Neural Robotics - A New Perspective
AAAI Robotics Fellowship 2016

Peter Ondruska
Mobile Robotics Group, University of Oxford

1 Introduction

This proposal describes a new paradigm for building intelligent robots based on novel usage of modern end-to-end machine learning methods as a substitution for traditional multi-stage robotics pipelines.

One of the major problems with the traditional robotics pipeline designs (Figure 1) is that they require a significant amount of hand-engineering, which imposes strong restrictions on the robot's capabilities. This is because every step of the pipeline (e.g. sensing, managing internal world representation or planning) is usually designed separately, and includes it's own simplifications, prior knowledge and assumptions which often do not hold in the real world. This results in a sub-optimal system which in many cases is unable to robustly handle the high complexity of the robot's operating environment.

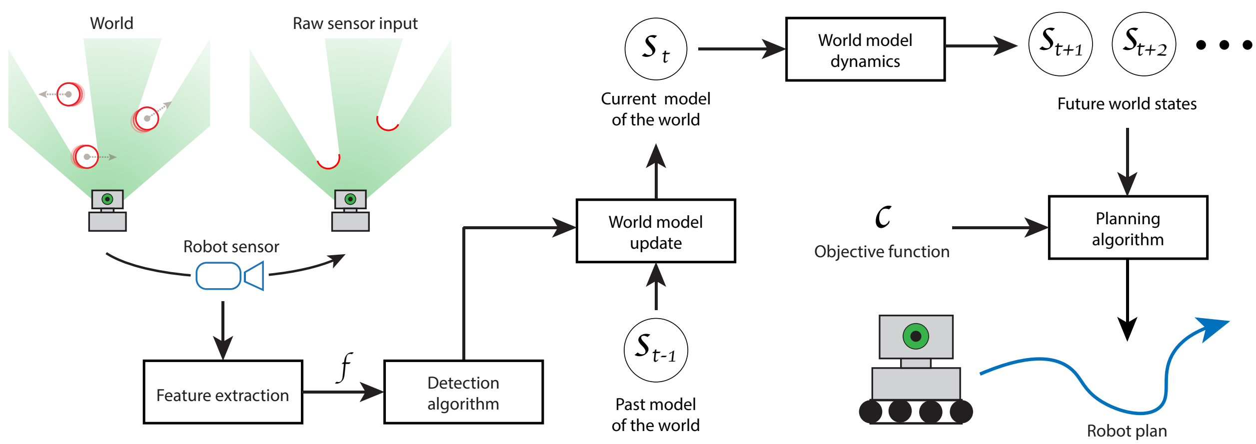

Figure 1: An example of a traditional robotics pipeline. Several hand-designed data processing stages necessary for the operation of the robot result in the sub-optimality of the whole system.

Here we propose a replacement for the traditional robotics pipelines. Our novel approach uses highly expressive machine learning methods for learning how to plan and perceive, thus avoiding the restrictions normally imposed on the internal representations of the robotic systems as described above. Our approach allows the robot to infer optimal procedures for sensing, internal world model updates and planning directly from the sensory data. The main benefit of our approach is that it reduces the amount of hand-engineering and prior knowledge injected into the pipeline, thus removing the constraints that come with them. This allows us to build more capable and powerful robots.

The key element of our solution is *DeepTracking* [Ondruska and Posner, 2016], which is a recurrent neural network architecture that is capable of learning world dynamics and predicting future world states directly from raw sensory data. We propose using this network as an oracle, which would determine feasible plans and later rank them using the objective function inferred from the human demonstrations using *Deep Inverse Reinforcement Learning* [Wulfmeier, Ondruska and Posner, 2015]. We believe that our solution can be an important milestone towards building more intelligent machines by overcoming the most daunting limitations of the current robotics designs.

2 Deep Tracking

DeepTracking [Ondruska and Posner, 2016] is the first end-to-end object tracking approach which directly maps from raw sensory input to object trajectories in the sensory space without requiring any feature engineering or system identification (e.g. plant or sensor models). Specifically, the system accepts a stream of raw sensory data at one end and, in real-time, produces an estimate of the entire state of the environment, including occluded objects, at the output (Figure 2).

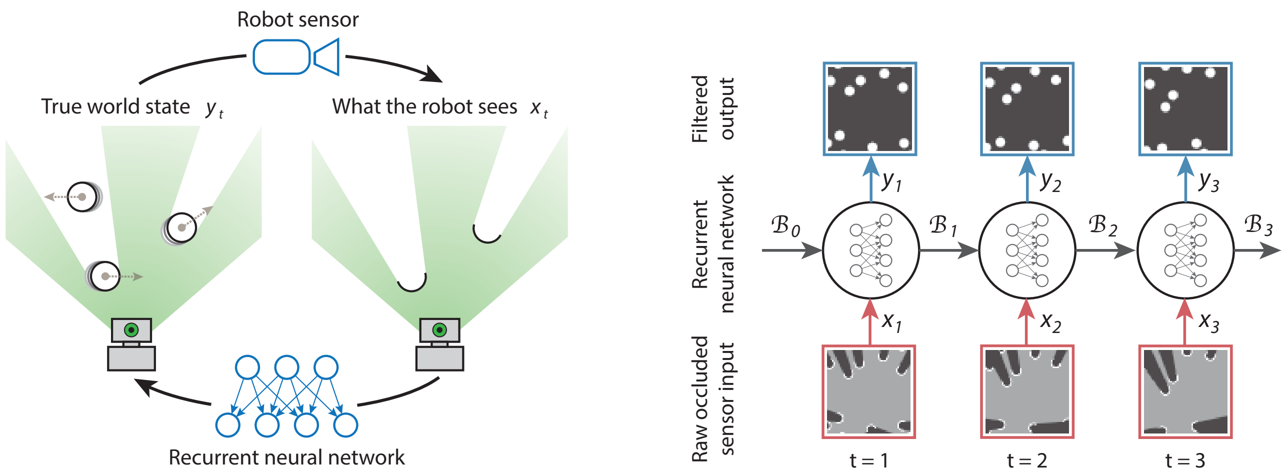

Figure 2: Robot's sensors provide only partial observations of the surrounding environment. DeepTracking leverages a recurrent neural network to effectively reveal occluded parts of a scene by learning to track objects from raw sensor data – thereby effectively reversing the sensing process.

This is achieved by framing the problem as a deep learning task and exploiting sequence models in the form of recurrent neural networks to learn a mapping from sensory measurements to object trajectories. In order to learn this mapping the network must learn to capture the dynamics of the world, before exploiting this information to simulate the occluded parts of the visual scene. Our model is able to learn the correct mapping without access to the ground-truth annotations, using an unsupervised learning method based on spatio-temporal dropout and raw occluded sensory data. This means that in order to train our model we only require the easily available raw sensory data, rather than the more expensive to attain supervised ground truth labelled data.

The success of our proposed method was demonstrated using a synthetic dataset designed to mimic a common task within robotics - tracking objects in 2D laser data: https://www.youtube.com/watch?v=pG3BBzGGgew The video demonstrates that the network learnt to track many dynamic objects despite occlusions and the presence of sensory noise. The network also learnt to predict the future states of the world at any point in time using the raw occluded sensory observations as input.

We expect that adding robot control signal as an additional input to the network will enable the model to learn the consequences of the robot's actions. In other words, the network will learn to predict the state of the world in response to the robot's actions. This step is critical to building an end-to-end machine learning-based robotics pipeline, and is discussed in the next section.

3 Neural Pipeline

The proposed pipeline builds on top of the Deep Tracking algorithm described above. Deep Tracking learns to act as an oracle that can predict the future world states. Such predictions effectively act as a world simulator capable of evaluating any possible robot plan, thus providing the robot with a simple yet effective planning paradigm.

The basic idea is to generate a large number of feasible kinematic plans for a robot to perform. This can be done through either an exhaustive generation of all possible sequences of control signals, or by using more elaborate methods. These plans would in turn be evaluated using the Deep Tracking network, whereby the network would generate a sequence of future world states $s_{t+1}, s_{t+2}, s_{t+3}, \ldots$ when following any given plan of actions. Finally, the cost of each sequence would be computed, hence capturing the feasibility and value of the plan and determining its ranking. The entire system is displayed in Figure 3.

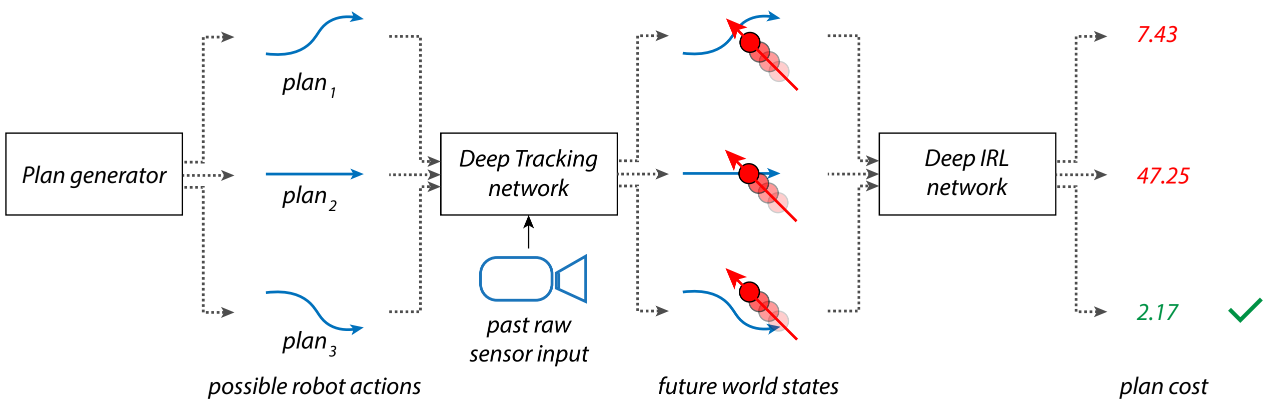

Figure 3: Proposed Neural Robotics pipeline. The Plan Generator synthesises a set of possible plans of actions. Deep Tracking Network acts as an oracle capable of predicting the future world states in response to any given plan of actions. The plan with the lowest cost as evaluated by the Deep Inverse Reinforcement Learning (IRL) network is then selected for execution.

We propose using *Deep Inverse Reinforcement Learning* (IRL) [Wulfmeier, Ondruska and Posner, 2015] for evaluating the different plans of actions as described above. Using deep IRL would avoid the need for designing such an evaluation function by hand. Instead, a separate neural network would be specifically trained to infer the desired objective function using a supervised dataset of demonstrations, where a human in the charge of the control signal.

The advantage of our proposed robotics pipeline design is that the *Deep Tracking* network can be trained using vast amounts of purely unsupervised data to learn *how the world works* and how to remove occlusions. Then the network for *Deep IRL* can exploit the outputs of the Deep Tracking framework to be trained only how to identify *which world states are desirable* without considering occlusions. This is a much simpler task, which can be learnt using relatively little amount of supervised data.

References

[1] Peter Ondruska and Ingmar Posner, *Deep Tracking: Seeing Beyond Seeing Using Recurrent Neural Networks*, in The Thirtieth AAAI Conference on Artificial Intelligence (AAAI), Phoenix, Arizona USA, 2016.

[2] Markus Wulfmeier, Peter Ondruska and Ingmar Posner: *Deep Inverse Reinforcement Learning*, NIPS workshop, 2015.