Abstract

Much work on scene understanding have focused on understanding what objects are, but less work has focused on understanding how they behave. A physics engine explicitly captures the physical intuition required to reason about the scene with appropriate constraints and assumptions, but the inability to differentiate through the physics engine limits its flexibility to model new scenarios. Neural network architectures can be thought of as differentiable programs that may be able to overcome this inflexibility. As a step toward using neural models to emulate physics engines, I built a predictive model of rigid body dynamics in Newtonian worlds with a variable number of objects. I present some preliminary results on the model's capacity to scale with the number of objects in a scene and to learn intuitions of physics that can be transferred to worlds with different numbers of objects.