2019 International Conference on Robotics and Automation (ICRA)
Palais des congres de Montreal, Montreal, Canada, May 20-24, 2019

Every Hop is an Opportunity: Quickly Classifying and Adapting to Terrain During Targeted Hopping

Alexander H. Chang[1], Christian Hubicki[2], Aaron Ames[3] and Patricio A. Vela[1]

Abstract— **Practical use of robots in diverse domains requires programming for, or adapting to, each domain and its unique characteristics. Failure to do so compromises the ability of the robot to achieve task-relevant objectives. Here we describe how the learned terrain reaction force profiles of a hopping robot serve the additional objectives of classifying terrain and quickly learning control strategies to accomplish a jumping task on novel terrain. We show that the reaction forces experienced during closed-loop jumping are sufficient to discriminate between three different terrain types (granular, trampoline, and rigid) when using the learned models as discriminators. Building on this, we show that applying the classification to unknown terrain types leads to faster task completion, where the task objective is to meet a specific jump height. The classification experiments, utilizing real-world jumping data, achieve 95% prediction accuracy. The online learning experiments leverage simulation as there is more control over the terrain properties. Terrain-informed learning achieves the target hop heights more than 2x faster than without terrain knowledge when the prediction is correct, and 1.5x faster when the prediction is incorrect. Thus, applying the closest approximately known terrain knowledge facilitates low shot learning when hopping on unknown terrain.**

I. Introduction

Legged robotic platforms are uniquely advantaged with respect to their traditionally wheeled or tracked counterparts. In particular, the former are well-equipped to traverse commonly-occurring environmental impediments and man-made environmental features that the latter cannot. Accurate models of robot-terrain interactions, however, are critical to motion planning in legged systems–even small, uncorrected errors may lead to catastrophic consequences. Despite the availability of terradynamic models for describing terrain, many environments remain undescribed or do not have *a priori* known parameters, especially in the context of field deployment. Missing terrain knowledge presents challenges for competent application of legged robotic platforms in many practical mission scenarios. This paper aims to address the challenge of incomplete terrain information by employing prior, approximate knowledge for fast, low-shot learning on novel terrain.

1: The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. Email: alexander.h.chang@gatech.edu, pvela@gatech.edu

2: The author is with the Dept. of Mechanical Engineering, Florida A&M University-Florida State University, Tallahassee, FL, USA. Email: hubicki@eng.famu.fsu.edu

3: The author is with the Dept. of Mechanical and Civil Engineering, California Institute of Technology, Pasadena, CA, USA. Email: ames@caltech.edu

This work was supported by NSF Award #1544857.

For bipedal and quadrupedal systems, incorrectly modeled terrain forces can compromise critical stability criteria, leading to falls. For hexapodal systems, mismatch between actual terrain dynamics and those assumed for control lead to inefficient or ineffective locomotion. Variations in encountered terrain dynamics have warranted the application of differently tuned control strategies to address a pre-determined set of unique foot-ground interaction forces [1], [2]. Controllers optimized for a particular terrain, however, may quickly become sub-optimal or unstable as terrain parameters experience geographical variations; terrain category itself may change quite abruptly, as exemplified by soil-sand boundaries found near beaches. An expeditious manner of modeling foot-ground interactions online, during the course of locomotion, is warranted. Furthermore, an ability to quickly distinguish among disparate terrain categories presents the opportunity to expedite terrain characterization and task-specific control synthesis.

A. Contribution

Our prior work [1] described a fast learning strategy for achieving hopping on unknown terrain with a specified jump-height task objective. Attempts to meet the objective provide learning data from which to recover the unknown terrain reaction forces and to meet the specified task objective. Inspired by [3], [4], where terrain reaction force models appeared to have discriminative properties, we now ask the question: *Are the learnt terrain reaction force models sufficient to discriminate between terrain types?* We, furthermore, study the follow-on inquiry: *Does the learning rate on novel terrain improve when approximate knowledge of the nearest available terrain category informs the learning process?*

We show that the answer is affirmative in both cases. The value of the work is that terrain reaction forces learned in the context of control can serve the secondary role of classifying ground types as well as knowledge transfer for rapid (e.g. *low-shot*) learning of similar but novel terrains. This supports the conclusion that learning for control may assist with approximate categorization of–and rapid adaptation to–novel, diverse domains. Task-based learning, control, and classification are shown to benefit from shared models.

B. Relevant Work

Terradynamic estimation, and subsequently informed control, occupies a significant role particularly in the context of autonomous vehicles where operator experience and training cannot be leveraged for corrections. Novel techniques must

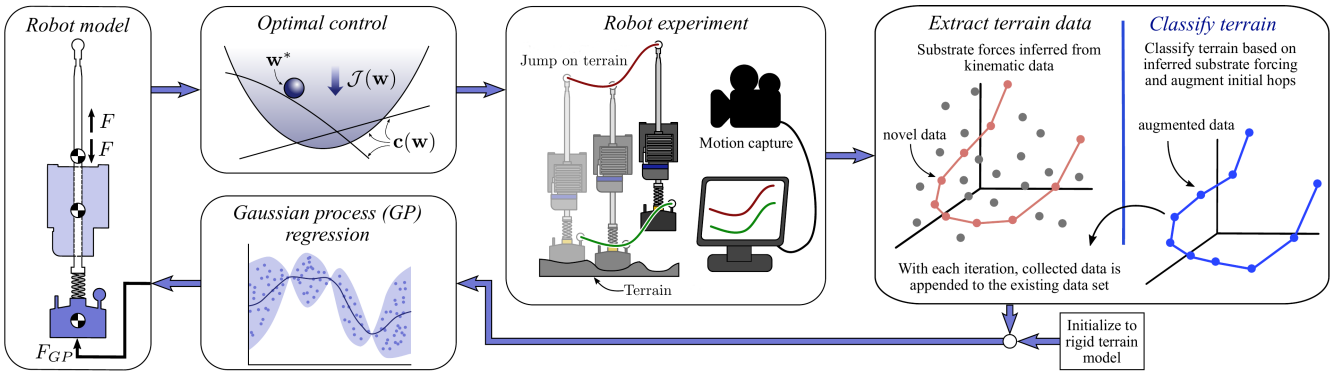

Fig. 1: Flowchart depicting the iterative optimal control/learning framework in Algorithm 1. With each jump, the algorithm regresses an updated terrain model, and then uses the learned model to synthesize and execute a new optimal control signal. New contributions are illustrated by the *Classify Terrain* sub-block, to the right.

be applied to estimate consequences of uncertain vehicle-terrain parameters in-situ [5]–[7]. Terrain classification carries varying degrees of criticality for different robotic platforms. In the context of legged machines, knowledge of relevant terradynamics is critical to safe and efficient operation. Mismatches in relevant terrain and applied control strategies lead to falls in the context of bipedal or quadrupedal platforms while entailing severe performance degradation for other legged mechanisms. Terrain classification approaches, in the context of mobile robotics, have relied on a variety of available sensing modalities and task-oriented affordances.

Exteroceptive Techniques. Visual information has been a primary choice to aid slip and traversability estimation in extraterrestrially-motivated vehicles [7]–[9]. Fusion of visual sensing modalities with additional tactile information has additionally been applied to enhance terrain classification accuracy [10]. Other perceptual modalities have demonstrated potential to classify terrain with varying levels of granularity. Passive acoustic information was demonstrated to distinguish not only material composition but a variety of hard, soft and granular surfaces [3], [11]. Likewise, thermal imaging aids the noninvasive assessment of granular terrain properties impacting traversability [12], [13]. Augmentation with tactile-force sensing enables distinguishing categorically different locomotion surfaces during the course of mission execution, without the need for dedicated probing procedures [14].

Proprioceptive Techniques. In the specific context of legged mechanisms, measurement of joint force and torque are conveniently exploited, whereby prescribed limb motions allow various properties of the locomotion substrate to be measured. The information obtained reveals terrain properties such as stiffness, friction and shape [2], [4], [15]. Exploratory motion strategies are then crafted to inspect surrounding terrain for safety-critical information [4], [16]. Terrain classification and characterization afford robotic platforms the capacity to distinguish un-traversable terrain as well as assess traversability prior to mission execution.

Proprioceptive sensing augments exteroceptive modalities to yield enhanced classification results. Vibration-based classifiers, for example, produce results that have been merged with, or applied to train, vision-based terrain classifiers deployed on vehicles intended for unmanned, unmonitored operational situations [17]–[19].

Task-Concurrent Perception. Exploratory procedures dedicated to probing the environment may be time and task prohibitive. In these scenarios, classification strategies employing sensing modalities already accessible and utilized for normal operation are favorable. They rely on the ability to measure disturbances from a nominal model of operation during the course of locomotion [2], [20], Measured disturbance patterns effectively encode terrain characteristics of interest that, when classified or characterized, inform selection of appropriate control strategies or tuning.

Classification Approach. Support Vector Machines (SVM) [10], [11], [14] and neural networks [20], [21] are the predominant classification approaches applied toward terrain identification for mobile robotic applications.

We pursue an alternative classification approach whereby an iterative, incremental learning strategy, based on Gaussian Processes (GPs), is extended. We leverage a pre-existing knowledge base, both as a foundation for accurate terrain classification as well as a mechanism facilitating low-shot learning of initially unknown robot-terrain dynamics.

II. Learning to Hop

An online, iterative optimal control synthesis and learning framework was previously presented [1] involving a cyclic procedure (Fig. 1) to learn the terrain forcing model and generate an optimal motion plan. Beginning with a solid terrain reaction force model, the process generates an optimal trajectory to meet a target hopping height. Applying it to simulated granular media (GM) and collecting the dynamic response provides data for training a Gaussian process (GP) model of the terrain forcing. Utilizing the model, within the optimal trajectory synthesis algorithm, leads to improved task performance, with the hopper meeting the target objectives within 2-3 control-learning iterations. Experiments employing this iterative control-learning procedure were subsequently conducted, demonstrating the procedure on categorically distinct surface types: *solid ground*, *trampoline*

and *granular media*. Here, we briefly review the procedure and underlying components; we refer the reader to [1] for comprehensive detail.

A. System Dynamics

The one-dimensional hopper, illustrated in Fig. 2(a) is modeled by three body masses: a linear motor, a thrust rod, and a foot [22], [23]. Jumps are achieved by applying a force between the motor and thrust rod which, in turn, transmits forcing to the foot through the connecting spring. The closed-form system dynamics are,

$$\ddot{x}_f = -g + \frac{1}{m_f}\left[k(\alpha - \bar{\alpha}) + c\dot{\alpha}\right] + \frac{1}{m_f}F_{sub} \tag{1a}$$

$$\ddot{\alpha} = -\left[\frac{m_f + m_r + m_m}{m_f(m_r + m_m)}\left(k(\alpha - \bar{\alpha}) + c\dot{\alpha}\right) + \frac{m_m}{m_r + m_m}u\right] - \frac{1}{m_f}F_{sub}. \tag{1b}$$

Masses of the motor, rod, and foot are m_m, m_r, and m_f, respectively. The value $\bar{\alpha}$ is the resting length of the spring (absent external forcing). The system state is $\boldsymbol{x} = [x_f\ \dot{x}_f\ \alpha\ \dot{\alpha}]^T \in \mathbb{R}^4$, where x_f and α denote the spatial position of the foot and relative position of the thrust rod with respect to the foot. Motor acceleration, $u = \ddot{\beta}$, serves as the control input, where β is the relative position of the motor with respect to the rod. Terrain reaction forcing is modeled as $F_{sub} > 0$ during the stance phase of the hybrid system. This quantity is initially unknown and presumed to model solid ground. During the flight phase $F_{sub} = 0$. Spring stiffness, k, and gravitational acceleration, g, are denoted.

B. Optimal Control

The task objective is to achieve a target jump height, as illustrated in Fig. 2(b). The jump height is the difference between initial rod height during the stance phase, x_r^0, and the rod's highest point during the flight phase, x_r^f, where the rod height is $x_r = x_f + \alpha$. Optimal trajectory synthesis methods generate an optimal control signal, u_{opt}, to meet the task. Here, we use a direct collocation approach to transcribe the trajectory optimization problem into a nonlinear program (NLP) [23]. Task objectives, actuator limits, system dynamics and other task-specific requirements are encoded as hard constraints defining the optimization problem. The optimal control synthesis relies on **(1)** an *a priori* robot model and **(2)** a GP-based model of the ground substrate forcing.

C. Gaussian Process-based Terrain Model

The unknown quantity in the system, F_{sub}, is modeled as the summation of multiple components,

$$F_{sub}(\boldsymbol{x}) = F_{sub}^{sg}(\boldsymbol{x}) + \mu(\boldsymbol{x}) + r^T(\boldsymbol{x})R^{-1}(\hat{\boldsymbol{y}} - \mu(\hat{\boldsymbol{x}})), \tag{2}$$

where $F_{sub}^{sg}(\boldsymbol{x})$ is the ideal solid ground forcing model, a known function of the hopper state [1]. The latter two terms model the "defect" in the predicted substrate forcing, $F_{sub}(\boldsymbol{x})$, with respect to the solid ground model. They are initially set to the zero function.

During hopping trials, measured state data passed through a Kalman smoother estimates the accelerations to then recover, from (1), the unknown forcing signal, $\hat{F}_{sub}$. Collected state data is denoted, $\hat{\boldsymbol{x}} = [\hat{\boldsymbol{x}}_1\ \hat{\boldsymbol{x}}_2\ \ldots\ \hat{\boldsymbol{x}}_p]^T$, where $\hat{\boldsymbol{x}}_i \in \mathbb{R}^4$. Associated measurements of the "defect" are maintained as, $\hat{y}_i = \Delta F_{sub}(\hat{\boldsymbol{x}}_i) = \hat{F}_{sub} - F_{sub}^{sg}(\hat{\boldsymbol{x}}_i)$. Together, $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}} = \{\hat{y}_i\} \in \mathbb{R}^p$ comprise a set of p training samples, whereby the latter two terms of (2) are learned using a composite Gaussian process (CGP) [24]. The function $\mu(\boldsymbol{x})$ captures the global trend in the training data as a linear combination of basis functions, $f(\boldsymbol{x}) = [1, x_f, \dot{x}_f, \alpha, \dot{\alpha}]^T$, such that $\mu(\boldsymbol{x}) = f(\boldsymbol{x})^T\boldsymbol{\Theta}$, with coefficients $\boldsymbol{\Theta} \in \mathbb{R}^5$. The last term in (2) describes $Z(\boldsymbol{x}) \sim \mathrm{GP}(\kappa)$, a non-stationary, zero-mean GP applied to the residual deviations from the global trend evaluation, $\mu(\hat{\boldsymbol{x}}) \in \mathbb{R}^p$. A squared exponential kernel, $\kappa(\boldsymbol{x}, \boldsymbol{x}')$, was selected to model covariance in the collected data based on proximity between any two elements, $\boldsymbol{x}$ and $\boldsymbol{x}'$. The function R is the covariance matrix such that $R_{i,j} = \kappa(\hat{\boldsymbol{x}}_i, \hat{\boldsymbol{x}}_j)$ and $r^T(\boldsymbol{x}) = [\kappa(\boldsymbol{x}, \hat{\boldsymbol{x}}_1)\ \kappa(\boldsymbol{x}, \hat{\boldsymbol{x}}_2)\ \ldots\ \kappa(\boldsymbol{x}, \hat{\boldsymbol{x}}_p)]^T$.

Algorithm 1 Iterative Optimal Control and Learning

procedure CONTROL AND LEARN(hopperControlTask)
  training set, $\Gamma \leftarrow \emptyset$
  initialize CGP
  while $\neg$TASKCOMPLETE(hopperControlTask) **do**
    generate optimal control, u_{opt}, using CGP-based model, (1) and (2)
    apply u_{opt} on true terrain
    measure hopper position trajectory
    currentHopHeight $\leftarrow \max(x_f(t) + \alpha(t))$
    apply Kalman smoothing to recover $\hat{x}$ and $(\ddot{x}_f, \ddot{\alpha})$
    compute $\hat{y}_i = \Delta\hat{F}_{sub}(\hat{x}_i)$ from smoother data
    if not classified **then** classify terrain per (3).
    end if
    if less than 2 hops **then**
      append Γ with new data: $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}) \bigcup (\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}^{c^*})$
    else
      append Γ with new data: $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$.
    end if
    $\bar{\Gamma} \leftarrow$ REDUCE(Γ)
    update global trend, $\mu(x)$
    train GP with data: $(\hat{x}_i,\ \hat{y}_i - \mu(\hat{x}_i))$
  end while
end procedure

* blue text denotes low-shot learning algorithm modification

D. Iterative Control-Learning Procedure

Attempts to control a robotic hopper to optimally accomplish the hopping task yield novel learning experiences. These collective experiences enrich a CGP-based dynamical model of the system, in turn, better informing follow-on control synthesis attempts. The *original* Algorithm 1, given by Algorithm 1 excluding the blue text, delineates this iterative learning procedure (as depicted in Fig. 1). The optimal control solver employs the GP-based dynamical model to synthesize an optimal control strategy accomplishing the specified task. The synthesized control is applied to the robotic hopper. If the control objective is not attained, measurements taken during task execution generate additional

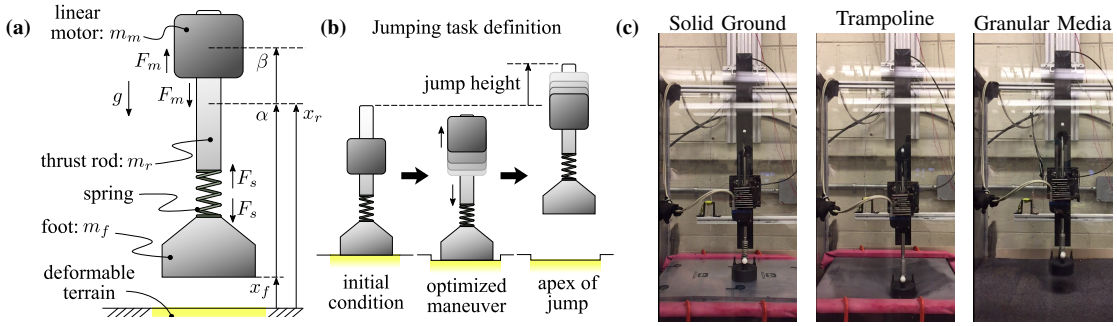

Fig. 2: **(a)** A visualization and labeling of the one-dimensional (linear vertical-only) jumping robot model. **(b)** An illustration of the desired jumping task achieved by controlling the motor mass to meet a target jump height. **(c)** Images of the three hopping terrain types: solid ground, trampoline, and granular media.

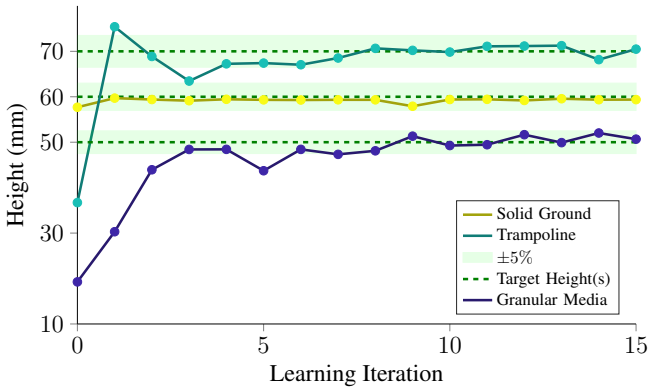

Fig. 3: Jump heights converge to target heights (dark green dashed lines) as terrain substrate forcing is learnt using the original Algorithm 1. Experiments were conducted on three different surface types using the robotic hopper apparatus in [23], depicted in Fig. 2(c).

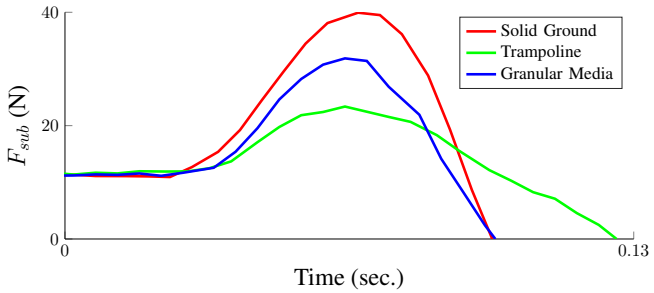

Fig. 4: A motor trajectory, designed to achieve a 60 mm target jump height, was synthesized assuming a solid ground model of the environment. The resulting terrain reaction forces differ based on the terrain type.

training data. The training set, Γ, comprising tuples of prior measured data, $(\hat{\boldsymbol{x}}_i, \hat{y}_i)$, is appended with this new experiential data and the GP-based model is re-trained.

E. Experimental Application

The original Algorithm 1 was applied to an experimental robotic hopper apparatus for different surface types. A rigid aluminum plate served as the solid ground substrate, modeling an undeformable surface. A trampoline surface was constructed from an elastic rubber-like sheet anchored tightly to a rigid hollow frame. A bed of poppy seeds comprised the granular media terrain [25]. Fig. 3 illustrates results for 3 targeted jump heights on different surfaces: 50 mm (*granular media*), 60 mm (*solid ground*) and 70 mm (*trampoline*). Measured hopper trajectories reach targeted jump heights, within $\pm 5\%$, after 8-10 control-learning iterations. Subsequent hopping iterations remained with $\pm 5\%$ of the target height. These results were reproduced for jump heights in the intervals $[40, 80]$, $[60, 90]$, and $[30, 60]$, for solid ground, trampoline and granular media ground types, respectively. The intervals are achievable hop heights for the terrain type.

III. Classifying Terrain

Given that a mechanism for the hopping system to quickly learn to meet the hop-height task objective exists, we now explore the value of the learnt substrate force models towards terrain classification. In particular, we show that the substrate forcing models per terrain type can serve to recognize the terrain category even when the terrain parameters differ from those learnt. Classification is performed by comparing the forcing predicted by the learnt models with the actual forcing. We leverage a collection of over 800 robotic hopper trajectories measured for various synthesized input control signals, on each of the three categorically distinct surfaces.

A. Hopper-Terrain Data Sets

Trajectories associated with experimental robotic jumps were collected and organized with respect to 3 terrain types: *solid ground*, *trampoline* and *granular media*. Fig. 4 illustrates recovered terrain reaction forcing applied by each surface type during the course of a jump. The hopper apparatus was driven by an identical control signal for each. The characteristics of each terrain manifest in distinct terrain reaction forcing profiles exerted on the foot of the hopper during the stance phase of each jump.

In total, we organized a collected set of 950 jump experiments according to the terrain type on which each was conducted. We furthermore designated $\mathcal{D}^{\textbf{official}}$, a subset of these hopper trajectories collected during repeated execution of the original Algorithm 1 for several targeted jump heights over each terrain type. Each experimental trial of the original Algorithm 1 comprised 15 control-learning iterations; a trial of the algorithm was conducted for each of the following

terrain type-target height configurations: *solid ground* (40, 50, 60, 70, 80 mm), *trampoline* (60, 70, 80, 90 mm) and *granular media* (30, 40, 50, 60 mm).

Training Data. A subset of trajectories in $\mathcal{D}^{\textbf{official}}$ was organized and denoted, $\mathcal{D}^{\textbf{official}}_{\text{train}}$. This subset comprised trajectories measured during learning iterations 0-2, for solid ground, and 0-8, for the other terrain types (i.e. iterations entailing the majority of advancement in the learnt model). In total, we have $|\mathcal{D}^{\textbf{official}}_{\text{train}}| = 87$ labeled and measured hopper trajectories encompassing experiments for the 3 terrain types, and driven by widely differing input control signals synthesized in the course of achieving the different target jump heights.

Test Data. The remaining 121 trajectories in $\mathcal{D}^{\textbf{official}}$ were organized as validation data set, $\mathcal{D}^{\textbf{official}}_{\text{test}}$. These correspond to iterations 9-15 (or 3-15 for solid ground) of the iterative control-learning procedure, associated with the same experimental trials that yielded data in $\mathcal{D}^{\textbf{official}}_{\text{train}}$. An additional collection of 742 jump experiments over the different terrain types was also available to leverage as test data. This set, designated $\mathcal{D}^{\textbf{unofficial}}_{\text{test}}$, comprises jumping trajectories measured during troubleshooting of the experimental setup as well as exploration of Algorithm 1 variations. It encompasses measurements subject to greater tracking error and noise, variation in sampling rate as well as variation of surface parameters (e.g. replacement of old poppy seeds, for GM, as well as wear and replacement of trampoline elastic material). We re-emphasize the mutual exclusivity of designated training and test data sets, $\mathcal{D}^{\textbf{official}}_{\text{train}} \cap \mathcal{D}^{\textbf{official}}_{\text{test}} \cap \mathcal{D}^{\textbf{unofficial}}_{\text{test}} = \emptyset$.

B. Classification Approach

Presume, now, that the hopper is jumping on unknown terrain. We present a terrain classifier employing the input-output data, $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$, typically recovered after a hop for learning purposes. Prior to learning, however, the classifier will first attempt to determine the surface type over which the robotic hopper is operating. Identification of the terrain will enable selection of a pre-existing knowledge base leveraged to expedite learning and, in turn, achieve a desired control objective in fewer control-learning iterations (§IV).

Training. The hopping data comprising the training data set $\mathcal{D}^{\textbf{official}}_{\text{train}}$ is used to train a set of CGPs labeled by terrain type, $\{\mathcal{CGP}^{sg}, \mathcal{CGP}^{tramp}, \mathcal{CGP}^{gm}\}$. Each CGP is trained using measured hopper trajectories collected from experiments on a single terrain type. These terrain-specific CGPs model the forcing profile exerted by each terrain type as a function of the hopper state space and accelerations. In particular, they are informed by $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$ measurements collected during early iterations of the Algorithm 1 when the learned model of terrain forcing has yet to mature and different control strategies are attempted (effectively exploring F_{sub}).

Procedure. Given that the reaction force profiles appear to be unique for each terrain type (Fig. 4), we use them to distinguish between the different terrain. In addition to the recovered input-output data $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$, we generate predicted input-output data $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}^{\boldsymbol{c}})$ where $c \in \mathcal{C} = \{sg, tramp, gm\}$ based on $\mathcal{CGP}^{c}$ evaluated on the set $\hat{\boldsymbol{x}}$ of trajectory data.

TABLE I: Confusion Matrix $\left[\text{Test Set, } \mathcal{D}^{\textbf{official}}_{\text{test}}\right]$: Terrain classifier validation using data collected in the same experimental trials from which training data originated.

	Ground Truth		
Prediction	sg	$tramp$	gm
sg	100% (65)	0%	0%
$tramp$	0%	100% (28)	0%
gm	0%	0%	100% (28)

TABLE II: Confusion Matrix $\left[\text{Test Set, } \mathcal{D}^{\textbf{unofficial}}_{\text{test}}\right]$: Terrain classifier evaluation using wider set of test data comprising 710 hopper trajectories, measured over each terrain types and driven by varying input control strategies.

	Ground Truth		
Prediction	sg	$tramp$	gm
sg	96.9% (250)	0.7% (2)	2.6% (5)
$tramp$	0%	92.8% (271)	0%
gm	3.1% (8)	6.5% (19)	97.4% (187)

Considering output data $\hat{\boldsymbol{y}}$, $\hat{\boldsymbol{y}}^{\boldsymbol{c}}$ as discrete measurements of continuous time signals, we treat the L^2-norm of the difference between the recovered signal and the predicted signal, for each class, as a similarity score. The minimal similarity score determines the terrain category:

$$c^* = \arg\min_{c \in \mathcal{C}} ||\hat{y}(t) - \hat{y}^c(t)||_{\mathcal{L}_2}. \tag{3}$$

Validation. The designated test sets, $\mathcal{D}^{\textbf{official}}_{\text{test}}$ and $\mathcal{D}^{\textbf{unofficial}}_{\text{test}}$, were input to the classifier. The outcomes for the two sets are given in the confusion matrices of Tables I and II, respectively. The numbers in parenthesis are the numerical quantities associated to the percentages. When evaluating hops on terrain matching that from the training set, the accuracy is 100%. When applying to more variable input data reflecting different instances of the same terrain types, the classifier had an overall accuracy of 95%. Per terrain classification accuracy associated remained well above 90%.

IV. Lower-Shot Learning through Knowledge Transfer

We modify Algorithm 1 to integrate the first-hop terrain classification outcome. The classification information alleviates the challenge of modeling initially unknown terrain dynamics; the task specializes, instead, to characterizing a particular terrain type. We capitalize on terrain category by leveraging the pre-existing terrain-specific CGP model to expedite learning. The resulting iterative learning system exhibits low-shot learning, as fewer hops are required to achieve the specified task objective, even when misclassified.

Low-Shot Learning Approach. After classification, rather than learn a custom terrain forcing CGP model with the dataset $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$ dataset, the learning dataset is augmented with the predicted tuples $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}^{\boldsymbol{c}^*})$. After the second hop, the same data augmentation is performed. The CGP training process of the modified Algorithm 1, which includes the blue text, thereby receives the amount of data equivalent to four hops during the first two hops. Predictions associated with the selected classifier CGP transfer knowledge and

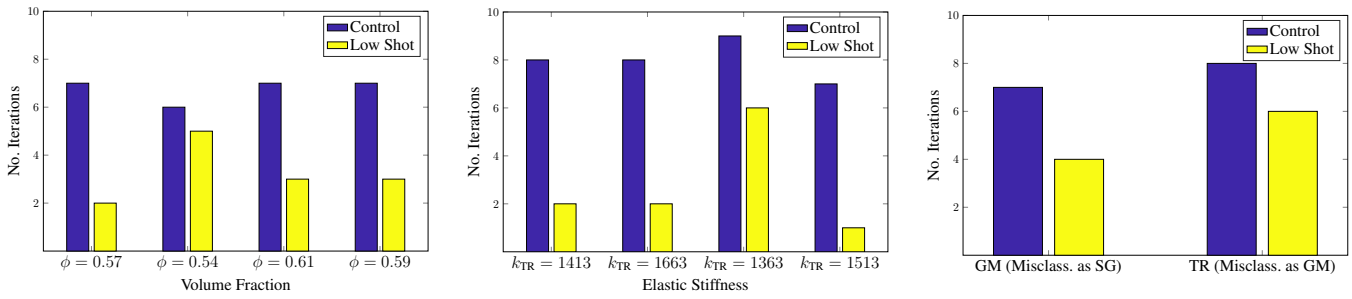

Fig. 5: **Low-Shot Learning.** Number of control-learning iterations required prior to achieving targeted jump height. **(a)** *Granular Media*: Several variants of GM terrain were generated by varying the terrain's volume fraction, ϕ. **(b)** *Trampoline Surface*: Surface variants were created by adjusting effective elastic stiffness, k_{TR}. **(c)** *Terrain Misclassification*: In cases where terrain was misclassified during Iteration 0, low-shot learning continued to be accomplished. More control-learning iterations were required to achieve targeted jump height compared to cases absent of misclassification. However, fewer iterations were needed relative to the counterpart control experiments.

expedite learning of the initially unknown terrain dynamics. Subsequent hops do not receive data augmentation. GP updates halt when the hopper achieves the target height.

A. Evaluation Methodology

Evaluation of classification-enabled, lower-shot learning is illustrated through several simulated experiments. Simulation presents the opportunity to vary terrain-specific parameters in addition to the terrain category. Both the granular media and trampoline surfaces are each specialized into 4 different variants; this is accomplished by varying volume fraction, for GM, and effective spring stiffness for the trampoline. No additional variants are derived for the solid ground model. For each terrain variant, random noise is injected into the measured hopper trajectories such that learning requires 7-9 iterations of the original Algorithm 1 to meet the hop height objective, much like that required in the real experiments (Fig. 3). A target jump height of 60 mm was specified for each experiment. The proposed classifier and data augmentation learning strategy was applied and compared to the original, non-augmented learning strategy.

We expect experiments on the GM and trampoline terrain variants to leverage relevant terrain knowledge bases and, in so doing, demonstrate lower-shot learning. After the initial jump in Iteration 0, only a handful more control-learning iterations should be required prior to the hopper achieving its target objective. This would be in stark contrast to baseline experiments, where the same target objective was instead achieved within 7-9 iterations.

B. Experimental Simulation

Granular Media. We simulate granular media terrain with the validated physical equations presented in [22]. They have been regularly applied to model poppy seeds [1], [23]. To create variants of this granular surface, we adjust the volume fraction (ie. compaction ratio), ϕ, describing the density of the granular terrain. The permissible volume fraction range, over which the model is valid, is small. However, small changes to this parameter greatly impact the simulated hopper trajectories. Fig. 5 plots the learning results on the 4 GM terrain variants. In each case, the terrain-informed learning procedure (yellow) expedites the learning process and facilitates achievement of the targeted jump height in fewer iterations, as compared to the baseline experiments (blue). In most cases, the low-shot learning procedures halves the number of iterations (averages a 2.2x speed-up).

Trampoline. Similar results hold for the simulated trampoline surface. We model this surface as a nonlinear spring and modify its effective elastic stiffness, k, to generate 4 variants for testing. Fig. 5 presents results on this surface. In all outcomes, the low-shot learning procedure allows the robotic hopper to achieve the target height in fewer iterations than the baseline method (averages a 2.9x speed-up). In several cases, the number of iterations executed reduced by 75%.

Misclassification. We assess performance of the low-shot learning procedure in the two misclassification scenarios most likely to occur based on the confusion matrix of Table II. In the first, a true GM surface is misclassified as solid ground; in the second a true trampoline surface is misclassified as GM. Fig. 5 presents the experimental outcomes. Despite terrain misclassification, the data augmented learning procedure still facilitated fast task completion. Though the misclassification degrades learning, it is still enhanced relative to the uninformed baseline approach (1.5x speed-up).

V. Conclusion

We demonstrated that the same external forcing model learnt to achieve task-specific control objectives can serve to distinguish different categorical models that lead to different forcing signals during operation. Furthermore, within the context of learning to hop to a target height over an unknown ground substrate, we showed that employing the learnt models for data augmentation during learning improves the learning process. The augmented learning system exhibits low-shot behavior relative to the uninformed learning system. It speeds up learning by more than a factor of 2 when correctly classified, and of 1.5 when incorrectly classified. The work shows that models learnt for control can serve the dual purpose of classifying terrain and informing low-shot learning of terrain substrate models even when approximate.

REFERENCES

[1] A. Chang, C. Hubicki, J. Aguilar, D. Goldman, A. Ames, and P. Vela, "Learning to jump in granular media: Unifying optimal control synthesis with gaussian process-based regression," in *IEEE ICRA*, Singapore, May 2017, pp. 2154–2160.

[2] W. Bosworth, J. Whitney, , and N. Hogan, "Robot locomotion on hard and soft ground: Measuring stability and ground properties in-situ," in *IEEE ICRA*, May 2016, pp. 3582–3589.

[3] E. Krotkov, "Active perception for legged locomotion: every step is an experiment," in *IEEE International Symposium on Intelligent Control*, Sep. 1990, pp. 227–232 vol.1.

[4] E. Krotkov, "Robotic perception of material," in *International Joint Conference on Artificial Intelligence*, Montreal, Quebec, Canada, 1995, pp. 88–94. [Online]. Available: http://dl.acm.org/citation.cfm?id=1625855.1625867

[5] T. M. Dar, "Vehicle-terrain parameter estimation for small-scale robotic tracked vehicle," Ph.D. dissertation, University of Texas at Austin, Austin, TX, 2010.

[6] J. Hidalgo-Carrió, D. Hennes, J. Schwendner, and F. Kirchner, "Gaussian process estimation of odometry errors for localization and mapping," in *IEEE ICRA*, Singapore, May 2017, pp. 5696–5701.

[7] C. Cunningham, M. Ono, I. Nesnas, J. Yen, and W. L. Whittaker, "Locally-adaptive slip prediction for planetary rovers using gaussian processes," in *IEEE ICRA*, Singapore, May 2017, pp. 5487–5494.

[8] L. Berczi, I. Posner, and T. D. Barfoot, "Learning to assess terrain from human demonstration using an introspective gaussian-process classifier," in *IEEE ICRA*, May 2015, pp. 3178–3185.

[9] A. Angelova, L. Matthies, D. Helmick, and P. Perona, "Slip prediction using visual information," in *Robotics: Science and Systems*, Philadelphia, USA, August 2006.

[10] K. Walas, "Terrain classification and negotiation with a walking robot," *Journal of Intelligent & Robotic Systems*, vol. 78, no. 3-4, pp. 401–423, June 2015.

[11] J. Christie and N. Kottege, "Acoustics based terrain classification for legged robots," in *IEEE ICRA*, May 2016, pp. 3596–3603.

[12] C. Cunningham, I. Nesnas, and W. L. Whittaker, "Terrain traversability prediction by imaging thermal transients," in *IEEE ICRA*, May 2015, pp. 3947–3952.

[13] C. Cunningham, W. Whittaker, and I. Nesnas, "Detecting loose regolith in lunar craters using thermal imaging," in *ASCE Conference on Engineering, Science, Construction, and Operations in Challenging Environments*, Orlando, FL, April 2016, pp. 16–26.

[14] X. A. Wu, T. M. Huh, R. Mukherjee, and M. Cutkosky, "Integrated ground reaction force sensing and terrain classification for small legged robots," *IEEE Robotics and Automation Letters*, vol. 1, no. 2, pp. 1125–1132, July 2016.

[15] M. A. Hoepflinger, C. D. Remy, M. Hutter, L. Spinello, and R. Siegwart, "Haptic terrain classification for legged robots," in *IEEE ICRA*, May 2010, pp. 2828–2833.

[16] P. R. Sinha and R. K. Bajcsy, "Implementation of an active perceptual scheme for legged locomotion of robots," in *IEEE IROS*, vol. 3, Nov 1991, pp. 1518–1523.

[17] K. Otsu, M. Ono, T. J. Fuchs, I. Baldwin, and T. Kubota, "Autonomous terrain classification with co- and self-training approach," *IEEE Robotics and Automation Letters*, vol. 1, no. 2, pp. 814–819, July 2016.

[18] I. Halatci, C. A. Brooks, and K. Iagnemma, "Terrain classification and classifier fusion for planetary exploration rovers," in *IEEE Aerospace Conference*, March 2007, pp. 1–11.

[19] C. A. Brooks and K. Iagnemma, "Self-supervised terrain classification for planetary surface exploration rovers," *Journal of Field Robotics*, vol. 29, no. 3, pp. 445–468, May 2012.

[20] C. Ordonez, J. Shill, A. Johnson, J. Clark, and E. Collins, "Terrain identification for RHex-type robots," in *Proc. SPIE*, vol. 8741, 2013, pp. 8741–12. [Online]. Available: https://doi.org/10.1117/12.2016169

[21] L. Ojeda, J. Borenstein, G. Witus, and R. Karlsen, "Terrain characterization and classification with a mobile robot," *Journal of Field Robotics*, vol. 23, no. 2, pp. 103–122, Feb. 2006.

[22] J. Aguilar and D. I. Goldman, "Robophysical study of jumping dynamics on granular media," *Nature Physics*, vol. 12, pp. 278–184, 2016.

[23] C. M. Hubicki, J. J. Aguilar, D. I. Goldman, and A. D. Ames, "Tractable terrain-aware motion planning on granular media: An impulsive jumping study," in *IEEEIROS*, Daejon, Korea, Oct 2016, pp. 3887–3892.

[24] S. Ba and V. R. Joseph, "Composite gaussian process models for emulating expensive functions," *The Annals of Applied Statistics*, vol. 6, no. 4, pp. 1838–1860, 2012.

[25] C. Li, T. Zhang, and D. I. Goldman, "A terradynamics of legged locomotion on granular media," *Science*, vol. 339, no. 6126, pp. 1408–1412, 2013.