudio's implementation of the Instant Neural Graphics Primitives (Instant-NGP) model [4] was used also and utilizes the following key parameters: grid-resolution (128), resolution of the grid used for the field; grid-levels (4), levels of the grid used for the field; and max-res (2048), maximum resolution of the hashmap for the base multi-layer perceptron (MLP).

Nerfacto. Nerfacto is a unified approach within NerfStudio that combines elements from various research papers to create a fast and high-quality rendering method [11]. It incorporates several key components:

- Pose refinement: This improves the accuracy of the camera poses, leading to a more realistic rendered scene.
- Piecewise sampler: This sampling strategy allocates samples throughout the scene, prioritizing areas with more objects and detail.
- Proposal sampler: This sampler further refines the sample locations by focusing on areas that significantly impact the final image. It utilizes a density function, implemented with a small fused MLP and hash encoding, to identify these crucial regions.

[4] https://github.com/nerfstudio-project/nerfstudio/

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23

– Density field: This field represents a coarse estimation of the scene's density, guiding the sampling process. It is also implemented with a small fused MLP and hash encoding.

Similar to the previous models, Nerfacto utilizes the following common parameters: grid-resolution (128), resolution of the grid used for the field; grid-levels (4), the number of levels within the hierarchical grid structur; and max-res (2048), maximum resolution of the hashmap for the base MLP.

All models share additional default parameters for training:

– the number of steps between saves (default: 2000), i.e. frequency at which the model checkpoints are saved during training,
– the maximum number of iterations (default: 30000), i.e. training iterations,
– the number of steps between batches (default: 500), i.e. the number of training steps before updating the model with a new batch of data.

3.3 Metrics

The paper presents various metrics to evaluate the performance of NERF models in reconstructing 3D representations of damaged buildings. These metrics include:

– SSIM (Structural Similarity Index Measure): measures the visual quality of the reconstructed scene compared to the ground truth. A higher SSIM value indicates better visual quality.
– PSNR (Peak Signal-to-Noise Ratio): measures the peak signal power relative to the noise power, with higher values indicating better fidelity to the original image.
– LPIPS (Learned Perceptual Image Patch Similarity): measures the perceptual similarity between the reconstructed image and the ground truth. A low LPIPS score means that image patches are perceptual similar.
– FPS (Frames Per Second): represents the rendering speed of the model, indicating how many frames can be generated per second.
– NRS (Number of Rays per Second): indicates the computational efficiency of the model, capturing the number of rays the model can process per second.

3.4 Workflow

The initial step involved collecting drone footage capturing the aftermath of Russian aggression in Ukraine. Due to the ongoing conflict, the primary method for acquiring this data was utilizing existing videos from diverse sources, mainly from YouTube. However, in practical scenarios, capturing high-resolution videos using drone cameras would be preferable to ensure comprehensive coverage of all relevant angles and details.

Once the video data is acquired, frames are extracted at a predetermined frame rate. In this work, a frame rate of 2 frames per second (fps) was chosen to

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23

achieve a balance between capturing sufficient scene information and maintaining dataset size within manageable limits.

While all extracted frames contribute to the dataset, they vary in informational value. To optimize representation and computational efficiency, frames capturing unique viewpoints or revealing crucial scene details were prioritized. Prioritized frames were those that maintained clarity and focus on the main object, discarding blurred frames or those not primarily focused on the main objects. This selection process aimed to ensure the dataset's quality and relevance for subsequent analyses and model training. Additionally, frames from multiple videos were combined to increase dataset size and viewpoint diversity, enhancing model robustness. This comprehensive approach aimed to create a rich and diverse dataset capable of supporting various analyses and training tasks effectively.

The next step involved processing the prepared frames with the COLMAP which is an open-source general-purpose Structure-from-Motion (SfM) and Multi-View Stereo (MVS) pipeline with a graphical and command-line interface[5]. It serves two primary purposes:

1. Camera Parameter Estimation: COLMAP helped estimate the intrinsic and extrinsic camera parameters associated with each extracted frame. These parameters are crucial for NeRF training, as they relate pixel locations within the images to their corresponding 3D world coordinates.
2. Feature Extraction and Matching: COLMAP was also employed to extract distinctive features from each frame and establish robust matches between corresponding features across different viewpoints. These feature matches provide valuable information about the scene's geometry and facilitate the subsequent NeRF training process.

Though COLMAP scripts usually require model-specific adjustments, NeRFStudio allows a single script for all models within the framework, improving efficiency and simplifying data preparation. Following COLMAP processing, the NeRFStudio framework was used for training. This framework offers various NeRF model implementations, facilitating the exploration of different approaches. Three specific models were employed in this work:

- Instant-NGP. This model was chosen for its fast training speed, achieved through efficient multi-resolution representations.
- SplatFacto. This model emphasizes rendering efficiency, utilizing pre-computed 3D Gaussians for real-time rendering capabilities.
- NeRFacto. This model combines multi-resolution encoding with neural surface rendering, aiming to achieve high-fidelity reconstructions.

Each of these models was trained on all prepared datasets within the NeRFStudio environment. This multi-model approach allowed the comparison of performance and characteristics of each method in the context of reconstructing the specific scene captured in the drone footage.

[5] https://colmap.github.io/

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23

The tools provided within the NeRFStudio framework were leveraged to evaluate the performance of trained models. Before initiating training, each dataset was split into dedicated training and testing sets. This ensured a fair evaluation process where models were assessed on previously unseen frames.

Various evaluation metrics are provided by NeRFStudio and calculated on a per-frame basis. These metrics typically include standard image quality measures like PSNR, SSIM, and LPIPS, offering different perspectives on reconstruction accuracy and visual fidelity. After calculating these metrics for all frames, the mean and standard deviation are computed across the entire test set. This aggregation provides a summary of the overall model performance, indicating both the central tendency and the variability across frames. The training and testing trials were performed in Google Colab environment[6] with access to a graphics processing unit NVIDIA Tesla T4 with 16 GB of video random access memory and 12 GB of system random access memory.

4 Results

After several attempts the mean and standard deviation values of metrics were measured for Instant-NGP (Table 1), Nerfacto (Table 2), and SplatFacto (Table 3) models and visualized in the plots below (Fig.6-7). The maximal values for SSIM, PSNR, LPIPS, FPS, NRS, and minimal values for LPIPS are emphasized by **bold** font.

Based on these results (Table 1, 2, and 3), one can make the following observations regarding the correlations between metrics and damage levels. There is no clear trend between the damage level and the values of SSIM, PSNR, LPIPS, FPS, or NRS across the different objects. While there might be slight variations in the means for each damage level, the standard deviations suggest significant overlap between different damage categories for each metric. This indicates that reconstruction quality metrics alone might not be reliable indicators of damage severity.

Table 1. Mean and standard deviation values of metrics for Instant-NGP model.

Object	Damage	SSIM	PSNR	LPIPS	FPS	NRS
Borodyanka	4	0.38±0.09	15.4±1.8	0.79±0.06	0.10±0.01	9.3E4±1.1E4
Hotel	3	**0.66±0.08**	20.8±2.1	**0.31±0.07**	0.07±0.01	6.7E4±0.8E4
Residential	2	0.45±0.09	18.2±1.5	0.55±0.10	0.064±0.004	5.9E4±0.4E4
Bridge	mixed	0.49±0.12	**21.6±1.5**	0.56±0.06	0.08±0.02	7.5E4±1.9E4
Cathedral	0	0.63±0.08	20.6±2.3	0.44±0.10	**0.13±0.02**	**11.9E4±1.7E4**

[6] https://colab.research.google.com/

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23

Table 2. Mean and standard deviation values of metrics for Nerfacto model.

Object	Damage	SSIM	PSNR	LPIPS	FPS	NRS
Borodyanka	4	0.38±0.13	15.4±2.4	0.50±0.14	0.13±0.01	11.4E4±8.3E4
Hotel	3	**0.55±0.07**	**18.8±1.4**	**0.28±0.06**	**0.21±0.02**	**19.2E4±1.9E4**
Residential	2	0.38±0.09	16.9±1.8	0.51±0.11	0.13±0.01	11.6E4±6.6E4
Bridge	mixed	0.37±0.14	**18.8±1.6**	0.56±0.05	0.13±0.01	11.9E4±7.3E4
Cathedral	0	0.50±0.10	18.5±2.2	0.44±0.11	0.12±0.01	11.4E4±9.6E4

Table 3. Mean and standard deviation values of metrics for SplatFacto model.

Object	Damage	SSIM	PSNR	LPIPS	FPS	NRS
Borodyanka	4	0.42±0.11	16.1±2.4	0.50±0.07	0.55±0.09	49.0E4±8.8E4
Hotel	3	**0.93±0.01**	**29.0±1.4**	**0.06±0.01**	**0.63±0.08**	**57.6E4±7.5E4**
Residential	2	0.75±0.18	23.3±4.1	0.18±0.11	0.55±0.08	50.8E4±8.0E4
Bridge	mixed	0.78±0.08	23.9±3.2	0.28±0.14	0.56±0.10	51.6E4±9.1E4
Cathedral	0	0.76±0.13	23.1±3.6	0.24±0.09	0.60±0.10	55.0E4±8.6E4

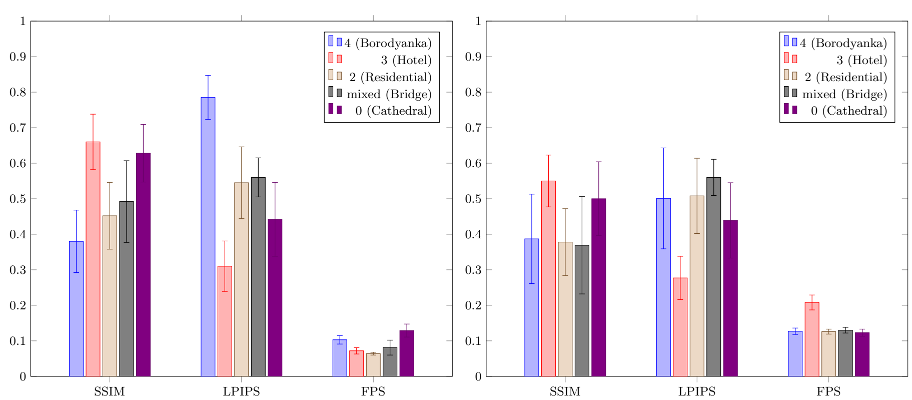

Fig. 6. Graphic visualization of some metrics for the objects with various damage levels for Instant-NGP model from Table 1 (left) and Nerfacto model from Table 2 (right).

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23

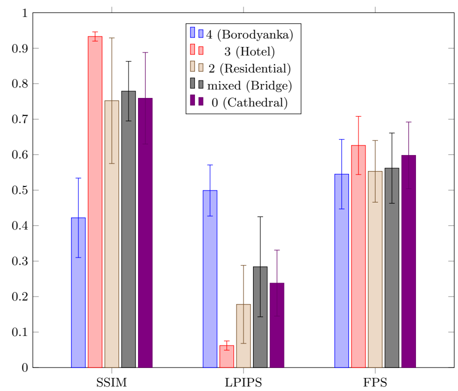

Fig. 7. Graphic visualization of some metrics for the objects with various damage levels for SplatFacto model from Table 3.

5 Discussion

This chapter delves deeper into the performance of the three rendering models: Splatfacto, Nerfacto, and Instant-NGP. The models were assessed under varying object damage levels (0: no damage, 4: high damage) using both visual quality metrics (SSIM, LPIPS, PSNR) and rendering speed metrics (FPS, rays per second).

Based on the results (Table 1, 2, and 3), one can make the following observations regarding the correlations between metrics and damage levels for various models that are summarized in Tables 4-8 and visualizations (Fig.8-10).

Table 4. Mean and standard deviation values of metrics for damage level 0 ("No Damage").

Model	SSIM	PSNR	LPIPS	FPS	NRS
Instant-NGP	0.63±0.08	20.6±2.3	0.44±0.10	0.13±0.02	11.8E4±1.7E4
Nerfacto	0.50±0.10	18.5±2.2	0.44±0.11	0.12±0.01	11.4E4±9.6E4
SplatFacto	**0.76±0.13**	**23.1±3.6**	**0.24±0.09**	**0.60±0.09**	**55.0E4±8.6E4**

Table 5. Mean and standard deviation values of metrics for the damage level 2.

Model	SSIM	PSNR	LPIPS	FPS	NRS
Instant-NGP	0.45±0.09	18.2±1.5	0.54±0.10	0.064±0.004	5.9E4±3.9E4
Nerfacto	0.38±0.09	16.9±1.8	0.51±0.11	0.126±0.007	11.6E4±6.6E4
SplatFacto	**0.75±0.18**	**23.3±4.1**	**0.18±0.11**	**0.55±0.09**	**50.9E4±8.0E4**

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23

Table 6. Mean and standard deviation values of metrics for the damage level 3.

Model	SSIM	PSNR	LPIPS	FPS	NRS
Instant-NGP	0.66±0.08	20.8±2.1	0.31±0.07	0.07±0.01	6.7E4±7.9E4
Nerfacto	0.55±0.07	18.8±1.4	0.27±0.06	0.21±0.02	19.2E4±1.9E4
SplatFacto	**0.93±0.01**	**29.0±1.4**	**0.06±0.01**	**0.63±0.08**	**57.6E4±7.5E4**

Table 7. Mean and standard deviation values of metrics for the damage level 4.

Model	SSIM	PSNR	LPIPS	FPS	NRS
Instant-NGP	0.38±0.09	15.5±1.8	0.79±0.06	0.10±0.01	9.3E4±1.1E4
Nerfacto	0.39±0.13	15.4±2.4	0.50±0.14	0.13±0.01	11.4E4±8.3E4
SplatFacto	**0.42±0.11**	**16.1±2.4**	**0.50±0.07**	**0.55±0.01**	**49.0E4±8.9E4**

SplatFacto achieves the highest SSIM and PSNR scores for all damage levels. This indicates that SplatFacto generally produces reconstructions that best preserve the structural similarity between the original scene and the reconstructed one. Nerfacto is the worst one and Instant-NGP is intermediate between SplatFacto and Nerfacto.

SplatFacto demonstrates the most consistent and best performance in terms of the lowest LPIPS (perceptual similarity) scores across all damage levels. This indicates that for all damage scenarios SplatFacto generally produces reconstructions with the highest perceptual similarity to the ground truth.

FPS performance of models across damage levels demonstrate that SplatFacto achieves the highest FPS consistently across all damage levels. This indicates that SplatFacto can generate reconstructions at a significantly faster rate compared to other models. Nerfacto has the lowest FPS across all damage levels and it might be computationally expensive and unsuitable for real-time applications. Instant-NGP falls between SplatFacto and Nerfacto and shows a moderate improvement over Instant-NGP but doesn't reach the speed of SplatFacto.

Finally, SplatFacto consistently outperforms other models, Instant-NGP and Nerfacto, in both visual quality and rendering speed across all damage levels. It excels at preserving fine details and achieves impressive rendering speeds compared to Nerfacto and Instant-NGP.

Table 8. Mean and standard deviation values of metrics for the damage level "mixed".

Model	SSIM	PSNR	LPIPS	FPS	NRS
Instant-NGP	0.49±0.12	21.6±1.5	0.56±0.06	0.08±0.02	7.5E4±1.9E4
Nerfacto	0.37±0.14	18.8±1.6	0.56±0.05	0.13±0.01	11.9E4±7.3E4
SplatFacto	**0.78±0.08**	**23.9±3.2**	**0.28±0.14**	**0.56±0.10**	**51.7E4±9.1E4**

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23

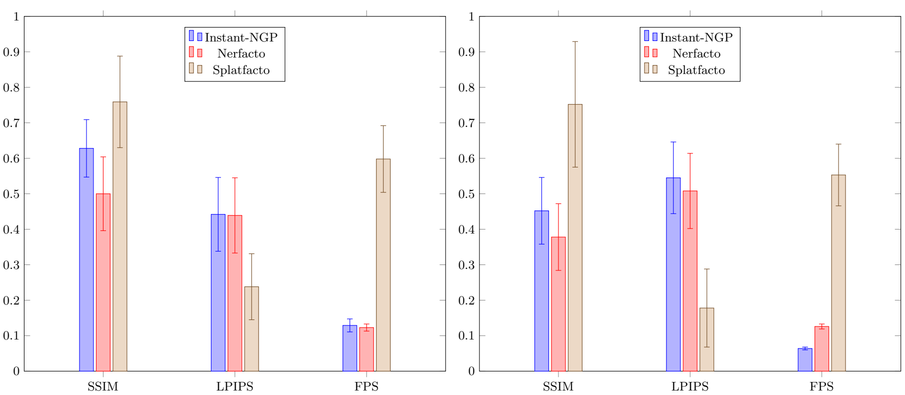

Fig. 8. Graphic visualization of some metrics for various models for the damage level "No Damage" (left) from Table 4 and 2 (right) Table 5.

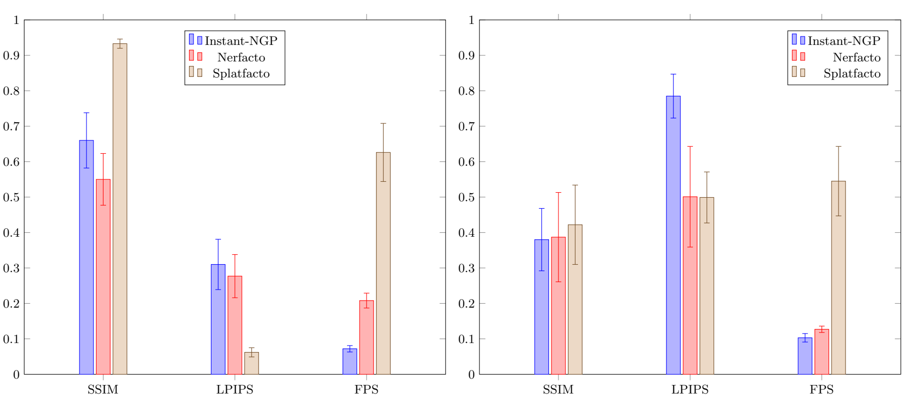

Fig. 9. Graphic visualization of some metrics for various models for the damage level 3 (left) from Table 6 and the damage level 4 (right) from Table 7.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23

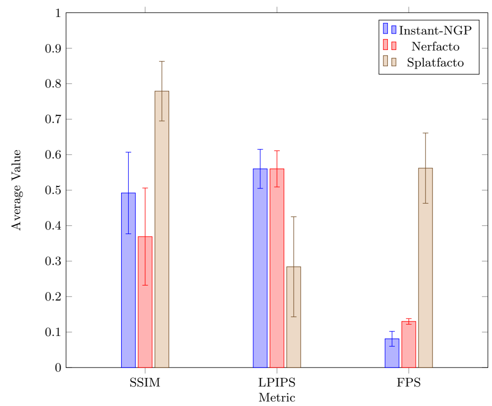

Fig. 10. Graphic visualization of some metrics from Table 8 for various models for the damage level "mixed".

6 Conclusions

The study extensively tests and compares three modern NERF models on a dataset of drone footages capturing the aftermath of Russian attacks in Ukraine. Each model is tasked with reconstructing 3D representations of destroyed buildings, offering insights into the spatial and structural aspects of the damage. The study successfully demonstrates the potential of NERF models for 3D reconstruction of damaged structures in conflict zones. The findings suggest that Splatfacto offers a compelling combination of high-quality reconstructions (measured by SSIM, PSNR, and LPIPS), fast rendering speed and computational efficiency. The paper acknowledges that the current research is in its initial stages, and further development is needed to improve the accuracy and efficiency of NERF models for real-world applications.

The research highlights several potential areas for future research. The impact of training data variations should be investigated, especially, sensitivity of the model performance with different datasets capturing diverse damage scenarios, lighting conditions, and viewpoints. The usage of various advanced NERF models, such as K-planes [8], TensoRF [12], Neuralangelo [7], and others, could be explored to understand how different model architectures could improve the reconstruction quality while maintaining efficiency. Also additional modalities (auxiliary data such as depth maps or semantic segmentation) should be considered for incorporation to enhance the reconstruction accuracy. The very important aspect is related to integration of this approach in real-time reconstruction systems, especially, in the context of optimisation of NERF models for real-time applications, enabling faster generation of 3D representations for immediate damage assessment or monitoring purposes.

By addressing these future research directions, one can further refine NERF technology for comprehensive documentation and analysis of conflict zones, po-

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23

tentially aiding humanitarian efforts and historical preservation. Overall, the study highlights the promising potential of NERF technology for revolutionizing the way one can use to document and analyze conflict zones, offering valuable insights for future research and real-world applications.

References

1. Deepak Rawtani, Gunjan Gupta, Nitasha Khatri, Piyush K Rao, and Chaudhery Mustansar Hussain. Environmental damages due to war in ukraine: A perspective. *Science of The Total Environment*, 850:157932, 2022.
2. Paulo Pereira, Ferdo Bašić, Igor Bogunovic, and Damia Barcelo. Russian-ukrainian war impacts the total environment. *Science of The Total Environment*, 837:155865, 2022.
3. 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. *Communications of the ACM*, 65(1):99–106, 2021.
4. 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.
5. Ting-Hu Yu, Yu Feng, Sida Lai, Mengqi Li, Jinxiang Zhou, and Hui Bao. Mip-nerf: A multiscale representation for anti-aliasing neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12868–12877, 2021.
6. Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. Mip-nerf 360: Unbounded anti-aliased neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5470–5479, 2022.
7. Zhaoshuo Li, Thomas Müller, Alex Evans, Russell H Taylor, Mathias Unberath, Ming-Yu Liu, and Chen-Hsuan Lin. Neuralangelo: High-fidelity neural surface reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8456–8465, 2023.
8. Sara Fridovich-Keil, Giacomo Meanti, Frederik Rahbæk Warburg, Benjamin Recht, and Angjoo Kanazawa. K-planes: Explicit radiance fields in space, time, and appearance. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12479–12488, 2023.
9. Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4), 2023.
10. Matthew Tancik, Ethan Weber, Evonne Ng, Ruilong Li, Brent Yi, Terrance Wang, Alexander Kristoffersen, Jake Austin, Kamyar Salahi, Abhik Ahuja, et al. Nerfstudio: A modular framework for neural radiance field development. In *ACM SIGGRAPH 2023 Conference Proceedings*, pages 1–12, 2023.
11. Nerfstudio Team. Nerfacto, 2022. `https://docs.nerf.studio/nerfology/methods/nerfacto.html`.
12. Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, and Hao Su. Tensorf: Tensorial radiance fields. In *European Conference on Computer Vision*, pages 333–350. Springer, 2022.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63751-3_23