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

Guillaume Duret, Frederik Heller*, Danylo Mazurak*,*
Tim Schneider, Alap Kshirsagar, Florence Zara, Jan Peters and Liming Chen

ICRA
IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION

Laboratoire d'InfoRmatique en Image et Systèmes d'information
LIRIS UMR 5205 CNRS / INSA de Lyon / Université Claude Bernard Lyon 1 / Université Lumière Lyon 2 / Ecole Centrale de Lyon

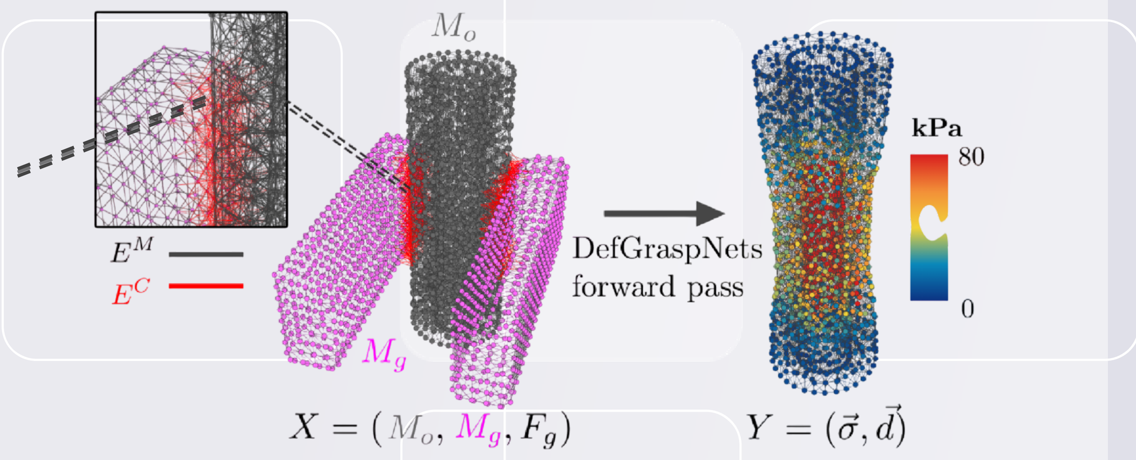
Motivation

- More and more requirement in large scale dataset
- For Soft body, FEM remains the reference but can be slow
- GNN have shown good results in physics simulation
- Growing need of data using Visual Tactile Sensors

Related works and GNN advantages

1) DefGrapNet [1]

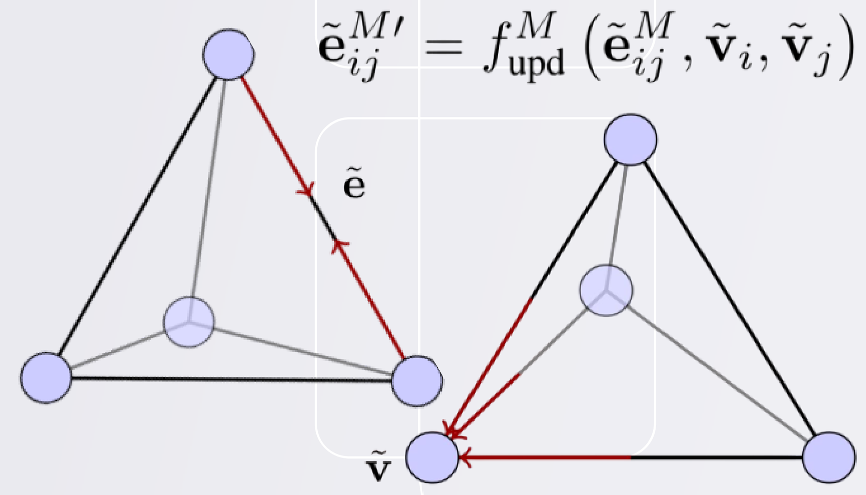
- Dataset generated with FEM simulator DefGraspSim [2]
- Object and gripper before grasp are expressed as input graph
- Mesh vertices are graph nodes, mesh edges graph edges



2) Encode-Process-Decode GNN [3]

- Multi-Layer Perceptrons (MLPs) encode node and edge features to common 128D latent space
- Message passing propagates information:

$$\tilde{\mathbf{e}}_{ij}^{M\prime} = f_{\text{upd}}^{M}\left(\tilde{\mathbf{e}}_{ij}^{M}, \tilde{\mathbf{v}}_i, \tilde{\mathbf{v}}_j\right)$$



$$\tilde{\mathbf{v}}_i' = f_{\text{upd}}^{v}\left(\tilde{\mathbf{v}}_i, \sum_j \tilde{\mathbf{e}}_{ij}^{M}, \sum_j \tilde{\mathbf{e}}_{ij}^{W}\right) + \tilde{\mathbf{v}}_j$$

- Decode deformation and stress prediction at each node

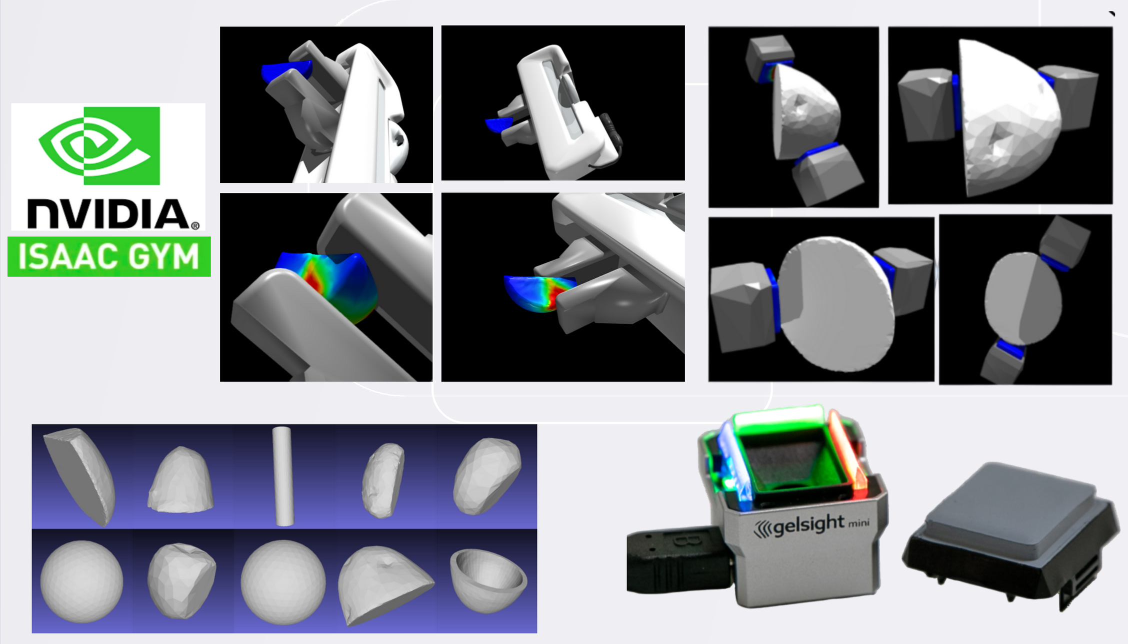
3) Graph Neural Network advantages

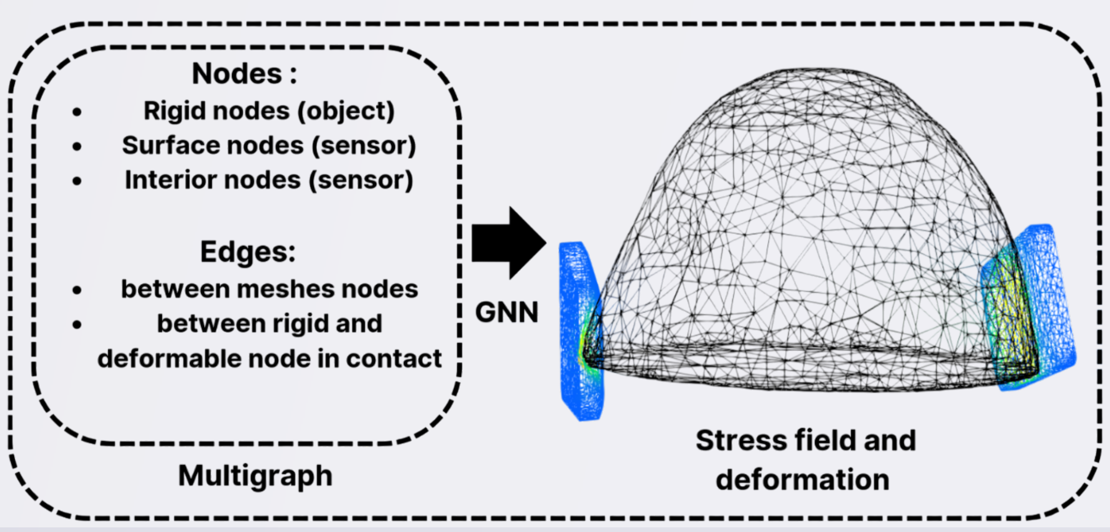
- **Natural representation :** 3D scene with meshes and spatial relationships
- **Local information :** Message passing allows information to propagate locally, mimicking traditional FEM simulation and real continuum mechanics.
- **Inference takes a few ms**
- **End-to-End Differentiability :** Control optimization, parameter estimation and inverse design

Limitation

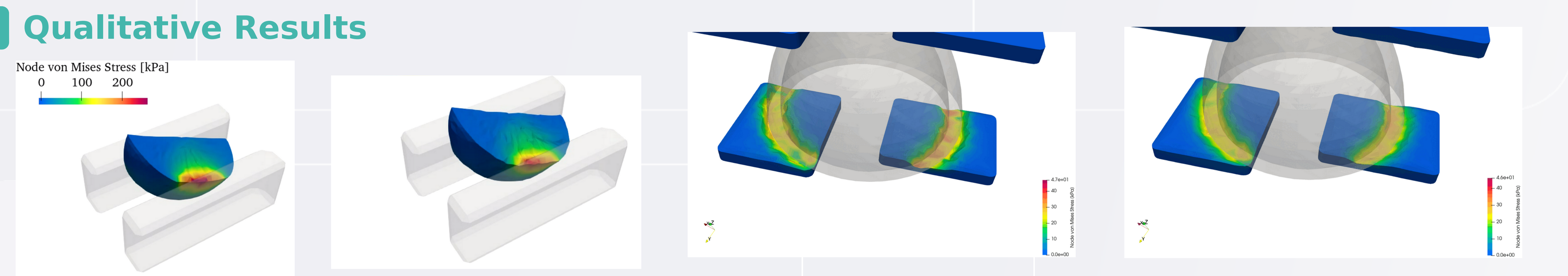
- Slow propagation through network hurts edge cases
- No preprocessing for DefGraspSim data
- No Visual tactile sensors

Soft body Grasping and Visual tactile sensor dataset generation :





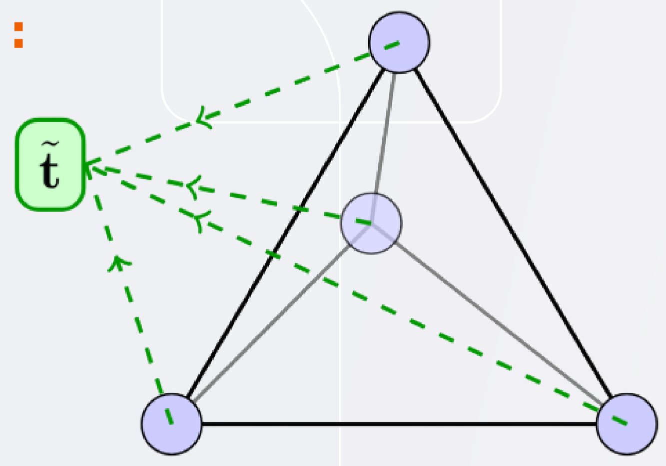
Qualitative Results



Graph neural network Improvements :

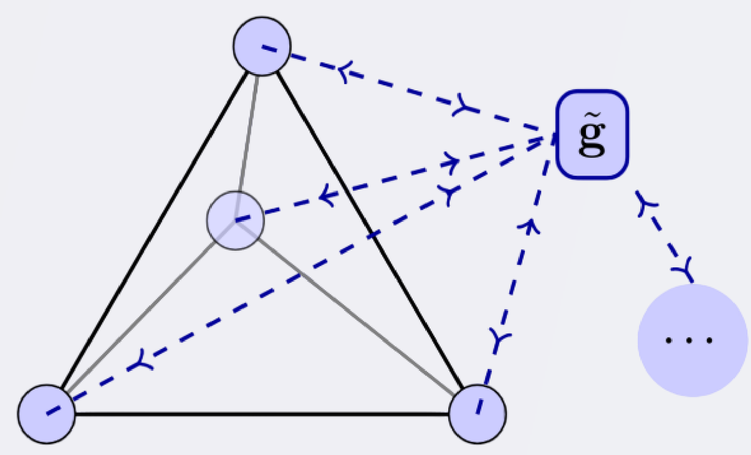
1) Tetrahedron features :

$$\tilde{\mathbf{t}}_i' = f_{\text{upd}}^{t}\left(\tilde{\mathbf{t}}_i, \{\tilde{\mathbf{v}}_j\}_{j\in\mathbb{T}_i}\right) + \tilde{\mathbf{t}}_i$$



- Better physical meaning
- GNN is informed about mesh tetrahedrons by giving a tetrahedrons set as input, similar to the edge set
- Decode stress prediction at tetrahedron

2) Global features :



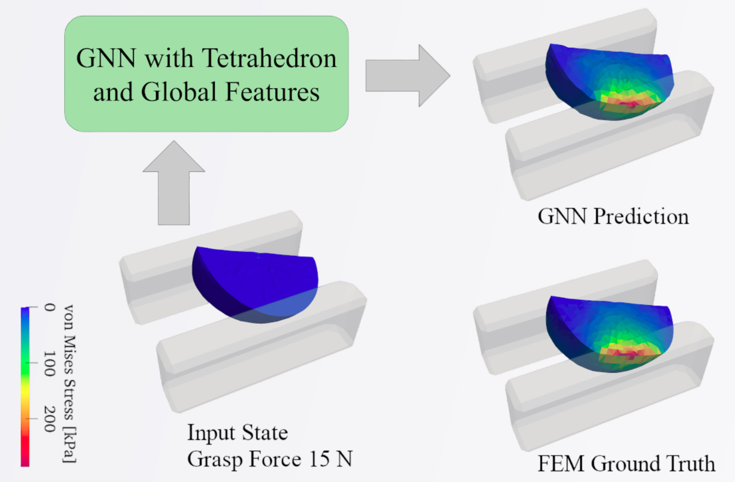
$$\tilde{\mathbf{g}}' = f_{\text{upd}}^{g}\left(\frac{1}{\#\text{nodes}}\sum_{i=1}^{\#\text{nodes}} \tilde{\mathbf{v}}_i\right) + \tilde{\mathbf{g}}$$

$$\tilde{\mathbf{v}}_i' = f_{\text{upd}}^{v}\left(\tilde{\mathbf{v}}_i, \sum_j \tilde{\mathbf{e}}_{ij}, \tilde{\mathbf{g}}\right) + \tilde{\mathbf{v}}_i$$

- Architectural extension to the GNN
- In message passing, global feature receives information from all nodes in graph, and sends to all nodes
- Act as shortcut for globally relevant information to be propagated through the graph

Evaluations on Soft body Grasping

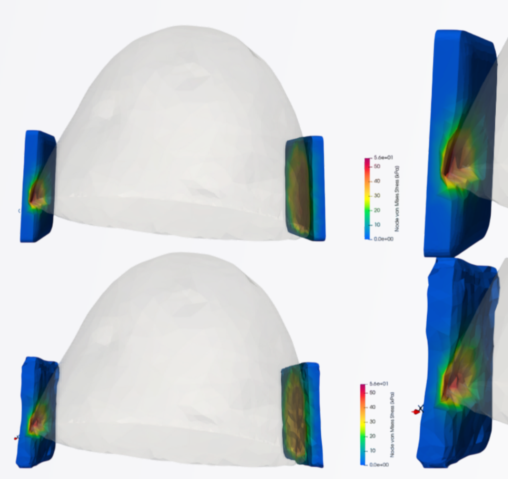
Object MAPE in %	8polygon06 u	8polygon06 σ	cylinder07 u	cylinder07 σ	lemon01 u	lemon01 σ	potato2 u	potato2 σ	sphere03 u	sphere03 σ	strawberry01 u	strawberry01 σ	u	σ
DefGraspNets [1]	3.24	1.52	3.00	3.23	4.66	1.39	**3.05**	2.46	**2.37**	1.37	3.34	**1.12**	3.28	1.85
Tet. Features	3.86	1.91	2.93	2.37	4.51	1.92	3.34	1.75	2.71	**1.18**	4.05	4.57	3.57	2.28
Tet. + Global Feat.	**2.04**	**0.93**	**1.84**	**1.04**	**3.48**	**0.77**	3.12	**0.53**	3.24	1.21	**3.16**	3.40	**2.81**	**1.31**



Evaluations on Visual Tactile Sensors Grasping

Object	Trans.	Def. MAE	Stress MAE
Potato	✓	**6.57e-05**	**372.7**
	–	2.92e-04	382.8
Apple	✓	**7.20e-05**	**370.5**
	–	2.97e-04	427.9
Lemon	✓	**5.40e-05**	**212.1**
	–	2.38e-04	265.6

Object Set	Trans.	Def. MAE	Stress MAE
Known-10	✓	**6.30e-05**	**360.3**
	–	2.69e-04	420.2
Known-63	✓	**7.90e-05**	**352.0**
	–	2.54e-04	365.0
Unknown-12	✓	**8.54e-05**	**6.19e2**
	–	2.81e-04	6.43e2



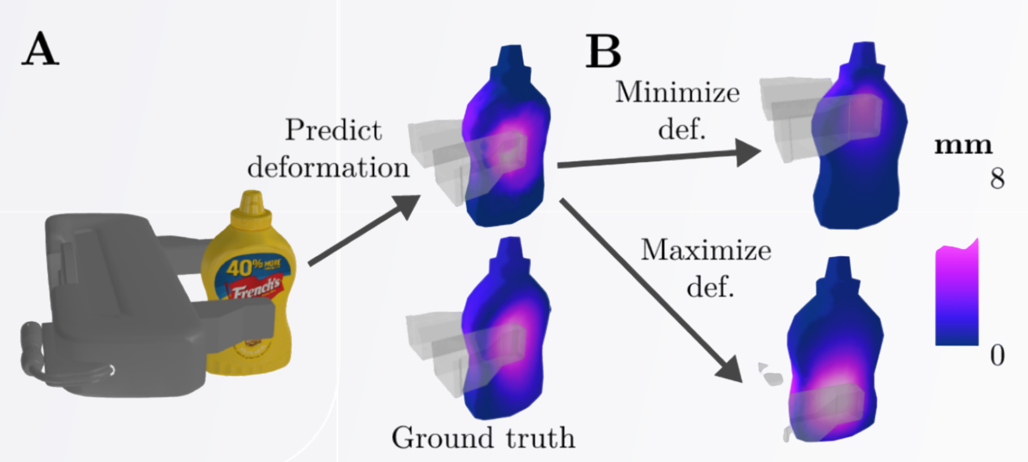
Method	Inference Time	Speedup vs. FEM	Hardware
FEM Simulation	100–200 s	$1\times$	GPU
GNN (Ours)	**5–50 ms**	$\mathbf{10^3}$**–**$\mathbf{10^4}\times$	GPU

Contributions

- Ful PyTorch re-implementation of the DefGraspNets baseline
- Tetrahedron and global features improvement
- Extend DefgraspSim to Visual tactile sensors
- Extensive evaluation of the GNN on visual tactile sensor, including generalization to news objects
- First application of GNN on Visual tactile sensors
- Evaluation of the new features on Soft body grasping

Future works

- Extend the Graph Neural network to new Tactile sensors
- Extend the dataset generation (More objects, soft body parameters, Gravity, Soft-Soft iteractions...)
- Use diffrentaiblity to parameter finding for Control optimization



- Grasping pose optimization through differentiability

References

[1] I. Huang, Y. Narang, R. Bajcsy, F. Ramos, T. Hermans, and D. Fox, "DefGraspNets: Grasp planning on 3D fields with graph neural nets," in 2023 IEEE International Conference on Robotics and Automation (ICRA).IEEE, 2023, pp. 5894–5901.
[2] 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.
[3] T. Pfaff, M. Fortunato, A. Sanchez-Gonzalez, and P. Battaglia, "Learning mesh-based simulation with graph networks," in International conference on learning representations, 2020.

LIRIS cnrs Lyon 1 IAS