EDO-Net: Learning Elastic Properties of Deformable Objects from Graph Dynamics

Alberta Longhini*[1], Marco Moletta*[1], Alfredo Reichlin[1], Michael C. Welle[1], David Held[2], Zackory Erickson[2], and Danica Kragic[1]

***Abstract*— In this work we present a framework for learning graph dynamics of deformable objects that generalizes to unknown physical properties. Our key insight is to extract a latent representation of elastic physical properties of cloth-like deformable objects from observations recorded from a pulling interaction. EDO-Net (*Elastic Deformable Object - Net*), jointly learns an *adaptation* module, and a *forward-dynamics* module. The former is responsible for extracting the latent representation of the physical properties of the object, while the latter leverages the latent representation to predict future states of cloth-like objects represented as graphs. We show both in simulation and in the real world that our proposed method generalizes its dynamics predictions to *unknown* physical properties of deformable objects.**

I. Introduction and Related Work

Manipulation of deformable objects is a fundamental skill towards folding clothes, assistive dressing, wrapping or packaging [1], [2], [3]. In these scenarios, deformables are subject to variations of physical properties such as mass, friction, density, or elasticity, that influence the dynamics of the manipulation [4], [5]. The complexity of the problem arises from the following two factors characterizing deformable objects [6]: *i)* their state is high dimensional and difficult to represent canonically; *ii)* their interaction dynamics are often non-linear and influenced by physical properties usually not known a priori.

To address *i)*, analytical models often employ particle-based representations such as graphs extracted from point clouds [7], [8]. These representations, combined with current advancements in Graph Neural Network (GNN), have shown promising results in learning complex physical systems [9], [10], [11]. However, current methods assume that the physical properties are known a priori, which may not hold when robots operate in human environments. Thus, addressing problem *ii)* is of fundamental importance. The field of *intuitive physics* [12] tackles this challenge by learning predictive models which distill knowledge about the physical properties from past experience and interaction observations [13]. Intuitive physics mainly focused on rigid objects, but new data-driven techniques suggest that interactions with deformable objects (e.g. whipping or pulling) might be relevant for learning their intuitive physics model [6], [14].

*Contributed equally (listed in alphabetical order)

[1]The authors are with the Robotics, Perception and Learning Lab, EECS, at KTH Royal Institute of Technology, Stockholm, Sweden albertal, moletta, alfrei, mwelle, dani@kth.se

[2]The authors are with are with Carnegie Mellon University, Pittsburgh, USA dheld, zerickso@andrew.cmu.edu

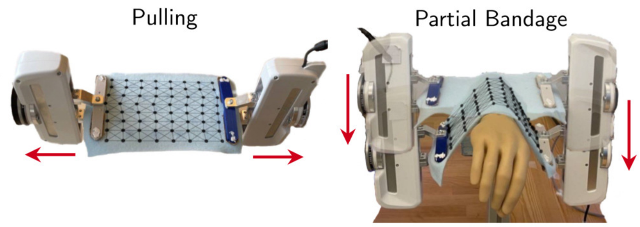

Fig. 1: A pulling interaction is leveraged by EDO-Net to explore the elastic properties of the object, which improves the performance in subsequent tasks such as partial bandage.

In this work, we study the problem of learning graph dynamics of deformable objects that generalize to objects with unknown physical properties. In particular, we focus on elastic properties of cloth-like deformable objects, such as textiles, that we explore through a pulling interaction (Fig. 1). We propose EDO-Net (*Elastic Deformable Object - Net*), a model trained on a large variety of samples with different elastic properties, without relying on ground-truth labels of these properties. EDO-Net jointly learns an *adaptation* module, responsible for extracting a latent representation of the physical properties of the object, and a *forward-dynamics* module, that leverages the latent representation to predict future states, represented as graphs. We evaluate our approach both in simulation and in the real world, showing how EDO-Net accurately predicts the future states of a deformable object with unseen physical properties. In summary, our contributions are:

- EDO-Net, a model to learn graph dynamics of cloth-like deformable objects and a latent representation of their physical properties without explicit supervision;
- a procedure to train EDO-Net on a large variety of samples with different elastic properties, enabling generalization to objects with *unknown* physical properties;

II. Problem Formulation

In our formulation, we refer to the object's elastic properties as $\mathcal{T}_i \sim \mathcal{T}$, where $\mathcal{T}$ is the distribution of all possible physical properties. We explore $\mathcal{T}_i$ by collecting a sequence of observations O^i through an Exploratory Action (EA) [4], [5]. An *adaptation module* is responsible for extracting a latent representation z_i of the physical properties $\mathcal{T}_i$ from the observations O^i, which can be subsequently leveraged by a *forward dynamics module* to generalize its predictions across different $\mathcal{T}_i \sim \mathcal{T}$. We define the state of a deformable object with physical properties $\mathcal{T}_i$ as a graph $G^i = (V^i, E^i)$ with nodes $v \in V^i$ and edges $e \in E^i$. The features of the

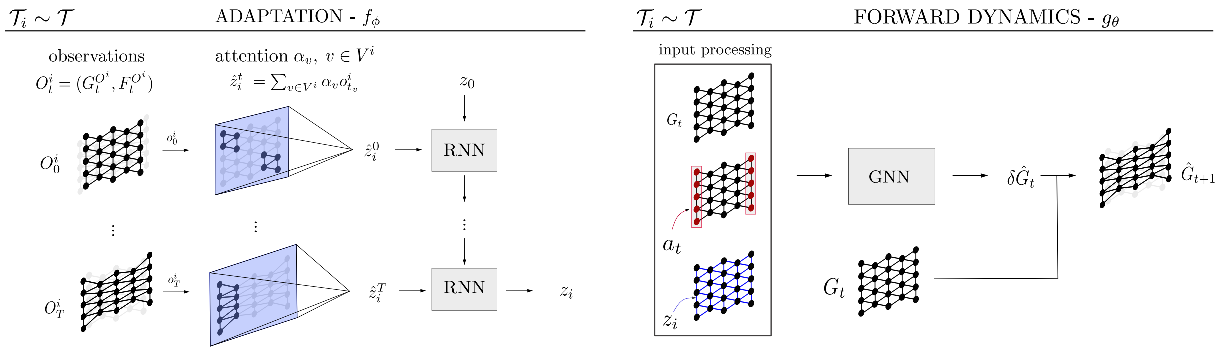

Fig. 2: Scheme of the overall model. Given a deformable object $\mathcal{T}_i$ with *unknown* physical properties, the adaptation module f_ϕ updates the initialization z_0 of the latent representation of the physical properties $\mathcal{T}_i$ from sequences of observations $O_t^i|_{t=1,\ldots,T}$ processed by an attention layer and a RNN. In a second phase, the forward dynamics module g_θ, implemented as a GNN, uses z_i obtained from the adaptation module to predict future states $\hat{G}_t$ of the deformable object.

node v describe the 3D Cartesian position of the nodes, while the features of the edge e characterize the elastic relationship among nodes. Given these, the aim of EDO-Net is to learn a graph dynamics model of cloth-like deformable objects g_θ, wherein instead of predicting the next state, the focus is on learning the delta displacement of the cloth $\delta\hat{G}_t^i$. In particular, the input of the dynamics model consists on a latent representation z_i of the underlying physical properties $\mathcal{T}_i$, the current state G_t^i and the robot control action a_t:

$$\delta\hat{G}_t^i = g_\theta(G_t^i, a_t, z_i). \tag{1}$$

The latent representation z_i can be obtained through a learned function f_ϕ that takes as input a sequence of observations O^i and an initialization z_0 of the representation:

$$z_i = f_\phi(O^i, z_0), \tag{2}$$

where the initialization z_0 is learned together with the model's parameters θ and ϕ. In what follows, we will describe in detail the method to implement and train the graph dynamics g_θ and adaptation f_ϕ functions, respectively.

III. Method

An overview of the proposed EDO-Net is shown in Fig. 2. In particular, for each deformable object with unknown physical properties $\mathcal{T}_i$, the robot has to adapt the initialization z_0 by using a sequence of exploratory observations $O^i = O_t^i\ |_{t=1,\ldots,T}$ where each $O_t^i = (G_t^{O^i}, F_t^{O^i})$ consists of the object state $G_t^{O^i}$, extracted from point clouds as done in [15], and the force $F_t^{O^i}$, recorded from the robot sensors at time t. From O^i, the adaptation module f_ϕ first extracts a latent representation z_i of the physical properties $\mathcal{T}_i$. The implementation of the adaptation function f_ϕ is the following: for each observation O_t^i, we encode $(G_t^{O^i}, F_t^{O^i})$ into a latent embedding o_t^i through a Multi-Layer Perceptron (MLP). We subsequently obtain an estimate $\hat{z}_i^t \in \mathbb{R}^p$ of z_i from o_t^i by learning a node's aggregation function through an attention layer. The extracted representation z_i is subsequently used in the forward dynamics module g_θ to obtain accurate predictions of the future states of $\mathcal{T}_i$ conditioned on different interactions a_t. We model the forward graph dynamics g_θ with a GNN conditioned on the latent representation z_i of the physical properties $\mathcal{T}_i$, where z_i is integrated as features of the edges of the input graph as shown in the *input processing* block in Fig. 2. We train the model g_θ to predict state differences δG_t^i, receiving as input the control action of the robotic manipulator a_t, the initial state of the object G_t^i and the latent representation z_i. We focus on the scenario where the physical properties $\mathcal{T}_i$ are not directly observable from the initial state of the object.

Training Loss: The overall model can be learned using a dataset of exploratory observations $\mathcal{D}^O = \{\mathcal{D}^{O^i}\}_{\mathcal{T}_i \sim \mathcal{T}}$ and a dataset of interactions $\mathcal{D} = \{\mathcal{D}_i\}_{\mathcal{T}_i \sim \mathcal{T}}$. The parameters ϕ, θ and the initialization z_0 can be optimized using a loss on the prediction of the state difference $\delta\hat{G}_t^i$ obtained from g_θ for each training sample with physical properties $\mathcal{T}_i \sim \mathcal{T}$. The loss function $\mathcal{L}$ can be defined as follows:

$$\mathcal{L} = \mathbb{E}_{\substack{\mathcal{T}_i \sim \mathcal{T} \\ G_t^i, a_t, \delta G_t^i \sim \mathcal{D}_i}} \left[d(\delta G_t^i, g_\theta(G_t^i, a_t, z_i)) \right], \tag{3}$$

where $z_i = f_\phi(O^i, z_0)$ with $O^i \sim \mathcal{D}^{O^i}$, and d is the Mean-Squared Error (MSE) between the ground truth state displacement of the deformable object and the model's prediction. Equation 3 optimizes the parameters θ to learn a forward dynamic model conditioned on different representations z_i of physical properties $\mathcal{T}_i$, implicitly driving the parameters ϕ to learn to encode z_i of different samples without supervision from ground truth labels of the physical parameters. Moreover, training across multiple $\mathcal{T}_i \sim \mathcal{T}$ enforces the model to learn how to generalize to deformable objects with unknown physical properties.

IV. Experiments

In this section we evaluate the performance of EDO-Net, regarding its capabilities to generalise dynamic predictions to deformable objects with different elastic properties. To this aim,we analyze *EDO-Net* dynamics predictions both in simulation and real-world environments, testing the model over a set of deformable objects with *unseen* elastic physical properties $\mathcal{T}_i \sim \mathcal{T}$.

Experimental Setup: we carry out the experiments in the real world in the *Partial Bandage* environment shown in Fig. 1, where the robot in the initial phase performs the

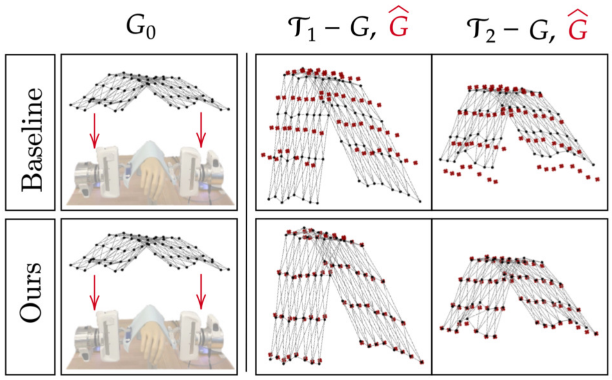

Fig. 3: Qualitative evaluation of the graph dynamics predictions $\hat{G}_t$ obtained by EDO-Net and the NC baseline starting from the inital graph G_0. For each environment we select two elastic samples with different physical properties $\mathcal{T}_1$, $\mathcal{T}_2$.

pulling EA, while in the second phase it pulls the cloth downward over a human arm where the action $a_t \in [0, F_{max}]$ specifies the perceived force at the end-effector that the robot needs to achieve while bandaging the arm. We collect the real-world pulling Exploratory Action (EA) and the interaction trajectories on 40 textile samples with different elastic properties where the dataset characteristics correspond to the one in prior work [16]. For the simulation experiments we use Pybullet [17], [18], in which we replicate the real-world set-up creating two free-floating Franka-Emika Panda end-effectors equipped with Force/Torque sensors and varying the *stiffness* and the *bending* parameters of the simulator. We separately train the models for the simulation and real-world scenarios, leaving sim-to-real evaluation for future work.

Baselines: We compare EDO-Net with a *Non-Conditioned* (NC) baseline model, which trains the forward model g_θ without conditioning on z_i. We also consider an ablation of EDO-Net trained on a single exploratory observation, rather than a sequence of interactions, which we denote by EDO1. Moreover, we include two oracle models in simulation to set an upper-bound performance for the tasks: one *Oracle Forward* model conditioned on the ground-truth simulation parameters (OF), and an *Oracle Supervised* forward model (OS), trained with an additional supervised loss term over z_i, to directly predict the ground-truth simulation parameters during the training procedure.

A. Generalization to Unseen Physical Properties

In this section, we evaluate the generalisation capabilities of EDO-Net. We consider both simulation and real-world environments, and we perform quantitative and qualitative tests of the model over a set of deformable objects with elastic physical properties $\mathcal{T}_i \sim \mathcal{T}$ unseen during training. We evaluated the model's performance by computing the MSE of the model's predictions with respect to the ground truth for each testing sample. We compare the performance of our model with respect to the NC and the EDO1 baselines. In simulation we also compare to OS and OF. In Table I we report the mean and standard deviation of the MSEs evaluated across all the testing samples with physical properties $\mathcal{T}_i \sim \mathcal{T}$. In all scenarios, *EDO-Net* outperforms the baseline models both in terms of the average error and the standard deviation across samples with different elastic properties. The high standard deviation of the NC model is due to the large difference between the average elastic behavior and the extreme (rigid/elastic) ones. Moreover, *EDO-Net* achieves comparable performances with respect to OF. Qualitative visualizations of the relevance of our proposed method are shown in Fig. 3. We can observe how the NC baseline does not distinguish among samples with different physical properties ($\mathcal{T}_1$ and $\mathcal{T}_2$), hindering its capability of predicting the outcome of the robot control actions. On the other hand, *EDO-Net* successfully leverages the latent representations (z_1 and z_2) provided by the adaptation module f_ϕ.

TABLE I: Generalisation results of EDO-Net and the baselines in the simulated and real-world environments (in normalized units), with T=5. Lower is better.

Model	MSE ($\times 10^{-3}$) *Partial Bandage* simulation	MSE ($\times 10^{-3}$) *Partial Bandage* real world
NC	29.60 ± 65.29	59.37 ± 57.50
EDO1	0.260 ± 0.197	3.046 ± 1.603
EDO-Net	0.151 ± 0.125	1.481 ± 0.500
OS	0.992 ± 1.480	–
OF	0.122 ± 0.194	–

V. Discussion and Conclusions

We presented EDO-Net, a data-driven model that learns a latent representation of physical properties of cloth-like deformable objects to generalize graph-dynamic predictions to objects with unseen physical properties. We assessed both in simulation and real world how conditioning the forward dynamics model to the latent representation z_i helps in generalizing over unseen physical properties. It can be shown that it is possible to decode the ground truth physical properties $\mathcal{T}_i$ of the deformable object from the latent representation z_i with a weak learner, suggesting that the loss in Eq. 3 implicitly trains the model to learn a latent representation of the physical properties without explicit supervision from the ground truth labels [15]. Moreover, the latent representation z_i can be transferred to different environments where elasticity matters (e.g. lifting an elastic bag), or to different downstream tasks such as learning an inverse model to predict the control action between two states [15].

A substantial part of the research in manipulating cloth-like deformable objects focuses on robotic tasks like cloth folding [19], [1], [20], cloth smoothing [21], [8], [22], assisted dressing [2], [23], [24], and bedding manipulation [25], [26]. These tasks require considering physical properties such as elasticity, stiffness, and mass, as they significantly impact the robot's manipulation strategy. However, existing methods fail to account for variations in these properties. By leveraging the latent representation of EDO-Net, these methods could adapt their manipulation strategies to different elastic properties, enhancing generalization and performance. Exploring this direction is part of our future work.

REFERENCES

[1] Y. Avigal, L. Berscheid, T. Asfour, T. Kröger, and K. Goldberg, "Speedfolding: Learning efficient bimanual folding of garments," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 1–8.

[2] Z. Erickson, H. M. Clever, G. Turk, C. K. Liu, and C. C. Kemp, "Deep haptic model predictive control for robot-assisted dressing," in *2018 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2018, pp. 4437–4444.

[3] D. Seita, P. Florence, J. Tompson, E. Coumans, V. Sindhwani, K. Goldberg, and A. Zeng, "Learning to rearrange deformable cables, fabrics, and bags with goal-conditioned transporter networks," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 4568–4575.

[4] Z. Xu, J. Wu, A. Zeng, J. B. Tenenbaum, and S. Song, "Densephysnet: Learning dense physical object representations via multi-step dynamic interactions," *arXiv preprint arXiv:1906.03853*, 2019.

[5] C. Wang, S. Wang, B. Romero, F. Veiga, and E. Adelson, "Swingbot: Learning physical features from in-hand tactile exploration for dynamic swing-up manipulation," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 5633–5640.

[6] C. Chi, B. Burchfiel, E. Cousineau, S. Feng, and S. Song, "Iterative residual policy: for goal-conditioned dynamic manipulation of deformable objects," *arXiv preprint arXiv:2203.00663*, 2022.

[7] Z. Weng, F. Paus, A. Varava, H. Yin, T. Asfour, and D. Kragic, "Graph-based task-specific prediction models for interactions between deformable and rigid objects," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 5741–5748.

[8] X. Lin, Y. Wang, Z. Huang, and D. Held, "Learning visible connectivity dynamics for cloth smoothing," in *Conference on Robot Learning*. PMLR, 2022, pp. 256–266.

[9] A. Sanchez-Gonzalez, J. Godwin, T. Pfaff, R. Ying, J. Leskovec, and P. Battaglia, "Learning to simulate complex physics with graph networks," in *International Conference on Machine Learning*. PMLR, 2020, pp. 8459–8468.

[10] T. Pfaff, M. Fortunato, A. Sanchez-Gonzalez, and P. W. Battaglia, "Learning mesh-based simulation with graph networks," *arXiv preprint arXiv:2010.03409*, 2020.

[11] A. Sanchez-Gonzalez, N. Heess, J. T. Springenberg, J. Merel, M. Riedmiller, R. Hadsell, and P. Battaglia, "Graph networks as learnable physics engines for inference and control," in *International Conference on Machine Learning*. PMLR, 2018, pp. 4470–4479.

[12] M. McCloskey, "Intuitive physics," *Scientific american*, vol. 248, no. 4, pp. 122–131, 1983.

[13] J. Wu, I. Yildirim, J. J. Lim, B. Freeman, and J. Tenenbaum, "Galileo: Perceiving physical object properties by integrating a physics engine with deep learning," *Advances in neural information processing systems*, vol. 28, 2015.

[14] A. Longhini, M. C. Welle, I. Mitsioni, and D. Kragic, "Textile taxonomy and classification using pulling and twisting," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 7564–7571.

[15] A. Longhini, M. Moletta, A. Reichlin, M. C. Welle, D. Held, Z. Erickson, and D. Kragic, "Edo-net: Learning elastic properties of deformable objects from graph dynamics," *arXiv preprint arXiv:2209.08996*, 2022.

[16] A. Longhini, M. Moletta, A. Reichlin, M. C. Welle, A. Kravberg, Y. Wang, D. Held, Z. Erickson, and D. Kragic, "Elastic context: Encoding elasticity for data-driven models of textiles," 2022. [Online]. Available: https://arxiv.org/abs/2209.05428

[17] E. Coumans and Y. Bai, "Pybullet, a python module for physics simulation for games, robotics and machine learning," http://pybullet.org, 2016–2021.

[18] Z. Erickson, V. Gangaram, A. Kapusta, C. K. Liu, and C. C. Kemp, "Assistive gym: A physics simulation framework for assistive robotics," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 10 169–10 176.

[19] T. Weng, S. M. Bajracharya, Y. Wang, K. Agrawal, and D. Held, "Fabricflownet: Bimanual cloth manipulation with a flow-based policy," in *Conference on Robot Learning*. PMLR, 2022, pp. 192–202.

[20] G. Salhotra, I.-C. A. Liu, M. Dominguez-Kuhne, and G. S. Sukhatme, "Learning deformable object manipulation from expert demonstrations," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 8775–8782, 2022.

[21] R. Hoque, D. Seita, A. Balakrishna, A. Ganapathi, A. K. Tanwani, N. Jamali, K. Yamane, S. Iba, and K. Goldberg, "Visuospatial foresight for physical sequential fabric manipulation," *Autonomous Robots*, vol. 46, no. 1, pp. 175–199, 2022.

[22] H. Ha and S. Song, "Flingbot: The unreasonable effectiveness of dynamic manipulation for cloth unfolding," in *Conference on Robot Learning*. PMLR, 2022, pp. 24–33.

[23] A. Kapusta, Z. Erickson, H. M. Clever, W. Yu, C. K. Liu, G. Turk, and C. C. Kemp, "Personalized collaborative plans for robot-assisted dressing via optimization and simulation," *Autonomous Robots*, vol. 43, pp. 2183–2207, 2019.

[24] Z. Erickson, H. M. Clever, V. Gangaram, G. Turk, C. K. Liu, and C. C. Kemp, "Multidimensional capacitive sensing for robot-assisted dressing and bathing," in *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*. IEEE, 2019, pp. 224–231.

[25] K. Puthuveetil, C. C. Kemp, and Z. Erickson, "Bodies uncovered: Learning to manipulate real blankets around people via physics simulations," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1984–1991, 2022.

[26] D. Seita, N. Jamali, M. Laskey, A. K. Tanwani, R. Berenstein, P. Baskaran, S. Iba, J. Canny, and K. Goldberg, "Deep transfer learning of pick points on fabric for robot bed-making," in *Robotics Research: The 19th International Symposium ISRR*. Springer, 2022, pp. 275–290.