

My Vision for Robotics

Mihir Prabhudesai

April 2023

Last few years of developments in AI has shown us data is the king. Using a larger dataset for the task you want to solve gives a **guaranteed** improvement in performance. Furthermore, methods that consume this large data in an efficient manner are crucial. Specifically, deep learning based methods have shown to be very efficient in learning from vast amount of unstructured data. Now there is a lot of data for tasks such as NLP and vision, however there is not a lot of data for robotics.

This makes us question on how can we make progress in the field of robotics. One simple way is to manually collect a lot of kinesthetic robot data for a lot of tasks, however this will require a huge investment in terms of cost, and we cant afford this as it's unclear if the cost is worth the ROI. So the main research question here is , how can we reduce the cost for solving robotics?

There are many ways of reducing the cost. I think the important ones are:

- Exploiting large datasets:
- Exploring data-efficient learning methods

In the Section below i'll go into depth for each of them.

1 Exploiting large datasets

In robotics there are many different types of learning paradigm such as Learning from robot or human kinesthetic data, Learning from simulation, Learning from passive data. It's an important design choice for which one of them we should build learning algorithms?. I think the answer to this question is the setup in which the number of datapoints increases free of cost. At this point in time, a very

clear answer to this is passive data. As humans share videos, with the intent of communication or entertainment, we end up getting a lot of data of them doing different actions to solve tasks. However currently for kinesthetic data or simulators, we need to explicitly invest money to get such data, thus making them bad prospects for learning methods. Note that this **might** change in the future, ones Meta or some other company find a way to sell Virtual reality headsets to consumers by making them more compelling. Until then i believe we should build learning methods that can exploit passive videos.

A critical question still remains on how do we transfer relevant information from passive videos to robot policies. This involves design choices at various levels, in the section below i'll try to give my thoughts on how can we make current learning algorithms more efficient so that they can better exploit data.

2 Exploring data-efficient learning methods

Lesser the inductive biases the better. If this were entirely true then currently we would be using a large MLP over all the image pixels. However this is untrue, what is true is that we need to find the right set of inductive biases that scale with data. I think there are two methods that can help us get more from less data.

2.1 Inducing structure during learning

It is well-known in the literature that when we induce structure during the learning process, it reduces the amount of data required. However there are a few caveats such as inducing the wrong structure would definitely make the learning worse. Thus one needs to be careful while inducing structure.

2.2 Trading off computation with data

When Humans do not have access to data, they seem to trade off data with computation. For example, when encountered a difficult problem, a human reasons to solve the problem. This ability to reason and think deeply about a problem is missing in current deep neural networks.