Intuitive Physics and Physical Reasoning in Deep Learning

PhD proposal by: Rodionov Danila
Supervisor : Zagoruyko Sergey

,

1 Introduction

The field of deep learning algorithms has made significant strides in areas like image recognition, natural language processing, and reinforcement learning. However, one critical aspect where machine learning still lags behind human cognition is the ability to understand and reason about real-world physical interactions. Humans possess an innate "intuitive physics engine" that allows them to predict outcomes of physical events with remarkable accuracy, even when they lack formal training in physics. This capability enables us to navigate the world safely, manipulate objects effectively, and solve problems involving physical phenomena.

Deep learning models, despite their impressive performance on various tasks, struggle to replicate this level of physical reasoning. Current approaches often rely on hand-crafted simulations ([1], [2]), explicit physical systems and model paradigms([3], [4], [5], [6]) or human-sensitivity based paradigms like Violation of Expectation (VoE)([7]), which limit their scalability and adaptability to novel situations and bring many obstacles to the ability to generalize. Such models fail to capture the nuances of human-like intuition, which makes them difficult to apply in systems that interact with real world physics. Addressing these limitations requires developing new methodologies that combine probabilistic reasoning, representation learning, and advanced neural architectures.

2 Problem Statement

The recent interest in ensuring robustness of Machine Learning models lies in provable methods for neural networks reasoning skill. There is a wide area of research considering certified reasoning skill in text-to-text generation([8], [9], [10]) and text-to-mage[11] generation of several open-source models. In fact, they only provide methods for alignment in "common sense" and limit deep learning models to learn "intuitive physics" understanding. It leads to some obstacles in implementation of deep learning in real-world systems. Existing model approaches brought some determinism to this area and systematized it. However, it still remains underexplored because of their limitations which must be resolved.

Specifically:

- Existing models perform well on specific benchmarks but struggle to generalize to unseen scenarios.

- Many approaches require predefined object properties or physics rules, limiting flexibility.
- Most models excel at short-term predictions but falter when forecasting over extended periods.
- Probabilistic reasoning remains underutilized, leading to brittle predictions in ambiguous contexts.
- Incorporating prior knowledge about physics laws into neural architectures remains challenging.

To this end, in this research, the existing models for physical reasoning will be studied and novel methods for intuitive physical reasoning in deep learning will be derived.

3 Objectives

The main goal of this research is to propose a new framework for enhancing physical reasoning capabilities in deep learning models. Complementary objectives include:

1. Identifying classes of physical reasoning tasks where current models are vulnerable and proposing certified robustness methodologies to address these vulnerabilities([12], [13], [14]).
2. Publishing papers of all achieved tasks (Core-A*, Nature Science Index journals).
3. Releasing open-source code for public use and reproducibility.
4. Providing evaluation results for the models used in real commercial cases.

4 Work Plan

The work will be distributed over three years as follows:

1. Year 1: Literature Review and Theoretical Foundations

   Conduct a thorough review of existing literature on intuitive physics, physical reasoning, and related benchmarks. Identify suitable deep learning architectures for physical reasoning tasks. Establish theoretical foundations for certified robustness guarantees in physical reasoning models. Analysis of the weaknesses of existing models on specific benchmarks.

2. Year 2: Development and Practical Testing.

   Design and implement a framework for evaluating robustness guarantees in physical reasoning models. Test the framework on benchmarks and real-world datasets. Refine techniques based on experimental results.

3. Year 3: Optimization, Validation, and Dissemination.

   Optimize the framework for scalability and efficiency. Validate the pipeline on diverse real-world scenarios, obtained from real-world physics-included domains (robotics, autonomous systems). Additionally, write the dissertation and publication.

5 Potential Impacts

As deep learning models have been increasingly used by humans for various purposes, their reasoning ability in real-world understanding become a great concern to the public. Successful completion of this research promises numerous benefits: advancing AI capabilities by enabling machines to reason about physical interactions at a human-like level, enabling safer and more reliable interaction between humans and machines in real-world environments, providing foundational contributions to the fields of robotics, autonomous systems, and computer vision. Also it will be the catalyst for motivating future research on integrating symbolic reasoning and neural computation for complex problem-solving.

6 References

1. A. Lerer et al., "PHYRE: A New Benchmark for Physical Reasoning," NeurIPS 2019.
2. R. Ronan et al., "IntPhys 2019: A Benchmark for Visual Intuitive Physics Understanding," 2019.
3. H. Qi et al., "Learning Long-term Visual Dynamics with Region Proposal Interaction Networks," ICLR 2021.
4. J. Wu et al., "Modeling Expectation Violation in Intuitive Physics with Coarse Probabilistic Object Representations," NeurIPS 2019.
5. Eltayeb Ahmed, Anton Bakhtin, Laurens van der Maaten, Rohit Girdhar "Physical Reasoning Using Dynamics-Aware Models," 2021
6. S. Ehrhardt, A M. Niloy, J. Mitra, A. Vedaldi "Learning a physical long-term predictor," 2017
7. S. Wang et al., "X-VoE: Measuring eXplanatory Violation of Expectation in Physical Events," 2023.
8. J. Schulman, S. Levine, P. Moritz, M. I. Jordan, P. Abbeel, "Trust Region Policy Optimization, " 2015.
9. J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, "Proximal Policy Optimization Algorithms," 2017.
10. R. Rafailov, A. Sharma, E. Mitchell, S. Ermon, Ch. D. Manning, C. Finn, "Direct Preference Optimization: Your Language Model is Secretly a Reward Model, " 2023
11. I. Kajić, O. Wiles, I. Albuquerque, M. Bauer, S. Wang, J. Pont-Tuset, A. Nematzadeh "Evaluating Numerical Reasoning in Text-to-Image Models," 2024
12. Andrew Melnik et al. "Benchmarks for Physical Reasoning AI," 2023
13. M. Lin et al., "LLMPHY: COMPLEX PHYSICAL REASONING USING LARGE LANGUAGE MODELS AND WORLD MODELS," 2023.
14. Q. Garrido et al. "Intuitive physics understanding emerges from self-supervised pretraining on natural videos", 2025