Model-agnostic Measure of Generalization Difficulty

Akhilan Boopathy, Kevin Liu, Jaedong Hwang,
Shu Ge, Asaad Mohammedsaleh, Ila Fiete

MCGOVERN INSTITUTE

CENTER FOR Brains Minds+ Machines

Benchmarks drive new machine learning architectures

ImageNet

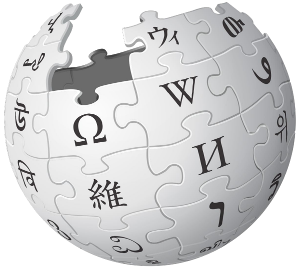
Atari

Large Language Corpora

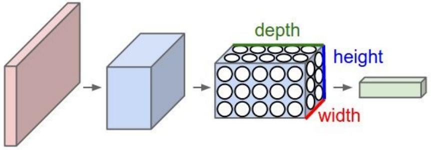
Benchmarks drive new machine learning architectures

ImageNet

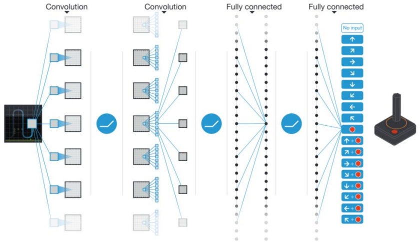
Atari



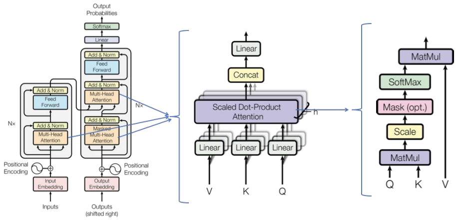
Large Language Corpora



Large-scale Convolutional Neural Networks



Deep Reinforcement Learning



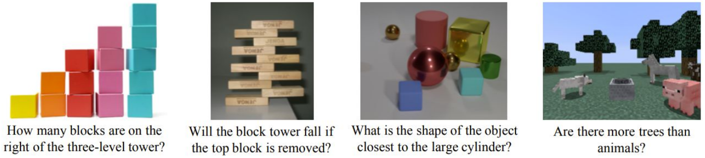
Large-scale Transformers

What are good tasks/benchmarks?

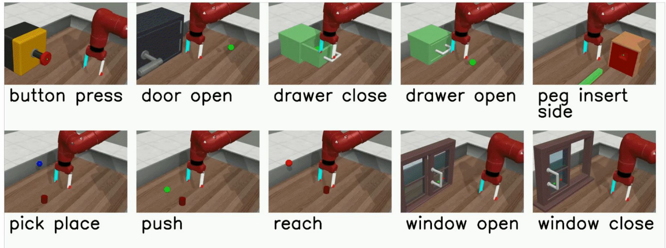
ImageNet

Atari

Large Language Corpora



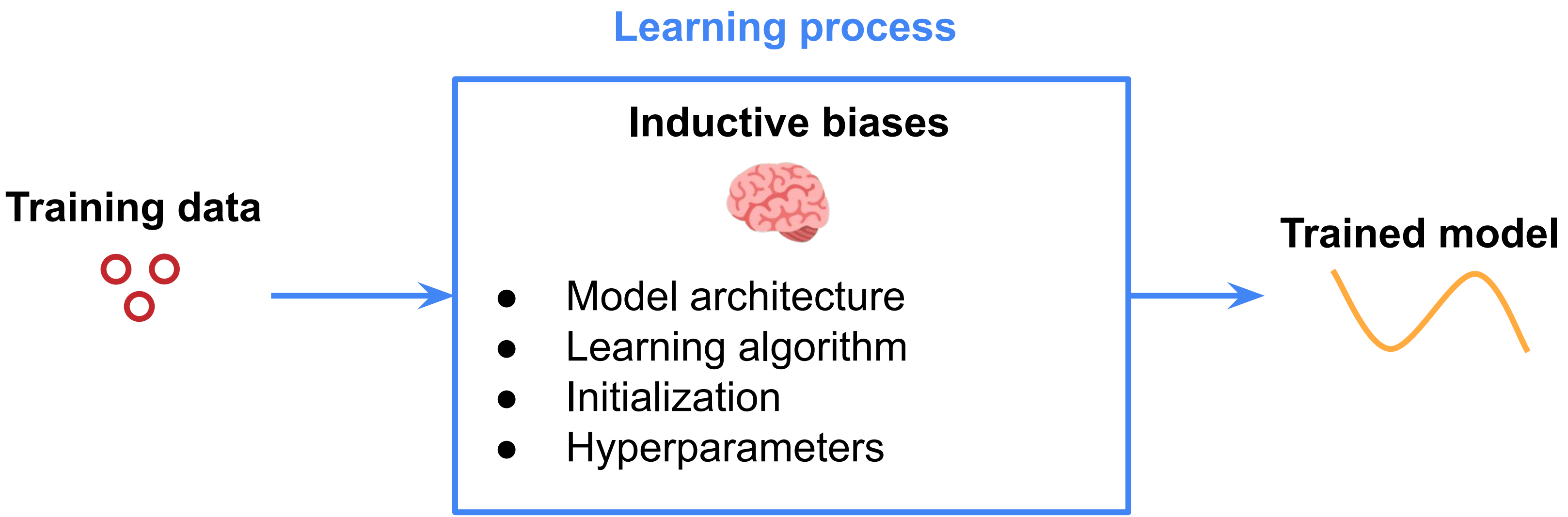
CLEVR



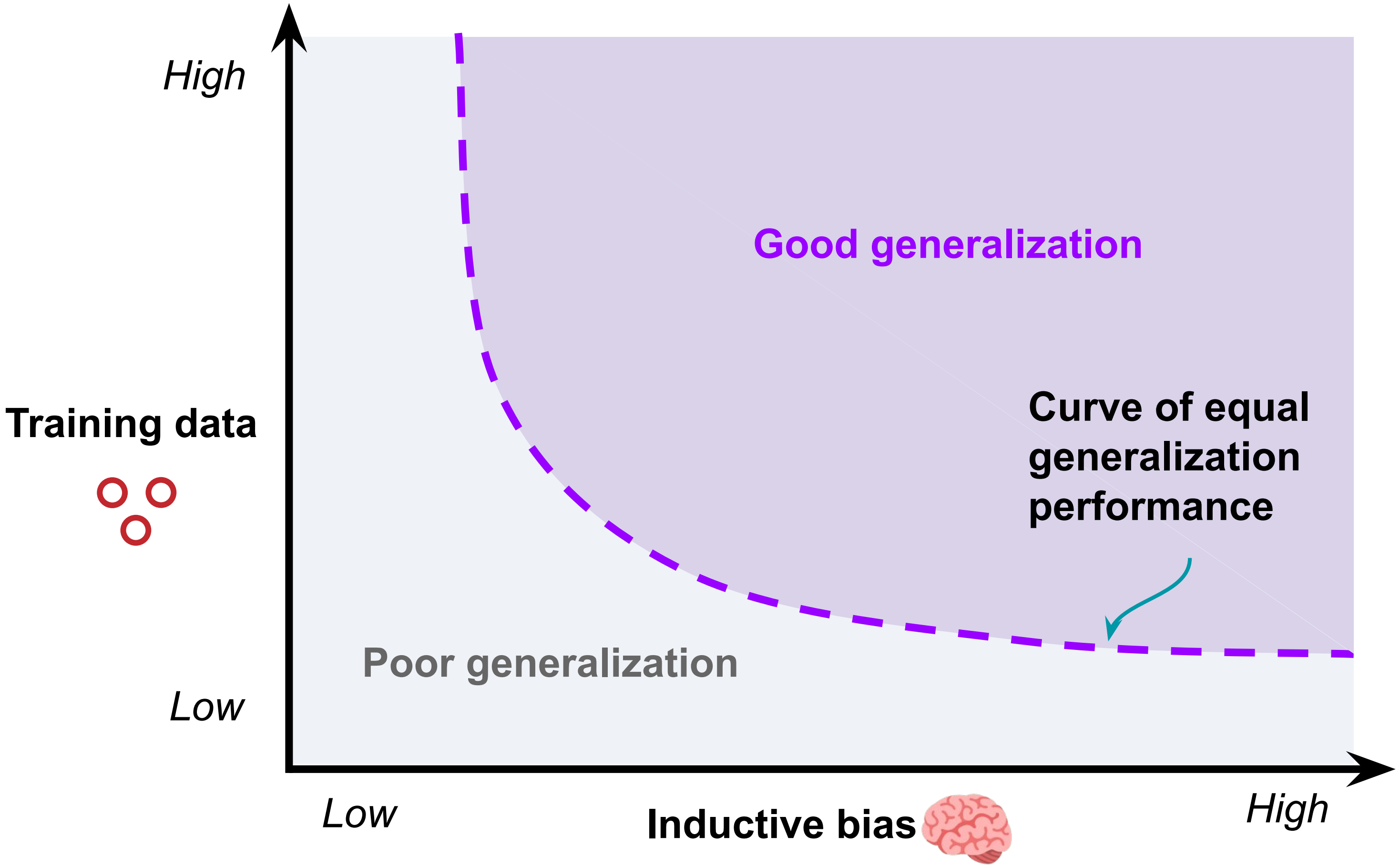
Meta-world

- How can we evaluate the difficulty of these benchmarks? Which ones will encourage the development of more generalizable inductive biases?

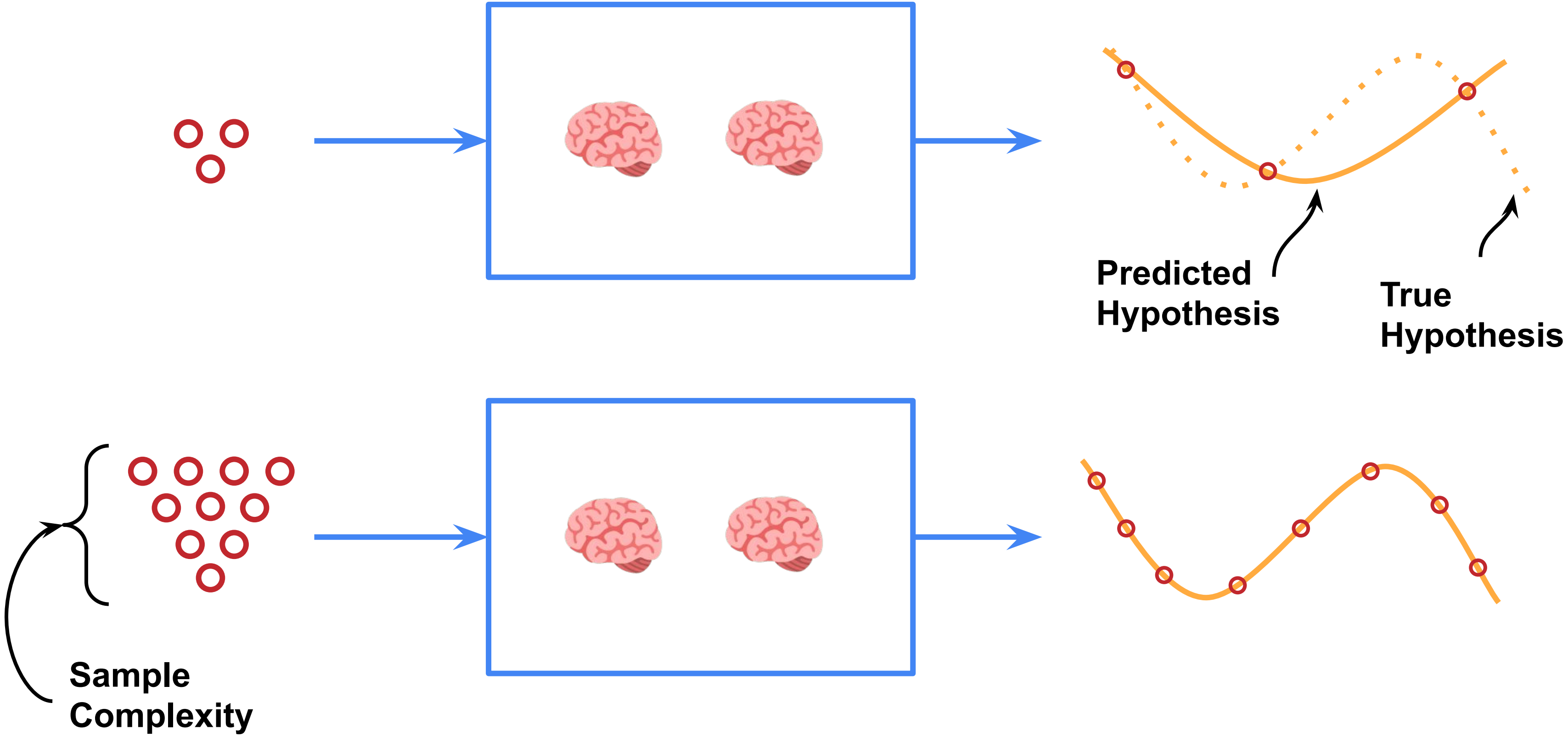
Generalizing on a task requires both training data and inductive biases



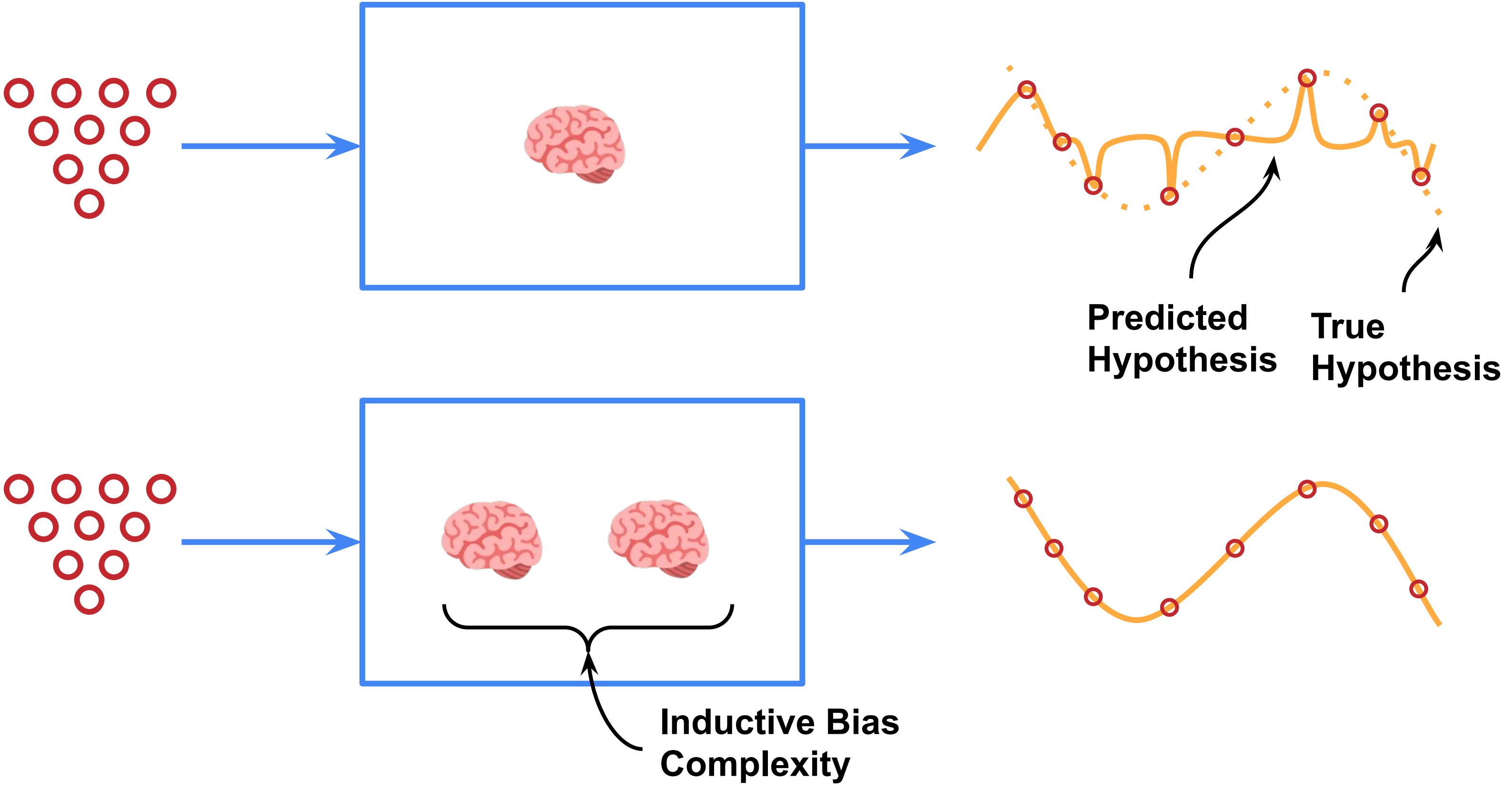
Training data can be traded-off with inductive bias



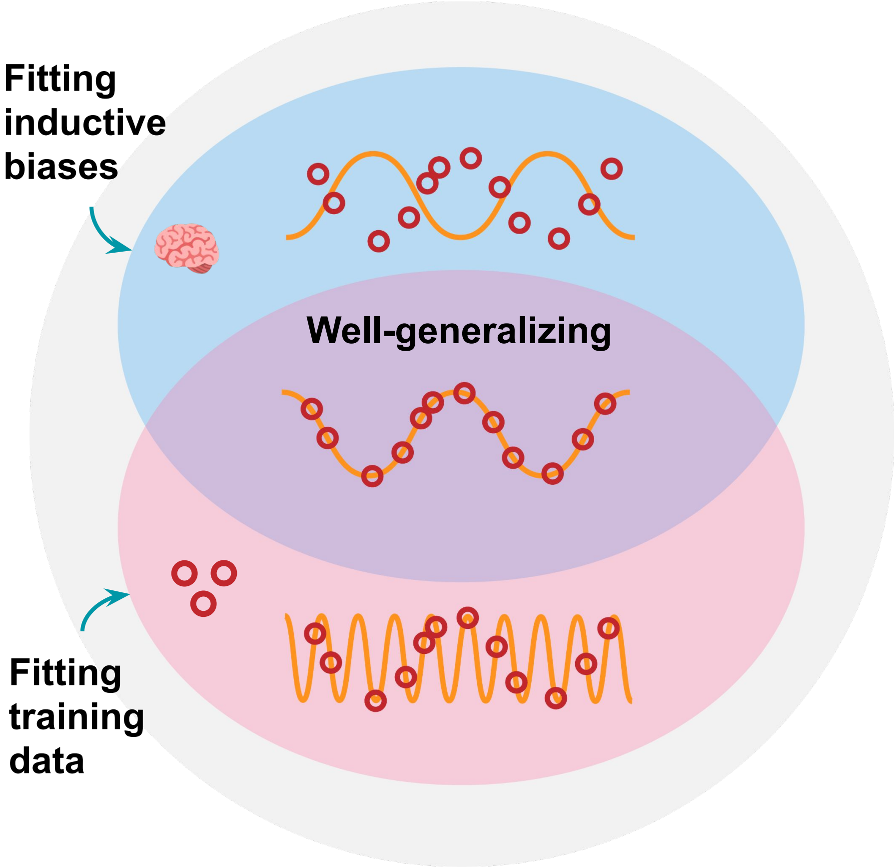
Sample complexity quantifies the amount of data needed to generalize



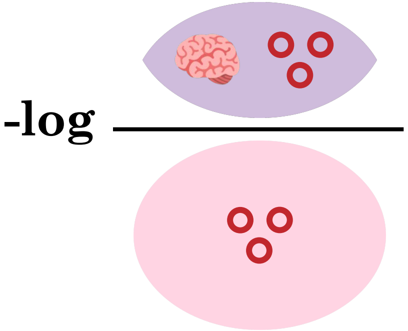
Inductive bias complexity quantifies the amount of inductive bias needed to generalize



How to quantify the inductive bias is required to solve a task?



- How much information does the inductive bias provide about correct hypothesis?
- Information content of inductive biases relates to the amount inductive bias shrinks the hypothesis space:



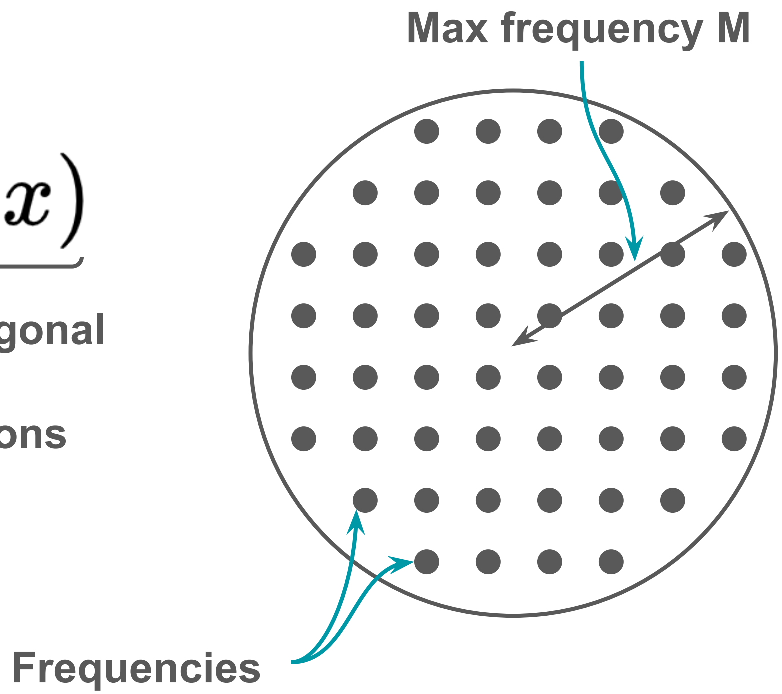
Setting a reasonable hypothesis space

- General hypothesis space:

$$f(x;\theta) = \sum_{\omega\in\Omega} \theta_\omega \underbrace{u_\omega(x)}_{\text{Orthogonal basis functions}}$$

Discrete set of frequencies

$$dim(\Theta) \propto |\Omega| \propto M^m$$



Inductive bias complexity scales exponentially with intrinsic input dimension

$$\tilde{I} = (M^m - Dn)[C - \log \varepsilon - \frac{1}{m} \log n]$$

Approximated task difficulty bound

Max frequency

Training set size

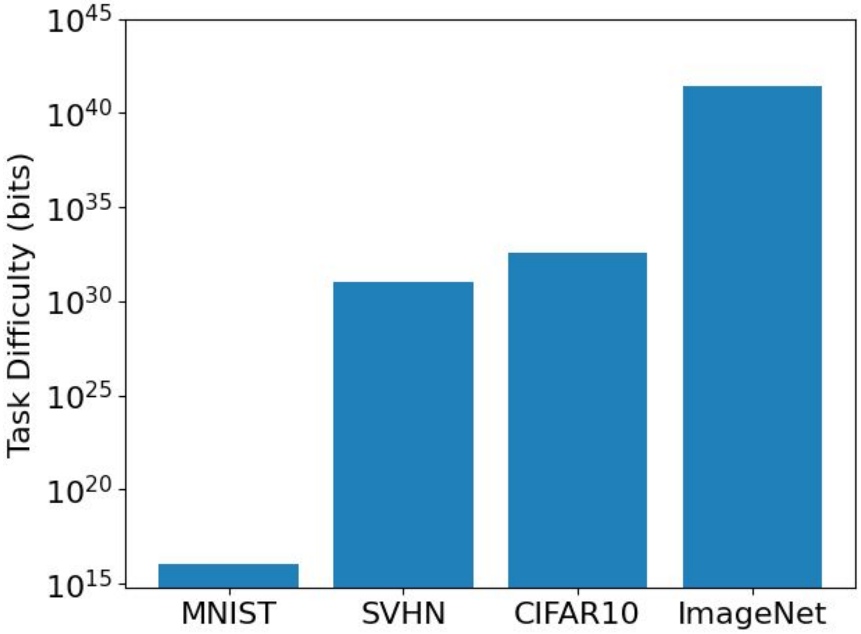
Error rate

Intrinsic data dimensionality

- Task difficulty decreases <u>linearly/logarithmically</u> with training set size
- Task difficulty decreases <u>logarithmically</u> with desired error rate
- Task difficulty increases <u>polynomially</u> with max frequency (i.e. data resolution)
- Task difficulty increases <u>exponentially</u> with intrinsic data dimensionality

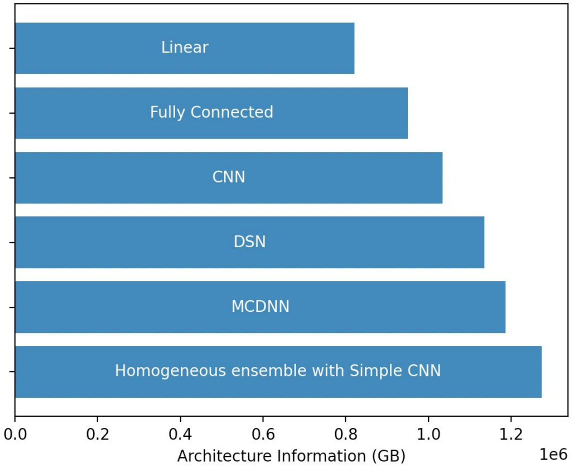
Quantifying difficulty of image classification benchmarks

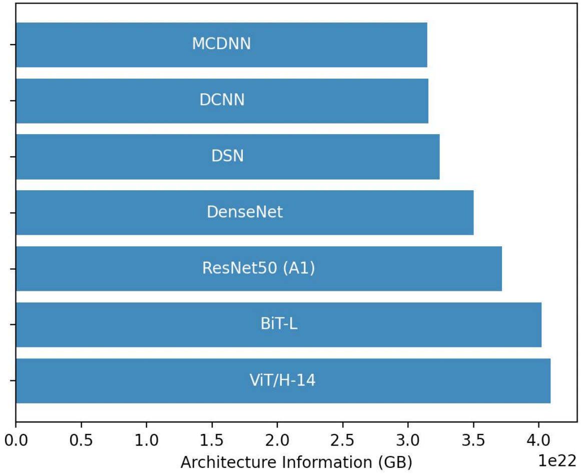
Difficulty to achieve SOTA accuracy

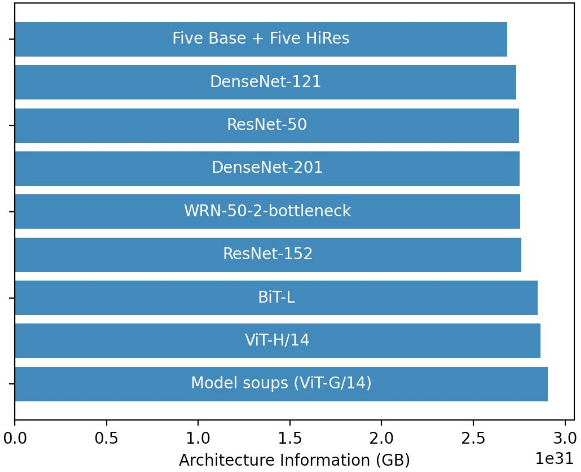


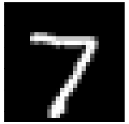
- Datasets of higher intrinsic dimensionality are more difficult
- Task difficulties are large: typical model classes provide high inductive bias

Harder datasets extract more inductive bias from a fixed architecture





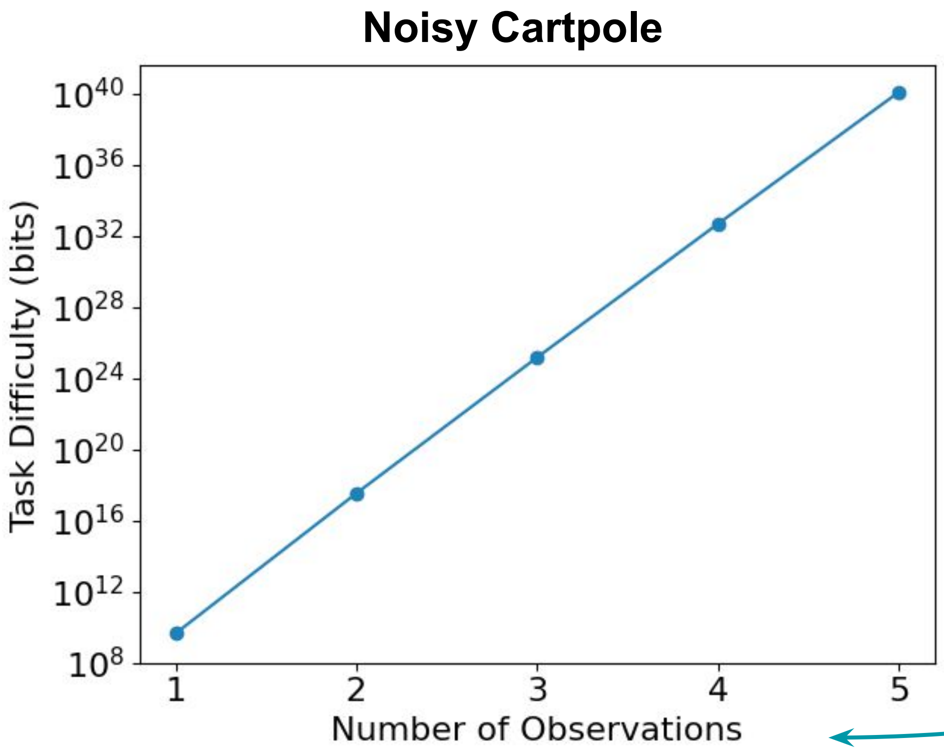


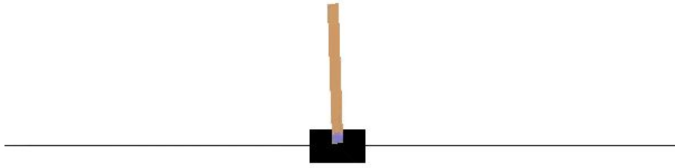
MNIST 

CIFAR-10 

ImageNet 

In RL, tasks with noisier observations require *exponentially* more inductive bias to generalize on

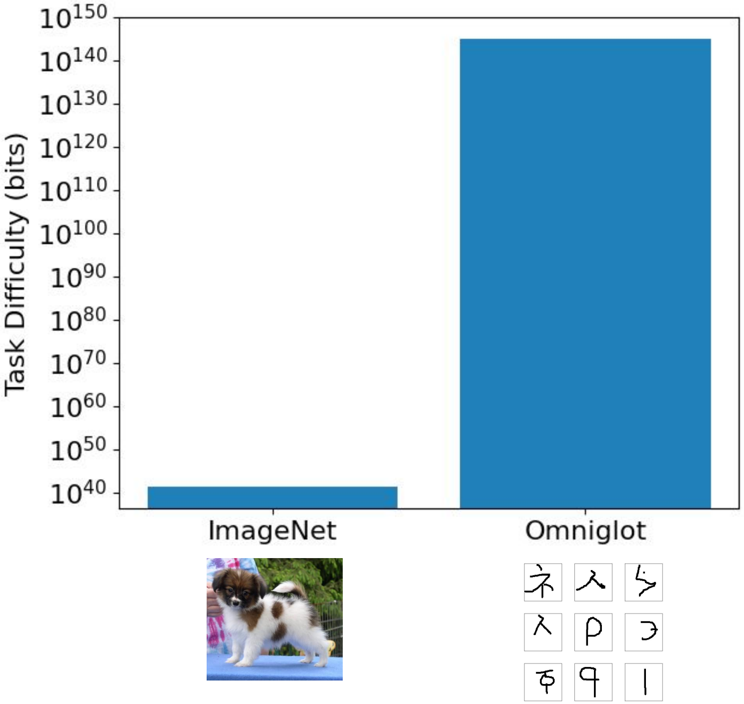




Observations required to determine state

Meta-learning tasks are dramatically more difficult than supervised learning

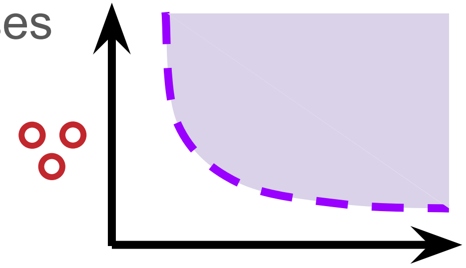
Difficulty to achieve SOTA accuracy



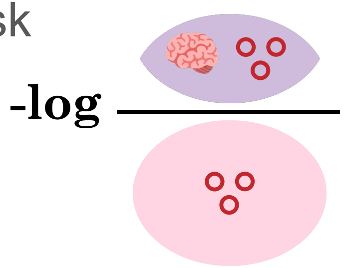
- Meta-learning requires generalizing over a very high dimensional space

Conclusion

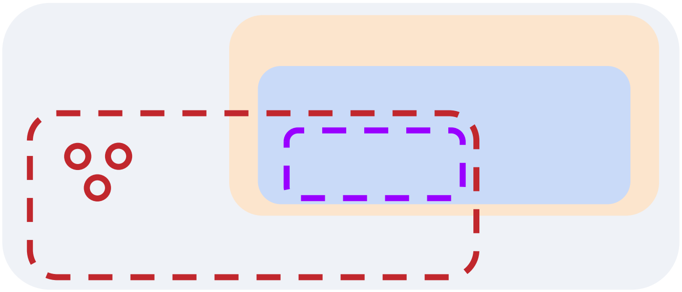
- Generalizing on a task requires both training data and inductive biases



- Task difficulty is information content of inductive biases required to solve a task



- Typical architectures encode vast amounts of inductive bias



- Higher intrinsic dimension tasks require more inductive bias