Adaptive Online Learning for Human-Robot Teaming in Dynamic Environments

Alexander D. Wissner-Gross
Gemedy, Inc. &
Institute for Applied Computational Science,
Harvard University
Cambridge, MA, USA
alexwg@post.harvard.edu

Noah Weston
U.S. Army Research Laboratory,
Aberdeen Proving Ground, MD, USA
noah.d.weston.civ@mail.mil

Manuel Vindiola
U.S. Army Research Laboratory
Aberdeen Proving Ground, MD, USA
manuel.m.vindiola.civ@mail.mil

***Abstract*—Robotic and vehicular autonomy in contested, dynamic environments has historically been limited to teleoperation and simple programmed behaviors due to the low survivability of available AI and machine-learning techniques in the face of novel situations. Here we report that recent few-shot machine-learning models trained using interactive, human-centered, vehicular simulations can enable collaborative learning that is both adaptive (dynamically recognizing unfamiliar environmental conditions) and online (learning at each time step). Specifically, we show that our human-machine teaming approach enables simulated vehicles to anticipate novel adversities imposed in real time, both externally by their terrain and internally by their own mechanics, using only images captured by their front-facing cameras. We conclude by discussing the implications of our work for enhancing the future survivability of human-robot teams in large-scale, cluttered, contested environments.**

Keywords—autonomy, survivability, few-shot learning, online learning, adaptive learning, machine learning, deep learning, autonomous vehicles, human-machine teaming

I. Introduction

Robotic and vehicular autonomy in contested, dynamic environments has historically been limited to teleoperation and simple programmed behaviors due to the low survivability of available AI and machine-learning techniques in the face of novel situations [1]. Ideally, robots and vehicles in such environments would be capable of Adaptive Online Learning (AOL) of environmental dynamics given limited prior observations, in that they would both:

- dynamically recognize unfamiliar environmental conditions (i.e., perform adaptive learning); and
- learn at each time step (i.e., perform online learning).

From an alternative perspective, the goal of AOL can be framed as learning environmental dynamics in a manner that is efficient in both:

- space, via a batch size of 1; and
- time, via one- or few-shot learning.

Here we study the problem of AOL for the application of enabling simulated vehicles to anticipate novel adversities imposed in real time, both externally by their terrain and internally by their own mechanics, using only images captured by their front-facing cameras. Specifically, we attempt to predict in an adaptive and online manner whether simulated vehicle wheels will slip on terrain at high initial speeds during late forward acceleration given prior images of terrain observed from the vehicle at low initial speeds during early forward acceleration.

Prior approaches for similar problems have included applying support vector machines to vector embeddings computed from the average of small image patches represented in a sparse basis computed from a set of exemplary images (as in, e.g., D4L [2]) and performing nonlinear regression on local receptive fields (as in, e.g., [3]). Such sparse dictionary-learning approaches to image classification have, however, generally been outperformed by more recent deep-learning techniques in their ability to encode higher-level features [4]. Unsurprisingly, therefore, deep learning has also been applied to the problem of predicting terrain friction from front-facing vehicular imagery [5], although not yet in an adaptive and online context, to our knowledge.

Moreover, to achieve the "best of both worlds" between nonparametric models that tend to adapt quickly but generalize poorly and parametric models like convolutional networks that tend to adapt slowly but generalize well [6], it is helpful to decompose the AOL problem statement into two generalized steps:

- an "online" step, in which dimensionality reduction of observations into an embedded space is performed; and
- an "adaptive" step, which provides a method for comparing new observations to old ones.

Under previous approaches, both of those steps have been hand-engineered, resulting in theoretically suboptimal performance since it is unlikely that either step is a local optimum. A potentially improved technique is the Matching Networks approach [6], which we adopt here, under which end-to-end differentiable learning is performed across both steps to ensure local optimality of the full solution.

II. Theory

Under the Matching Networks model for few-shot learning as applied to our problem, the task of predicting terrain classifications $\hat{y}$ is posed as an end-to-end learning problem in which the embeddings of k prior support images x_i and a novel image $\hat{x}$ are learned at the same time as a differentiable neural

978-1-7281-4732-1/19/$31.00 ©2019 IEEE

attention function $a(\hat{x}, x_i)$ applied to the image labels y_i from the support set:

$$\hat{y} = \sum_{i=1}^{k} a(\hat{x}, x_i) y_i, \quad (1)$$

where the attention function a is defined as a softmax over cosine distances c between separate neural embeddings $f_\theta(\hat{x})$ of the novel image and $g_\theta(x_i)$ of the support images parameterized by θ,

$$a(\hat{x}, x_i) = e^{c(f(\hat{x}), g(x_i))} / \sum_{j=1}^{k} e^{c(f(\hat{x}), g(x_j))}, \quad (2)$$

and the training objective is to minimize the error in predicting the labels for a batch set B conditioned on a support set S, where both sets are sampled from a label set L:

$$\theta = \arg\max_\theta E_{S\sim L, B\sim L}\left[\sum_{(x,y)\in B} \log P_\theta(y|x, S)\right]. \quad (3)$$

Since the support set of images and labels can be chosen to consist of only one or a few examples, this model can be viewed as a few-shot, differentiable analogue to k-nearest-neighbors estimation, as well as a form of meta-learning.

III. Experiment

A. Simulation Environment

To evaluate the performance of our approach, we used the Autonomous Navigation Virtual Environment Laboratory (ANVEL) version 3.5 environment [7] to simulate a vehicle driving through varying terrain conditions. For this initial investigation, we simulated a “Generic 4x4” vehicle, with an "API Camera" sensor positioned on its front fender facing forward and down toward upcoming terrain, as shown in Fig. 1. For the environment, we simulated daylight driving conditions. For the terrain, we simulated a flat concrete surface covered with irregular stripes of ice, as shown in Fig. 2. For driving control, we set the vehicle to constant half (50%) throttle, with straight (0%) steering and no (0%) braking, and an overall mechanical simulation timestep of 10 ms.

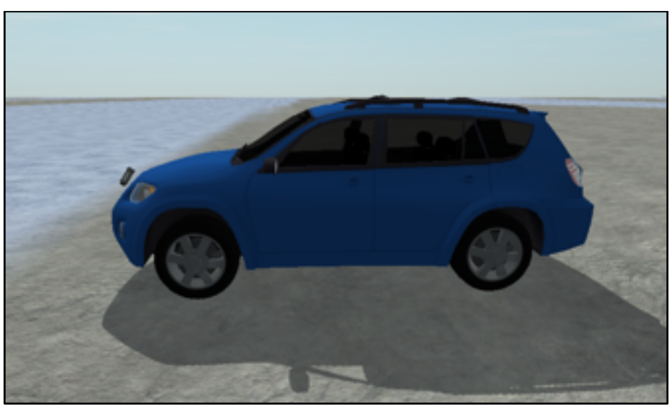

Fig. 1. Simulated vehicle with front camera pointing downward, approaching a terrain transition.

In this simulated environment, every 100 ms, we captured both a camera image and the mean of the left-front and right-front wheel SAE slip ratios [8]. Simulated adversities were imposed both externally to the vehicle, via varying terrain materials with different appearances and amounts of slipperiness, and internally to vehicle, via mechanical interactions and resonances between individual vehicle components, including each of its four wheels, its suspension system, and its carriage suspension.

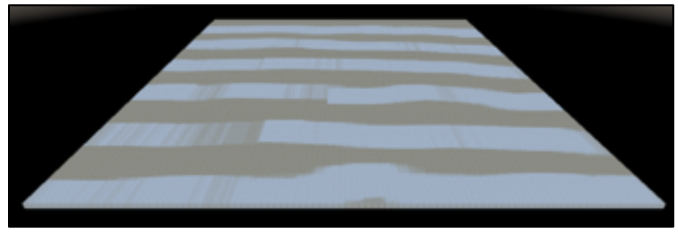

Fig. 2. Simulated terrain in ANVEL consisting of irregular ice strips (light blue) overlaid on concrete (gray).

B. Dataset Generation

We ran the simulation for 18 simulated seconds and grouped the captured terrain images into “high-slip” and “low-slip” classes corresponding to measured slip ratios above and below 0.28, respectively. We then partitioned the overall data into training (background) and test (evaluation) sets corresponding to respective simulation times during and after the first 10 s, as shown in Fig. 3, since our initial focus was to study adaptation from early-acceleration conditions to late-acceleration conditions. Future studies may explore other forms of adaption, including to novel terrain materials. We then rescaled the images from 640×480 to 84×84 pixels and converted them to grayscale.

Manual examination of the collected images suggested the importance of learning representations for high-level global features, and not simply local patch-based textures. For example, see Fig. 4, in which we show two captured images from nearly symmetric high-to-low-slip and low-to-high-slip transitions, in which the fact that the vehicle has nearly completed passage over an ice stripe (ice appears at bottom of image instead of top) provides a high-level contextual clue for momentum-enhanced slippage that might be missed with a local analysis approach.

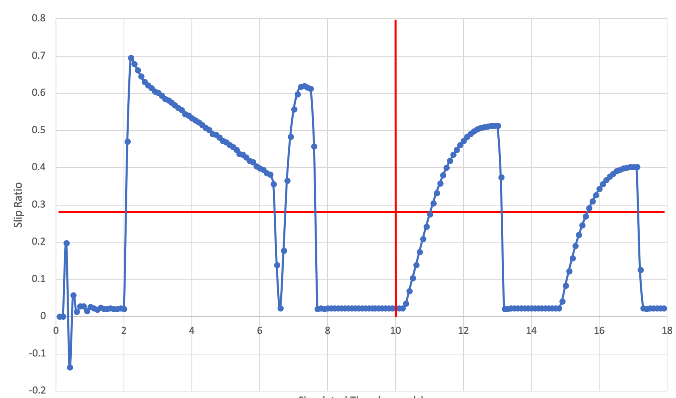

Fig. 3. Measured slip ratios over time as simulated vehicle traversed ice stripes. The data are divided into an early-acceleration training set (left of red vertical line) and late-acceleration test set (right of red vertical line), and further subdivided into low-slip (below red horizontal line) and high-slip (above red horizontal line) classes.

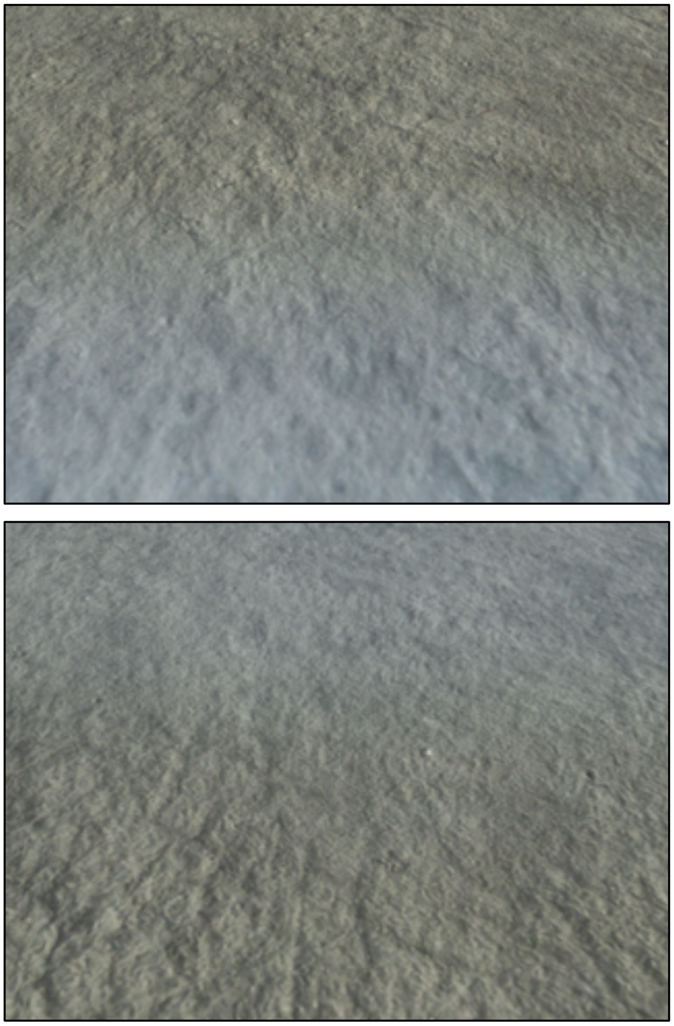

Fig. 4. Examples of images collected by simulated camera in which global context is required to distinguish between high-slip (top) conditions at the completion of driving over an ice stripe and low-slip (bottom) conditions at the completion of driving over a cement stripe.

C. Training and Model Architecture

We evaluated multiple values of the number N_{train} of support samples per class for training, the number N_{test} of support samples per class for validation, the number Q_{train} of query samples per class for training, and the number Q_{test} of query samples per class for validation: $(N_{\text{train}}, N_{\text{test}}, Q_{\text{train}}, Q_{\text{test}}) \in \{(1,1,1,1), (2,2,2,2), (4,4,4,4)\}$. For our model architecture, we followed the single-channel Omniglot design from [6], with neural encoders $f = g$ implemented as a stack of 4 modules, each consisting of a 3×3 convolution with 64 filters, batch normalization, ReLU activations, and 2×2 max pooling with a stride of 2 [9].

IV. Results

We observed improved categorical prediction accuracy ("high-slip" versus "low-slip" class) of 85% using Matching Networks on our dataset compared to the 72.5% reported for the D4L technique [2] on its robotically generated IRA dataset after a comparable 1,000, one-shot iterations (see Fig. 5). Our model's accuracy improved to more than 97% after 10,000 one-shot iterations, although some of this performance may be attributable to overfitting on a limited dataset. Benchmarking categorical-prediction performance against image datasets not collected by real or simulated vehicles, such as the popular Brodatz texture database [10], is reserved for future investigations.

V. Conclusions

Here we demonstrated the application of few-shot machine learning to human-centered vehicular training data that we synthesized using a high-fidelity simulation, but which could alternatively have been gathered from actual human driving experiments. Our approach enabled simulated vehicles to predict novel adversities imposed in real time, both externally by their terrain and internally by their own mechanics, using only images captured by their front-facing cameras.

We anticipate that, in the near future, humans and robots will need to be able to operate as teams in large-scale, cluttered, and contested environments. Our results suggest a path to improving the survivability of those teams, namely through adaptive online learning of the dynamics of those environments from the perspective of humans by nearby robots and vehicles, whose predictions can then be fed to back to humans in real time from improved decision making.

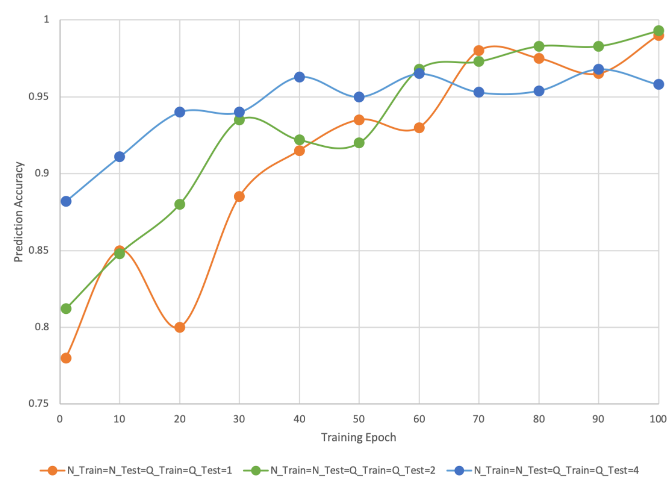

Fig. 5. Training curves of our model with various hyperparameter choices.

References

[1] A. Kott, "Challenges and characteristics of intelligent autonomy for Internet of Battle Things in highly adversarial environments," AAAI Spring Symposium Series, pp. 145-151, 2018.

[2] A. Koppel, G. Warnell, E. Stump, and A. Ribeiro, "D4L: Decentralized Dynamic Discriminative Dictionary Learning," IEEE Transactions on Signal and Information Processing over Networks, vol. 3, pp. 728-743, December 2017.

[3] A. Angelova, L. Matthies, D. Helmick, P. Perona, "Learning and prediction of slip from visual information," Journal of Field Robotics, vol. 24, pp. 205-231, 2007.

[4] S. Tariyal, A. Majumdar, R. Singh, M. Vatsa, "Deep dictionary learning," IEEE Access, vol. 4, pp. 2169-3536, 9 December 2016.

[5] S. Roychowdhury, M. Zhao, A. Wallin, N. Ohlsson, M. Jonasson, "Machine learning models for road surface and friction estimation using

front-camera images," International Joint Conference on Neural Networks, pp. 1-8, 2018.

[6] O. Vinyals, C. Blundell, T. Lillicrap, K. Kavukcuoglu, and D. Wierstra, "Matching networks for one shot learning," Advances in Neural Information Processing Systems, vol. 29, pp. 3630-3638, 2016.

[7] P. J. Durst, C. T. Goodin, C. L. Summins, B. Q. Gates, G. B. McKinley, T. R. George, "A real-time, interactive simulation environment for unmanned ground vehicles: the Autonomous Navigation Virtual Environment Laboratory (ANVEL)," Fifth International Conference on Information and Computing Science, pp. 7-10, July 2012.

[8] N. D. Smith, "Understanding parameters influencing tire modeling," Department of Mechanical Engineering, Colorado State University, Formula SAE Platform, 2004.

[9] O. Knagg, "Advances in few-shot learning: reproducing results in PyTorch," Towards Data Science, 2018.

[10] P. Brodatz, "Textures: A Photographic Album for Artists and Designers," Dover, 1996.