WireSeg-32K: A Physics-Grounded Synthetic Dataset for Wire Instance Segmentation Enabled by DeformX

Anonymous CVPR submission

Paper ID 0000

Abstract

Deformable linear objects such as wires and cables are difficult to segment because they are thin, highly deformable, and frequently self-occluded, while large-scale instance-level annotations are expensive to obtain in real scenes. Existing resources either focus on cable tracing or semantic segmentation under constrained settings, or generate visually plausible images without physically grounded wire deformation. We present WireSeg-32k, a synthetic dataset for wire instance segmentation with 32,000 RGB images, per-wire instance masks, depth maps, difficulty splits, and a complementary real-world test set with manual annotations. To generate this dataset, we develop DeformX, a co-simulation pipeline that couples Cosserat-rod wire dynamics with photorealistic Isaac Sim rendering, enabling physically plausible, contact-consistent wire shapes, CAD-based wire assets, and diverse visually grounded scenes. As a simple baseline, LoRA fine-tuning SAM3 on WireSeg-32k alone improves real-world mAP@75 by 10.2% over the off-the-shelf model, showing that physically grounded synthetic data can transfer to real wire perception.

1. INTRODUCTION

Deformable linear objects (DLOs) such as wires, cables, and ropes are common in robotics and industrial environments, including wire harness assembly, cable routing, maintenance, and inspection [5, 7]. Reliable visual perception is a prerequisite for these tasks, yet wire instance segmentation remains difficult: wires are thin, visually repetitive, frequently occluded, and expensive to annotate at scale in cluttered scenes. As a result, large real-world datasets with instance-level wire labels remain scarce, making synthetic data an attractive path toward scalable supervision.

Existing datasets and synthetic pipelines only partially address this need. HANDLOOM [6] provides large-scale synthetic data for cable tracing, but its supervision is trace-based and its setting focuses on grayscale, semi-planar cable images. Fresnillo et al. [3] generate 25k photorealistic cable images that are effective for semantic segmentation, but do not provide per-wire instance masks. Earlier auto-generated wire datasets [8] and perception methods such as FASTDLO [1] and ISCUTE [2] have advanced DLO segmentation, but the data landscape remains fragmented across semantic masks, cable traces, or constrained scene assumptions. For a dataset aimed at modern wire perception, we need instance-level labels, visually grounded scenes, and wire configurations that are not only diverse in appearance but also plausible under gravity, contact, and self-interaction.

In this work, we center the paper on *WireSeg-32k*, a synthetic dataset designed to fill this gap. WireSeg-32k contains 32,000 rendered images from 300+ simulation runs across tabletop, hanging-wire, and data-center scenarios, together with per-wire instance masks, dense depth maps, and easy/medium/hard splits. To generate it, we develop *DeformX*, a co-simulation framework that combines a Cosserat rod engine with Isaac Sim so that dataset generation benefits from both physically grounded DLO dynamics and photorealistic rendering. DeformX is therefore the tool that enables the dataset, rather than the final end product of the paper.

We make three contributions:

- **WireSeg-32k:** A 32k-image synthetic wire instance segmentation dataset with per-wire masks, depth maps, difficulty splits, and a manually annotated real-world test set.
- **DeformX for dataset generation:** A co-simulation pipeline for generating physically plausible and visually realistic wire data using Cosserat-rod dynamics, CAD-compatible assets, and photorealistic rendering.
- **A simple transfer baseline:** A lightweight LoRA fine-tuning recipe on SAM3 trained only on WireSeg-32k that improves real-world mAP@75 by 10.2%, demonstrating sim-to-real value even without real-image training.

2. Method: DeformX

Physics model. To generate realistic wire configurations, DeformX models each DLO as a Cosserat rod, a slender-body

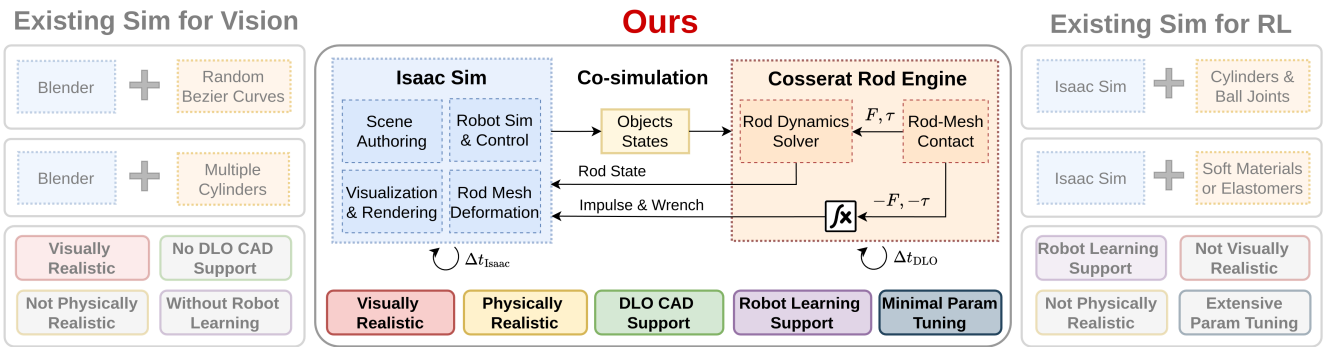

Figure 1. Overview of our co-simulation framework. Compared to existing simulators for vision (left) and RL (right), our framework jointly achieves visual realism, physical accuracy, DLO CAD support, and robot learning support with minimal parameter tuning by combining Isaac Sim and a dedicated Cosserat rod engine.

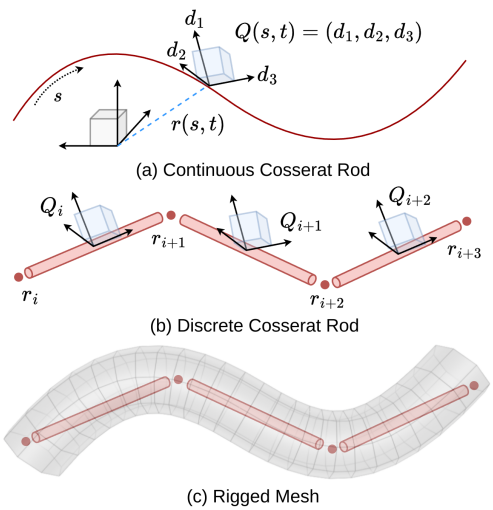

Figure 2. Cosserat rod modeling of DLOs. (a) Continuous Cosserat rod representation; (b) Discrete Cosserat rod representation; (c) Mesh skinned to the discrete Cosserat rod for realistic visualization.

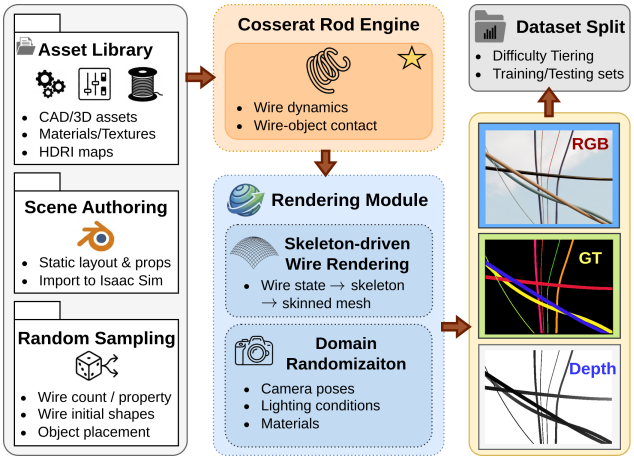

Figure 3. Synthetic wire segmentation dataset generation pipeline. Randomized scenes are built from asset libraries. Wires are simulated with a Cosserat rod engine, and visualized as skinned meshes. RGB images, segmentation masks, and depth maps are generated and split by difficulty levels.

continuum representation that captures stretching, shearing, bending, and twisting [4]. In the continuous view, a rod is parameterized by a centerline $\mathbf{r}(s,t) \in \mathbb{R}^3$ and an orthonormal material frame $\mathbf{Q}(s,t) \in SO(3)$ along arc length s. In practice, we discretize the rod into vertices and segment frames, and integrate its dynamics in a dedicated rod engine. Compared with procedural curves or rigid-link approximations commonly used in synthetic wire generation [3, 8], this gives physically interpretable control over gravity-driven sagging, bending stiffness, torsion, and quasi-static shape transitions. These properties matter for data generation because the visual ambiguity of wire instance segmentation is often created by physically plausible crossings, draping, and self-occlusion rather than by arbitrary curve sampling alone.

Free-form mesh contact. Realistic wire data requires interaction with cluttered geometry. While existing Cosserat rod simulators provide strong mechanics, they do not natively support contact-rich synthetic scenes with arbitrary CAD assets. DeformX therefore adds free-form mesh contact instead of limiting interactions to planes or analytic primitives. Building on a penalty-based formulation, we compute closest-point distances between rod nodes and mesh surfaces, then apply repulsive normal forces with stiffness and damping, while returning reaction impulses to the corresponding rigid objects. To accelerate contact queries, meshes are organized with broad-phase pruning, and a repulsion distance is introduced to reduce deep penetration at larger time steps. This allows wires to drape over structures, slide along supports, wrap around obstacles, and form realistic multi-point contacts common in cable scenes.

Mesh skinning and co-simulation. The rod state is converted into a renderable wire through skeleton-driven mesh skinning: a smooth tubular or CAD-derived mesh is bound to the simulated rod vertices and deformed according to the rod state. This preserves physical consistency while producing visually realistic cable geometry compatible with asset

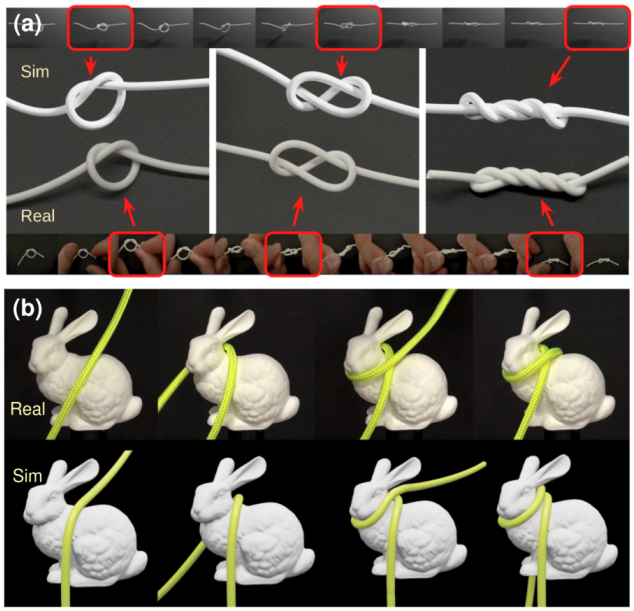

Figure 4. Sim2real comparison for DLOs. (a) We tie a trefoil knot on a wire, apply continuous twists to induce shape transitions, and qualitatively compare the representative configurations between simulation (top) and real experiments (bottom). (b) shows flexible rope (scarf) wrapping around a bunny with multi-point contact, sliding, and self-contact.

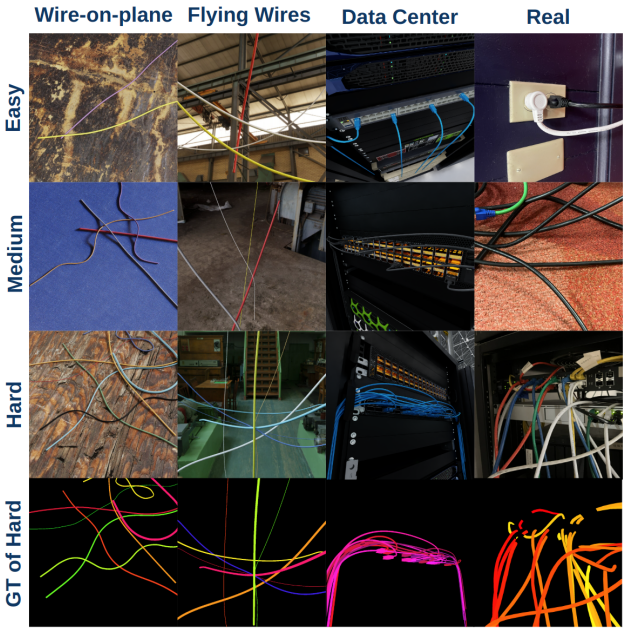

Figure 5. Representative images from three settings across Easy/Medium/Hard, with Hard ground-truth instance masks.

libraries and imported DLO meshes. DeformX couples the rod engine to Isaac Sim through a lightweight multi-rate interface: the rod solver advances the DLO at fine substeps, while Isaac Sim updates rigid bodies, cameras, lighting, and rendering at a coarser rate. This lets users manipulate physically realistic DLOs in photorealistic scenes while relying on a dedicated rod solver underneath. The same interface supports both interactive scene authoring and large-scale headless data generation.

Data-generation pipeline. Figure 3 summarizes the pipeline used to build WireSeg-32K. We first assemble scene templates from publicly available assets and Isaac Sim environments, then sample scene objects, textures, camera viewpoints, and lighting conditions. For each scene instance, we generate wire configurations by varying the number of wires, lengths, radii, initial shapes, and material parameters, and simulate them with DeformX while the rest of the scene is managed by Isaac Sim. The resulting rod states drive skinned meshes for rendering, and because labels remain tied to the simulated DLO identities throughout the pipeline, we can export RGB images, per-wire instance masks, and depth maps consistently even under heavy deformation, occlusion, and clutter. In short, DeformX is not the dataset itself; it is the tool that makes it possible for users to manipulate physically realistic DLOs and systematically synthesize wire data at scale.

Physics sanity checks for data realism. We validates DeformX with two real-world demonstrations that motivate its use for dataset generation. In the first demo, a trefoil knot is twisted and undergoes qualitative shape transitions that match the real wire, showing that the simulator reproduces torsion-bending coupling rather than merely rendering plausible static shapes. In the second demo, a flexible scarf is pulled around a rigid bunny, inducing sliding contact, self-contact, and geometry-conforming draping. As shown in Fig. 4, the simulated configurations remain stable, avoid unrealistic penetration, and qualitatively follow the real trajectories. These demonstrations matter because they directly target the kinds of crossings, wraps, and occlusions that later appear in the synthetic segmentation benchmark.

3. WireSeg-32K Dataset

Overview. WireSeg-32K is designed around three principles: high diversity in wire configurations, a broad spectrum of perceptual difficulty, and scenario categories aligned with realistic deployment settings. Using DeformX, we vary wire count, geometry, material properties, and initial states, then render the resulting scenes in Isaac Sim with randomized viewpoints, lighting, clutter, and textures. Because the labels are derived from simulated DLO identities rather than manual tracing, the pipeline provides consistent per-wire instance masks even for thin, overlapping, and self-occluding structures.

Categories and annotations. The dataset contains three scenario categories: *wire-on-plane*, a canonical tabletop setup; *flying wires*, which emphasizes gravity-driven dangling behavior; and *data center*, which places cables in cluttered rack-like environments. In total, WireSeg-32K contains 32,000

Model	Hard (Syn)		Medium (Syn)		Easy (Syn)		Total (Syn)		Total (Real)	
	F1@75	mAP@75	F1@75	mAP@75	F1@75	mAP@75	F1@75	mAP@75	F1@75	mAP@75
SAM3 (Base)	0.179	0.066	0.446	0.310	0.803	0.735	0.409	0.290	0.296	0.157
SAM3 + LoRA	0.225	0.102	0.512	0.404	0.850	0.816	0.465	0.365	0.314	0.173
Δ (LoRA–Base)	+25.7%	+54.5%	+14.8%	+30.3%	+5.9%	+11.0%	+13.7%	+25.7%	+6.1%	+10.2%

Table 1. SAM3 performance with and without LoRA fine-tuning. For each subset, we report dataset-level F1@75 and COCO mAP@75 (higher is better) on the Easy/Medium/Hard tiers, the full synthetic set, and a held-out real test set. The last row reports the percent improvement from LoRA fine-tuning.

Dataset	Physics Realism	Instance Label	CAD Support	Visually-grounded Scenes	Num. Images
HANDLOOM [6]	×	×	×	×	30k
FASTDLO [1]	×	✓	×	×	32k
Fresnillo et al. [3]	✓	×	×	✓	25k
Zanella et al. [8]	×	×	×	×	28.5k
ISCUTE [2]	×	✓	✓	✓	28k
WireSeg-32k (Ours)	✓	✓	✓	✓	32k

Table 2. Binary comparison between existing DLO datasets and our WireSeg-32k dataset. "Physics realism" indicates whether data generation uses physics-based simulation of DLO deformation (e.g., gravity, wind, and collisions). "Instance label" indicates whether per-object instance masks are provided (trace-only supervision is counted as ×). "CAD support" indicates whether DLOs can be represented as free-form meshes (enabling direct CAD import). "Visually grounded scenes" indicates whether cables are rendered within realistic, image-based backgrounds rather than on plain or purely synthetic backdrops.

rendered images from more than 300 independent simulation runs. Each sample includes RGB, per-wire instance masks, and depth. We also provide a complementary real test set of 300 in-the-wild images with manual instance annotations to evaluate synthetic-to-real transfer.

Difficulty splits and positioning relative to prior work. Following the full paper, we stratify the synthetic set by difficulty using SAM3 performance on generated images: samples with AP@75 < 0.3 are labeled hard, those with AP@75 > 0.6 are labeled easy, and the remainder are labeled medium. This split is useful because difficulty is not determined only by background appearance; it also reflects wire density, crossings, illumination, and thin-object ambiguity. Compared with prior wire datasets and synthetic pipelines [1–3, 6, 8], WireSeg-32K is designed specifically to combine instance-level labels, physically grounded DLO deformation, photorealistic scene context, and CAD-compatible rendering. That combination is what makes it suitable not only for synthetic training, but also for studying sim-to-real transfer on thin deformable objects.

4. Vision Benchmark

To assess the value of WireSeg-32K beyond dataset construction, we next evaluate whether it can provide effective supervision for learning-based wire perception. As a simple benchmark, we test whether training on WireSeg-32K improves wire instance segmentation on both synthetic images and a held-out real-world test set.

We fine-tune Segment Anything Model 3 (SAM3) using lightweight LoRA adapters. Following the full paper, we insert LoRA modules into the attention projections of the vision encoder and mask decoder (`q_proj`, `k_proj`, and `v_proj`), using rank 16, $\alpha = 32$, and dropout 0.05, while keeping all remaining SAM3 parameters frozen for parameter-efficient and stable adaptation. Training is performed for 5 epochs on WireSeg-32K with a learning rate of 1×10^{-5} and an effective batch size of 16.

As a foundation-model baseline, we also evaluate off-the-shelf SAM3 in text-prompt mode using the prompt `cable`. For evaluation, we report dataset-level F1@75, defined as instance F1 with one-to-one matching at IoU ≥ 0.75, together with COCO mAP@75. We compare performance across the easy, medium, and hard synthetic subsets, the full synthetic benchmark, and the held-out real test set.

Table 1 shows that LoRA fine-tuning consistently improves performance over the off-the-shelf baseline across all subsets, including real images. The gains are especially clear on harder scenes with crossings, clutter, and thin-object ambiguities, suggesting that the diversity and physical realism of WireSeg-32K provide useful supervision for challenging wire perception.

5. Conclusion

We presented WireSeg-32K, a physics-grounded synthetic dataset for wire instance segmentation, together with DeformX, the simulation tool used to generate it. By combining Cosserat rod dynamics, free-form mesh contact, lightweight co-simulation, and mesh-skinned rendering, DeformX enables users to manipulate physically realistic DLOs and synthesize labeled data in realistic scenes. Initial experiments with SAM3 show measurable synthetic-to-real transfer, suggesting that WireSeg-32K is a useful benchmark for perception on thin deformable objects.

References

[1] Alessio Caporali, Kevin Galassi, Riccardo Zanella, and Gianluca Palli. FASTDLO: Fast deformable linear objects instance segmentation. *IEEE Robotics and Automation Letters*, 7(4): 9075–9082, 2022. 1, 4

[2] Shir Kozlovsky, Omkar Joglekar, and Diego Di Castro. ISCUTE: Instance segmentation of cables using text embedding. arXiv preprint arXiv:2402.11996, 2024. 1, 4

[3] Pablo Malvido Fresnillo, Wadah M. Mohammed, Sangeetha Vasudevan, Jose Antonio Perez-Garcia, and Jose Luis Martinez-Lastra. Generation of realistic synthetic cable images to train deep learning segmentation models. *Machine Vision and Applications*, 35(4):84, 2024. 1, 2, 4

[4] Noel Naughton, Junang Sun, Aykut Tekinalp, Tejaswin Parthasarathy, Girish Chowdhary, and Mattia Gazzola. Elastica: A compliant mechanics environment for soft robotic control. *IEEE Robotics and Automation Letters*, 6(2):3389–3396, 2021. 2

[5] Jose Sanchez, Juan-Antonio Corrales, Belhassen-Chedli Bouzgarrou, and Youcef Mezouar. Robotic manipulation and sensing of deformable objects in domestic and industrial applications: A survey. *The International Journal of Robotics Research*, 37(7):688–716, 2018. 1

[6] Vainavi Viswanath, Kaushik Shivakumar, Mallika Parulekar, Jainil Ajmera, Justin Kerr, Jeffrey Ichnowski, Roy Cheng, Thomas Kollar, and Ken Goldberg. HANDLOOM: Learned tracing of one-dimensional objects for inspection and manipulation. In *Proceedings of the 7th Conference on Robot Learning (CoRL)*, pages 341–357, 2023. 1, 4

[7] Hao Wang, Omkar Salunkhe, Walter Quadrini, Dan Lämkull, Fredrik Ore, Mélanie Despeisse, Luca Fumagalli, Johan Stahre, and Björn Johansson. A systematic literature review of computer vision applications in robotized wire harness assembly. *Advanced Engineering Informatics*, 62:102596, 2024. 1

[8] Riccardo Zanella, Alessio Caporali, Kalyan Tadaka, Daniele De Gregorio, and Gianluca Palli. Auto-generated wires dataset for semantic segmentation with domain-independence. In *International Conference on Computer, Control and Robotics (ICCCR)*, pages 292–298, 2021. 1, 2, 4