UniTac: Whole-Robot Touch Sensing Without Tactile Sensors

Anonymous CVPR submission

Paper ID 9

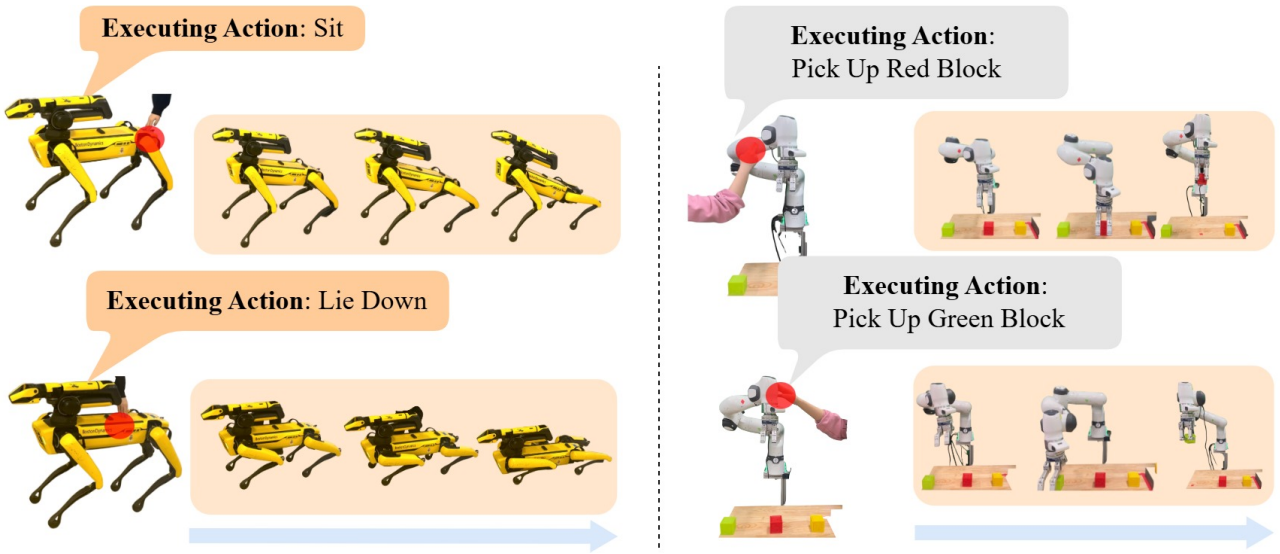

Figure 1. **Interactions achieved using UniTac**. UniTac achieves whole-robot touch sensing without using any tactile sensors and empowers applications such as patting the quadruped for canine-inspired responses or touch-based instructions for manipulation. Our method can be applied to robots with different embodiment types, including quadrupeds and arms.

Abstract

Touch sensing allows robots to better understand spatial interactions with humans and unstructured environments. However, most commercial robots are not equipped with tactile skins, making it challenging to achieve even basic touch-sensing functions, such as contact localization. We present UniTac, a data-driven whole-body touch-sensing approach that uses only proprioceptive joint sensors and does not require the installation of additional sensors. Our approach enables a robot equipped solely with joint sensors to localize contacts. Our goal is to: (1) democratize touch sensing through the modality of proprioception; (2) provide HRI researchers with an off-the-shelf tool for touch-enabled interaction and manipulation. We validate our approach on two platforms: the Franka Emika FR3 robot arm and the Boston Dynamics Spot quadruped. On FR3, we can localize contact to within 8.0 centimeters, and on Spot, we can localize to within 7.2 centimeters without adding any additional sensors to the robot. Our code and model will be released upon paper acceptance.

1. Introduction

Commercial robots are becoming increasingly capable in diverse spatial constraints. We now have bipedal/quadrupedal robots that can walk or run in challenging environments [2, 31, 34], and robot arms that assemble products with precision [11]. Despite the impressive capabilities of these robots, they lack a critical aspect of animal behavior: physical interaction through touch (Fig. 1). This compromises their ability to sense physical interactions with objects and humans. Consider how a simple pat can convey trust or instruction when interacting with a person or an animal [9]. A similar level of nuanced touch-based communication is currently out of reach for most robots. This limitation is largely due to the absence or difficulty in endowing robots with touch-sensing capabili-

ties. Although hardware and software advances for touch sensing have been made, the practical challenges of integrating tactile sensors into robots has prevented widespread use [6, 35].

Touch sensing is essential for a variety of tasks, including recovery from unintended collisions [15, 25], classification of types of human touch [18], improving visual state estimation [21, 23, 24], provision of social support for the elderly [5], and aid in minimally invasive surgery [33]. These interactions have been supported by installing dedicated tactile sensors, including sensors on robot hands [19, 20, 22], and full-body tactile skins, either rigid [4, 8, 15, 18] or soft [12, 14, 16, 26, 28, 29, 32, 38]. Despite their usefulness, installing these tactile sensors is error-prone and cumbersome. Rigid sensors tend to compromise the robot's dexterity and mechanical flexibility, while soft sensors are prone to damage and produce errors due to self-contact at joints [3]. Moreover, the integration of tactile sensors involves high costs and complex considerations such as calibration, power supply, wiring, and communication infrastructure [6, 37]. Alternative sensing modalities that enable robots to feel touch without dedicated tactile sensors would address these issues.

In this paper, we propose **UniTac**, a unified method to enable **whole-robot touch sensing capabilities across different robots without tactile sensors**. Our approach is applicable to various robot platforms and leverages data from *only existing sensors*. Specifically, we leverage the modality of robot proprioception by using torque and position data from joint sensors, which are readily available on most commercial robots. Different contact patterns generate distinguishable proprioceptive feedback, which can be used to infer the location of contact. Unlike model-based approaches [10, 13, 27, 40], which depend on physical models and demand extensive expert tuning for each robot platform, our technique is entirely data-driven. While prior data-driven methods [25, 39] depend on simulated data - necessitating the construction of bespoke simulations for each robot - we train a neural network on real-world joint sensor data to directly predict contact location in real time, thereby eliminating simulation designs and the sim-to-real gap. Notably, our efficient data collection process requires as little as 2.5 hours, yet it is sufficient for robust real-world whole-robot touch sensing.

To validate our approach, we collected datasets from both a robot arm (Franka Research 3 or FR3) and a quadruped platform (Boston Dynamic Spot), applying touch at sampled points (Fig. 4) across the robots' surfaces. We introduced random perturbations in joint positions between data collections to ensure robustness. Using only the joint sensors, UniTac is able to localize touches within an average of 7.2 centimeters on Spot and 8.0 centimeters on FR3 at a runtime speed of 2,000 Hz. We further conducted experiments on a Spot robot not used during data collection. Results confirm the transferability of UniTac on different robot instances with the same type. This demonstrates the capability of proprioception (i.e. joint angles and joint torques) to equip robots with spatial perception of touch.

We demonstrate potential applications of UniTac in physical Human-Robot Interaction (pHRI) [7]. For instance, with live predictions from UniTac, a quadruped robot (Spot) can react to human touch online, while an FR3 robotic arm can perform manipulation tasks based on the location of human contact (Fig. 1).

In summary, our contributions are as follows:

- We present a data-driven model, UniTac, that leverages built-in joint torque sensors to achieve live whole-body touch sensing across various robot platforms, eliminating the need for dedicated tactile sensors.
- UniTac demonstrates generalizability across multiple robot instances with the same type, allowing a wider community to use it as an off-the-shelf interface directly.
- We demonstrate potential applications in touch-based human-robot interaction, including scenarios such as bio-inspired quadruped choreography.

2. Related Works

Prior works have studied the problem of using proprioceptive feedback for touching sensing. These works localize the contact on the robot body and can be categorized into two main directions: model-driven and data-driven.

2.1. Model-Driven Approaches

Model-driven methods [10, 13, 27, 40] leverage explicit physical models to interpret sensor data for contact localization. Manuelli and Tedrake [27] introduced the contact particle filter, which formulates contact localization as a convex quadratic optimization problem integrated with particle filtering and demonstrated on a simulated humanoid. Iskandar et al. [13] derived momentum-based equations from redundant joint force-torque sensor data, achieving high-resolution touch sensing without additional tactile hardware. Although these approaches eliminate the need for extensive data collection, they demand open sourced kinematic models for each specific robot platform and configuration, thereby limiting their adaptability. Furthermore, if the robot's payload changes, the model and model parameters will need to be updated by an expert, in contrast to our approach, which merely requires collecting an updated dataset within a few hours' time.

2.2. Data-Driven Approaches

Data-driven methods offer an alternative by learning contact localization directly from sensor data. Seminal works [25, 39] have demonstrated this approach by simulating proprioceptive feedback. Zwiener et al. [39] employed random

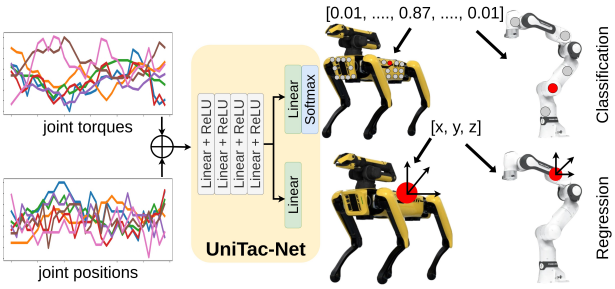

Figure 2. **Design of UniTac-Net.** UniTac-Net is a four-layer MLP with either a regression or a classification output head. It takes proprioceptive feedback (joint torques and positions) as input and predicts the contact (if any).

forests and multi-layer perceptrons (MLP) to classify contacts at pre-selected points on the Jaco arm, while Liang and Kroemer [25] trained a neural network to localize contact on the FR3 arm by projecting contact points onto mesh surfaces. Both studies follow the sim-to-real paradigm, in which they carefully design a simulation, train the model using simulated data, and deploy it on a single robot arm in the real world. In contrast, our work collects data exclusively from real-world platforms, thereby closing the sim-to-real gap and eliminating the need for customizing simulation environments. Moreover, our approach is applicable to various embodiments, including both robot arms and quadrupeds, and provides a straightforward interface for HRI researchers in developing downstream capabilities.

3. Method

UniTac localizes touch on the robot's surface using only proprioceptive feedback, exploiting the principle that contact location uniquely affects joint torques. We first randomly sample a preset number of n points on the surface of the robot mesh and define them as the ground truth contact locations. We collect joint data during contact at each point multiple times by varying joint configurations, and construct a dataset $\mathcal{D} = \{d_1, d_2, \ldots, d_k\}$ with k samples. A detailed process for contact collection will be explained in the next section. Each data tuple is $d_i = (p_i, q_i, \tau_i)$, where $p \in \mathbb{R}^3$ is the ground truth contact location. $q \in \mathbb{R}^{DoF}$ and $\tau \in \mathbb{R}^{DoF}$ are the joint positions and torques, respectively. We build a contact localization model that maps the proprioceptive signal - joint positions and torques (q and τ) - to the contact coordinate (p), defined in the robot frame, using a neural network, namely **UniTac-Net** (Fig. 2). Contact localization can be treated as either a regression or a classification problem [25, 30], which differ in the output head of the neural network.

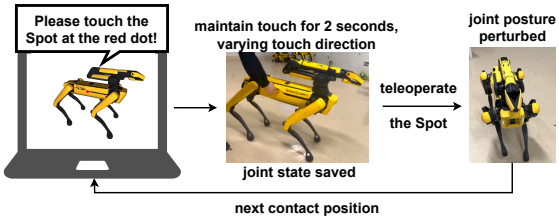

Figure 3. **Data collection pipeline.** Instruction on touch location is provided as a red marker on a URDF mesh visualization, which can be accurately mapped onto the actual robot.

3.1. Regression

Our regression approach directly predicts the continuous 3D coordinates of the contact point. For the architecture, we use a four-layer MLP with ReLU activations. The final layer has a size of three (x, y, z coordinates) and no activation function. A "no-contact" state is represented by the point (0, 0, 0). The model is trained using the mean squared error loss between ground truth p and predicted $\hat{p}$ as the loss. Compared to the classification model, the regression method avoids the inherent discretization error, potentially providing more precise localization. However, it may be more sensitive to noise in sensor readings, necessitating careful regularization and filtering.

3.2. Classification

When we treat it as a classification problem, we predict one of the pre-defined n points (as classes) with an additional "no-contact" class. The ground truth labels are obtained by mapping each ground truth contact location p into a one-hot encoding c with size of $(n + 1)$. The classification model shares the same overall architecture as the regression model up to the final layer. In this case, the final output layer uses a softmax activation to produce a probability distribution over the classes. We compute the cross-entropy loss between the ground truth class index c and the one-hot encoding for the predicted class index $\hat{c}$.

The classification method is robust against minor variations in joint signals because it maps inputs to a limited set of well-defined contact locations. However, it discretizes the contact space, which may limit precision if the contact point falls between the sampled points. The choice of the number of classes $(n+1)$ plays a crucial role here: increasing it can improve spatial resolution but may also make the model more complex and prone to overfitting.

4. Experiments

We validate the effectiveness of UniTac on two platforms with distinct morphologies: the Spot quadruped from Boston Dynamics and the Franka Research 3 robotic arm from Franka Robotics. In this section, we first introduce

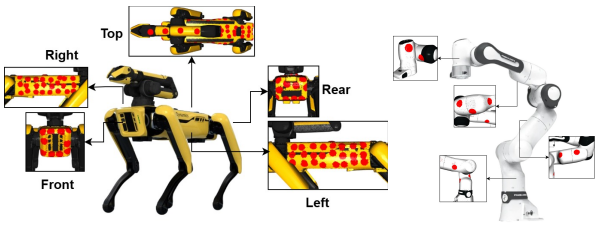

Figure 4. **Sampled contact points for data collection.** We sample 104 points on Spot (left) and 10 points on FR3 (right). Dense sampling on Spot covers the whole robot except for the legs, while the sparser sampling on FR3 covers each link.

our data collection process. Based on the collected dataset, we conduct experiments to evaluate the performance of our contact localization model. In the end, we demonstrate pHRI applications based on online deployment of UniTac.

4.1. Data Collection

Data is first collected when no contact is made with the robot, in which case the robot remains stationary. Meanwhile, the visualizer window shows the robot mesh to prepare the user for the following steps. The user will then collect each selected contact location, where the visualizer shows the robot mesh with a red marker on the ground truth touch location (Fig. 3). For data augmentation purposes, when collecting each sample, the user may vary the direction and magnitude of the touch force they apply as long as the touch location is kept the same. Spot data consists of joint state data during which contact is applied at 104 different touch locations throughout its whole body, collected over 50 sets of different joint configurations for 19 joints. We select 10 different touch locations to cover all arm links of FR3, over 25 sets of different joint configurations (Fig. 4).

4.1.1. Spot

Spot (with robot arm) has a total of 19 degrees of freedom (DoF): 12 joints for the legs (3 per leg), 6 for the arm, and 1 for the gripper. In our experiments, we focused on detecting where contact occurs on the robot's body and arm. The legs are ignored for safety reasons, and whole-body touch sensing will refer to touch sensing on the body and arm for the Spot. We sampled 104 contact points on the robot. These points are distributed as follows: 26 points each on the left, right, and top surfaces; 12 points on the back; 10 on the front; and 4 on the arm, as shown in Fig. 4 (left). The robot is kept stationary for each sample. To capture different configurations, we intentionally randomize the joint positions by teleoperating the robot into various postures (including different standing heights). Joint states for Spot are recorded using the Boston Dynamics Clients API at 60Hz.

4.1.2. FR3

FR3 model has 7 DoF. Unlike the quadruped Spot, the FR3 is more challenging for contact localization because it has an open kinematic chain with a fixed base, i.e. fewer joints provide non-zero torque offsets as the contact gets closer to the base. The robot arm joints are arranged sequentially, as opposed to the quadruped. We sample 10 points on the surface of FR3, as shown in Fig. 4 (right). The arm is assumed to be equipped with a fixed end-effector that remains steady throughout the experiments. (Changing the end-effector might lead to worse performance, since our model was trained using this specific configuration.) Similar to Spot, we perturb the joint positions after each collection. Joint torques and joint positions for FR3 are recorded using `libfranka` SDK at 30Hz.

4.2. Training

We preprocess the data by normalizing the joint positions q and torques τ to the range of -1 to 1. This normalization, performed for each dimension across the entire dataset, helps standardize the input and speeds up the convergence of the training process. Both models are trained using the Adam optimizer [17] with a fixed learning rate of 2.5×10^{-3}. Training is conducted over 30 epochs with a batch size of 256.

4.3. Contact Localization

UniTac-Net runs at the frequency of 2,000 Hz on an NVIDIA RTX 3090 GPU. We apply an exponential moving average (EMA) filter on model predictions with a smoothing factor of 0.1 and a sliding window length of 40. This EMA filter is used to filter noise in the output and improve temporal stability.

Our approach is purely data-driven and doesn't require any modeling related to the embodiment, while prior works require robot's physical model and don't have open sourced codebase [10, 13, 27, 39], we cannot perform direct comparison.

4.3.1. Quantitative Results

We use two metrics to evaluate the performance of our model: L2 norm and accuracy. L2 norm is defined as the Euclidean distance $||p - \hat{p}||_2$ between the predicted position $\hat{p}$ and ground truth contact position p. Accuracy (Acc) is calculated as the percentage of predictions whose Euclidean distance from the ground truth is within a threshold ϵ:

$$\text{Acc} = \frac{1}{N}\sum_{i=1}^{N} \mathbb{I}(\|p_i - \hat{p}_i\|_2 \leq \epsilon), \tag{1}$$

where N represents the number of samples.

We compare our regression model with the classification model (Tab. 1). When calculating the L2 norm for the classification model, the one-hot encodings are mapped back

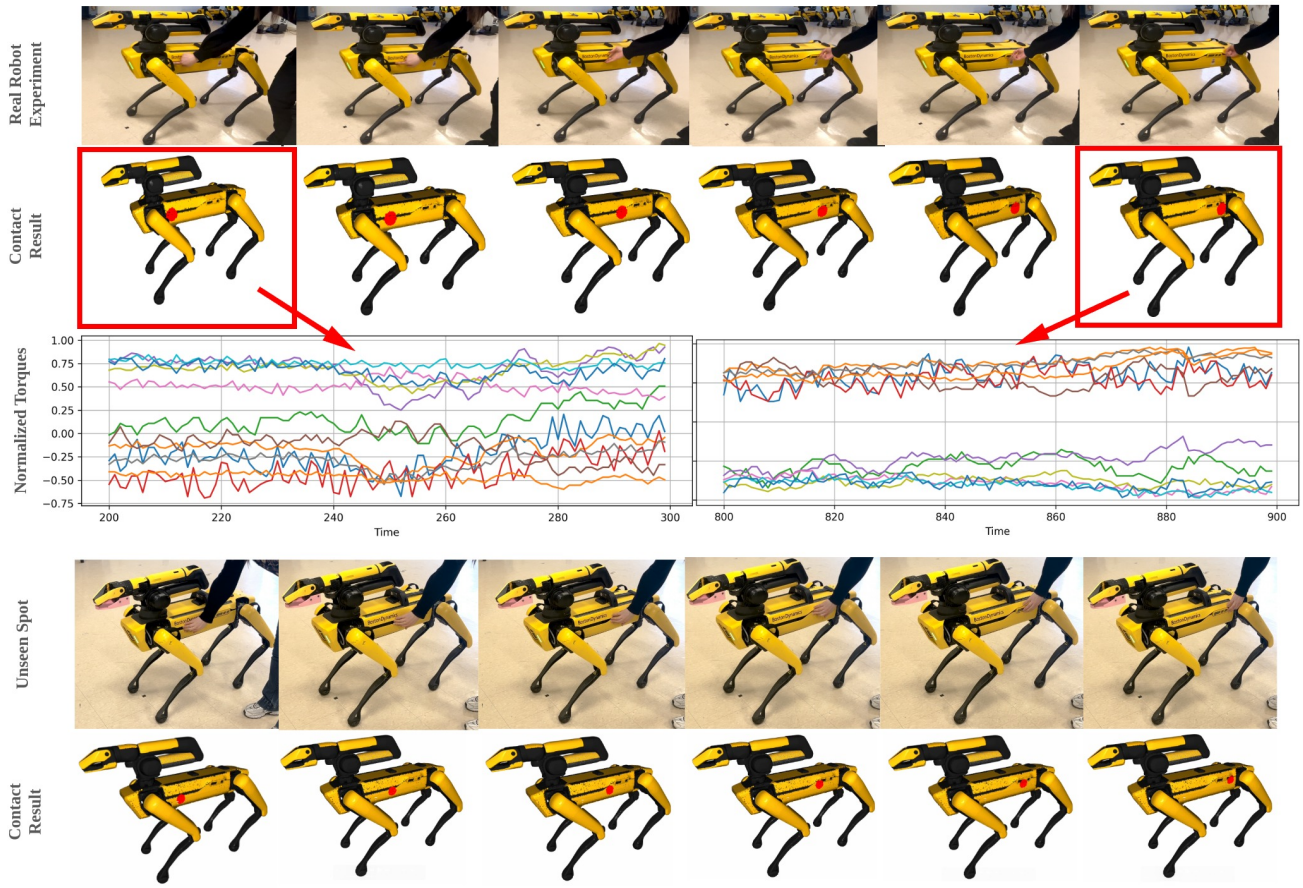

Figure 5. **Live contact localization on Spots.** Top row: A human applies touch to the robot. Middle row: The system localizes the contact point on the robot's mesh. Bottom row: Normalized joint torque changes are displayed (different colors indicate distinct joint sensors). The last two rows show that similar localization performance is achieved on *an unseen Spot* not used during training.

Method	Spot		FR3	
	Acc (%) ↑	L2 (cm) ↓	Acc (%) ↑	L2 (cm) ↓
KNN Classifier	34.3	24.1	55.8	20.3
KNN Regressor	45.9	19.4	43.3	17.2
Classification	54.9	13.7	53.7	14.8
Regression	**86.5**	**7.2**	**83.5**	**8.0**

Table 1. **Comparison of model choices**. We compare the performance of our regression and classification models and KNN baselines.

to the ground truth 3D coordinates of the contact positions. We partition the dataset into 80% for training and 20% for validation, and compare our regression model with the classification model, each evaluated alongside its corresponding k-nearest neighbors (KNN) baseline (with $k = 3$, Tab. 1). On Spot, the regression model achieves 86.5% accuracy (at a threshold of $\epsilon = 12$), with an average L2 error of 7.2 cm, compared to 54.9% accuracy and a 13.7 cm L2 error for the classification model. Similarly, on FR3, the regression approach attains 83.5% accuracy with an 8.0 cm L2 error, while the classification method reaches only 53.7% accuracy with a 14.8 cm L2 error. Each method also outperforms its KNN baseline, demonstrating the effectiveness of our network design. Performance differences across robot platforms stem from differences in the number of sampled ground truth points and the amount of collected data. These results demonstrate that the regression model outperforms the classification model by approximately 30 percentage points in accuracy and achieves a reduction of around 6.5 cm in L2 error. This suggests that avoiding discretization leads to more precise contact localization, especially in scenarios with noisy sensor readings. Fig. 7 plots the model accuracy as a function of the threshold ϵ.

4.3.2. Qualitative Results

We demonstrate qualitative results of online contact localization on Spot (Fig. 5) and FR3 (Fig. 6) using our regression model. On the Spot, we slide our touch horizontally along the left side of the body and visualize the live contact localization prediction results. Similarly, on the FR3 robot arm, we demonstrate by touching different links. The re-

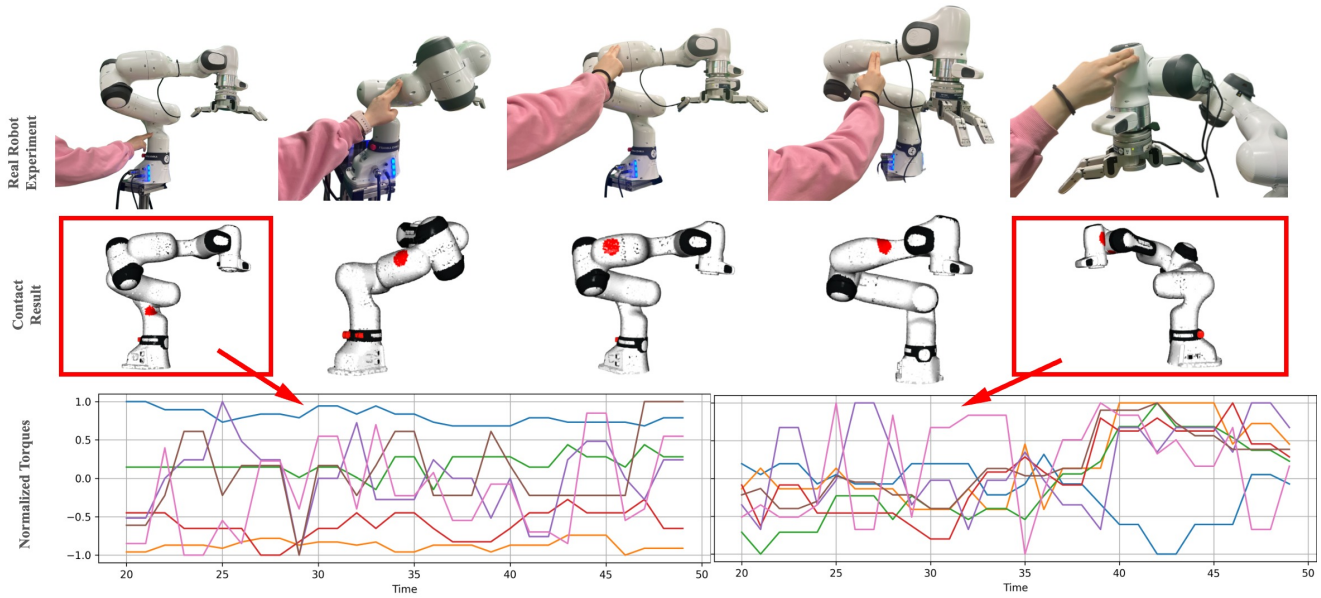

Figure 6. **Live contact localization on FR3 robot arm.** Similar to the Spot figure above, we show the actual human touch, localization result, and torque value change. The torque plots show the unique signature look for each touch location, which we utilized to localize the contact. More results are provided in our supplementary video.

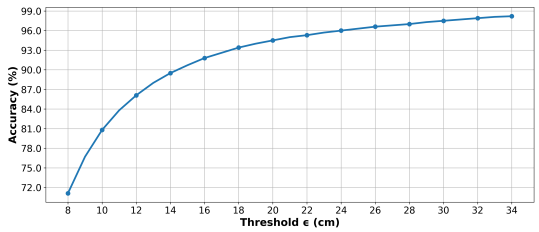

Figure 7. **Model accuracy with respect to threshold** ϵ. We evaluate the accuracy of UniTac-Net using various threshold values.

sults demonstrate that our model could accurately localize rapidly changing contacts online.

4.3.3. Generalization

We show that UniTac-Net generalizes to different instances of the same robot model without additional retraining. As illustrated in the last two rows of Fig. 5, we slide our touch horizontally along the left side of a Spot robot (not used for training data collection). The contact localization results exhibit similar accuracies with the results on a seen robot. Since different instances of the same robot model can operate with varying payloads, this result demonstrates the capability of UniTac-Net to generalize across instances within the same robot model.

Meanwhile, during data collection, the Spot's joint configuration is perturbed either by adjusting its standing height or by teleoperating it to different locations, resulting in small variations in joint angles. With an 80% to 20% train test split, the dataset captures a range of quadruped postures, and the quantitative and qualitative results presented above demonstrate that UniTac-Net generalizes effectively across these posture variations.

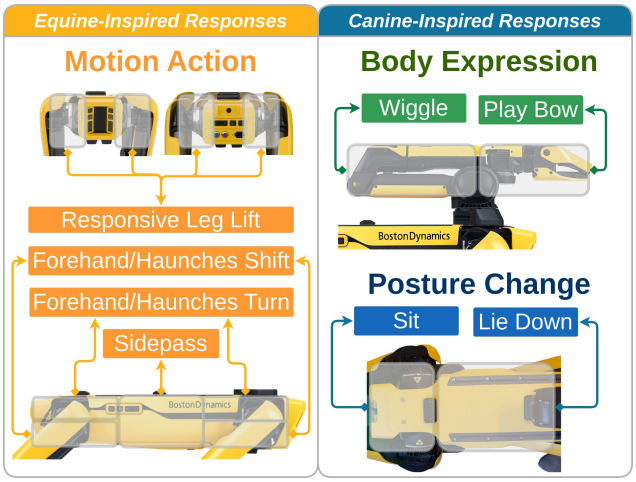

Figure 8. **Quadruped interactions design**. Touch-based interactions on Spot are categorized into three response types: motion actions (equine-inspired), posture changes, and body expressions (both canine-inspired)

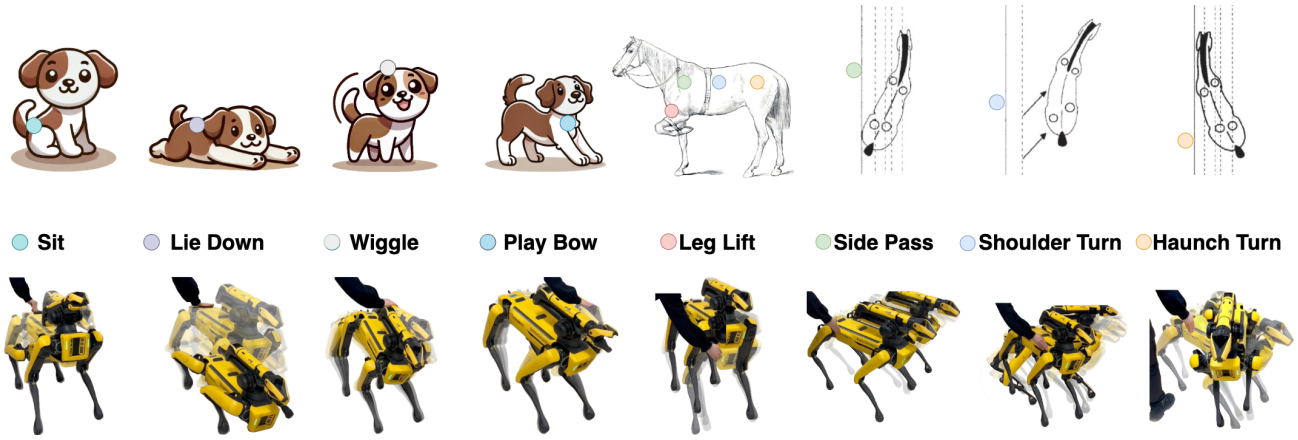

Figure 9. **pHRI deployment on Spot.** The first row illustrates the inspiration from human-animal interactions, showcasing how dogs and horses respond to touch cues (colored dots). The second row depicts the corresponding robotic responses in deployment on Spot. Video illustrations can be found in our supplementary material.

5. pHRI Applications

In the above sections, we presented how UniTac offers a pipeline to equip a quadruped or robot arm with whole-robot touch sensing capabilities. Tactile sensing enables more natural and intuitive pHRI by allowing robots to respond to physical touch. In this section, we further demonstrate potential applications enabled by such whole-robot touch-sensing abilities.

5.1. Quadruped Robot

Humans have an innate understanding of animal behavior. Mimicking natural behaviors allows for more fluid interactions in scenarios like guiding, assisting, or responding to human intentions [36]. To illustrate how a robot can behave and interact with humans in a way similar to animals with actual touch-sensing skins, we show quadruped choreography based on localized touch. This interaction design is inspired by ethologically relevant interactions observed in human-animal communication, particularly with social animals like dogs and horses.

5.1.1. Design

Using Spot, we program primitive actions using its Choreography SDK. Based on the predictions from UniTac-Net, we segment the robot's body into distinct regions, each triggering a specific action (Fig. 8). We divide Spot actions into three categories: 1) Motion actions; 2) Posture change; and 3) Body expression. As a standardization of terminology, we refer to the five sides of its body as front, rear, left, right, and top. All motion actions are inspired by human-equine interactions [1], while posture changes and body expression movements are derived from human-canine interactions.

Motion Actions. Motion actions involve whole-body movement and are based on horse training techniques that use physical cues to guide motion. These actions respond to touch on the side faces of the robot.

- *Turning on the Forehand* is triggered by touching the upper frontal section, prompting Spot to turn in the opposite direction by stepping with its front legs while its hind legs step in place.
- *Turning on the Haunches* occurs when the upper dorsal section is touched, making Spot step sideways while keeping its front legs stationary.
- *Shifting on Forehand/Haunches* results from touch on the lower frontal or dorsal sections, causing a weight shift in the opposite direction, mimicking a horse's response to abdominal pressure.
- *Sidepassing* is activated by touch on the mesial section, prompting a lateral step while maintaining body alignment.
- *Leg Lifting* occurs when the front or rear sections are touched, making Spot lift the corresponding leg, similar to a horse raising its hoof when handled by an equestrian.

Posture Change. We define posture change as Spot adjusting its body posture from a standing position, inspired by canine social signaling. These responses are associated with touch on the top side of Spot's body, excluding the front half, which is blocked when the arm is stowed.

- *Lying Down* is triggered by touching the middle section, causing Spot to fully lower itself.
- *Sitting* occurs when the rear section near the hip is touched, prompting Spot to lower its hindquarters, similar to a dog sitting when patted on the hip.

Body Expression. Body expressions represent robot actions that replicate a canine's expressive responses through movement. These behaviors are triggered by touch on Spot's arm, which corresponds to the head and upper torso of a canine.

- *Wiggle* occurs when the arm is touched, causing Spot to sway its body, similar to a puppy reacting to a pat on the

Figure 10. **Robot arm interactions design.** With UniTac, we can assign "virtual buttons" for operators to provide manipulation instructions at no cost.

neck.

- *Play Bow* is triggered by touch near the gripper, making Spot sway while opening its gripper, mimicking a dog playfully bowing to welcome a friendly pat.

5.1.2. Spot Demonstration

The quadruped choreography results show promising applications that are enabled by UniTac as shown in Fig. 9.

5.2. Robotic arm

Equipping robot arms with touch sensing capabilities provides significant benefits across multiple domains, improving safety, adaptability, and usability in human-robot interaction and manipulation tasks [3]. Joint torque sensors are available in many robot arms. With touch sensing capabilities, one potential application is replacing physical buttons with virtual buttons to provide manipulation instructions.

5.2.1. Design

Similar to seminal work [13], we focus on human-robot interaction in manipulation tasks, where the robot selects and picks up one of three blocks based on touch input.

We randomly select three locations on the FR3 to attach green, red, and yellow stickers. The green sticker is attached at a lower joint link close to FR3's base. The red sticker is attached to its joint link at medium height, and the yellow sticker is attached close to the end effector. We use these stickers to mimic "buttons" that will activate the Franka robot to pick up blocks of the same color from the desk in front of it (Fig. 10).

5.2.2. FR3 Demonstration

We demonstrate the FR3 tactile interactions in Fig. 11. Each row corresponds to pressing a different "button", where Franka is prompted to pick up the green, red, and yellow blocks, respectively.

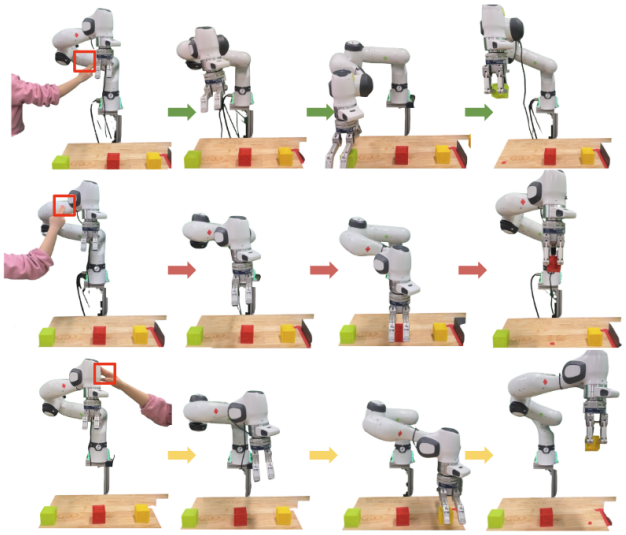

Figure 11. **pHRI deployment on the FR3 robot arm.** Each row illustrates the FR3 picking up a block of a different color.

6. Limitations & Future Work

Our work is currently limited to single contact localization, and does not estimate contact forces or generalize to dynamic postures. Nevertheless, our results suggest a promising direction for data-driven, proprioception-based touch sensing, which could enable robot physical interaction and manipulation without dedicated tactile sensors. We plan to address these limitations in future work by extending the method to more comprehensive spatial and interactive settings.

7. Conclusion

We present UniTac, a whole-robot touch sensing method that uses only built-in joint sensors for online contact localization. Evaluated on both FR3 arm and Spot quadruped, our data-driven approach achieves average localization errors of about 7–8 cm at 2,000 Hz, generalizing well across different robot instances. Furthermore, our pHRI demonstrations - such as quadruped choreography and robotic arm manipulation via virtual touch buttons - highlight the practical benefits of our approach for natural human-robot interactions. UniTac offers an easy-to-deploy alternative to dedicated tactile hardware, paving the way for more natural human-robot interactions. In our future work, we aim to scale up data collection to support more robust and multi-contact predictions, as well as generalization to dynamic postures.

References

[1] Chapter 6 - Communication. In *Equine Behavior*, pages 151–163. W.B. Saunders, Oxford, 2004. 7

[2] Ananye Agarwal, Ashish Kumar, Jitendra Malik, and Deepak Pathak. Legged Locomotion in Challenging Terrains using Egocentric Vision, 2022. arXiv:2211.07638 [cs]. 1

[3] Brenna D. Argall and Aude G. Billard. A survey of Tactile Human–Robot Interactions. *Robotics and Autonomous Systems*, 58(10):1159–1176, 2010. 2, 8

[4] Raunaq Bhirangi, Venkatesh Pattabiraman, Enes Erciyes, Yifeng Cao, Tess Hellebrekers, and Lerrel Pinto. AnySkin: Plug-and-play Skin Sensing for Robotic Touch, 2024. arXiv:2409.08276 [cs]. 2

[5] Alexis E. Block, Sammy Christen, Roger Gassert, Otmar Hilliges, and Katherine J. Kuchenbecker. The Six Hug Commandments: Design and Evaluation of a Human-Sized Hugging Robot with Visual and Haptic Perception. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, pages 380–388, Boulder CO USA, 2021. ACM. 2

[6] Ravinder S. Dahiya, Giorgio Metta, Maurizio Valle, and Giulio Sandini. Tactile Sensing—From Humans to Humanoids. *IEEE Transactions on Robotics*, 26(1):1–20, 2010. 2

[7] Agostino De Santis, Bruno Siciliano, Alessandro De Luca, and Antonio Bicchi. An atlas of physical human–robot interaction. *Mechanism and Machine Theory*, 43(3):253–270, 2008. 2

[8] Manel Frigola, Alicia Casals, and Josep Amat. Human-Robot Interaction Based on a Sensitive Bumper Skin. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 283–287, 2006. ISSN: 2153-0866. 2

[9] Alberto Gallace and Charles Spence. The science of interpersonal touch: an overview. *Neuroscience & Biobehavioral Reviews*, 34(2):246–259, 2010. 1

[10] Sami Haddadin, Alessandro De Luca, and Alin Albu-Schäffer. Robot Collisions: A Survey on Detection, Isolation, and Identification. *IEEE Transactions on Robotics*, 33 (6):1292–1312, 2017. 2, 4

[11] Sami Haddadin, Sven Parusel, Lars Johannsmeier, Saskia Golz, Simon Gabl, Florian Walch, Mohamadreza Sabaghian, Christoph Jähne, Lukas Hausperger, and Simon Haddadin. The Franka Emika Robot: A Reference Platform for Robotics Research and Education. *IEEE Robotics & Automation Magazine*, 29(2):46–64, 2022. 1

[12] Hiroshi Ishiguro, Tetsuo Ono, Michita Imai, Takeshi Maeda, Takayuki Kanda, and Ryohei Nakatsu. Robovie: an interactive humanoid robot. *Industrial Robot: An International Journal*, 28(6):498–504, 2001. 2

[13] Maged Iskandar, Alin Albu-Schäffer, and Alexander Dietrich. Intrinsic sense of touch for intuitive physical human-robot interaction. *Science Robotics*, 9(93):eadn4008, 2024. Publisher: American Association for the Advancement of Science. 2, 4, 8

[14] H. Iwata and S. Sugano. Human-robot-contact-state identification based on tactile recognition. *IEEE Transactions on Industrial Electronics*, 52(6):1468–1477, 2005. Conference Name: IEEE Transactions on Industrial Electronics. 2

[15] H. Iwate, H. Hoshino, T. Morita, and S. Sugano. Human-humanoid physical interaction realizing force following and task fulfillment. In *Proceedings. 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2000)*, pages 522–527 vol.1, 2000. 2

[16] T. Kanda, H. Ishiguro, T. Ono, M. Imai, and R. Nakatsu. Development and evaluation of an interactive humanoid robot "Robovie". In *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, pages 1848–1855 vol.2, 2002. 2

[17] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization, 2017. arXiv:1412.6980 [cs]. 4

[18] Seong-Yong Koo, Jong Gwan Lim, and Dong-soo Kwon. *Online touch behavior recognition of hard-cover robot using temporal decision tree classifier*. 2008. Journal Abbreviation: Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN Pages: 429 Publication Title: Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN. 2

[19] Mike Lambeta, Po-Wei Chou, Stephen Tian, Brian Yang, Benjamin Maloon, Victoria Rose Most, Dave Stroud, Raymond Santos, Ahmad Byagowi, Gregg Kammerer, Dinesh Jayaraman, and Roberto Calandra. DIGIT: A Novel Design for a Low-Cost Compact High-Resolution Tactile Sensor With Application to In-Hand Manipulation. *IEEE Robotics and Automation Letters*, 5(3):3838–3845, 2020. 2

[20] Mike Lambeta, Tingfan Wu, Ali Sengul, Victoria Rose Most, Nolan Black, Kevin Sawyer, Romeo Mercado, Haozhi Qi, Alexander Sohn, Byron Taylor, Norb Tydingco, Gregg Kammerer, Dave Stroud, Jake Khatha, Kurt Jenkins, Kyle Most, Neal Stein, Ricardo Chavira, Thomas Craven-Bartle, Eric Sanchez, Yitian Ding, Jitendra Malik, and Roberto Calandra. Digitizing Touch with an Artificial Multimodal Fingertip, 2024. arXiv:2411.02479. 2

[21] Hongyu Li, Snehal Dikhale, Soshi Iba, and Nawid Jamali. ViHOPE: Visuotactile In-Hand Object 6D Pose Estimation with Shape Completion. *IEEE Robotics and Automation Letters*, 8(11):6963–6970, 2023. arXiv:2309.05662 [cs]. 2

[22] Hongyu Li, Snehal Dikhale, Jinda Cui, Soshi Iba, and Nawid Jamali. HyperTaxel: Hyper-Resolution for Taxel-Based Tactile Signals Through Contrastive Learning. In *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7499–7506, 2024. ISSN: 2153-0866. 2

[23] Hongyu Li, James Akl, Srinath Sridhar, Tye Brady, and Taskin Padir. ViTa-Zero: Zero-shot Visuotactile Object 6D Pose Estimation, 2025. arXiv:2504.13179. 2

[24] Hongyu Li, Mingxi Jia, Tuluhan Akbulut, Yu Xiang, George Konidaris, and Srinath Sridhar. V-HOP: Visuo-Haptic 6D Object Pose Tracking. In *Robotics: Science and Systems XXI*, 2025. arXiv:2502.17434 [cs]. 2

[25] Jacky Liang and Oliver Kroemer. Contact Localization for Robot Arms in Motion without Torque Sensing. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021. 2, 3

[26] Karl F MacDorman, Takashi Minato, Michihiro Shimada, Shoji Itakura, Stephen Cowley, Hiroshi Ishiguro, et al. Assessing human likeness by eye contact in an android testbed. In *Proceedings of the XXVII annual meeting of the cognitive science society*, pages 21–23, 2005. 2

[27] Lucas Manuelli and Russ Tedrake. Localizing external contact using proprioceptive sensors: The Contact Particle Filter. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5062–5069, 2016. ISSN: 2153-0866. 2, 4

[28] Noriaki Mitsunaga, Takahiro Miyashita, Hiroshi Ishiguro, Kiyoshi Kogure, and Norihiro Hagita. Robovie-IV: A Communication Robot Interacting with People Daily in an Office. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5066–5072, 2006. ISSN: 2153-0866. 2

[29] Takahiro Miyashita, Taichi Tajika, Hiroshi Ishiguro, Kiyoshi Kogure, and Norihiro Hagita. Haptic Communication Between Humans and Robots. In *Robotics Research*, pages 525–536. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. Series Title: Springer Tracts in Advanced Robotics. 2

[30] Artem Molchanov, Oliver Kroemer, Zhe Su, and Gaurav S. Sukhatme. Contact localization on grasped objects using tactile sensing. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 216–222, Daejeon, South Korea, 2016. IEEE. 3

[31] Yunlong Song, Sangbae Kim, and Davide Scaramuzza. Learning Quadruped Locomotion Using Differentiable Simulation, 2024. arXiv:2403.14864. 1

[32] Taichi Tajika, Takahiro Miyashita, Hiroshi Ishiguro, and Norihiro Hagita. Reducing influence of robot's motion on tactile sensor based on partially linear model. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 512–517, 2008. ISSN: 2153-0866. 2

[33] Mohsin I. Tiwana, Stephen J. Redmond, and Nigel H. Lovell. A review of tactile sensing technologies with applications in biomedical engineering. *Sensors and Actuators A: Physical*, 179:17–31, 2012. 2

[34] Giorgio Valsecchi, Ruben Grandia, and Marco Hutter. Quadrupedal Locomotion on Uneven Terrain With Sensorized Feet. *IEEE Robotics and Automation Letters*, 5(2): 1548–1555, 2020. 1

[35] Longteng Yu and Dabiao Liu. Recent Progress in Tactile Sensing and Machine Learning for Texture Perception in Humanoid Robotics. *Interdisciplinary Materials*, 4(2):235–248, 2025. 2

[36] Lishuang Zhan, Yancheng Cao, Qitai Chen, Haole Guo, Jiasi Gao, Yiyue Luo, Shihui Guo, Guyue Zhou, and Jiangtao Gong. Enable Natural Tactile Interaction for Robot Dog based on Large-format Distributed Flexible Pressure Sensors, 2023. arXiv:2303.07595 [cs]. 7

[37] Can Zhao, Jieji Ren, Hexi Yu, and Daolin Ma. In-situ Mechanical Calibration for Vision-based Tactile Sensors. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 10387–10393, 2023. 2

[38] Hua Zhong, Yaxi Wang, Jiahao Xu, Yu Cheng, Sicong Liu, Jia Pan, Wenping Wang, and Zheng Wang. Dynamics-Oriented Underwater Mechanoreception Interface for Simultaneous Flow and Contact Perception. *Advanced Intelligent Systems*, 7(1):2400492, 2025. 2

[39] Adrian Zwiener, Christian Geckeler, and Andreas Zell. Contact Point Localization for Articulated Manipulators with Proprioceptive Sensors and Machine Learning. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 323–329, 2018. 2, 4

[40] Adrian Zwiener, Richard Hanten, Cornelia Schulz, and Andreas Zell. ARMCL: ARM Contact point Localization via Monte Carlo Localization. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7105–7111, 2019. ISSN: 2153-0866. 2