

# SlipSense: Multimodal Sensing for Online Slip Detection in Legged Robots

Iris Szu-Yao Liu<sup>1,2</sup>, Chien Chern Cheah<sup>1</sup>, Meng Yee (Michael) Chuah<sup>2</sup>

**Abstract**—Legged robots rely on accurate ground interaction awareness to traverse variable terrains, such as slippery surfaces. Existing slip detection methods often rely on kinematics and proprioception, which lack the sensitivity to detect early-stage slips that occur prior to catastrophic instability. Thus, this paper presents SlipSense, a novel framework for online force-based slip detection using a custom lightweight sensorized foot for quadrupeds to detect slip. The framework integrates a multimodal sensor design with a LSTM-based model to infer ground reaction forces and detect slip-indicative anomalies during locomotion. The proposed framework is deployed on a Unitree Go1 quadruped to demonstrate blind online slip detection over a slippery terrain. Our method detects early-stage slips down to an average displacement of  $24.1 \pm 6.4\text{mm}$  with an overall accuracy of 85.9%. This represents a 3.3-fold finer detection resolution and a 24% relative accuracy improvement over a standard kinematic baseline that uses foot velocity inferred through state estimation. The work in this paper serves as a foundation for force-aware gait adaptation in legged robotic locomotion, allowing future controllers to estimate terrain friction and adjust constraints, thus improving the overall stability of the system.

## I. INTRODUCTION

Recent years have seen an increasing number of legged robots deployed for navigating complex environments such as for exploration or industrial automation. These systems are often subject to highly dynamic and unpredictable environments that may involve terrains with varying friction. This remains a challenge to achieve safe and efficient locomotion in versatile terrains due to the risk of slip and instability. In such cases, multidimensional ground reaction forces (GRFs) are a crucial source of information for robust legged robot control, as it provides direct insights into the robot's interaction with the environment. Conventional approaches use a combination of proprioception and state estimation to perform GRF estimation [1] [2]. However, these methods often rely on precise modeling of leg mechanics and are highly sensitive to unmodeled inertial effects or joint friction. Consequently, slip detection strategies built on this principle, such as those that infer slip from foot velocity [3] [4], must wait for a significant slip to occur before its signal can be distinguished from the noise of normal locomotion. This leads to a compromise of setting a high detection threshold to prevent false positives at the cost of insensitivity to early-stage slips that precede unrecoverable stability. Alternatively,

This research is supported by the Home Team Science and Technology Agency (HTX).

<sup>1</sup> School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

<sup>2</sup> Institute for Infocomm Research, Agency for Science Technology and Research, Singapore

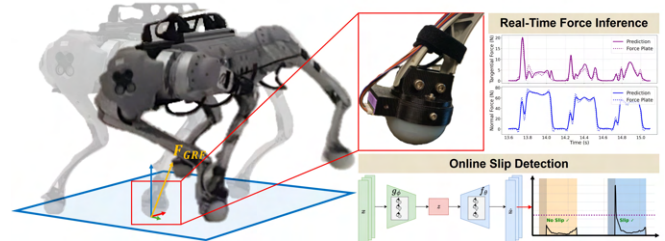


Fig. 1. SlipSense: multimodal online slip detection for legged robots using a custom foot-mounted force sensor. The hardware is integrated with dual-temporal learning models to perform real-time force inference and slip-indicative anomaly detection. Figure shows snapshots of robot becoming unstable after slip.

vision-based terrain mapping has also been explored [5], but it relies on accurate elevation maps and fails under occlusion or poor visibility, which can severely degrade stability. Therefore, it is desirable to integrate a direct force sensing system into legged robots to perform real-time GRF estimation. However, practical integration of foot-mounted sensors face challenges such as durability, bulkiness, and the ability to operate under dynamic locomotion. These hardware challenges have limited the adoption of sensorized feet into dynamic locomotion applications, despite their potential for enhancing stability.

To address these limitations, we propose a framework which integrates a lightweight, multimodal force sensor with a learning model to provide real-time force inference and high-sensitivity slip detection. To the best of our knowledge, this is the first work to present a deployed multimodal force-based slip detection framework under dynamic quadruped gait, which can serve as a foundation for future force-aware locomotion control. Our contributions are summarized as follows:

- **SlipSense (Fig. 1), a complete framework for high-sensitivity slip detection:** we introduce an end-to-end system that integrates a custom multimodal force sensor with a LSTM-based learning model, enabling the detection of early slip events that cannot be detected by conventional kinematic-only methods.
- **A data-efficient anomaly detection approach for slip detection training:** we frame the slip detection as a one-class self-supervised anomaly detection problem, mitigating the challenge of acquiring accurate, well-labeled slip data for training our model.
- **Experimental validation on a real quadruped demonstrates a 3.3-fold improvement in detection**

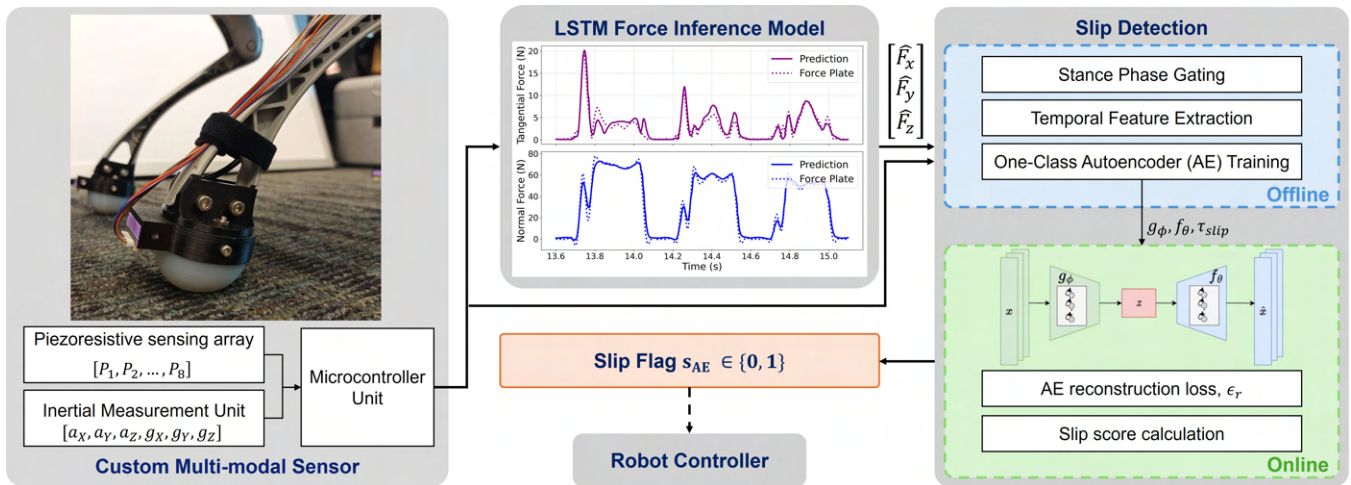


Fig. 2. Overview of the proposed SlipSense framework for online slip detection in quadruped robots. The system consists of three modules: (1) a lightweight multimodal footpad sensor combining pressure and inertial measurements, (2) a temporal LSTM-based model that reconstructs 3D GRFs in real-time, and (3) a self-supervised anomaly detection module that fuses kinetic and kinematic features to identify slip events. Integration with the robot’s controller is for future work.

**resolution:** we deployed our framework for blind robot traversal over terrains of varying friction and achieved a higher accuracy and ability to detect smaller slips compared to a kinematic baseline using inferred foot velocity through state estimation. This was validated through a motion capture system.

## II. RELATED WORK

Ground reaction force (GRF) sensing plays a critical role in legged robot control, providing low-level feedback for balance, adaptation, and interaction with diverse terrains. Legged robotic force sensing methods fall under two categories, namely state estimation-based and direct force sensing. State estimation-based sensing uses sensorless approaches such as inference from joint torques [1] [2], extended state observers [6], and generalized momentum observers [7]. These approaches often struggle to detect fine sensor-surface contact changes [8] and are commonly contaminated by noise due to electromagnetic interference. Direct force sensing methods adopted in studies include piezoresistive pressure [9] [10], capacitive [11] [12], optical [8] [13], and strain-gauge based force torque (F/T) sensors [14]. Sensorized feet in legged robots offer potential in enabling tasks such as terrain characterization [11] [15], friction estimation [13], slip detection [16] [17], and edge detection [18] [19]. The focus of this paper will be on legged robot slip detection. Slip detection is essential for maintaining stability during locomotion, especially on low-friction or uncertain terrains and often fall under two categories: direct force-based and kinematic-based methods.

Force-based approaches rely on additive sensors integrated into the foot which directly estimates the ground contact forces. An example is presented in [16] where a MEMS slip sensor combined with a force sensor and shock absorber is used to measure slipperiness. However, this design is limited to vertical forces up to only 20N and introduced

significant inertia due to bulkiness. Alternatively, Park et al. [17] proposed an anti-skid foot design that detects slip when excessive tangential force is applied to the foot, triggering the release of anchoring spines to prevent slip. Slip detection was demonstrated in [9] using a piezoresistive stress field sensor on a detached Mini Cheetah leg in a quasi-static setup. More generally, previous studies highlight the fragility and limited durability of additive foot force sensors as a key bottleneck for practical force-based slip detection [3] [20]. Thus, while force-based methods offer greater sensitivity, no existing design has demonstrated the combination of a lightweight, high force-range, multi-axis sensor reliable for deployment on a dynamic quadruped for slip detection. These works are often limited in practical integration due to two key challenges: fragility under high-impact and adverse inertial effects from bulky designs.

Alternatively, kinematic-based slip detection methods rely on proprioceptive information, where slippage is inferred from foot motion, typically using joint encoders and IMUs [21] [22]. Commonly, these methods adopt a velocity-level slip detection, such as in [3], where single and multiple leg slippage are detected by estimating the foot velocity in the body and world frame, respectively. Similarly in [20], slip is detected when the norm of the difference between the desired and actual foot velocities in the body frame exceeds a certain threshold. Yan et al. [4] fused the foot displacement and velocity to estimate the slip state. As another approach, the authors in [23] mounted an IMU at the foot to probabilistically determine the contact state stability. In [24], a deep learning approach was taken and a feature set comprising proprioceptive information was used to predict slips in simulation. A common limitation across all kinematic methods is the reliance on indirect, model-dependent data. Error accumulation from IMU integration and state estimation drift results in insensitivity to the early-stage slips that precede larger, unrecoverable slips.

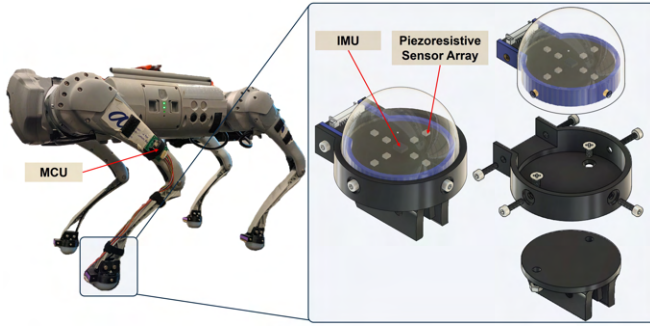


Fig. 3. Our custom multimodal footpad sensor mounted onto a Unitree Go1 quadruped. The exploded diagram reveals the modularity of the footpad mounting design and the internal PCB components.

In summary, the literature reveals a critical trade-off in legged robot slip detection methods: force-based methods offer high sensitivity but are fragile and impractical for dynamic deployment, while kinematic methods allow for easier integration but are insensitive to early-stage slip events. This presents a research gap for a system that can achieve both high-fidelity detection and demonstrated robust real-world deployment. To address this, we propose our end-to-end multimodal slip detection framework, named SlipSense, which fuses force and inertial information and leverages them to reliably detect slip, enabling finer sensitivity and more robust online deployment.

### III. INTEGRATED FORCE INFERENCE AND SLIP DETECTION FRAMEWORK

The overview of our proposed framework, SlipSense, is illustrated in Fig. 2. The pipeline begins with our custom multimodal footpad sensor which outputs raw piezoresistive pressure and inertial measurements. A temporal force inference model processes the sequence of piezoresistive measurements to reconstruct the 3D GRFs,  $[F_X, F_Y, F_Z]$ . Coupled with the inertial measurements, the feature set is fed into the slip detection module, which leverages a one-class anomaly detection model to distinguish between stable contacts and anomalous slip behavior.

### IV. SENSORIZED FOOT AND FORCE INFERENCE MODEL

In the following section, we introduce our custom-made force sensor design and the temporal force inference model for predicting GRFs in real-time. This serves as the initial stage of the pipeline for the downstream slip detection task.

#### A. Force Sensor Design

Our force sensor (shown in Fig. 3) is based on the stress field sensing concept presented in [9] [25], which uses an array of piezoresistive sensors embedded within a dome of rubber to reconstruct the 3D forces from a load. However, to enable robust slip detection during dynamic locomotion, we introduced two key modifications:

- **IMU Integration:** An IMU chip is placed at the center of the custom PCB board, surrounded by an array of eight piezoresistive pressure sensors arranged in two

concentric diamonds. This arrangement ensures that kinematic and kinetic data are captured from the same rigid body.

- **Enhanced Durability:** Similar to in [26], we remove the piezoresistive sensor caps to expose the diaphragm, thus increasing the sensitivity. We automated this process using a 5-axis CNC machine to precisely mill off the sensor caps. To ensure the exposed diaphragm is protected from repeated loads, we pot the sensor cavities with a harder Shore 40A silicone before overmolding with a softer Shore 30A elastomer. The fabricated sensor has a diameter of 45mm and weighs 65g, which is an insignificant increase compared to the original Unitree Go1 foot weighing 51g.

#### B. Force Inference Model

To map the raw sensor array readings to contact forces, a learning-based approach was chosen to generalize across the various loading conditions of dynamic locomotion. To model the material viscoelastic behavior introduced by the elastomeric material, we use a Long Short Term Memory (LSTM) model to use historical deformation information to predict the current state of the output. Previous works using static models such as Gaussian Process Regression or ANN [9], often struggle to predict forces at the upper range, due to the more complex material behavior such as hysteresis and stress relaxation. The sequence-to-vector mapping function,  $F$ , of our model is defined as:

$$[\mathbf{x}^{(t-S+1)}, \dots, \mathbf{x}^{(t)}] \xrightarrow{F} \hat{\mathbf{y}}^{(t)} \quad (1)$$

where  $S$  is the length of the input sequence of sensor readings and each  $\mathbf{x}^{(i)} \in \mathbb{R}^{D_{in}}$  is a vector of sensor measurements at timestep  $i$ . The model's objective is to predict the corresponding GRF vector  $\hat{\mathbf{y}}^{(t)} \in \mathbb{R}^{D_{out}}$  at the final timestep  $t$ . For this work, the input dimension  $D_{in} = 8$ , corresponding to the eight pressure sensors, and the output dimension  $D_{out} = 3$ , for the force components,  $[F_x, F_y, F_z]$ . The input sequence length was set to  $S = 50$ , which, given our sensor sampling rate of 1000Hz, represents a 50ms history of the contact event. To collect a diverse set of training data, a 5-axis Pocket NC V2-10 CNC machine was used. A commercial ATI Industrial Automation 6-axis force torque (F/T) sensor was mounted on the bed of the CNC machine, while the footpad sensor was mounted on the spindle. The CNC was programmed to apply a wide range of controlled trajectories such as normal, shear, and roll motions to simulate the force profiles of locomotion. We obtained  $2.3 \times 10^6$  data points for our training dataset.

To evaluate the performance of the proposed force inference model, we integrate the sensor onto a Unitree Go1 quadruped. Fig. 4 compares the model predictions against ground-truth measurements from a Kistler 9260AA6 force plate during consecutive trotting motions. Our force inference latency is 0.24ms and we run the module at 500Hz. The results show good alignment between predicted and measured forces, demonstrating that the inferred forces are

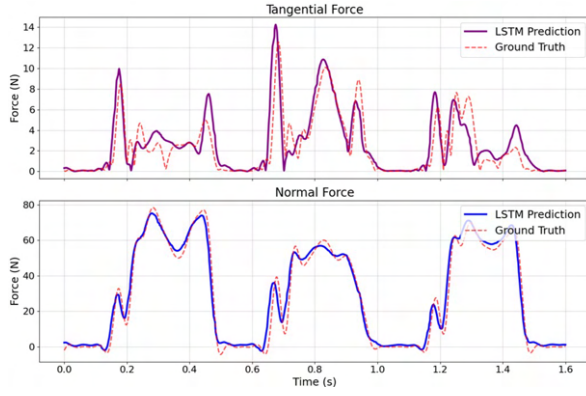


Fig. 4. Comparison of simultaneous tangential and normal force predictions from LSTM model with ground truth force plate data.

sufficiently accurate to serve as a reliable input for the downstream slip detection task.

## V. ANOMALY DETECTION FOR ONLINE SLIP DETECTION

The next component of our SlipSense framework leverages the force estimates from our force inference model to learn a robust representation of normal gait, enabling the detection of slips as anomalies from this learned behavior. This section discusses the physical principles of slip, our one-class learning methodology, and the real-time implementation. A summary of the slip detection framework is shown in Fig. 5.

The interaction between the robot’s foot and the ground can be defined as a stable foothold (stick) or a slip state. According to the Coulomb friction model, a stable foothold is maintained as long as the total tangential force,  $F_T = \sqrt{F_x^2 + F_y^2}$ , exerted by the leg remains less than the maximum allowable static friction force, as described by the equation:

$$F_T < \mu_s F_N \quad (2)$$

where  $\mu_s$  is the static coefficient of friction and  $F_N$  is the normal force. Conversely, incipient slip occurs at the moment the stick condition is violated, i.e. when  $F_T \geq \mu_s F_N$ , which is the point we are interested in detecting via our multimodal method. A loss of traction involves a rapid transition from static to kinetic friction, in the form of high-frequency fluctuations and large transient spikes in the shear force derivatives ( $\frac{dF_x}{dt}$ ,  $\frac{dF_y}{dt}$ ) [3]. These sharp changes in force profiles indicate an unstable contact occurring.

From a kinematics point of view, a stable foothold is defined by zero velocity of the foot relative to the ground. The foot acceleration in the world frame should therefore remain near zero, accounting only for minor vibrations and noise. We leverage kinematics data to detect fluctuations in the tangential acceleration which represent an unstable foothold. To achieve this, we utilize IMU data, which is transformed from the foot’s body frame to the world frame via a Madgwick filter. Significant, non-zero accelerations in the tangential plane during stance hint towards an unstable foothold. Our methodology, therefore, is built on the hypothesis that the onset of slip manifests as a combined anomaly

in both the kinematic (inertial) and kinetic (force derivative) signals. We fuse these modalities to create a feature set for learning the characteristics of contact stability.

Acquiring a large and diverse dataset of labeled slip events is difficult and often impractical due to the dangerous and inconsistent nature of the event. To overcome this challenge, we frame slip detection as a one-class classification (or anomaly detection) problem. This approach allows us to train a model exclusively on an abundance of easily-collected data of the robot walking with stable contacts on high friction terrains. Our stable contacts dataset was collected across a variety of conditions to ensure the model learns a comprehensive representation of stability. This included multiple speeds, changes in direction, in-place rotations, and traversal over different high-traction terrains.

The core of our methodology is an LSTM Autoencoder (AE), trained via self-supervised learning on this stable contacts dataset. The training pipeline is as follows:

- **Stance Phase Extraction:** By setting a normal force threshold,  $F_z > 5N$ , we segment the time series data to isolate the stance phases.
- **Feature Engineering:** For each time step, a feature vector,  $x_t$ , is constructed and involves a fusion of kinematics (inertial data) and kinetics (shear force derivatives) information to capture the contact dynamics:

$$x_t = [a_x, a_y, a_z, g_x, g_y, g_z, \frac{dF_x}{dt}, \frac{dF_y}{dt}] \in \mathbb{R}^d$$

where  $d = 8$  is the feature dimension.

- **Windowing:** The feature vectors are segmented into moving windows of 50ms ( $T = 25$  samples at 500 Hz). This duration is short enough for real-time detection while being sufficiently long to capture temporal patterns within the window. Our feature window is defined as:

$$\mathbf{X} = (x_1, x_2, \dots, x_T) \in \mathbb{R}^{T \times d} \quad (3)$$

- **Self-supervised Training:** The autoencoder is trained to reconstruct these feature windows. The LSTM encoder,  $g_\phi$ , maps the input sequence,  $\mathbf{X}$ , into a low-dimensional latent embedding,  $z \in \mathbb{R}^k$ , where  $k$  is the embedding dimension. We use the encoder’s final hidden state,  $h_T$ , as the latent representation,  $z = h_T \in \mathbb{R}^k$ .

The decoder,  $f_\theta$ , then attempts to reconstruct the original input sequence,  $\hat{\mathbf{X}} = (\hat{x}_1, \dots, \hat{x}_T)$ , from the latent embedding,  $z$ . At each time step, the decoder uses a dense layer to output the final reconstructed feature vector:

$$\hat{x}_t = W_y h'_t + b_y \quad (4)$$

where  $h'_t \in \mathbb{R}^k$  is the hidden state output of the decoder LSTM at time step  $t$ .  $W_y$  and  $b_y$  are the weight matrix and the bias vector of the dense layer, respectively. The model is trained by minimizing the L1 reconstruction loss, which computes the mean absolute error between the original and reconstructed sequences over all time steps:

$$\mathcal{L}_{\text{rec}} = \frac{1}{T} \sum_{t=1}^T \|x_t - \hat{x}_t\|_1 \quad (5)$$

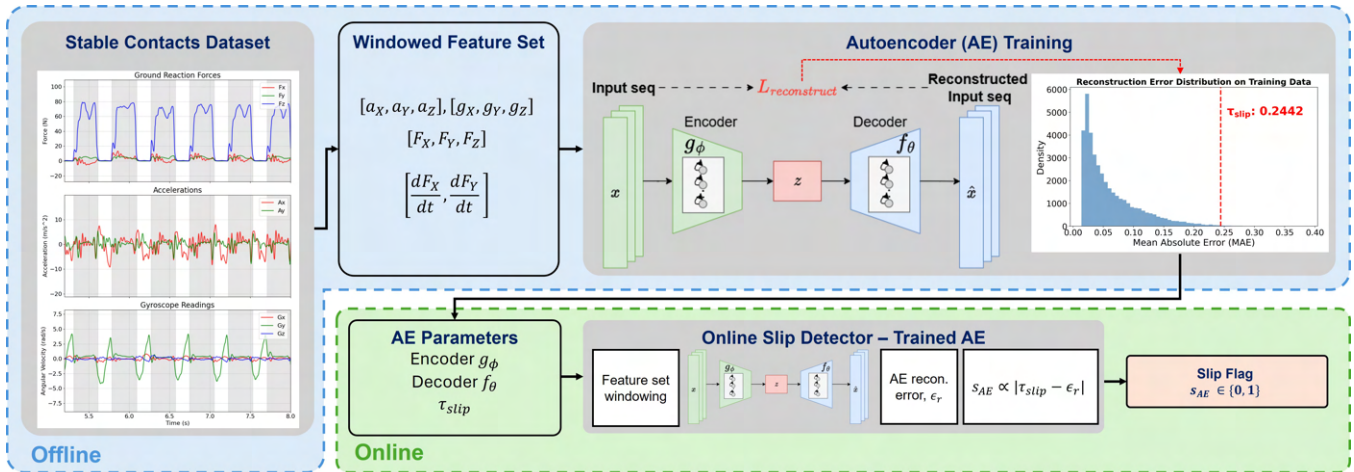


Fig. 5. The proposed slip detection framework, illustrating the offline training pipeline and online deployment. This framework consists of three key modules: (1) Data collection of stable contacts and extraction of windowed features. (2) Offline self-supervised training of an LSTM AE to learn a stable contact model. The 99.5<sup>th</sup> percentile of the reconstruction error distribution on a validation set is used to determine the threshold for slip classification,  $\tau_{slip}$ . (3) Online anomaly detection using the trained AE model where the real-time reconstruction error,  $\epsilon_r$ , is computed and compared against  $\tau_{slip}$  to generate a binary slip flag,  $s_{AE}$ .

where  $\|v\|_1 = \sum_j |v_j|$  is the L1 norm.

After training on the stable contacts dataset, the 99.5<sup>th</sup> percentile of the reconstruction errors (as shown in Fig. 5) on a validation set is then calculated and stored as the anomaly detection threshold,  $\tau_{slip}$ , for determining the slip label,  $s_{AE}$ :

$$s_{AE} = \begin{cases} 1, & \text{if } \epsilon_r(t) > \tau_{slip}, \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

where  $\epsilon_r(t)$  is the real-time L1 reconstruction error, calculated as:

$$\epsilon_r(t) = \frac{1}{T} \sum_{i=1}^T \|\mathbf{x}_{t-T+i} - \hat{\mathbf{x}}_{t-T+i}\|_1 \quad (7)$$

## VI. EXPERIMENTAL VALIDATION OF SLIPSENSE ON A QUADRUPED ROBOT

The trained AE model is deployed to evaluate the performance of slip detection on a real quadruped robot. This section introduces the experimental setup, the ground truth labeling methodology, the baseline model, and the experimental results.

### A. Experimental Setup

The experimental validation was conducted in a motion capture lab equipped with a Qualisys system running at 200Hz, using a mix of Miquis M3 and Arqus A12 cameras (15 cameras in total) with resolutions of 2MP and 12MP, respectively. The floor was integrated with two Kistler 9260AA6 force plates sampling at 2kHz, providing ground-truth GRFs. The setup is shown in Fig. 6. The sensor was mounted on all four feet of the Unitree Go1 quadruped. However, for clarity of evaluation and to isolate the sensing and inference pipeline, only the front-left foot was used for real-time force inference and slip detection in this study. The robot was tasked with trotting across transitioning terrains

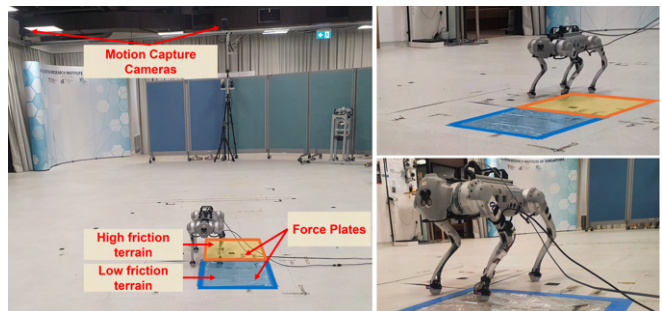


Fig. 6. Experimental setup in motion capture lab with force plates. The setup involves a setup with two force plates with a non-slippery terrain (orange) and a slippery terrain (blue).

consisting of a high-friction laboratory floor ( $\mu_s \approx 0.8$ ) and a low-friction patch created using a vinyl sheet sprayed with WD40 lubricant ( $\mu_s \approx 0.1$ ) with speeds up to  $0.3m/s$ . Three reflective markers were placed on the front-left leg (one on the foot and two on the shank forming a triangular plane) to track the relative orientation between the sensor frame and the force plate frame. The marker directly on the foot was used for extracting foot positions. During trials, we recorded motion capture kinematics, force plate GRFs, force sensor readings, and slip detection outputs from both our method and a kinematic baseline.

### B. Ground Truth Labeling

To obtain ground truth slip labels, we leveraged the motion capture system to track the tangential foot position during stance. Note that in similar works such as [4], the ground truth labels were simply determined based on whether the foot is on the slippery patch or not. However, this may result in false labeling as there are cases where slipping occurs on the non-slippery patch due to residue, or no slipping on the slippery patch. In our work, a stance is labeled as slip if the

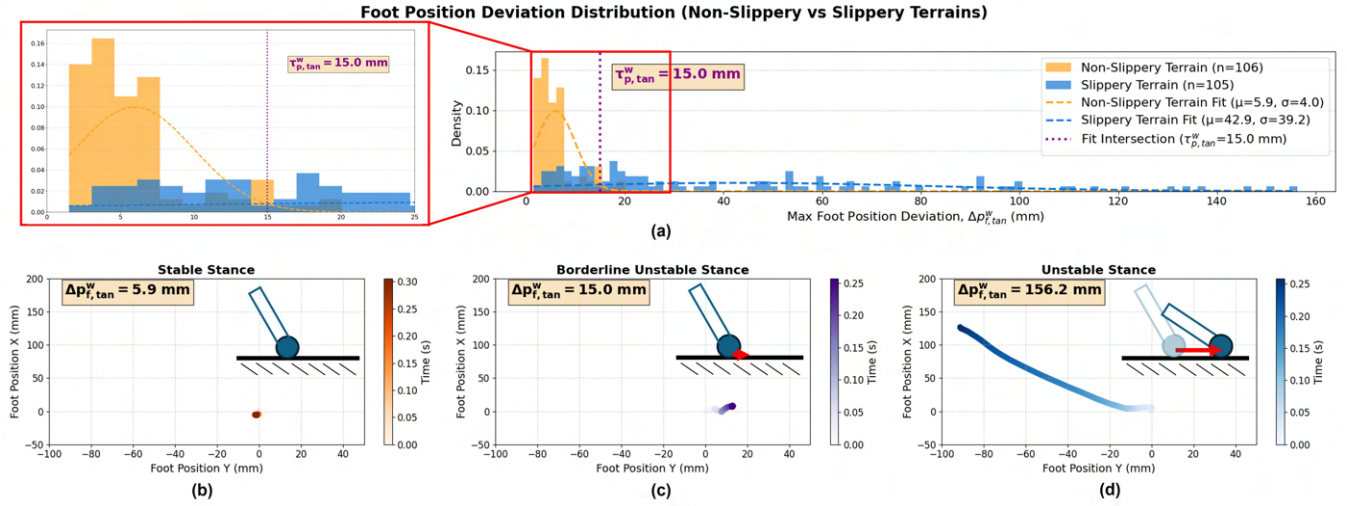


Fig. 7. Data-driven ground truth labeling and visualization of different slips. (a) distributions of foot position deviations,  $\Delta p_{f,tan}^w$ , using motion capture for non-slippery (orange) and slippery (blue) terrains and the determined threshold,  $\tau_{p,tan}^w$ . Single stance foot positions over time are shown in (b) stable foothold ( $\Delta p_{f,tan}^w < \tau_{p,tan}^w$ ), (c) borderline slip stance ( $\Delta p_{f,tan}^w = \tau_{p,tan}^w$ ), and (d) extreme slip case ( $\Delta p_{f,tan}^w > \tau_{p,tan}^w$ ).

deviation of the foot position from its initial contact point,  $\Delta p_{f,tan}^w$ , exceeds a threshold, as calculated from:

$$\Delta p_{f,tan}^w = \sqrt{(x_f^w(t) - x_f^w(t_0))^2 + (y_f^w(t) - y_f^w(t_0))^2}, \quad (8)$$

where  $t_0$  denotes the foot touchdown. The ground truth slip label,  $s$ , is therefore defined as:

$$s = \begin{cases} 1, & \text{if } \Delta p_{f,tan}^w > \tau_{p,tan}^w, \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

To determine the threshold,  $\tau_{p,tan}^w$ , we adopt a data-driven approach (Fig. 7(a)). We collected foot position data over 211 stances across both non-slippery (orange) and slippery terrains (blue). For each stance, we computed  $\Delta p_{f,tan}^w$ . The resulting distributions are plotted and fitted with Gaussian curves, and the intersection point between the two distributions is chosen as the threshold for ground truth labeling. Fig. 7(b)-(d) further illustrates example stance trajectories, showing a stable stance, a borderline unstable stance, and an unstable stance with over 150mm of foot displacement. Given the difficulty in explicitly labeling slip, this approach provides a robust, data-driven method for generating reliable slip labels to validate our framework.

### C. Baseline Slip Detection Model

To assess the performance of our slip detection framework, we implemented a baseline model using a common proprioceptive method based on kinematics. The core principle, discussed in works such as [3], is that a stable foothold should exhibit near-zero velocity. This world-frame velocity check serves as a point of comparison and its components are integrated in various proprioceptive detectors [4] [20]. To ensure a rigorous and fair comparison, the baseline's velocity threshold was determined using the same data-driven, one-class methodology and stable contact dataset used for our

SlipSense framework. The foot velocity in the world frame,  $\mathbf{v}_f^w$ , is computed as

$$\mathbf{v}_f^w = \mathbf{v}_b^w + \mathbf{R}_b^w (\mathbf{v}_f^b + \boldsymbol{\omega}^b \times \mathbf{p}_f^b) \quad (10)$$

where  $\mathbf{v}_b^w$  is the robot base linear velocity in world frame,  $\mathbf{R}_b^w$  is the rotation matrix from base to world frame obtained from the base IMU quaternion,  $\mathbf{v}_f^b$  is the foot velocity expressed in the base frame,  $\boldsymbol{\omega}^b$  is the base angular velocity, and  $\mathbf{p}_f^b$  is the foot position in the base frame. The tangential component is then extracted as

$$v_{f,tan}^w = \sqrt{(v_{f,x}^w)^2 + (v_{f,y}^w)^2}. \quad (11)$$

We extract  $v_{f,tan}^w$  values across stance phases and set the 99.5<sup>th</sup> percentile as the slip threshold,  $\tau_{vel}$ . During inference, slip is flagged whenever  $v_{f,tan}^w$  exceeds this threshold, as represented by:

$$s_{baseline} = \begin{cases} 1, & \text{if } v_{f,tan}^w > \tau_{vel}, \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

### D. Results: Online Slip Detection

We experimentally validate the performance of our SlipSense framework against the baseline. For over 200 stance sequences, we observe an accuracy of 85.9% on the multimodal SlipSense feature set which significantly outperforms the kinematics baseline's 69.3%, as shown in Table I. An ablation study was performed to provide insight into the necessity of multimodal fusion. The AE model was trained using two feature set variants – force-only and inertial-only. The force-only model, while more accurate than the baseline, struggled to differentiate between slip-induced force behavior and intended initial ground impact effects, limiting its accuracy to 73.7%. Conversely, the inertial-only model achieved a higher accuracy of 83.9%, demonstrating the benefit of gathering inertial information right at the end

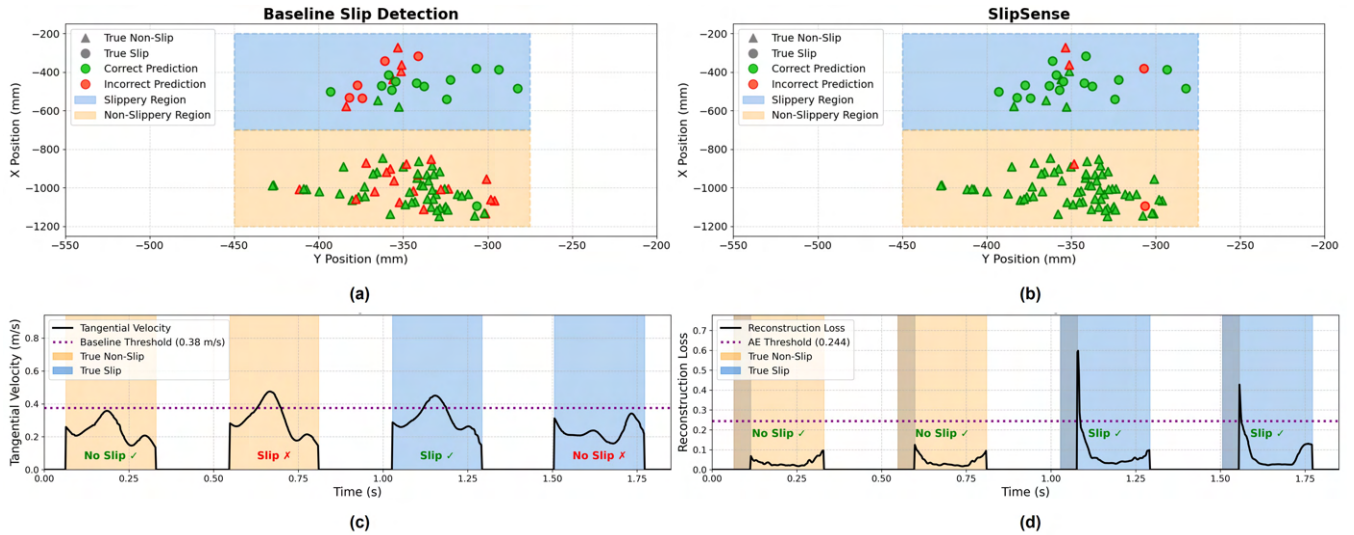


Fig. 8. Experimental results comparison between baseline and the proposed multimodal SlipSense slip detection models. Shown are 100 stances and the corresponding correct/incorrect predictions for (a) baseline model and (b) SlipSense. The time series plot over four stances are shown and the corresponding thresholds plotted for the (c) baseline and (d) SlipSense models.

TABLE I

SLIP DETECTION PERFORMANCE AND RESOLUTION ACROSS MODELS.

Model	Accuracy (%)	Smallest Slips <sup>1</sup> (mm)	Min. Detected Slip (mm)
Baseline Detector	69.3	80.8 ± 35.6	16.5
SlipSense (Force-only)	73.7	<b>20.7 ± 4.2</b>	<b>15.4</b>
SlipSense (IMU-only)	83.9	37.3 ± 15.5	15.6
SlipSense (Multimodal)	<b>85.9</b>	24.1 ± 6.4	<b>15.4</b>

<sup>1</sup>Mean ± Std. Dev. over the 30 smallest correctly classified slips.

effector, but its sensitivity was limited. By capturing the high-frequency transient spikes as foot slippage occurs, the force derivatives enable the detection of early slips, justifying the multimodal fusion.

Fig. 8(a)-(b) shows 100 stance phases of the robot traversing over two terrains of different friction to qualitatively evaluate the performance of the two models. The baseline model exhibits two primary failure modes. Firstly, a significant number of false positives on the non-slippery terrain is observed, where stable stances are incorrectly flagged as slip. This is attributed to noise in the world frame foot velocity estimate and the inability of the model to differentiate between a slip and intended ground impact. Secondly, a large proportion of false negatives is observed, in particular for low-displacement slips that occur near the ground truth 15mm boundary. In contrast, our SlipSense demonstrates superior performance over the baseline, with a significant reduction in false positives and negatives. A comparison of the time-series plots in Fig. 8(c)-(d) shows our model's reconstruction error remains low during non-slip stances and spikes under the slip conditions. Over four consecutive stances, our model successfully classified the slip state, while the baseline model resulted in a false positive and false negative.

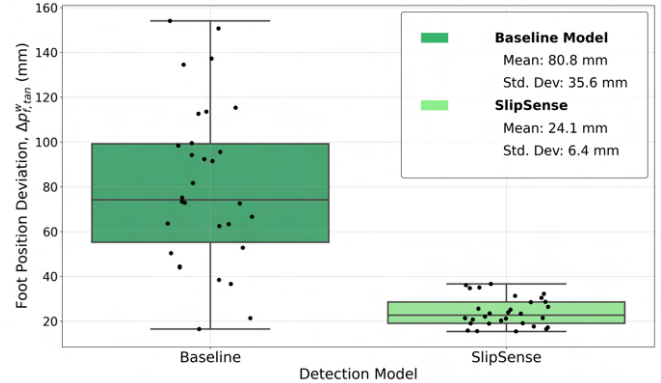


Fig. 9. Box plot showing the smallest 30 slip stances accurately classified for the baseline and SlipSense models.

A key contribution of our work is the improved resolution of slip detection, as shown in Table I and Fig. 9. Over the 30 smallest correctly detected slip stances, which represents the best case scenario for each model, the baseline model was only able to detect slips with an average minimum displacement of  $80.8 \pm 35.6$ mm, which represents significant gross slips and large variability. Our multimodal SlipSense model, however, achieved an average minimum detected slip of  $24.1 \pm 6.4$ mm; a 3.3-fold improvement from baseline. The significantly lower mean and standard deviation demonstrates its superior ability to detect smaller slips more consistently. Detection of slips with 24.1mm foot displacement provides a crucial early warning to the downstream controller, enabling ample time for recovery and mitigation strategies before catastrophic instability.

## VII. CONCLUSION AND FUTURE WORK

This paper presents SlipSense, an end-to-end framework for online slip detection using a custom multimodal force

sensor. By fusing kinetic and kinematic data within a self-supervised anomaly detection model, our framework demonstrates a 3.3-fold improvement in slip detection resolution and higher overall accuracy over a velocity-level kinematic baseline.

While promising for integration into terrain-aware controllers, we acknowledge limitations that provide clear avenues for future research. The model was validated on rigid, homogeneous surfaces with a set trotting gait. Extension to unstructured ground or inclines requires improved generalization in the one-class learning. Secondly, our current slip detection involves a binary classification task in which a data-driven labeling method is adopted, which introduces ambiguity near the threshold. The AE model shows significantly more errors near this threshold. Thus, we look to create a third class of incipient slip near the boundary to better represent the labeling function.

Future work will involve optimizing the force inference model to improve the tangential force estimates which will improve the downstream slip detection task overall. SlipSense will be extended to perform online terrain friction estimation which can be integrated into a robot controller to perform slip recovery and terrain-aware gait adaptation. Finally, integration of all four foot force sensors will also be explored. The overall paper forms the foundation for future force-based gait adaptation control.

#### ACKNOWLEDGMENTS

Iris Szu-Yao Liu is an A\*STAR scholar under the Singapore International Graduate Award (SINGA) framework. The authors would like to thank Yuehang (Aria) She and Alex Lim Lek Syn for assisting with the experiments.

#### REFERENCES

- [1] Gerardo Bledt, Patrick M Wensing, Sam Ingersoll, and Sangbae Kim. Contact model fusion for event-based locomotion in unstructured terrains. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4399–4406. IEEE, 2018.
- [2] Will Bosworth, Sangbae Kim, and Neville Hogan. The MIT super mini cheetah: A small, low-cost quadrupedal robot for dynamic locomotion. In *2015 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pages 1–8. IEEE, 2015.
- [3] Michele Focchi, Victor Barasuol, Marco Frigerio, Darwin G Caldwell, and Claudio Semini. Slip detection and recovery for quadruped robots. *Robotics Research: Volume 2*, pages 185–199, 2017.
- [4] Chengzhen Yan, Jiahu Qin, Qingchen Liu, and Qichao Ma. Slip detection and recovery for quadruped robots via orthogonal decomposition. *IEEE Transactions on Industrial Electronics*, 2024.
- [5] Péter Fankhauser, Marko Bjelonic, C Dario Bellicoso, Takahiro Miki, and Marco Hutter. Robust rough-terrain locomotion with a quadrupedal robot. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5761–5768. IEEE, 2018.
- [6] Linping Chan, Fazel Naghdy, and David Stirling. Extended active observer for force estimation and disturbance rejection of robotic manipulators. *Robotics and Autonomous Systems*, 61(12):1277–1287, 2013.
- [7] Zhang Cong, An Honglei, Chengye Wu, Lin Lang, Qing Wei, and Ma Hongxu. Contact force estimation method of legged-robot and its application in impedance control. *IEEE Access*, 8:161175–161187, 2020.
- [8] Elizabeth A Stone, Nathan F Lepora, and David AW Barton. Walking on TacTip toes: A tactile sensing foot for walking robots. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 9869–9875. IEEE, 2020.
- [9] Lindsay Epstein, Andrew SaLoutos, Donghyun Kim, and Sangbae Kim. Bi-modal hemispherical sensors for dynamic locomotion and manipulation. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7381–7381. IEEE, 2020.
- [10] Felix Ruppert and Alexander Badri-Spröwitz. Foottile: a rugged foot sensor for force and center of pressure sensing in soft terrain. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4810–4816. IEEE, 2020.
- [11] X Alice Wu, Tae Myung Huh, Rudranarayan Mukherjee, and Mark Cutkosky. Integrated ground reaction force sensing and terrain classification for small legged robots. *IEEE Robotics and Automation Letters*, 1(2):1125–1132, 2016.
- [12] X Alice Wu, Tae Myung Huh, Aaron Sabin, Srinivasan A Suresh, and Mark R Cutkosky. Tactile sensing and terrain-based gait control for small legged robots. *IEEE Transactions on Robotics*, 36(1):15–27, 2019.
- [13] Ziwu Song, Chenchang Li, Zhentan Quan, Shilong Mu, Xiaosa Li, Ziyi Zhao, Wanxin Jin, Chenye Wu, Wenbo Ding, and Xiao-Ping Zhang. Tactid: High-performance visuo-tactile sensor-based terrain identification for legged robots. *IEEE Sensors Journal*, 2024.
- [14] 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.
- [15] Xiaofeng Guo, Bryan Blaise, Jennifer Molnar, Jeremiah Coholich, Shantanu Padte, Ye Zhao, and Frank L Hammond. Soft foot sensor design and terrain classification for dynamic legged locomotion. In *2020 3rd IEEE International Conference on Soft Robotics (RoboSoft)*, pages 550–557. IEEE, 2020.
- [16] Taiyu Okatani and Isao Shimoyama. Evaluation of ground slipperiness during collision using MEMS local slip sensor. In *2019 IEEE 32nd International Conference on Micro Electro Mechanical Systems (MEMS)*, pages 823–825. IEEE, 2019.
- [17] Jaecun Park and Hae-Won Park. Design of anti-skid foot with passive slip detection mechanism for conditional utilization of heterogeneous foot pads. *IEEE Robotics and Automation Letters*, 4(2):1170–1177, 2019.
- [18] Thijs Van Hauwermeiren, Anatolii Sianov, Annelies Coene, and Guillaume Crevecoeur. Integrated barometric pressure sensors on legged robots for enhanced tactile exploration of edges. *IEEE Robotics and Automation Letters*, 2024.
- [19] Georg Wiedebach, Sylvain Bertrand, Tingfan Wu, Luca Fiorio, Stephen McCrory, Robert Griffin, Francesco Nori, and Jerry Pratt. Walking on partial footholds including line contacts with the humanoid robot atlas. In *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, pages 1312–1319. IEEE, 2016.
- [20] Ylenia Nisticò, Shamel Fahmi, Lucia Pallottino, Claudio Semini, and Geoff Fink. On slip detection for quadruped robots. *Sensors*, 22(8):2967, 2022.
- [21] Michael Bloesch, Christian Gehring, Péter Fankhauser, Marco Hutter, Mark A Hoepflinger, and Roland Siegwart. State estimation for legged robots on unstable and slippery terrain. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 6058–6064. IEEE, 2013.
- [22] Sangli Teng, Mark Wilfried Mueller, and Koushil Sreenath. Legged robot state estimation in slippery environments using invariant extended kalman filter with velocity update. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3104–3110. IEEE, 2021.
- [23] Michael Maravagakis, Despina-Ekaterini Argiropoulos, Stylianos Piperakis, and Panos Trahanias. Probabilistic contact state estimation for legged robots using inertial information. *arXiv preprint arXiv:2303.00538*, 2023.
- [24] Peng Sun, Junjie Qiang, Letian Qian, and Xin Luo. Learning slip detection for agile locomotion of quadruped robots. In *2023 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 1–6. IEEE, 2023.
- [25] Meng Yee Michael Chuah, Lindsay Epstein, Donghyun Kim, Juan Romero, and Sangbae Kim. Bi-modal hemispherical sensor: A unifying solution for three axis force and contact angle measurement. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7968–7975. IEEE, 2019.
- [26] Meng Yee Michael Chuah and Sangbae Kim. Enabling force sensing during ground locomotion: A bio-inspired, multi-axis, composite force sensor using discrete pressure mapping. *IEEE Sensors Journal*, 14(5):1693–1703, 2014.