

Robotic perception and manipulation of deformable linear objects: A survey

The International Journal of
Robotics Research
2026, Vol. 0(0) 1–44
© The Author(s) 2026



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/02783649261432253
journals.sagepub.com/home/ijr



Alessio Caporali¹ , Ignacio Cuiral-Zueco² ,
Gonzalo López-Nicolás³  and Gianluca Palli¹ 

Abstract

Deformable linear objects (DLOs) are widely encountered in everyday life, taking forms such as plastic tubes, wires, ropes, and cables. They are prevalent across diverse settings, including industrial, domestic, and medical environments, as well as in outdoor applications like electric power lines, subaquatic cables, and aerial transport systems. These objects are termed *deformable* due to their ability to undergo significant shape changes under external forces, and *linear* because their length vastly exceeds their cross-sectional dimensions. Despite their importance and widespread presence, developing robotic systems capable of interacting with DLOs poses numerous challenges. This survey presents a comprehensive review of the state-of-the-art methods developed over the past decade to address these challenges. It covers key areas including physical and data-driven modeling techniques, simulation environments, perception approaches based on vision and tactile sensing, as well as strategies for estimation, planning, and control. It also reviews common manipulation tasks such as grasping, shaping, routing, knotting, suturing, and transport. The survey concludes with a critical discussion of current limitations and outlines promising directions for future research.

Keywords

autonomy for mobility and manipulation, computer vision for automation, deep learning in grasping and manipulation, manipulation and grasping, manipulation planning, perception for grasping and manipulation, RGB-D perception, visual perception and learning

1. Introduction

Deformable Linear Objects (DLOs) are a class of elongated Deformable Objects (DOs) that include items such as cables, ropes, and tubes. The term *deformable* emphasizes their ability to undergo significant shape changes in response to external forces, while *linear* reflects their geometry, that is, their length greatly exceeds their cross-sectional dimensions. Importantly, DLOs exhibit complex, highly nonlinear behavior, making their modeling and manipulation particularly challenging. In the literature, DLOs are often classified as *uniparametric* DOs (Sanchez et al., 2018).

DLOs play an essential role in a wide range of practical applications across multiple domains. They are commonly encountered in domestic environments, where they appear as cables, ropes, and wires. In industrial sectors such as automotive (Jiang et al., 2011; Trommnau et al., 2019) and aerospace (Shah et al., 2018), DLOs are present not only as individual electrical cables and wires but also as complex branched structures composed of multiple interconnected elements, such as wiring harnesses and hose bundles. These are often referred to in the literature as Deformable Multi-Linear Objects (DMLOs) or Branched Deformable Linear

Objects (BDLOs) (Caporali et al., 2025; Zürm et al., 2023b). In the healthcare domain, DLOs also appear in the form of surgical materials such as suture threads (Lu et al., 2022). Despite their widespread presence, automating processes involving DLOs remains a significant challenge (Trommnau et al., 2019), largely due to the limited availability of effective robotic solutions for accurately perceiving and manipulating these highly flexible and deformable objects.

¹Department of Electrical, Electronic and Information Engineering, University of Bologna, Bologna, Italy

²Department of Image, Perception Systems, and Robotics (ISPR), Université Clermont Auvergne, CNRS, Clermont Auvergne INP, Institut Pascal, Clermont-Ferrand, France

³Department of Computer Science and Systems Engineering, Instituto de Investigación en Ingeniería de Aragón, Universidad de Zaragoza, Zaragoza, Spain

Corresponding author:

Alessio Caporali, Department of Electrical, Electronic and Information Engineering, University of Bologna, Viale Risorgimento 2, Bologna 40136, Italy.

Email: alessio.caporali@unibo.it

Throughout this survey, the term DLOs is used broadly to encompass common objects such as ropes, alongside more specialized types like DMLOs and suture threads. Specific distinctions are made only when necessary to highlight key differences relevant to the discussion.

Although interest in DLOs has been growing, the literature still lacks a dedicated and comprehensive survey addressing their unique challenges. Existing reviews primarily focus on the broader category of DOs, often emphasizing planar or volumetric DOs (Arriola-Rios et al., 2020; Hou et al., 2019; Jiménez, 2012; Sanchez et al., 2018; Yin et al., 2021; Zhu et al., 2022). Among these, only Jiménez (2012) and Sanchez et al. (2018) explicitly address DLOs in detail. The former offers a limited discussion, focusing solely on model-based planning strategies. The latter presents a classification of DOs based on physical and geometric aspects, and includes DLO-specific challenges in modeling, perception, and manipulation. However, the coverage remains limited and outdated, lacking recent advancements in the rapidly evolving field of DLO research.

This review provides the reader with a well-structured overview of the current literature and state-of-the-art approaches to the modeling, perception, and manipulation of DLOs, offering valuable insights to both newcomers and experienced researchers in the field.

The literature search for this work followed the guidelines of the preferred reporting items for systematic reviews and meta-analyses (PRISMA) approach (Page et al., 2021), encompassing identification, screening, eligibility, and inclusion stages. In the identification stage, a comprehensive search was conducted across electronically indexed databases, followed by manual searches of indexed conference and journal papers, as well as the bibliographies of identified articles, to ensure thorough coverage and mitigate biases from automatic-only searches. This survey included, reviewed, and classified more than 260 articles.

The remainder of this survey is structured as follows. Section 2 discusses modeling aspects of DLOs, followed by perception methods in Section 3. Challenges related to estimation, planning, and control are discussed in Section 4. Key manipulation tasks, including shaping, routing, and unknotting, are examined in Section 5. A discussion of current limitations and promising directions for future research is outlined in Section 6. Finally, Section 7 concludes the survey. An overview of the survey's structure is illustrated in Figure 1, providing a visual map of the main topics and their connections to help guide the reader through the subsequent sections.

2. Modeling

In this survey, the term *model* refers to both classical physical formulations and other representations of DLO behavior, including geometric and data-driven approaches. This section begins by introducing the physical formulation for DLO modeling based on Cosserat rod theory, which will

serve as reference and baseline for characterizing and comparing other widely used models in the literature in this survey's proposed DLO model classification. For a broader overview of slender-object modeling techniques, readers are referred to the review by Lv et al. (2020) whose scope is primarily oriented toward non-robotic contexts, in contrast to the robotics-centered perspective of this survey.

2.1. DLO baseline model: Cosserat rod theory

DLOs are continuous mechanical systems, for which the Cosserat rod theory (Antman, 1972) offers a particularly complete physical formulation. It models both the 3D position and orientation of the DLO's centerline through a director field, and accounts for bending, torsion, shear, and axial deformations via constitutive relations.

2.1.1. Cosserat rod formulation. In Cosserat rod theory, DLOs are modeled as continuous bodies through a curve $\mathbf{x}(s, t) \in \mathbb{R}^3$ parameterized by arc-length $s \in [0, L]$ and time t . The dynamic equilibrium equations governing the motion and deformation of the DLO are derived from the balance of forces and moments along the rod, resulting, respectively, in differential equations:

$$\rho \frac{\partial^2 \mathbf{x}}{\partial t^2} = \frac{\partial \mathbf{n}}{\partial s} + \mathbf{f}_d(s, t) + \sum_i \mathbf{f}_p^{(i)}(t), \quad \text{and} \quad (1)$$

$$\mathbf{0} = \frac{\partial \mathbf{m}}{\partial s} + \mathbf{t} \times \mathbf{n} + \mathbf{c}_d(s, t) + \sum_i \mathbf{c}_p^{(i)}(t), \quad (2)$$

where ρ is the linear mass density, \mathbf{n} denotes the internal restoring forces, and \mathbf{m} refers to the internal moments (couples). Forces $\mathbf{f}_d(s, t)$ and $\mathbf{f}_p^{(i)}(t)$ correspond to distributed and point external forces, respectively, while $\mathbf{c}_d(s, t)$ and $\mathbf{c}_p^{(i)}(t)$ denote distributed and point external moments. The tangent vector along the centerline is defined as $\mathbf{t} = \partial \mathbf{x} / \partial s$.

The differential equations (1) and (2) incorporate *constitutive laws* that describe how internal forces and moments arise from deformations. That is, these relations connect strains to stresses based on the object's physical properties, describing how bending, twisting, stretching, and shearing generate internal forces and moments depending on the rod's material properties and geometric characteristics. For slender, isotropic, and linearly elastic rods (common assumptions for DLOs), the *constitutive laws* are typically expressed as:

$$\mathbf{n} = \mathbf{K}_n(\boldsymbol{\nu} - \boldsymbol{\nu}_0), \quad \mathbf{m} = \mathbf{K}_m(\boldsymbol{\mu} - \boldsymbol{\mu}_0), \quad (3)$$

where $\boldsymbol{\nu}$ and $\boldsymbol{\mu}$ denote the strain fields along the rod, including axial, shear, curvature, and twist strains, while $\boldsymbol{\nu}_0$ and $\boldsymbol{\mu}_0$ represent the corresponding reference (undeformed) strain states. The stiffness matrices are defined as:

$$\mathbf{K}_n = \text{diag}(GA, GA, EA), \quad \mathbf{K}_m = \text{diag}(EI_1, EI_2, GJ), \quad (4)$$

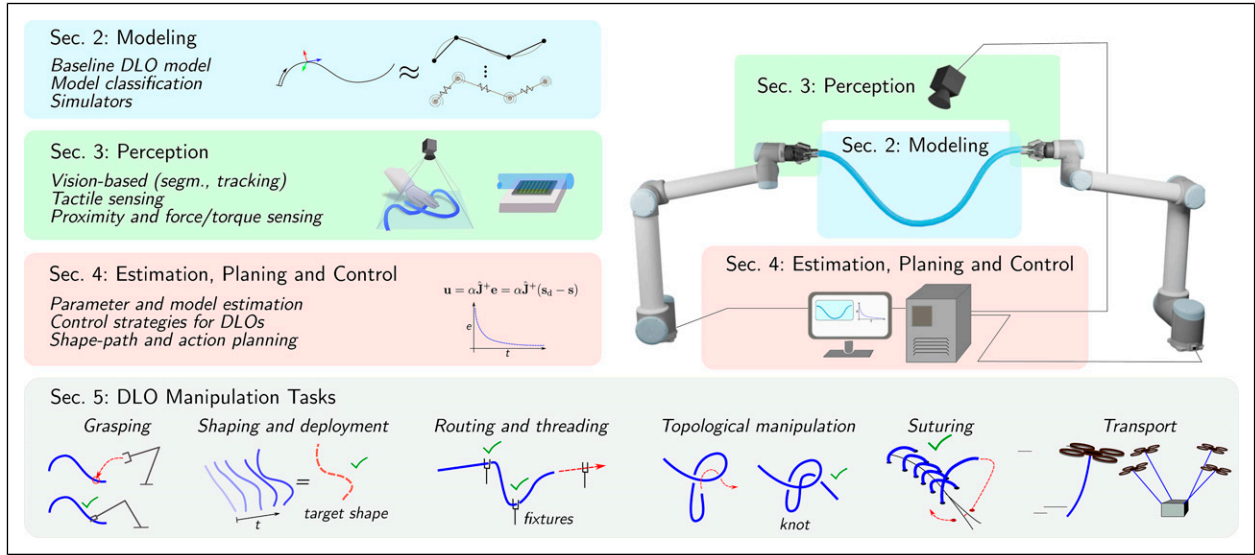


Figure 1. General overview of the survey’s contents. Section 2 reviews the current literature on DLO modeling, beginning with an introduction to the Cosserat physical model as a baseline, followed by an analysis of existing modeling approaches and a summary of recent trends in DLO simulators. Section 3 covers various perception methods for DLOs, including vision-based tasks such as segmentation and tracking, tactile sensing, and additional modalities like proximity and force/torque sensing. Section 4 examines DLO-specific techniques for parameter and model estimation, as well as control and planning strategies, providing a comprehensive classification. These three sections serve as background for a task-oriented analysis presented in Section 5, which offers an overview of DLO manipulation methods for tasks including shape control, routing, and transport.

with E being the Young’s modulus, G the shear modulus, A the cross-sectional area, I_1 and I_2 the second moments of area, and J the polar moment of inertia. Typically, \mathbf{K}_n and \mathbf{K}_m in (3) are assumed to be constant along the DLO domain. For cases involving variable stiffness along $\mathbf{x}(s, t)$, that is, $\mathbf{K}_n(s)$ and $\mathbf{K}_m(s)$, the reader is referred to a more general formulation of Tummers et al. (2023). Another clear and concise discussion of the Cosserat model, particularly focused on the kinematic aspects, is provided in Rucker and Webster (2011).

2.1.2. Boundary conditions. Solving differential equations (1) and (2) requires boundary conditions (BCs) that *establish* the model within a specific manipulation scenario, thereby defining the physical context for the model. In the context of DLOs, boundary conditions typically fall into two main categories: *actuation*, which describes how the object is controlled or manipulated, and *environmental* constraints, representing interactions with external elements such as surfaces, fixtures, or obstacles.

2.1.2.1. Actuation. Refers to boundary conditions that can be actively modified during manipulation. Typical setups include single-end grasping (one grasped end, the other free), dual-end grasping (both ends are grasped, i.e., clamped DLO), and variable contact points (e.g., a DLO pushed on a table at different points). These impose specific pose and force constraints and determine the numerical strategy for solving (1, 2). In single-end cases, the problem is usually solved via shooting methods, while dual-end cases

are formulated as two-point boundary value problems (BVPs) and can be solved, for example, through spectral collocation. Specialized DLO models with BCs arising from continuous actuation along the domain, such as tendon-driven actuation, are also common in related fields like soft robotics (e.g., Tummers et al., 2023).

2.1.2.2. Environmental constraints. Refer to external factors that affect the DLO but cannot be actively modified during DLO manipulation. These typically include factors such as distributed forces like gravity, friction and contacts with surfaces (Jilani et al., 2025), fixtures, and obstacles in the environment.

2.2. Classification of DLO models

The continuous formulation of the Cosserat rod model, as outlined in Section 2.1, provides a strong physical foundation. However, in the practical context of DLO manipulation, it presents several challenges. Solving equations (1) and (2) can be computationally intensive, particularly when simulating complex interactions or real-time scenarios. Additionally, the accuracy of Cosserat’s model is highly dependent on precise material properties and detailed geometric data of the DLO (3), such as length or cross-sectional areas and shapes. These limitations have led to the use of alternative models for DLO manipulation, which are classified in this survey based on criteria designed to enhance both clarity and practicality for the reader. As

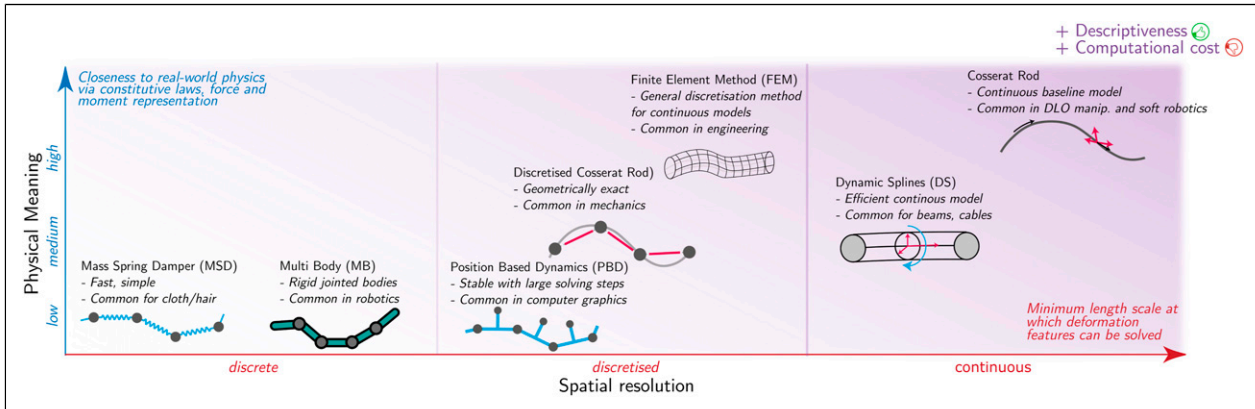


Figure 2. Comparison of different modeling approaches for DLOs, evaluated according to their physical fidelity and spatial resolution. This overview highlights key trade-offs between model accuracy and descriptive power versus computational efficiency, guiding the selection of appropriate models for different robotic applications.

illustrated in Figure 2, different modeling approaches for DLOs vary significantly in terms of physical realism and computational complexity.

2.2.1. Spatial resolution. It refers to the level of granularity with which a model represents the geometry and deformation of a DLO along its length. Therefore, spatial resolution affects both model accuracy and computational cost. Models can be categorized into:

- **Continuous:** They describe the deformation of DLOs over a continuous spatial domain. Continuous physical models offer high accuracy but typically imply high computational cost. Examples include the already discussed Cosserat rod formulation (Antman, 1972), along with particularizations of the Cosserat rod like the Kirchhoff rod (Bretl and McCarthy, 2014) for inextensible, non-shearable rods, or the Timoshenko rod, which accounts for shear in thick inextensible beams. Among these and other classic beam and rod models, other continuous representations for DLO description and modeling include the use of parametric curves such as dynamic splines (DS) (Palli, 2020; Theetten et al., 2008; Valentini and Pennestrì, 2011), Fourier series (Zhu et al., 2018), or curvature-based analysis (Cuiral-Zueco et al., 2023).
- **Discretized:** These models approximate continuous deformation by discretizing the DLO's spatial domain, maintaining correspondence with continuous models either through inherent equivalence or by converging to the continuous solution as the discretization becomes finer. Discretized models typically provide a balance between physical accuracy and computational efficiency. Examples include discretized Cosserat rod formulations through variational methods (Jung et al., 2011), geometrically exact

Cosserat through multi-body dynamics (Lang et al., 2011), Cosserat numerical space and time discretizations (Gazzola et al., 2018), as well as other conventional constitutive model approximations like the finite element method (FEM) (Duenser et al., 2018; Koessler et al., 2021).

- **Discrete:** While some analogies to continuous models may exist, discrete deformation models are inherently defined on a discrete structure and are not derived from the domain discretization of a continuous formulation. They are typically computationally efficient, but are often less physically realistic. This category includes physics-inspired models, such as the mass-spring-damper (MSD) model (Almaghout et al., 2024; Caporali et al., 2024b; Lv et al., 2017; Yu et al., 2023a), multi-body (MB) model (Servin and Lacoursiere, 2008; Yang et al., 2021), or purely geometric models, like the as-rigid-as-possible (ARAP) model (Sorkine and Alexa, 2007). Alternatively, projective dynamics (Bouaziz et al., 2023) and position-based dynamics (PBD) approaches (Bender et al., 2015) achieve a discrete analogy of Cosserat models (not geometrically exact, as opposed to, e.g., Lang et al. (2011)) by including the preservation of the angular momentum (Soler et al., 2018), through the introduction of ghost points for orientation constraints (Umetani et al., 2014), by incorporating quaternions for orientation in the PBD formulation (Kugelstadt and Schömer, 2016), or even achieving differentiable models through compliant PBD (Liu et al., 2023).

2.2.2. Static/quasi-static vs dynamic models. In DLO manipulation, many setups and tasks are quasi-static, assuming negligible inertial effects (e.g., a foam rod grasped by two grippers), while others require full dynamic modeling (e.g., a rope rapidly shaken by a robot). According to this criterion, models are classified into:

- **Static/quasi-static models:** They assume the system is in static equilibrium at each given time, neglecting inertial effects. These models are typically used when DLO deformations are slow and the object's state only (or mostly) changes under actuation. Examples include static Cosserat formulations (Jung et al., 2011; Rucker and Webster III, 2011; Tummers et al., 2023), quasi-static Kirchhoff rod models (Bretl and McCarthy, 2014), first order data-driven deformation Jacobian approximations (Zhu et al., 2018), quasi-static MSD (Almaghout et al., 2024; Caporali et al., 2024b), or static geometric models like Sorkine and Alexa (2007).
- **Dynamic models:** They model deformation behaviors that depend not only on the current state but also on the object's previous states. These models consider the time-dependent response of the object, typically including acceleration, velocity, or forces acting over time. They are typically governed by second-order differential equations that describe dynamic behavior. Examples of dynamic models include dynamic Cosserat rod formulations (Gazzola et al., 2018; Lang et al., 2011), PBD models (Kugelsstadt and Schömer, 2016), MB models (Servin and Lacoursiere, 2008), and other classical approaches such as dynamic MSD models (Cui et al., 2022). A notable subset of dynamic models addresses *fast* dynamics, which involve rapid motions where effects often neglected in slower scenarios become significant. These include aerodynamic forces (Shen et al., 2025), reduced influence of gravity due to high speeds (Yamakawa et al., 2013), and other fast transient phenomena, as studied in Zhang et al. (2021a) and Chi et al. (2024a).

2.2.3. *Physics-based vs physics-inspired vs empirical/heuristic.* This classification differentiates DLO models according to the degree to which they rely on fundamental physical principles. According to this criterion, models are classified into:

- **Physics-based:** Grounded in fundamental physical laws (e.g., continuum mechanics and deformation energy principles), they incorporate explicit constitutive relations that describe mechanical quantities such as stress, strain, stiffness, and internal forces and moments. Their parameters, such as Young's modulus, damping coefficients, or mass, are physically meaningful, enabling direct interpretability and calibration from experimental material data. These models are particularly suitable for modeling force inputs and boundary conditions (e.g., continuously distributed loads). Representative examples include the Cosserat model (Gazzola et al., 2018; Jung et al., 2011; Lang et al., 2011; Tummers et al., 2023), the quasi-static Kirchhoff rod model (Bretl and McCarthy, 2014), or the Timoshenko rod model, which accounts for shear deformation in thicker DLOs.
- **Physics-inspired:** These models are informed by physical intuition but do not explicitly formulate fundamental mechanical laws. They typically omit rigorous constitutive relations and force or moment equilibrium, instead approximating mechanical behavior through simplified formulations such as deformation energy or nodal balance (e.g., MSD systems). Parameters generally lack a direct and measurable link to material properties, reducing physical interpretability while offering computational simplicity and flexibility. Examples include PBD Cosserat approximations (Kugelsstadt and Schömer, 2016; Umetani et al., 2014), MBD (Servin and Lacoursiere, 2008; Yang et al., 2021), DS (Palli, 2020; Theetten et al., 2008; Valentini and Pennestri, 2011), and MSD formulations (Almaghout et al., 2024; Caporali et al., 2024b; Lv et al., 2017; Yu et al., 2023a).
- **Empirical/heuristic:** These models are based purely on observed geometric behavior or approximations, without explicit consideration of mechanical deformation principles. They model shape evolution or kinematics effectively but do not capture internal mechanical states (e.g., forces, moments, or strains), limiting their ability to handle physical interactions such as contact, gravity, or friction in a principled way. Examples include geometry-based approaches like (Aghajanzadeh et al., 2022b; Sorkine and Alexa, 2007) and online-updated deformation Jacobian estimations (Cuiral-Zueco et al., 2023; Zhu et al., 2018). While such models can be extended with additional layers to infer mechanical quantities, these are not intrinsic to their formulation.

2.2.4. *Numerical versus data-driven.* This classification considers the computational approach typically used by models to simulate or predict DLO behavior. Analytical (closed-form) solutions are disregarded here due to their rarity in practical DLO models.

- **Numerical methods:** They are typically used for physically grounded models and approximate their governing equations through techniques such as ODE integration, BVP shooting, Gauss–Newton optimization, or finite difference schemes (Jung et al., 2011; Tummers et al., 2023). Numerical methods also include energy minimization for discretized physically-based models (Bender et al., 2015) or geometrically inspired models (Sorkine and Alexa, 2007). Models solved through numerical methods are grounded in mathematical representations of physical or geometric behavior and thus can be applied to a wide range of DLO manipulation scenarios and constraints.
- **Data-driven methods:** They model DLO behavior by learning patterns directly from empirical or simulated data, using statistical approaches (Aghajanzadeh et al., 2022a; Cuiral-Zueco et al., 2023; Navarro-Alarcon et al., 2013; Zhu et al., 2018) or machine learning techniques (Huo et al., 2022; Laezza and Karayiannidis,

2021; Yang et al., 2022a; Zakaria et al., 2022), which are generally more data-intensive. These methods prioritize low online computational cost and typically require minimal prior knowledge of the underlying physics. While capable of capturing complex, nonlinear dynamics implicitly, their performance is highly sensitive to the quality, diversity, and representativeness of the training data. Consequently, they may struggle to generalize across varying tasks or environmental conditions.

- **Hybrid approaches:** They combine elements from both numerical and data-driven methods. These models typically embed physical laws or constraints into data-driven frameworks to improve generalization and robustness. Examples include machine learning for parameter estimation (Caporali et al., 2024b), or data-acquisition from physical models for data-driven Jacobian estimation (Artinian et al., 2024). A prominent example is physics-informed neural networks, which incorporate governing equations directly into the training process (Bensch et al., 2024). Other hybrid approaches may use simulation data to pre-train models or augment learning with differentiable physics modules or simulators (Liu et al., 2023), or exploit graph neural networks to initialize iterative solvers (Shao et al., 2021).

2.3. Simulation of DLOs

Several physics-based simulators support DLO modeling, each providing unique numerical methods, levels of physical realism, and degrees of integration with robotic platforms. Below is an overview of some of the most prominent simulators commonly employed in DLO research, highlighting their core modeling approaches and key references (see Table 1 for a summary):

- *Bullet* (Coumans, 2025) and *MuJoCo* (Todorov et al., 2012): These are widely adopted open-source physics engines that employ PBD-like solvers for real-time simulations. In *Bullet*, DLOs can be represented either as FEM-like soft bodies or as chains of cylindrical segments connected by 6D springs, mimicking MB

models. *MuJoCo* offers two main representations: *cables*, which model inextensible rods with bending and twisting stiffness using a geometrically exact formulation discretized into capsules or boxes; and *1D flex*, recommended for simulating extensible strings under tension, such as rubber bands.

- *AGX Dynamics*¹: Proprietary physics simulation platform that employs a hybrid constraint-based solver for real-time simulations. It offers specialized modules such as *Wires*, for simulating long, bendable structures under extreme tension and in large-scale scenarios, and *Cables*, which capture elastic deformations as well as plasticity, using sequences of rigid bodies connected by constraints.
- *Obi*²: Real-time particle-based physics engine that uses extended PBD to simulate deformable objects, available as a plugin for Unity. It supports rod simulations with stretch/shear and bend/twist constraints, and rope simulations with distance and bend constraints.
- *FleX* (Macklin et al., 2014): Open-source GPU-accelerated particle-based simulator that represents all objects, including DLOs, as particle systems governed by PBD.
- *IsaacLab* (Mittal et al., 2025): A GPU-accelerated simulation and robot-learning framework built on Isaac Sim and PhysX, and the successor to Isaac Gym. It provides parallel physics, photorealistic rendering, and rich multi-modal sensor simulation, together with tools for domain randomization and large-scale data collection. While not DLO-specific, IsaacLab exposes PhysX’s GPU-accelerated FEM soft-body capabilities, enabling simulation of cables and other deformable linear structures within robotic environments.
- *Elastica* (Naughton et al., 2021): Open-source software package designed to numerically solve systems made up of collections of Cosserat rods. While not a general-purpose simulator and not real time, it offers physically accurate modeling of slender structures.

3. Perception

DLOs perception is primarily achieved using vision-based (Section 3.1) or tactile sensors (Section 3.2). Additionally,

Table 1. Overview of widely used physics-based frameworks for DLO simulation in robotics, highlighting their core DLO modeling approaches and representative references.

Simulator	Approach	DLO references
Bullet	MB	Yang et al. (2021)
	FEM	Seita et al. (2021); Zakaria et al. (2022); Daniel et al. (2024)
Mujoco	PBD	Chi et al. (2024a); Zhaole et al. (2024)
AGX Dynamics	MB (elastoplastic)	Laezza and Karayiannidis (2021); Laezza et al. (2021); Yang et al. (2022a,b)
Obi (+Unity)	Extended PBD	Yu et al. (2022, 2023a, 2025); Lv et al. (2023); Luo and Demiris (2025)
FleX	Particle-based	Pecyna et al. (2022); Ma et al. (2022)
IsaacLab (PhysX)	FEM (soft bodies)	Kamaras and Ramamoorthy (2025); Govoni et al. (2025)
Elastica	Cosserat rod	Caporali and Palli (2025)

less common proximity sensors and force/torque sensors are reviewed in Section 3.3. An overview of the section’s structure is provided in Figure 3. Perception-related challenges and future research directions are discussed in Section 6.1.

3.1. Vision-based perception

Vision-based perception is the most common sensing modality for DLOs, thanks to the availability of diverse sensors and cameras that integrate seamlessly with robotic platforms. The main perception tasks for DLOs include *Data-driven Segmentation* (Section 3.1.1), *2D Shape Estimation* (Section 3.1.2), *3D Shape Estimation* (Section 3.1.3), and *Tracking* (Section 3.1.4). Sections 3.1.5 and 3.1.6 offer a brief overview of vision-based perception methods specifically developed for *DMLOs* and *suture threads*. Finally, Section 3.1.7 presents vision-related *Emerging Tasks*.

3.1.1. Data-driven segmentation. Among various vision tasks, some provide limited utility for DLO manipulation. For instance, while object detection can generate bounding boxes around DLOs, these often fail to capture precise information about the DLO shape and configuration, details crucial for manipulation. In contrast, *semantic segmentation*, and particularly *instance segmentation*, delivers richer pixel-level information by accurately identifying DLO regions within an image, as illustrated in Figure 4.

Semantic segmentation involves classifying each pixel of an image with a unique class, for example, the DLO class or the background class. *Instance segmentation* further

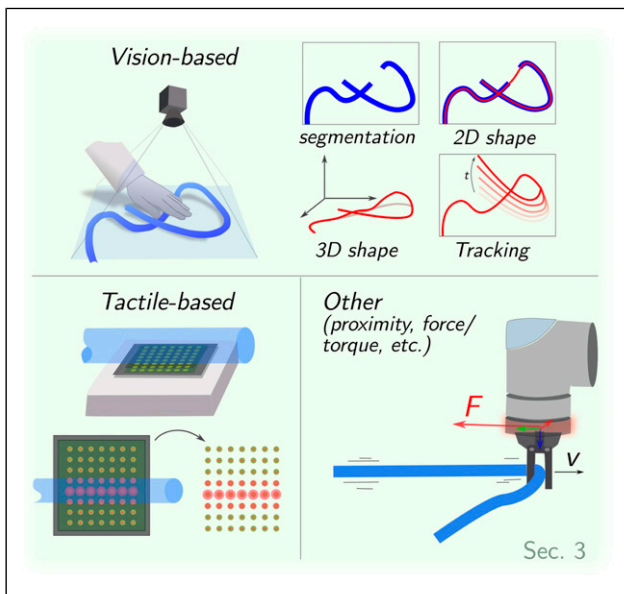


Figure 3. The three main DLO perception method groups analyzed in Section 3: vision-based, tactile, and other sensor-based (e.g., proximity, force/torque).

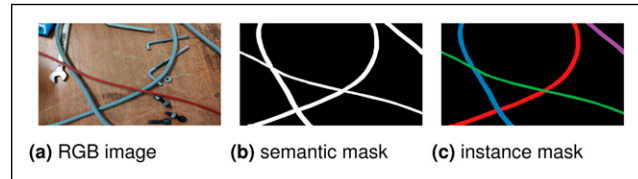


Figure 4. Example of semantic and instance segmentation tasks involving four DLOs in an industrial scenario.

distinguishes individual DLO objects by assigning unique identifiers to each instance, allowing differentiation among multiple DLOs in the scene.

Traditional segmentation techniques, like color-based thresholding or background difference, rely on substantial assumptions about the scene, making them usually unsuitable as general solutions. Recently, deep learning approaches have demonstrated their viability in effectively solving some of the segmentation challenges related to DLOs, driving research toward data-driven segmentation of DLOs (Caporali et al., 2023c; Dai et al., 2022; Dirr et al., 2023; Huang et al., 2024; Jin et al., 2022; Song et al., 2019; Sun et al., 2024; Wu et al., 2022b; Zanella et al., 2021).

3.1.1.1. Dataset generation. The key issue in deep learning approaches increasingly revolves around the challenge of gathering and labeling large amounts of data for training purposes. Several works rely on manual annotation procedures to generate a training dataset, as seen in Wu et al. (2022b); Dai et al. (2022), Song et al. (2019), and Huang et al. (2024). However, the manual process is notoriously tedious, inaccurate, time-consuming, and not scalable. Furthermore, as visual perception tasks grow more complex (such as in DLO segmentation), the annotation effort becomes increasingly slow and challenging.

Some works have focused on investigating dataset-generation approaches that require minimal or, ideally, zero human intervention. In Jin et al. (2022), a self-supervised method is presented that collects training images by moving a camera mounted on a robot arm. Initial labels are generated using color thresholding on a high-contrast DLO, and a deep-learning segmentation network trained on augmented data is employed to enhance the estimator’s ability to generalize across varying cable colors and backgrounds.

A similar approach is investigated in Zanella et al. (2021), which proposes a two-phase data labeling method for semantic segmentation: first, foreground masks are created using color difference between the DLO and background; second, these masks are combined with synthetic backgrounds to form the training dataset. Initial images are collected with minimal human effort by moving the DLO against a uniform background, and the dataset is then augmented to improve generalization to new scenes.

Compared to Jin et al. (2022), the method in Zanella et al. (2021) enables generalization across the general DLO class and is not limited to specific scenes, due to the use of synthetic backgrounds. The main limitations of both Jin et al. (2022) and Zanella et al. (2021) are (1) susceptibility to incorrect labels due to color separation from video, which is sensitive to lighting and shadows, necessitating validation; and (2) the need for human intervention in data gathering, particularly for the movement and deformation of DLOs.

To reduce human involvement in dataset generation, alternative approaches advocate fully synthetic processes using rendering engines (e.g., Blender) to create photo-realistic datasets (Caporali et al., 2023c; Dirr et al., 2023; Fresnillo et al., 2024). Synthetic images offer the added benefit of automatically generating accurate and error-free labels for both semantic and instance segmentation. Although labeling is eliminated, significant time is still required to set up and implement the synthetic scene generation pipeline. Additionally, the use of synthetic data raises concerns about the *domain gap* between simulated and real-world environments.

To mitigate the *domain gap*, Caporali et al. (2023c) propose a weakly supervised method that leverages keypoint annotations on real DLO images captured from multiple viewpoints. Then, a neural network refines the sparse keypoint-based annotations into dense segmentation labels. Although these methods effectively capture real-world details, they require multiple annotations to account for scene variability, limiting their scalability. Nevertheless, small-scale real datasets remain valuable when combined with synthetic data to reduce the *domain gap* (Caporali et al., 2023c).

3.1.1.2. Semantic segmentation. The semantic segmentation of DLOs is performed using off-the-shelf deep learning models based on convolutional neural networks (CNNs): UNet (Ronneberger et al., 2015) in Dai et al. (2022) and Jin et al. (2022); FCN (Long et al., 2015) in Wu et al. (2022b); and DeepLabV3+ (Chen et al., 2018) in Zanella et al. (2021), Dai et al. (2022), and Caporali et al. (2023c). A CNN-based encoder-decoder architecture is proposed in Huang et al. (2024) and Song et al. (2019).

Concerning real-world and synthetic datasets, both Dai et al. (2022) and Caporali et al. (2023c) evaluate their respective synthetic datasets against the *electrical wires dataset* (featuring real DLOs images but with synthetic backgrounds) released by Zanella et al. (2021). From the comparisons, synthetic images emerge as a viable alternative to real-world image labeling. Moreover, mixing synthetic images with real-world images is shown to improve segmentation performance compared to the synthetic-only case (Caporali et al., 2023c).

3.1.1.3. Instance segmentation. The pipelines introduced in Caporali et al. (2023c) and Dirr et al. (2023) enable

instance-wise mask generation using fully synthetic approaches. When applied to instance segmentation models (e.g., YOLACT (Bolya et al., 2019) in Caporali et al. (2023c) and SOLOv2 (Wang et al., 2020) in Dirr et al. (2023)), performance is generally weaker than in semantic segmentation tasks, particularly in scenarios where different DLOs intersect. This underscores the need for further research into fully deep-learning-based instance segmentation methods tailored to DLOs, or the exploration of alternative strategies, as discussed in Section 3.1.2.

3.1.2. 2D shape estimation. Accurately estimating the shape of a DLO is crucial for effective manipulation. As a result, many vision-based algorithms have been developed to robustly extract the DLO's state, typically represented as a sequence of keypoints that describe its shape and configuration. This section focuses on methods for 2D shape estimation, while 3D shape estimation techniques are covered separately in Section 3.1.3.

Table 2 provides an overview of the main approaches for *2D Shape Estimation*. The table serves as a convenient reference, highlighting the data-driven nature of each method, along with its inference speed, pre-processing requirements, simplification and DLOs crossing strategies, and core procedures. The listed methods and their classification are further analyzed in the following discussion.

3.1.2.1. Pre-processing. The majority of the approaches assume the utilization of a pre-processing step to generate a semantic segmentation mask of the scene. This mask is typically a binary image, where pixels corresponding to the DLO are labeled as 1, and background pixels as 0. This step is common across methods such as Caporali et al. (2022a,b, 2023b), where a semantic segmentation network is employed (see Section 3.1.1), color-based thresholding methods as in Choi et al. (2023) and Keipour et al. (2022a), or depth-thresholding approaches as in the case of the RGB-D camera in Kicki et al. (2023).

In contrast, some approaches necessitate initialization with endpoint locations. Specifically, endpoints are supplied by specific CNN-based object detection networks in De Gregorio et al. (2018a) and Viswanath et al. (2023). Similarly, external knowledge of the scene structure is harnessed in Fresnillo et al. (2023) to initialize the algorithm.

Alternatively, Huang et al. (2024) uses a dedicated segmentation network to extract a gradient map of the DLOs directly from RGB images, eliminating separate segmentation and endpoint detection steps. However, its generalizability beyond training-like scenarios needs careful assessment (see Section 3.1.1). Yan et al. (2020) avoid both segmentation and endpoint detection but relies heavily on strong contrast between the DLO and background in its self-supervised process.

3.1.2.2. Image simplification. The methods summarized in Table 2 propose various innovative solutions for DLO shape

Table 2. Summary of the main literature on 2D shape estimation of DLOs via vision-based sensors. The methods are presented in chronological order from the oldest (first row) to the newest (last row). Time inference scale: several seconds (*), real time (****).

Reference	Acronym/first author	Data-driven	Code available	Inference time	Pre-processing	Image simplification	Core procedure	Crossing order determination
De Gregorio et al. (2018a)	Ariadne		✓	*	Endpoints detection	Superpixels	Tracing	-
Yan et al. (2020)	Yan et al.	✓		-	-	-	Hierarchical update	-
Keipour et al. (2022a)	Keipour et al.			**	Segmentation	Skeleton	Merging	-
Caporali et al. (2022b)	Ariadne+	✓	✓	**	Segmentation	Superpixels	Merging	Patch classifier
Caporali et al. (2022a)	FASTDLO	✓	✓	***	Segmentation	Skeleton	Merging	Color deviation
Huang et al. (2024)	Huang et al.	✓		-	-	Skeleton	Tracing	Patch classifier
Kicki et al. (2023)	DLOFTBs		✓	****	Segmentation	Skeleton	Merging	-
Caporali et al. (2023b)	RT-DLO	✓	✓	****	Segmentation	Graph-based	Merging	Color deviation
Choi et al. (2023)	mBEST		✓	****	Segmentation	Skeleton	Merging	Color deviation
Fresnillo et al. (2023)	Fresnillo et al.		✓	*	-	-	Tracing	-
Viswanath et al. (2023)	HANDLOOM	✓	✓	*	Endpoints detection	Skeleton	Tracing	Patch classifier

estimation, yet many approaches share common image processing strategies. A prevalent approach is to reduce image complexity using either *superpixels* or *skeleton*-based techniques. To highlight the differences between these two strategies, Figure 5 presents a side-by-side comparison of their application to the same input image.

Superpixel segmentation, as used in De Gregorio et al. (2018a) and Caporali et al. (2022b), groups pixels with similar properties into coherent regions. De Gregorio et al. (2018a) employ the SLIC algorithm (Achanta et al., 2012), while Caporali et al. (2022b) use MaskSLIC (Irving, 2016), which incorporates a binary mask to restrict focus to targeted areas within the image.

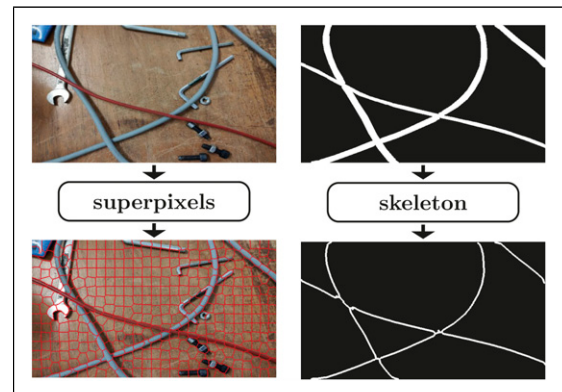
Skeletonization is an alternative approach consisting of a thinning procedure performed on a binary mask. It is a widely chosen method for mask-based simplification (Caporali et al., 2022a; Choi et al., 2023; Huang et al., 2024; Keipour et al., 2022a; Kicki et al., 2023; Viswanath et al., 2023). Its popularity is attributed to several key properties: (1) after the skeleton operation, both the connectivity and general topology of the DLOs are preserved; (2) since the segments are only 1 pixel wide, traversals along segments are not prone to path ambiguity; and (3) fast implementations are feasible. Among various algorithms, one of the most frequently used methods is Zhang and Suen (1984). Applied to a binary mask, the skeleton approach is quite sensitive to the mask quality.

Unlike superpixel- and skeleton-based methods, Caporali et al. (2023b) use a graph-based representation where nodes are generated by dilating the distance

transform (Borgefors, 1986) and applying farthest point sampling (Qi et al., 2017), improving robustness to poor masks. Methods like Yan et al. (2020) and Fresnillo et al. (2023) bypass any form of image simplification and instead process raw images directly.

3.1.2.3. Main procedure. Among the examined algorithms of Table 2, two core processes consistently appear: *tracing* and *merging*. For clarity, these processes are illustrated in Figure 6.

The *tracing* process iteratively extends the existing path of a DLO by adding new segments, building upon previously traced portions. In contrast, *merging* algorithms

**Figure 5.** Superpixel and skeleton-based image simplification techniques. The first is applied directly to RGB images. The latter requires a binary mask.

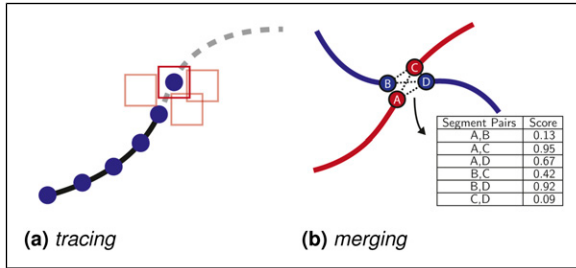


Figure 6. Illustration of DLO shape estimation procedures based on *tracing* and *merging* approaches. In *tracing*, the DLO curve is expanded incrementally by adding small, locally guided segments one at a time. In contrast, *merging* combines larger, pre-estimated segment hypotheses into longer, coherent curves based on similarity scores or geometric compatibility.

combine smaller, independent segment estimates into a unified detection. Although similar in goal, the key difference lies in scale and dependency: *tracing* operates locally and sequentially, while *merging* handles fewer, independent operations that can be performed in any order, often in parallel. This flexibility offers notable advantages in reducing inference time.

Several algorithms leverage forms of *tracing* to extract paths (De Gregorio et al., 2018a; Fresnillo et al., 2023; Huang et al., 2024; Viswanath et al., 2023). For example, De Gregorio et al. (2018a) generate candidate paths by tracing through superpixels from endpoints, selecting paths based on color, curvature, and distance. Fresnillo et al. (2023) trace in both forward and backward directions for increased robustness. Huang et al. (2024) use a skeleton map to trace between endpoints, while Viswanath et al. (2023) introduce a data-driven approach, using a UNet-based trace predictor to produce probability heatmaps that guide the tracing process.

Merging-based algorithms use cost functions to evaluate and merge candidate segment pairs (Choi et al., 2023; Kiki et al., 2023; Keipour et al., 2022a; Caporali et al., 2022a, 2022b, 2023b). Across these approaches, the choice of metric varies substantially. A data-driven metric in Caporali et al. (2022b) employs a CNN to assess superpixel similarity, whereas Caporali et al. (2022a) use a similarity network to compare sampled feature vectors. In contrast, analytical metrics such as curvature, distance, and shape smoothness are adopted in Choi et al. (2023), Keipour et al. (2022a), and Caporali et al. (2023b), with Choi et al. (2023) introducing a merging criterion based on bending energy.

In practice, merging is often performed over an intermediate structural representation of the object. A common implementation relies on a skeleton map, as initial segments can be formed by linking skeleton pixels with two neighbors, then resolving intersections via merging. This approach is employed by Keipour et al. (2022a), Choi et al. (2023), and Caporali et al. (2022a). In Caporali et al. (2022b), superpixels replace the skeleton map but apply a

similar merging strategy exploiting the segmentation masks. In Caporali et al. (2023b), the sparsity of nodes prevents standard merging, so each node is merged individually with its neighbors. Unlike tracing, this merging is performed concurrently across all nodes, based on both their similarity and spatial proximity.

Unlike merging and tracing methods, Yan et al. (2020) use a neural network to hierarchically update DLO segment endpoints and predict new center points, progressively increasing the granularity of the shape representation.

3.1.2.4. Crossing order determination. In cases of crossings or intersections, for example, whether between different DLOs or loops of the same DLO, determining which segment lies on top is essential for manipulation tasks, such as deciding which DLO to move (see Section 5.4). The two common strategies for this, illustrated in Figure 7, are data-driven patch classifier methods and analytical color deviation techniques.

CNN classifiers based on ResNet architectures analyze masked image crops centered on the crossing region to predict which segment lies on top (Caporali et al., 2022b; Viswanath et al., 2023). To enhance robustness and ensure rotation invariance, the crops are often rotated or oriented so that the DLO segments are aligned consistently before classification (Huang et al., 2024).

Alternatively, analytical methods determine order by comparing the color variability along each path near the crossing, selecting the segment with lower RGB channel variance as the top (Caporali et al., 2022a; 2023b). This approach is further refined by using blurred images to reduce glare effects (Choi et al., 2023).

3.1.3. 3D shape estimation. While estimating the 2D shape of a DLO (see Section 3.1.2) provides valuable information, it is often insufficient for effective grasping and manipulation. The ultimate objective is to recover the DLO’s configuration in 3D space. However, direct 3D shape estimation remains

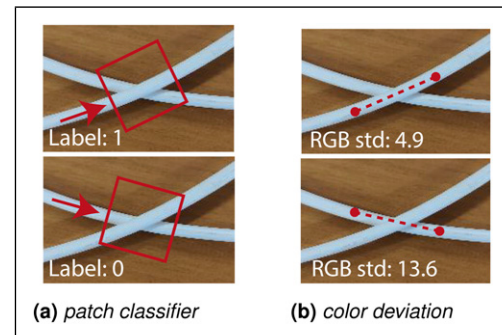


Figure 7. Illustration of two strategies for determining the crossing order of DLOs: (a) *data-driven patch classifier*, which predicts the top segment from a cropped image region; (b) *color deviation* method, which infers the top segment by comparing color variability along the intersecting paths.

less explored than its 2D counterpart. Indeed, a common strategy involves first estimating the 2D shape and subsequently projecting it into Cartesian space using depth information (Kicki et al., 2023; Sun et al., 2024).

One reason for the limited number of 3D shape estimation methods is the challenge posed by current sensing technologies. As highlighted in a benchmark of 3D camera systems for DLO perception (Cop et al., 2021), only high-end depth sensors can reliably capture the shape of thin, cylindrical objects like DLOs. These high-performance cameras tend to be bulky—making them difficult to mount on robot end-effectors—and their cost restricts widespread use in research settings. In contrast, popular robotic cameras such as the Kinect Azure and Intel RealSense often struggle to detect DLOs with diameters under 1 cm.

To mitigate the impact of depth sensor noise, Sun et al. (2024) incorporate a smoothness prior using a discrete elastic rod model (see Section 2). Reliable 3D detection of DLOs is also explored in Caporali et al. (2023a), which leverages a multi-view stereo approach using a single 2D camera in combination with 2D shape detection algorithms (see Section 3.1.2). This method proves effective for reconstructing the 3D shape of DLOs for grasping and manipulation tasks. However, achieving accurate results requires detecting multiple, closely spaced DLO segments, making the process time-consuming. Additionally, the approach is limited to static scenes, restricting its use in dynamic environments.

3.1.4. Tracking. DLOs are theoretically characterized by an infinite number of degrees of freedom, which makes tracking challenging, particularly under real-time constraints. In practice, they are discretized into a finite set of key nodes (as described in Section 2), and tracking (in its basic form) reduces to estimating the positions of these nodes over time while handling potential occlusions (as in Figure 8). These occlusions may result from self-occlusion within the DLO or from external factors, such as interactions with a robotic manipulator.

A wide range of methods have been proposed to address the challenges of DLO tracking, which vary in their modeling assumptions, algorithmic foundations, and robustness to occlusions. Table 3 categorizes these methods into two main groups: registration-based approaches, primarily built upon non-rigid registration techniques such as coherent point drift (CPD; Myronenko and Song, 2010) and global-local topology preservation (GLTP; Ge et al., 2014), and learning-based methods.

3.1.4.1. Registration-based methods. A widely adopted approach to DLO tracking formulates the task as a point-set registration problem, leveraging algorithms such as CPD and GLTP. In CPD, registration is framed as a probability density estimation problem, where one point set—the

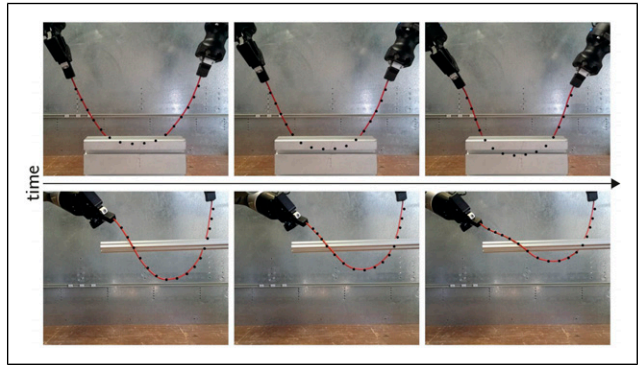


Figure 8. Example of DLO tracking during manipulation. The DLO key nodes are tracked despite partial visibility and environment interactions.

Gaussian mixture model (GMM) centroids, typically represents the estimated positions of key nodes along the DLO, and the other set consists of observed data points from the current camera frame. A key feature of CPD is the enforcement of coherent motion among the GMM centroids, ensuring smooth and physically plausible deformations (Yuille and Grzywacz, 1989). GLTP extends CPD by introducing a local regularization term based on locally linear embedding (Roweis and Saul, 2000), complementing CPD’s global regularization.

While CPD and GLTP are effective for non-rigid registration, they do not inherently incorporate physical constraints or domain knowledge specific to deformable objects. To address this limitation, several DLO tracking approaches have extended these algorithms by integrating physics-based simulators or introducing regularization techniques tailored to DLO behavior. As summarized in Table 3, each registration-based approach enhances the core registration pipeline with DLO-specific constraints or priors to improve tracking accuracy and robustness. A common direction involves embedding physical knowledge, either through a physics simulator (Luo and Demiris, 2025; Tang et al., 2017; Tang and Tomizuka, 2022) or an analytical model from Section 2 (Wang and Yamakawa, 2022). Alternatively, constraints based on stretching limits (Chi and Berenson, 2019) or geometric/topological properties (Wang et al., 2021; Xiang et al., 2023) are employed to improve robustness and accuracy.

3.1.4.2. Learning-based methods. Recent work in DLO tracking has explored data-driven approaches (Caporali and Palli, 2025; Lv et al., 2023; Yang et al., 2022c), as shown at the bottom of Table 3. These methods aim to overcome challenges such as high dimensionality, occlusions, or the need for explicit physical modeling.

In Yang et al. (2022c), an autoencoder is used to learn a low-dimensional embedding of DLO states, enabling efficient tracking via particle filtering in the latent space. This approach captures physically plausible behaviors directly

Table 3. Overview of DLO tracking methods grouped by non-rigid registration techniques and learning-based approaches.

Group	Method	Based on	Key aspects
Registration-based	CPD + Physics (Tang et al., 2017)	CPD	DLO physical simulator
	SPR (Tang and Tomizuka, 2022)	CPD + physics	Regularization using local structure and global topology
	CDCPD (Chi and Berenson, 2019)	GLTP	Enforces DLO stretching limits; robust recovery from tracking failures
	CDCPD2 (Wang et al., 2021)	CDCPD	Handles tip, self, and severe occlusions via constraint incorporation
	CPD + FEM (Wang and Yamakawa, 2022)	CPD	FEM model integrating local structure, global topology, and material properties
	TrackDLO (Xiang et al., 2023) TSL (Luo and Demiris, 2025)	GLTP CDCP2	Addresses tip and self-occlusions using geodesic distances Shoelace tracking
Learning-based	Yang et al. (2022c)	–	Low-dimensional embedding space via autoencoder
	Lv et al. (2023)	–	Two-branch encoding network combined with modified CPD
	Caporali and Palli (2025)	–	Multi-view triangulation combined with Cosserat model of DLO behavior

from data, without requiring a physical simulator or regularization during deployment. Lv et al. (2023) employ a PointNet++ encoder to extract features from input point clouds, followed by a two-branch fusion strategy: a regression branch that models global DLO topology, and a voting branch that estimates local geometric offsets. Thus, a modified CPD algorithm fuses both branches.

A key advantage of these learning-based methods is their independence from initial DLO state estimates, simulators, or hand-crafted constraints at inference time. Instead, they encode physical priors during training. For instance, Yang et al. (2022c) use synthetic data matching specific DLO properties, though performance may degrade when real-world behavior deviates from the training distribution. In contrast, Lv et al. (2023) apply domain randomization to improve generalization to real-world scenarios.

Tracking methods often require high-quality depth or point cloud data, difficult to obtain given the small size of DLOs (Cop et al., 2021), and depend on pre-segmented inputs, which are challenging to acquire outside controlled settings. To address these issues, Caporali and Palli (2025) propose using multiple 2D images for segmentation in cluttered scenes (see Sections 3.1.2 and 3.1.3) combined with a learned physics-based DLO model to handle occlusions, enabling estimation and tracking of the 3D DLO shape during manipulation. However, this approach depends on knowledge of the robot’s actions and has limited use of tracking history.

3.1.5. Vision-based perception of suture threads. A notable subclass of DLOs is represented by *suture threads*, which are extensively studied in the field of surgical robotics. These threads are typically inextensible, have very small diameters, and are often connected to a curved metal needle. Suture threads pose unique challenges for vision-based perception due to their extremely thin structure. Indeed, in typical surgical

imaging setups, the thread amounts to roughly 0.25% of the total image width (Joglekar et al., 2023). This makes them significantly more difficult to detect and track than larger deformable objects such as ropes or cables. An overview of perception methods for suture threads is provided here, while a broader discussion on robotic suturing for manipulation tasks is presented in Section 5.5.

Various methods have been proposed for reconstructing thread geometry from visual input, including stereo-based curve fitting using non-uniform rational B-splines (Jackson et al., 2018; Schorp et al., 2023), shortest path computations between thread endpoints (Lu et al., 2020, 2022), and minimum variation splines (Joglekar et al., 2023). These methods typically incorporate prior knowledge of thread continuity and curvature to compensate for weak visual cues. Many treat the problem as a curve-fitting task (Jackson et al., 2018; Schorp et al., 2023). Some connections with the methods discussed in Section 3.1.2 are also present, for example, the cost-based path growth in Jackson et al. (2018) closely resembles the *tracing* approach.

Learning-based techniques have also been introduced to enhance suture perception. Lu et al. (2020) apply transfer learning to improve generalization, while Lu et al. (2022) propose a semi-supervised segmentation approach to improve thread detection with limited labeled data.

Some systems require manual initialization, such as seeding a starting point (Jackson et al., 2018; Lu et al., 2022). In contrast, more recent methods achieve fully automatic suture reconstruction without user input (Joglekar et al., 2023). Both Joglekar et al. (2023) and Lu et al. (2022) also address a critical challenge in stereo-based perception of false correspondences across stereo image pairs, especially due to the thread being tangent to the epipolar line (Joglekar et al., 2023).

A key challenge in suture thread perception is the lack of standardized benchmarks and public datasets (see Section

6.2), which are especially essential in data-scarce clinical settings where real-world samples are limited (Joglekar et al., 2023).

3.1.6. Vision-based perception of DMLOs. As a subclass of DLOs, DMLOs share many of the same properties but are uniquely characterized by the presence of *branch points*, where two or more linear components converge (Caporali et al., 2025). Extending DLOs vision-based perception methods to DMLOs introduces additional challenges, primarily due to the complexity of bifurcations. Moreover, the integration of rigid elements such as plugs, clips, and connectors further complicates the visual processing and interpretation of these structures.

Most research on DMLOs centers on automotive wiring harnesses, highlighting the need for advanced automation. Nguyen and Franke (2021) use data-driven segmentation methods (see Section 3.1.1) for optical inspection, though with limited manually annotated data. In follow-up work (Nguyen et al., 2022), synthetic data from CAD models is introduced, demonstrating its effectiveness and benefits over limited real data. In Kicki et al. (2021), DMLO branch classification is explored using a manually annotated, small-scale dataset, with data augmentation applied to counter limited data availability.

Several works represent DMLOs as graphs generated around branch points (Zürn et al., 2022, 2023b, 2023a). For instance, Zürn et al. (2023b) estimate correspondences between a known (CAD-based) directed topology and an image-derived undirected graph. Zürn et al. (2022) introduce a DMLO tracking method using rigid and non-rigid registration but assumes non-overlapping configurations. Zürn et al. (2023a) address branch point detection with a data-driven method and semi-manual annotation, though its evaluation is limited to a single user and DMLO type.

Caporali et al. (2025) present a learning-based graph method for representing DMLO topology using graph neural networks trained on synthetic data. The method’s effectiveness is demonstrated in a dual-arm disentangling task (see Section 5.4). However, since it relies solely on the segmented mask of the scene, it remains vulnerable to significant errors caused by mask inaccuracies.

Despite significant progress in DMLO perception, future research should focus on developing systems that are not only accurate but also adaptable, scalable, and robust to the inherent variability of real-world environments, as discussed in Section 6.2.

3.1.7. Emerging vision-related tasks. Recent advances in computer vision have enabled the exploration of novel tasks in the context of DLO perception.

3.1.7.1. Multi-modal segmentation. Few works have explored multi-modal approaches for DLO segmentation. The segment anything model (SAM) (Kirillov et al., 2023) has been applied in zero-shot settings. For

example, Sun et al. (2024) prompt SAM using generic keywords such as “ropes” or “cables,” but require several post-processing steps to achieve satisfactory segmentation masks. Caporali et al. (2024a) leverage both visual and textual modalities to segment only the target DLO. It introduces two main improvements over prior works: (1) task-specific prompts for accurate target-object segmentation, and (2) a lightweight, real-time capable architecture, unlike the larger foundation models.

3.1.7.2. Interactive segmentation. It is the process of leveraging forceful physical interactions with objects to enhance and inform the perception process (Bohg et al., 2017; Weng et al., 2024). Holešovský et al. (2024) propose an optical flow-based approach for segmenting moving DLOs, inspired by how humans use motion—specifically poking—to distinguish and separate tangled cables. They also introduce an automatically annotated dataset with instance and motion ground truth. Due to reliance on flow magnitude thresholding, the method may merge multiple moving cables. To address this, Holešovský et al. (2025) incorporate motion correlation and interactive grasping strategies to improve accuracy. The approach is evaluated on small motions and thick DLOs, given vision limitations in more complex scenarios (see Section 3.1.3).

3.2. Tactile-based perception

Vision-based perception can be challenging in tight spaces and with occlusions. Moreover, when the robot is actively manipulating the object, having additional information on the grasp itself becomes crucial. Tactile sensors represent a valuable alternative to overcome the limitations of vision-based perception of DLOs. Concerning DLOs’ touch perception, three types of tactile sensor technologies are usually employed: photoreflector-based (Cirillo et al., 2021b; Palli and Pirozzi, 2019; Pirozzi and Natale, 2018; Zanella et al., 2019), camera-based (She et al., 2021; Wilson et al., 2023), and capacitive (Monguzzi et al., 2023; 2024a).

Photoreflector tactile sensors are employed in Pirozzi and Natale (2018) and Palli and Pirozzi (2019) for automatic wiring tasks. Assuming a known DLO diameter, the tip of a grasped DLO is estimated by modeling the grasped section with a quadratic function and the external section with a linear segment. In contrast, Cirillo et al. (2021b) propose a data-driven method to estimate the DLO diameter, leveraging both tactile measurements and gripper closure levels to classify diameters and generalize across varying grasp levels. Additionally, tactile sensing is also applied to estimate external forces acting on the grasped DLO through a data-driven approach based on RNNs (Zanella et al., 2019).

Camera-based tactile sensors allow for the estimation of pose and friction force (She et al., 2021; Wilson et al., 2023). Pose estimation is achieved through processing the depth

image and applying principal component analysis (PCA). The friction force is determined by computing the marker flow on the tactile surface, with the estimated displacement assumed to be proportional to the friction force. Additionally, [She et al. \(2021\)](#) introduce grasp quality, assessed by evaluating the area of the tactile imprint in relation to a predefined threshold boundary.

Capacitive tactile sensors are employed in [Monguzzi et al. \(2023, 2024a\)](#), where capacitance measurements are exploited to estimate DLO diameter and alignment.

Both photoreflector and capacitive-based tactile sensors provide low-resolution output (usually a 4×4 or 5×5 map), unlike high-resolution camera-based sensors that generate depth-image-like outputs (see [Figure 9](#)). However, photoreflector and capacitive sensors are generally more compact and slim, making them better suited for use in confined or tight spaces where larger camera-based sensors may not fit.

Several works combine vision and tactile sensing to overcome their individual limitations in DLO manipulation. [De Gregorio et al. \(2018b\)](#) employ vision for tip detection and tactile sensors to assess grasp quality. In [Pecyna et al. \(2022\)](#), vision and tactile data are integrated within a manipulation framework, highlighting the importance of fusing both sensing modalities.

3.3. Proximity and force/torque sensing

This section discusses two relatively underexplored sensory modalities for robotic perception of DLOs: proximity sensing and force/torque sensing.

3.3.1. Proximity sensing. Leveraging recent advancements in time-of-flight sensors, [Cirillo et al. \(2021a\)](#) present a proximity sensor enabling pre-touch sensing to improve depth accuracy in thin DLO perception and proposes a 3D scanning method to reconstruct DLO shapes as point clouds for grasping. However, this approach faces three main

limitations: it is restricted to uncluttered environments where DLOs are well separated; it requires the sensor to be positioned very close to the object (necessitating an initial rough estimate of the DLO's location); and it suffers from low reconstruction speed.

3.3.2. Force/torque sensing. For purely elastic DLOs, each equilibrium configuration of a Kirchhoff elastic rod corresponds to a unique point in a subset of \mathbb{R}^6 , fully characterized by the force and torque exerted at the base of the DLO ([Bretl and McCarthy, 2014](#)). Building on this foundational concept, several works ([Mishani and Sintov, 2021, 2023](#)) develop manipulation frameworks that rely exclusively on force/torque (F/T) sensing to estimate the shape of a DLO held by a dual-arm robotic system, achieving 3D shape estimation similar to that in [Section 3.1.3](#), but without relying on vision-based sensing.

Initially introduced in [Mishani and Sintov \(2021\)](#), the framework assumes a quasi-static DLO in a straight, undeformed configuration with high stiffness and inextensibility. Leveraging the mapping between F/T measurements and DLO configurations derived from [Bretl and McCarthy \(2014\)](#), a neural network is trained to predict the shape from sensor data. However, this method requires prior estimation of the DLO's mechanical properties and exhibits reduced accuracy when the underlying assumptions are violated. To address these challenges, [Mishani and Sintov \(2023\)](#) propose an enhanced approach using an autoencoder neural network that directly maps F/T sensor readings to DLO shapes, significantly improving both estimation accuracy and computational efficiency. A key limitation remains the necessity to retrain the model for each new DLO.

4. Estimation, control, and planning for DLO manipulation

This section covers three core components of DLO manipulation: *Estimation* ([Section 4.1](#)), *Control* ([Section 4.2](#)), and *Planning* ([Section 4.3](#)). An overview of their roles and interconnections is illustrated in [Figure 10](#).

4.1. DLO state and model estimation techniques

In DLO manipulation, estimation refers to the process of identifying physical or kinematic models, such as *deformation Jacobians* or *material parameters*, that capture the behavior of deformable objects. Reliable estimation improves the fidelity of DLO models for planning and control, and is critical for narrowing the gap between simulation and reality.

Discrepancies between simulated and real-world DLO behavior often result from idealized assumptions in physical modeling, such as neglecting dynamic effects, contact interactions, or non-homogeneous material properties. These

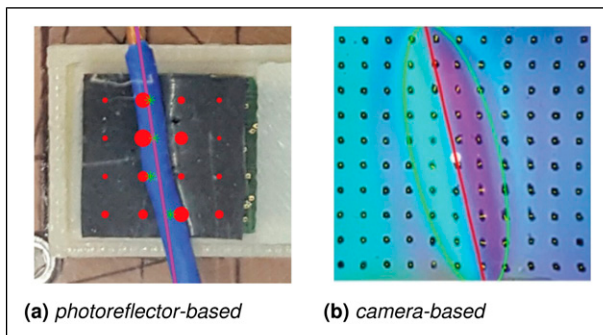


Figure 9. Photoreflector-based and camera-based tactile sensors. Images courtesy of (a) [Palli and Pirozzi \(2019\)](#) and (b) [Wilson et al. \(2023\)](#).

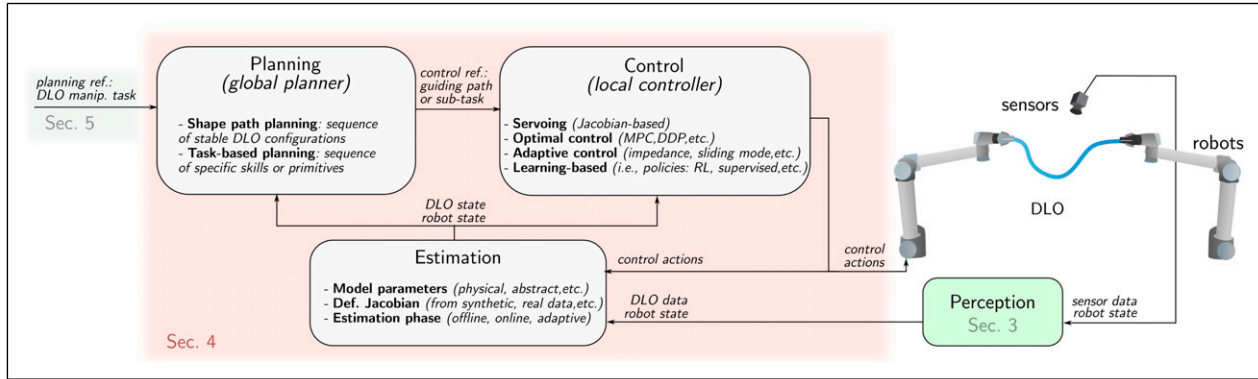


Figure 10. Overview of a representative DLO manipulation pipeline, highlighting the core components studied in this section: estimation, which infers physical properties or local models from data; control, which regulates the DLO’s state during interaction using feedback; and planning, which generates action sequences or desired DLO configurations to achieve task-level goals.

omissions lead to inaccurate predictions even when using high-resolution or computationally expensive models.

To improve alignment with real-world behavior, two estimation approaches are commonly used. The first involves identifying unknown physical parameters—such as stiffness, damping, or friction—to *calibrate* analytical models (Section 2). The second constructs local approximations of DLO behavior directly from data, for instance by estimating *deformation Jacobians* that map manipulator motions to object deformations.

Estimation strategies can be implemented *offline*, based on data obtained before the manipulation process; *online*, by continuously updating model estimates during task execution; or through *adaptive* schemes that selectively refine offline estimates based on observed discrepancies.

4.1.1. Model parameters. The models, described in Section 2, utilize several parameters to characterize the behavior of DLOs. These parameters may have a direct physical interpretation, such as length, mass, or bending stiffness, or they may serve primarily to adjust the model’s response under varying conditions, even if they are not directly measurable or lack an intuitive physical interpretation.

Certain parameters, such as length or mass, can be measured non-invasively and with relative ease. In contrast, material properties like Young’s modulus or damping coefficients require specialized and more invasive tests that are often impractical in robotic setups. Even with accurate identification, model approximations and measurement noise limit the prediction and simulation-to-reality fidelity of DLO behavior (Hermansson et al., 2016).

To improve model accuracy, a common strategy is to optimize model parameters to minimize the discrepancy between simulated deformations and reference shapes observed in the real world. Various optimization techniques have been explored for this purpose. Heuristic and gradient-free methods are commonly used, such as heuristic search (Lv et al., 2022; Tong et al., 2024), the cross-entropy method

(Yan et al., 2020), evolution strategies (Lim et al., 2022; Yang et al., 2022a; Zhang et al., 2024a), particle swarm optimization (Yu et al., 2025), and Bayesian optimization (Zhang et al., 2024a). In contrast, gradient-based approaches are used in works that exploit differentiable models or simulation environments (Caporali et al., 2024b; Liu et al., 2023).

Beyond parameter tuning, recent studies also emphasize the importance of action and trajectory design in the estimation process. For instance, Yu et al. (2025) proposed a specific trajectory in which the DLO shape is affected by twisting and gravity effects, enabling more accurate identification. Similarly, Zhang et al. (2024a) introduce action sequences aimed at maximizing the object’s displacement.

4.1.2. Deformation Jacobian estimation. As an alternative to the *pre-computed* models of Section 2, a widely adopted strategy is to estimate a *local deformation model*, often referred to as the *deformation Jacobian*. This model captures the relationship between local changes in actuation and changes in the DLOs state.

The *deformation Jacobian* can be derived either from *simulation data*, or through *real-world interactions* (Artinian et al., 2024; Navarro-Alarcon et al., 2013). In some cases, analytical expressions of the *deformation Jacobian* are available for specific formulations, such as the ARAP model (Shetab-Bushehri et al., 2023). While some methods estimate the *deformation Jacobian* purely from data, others assume a predefined structure on the matrix and estimate parameters within that structure. For example, Berenson (2013) and Wang et al. (2015) assume that the robot’s influence on the object decays exponentially with the distance between the end-effector and the manipulated point, and focus on estimating the decay rate.

In contrast to *pre-computed* models, *deformation Jacobians* are typically valid only within a narrow range of configurations. As such, they are most effective for local, small-scale deformation tasks and are not suited for long-

horizon predictions or transfer to different tasks (Yu et al., 2022). Despite this limitation, *deformation Jacobian* models are attractive due to their ability to capture object-specific behavior without requiring accurate physical modeling.

4.1.3. Modes of estimation: Offline, online, and adaptive. In many works, *model parameter* estimation is computationally intensive and typically performed *offline*, either before task execution (Lv et al., 2022; Zhang et al., 2024a) or during the generation of training datasets for learning-based methods (Yan et al., 2020; Yang et al., 2022a). An alternative is the method proposed in Caporali et al. (2024b), which performs online parameter estimation in parallel with task execution.

Similarly, *deformation Jacobians* can be estimated either *offline* or *online*. Online estimation strategies include incremental updates using the *Broyden update rule* (Navarro-Alarcon et al., 2013) or adaptive updates schemes such as those in Qi et al. (2021).

Beyond traditional estimation, several learning-based approaches address the sim-to-real gap through *online adaptation*. For instance, in Yu et al. (2022, 2025), a globally learned *deformation Jacobian* serves as a coarse approximation, which is then refined online via gradient-based updates using a sliding window of recent observations. Similarly, Wang et al. (2022) introduce a hybrid approach where an offline model is augmented with a local linear residual correction, computed online to enhance prediction accuracy. When source (e.g., simulation) and target (e.g., real-world) environments differ in specific regions, Mitrano et al. (2023) introduce a method that dynamically reweights dataset samples based on a learned similarity metric, enabling “targeted” model adaptation.

Assessing the reliability of a model and determining whether adaptation is required is another critical aspect of minimizing the sim-to-real gap. To this end, Mitrano et al. (2021) propose a learned classifier that predicts the reliability of a trained dynamics model. However, reliability assessment, fault detection, and recovery remain underexplored, as discussed in Section 6.2.

4.2. Control methods for DLO manipulation

In the context of DLO manipulation, control refers to the use of feedback information to regulate the object’s state during interaction. Control is critical for DLOs due to the difficulty of accurately modeling their highly nonlinear and typically underactuated behavior. It also becomes highly relevant to compensate for uncertainty and external disturbances during manipulation.

The aspects typically subject to control include the DLO’s shape, the position of relevant feature points, or the regulation of contact forces. This section provides an overview of the most commonly employed control

strategies in DLO manipulation. These control strategies will be framed within the task-oriented classification of methods in Section 5.

4.2.1. Servoing. This control paradigm uses the deformation Jacobian matrix (or interaction matrix, using servoing terminology) to relate control inputs directly to changes in the DLO state:

$$\dot{\mathbf{s}} = J(\mathbf{s}, \mathbf{u}) \approx \hat{\mathbf{J}}\mathbf{u}, \quad (5)$$

where $\mathbf{s} \in \mathbb{R}^N$ is the DLO state, $\mathbf{u} \in \mathbb{R}^M$ is the control input, and $J: \mathbb{R}^N \times \mathbb{R}^M \rightarrow \mathbb{R}^N$ is the deformation Jacobian, locally approximated by the matrix $\hat{\mathbf{J}} \in \mathbb{R}^{N \times M}$ through estimation methods (Section 4.1.2).

In the Jacobian-based control context, the DLO state \mathbf{s} may include geometric features such as keypoint positions (Artinian et al., 2024), visual representations such as image points (Aghajanzadeh et al., 2022b; Navarro-Alarcon et al., 2013), curves (Qi et al., 2023) and splines (Lagneau et al., 2020), or contours (Cuiral-Zueco et al., 2023; Zhu et al., 2021), among other representations.

A conventional feedback control law is given by

$$\mathbf{u} = \hat{\alpha} \hat{\mathbf{J}}^+ \mathbf{e} = \hat{\alpha} \hat{\mathbf{J}}^+ (\mathbf{s}_d - \mathbf{s}), \quad (6)$$

where $\mathbf{s}_d \in \mathbb{R}^N$ is the desired state, $\mathbf{e} = \mathbf{s}_d - \mathbf{s}$ is the state error, $\alpha > 0$ is a control gain, and $\hat{\mathbf{J}}^+$ denotes the Moore-Penrose pseudoinverse (or the true inverse if $N = M$, in either case $\hat{\mathbf{J}}$ needs to be full rank). Additional terms can be incorporated into (6), for example, collision avoidance as in Berenson (2013).

Since DLOs are typically underactuated ($N \gg M$), $\hat{\mathbf{J}}$ is a tall matrix. As a result, the pseudoinverse $\hat{\mathbf{J}}^+$ presents a non-trivial nullspace, meaning that controllers (6) cannot reduce errors $\mathbf{e} \in \ker(\hat{\mathbf{J}}^+)$. To mitigate these issues, the state \mathbf{s} is often defined as a projection of a larger state into a lower-dimensional space using techniques such as Principal Component Analysis (Zhu et al., 2021), or truncated Fourier descriptors (Zhu et al., 2018). Furthermore, numerical stability in equation (6) can be enhanced by applying regularization techniques to the Jacobian matrix, for example, using the Tikhonov regularization as in Zhu et al. (2021) and Cuiral-Zueco and López-Nicolás (2025). Under the common assumption that $\hat{\mathbf{J}}$ approximates J not too coarsely, the control system (6) achieves local asymptotic stability with exponential convergence.

4.2.2. Optimal control. It formulates the DLO manipulation task as a trajectory optimization problem, aiming to regulate the shape or the position of key features over a finite horizon while minimizing a predefined cost. This cost typically balances control effort with task-related objectives and may also penalize internal stress

(Aghajanzadeh et al., 2022c). Besides methods with quasi-static models (Aghajanzadeh et al., 2022c; Azad et al., 2023), second-order (dynamic) models have been employed in trajectory optimization methods, including Newton solvers and differential dynamic programming (DDP), both of which have been compared for optimal trajectory generation in Zimmermann et al. (2021).

Model predictive control (MPC) has gained popularity in the context of DLO manipulation, particularly when paired with learned dynamics models. For example, Wang et al. (2022) incorporate an offline-trained model into the MPC formulation as a constraint within a trust region. Similarly, Yang et al. (2022b) integrate online model learning with MPC, enabling continuous adaptation and model refinement during execution. Other MPC-based variants include Ma et al. (2022), where a graph-based MPC framework applied to a sparse set of learned keypoints is proposed, and Yu et al. (2023a), which introduces a single-step MPC controller used as a local feedback module within a broader planning framework. Serving as a tracker for the planning strategy in Yu et al. (2025), MPC control is employed along with an adaptive Jacobian model, allowing for collision avoidance and over-stretch constraints. As per approaches employing nonlinear MPC, Shen et al. (2025) apply proper orthogonal decomposition (POD) to reduce the dimension of a PDE–ODE model for a quadrotor with a hanging cable, enabling position and shape tracking.

4.2.3. Adaptive control. In this paradigm, the aim is to adjust controller parameters online to account for model uncertainty or variability in the object’s properties. In Qi et al. (2023), an adaptive controller is proposed for shape regulation based on B-spline and NURBS representations, allowing the system to account for dynamic changes in the DLO geometry. Chen et al. (2023) develop an adaptive impedance controller for clip fixing tasks, enabling stable interaction under varying contact conditions. In Aghajanzadeh et al. (2022a), the authors present an adaptive feature-space controller with formal Lyapunov stability guarantees, allowing feedback regulation of keypoints even under uncertain or time-varying dynamics. Another adaptive strategy is presented in Qi et al. (2021), which exploits sliding mode control with two variants, a linear and a finite-time method, with an online-updated adaptive term.

4.2.4. Learning-based control. These strategies are typically formulated in terms of a *policy*, a learned function or model that generates control signals to achieve a desired goal or to maximize a reward function within a given environment. These approaches are most commonly applied to *action selection* problems (Nair et al., 2017; Wang et al., 2019). The objective is to learn an *action policy* that, given current and goal observations of the system, produces actions that guide the system from its initial state toward the target state.

The observations (i.e., inputs) often come from raw sensory inputs like images (Nair et al., 2017; Wang et al., 2019), though more compact representations, such as DLO states, are also used (Zanella and Palli, 2021).

The output of the *policy* can take various forms depending on the control task. For instance, the predicted action may be a target position in Cartesian or image space (Nair et al., 2017; Wang et al., 2019; Zanella and Palli, 2021), or it may represent a velocity command (Daniel et al., 2024; Laezza and Karayiannidis, 2021).

Policy learning is often framed within standard reinforcement learning (RL) paradigms, where the agent interacts with the environment and improves based on reward signals (Daniel et al., 2024; Laezza and Karayiannidis, 2021; Zanella and Palli, 2021). Alternatively, *policies* can be learned through supervised learning approaches that leverage expert demonstrations or offline datasets (Nair et al., 2017; Seita et al., 2021; Wang et al., 2019).

4.3. Planning approaches in DLO manipulation

Planning approaches for DLO manipulation are broadly categorized into two main types, based on how manipulation is represented and structured: *Shape Path* and *Action-Based* planning. Though not mutually exclusive, methods involving both are classified by their main central component.

Shape Path Planning focuses on generating sequences of stable DLO configurations that connect an initial and a goal state, typically grounded in quasi-static models and supported by local controllers for trajectory tracking.

Action-Based Planning, on the other hand, formulates manipulation as a sequence of discrete task-oriented actions, also referred to as “skills” or “primitives,” such as grasping, sliding, or clipping. These are low-level, reusable, and parameterized actions that perform a specific movement or interaction. They can be sequenced or combined to achieve higher-level manipulation tasks.

4.3.1. Shape path planning. These approaches generate energy-minimized deformation trajectories for DLOs by exploring geometric or physical configuration spaces using sampling-based planners such as rapidly-exploring random tree (RRT) or bidirectional variants (BiRRT).

As one of the pioneering works, Moll and Kavraki (2006) introduced a planner that operates entirely within the space of minimal energy curves, that is, stable configurations under manipulation constraints. An adaptive representation and a local planner connect energy-minimizing states, producing smooth, physically plausible paths for applications such as routing or surgical suturing (Sections 5.3 and 5.5).

The Kirchhoff elastic rod model, detailed in Section 2.2, is the fundamental building block of several works (Roussel et al., 2015, 2019; Sintov et al., 2020; Wu et al., 2022a). It is used to sample valid configurations for DLOs. The static

equilibrium of a DLO is defined as an optimal control solution, considering the configuration space of a one-end-fixed Kirchhoff elastic rod as a six-dimensional manifold, which is suitable for using sampling-based planning algorithms (Bretl and McCarthy, 2014). However, the direct application of this formulation for sampling-based planning becomes computationally intensive, as highlighted by Sintov et al. (2020), and is limited to collision-free configurations. Roussel et al. (2015, 2019) extend the model with dynamic simulation, allowing contact interactions, including sliding, to traverse narrow passages. To avoid costly on-the-fly integration, Sintov et al. (2020) pre-compute a roadmap of elastic rod equilibrium shapes tightly coupled with robot joint configurations. A constrained BiRRT in this combined space enables rapid dual-arm manipulation planning without repeatedly solving differential equations. Wu et al. (2022a) combine a differentiable Kirchhoff rod model with a configuration distance descent strategy to iteratively guide the manipulated end along a predefined six-dimensional equilibrium manifold track, significantly improving convergence and success rates compared to sampling only or straight line approaches.

The Cosserat rod model is applied in Golestaneh et al. (2024) to represent multi-agent formations, formulating the planning task as a partial differential equation constrained optimal control problem, solved via nonlinear programming. Similarly, Azad et al. (2023) define minimal elastic energy trajectories by optimizing the Cosserat rod model to generate the commands required for desired deformations.

To improve computational efficiency, several works adopt simplified DLO models (Guo et al., 2020; Monguzzi et al., 2025; Yu et al., 2025). Guo et al. (2020) use a geometric spline representation combined with classical minimal-energy theory under quasi-static assumptions, enabling fast path planning in constrained environments. Yu et al. (2025) instead use a simplified discrete elastic rod model within a dual-arm framework, combined with a constrained BiRRT planner. Both Yu et al. (2023a) and Monguzzi et al. (2025) utilize a MSD model. The latter further incorporates clip constraints, particularly important for routing tasks involving fixed anchoring points. Learning-based models have also been applied to DLO planning to improve efficiency by approximating complex dynamics. For example, McConachie et al. (2020) propose planning in a reduced state space using a learned dynamics model from data obtained by simulation.

Simplified models inherently introduce approximations that may diverge from real DLO behavior, potentially resulting in shape inaccuracies or unforeseen collisions. Thus, in Yu et al. (2025), the resulting coarse paths guide a local model predictive controller (Section 4.2.2) to track deformation trajectories. Instead, Guo et al. (2022) propose a deviation-aware replanning strategy that monitors execution discrepancies, classifies their severity, and applies local corrections using potential fields. These corrections are then

smoothly merged back into the original plan using a time-decaying fusion policy, enhancing execution robustness. McConachie et al. (2020) use a learned classifier to determine the reliability of the approximated learned model in comparison to the real system. The role of the classifier is to guide the planner by discouraging actions whose approximation lacks reliability. While connected to the strategies discussed in Section 4.1.3, these approaches are specifically tailored for planning.

4.3.2. Action-based planning. The main idea is to decompose DLO manipulation into a sequence of discrete *primitive actions*, such as alignment, clipping, or Reidemeister moves (Section 5.4), and rely on simplified object models for planning.

In these settings, a hierarchical planning scheme is often adopted, where a high-level planner selects among available primitive actions, and a low-level controller executes motion to achieve the resulting sub-goals (Huo et al., 2022; Shah et al., 2018). Central to this hierarchy is the design of the high-level planner, which determines the sequence of actions based on the current task state.

Heuristic or rule-based strategies are frequently exploited. For example, a heuristic planner is deployed in Waltersson et al. (2022) trying to solve the DLO routing across several fixtures. When a plan fails, a genetic algorithm is employed to find a recovery sequence. The high-level plan is then translated into joint-level motions, while vision modules track the DLO and environmental features. Similarly, Shah et al. (2018) propose a planner that sequences clamp and grip actions to respect link constraints and place cables under gravity. Viswanath et al. (2021) approach unknotting with a graph-based planner that uses image-predicted keypoints to select actions that remove crossings and control slack.

The decision-making process is typically guided by visual observations, including visual input of the DLO state (Chen et al., 2023; Viswanath et al., 2021), fixture positions (Waltersson et al., 2022), fixture contact state (Huo et al., 2022), or fixture contact level indicators (Zhu et al., 2019).

Other recent advances explore learning-based approaches for high-level action selection. For instance, Luo et al. (2024) propose a deep neural policy that selects manipulation primitives based on multi-camera visual embeddings and the history of previously executed actions.

Some works have investigated alternative representations of the task space to significantly simplify and accelerate the planning process. For example, Keipour et al. (2022b) encode DLO configurations as sequences of convex subspaces via spatial decomposition, enabling planning using modified dynamic programming. In contrast, Jin et al. (2022) use a compact spatial vector encoding cable-fixture relations, enabling action selection via incremental state changes. Both methods leverage simplified, discrete representations to model DLO configurations, facilitating

efficient and generalizable planning without relying on exact geometric correspondence.

5. Manipulation tasks

This section builds upon the previous discussions on DLO modeling (Section 2), perception (Section 3), and estimation, control, and planning (Section 4), and categorizes the existing DLO manipulation literature based on the specific manipulation tasks addressed (see Figure 11). The primary tasks identified are: *Grasping* (Section 5.1), *Shaping and Deployment* (Section 5.2), *Routing and Threading* (Section 5.3), *Topological Manipulation* (Section 5.4), *Suturing* (Section 5.5), and *Transport* (Section 5.6). Each task-specific subsection not only reviews representative approaches but also highlights open challenges and gaps in the literature. A broader discussion of manipulation-related issues and future research directions is provided in Section 6.2.

5.1. Grasping

In the context of DLO manipulation, *grasping* refers to the process of establishing a controlled contact between a manipulator and the deformable object to constrain its motion and enable purposeful interaction. Unlike rigid object grasping, DLO grasping must account for compliance, shape variability, and the potential for deformation during contact, often requiring strategies that ensure stability without inducing unwanted strain or slippage. Although grasping is a fundamental component across diverse DLO manipulation tasks, it remains considerably underexplored, as most works assume predefined grasp points and stable contact conditions.

The problem of grasping can be decomposed into two related but distinct challenges: *where* to grasp, usually addressed as *grripper positioning* (Cuiral-Zueco et al., 2022);

and *how* to grasp, thus concerning grasp quality and stability (Roa and Suárez, 2015).

Gripper positioning for deformable objects focuses on identifying optimal grasp placements that facilitate related tasks such as shape control (Section 5.2) or topological manipulation (Section 5.4). For these tasks, parallel yaw grippers are widely used for both shaping (Yu et al., 2022), disentangling (Caporali et al., 2025; Lui and Saxena, 2013), or bin picking in cluttered environments (Zhang et al., 2022, 2024b, 2024c; Dirr et al., 2024).

Conversely, the grasp strategy, that is, *how to grasp*, is particularly critical in tasks such as routing (specifically DLO following, see Section 5.3), where grasp stability is more easily compromised during task execution. Despite parallel-jaw grippers still being widely deployed (She et al., 2021), dexterous hands are gaining traction due to the increased versatility (Yu et al., 2024).

Moreover, several transport-related tasks (Section 5.6) have explored alternatives to direct grasping. In particular, non-prehensile transport methods (see Section 5.6.2) leverage the compliant properties of DLOs to manipulate external objects through indirect interactions—such as dragging, wrapping, or tethering—without the need for rigid attachment (Zhi et al., 2024). These approaches are especially advantageous in environments where grasping is difficult, costly, or infeasible. A discussion on the role and potential of non-prehensile DLO manipulation is presented in Section 6.2.

5.2. Shaping and deployment

Shaping involves manipulating a DLO from an initial configuration to a desired target shape using one or more robotic manipulators (Cuiral-Zueco and López-Nicolás, 2024), as illustrated in Figure 12. A closely related task is

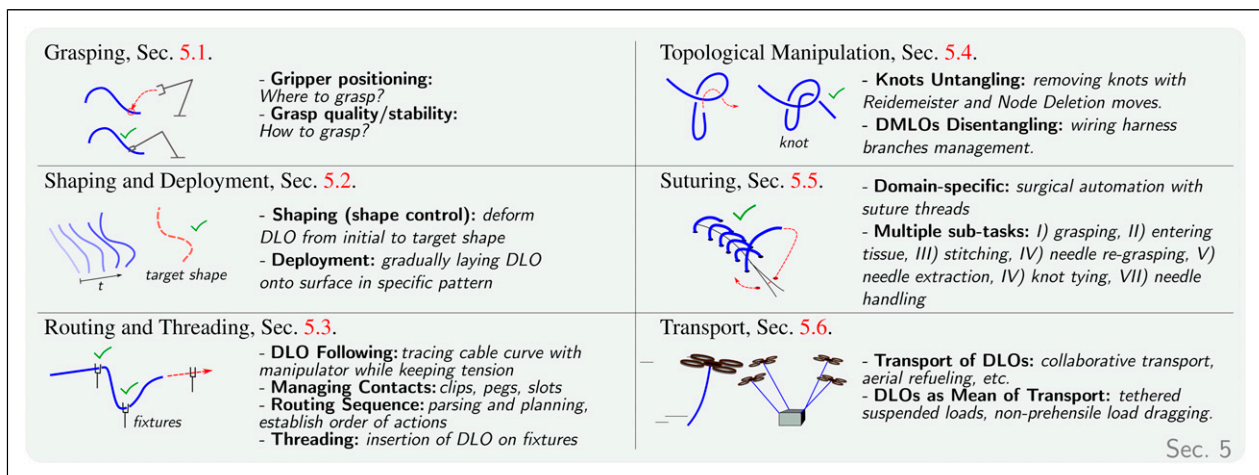


Figure 11. Overview of main DLO manipulation tasks identified in the literature: *Grasping* (Section 5.1), *Shaping and Deployment* (Section 5.2), *Routing and Threading* (Section 5.3), *Topological Manipulation* (Section 5.4), *Suturing* (Section 5.5), and *Transport* (Section 5.6). These categories build upon prior discussions of DLO modeling (Section 2), perception (Section 3), and estimation, control, and planning (Section 4).

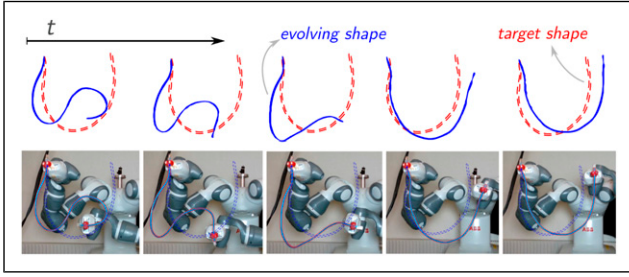


Figure 12. Shape control process where an ABB YuMi-IRB 14000 manipulates a blue Ethernet cable from an S-shape to a U-shape (experiment from Cuiral-Zueco et al. (2023)).

DLO deployment, which concerns the objective of gradually laying the DLO onto a surface following a specified shape or pattern (Lv et al., 2022; Tong et al., 2024).

Shaping and deployment tasks, following Cuiral-Zueco and López-Nicolás (2024), can be expressed in the standard form:

$$\min_{u(\cdot)} E(\mathcal{S}(t, u(t)), \Pi(\mathcal{S}_d)), \quad (7)$$

where $\mathcal{S}(t, u(t))$ represents the object’s shape under actions $u(t)$, \mathcal{S}_d is the desired target shape, map $\Pi(\mathcal{S}_d) = \mathcal{S}$ defines domain correspondences between \mathcal{S} and \mathcal{S}_d (e.g., feature matching), and $E(\cdot, \cdot)$ is a shape error metric (e.g., Procrustes distance, L_2 curvature norm, etc.). Problem (7) is considered solved when $E(\mathcal{S}(t, u(t)), \Pi(\mathcal{S}_d)) = 0$ (in the asymptotic sense, $\lim_{t \rightarrow \infty} E(\cdot, \cdot) = 0$), or when E reaches the minimum value that is feasible given the object’s physical constraints and system actuation capabilities. Most existing works assume the evolution of $\mathcal{S}(t, u(t))$ to be quasi-static, that is, $\mathcal{S}(t, u(t)) = \mathcal{S}(u(t))$ (see Section 2.2.2); accordingly, Section 5.2.1 focuses on quasi-static shape control and deployment methods. In contrast, some works consider highly dynamic (non-quasi-static) DLO systems, where the DLO’s shape, or specific parts such as a rope’s tip, are manipulated to achieve *casting*, thereby addressing *high-speed shape control*. Due to the distinct nature of these casting-based approaches compared to conventional quasi-static shape control, a dedicated *casting* section is provided in Section 5.2.2.

5.2.1. Quasi-static shaping and deployment. Table 4 provides a comprehensive overview of the main literature addressing DLO shaping tasks. Among the various aspects considered, a key distinction is made based on the behavior of the DLO, that is, how its shape changes in response to manipulation by the robotic arm. The DLO behavior can be broadly classified as follows³:

- **Elastic:** The DLO deforms under force but returns to its original shape when the force is removed (Yu et al., 2022) (e.g., stiff wires, plastic tubes).

- **Plastic:** The DLO retains permanent deformation after manipulation (Yan et al., 2020) (e.g., soft wires or ropes with low stiffness when manipulated on frictional surfaces).
- **Elastoplastic:** The DLO behaves elastically up to a yield point, after which it deforms plastically (Laezza and Karayiannidis, 2021). This hybrid behavior increases modeling complexity (e.g., electrical cables).

The analysis of Table 4 reveals several recurring patterns in the literature. While earlier works often relied on image-based inputs, recent approaches increasingly provide the DLO state directly, either by adopting simplifying assumptions about perception or by leveraging learning-based methods, as discussed in Section 3. Most studies focus on *elastic* DLOs, whereas DLOs exhibiting *plastic* responses were more prevalent in earlier methods. Notably, DLOs with *elastoplastic* behavior have been investigated only in Laezza and Karayiannidis (2021). An interesting object type is manipulated in Qi et al. (2021), where shape servoing of composite rigid-deformable objects—including uniform and joint-like connected DLOs—is performed through contour moments analysis. The majority of shaping tasks are carried out in *2D environments*, although a growing number of recent works address the complexities of *3D manipulation*. The exploitation of contact-rich interactions or environmental constraints remains relatively rare, with only a few works explicitly incorporating them into the shaping process (Huang et al., 2023; Huo et al., 2022; Tang et al., 2024). Human-in-the-loop shaping has been analyzed only in Zhou et al. (2024).

Action types generally fall into two main categories: *velocity-based* actions, typically used in servoing controllers, and *pick-and-place* strategies, which are often paired with learning-based models or policy-driven controllers. The *Exploited DLO Model* column in Table 4 refers to the specific representation of the DLO that informs the control policy, that is, capturing how the DLO is expected to respond to a given action. These models may include analytical physics-based formulations, learned approximations aimed at accelerating prediction and control such as multi-layer perceptrons (MLPs), graph neural networks (GNNs), and recurrent neural networks (RNNs), estimated deformation Jacobians (see Section 4.1.2), or hybrid combinations. The choice of the exploited DLO model is closely tied to the applied control strategy, which predominantly includes *servoing*, *sampling-* or *gradient-based MPC*, or *learned policies*. Importantly, in the case of learned policies, for example, those trained via reinforcement learning (RL) or behavior cloning (BC), actions are predicted directly from input observations without relying on an explicit model of DLO behavior. As such, these entries are left empty in the *Exploited DLO Model* column, as no internal representation is used during control. The same

Table 4. Summary of the surveyed literature on DLOs shaping. State estimation combined with online Jacobian modeling via least squares (LS) emerges as the predominant approach. Learning-based methods are most commonly integrated within short-horizon MPC frameworks. Elastoplastic behavior and contact-rich manipulation remain largely underexplored.

References	Perception input	DLO behavior	Env	Robot setup	Action type	Exploited DLO model	Control method	Type exp
Nair et al. (2017)	Image (RGB)	Plastic	2D	Single + fix	Pick position + displacement	–	Learned policy (self-supervised BC)	Real
Zhu et al. (2018)	State	Elastic	2D	Dual	Velocities	Jacobian (online)	Servoing	Real
Jin et al. (2019)	State	Elastic	2D	Dual	Velocities	Jacobian (robust LS, online)	Servoing	Real
Wang et al. (2019)	Image (grayscale)	Plastic	2D	Single	Pick position + displacement	–	Learned policy (self-supervised BC)	Sim, real
Lagneau et al. (2020)	State	Elastic	3D	Dual	Velocities	Jacobian (weighted LS, online)	Servoing	Real
Sundaresan et al. (2020)	Depth descriptor	Plastic	2D	Single	Pick-and-place	–	Greedy heuristic geometric policy	Real
Yan et al. (2020)	State	Plastic	2D	Single	Pick-and-place	Learned bi-LSTM	MPC (sampling-based)	Sim, real
Laezza and Karayiannidis (2021)	State	Elastoplastic	3D	Dual	Velocity + hinge/lock constraint	–	Learned policy (RL)	Sim
Lee et al. (2021)	Image (binary)	Plastic	2D	Single	Pick-and-place (image-space)	Learned image-space predictive model	Cost function minimization	Sim, real
Khalifa and Palli (2022)	State	Plastic	2D	Single	Pick-and-place	Dynamic splines	Cost function minimization	Sim
Seita et al. (2021)	Image (RGB)	Plastic	2D	Single	Pick-and-place	–	Learned policy (supervised BC)	Sim
Yang et al. (2021)	State	Elastic	3D	Single + fix	Displacements	Learned bi-LSTM + interaction network	MPC (gradient-based)	Sim, real
Zanella and Palli (2021)	State	Plastic	2D	Single	Pick-and-place	–	Learned policy (RL)	Real
Zhang et al. (2021b)	Image (binary)	Plastic	2D	Single	Pick position + displacement	Learned model	MPC (sampling-based)	Real
Zhu et al. (2021)	State	Elastic	2D	Single + fix	Displacements	Jacobian (online receding horizon)	Servoing	Sim, real
Aghajanzadeh et al. (2022c)	State	Elastic	2D	Single + fix	Velocities	Jacobian (ARAP, online)	Optimal control law	Sim, real
Aghajanzadeh et al. (2022b)	Keypoints	Elastic	2D	Single + fix	Velocities	Jacobian (ASAP, offline)	Servoing	Sim, real
Aghajanzadeh et al. (2022a)	Keypoints	Elastic	2D	Single + fix	Velocities	–	Adaptive servoing	Sim, real

(continued)

Table 4. (continued)

References	Perception input	DLO behavior	Env	Robot setup	Action type	Exploited DLO model	Control method	Type exp
Huo et al. (2022)	State	Plastic	2D	Dual + contacts	Pick-and-place	–	Heuristic motion primitives	Real
Lv et al. (2022)	State	Elastic	3D	Single/dual	Velocities	Discrete elastic rod	Optimization-based controller (deployment task)	Sim, real
Ma et al. (2022)	Keypoints	Plastic	2D	Single	Pick-and-place	Learned GNN + RNN	MPC (sampling-based with learned reward)	Sim, real
Yu et al. (2022)	State	Elastic	3D	Dual	Velocities	Learned GNN (online adaptation)	Optimization-based adaptive controller	Sim, real
Zakaria et al. (2022)	State	Elastic	3D	Single + fix	Linear velocities	–	Learned policy (RL)	Sim
Wang et al. (2022)	State	Elastic	2D	Dual	Linear velocities	Learned GNN (online adaptation)	MPC (gradient-based)	Sim, real
Daniel et al. (2024)	State	Elastic	3D	Single + fix	Velocities	–	Learned policy (RL)	Sim, real
Huang et al. (2023)	State	Elastic	2D	Dual + contacts	Pick-and-place	Learned GNN	MPC (gradient-based)	Sim, real
Qi et al. (2023)	State	Elastic	2D	Single + fix	Velocities	Jacobian (UKF, online)	Servoing (with gradient-based gains optimization)	Real
Shetab-Bushehri et al. (2023)	State	Elastic	3D	Dual	Velocities	Jacobian (analytical ARAP, online)	Servoing	Real
Tong et al. (2024)	State	Elastic	3D	Single	Positions	Learned MLP	Optimization-based controller (deployment task)	Real
Artinian et al. (2024)	State	Elastic	3D	Dual	Velocities	Jacobian (analytical Cosserat, online)	Servoing	Real
Caporali et al. (2024b)	State	Plastic	2D	Singe	Pick-and-place	Learned MLP	Cost function minimization (gradient-based)	Real
Szymko et al. (2024)	State	Elastic	3D	Dual	3D velocities	Jacobian (recursive LS, online)	Servoing	Real
Tang et al. (2024)	State	Elastic	3D	Dual + constraints	Velocities	Learned MLP + Jacobian (online)	MPC (gradient-based with post-process safety filter)	Sim, real
Zhou et al. (2024)	State	Elastic	3D	Single + human	Linear velocities	Jacobian (latent derivation, online)	Servoing (with sliding mode control)	Real
Zhang et al. (2024a)	Pointcloud	Plastic	2D	Single	Pusher positions	Learned GNN	MPC (sampling-based)	Sim, real
Gu et al. (2025)	State	Elastic	2D	Single + fix	Velocities	Learned GNN	MPC (gradient-based)	Sim, real

applies to approaches using heuristic or rule-based motion primitives that do not explicitly incorporate any predictive model of the DLO response. This terminology aligns with the standard convention in the field, where a manipulation method is considered model-based only if it explicitly learns or utilizes a *model* (defined as in Section 2) to determine the manipulation actions.

Earlier works often favored RL and BC techniques, while the use of learned approximations of analytical models (such as neural networks trained to mimic physical dynamics) is becoming increasingly popular in recent approaches. While early studies primarily focused on simulation-only evaluations, a growing number of recent works demonstrate and validate their approaches in both *simulation and real-world scenarios*.

A strong correlation emerges between *plastic* material response, 2D environments, single-arm robot setups, pick-and-place actions, and learned models, characteristic of shaping tasks where the DLO is manipulated over a supporting surface. Conversely, the combination of *elastic* DLOs, dual-arm setups, and velocity-based actions is commonly associated with servoing tasks, which are prevalent in both 2D and 3D environments.

5.2.2. Casting and high-speed shaping. Dynamic manipulation of DLOs involves applying fast⁴, time-dependent motions to produce complex behaviors. This contrasts with quasi-static manipulation. These motions are typically executed open loop once generated.

Typical tasks in this domain include *whipping* (Chi et al., 2024a; Lim et al., 2022; Zimmermann et al., 2021), which generally involve free-end DLOs, and *vaulting, knocking, and weaving* (Zhang et al., 2021a), which are performed with fixed-end cables. *Whipping* (often addressed also as *casting*) is a dynamic manipulation task where a robot quickly moves one end of a DLO to generate high-speed motion that travels along it, using the object’s elasticity and inertia to control the free end and reach targets beyond the robot’s immediate reach. The latter tasks involve dynamically manipulating a DLO to (1) vault over an obstacle, (2) knock an object off an obstacle, and (3) weave the DLO between multiple obstacles.

In terms of modeling, an algebraic deformation model that assumes negligible gravity effects due to fast motion and models the DLO as a series of joints following the robot with a constant delay is proposed in Yamakawa et al. (2013). On the other hand, Zimmermann et al. (2021) study the dynamic manipulation of free-end beams using FEM models combined with optimal control techniques for trajectory optimization.

Given the complexity of dynamic DLO manipulation, learning-based approaches are commonly applied, often relying on low-dimensional, parameterized action spaces to simplify control and training: two sweeping arcs (Lim et al., 2022), apex point (Zhang et al., 2021a), and two joint angles

(Chi et al., 2024a). Maximum velocity is also treated as a learnable parameter, and the task space is often constrained to 2D (Chi et al., 2024a; Lim et al., 2022). Resetting the object’s initial state before each action ensures consistency and repeatability (Lim et al., 2022; Zhang et al., 2021a).

Across all approaches, simulation plays a critical role, either for bridging the sim-to-real gap (Lim et al., 2022) or for bootstrapping the learning process (Zhang et al., 2021a). Iterative refinement strategies enable easy online adaptation to system changes (Chi et al., 2024a).

5.2.3. Existing literature gaps. Research on DLO shaping is predominantly confined to well-established setups, such as pick-and-place manipulation of plastic DLOs on planar surfaces or dual-arm manipulation of elastic DLOs. These scenarios often assume simplified environments without obstacles or complex interactions. However, in real-world applications (and in human manipulation), contact-rich interactions and environmental constraints play a crucial role, especially given the underactuated nature of DLO shape control (Huang et al., 2023; Huo et al., 2022).

A second gap lies in the predominant elastoplastic behavior of real-world DLOs, such as electrical cables, which received limited attention in existing research (Laezza and Karayiannidis, 2021).

Many existing approaches assume the desired target shape is reachable and valid without explicitly verifying these conditions, often relying on heuristic checks instead. This highlights a significant gap in the current research, where formal reachability analysis is lacking.

5.3. Routing and threading

Routing involves systematically arranging DLOs to conform to a target configuration while establishing contact with environment objects. A key aspect of routing strategies is the use of *fixtures* (e.g., clips or jigs), which anchor the manipulated DLO in place, and *contact or pivoting points*, which facilitate tension control and enable smooth directional changes of the DLO along the routing path. The key elements involved in the routing process are illustrated in Figure 13. Threading is a related sub-task frequently encountered in practical applications, such as threading a needle or inserting wires into industrial assemblies. It involves guiding the DLO through a designated *hole* or *eyelet* in the environment, typically demanding higher precision despite the intrinsic DLO flexibility.

Unlike the more general shaping task of Section 5.2, routing also requires managing the sliding motion of the DLO along the gripper fingers (as discussed in Section 5.1), as well as executing precise insertions into fixing points/holes and interacting with pivoting elements along the path.

The current literature on DLO routing and threading is summarized in Table 5, which categorizes existing approaches based on sensing modalities, DLO characteristics,

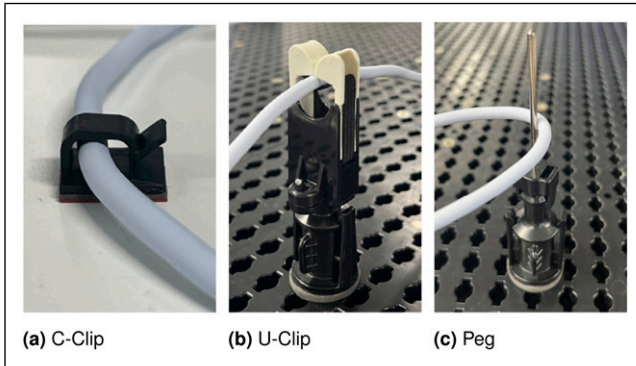


Figure 13. Examples of routing elements. Images courtesy of Chen et al. (2023).

robotic configurations, and primary task. Most methods follow a form of action-based planning as introduced in Section 4.3.2, typically structured around the scripting or learning of reusable motion primitives, for example, *clipping* or *insertion*. Despite the diversity of approaches, four predominant sub-tasks can be identified and are analyzed in the following subsections: *DLO following* (Section 5.3.1), *managing contacts* (Section 5.3.2), *routing sequence parsing and planning* (Section 5.3.3) and *threading* (Section 5.3.4).

Regarding benchmarks (see connected discussion in Section 6.2), the NIST Assembly Task Board (NIST, 2025) has emerged as a standardized setup for evaluating robotic capabilities in routing tasks. Although it represents a simplified scenario, it is increasingly adopted in research to support reproducibility and comparative evaluation, as in Keipour et al. (2022b) and Zhang et al. (2024d). Beyond conventional routing tasks, the NIST board has also been employed to study more complex manipulation settings (Luo et al., 2025).

5.3.1. DLO following. It involves grasping one end of the cable and manipulating the gripper to trace its contour while maintaining tension by securing the opposite end, either with a fixture or a second grasp.

In this sub-task, tactile sensing (see Section 3.2) is commonly used (e.g., She et al. (2021); Hellman et al. (2017); Pecyna et al. (2022)) as it provides localized feedback that enables dynamic grip adjustments and precise alignment during sliding. Among tactile sensors, high-resolution, image-based sensors like GelSight (Yuan et al., 2017) are used by both She et al. (2021) and Wilson et al. (2023), offering detailed measurements of normal force, shear, and torque. Alternative approaches include capacitive-based sensors (Monguzzi et al., 2023;

Table 5. Summary of the main literature on DLO routing and threading. The table highlights a strong reliance on tactile/force sensors for local execution tasks (Following, Contacts) and vision for global Planning. Notably, the field is heavily skewed toward 2D planar setups, leaving full 3D routing and dynamic harnessing relatively underexplored.

References	External sensor	DLO behavior	Env	Robot setup	Task
Huang et al. (2015)	Vision	Plastic	2D	Custom	Threading
Wang et al. (2015)	–	Elastic	3D	Single	Threading
Hellman et al. (2017)	Tactile	Plastic	2D	Single + fix	Following
De Gregorio et al. (2018b)	Vision, tactile	Elastic	3D	Single	Threading
Zanella et al. (2019)	Tactile	Elastic	3D	Single	Threading
Zhu et al. (2019)	Vision	Plastic	2D	Dual	Contacts
Galassi and Palli (2021)	Tactile	Plastic	2D	Single + fix	Following, contacts, planning
She et al. (2021)	Tactile	Plastic	2D	Single + fix	Following
Keipour et al. (2022b)	Vision	Plastic	2D	Single	Planning
Jin et al. (2022)	Vision	Plastic	2D	Single + fix	Planning
Pecyna et al. (2022)	Vision, tactile	Plastic	2D	Single + fix	Following
Süberkrüb et al. (2022)	Force	Plastic	2D	Single, dual	Following, contacts
Chen et al. (2023)	Vision, force	Plastic	2D	Dual	Contacts, planning
Monguzzi et al. (2023)	Tactile	Elastic	3D	Single + fix	Following, contacts
Yu et al. (2023b)	Tactile	Elastic	2D	Single	Threading
Wilson et al. (2023)	Vision, tactile	Plastic	2D	Single + fix	Following, contacts, planning
Chen et al. (2024)	Force	Plastic	2D	Dual	Contacts
Li and Choi (2024)	Vision	Elastic	2D	Single	Threading
Luo et al. (2024)	Vision	Plastic	2D	Single + fix	Contacts, planning
Monguzzi et al. (2024a)	Tactile	Elastic	2D	Single + fix	Following
Monguzzi et al. (2024b)	–	Elastic	2D	Single + fix	Following
Yu et al. (2024)	Tactile	Plastic	2D	Single	Following
Zhang et al. (2024d)	Force	Plastic	2D	Dual	Following, contacts
Li et al. (2025)	Vision	Elastic	2D	Dual	Threading

2024a) and optoelectronic sensors (Galassi and Palli, 2021), which trade spatial resolution for higher update rates, supporting faster control loops. Galassi and Palli (2021) combine tactile sensors with force/torque sensing, Zhang et al. (2024d) use only force sensing, while Monguzzi et al. (2024b) explore the use of internal joint torque signals as a form of proprioceptive feedback. Instead, vision-based sensing is generally avoided.

Motion primitives to trace along the DLO are often either scripted (using predefined actions such as *slide*, *grasp*, or *reorient*, e.g., Süberkrüb et al. (2022)) or learned (RL policies, e.g., Pecyna et al. (2022); Hellman et al. (2017)). Monguzzi et al. (2023) propose tactile-driven skills such as local diameter estimation, 3D alignment, and adaptive sliding based on local predictions of the DLO shape. This approach is expanded in Monguzzi et al. (2024a) by considering collisions and a global instead of local DLO shape. In contrast, Monguzzi et al. (2024b) eliminate the need for explicit contact sensing by leveraging a compliant last robot joint while following an estimated local DLO shape. RL methods (e.g., Pecyna et al. (2022); Hellman et al. (2017)) show promising results for developing adaptive policies, particularly when leveraging multi-modal sensory feedback. Unlike the general approaches, a learned dynamic model of the wire behavior under force feedback is combined with an MPC strategy in Zhang et al. (2024d).

Specifically in terms of low-level tactile-driven control strategies, the task is frequently divided into DLO pose and grip controllers, as in She et al. (2021). These controllers jointly regulate the gripper’s position and applied force relative to the DLO, ensuring smooth path following while preventing slippage or buckling. A key simplification of She et al. (2021) is the horizontal gripper orientation, which removes gravity effects. In contrast, Galassi and Palli (2021) employ a tactile-based correction within a vertical grasp, while Yu et al. (2024) introduce a “V-shaped” grasping strategy using a robotic hand. A simple threshold-based method is used to maintain the DLO centered during sliding in Wilson et al. (2023).

5.3.2. Managing contacts. *Clips*, *pegs*, and *slots* play a crucial role in routing tasks. To handle these, specific motion primitives are typically designed to account for contact interactions. Most of these primitives are scripted or heuristic in nature. Luo et al. (2024) propose a learned slot-insertion policy. Importantly, these primitives need to be orchestrated by a planner, see Section 4.3.2. The choice of motion strategy and sensing modality often depends on the type of contact object. Manipulating clips generally requires the DLO to be tensioned during insertion (Galassi and Palli, 2021; Zhang et al., 2024d), whereas pegs impose a less stringent requirement, and slots typically do not require tensioning at all. Below is a summarized overview of motion strategies and sensing for each contact type (see also Figure 13):

- *Pegs* are addressed in Zhu et al. (2019) through a vision-based angular contact mobility index that quantifies DLO–obstacle interactions, or via scripted primitives that incorporate tactile sensing (Galassi and Palli, 2021; Wilson et al., 2023).
- Routing through *slots* is explored in Luo et al. (2024), where an imitation learning framework is used to train a slot-insertion policy from visual feedback collected via multiple cameras. Additionally, Wilson et al. (2023) propose a heuristic “weaving slot” primitive featuring a wiggling motion guided by tactile feedback.
- For *clips*, force sensing is commonly used to monitor the DLO state in conjunction with scripted motion primitives (Galassi and Palli, 2021; Süberkrüb et al., 2022; Zhang et al., 2024d). In Chen et al. (2023), a dedicated clipping primitive is developed using threshold-based control on force signals to regulate motion during execution. Building on this, Chen et al. (2024) propose an enhanced set of contact indicators to improve detection accuracy.

Notably, Süberkrüb et al. (2022) also introduce a feature point estimation algorithm based on Kalman filtering, which identifies fixture locations, such as clips or jigs, by fusing multiple observations of force readings from a tensioned DLO.

5.3.3. Routing sequence parsing and planning. The placement of fixtures (i.e., determining where they need to be positioned) and their routing sequence are important aspects of the task. High-level task parsing is addressed in Wilson et al. (2023), exploiting a simple processing of visual observations. Chen et al. (2023) tackle fixture placement by optimizing their positions based on a target DLO shape. The approach minimizes deviation from the desired shape while maintaining adequate spacing between consecutive fixtures to ensure smooth execution by the dual-arm robot setup.

Regarding the representation of the routing problem, both Keipour et al. (2022b) and Jin et al. (2022) focus on modeling the environment to support planning. Keipour et al. (2022b) employ a convex decomposition of space to encode the DLO configuration, simplifying the planning process. In contrast, Jin et al. (2022) model the relative spatial relationships between DLOs and fixtures to facilitate efficient data collection and learning of three motion primitives for routing.

5.3.4. Threading. It requires precise manipulation of the DLO and accurate localization of both its tip and the target hole or eyelet.

To localize the DLO tip (or “tail-end”), some authors have proposed vision-based strategies (De Gregorio et al., 2018b; Wang et al., 2015), while others have leveraged tactile sensing by following the DLO shape (Yu et al.,

2023b), similar to the approach described in Section 5.3.1. The location of the target hole or eyelet is typically assumed to be known in advance or determined using fiducial markers (Li and Choi, 2024). Yu et al. (2023b) leverage the same (camera-based) tactile sensor to localize the needle eyelet by employing a pre-trained (image-based) foundation model (see Section 6.1).

Grasping plays a crucial role in this task, as sufficient “slack” between the grasp point and the tip is necessary for successful insertion. Li and Choi (2024) parametrize the grasp point based on the DLO flexibility. When threading requires pulling the DLO further through the hole, re-grasping is typically employed (Li et al., 2025; Wang et al., 2015).

The actual insertion is performed via quite diverse strategies. In Wang et al. (2015), a diminishing rigidity Jacobian (Section 4.1.2) is used in combination with a virtual vector field for tip guidance. Zanella et al. (2019) explicitly address the deformability of DLOs by using a data-driven controller that adjusts insertion orientation in real time based on estimated force components from tactile feedback. Similarly, De Gregorio et al. (2018b) use tactile sensors with a data-driven regressor to evaluate grasp quality and predict insertion collisions.

RL has also been applied in recent work, although in simplified scenarios. In Yu et al. (2023b), a goal-conditioned tactile-driven policy that learns to output low-dimensional end-effector displacements from segmented tactile observations is proposed. However, this method relies on the eyelet being in continuous contact with the tactile surface, limiting its real practicality. In Li and Choi (2024), a policy conditioned on the DLO flexibility is used to produce two spatial waypoints for insertion. In a follow-up work, Li et al. (2025) leverage RL agents producing expert demonstrations to train a diffusion policy capable of both insertion and pulling. Both approaches are restricted to simplified 2D environments.

Unlike the previously discussed approaches, high-speed manipulation is used to generate centrifugal force during thread rotation, effectively transforming the task into a simplified peg-in-hole insertion (Huang et al., 2015). However, the method is limited to 2D and requires a custom robotic setup.

5.3.5. Existing literature gaps. Many existing approaches rely on hand-crafted motion primitives or models trained on specific types of DLOs (Wilson et al., 2023; Zhang et al., 2024d). This specialization may limit generalization, particularly when handling DLOs with varying diameter or stiffness properties.

Manipulation scenarios are frequently simplified through assumptions of quasi-static dynamics or restriction to planar environments (Jin et al., 2022; Keipour et al., 2022b; Yu et al., 2023b). These abstractions omit critical aspects of real-world 3D routing tasks, including complex spatial

configurations, occlusions, and obstacle interactions. In contrast, Luo et al. (2025) move toward more realistic routing scenarios by addressing a challenging belt-threading and tensioning task involving a *closed* DLO, in which a reinforcement-learning-derived policy enables coordinated dual-arm task execution.

In perception, visual and tactile sensing are often deployed in isolated stages rather than being fused and utilized concurrently in real time (Wilson et al., 2023), which limits responsiveness and robustness in scenarios that may benefit from multi-modal sensing (Pecyna et al., 2022).

For contact-rich tasks involving *clips*, current strategies typically employ simplified geometries such as circular holes or loosely constrained channels (Galassi and Palli, 2021; Süberkrüb et al., 2022; Wilson et al., 2023; Zhang et al., 2024d), which generate minimal contact forces. However, more realistic *clip* designs, such as those capable of elastic deformation during insertion, introduce complex contact dynamics that remain underexplored (Chen et al., 2023).

Finally, repeated mechanical interactions such as tensioning or insertion of the DLO can induce material fatigue and degradation over time (Zanella et al., 2019; Zhang et al., 2024d). This raises concerns about long-term reliability and operational safety in real-world deployments, which are not addressed in current literature.

5.4. Topological manipulation

Topological manipulation focuses on the challenging task of untangling knots formed by one or more DLOs as well as disentangling branches of DMLOs. This task presents a complex challenge at the intersection of perception, planning, and manipulation. The key difficulty lies in perceiving and representing the knot or tangle structure, then determining a sequence of actions to simplify and ultimately untangle it.

Table 6 summarizes key literature in robotic topological manipulation, divided between knot untangling in DLOs (top part, see Section 5.4.1) and *DMLO* disentanglement (bottom part, see Section 5.4.2).

5.4.1. Knots untangling. Most solutions proposed in the literature to automatically unknot DLOs draw inspiration from knot theory, a branch of topology concerned with the mathematical properties of knots. Two key concepts commonly employed are the *Dowker–Thistlethwaite* (DT) notation and *Reidemeister* moves (Lui and Saxena, 2013). DT notation encodes knots as a linear sequence of integers, providing a compact, symbolic representation of crossings. In contrast, *Reidemeister* moves define three elementary primitives to remove intersections in the structure. Together, these tools form the basis of many state estimation and manipulation strategies in robotic knot untangling.

Early approaches typically assumed semi-planar, loosely tangled knots with visible ends, as in Lui and Saxena (2013),

Table 6. Summary of the key literature on robotic topological manipulation, with knot untying in DLOs (top) and DMLOs disentangling (bottom). The table highlights each work’s key perception and manipulation contributions. Generally, single-DLO untying relies on graph-based topological abstractions and geometric moves (e.g., Reidemeister moves), whereas multi-object disentangling (DMLOs) additionally requires interactive perception and dynamic agitation strategies to overcome severe occlusion.

Reference	Robot setup	# DLOs/ DMLOs	Perception/state representation method	Manipulation strategy
Lui and Saxena (2013)	Dual-arm (PR2)	1	Linear graph with DT notation (semi-planar)	Score-based (Reidemeister moves, node deletion move)
Grannen et al. (2021)	Dual-arm (dVRK)	1	Learned knot detector + keypoint regressor	Sequential node deletion move + Reidemeister move
Sundaresan et al. (2021)	Dual-arm (dVRK)	1	Linear graph (non-planar extension)	Recovery moves (wedged, recentering)
Viswanath et al. (2021)	Dual-arm (dVRK)	> 1	–	Cable extraction move
Shivakumar et al. (2023)	Dual-arm (ABB YuMi)	1	Interactive perception (partial observability)	Interactive untying moves
Huang et al. (2024)	Dual-arm (KUKA LBR)	≫ 1	Learned topological representation (partial observability)	Reidemeister + multi-cable moves
Zhang et al. (2022)	Single-arm (Nextage)	≫ 1	Learned grasp policy (bin picking)	Helix + spinning moves
Zhang et al. (2024c)	Dual-arm (Nextage)	≫ 1	–	Swing + re-grasping strategy
Caporali et al. (2025)	Dual-arm (Franka Emika)	1	GNN-based topological embedding	Circular move

where the *Node Deletion* move is introduced. This move involves pulling a DLO out of an under-crossing intersection and is used in conjunction with classical *Reidemeister* moves to simplify tangled configurations (see Figure 14). The authors also propose a sufficient condition for entanglement, based on crossing transitions along the rope.

Building on Lui and Saxena (2013), later works refine several key aspects. Grannen et al. (2021) propose a sequential manipulation strategy that combines *Node Deletion* and *Reidemeister* moves, achieving a monotonic reduction in the number of crossings until the rope is untangled. Avoiding an explicit topological representation (e.g., linear graphs with DT notation as in Lui and Saxena (2013)), a learned perception system composed of a knot detector and a keypoint regressor is also proposed.

Further advancements are proposed by Sundaresan et al. (2021), which extends the linear graph representation of Lui and Saxena (2013) to handle non-planar configurations, enabling a generalization to more complex 3D entanglements. Additionally, they improve upon Grannen et al.

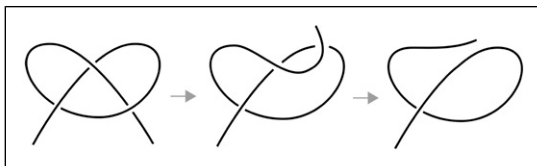


Figure 14. Process of untying a DLO using a *node deletion* followed by a *Reidemeister* move.

(2021) by introducing a coarse-to-fine refinement strategy for keypoint predictions, significantly reducing grasping errors due to near misses, and a set of *Recovery* moves. Viswanath et al. (2021) further extend this work by considering multiple DLOs in the scene (“intra-cable” and “inter-cable” crossings) and introducing a new manipulation primitive: the *Cable Extraction* move. These works focus on *dense* knots as opposed to the loose ones of Lui and Saxena (2013), thanks to the capabilities of the employed robotic setup (see Table 6).

Recent works have extended topological manipulation to partially observable scenes, relaxing the assumption that DLO ends must be visible. In Huang et al. (2024), *Reidemeister* moves are combined with a novel *Multi-Cable* move to handle loosely entangled knots, similar to the setup in Lui and Saxena (2013). The focus is on learning a robust topological state representation of multiple DLOs using deep learning. In contrast, Shivakumar et al. (2023) address the single-DLO case and introduce interactive perception through modified manipulation primitives that actively explore the tangled DLO’s configuration.

With a quite diverse approach, Yamakawa et al. (2013) explore high-speed dynamic knot tying, where knots are formed by rapidly whipping cables upward and leveraging self-collisions of the DLO to achieve the final tie.

5.4.2. DMLOs disentangling. A practical application of DLOs disentangling is exemplified by the challenge of manipulating DMLO, for example, spreading the wiring

harness branches free of intersections (Caporali et al., 2025), or extracting a wiring harness from a bin (Zhang et al., 2022).

The complex structure of a wire harness introduces additional complexities in both robotic bin picking and untangling, given its multi-branched configuration and the coexistence of deformable and rigid components (e.g., connectors and clips).

For DMLOs bin picking, Zhang et al. (2022) propose two distinct motion primitives for disentangling (i.e., a *Helix* and *Spinning* moves) combined with a learned perception system that leverages active learning to predict grasp points and estimate grasp success probabilities. Building on this, Zhang et al. (2024c) introduce a dual-arm closed-loop framework that enhances system robustness and accuracy.

To remove intersections among DMLO branches, Caporali et al. (2025) propose a topological representation constructed using a GNN from a single image of the scene. Based on the extracted topology, a dual-arm manipulation primitive is executed using a circular motion strategy that satisfies the topological constraints of the manipulated DMLO.

5.4.3. Existing literature gaps. Current methods predominantly rely on passive visual perception. This approach often struggles with tightly tangled knots or complex DMLOs entanglements. In contrast, active perception strategies remain underexplored and have so far been applied to single DLOs with loosely tied, simple knots (Shivakumar et al., 2023). Learning-based perception strategies are also usually trained on synthetic or specific real-world data over simplified scenarios (Grannen et al., 2021; Huang et al., 2024), making generalization to real-world scenarios hard.

Grasping and manipulation in dense, high-friction DLO configurations remain challenging. Common failure modes include missed or unintended multi-object grasps, as well as slippage during manipulation (Shivakumar et al., 2023; Zhang et al., 2022). Tactile sensors, which could potentially verify grasp quality and monitor the manipulation process in real time, are currently not exploited.

5.5. Suturing

Suturing is a highly complex and domain-specific task within surgical automation that integrates several DLO-related *sub-tasks*, such as routing, threading, and knot tying (Sections 5.3 and 5.4). However, its uniquely constrained clinical environment, specialized robotic setups and DLO characteristics (i.e., suture *thread*), and strict procedural demands distinguish it from more general DLO manipulation scenarios, as shown in Figure 15. These distinct characteristics motivate treating suturing as a dedicated manipulation task.

The suturing process can be broken down into seven steps, paraphrased from Pedram et al. (2021): (I) grasping

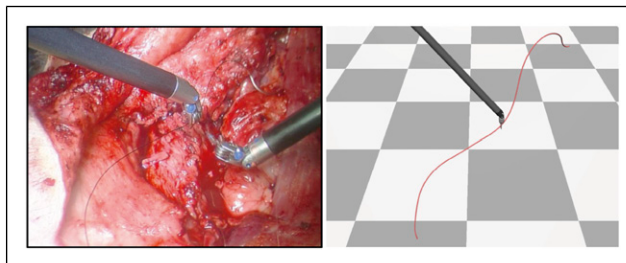


Figure 15. Suturing task with dVRK and reconstructed 3D suture thread. Images courtesy of Joglekar et al. (2023).

the needle with the inserting arm, (II) moving toward the wound and entering the tissue perpendicularly, (III) stitching, (IV) grasping the needle with the extracting arm, (V) extracting the needle, (VI) knot tying and grasping the needle with the extracting arm, and (VII) handing off the needle to the inserting arm.

Existing approaches typically focus on specific subsets of these steps. For example, Pedram et al. (2021) cover steps II to V and VII, whereas Lu et al. (2019) focus exclusively on knot tying (step VI).

Regardless of the suturing process focus, most approaches address thread perception as a central component of their frameworks, as discussed in Section 3.1.5. Some works further link thread perception with grasp point estimation: Joglekar et al. (2023) introduce a perception-based confidence map for grasping, while Lu et al. (2022) compute optimal grasping poses directly. Another important perceptual focus is needle detection and pose estimation, as explored by Sen et al. (2016) and Pedram et al. (2021). The former also addresses automatic needle size selection as part of the task-oriented setup. Given the specialized requirements of needle and thread handling, several works have designed custom grippers to facilitate precise grasping and alignment (Jackson et al., 2018; Sen et al., 2016).

Regarding action planning and control, the typical approach is often optimization-based, including sequential convex programming (Sen et al., 2016), linear-quadratic control (Lu et al., 2019), nonlinear constrained optimization (Pedram et al., 2021), and MPC (Marra et al., 2024).

Both the manipulation and perception approaches in the literature are typically validated using real surgical robotic systems, with the da Vinci Research Kit (dVRK) being the most commonly used platform. An overview of recent suturing methods, including robot setup and task focus, is presented in Table 7.

5.5.1. Existing literature gaps. While recent advances in autonomous suturing are promising, real surgical environments still pose significant challenges. Perception is often affected by poor lighting and tissue reflections, suspended particles (e.g., from laparoscopic insufflation), tissue deformation, surface motion (as considered in Jackson et al., 2018), and prolonged occlusions from surgical instruments.

Table 7. Summary of representative literature on robotic suturing. Subtask indices (I–VII) refer to the breakdown in Section 5.5. The table highlights a standard reliance on stereo vision for high-precision depth estimation and optimization-based control (e.g., MPC) to enforce safety constraints during tissue interaction. Furthermore, most works decouple the problem, focusing either on robust perception (Subtask I) or the execution of stitching mechanics (Subtasks II–V), with knot tying (VI) remaining an outlier.

References	Vision setup	Robot setup	Subtasks (I–VII)	Key focus	Method type
Sen et al. (2016)	Stereo	dVRK, custom gripper	I–V	Needle tracking, multi-throw suturing, needle size selection	Sequential convex programming
Jackson et al. (2018)	Stereo	ABB IRB 140, custom gripper	–	3D thread tracking	NURBS model, optimized for image reprojection
Lu et al. (2019)	Monocular	MP-285, MPC-200 controller	VI	2D thread perception, automatic knot tying	Template matching, ad-hoc planner, LQ controller
Lu et al. (2020)	Stereo	UR, dVRK	I	3D thread reconstruction	Transfer learning using legacy surgical data, shortest path 3D reconstruction
Pedram et al. (2021)	Stereo	Raven IV surgical system	II–V, VII	Needle pose estimation, autonomous suturing	Constrained nonlinear optimization path planner
Lu et al. (2022)	Stereo	dVRK	I	Grasp pose estimation via 3D thread perception	Semi-supervised segmentation, sliding pairing, shortest path computation
Joglekar et al. (2023)	Stereo	dVRK	I	Confidence-map based grasping via 3D thread perception	Minimum variation spline (MVS) smoothing optimization
Schorp et al. (2023)	Stereo	dVRK	–	3D thread reconstruction from 2D detection	Self-supervised 2D thread segmentation, stereo triangulation, NURBS reconstruction
Marra et al. (2024)	–	dVRK	III–V, VII	Autonomous stitching control	MPC with kinematics and safety constraints

On the manipulation side, most approaches address specific sub-tasks, with fewer providing full end-to-end solutions. Effectively handling dynamic tissue behavior across the entire suturing process remains an open challenge.

5.6. Transport

Transport tasks involving DLOs can be broadly classified into two categories: transporting the DLO itself (Figure 16(a) and Section 5.6.1), and using DLOs as a means to transport external loads—either through direct physical coupling, as in cable-suspended systems, or via non-prehensile interactions without rigid attachment (Figure 16(b) and Section 5.6.2).

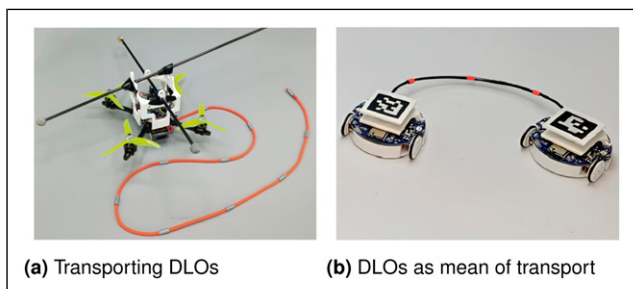


Figure 16. Representative examples of transport tasks involving DLOs. Images courtesy of (a) Shen et al. (2025) and (b) Zhi et al. (2024).

Table 8 summarizes key literature on DLO transport⁵, organized according to the above introduced transport types, sensing modality, DLO characteristics (aligned with criteria from Section 5.2), operational environment, robotic setup, and method’s approach.

5.6.1. Transporting DLOs. Mobile robots (e.g., wheeled robots) and aerial systems (e.g., drones) have demonstrated the capability to transport DLOs using cooperative and autonomous strategies. Some methods involve pairs of drones collaboratively carrying flexible cables or beams (Chen et al., 2021; Xu et al., 2025), while others rely on single-drone solutions guided by second-order modeling and nonlinear MPC (Shen et al., 2025). In simulated scenarios, teams of aerial robots manipulate hoses using techniques such as time-varying linear quadratic regulation (LQR) (Kotaru and Sreenath, 2020), feedback-integrated planning (Gabellieri and Franchi, 2023), or adaptive proportional-derivative (PD) control with fuzzy error modeling (Estevez et al., 2017). As a related challenge, aerial refueling involves transporting and aligning hose-drogue systems, which can be achieved through approaches such as dynamic surface control (Song and Huang, 2022) or oscillation suppression with boundary-control strategies (Liu et al., 2017).

5.6.2. DLOs as mean of transport. DLOs have been increasingly explored as flexible media for *tethered* and *non-prehensile* transport in both aerial and ground-based robotic

Table 8. Overview of selected literature on DLO transport. The table reveals a domain-dependent split: aerial works (e.g., Kotaru and Sreenath (2020); Xu et al. (2025)) predominantly prioritize dynamic stability via LQR or adaptive control and high-frequency state estimation (IMU/motion-capture), whereas ground-based approaches (e.g., Su et al. (2022); Zhi et al. (2024)) typically rely on optimization-based methods to manage interaction constraints and obstacle avoidance via exteroceptive sensing.

References	External sensor	DLO behavior	Environment	Robot setup	Transport task	Control method
Estevez et al. (2017)	–	Elastic (cable)	3D	Sim, multiple aerial vehicles	Aerial transport of DLOs	Adaptive PD controller with fuzzy error modeling
Liu et al. (2017)	–	Elastic (hose–drogue)	2D (transv.)	Sim, 1 aerial vehicle	Aerial refueling	Boundary-control scheme
Kotaru et al. (2018)	–	Elastic (flexible hose)	3D	Sim, multiple aerial vehicles	Suspended load transport with DLOs	Variation-based linearized equations, finite-horizon LQR
Kotaru and Sreenath (2020)	–	Elastic (flexible hose)	3D	Sim, multiple aerial vehicles	Aerial transport of DLOs	Linear time-varying LQR
Chen et al. (2021)	OptiTrack	Elastic (rigid beams)	3D	Real, 2 aerial vehicles	Aerial transport of DLOs	Linearized model and LQR
Song and Huang (2022)	–	Elastic (hose–drogue)	3D	Sim, 1 aerial vehicle	Aerial refueling	Dynamic surface control + extended state observer
Su et al. (2022)	LiDAR (SLAM)	Plastic (long net)	2D (planar)	Real, 2 linked mobile robots	Non-prehensile transport of objects with DLO	Iterative optimization for collision-free trajectories
Gabellieri and Franchi (2023)	–	Elastic (cables)	3D	Sim, 2 aerial vehicles	Aerial transport of DLOs	Backward iteration with feedback integral term
Huang and Zhang (2024)	–	Plastic (ball–string–ball struct.)	2D (planar)	Sim, 2 mobile robots	Non-prehensile transport of objects with DLO	Optimization-based controller
Zhi et al. (2024)	Camera (tag detection)	Elastic (flexible tube)	2D (planar)	Real, 2 linked mobile robots	Non-prehensile transport of objects with DLO	2-step (enveloping + transport), MPC, react. obst. avoidance
Shen et al. (2025)	IMU + motion-capture	Plastic (rope)	3D	Real, 1 aerial vehicle	Aerial transport of DLOs	POD reduced-order model, nonlinear MPC
Xu et al. (2025)	IMU + motion-capture	Elastic (flexible cable)	3D	Real, 2 aerial vehicles	Aerial transport of DLOs	Adaptive controller, Lyapunov stability, obst. avoidance
Sun et al. (2025)	IMU + motion-capture	Elastic (cable)	3D	Real, 4 aerial vehicles	Suspended load transport with DLOs	Finite-time optimal, obst. avoidance, wind-robust

systems. In aerial contexts, cable-suspended load transport has been demonstrated in simulation with flexible hoses manipulated by multiple drones, modeled through variation-based linearization and finite-horizon LQR (Kotaru et al., 2018). In Sun et al. (2025), several drones collaboratively transport suspended loads in agile tasks exploiting a trajectory-based framework that solves the whole-body kinodynamic motion planning problem online, successfully achieving obstacle avoidance and robustness against over 5 [m/s] winds. Aside from aerial transport, ground transport systems have achieved non-prehensile transport by enveloping objects with elastic tubes, coordinated via a two-stage MPC framework with reactive obstacle avoidance (Zhi et al., 2024), or by dragging a long net between mobile

robots using iterative optimization for collision-free motion planning (Su et al., 2022). In simulation, soft structures such as a ball–string–ball mechanism have been employed for gathering and transporting objects, exploiting optimization-based control strategies (Huang and Zhang, 2024). These non-prehensile approaches (Huang and Zhang, 2024; Su et al., 2022; Zhi et al., 2024) demonstrate the versatility of DLOs as compliant transport tools, enabling object manipulation without the need for direct grasping, unlike traditional grasp-based methods (see Section 5.1).

5.6.3. Existing literature gaps. Existing DLO-related transport strategies are primarily tested in controlled indoor environments, limiting their validation to structured settings

such as warehouses. In contrast, real-world transport often occurs outdoors, where wind, uneven terrain, and dynamic obstacles present significant challenges. In these conditions, DLO perception becomes highly unreliable due to sensor limitations, making perception-light methods such as Chen et al. (2021) particularly appealing.

Energy consumption is another often-overlooked factor, yet it critically affects the autonomy of mobile robots. This highlights the value of energy-aware control approaches, such as MPC-based (Shen et al., 2025; Zhi et al., 2024) and LQR-based methods (Chen et al., 2021; Kotaru et al., 2018; Kotaru and Sreenath, 2020), which can improve efficiency during transport.

Another key limitation is the common assumption that DLOs are already attached to drones or that loads are pre-attached to DLOs. However, the grasping/attachment process is nontrivial to automate in practice. Thus, non-prehensile methods (Huang and Zhang, 2024; Su et al., 2022; Zhi et al., 2024) offer a promising alternative by avoiding these attachment challenges, enabling more direct applicability in real-world scenarios. However, they are constrained by terrain conditions, typically requiring flat surfaces suitable for object dragging.

6. Discussion and future directions

Current DLO manipulation methods still struggle with generalization and robustness in the real world. For

example, *shaping* techniques often assume contact-free environments, ignoring the elastoplastic behavior of real DLOs like electrical cables, and skip formal reachability analysis. *Routing* approaches frequently depend on task-specific motion primitives and are restricted to quasi-static, mostly planar environments, overlooking the complexities of clip-based interactions and long-term mechanical stress on the DLO. *Topological manipulation* is still dominated by passive visual perception, which performs poorly in the presence of dense entanglements and tight knots, resulting in frequent grasping failures and limited robustness. Most *suturing* approaches target isolated sub-tasks without offering robust end-to-end pipelines capable of managing tissue deformation, occlusions, and surface motion throughout the full procedure. Lastly, in *transport*, methods are typically validated in structured indoor settings and often assume pre-attached payloads, which limits their adaptability to outdoor or dynamic contexts.

Together, these limitations highlight the need for future research to address challenges across multiple fronts: advanced perception (Section 6.1), high-level and low-level manipulation strategies (Section 6.2), robust failure detection and recovery (Section 6.3), and scalable data collection and curation (Section 6.4). This broader perspective on next-generation DLO manipulation is reflected in the framework illustrated in Figure 17, which is intended as a guiding example rather than an exhaustive architectural specification.

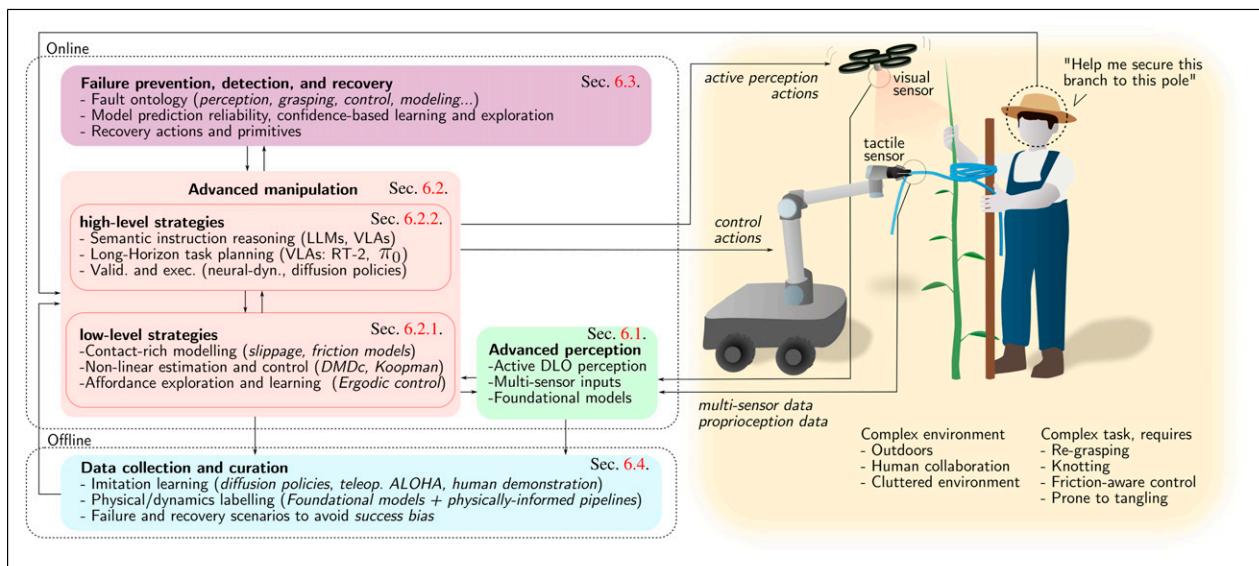


Figure 17. An illustrative full-system framework that addresses some of the most critical challenges in next-generation DLO manipulation. This flowchart outlines a system capable of handling unstructured, real-world complex DLO manipulation tasks (such as the process, shown in the illustration, of securing a branch to a pole in an agricultural setting). It links scalable DLO data-generation pipelines (bottom block, Section 6.4) directly to real-time online manipulation strategies. The runtime loop fuses advanced perception (right block, Section 6.1) with hierarchical manipulation strategies (middle block, Section 6.2), integrating high-level semantic and long-horizon planning (Section 6.2.2) with contact-rich, affordance-aware control (Section 6.2.1). Identifying a major future challenge, the framework explicitly includes fault prevention, detection, and recovery (top block, Section 6.3) as a critical requirement for transferring DLO manipulation research to real-world applications.

6.1. Advanced DLO perception

Reliable DLO manipulation increasingly requires force, tactile, and proprioceptive feedback. Vision alone struggles with real-world lighting, reflections, occlusions, and the visually uniform, thin nature of many DLOs. Even in human manipulation, vision mainly guides initial grasp selection, while successful handling depends on touch, force feedback, and purposeful physical interaction. Interactive perception (Bohg et al., 2017; Weng et al., 2024), where the robot actively probes and manipulates the object to gain information, naturally complements passive sensing and enhances robustness. Additionally, advances in sensorless and tactile strategies, such as those proposed in Monguzzi et al. (2023, 2024b), offer a promising path toward reliable, application-ready systems. While vision remains essential for grasp planning and state estimation, physical interaction sensing is a requisite for true application-ready systems.

Another issue is temporal resolution. Most methods reviewed here operate at low frequencies, restricting them to quasi-static tasks. Interestingly, this is also reflected in the manipulation tasks, usually exploiting quasi-static settings (see Section 6.2). Higher processing rates could simplify perception by minimizing inter-frame visual changes and enable dynamic manipulation. This requires optimizing current architectures or adopting new data sources, such as event cameras or reduced-order observers from force data.

Finally, DLO perception is usually task-specific, diverging from conventional setups that rely on large, general-purpose public datasets, like ImageNet (Deng et al., 2009). Foundation Models, pre-trained on extensive and diverse datasets, are increasingly being applied to robotic tasks (Firoozi et al., 2025). Their influence is now beginning to extend to DLO perception as well, for example, through text-driven segmentation (Sun et al., 2024) or the digital-twin reconstruction pipeline proposed in Jiang et al. (2025). The latter provides a compelling demonstration of foundation models-based DLO perception, reconstructing a physically accurate, simulation-ready digital twin of a real rope from sparse RGB-D video using pre-trained foundation models in zero-shot settings. However, the object is intentionally thick, which simplifies 3D perception of DLOs and also mitigates the difficulty of current foundation models when dealing with DLO-like thin objects, for example, under-representation in internet-scale data and difficulty with small-scale and thin details. Despite these limitations, research in this direction can substantially advance DLO perception by reducing reliance on task-specific data collection and manual annotation. Indeed, the strong generalization capabilities of foundation models can help bridge the real-world DLO variability gap and improve the robustness and adaptability of future DLO perception systems.

6.2. Manipulation of DLOs: What's next

6.2.1. Adaptive and contact-rich control in unstructured environments. To move beyond quasi-static, fixed-grasp

setups, the field must tackle contact-rich interactions like slippage, re-grasping, and environmental friction. A critical barrier is the accurate modeling of hybrid stick-slip transitions. While recent soft robotics research has progressed by incorporating contact constraints into analytical models like Cosserat rods (Jilani et al., 2025; Wiese et al., 2023), these high-fidelity formulations remain inherently nonlinear and computationally intensive, often limiting their utility for robust, real-time control.

Data-driven nonlinear control offers a faster alternative. Techniques like the Koopman operator and dynamic mode decomposition with control (DMDc) (Brunton et al., 2016; Kaiser et al., 2018, 2021) can map complex DLO dynamics into linear observable spaces. Paired with nonlinear model predictive control (NMPC) (Folkestad and Burdick, 2021; Korda and Mezić, 2020), these methods can manage high-dimensional DLO systems at industrial speeds.

In unstructured environments, however, physical properties are often unknown. Active learning and ergodic exploration (Abraham et al., 2021; Saviolo et al., 2023) allow the system to autonomously probe the object, gauging stiffness or friction. This active data gathering enables online adjustments, such as controlled sliding or re-grasping.

Validating these complex behaviors requires standardized hardware benchmarks like the NIST (2025) task boards (Luo et al., 2025; Qi et al., 2026) and shift from binary success criteria to continuous performance metrics (Laezza et al., 2021).

Lastly, human-robot collaboration with DLOs (Zhou et al., 2024) demands dexterity that quasi-static methods cannot provide. Adapting to human movements requires the robot to safely slip, regrasp (Zhaole et al., 2024), and manage varying grasp orientations (Yu et al., 2024).

6.2.2. Long-horizon planning and foundational intelligence. Complex DLO tasks require reasoning that inter-relates high-level semantic understanding with low-level physical interactions. While vision-language-action (VLA) models like π_0 (Black et al., 2024) and RT-2 (Zitkovich et al., 2023) are promising, their precision in DLO manipulation is unproven. A safer approach decouples reasoning from execution (Qi et al., 2026): Large language models (LLMs) generate symbolic sub-goals (e.g., “grasp the cable end”), which physics-aware planners then execute.

Validating these symbolic plans requires accurate deformation modeling to cross the sim-to-real gap. Physics-informed methods like PhysTwin (Jiang et al., 2025) and Particle-Grid Neural Dynamics (Zhang et al., 2025) learn deformable physics directly from RGB-D video. By building simulation-ready digital twins, these methods let planners verify long-horizon sequences using real-world physics priors instead of manually tuned simulators (Xiang et al., 2025).

Finally, executing these plans requires policies that handle complex, multi-modal actions. Diffusion policies (Chi et al., 2025) are highly effective here, but their success depends on in-domain data. The scarcity of datasets capturing the full physical state-space of DLOs remains a severe bottleneck (see Section 6.4).

6.3. Fault prevention, detection and recovery

For real-world deployment, fault detection and recovery must be a core system capability, not an afterthought. Research should start by establishing a fault ontology for DLOs, categorizing failures across perception (occlusion, tracking loss), grasping (slippage), control (instability), and modeling (parameter mismatch). Some work has begun here: Mitrano et al. (2021) predict the reliability of learned DLO models, and Sundaesan et al. (2021) provide recovery strategies for tangled DLOs.

Because learning-based methods struggle with out-of-distribution events, integrating uncertainty estimation (Amini et al., 2020; Kendall and Gal, 2017) is vital for detecting impending failures. Similarly, during online adaptation, incorporating ergodic exploration from equilibrium with stability guarantees (Abraham et al., 2021) can actively enhance failure avoidance during system exploration.

Developing dynamic primitives for fault recovery is another critical area. Humans resolve tangles using dynamic motions like impulsive pulls or whipping. While some primitive-based untangling exists (Zhang et al., 2022), formalizing these into a taxonomy would give supervisors (like VLMs in the RACER framework (Dai et al., 2025)) the vocabulary to command complex, high-frequency maneuvers to fix entanglements. Multi-robot strategies (Aranda et al., 2025; Herguedas et al., 2019) can also mitigate local failures, using physical cues (like tension along a cable) to coordinate recovery when communication is poor.

6.4. Data collection and curation challenges

Recent advances in imitation learning, particularly using Diffusion policies (Chi et al., 2025), have shown that model-free approaches can master complex skills with relatively few demonstrations. Teleoperation frameworks like ALOHA (Zhao et al., 2023, 2025) can enable the learning of precise DLO-related tasks—such as cable routing or shoe lacing—given sufficient demonstrations. Complementing these hardware-intensive approaches, the Universal Manipulation Interface (UMI) provides a scalable alternative by leveraging in-the-wild human demonstrations through hand-held grippers (Chi et al., 2024b). Further reducing hardware requirements, human video-based learning methods directly translate RGB-D videos into robot supervision (Lepert et al., 2025). However, a critical bias limits the utility of these general-purpose data-collection pipelines. Since large-scale datasets naturally lean toward conventional, *easy-to-record*

scenarios, it systematically underrepresents challenging environments (such as underwater cable inspection or highly congested DMLO routing) precisely where automation offers great value. As a result, these models often fail to generalize to DLOs with different physical properties, necessitating ad-hoc data collection for new materials or environmental variations.

One of the main causes of this poor generalization is that DLO dynamics are governed by latent physical properties (such as stiffness and friction) that are often visually imperceptible. To enable robust generalization, data collection must pivot toward *multi-modal* pipelines, synchronizing visual kinematics with high-frequency force and tactile feedback to resolve these invisible parameters. Novel hardware approaches, such as sensorized hydrogels (Hardman et al., 2021), offer a promising avenue for dataset generation by directly capturing mechanical deformation both internally and at the contact surface of the DLO. Alternatively, emerging foundation models address the physical annotation bottleneck computationally. By combining dense point tracking from models like CoTracker3 (Karaev et al., 2024) with physically-informed pipelines like PhysTwin (Jiang et al., 2025), it becomes possible to semi-automate labeling *in the wild*, effectively creating valid generators for dense, physics-informed datasets.

Finally, current datasets suffer from a “success bias,” as they consist almost exclusively of smooth expert demonstrations. Robust DLO manipulation, however, requires recovering from inherent instability and failure modes like snagging or entanglement. Because experts naturally avoid these errors, models trained on such data lack the support to handle out-of-distribution failures. Future datasets must therefore explicitly include *failure and recovery trajectories*, teaching the policy not only the nominal path but also how to untangle or re-route when the primary strategy fails.

7. Conclusions

This survey has reviewed the growing body of work on the robotic perception and manipulation of Deformable Linear Objects (DLOs). Designed as both an introduction for newcomers and a roadmap for experienced researchers, the literature reveals a clear takeaway: while DLO manipulation is a highly active and advancing area of research, it has not yet reached the maturity required for widespread industrial use.

To understand this maturity gap compared to established domains like rigid manipulation or navigation, we must reevaluate what constitutes a “baseline.” In mature fields, capabilities like *tactile sensing* and *robustness to unstructured environments* are often treated as advanced upgrades to an already functional system. For DLO manipulation, however, these capabilities are not optional enhancements; they are fundamental prerequisites for achieving even *basic* reliability.

Consider, for example, the coupling between tactile sensing and deformation. Simply touching a DLO to measure its state can inadvertently actuate it. This physical contact perturbs the object, significantly increasing measurement uncertainty. This problem peaks in near-singular configurations (such as buckling, *pigtailing*, or transitioning from slack to taut regimes). In these states, visual feedback cannot detect accumulated mechanical stress, making tactile data crucial; yet, it is exactly in these regimes that the DLO is most sensitive to touch.

Moreover, DLOs are inherently unstructured. Even in controlled environments, a DLO can generate complex and cluttered conditions *on its own*, including self-occlusions, knots, tangles, and high-friction self-contacts. While other robotics domains usually attribute *unstructured conditions* to external factors (like poor lighting or environmental obstacles in navigation), DLOs generate extreme complexity purely through their own geometry and deformation. Thus, robustness against these intrinsic dynamics is mandatory.

The main challenge to establishing standard, replicable baselines is this intrinsic physical complexity. DLOs are highly nonlinear underactuated systems plagued by singularities and abrupt behavioral shifts. Their configuration space explodes when accounting for varying shapes, materials, and grasp points. Because of this, purely data-driven approaches are rarely sufficient. Typically, black-box learning models (like foundation models) lack the interpretability and provable guarantees needed for industrial safety. Moving forward, the most promising path is to fuse the adaptability and context-awareness of modern machine learning with the rigorous stability and safety guarantees of classical mechanics and systems analysis.

ORCID iDs

Alessio Caporali  <https://orcid.org/0000-0001-9522-4231>
 Ignacio Cuiral-Zueco  <https://orcid.org/0000-0001-5663-7033>
 Gonzalo López-Nicolás  <https://orcid.org/0000-0001-9347-5969>
 Gianluca Palli  <https://orcid.org/0000-0001-9457-4643>

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Horizon Europe project *IntelliMan - AI-Powered Manipulation System for Advanced Robotic Service, Manufacturing and Prosthetics* [grant number 101070136]. Alessio Caporali is funded by FSE+ 2021–2027 under a research contract per Law 240/2010, Art. 24(3)(a), and D.G.R. 693/2023 (REF. PA: 2023-20090/RER - CUP: J19J23000730002). This work was also supported via project REMAIN S1/1.1/E0111 (Interreg Sudoe Programme, ERDF) and projects PID2021-124137OB-I00 and PID2024-159279OB-I00 funded by MICIU/AEI/10.13039/501100011033 and by ERDF/EU, and ANR-25-PERO-0003 PEPR Robotique - PC DRMI (Dexterous Robotic Manipulation for Industry).

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

1. <https://www.algoryx.se/agx-dynamics/>.
2. <https://obi.virtualmethodstudio.com/>.
3. The terms elastic, plastic, and elastoplastic are used here as qualitative, practitioner-oriented labels rather than strict material-science definitions.
4. Referring to motions performed far from static equilibrium, where the DLO response is highly sensitive to timing and acceleration, and inertial effects are not negligible.
5. Note that this classification focuses on methods that explicitly consider the deformable nature of DLOs; over-simplified approaches like tensioned DLOs modeled as a single pendulum or single mass-spring-damper system, while efficient for suspended load scenarios, are considered outside this survey's scope.

References

- Abraham I, Prabhakar A and Murphey TD (2021) An ergodic measure for active learning from equilibrium. *IEEE Transactions on Automation Science and Engineering* 18(3): 917–931. <https://doi.org/10.1109/tase.2020.3043636>
- Achanta R, Shaji A, Smith K, et al. (2012) Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34(11): 2274–2282. <https://doi.org/10.1109/TPAMI.2012.120>
- Aghajanzadeh O, Aranda M, Corrales Ramon JA, et al. (2022a) Adaptive deformation control for elastic linear objects. *Frontiers in Robotics and AI* 9: 868459. <https://doi.org/10.3389/frobt.2022.868459>
- Aghajanzadeh O, Aranda M, López-Nicolás G, et al. (2022b) An offline geometric model for controlling the shape of elastic linear objects. In: 2022 IEEE/RSJ international conference on intelligent robots and systems (IROS), Kyoto, Japan, 23–27 October 2022. IEEE, pp. 2175–2181.
- Aghajanzadeh O, Picard G, Ramon JAC, et al. (2022c) Optimal deformation control framework for elastic linear objects. In: 2022 IEEE 18th international conference on automation science and engineering (CASE), Mexico City, Mexico, 20–24 August 2022. IEEE, pp. 722–728.
- Almaghout K, Cherubini A and Klimchik A (2024) Robotic co-manipulation of deformable linear objects for large deformation tasks. *Robotics and Autonomous Systems* 175: 104652. <https://doi.org/10.1016/j.robot.2024.104652>
- Amini A, Schwarting W, Soleimany A, et al. (2020) Deep evidential regression. *Advances in Neural Information Processing Systems* 33: 14927–14937.
- Antman SS (1972) The theory of rods. In: *Linear Theories of Elasticity and Thermoelasticity: Linear and Nonlinear Theories of Rods, Plates, and Shells*. Springer, pp. 641–703.

- Aranda M, Cuiral-Zueco I and López-Nicolás G (2025) Distributed control of flexible chained multiagent formations. *IEEE Control Systems Letters* 9: 2018–2023. <https://doi.org/10.1109/lcsys.2025.3590428>
- Arriola-Rios VE, Guler P, Ficuciello F, et al. (2020) Modeling of deformable objects for robotic manipulation: a tutorial and review. *Frontiers in Robotics and AI* 7: 82. <https://doi.org/10.3389/frobt.2020.00082>
- Artinian A, Amar FB and Perdereau V (2024) Closed-loop shape control of deformable linear objects based on Cosserat model. *IEEE Robotics and Automation Letters* 9(10): 8746–8753. <https://doi.org/10.1109/ra.2024.3451368>
- Azad A, Quentin H, Faiz BA, et al. (2023) Optimal cosserat-based deformation control for robotic manipulation of linear objects. In: 2023 IEEE/ASME international conference on advanced intelligent mechatronics (AIM), Seattle, Washington, USA, 27 June - 1 July 2023. IEEE, pp. 381–388.
- Bender J, Müller M and Macklin M (2015) Position-based simulation methods in computer graphics. *Eurographics (Tutorials)*. The Eurographics Association, 8.
- Bensch M, Job TD, Habich TL, et al. (2024) Physics-informed neural networks for continuum robots: towards fast approximation of static cosserat rod theory. In: 2024 IEEE international conference on robotics and automation (ICRA), Yokohama, Japan, 13–17 May 2024. IEEE, pp. 17293–17299.
- Berenson D (2013) Manipulation of deformable objects without modeling and simulating deformation. In: 2013 IEEE/RSJ international conference on intelligent robots and systems, Tokyo, Japan, 03–07 November 2013. IEEE, pp. 4525–4532.
- Black K, Brown N, Driess D, et al. (2024) $\pi 0$: a vision-language-action flow model for general robot control. *arXiv preprint arXiv:2410.24164*.
- Bogh J, Hausman K, Sankaran B, et al. (2017) Interactive perception: leveraging action in perception and perception in action. *IEEE Transactions on Robotics* 33(6): 1273–1291. <https://doi.org/10.1109/tro.2017.2721939>
- Bolya D, Zhou C, Xiao F, et al. (2019) YOLACT: real-time instance segmentation. In: Proceedings of the IEEE/CVF international conference on computer vision, Seoul, Korea, 27 October 2019–02 November 2019, pp. 9157–9166.
- Borgefors G (1986) Distance transformations in digital images. *Computer Vision, Graphics, and Image Processing* 34(3): 344–371. [https://doi.org/10.1016/s0734-189x\(86\)80047-0](https://doi.org/10.1016/s0734-189x(86)80047-0)
- Bouaziz S, Martin S, Liu T, et al. (2023) Projective dynamics: fusing constraint projections for fast simulation. *Seminal Graphics Papers: Pushing the Boundaries* 2: 787–797.
- Bretl T and McCarthy Z (2014) Quasi-static manipulation of a kirchhoff elastic rod based on a geometric analysis of equilibrium configurations. *The International Journal of Robotics Research* 33(1): 48–68. <https://doi.org/10.1177/0278364912473169>
- Brunton SL, Proctor JL and Kutz JN (2016) Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* 113(15): 3932–3937. <https://doi.org/10.1073/pnas.1517384113>
- Caporali A and Palli G (2025) Robotic manipulation of deformable linear objects via multi-view model-based visual tracking. *IEEE* 30(5): 3966–3977. <https://doi.org/10.1109/tmech.2025.3562295>
- Caporali A, Zanella R, De Gregorio D, et al. (2022b) Ariadne+: deep learning-based augmented framework for the instance segmentation of wires. *IEEE Transactions on Industrial Informatics* 18(12): 8607–8617. <https://doi.org/10.1109/tii.2022.3154477>
- Caporali A, Galassi K and Palli G (2024a) DLO perceiver: grounding large language model for deformable linear objects perception. *IEEE Robotics and Automation Letters* 9(12): 11385–11392. Available at: <https://doi.org/10.1109/ra.2024.3491428>
- Caporali A, Pantano M, Janisch L, et al. (2023c) A weakly supervised semi-automatic image labeling approach for deformable linear objects. *IEEE Robotics and Automation Letters* 8(2): 1013–1020. <https://doi.org/10.1109/ra.2023.3234799>
- Caporali A, Galassi K and Palli G (2023a) Deformable linear objects 3D shape estimation and tracking from multiple 2D views. *IEEE Robotics and Automation Letters* 8(6): 3852–3859. Available at: <https://doi.org/10.1109/ra.2023.3273518>
- Caporali A, Kicki P, Galassi K, et al. (2024b) Deformable linear objects manipulation with online model parameters estimation. *IEEE Robotics and Automation Letters* 9(3): 2598–2605. <https://doi.org/10.1109/ra.2024.3357310>
- Caporali A, Galassi K, Žagar BL, et al. (2023b) RT-DLO: real-time deformable linear objects instance segmentation. *IEEE Transactions on Industrial Informatics* 19(11): 11333–11342. Available at: <https://doi.org/10.1109/tii.2023.3245641>
- Caporali A, Galassi K, Zanella R, et al. (2022a) FASTDLO: fast deformable linear objects instance segmentation. *IEEE Robotics and Automation Letters* 7(4): 9075–9082. Available at: <https://doi.org/10.1109/ra.2022.3189791>
- Caporali A, Galassi K, Zanella R, et al. (2025) GNN topology representation learning for deformable multi-linear objects dual-arm robotic manipulation. *IEEE Transactions on Automation Science and Engineering* 22: 14738–14751. Available at: <https://doi.org/10.1109/tase.2025.3562231>
- Chen LC, Zhu Y, Papandreou G, et al. (2018) Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European conference on computer vision (ECCV), Munich, Germany, 8–14 September 2018, pp. 801–818.
- Chen T, Shan J and Liu HH (2021) Cooperative transportation of a flexible payload using two quadrotors. *Journal of Guidance, Control, and Dynamics* 44(11): 2099–2107. <https://doi.org/10.2514/1.g005914>
- Chen K, Bing Z, Wu F, et al. (2023) Contact-aware shaping and maintenance of deformable linear objects with fixtures. In: 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Detroit, MI, 1–5 October 2023, pp. 1–8.

- Chen K, Bing Z, Wu Y, et al. (2024) Real-time contact state estimation in shape control of deformable linear objects under small environmental constraints. In: 2024 IEEE international conference on robotics and automation (ICRA), Yokohama, Japan, 13-17 May 2024. IEEE, pp. 13833–13839.
- Chi C and Berenson D (2019) Occlusion-robust deformable object tracking without physics simulation. In: 2019 IEEE/RSJ international conference on intelligent robots and systems (IROS), Macau, China, 03–08 November 2019. IEEE, pp. 6443–6450.
- Chi C, Burchfiel B, Cousineau E, et al. (2024a) Iterative residual policy: for goal-conditioned dynamic manipulation of deformable objects. *The International Journal of Robotics Research* 43(4): 389–404. <https://doi.org/10.1177/02783649231201201>
- Chi C, Xu Z, Pan C, et al. (2024b) Universal manipulation interface: In-the-wild robot teaching without in-the-wild robots. In: Proceedings of robotics: science and systems (RSS), Delft, Netherlands, 15–19 July 2024.
- Chi C, Xu Z, Feng S, et al. (2025) Diffusion policy: visuomotor policy learning via action diffusion. *The International Journal of Robotics Research* 44(10-11): 1684–1704. <https://doi.org/10.1177/02783649241273668>
- Choi A, Tong D, Park B, et al. (2023) mBEST: realtime deformable linear object detection through minimal bending energy skeleton pixel traversals. *IEEE Robotics and Automation Letters* 8(8): 4863–4870. Available at: <https://doi.org/10.1109/LRA.2023.3290419>
- Cirillo A, Laudante G and Pirozzi S (2021a) Proximity sensor for thin wire recognition and manipulation. *Machines* 9(9): 188. <https://doi.org/10.3390/machines9090188>
- Cirillo A, Laudante G and Pirozzi S (2021b) Tactile sensor data interpretation for estimation of wire features. *Electronics* 10(12): 1458. <https://doi.org/10.3390/electronics10121458>
- Cop KP, Peters A, Žagar BL, et al. (2021) New metrics for industrial depth sensors evaluation for precise robotic applications. In: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Prague, Czech Republic, 27 September - 1 October 2021. IEEE, pp. 5350–5356.
- Coumans E (2025) Bullet physics SDK: real-time collision detection and multi-physics simulation. <https://github.com/bulletphysics/bullet3>
- Cui Z, Ma W, Lai J, et al. (2022) Coupled multiple dynamic movement primitives generalization for deformable object manipulation. *IEEE Robotics and Automation Letters* 7(2): 5381–5388. <https://doi.org/10.1109/lra.2022.3156656>
- Cuiral-Zueco I and López-Nicolás G (2024) Taxonomy of deformable object shape control. *IEEE Robotics and Automation Letters* 9(10): 9015–9022. <https://doi.org/10.1109/LRA.2024.3455770>
- Cuiral-Zueco I and López-Nicolás G (2025) Time consistent surface mapping for deformable object shape control. *IEEE Transactions on Automation Science and Engineering* 22: 11099–11111. <https://doi.org/10.1109/tase.2025.3529180>
- Cuiral-Zueco I, López-Nicolás G and Araujo H (2022) Gripper positioning for object deformation tasks. In: 2022 international conference on robotics and automation (ICRA), Philadelphia, PA, USA, 23–27 May 2022, pp. 963–969.
- Cuiral-Zueco I, Karayiannidis Y and López-Nicolás G (2023) Contour based object-compliant shape control. *IEEE Robotics and Automation Letters* 8(8): 5164–5171. <https://doi.org/10.1109/lra.2023.3292617>
- Dai C, Shan G, Liu H, et al. (2022) Robotic manipulation of sperm as a deformable linear object. *IEEE Transactions on Robotics* 38(5): 2799–2811. <https://doi.org/10.1109/tro.2022.3158200>
- Dai Y, Lee J, Fazeli N, et al. (2025) RACER: rich language-guided failure recovery policies for imitation learning. In: 2025 IEEE international conference on robotics and automation (ICRA), Vienna, Austria, 1 - 6 June 2025. IEEE, pp. 15657–15664.
- Daniel M, Magassouba A, Aranda M, et al. (2024) Multi actor-critic DDPG for robot action space decomposition: a framework to control large 3D deformation of soft linear objects. *IEEE Robotics and Automation Letters* 9(2): 1318–1325. Available at: <https://doi.org/10.1109/lra.2023.3342672>
- De Gregorio D, Palli G and Di Stefano L (2018a) Let's take a walk on superpixels graphs: deformable linear objects segmentation and model estimation. In: Asian conference on computer vision, Perth, Western Australia, 2 - 6 December 2018, Springer, pp. 662–677.
- De Gregorio D, Zanella R, Palli G, et al. (2018b) Integration of robotic vision and tactile sensing for wire-terminal insertion tasks. *IEEE Transactions on Automation Science and Engineering* 16(2): 585–598. <https://doi.org/10.1109/tase.2018.2847222>
- Deng J, Dong W, Socher R, et al. (2009) Imagenet: a large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition, Miami, FL, USA, 20–25 June 2009. IEEE, pp. 248–255.
- Dirr J, Gebauer D, Yao J, et al. (2023) Automatic image generation pipeline for instance segmentation of deformable linear objects. *Sensors* 23(6): 3013. <https://doi.org/10.3390/s23063013>
- Dirr J, Xu C, Zeller J, et al. (2024) Bin picking of deformable linear objects using object-oriented grasp planning. *Procedia CIRP* 130: 810–815. <https://doi.org/10.1016/j.procir.2024.10.169>
- Duenser S, Bern JM, Poranne R, et al. (2018) Interactive robotic manipulation of elastic objects. In: 2018 IEEE/RSJ international conference on intelligent robots and systems (IROS), Madrid, Spain, 1-5 October 2018. IEEE, pp. 3476–3481.
- Estevez J, Graña M and Lopez-Guede JM (2017) Online fuzzy modulated adaptive PD control for cooperative aerial transportation of deformable linear objects. *Integrated Computer-Aided Engineering* 24(1): 41–55. Available at: <https://doi.org/10.3233/ica-160530>
- Firoozi R, Tucker J, Tian S, et al. (2025) Foundation models in robotics: applications, challenges, and the future. *The International Journal of Robotics Research* 44(5): 701–739. <https://doi.org/10.1177/02783649241281508>

- Folkestad C and Burdick JW (2021) Koopman NMPC: Koopman-Based learning and nonlinear model predictive control of control-affine systems. In: 2021 IEEE international conference on robotics and automation (ICRA), Xi'an, China, 30 May 2021–05 June 2021. IEEE, pp. 7350–7356.
- Fresnillo PM, Vasudevan S, Mohammed WM, et al. (2023) An approach based on machine vision for the identification and shape estimation of deformable linear objects. *Mechatronics* 96: 103085. <https://doi.org/10.1016/j.mechatronics.2023.103085>
- Fresnillo P, Mohammed WM, Vasudevan S, et al. (2024) Generation of realistic synthetic cable images to train deep learning segmentation models. *Machine Vision and Applications* 35(4): 84. <https://doi.org/10.1007/s00138-024-01562-y>
- Gabellieri C and Franchi A (2023) Differential flatness and manipulation of elasto-flexible cables carried by aerial robots in a possibly viscous environment. In: 2023 International Conference on Unmanned Aircraft Systems (ICUAS), Warsaw, Poland, 06–09 June 2023. IEEE, pp. 963–968.
- Galassi K and Palli G (2021) Robotic wires manipulation for switchgear cabling and wiring harness manufacturing. In: 2021 4th IEEE international conference on industrial cyber-physical systems (ICPS), Victoria, BC, Canada, 10–12 May 2021. IEEE, pp. 531–536.
- Gazzola M, Dudte LH, McCormick AG, et al. (2018) Forward and inverse problems in the mechanics of soft filaments. *Royal Society Open Science* 5(6): 171628. <https://doi.org/10.1098/rsos.171628>
- Ge S, Fan G and Ding M (2014) Non-rigid point set registration with global-local topology preservation. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, Columbus, OH, USA, 23–28 June 2014, pp. 245–251.
- Golestaneh AF, Hammond M and Cichella V (2024) Scalable optimal motion planning for multi-agent systems by cosserat theory of rods. *IEEE Control Systems Letters* 8: 1391–1396. <https://doi.org/10.1109/lcsys.2024.3412881>
- Govoni A, Zubair N, Soprani S, et al. (2025) Performance analysis of a mass-spring-damper deformable linear object model in robotic simulation frameworks. In: *European Robotics Forum*. Springer, pp. 187–192.
- Grannen J, Sundaresan P, Thananjeyan B, et al. (2021) Untangling dense knots by learning task-relevant keypoints. In: Conference on Robot Learning (CORL), London, UK, 08–11 November 2021. PMLR, pp. 782–800.
- Gu F, Sang H, Zhou Y, et al. (2025) Learning graph dynamics with interaction effects propagation for deformable linear objects shape control. *IEEE Transactions on Automation Science and Engineering* 22: 10881–10892.
- Guo J, Zhang J, Wu D, et al. (2020) An algorithm based on bi-directional searching and geometric constrained sampling for automatic manipulation planning in aircraft cable assembly. *Journal of Manufacturing Systems* 57: 158–168. <https://doi.org/10.1016/j.jmsy.2020.08.015>
- Guo J, Zhang J, Wu D, et al. (2022) A local manipulation path replanning algorithm on deformable linear objects for collisions resulted from model deviation. *Journal of Manufacturing Systems* 65: 362–377. <https://doi.org/10.1016/j.jmsy.2022.09.015>
- Hardman D, Hughes J, Thuruthel TG, et al. (2021) 3D printable sensorized soft gelatin hydrogel for multi-material soft structures. *IEEE Robotics and Automation Letters* 6(3): 5269–5275. <https://doi.org/10.1109/lra.2021.3072600>
- Hellman RB, Tekin C, Van Der Schaar M, et al. (2017) Functional contour-following via haptic perception and reinforcement learning. *IEEE transactions on haptics* 11(1): 61–72. <https://doi.org/10.1109/TOH.2017.2753233>
- Herguedas R, López-Nicolás G, Aragüés R, et al. (2019) Survey on multi-robot manipulation of deformable objects. In: 2019 24th IEEE International conference on emerging technologies and factory automation (ETFA), Zaragoza, Spain, 10–13 September 2019. IEEE, pp. 977–984.
- Hermansson T, Vajedi S, Forsberg T, et al. (2016) Identification of material parameters of complex cables from scanned 3D shapes. *Procedia CIRP* 43: 280–285. Available at: <https://doi.org/10.1016/j.procir.2016.02.009>
- Holešovský O, Škoviera R and Hlaváč V (2024) Movingcables: moving cable segmentation method and dataset. *IEEE Robotics and Automation Letters* 9(8): 6991–6998. <https://doi.org/10.1109/lra.2024.3416800>
- Holešovský O, Škoviera R and Hlaváč V (2025) Interactive robotic moving cable segmentation by motion correlation. *IEEE Robotics and Automation Letters* 10(7): 7420–7427. <https://doi.org/10.1109/lra.2025.3574960>
- Hou YC, Sahari KSM and How DNT (2019) A review on modeling of flexible deformable object for dexterous robotic manipulation. *International Journal of Advanced Robotic Systems* 16(3): 1729881419848894. <https://doi.org/10.1177/1729881419848894>
- Huang Y and Zhang S (2024) Cooperative object transport by two robots connected with a ball-string-ball structure. *IEEE Robotics and Automation Letters* 9(5): 4313–4320. <https://doi.org/10.1109/lra.2024.3379802>
- Huang S, Yamakawa Y, Senoo T, et al. (2015) Robotic needle threading manipulation based on high-speed motion strategy using high-speed visual feedback. In: 2015 IEEE/RSJ international conference on intelligent robots and systems (IROS), Hamburg, Germany, 28 September - 02 October 2015. IEEE, pp. 4041–4046.
- Huang Y, Xia C, Wang X, et al. (2023) Learning graph dynamics with external contact for deformable linear objects shape control. *IEEE Robotics and Automation Letters* 8(6): 3892–3899. <https://doi.org/10.1109/lra.2023.3264764>
- Huang X, Chen D, Guo Y, et al. (2024) Untangling multiple deformable linear objects in unknown quantities with complex backgrounds. *IEEE Transactions on Automation Science and Engineering* 21(1): 671–683. <https://doi.org/10.1109/tase.2023.3233949>
- Huo S, Duan A, Li C, et al. (2022) Keypoint-based planar bimanual shaping of deformable linear objects under environmental

- constraints with hierarchical action framework. *IEEE Robotics and Automation Letters* 7(2): 5222–5229. <https://doi.org/10.1109/lra.2022.3154842>
- Irving B (2016) maskSLIC: regional superpixel generation with application to local pathology characterisation in medical images. *arXiv preprint arXiv:1606.09518*.
- Jackson RC, Yuan R, Chow DL, et al. (2018) Real-time visual tracking of dynamic surgical suture threads. *IEEE Transactions on Automation Science and Engineering* 15(3): 1078–1090. <https://doi.org/10.1109/TASE.2017.2726689>
- Jiang X, Koo K, Kikuchi K, et al. (2011) Robotized assembly of a wire harness in a car production line. *Advanced Robotics* 25(3-4): 473–489. <https://doi.org/10.1163/016918610x551782>
- Jiang H, Hsu HY, Zhang K, et al. (2025) Phystwin: physics-informed reconstruction and simulation of deformable objects from videos. In: Proceedings of the IEEE/CVF international conference on computer vision (ICCV), Honolulu, HI, USA, 19-20 October 2025.
- Jilani R, Villard PF and Kerrien E (2025) Quasi-static cosserat rods in contact with implicit surfaces. *IEEE Robotics and Automation Letters* 10(7): 6536–6543. <https://doi.org/10.1109/lra.2025.3570131>
- Jiménez P (2012) Survey on model-based manipulation planning of deformable objects. *Robotics and Computer-Integrated Manufacturing* 28(2): 154–163. <https://doi.org/10.1016/j.rcim.2011.08.002>
- Jin S, Wang C and Tomizuka M (2019) Robust deformation model approximation for robotic cable manipulation. In: 2019 IEEE/RSJ international conference on intelligent robots and systems (IROS), Macau, SAR, China, 3-8 November 2019. IEEE, pp. 6586–6593.
- Jin S, Lian W, Wang C, et al. (2022) Robotic cable routing with spatial representation. *IEEE Robotics and Automation Letters* 7(2): 5687–5694. <https://doi.org/10.1109/lra.2022.3158377>
- Joglekar N, Liu F, Orosco R, et al. (2023) Suture thread spline reconstruction from endoscopic images for robotic surgery with reliability-driven keypoint detection. In: 2023 IEEE international conference on robotics and automation (ICRA), London, UK, 29 May 2023–02 June 2023, pp. 4747–4753.
- Jung P, Leyendecker S, Linn J, et al. (2011) A discrete mechanics approach to the cosserat rod theory—part 1: static equilibria. *International Journal for Numerical Methods in Engineering* 85(1): 31–60. <https://doi.org/10.1002/nme.2950>
- Kaiser E, Kutz JN and Brunton SL (2018) Sparse identification of nonlinear dynamics for model predictive control in the low-data limit. *Proceedings of the Royal Society A* 474(2219): 20180335. <https://doi.org/10.1098/rspa.2018.0335>
- Kaiser E, Kutz JN and Brunton SL (2021) Data-driven discovery of Koopman eigenfunctions for control. *Machine Learning: Science and Technology* 2(3): 035023. Available at: <https://doi.org/10.1088/2632-2153/abf0f5>
- Kamaras G and Ramamoorthy S (2025) Distributional treatment of real2sim2real for object-centric agent adaptation in vision-driven DLO manipulation. *IEEE Robotics and Automation Letters* 10(8): 8075–8082. Available at: <https://doi.org/10.1109/lra.2025.3581744>
- Karaev N, Makarov I, Wang J, et al. (2024) Cotracker3: simpler and better point tracking by pseudolabelling real videos. *arXiv preprint arXiv:2410.11831*.
- Keipour A, Bandari M and Schaal S (2022a) Deformable one-dimensional object detection for routing and manipulation. *IEEE Robotics and Automation Letters* 7(2): 4329–4336. <https://doi.org/10.1109/lra.2022.3146920>
- Keipour A, Bandari M and Schaal S (2022b) Efficient spatial representation and routing of deformable one-dimensional objects for manipulation. In: 2022 IEEE/RSJ international conference on intelligent robots and systems (IROS), Kyoto, Japan, 23-27 October 2022. IEEE, pp. 211–216.
- Kendall A and Gal Y (2017) What uncertainties do we need in Bayesian deep learning for computer vision? *Advances in Neural Information Processing Systems* 30: 5580–5590.
- Khalifa A and Palli G (2022) New model-based manipulation technique for reshaping deformable linear objects. *The International Journal of Advanced Manufacturing Technology* 118(11): 3575–3583. <https://doi.org/10.1007/s00170-021-08107-x>
- Kicki P, Bednarek M, Lembicz P, et al. (2021) Tell me, what do you see? Interpretable classification of wiring harness branches with deep neural networks. *Sensors* 21(13): 4327. <https://doi.org/10.3390/s21134327>
- Kicki P, Szymko A and Walas K (2023) DLOFTBs – fast tracking of deformable linear objects with b-splines. In: 2023 IEEE international conference on robotics and automation (ICRA), London, UK, 29 May 2023–02 June 2023, pp. 7104–7110.
- Kirillov A, Mintun E, Ravi N, et al. (2023) Segment anything. In: Proceedings of the IEEE/CVF international conference on computer vision, Honolulu, Hawaii, 19–23 October 2025, pp. 4015–4026.
- Koessler A, Filella NR, Bouzgarrou BC, et al. (2021) An efficient approach to closed-loop shape control of deformable objects using finite element models. In: 2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, pp. 1637–1643.
- Korda M and Mezić I (2020) Koopman model predictive control of nonlinear dynamical systems. In: *The Koopman Operator in Systems and Control: Concepts, Methodologies, and Applications*. Springer, pp. 235–255.
- Kotaru P and Sreenath K (2020) Multiple quadrotors carrying a flexible hose: dynamics, differential flatness and control. *IFAC-PapersOnLine* 53(2): 8832–8839. <https://doi.org/10.1016/j.ifacol.2020.12.1396>
- Kotaru P, Wu G and Sreenath K (2018) Differential-flatness and control of quadrotor (s) with a payload suspended through flexible cable(s). In: 2018 Indian Control Conference (ICC). IEEE, pp. 352–357.
- Kugelstadt T and Schömer E (2016) Position and orientation based cosserat rods. *Symposium on Computer Animation* 169: 178.
- Laezza R and Karayiannidis Y (2021) Learning shape control of elastoplastic deformable linear objects. In: 2021 IEEE

- International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 4438–4444.
- Laezza R, Gieselmann R, Pokorny FT, et al. (2021) Reform: a robot learning sandbox for deformable linear object manipulation. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 4717–4723.
- Lagneau R, Krupa A and Marchal M (2020) Automatic shape control of deformable wires based on model-free visual servoing. *IEEE Robotics and Automation Letters* 5(4): 5252–5259. <https://doi.org/10.1109/lra.2020.3007114>
- Lang H, Linn J and Arnold M (2011) Multi-body dynamics simulation of geometrically exact cosserat rods. *Multibody System Dynamics* 25(3): 285–312. <https://doi.org/10.1007/s11044-010-9223-x>
- Lee R, Hamaya M, Murooka T, et al. (2021) Sample-efficient learning of deformable linear object manipulation in the real world through self-supervision. *IEEE Robotics and Automation Letters* 7(1): 573–580. <https://doi.org/10.1109/lra.2021.3130377>
- Lepert M, Fang J and Bohg J (2025) Phantom: training robots without robots using only human videos. In: 9th annual conference on robot learning, Seoul, Korea, 27–30 September 2025.
- Li M and Choi C (2024) Learning for deformable linear object insertion leveraging flexibility estimation from visual cues. In: *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 5183–5189.
- Li M, Yu H and Choi C (2025) Routing manipulation of deformable linear object using reinforcement learning and diffusion policy. *2025 IEEE International Conference on Robotics and Automation (ICRA), Atlanta, GA, USA, 19-23 May 2025*. IEEE. IEEE.
- Lim V, Huang H, Chen LY, et al. (2022) Real2sim2real: self-supervised learning of physical single-step dynamic actions for planar robot casting. In: *2022 international conference on robotics and automation (ICRA)*, Philadelphia, PA, USA, 23–27 May 2022, pp. 8282–8289.
- Liu Z, Liu J and He W (2017) Modeling and vibration control of a flexible aerial refueling hose with variable lengths and input constraint. *Automatica* 77: 302–310. <https://doi.org/10.1016/j.automatica.2016.11.002>
- Liu F, Su E, Lu J, et al. (2023) Robotic manipulation of deformable rope-like objects using differentiable compliant position-based dynamics. *IEEE Robotics and Automation Letters* 8(7): 3964–3971. <https://doi.org/10.1109/lra.2023.3264766>
- Long J, Shelhamer E and Darrell T (2015) Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, Boston, MA, USA, 07–12 June 2015, pp. 3431–3440.
- Lu B, Chu HK, Huang KC, et al. (2019) Vision-based surgical suture looping through trajectory planning for wound suturing. *IEEE Transactions on Automation Science and Engineering* 16(2): 542–556. <https://doi.org/10.1109/tase.2018.2840532>
- Lu B, Chen W, Jin YM, et al. (2020) A learning-driven framework with spatial optimization for surgical suture thread reconstruction and autonomous grasping under multiple topologies and environmental noises. In: *2020 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, Las Vegas, NV, USA, 24 October 2020–24 January 2021, pp. 3075–3082.
- Lu B, Li B, Chen W, et al. (2022) Toward image-guided automated suture grasping under complex environments: a learning-enabled and optimization-based holistic framework. *IEEE Transactions on Automation Science and Engineering* 19(4): 3794–3808. <https://doi.org/10.1109/TASE.2021.3136185>
- Lui WH and Saxena A (2013) Tangled: learning to untangle ropes with RGB-D perception. *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 837–844.
- Luo H and Demiris Y (2025) TSL: tracking deformable linear objects for bimanual shoe lacing. *IEEE Robotics and Automation Letters* 10(8): 8212–8219. Available at: <https://doi.org/10.1109/lra.2025.3583476>
- Luo J, Xu C, Geng X, et al. (2024) Multi-stage cable routing through hierarchical imitation learning. *IEEE Transactions on Robotics* 40: 1476–1491. <https://doi.org/10.1109/tro.2024.3353075>
- Luo J, Xu C, Wu J, et al. (2025) Precise and dexterous robotic manipulation via human-in-the-loop reinforcement learning. *Science Robotics* 10(105): eads5033. <https://doi.org/10.1126/scirobotics.ads5033>
- Lv N, Liu J, Ding X, et al. (2017) Physically based real-time interactive assembly simulation of cable harness. *Journal of Manufacturing Systems* 43: 385–399. <https://doi.org/10.1016/j.jmsy.2017.02.001>
- Lv N, Liu J, Xia H, et al. (2020) A review of techniques for modeling flexible cables. *Computer-Aided Design* 122: 102826. <https://doi.org/10.1016/j.cad.2020.102826>
- Lv N, Liu J and Jia Y (2022) Dynamic modeling and control of deformable linear objects for single-arm and dual-arm robot manipulations. *IEEE Transactions on Robotics* 38(4): 2341–2353. <https://doi.org/10.1109/tro.2021.3139838>
- Lv K, Yu M, Pu Y, et al. (2023) Learning to estimate 3-D states of deformable linear objects from single-frame occluded point clouds. *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 7119–7125.
- Ma X, Hsu D and Lee WS (2022) Learning latent graph dynamics for visual manipulation of deformable objects. In: *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 8266–8273.
- Macklin M, Müller M, Chentanez N, et al. (2014) Unified particle physics for real-time applications. *ACM Transactions on Graphics* 33(4): 1–12. <https://doi.org/10.1145/2601097.2601152>
- Marra P, Hussain S, Caianiello M, et al. (2024) MPC for suturing stitch automation. *IEEE Transactions on Medical Robotics and Bionics* 6(4): 1468–1477. Available at: <https://doi.org/10.1109/tmrb.2024.3472796>
- McConachie D, Power T, Mitrano P, et al. (2020) Learning when to trust a dynamics model for planning in reduced state spaces.

- IEEE Robotics and Automation Letters* 5(2): 3540–3547. <https://doi.org/10.1109/lra.2020.2972858>
- Mishani I and Sintov A (2021) Real-time non-visual shape estimation and robotic dual-arm manipulation control of an elastic wire. *IEEE Robotics and Automation Letters* 7(1): 422–429. <https://doi.org/10.1109/lra.2021.3128707>
- Mishani I and Sintov A (2023) Learning configurations of wires for real-time shape estimation and manipulation planning. *Engineering Applications of Artificial Intelligence* 121: 105967. <https://doi.org/10.1016/j.engappai.2023.105967>
- Mitrano P, McConachie D and Berenson D (2021) Learning where to trust unreliable models in an unstructured world for deformable object manipulation. *Science Robotics* 6(54): eabd8170. <https://doi.org/10.1126/scirobotics.abd8170>
- Mitrano P, LaGrassa A, Kroemer O, et al. (2023) Focused adaptation of dynamics models for deformable object manipulation. In: *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 5931–5937.
- Mittal M, Roth P, Tigue J, et al. (2025) Isaac lab: a GPU-accelerated simulation framework for multi-modal robot learning. *arXiv preprint arXiv:2511.04831*.
- Moll M and Kavraki LE (2006) Path planning for deformable linear objects. *IEEE Transactions on Robotics* 22(4): 625–636. <https://doi.org/10.1109/tro.2006.878933>
- Monguzzi A, Pelosi M, Zanchettin AM, et al. (2023) Tactile based robotic skills for cable routing operations. In: *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 3793–3799.
- Monguzzi A, Mantegna N, Zanchettin AM, et al. (2024a) Potential field-based online path planning for robust cable routing. In: *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 7558–7564.
- Monguzzi A, Zanchettin AM and Rocco P (2024b) Sensorless robotized cable contour following and connector detection. *Mechatronics* 97: 103096. <https://doi.org/10.1016/j.mechatronics.2023.103096>
- Monguzzi A, Dotti T, Fattorelli L, et al. (2025) Optimal model-based path planning for the robotic manipulation of deformable linear objects. *Robotics and Computer-Integrated Manufacturing* 92: 102891. <https://doi.org/10.1016/j.rcim.2024.102891>
- Myronenko A and Song X (2010) Point set registration: coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32(12): 2262–2275. <https://doi.org/10.1109/TPAMI.2010.46>
- Nair A, Chen D, Agrawal P, et al. (2017) Combining self-supervised learning and imitation for vision-based rope manipulation. In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 2146–2153.
- Naughton N, Sun J, Tekinalp A, et al. (2021) Elastica: a compliant mechanics environment for soft robotic control. *IEEE Robotics and Automation Letters* 6(2): 3389–3396. <https://doi.org/10.1109/lra.2021.3063698>
- Navarro-Alarcon D, Liu YH, Romero JG, et al. (2013) Model-free visually servoed deformation control of elastic objects by robot manipulators. *IEEE Transactions on Robotics* 29(6): 1457–1468. <https://doi.org/10.1109/tro.2013.2275651>
- Nguyen HG and Franke J (2021) Deep learning-based optical inspection of rigid and deformable linear objects in wiring harnesses. *Procedia CIRP* 104: 1765–1770. <https://doi.org/10.1016/j.procir.2021.11.297>
- Nguyen HG, Habiboglu R and Franke J (2022) Enabling deep learning using synthetic data: a case study for the automotive wiring harness manufacturing. *Procedia CIRP* 107: 1263–1268. <https://doi.org/10.1016/j.procir.2022.05.142>
- NIST (2025) Assembly performance metrics and test methods. <https://www.nist.gov/el/intelligent-systems-division-73500/robotic-grasping-and-manipulation-assembly/assembly> (Accessed 16 July 2025).
- Page MJ, McKenzie JE, Bossuyt PM, et al. (2021) The prisma 2020 statement: an updated guideline for reporting systematic reviews. *International Journal of Surgery* 88: 105906. <https://doi.org/10.1016/j.ijso.2021.105906>
- Palli G (2020) Model-based manipulation of deformable linear objects by multivariate dynamic splines. In: *2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS)*. IEEE, pp. 520–525.
- Palli G and Pirozzi S (2019) A tactile-based wire manipulation system for manufacturing applications. *Robotics* 8(2): 46. <https://doi.org/10.3390/robotics8020046>
- Pecyna L, Dong S and Luo S (2022) Visual-tactile multimodality for following deformable linear objects using reinforcement learning. In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 3987–3994.
- Pedram SA, Shin C, Ferguson PW, et al. (2021) Autonomous suturing framework and quantification using a cable-driven surgical robot. *IEEE Transactions on Robotics* 37(2): 404–417. <https://doi.org/10.1109/tro.2020.3031236>
- Pirozzi S and Natale C (2018) Tactile-based manipulation of wires for switchgear assembly. *IEEE* 23(6): 2650–2661. <https://doi.org/10.1109/tmech.2018.2869477>
- Qi CR, Yi L, Su H, et al. (2017) Pointnet++: deep hierarchical feature learning on point sets in a metric space. *Advances in Neural Information Processing Systems* 30: 5105–5114.
- Qi J, Ma G, Zhu J, et al. (2021) Contour moments based manipulation of composite rigid-deformable objects with finite time model estimation and shape/position control. *IEEE/ASME transactions on mechatronics* 27(5): 2985–2996. <https://doi.org/10.1109/tmech.2021.3126383>
- Qi J, Ran G, Wang B, et al. (2023) Adaptive shape servoing of elastic rods using parameterized regression features and auto-tuning motion controls. *IEEE Robotics and Automation Letters* 9(2): 1428–1435. <https://doi.org/10.1109/lra.2023.3346758>
- Qi J, Lu L, Wang F, et al. (2026) LLM-driven symbolic planning and hierarchical imitation learning for long-horizon deformable object assembly. *Robotics and Computer-Integrated Manufacturing* 97: 103096. <https://doi.org/10.1016/j.rcim.2025.103096>

- Roa MA and Suárez R (2015) Grasp quality measures: review and performance. *Autonomous Robots* 38: 65–88. <https://doi.org/10.1007/s10514-014-9402-3>
- Ronneberger O, Fischer P and Brox T (2015) U-Net: convolutional networks for biomedical image segmentation. In: *Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, pp. 234–241.
- Roussel O, Borum A, Taix M, et al. (2015) Manipulation planning with contacts for an extensible elastic rod by sampling on the submanifold of static equilibrium configurations. In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 3116–3121.
- Roussel O, Fernbach P and Taix M (2019) Motion planning for an elastic rod using contacts. *IEEE Transactions on Automation Science and Engineering* 17(2): 670–683. <https://doi.org/10.1109/tase.2019.2941046>
- Roweis ST and Saul LK (2000) Nonlinear dimensionality reduction by locally linear embedding. *Science* 290(5500): 2323–2326. <https://doi.org/10.1126/science.290.5500.2323>
- Rucker DC and Webster RJ III (2011) Statics and dynamics of continuum robots with general tendon routing and external loading. *IEEE Transactions on Robotics* 27(6): 1033–1044. <https://doi.org/10.1109/tro.2011.2160469>
- Sanchez J, Corrales JA, Bouzgarrou BC, et al. (2018) Robotic manipulation and sensing of deformable objects in domestic and industrial applications: a survey. *The International Journal of Robotics Research* 37(7): 688–716. <https://doi.org/10.1177/0278364918779698>
- Saviolo A, Frey J, Rathod A, et al. (2023) Active learning of discrete-time dynamics for uncertainty-aware model predictive control. *IEEE Transactions on Robotics* 40: 1273–1291. <https://doi.org/10.1109/tro.2023.3339543>
- Schorp V, Panitch W, Shivakumar K, et al. (2023) Self-supervised learning for interactive perception of surgical thread for autonomous suture tail-shortening. In: *2023 IEEE 19th international conference on automation science and engineering (CASE)*, Auckland, New Zealand, 26–30 August 2023, pp. 1–6.
- Seita D, Florence P, Tompson J, et al. (2021) Learning to rearrange deformable cables, fabrics, and bags with goal-conditioned transporter networks. In: *2021 IEEE international conference on robotics and automation (ICRA)*. IEEE, pp. 4568–4575.
- Sen S, Garg A, Gealy DV, et al. (2016) Automating multi-throw multilateral surgical suturing with a mechanical needle guide and sequential convex optimization. In: *2016 IEEE international conference on robotics and automation (ICRA)*. IEEE, pp. 4178–4185.
- Servin M and Lacoursiere C (2008) Rigid body cable for virtual environments. *IEEE Transactions on Visualization and Computer Graphics* 14(4): 783–796. <https://doi.org/10.1109/TVCG.2007.70629>
- Shah A, Blumberg L and Shah J (2018) Planning for manipulation of interlinked deformable linear objects with applications to aircraft assembly. *IEEE Transactions on Automation Science and Engineering* 15(4): 1823–1838. <https://doi.org/10.1109/tase.2018.2811626>
- Shao H, Kugelstadt T, Hädrich T, et al. (2021) Accurately solving rod dynamics with graph learning. *Advances in Neural Information Processing Systems* 34: 4829–4842.
- She Y, Wang S, Dong S, et al. (2021) Cable manipulation with a tactile-reactive gripper. *The International Journal of Robotics Research* 40(12-14): 1385–1401. <https://doi.org/10.1177/02783649211027233>
- Shen Y, Franchi A and Gabellieri C (2025) Aerial robots carrying flexible cables: dynamic shape optimal control via spectral method model. *IEEE Transactions on Robotics* 41: 3162–3182. <https://doi.org/10.1109/tro.2025.3562459>
- Shetab-Bushehri M, Aranda M, Mezouar Y, et al. (2023) Lattice-based shape tracking and servoing of elastic objects. *IEEE Transactions on Robotics* 40: 364–381. <https://doi.org/10.1109/tro.2023.3331596>
- Shivakumar K, Viswanath V, Gu A, et al. (2023) SGTm 2.0: autonomously untangling long cables using interactive perception. *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 5837–5843.
- Sintov A, Macenski S, Borum A, et al. (2020) Motion planning for dual-arm manipulation of elastic rods. *IEEE Robotics and Automation Letters* 5(4): 6065–6072. <https://doi.org/10.1109/lra.2020.3011352>
- Soler C, Martin T and Sorkine-Hornung O (2018) Cosserat rods with projective dynamics. In: *Computer Graphics Forum*. Wiley Online Library, Vol. 37, 137–147. <https://doi.org/10.1111/cgf.13519>
- Song M and Huang P (2022) Dynamics and anti-disturbance control for tethered aircraft system. *Nonlinear Dynamics* 110(3): 2383–2399. <https://doi.org/10.1007/s11071-022-07742-7>
- Song Y, Yang K, Jiang X, et al. (2019) Vision based topological state recognition for deformable linear object untangling conducted in unknown background. In: *2019 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, pp. 790–795.
- Sorkine O and Alexa M (2007) As-rigid-as-possible surface modeling. In: *Symposium on Geometry processing*, Vol. 4, 109–116. Citeseer.
- Su Y, Jiang Y, Zhu Y, et al. (2022) Object gathering with a tethered robot duo. *IEEE Robotics and Automation Letters* 7(2): 2132–2139. <https://doi.org/10.1109/lra.2022.3141828>
- Süberkrüb F, Laezza R and Karayiannidis Y (2022) Feel the tension: manipulation of deformable linear objects in environments with fixtures using force information. In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 11216–11222.
- Sun Z, Zhou H, Nanbo L, et al. (2024) A robust deformable linear object perception pipeline in 3D: from segmentation to reconstruction. *IEEE Robotics and Automation Letters* 9(1): 843–850.
- Sun S, Wang X, Sanalidro D, et al. (2025) Agile and cooperative aerial manipulation of a cable-suspended load. *Science*

- Robotics* 10(107): eadu8015. <https://doi.org/10.1126/scirobotics.adu8015>
- Sundaresan P, Grannen J, Thananjeyan B, et al. (2020) Learning rope manipulation policies using dense object descriptors trained on synthetic depth data. In: 2020 IEEE international conference on robotics and automation (ICRA), Paris, France, 31 May - 31 August 2020. IEEE, pp. 9411–9418.
- Sundaresan P, Grannen J, Thananjeyan B, et al. (2021) Untangling dense non-planar knots by learning manipulation features and recovery policies. In: Proceedings of robotics: science and systems (RSS), Virtual, 12–16 July 2021.
- Szymko A, Kicki P and Walas K (2024) Calibrationless bimanual deformable linear object manipulation with recursive least squares filter. *IEEE Access* 12: 126707–126716. <https://doi.org/10.1109/access.2024.3438624>
- Tang T and Tomizuka M (2022) Track deformable objects from point clouds with structure preserved registration. *The International Journal of Robotics Research* 41(6): 599–614. <https://doi.org/10.1177/0278364919841431>
- Tang T, Fan Y, Lin HC, et al. (2017) State estimation for deformable objects by point registration and dynamic simulation. In: 2017 IEEE/RSJ international conference on intelligent robots and systems (IROS), Vancouver, Canada, 24–28 September 2017. IEEE, pp. 2427–2433.
- Tang Y, Chu X, Huang J, et al. (2024) Learning-based MPC with safety filter for constrained deformable linear object manipulation. *IEEE Robotics and Automation Letters* 9(3): 2877–2884. Available at: <https://doi.org/10.1109/lra.2024.3362643>
- Theetten A, Grisoni L, Andriot C, et al. (2008) Geometrically exact dynamic splines. *Computer-Aided Design* 40(1): 35–48. <https://doi.org/10.1016/j.cad.2007.05.008>
- Todorov E, Erez T and Tassa Y (2012) Mujoco: a physics engine for model-based control. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, pp. 5026–5033.
- Tong D, Choi A, Qin L, et al. (2024) Sim2real neural controllers for physics-based robotic deployment of deformable linear objects. *The International Journal of Robotics Research* 43(6): 791–810. <https://doi.org/10.1177/02783649231214553>
- Trommnau J, Kühnle J, Siegert J, et al. (2019) Overview of the state of the art in the production process of automotive wire harnesses, current research and future trends. *Procedia CIRP* 81: 387–392. <https://doi.org/10.1016/j.procir.2019.03.067>
- Tummers M, Lebastard V, Boyer F, et al. (2023) Cosserat rod modeling of continuum robots from newtonian and lagrangian perspectives. *IEEE Transactions on Robotics* 39(3): 2360–2378. <https://doi.org/10.1109/tro.2023.3238171>
- Umetani N, Schmidt R and Stam J (2014) Position-based elastic rods. In: Proceedings of the ACM SIGGRAPH/Eurographics symposium on computer animation, pp. 21–30.
- Valentini PP and Pennestrì E (2011) Modeling elastic beams using dynamic splines. *Multibody System Dynamics* 25: 271–284. <https://doi.org/10.1007/s11044-010-9232-9>
- Viswanath V, Grannen J, Sundaresan P, et al. (2021) Disentangling dense multi-cable knots. In: 2021 IEEE/RSJ international conference on intelligent robots and systems (IROS), Prague, Czech Republic, 27 September - 1 October 2021. IEEE, pp. 3731–3738.
- Viswanath V, Shivakumar K, Parulekar M, et al. (2023) HAND-LOOM: learned tracing of one-dimensional objects for inspection and manipulation. In: 7th annual conference on robot learning, Atlanta, Georgia USA, 6–9 Nov 2023.
- Waltersson GA, Laezza R and Karayiannidis Y (2022) Planning and control for cable-routing with dual-arm robot. In: 2022 international conference on robotics and automation (ICRA), Philadelphia, Pennsylvania, USA, 23–27 May 2022. IEEE, pp. 1046–1052.
- Wang T and Yamakawa Y (2022) Real-time occlusion-robust deformable linear object tracking with model-based Gaussian mixture model. *Frontiers in Neurorobotics* 16: 886068. <https://doi.org/10.3389/fnbot.2022.886068>
- Wang W, Berenson D and Balkcom D (2015) An online method for tight-tolerance insertion tasks for string and rope. In: 2015 IEEE international conference on robotics and automation (ICRA), Seattle, Washington, USA, 26–30 May 2015. IEEE, pp. 2488–2495.
- Wang A, Kurutach T, Liu K, et al. (2019) Learning robotic manipulation through visual planning and acting. *arXiv preprint arXiv:1905.04411*.
- Wang X, Zhang R, Kong T, et al. (2020) SOLOv2: dynamic and fast instance segmentation. *Advances in Neural Information Processing Systems* 33: 17721–17732.
- Wang Y, McConachie D and Berenson D (2021) Tracking partially-occluded deformable objects while enforcing geometric constraints. In: 2021 IEEE international conference on robotics and automation (ICRA), Xi'an, China, 30 May - 5 June 2021. IEEE, pp. 14199–14205.
- Wang C, Zhang Y, Zhang X, et al. (2022) Offline-online learning of deformation model for cable manipulation with graph neural networks. *IEEE Robotics and Automation Letters* 7(2): 5544–5551. <https://doi.org/10.1109/lra.2022.3158376>
- Weng Z, Zhou P, Yin H, et al. (2024) Interactive perception for deformable object manipulation. *IEEE Robotics and Automation Letters* 9(9): 7763–7770. <https://doi.org/10.1109/lra.2024.3431943>
- Wiese M, Berthold R, Wangenheim M, et al. (2023) Describing and analyzing mechanical contact for continuum robots using a shooting-based cosserat rod implementation. *IEEE Robotics and Automation Letters* 9(2): 1668–1675. <https://doi.org/10.1109/lra.2023.3346272>
- Wilson A, Jiang H, Lian W, et al. (2023) Cable routing and assembly using tactile-driven motion primitives. In: 2023 IEEE international conference on robotics and automation (ICRA), London, UK, 29 May 2023–02 June 2023, pp. 10408–10414.
- Wu S, Zhang J and Wu D (2022a) Equilibrium manipulation planning for a soft elastic rod considering an external distributed force and intrinsic curvature. *IEEE Robotics and*

- Automation Letters* 7(4): 11442–11449. <https://doi.org/10.1109/lra.2022.3199823>
- Wu W, Zhu Y, Zheng X, et al. (2022b) A novel cable-grasping planner for manipulator based on the operation surface. *Robotics and Computer-Integrated Manufacturing* 73: 102252. <https://doi.org/10.1016/j.rcim.2021.102252>
- Xiang J, Dinkel H, Zhao H, et al. (2023) TrackDLO: tracking deformable linear objects under occlusion with motion coherence. *IEEE Robotics and Automation Letters* 8(10): 6179–6186. Available at: <https://doi.org/10.1109/lra.2023.3303710>
- Xiang J, Lv Z, Xu S, et al. (2025) Structured 3D latents for scalable and versatile 3D generation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR), Nashville, TN, USA, 10–17 June 2025, pp. 21469–21480.
- Xu J, Gao L, Fierro R, et al. (2025) Airbender: adaptive transportation of bendable objects using dual UAVS. *IEEE Robotics and Automation Letters* 10(3): 2790–2797. <https://doi.org/10.1109/lra.2025.3536276>
- Yamakawa Y, Namiki A and Ishikawa M (2013) Dynamic high-speed knotting of a rope by a manipulator. *International Journal of Advanced Robotic Systems* 10(10): 361. <https://doi.org/10.5772/56783>
- Yan M, Zhu Y, Jin N, et al. (2020) Self-supervised learning of state estimation for manipulating deformable linear objects. *IEEE Robotics and Automation Letters* 5(2): 2372–2379. <https://doi.org/10.1109/lra.2020.2969931>
- Yang Y, Stork JA and Stoyanov T (2021) Learning to propagate interaction effects for modeling deformable linear objects dynamics. In: 2021 IEEE international conference on robotics and automation (ICRA), Xi'an, China, 30 May - 5 June 2021. IEEE, pp. 1950–1957.
- Yang Y, Stork JA and Stoyanov T (2022a) Learning differentiable dynamics models for shape control of deformable linear objects. *Robotics and Autonomous Systems* 158: 104258. <https://doi.org/10.1016/j.robot.2022.104258>
- Yang Y, Stork JA and Stoyanov T (2022b) Online model learning for shape control of deformable linear objects. In: 2022 IEEE/RSJ international conference on intelligent robots and systems (IROS), Kyoto, Japan, 23–27 October 2022. IEEE, pp. 4056–4062.
- Yang Y, Stork JA and Stoyanov T (2022c) Particle filters in latent space for robust deformable linear object tracking. *IEEE Robotics and Automation Letters* 7(4): 12577–12584. <https://doi.org/10.1109/lra.2022.3216985>
- Yin H, Varava A and Kragic D (2021) Modeling, learning, perception, and control methods for deformable object manipulation. *Science Robotics* 6(54): eabd8803. <https://doi.org/10.1126/scirobotics.abd8803>
- Yu M, Lv K, Zhong H, et al. (2022) Global model learning for large deformation control of elastic deformable linear objects: an efficient and adaptive approach. *IEEE Transactions on Robotics* 39(1): 417–436. <https://doi.org/10.1109/tro.2022.3200546>
- Yu M, Lv K, Wang C, et al. (2023a) A coarse-to-fine framework for dual-arm manipulation of deformable linear objects with whole-body obstacle avoidance. In: 2023 IEEE international conference on robotics and automation (ICRA), London, UK, 29 May 2023–02 June 2023. IEEE, pp. 10153–10159.
- Yu Z, Xu W, Yao S, et al. (2023b) Precise robotic needle-threading with tactile perception and reinforcement learning. In: Conference on robot learning, Atlanta, Georgia USA, 6–9 Nov 2023. PMLR, pp. 3266–3276.
- Yu M, Liang B, Zhang X, et al. (2024) In-hand following of deformable linear objects using dexterous fingers with tactile sensing. In: 2024 IEEE/RSJ international conference on intelligent robots and systems (IROS), Abu Dhabi, United Arab Emirates, 14–18 October 2024, pp. 13518–13524.
- Yu M, Lv K, Wang C, et al. (2025) Generalizable whole-body global manipulation of deformable linear objects by dual-arm robot in 3-D constrained environments. *The International Journal of Robotics Research* 44(4): 607–639. Available at: <https://doi.org/10.1177/02783649241276886>
- Yuan W, Dong S and Adelson EH (2017) Gelsight: high-resolution robot tactile sensors for estimating geometry and force. *Sensors* 17(12): 2762. <https://doi.org/10.3390/s17122762>
- Yuille AL and Grzywacz NM (1989) A mathematical analysis of the motion coherence theory. *International Journal of Computer Vision* 3(2): 155–175. <https://doi.org/10.1007/bf00126430>
- Zakaria MHD, Aranda M, Lequière L, et al. (2022) Robotic control of the deformation of soft linear objects using deep reinforcement learning. In: 2022 IEEE 18th international conference on automation science and engineering (CASE), Mexico City, Mexico, 20–24 August 2022. IEEE, pp. 1516–1522.
- Zanella R and Palli G (2021) Robot learning-based pipeline for autonomous reshaping of a deformable linear object in cluttered backgrounds. *IEEE Access* 9: 138296–138306. <https://doi.org/10.1109/access.2021.3118209>
- Zanella R, De Gregorio D, Pirozzi S, et al. (2019) DLO-In-Hole for assembly tasks with tactile feedback and LSTM networks. In: 2019 6th international conference on control, decision and information technologies (CoDIT), Paris, France, 23–26 April 2019. IEEE, pp. 285–290.
- Zanella R, Caporali A, Tadaka K, et al. (2021) Auto-generated wires dataset for semantic segmentation with domain-independence. In: 2021 international conference on computer, control and robotics (ICCCR), Shanghai, China, 8–10 January 2021. IEEE, pp. 292–298.
- Zhang TY and Suen CY (1984) A fast parallel algorithm for thinning digital patterns. *Communications of the ACM* 27(3): 236–239. <https://doi.org/10.1145/357994.358023>
- Zhang H, Ichnowski J, Seita D, et al. (2021a) Robots of the lost arc: self-supervised learning to dynamically manipulate fixed-endpoint cables. In: 2021 IEEE international conference on robotics and automation (ICRA), Xian, China, 30 May - 5 June 2021. IEEE, pp. 4560–4567.

- Zhang W, Schmeckpeper K, Chaudhari P, et al. (2021b) Deformable linear object prediction using locally linear latent dynamics. In: 2021 IEEE international conference on robotics and automation (ICRA), Xi'an, China, 30 May - 5 June, 2021. IEEE, pp. 13503–13509.
- Zhang X, Domae Y, Wan W, et al. (2022) Learning efficient policies for picking entangled wire harnesses: an approach to industrial bin picking. *IEEE Robotics and Automation Letters* 8(1): 73–80. <https://doi.org/10.1109/lra.2022.3222995>
- Zhang K, Li B, Hauser K, et al. (2024a) Adaptigraph: material-adaptive graph-based neural dynamics for robotic manipulation. In: Proceedings of robotics: science and systems (RSS), Delft, Netherlands, 15–19 Jul 2024.
- Zhang L, Bai K, Li Q, et al. (2024b) A collision-aware cable grasping method in cluttered environment. In: 2024 IEEE International Conference on Robotics and Automation (ICRA), Yokohama, Japan, 13-17 May 2024. IEEE, pp. 2126–2132.
- Zhang X, Domae Y, Wan W, et al. (2024c) A closed-loop bin picking system for entangled wire harnesses using bimanual and dynamic manipulation. *Robotics and Computer-Integrated Manufacturing* 86: 102670. <https://doi.org/10.1016/j.rcim.2023.102670>
- Zhang X, Lin HC, Zhao Y, et al. (2024d) Harnessing with twisting: single-arm deformable linear object manipulation for industrial harnessing task. In: 2024 IEEE/RSJ international conference on intelligent robots and systems (IROS), Abu Dhabi, United Arab Emirates, 14-18 October 2024. IEEE, pp. 4069–4075.
- Zhang K, Li B, Hauser K, et al. (2025) Particle-grid neural dynamics for learning deformable object models from RGB-D videos. In: *Proceedings of robotics: science and systems (RSS)*.
- Zhao TZ, Kumar V, Levine S, et al. (2023) Learning fine-grained bimanual manipulation with low-cost hardware. In: Proceedings of robotics: science and systems, Daegu, Republic of Korea, 10–14 July 2023.
- Zhao TZ, Tompson J, Driess D, et al. (2025) ALOHA unleashed: a simple recipe for robot dexterity. In: Conference on robot learning, Seoul, Republic of Korea, 27–30 September 2025. PMLR, pp. 1910–1924.
- Zhaole S, Zhu J and Fisher RB (2024) DexDLO: learning goal-conditioned dexterous policy for dynamic manipulation of deformable linear objects. In: 2024 IEEE international conference on robotics and automation (ICRA), Yokohama, Japan, 13-17 May 2024. IEEE, pp. 16009–16015.
- Zhi H, Zhang B, Qi J, et al. (2024) Non-prehensile object transport by nonholonomic robots connected by linear deformable elements. *IEEE Robotics and Automation Letters* 9(10): 8651–8658. <https://doi.org/10.1109/lra.2024.3440096>
- Zhou P, Zheng P, Qi J, et al. (2024) Reactive human–robot collaborative manipulation of deformable linear objects using a new topological latent control model. *Robotics and Computer-Integrated Manufacturing* 88: 102727. <https://doi.org/10.1016/j.rcim.2024.102727>
- Zhu J, Navarro B, Fraise P, et al. (2018) Dual-arm robotic manipulation of flexible cables. In: 2018 IEEE/RSJ international conference on intelligent robots and systems (IROS), Madrid, Spain, 01–05 October 2018. IEEE, pp. 479–484.
- Zhu J, Navarro B, Passama R, et al. (2019) Robotic manipulation planning for shaping deformable linear objects with environmental contacts. *IEEE Robotics and Automation Letters* 5(1): 16–23. <https://doi.org/10.1109/lra.2019.2944304>
- Zhu J, Navarro-Alarcon D, Passama R, et al. (2021) Vision-based manipulation of deformable and rigid objects using subspace projections of 2D contours. *Robotics and Autonomous Systems* 142: 103798. <https://doi.org/10.1016/j.robot.2021.103798>
- Zhu J, Cherubini A, Dune C, et al. (2022) Challenges and outlook in robotic manipulation of deformable objects. *IEEE Robotics and Automation Magazine* 29(3): 67–77. <https://doi.org/10.1109/mra.2022.3147415>
- Zimmermann S, Poranne R and Coros S (2021) Dynamic manipulation of deformable objects with implicit integration. *IEEE Robotics and Automation Letters* 6(2): 4209–4216. <https://doi.org/10.1109/lra.2021.3066969>
- Zitkovich B, Yu T, Xu S, et al (2023) RT-2: vision-language-action models transfer web knowledge to robotic control. In: Conference on robot learning (CoRL), Atlanta, Georgia USA, 6 to 9 November 2023. PMLR, pp. 2165–2183.
- Zürn M, Wnuk M, Schneider A, et al. (2022) Localization and tracking of deformable linear objects with self organizing maps. In: ISR Europe 2022; 54th international symposium on robotics, Munich, Germany, 20-21 June 2022, pp. 1–9.
- Zürn M, Kienzlen A, Klingel L, et al. (2023a) Deep learning-based instance segmentation for feature extraction of branched deformable linear objects for robotic manipulation. In: *2023 IEEE 19th International Conference on Automation Science and Engineering (CASE)*. IEEE, pp. 1–6.
- Zürn M, Wnuk M, Lechler A, et al. (2023b) Topology matching of branched deformable linear objects. In: 2023 IEEE international conference on robotics and automation (ICRA), London, UK, 29 May 2023–02 June 2023. IEEE, pp. 7097–7103.