Doctoral thesis/dissertation Digest Form

Thesis/dissertation Title
Domain-Adaptive Robust Training and Deployment in Real-World Noisy Data-Driven Robotics Applications

Student’s Name
Wendyam Eric Lionel Ilboudo

Approved Digest
Chapter 1: Introduction
This study focuses on enhancing the robustness of data-driven robotics against noisy data and distribution shifts across the training, evaluation, and deployment phases. For the training phase, adaptive stochastic gradient descent algorithms were developed to mitigate the impact of outliers, enabling effective learning from datasets with varying complexities. In the deployment phase, a novel reinforcement learning framework was introduced, yielding three algorithms that equip policies with domain uncertainty awareness. This framework, inspired by multi-objective reinforcement learning, facilitates efficient transitions from simulated environments to real-world applications. The work aims to make robots more adaptable, efficient, and versatile, addressing challenges in real-world applications.

Chapter 2:
In this chapter, we derive and evaluate the adaptive robust exponential moving average (EMA)-based SGD algorithms, starting with the At-momentum which is then integrated to the Adam optimizer as At-Adam for the evaluations. Specifically, the At-momentum is derived from the t-momentum algorithm by making use of a modified version of the Direct incremental degrees of freedom estimation algorithm proposed by Aeschliman et al. using properties of the Student’s t-distribution as an elliptical distribution. In order to make use of this algorithm in the At-momentum, we make modifications that reduce the computational cost and makes it more suited for the stochastic gradient descent algorithm. Following the At-momentum, we then derive the AdaTerm algorithm which provides a more unified approach to the Student’s

t-distribution approach to robustness. Indeed, the AdaTerm algorithm is fully derived from the maximum log-likelihood approach to estimating the parameters of the Student’s distribution, giving it better stability guarantees and better robustness adaptiveness.

After the mathematical derivations, we then evaluate the At-momentum and AdaTerm algorithms on typical machine learning tasks, ranging from behavioral cloning and policy distillation to classification and regression tasks and reinforcement learning tasks. These experiments showed that the At- momentum integrated to the Adam optimizer, i.e. At-Adam, and the AdaTerm optimizer both interpolate between or prevail over the performance of non-robust optimizers (such as Adam) and optimizers with fixed robust (i.e., unable to automatically change their robustness in response to the training data) when encountering practical and noisy problems.

Chapter 3:

In this chapter, we tackle the Sim-to-Real transfer problem of the deployment phase by putting forward a new reinforcement learning (RL) framework called Pseudo Multi- Objective Markov Decision Process (PMOMDP) which is based on the similitudes between the multi-objective problem and the domain randomization problem. Using this framework, we derive three main RL algorithms for solving the full Convex Coverage Set (CCS), which defines the set of optimal uncertainty aware policies under the domain randomization setting. Specifically, these three algorithms use the concept of Universal Policies (UP) in order to learn a set of policies conditioned on the system identification uncertainty. Thanks to this, these policies can interpolate their behavior from a conservative policy when the uncertainty is large to a more optimal policy when the target system has been properly identified.

After the mathematical derivations of the algorithms, they are integrated with the soft-actor-critic (SAC) algorithm and then evaluated on four Mujoco environments in order to measure the Sim-to-Sim performance under various simulation parameters and uncertainty levels. Eventually, the study combines uncertainty-aware policies with a probabilistic system identification strategy and shows how the performance improves as the uncertainty decreases in both simulation and real-world D’Claw robot evaluations.

Chapter 4:

In this chapter, we discuss the proposed algorithms, their limitations, strengths and possible improvements. Specifically, we draw attention on the disadvantages of using adaptive robust optimizers with respect to problems where the quality of the dataset is already known in advance, but emphasize the fact that necessity of adaptive algorithms lies in not being able to have such knowledges in typical machine learning applications. We also discuss the scalability of the algorithms proposed for learning uncertainty-aware policies and their generalizability to other type of RL tasks beside continuous states and actions tasks, with continuous and dense reward functions. We then mention various directions for improving the proposals and a specific extension for dealing with environments with noisy system processes.

Summary:

The fusion of robotics and data-driven learning sparks a profound shift in the functionality of robots in real-world scenarios. This transition introduces challenges related to adapting robots to learn from imperfect data and excel in diverse environments.

This study concentrates on enhancing the robustness of data-driven robotics against noisy or corrupted data and shifts in data distribution, crucial challenges across the training, evaluation, and deployment phases. In the training phase, where data quality profoundly influences learning efficacy, we introduce adaptive stochastic gradient descent algorithms. These algorithms dynamically adjust their robustness to varying complexities in datasets, enabling robots to learn effectively from diverse and imperfect real-world data.

In the deployment phase, data-driven controllers, particularly from reinforcement learning, often face distribution shifts in novel environments. We present a novel reinforcement learning framework, yielding three algorithms that imbue policies with domain uncertainty awareness. Drawing an analogy between multi-objective RL and multi-domain RL from domain randomization, our framework generates adaptable policies for different environments and system uncertainties, facilitating seamless transitions from simulation to real-world applications.

In essence, our work aims to elevate robots' adaptability by improving their ability to learn from real-world data and navigate varying environmental conditions. This advancement contributes to making robots more versatile and efficient in practical applications. Future investigations could extend uncertainty-awareness to other areas beyond RL, explore applications of developed algorithms in universal model training, and assess adaptive robust gradient descent methods in per-data-point gradient computations.