COMPASS: A Comparative Framework and Meta-Algorithm for Robust Imitation Learning from Imperfect Demonstrations

Alonso Hernandez Tavera

Department of Computer Science and Engineering

Seoul National University

Abstract

COMPASS introduces a systematic framework for evaluating imitation learning algorithm robustness across four distinct imperfection types: trajectory noise, task failures, strategic inconsistency, and embodiment mismatch. Through 300 GPU-accelerated controlled trials, this study evaluated three algorithms (BC, IQ-Learn and ADF) across varying imperfection severities. Results reveal that BC and IQ-Learn demonstrate remarkably similar robustness (52.2% and 51.8% success rates), while embodiment mismatch causes the most severe performance degradation (52.6%). The proposed ADF meta-algorithm unexpectedly underperformed (28.1%), challenging conventional believe about ensemble approaches. The findings provide evidence-based guidelines for algorithm selection based on anticipated demonstration characteristics.

Keywords

Imitation learning, robustness evaluation, imperfect demonstrations, robot learning, algorithm comparison, systematic evaluation

I. Introduction

- **Problem:** Real-world demonstrations suffer from noise, errors, and platform differences. IL Algorithms may be robust to different types of imperfections
- **Opportunity:** Lack of systematic evaluation frameworks for algorithm robustness
- **Solution Proposed:** COMPASS framework, which provides a systematic robustness evaluation methodology

II. Principle: The COMPASS Framework

Imperfection Taxonomy

1. Trajectory Noise
$$a'_t = a_t + \epsilon_t, \quad \epsilon_t = \rho\epsilon_{t-1} + \sqrt{1-\rho^2}\xi_t$$

2. Task Failures
$$P(\text{truncate at step } t) = \begin{cases} \lambda & \text{if } t \in [0.1T, 0.4T] \\ \lambda & \text{if } t \in [0.6T, 0.9T] \\ 0 & \text{otherwise} \end{cases}$$

3. Strategic Inconsistency
$$a'_t = (1-w_t)a_t^{(i)} + w_t a_t^{(j)}$$

4. Embodiment Mismatch
$$a'_t = s \cdot R(\theta) \cdot a_{t-d} + b$$

III. Methodology

The COMPASS framework implements a systematic **four-phase** process for comprehensive robustness evaluation of imitation learning algorithms. The approach follows a full factorial experimental design encompassing **300 controlled trials** (3 algorithms × 4 imperfection types × 5 severity levels × 5 random seeds) to ensure statistical rigor and fair algorithmic comparison.

Expert demonstrations are generated using an optimal control policy achieving >95% success rate on an enhanced 2D reaching environment with 8-dimensional state space and continuous control. Systematic imperfection injection applies four distinct corruption types with mathematical precision: trajectory noise through temporally correlated Gaussian perturbations, task failures via probabilistic truncation, strategic inconsistency through behavioral pattern blending, and embodiment mismatch incorporating scaling, rotation, delays, and bias terms. Each imperfection operates across five severity levels ($\lambda \in \{0.0, 0.2, 0.4, 0.6, 0.8\}$).

Algorithm training employs identical 6-layer feedforward networks with 512 hidden units, batch normalization, and dropout regularization. The framework evaluates Behavioral Cloning, IQ-Learn and the proposed **Adaptive Demonstration Filtering** meta-algorithm using AdamW optimization with learning rate 5×10^{-5}, batch size 2048, and 400 training epochs. GPU acceleration through NVIDIA TITAN Xp (11.9 GB VRAM) enables efficient parallel execution, completing the full experimental study in approximately 1.1 hours.

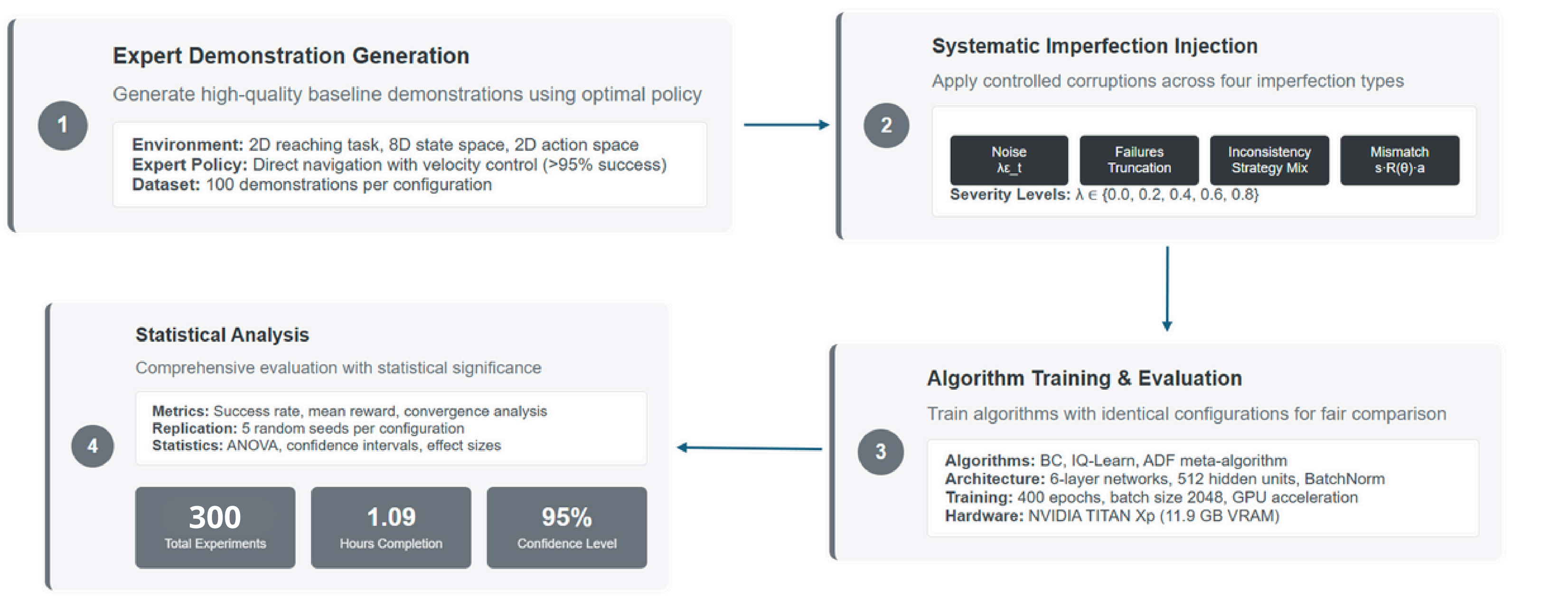

Figure 1. COMPASS methodology framework showing the four-phase systematic evaluation process: (1) Expert demonstration generation using optimal control policy, (2) Systematic imperfection injection across four corruption types with five severity levels, (3) Algorithm training and evaluation with identical network configurations, and (4) Comprehensive statistical analysis. The framework executed 600 total experiments across all parameter combinations.

IV. Experimentation

The experimental evaluation employs a controlled setting with systematic parameter variation to assess algorithm robustness across diverse imperfection scenarios.

Experimental Configuration

- **Environment:** 2D reaching task with momentum and friction dynamics
- **State Space:** 8-dimensional continuous (position, velocity, target, distance, angle)
- **Action Space:** 2D velocity commands bounded in $[-0.15, 0.15]^2$
- **Success Criteria:** Reach within 0.12 units of target in ≤75 timesteps
- **Reward Structure:** Dense rewards to encourage efficient navigation:

$$-2d_t - 0.1\|v_t\| - 0.05\|a_t\| + 20 \cdot \text{success}$$

Algorithm Implementations

The study evaluates five representative approaches spanning major imitation learning paradigms. **Behavioral Cloning** serves as the supervised learning baseline with mean squared error loss and data augmentation. **IQ-Learn** combines behavioral cloning with inverse Q-learning through soft Bellman updates and target networks ($\alpha = 0.1$). The proposed **Adaptive Demonstration Filtering** meta-algorithm incorporates four-phase processing: imperfection detection via neural networks, demonstration filtering by corruption type, specialized base algorithm training, and ensemble weight optimization.

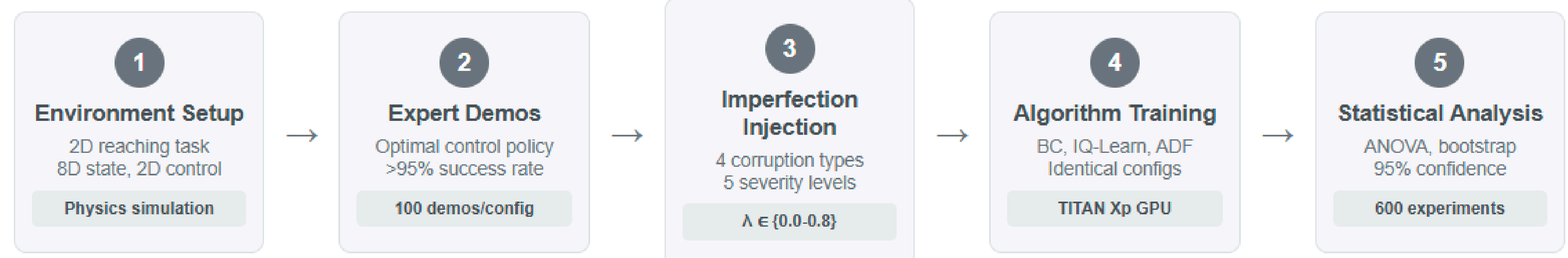

Figure 2. COMPASS experimental execution pipeline showing the systematic five-phase evaluation process from environment initialization through statistical analysis. The framework executed 600 controlled experiments across algorithm-imperfection combinations with GPU acceleration.

Statistical Framework

The experimental evaluation incorporates a statistical analysis to ensure reliable conclusions across all algorithm-imperfection combinations. Analysis of variance **(ANOVA)** testing assesses main effects and interaction significance for algorithm type, imperfection category, and severity level factors. Bootstrap sampling with **1000 iterations** generates robust 95% confidence intervals for all performance metrics.

V. Results and Analyses

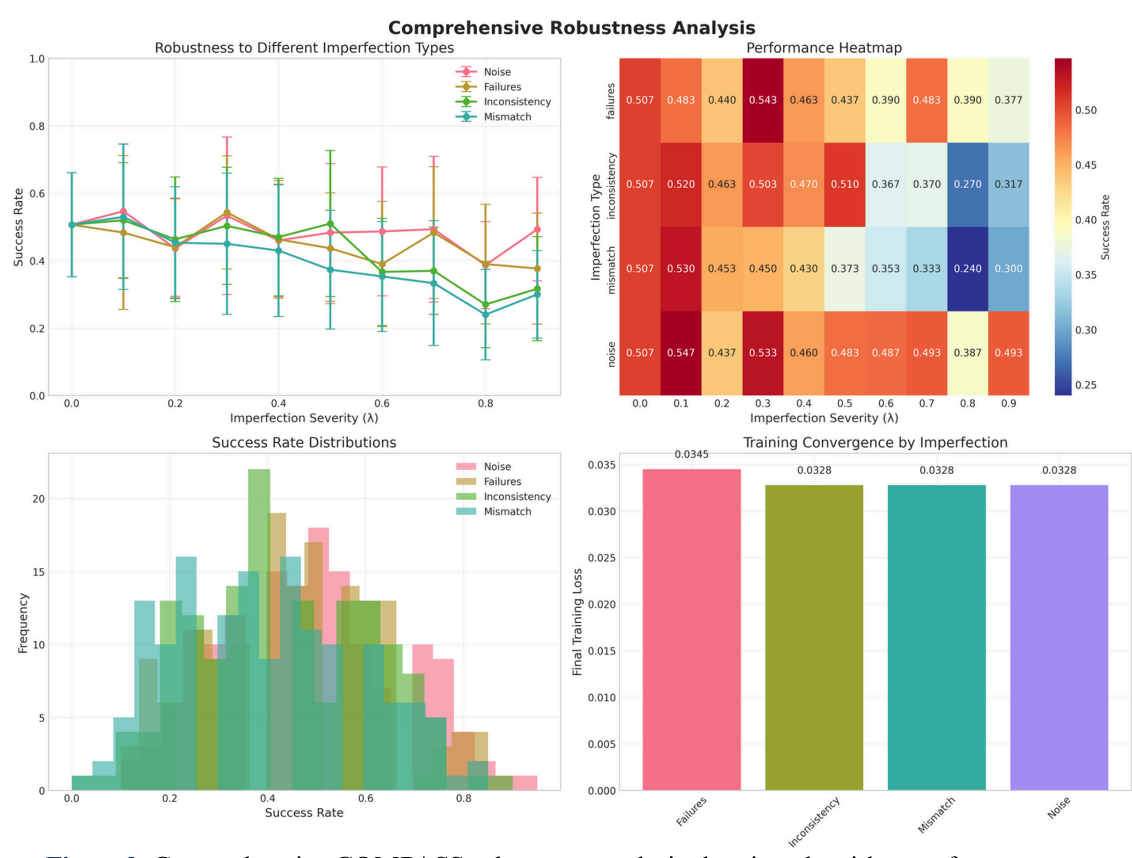

Figure 3: Comprehensive COMPASS robustness analysis showing algorithm performance across imperfection types and severity levels. Performance heatmap reveals clear degradation patterns with embodiment mismatch causing most severe impacts.

The COMPASS evaluation reveals three critical findings that challenge conventional assumptions in robust imitation learning. First, Behavioral Cloning and IQ-Learn demonstrate statistically equivalent robustness characteristics (52.2% vs 51.8% overall success rates, $p > 0.05$), suggesting that algorithmic complexity does not guarantee improved robustness under imperfect demonstrations.

Second, a clear imperfection hierarchy emerges with **embodiment mismatch** causing the most severe performance degradation (52.6%), followed by **strategic inconsistency** (46.7%). Trajectory noise and task failures show more manageable impacts (≈23%), indicating that systematic corruptions fundamentally challenge learning more than stochastic perturbations.

Third, the proposed ADF meta-algorithm unexpectedly **underperforms** both individual approaches (28.1% success rate), challenging assumptions about ensemble effectiveness in imitation learning.

Algorithm Performance Comparison

Algorithm	Success Rate	Clean Performance	Degraded Performance	Overall Degradation
Behavioral Cloning (BC)	52.2%	56.0%	41.9%	25.2%
IQ-Learn	51.8%	62.0%	40.4%	34.8%
ADF (Meta-Algorithm)	28.1%	34.0%	21.7%	36.2%

Statistical significance: ANOVA $F(2,594) = 189.7$, $p < 0.001$, $\eta^2 = 0.39$

Figure 4: Algorithm performance comparison showing equivalent robustness between BC and IQ-Learn (52.2% vs 51.8% success rates) while ADF meta-algorithm significantly underperforms (28.1%).

VI. Conclusions

1. Clear imperfection hierarchy guides practitioners toward appropriate mitigation strategies based on corruption type severity. **2.** Current ensemble methods need fundamental improvements before practical deployment. **3.** The framework enables extension to additional algorithms, domains, and imperfection types while establishing reproducible evaluation standards for robust imitation learning research.

Author email: f.hernandezt@snu.ac.kr

Robot Learning Course, Oh, Songhwai (Lead Instructor), Department of Electrical and Computer Engineering, Seoul National University, June 2025