

1 Proposal

The end goal is to achieve good, aggressive control at the limits from high-dimensional sensor data. Initially, that high-dimensional data will be camera images; in the fullness of time, maybe that also includes LiDAR data, IMU data, etc.

What I'm proposing is model-predictive control (i.e., MPPI) from pixels. During training, when we have access to ground-truth state, we can train an environment model that maps input images to the state variables that we need; a high-level proposal for the design of a deep learning architecture is as follows:

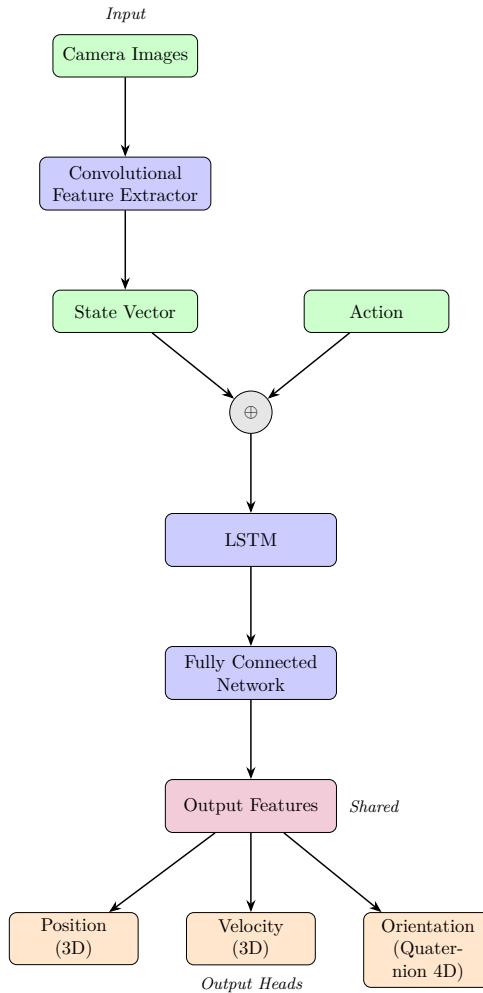


Figure 1: Proposed deep learning architecture

This is a high-level diagram, and the details will probably need to be fleshed

out as results come in. The output is the state vector necessary for model-predictive control, with which we can run the model-predictive controller.