



Learning to detect slip through tactile estimation of the contact force field and its entropy properties^{☆,☆☆}

Xiaohai Hu^a, Aparajit Venkatesh^a, Yusen Wan^a, Guiliang Zheng^a, Neel Jawale^a, Navneet Kaur^a, Xu Chen^{a,*}, Paul Birkmeyer^b

^a Department of Mechanical Engineering, University of Washington, United States of America

^b Amazon Robotics, United States of America

ARTICLE INFO

Keywords:

Tactile sensing
Slip detection
Object manipulation
Machine learning
Control system

ABSTRACT

Slip detection during object grasping and manipulation plays a vital role in object handling. Visual feedback can help devise a strategy for grasping. However, for robotic systems to attain a proficiency comparable to humans, integrating artificial tactile sensing is increasingly essential, especially in consistently handling unfamiliar objects. We introduce a novel physics-informed, data-driven approach to detect slip continuously for control-oriented tasks. Our work leverages the inhomogeneity of tactile sensor readings during slip events to develop distinct features and formulates slip detection as a classification problem. We test multiple data-driven models on 10 common objects under different loading conditions, textures, and materials to evaluate our approach. The resulting best classification algorithm achieves a high average accuracy of 95.61%. Practical application in dynamic robotic manipulation demonstrates the effectiveness of the proposed real-time slip detection and prevention.

1. Introduction

Tactile sensing, an essential sensory modality in humans, is pivotal during object manipulation and grasping [1]. Such sensing allows one to discern various object properties such as stiffness, weight, and surface texture. Humans can adjust their grip force independent of visual feedback, highlighting the crucial role of tactile sensing in this process. Johansson et al. in their 1984 study [2], demonstrated how humans utilize receptors in glabrous skin and sensorimotor memory for automatic precision grip control when handling objects with different surface textures. This research underscores a synergistic use of tactile feedback and sensorimotor memory in humans to predict and adjust to potential slip.

In robotic systems, incorporating tactile-based slip detection has proven vital in enhancing a robot's grip stability across varied operational dynamics (cf. Section 2). The focus of this paper is on identifying key slip characteristics and developing a framework for real-time detection and control of slips. Fig. 1 depicts the overall structure of the proposed slip-learning workflow. To achieve these objectives, our research introduces several innovations in the field of tactile-based slip detection and control:

- (1) Physics-aware analysis: Our approach integrates the displacement field at the object–gripper contact surface and its first derivatives, reflecting the friction field and its variations.
- (2) Entropy-based slip indication and its derivative: We incorporate entropy as a measure of randomness in the distributed marker displacements from tactile images, serving as a reliable indicator of potential slip events. Furthermore, we reveal the importance of the entropy derivative in understanding dynamic changes in the object–gripper interaction, allowing for more sensitive and timely indication of slip initiation.
- (3) Reduced reliance on prior knowledge of objects: Similar to how humans can still detect slippage with their eyes closed, our approach focuses on the interface interaction forces without requiring extensive prior knowledge of the object and its grasping conditions. This makes it potentially valuable for real-world applications where complete information is often unavailable or impractical to obtain.
- (4) Real-time detection and control: We develop a framework that enables swift corrective actions based on our slip detection

[☆] This paper was recommended for publication by Associate Editor Dr. Oliver Sawodny.

^{☆☆} This research was supported by a UW + Amazon Science Hub Gift-funded Robotic Research Project.

* Corresponding author.

E-mail addresses: huxh@uw.edu (X. Hu), venkat11@uw.edu (A. Venkatesh), yusenwan@uw.edu (Y. Wan), guiliang@uw.edu (G. Zheng), neelj42@uw.edu (N. Jawale), navneet@uw.edu (N. Kaur), chx@uw.edu (X. Chen), paubirkm@amazon.com (P. Birkmeyer).

<https://doi.org/10.1016/j.mechatronics.2024.103258>

Received 15 April 2024; Received in revised form 20 August 2024; Accepted 23 September 2024

Available online 1 October 2024

0957-4158/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

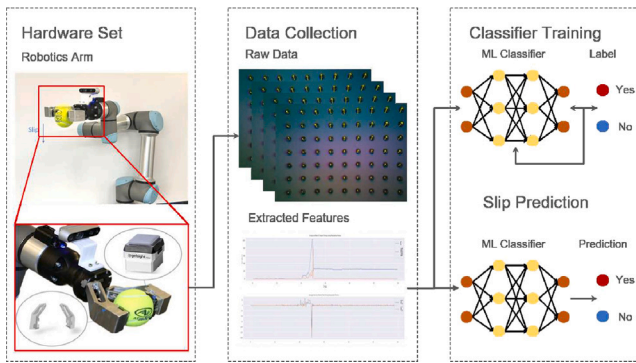


Fig. 1. Overview of the proposed hardware and software configurations for slip learning. Slip experiments were conducted using a UR5e robotic manipulator equipped with a Robotiq parallel gripper, Gelsight tactile sensors and their custom housing for data acquisition. Subsequent feature extraction and machine learning distill slip instances and predict occurrences of slippage.

method, enhancing the robot's ability to maintain stable grasps across varied operational dynamics.

These innovations collectively elevate the performance and application of slip detection in robotic systems. Our adoption of the entropy derivative, in particular, significantly enhances our detection system's performance by directly quantifying the rate at which disorder changes, providing an essential metric for the early detection of slips. In the remainder of the paper, we will first introduce the basic working principle of the tactile sensor and related works in Section 2, followed by the hardware setup and the extraction of features from tactile images in Section 3. Section 4 introduces how we collect the data and the classification methods we chose. We will then present the results given by various data-driven methods, comparison with existing method and validate the generalizability and future selection efficiency evaluation in Section 5. To demonstrate the real-world applicability of our proposed method, we illustrate its deployment in a robotic system that has been integrated with our slip detection and prevention algorithm, particularly emphasizing its performance during dynamic manipulation tasks. Section 6 concludes the paper and discusses the future work.

2. Background and related work

2.1. Tactile sensor

The growth of tactile sensing is pivotal in robotic dexterity. Tactile sensing can be done intrinsically or extrinsically [3]. Examples of intrinsic tactile sensing include measuring contact forces via joint torques [4, 5] or transmission cable tension in equipped manipulators [6]. Extrinsic sensors, mounted on the robot hand's exterior, use such technologies as piezoelectric sensors for force and pressure detection [7,8]. Multi-modal sensors, such as BioTac [9], provide varied data like force and temperature, but may lack resolution in complex force environments. Optical tactile sensors like Tactip [10] and others [11] provide representations of surface interactions for deducing force. Among these, the GelSight sensor stands out for its precise measurement of contact surface geometry [12,13]. The GelSight sensor utilizes a camera to record the deformation of a reflective, gel-coated transparent elastomer under applied force. This process generates images during surface contact, facilitating the creation of a depth map for the contact region. Specifically, our project employs the GelSight Mini sensor, featuring marker dots on its cartridge.

2.2. Slip detection in robotics

Slip detection has been a focus in robotics for decades, with various methods proposed over the years. Starting with the skin acceleration



Fig. 2. Design of adapter for housing Gelsight tactile sensors. A 45-degree flange extension was designed for the end-effector to extend the opening distance of the gripper to 85 mm. The adapters are mounted on the Robotiq Hand-e adaptive gripper through four M2.5 screws.

sensor for slip and texture detection by Howe et al. in 1989 [14], the field has evolved significantly. By 2004, methods like Ikeda et al.'s camera-based slip detection [15], and by 2012, Maldonado et al.'s fingertip sensing for object characteristics [16], have been introduced. Recent advancements primarily employ tactile sensing. Veiga et al. combined traditional tactile sensors with machine learning for a 75% accurate slip detection in 2015 [17]. Later developments include James et al.'s vision-based tactile sensors in 2018 [18], Dong et al.'s incipient slip detection in 2018 [19], and Li et al.'s visual-tactile deep neural network in 2019 [20]. Griffa et al.'s 2022 study used deep neural networks for force distribution classification [11], while Juddy et al. explored soft force sensors for deformable object grasping [21]. Yuan et al. [22] concentrated on processing tactile data. Their investigation revealed a correlation between the irregularity of gel deformation displacement and the onset of slip, signifying an early stage of entropy-inspired slip detection research. Despite these advancements, broader adoption of entropy in slip detection remains underexplored, particularly within the domain of machine learning. Drawing from statistical properties of a vector field, the proposed method takes a step change to incorporate not only entropy but also the derivative of entropy as core features during a slip. Such insight turns out pivotal for successfully applying machine learning classifiers in slip detection tasks.

3. Hardware setup and tactile data

3.1. Hardware setup

Our hardware setup includes a UR5e robotic arm, a Robotiq Hand-E gripper, and custom adapters for two Gelsight tactile sensors. The sensors are incorporated into tailored parallel end-effectors on the Robotiq Hand-e gripper and secured with screws. As shown in Fig. 2, the end-effectors were designed to extend outwards along the movement directions of the gripper to create a large operational range of the system for a wide variety of target objects. The end-effector adapters, machined from 6061-T6 Aluminum for strength and stability, enable firm grasping of diverse objects. The Gelsight mini sensor can capture high-resolution imprints of the contact surface, with a resolution of 3840×2160 . The sensor is coated with a Lambertian silicone gel layer and has a surface area of $18.6(H) \times 14.3(V)$ mm². As the primary objective of this paper is to learn the contact force field during slip, our slip detection approach focuses on analyzing data from a single sensor and depth camera images were not utilized in this current study. All equipment was operated using Ubuntu 20.04 on a PC equipped with an Intel Core i5-10210U CPU at a clock speed of $1.60 \text{ GHz} \times 8$.

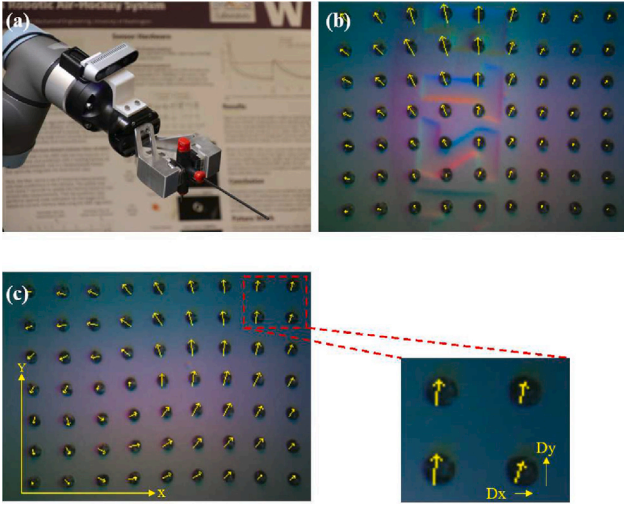


Fig. 3. (a) Gripper grasping a T-handle hex key, (b) the displacement of individual markers overlaid on the tactile image, and (c) zoomed-in section of GelSight image, illustrates the gel deformation through arrows resulting from contact, denoted as D_{x_i} and D_{y_i} , and referred to from hereon as the displacement field of the markers.

3.2. Translating tactile frames to features

The camera within the GelSight tactile sensor captures the surface deformation dynamics as video sequences. This process is enhanced by the presence of 63 black markers, strategically arranged in a 7×9 matrix on the sensor's surface. These markers are crucial for enabling precise measurements of surface deformation. By observing the displacement of these markers over time, we can accurately determine the extent of deformation. We process each video frame and compare it to the initial frame to assess gel deformation. To quantitatively depict this deformation, yellow arrows are drawn on each frame (Fig. 3). The specific methodologies employed are enumerated as follows: We use the canny filter for marker selection and OpenCV's SimpleBlobDetector algorithm to determine their centers. By comparing marker positions from the initial to current frames, we compute each marker's average displacement. In a stable grasp, the gel moves uniformly under the shear force indicated by evenly distributed arrows. However, during slip, this uniformity is disrupted, leading to increased entropy and inhomogeneity, as well as a significant rise in entropy derivative. This increase in the derivative of entropy turns out to be a valuable feature for the classifier in subsequent stages.

Feature \bar{V}_x and \bar{V}_y : In addition to the raw displacement field, we have observed that changes in the vector field magnitude serve as significant indicators of slip. Given the sampling rate of 25 Hz for the tactile sensor, we introduce features that capture the change in magnitude between consecutive frames of tactile samples. These features, averaged over horizontal and vertical directions, can be analogously regarded as “velocities” of the tactile surface. This approach aligns with the findings reported by Jame [23], who demonstrated the efficacy of velocity-based features in slip detection. To quantify these changes, we define the discrete-time mean velocity features \bar{V}_x and \bar{V}_y for the x and y components. For the x -component:

$$v_{x_i}(t) = f \cdot \Delta D_{x_i}(t) \quad (1)$$

$$\bar{V}_x(t) = \frac{1}{n} \sum_{i=1}^n v_{x_i}(t) \quad (2)$$

Similarly, for the y -component:

$$v_{y_i}(t) = f \cdot \Delta D_{y_i}(t) \quad (3)$$

$$\bar{V}_y(t) = \frac{1}{n} \sum_{i=1}^n v_{y_i}(t) \quad (4)$$

Here, f (=25 Hz) is the sampling frequency, n (=63) represents the number of data points, $D_{x_i}(t)$ and $D_{y_i}(t)$ represent the positions of the i th data point at time t , and $\Delta D_{x_i}(t)$ is defined as $D_{x_i}(t) - D_{x_i}(t - \delta t)$, where δt is the time step between consecutive measurements

Feature E and $\frac{\delta E}{\delta t}$: The marker flow displacement inhomogeneity, quantified as entropy, serves as a key metric in our slip detection approach. We define this entropy at time t , $E(\Gamma(t))$, mathematically as:

$$E(\Gamma(t)) = - \int_{\Gamma(t)} p(r(t)) \log p(r(t)) dr \quad (5)$$

Our analysis employs the discrete version of Shannon entropy, a choice validated by prior research such as Yuan et al. [22]. We define the histogram Γ to represent the frequency distribution of the displacement field magnitude. The magnitude γ is calculated as:

$$\gamma = \xi \sqrt{D_x^2 + D_y^2} \quad (6)$$

where D_x and D_y represent the displacements in the x and y directions, respectively. The magnitude is scaled by a factor ξ of 10, transforming the original range of 0–3 mm to 0–30 units. This scaling facilitates more granular categorization in the histogram. We classify γ into 30 categories using the floor function, (0, 1, ..29), such that each arrow length falls into one of these categories. The corresponding probability density function is represented by $p(\gamma)$. Fig. 4 illustrates this concept, depicting a tactile image with 63 arrows of varying lengths and its corresponding histogram. In this particular instance, the data naturally falls into 15 distinct bins. Our discrete entropy calculation is thus:

$$E(\Gamma(t)) = - \sum_{i=1}^n p(\gamma_i) \log p(\gamma_i) \quad (7)$$

where n represents the number of observed distinct states (e.g., 15 in Fig. 4), $\gamma_1, \gamma_2, \dots, \gamma_n$ are the unique states, and $p(\gamma_1), p(\gamma_2), \dots, p(\gamma_n)$ their associated probabilities.

When an object begins to slip, the displacement field becomes more inhomogeneous due to the non-uniform contact forces that arise throughout the contact surface as the object moves. The inhomogeneity is more significant around the edges of the contact region, resulting in a non-uniform displacement field and increased entropy, as illustrated in Fig. 5.

In Fig. 5, we see that the entropy increases when the grippers initially come into contact with the object, and this value remains relatively constant as long as the object is securely grasped. However, when the object begins to slip, a sharp increase in entropy is observed. Thus, it is evident that reasonably high entropy values can exist even when an object is securely grasped. Additionally, the entropy remains almost constant when a secure grasp is established. To further enhance the classification of slip, we propose integrating the rate of change of entropy as another feature to feed into the classifier. The rate of change of entropy is calculated as follows:

$$\frac{\delta E}{\delta t} \approx f \cdot [E(\Gamma(t)) - E(\Gamma(t - \Delta t))] \quad (8)$$

where E is the entropy and Δt is the sampling time. Since entropy varies with object characteristics, we trained a classifier using data from objects of diverse shapes, sizes, and materials to categorize slips effectively.

4. Method

4.1. Data acquisition

Our data acquisition protocol was designed to capture a comprehensive range of object manipulation scenarios, with a focus on

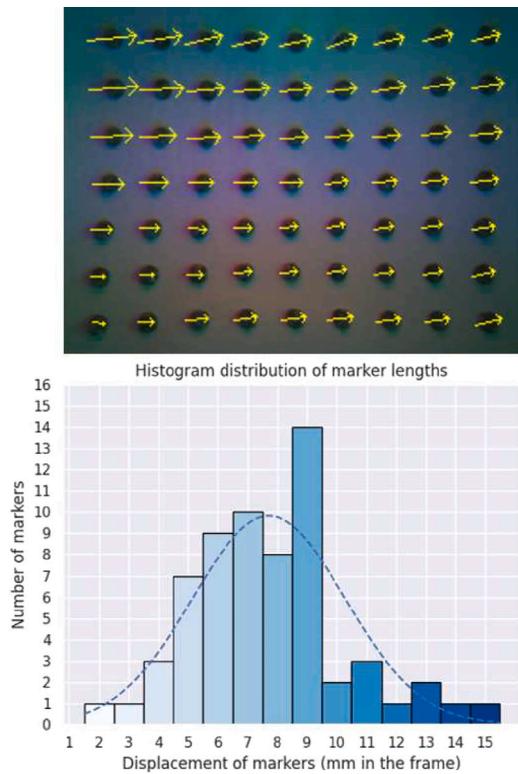


Fig. 4. Histogram illustrating the distribution of marker flow. Each bar represents a distinct state, with its height indicating the relative occurrence of that state within the overall distribution.

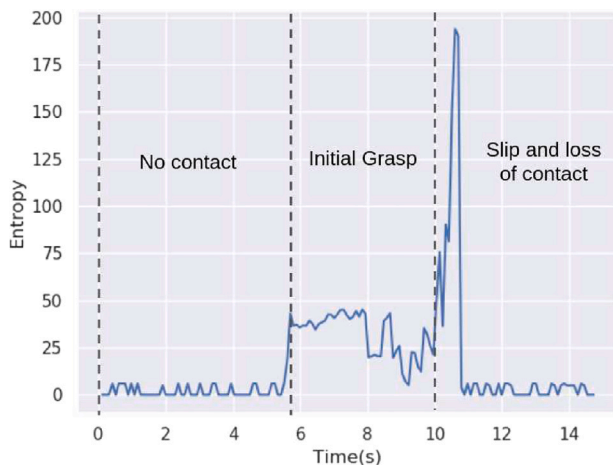


Fig. 5. A slip trial was conducted to illustrate the change of entropy from the no contact to object, through the initial grasp, to the incipient slip, and ultimately to the loss of contact to object. Initially, a rectangular cardboard box was held between the grippers (embedded with GelSight sensor) of the robot, at this stage the entropy was almost zero. The grippers were brought closer together till there was an initial contact and a gentle grasp of the box, leading to a notable increase and subsequent stabilization of entropy. Following this, the robotic arm was maneuvered in a manner that induced slippage of the object. At the moment of slip, a sharp spike in entropy was observed and then the entropy returned zero once the robotic gripper with the tactile sensor lost complete contact.

detecting object slippage. We selected 10 objects of varying materials and shapes commonly encountered. To enhance the generalizability of our findings, we positioned these objects at diverse locations on a table. The data collection process followed a systematic approach. First, we devised a predetermined grasping sequence incorporating 12

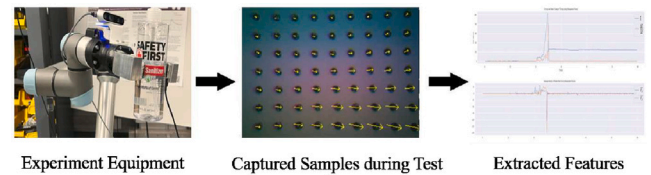


Fig. 6. Illustration of the data acquisition process.

distinct motion primitives. These included translational and rotational movements across all axes of the Cartesian coordinate system, as well as complex planar and spherical motions. This diverse set of manipulations was specifically designed to elicit a broad spectrum of potential slip scenarios. For each trial, the manipulator was programmed to approach the object with precision, grasp it, and lift it to a standardized position to eliminate initial condition variability. Only after confirming a secure grasp did we initiate the pre-defined manipulation sequence. Throughout this process, we continuously collected tactile sensor data (Fig. 6). To induce slip, we employed multiple methods. These included the manual application of incremental external loads to the object and the use of a precision rotary stage for controlled angular displacements. This comprehensive approach enabled the capture of both translational and rotational slip events. We also applied external forces in our non-slip dataset to ensure our model could distinguish between force application and actual slippage. Our data labeling process involved a meticulous frame-by-frame manual classification of the tactile sensor output as either “slip” or “non-slip”. We paid particular attention to transition points where the object state changed from static to slip. When such a transition was identified, we precisely marked the exact frame at which it occurred. All frames preceding this transition point were labeled as non-slip, while those following it were categorized as slip. This labeling continues until the object slips out of grasp entirely, at which point the remaining data is truncated. To ensure data integrity, we conducted post-collection verification through thorough review of the collected data and corresponding high-speed video recordings. Instances of unsuccessful grasps were identified and excluded from the dataset.











4.2. Classification method

Following the feature extraction phase, we utilized machine learning classifiers to identify the states of object grasping. Given the inherent nonlinearity in the task, we experimented with various classifiers to understand the intricate slip dynamics. We selected four algorithms:

- Support Vector Machine (SVM): $g(x) = \text{sign}(w^T x + b)$.
- Random Forest (RF): $g(x) = \frac{1}{T} \sum_{t=1}^T h_t(x)$.
- K-Nearest Neighbor (KNN): $g(x) = \text{mode}(y_i : x_i \in N_k(x))$.
- Decision Tree (DT): $g(D, A) = H(D) - H(D|A)$.

In this framework, the function $g(x)$ represents the probability that the input x belongs to a particular class. For Support Vector Machine (SVM), the parameters w and b denote the weight vector and the bias, respectively. In the context of Random Forest (RF), $h_t(x)$ is the output of the t th decision tree. The K-Nearest Neighbors (KNN) algorithm determines the most common class among the k closest neighbors to x through the use of the mode function. Decision Trees (DT) assess the significance of features by computing information gain, $g(D, A)$, which involves calculating the difference between the information entropy, $H(D)$, and the conditional entropy, $H(D|A)$. In the next section, we will assess these classifiers' accuracy using our extracted features.

Table 1
Performance of the developed slip detection for objects with different materials and surface characteristics.

| | Screw drive | Tennis ball | Toy duck | Mouse | Box | High-lighter | Toy bear | Scrub sponges | Toy owl | Floss | Total |
|-----------|---|---|---|---|---|---|---|---|---|---|--------|
| |  |  |  |  |  |  |  |  |  |  | |
| Accuracy | 97.17% | 94.49 % | 98.82 % | 93.60 % | 98.43 % | 97.75 % | 98.97% | 93.92% | 83.87% | 97.13% | 95.42% |
| Precision | 95.21% | 93.58% | 98.36% | 89.66% | 98.01% | 99.18% | 98.37% | 89.16% | 93.94% | 100.00% | 95.55% |
| Recall | 99.29% | 93.58% | 98.36% | 97.01% | 98.67% | 96.03% | 99.18% | 98.50% | 88.57% | 93.13% | 96.23% |
| F1 Score | 97.20% | 92.73% | 98.36% | 93.19% | 98.34% | 97.58% | 98.78% | 93.57% | 91.18% | 96.44% | 95.74% |

5. Experiment results

Experiments are performed to analyze the proposed new features. To thoroughly gauge their efficacy, we juxtapose these features against basic alternatives such as \vec{V}_x and \vec{V}_y . These experiments consist of two parts: Part I focuses on classification accuracy, and Part II sheds light on generalization capabilities. Additionally, we visualize the effectiveness of feature selection, examine the computational time of inference, and provide a demonstration of using the proposed method to slide a book out of a shelf while preventing slip.

5.1. Experiment sets

We implemented the chosen classifiers via Python's scikit-learn library. Our dataset includes approximately 14,000 samples among ten different objects. Here, a "sample" refers to a sequence of tactile images extracted from the sensor's video stream. From the samples, we derive features for classifier training. The features of one sample consist of one feature vector, which is the input of the machine learning model. Thus, one sample has one feature vector and there are 14,000 feature vectors in the dataset. To assess the accuracy and generalization of the proposed features, we conducted two sets of experiments.

In our initial experiment, we employed four machine learning classifiers to evaluate the efficacy of our proposed method in comparison to existing approaches. To comparative analysis, we implemented the velocity-based method developed by James et al.'s [23] and adapted the research of Yuan et al.'s [22] by focusing on the entropy feature they proposed. James et al.'s approach primarily utilizes velocity-related information for grasping state classification. We extended this concept by incorporating the entropy feature (E) proposed by Yuan et al. into what we term the "entropy-baseline" method. This method employs features including mean velocity components \vec{V}_x , \vec{V}_y , and entropy (E). Our proposed method builds upon these foundations by introducing $\frac{\delta E}{\delta t}$ as a novel feature, ultimately utilizing \vec{V}_x , \vec{V}_y , E, and $\frac{\delta E}{\delta t}$ in our classifiers. By consolidating all samples, we adopted a *five-fold cross-validation* technique. This method splits the samples into five sections, each holding 2800 samples. During each iteration, four sections were designated for classifier training, leaving one for evaluation. The final accuracy was the average over the five cycles.

In the second experiment, we focus on KNN, the best-performing classifier from the first experiment, to gauge the feature's adaptability across various objects. In this case, we divided the samples based on object categories, resulting in ten distinct subsets of samples. Here, samples were categorized based on object types, creating ten unique sample subsets. In each test, nine subsets were used for training and the tenth for evaluation. It is important to highlight that the classifiers were tasked with categorizing samples from unfamiliar objects not present in the training dataset. The accuracy thus reflects the feature's adaptability.

Accuracy assessment involved four metrics: accuracy, recall, precision, and F1 score. Accuracy signifies the percentage of correctly classified instances relative to the total number of instances. Precision measures the classifier's ability to distinguish true positives from false

Table 2
Comparative analysis.

| Methods | Metrics | DT | RF | SVM | KNN |
|--------------|-----------|-------|-------|-------|-------|
| James et al. | Accuracy | 82.20 | 84.22 | 89.30 | 80.25 |
| | Precision | 94.72 | 92.63 | 83.30 | 94.25 |
| | Recall | 75.91 | 79.56 | 93.58 | 74.66 |
| | F1 | 83.94 | 85.15 | 88.13 | 82.68 |
| E baseline | Accuracy | 92.73 | 94.53 | 84.85 | 93.43 |
| | Precision | 91.95 | 94.06 | 74.42 | 91.74 |
| | Recall | 92.78 | 94.45 | 92.37 | 94.34 |
| | F1 | 92.35 | 94.26 | 82.42 | 93.02 |
| Our methods | Accuracy | 93.14 | 95.25 | 94.88 | 95.61 |
| | Precision | 92.54 | 95.22 | 91.19 | 95.07 |
| | Recall | 93.07 | 94.86 | 97.94 | 95.76 |
| | F1 | 92.80 | 95.04 | 94.44 | 95.40 |

positives, with both precision and accuracy ideally being high for a perfect classifier. Recall assesses the classifier's capacity to distinguish true positives from false negatives. The F1 score, the harmonic mean of precision and recall, offers a balanced assessment of model performance.

For hyperparameter tuning, a grid search was conducted during training for each classification algorithm. The best results were observed with KNN classifier, setting the nearest neighbor hyperparameter to 1, and SVM using the RBF kernel and a regularization parameter of 1. Hyperparameter adjustments did not enhance the performance of the RF and DT classifiers, so their default settings were maintained.

5.2. Experiment I result: Accuracy of slip detection

The outcomes of Experiment I, as detailed in Table 2, reveal that incorporating entropy and its rate of change significantly enhances classifier accuracy across all models. In comparison to James et al.'s method, the entropy-baseline method showed substantial accuracy improvements for DT, RF, and KNN by 10.53%, 10.31%, and 13.18%, respectively. Similar trends were observed in other performance metrics, including precision, recall, and F1 score. Our proposed method, which integrates the novel feature $\frac{\delta E}{\delta t}$, further improved classification accuracy compared to the entropy-baseline method. Specifically, we observed enhancements of 0.41%, 0.7%, 10.03%, and 2.18% for DT, RF, SVM and KNN. Moreover, our method consistently outperformed James et al.'s approach across all evaluated metrics.

Table 1 showcases the classification accuracy of different objects using KNN. Most objects exhibit an accuracy greater than 95%, with only the owl falling below 90%. The results show that the performance of the classifier for each category is balanced, and the proposed features have broader usage rather than can be used on specific objects.

5.3. Experiment II result: Generalizability of the method

The results of Experiment II, presented in Table 3, involved training classifiers with samples from nine objects and testing them with samples from a different, single object each time. In the table, the object

Table 3
Cross-validation results for 10 different objects.

| Feature set | Metric | Experiment index and valid object | | | | | | | | | |
|-------------------------------------|-------------|-----------------------------------|-------|-------|-------|-------|-------------|-------|--------|-------------|--------|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| | | Ball | Bear | Box | Duck | Floss | Highlighter | Mouse | Owl | Screwdriver | Sponge |
| \bar{V}_x and \bar{V}_y | Accuracy % | 42.92 | 47.93 | 45.95 | 48.79 | 40.13 | 49.29 | 42.69 | 79.27 | 53.5 | 38.46 |
| | Precision % | 89.81 | 88.82 | 75.38 | 90.1 | 91.78 | 86.5 | 89.42 | 77.43 | 89.42 | 88.2 |
| | Recall % | 45.12 | 44.76 | 44.04 | 42.87 | 40.94 | 48.29 | 44.96 | 100.00 | 53.91 | 40.54 |
| | F1% | 60.06 | 59.53 | 55.6 | 58.1 | 56.62 | 61.98 | 59.84 | 87.28 | 67.26 | 55.55 |
| All features | Accuracy % | 71.24 | 80.72 | 77.84 | 89.07 | 61.81 | 84.02 | 70.44 | 91.06 | 58.37 | 84.63 |
| | Precision % | 93.52 | 94.57 | 92.31 | 96.25 | 91.17 | 93.09 | 90.41 | 90.27 | 96.96 | 91.15 |
| | Recall % | 63.52 | 70.64 | 68.89 | 80.05 | 53.00 | 77.82 | 63.34 | 100.00 | 56.43 | 77.54 |
| | F1% | 75.66 | 80.87 | 78.90 | 87.41 | 67.03 | 84.77 | 74.49 | 94.88 | 71.34 | 83.80 |
| E and $\frac{\delta E}{\delta t}$ | Accuracy % | 91.67 | 86.91 | 91.09 | 97.58 | 93.16 | 89.27 | 87.21 | 91.73 | 97.46 | 90.66 |
| | Precision % | 90.43 | 88.18 | 90.35 | 94.45 | 92.15 | 91.99 | 91.26 | 91.00 | 95.90 | 89.60 |
| | Recall % | 91.99 | 82.63 | 89.85 | 99.37 | 91.81 | 86.43 | 83.48 | 100.00 | 99.32 | 89.04 |
| | F1% | 91.20 | 85.32 | 90.10 | 96.85 | 91.98 | 89.13 | 87.20 | 95.29 | 97.58 | 89.32 |

Table 4
Performance metrics for unseen objects.

| Unseen object | Accuracy % | Precision % | Recall % | F1 % |
|------------------|------------|-------------|----------|-------|
| Book | 71.69 | 88.45 | 59.17 | 70.91 |
| Raccoon | 93.73 | 97.03 | 91.02 | 93.93 |
| Contact solution | 86.66 | 95.09 | 80.32 | 87.08 |

name indicates the test object. For instance, “box” implies training with samples from all objects except the box and testing with box samples.

The findings suggest classifiers perform better with entropy (E) and its rate of change ($\frac{\delta E}{\delta t}$) than with \bar{V}_x and \bar{V}_y when encountering unfamiliar objects. A comparison of Tables 2 and 3 shows a decline in accuracy when facing new objects. However, when using E and $\frac{\delta E}{\delta t}$ as features, the accuracy only decreases less than 6% compared with the accuracy in experiment I. In contrast, reliance on \bar{V}_x and \bar{V}_y leads to a more substantial decrease in accuracy; for example, accuracy for “sponge” drops to 38.46% and “mouse” to 42.69%. The decline is also notable when all features are used together.

These results lead to the conclusion that E and $\frac{\delta E}{\delta t}$ offer better generalization capabilities. They can be directly applied to new objects without substantial performance reduction. Conversely, the generalizability of \bar{V}_x and \bar{V}_y is limited and can negatively impact classification. The use of all features combined does not enhance model performance compared to using only E and $\frac{\delta E}{\delta t}$. Additionally, we assessed the performance of the trained classifier on unseen objects, as shown in Table 4. The results indicate that the classifier maintains good accuracy with unfamiliar objects.

5.4. Feature visualization and analysis

As shown in Fig. 7, we employ t-SNE to visualize the extracted features, where each point represents a sample characterized by its respective features. In Fig. 7(a), the division between E and $\frac{\delta E}{\delta t}$ is clearly evident. Despite some overlap, samples from different classes predominantly occupy distinct areas. Conversely, Fig. 7(b) illustrates that the visualization of \bar{V}_x and \bar{V}_y is less distinct. The points representing “static” and “slip” classes significantly overlap, indicating less effective separation. This visualization suggests that E and $\frac{\delta E}{\delta t}$ lead to more pronounced classification behavior.

5.5. Computational time of inference

During each of the 10 trials, we captured 200 prediction values and evaluated the classifiers’ inference times using Python’s `time.time()` function. With the GelSight mini tactile sensor operating at 25 FPS, there is a 40 ms window for each frame processing. The average inference times and their usage percentages are:

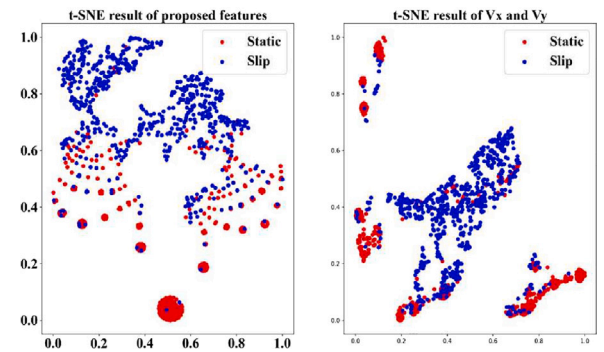


Fig. 7. t-SNE visualization of features: (Left) Entropy (E) and its rate of change ($\frac{\delta E}{\delta t}$); (Right) \bar{V}_x and \bar{V}_y .

- Support Vector Machine (SVM): 0.33 ms (0.825% of frame time).
- Random Forest (RF): 0.54 ms (1.35% of frame time).
- K-Nearest Neighbor (KNN): 0.94 ms (2.35% of frame time).
- Decision Tree (DT): 0.29 ms (0.725% of frame time).

All classifiers efficiently fit within the available frame processing time, making them suitable for real-time tactile image processing with the GelSight mini sensor.

5.6. Sliding out a book from a shelf

This experiment involved a UR5e robot equipped with a Robotiq parallel gripper, tasked with retrieving and repositioning a book from a shelf. The robot had no prior knowledge of the book’s weight or stiffness, necessitating a cautious approach with minimal initial clamping force to prevent potential damage. The task began with positioning the parallel gripper adjacent to the bookshelf without making initial contact with the book. The gripper was then maneuvered towards the book, and a safe grasp was executed using appropriate parameters, which entailed a relatively low grasping force. No slip occurred during the pre-manipulation phase. However, during the subsequent stage of extracting the book, the absence of a slip detection method and corresponding prevention led to slippage, hindering task completion.

Repeating the experiment with our slip detection algorithm and integrating a slip-prevention force control into the robotic system showed significant improvement. Inspired by work in [17,24], we developed a control system combining a Proportional–Derivative (PD) controller with an integrator-structured adjustment. This system utilizes the mean marker displacement (\bar{D}) as feedback, regulating it to a target value (D_t). The integrator adjusts D_t based on the slip detector result (S):

$$D_t[n] = D_t[n-1] + \alpha \cdot S \quad (9)$$

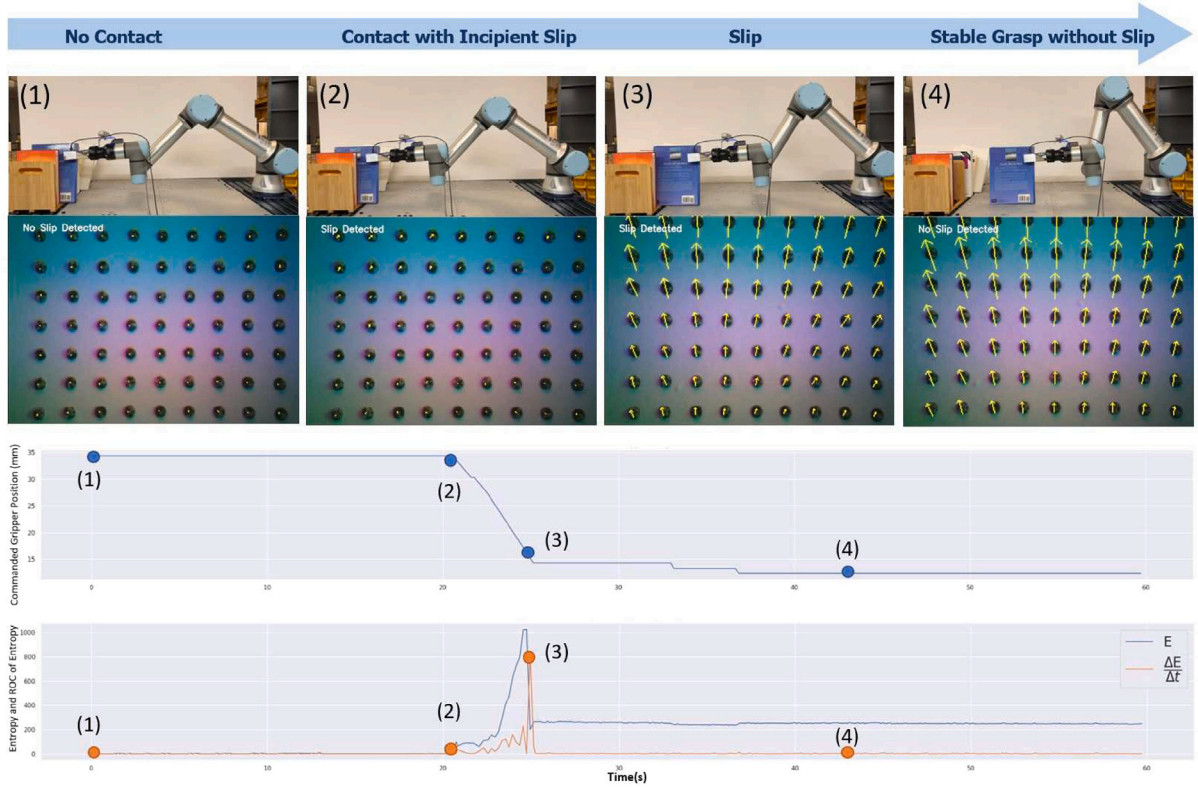


Fig. 8. A demonstration of sliding a book out of a shelf is presented, consisting of multiple stages of grasping. The initial row of images portrays the progressive grasping stages, commencing from (1) a static initial grasp, advancing towards (2) an incipient slip at the start of manipulation, further to (3) an actual slip, and culminating at (4) a stable grasp. The subsequent row exhibits the data obtained from a tactile sensor and the corresponding real-time detection of slip. The third row displays the command gripper distance to forestall slippage, whereas the fourth row depicts E and $\frac{\delta E}{\delta t}$ throughout the grasping procedure.

where S is 1 for detected slip and 0 for non-slip, and α is a fixed factor. To regulate the clamping force, we implement a position control strategy. It is noteworthy that the choice of position control is influenced by the gripper's lack of direct force adjusting capability and is subject to the object's impedance characteristics [25]. The incremental control input Δu is expressed as:

$$\Delta u[n] = K_p \cdot e[n] + K_d \cdot (e[n] - e[n-1]) \quad (10)$$

where $e[n] = \bar{D}[n] - D_t[n]$, K_p and K_d are proportional and derivative coefficients, respectively. Upon slip detection, the algorithm incrementally increases the gripper position value by Δu , typically 2–5 units per control cycle, depending on slip magnitude. Given the gripper's position control API range (0–255 units) corresponding to a clamping force range of 20–185 N, this translates to approximately 0.6–3 N increase in clamping force per cycle. The control system along with the frame fresh rate operate at 25 Hz, with each cycle lasting 40 ms. The algorithm promptly detected the initial slip during book extraction. Despite the increase in grasping force, the smooth book cover and the drag exerted by the gripper caused the incipient slip to increase. The slip-prevention algorithm responded by gradually increasing the grasping force until a stable grasp was achieved, which was maintained for a brief duration to ensure secure gripping of the book. While maintaining the grasping force, the robot slid the book out of the bookshelf and adjusted the force as necessary to prevent further slips. After the book is extracted, the grasping force is gradually reduced to release the book. With the slip detection and prevention algorithms implemented, the grasp was continuously sustained, and the manipulation task was accomplished using the same set of initial grasping parameters. Fig. 8 shows the grasping stages of the book retrieval, along with corresponding E and $\frac{\delta E}{\delta t}$. The figure indicates that E and $\frac{\delta E}{\delta t}$ were almost negligible when the book is held with safe grasping parameters. However, as the

manipulation process begins, E and $\frac{\delta E}{\delta t}$ slightly increased, indicating the need for a grip adjustment. As the manipulation process proceeds, those values rose, indicating the occurrence of more slips. To prevent slip, the gripping pose was further modified until the E reached a constant value and $\frac{\delta E}{\delta t}$ approached zero. Notably, the E value is higher at the end of the manipulation process than at the start, which can be attributed to the dynamic forces acting on the book during the manipulation process, resulting in non-homogeneity in the marker field. The classifier achieved a 71.69% accuracy in categorizing the book's grasping state. This experiment demonstrates that monitoring changes in entropy profile serves as an effective method for slip detection and prevention, thereby enhancing the success rate of manipulation tasks.

6. Conclusions and future work

This paper proposes a novel approach for continuous slip detection using modern optical tactile sensors. Our method employs a physics-informed, data-driven strategy that leverages the distributed contact force field, its entropy, and the rate of change of entropy extracted from tactile sensors.

The proposed method facilitates the monitoring of incipient slip and the prediction of slip occurrences during object manipulation tasks. Furthermore, we delineate a control strategy to mitigate slip events upon detection. Upon training the slip detection classifier with a sufficient corpus of objects across diverse categories, the method demonstrates potential for generalization to slip detection and prevention on previously unencountered objects.

Future work includes developing algorithms to control the slip of objects while performing manipulation and grasping tasks. With the identified good indicators of slip, work can be done to determine how these parameters can be utilized to control slip. Further research can be done to improve the synchronization of the sensor data from two

input sensors [26]. Finally, exploring how to merge data from sparse sensing to build an enhanced model for detecting and preventing slips presents a compelling opportunity.

CRedit authorship contribution statement

Xiaohai Hu: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Data curation, Conceptualization. **Aparajit Venkatesh:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation. **Yusen Wan:** Writing – review & editing, Validation, Software, Data curation. **Guiliang Zheng:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation. **Neel Jawale:** Writing – review & editing, Visualization, Software. **Navneet Kaur:** Writing – review & editing, Validation. **Xu Chen:** Writing – review & editing, Supervision, Resources, Funding acquisition, Formal analysis, Conceptualization. **Paul Birkmeyer:** Writing – review & editing, Supervision, Project administration, Methodology.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Xu Chen reports financial support was provided by Amazon. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors thank Siyuan Dong and Yu She for their input in using the sensor. We thank Hui Xiao for the suggestions and discussion. We also appreciate Michael Wolf's coordination and feedback along the way. This work was supported by a UW+Amazon Science Hub Gift-funded Robotic Research Project.

References

- [1] Luo Shan, Bimbo Joao, Dahiya Ravinder, Liu Hongbin. Robotic tactile perception of object properties: A review. *Mechatronics* 2017;48:54–67, URL <https://www.sciencedirect.com/science/article/pii/S0957415817301575>.
- [2] Johansson RS, Westling G. Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects. *Exp Brain Res* 1984;56:550–64.
- [3] Xia Ziwei, Deng Zhen, Fang Bin, Yang Yiyong, Sun Fuchun. A review on sensory perception for dexterous robotic manipulation. *Int J Adv Robot Syst* 2022;19(2).
- [4] Butterfass J, Grebenstein M, Liu H, Hirzinger G. DLR-hand II: next generation of a dextrous robot hand. In: *Proceedings 2001 ICRA. IEEE international conference on robotics and automation. Vol. 1, 2001, p. 109–14, vol.1.*
- [5] Tsetserukou D, Tadakuma R, Kajimoto H, Tachi S. Optical torque sensors for implementation of local impedance control of the arm of humanoid robot. In: *Proceedings 2006 IEEE international conference on robotics and automation. 2006, p. 1674–9.*
- [6] Jeong Seok Hwan, Kim Kyung-Soo, Kim Soohyun. Designing anthropomorphic robot hand with active dual-mode twisted string actuation mechanism and tiny tension sensors. *IEEE Robot Autom Lett* 2017;2(3):1571–8.
- [7] Wettels Nicholas, Santos Veronica, Johansson Roland, Loeb Gerald. Biomimetic tactile sensor array. *Adv Robot* 2008;22:829–49.
- [8] Costanzo Marco. Control of robotic object pivoting based on tactile sensing. *Mechatronics* 2021;76:102545, URL <https://www.sciencedirect.com/science/article/pii/S095741582100043X>.
- [9] Ravi Balasubramanian Veronica J Santos. *The human hand as an inspiration for robot hand development. Springer tracts in advanced robotics, New York: Springer; 2014.*

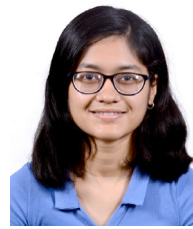
- [10] Ward-Cherrier Benjamin, Pestell Nicholas, Cramphorn Luke, Winstone Benjamin, Giannaccini Maria Elena, Rossiter Jonathan, Lepora Nathan F. The TacTip family: Soft optical tactile sensors with 3D-printed biomimetic morphologies. *Soft Robot* 2018;5(2):216–27, PMID: 29297773.
- [11] Griffa Pietro, Sferrazza Carmelo, D'Andrea Raffaello. Leveraging distributed contact force measurements for slip detection: a physics-based approach enabled by a data-driven tactile sensor. In: *2022 international conference on robotics and automation. ICRA, 2022, p. 4826–32.*
- [12] Yuan Wenzhen, Dong Siyuan, Adelson Edward H. GelSight: High-resolution robot tactile sensors for estimating geometry and force. *Sensors* 2017;17(12). URL <https://www.mdpi.com/1424-8220/17/12/2762>.
- [13] Dong Siyuan, Yuan Wenzhen, Adelson Edward H. Improved GelSight tactile sensor for measuring geometry and slip. In: *2017 IEEE/RSJ international conference on intelligent robots and systems. IROS, 2017, p. 137–44.*
- [14] Howe RD, Cutkosky MR. Sensing skin acceleration for slip and texture perception. In: *Proceedings, 1989 international conference on robotics and automation. Vol.1, 1989, p. 145–50.*
- [15] Ikeda A, Kurita Y, Ueda J, Matsumoto Y, Ogasawara T. Grip force control for an elastic finger using vision-based incipient slip feedback. In: *2004 IEEE/RSJ international conference on intelligent robots and systems. IROS, Vol. 1, 2004, p. 810–5, vol.1.*
- [16] Maldonado Alexis, Alvarez Humberto, Beetz Michael. Improving robot manipulation through fingertip perception. In: *2012 IEEE/RSJ international conference on intelligent robots and systems. 2012, p. 2947–54.*
- [17] Veiga Filipe, van Hoof Herke, Peters Jan, Hermans Tucker. Stabilizing novel objects by learning to predict tactile slip. In: *2015 IEEE/RSJ international conference on intelligent robots and systems. IROS, 2015, p. 5065–72.*
- [18] James Jasper Wollaston, Pestell Nicholas, Lepora Nathan F. Slip detection with a biomimetic tactile sensor. *IEEE Robot Autom Lett* 2018;3(4):3340–6.
- [19] Dong Siyuan, Ma Daolin, Donlon Elliott, Rodriguez Alberto. Maintaining grasps within slipping bounds by monitoring incipient slip. In: *2019 international conference on robotics and automation. ICRA, 2019, p. 3818–24.*
- [20] Li Jianhua, Dong Siyuan, Adelson Edward H. Slip detection with combined tactile and visual information. In: *2018 IEEE international conference on robotics and automation. ICRA, 2018, p. 7772–7.*
- [21] Judd Euan, Aksoy Bekir, Digumarti Krishna Manaswi, Shea Herbert, Floreano Dario. Slip anticipation for grasping deformable objects using a soft force sensor. In: *2022 IEEE/RSJ international conference on intelligent robots and systems. IROS, 2022, p. 10003–8.*
- [22] Yuan Wenzhen, Li Rui, Srinivasan Mandayam A, Adelson Edward H. Measurement of shear and slip with a GelSight tactile sensor. In: *2015 IEEE international conference on robotics and automation. ICRA, 2015, p. 304–11.*
- [23] James Jasper Wollaston, Lepora Nathan F. Slip detection for grasp stabilization with a multifingered tactile robot hand. *IEEE Trans Robot* 2021;37(2):506–19.
- [24] Y. She S Wang, S. Dong A Rodriguez, Adelson E. Cable manipulation with a tactile-reactive gripper. In: *Robotics: science and systems. RSS, 2020.*
- [25] Jiang Jingang, Huang Zhiyuan, Bi Zhuming, Ma Xuefeng, Yu Guang. State-of-the-art control strategies for robotic PiH assembly. *Robot Comput-Integr Manuf* 2020;65:101894, URL <https://www.sciencedirect.com/science/article/pii/S0736584519302418>.
- [26] Hu Xiaohai, Chu Thomas L, Chen Xu. State-space system identification beyond the nyquist frequency with collaborative non-uniform sensing data. In: *2024 American Control Conference (ACC). 2024, p. 5400–5.*



Xiaohai Hu received his Bachelor of Engineering degree from Zhejiang University in 2021 and went on to complete a Master of Mechanical Engineering degree with a minor in Data Science from the University of Washington in 2023. He is currently a Ph.D. candidate in Robotics and Control at the Mechatronics, Automation, and Control Systems Laboratory, situated within the Department of Mechanical Engineering at the University of Washington. His doctoral research focuses on the interdisciplinary convergence of robotics and control systems, with a particular emphasis on system identification, optimal control, and robotic perception. His objective is to improve the efficiency and adaptability of robotic systems operating in intricate environments.



Aparajit got his BS in Mechanical Engineering from Purdue University in 2019. After graduation he worked for 2 years at ZF Group as a Mechatronics Engineer developing steering systems for autonomous and semi-autonomous Class 6 trucks, RVs and buses. He got his MS in Mechanical Engineering from University of Washington in 2023, working under the supervision of Dr. Xu Chen. During his MS he worked on developing algorithms for slip detection and slip prevention during robotic manipulation using tactile feedback. He is currently working at PACCAR as a Vehicle Controls Engineer.



Navneet Kaur obtained her Bachelor of Technology degree in Mechanical Engineering with a minor in Robotics from the Indian Institute of Technology Gandhinagar (IIT Gandhinagar). Currently, she is pursuing her Master's degree in Mechanical Engineering at the University of Washington (UW). Her research interests lie in the areas of feedback and feedforward controls and utilizing tactile sensing for gripper control.



Yusen Wan received the B.S. degree in mechanical design, manufacturing and automation from the Huazhong University of Science and Technology (HUST), Wuhan, in 2020, and the master's degree in mechanical engineering from HUST in 2023. And now he is currently pursuing the Ph.D. degree in the Mechatronics, Automation, and Control Systems Laboratory, Department of Mechanical Engineering, University of Washington. His research interests include active sensing, surface defect detection, motion planning, and 3D reconstruction.



Xu Chen is an Associate Professor and holds the Bryan T. McMinn Endowed Research Professorship of Mechanical Engineering at the University of Washington (UW), Seattle. He obtained his Ph.D. degree in mechanical engineering from the University of California, Berkeley in 2013, and his bachelor's degree in mechanical engineering from Tsinghua University, China in 2008. He researches into dynamic systems, controls, and robotics, to better understand and engineer smart manufacturing (e.g., with feedback controls, lasers, machine vision, and nondestructive inspection). He also serves as Director of the Boeing Advanced Research Collaboration at the UW — an interdisciplinary strategic Boeing-UW partnership for the future of flight. Xu Chen is an alumnus of the National Academy of Engineering's 2023 Frontiers of Engineering Symposium, a recipient of the National Science Foundation CAREER Award, the SME Sandra L. Bouckley Outstanding Young Manufacturing Engineer Award, the Mechatronic Systems Outstanding Young Researcher Award from the International Federation of Automatic Control (IFAC) Technical Committee on Mechatronic Systems, and the Young Investigator Award from ISCIE / ASME International Symposium on Flexible Automation.



Guiliang received his Bachelor of Science in Mechanical Engineering, along with a minor in Electrical Engineering, from Purdue University in December 2019. After graduation, Guiliang worked for Cummins as a mechanical design engineer for two years, where he worked on integrating electrified powertrain systems into existing commercial trucks. Guiliang is currently a graduate student in the Mechanical Engineering Department at the University of Washington. His research interests include data-driven control, machine learning, and applied optimal control theory, all within the context of autonomous vehicles and robots.



Neel Jawale earned a Bachelor of Technology (B.Tech) degree in Mechanical Engineering from the Vellore Institute of Technology (VIT) in India. His capstone thesis focused on the design and intelligent control of quadrotors for disaster mitigation applications. After graduation, he worked as an Associate Technical Consultant at Oracle, India, within the Financial Services Group. He is currently pursuing a Master of Science degree in Mechanical Engineering at the University of Washington (UW) in Seattle. His research interests include slip-aware tactile feedback-driven gripper control, whole-body control of manipulators, and the integration of Reinforcement Learning with Model Predictive Control for trajectory optimization. His work aims to enable reactive, contact-aware manipulation in real-world scenarios.



Paul Birkmeyer earned his Ph.D. from the University of California, Berkeley in 2013. He is currently employed in the Robotics AI division of Applied Science at Amazon.