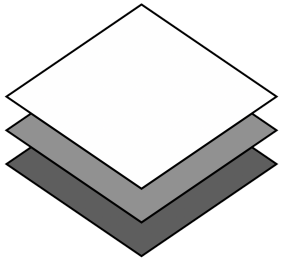
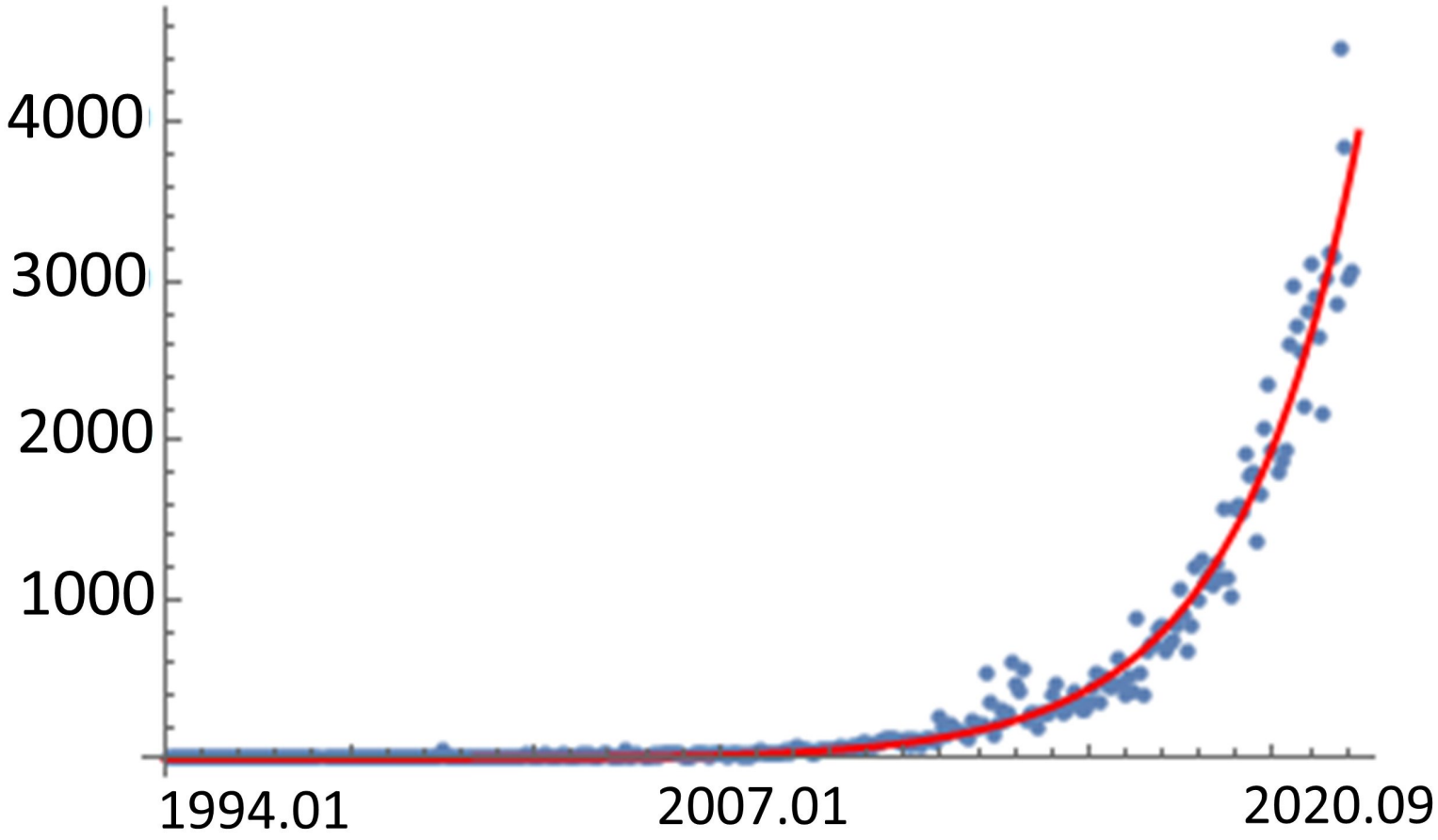


Full Stack Deep Learning

Research Directions

Pieter Abbeel, Sergey Karayev, Josh Tobin

ML+AI arXiv papers per month



Outline

- Sampling of research directions
- Overall research theme
- How to keep up

Many Exciting Directions in AI

- Unsupervised Learning
- Reinforcement Learning
- Unsupervised RL
- Meta-Reinforcement Learning
- Few-Shot Imitation
- Domain Randomization
- DL for Science and Engineering
- Mitigating Bias
- Multi-modal Learning
- Architecture Search
- Value Alignment
- Scaling Laws
- Human-in-the-Loop
- Explainability

Many Exciting Directions in AI

- Unsupervised Learning
- Reinforcement Learning
- Unsupervised RL
- Meta-Reinforcement Learning
- Few-Shot Imitation
- Domain Randomization
- DL for Science and Engineering

- Mitigating Bias
- Multi-modal Learning
- Architecture Search
- Value Alignment
- Scaling Laws
- Human-in-the-Loop
- Explainability

Many Exciting Directions in AI

- ***Unsupervised Learning***
- Reinforcement Learning
- Unsupervised RL
- Meta-Reinforcement Learning
- Few-Shot Imitation
- Domain Randomization
- DL for Science and Engineering

- Mitigating Bias
- Multi-modal Learning
- Architecture Search
- Value Alignment
- Scaling Laws
- Human-in-the-Loop
- Explainability

Recall: Deep Supervised Learning

- Works!
- BUT: requires so much annotated data

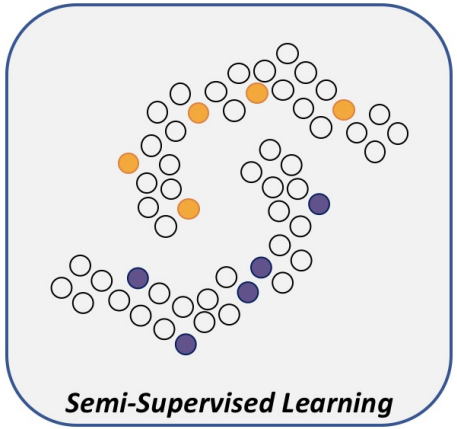
So: Can we learn with less labels?

Yes!

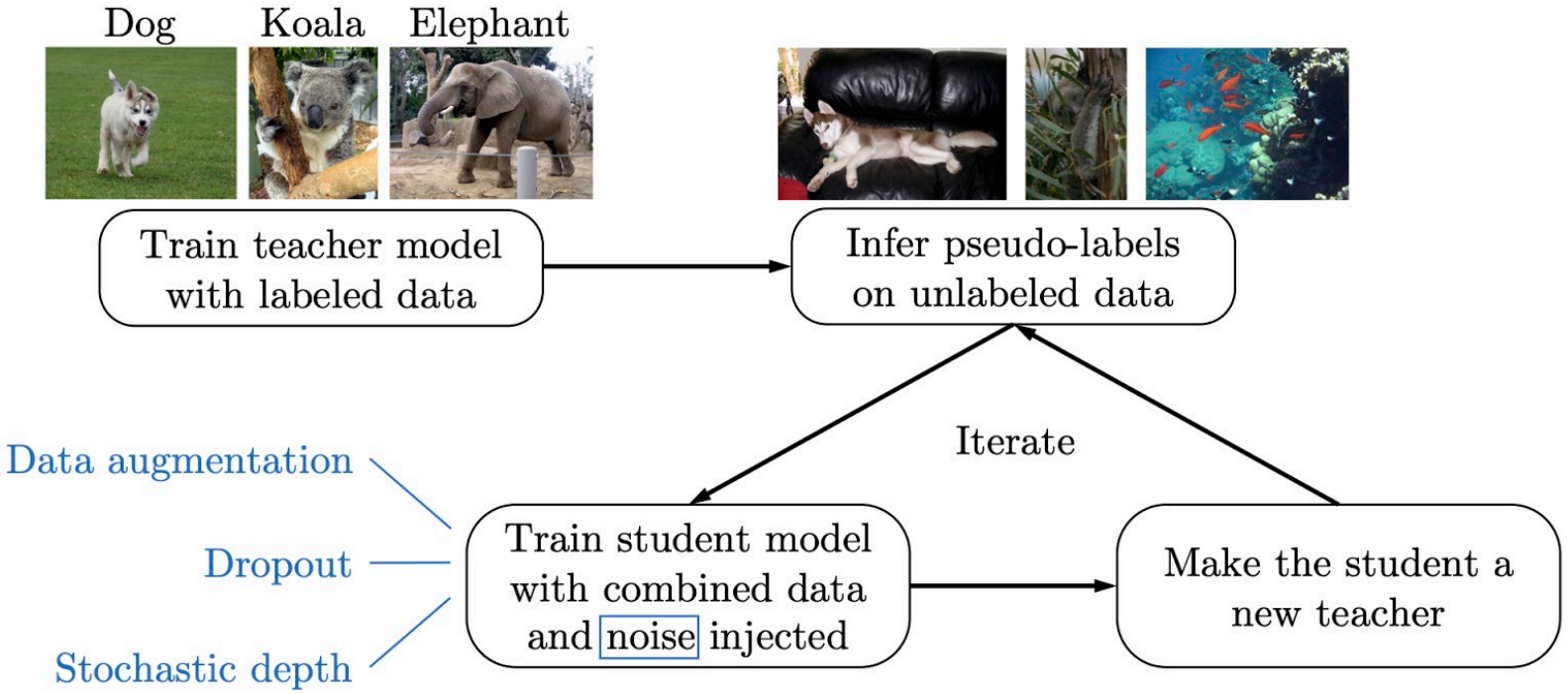
- ***Deep semi-supervised learning***
- Deep unsupervised learning

Semi-supervised Learning

- Assumption:
  - Classification problem
  - Each data-point belongs to one of the classes
- Toy Example: “two moons”

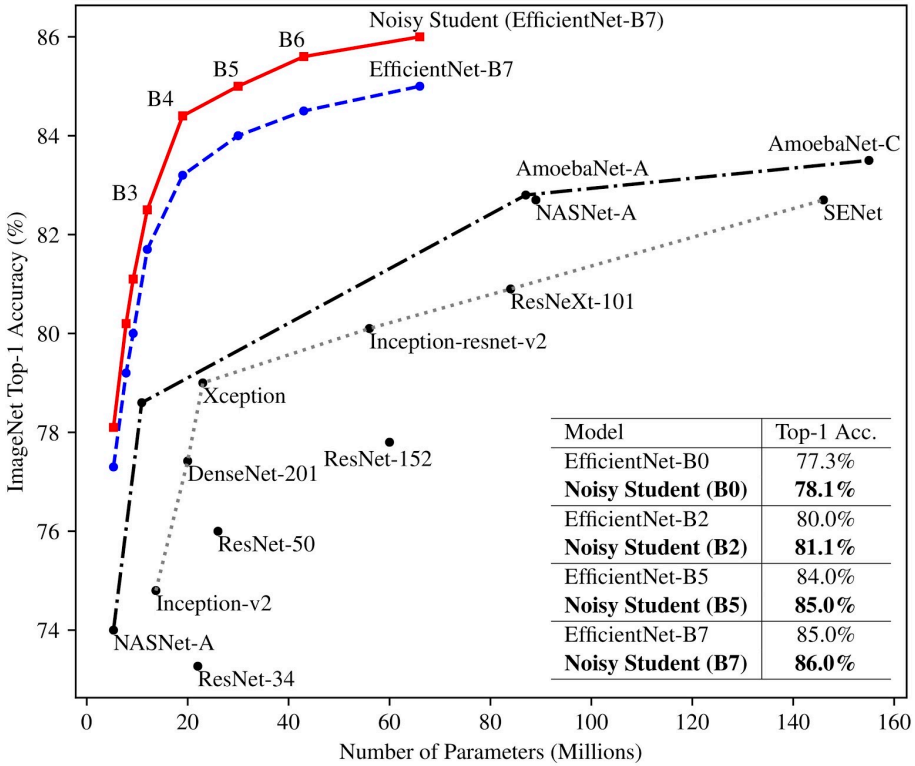


Noisy-Student



[Xie et al, 2020]

Noisy-Student -- Results



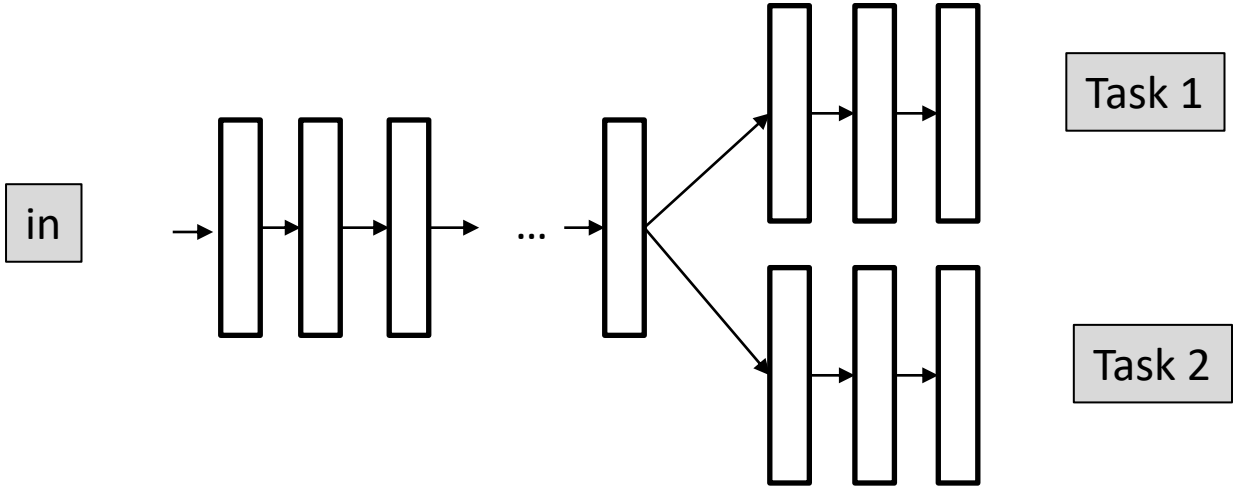
Model	Top-1 Acc.
EfficientNet-B0	77.3%
Noisy Student (B0)	**78.1%**
EfficientNet-B2	80.0%
Noisy Student (B2)	**81.1%**
EfficientNet-B5	84.0%
Noisy Student (B5)	**85.0%**
EfficientNet-B7	85.0%
Noisy Student (B7)	**86.0%**

So: Can we learn with less labels?

Yes!

- Deep semi-supervised learning
- ***Deep unsupervised learning***

Transfer with Multi-headed Networks



Deep Unsupervised Learning

- Key hypothesis:

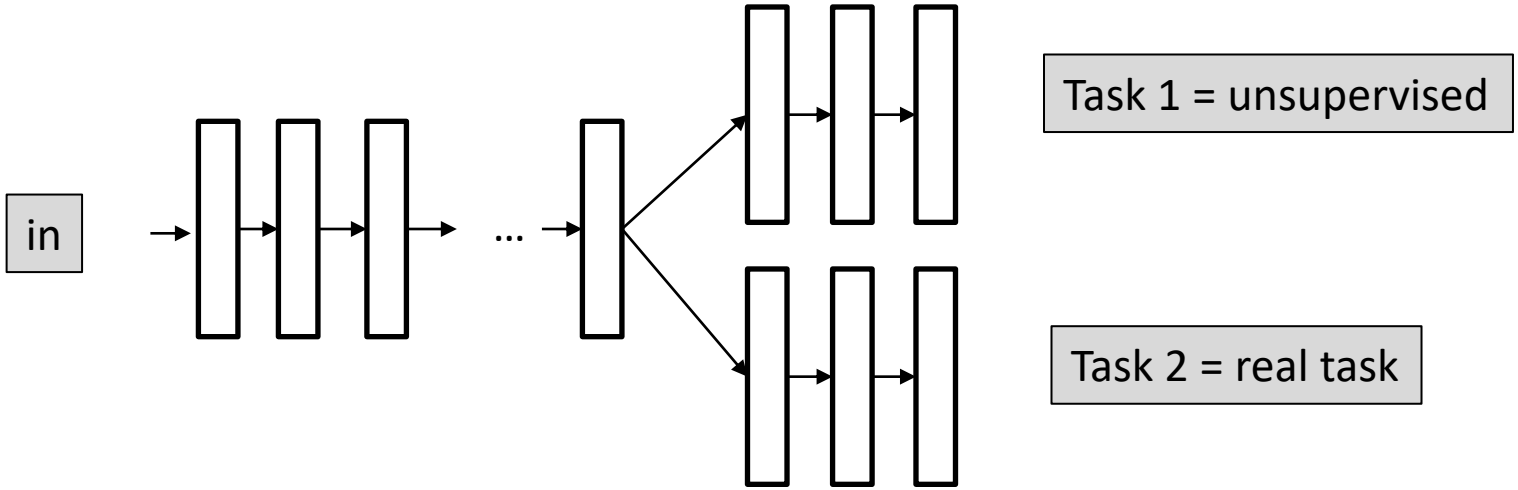
Task 1

- IF neural network smart enough to predict:
  - Next frame in video
  - Next word in sentence
  - Generate realistic images
  - ``Translate'' images
  - …

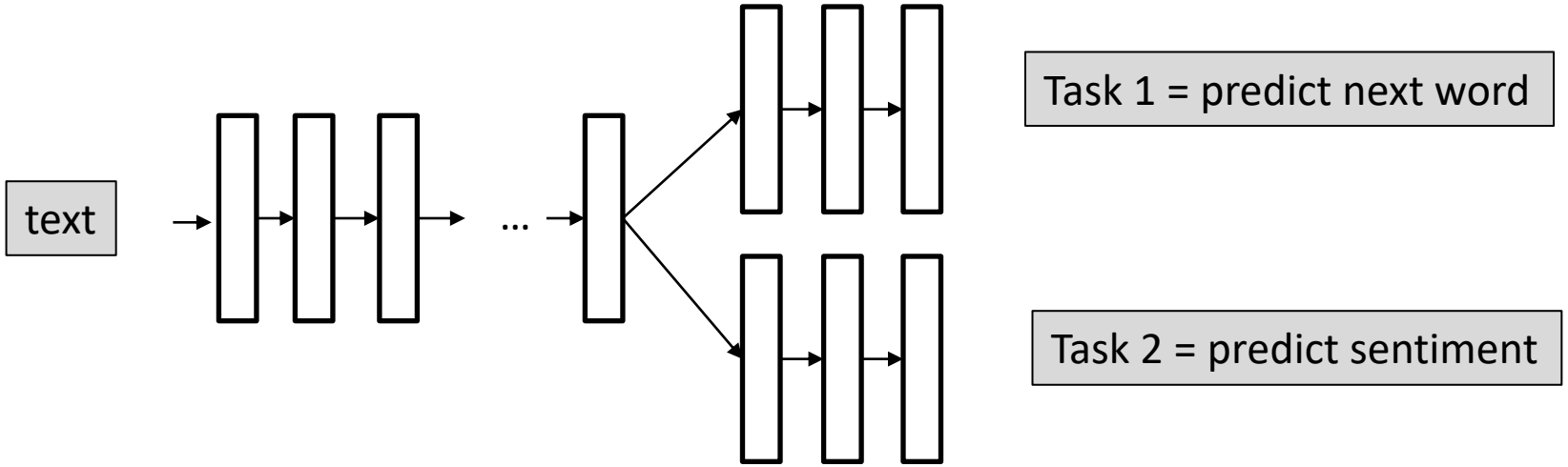
Task 2

- THEN same neural network is ready to do Deep Supervised Learning from very small data-set

Transfer from Unsupervised Learning



Example Setting



Next Word Prediction (OpenAI's GPT-2)

Prompt: A train carriage containing controlled nuclear materials was stolen in Cincinnati today. Its whereabouts are unknown.

GPT-2: *The incident occurred on the downtown train line, which runs from Covington and Ashland stations. In an email to Ohio news outlets, the U.S. Department of Energy said it is working with the Federal Railroad Administration to find the thief.*
'The theft of this nuclear material will have significant negative consequences on public and environmental health, our workforce and the economy of our nation,' said Tom Hicks, the U.S. Energy Secretary, in a statement. 'Our top priority is to secure the theft and ensure it doesn't happen again.'
The stolen material was taken from the University of Cincinnati's Research Triangle Park nuclear research site, according to a news release from Department officials.
The Nuclear Regulatory Commission did not immediately release any information.
According to the release, the U.S. Department of Energy's Office of Nuclear Material Safety and Security is leading that team's investigation.
'The safety of people, the environment and the nation's nuclear stockpile is our highest priority,' Hicks said. 'We will get to the bottom of this and make no excuses.'

Next Word Prediction (OpenAI's GPT-2)

Prompt: Recycling is good for the world, no, you could not be more wrong.

GPT-2: Recycling is NOT good for the world. It is bad for the environment, it is bad for our health, and it is bad for our economy. I'm not kidding. Recycling is not good for the environment. It is destructive to the earth and it is a major contributor to global warming. Recycling is not good for our health. It contributes to obesity and diseases like heart disease and cancer. Recycling is bad for our economy. It increases the cost of a product, and in turn, the price of everything that is made with that product. Recycling is not good for our nation. We pay a tremendous price for the privilege of having the world's most advanced and efficient recycling system. Recycling is a huge, colossal waste of time, energy, money, and resources.

Text Generatio… …'s GPT-2)

SYSTEM PROMPT (HUMAN-WRITTEN)

In a shocking fi… remote, previou… surprising to the… English.

MODEL COMPLETION (MACHINE-WRITTEN, 10 TRIES)

The scientist nam… Ovid's Unicorn. Th… previously unknown…

Now, after almost tw… odd phenomenon is fi…

Dr. Jorge Pérez, an ev… La Paz, and several co… Mountains when they fou… or humans. Pérez noticed… a natural fountain, surr… snow.

Pérez and the others then … the time we reached the top… with some crystals on top,"…

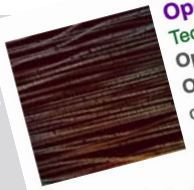
Pérez and his friends were as… These creatures could be seen … too much to see them — they we… horns.

…izarre creatures the scientists discovered … spoke some fairly regular English. Pérez … example, that they have a common …e a dialect or dialectic."

…e unicorns may have originated in …ls were believed to be descendants of a …ed there before the arrival of humans …rica.

… unclear, some believe that perhaps …n a human and a unicorn met each …ivilization. According to Pérez, …ts seem to be quite common."

… that it is likely that the only …ns are indeed the descendants of … "But they seem to be able to … which I believe is a sign of … social organization," said the …



OpenAI built a text generator so good, it's considered too dangerous…
TechCrunch - 17 Feb 2019
OpenAI built a text generator **so** good, it's considered **too dangerous to release** …
OpenAI said its new natural **language model**, GPT-2, was trained to … said, it's only **releasing** a smaller version of the **language model**, citing its …
Scientists Developed an AI **So** Advanced They Say It's **Too Dangerous** …
ScienceAlert - 18 Feb 2019

AI text writing technology **too dangerous to release**, creators claim
The Drum - 17 Feb 2019
This technology could 'absolutely devastate' the internet as we know it
NEWS.com.au - 17 Feb 2019
This AI is **so** good at writing that its creators won't let you use it
In-Depth - CNN - 18 Feb 2019
Lord of The Rings, Celebrity Gossip: This AI is **So** Good at Writing That …
In-Depth - News18 - 18 Feb 2019

View all

When Is Technology Too Dangerous to Release to the Public?
Slate Magazine - 22 Feb 2019
If your knowledge of the model, called GPT-2, came solely on headlines … U.K. read, "Elon Musk-Founded **OpenAI** Builds Artificial Intelligence **So** … had trained a **language model** using text from 8 million webpages to predict …
AI Weekly: Experts say **OpenAI's** controversial **model** is a potential …
In-Depth - VentureBeat - 22 Feb 2019

View all

OpenAI's Text Model so Disruptive it's Deemed Too Dangerous To …
Computer Business Review - 15 Feb 2019
OpenAI's Text Model **so** Disruptive it's Deemed **Too Dangerous To Release** …
OpenAI has declined to **release** the full research due to concerns over … We've trained an unsupervised **language model** that can generate …
New AI fake text generator may be **too dangerous to release**, say …
Highly Cited - The Guardian - 14 Feb 2019

View all

[Radford et al, 2019]

Unsupervised Sentiment Neuron

This is one of Crichton's best books. The characters of Karen Ross, Peter Elliot, Munro, and Amy are beautifully developed and their interactions are exciting, complex, and fast-paced throughout this impressive novel. And about 99.8 percent of that got lost in the film. Seriously, the screenplay AND the directing were horrendous and clearly done by people who could not fathom what was good about the novel. I can't fault the actors because frankly, they never had a chance to make this turkey live up to Crichton's original work. I know good novels, especially those with a science fiction edge, are hard to bring to the screen in a way that lives up to the original. But this may be the absolute worst disparity in quality between novel and screen adaptation ever. The book is really, really good. The movie is just dreadful.

[Radford et al, 2017]

Benchmarks

TASK	**Common Sense Reasoning**: resolution of an ambiguous pronoun
DATASET	Winograd Schema Challenge
EXAMPLES	*The trophy doesn't fit into the brown suitcase because it is too large.*
	Correct answer: *it = trophy* **Model answer**: `it = trophy`
	The trophy doesn't fit into the brown suitcase because it is too small.
	Correct answer: *it = suitcase* **Model answer**: `it = suitcase`

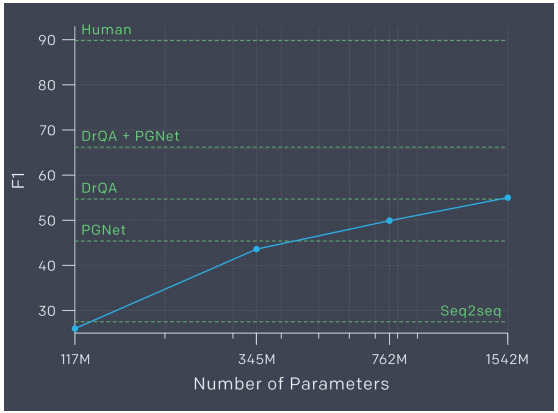
Benchmarks

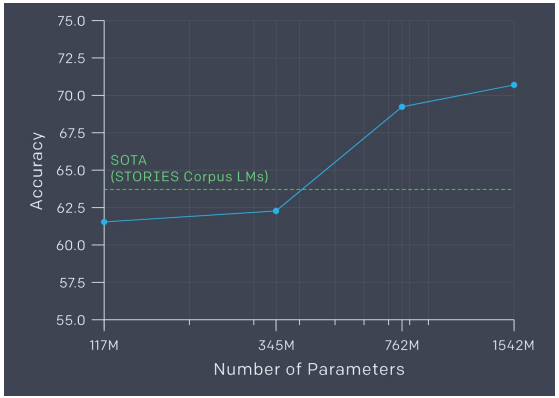
TASK	**Question Answering**
DATASET	Natural Questions
EXAMPLES	*Who wrote the book the origin of species?*
	Correct answer: *Charles Darwin* **Model answer**: `Charles Darwin`
	What is the largest state in the U.S. by land mass?
	Correct answer: *Alaska* **Model answer**: `California`

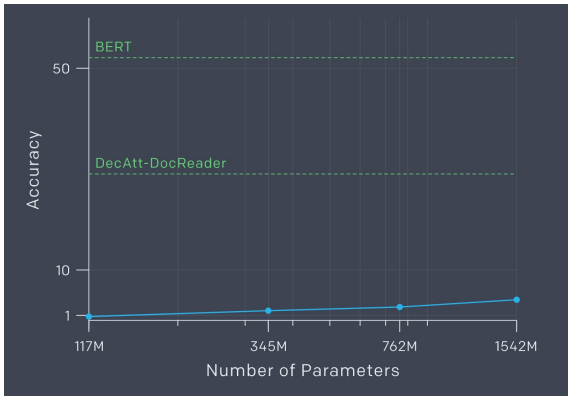
Benchmarks

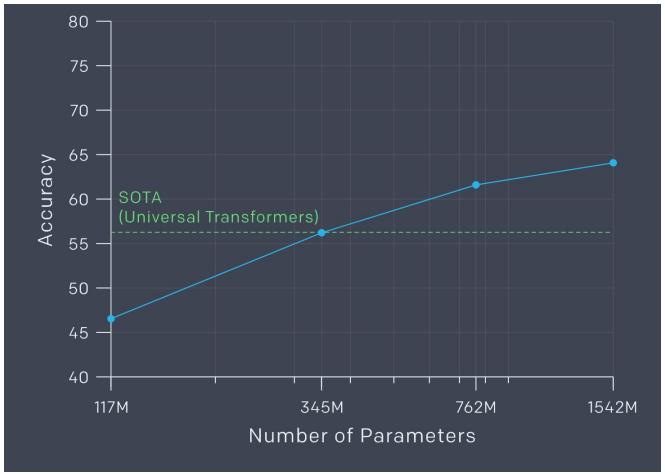
DATASET	METRIC	OUR RESULT	PREVIOUS RECORD	HUMAN
Winograd Schema Challenge	accuracy (+)	**70.70%**	63.7%	92%+
LAMBADA	accuracy (+)	**63.24%**	59.23%	95%+
LAMBADA	perplexity (-)	**8.6**	99	~1-2
Children's Book Test Common Nouns (validation accuracy)	accuracy (+)	**93.30%**	85.7%	96%
Children's Book Test Named Entities (validation accuracy)	accuracy (+)	**89.05%**	82.3%	92%
Penn Tree Bank	perplexity (-)	**35.76**	46.54	unknown
WikiText-2	perplexity (-)	**18.34**	39.14	unknown

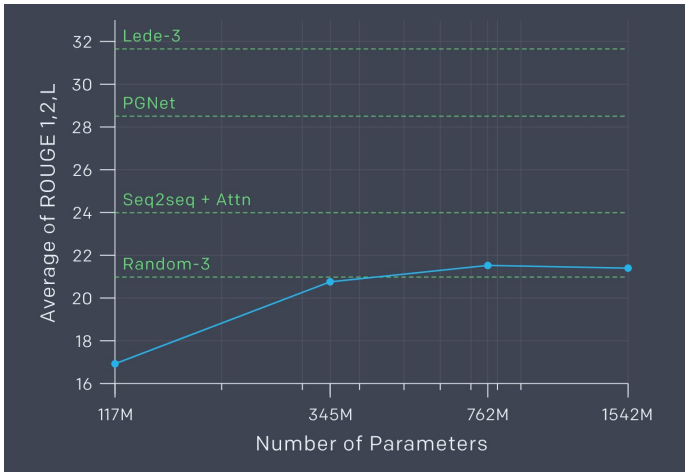
Scaling


Human
DrQA + PGNet
DrQA
PGNet
Seq2seq
F1
90
80
70
60
50
40
30
117M
345M
762M
1542M
Number of Parameters


SOTA
(STORIES Corpus LMs)
Accuracy
75.0
72.5
70.0
67.5
65.0
62.5
60.0
57.5
55.0
117M
345M
762M
1542M
Number of Parameters

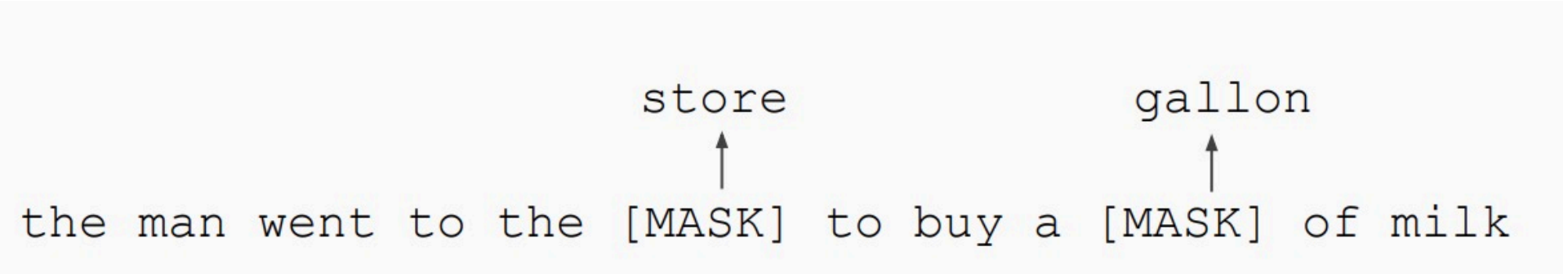

BERT
DecAtt-DocReader
Accuracy
50
10
1
117M
345M
762M
1542M
Number of Parameters


SOTA
(Universal Transformers)
Accuracy
80
75
70
65
60
55
50
45
40
117M
345M
762M
1542M
Number of Parameters


Lede-3
PGNet
Seq2seq + Attn
Random-3
Average of ROUGE 1,2,L
32
30
28
26
24
22
20
18
16
117M
345M
762M
1542M
Number of Parameters

BERT

- Main idea:
  - GPT-2: predict next word / token
  - BERT: predict a word / token that was removed



BERT

Sentence A = The man went to the store.
Sentence B = He bought a gallon of milk.
Label = IsNextSentence

Sentence A = The man went to the store.
Sentence B = Penguins are flightless.
Label = NotNextSentence

BERT

GLUE Results

System	MNLI-(m/mm) 392k	QQP 363k	QNLI 108k	SST-2 67k	CoLA 8.5k	STS-B 5.7k	MRPC 3.5k	RTE 2.5k	**Average** -
Pre-OpenAI SOTA	80.6/80.1	66.1	82.3	93.2	35.0	81.0	86.0	61.7	74.0
BiLSTM+ELMo+Attn	76.4/76.1	64.8	79.9	90.4	36.0	73.3	84.9	56.8	71.0
OpenAI GPT	82.1/81.4	70.3	88.1	91.3	45.4	80.0	82.3	56.0	75.2
$\text{BERT}_{\text{BASE}}$	84.6/83.4	71.2	90.1	93.5	52.1	85.8	88.9	66.4	79.6
$\text{BERT}_{\text{LARGE}}$	**86.7/85.9**	**72.1**	**91.1**	**94.9**	**60.5**	**86.5**	**89.3**	**70.1**	**81.9**

MultiNLI

Premise: Hills and mountains are especially sanctified in Jainism.
Hypothesis: Jainism hates nature.
Label: Contradiction

CoLa

Sentence: The wagon rumbled down the road.
Label: Acceptable

Sentence: The car honked down the road.
Label: Unacceptable

What Is the Google BERT Search Algorithm Update?

JOYDEEP BHATTACHARYA

Google BERT stands for **Bidirectional Encoder Representations from Transformers** and is an update to the core search algorithm aimed at improving the language understanding capabilities of Google.

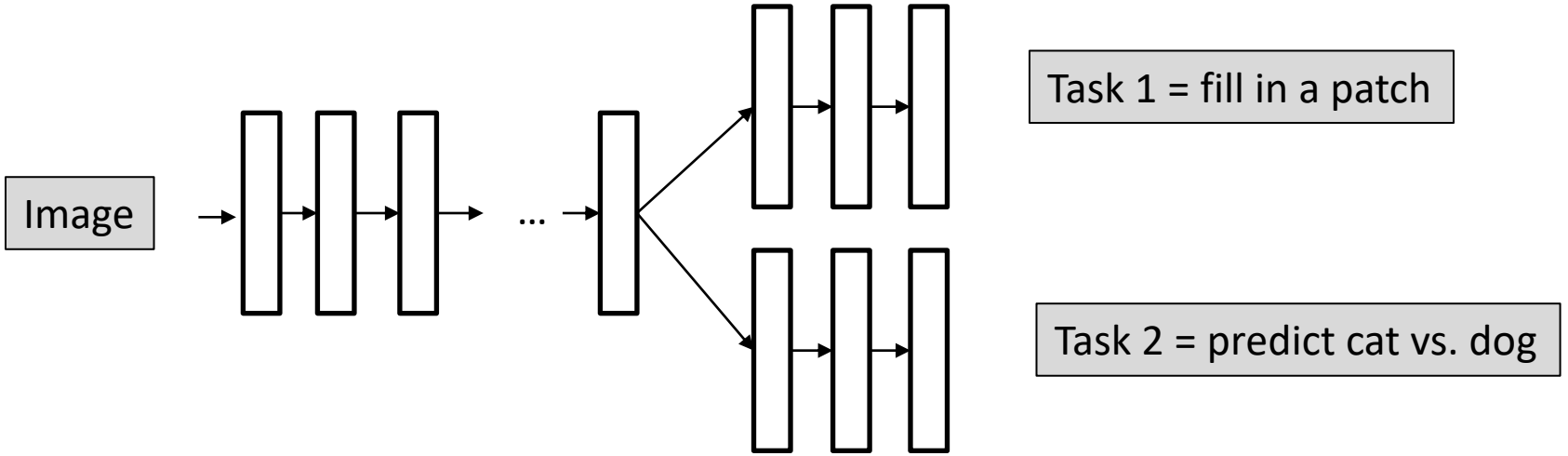
BERT is one of the biggest updates that Google has made since RankBrain in 2015 and has proven successful in comprehending the intent of the searcher behind a search query.

Downstream Tasks - NLP (BERT Revolution)

Rank	Name	Model	URL	Score	CoLA	SST-2	MRPC	STS-B	QQP	MNLI-m	MNLI-mm	QNLI	RTE	WNLI
1	T5 Team - Google	T5		90.3	71.6	97.5	92.8/90.4	93.1/92.8	75.1/90.6	92.2	91.9	96.9	92.8	94.5
2	ERNIE Team - Baidu	ERNIE		90.0	72.2	97.5	93.0/90.7	92.9/92.5	75.2/90.8	91.2	90.8	96.0	90.9	94.5
3	Microsoft D365 AI & MSR AI & GATECH	MT-DNN-SMART		89.9	69.5	97.5	93.7/91.6	92.9/92.5	73.9/90.2	91.0	90.8	99.2	89.7	94.5
4	王玮	ALICE v2 large ensemble (Alibaba DAMO NLP)		89.7	73.2	97.1	93.9/91.9	93.0/92.5	74.8/91.0	90.8	90.6	95.9	87.4	94.5
5	Microsoft D365 AI & UMD	FreeLB-RoBERTa (ensemble)		88.4	68.0	96.8	93.1/90.8	92.3/92.1	74.8/90.3	91.1	90.7	95.6	88.7	89.0
6	Junjie Yang	HIRE-RoBERTa		88.3	68.6	97.1	93.0/90.7	92.4/92.0	74.3/90.2	90.7	90.4	95.5	87.9	89.0
7	Facebook AI	RoBERTa		88.1	67.8	96.7	92.3/89.8	92.2/91.9	74.3/90.2	90.8	90.2	95.4	88.2	89.0
8	Microsoft D365 AI & MSR AI	MT-DNN-ensemble		87.6	68.4	96.5	92.7/90.3	91.1/90.7	73.7/89.9	87.9	87.4	96.0	86.3	89.0
9	GLUE Human Baselines	GLUE Human Baselines		87.1	66.4	97.8	86.3/80.8	92.7/92.6	59.5/80.4	92.0	92.8	91.2	93.6	95.9
10	Stanford Hazy Research	Snorkel MeTaL		83.2	63.8	96.2	91.5/88.5	90.1/89.7	73.1/89.9	87.6	87.2	93.9	80.9	65.1

[https://gluebenchmark.com/leaderboard]

Unsupervised Learning in Vision

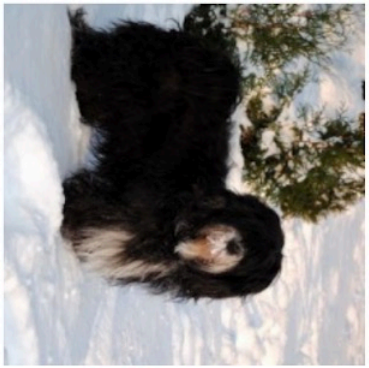


Predict Missing Patch

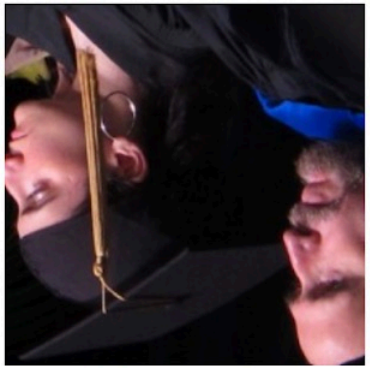
Solving Jigsaw Puzzles

Rotation Prediction

90° rotation



270° rotation



180° rotation

0° rotation

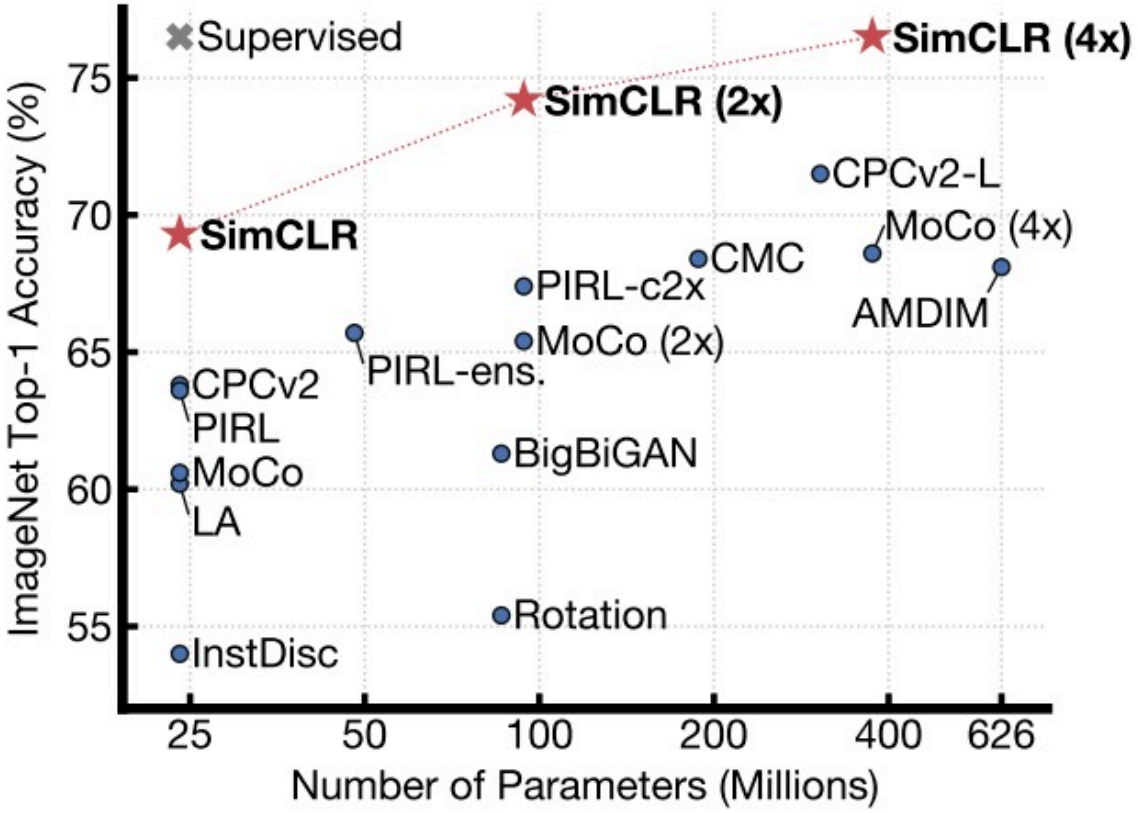
270° rotation

SimCLR and MoCo

attract

repel

SimCLR + linear classifier

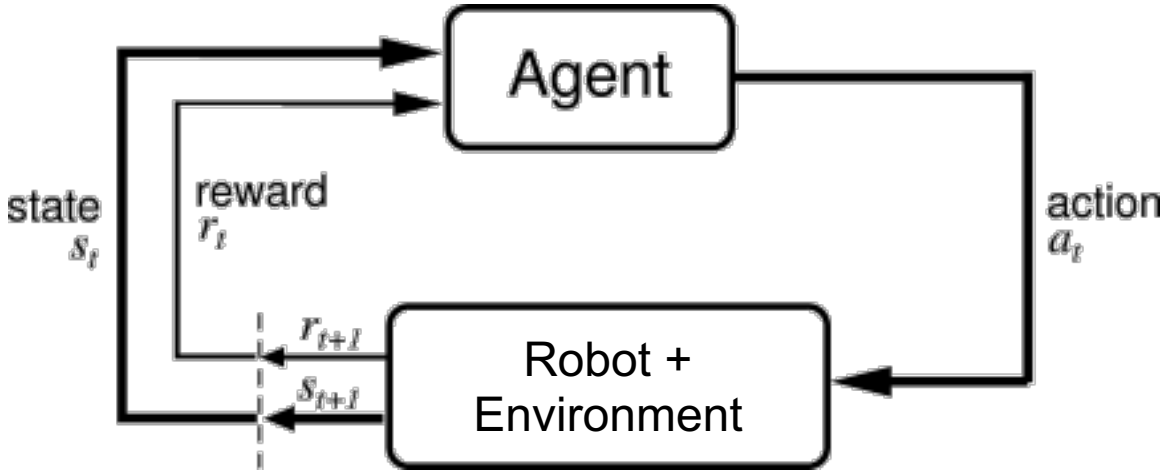


Many Exciting Directions in AI

- Unsupervised Learning
- ***Reinforcement Learning***
- Unsupervised RL
- Meta-Reinforcement Learning
- Few-Shot Imitation
- Domain Randomization
- DL for Science and Engineering

- Mitigating Bias
- Multi-modal Learning
- Architecture Search
- Value Alignment
- Scaling Laws
- Human-in-the-Loop
- Explainability

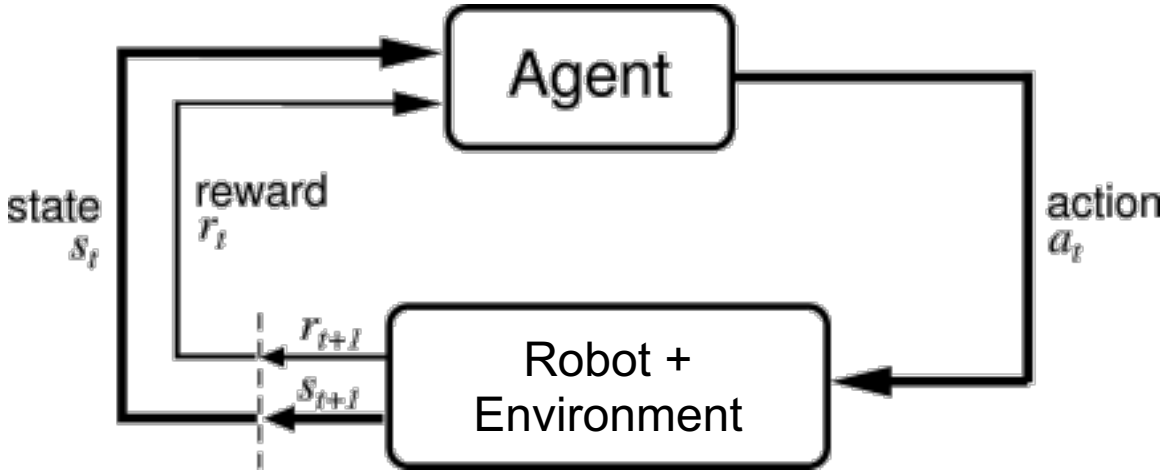
Reinforcement Learning (RL)



$$\max_{\theta} \quad \mathrm{E}[\sum_{t=0}^{H} R(s_t)|\pi_\theta]$$

[Image credit: Sutton and Barto 1998]

Reinforcement Learning (RL)

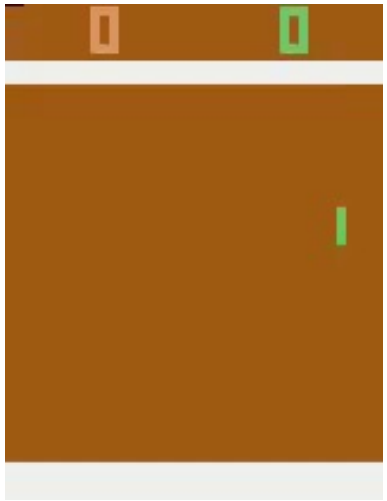
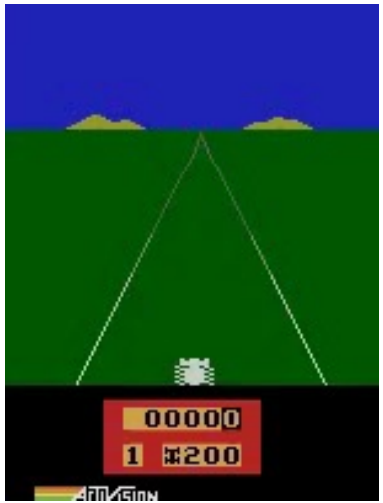


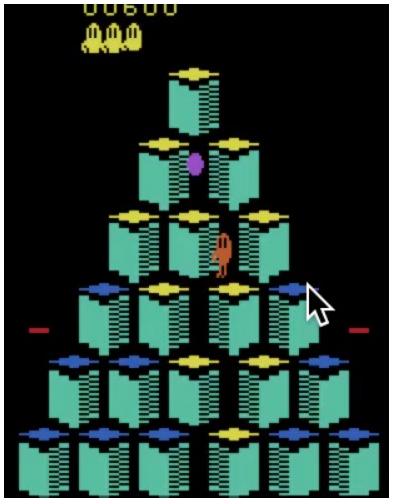
$$\max_{\theta} \quad \mathrm{E}[\sum_{t=0}^{H} R(s_t)|\pi_\theta]$$

- **Compared to supervised learning, additional challenges:**
  - **Credit assignment**
  - **Stability**
  - **Exploration**

[Image credit: Sutton and Barto 1998]

Deep RL: Atari

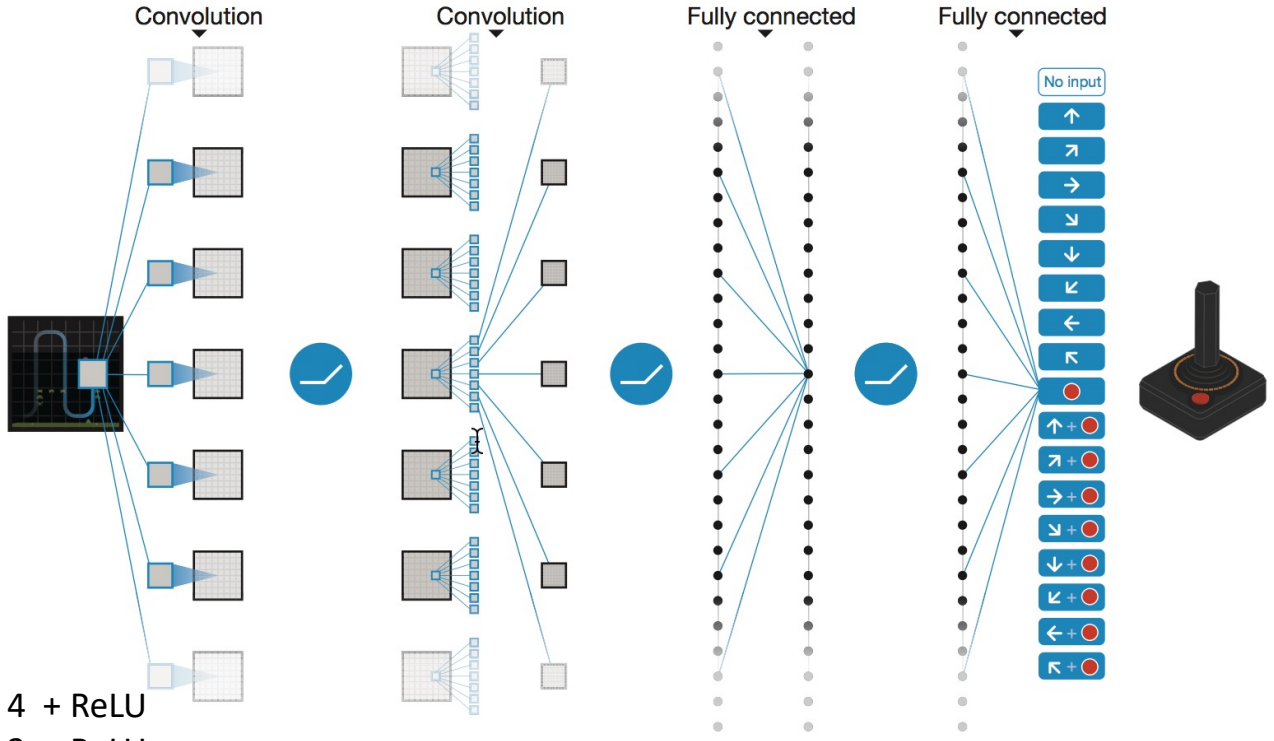






DQN Mnih et al, NIPS 2013 / Nature 2015

MCTS Guo et al, NIPS 2014; **TRPO** Schulman, Levine, Moritz, Jordan, Abbeel, ICML 2015; **A3C** Mnih et al, ICML 2016; **Dueling DQN** Wang et al ICML 2016; **Double DQN** van Hasselt et al, AAAI 2016; **Prioritized Experience Replay** Schaul et al, ICLR 2016; **Bootstrapped DQN** Osband et al, 2016; **Q-Ensembles** Chen et al, 2017; **Rainbow** Hessel et al, 2017; **Accelerated** Stooke and Abbeel, 2018; …

Deep Q-Network (DQN): From Pixels to Joystick Commands



32 8x8 filters with stride 4 + ReLU
64 4x4 filters with stride 2 + ReLU
64 3x3 filters with stride 1 + ReLU
fully connected 512 units + ReLU
fully connected output units, one per action

[Source: Mnih et al., Nature 2015 (DeepMind)]

Deep RL: Go

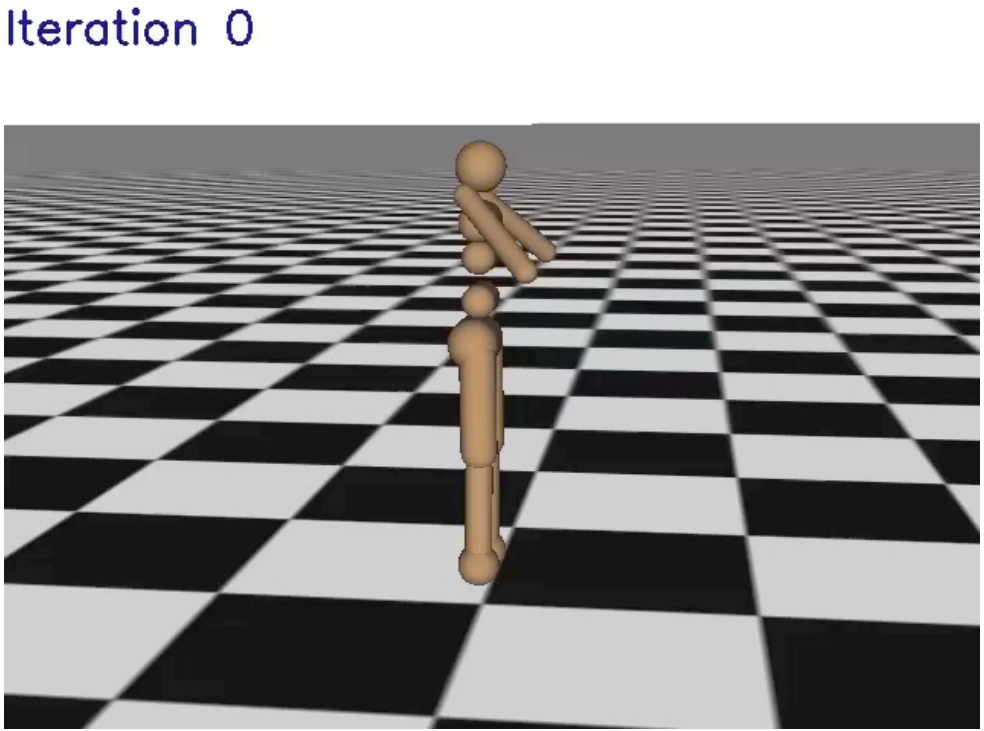
AlphaGo Silver et al, Nature 2015

AlphaGoZero Silver et al, Nature 2017

AlphaZero Silver et al, 2017

Tian et al, 2016; Maddison et al, 2014; Clark et al, 2015

Deep RL: Robot Locomotion



^ **TRPO** Schulman et al, 2015 + **GAE** Schulman et al, 2016

See also: **DDPG** Lillicrap et al 2015; **SVG** Heess et al, 2015; **Q-Prop** Gu et al, 2016; **Scaling up ES** Salimans et al, 2017; **PPO** Schulman et al, 2017; **Parkour** Heess et al, 2017;

Deep RL: Robot Locomotion

[Peng, Abbeel, Levine, van de Panne, 2018]

Deep RL Success: Dynamic Animation

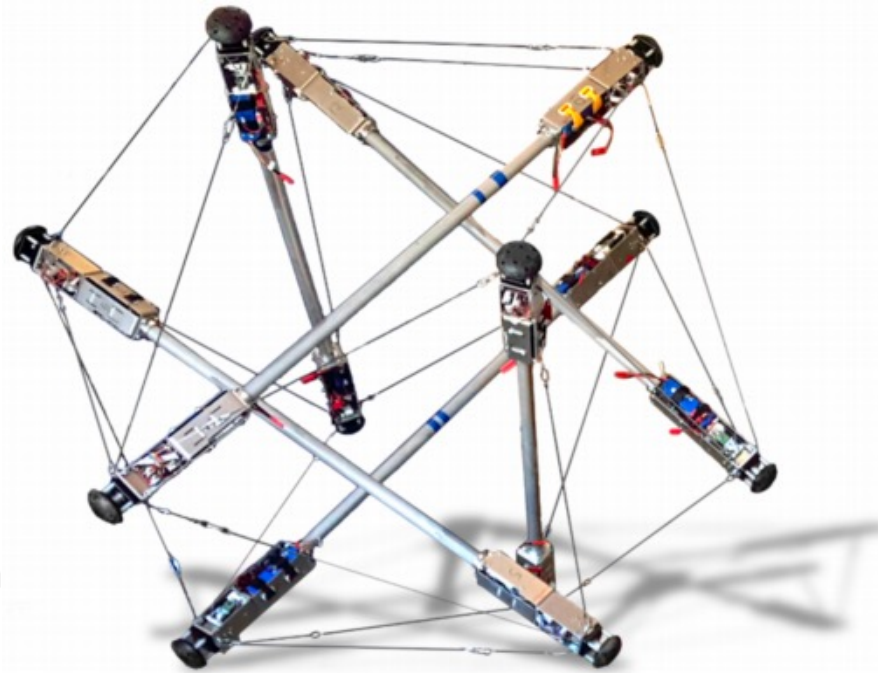
[Peng, Abbeel, Levine, van de Panne, 2018]

BRETT: Berkeley Robot for the Elimination of Tedious Tasks

[Levine*, Finn*, Darrell, Abbeel, JMLR 2016]

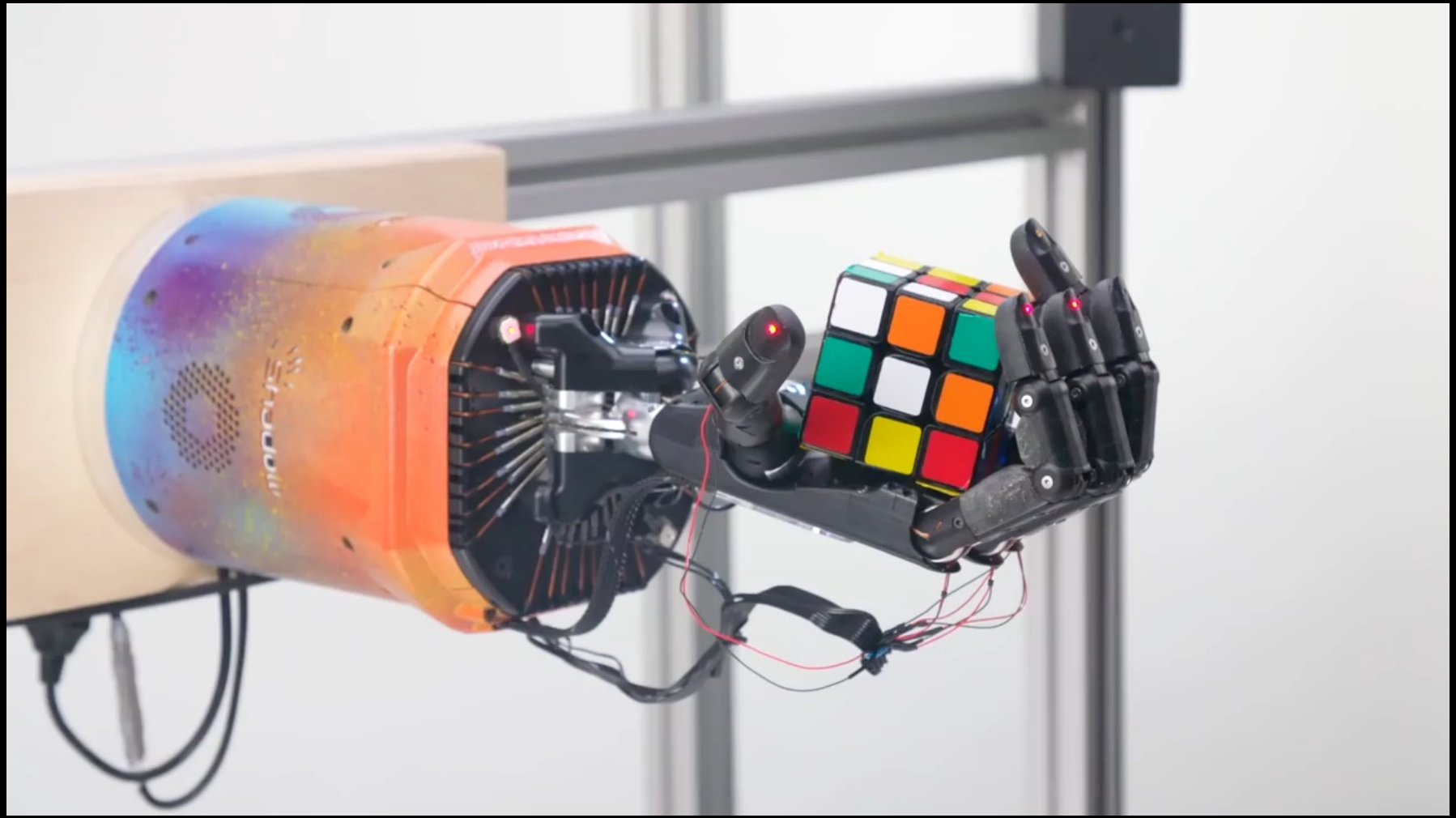
Tensegrity Robotics: NASA SuperBall

- Rigid rods connected by elastic cables
- Controlled by motors that extend / contract cables
- Properties:
  - Lightweight
  - Low cost
  - Capable of withstanding significant impact
- NASA investigates them for space exploration
- Major challenge: control



[Geng*, Zhang*, Bruce*, Caluwaerts, Vespignani, Sunspiral, Abbeel, Levine, ICRA 2017]

[Geng*, Zhang*, Bruce*, Caluwaerts, Vespignani, Sunspiral, Abbeel, Levine, ICRA 2017]



[OpenAI, 2019]

Covariant

Press January 29, 2020

A Warehouse Robot Learns to Sort Out the Tricky Stuff

WIRED

Press January 29, 2020

AI Helps Warehouse Robots Pick Up New Tricks

Press May 6, 2020

Logistics AI Startup Covariant Reaps $40 Million in Funding Round

This is the entry point of AI Robotics into the ***real*** world

The New York Times
29-January-2020

Autonomous Order Picking

Knapp Pick-It-Easy powered by Covariant Brain

Many Exciting Directions in AI

- Unsupervised Learning
- Reinforcement Learning
- ***Unsupervised RL***
- Meta-Reinforcement Learning
- Few-Shot Imitation
- Domain Randomization
- DL for Science and Engineering

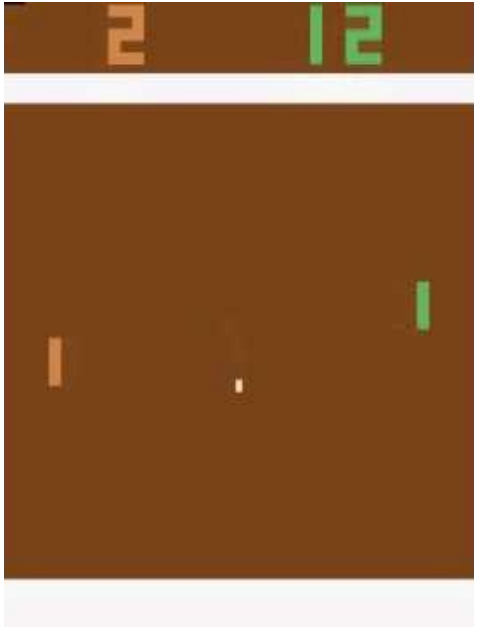
- Mitigating Bias
- Multi-modal Learning
- Architecture Search
- Value Alignment
- Scaling Laws
- Human-in-the-Loop
- Explainability

Mastery? Yes

Deep RL (DQN)

Score: **18.9**

vs.



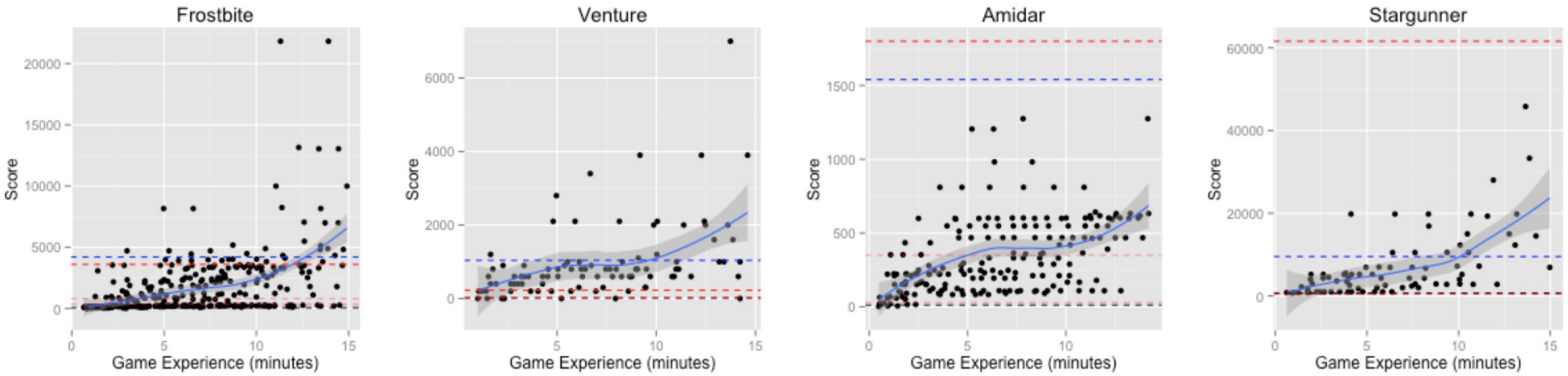
Human

Score: 9.3

How fast is the learning itself?

Humans vs. DDQN

Humans after 15 minutes tend to outperform DDQN after 115 hours



Black dots: human play

Blue curve: mean of human play

Blue dashed line: 'expert' human play

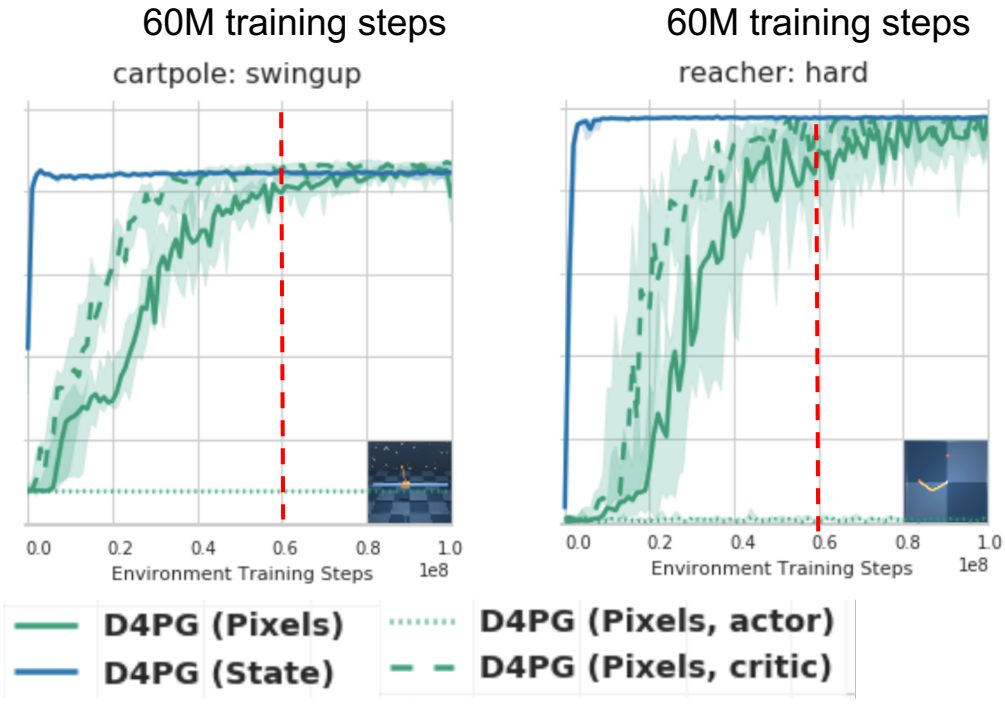
Red dashed lines:
DDQN after 10, 25, 200M frames
(~ 46, 115, 920 hours)

[Tsividis, Pouncy, Xu, Tenenbaum, Gershman, 2017]

How to bridge this gap?

Can visual RL achieve same data-efficiency as RL on state?

- State-based D4PG (blue) vs pixel-based D4PG (green)

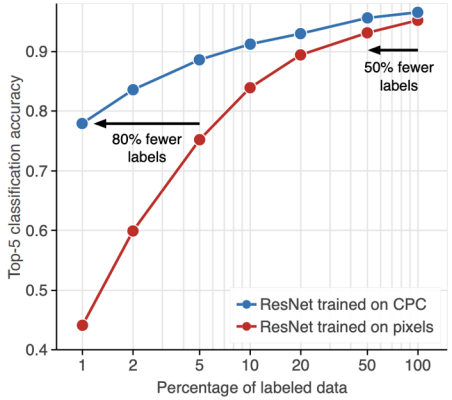


Pixel-based needs > 50M more training steps than state-based to solve same tasks

[Tassa et al., 2018] Tassa, Y., Doron, Y., Muldal, A., Erez, T., Li, Y., Casas, D.D.L., Budden, D., Abdolmaleki, A., Merel, J., Lefrancq, A. and Lillicrap, T DeepMind Control Suite, arxiv:1801.00690, 2018.

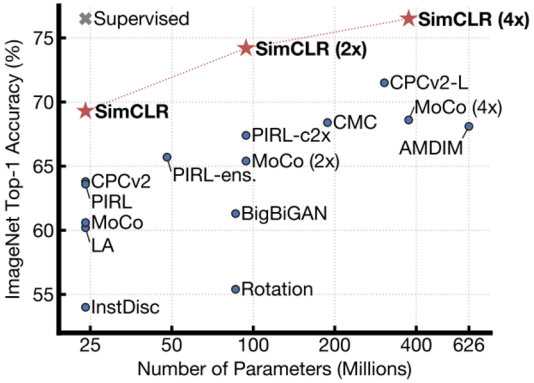
Contrastive learning: SOTA in computer vision

CPCv2 **top-5** ImageNet accuracy as function of labels



[Henaff, Srinivas et al., 2019]

SimCLR **top-1** ImageNet accuracy as function of # of parameters

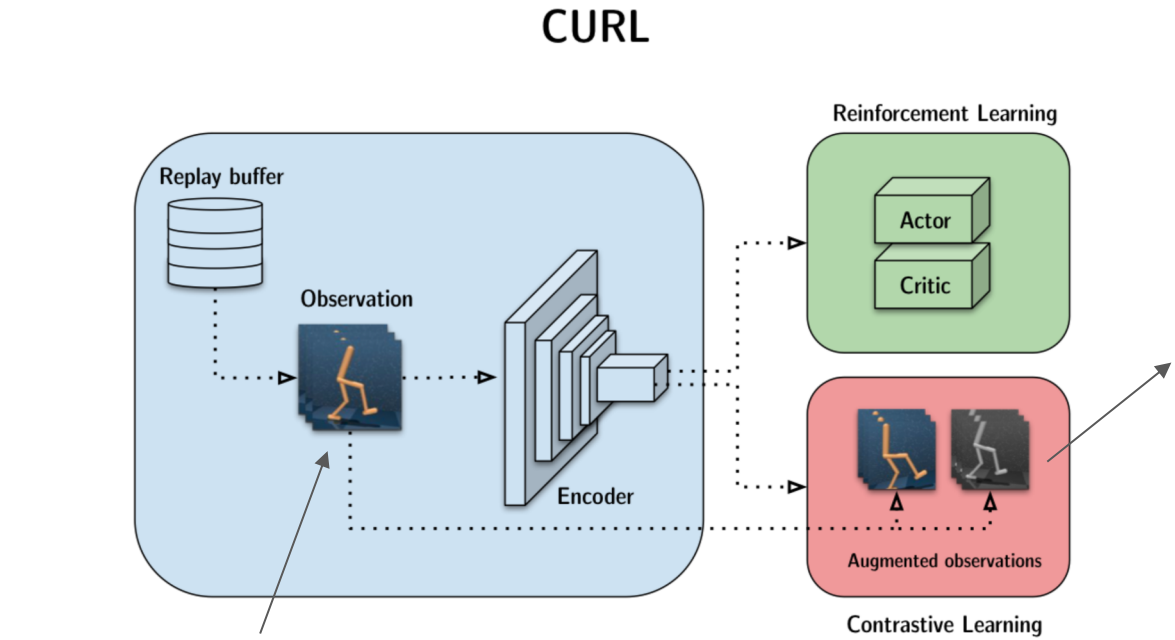


[Chen et al., 2020]

[Henaff et al., 2019] Olivier J. Hénaff, Aravind Srinivas, Jeffrey De Fauw, Ali Razavi, Carl Doersch, S. M. Ali Eslami, Aaron van den Oord Data-Efficient Image Recognition with Contrastive Coding arxiv:1905.09272, 2019.

[Chen et al., 2020] Chen, T., Kornblith, S., Norouzi, M. and Hinton, G. A Simple Framework for Contrastive Learning of Visual Representations arxiv:2002.05709, 2020.

Contrastive + RL



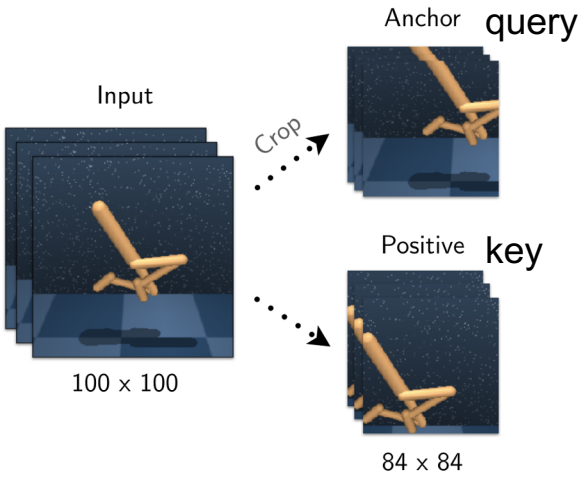
Need to define:

1. query / key pairs
2. similarity measure
3. architecture

Observations are stacked frames

[CURL: A Srinivas*, M Laskin*, P Abbeel, 2020]

1. Query / key pairs: random crop

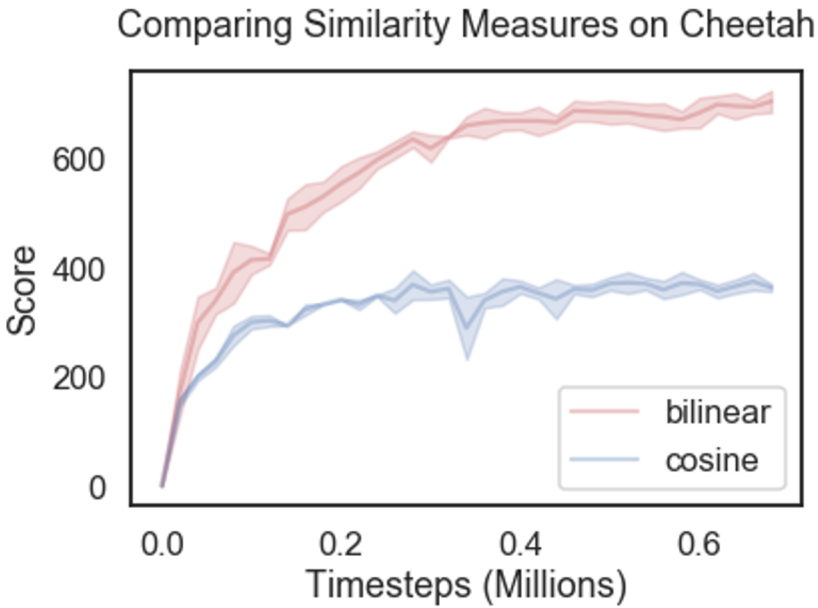


[CURL: A Srinivas*, M Laskin*, P Abbeel, 2020]

2. Bilinear inner product with learned weight matrix

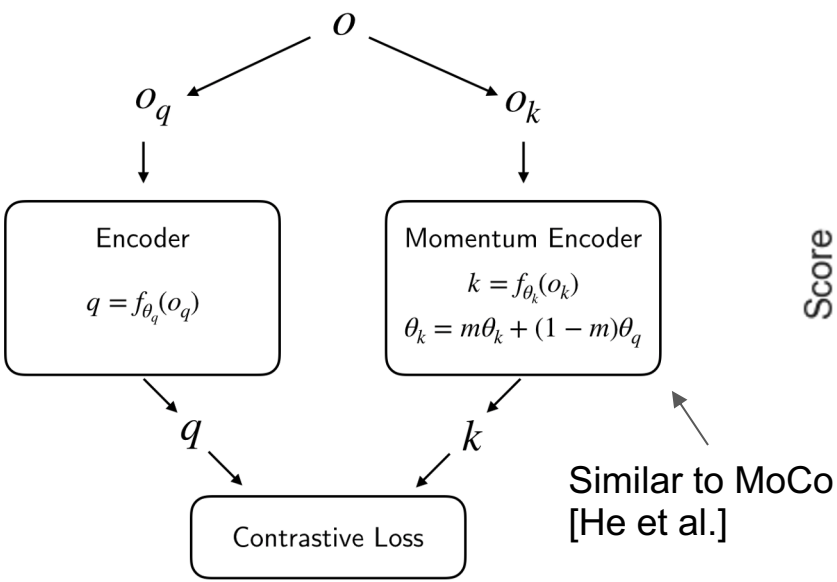
$$\begin{matrix} \text{logits} \\ \begin{bmatrix} q_0^T W k_0 & q_0^T W k_1 & \dots & q_0^T W k_j \\ q_1^T W k_0 & q_1^T W k_1 & \dots & q_1^T W k_j \\ \vdots & \vdots & \ddots & \vdots \\ q_j^T W k_0 & q_j^T W k_1 & \dots & q_k^T W k_j \end{bmatrix} \end{matrix} \quad \begin{matrix} \text{labels} \\ \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \end{matrix}$$

$$\mathscr{L}_q = \frac{\exp(q^T W k_+)}{\exp(q^T W k_+) + \sum_{i=0}^{K-1} \exp(q^T W k_i)}$$

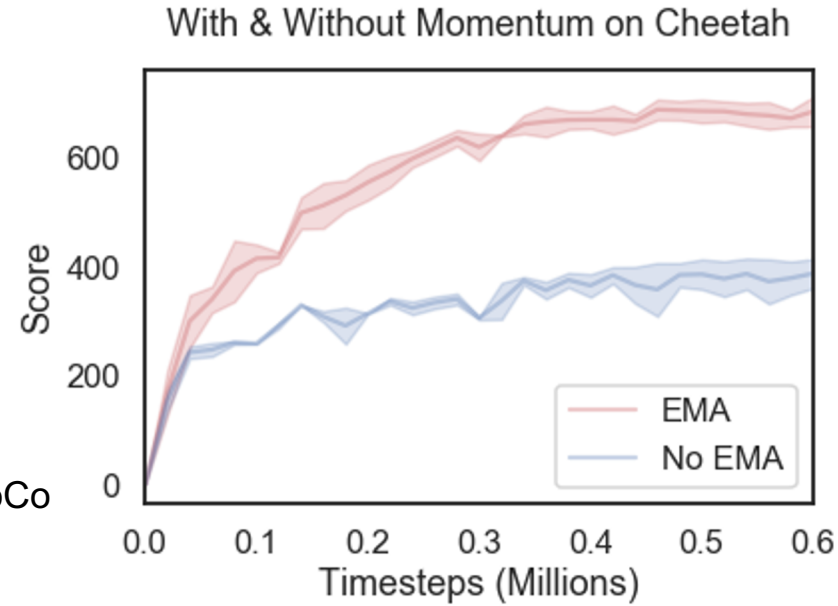


[CURL: A Srinivas*, M Laskin*, P Abbeel, 2020]

3. Keys encoded with momentum



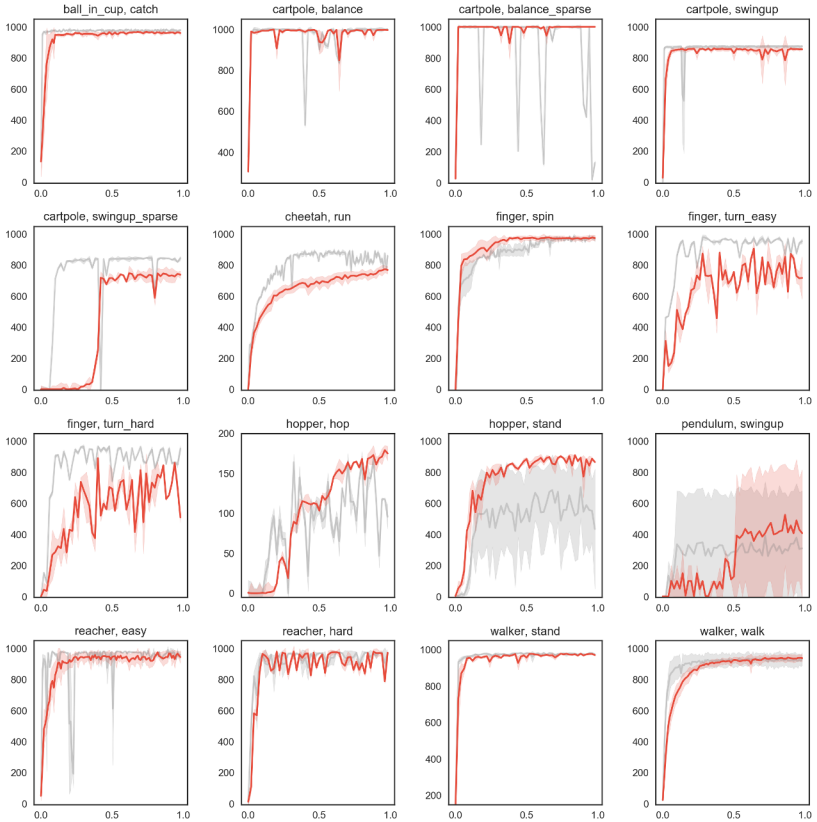
Similar to MoCo [He et al.]



[He et al., 2019] He, K., Fan, H., Wu, Y., Xie, S. and Girshick, R. Momentum Contrast for Unsupervised Visual Representation Learning arxiv:1911.05722, 2019.

CURL from pixels matches state-based SAC

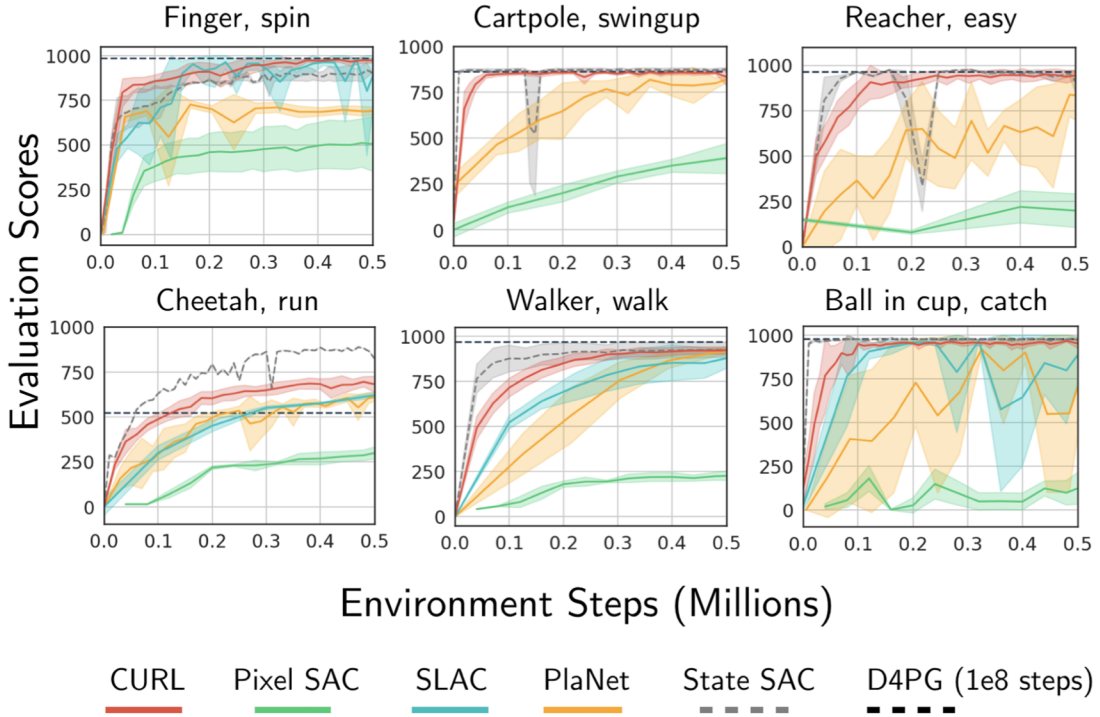
GRAY: SAC State
RED: CURL



[CURL: A Srinivas*, M Laskin*, P Abbeel, 2020]

Pieter Abbeel

CURL Comparison: DeepMind Control Suite



[CURL: A Srinivas*, M Laskin*, P Abbeel, 2020]

CURL Comparison: Atari

Atari performance benchmarked at 100K frames

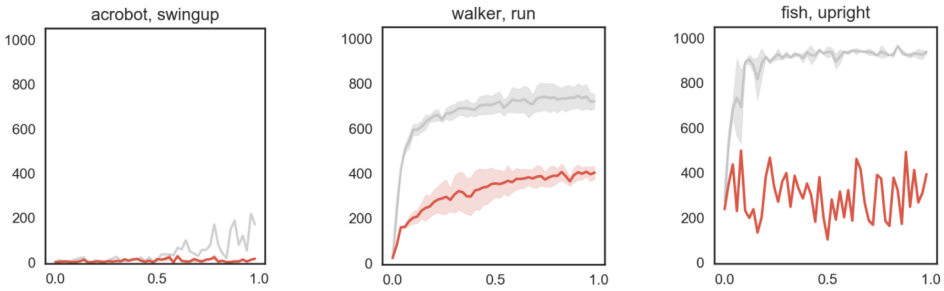
100K step scores	CURL Rainbow	SimPLe	Rainbow	Human	Random
Alien	**1148.2**	616.9	318.7	6875	184.8
Amidar	**232**	74.3	32.5	1676	11.8
Assault	473	**527.2**	231	1496	248.8
BattleZone	**11208**	4031.2	3285.71	37800	2895
Freeway	**27**	16.7	0	29.6	0
Frostbite	**924**	236.9	60.2	4335	74
Jamesbond	**400**	100.5	47.4	406.7	29.2
Qbert	**1352**	1288.8	123.46	13455	166.1
Seaquest	408	**683.3**	131.69	20182	61.1

[CURL: A Srinivas*, M Laskin*, P Abbeel, 2020]

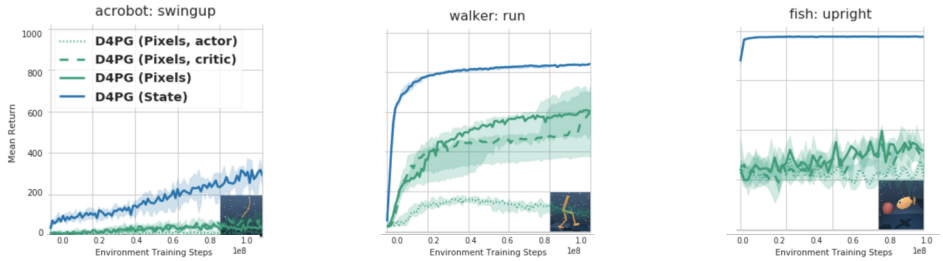
Hard Environments

GRAY: SAC State
RED: CURL

Environment training steps 1 = 1M



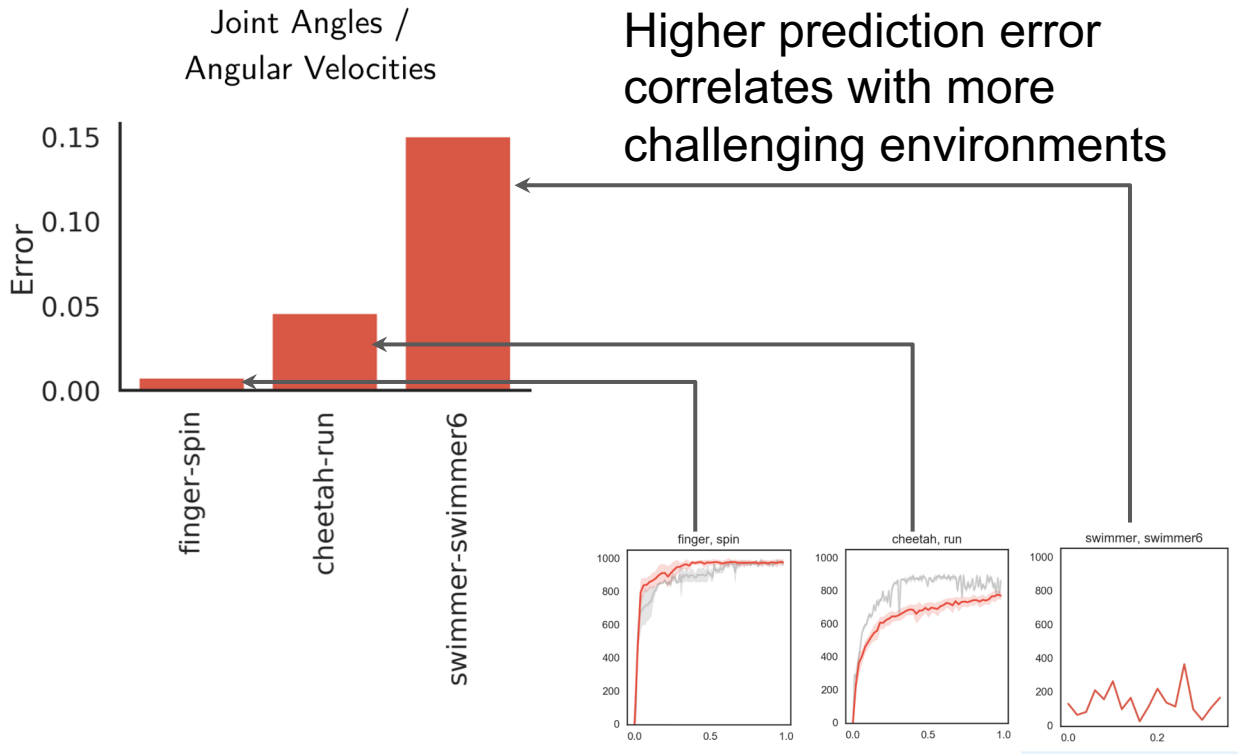
Environment training steps 1 = 100M



Observation:
similar struggle asymptotically for pixel-based Deep RL

[CURL: A Srinivas*, M Laskin*, P Abbeel, 2020]

Predicting state from pixels



[CURL: A Srinivas*, M Laskin*, P Abbeel, 2020]

Many Exciting Directions in AI

- Unsupervised Learning
- Reinforcement Learning
- Unsupervised RL
- ***Meta-Reinforcement Learning***
- Few-Shot Imitation
- Domain Randomization
- DL for Science and Engineering

- Mitigating Bias
- Multi-modal Learning
- Architecture Search
- Value Alignment
- Scaling Laws
- Human-in-the-Loop
- Explainability

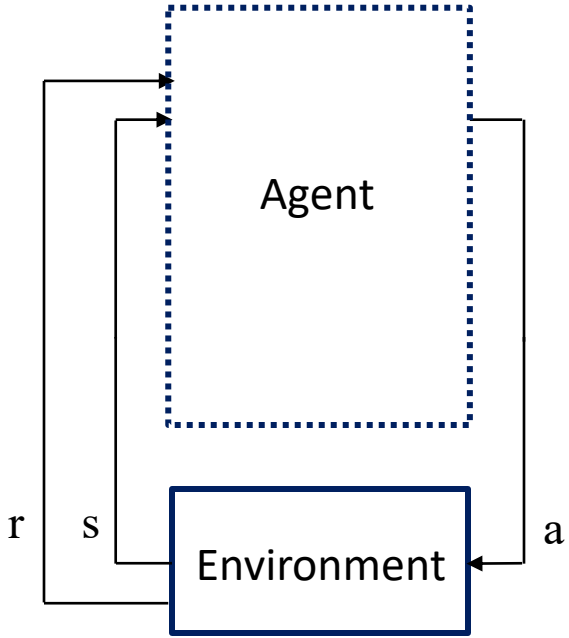
Starting Observations

- TRPO, DQN, A3C, DDPG, PPO, Rainbow, … are fully general RL algorithms
  - i.e., for any environment that can be mathematically defined, these algorithms are equally applicable

- Environments encountered in real world
  = tiny, tiny subset of all environments that could be defined
  (e.g. they all satisfy our universe's physics)

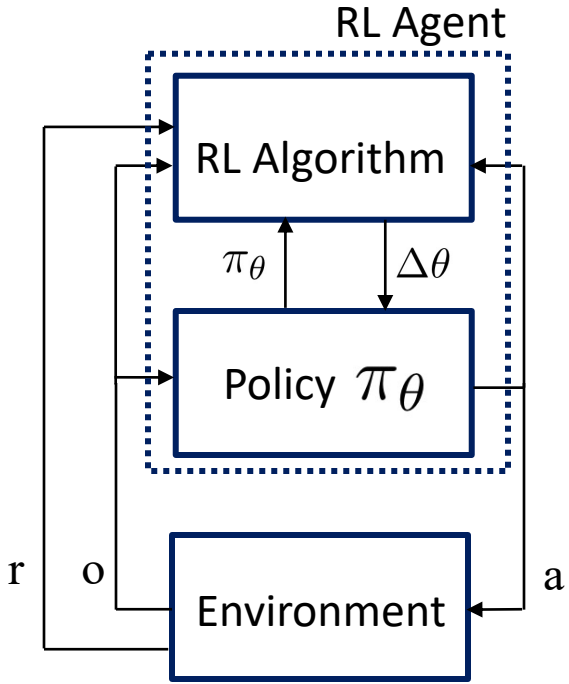
Can we develop "fast" RL algorithms that take advantage of this?

Reinforcement Learning



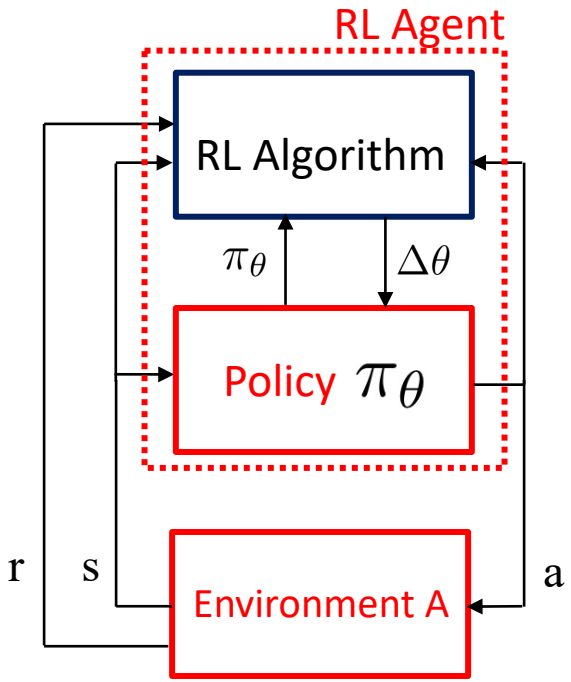
[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016] also: [Wang et al, 2016]

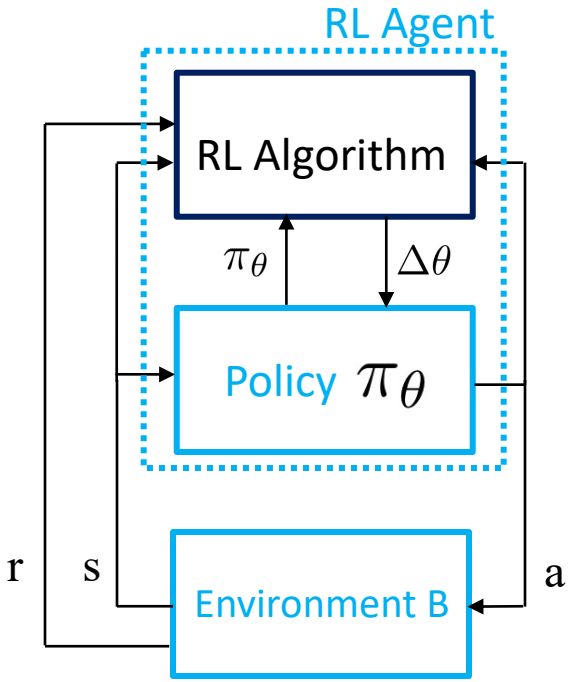
Reinforcement Learning



[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016] also: [Wang et al, 2016]

Reinforcement Learning

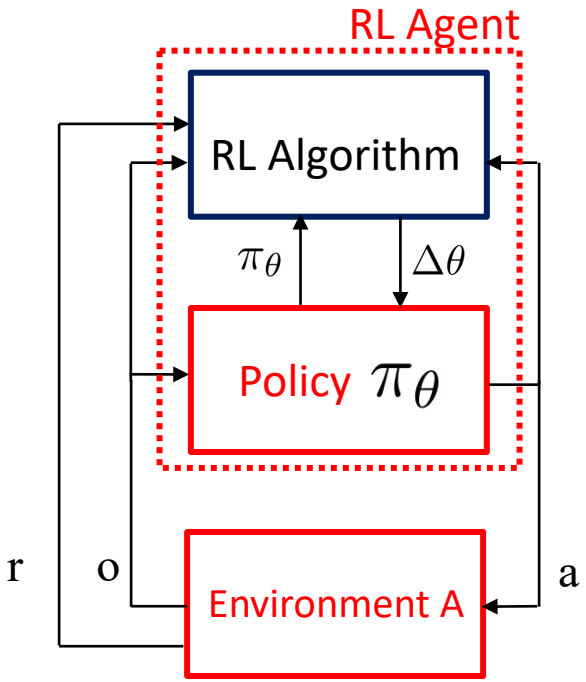


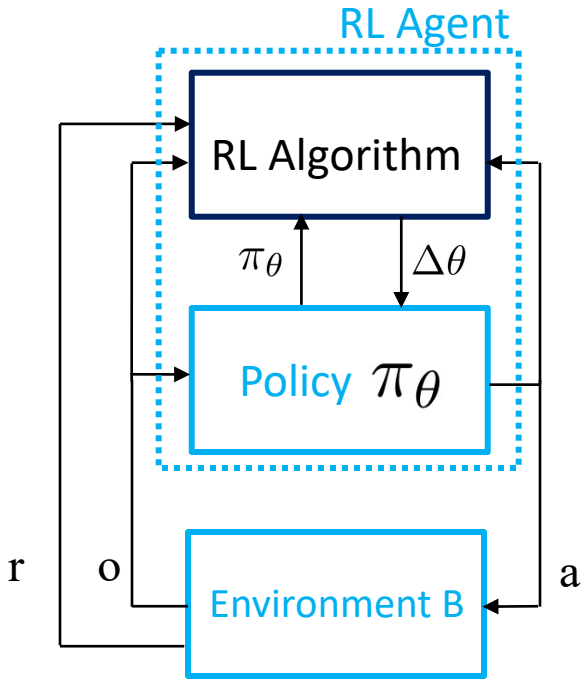


...

[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016] also: [Wang et al, 2016]

Reinforcement Learning





Traditional RL research:

- Human experts develop the RL algorithm
- After many years, still no RL algorithms nearly as good as humans…

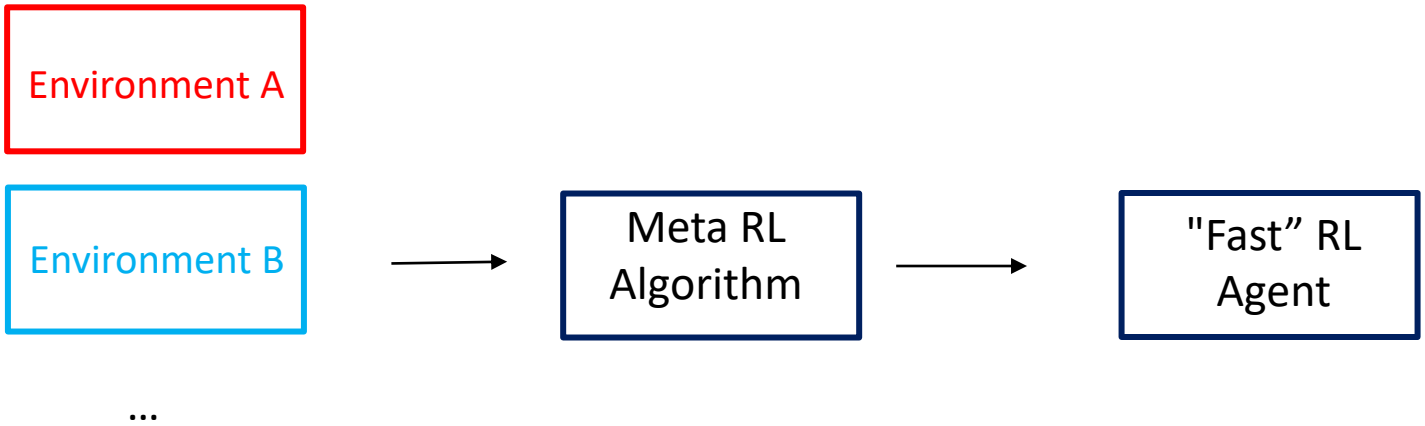
Alternative:

- Could we learn a better RL algorithm?
- Or even learn a better entire agent?

[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016] also: [Wang et al, 2016]

Meta-Reinforcement Learning

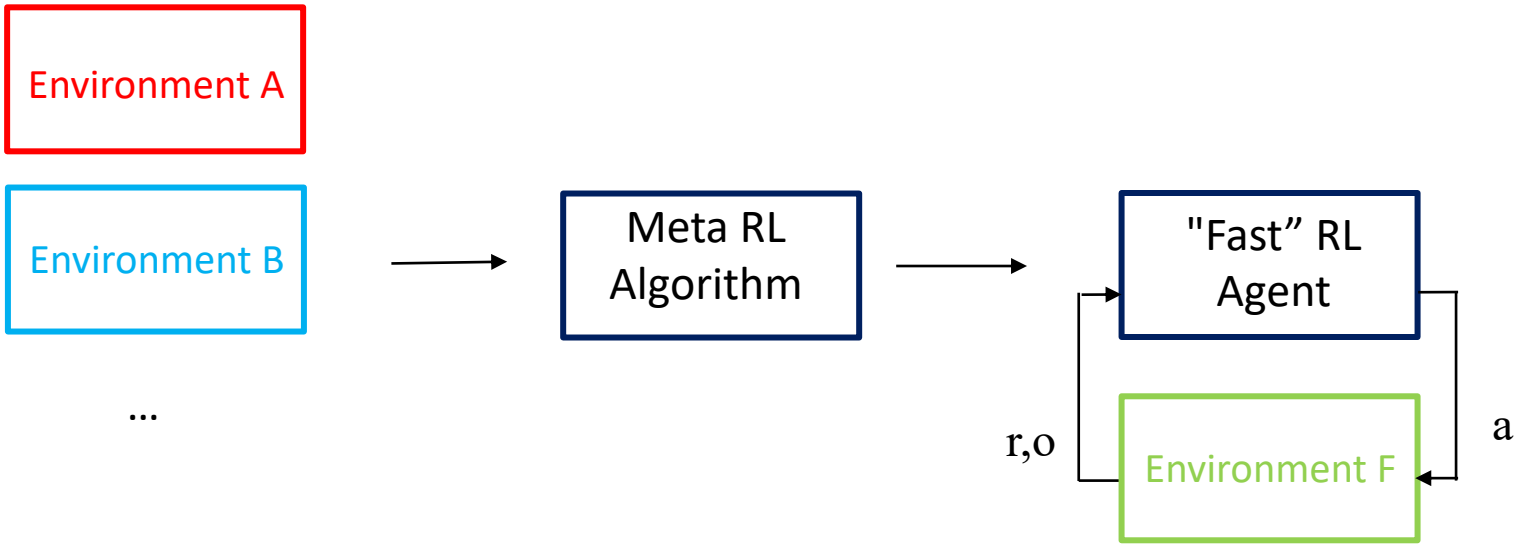
Meta-training environments



[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016] also: [Wang et al, 2016]

Meta-Reinforcement Learning

Meta-training environments

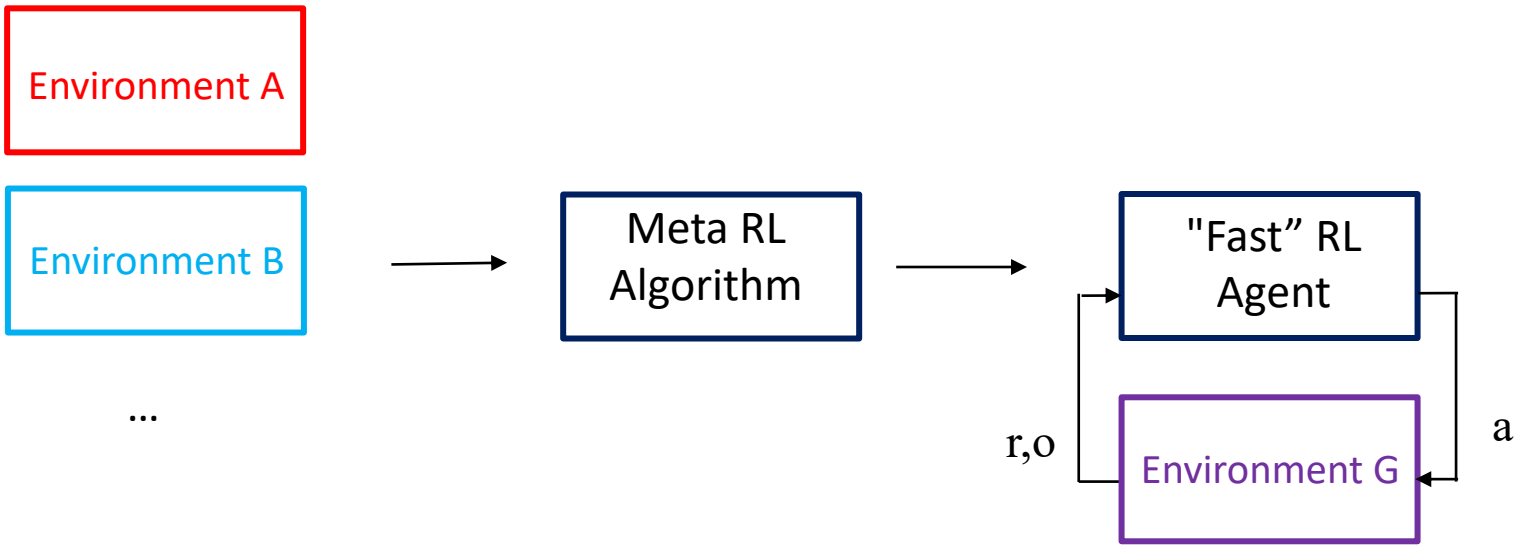


Testing environments

[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016] also: [Wang et al, 2016]

Meta-Reinforcement Learning

Meta-training environments

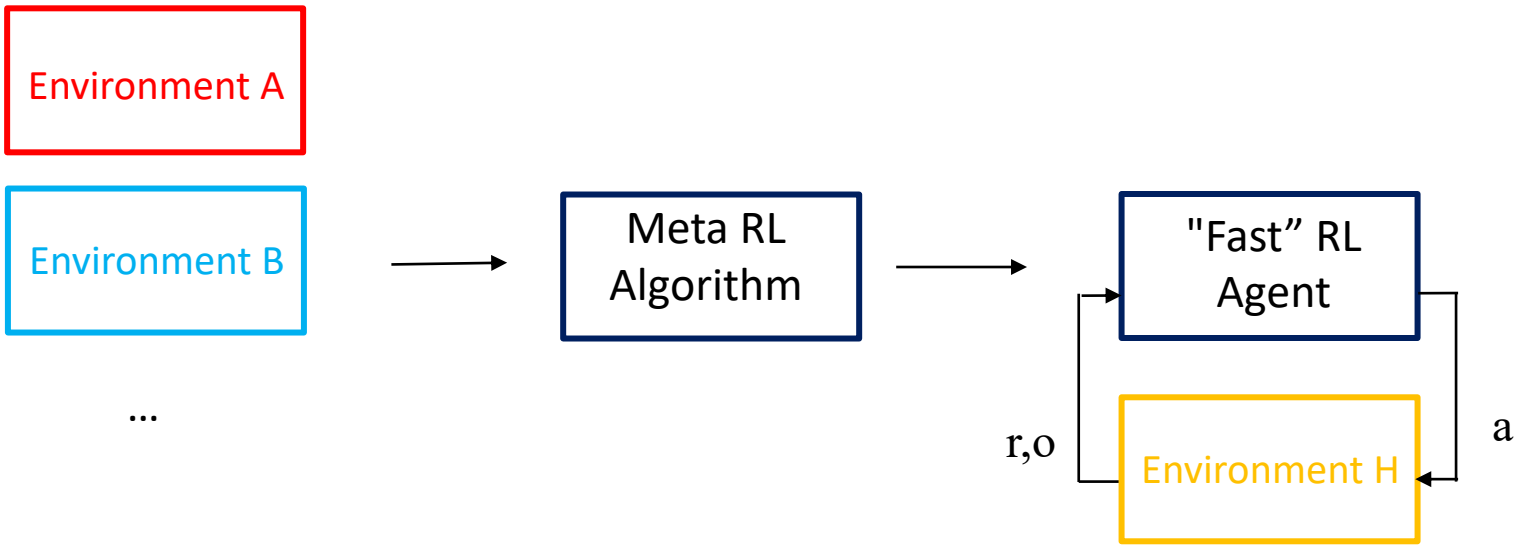


Testing environments

[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016] also: [Wang et al, 2016]

Meta-Reinforcement Learning

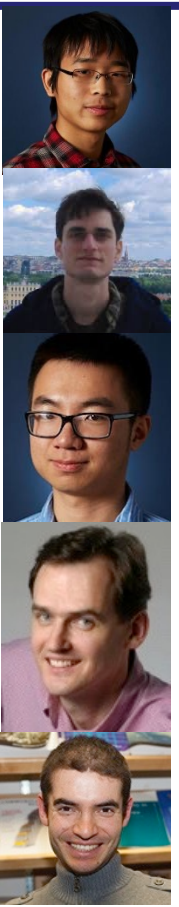
Meta-training environments



Testing environments

[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016] also: [Wang et al, 2016]

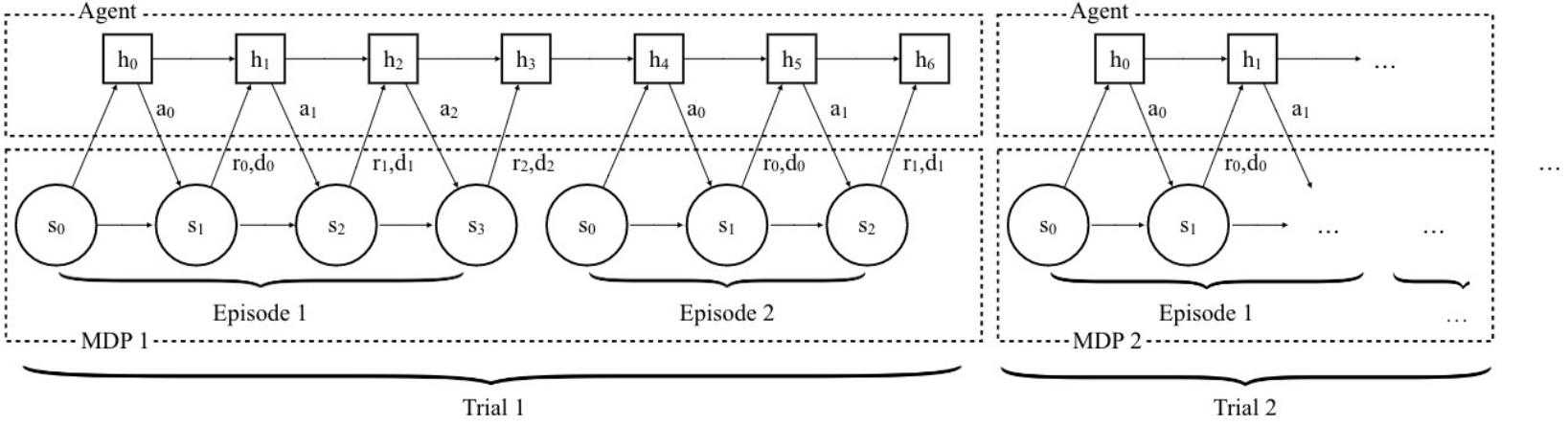
Formalizing Learning to Reinforcement Learn



$$\max_{\theta} \mathbb{E}_{M} \mathbb{E}_{\tau_M^{(k)}} \left[\sum_{k=1}^{K} R(\tau_M^{(k)}) \mid \text{RLagent}_{\theta} \right]$$

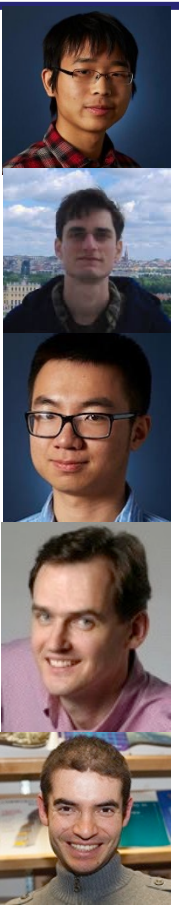
M : sample environment

$\tau_M^{(k)}$: k'th episode in environment M



[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016] also: [Wang et al, 2016]

Formalizing Learning to Reinforcement Learn



$$\max_{\theta} \mathbb{E}_{M} \mathbb{E}_{\tau_M^{(k)}} \left[\sum_{k=1}^{K} R(\tau_M^{(k)}) \mid \text{RLagent}_{\theta} \right]$$

M : sample MDP

$\tau_M^{(k)}$: k'th trajectory in MDP M

Meta-train:

$$\max_{\theta} \sum_{M \in M_{\text{train}}} \mathbb{E}_{\tau_M^{(k)}} \left[\sum_{k=1}^{K} R(\tau_M^{(k)}) \mid \text{RLagent}_{\theta} \right]$$

[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016] also: [Wang et al, 2016]

Representing $\mathrm{RLagent}_\theta$

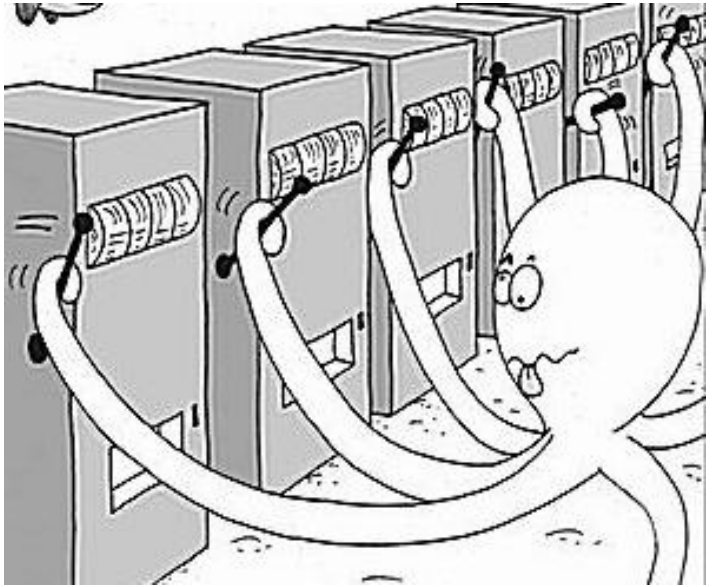
$$\max_\theta \sum_{M \in M_{\text{train}}} \mathbb{E}_{\tau_M^{(k)}} \left[\sum_{k=1}^{K} R(\tau_M^{(k)}) \mid \mathrm{RLagent}_\theta \right]$$

- RLagent = RNN = generic computation architecture
  - different weights in the RNN means different RL algorithm and prior
  - different activations in the RNN means different current policy
  - meta-train objective can be optimized with an existing (slow) RL algorithm

[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016] also: [Wang et al, 2016]

Evaluation: Multi-Armed Bandits

- Multi-Armed Bandits setting
  - Each bandit has its own distribution over pay-outs
  - Each episode = choose 1 bandit
  - Good RL agent should explore bandits sufficiently, yet also exploit the good/best ones
- Provably (asymptotically) optimal RL algorithms have been invented by humans: Gittins index, UCB1, Thompson sampling, …



[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016]

Evaluation: Multi-Armed Bandits

Setup	Random	Gittins	TS	OTS	UCB1	ϵ-Greedy	Greedy	RL^2
$n = 10, k = 5$	5.0	**6.6**	5.7	6.5	**6.7**	**6.6**	**6.6**	**6.7**
$n = 10, k = 10$	5.0	**6.6**	5.5	6.2	**6.7**	**6.6**	**6.6**	**6.7**
$n = 10, k = 50$	5.1	6.5	5.2	5.5	**6.6**	6.5	6.5	**6.8**
$n = 100, k = 5$	49.9	**78.3**	74.7	**77.9**	**78.0**	75.4	74.8	**78.7**
$n = 100, k = 10$	49.9	**82.8**	76.7	81.4	82.4	77.4	77.1	**83.5**
$n = 100, k = 50$	49.8	**85.2**	64.5	67.7	84.3	78.3	78.0	**84.9**
$n = 500, k = 5$	249.8	**405.8**	**402.0**	**406.7**	**405.8**	388.2	380.6	**401.6**
$n = 500, k = 10$	249.0	**437.8**	429.5	**438.9**	**437.1**	408.0	395.0	432.5
$n = 500, k = 50$	249.6	**463.7**	427.2	437.6	457.6	413.6	402.8	438.9

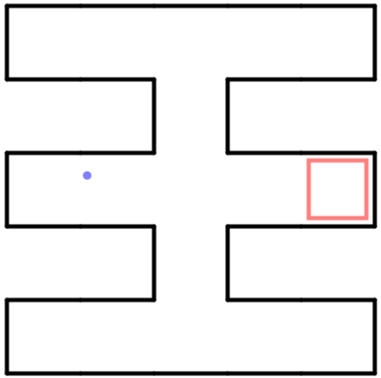
We consider Bayesian evaluation setting. Some of these prior works also have adversarial guarantees, which we don't consider here.

[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016]

Evaluation: Visual Navigation

Agent input: current image
Agent action: straight / 2 degrees left / 2 degrees right
Map just shown for our purposes, but not available to agent

Agent's view

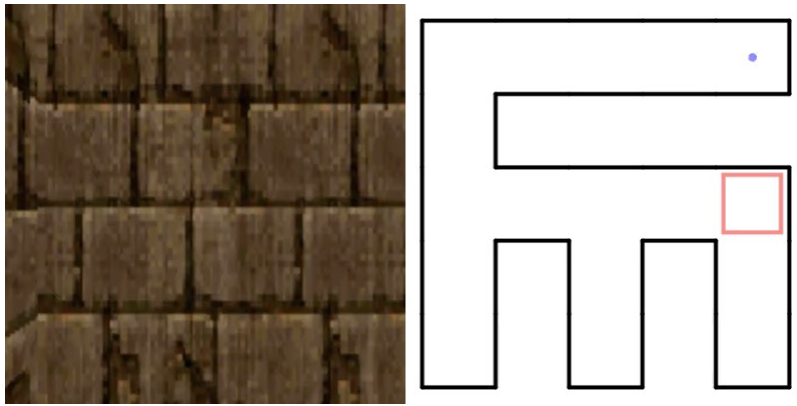


Maze

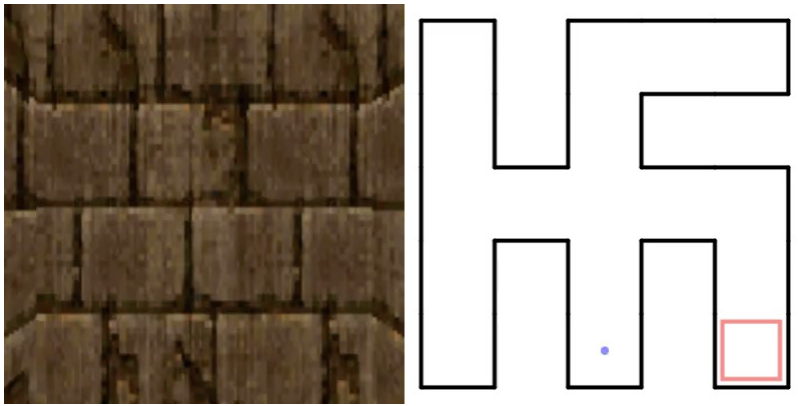
Related work: Mirowski, et al, 2016; Jaderberg et al, 2016; Mnih et al, 2016; Wang et al, 2016

Evaluation: Visual Navigation

Agent input: current image
Agent action: straight / 2 degrees left / 2 degrees right
Map just shown for our purposes, but not available to agent



Before learning-to-learn



After learning-to-learn

[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016]

Meta-Learning Curves

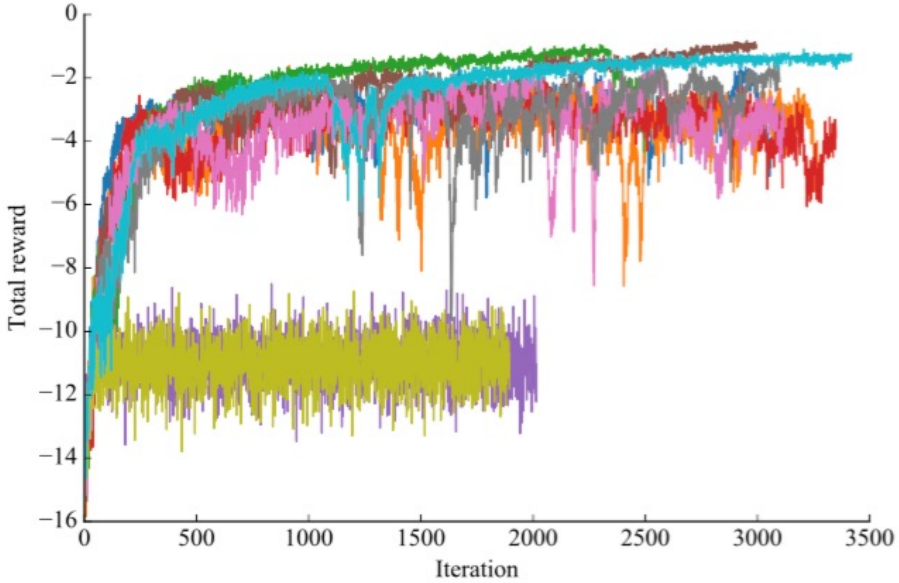


Figure 5: RL^2 learning curves for visual navigation. Each curve shows a different random initialization of the RNN weights. Performance varies greatly across different initializations.

[Duan, Schulman, Chen, Bartlett, Sutskever, Abbeel, 2016]

Meta Learning for RL

Task distribution: different environments

- Schmidhuber. Evolutionary principles in self-referential learning. (1987)
- Wiering, Schmidhuber. Solving POMDPs with Levin search and EIRA. (1996)
- Schmidhuber, Zhao, Wiering. Shifting inductive bias with success-story algorithm, adaptive Levin search, and incremental self-improvement. (MLJ 1997)
- Schmidhuber, Zhao, Schraudolph. Reinforcement learning with self-modifying policies (1998)
- Zhao, Schmidhuber. Solving a complex prisoner's dilemma with self-modifying policies. (1998)
- Schmidhuber. A general method for incremental self-improvement and multiagent learning. (1999)
- Singh, Lewis, Barto. Where do rewards come from? (2009)
- Singh, Lewis, Barto. Intrinsically Motivated Reinforcement Learning: An Evolutionary Perspective (2010)
- Niekum, Spector, Barto. Evolution of reward functions for reinforcement learning (2011)
- Duan et al., (2016) RL2: Fast Reinforcement Learning via Slow Reinforcement Learning
- Wang et al., (2016) Learning to Reinforcement Learn
- Finn et al., (2017) Model-Agnostic Meta-Learning (MAML)
- Mishra, Rohinenjad et al., (2017) Simple Neural AttentIve meta-Learner
- Frans et al., (2017) Meta-Learning Shared Hierarchies

Meta Learning for RL

Task distribution: different environments

- Schmidhuber. Evolutionary principles in self-referential learning. (1987)
- Wiering, Schmidhuber. Solving POMDPs with Levin search and EIRA. (1996)
- Schmidhuber, Zhao, Wiering. Shifting inductive bias with success-story algorithm, adaptive Levin search, and incremental self-improvement. (MLJ 1997)
- Schmidhuber, Zhao, Schraudolph. Reinforcement learning with self-modifying policies (1998)
- Zhao, Schmidhuber. Solving a complex prisoner's dilemma with self-modifying policies. (1998)
- Schmidhuber. A general method for incremental self-improvement and multiagent learning. (1999)
- Singh, Lewis, Barto. Where do rewards come from? (2009)
- Singh, Lewis, Barto. Intrinsically Motivated Reinforcement Learning: An Evolutionary Perspective (2010)
- Niekum, Spector, Barto. Evolution of reward functions for reinforcement learning (2011)
- Duan et al., (2016) RL2: Fast Reinforcement Learning via Slow Reinforcement Learning
- Wang et al., (2016) Learning to Reinforcement Learn
- ***Finn et al., (2017) Model-Agnostic Meta-Learning (MAML)***
- Mishra, Rohinenjad et al., (2017) Simple Neural AttentIve meta-Learner
- Frans et al., (2017) Meta-Learning Shared Hierarchies

Many Exciting Directions in AI

- Unsupervised Learning
- Reinforcement Learning
- Unsupervised RL
- Meta-Reinforcement Learning
- ***Few-Shot Imitation***
- Domain Randomization
- DL for Science and Engineering

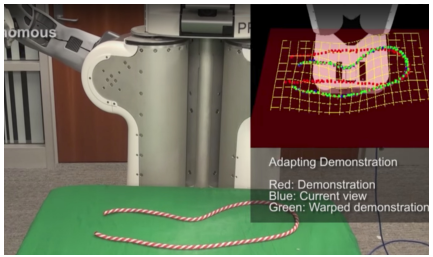
- Mitigating Bias
- Multi-modal Learning
- Architecture Search
- Value Alignment
- Scaling Laws
- Human-in-the-Loop
- Explainability

Imitation Learning in Robotics

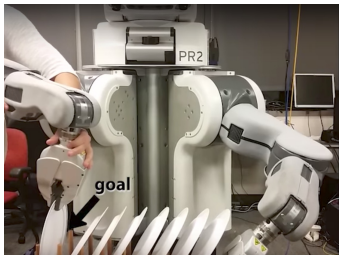
[Abbeel et al. 2008]

[Kolter et al. 2008]

[Ziebart et al. 2008]

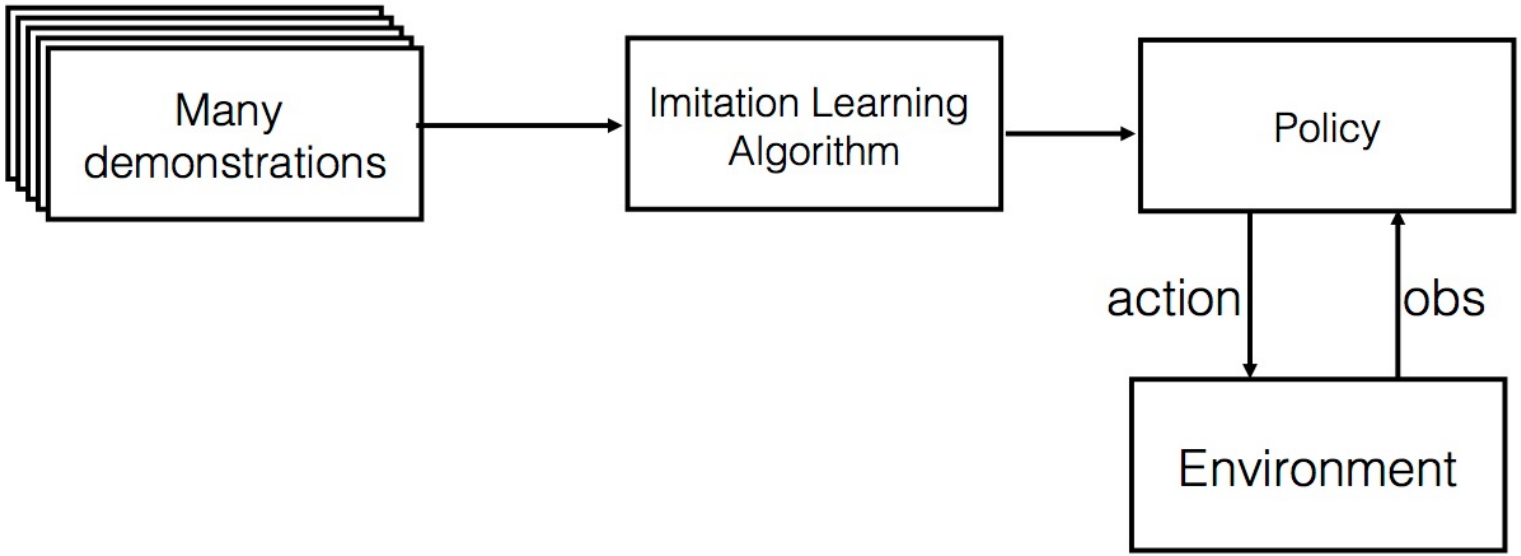


[Schulman et al. 2013]

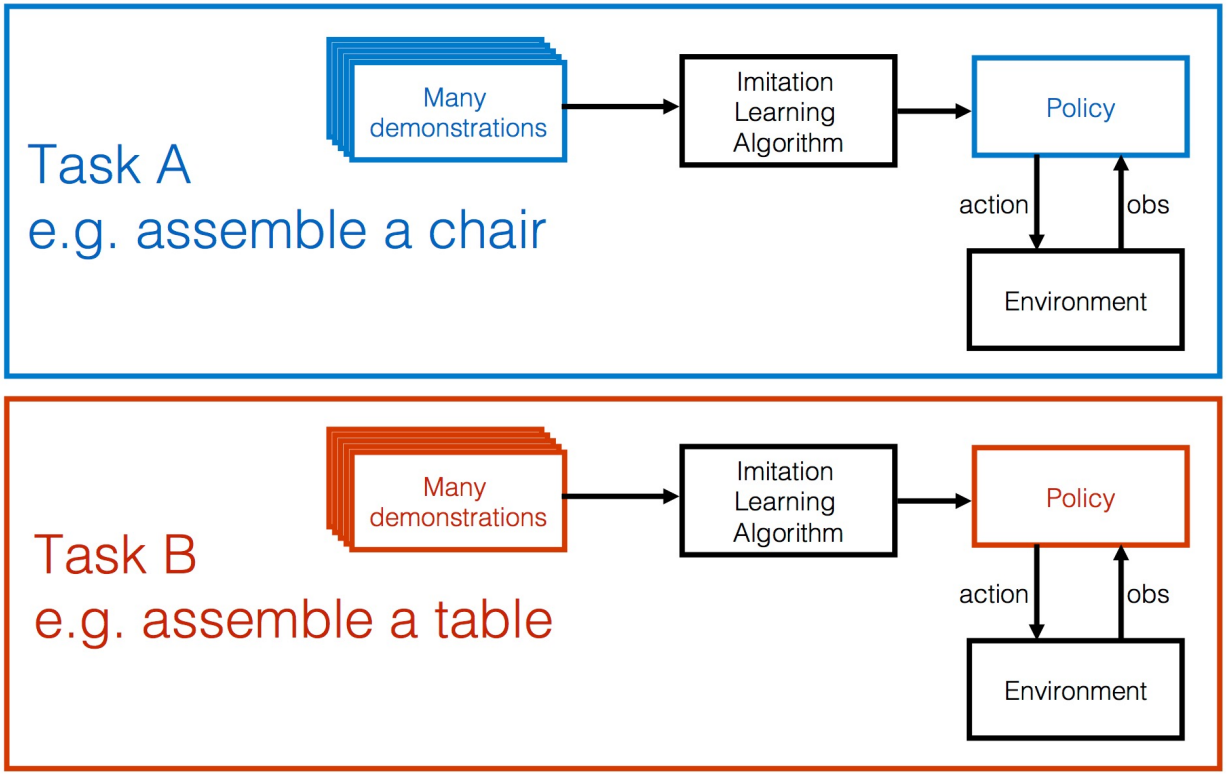


[Finn et al. 2016]

Imitation Learning

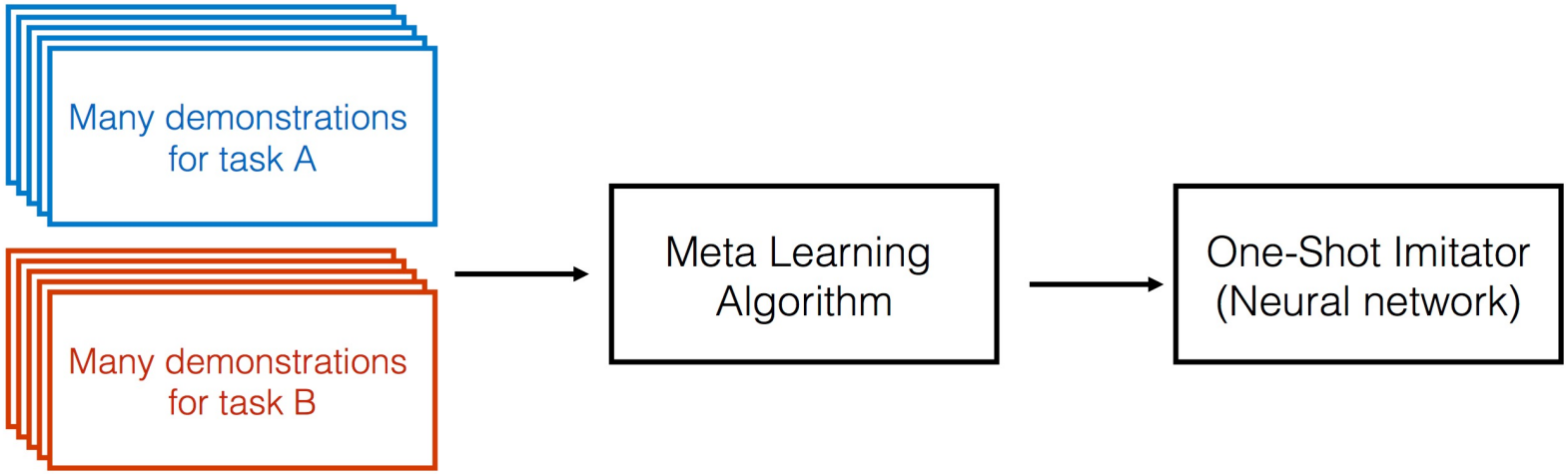


Imitation Learning



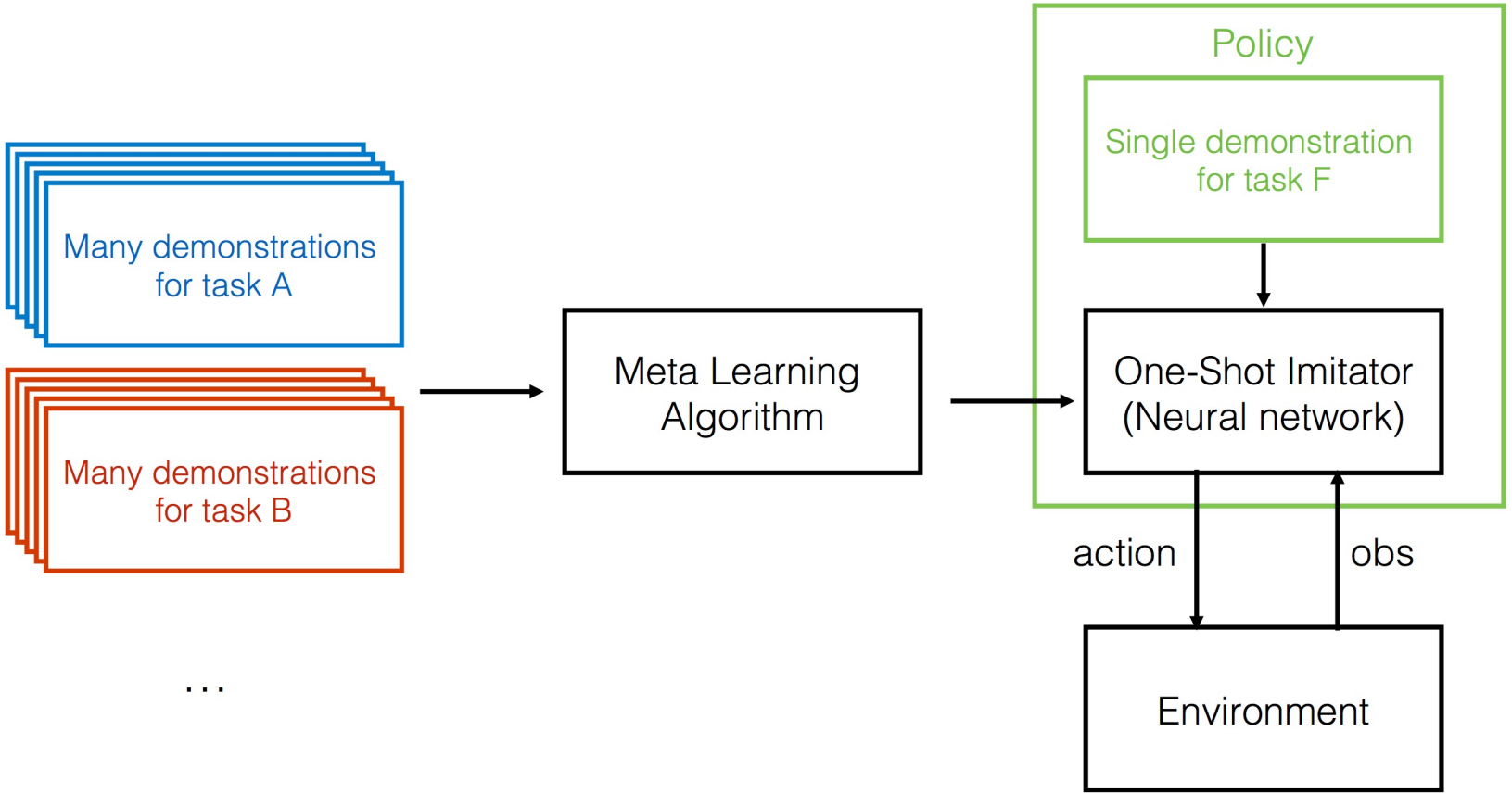
...

One-Shot Imitation Learning



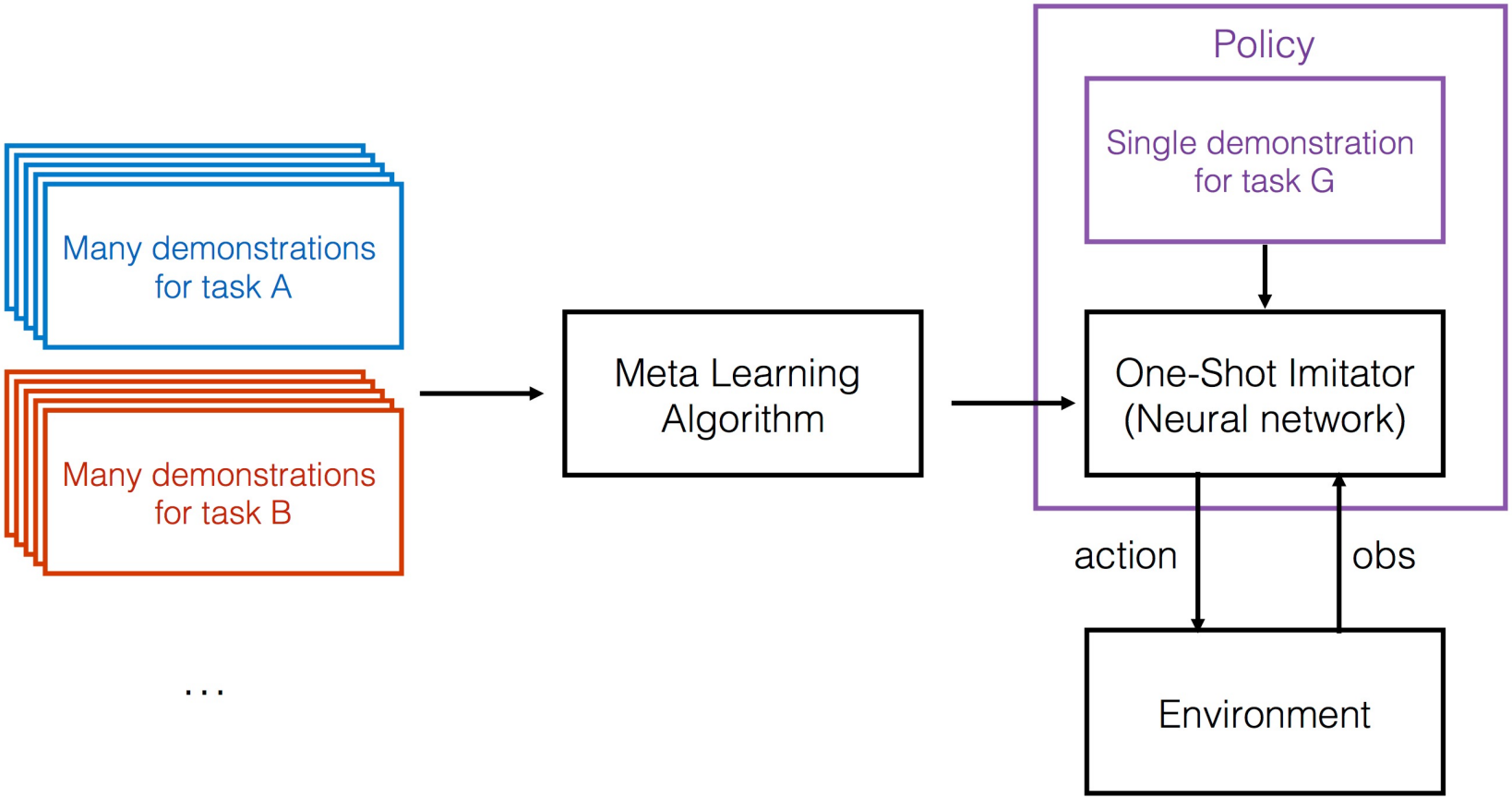
[Duan, Andrychowicz, Stadie, Ho, Schneider, Sutskever, Abbeel, Zaremba, 2017]

One-Shot Imitation Learning



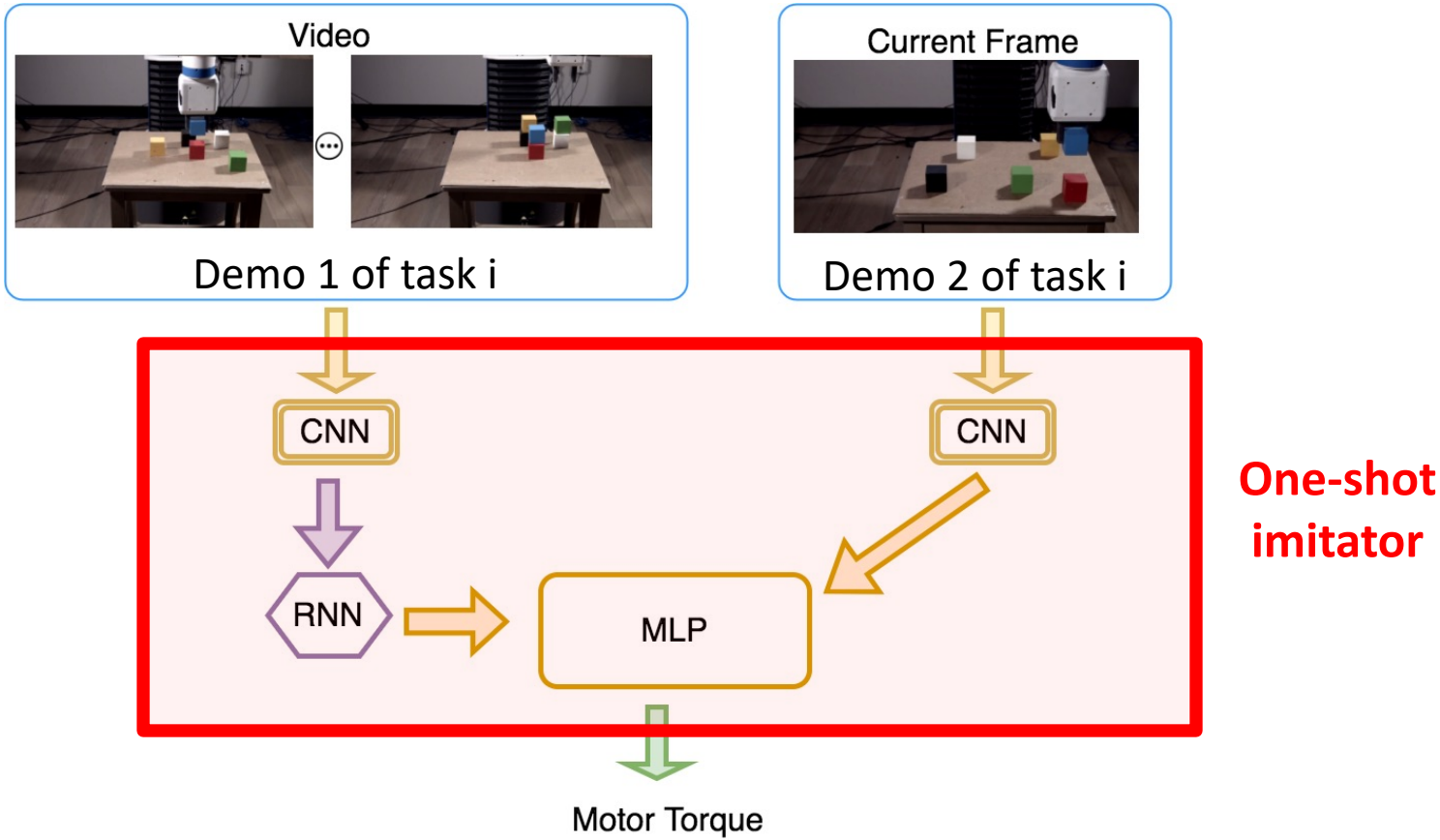
[Duan, Andrychowicz, Stadie, Ho, Schneider, Sutskever, Abbeel, Zaremba, 2017]

One-Shot Imitation Learning



[Duan, Andrychowicz, Stadie, Ho, Schneider, Sutskever, Abbeel, Zaremba, 2017]

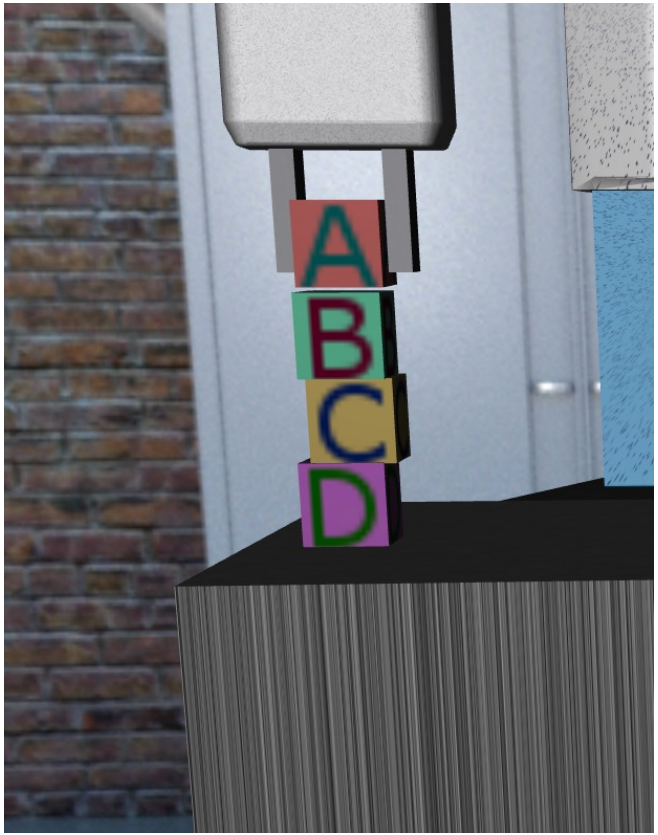
Learning a One-Shot Imitator



[Figure credit: Bradly Stadie]

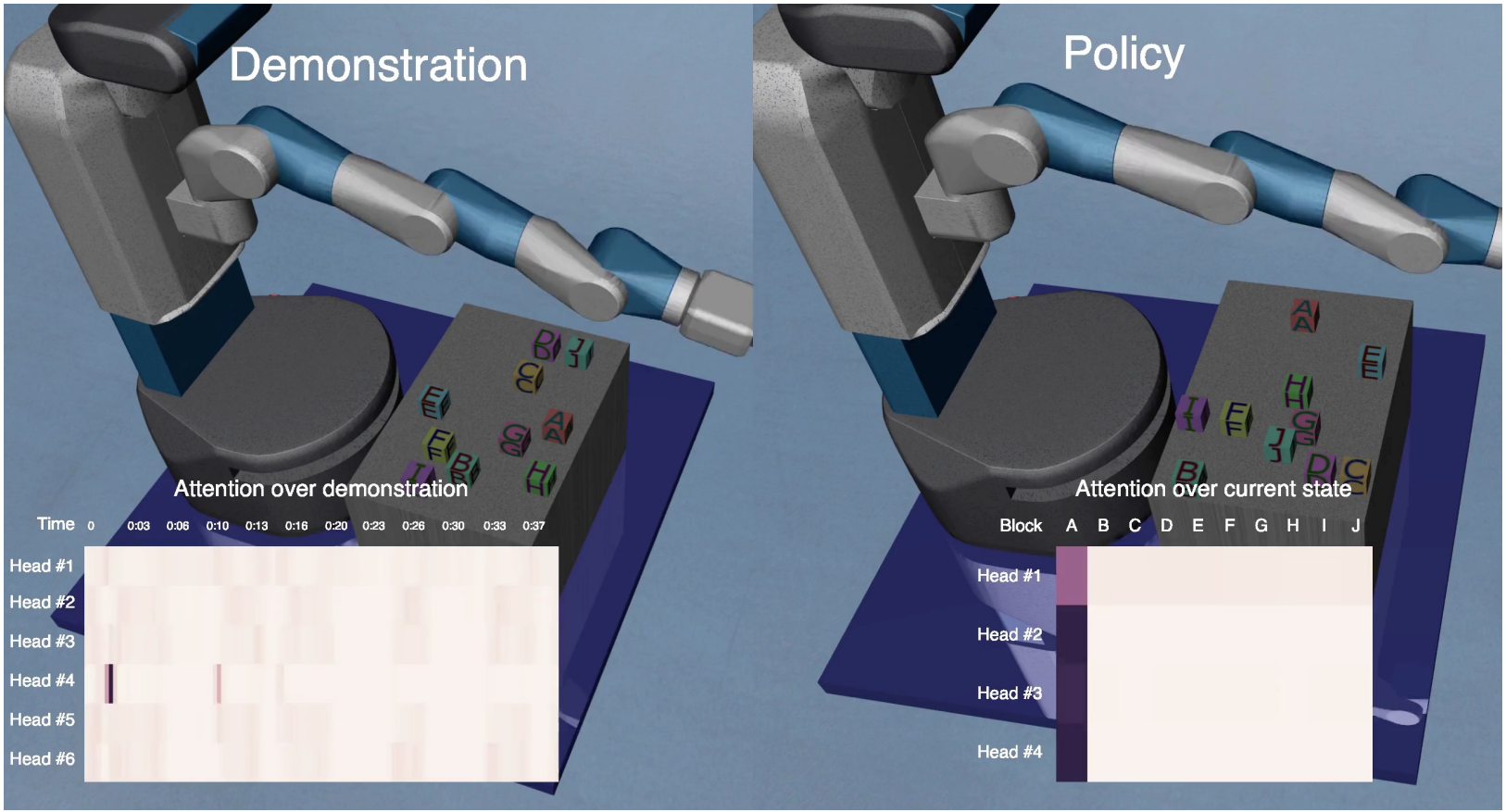
Proof-of-concept: Block Stacking

- Each task is specified by a desired final layout
  - Example: abcd
    - “Place c on top of d, place b on top of c, place a on top of b.”



[Duan, Andrychowicz, Stadie, Ho, Schneider, Sutskever, Abbeel, Zaremba, 2017]

Evaluation



[Duan, Andrychowicz, Stadie, Ho, Schneider, Sutskever, Abbeel, Zaremba, 2017]

Many Exciting Directions in AI

- Unsupervised Learning
- Reinforcement Learning
- Unsupervised RL
- Meta-Reinforcement Learning
- Few-Shot Imitation
- ***Domain Randomization***
- DL for Science and Engineering

- Mitigating Bias
- Multi-modal Learning
- Architecture Search
- Value Alignment
- Scaling Laws
- Human-in-the-Loop
- Explainability

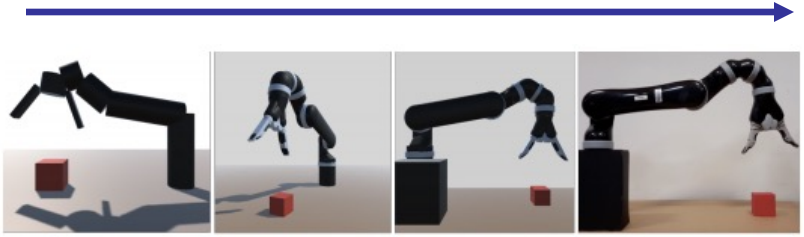
Motivation for Simulation

Compared to the real world, simulated data collection is…

- Less expensive
- Faster / more scalable
- Less dangerous
- Easier to label

How can we learn useful real-world skills in the simulator?

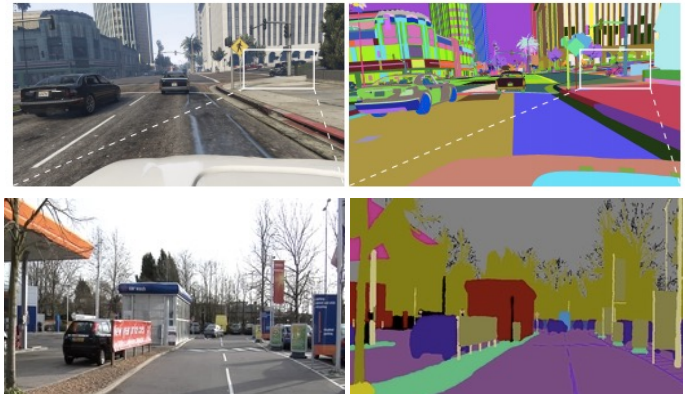
Approach 1 – Use Realistic Simulated Data



Simulation

Real world

Carefully match the simulation to the world [1,2,3,4]



GTA V

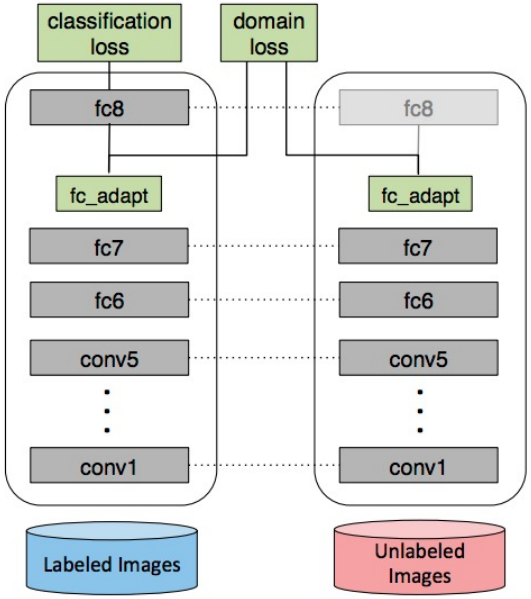
Real world

Augment simulated data with real data [5,6]

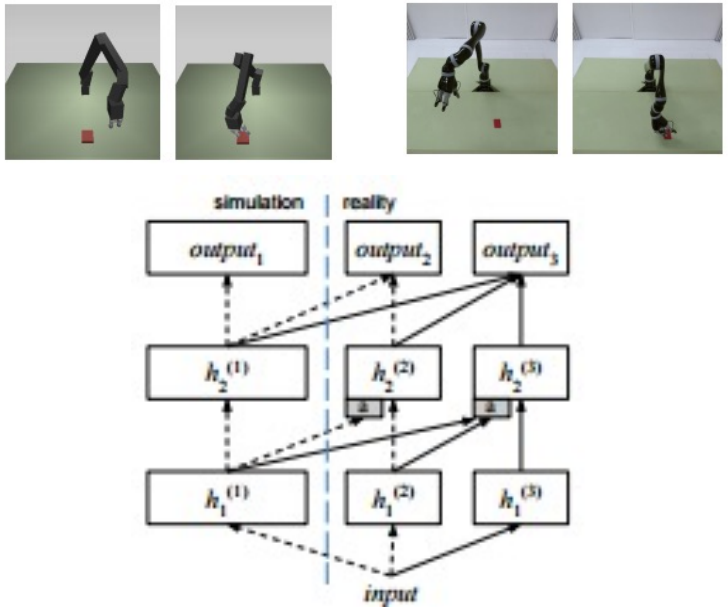
[1] Stephen James, Edward Johns. *3d simulation for robot arm control with deep q-learning (2016)*
[2] Johns, Leutenegger, Davision. Deep learning a grasp function for grasping under gripper pose uncertainty (2016)
[3] Mahler et al, Dex-Net 3.0 (2017)
[4] Koenemann et al. Whole-body model-predictive control applied to the HRP-2 humanoid. (2015)

[5] Stephan R Richter, Vibhav Vineet, Stefan Roth, and Vladlen Koltun. *Playing for data: Ground truth from computer games (2016)*
[6] Bousmalis et al. Using simulation and domain adaptation to improve efficiency of robotic grasping (2017)

Approach 2 – Domain Confusion / Adaptation

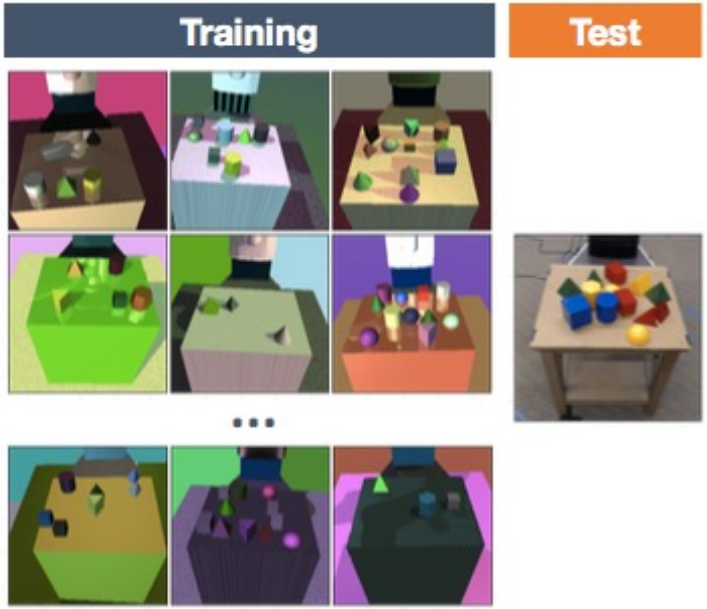


Eric Tzeng, Judy Hoffman, Ning Zhang, Kate Saenko, Trevor Darrell. Deep Domain Confusion: Maximizing for Domain Invariance. *arXiv preprint arXiv:1412.3474, 2014*.



Andrei A Rusu, Matej Vecerik, Thomas Rotho¨rl, Nicolas Heess, Razvan Pascanu, and Raia Hadsell. Sim-to-real robot learning from pixels with progressive nets. *arXiv preprint arXiv:1610.04286*, 2016.

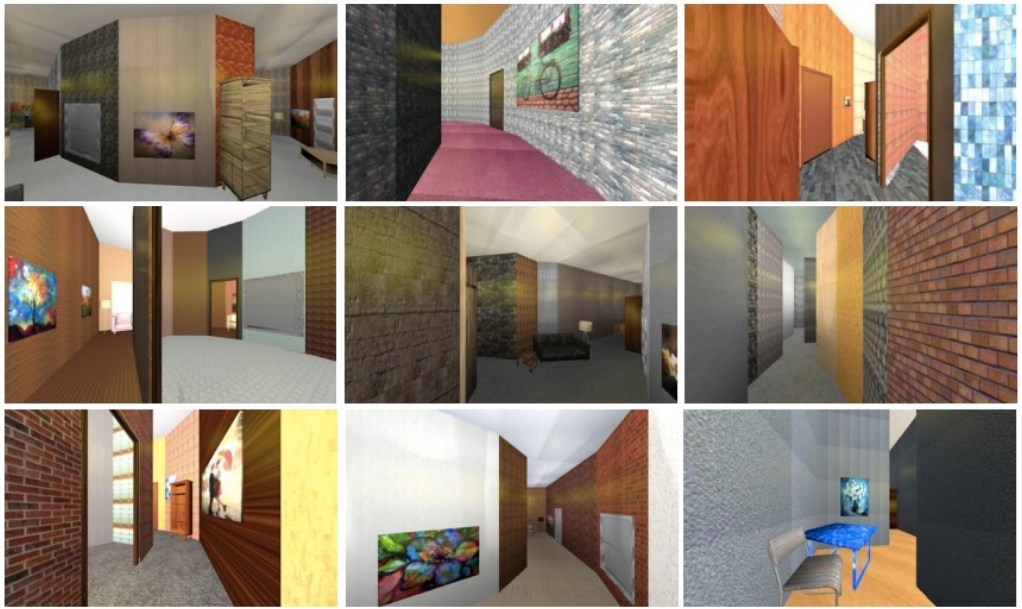
Approach 3 – Domain Randomization



If the model sees enough simulated variation, the real world may look like just the next simulator

Domain Randomization

(cad)^2 rl: Real Single-Image Flight Without a Single Real Image.



- Quadcopter collision avoidance
- ~500 semi-realistic textures, 12 floor plans
- ~40-50% of 1000m trajectories are collision-free

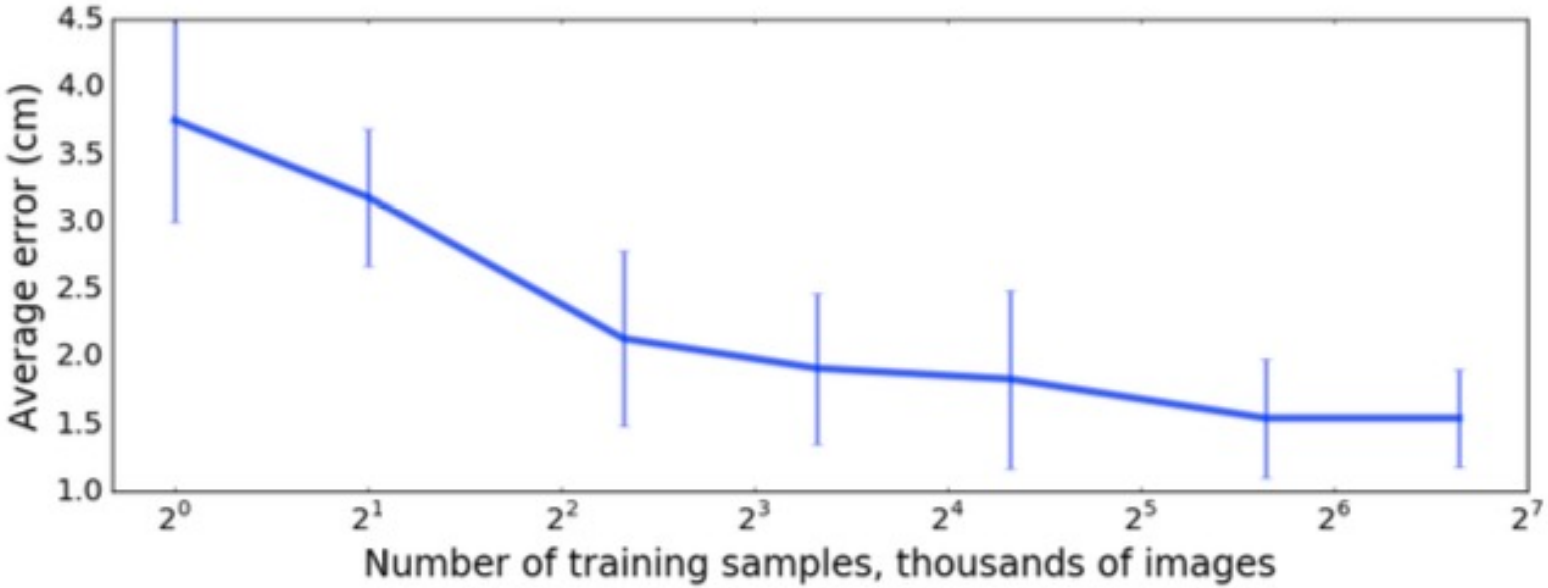
[3] Fereshteh Sadeghi and Sergey Levine. (cad)^2 rl: Real single-image flight without a single real image. *arXiv preprint arXiv:1611.04201*, 2016.

Domain Randomization for Pose Estimation

- Precise object pose localization
- 100K images with simple randomly generated textures

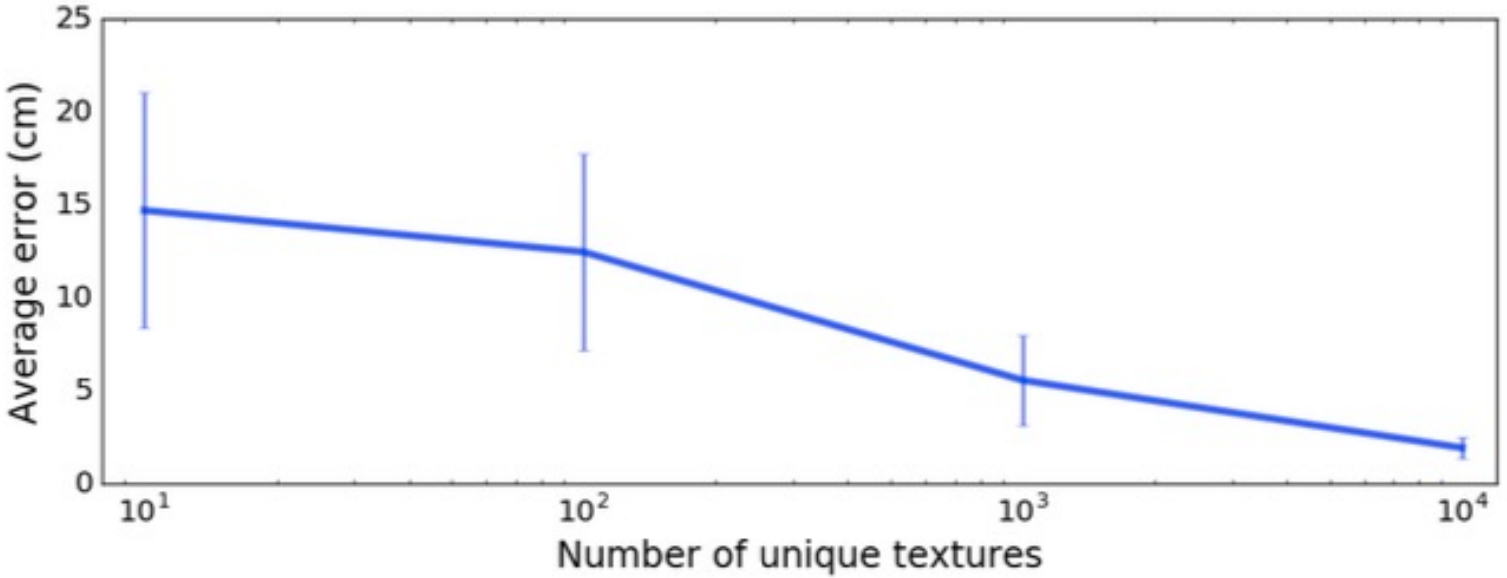
[Tobin, Fong, Ray, Schneider, Zaremba, Abbeel, 2017]

How does it work? More Data = Better



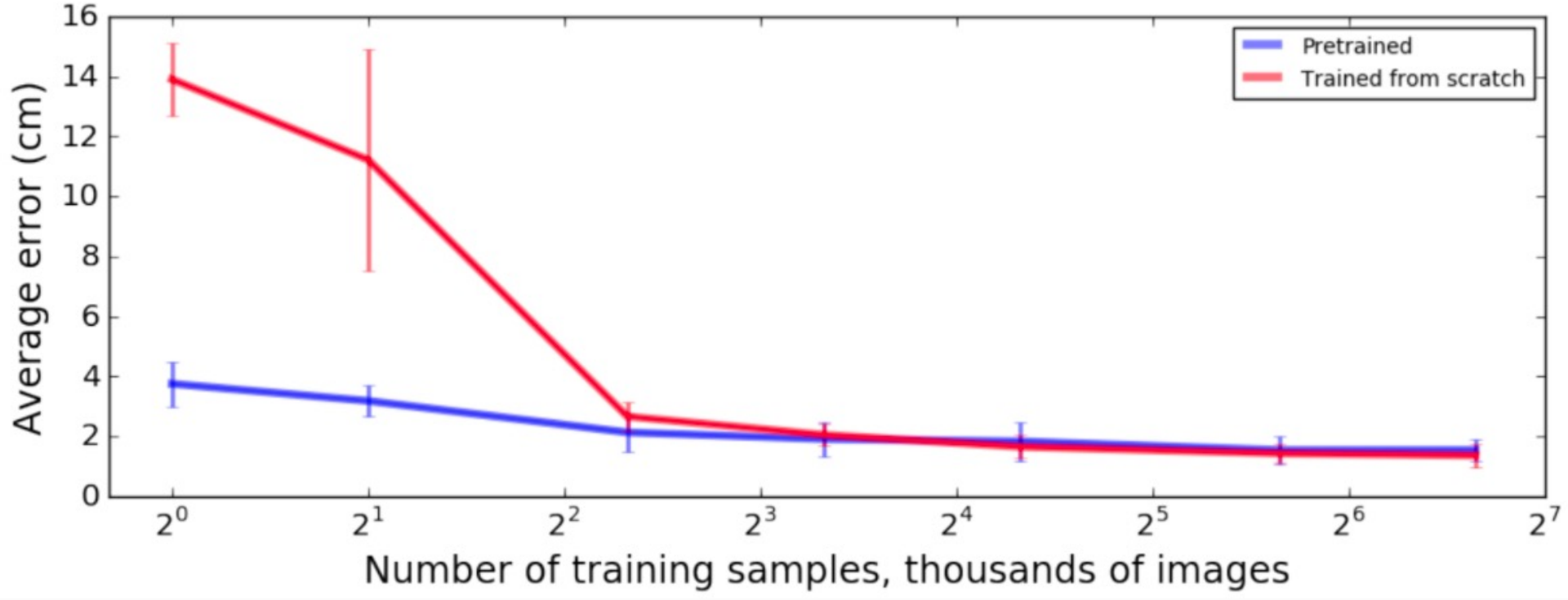
[Tobin, Fong, Ray, Schneider, Zaremba, Abbeel, 2017]

More Textures = Better



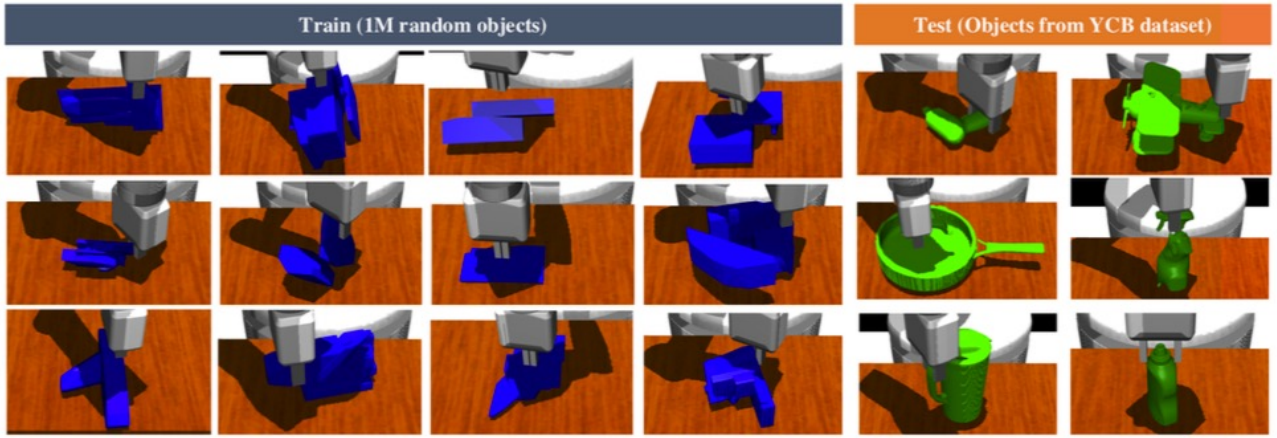
[Tobin, Fong, Ray, Schneider, Zaremba, Abbeel, 2017]

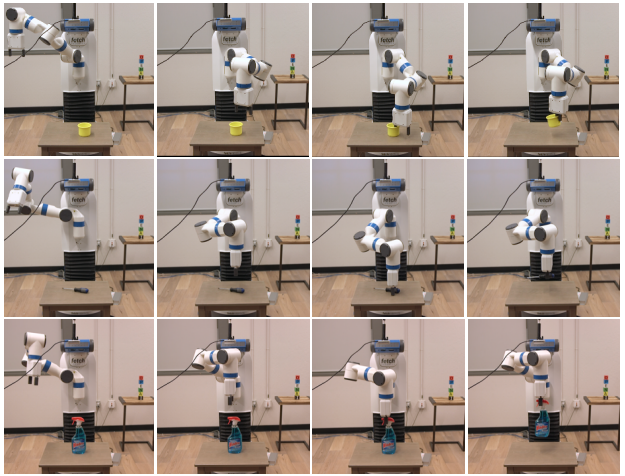
Pre-Training is not Necessary



[Tobin, Fong, Ray, Schneider, Zaremba, Abbeel, 2017]

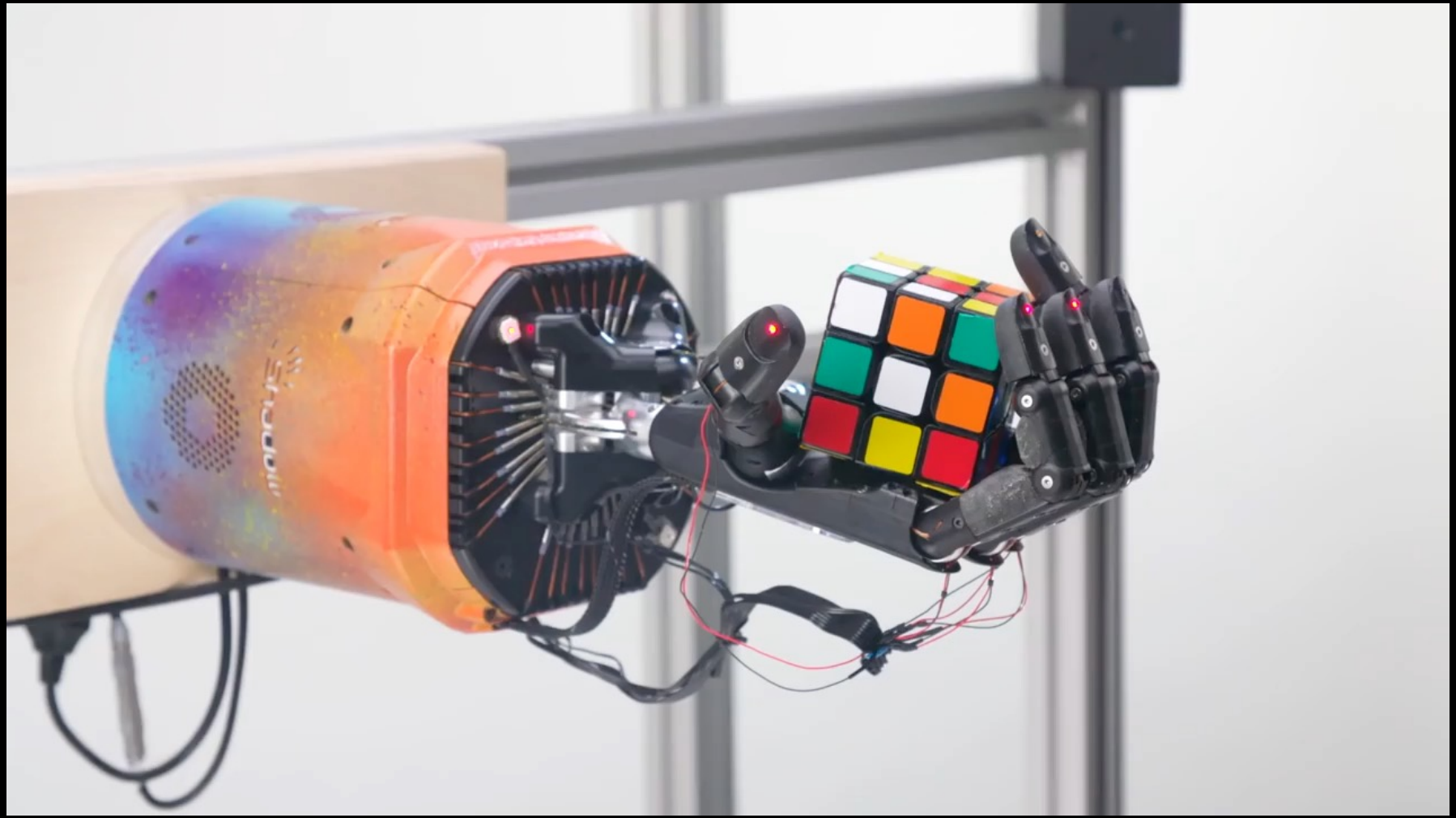
Domain Randomization for Grasping





Hypothesis: Training on a diverse array of procedurally generated objects can produce comparable performance to training on realistic object meshes.

[Tobin, Zaremba, Abbeel, 2017] [see also: Levine et al and Goldberg et al]



[OpenAI, 2019]

Many Exciting Directions in AI

- Unsupervised Learning
- Reinforcement Learning
- Unsupervised RL
- Meta-Reinforcement Learning
- Few-Shot Imitation
- Domain Randomization
- ***DL for Science and Engineering***

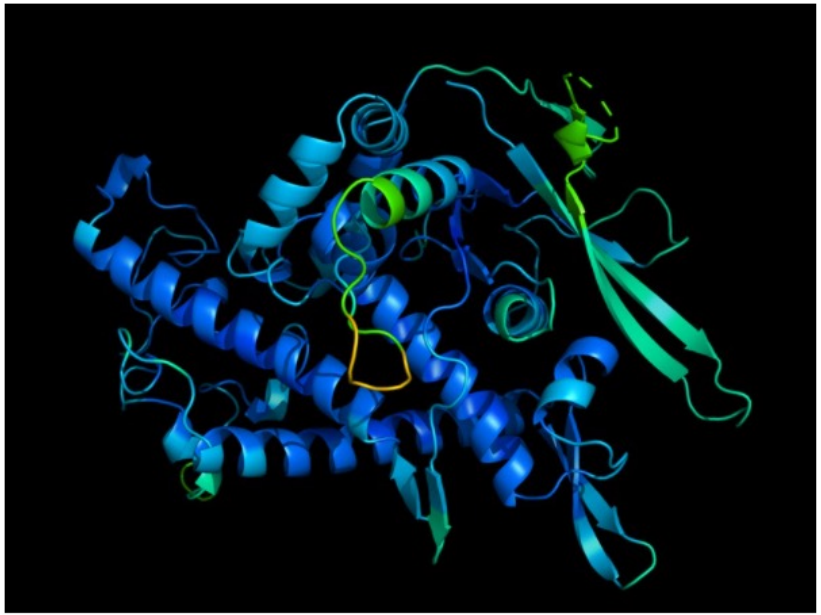
- Mitigating Bias
- Multi-modal Learning
- Architecture Search
- Value Alignment
- Scaling Laws
- Human-in-the-Loop
- Explainability

NEWS · 30 NOVEMBER 2020

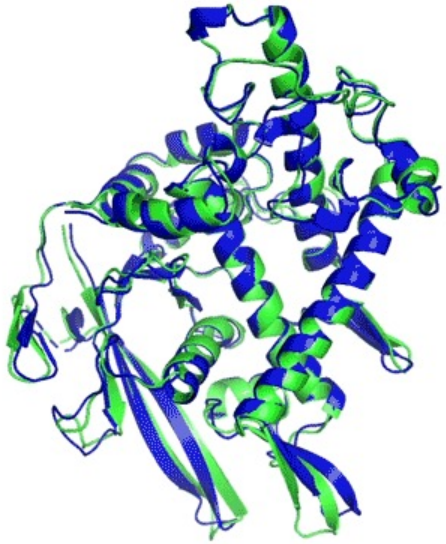
‘It will change everything’: DeepMind’s AI makes gigantic leap in solving protein structures

Google’s deep-learning program for determining the 3D shapes of proteins stands to transform biology, say scientists.

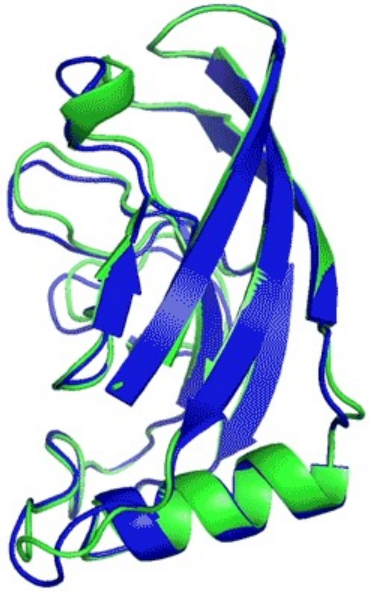
Ewen Callaway



A protein’s function is determined by its 3D shape. Credit: DeepMind



T1037 / 6vr4
90.7 GDT
(RNA polymerase domain)

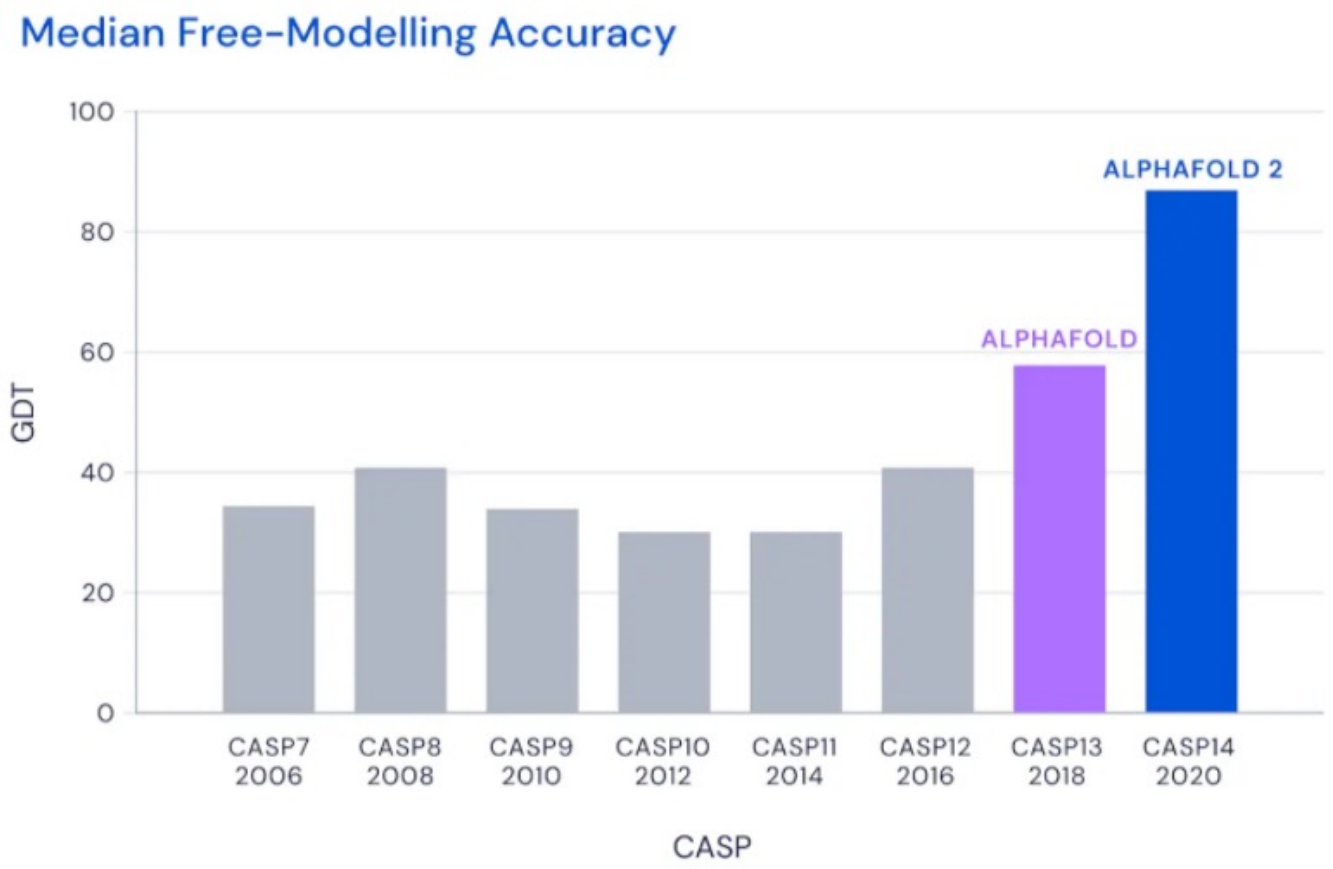


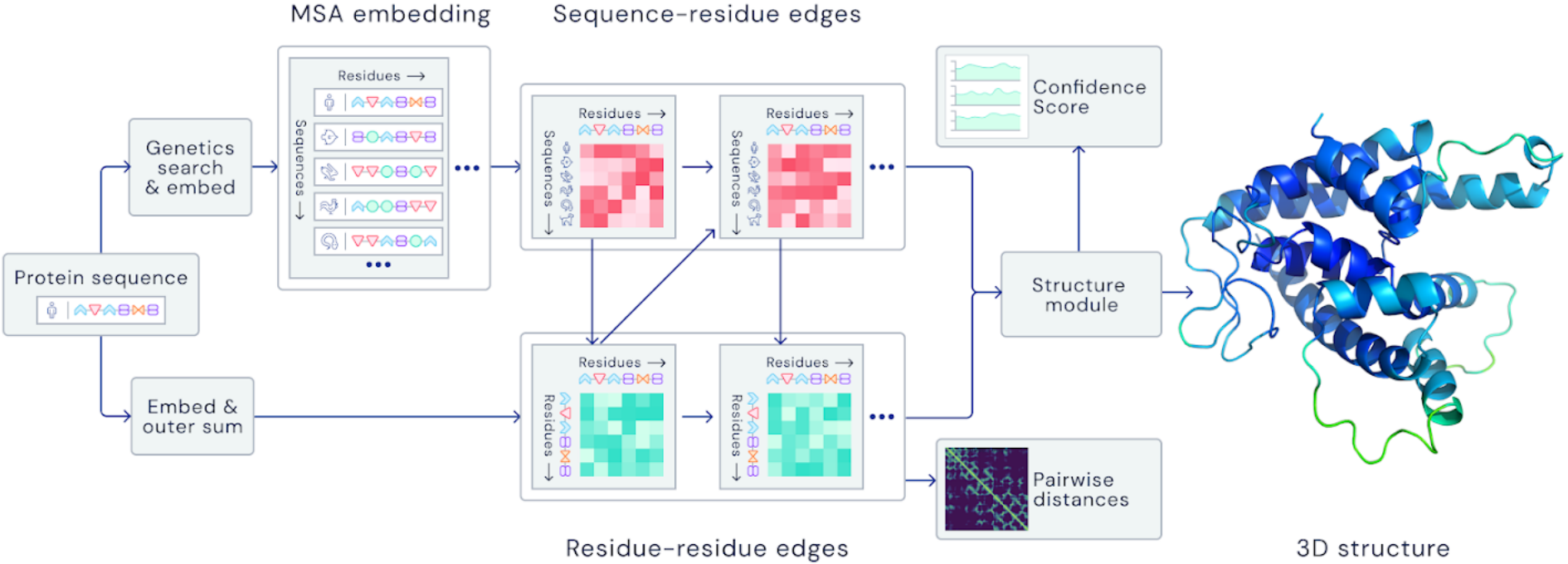
T1049 / 6y4f
93.3 GDT
(adhesin tip)

Experimental result

Computational prediction

CASP 2020 Competition




MSA embedding
Sequence-residue edges
Residues
Sequences
Genetics search & embed
Protein sequence
Embed & outer sum
Confidence Score
Structure module
Pairwise distances
Residue-residue edges
3D structure

Molecule Design

- DL to predict properties

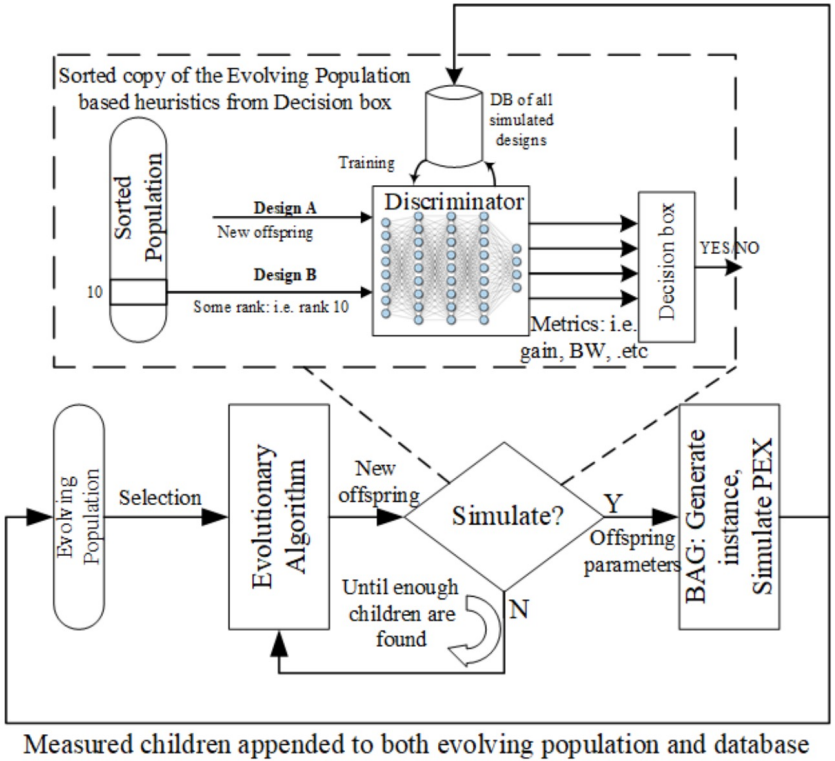
  → optimize molecule against desired properties

- Synthesis

  → DL to propose synthesis steps

Exploring Machine Learning Applications to Enable Next-Generation Chemistry
Jennifer Wei (Google)
https://docs.google.com/presentation/d/1zGoxOMWmid25hgtSgVkQYu1-GP9PT3EQ2_tzOlQZcF4/edit#slide=id.p1

A General Approach to Speed Up Design



BagNet: Berkeley Analog Generator with Layout Optimizer Boosted with Deep Neural Networks

K. Hakhamaneshi, N. Werblun, P. Abbeel, V. Stojanovic.

IEEE/ACM International Conference on Computer-Aided Design (ICAD), Westminster, Colorado, November 2019.

https://arxiv.org/abs/1907.10515

A General Approach to Speed Up Design

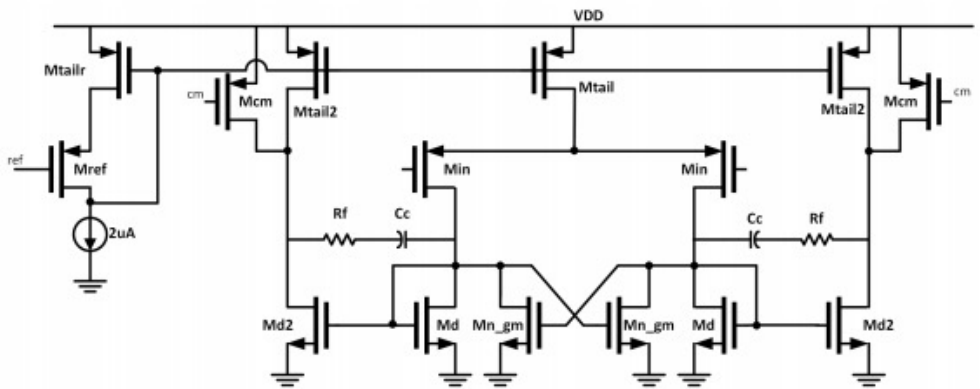


Figure 5: Two stage op-amp with negative g_m load

Table 3: Performance of expert design methodology and our approach

	Requirement	Expert	Ours
f_{unity}	>100 MHz	382 MHz	159 MHz
pm	> 60°	64°	75°
gain	>100 (for ours)	42	105

BagNet: Berkeley Analog Generator with Layout Optimizer Boosted with Deep Neural Networks
K. Hakhamaneshi, N. Werblun, P. Abbeel, V. Stojanovic.
IEEE/ACM International Conference on Computer-Aided Design (ICAD), Westminster, Colorado, November 2019.
https://arxiv.org/abs/1907.10515

A General Approach to Speed Up Design

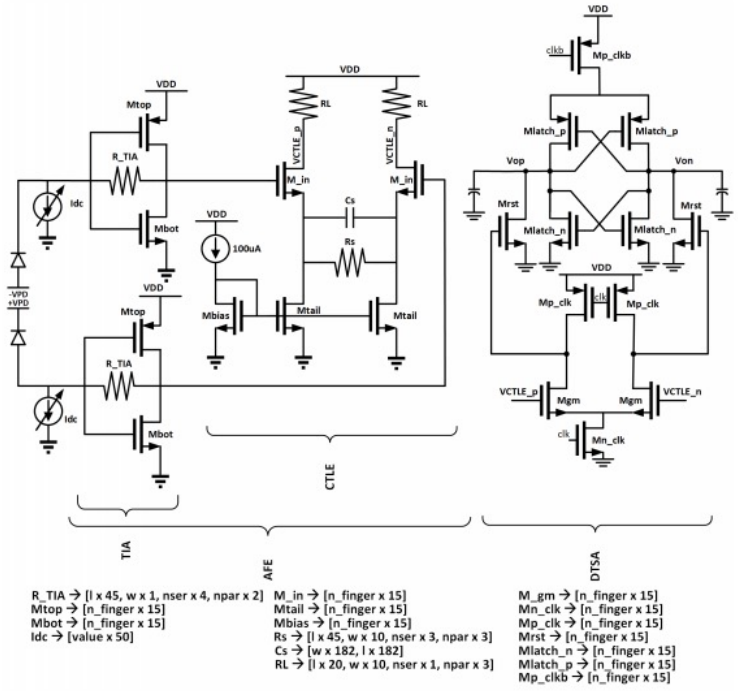


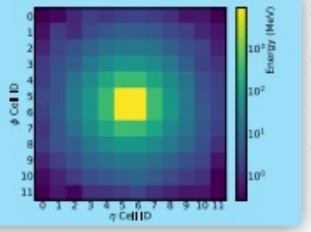
Figure 6: Optical receiver schematic

BagNet: Berkeley Analog Generator with Layout Optimizer Boosted with Deep Neural Networks
K. Hakhamaneshi, N. Werblun, P. Abbeel, V. Stojanovic.
IEEE/ACM International Conference on Computer-Aided Design (ICAD), Westminster, Colorado, November 2019.
https://arxiv.org/abs/1907.10515

GANs in High-Energy Physics

Accelerating simulations

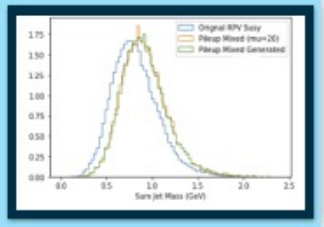
replace or augment physics simulator



M. Paganini, L. de Oliveira, B.P.N.

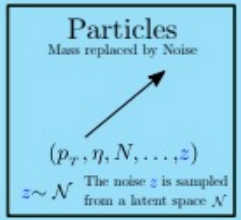
Saving disk space

replace libraries with on-the-fly generation



W. Bhimji, W. Blair, S. Farrell, B.P.N.

Unbinned, high-dimensional interpolation



J. Lin, W. Bhimji, B.P.N.

GANs for HEP
Ben Nachman
https://drive.google.com/file/d/1op6Q6OuVZvJ4VbtLkemi3oJWtSBx5FCc/view

Symbolic Math: Integrals and ODEs

Equation	Solution
$y' = \frac{16x^3 - 42x^2 + 2x}{(-16x^8 + 112x^7 - 204x^6 + 28x^5 - x^4 + 1)^{1/2}}$	$y = \sin^{-1}(4x^4 - 14x^3 + x^2)$
$3xy\cos(x) - \sqrt{9x^2\sin(x)^2 + 1}y' + 3y\sin(x) = 0$	$y = c\exp\left(\sinh^{-1}(3x\sin(x))\right)$
$4x^4yy'' - 8x^4y'^2 - 8x^3yy' - 3x^3y'' - 8x^2y^2 - 6x^2y' - 3x^2y'' - 9xy' - 3y = 0$	$y = \frac{c_1 + 3x + 3\log(x)}{x(c_2 + 4x)}$

Table 4: Examples of problems that our model is able to solve, on which Mathematica and Matlab were not able to find a solution. For each equation, our model finds a valid solution with greedy decoding.

[Lample and Charton, ICLR 2020]

Symbolic Math: Integrals and ODEs

	Integration (BWD)	ODE (order 1)	ODE (order 2)
Mathematica (30s)	84.0	77.2	61.6
Matlab	65.2	-	-
Maple	67.4	-	-
Beam size 1	98.4	81.2	40.8
Beam size 10	99.6	94.0	73.2
Beam size 50	99.6	97.0	81.0

[Lample and Charton, ICLR 2020]

Learn more about Deep Learning for Science/Engineering?

- **https://deepmind.com/blog/article/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology**
- **AlphaFold: Improved protein structure prediction using potentials from deep learning**
  Deepmind (Senior et al)
  https://www.nature.com/articles/s41586-019-1923-7
- **BagNet: Berkeley Analog Generator with Layout Optimizer Boosted with Deep Neural Networks**
  K. Hakhamaneshi, N. Werblun, P. Abbeel, V. Stojanovic.
  IEEE/ACM International Conference on Computer-Aided Design (ICAD), Westminster, Colorado, November 2019.
  https://arxiv.org/abs/1907.10515
- **Evaluating Protein Transfer Learning with TAPE**
  R. Rao, N. Bhattacharya, N. Thomas, Y, Duan, X. Chen, J. Canny, P. Abbeel, Y. Song
  https://www.biorxiv.org/content/10.1101/676825v1
- **Opening the black box: the anatomy of a deep learning atomistic potential**
  Justin Smith
  https://drive.google.com/file/d/1f1iiXKzxNbNz5lL5x2ob9xGYLHNHhxeU/view
- **Exploring Machine Learning Applications to Enable Next-Generation Chemistry**
  Jennifer Wei (Google)
  https://docs.google.com/presentation/d/1zGoxOMWmid25hgtSgVkQYu1-GP9PT3EQ2_tzOlQZcF4/edit#slide=id.p1
- **GANs for HEP**
  Ben Nachman
  https://drive.google.com/file/d/1op6Q6OuVZvJ4VbtLkemi3oJWtSBx5FCc/view
- **Deep Learning for Symbolic Mathematics**
  G. Lample and F. Charton
  https://openreview.net/pdf?id=S1eZYeHFDS
- **A Survey of Deep Learning for Scientific Discovery**
  Maithra Raghu, Eric Schmidt
  https://arxiv.org/abs/2003.11755

Outline

- Sampling of research directions
- ***Overarching research trend***
- How to keep up

Compute Increasing Rapidly

Companies Developing Deep Learning Chips

	Company	HQ	Story
Public	Ambarella	United States	Developing computer vision chips for autonomous cars
	AMD	United States	GPU based deep learning
	Facebook	United States	Forming a team to build SoCs and perhaps inference chips
	Google	United States	Custom designed TPU deployed in Google Cloud
	Intel	United States	Developing a Neural Network chip based on Nervana acquisition
	Nvidia	United States	Current market leader using GPU based deep learning
	Tesla	United States	Developing a custom AI chip for autonomous driving
Private	Bitmain	China	Top maker of Bitcoin mining chips
	Cambricon	China	China's state-backed startup with $1B+ valuation.
	Cerebras Systems	United States	Ex-AMD team backed by Benchmark Capital
	DeePhi	China	China based startup with a focus on video analysis
	GraphCore	United Kingdom	Building a 16nm deep learning chip for training and inference
	Groq	United States	Ex-Google TPU team backed by Social Capital
	Horizon Robotics	China	Ex-Baidu team. Embedded / computer vision focus
	KnuEdge	United States	Headed by former NASA CTO
	Mythic	United States	In-memory inference for IoT backed by DFJ
	Tenstorrent	Canada	Toronto based chip startup
	Thinci	United States	Computer vision / auto focus
	Wave Computing	United States	Makes data flow acceleration servers

Source: ARK Invest [20-April-2018]

Let’s Calibrate Compute Scale

Architecture	Num neurons	Num synapses

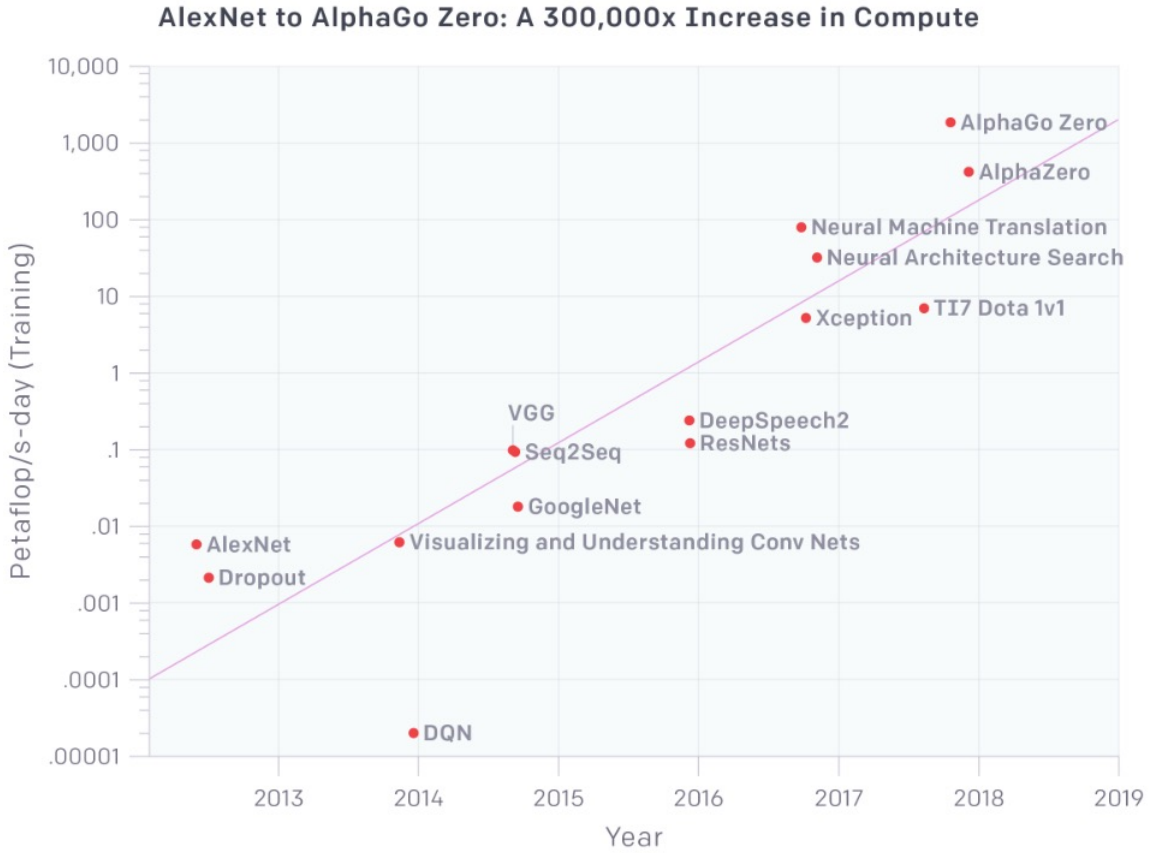
If each synapse is 1 FLOP (i.e., can fire / not fire once per second),

Then human brain requires 10^{15} flops = 1 petaflop.

NVIDIA’s DGX-2 = 2 petaflops in one server rack! ($400,000 price tag)

2 TPU-v3 slices = ~1 petaflop in Google Cloud at $16/hour ($5/hour if pre-emptible)

Compute per Major Result



Source: OpenAI

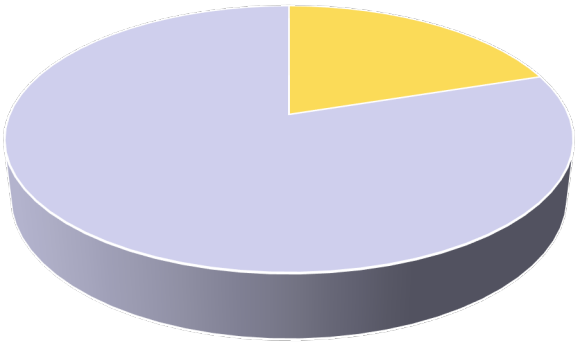
Implications for research agenda?

Key Enablers for AI Progress?

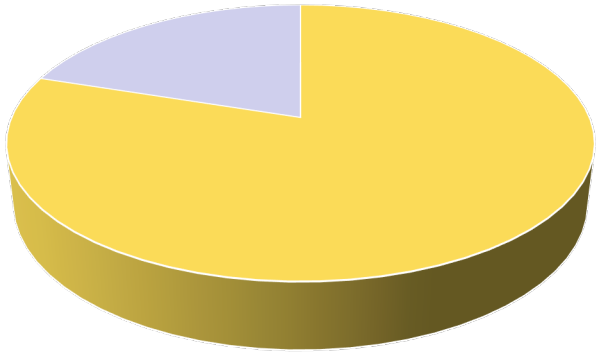
DATA

COMPUTE

HUMAN INGENUITY



Problem Territory 1

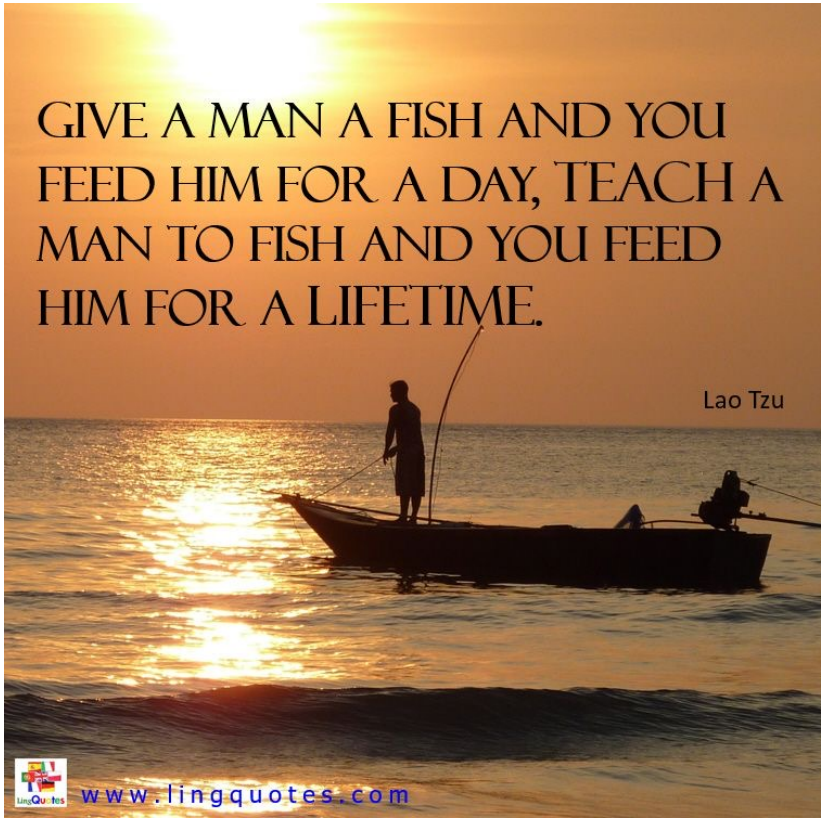


Problem Territory 2

E.g., Deep Learning to Learn

Outline

- Sampling of research directions
- Overarching research theme
- ***How to keep up***



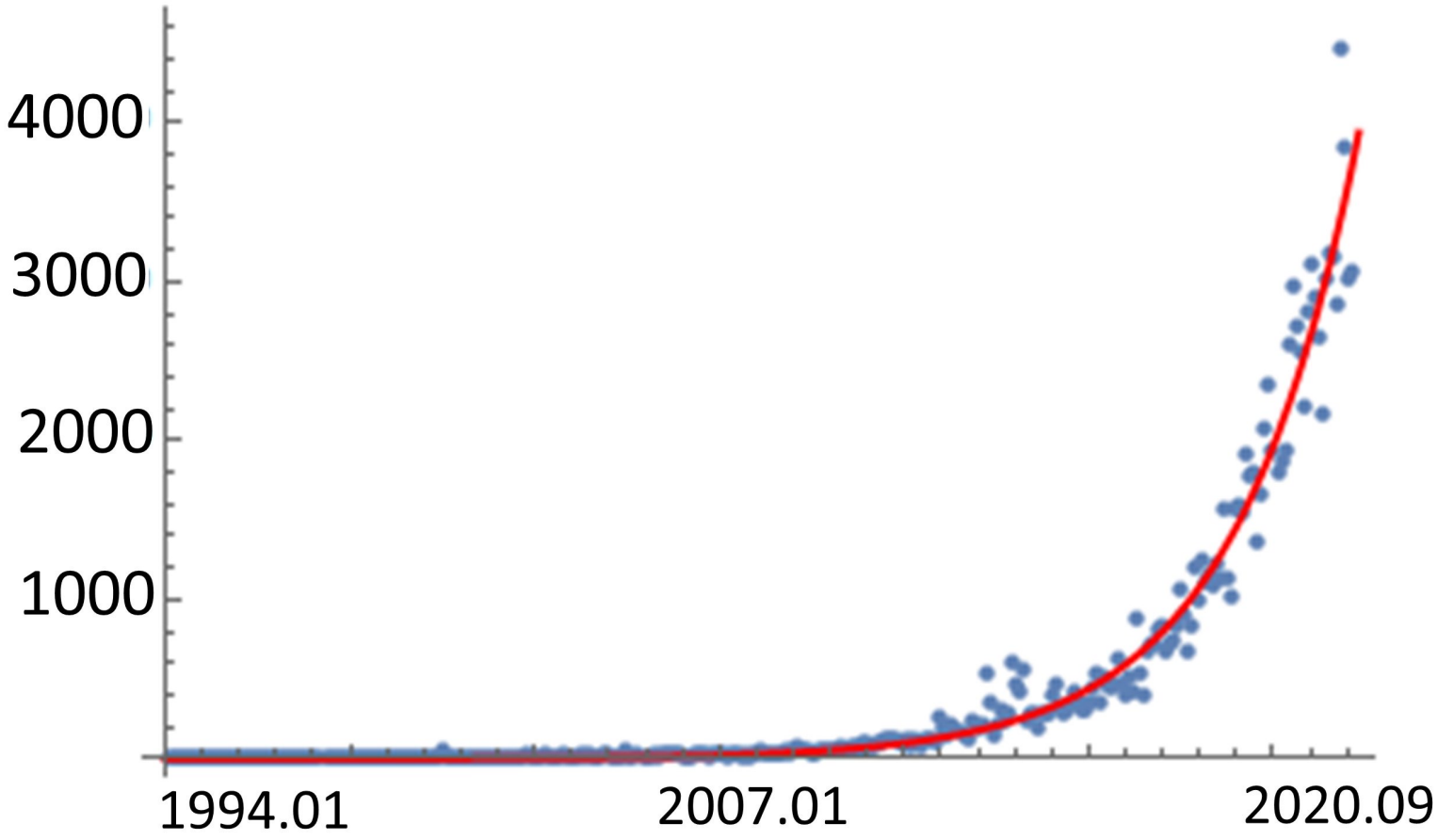
How to Keep Up

- (mostly) keeping up with (mostly) not reading papers
- when you decide to read papers

How to Keep Up

- ***(mostly) keeping up with (mostly) not reading papers***
- when you decide to read papers

ML+AI arXiv papers per month



Great Resources

- Tutorials at conferences (e.g., ICML, NeurIPS)
- Graduate courses and seminars
- Yannic Kilcher youtube channel – explains papers
- Two Minute Papers youtube channel
- The Batch – newsletter by Andrew Ng
- Import AI – newsletter by Jack Clark

How to Keep Up

- (mostly) keeping up with (mostly) not reading papers
- ***when you decide to read papers***

Which papers to read? (1/2)

- Tutorials at conferences (e.g., ICML, NeurIPS)
- Graduate courses and seminars
- Yannic Kilcher youtube channel – explains papers
- Two Minute Papers youtube channel
- The Batch – newsletter by Andrew Ng
- Import AI – newsletter by Jack Clark

Which papers to read? (2/2)

- Arxiv sanity (by Karpathy)

- Twitter
  - Jack Clark, Karpathy, Ian Goodfellow, hardmaru, smerity, pabbeel, …
- AI/DL Facebook Group: Very active group where members post anything from news articles to blog posts to general ML questions
- ML Subreddit

How to Read a Paper?

- Read the title, abstract, section headers, and figures
- Try and find slides or a video on the paper (these do not have to be by the authors).
- Read the introduction (Jennifer Widom)
  - What is the problem?
  - Why is it interesting and important?
  - Why is it hard? (e.g., why do naive approaches fail?)
  - Why hasn't it been solved before? (Or, what's wrong with previous proposed solutions? How does this paper differ?)
  - What are the key components of this approach and results? Any limitations?
- Skim the related work.
  - Is there any related work you are familiar with?
  - If so, how does this paper relate to those works?
- Skim the technical section.
  - Where are the novelties?
  - What are the assumptions?
- Read the experiments.
  - What questions are the experiments answering?
  - What questions are the experiments not answering?
  - What baselines do they compare against?
  - How strong are these baselines?
  - Is the experiment methodology sound?
  - Do the results support their claims?
- Read the conclusion/discussion.
- Read the technical section.
  - Read in “good faith” (i.e., assume the authors are correct)
  - Skip over confusing parts that don’t seem fundamental
  - If any important part is confusing, see if the material is in class slides or prior papers
- Read the paper as you see fit.

[source: Gregory Kahn]

Reading Group

- ***Read papers together with friends:***
  - Mode 1: everyone reads, then discusses
  - Mode 2: 1 (or 2) people read, give small tutorial to others

PS: Why do a PhD? (or not)

- Become one of the world's experts in a topic you really care about
- Technically deep and demanding
- Crudely: develop new tools/techniques rather than use existing tools/techniques

Thank you