Real-Time Simulation of Deformable Tactile Sensors in Robotic Grasping using Graph Neural Networks

Guillaume Duret[1,3], Danylo Mazurak[1], Frederik Heller[3], Florence Zara[2], Jan Peters[3] and Liming Chen[1]

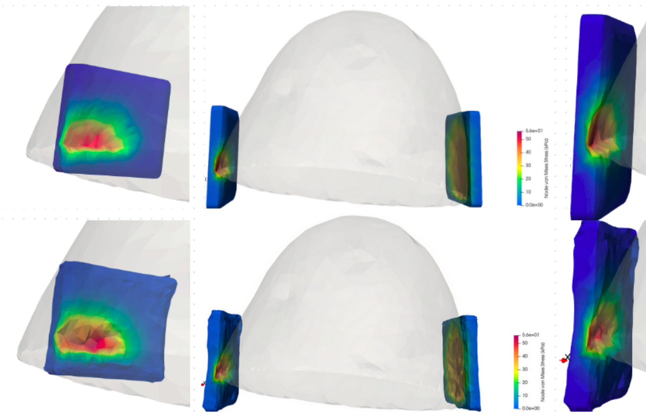

Fig. 1: Illustration of the ground truth stress and deformation of the tactile sensor across different non-grasping poses. The first row shows the ground truth, while the second row presents the prediction from the Graph Neural Network.

I. Introduction

Physical simulation plays a crucial role in the development of robotic manipulation methods, and its importance is even greater when dealing with visual tactile sensors in contact-rich scenarios. Despite years of research, simulating such sensors remains highly challenging—both in terms of the underlying physical dynamics and the rendering of tactile images. In this work, we focus exclusively on the *physical simulation* aspect, leaving the rendering problem outside the scope of our study.

Related work on visual tactile sensor simulation can be broadly divided into two categories: (i) *rigid-body* simulations and (ii) *soft-body* simulations. Soft-body approaches offer higher realism by capturing shear forces and deformations under contact with external objects. However, they are significantly more computationally expensive and orders of magnitude slower than rigid-body simulations. In contrast, rigid-body simulations prioritize execution speed, making them suitable for scenarios requiring large-scale data generation, such as reinforcement learning.

This work addresses the speed limitations of soft-body simulations by leveraging Graph Neural Networks (GNNs), which have been successfully applied to learning the physics of deformable objects [2]. We explore the use of GNN models for simulating grasping interactions with visual tactile sensors, achieving performance gains between 10^2 and 10^4 times faster than traditional FEM simulations, while predicting both deformation and stress on the sensor.

The main contributions of this paper are:

- The first application of GNN-based physics learning to visual tactile sensor simulation.
- A simulation framework capable of accelerating grasping simulations by a factor of 10^2–10^3 compared to FEM, while generalizing to unseen grasping poses.

II. Dataset Generation

For dataset generation2, we employ a Finite Element Method (FEM) simulation within Isaac Gym [4], which has been shown to offer an effective trade-off between simulation speed and quality [5]. While prior work has proposed FEM-based simulations [5] for visual tactile sensors pressed against indenters, these approaches primarily focus on static object–sensor interactions. In contrast, we adopt *Tacgrsp-Sim* [1], which enables the loading of predefined grasping poses for a given object and simulates parallel grasping scenarios involving rigid objects and deformable tactile sensors. The simulation uses a parallel gripper equipped with two visual tactile sensors—specifically, the GelSight Mini, one of the most widely used and commercially available sensors—facilitating reproducibility. It is worth noting that both the dataset generation pipeline and the subsequent GNN framework are compatible with arbitrary tactile sensors, provided that their meshes are available and can be converted into tetrahedral `.tet` files.

The simulation process follows a consistent procedure: (1) load the rigid object, (2) position the gripper at a specified grasping pose, (3) close the gripper until contact is detected, and (4) gradually increase the grasping force applied to the object until a target force N is reached. Data is recorded during phase 4. The simulation terminates once the predefined force threshold N is achieved. During each run, we save 50 simulation frames corresponding to progressively increasing grasping forces. For each frame, the simulator records (i) the node-by-node deformation of the

*This work was in part supported by the French Research Agency, l'Agence Nationale de Recherche (ANR), through the projects Learn Real (ANR-18-CHR3-0002-01), Chiron (ANR-20-IADJ-0001-01), Aristotle (ANR-21-FAI1-0009-01). It was granted access to the HPC resources of IDRIS under the allocation 2023-[AD011013894] and 2024-[AD011015271] made by GENCI.

[1]Centrale Lyon, CNRS, LIRIS, UMR5205, F-69130 Ecully, France, {guillaume.duret, liming.chen}@ec-lyon.fr

[2]UCBL, CNRS, LIRIS, UMR5205, F-69622 Villeurbanne, France, florence.zara@liris.cnrs.fr

[3]Intelligent Autonomous Systems Lab, Technical University of Darmstadt, 64289 Darmstadt, Germany, jan.peters@tu-darmstadt.de

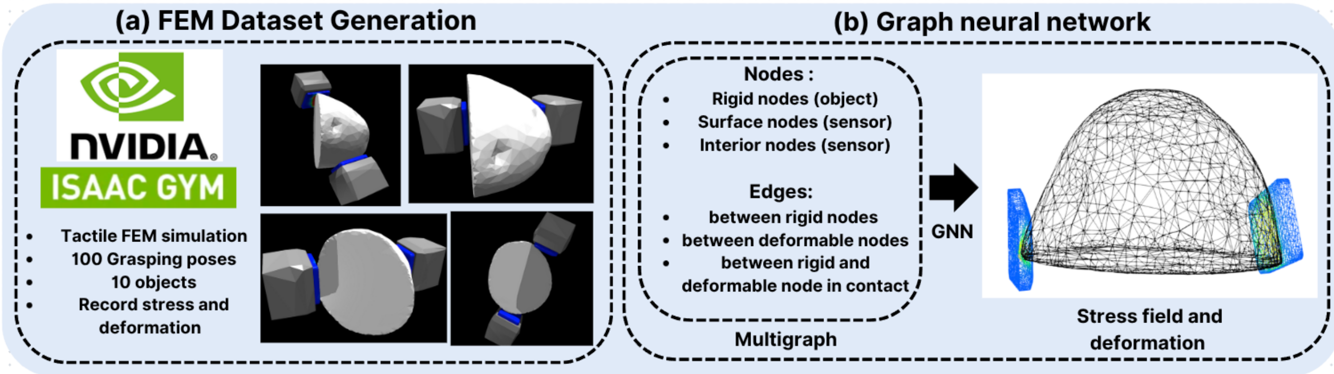

Fig. 2: Pipeline illustrating: (a) dataset creation using FEM simulation [4], [3], [1] across 10 different objects and 100 grasping poses per object, and (b) construction of the Graph Neural Network with node and edge descriptions to predict stress and deformation outputs.

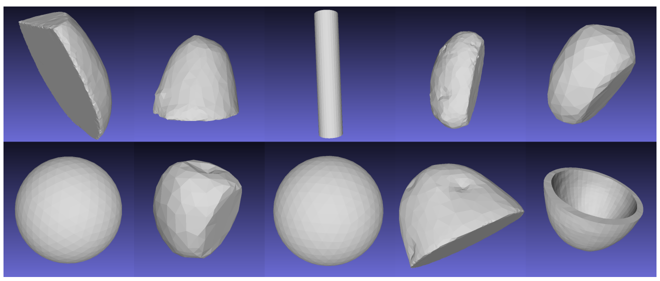

Fig. 3: The 10 objects used in the dataset

soft components, (ii) the rigid body pose, and (iii) the stress distribution within the soft body.

For the FEM parameters, we use pre-calibrated values of Poisson's ratio and Young's modulus for the GelSight Mini [5], in order to minimize the sim-to-real gap in deformation and stress patterns during grasping. The dataset used for GNN training is based on the *DefGraspNet* data [2], containing 10 objects, each with 100 grasping poses.

III. GNN

In this work, we employ a Graph Neural Network (GNN) as the central component for learning the interaction between a deformable gripper and a rigid object. Each node represents either a vertex of the deformable gripper (interior or surface of the tetrahedral mesh) or a vertex of the rigid object, while two relation types are modelled: mesh edges that connect neighbouring nodes within each structure to capture local geometric/mechanical dependencies, and contact edges that connect nearby object–gripper nodes to model contact as it can described in Fig 2. The network follows an Encode–Process–Decode design. Node features combine a compact geometric state, a node-type indicator, and a directional cue for expected finger motion (zero on object nodes, constant per finger on gripper nodes). Edge features are purely geometric on mesh edges (relative offsets and distances), and, on world edges, they additionally carry a scalar force signal associated with the applied grasp. During processing, several rounds of message passing compute edge-based messages from sender/receiver states and edge features and aggregate them at target nodes, allowing the model to approximate the propagation of deformation and stress through the deformable structure and across the contact interface. The decoder then outputs a 3D displacement field for the gripper nodes—added to a fixed reference state—and a non-negative stress field. The rigid object is treated distinctly: its nodes are not deformed and remain aligned with known rigid-body poses, ensuring that learning capacity is devoted to the gripper's non-linear behavior under contact. Training uses supervision from physics-based simulation, minimizing a normalized mean-squared error between predicted and ground-truth node displacements, combined with an error term on the stress field, promoting consistency between predicted kinematics and internal mechanics. Overall, the GNN provides a compact, expressive representation of deformable grasping, enabling accurate prediction of localized deformations and stress distributions that are essential for robust grasp planning and downstream control.

IV. Results and Analysis

To evaluate the proposed model, we conducted a series of experiments progressing from the simplest to the most challenging scenarios.

First, we trained the model on a *single object* (out of a set of eight) to verify its ability to accurately predict deformation and stress for *unseen grasping poses* of the *same object*. This step served as a proof of concept to confirm that the model can generalize to new grasp configurations when the object geometry remains constant.

Next, we performed a large-scale experiment on 10 objects3 to investigate the model's performance in a more challenging, data-limited regime. In this case, the goal was to evaluate the model's ability to predict deformation and stress for completely unseen grasping poses, even when trained on an early-generation dataset of limited size. The training data consisted of 80% of the grasping poses for each object, with the remaining 20% used for testing. Figure 1 shows that the predictions closely follow the ground truth with high accuracy.

V. Limitations and Future Work

This study presents two primary limitations that open avenues for future research. First, the generalization capability of the Graph Neural Network (GNN) remains to be thoroughly evaluated on novel objects and across different tactile sensor configurations. Extending the model to handle a broader range of geometries and sensor modalities would enhance its robustness and applicability in real-world scenarios. Second, the current GNN implementation does not account for gravitational effects, which can lead to inaccuracies in predicting force distribution across the tactile sensor.

VI. Conclusions

This work explores a promising approach using Graph Neural Networks (GNNs) to replace computationally expensive Finite Element Method (FEM) simulations, achieving speed improvements of over 1000×.

References

[1] G. Duret, F. Zara, J. Peters, and L. Chen, "Toward synthetic data generation for robotic tactile manipulations," in *Workshop on "Robot Embodiment through Visuo-Tactile Perception" - 2024 IEEE International Conference on Robotics and Automation (ICRA) Conference Workshop*, Yokohama, Japan, May 2024. [Online]. Available: https://hal.science/hal-04566202

[2] I. Huang, Y. Narang, R. Bajcsy, F. Ramos, T. Hermans, and D. Fox, "Defgraspnets: Grasp planning on 3d fields with graph neural nets," *arXiv preprint arXiv:2303.16138*, 2023.

[3] I. Huang, Y. Narang, C. Eppner, B. Sundaralingam, M. Macklin, R. Bajcsy, T. Hermans, and D. Fox, "Defgraspsim: Physics-based simulation of grasp outcomes for 3d deformable objects," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 6274–6281, 2022.

[4] V. Makoviychuk, L. Wawrzyniak, Y. Guo, M. Lu, K. Storey, M. Macklin, D. Hoeller, N. Rudin, A. Allshire, A. Handa, *et al.*, "Isaac gym: High performance gpu-based physics simulation for robot learning," *arXiv preprint arXiv:2108.10470*, 2021.

[5] C. Zhang, S. Cui, J. Hu, T. Jiang, T. Zhang, R. Wang, and S. Wang, "Tacflex: Multi-mode tactile imprints simulation for visuotactile sensors with coating patterns," *IEEE Transactions on Robotics*, 2025.