PhysGaia: A Physics-Aware Dataset of Multi-Body Interactions for Dynamic Novel View Synthesis

Anonymous Author(s)
Affiliation
Address
email

Abstract

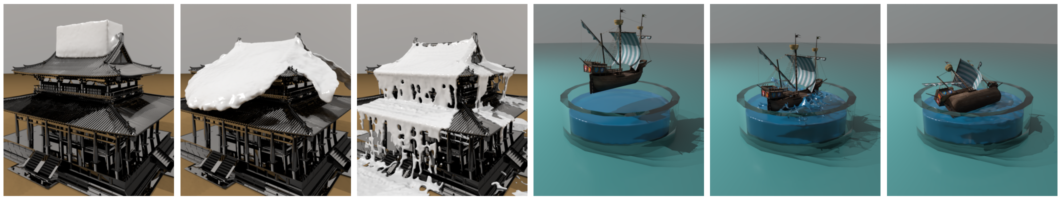
We introduce *PhysGaia*, a novel physics-aware dataset specifically designed for Dynamic Novel View Synthesis (DyNVS), encompassing both structured objects and unstructured physical phenomena. Unlike existing datasets that primarily focus on photorealistic reconstruction, *PhysGaia* is created to actively support physics-aware dynamic scene modeling. Our dataset provides complex dynamic scenarios with rich interactions among multiple objects, where they realistically collide with each other and exchange forces. Furthermore, it contains a diverse range of physical materials, such as liquid, gas, textile, and viscoelastic substances, which moves beyond the rigid bodies prevalent in existing datasets. All scenes in *PhysGaia* are faithfully generated to strictly adhere to physical laws, leveraging carefully selected material-specific physics solvers. To enable quantitative evaluation of physical modeling, our dataset provides essential ground-truth information, including 3D particle trajectories and physics parameters, e.g., viscosity. To facilitate research adoption, we also provide essential integration pipelines for using state-of-the-art 4D Gaussian splatting models with our dataset and report theirs results. By addressing the critical lack of datasets for physics-aware modeling, *PhysGaia* will significantly advance research in dynamic view synthesis, physics-based scene understanding, and deep learning models integrated with physical simulation–ultimately enabling more faithful reconstruction and interpretation of complex dynamic scenes.

1 Introduction

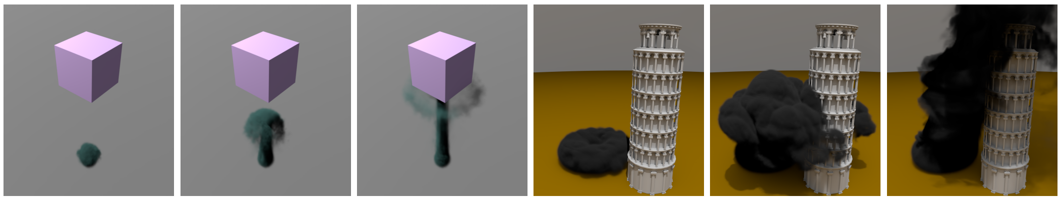
Since the introduction of Neural Radiance Fields (NeRF), Novel View Synthesis (NVS) algorithms based on deep learning have advanced rapidly. While early research primarily targeted static scenes, recent efforts have extended the scope to dynamic scene understanding to better support diverse and interactive AR/VR applications. This emerging direction, known as Dynamic Novel View Synthesis (DyNVS), aims to reconstruct 4D scenes from input videos and synthesize photorealistic images at novel viewpoints and time steps that are not seen during training.

The evolution of DyNVS has been closely tied to the development of suitable datasets. Early DyNVS datasets [1–3] primarily provided multiview training videos and typically involved only limited object motion. These datasets laid the foundation for early DyNVS research [2, 4–10], which focused on modeling scene dynamics by deforming canonical geometries over time. Recently, to better support AR/VR applications, datasets captured using handheld mobile devices have been introduced [6, 11, 12]. These datasets consist of monocular videos, and recent research has focused on mitigating overfitting under such conditions. Building on these datasets, subsequent work [13–19] has commonly adopted 4D Gaussian Splatting (4DGS) models, with a primary focus on resolving the inherent ambiguities of monocular video reconstruction and achieving photorealistic rendering.

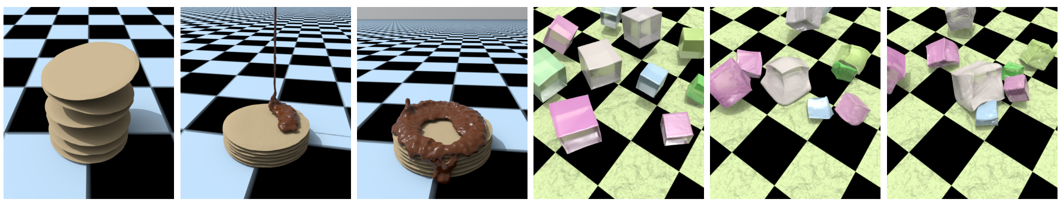
Submitted to 39th Conference on Neural Information Processing Systems (NeurIPS 2025). Do not distribute.



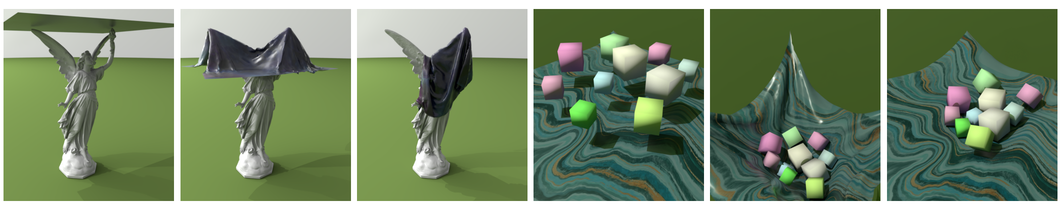
(a) Liquid with FLIP solver: *Hanok* (left) and *Ship* (right)



(b) Gas with Pyro solver: *Box-smoke* (left) and *Pisa* (right)



(c) Viscoelastic substances with MPM solver: *Pancake* (left) and *Jelly party* (right)



(d) Textile with Vellum solver: *Lucy* (left) and *Basin* (right)

Figure 1: Examples from the proposed physics-aware dataset, PhysGaia. They exhibit complex physical interactions between multiple objects composed of diverse materials such as liquid, gas, viscoelastic substance, and textile. This dataset will foster physics reasoning in dynamic scenes.

As DyNVS continues to advance, a natural next step is to move beyond photorealism and incorporate physical realism—enabling models not only to render how scenes look, but also to reason about how they behave. Recent pioneering works [20–24] have begun exploring this direction by integrating differentiable physics simulation into 4DGS frameworks. Despite this growing interest, current research largely remains limited to 4D generation or simplified DyNVS scenarios [24, 25], often restricted to single object or single material. Consequently, more complex senarios—such as multi-object interactions or the modeling of diverse physical materials like liquids and gases—are still significantly underexplored. In response, we introduce PhysGaia, a new benchmark designed to support and accelerate research in these emerging and challenging directions.

Our contributions are summarized as follows:

- We introduce PhysGaia, a physics-aware dataset featuring rich interactions among multiple objects and encompassing a wide range of physical materials, including liquids, gases, textiles, and viscoelastic substances, as illustrated in Figure 1.
- PhysGaia provides essential ground-truth information, such as 3D particle trajectories and physical parameters, enabling quantitative evaluation of physical modeling.
- We test existing state-of-the-art DyNVS methods on PhysGaia, revealing their fundamental limitations in achieving physical realism and demonstrating the potential for improvement in this field.

Table 1: Comparison in existing Dynamic Novel View Synthesis (DyNVS) datasets. "# total scene" denotes the number of scenes having test videos for DyNVS evaluation. "# physics" indicates the number of scenes exhibiting physical phenomena like gas, liquid, viscoelastic, or textiles. Unlike existing datasets, our dataset provides diverse scenes containing complex multi-object interactions. Also, our dataset contains the ground-truth physics information such as physics parameters like viscosity, and ground-truth 3D trajectories.

Datasets	**Scene Stats**			**Physics Information**		**Capture Type**
	# total scene	# physics	Interaction	Physics param.	3D Traj.	
Plenoptic [1]	6	1 (gas)	No	No	No	Multiview
D-NeRF [2]	8	1 (viscoelastic)	No	No	No	Monocular
NVIDIA Dynamic [3]	8	0	No	No	No	Monocular
Nerfies [6]	4	0	No	No	No	Monocular
HyperNeRF [11]	4	0	No	No	No	Monocular
DyCheck [12]	7	1 (textile)	No	No	No	Monocular
NeRF-DS [48]	8	0	No	No	No	Monocular
EvDNeRF [49]	6	0	No	No	No	Multiview
Synthetic Soccer [50]	3	0	No	No	No	Multiview
HDR-HexPlane [51]	8	0	No	No	No	Multiview
***PhysGaia* (Ours)**	**17**	**17**	Yes	Yes	Yes	Multiview/ Monocular

The rest of this paper is organized as follows. Section 2 reviews related work and highlights the uniqueness of our dataset. We discuss the main properties and potential research enabled by our dataset in Section 3. Section 4 details material-specific physics solvers used for dataset construction and methodologies for generating multi-object interaction scenarios. Section 5 shows our analysis, including evaluations of existing 4DGS methods on *PhysGaia*, We conclude the paper in Section 6.

2 Related Work

2.1 Dynamic Novel View Synthesis

In recent years, significant advances have been made in novel view synthesis [26–31]. Although this field initially have focused on static scene reconstruction but now it extended to handling dynamic scenes, which is known as Dynamic Novel View Synthesis (DyNVS). Early DyNVS methods were built on Neural Radiance Fields (NeRF) [4–10], which usually modeled scene dynamics either by implictly modeling with temporal inputs [4, 5] or directly estimating the time-wise deformation of canonical geometry through auxiliary neural networks [6–10]. Following the emergence of 3D Gaussian Splatting (3DGS)[31], recent DyNVS research has shifted toward Gaussian-based representations, leading to the development of 4D Gaussian Splatting (4DGS) [16, 32–43]. In 4DGS, an additional deformation network is employed to animate canonical Gaussian primitives over time, enabling efficient and high-quality modeling of dynamic scenes.

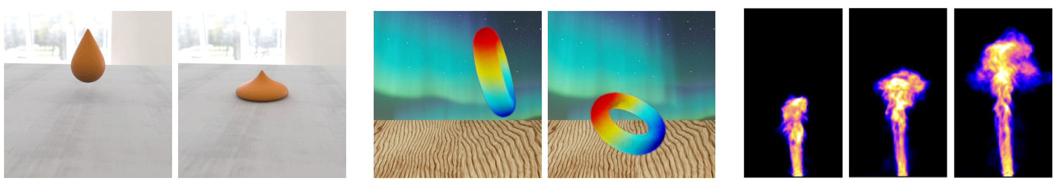
Modeling dynamic scenes with 4DGS is now leaning towards incorporating physical laws to govern motion. PhysGaussian [20] pioneers this by combining an MPM simulator with Gaussian Splatting, where each Gaussian primitive is handled as a particle within the MPM's particle-grid simulation. This work has inspired many subsequent studies [25, 24, 44–47] integrating physics-aware priors into 4DGS; however, these efforts remain largely confined to generation tasks [44–47], with only a few addressing Dynamic Novel View Synthesis (DyNVS) [25, 24]. Even for DyNVS, methods are typically limited to single objects or single materials, with most focusing on viscoelastic substances. Consequently, exploring physics-aware DyNVS research with rich object interactions with diverse physical materials remains underexplored, and we believe our dataset can serve as a foundation for this direction.

2.2 4D Datasets for Dynamic Novel View Synthesis

Table 1 provides a comparative overview of our dataset alongside existing multiview and monocular DyNVS datasets. The initial DyNVS datasets [1–3] primarily employed multiview configurations or captured scenes with very limited motion, typically involving mostly rigid objects. These early datasets paved the way for research into DyNVS, with subsequent work [2, 4–10] exploring paradigms of deforming canonical geometries over time to model scene dynamics.

Table 2: Comparison with physics-simulated datasets. "Multi-obj. Inter." denotes multi-object interaction. Unlike existing datasets, our PhysGaia offers rich object interactions, a diverse range of physical materials, and access to accurate physics parameters and ground truth 3D trajectories, making it a uniquely valuable resource for advancing physics-based understanding of dynamic scenes.

Datasets	**Multi-obj. Inter.**	**Materials**				**Physics Information**	
		Liquids	Gas	Textile	Viscoelastic	Physics params.	3D traj.
DNG [57]	Yes			✓		No	Yes
CLOTH4D [58]	Yes			✓		No	Yes
4D-DRESS [59]	Yes			✓		No	Yes
Rasheed *et al.* [60]	No			✓		Yes	No
Deng *et al.* [61]	No	✓				No	No
PAC-NeRF [62]	No	✓			✓	Yes	Yes
Spring-Gaus [25]	No				✓	Yes	Yes
ScalarFlow [63]	No		✓			Yes	Yes
PhysGaia (Ours)	Yes	✓	✓	✓	✓	Yes	Yes



(a) PAC-NeRF [62] (b) Spring-Gaus [25] (c) ScalarFlow [63]

Figure 2: Visualization of datasets most similar to our PhysGaia. While all of these datasets address physical phenomena, they are limited in several key aspects: limited coverage of physical materials, overly simplified dynamics, and an absence of rich multi-object interactions.

To advance DyNVS toward practical AR/VR applications, more user-friendly datasets, often captured using handheld mobile phones, were later introduced. Nerfies [6] pioneered handheld iPhone captures, though its scenes remained largely static scenarios. HyperNeRF [11] then introduced more rapid and varied motions, and DyCheck [12] further resolved camera teleportation issues seen in HyperNeRF. These monocular datasets have motivated research mitigating overfitting to training videos in these settings. Some approaches [13, 15, 14] focus on leveraging additional priors such as diffusion models [52], depth estimation model [53], and point trackers [54], while others [16, 14, 18, 17, 19] emphasize constraining deformation adopting techniques like motion factorization [55] or As-Rigid-As-Possible (ARAP) regularization [56]. Nonetheless, the primary objective across these datasets remains photorealistic reconstruction, with limited emphasis on physics-aware dynamic modeling [1].

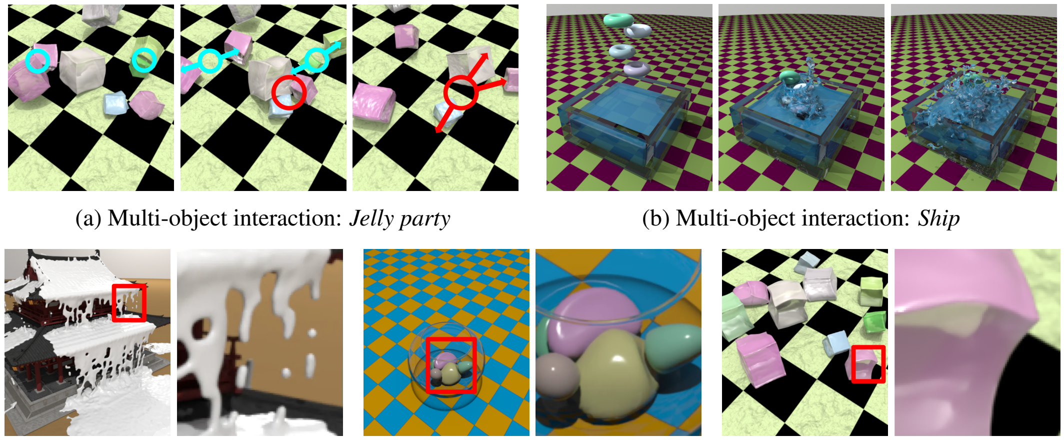
Unlike these existing DyNVS datasets, our PhysGaia offers the scenes containing multi-object interaction with diverse physics materials, as shown in Table 1. Although some datasets [1, 2, 12] include scenes with physical phenomena, these are very limited in number, not generated using accurate physics solvers, and lack complex multi-object interactions. In this context, PhysGaia occupies a unique position and holds strong potential to spur advancements in physics-aware dynamic scene modeling.

2.3 4D Datasets from Physics Simulator

To provide a comprehensive comparison, we also compare our PhysGaia with the 4D datasets generated via physics simulation, as summarized in Table 2. Since DyNVS tasks require multiview RGB imagery, we focus on the existing datasets that offer such data. Figure 2 visualizes several representative datasets, including PAC-NeRF [62], Spring-Gaus [25], and ScalarFlow [63].

Some existing datasets support multi-object interactions, but they either lack rich interaction dynamics or are limited to a single type of material. For instance, clothed human datasets [57–59] naturally exhibit textile-body interactions, yet are restricted to textile materials and human-centric motions. In terms of ground-truth physical information, a few datasets [62, 25, 63] provide both physical

[1]Note that some recent datasets have also been introduced, but they are typically tailored to specific scenarios, such as event cameras [49], HDR rendering [51], specular lighting effects [48], or human motion capture [50].



(a) Multi-object interaction: *Jelly party* (b) Multi-object interaction: *Ship*

(c) Splashing (non-locally rigid) (d) Specular Modeling (e) Refraction

Figure 3: Visualization of physics properties in *PhysGaia*. Alongside multi-object interactions, *PhysGaia* also includes various physical phenomena like splashing, refraction, and specular effects.

parameters and 3D trajectories, but these datasets are also typically confined to single-material, do not capture the kind of rich, multi-object interactions, as shown in Figure 2a, 2b, and 2c. PAC-NeRF [62], for instance, focuses on liquids and viscoelastic materials; however, its liquid scenarios are constrained to highly viscous flows, making them behaviorally similar to viscoelastic materials. Spring-Gaus [25] is restricted to viscoelasticity, while ScalarFlow [63] concentrates on gas. In contrast, our PhysGaia directly tackles these limitations by providing rich object interactions, a diverse range of physical materials, and access to accurate simulation parameters and trajectories, which makes it a uniquely valuable dataset for advancing physics understanding of dynamic scenes.

3 Dataset Properties and Research Impact

We propose *PhysGaia* to advance physically realistic reconstruction in DyNVS, moving beyond mere photorealism. As highlighted in Tables 1 and 2 within Section 2, our dataset uniquely features complex multi-object interactions involving a diverse range of physical materials, distinguishing it within the DyNVS landscape. We believe this comprehensive dataset holds significant potential to enhance the understanding of physics in dynamic scenes. In this section, we detail the specific properties of our dataset and discuss the future research directions it enables.

3.1 Dataset Properties

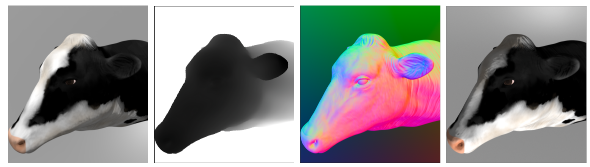
Complex scenarios with physics-aware dynamics Our *PhysGaia* dataset consists of 17 scenes with multi-object interaction, as visualized in Figure 1. We further visualize the detailed examples in Figure 3a and Figure 3b, which exhibit viscoelastic objects colliding leading rapid moving direction change and rigid-liquid surface hitting leading splashing, respectively. To ensure the scenes adhere to physical laws with accurately calculated force exchange among objects, we carefully select material-specific solvers: FLIP for liquids, Pyro for gases, Vellum for textiles, and MPM for viscoelastic materials. Further details on simulation configurations for handling multi-object interactions can be found in Section 4.

Beyond multi-object interactions, our dataset also exhibits various physical phenomena, such as non-locally rigid motion commonly observed in liquid and gas scenes, specular reflection, and refraction, as shown in Figures 3c, 3d, and 3e, respectively. These properties enhance the realism of our dataset and enable a wide range of downstream tasks and research applications.

Providing physics parameters In contrast to real-world video datasets, where the ease of capture is offset by inaccessible underlying physics, our simulated dataset offers complete access to all physical information. This ground-truth includes 3D particle trajectories and physics parameters such as viscosity, Young's modulus, Poisson's ratio, and temperature for gas scene[2]. This comprehensive provision enables precise evaluation of physical reasoning in dynamic scenes, directly facilitating the future research directions outlined in Subsection 3.2.

Supporting diverse DyNVS tasks Our dataset uniquely supports both multiview and monocular DyNVS. Unlike most multiview datasets [1, 49–51] with videos captured from fixed camera positions like CCTV, ours provides several moving monocular video sequences from independent trajectories as training videos. This allows diverse training configurations; using full sequences corresponds to multiview DyNVS task, while using a single sequence corresponds to monocular DyNVS task. The monocular setup matches the DyCheck [12] dataset, providing realistic handheld camera inputs without synthetic teleportation. For evaluation, we employ two static cameras with a large baseline.

Customizability We provide the complete simulation node graphs and their exact parameter settings used to build our dataset. These graphs encompass all relevant components, including physics solvers, source geometries, camera positions, lights, materials, and texture controls. Thus, by modifying these nodes, users can easily generate customized scenes or additional modalities such as depth maps, surface normals, and re-lighted images tailored to their specific downstream tasks, as shown in Figure 4.



(a) RGB (b) Depth (c) Normal (d) Re-lighted

Figure 4: Examples of diverse modalities that users can generate from the provided simulation node graphs.

Accessibility Our *PhysGaia* benchmark is designed for research-friendly and easy access, providing integration pipelines that enable the use of state-of-the-art 4D Gaussian Splatting models [14, 32–34] with our data. Furthermore, we include COLMAP-reconstructed point clouds for each scene. These provisions aim to facilitate the adoption of our dataset by researchers working with state-of-the-art DyNVS models.

3.2 Potential Research

This subsection highlights the potential impact of our *PhysGaia* dataset by outlining several promising future research directions it uniquely enables.

Physical reasoning of dynamic scenes Since our PhysGaia dataset provides ground-truth physics information, it facilitates precise evaluation of physical reasoning in dynamic scenes. For instance, ground-truth physics parameters like viscosity can be used to evaluate inverse physics estimation methods, where differentiable simulators are employed to optimize these parameters. Furthermore, unlike existing 4DGS research that primarily focuses on photorealism and relies on ground-truth RGB images, our dataset offers ground-truth 3D trajectories, enabling evaluation of the actual motion of individual Gaussian primitives in 4DGS. We believe this unique feature establishes our dataset as a valuable benchmark for developing and evaluating physics-aware DyNVS models.

Multi-object interaction While recent research integrates physics into DyNVS [24, 25], it largely remains limited to single materials and often single objects within scenes. As a result, the crucial aspect of physics reasoning for interactions between multiple objects in DyNVS – particularly the estimation of force exchange and deformation during contact – remains largely unexplored. We believe our novel dataset, specifically designed with complex multi-object interactions, will be instrumental in enabling significant future research in multi-physics modeling and adaptive representations for handling hybrid scenes.

Integration of material-specific physics solver The dominant approach for integrating physics into DyNVS algorithms currently involves adopting differentiable simulators [64–69], treating Gaussian primitives as particles in simulators. However, we emphasize that different physical phenomena are best captured by different physics solvers, a principle reflected in our dataset construction process

[2]Note that temperature itself is very crucial component when adopting smoke-related simulator to physical reasoning for buoyancy calculation

detailed in Section 4. This can guide researchers seeking to integrate more appropriate solvers tailored to specific material behaviors, such as fluids, cloth, smoke. For example, the FLIP [70] solver excels at simulating incompressible fluids due to its hybrid particle-grid representation, offering greater stability and realism compared to purely particle-based SPH-based solver. Similarly, thermodynamic effects like temperature and buoyancy are crucial for smoke simulation, typically represented using voxel-based grids. Integrating such volumetric solvers into particle-based frameworks like Gaussian splatting, however, remains a largely unexplored area.

4 Dataset Construction

This section details solver properties for liquids, gases, viscoelastic substances, and textiles, along with some simulation details for dynamics with multi-object interactions. Please refer to our supplemental document for other simulation configurations.

4.1 Common Setup

Simulator selection We select SideFX Houdini 20.5 as the foundation of our physics-informed data-generation pipeline because it integrates multiple physics solvers within a unified procedural environment. By sharing a common computational graph, it ensures consistent multi-material interactions under uniform boundary conditions. We access simulation data such as particle positions and flow fields on a per-frame basis via its Python API.

Rendering We render all frames at a resolution of 640×720, using NVIDIA OptiX denoiser and path tracing with 256 samples per pixel. Scenes are illuminated with 1–3 point lights (intensity: 600–4000), shadows are disabled in textile-focused scenes like *tube-flag* to emphasize geometry.

4.2 Liquid

For liquid scenes, we adopt the Fluid-Implicit Particle (FLIP) solver [70], a hybrid particle-grid method. FLIP maintains particle velocities throughout the simulation, using the grid solely to compute and apply forces such as pressure and viscosity. This approach preserves fine-scale, high-frequency particle velocities, crucial for modeling realistic and rapid liquid behavior while adhering to the Navier–Stokes equations. Although the Material Point Method (MPM) [64] solver can also model fluids, its direct velocity aggregation onto the grid limits its ability to capture highly dynamic fluid phenomena like splashing. While particle-only-based solver [71] are another option, FLIP is generally better suitable for incompressible fluids thanks to its hybrid grid-based representation.

Dynamic interactions For the *ice* and *hanok* scenes, where fluid spills onto fixed objects, we adopt the surface operator to simulate multi-object interaction. Since the interaction occurs primarily near the surface, this approach reduces computational overhead. In contrast, for the *ship* and *cereal* scenes–where objects fall into liquid, causing both fluid and objects to move and influence each other with force exchange–we use a dynamic operator to accurately model these more complex interactions.

4.3 Gas

For gas (smoke) simulation, we utilize the Pyro solver [72], which models the temperature field essential for accurately capturing buoyancy effects in gaseous materials. Pyro employs grid-based representations of density, velocity, and temperature, ensuring compliance with the Navier–Stokes equations governing fluid mechanics. Since the ground-truth motion is represented as a velocity field and storing full velocity fields can require up to 2GB per frame, we provide a subsampled set of particle trajectories per scene to facilitate efficient data storage and processing.

Dynamic interactions In the *pisa* scene, we reduced the voxel size from the default 0.1 to 0.05 to better capture the tower's intricate details and added a lateral wind of speed 2 to wrap the plume around it. For the other scenes involving this material, we used the default simulation settings.

4.4 Viscoelastic Substances

MPM [64] extends FLIP [73] to handle solid mechanics and is ideal for simulating chunk-based, viscoelastic substances like snow, jelly, and soil. It aggregates particle information on a grid, performs

Table 3: Quantitative results of existing 4D Gaussian Splatting models on the proposed PhysGaia dataset. While multiview setups generally offer better reconstruction performance than monocular ones, even multiview results achieve PSNR scores below 30. This highlights the substantial difficulty in reconstructing the complex multi-object interactions in our dataset.

Capture Type	Method	Liquid			Gas		
		PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
Monocular	D-3DGS [32]	22.7	**0.87**	**0.22**	21.9	**0.89**	**0.16**
	4DGS [33]	**24.2**	**0.87**	0.23	21.7	0.88	0.17
	STG [34]	19.2	0.72	0.39	**21.9**	0.85	0.24
	SOM [14]	18.2	0.71	0.54	20.5	0.83	0.28
Multiview	D-3DGS [32]	22.2	0.87	0.24	23.7	**0.91**	**0.13**
	4DGS [33]	**25.1**	**0.88**	**0.22**	24.2	0.89	0.17
	STG [34]	20.8	0.75	0.40	**25.0**	**0.91**	0.19
Capture Type	Method	Viscoelastic materials			Textile		
		PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
Monocular	D-3DGS [32]	**20.1**	**0.84**	**0.15**	22.1	0.83	**0.18**
	4DGS [33]	19.5	0.82	0.18	**24.9**	**0.84**	**0.18**
	STG [34]	13.6	0.63	0.40	21.9	**0.84**	0.21
	SOM [14]	12.0	0.53	0.49	19.3	0.78	0.25
Multiview	D-3DGS [32]	**22.2**	**0.89**	**0.10**	**27.7**	**0.90**	**0.12**
	4DGS [33]	21.0	0.85	0.15	26.6	0.87	0.15
	STG [34]	17.2	0.70	0.36	21.1	0.81	0.25

computation, and reprojects to particles–making it effective for capturing deformation and internal force propagation.

Dynamic interactions For the *pancake* scene, we reduce the grid size from the default 0.025 to 0.002 to more faithfully capture its thin-sheet dynamics, while other scenes use the default grid size. To suppress spurious artifacts that can arise from aggregating particle properties onto the grid, we increase the number of samples participating in node calculations by oversampling. The oversampling scales are set to 6, 2, 4, and 2 for the *bouncing balls*, *cow*, *jelly party*, and *pancake* scenes, respectively. Additionally, for the *bouncing balls* scene, we add a static bowl-shaped collider so that the falling balls rebound off both one another and the bowl's surface.

4.5 Textile

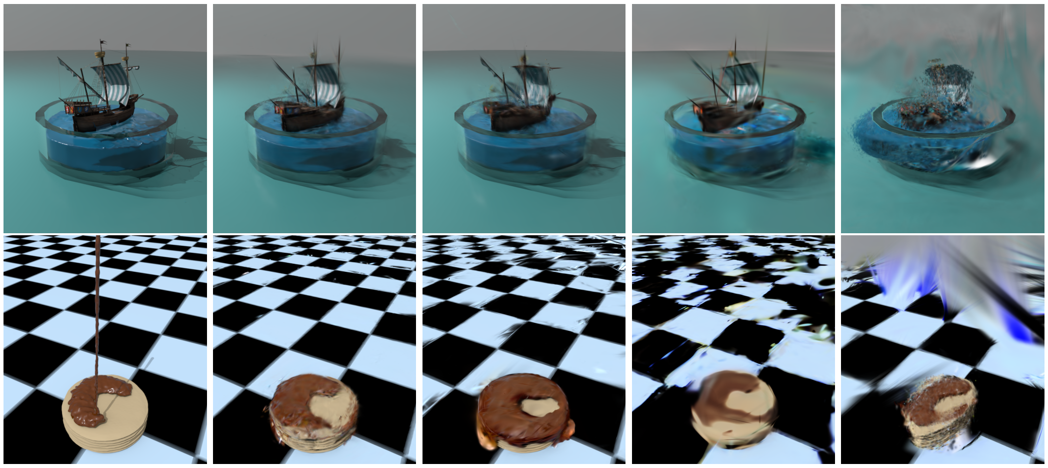
For textile materials, we adopt the Vellum solver [74], which is based on the Extended Position Based Dynamics (XPBD) framework [75]. XPBD improves upon classical Position Based Dynamics (PBD) [76] by integrating a Lagrange multiplier and its update. This effectively decouples material stiffness from the solver's time-step size and iteration count, making it a widely used method for simulating deformable objects, especially cloth.

Dynamic interactions To simulate interactions between objects and textiles–where both move and exchange forces, as seen especially in the *basin* scene–we employ the shape match constraint. This constraint helps maintain the overall shape of objects by driving points toward their rest configuration, allowing the material to preserve its structural integrity while still interacting dynamically with textiles and other objects. For the *lucy* scene, we increased the simulation sub-step count fivefold over the default to robustly handle collisions with the statue's complex geometry. To simulate wind-induced fluttering in the *flags*, *single-flag*, and *tube* scenes, we applied external forces using a POP Wind node that blows parallel to the ground plane.

5 Analysis

5.1 Existing Algorithms with *PhysGaia*

Implementation On the proposed *PhysGaia* dataset, we test the existing 4D Gaussian Splatting baselines: D-3DGS [32], 4DGS [33], STG [34], and SOM [14]. Except for SOM [14], which is specialized for monocular setup, all models were tested under both monocular and multiview setup. For evaluation, we adopt standard image quality metrics: peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM) [77], and learned perceptual image patch similarity (LPIPS) [78].



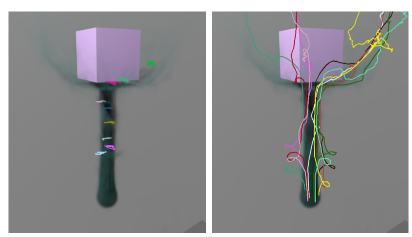
(a) GT image (b) D-3DGS [32] (c) 4DGS [33] (d) STG [34] (e) SOM [14]

Figure 5: Qualitative results of existing 4DGS methods on the *ship* and *pancake* scenes show that all methods frequently exhibit needle-like artifacts and under-reconstruct dynamic elements.

Quantitative and qualitative results Table 3 shows average quantitative results across material categories. While multiview setups generally offer better reconstruction performance than monocular ones, even multiview results achieve PSNR scores below 30. This highlights the difficulty of reconstructing complex multi-object interactions in our dataset, suggesting that incorporating physics-aware priors is crucial for accurate capture, rather than relying solely on RGB image fitting. SOM [14], which typically performs well on other datasets, does not consistently lead on ours as much as we expected. We presume its motion-factorization struggles with highly dynamic phenomena such as splashing and multiple interacting objects. Also, its heavy reliance on external modules like point trackers leads to performance drops when these modules fail in challenging scenes. Figure 5 shows qualitative results with monocular settings; all methods show need-like artifacts and some under-reconstruction of dynamic regions. Detailed results are available in the supplementary material.

5.2 Limitation on Physical Realism

We further analyze the inherent limitations of existing 4D Gaussian Splatting methods. Since Gaussian Splatting is an explicit representation, the motion of Gaussian primitives is expected to closely follow the actual trajectories of the scene. However, we observed that reconstructed trajectories often deviate from the ground truth, especially in liquid and gas scenes as shown in Figure 6. In such cases, photorealistic appearance can be achieved without faithfully modeling true motion due to their color similarity, allowing Gaussian primitives to remain near the surface and just fluctuate locally, rather than follow the actual upward movement.



(a) 4DGS [33] (b) Ground truth

Figure 6: (a) Trajectories of gaussian primitives from 4DGS [33] and (b) GT

6 Conclusion

We propose a novel physics-aware dataset, *PhysGaia*, specifically designed to understand physics in dynamic scenes, particularly for Dynamic Novel View Synthesis (DyNVS). Comprising 17 diverse scenes, our dataset captures complex multi-object interactions with a wide variety of materials. Each scene is faithfully generated using material-specific physics solvers, ensuring adherence to physical laws and providing rich ground-truth physics data including 3D particle trajectories and physics parameters. This ground truth data uniquely enables the evaluation of physics reasoning. We also test state-of-the-art DyNVS methods on *PhysGaia*, revealing their fundamental limitations in achieving physical realism and highlighting significant potential for improvement. We believe *PhysGaia* will be a critical resource that accelerates progress in physics-aware dynamic scene understanding.

References

[1] Li, T., Slavcheva, M., Zollhoefer, M., Green, S., Lassner, C., Kim, C., Schmidt, T., Lovegrove, S., Goesele, M., Newcombe, R., et al.: Neural 3d video synthesis from multi-view video. In CVPR. (2022)

[2] Pumarola, A., Corona, E., Pons-Moll, G., Moreno-Noguer, F.: D-nerf: Neural radiance fields for dynamic scenes. In CVPR. (2021)

[3] Yoon, J.S., Kim, K., Gallo, O., Park, H.S., Kautz, J.: Novel view synthesis of dynamic scenes with globally coherent depths from a monocular camera. In CVPR. (2020)

[4] Gao, C., Saraf, A., Kopf, J., Huang, J.B.: Dynamic view synthesis from dynamic monocular video. In ICCV. (2021)

[5] Du, Y., Zhang, Y., Yu, H.X., Tenenbaum, J.B., Wu, J.: Neural radiance flow for 4d view synthesis and video processing. In ICCV. (2021)

[6] Park, K., Sinha, U., Barron, J.T., Bouaziz, S., Goldman, D.B., Seitz, S.M., Martin-Brualla, R.: Nerfies: Deformable neural radiance fields. In ICCV. (2021)

[7] Cao, A., Johnson, J.: Hexplane: A fast representation for dynamic scenes. In CVPR. (2023)

[8] Fridovich-Keil, S., Meanti, G., Warburg, F.R., Recht, B., Kanazawa, A.: K-planes: Explicit radiance fields in space, time, and appearance. In CVPR. (2023)

[9] Shao, R., Zheng, Z., Tu, H., Liu, B., Zhang, H., Liu, Y.: Tensor4d: Efficient neural 4d decomposition for high-fidelity dynamic reconstruction and rendering. In CVPR. (2023)

[10] Fang, J., Yi, T., Wang, X., Xie, L., Zhang, X., Liu, W., Nießner, M., Tian, Q.: Fast dynamic radiance fields with time-aware neural voxels. In SIGGRAPH Asia. (2022)

[11] Park, K., Sinha, U., Hedman, P., Barron, J.T., Bouaziz, S., Goldman, D.B., Martin-Brualla, R., Seitz, S.M.: Hypernerf: A higher-dimensional representation for topologically varying neural radiance fields. ACM Trans. Graph. (2021)

[12] Gao, H., Li, R., Tulsiani, S., Russell, B., Kanazawa, A.: Monocular dynamic view synthesis: A reality check. In NeurIPS. (2022)

[13] Kim, M., Lim, J., Han, B.: Ua-4dgs: 4d gaussian splatting in the wild with uncertainty-aware regularization. In NeurIPS. (2024)

[14] Wang, Q., Ye, V., Gao, H., Austin, J., Li, Z., Kanazawa, A.: Shape of motion: 4d reconstruction from a single video (2024) arXiv preprint arXiv:2407.13764.

[15] Zhang, T., Gao, Q., Li, W., Liu, L., Chen, B.: Bags: Building animatable gaussian splatting from a monocular video with diffusion priors (2024)

[16] Huang, Y.H., Sun, Y.T., Yang, Z., Lyu, X., Cao, Y.P., Qi, X.: Sc-gs: Sparse-controlled gaussian splatting for editable dynamic scenes. In CVPR. (2024)

[17] Kwak, S., Kim, J., Jeong, J.Y., Cheong, W.S., Oh, J., Kim, M.: Modec-gs: Global-to-local motion decomposition and temporal interval adjustment for compact dynamic 3d gaussian splatting (2025) arXiv preprint arXiv:2501.03714.

[18] Kratimenos, A., Lei, J., Daniilidis, K.: Dynmf: Neural motion factorization for real-time dynamic view synthesis with 3d gaussian splatting. In ECCV. (2024)

[19] Cai, W., Ye, W., Ye, P., He, T., Chen, T.: Dynasurfgs: Dynamic surface reconstruction with planar-based gaussian splatting (2024) arXiv preprint arXiv:2408.13972.

[20] Xie, T., Zong, Z., Qiu, Y., Li, X., Feng, Y., Yang, Y., Jiang, C.: Physgaussian: Physics-integrated 3d gaussians for generative dynamics. In CVPR. (2024)

[21] Zhang, T., Yu, H.X., Wu, R., Feng, B.Y., Zheng, C., Snavely, N., Wu, J., Freeman, W.T.: Physdreamer: Physics-based interaction with 3d objects via video generation. In ECCV. (2024)

[22] Liu, S., Ren, Z., Gupta, S., Wang, S.: Physgen: Rigid-body physics-grounded image-to-video generation. In ECCV. (2024)

[23] Borycki, P., Smolak, W., Waczyńska, J., Mazur, M., Tadeja, S., Spurek, P.: Gasp: Gaussian splatting for physic-based simulations (2024) arXiv preprint arXiv:2409.05819.

[24] Jiang, H., Hsu, H.Y., Zhang, K., Yu, H.N., Wang, S., Li, Y.: Phystwin: Physics-informed reconstruction and simulation of deformable objects from videos (2025) arXiv preprint arXiv:2503.17973.

[25] Zhong, L., Yu, H.X., Wu, J., Li, Y.: Reconstruction and simulation of elastic objects with spring-mass 3d gaussians. In ECCV. (2024)

[26] Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: Nerf: Representing scenes as neural radiance fields for view synthesis. In ECCV. (2020)

[27] Chen, A., Xu, Z., Geiger, A., Yu, J., Su, H.: Tensorf: Tensorial radiance fields. In ECCV. (2022)

[28] Garbin, S.J., Kowalski, M., Johnson, M., Shotton, J., Valentin, J.: Fastnerf: High-fidelity neural rendering at 200fps. In ICCV. (2021)

[29] Wang, L., Zhang, J., Liu, X., Zhao, F., Zhang, Y., Zhang, Y., Wu, M., Yu, J., Xu, L.: Fourier plenoctrees for dynamic radiance field rendering in real-time. In CVPR. (2022)

[30] Müller, T., Evans, A., Schied, C., Keller, A.: Instant neural graphics primitives with a multiresolution hash encoding. In ACM TOG. (2022)

[31] Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G.: 3d gaussian splatting for real-time radiance field rendering. In ACM ToG. (2023)

[32] Yang, Z., Gao, X., Zhou, W., Jiao, S., Zhang, Y., Jin, X.: Deformable 3d gaussians for high-fidelity monocular dynamic scene reconstruction. In CVPR. (2024)

[33] Wu, G., Yi, T., Fang, J., Xie, L., Zhang, X., Wei, W., Liu, W., Tian, Q., Wang, X.: 4d gaussian splatting for real-time dynamic scene rendering. In CVPR. (2024)

[34] Li, Z., Chen, Z., Li, Z., Xu, Y.: Spacetime gaussian feature splatting for real-time dynamic view synthesis. In CVPR. (2024)

[35] Lin, Y., Dai, Z., Zhu, S., Yao, Y.: Gaussian-flow: 4d reconstruction with dynamic 3d gaussian particle. In CVPR. (2024)

[36] Lu, Z., Guo, X., Hui, L., Chen, T., Yang, M., Tang, X., Zhu, F., Dai, Y.: 3d geometry-aware deformable gaussian splatting for dynamic view synthesis. In CVPR. (2024)

[37] Guo, Z., Zhou, W., Li, L., Wang, M., Li, H.: Motion-aware 3d gaussian splatting for efficient dynamic scene reconstruction. In TCSVT. (2024)

[38] Liang, Y., Khan, N., Li, Z., Nguyen-Phuoc, T., Lanman, D., Tompkin, J., Xiao, L.: Gaufre: Gaussian deformation fields for real-time dynamic novel view synthesis. In WACV. (2025)

[39] Lei, J., Weng, Y., Harley, A., Guibas, L., Daniilidis, K.: Mosca: Dynamic gaussian fusion from casual videos via 4d motion scaffolds (2024) arXiv preprint arXiv:2405.17421.

[40] Duan, Y., Wei, F., Dai, Q., He, Y., Chen, W., Chen, B.: 4d-rotor gaussian splatting: towards efficient novel view synthesis for dynamic scenes. In SIGGRAPH. (2024)

[41] Waczynska, J., Borycki, P., Kaleta, J., Tadeja, S., Spurek, P.: D-miso: Editing dynamic 3d scenes using multi-gaussians soup. In NeurIPS. (2024)

[42] Liu, Q., Liu, Y., Wang, J., Lyv, X., Wang, P., Wang, W., Hou, J.: Modgs: Dynamic gaussian splatting from casually-captured monocular videos. In ICLR. (2025)

[43] Stearns, C., Harley, A., Uy, M., Dubost, F., Tombari, F., Wetzstein, G., Guibas, L.: Dynamic gaussian marbles for novel view synthesis of casual monocular videos. In SIGGRAPH. (2024)

[44] Jiang, Y., Yu, C., Xie, T., Li, X., Feng, Y., Wang, H., Li, M., Lau, H., Gao, F., Yang, Y., et al.: Vr-gs: A physical dynamics-aware interactive gaussian splatting system in virtual reality. In SIGGRAPH. (2024)

[45] Lin, Y., Lin, C., Xu, J., Mu, Y.: Omniphysgs: 3d constitutive gaussians for general physics-based dynamics generation. In ICLR. (2025)

[46] Huang, T., Zhang, H., Zeng, Y., Zhang, Z., Li, H., Zuo, W., Lau, R.W.: Dreamphysics: Learning physics-based 3d dynamics with video diffusion priors. In AAAI. (2025)

[47] Qiu, R.Z., Yang, G., Zeng, W., Wang, X.: Feature splatting: Language-driven physics-based scene synthesis and editing. In ECCV. (2024)

[48] Yan, Z., Li, C., Lee, G.H.: Nerf-ds: Neural radiance fields for dynamic specular objects. In CVPR. (2023)

[49] Bhattacharya, A., Madaan, R., Cladera, F., Vemprala, S., Bonatti, R., Daniilidis, K., Kapoor, A., Kumar, V., Matni, N., Gupta, J.K.: Evdnerf: Reconstructing event data with dynamic neural radiance fields. In WACV. (2024)

[50] Lewin, S., Vandegar, M., Hoyoux, T., Barnich, O., Louppe, G.: Dynamic nerfs for soccer scenes. In Multimedia Content Analysis in Sports. (2023)

[51] Wu, G., Yi, T., Fang, J., Liu, W., Wang, X.: Fast high dynamic range radiance fields for dynamic scenes. In 3DV. (2024)

[52] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In CVPR. (2022)

[53] Yang, L., Kang, B., Huang, Z., Xu, X., Feng, J., Zhao, H.: Depth anything: Unleashing the power of large-scale unlabeled data. In CVPR. (2024)

[54] Yang, Z., Du, Y., Sun, D., Jampani, V., Liu, C., Freeman, W.T., Tenenbaum, J.B., Wu, J.: Cotracker: Transformers for tracking any point. (2023) arXiv preprint arXiv:2303.06583.

[55] Tomasi, C., Kanade, T.: Shape and motion from image streams under orthography: A factorization method. International Journal of Computer Vision (1992)

[56] Sorkine, O., Alexa, M.: As-rigid-as-possible surface modeling. TOG (2007)

[57] Zhang, M., Wang, T.Y., Ceylan, D., Mitra, N.J.: Dynamic neural garments. TOG (2021)

[58] Zou, X., Han, X., Wong, W.: Cloth4d: A dataset for clothed human reconstruction. In CVPR. (2023)

[59] Wang, W., Ho, H.I., Guo, C., Rong, B., Grigorev, A., Song, J., Zarate, J.J., Hilliges, O.: 4D-DRESS: A 4d dataset of real-world human clothing with semantic annotations. In CVPR. (2024)

[60] Rasheed, A.H., Romero, V., Bertails-Descoubes, F., Wuhrer, S., Franco, J.S., Lazarus, A.: Learning to measure the static friction coefficient in cloth contact. In CVPR. (2020)

[61] Deng, Y., Yu, H.X., Wu, J., Zhu, B.: Learning vortex dynamics for fluid inference and prediction. In ICML. (2023)

[62] Li, X., Qiao, Y.L., Chen, P.Y., Jatavallabhula, K.M., Lin, M., Jiang, C., Gan, C.: Pac-nerf: Physics augmented continuum neural radiance fields for geometry-agnostic system identification. In ICLR. (2023)

[63] Marie-Lena Eckert, Kiwon Um, N.T.: Scalarflow: A large-scale volumetric data set of real-world scalar transport flows for computer animation and machine learning. In TOG. (2019)

[64] Sulsky, D., Chen, Z., Schreyer, H.L.: A particle method for history-dependent materials. Computer Methods in Applied Mechanics and Engineering (1994)

[65] Hu, Y., Li, T.M., Anderson, L., Ragan-Kelley, J., Durand, F.: Taichi: a language for high-performance computation on spatially sparse data structures. In TOG. (2019)

[66] Macklin, M.: Warp: A high-performance python framework for gpu simulation and graphics. https://github.com/nvidia/warp (March 2022) NVIDIA GPU Technology Conference (GTC).

[67] Authors, G.: Genesis: A universal and generative physics engine for robotics and beyond (2024)

[68] Hu, Y., Anderson, L., Li, T.M., Sun, Q., Carr, N., Ragan-Kelley, J., Durand, F.: Difftaichi: Differentiable programming for physical simulation. In ICLR. (2020)

[69] Hu, Y., Fang, Y., Ge, Z., Qu, Z., Zhu, Y., Pradhana, A., Jiang, C.: A moving least squares material point method with displacement discontinuity and two-way rigid body coupling. In TOG. (2018)

[70] Brackbill, J.U., Kothe, D.B., Ruppel, H.M.: Flip: A low-dissipation, particle-in-cell method for fluid flow. Computer Physics Communications (1988)

[71] Gingold, R.A., Monaghan, J.J.: Smoothed particle hydrodynamics: theory and application to non-spherical stars. Monthly Notices of the Royal Astronomical Society (1977)

[72] SideFX Software: Pyro solver. https://www.sidefx.com/docs/houdini/pyro/intro.html (2012)

[73] Brackbill, J.U.: Flip: A low-dissipation, particle-in-cell method for fluid flow. Journal of Computational Physics (1986)

[74] SideFX Software: Vellum solver. https://www.sidefx.com/docs/houdini/vellum/overview.html (2017)

[75] Macklin, M., Müller, M., Chentanez, N.: Xpbd: position-based simulation of compliant constrained dynamics. In Proceedings of the 9th International Conference on Motion in Games. (2016)

[76] Müller, M., Heidelberger, B., Hennix, M., Ratcliff, J.: Position based dynamics. Journal of Visual Communication and Image Representation (2007)

[77] Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. In TIP. (2004)

[78] Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In CVPR. (2018)

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [NA]

Justification: **[TODO]**

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: **[TODO]**

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyper-parameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [Yes]

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [NA]

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [NA] .

    Justification: **[TODO]**

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to

generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: **[TODO]**

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: Our dataset follows CreativeCommons-BY-NC. Please check our supplementary document for the details.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [Yes]

Justification: **[TODO]**

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification: **[TODO]**

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: **[TODO]**

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification: **[TODO]**

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.