DeepMind

Representation Learning Without Labels

S. M. Ali Eslami
@arkitus

OxML 2022

DeepMind

Acknowledgements

Slides prepared with
Irina Higgins, Danilo J. Rezende

With valuable input from
Mihaela Rosca, Shakir Mohamed, Alex Graves, Olivier Henaff, Brian McWilliams, Steven McDonagh, David Pfau, Jovana Mitrovic, Andrew Zisserman, Maria Tsimpoukelli

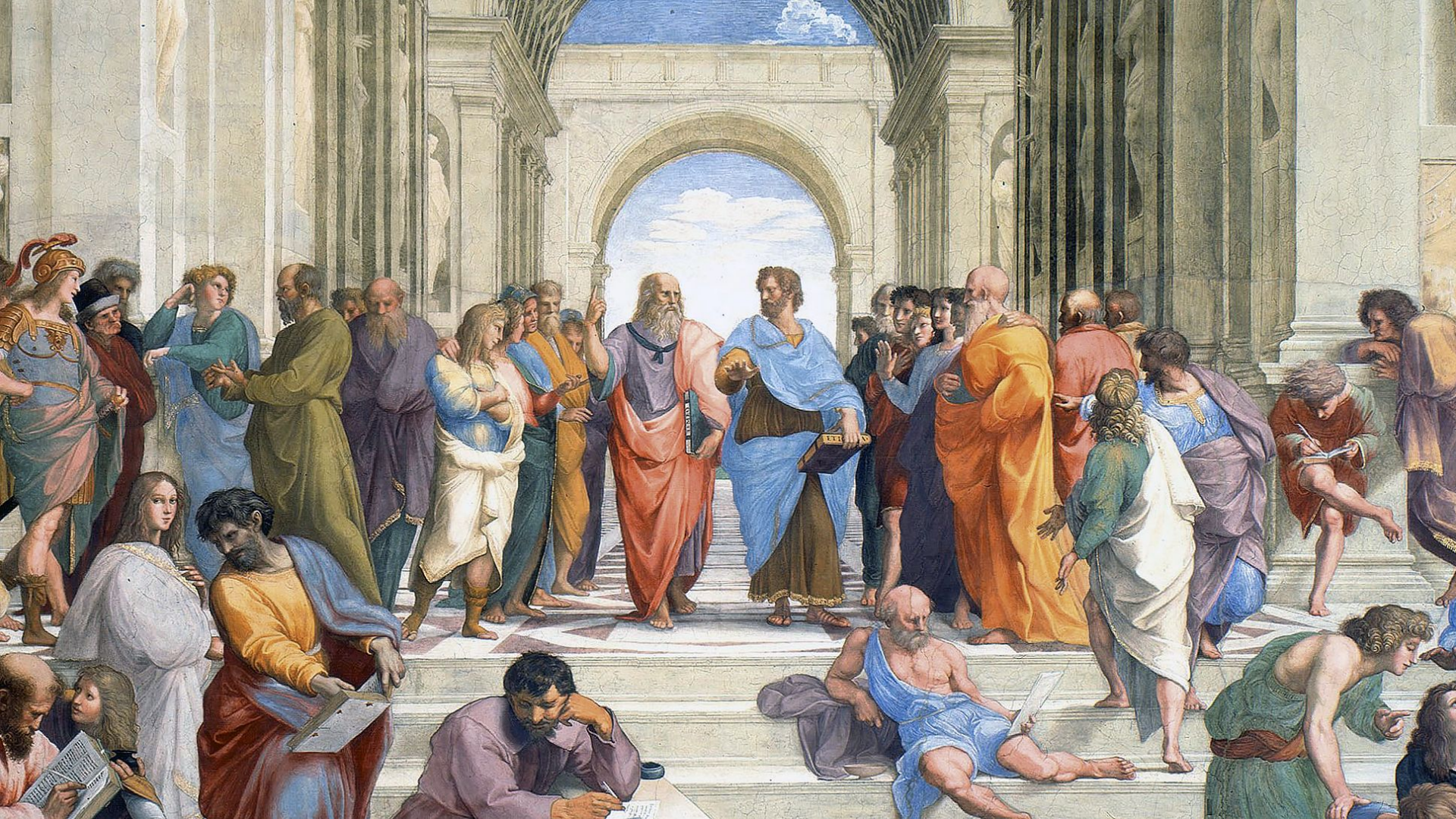
OxML 2022

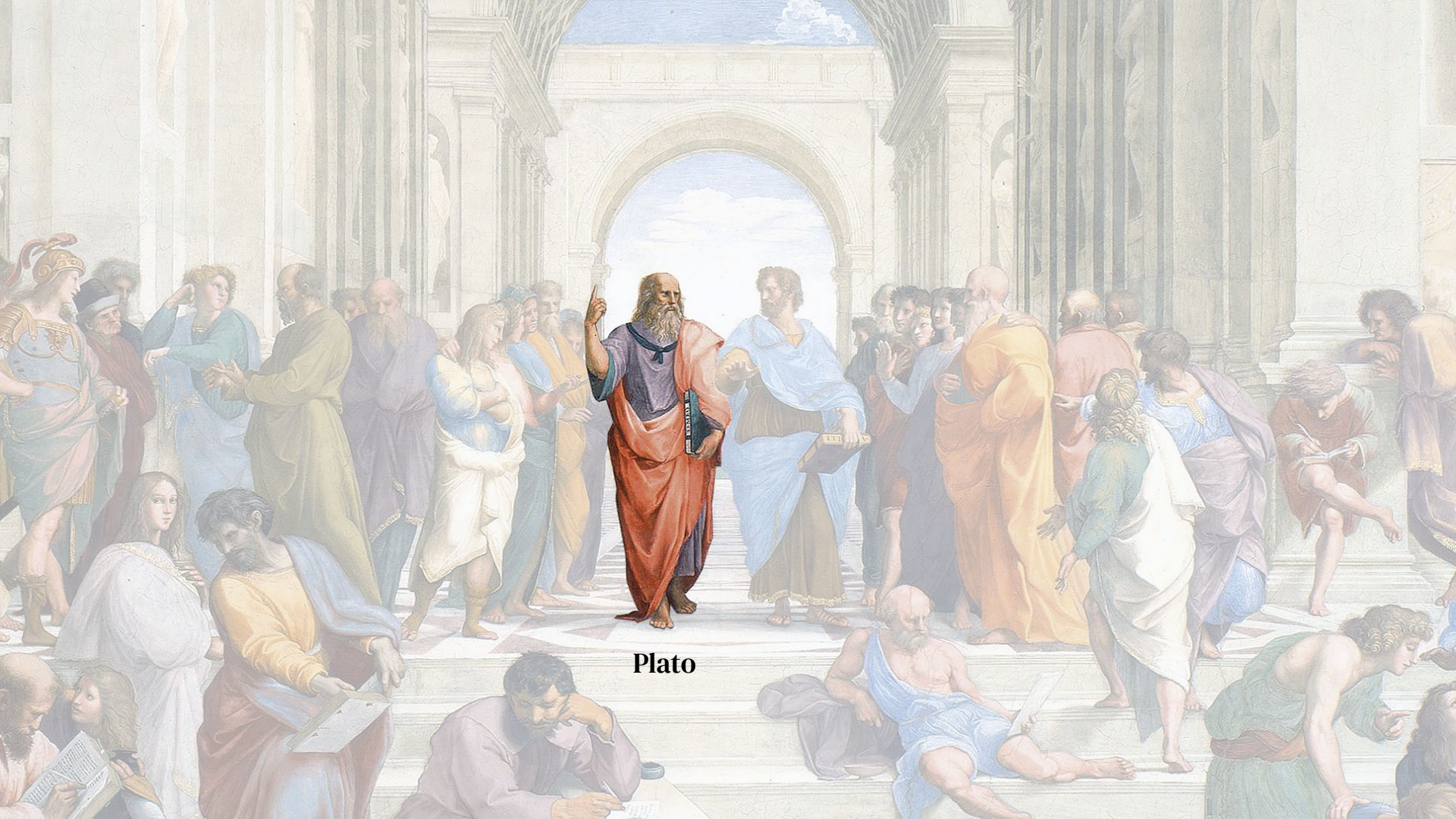
Disclaimers

- This is a huge topic with a vast, multi-disciplinary history
- I will inevitably miss important related work
- There are many views on the literature, this is one
- Not necessarily chronological
- No slide in isolation matters, it’s their relation to each other that matters most
- Emphasis on insight rather than technique

DeepMind

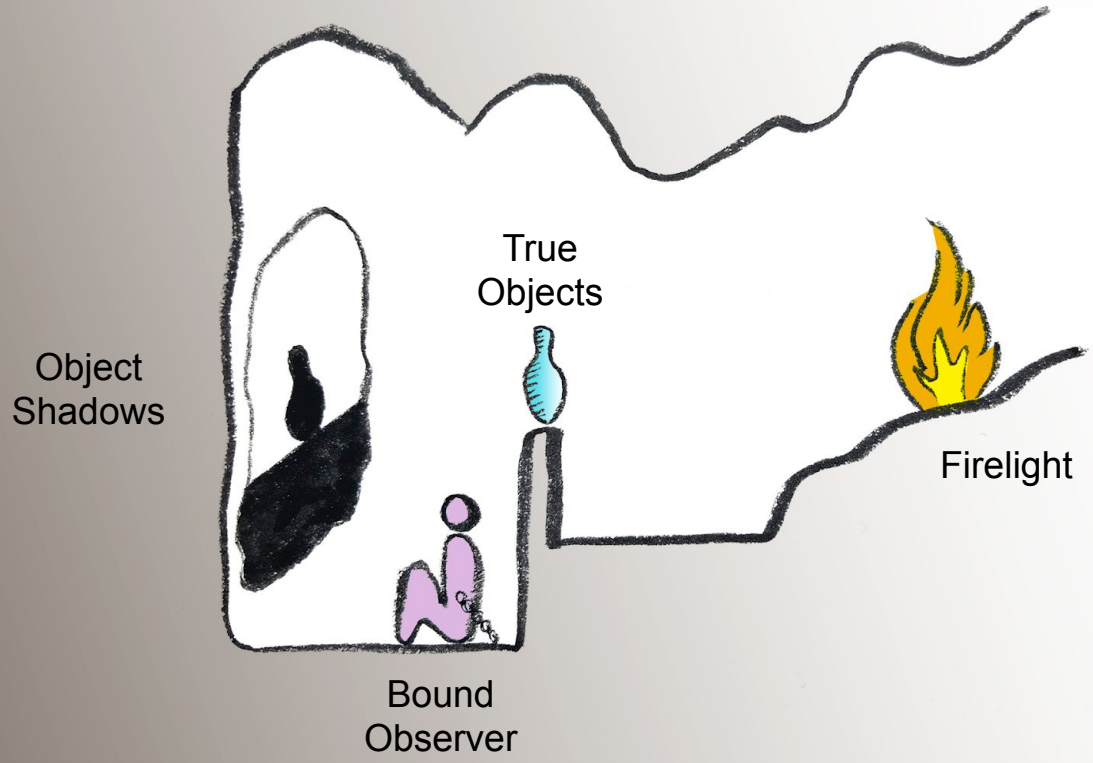
1 Introduction



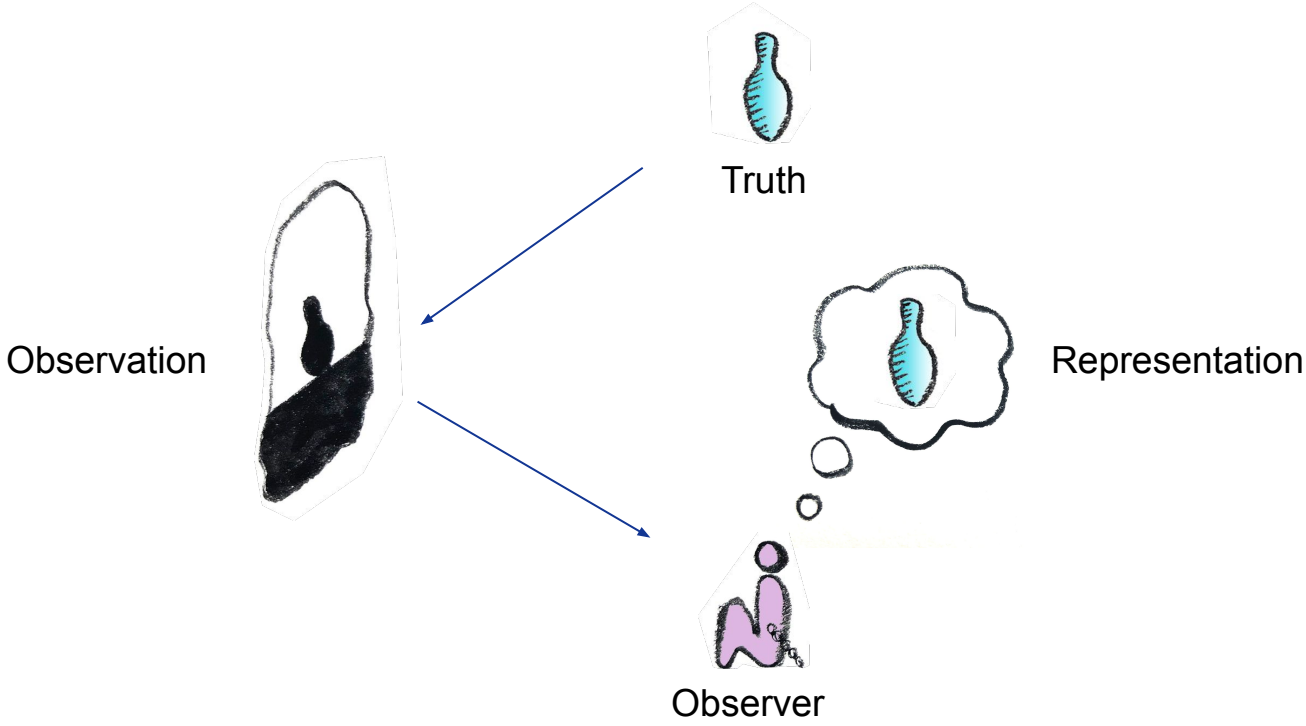


Plato

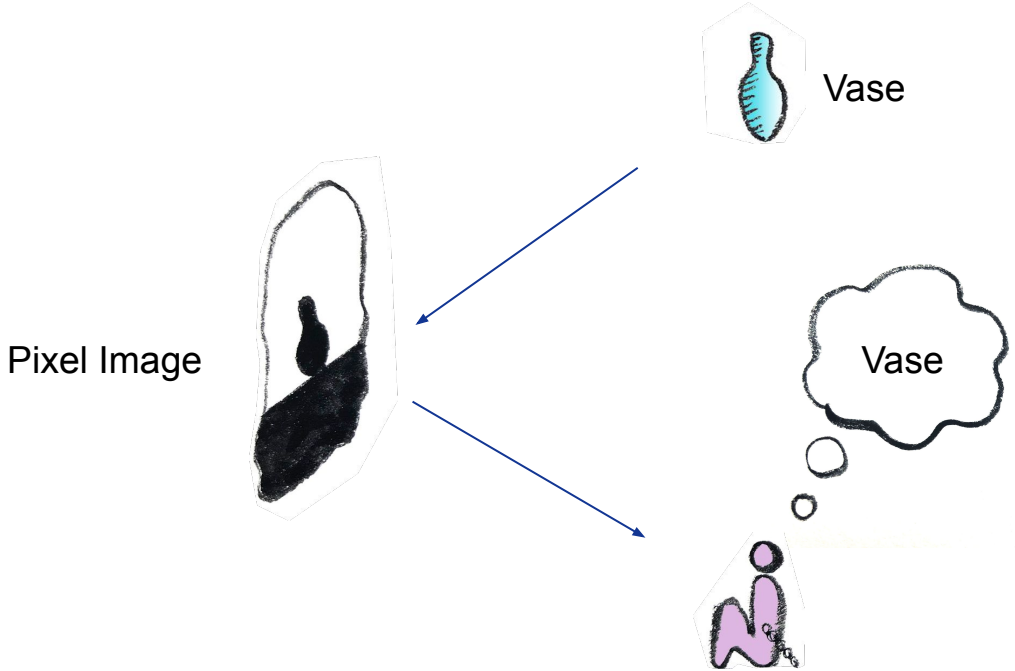
Plato's allegory of the cave



Plato's allegory of the cave

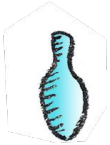


Desired understanding: simplistic view



The representation problem

Pixel Image



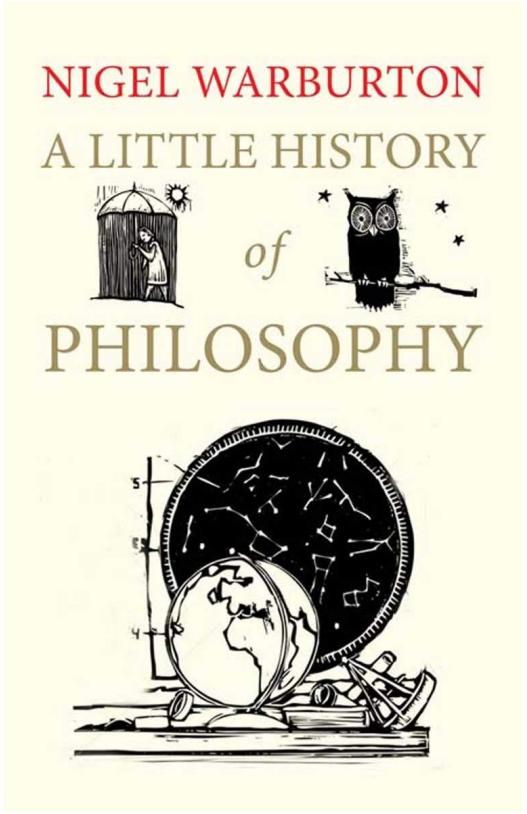
?

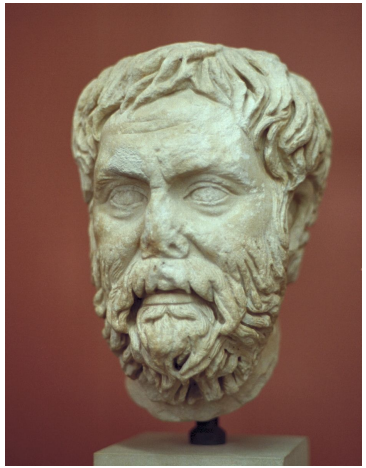
- **Class:** Vase
- **Shape:** Gourd
- **Colour:** Blue
- **Height:** 15cm tall
- **Weight:** 230g
- **Scratched:** Yes
- **Moving:** No
- And many more attributes...

- Which attributes?
- What formats?
- Partial observability?
- How quickly?
- Measure of success?

AI is a form of empirical philosophy

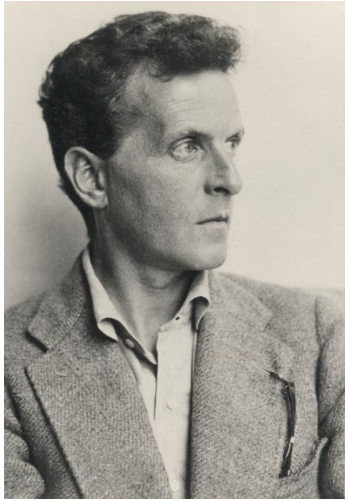
Hot take!





Phyrro

Descartes



Wittgenstein

“

Vision is a process that produces from images of the external world a description that is useful to the viewer and not cluttered by irrelevant information.

— Marr and Nishihara, 1978

Want to learn more?

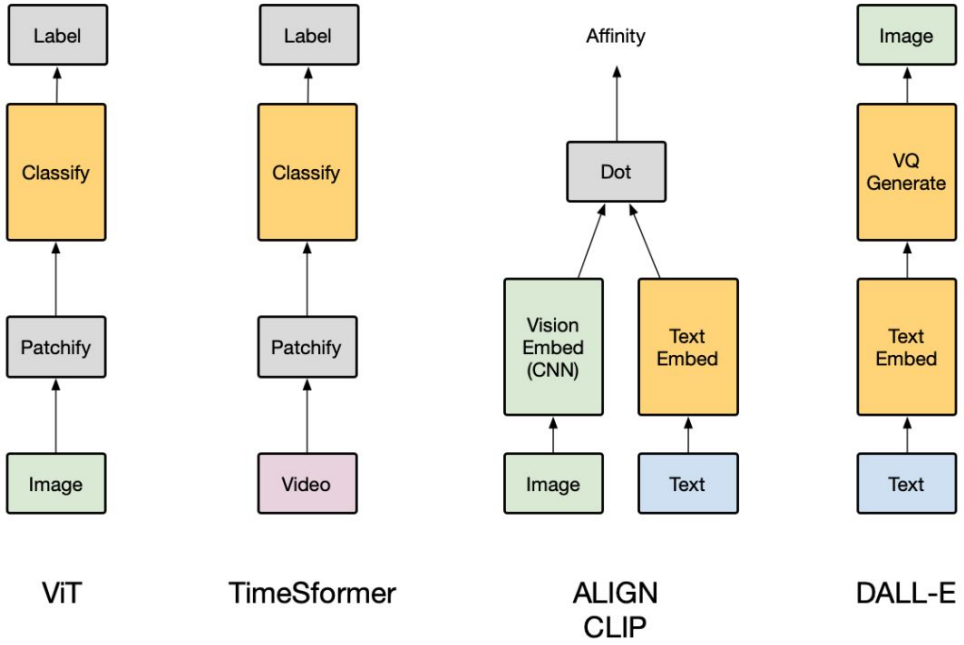
Representation and Recognition of the Spatial Organization of Three-Dimensional Shapes, Marr et al, Proc. R. Soc. Lond. (1978)

The supervised solution

1. Decide **which attributes** you care about
2. Decide the **format** for each attribute
3. Create a **large dataset** of (image, label) pairs
4. **Train a neural network** to predict labels

label= ‘Vase’ label= ‘Ball’ label= ‘Dog’

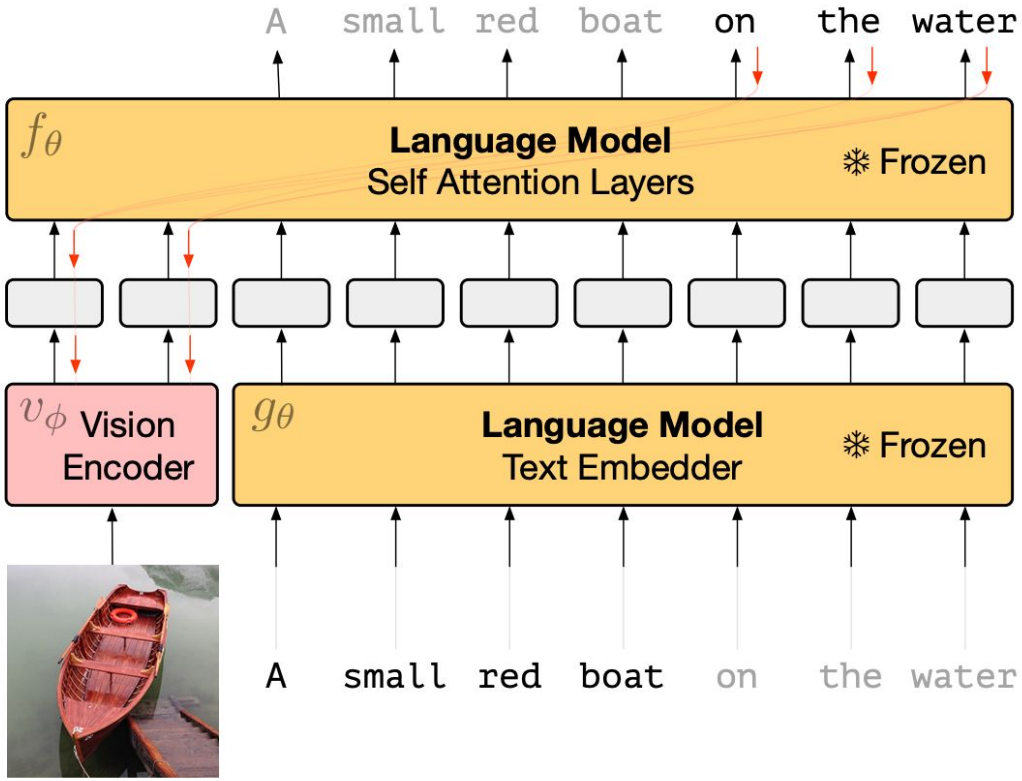
Just train with labels



Vision + Language Models

Want to learn more?

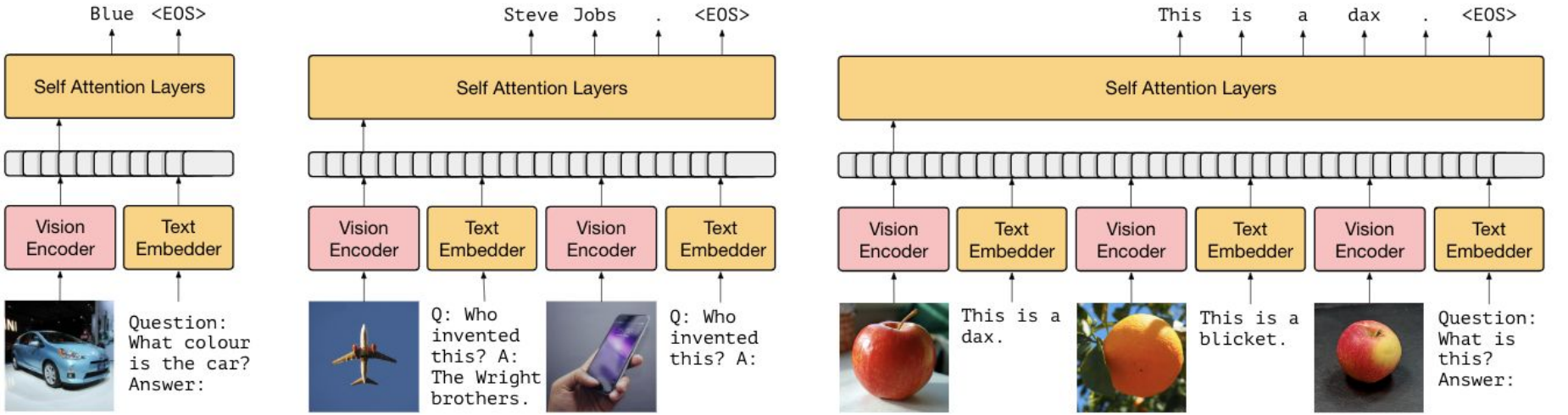
Multimodal Few-Shot Learning with Frozen Language Models, Tsimpoukelli et al, (2021)



Vision + Language Models

Want to learn more?

Multimodal Few-Shot Learning with Frozen Language Models, Tsimpoukelli et al, (2021)



(a) **0-shot VQA**

(b) **1-shot outside-knowledge VQA**

(c) **Few-shot image classification**

Evaluation

Want to learn more?

Multimodal Few-Shot Learning with Frozen Language Models, Tsimpoukelli et al, (2021)

This person is like 😄.

This person is like 😞.

This person is like

Model Completion

😱. <EOS>

This was invented by Zacharias Janssen.

This was invented by Thomas Edison.

This was invented by

Model Completion

the Wright brothers. <EOS>

With one of these I can drive around a track, overtaking other cars and taking corners at speed

With one of these I can take off from a city and fly across the sky to somewhere on the other side of the world

With one of these I can

Model Completion

break into a secure building, unlock the door and walk right in <EOS>

DeepMind's Flamingo

Color is "Yellow" and is written in green.

Color is "Black" and is written in yellow.

Great! Do you know the name of this test?

I think it is called the Stroop test.

What is in this picture?

It's a bowl of soup with a monster face on it

OpenAI's DALL-E 2

HeavensLastAngel
@HvnsLstAngel

"A still of Kermit The Frog in WALL-E (2008)" #dalle

HeavensLastAngel
@HvnsLstAngel

"A still of Kermit The Frog in Family Guy (2008)" #dalle

HeavensLastAngel
@HvnsLstAngel

"A still of Kermit The Frog in Star Wars (1977)"

The supervised solution

1. Decide **which attributes** you care about
2. Decide the **format** for each attribute
3. Create a **large dataset** of (image, label) pairs
4. **Train a neural network** to predict labels

label= ‘Vase’ label= ‘Ball’ label= ‘Dog’

The supervised solution

1. Decide **which attributes** you care about
2. Decide the **format** for each attribute
3. Create a **large dataset** of (image, label) pairs
4. **Train a neural network** to predict labels

label= ‘Vase’ label= ‘Ball’ label= ‘Dog’

Works extremely well on a diverse set of problems

However, it also raises several questions:

1. Who provides ground truth for the labels?
2. What if we don’t ‘know’ the groundtruth ourselves?
3. Which attributes are chosen for labelling?
4. Which attributes are ignored for labelling?
5. What biases do the labels propagate?
6. Do children learn purely from labels?
7. Do animals learn from labels at all?
8. **Can useful representations develop without labels?**

Can useful representations develop without labels?

If the answer is yes:

1. We learn more efficiently when we do gain access to labels

2. We still learn useful things when label collection is impossible

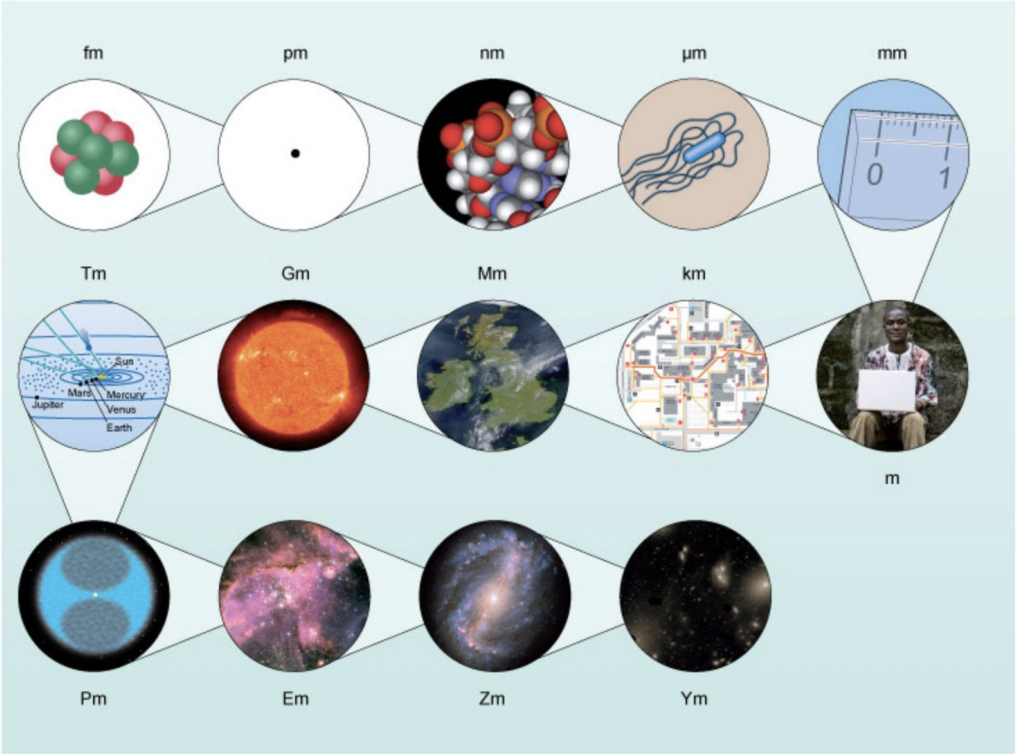


Figure 3 The scale of the Universe from atoms to galaxies. Each image is representative of the unit indicated. Each stage is 1000 times larger than the previous one. The smallest shown, femtometres (fm), is depicted by the nucleus of an oxygen atom. The next microscopic scale (pm) is still too small to show a whole atom. The next two scales (nm and μm) are represented by the diameter of deoxyribonucleic acid (DNA) and a bacterial cell respectively. You will be familiar with the scales millimetres (mm) to megametres (Mm). The Sun is at the scale of a gigametre (Gm). Moving outwards are the inner Solar System (Tm), Oort cloud (Pm), a nebula (Em), galaxy (Zm) and cluster of galaxies (Ym).

Human experience is vanishingly small

Hot take!

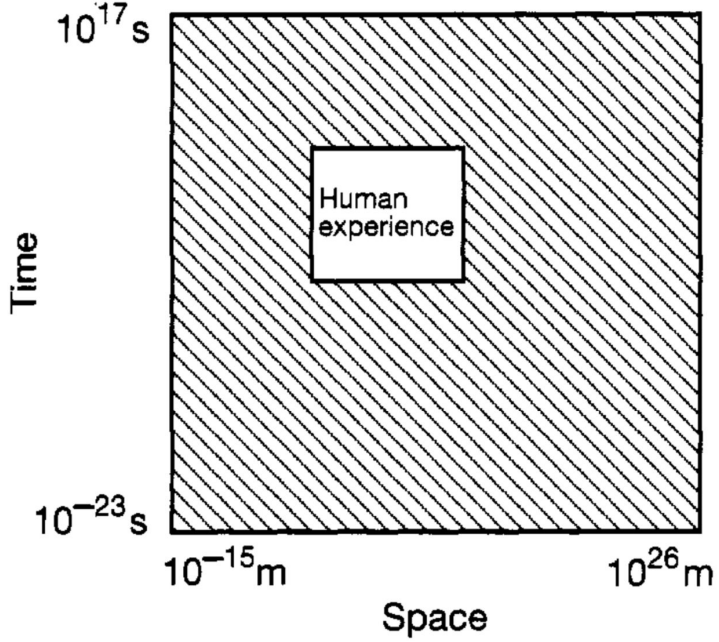


Figure 1.1: *Human experience of space and time in the physical world.*

DeepMind

2 Why Now?

History of representation learning

Want to learn more?

Some Studies in Machine Learning Using the Game of Checkers, Samuel, IBM Journal (1959)

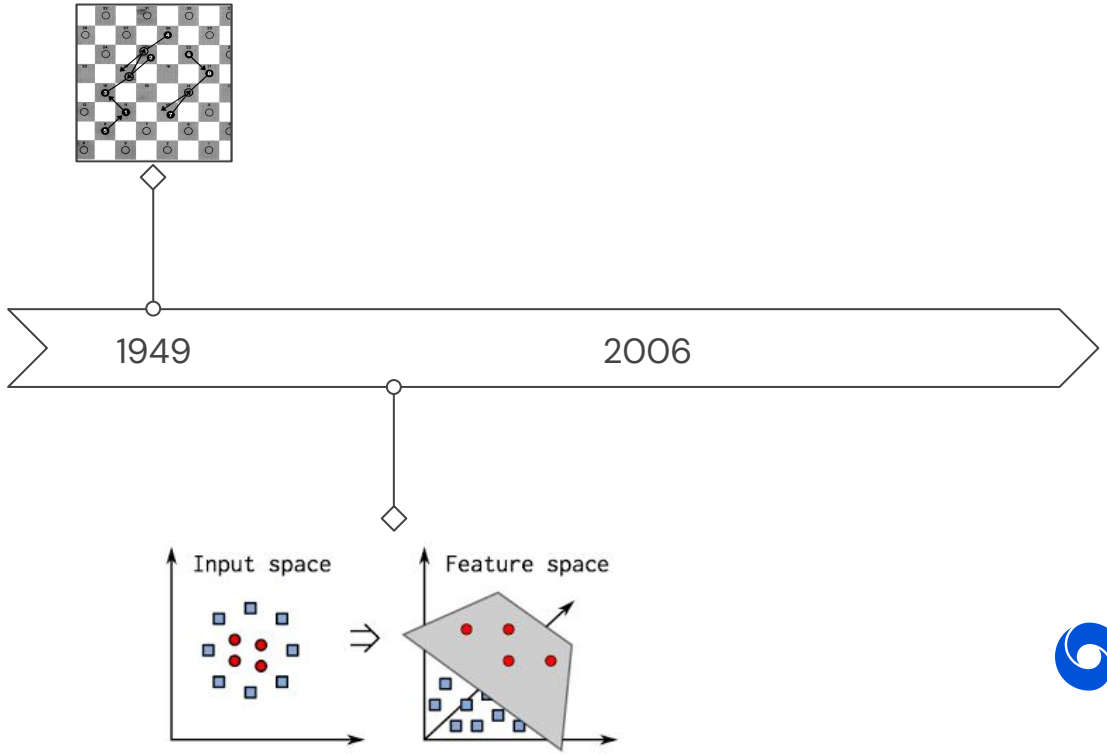
- Arthur Samuel coins the term “machine learning”

History of representation learning

Want to learn more?

Kernel Methods in Machine Learning, Hofmann et al, The Annals of Statistics (2008)

- Arthur Samuel coins the term “machine learning”
- Feature engineering and kernel methods

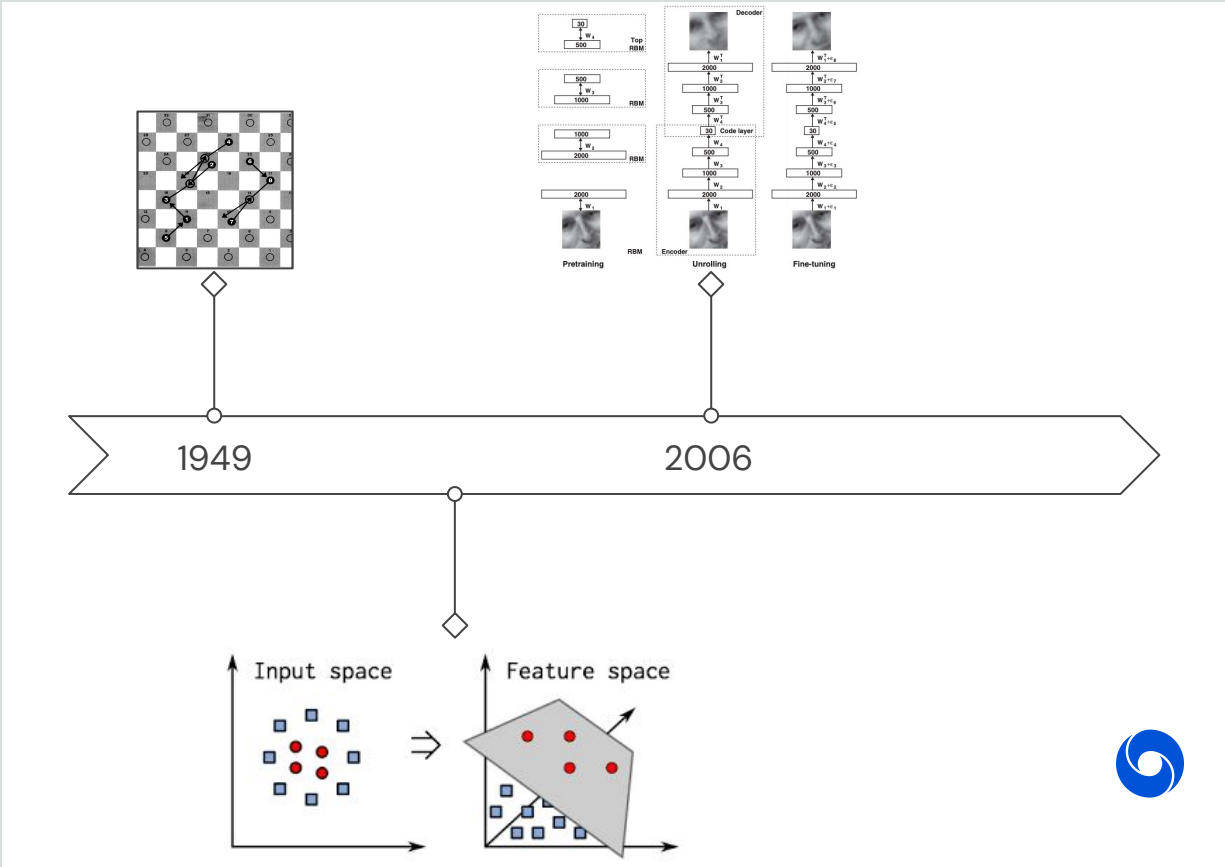


History of representation learning

Want to learn more?

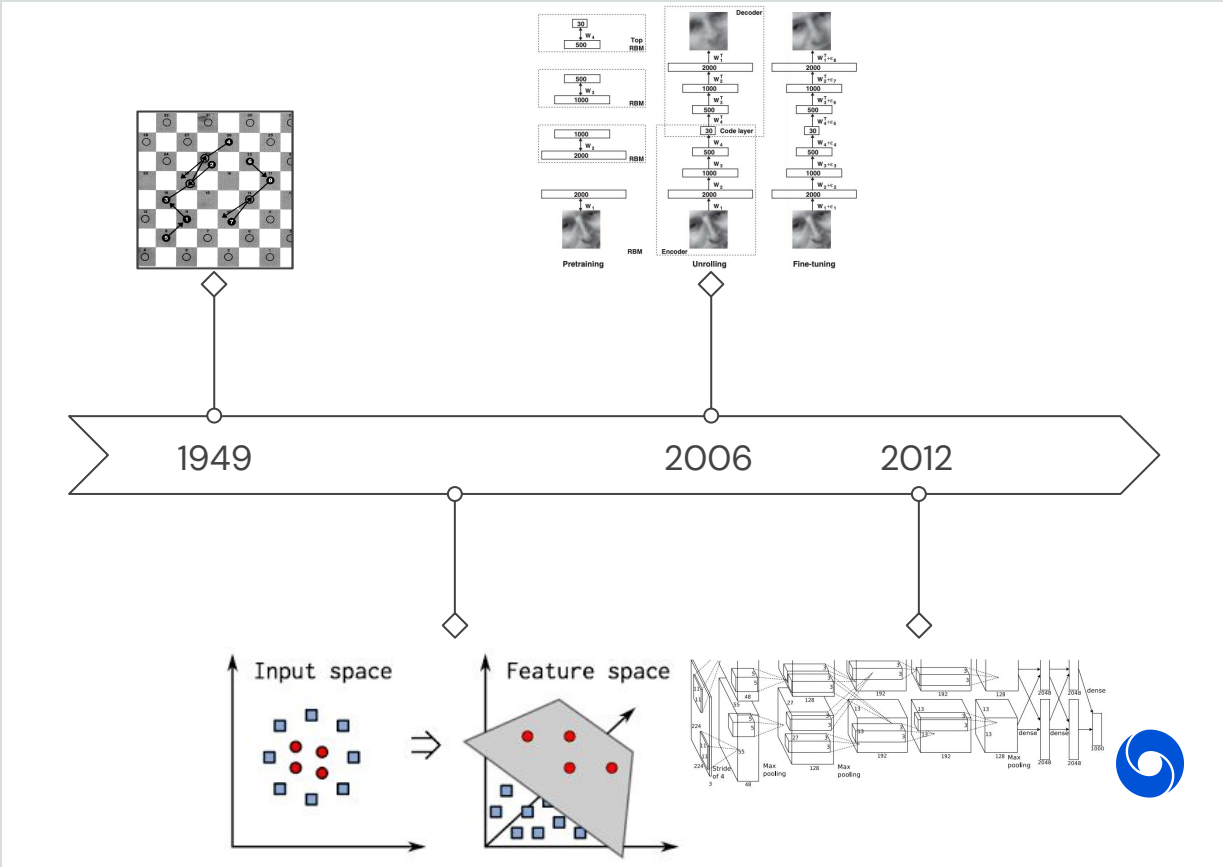
Reducing the Dimensionality of Data with Neural Networks, Hinton and Salakhutdinov, Science (2006)

- Arthur Samuel coins the term “machine learning”
- Feature engineering and kernel methods
- Restricted Boltzmann Machines used for initialising deep classifiers



History of representation learning

- Arthur Samuel coins the term “machine learning”
- Feature engineering and kernel methods
- Restricted Boltzmann Machines used for initialising deep classifiers
- AlexNet wins ImageNet challenge by a large margin with no unsupervised pre-training



Want to learn more?

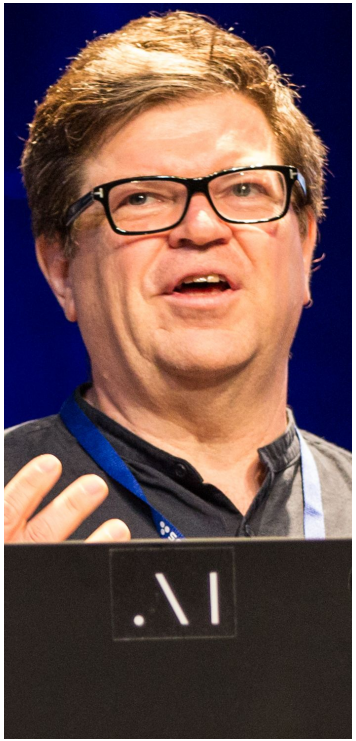
ImageNet Classification with Deep Convolutional Neural Networks, Krizhevsky et al, NeurIPS (2012)

Turing Award winners at AAAI 2020

> I always knew unsupervised learning was the right thing to do
>
> — Geoff Hinton

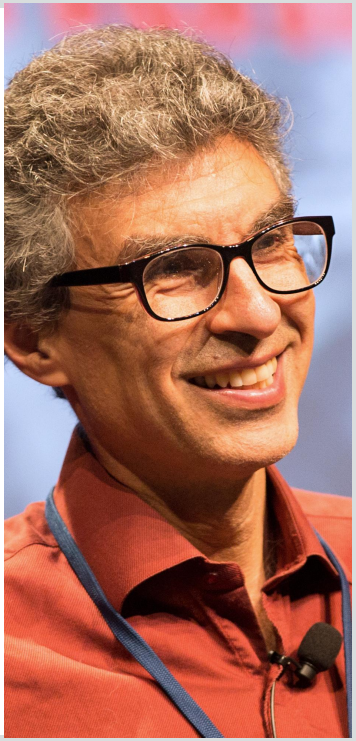
> Basically it's the idea of learning to represent the world before learning a task — and this is what babies do
>
> — Yann LeCun

> And so if we can build models of the world where we have the right abstractions, where we can pin down those changes to just one or a few variables, then we will be able to adapt to those changes because we don't need as much data, as much observation in order to figure out what has changed.
>
> — Yoshua Bengio



Jérémy Barande / Ecole polytechnique Université Paris-Saclay / CC BY-SA 2.0

Eviatar Bach / CC BY-SA



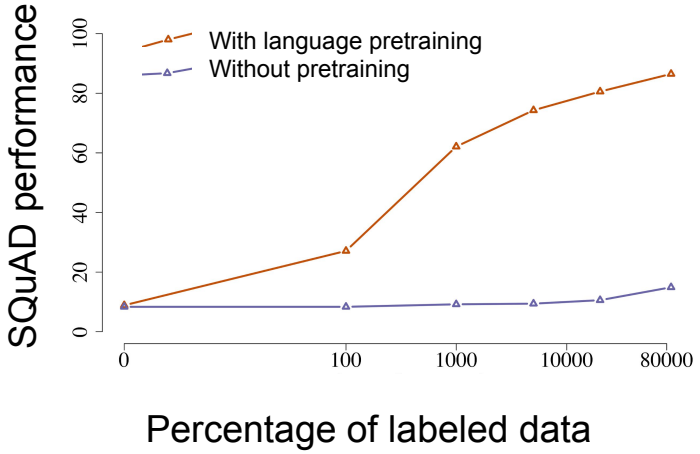
Jérémy Barande / Ecole polytechnique Université Paris-Saclay / CC BY-SA 2.0

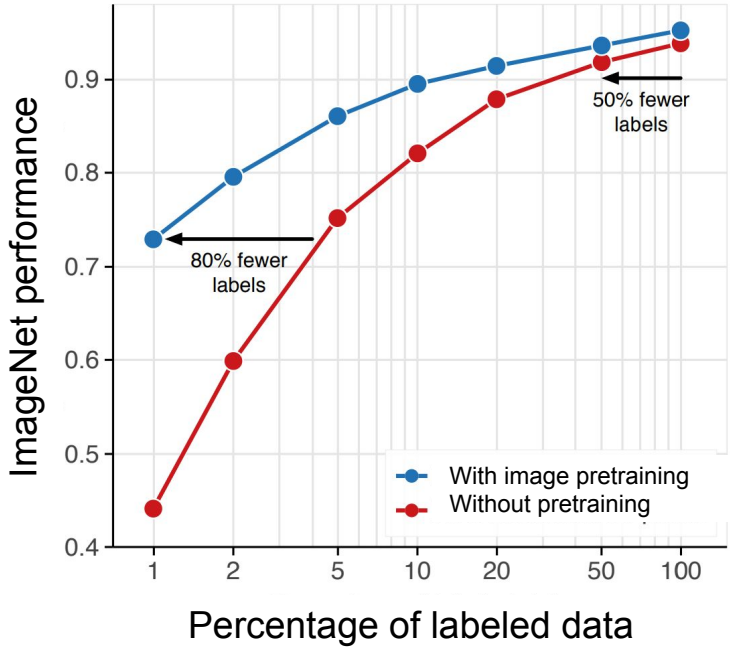
Inflection point

Want to learn more?

Learning and Evaluating General Linguistic Intelligence, Yogatama et al (2019)

Data-Efficient Image Recognition with Contrastive Predictive Coding, Olivier J. Hénaff et al, ICML (2020)





DeepMind

3 Building Blocks

The representation problem

Realm of **Images**

Representation
(Understanding)

Generation
(Imagination)

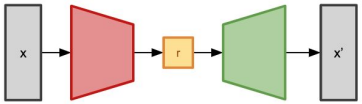
Realm of **Representations**

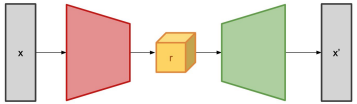
There is a zoo of models. Not all models matter.

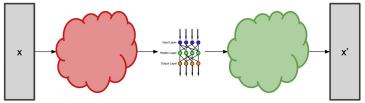
Hot take!

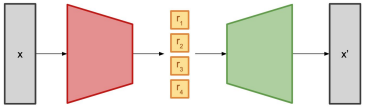
Curtis's Botanical Magazine, 1831

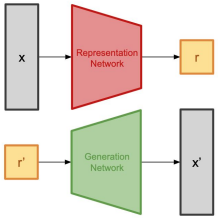
Model zoo

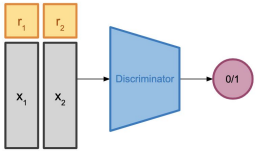

x
r
x'

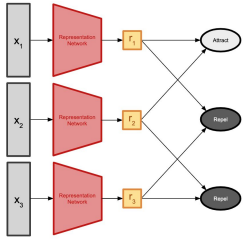

x
r
x'

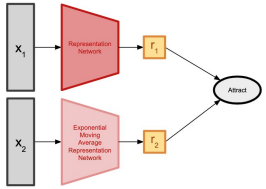

x
x'

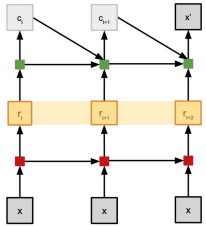

x
x'

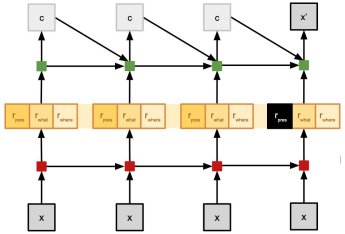

x
Representation Network
r
r'
Generation Network
x'

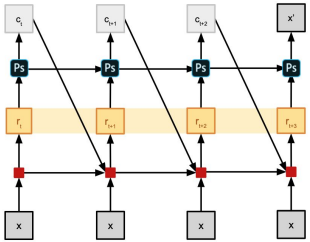

r1
r2
x1
x2
Discriminator
0/1


x1
x2
x3
Representation Network
r1
r2
r3
Attract
Repel
Repel

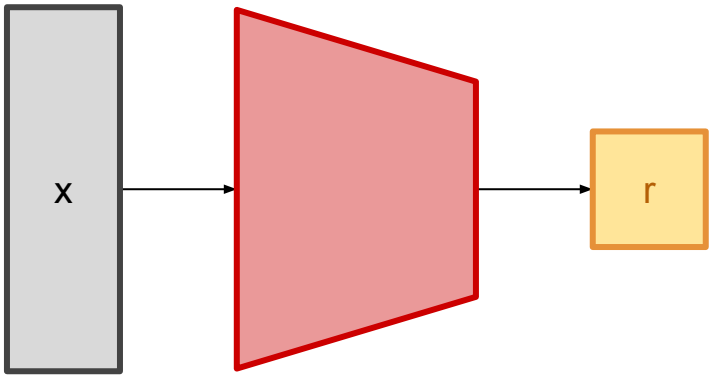

x1
x2
Representation Network
Exponential Moving Average Representation Network
r1
r2
Attract


x'
x
x
x


c
c
c
x'
x
x
x
x


Ps
Ps
Ps
Ps
x'
x
x
x
x

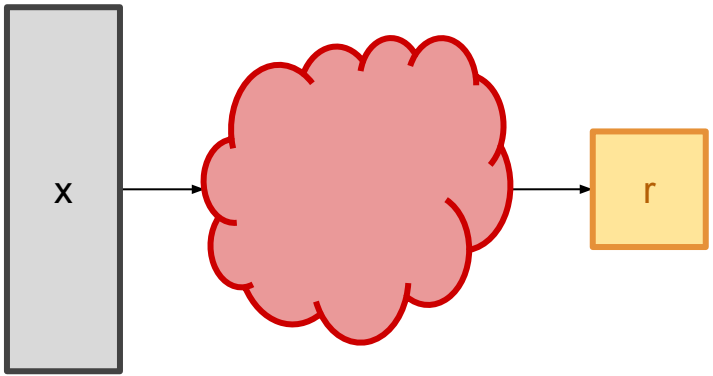
(Representation / Encoder / Inference) Networks



- **Size:** Smaller or larger than x
- **Structure:** Flat or interpretable
- **Type:** Continuous or discrete
- **Shape:** Fixed or variable
- **Disentangled** or not

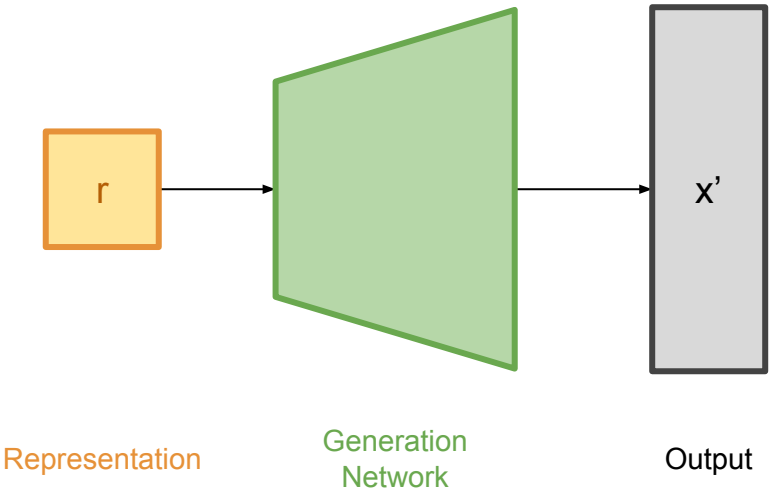
- Multi-layer perceptron
- ConvNet
- Transformer
- Recurrent neural net

(Representation / Encoder / Inference) Networks



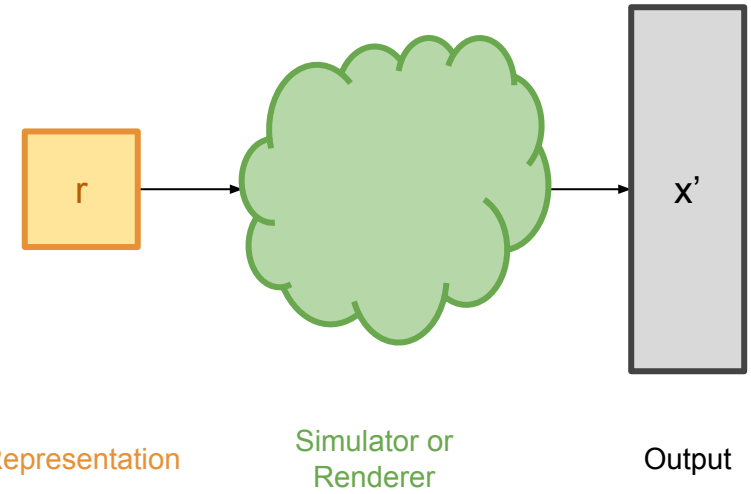
- Differentiable or not
- Interpretable or not
- Deterministic or stochastic

(Generation / Generator / Decoder) Networks



- Multi-layer perceptron
- DeconvNet
- Transformer
- Recurrent neural net

(Generation / Generator / Decoder) Networks



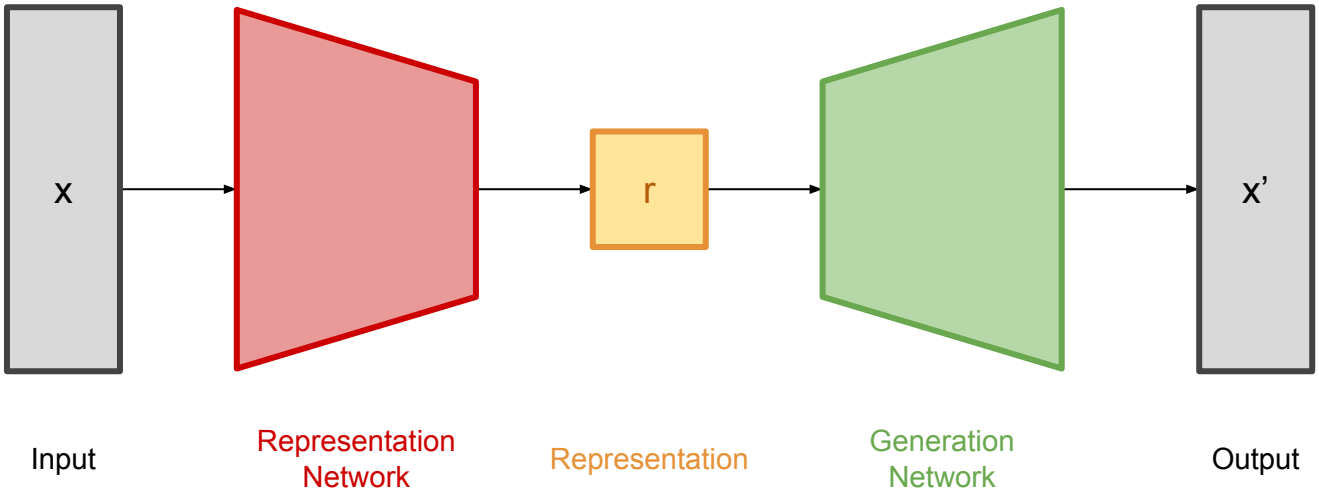
- Differentiable or not
- Interpretable or not
- Deterministic or stochastic

Autoencoders

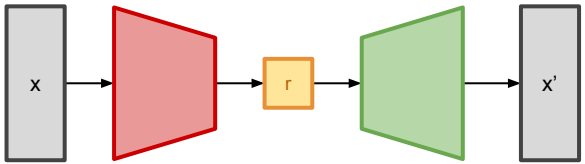
Want to learn more?

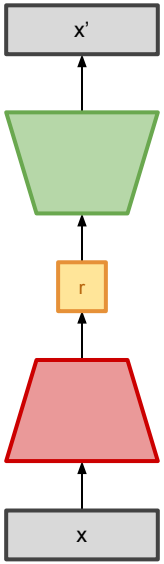
Auto-Encoding Variational Bayes, Kingma et al, ICLR (2014)

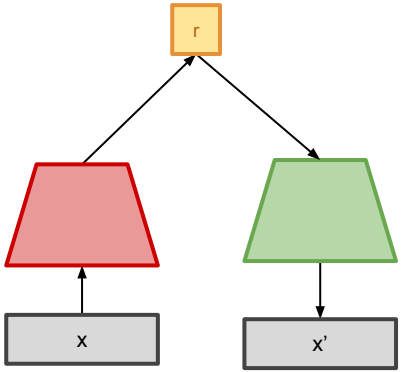
Stochastic backpropagation and approximate inference in deep generative models, Rezende et al, ICML (2014)



Autoencoder Graphics





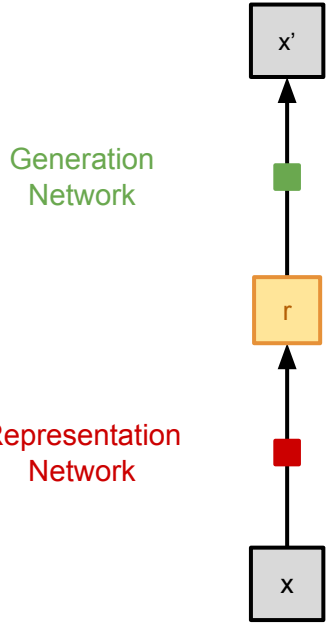


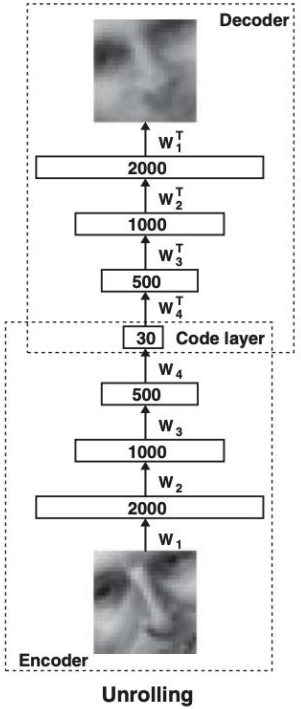
Autoencoders: What are they for?

- Density estimation
- Dimensionality reduction
- Image generation
- Denoising
- **Representation learning**

Want to learn more?

Reducing the Dimensionality of Data with Neural Networks, Hinton et al, Science (2006)



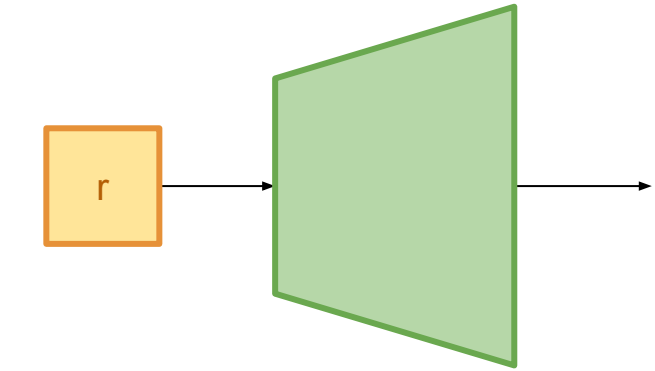


Unrolling

Autoencoders

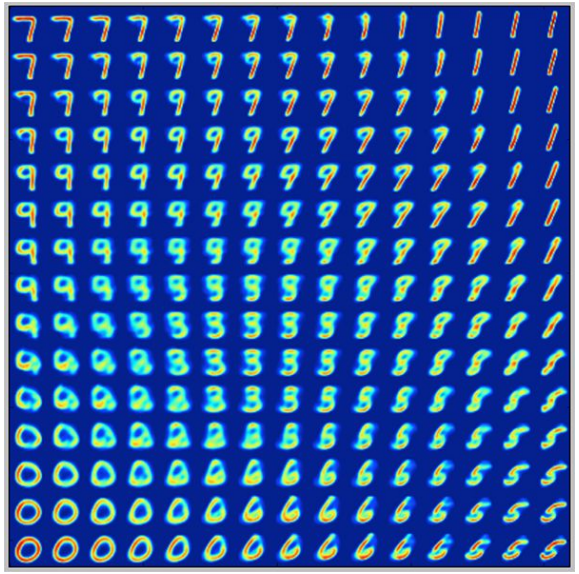
Want to learn more?

Building Autoencoders in Keras, Chollet (2016)

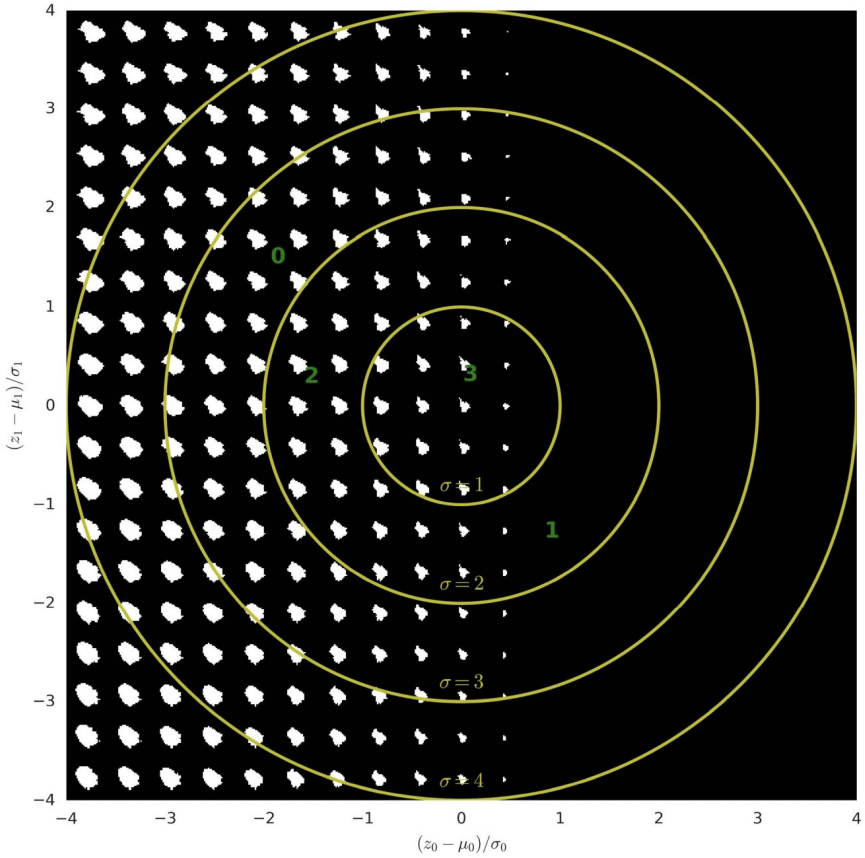


Representation

Generation Network



Conditional autoencoders



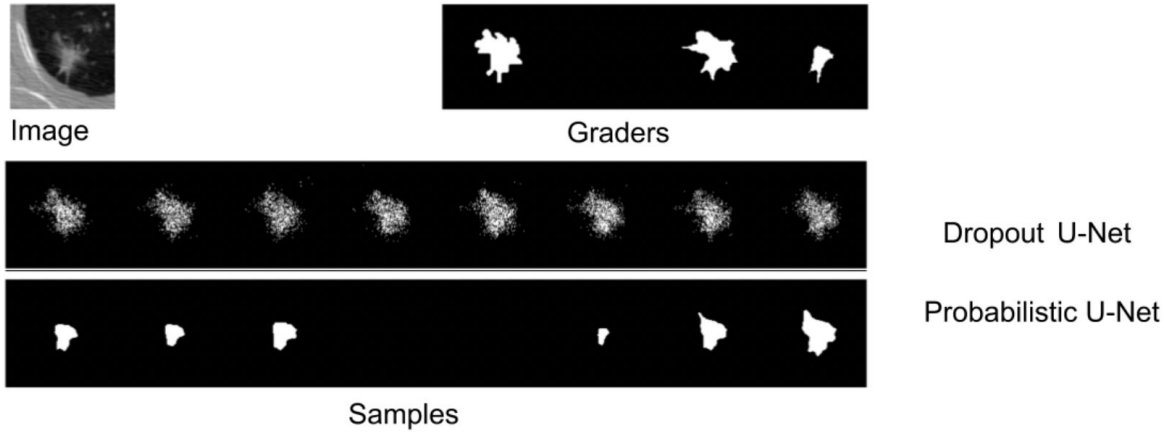
Want to learn more?

A Probabilistic U-Net for Segmentation of Ambiguous Images, Kohl et al (2018)

Conditional autoencoders

Want to learn more?

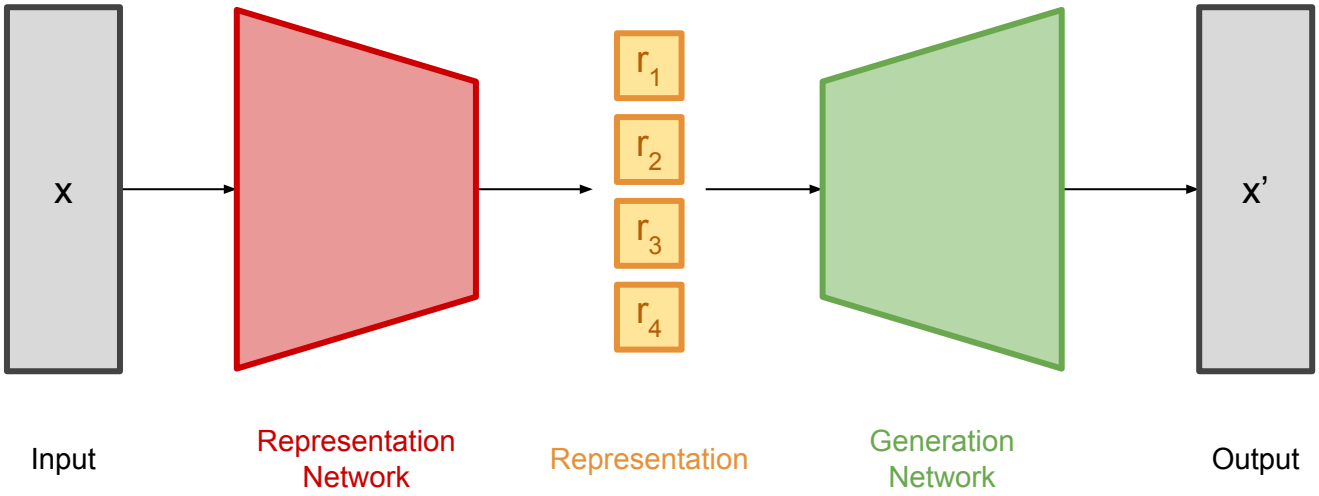
A Probabilistic U-Net for Segmentation of Ambiguous Images, Kohl et al (2018)



Disentangled Autoencoders

Want to learn more?

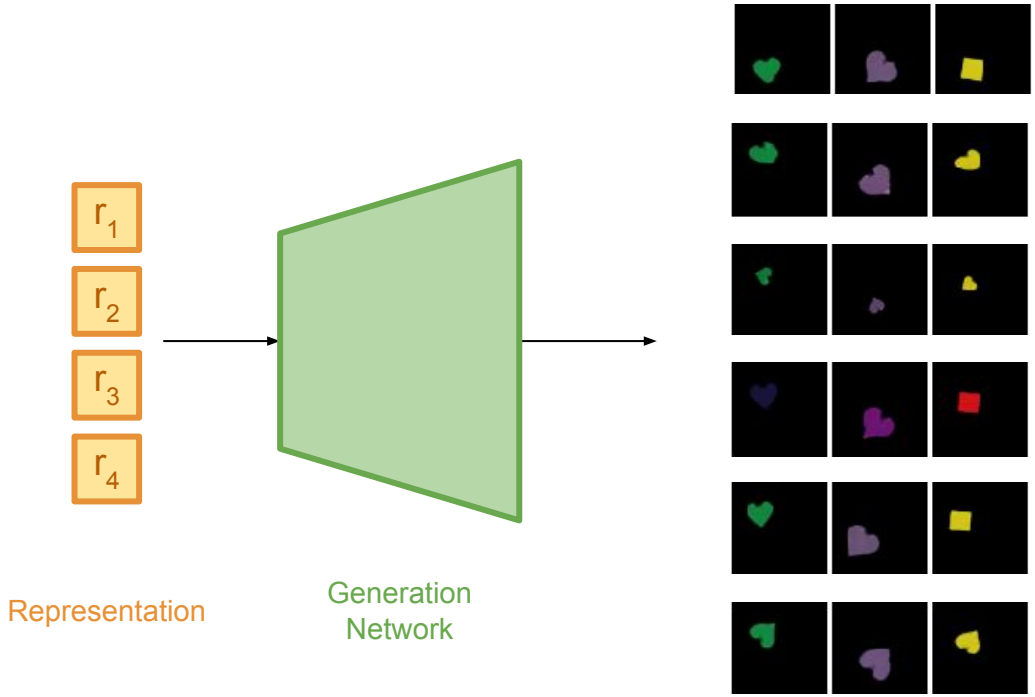
β-VAE: Learning Basic Visual Concepts with a Constrained Variational Framework, Higgins et al., ICLR 2017



Disentangled Autoencoders

Want to learn more?

β-VAE: Learning Basic Visual Concepts with a Constrained Variational Framework, Higgins et al., ICLR 2017



GIFs adapted from Chris Burgess

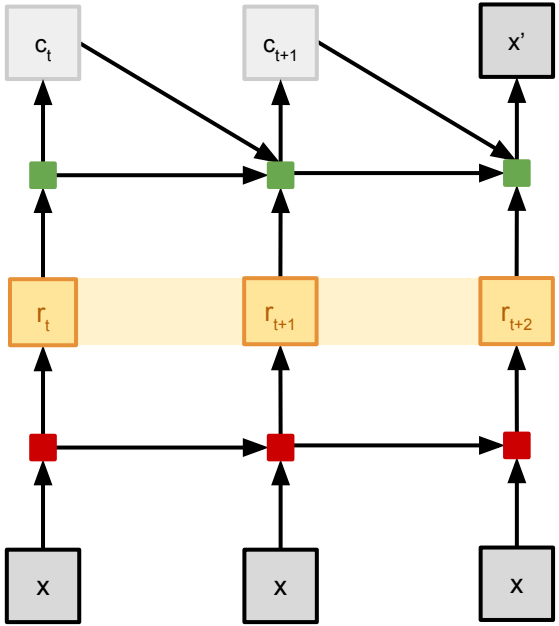
Sequential Autoencoders

Want to learn more?

Towards Conceptual Compression, Gregor et al, NeurIPS (2016)

Generation Network

Representation Network



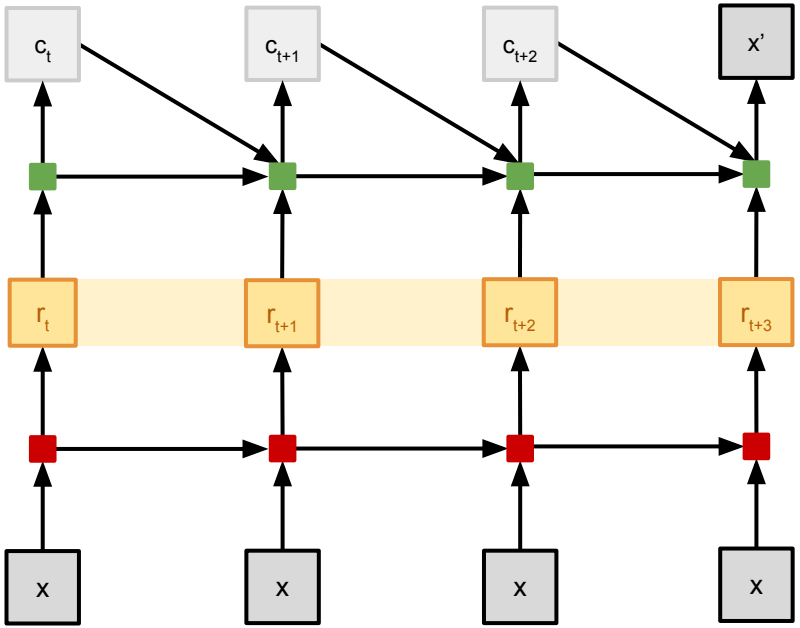
Generation RNN

Representation RNN

Sequential Autoencoders

Want to learn more?

Towards Conceptual Compression, Gregor et al, NeurIPS (2016)



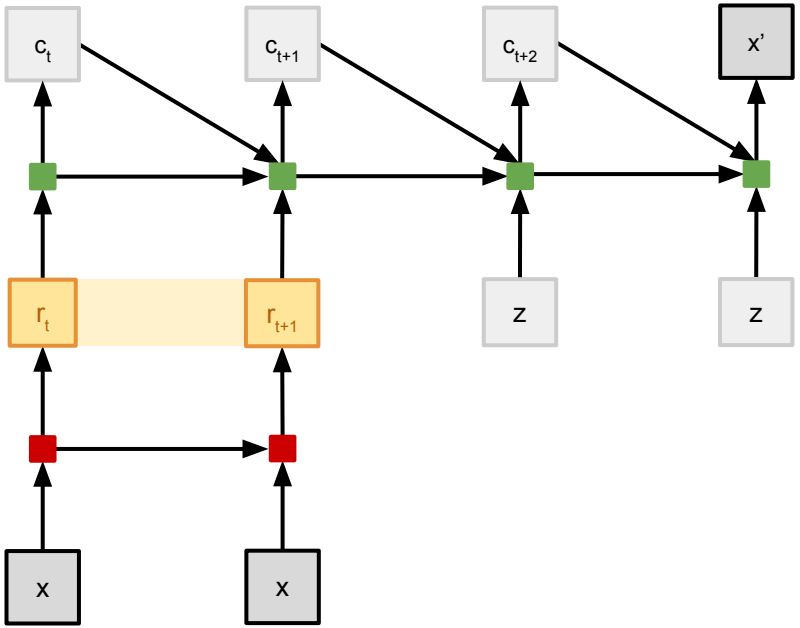
Generation RNN

Representation RNN

Sequential Autoencoders

Want to learn more?

Towards Conceptual Compression, Gregor et al, NeurIPS (2016)



Generation RNN

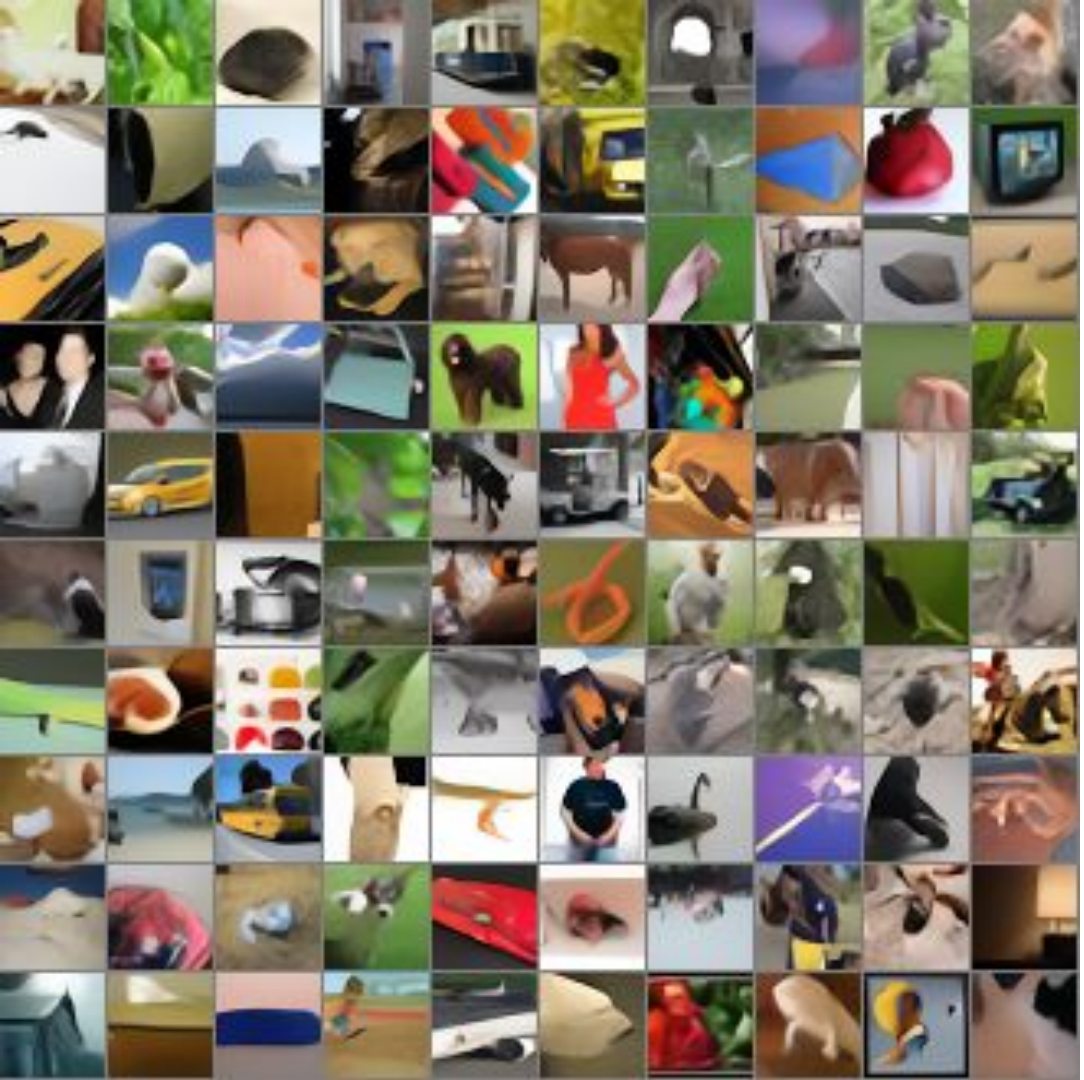
Representation RNN

76 bits

Original raw image:
24576 bits

112 bits

221 bits



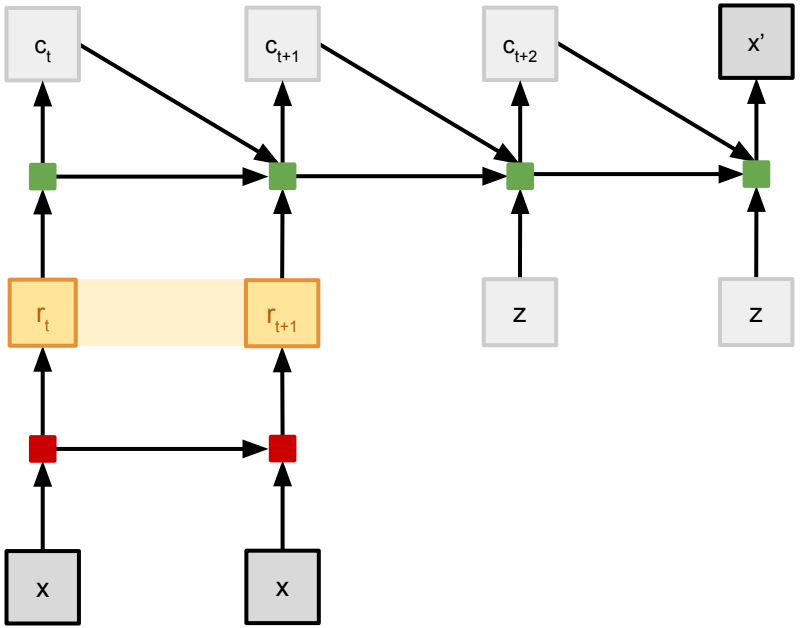
380 bits

2364 bits

Sequential Autoencoders

Want to learn more?

Towards Conceptual Compression, Gregor et al, NeurIPS (2016)



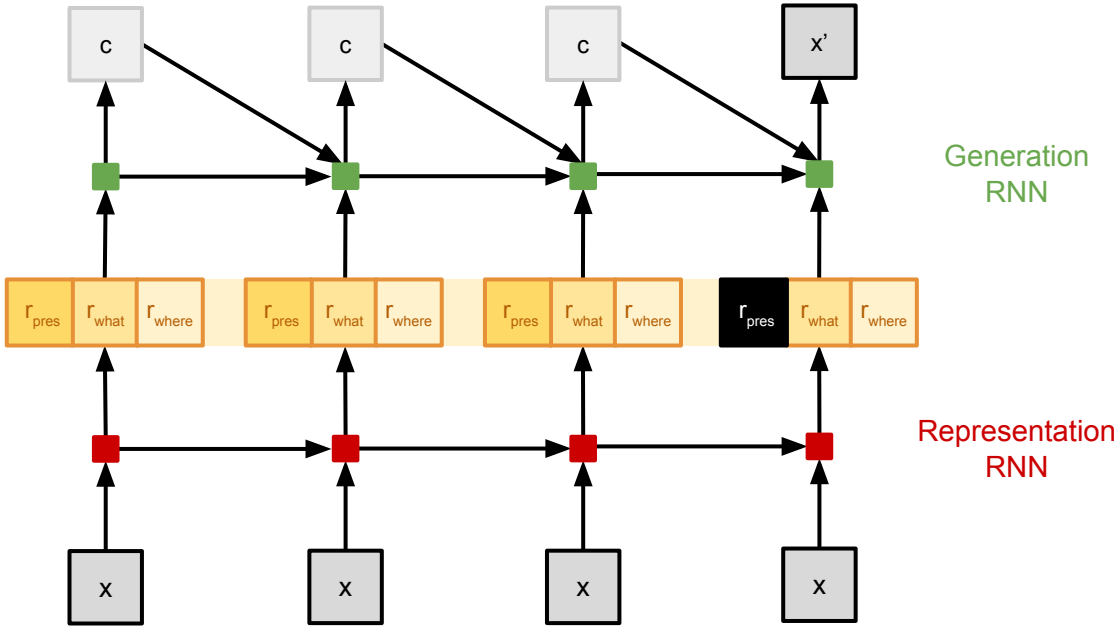
Generation RNN

Representation RNN

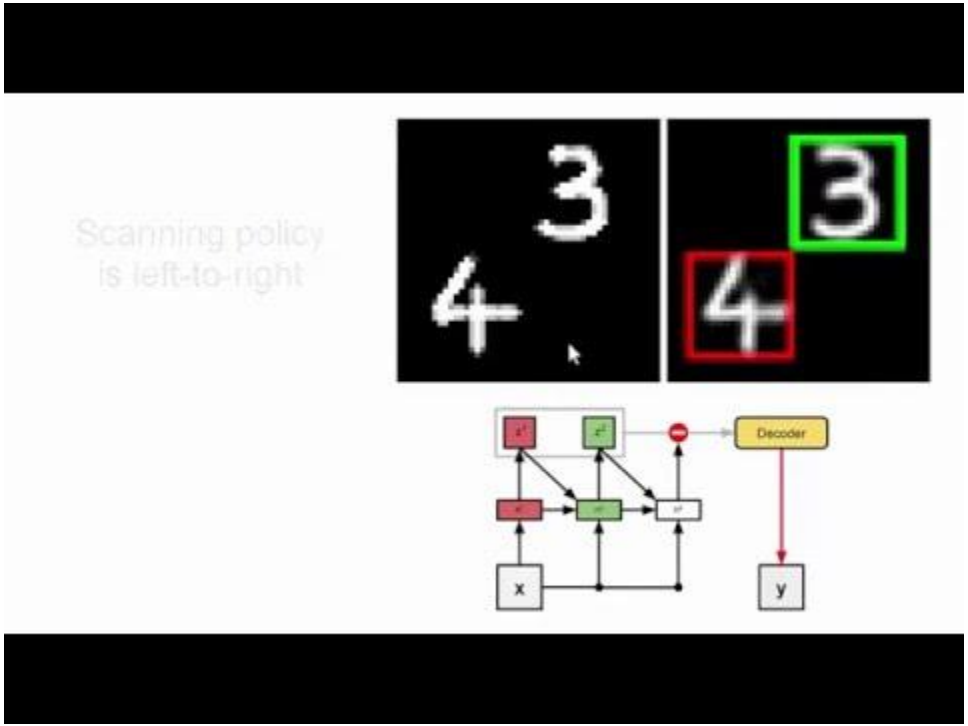
Variable Length, Interpretable, Sequential Autoencoders

Want to learn more?

Attend, Infer, Repeat, Eslami et al, NeurIPS (2016)



Variable Length, Interpretable, Sequential Autoencoders



Want to learn more?

Attend, Infer, Repeat, Eslami et al, NeurIPS (2016)

Structure is an illusion

Hot take!

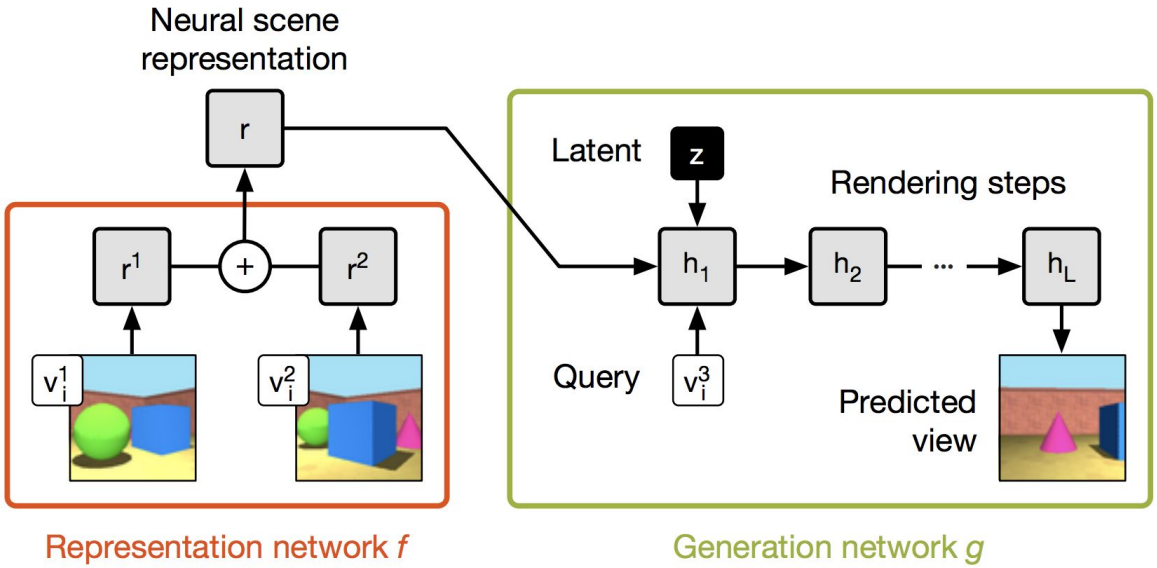
Structure is an illusion

Hot take!

Generative Query Networks

Want to learn more?

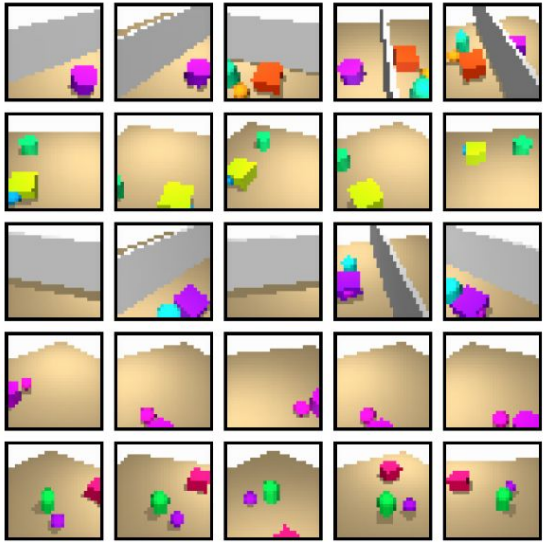
Neural scene representation and rendering, Eslami et al, Science (2018)



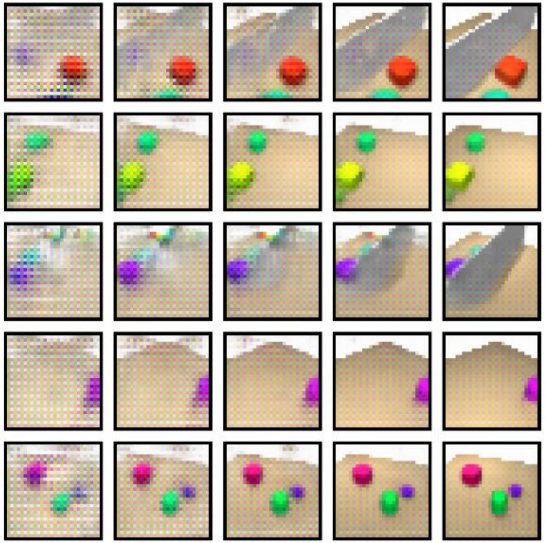
GQN: Accurate generation

Want to learn more?

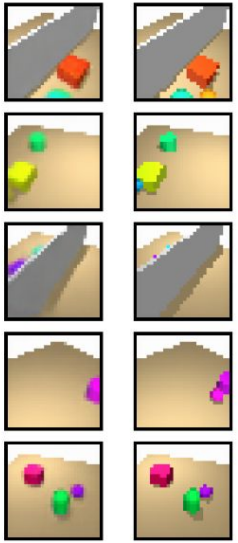
Neural scene representation and rendering, Eslami et al, Science (2018)



Observations



Generation steps

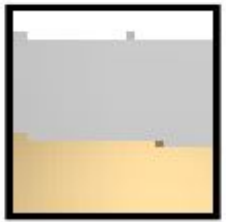


Pred Truth

GQN: Capturing uncertainty

Want to learn more?

Neural scene representation and rendering, Eslami et al, Science (2018)



Observation

Samples

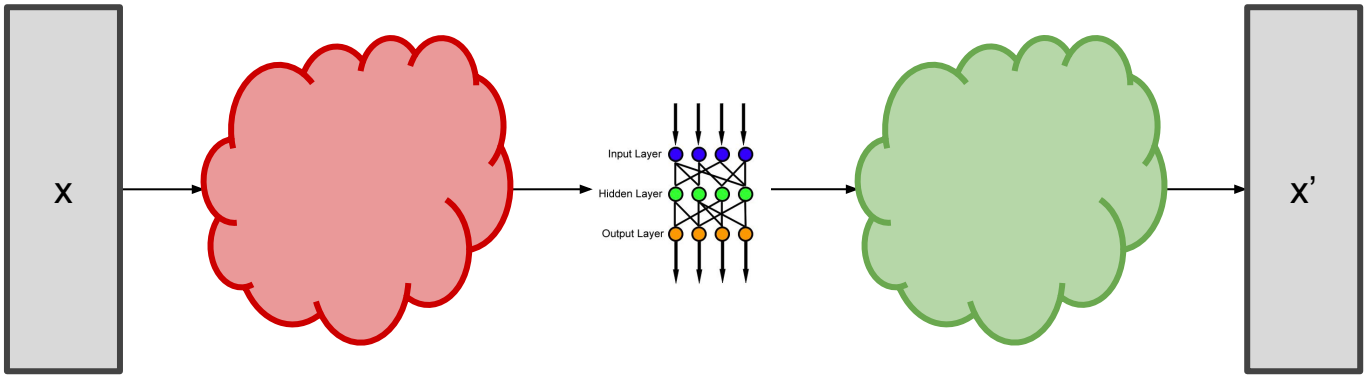
observations

ground truth | neural rendering | neural rendering | map

NeRF

Want to learn more?

NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, Mildenhall (2020)



Input

Inference Optimisation Algorithm

Neural Network Representation

Forward Propagation

Output

NeRF

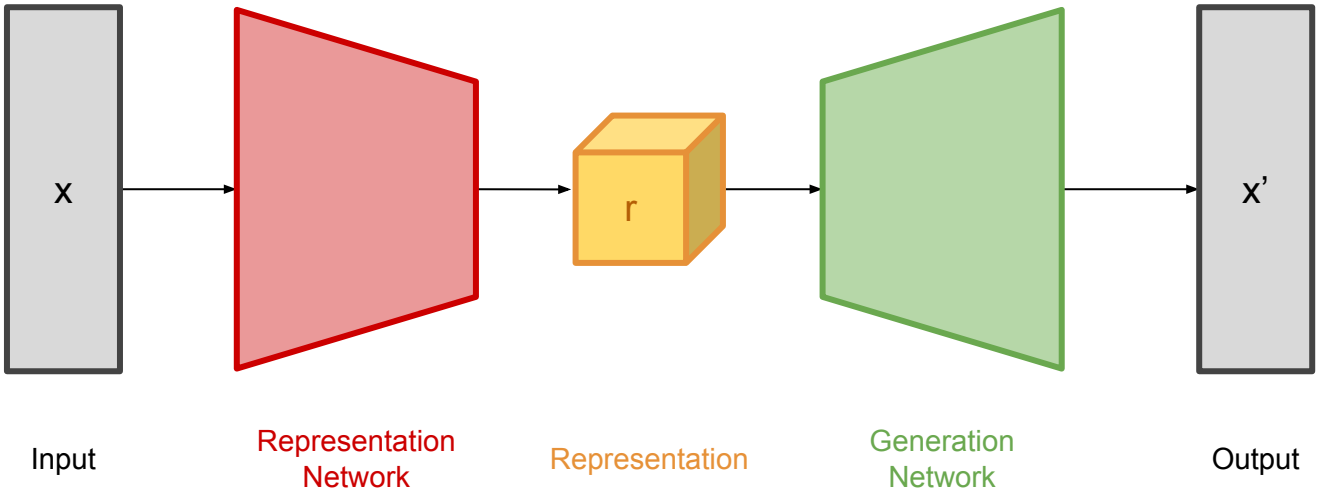
Want to learn more?

NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, Mildenhall (2020)

Voxel Autoencoders

Want to learn more?

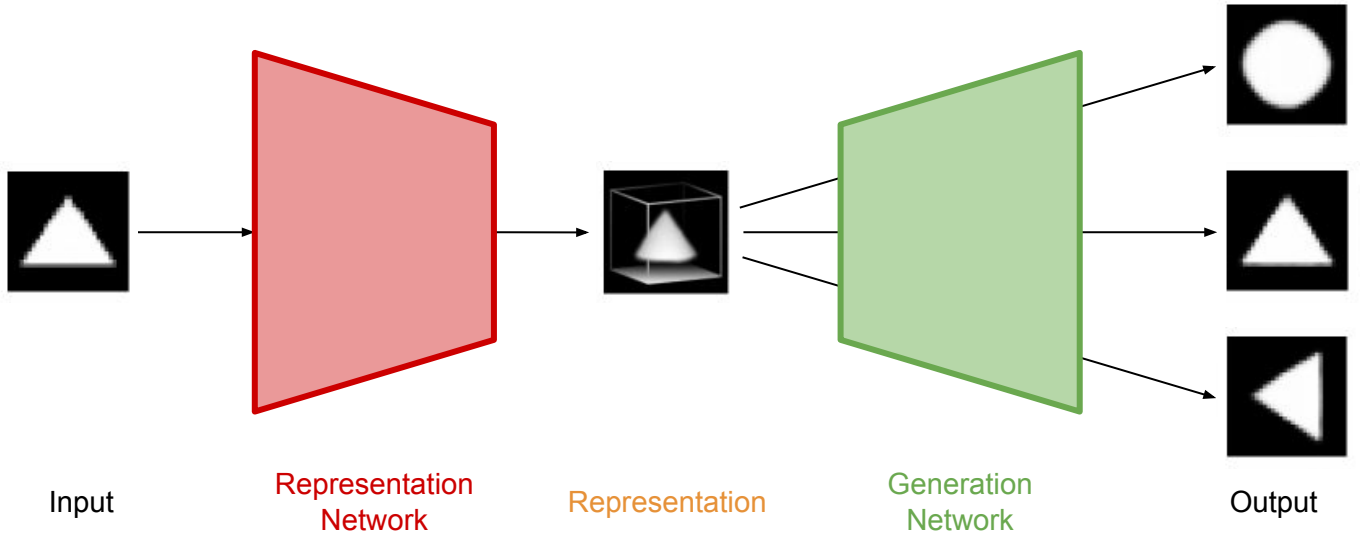
Unsupervised Learning of 3D Structure from Images, Rezende et al, NeurIPS (2016)



Voxel Autoencoders

Want to learn more?

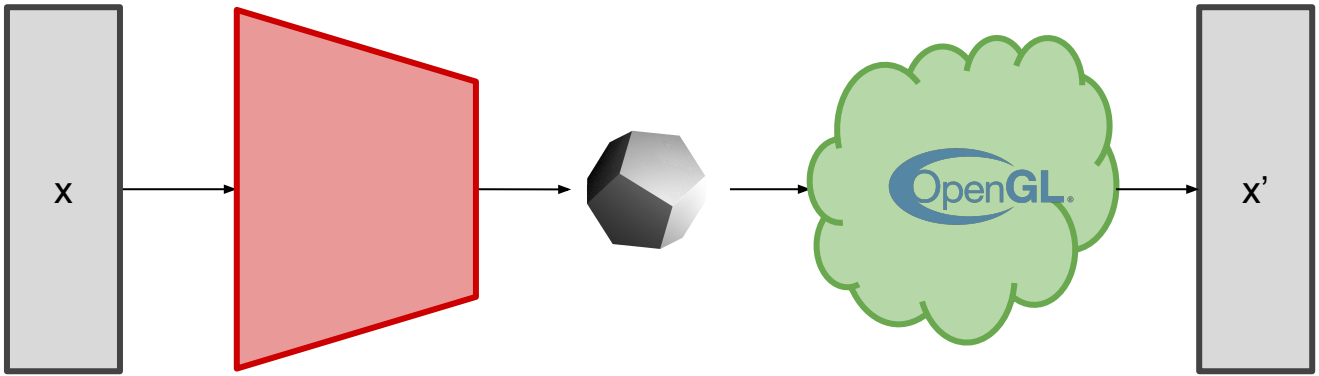
Unsupervised Learning of 3D Structure from Images, Rezende et al, NeurIPS (2016)



Mesh Autoencoders

Want to learn more?

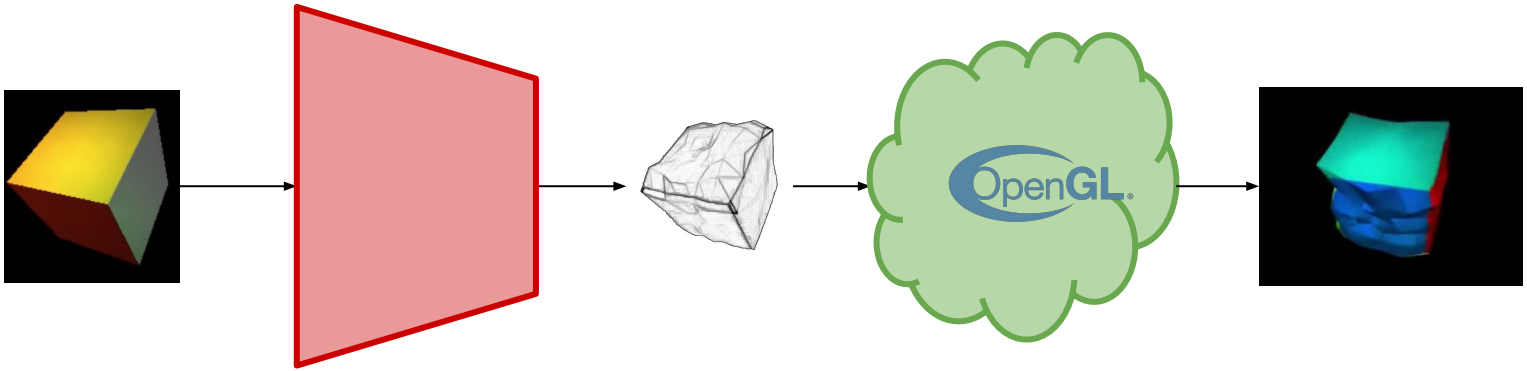
Unsupervised Learning of 3D Structure from Images, Rezende et al, NeurIPS (2016)



Mesh Autoencoders

Want to learn more?

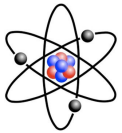
Unsupervised Learning of 3D Structure from Images, Rezende et al, NeurIPS (2016)



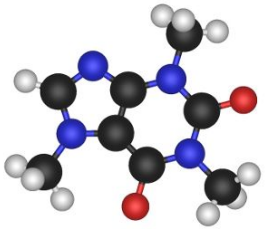
DeepMind

4 Intermission: Visual Protein Folding

Biology 101

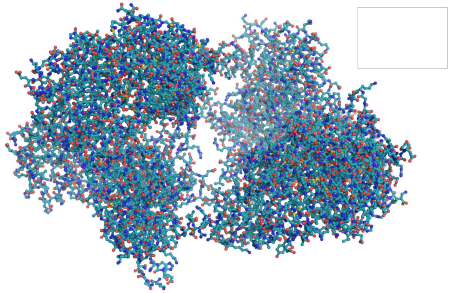


Atoms



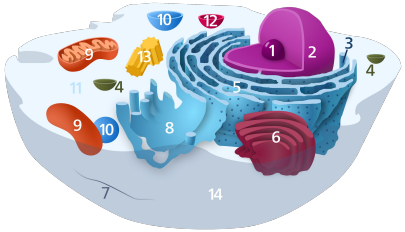
Molecules

Tens of atoms in a particular type of molecule called ‘amino acids’



Proteins

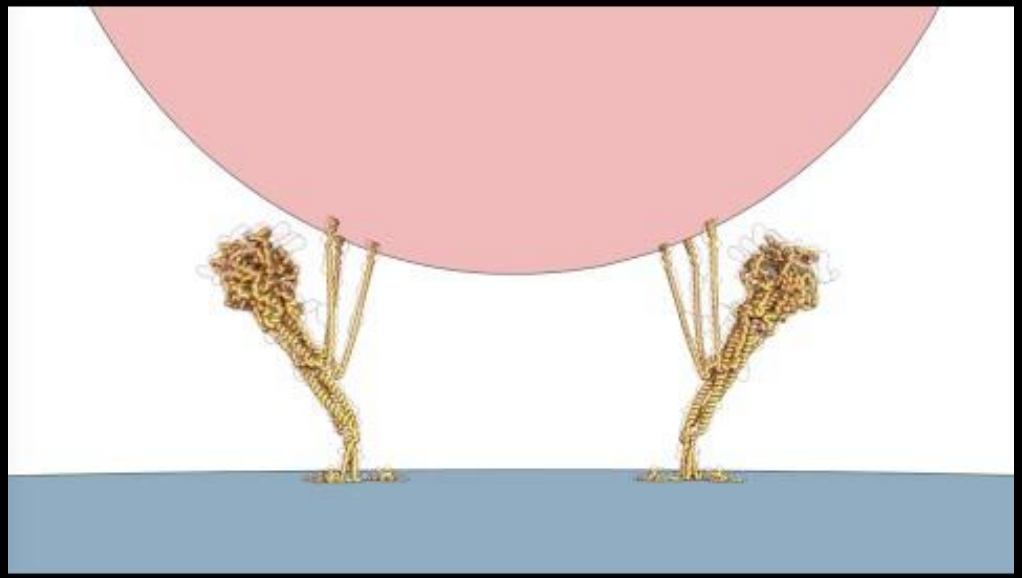
Hundreds to tens of thousands of amino acids in a typical protein



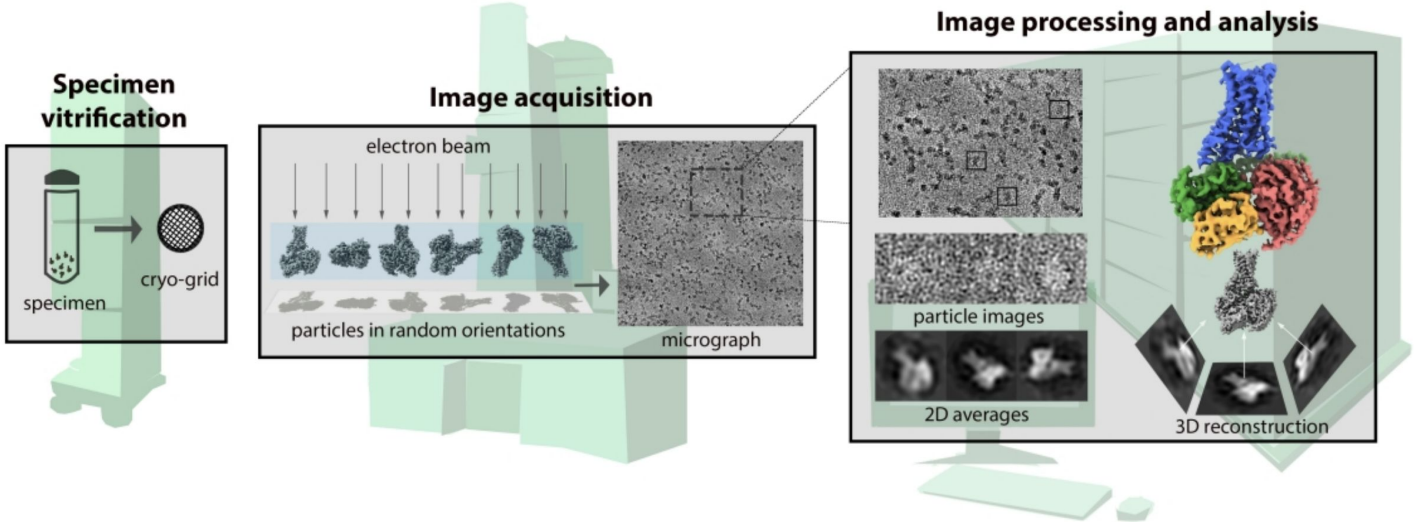
Organisms

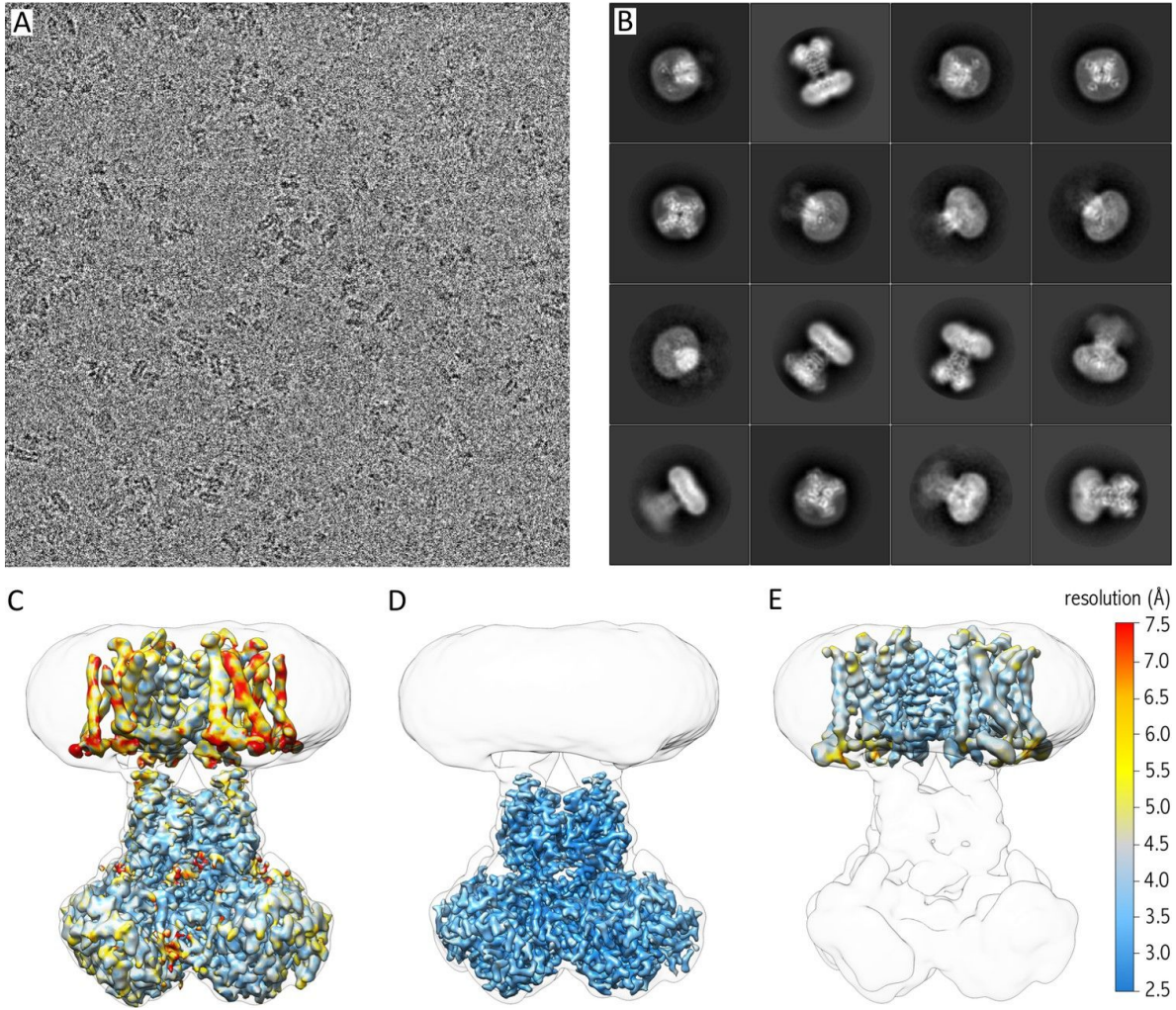
Tens of millions of proteins in a simple cell

Coronavirus spike proteins in action

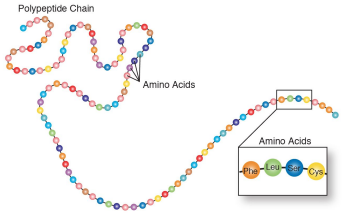


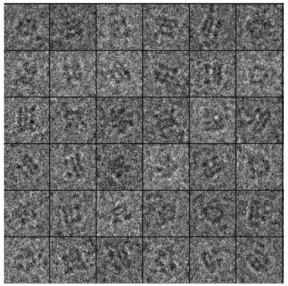
Cryo Electron Microscopy




A
B
C
D
E
resolution (Å)
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5

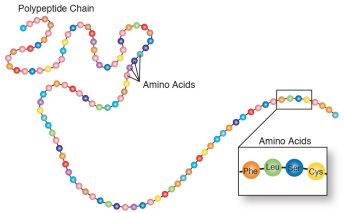
Traditional pipeline

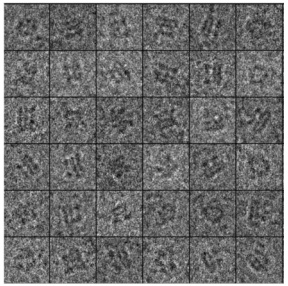


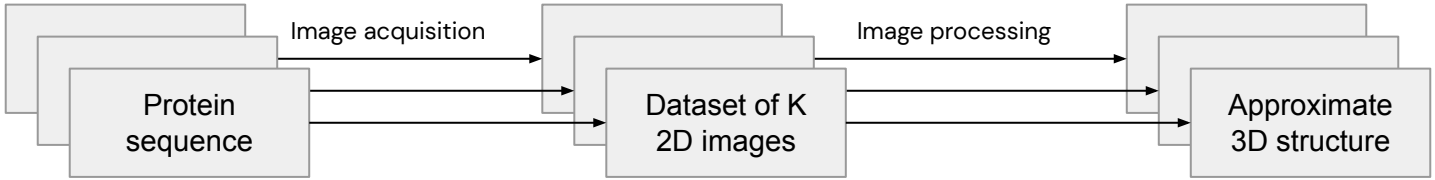


Single protein sequence → Image acquisition → Dataset of K 2D images → Image processing → Single approx. 3D structure

Traditional pipeline







N proteins

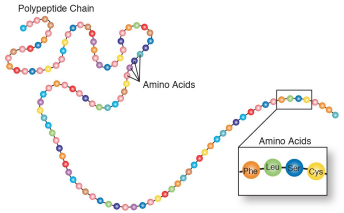
N datasets

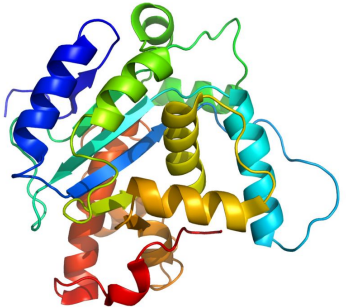
N structures

Supervised training: AlphaFold

Want to learn more?

Highly accurate protein structure prediction with AlphaFold, Jumper et al (2021)

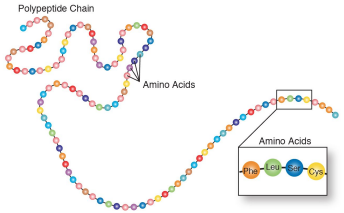


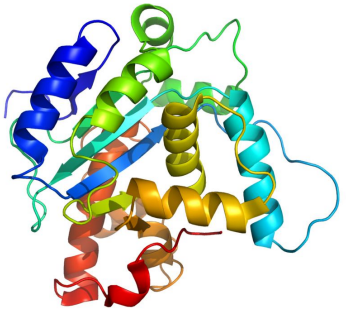


Protein sequence → Neural Net → Approximate 3D structure

1 model

Traditional pipeline





Protein sequence → Neural Net → Approximate 3D structure

Single structure is not best representation of protein

In reality, each protein appears as a distribution of structures in the dataset

Of course, no ground truth labels available

Hot take!

OPINION

Why AlphaFold won't revolutionise drug discovery

BY DEREK LOWE | 5 AUGUST 2022

Protein structure prediction is a hard problem, but even harder ones remain

Approximate 3D structure

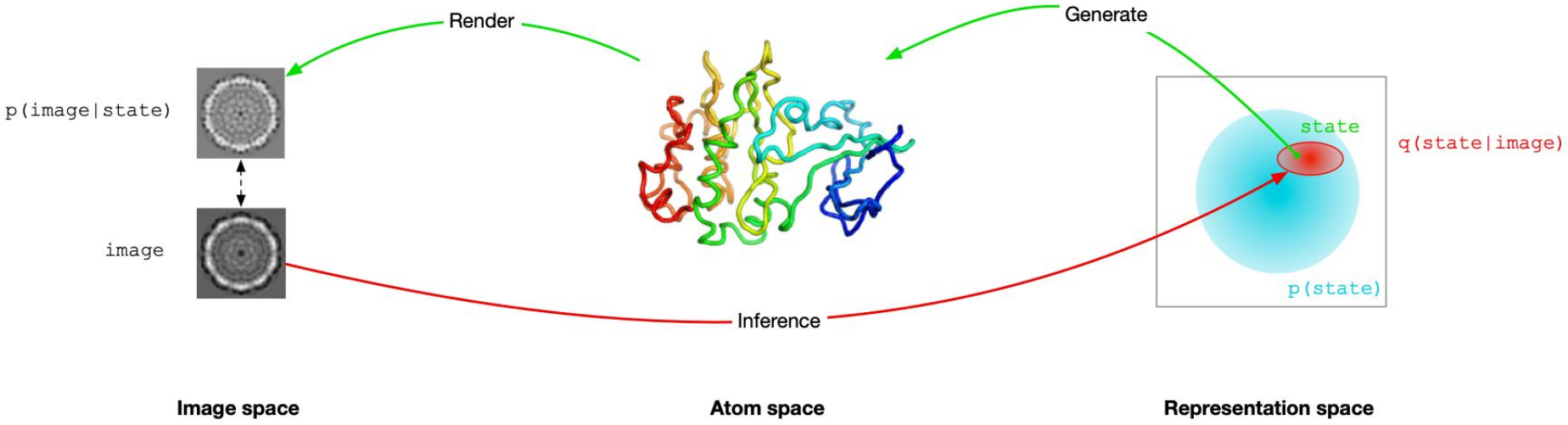
Single structure is not best representation of protein

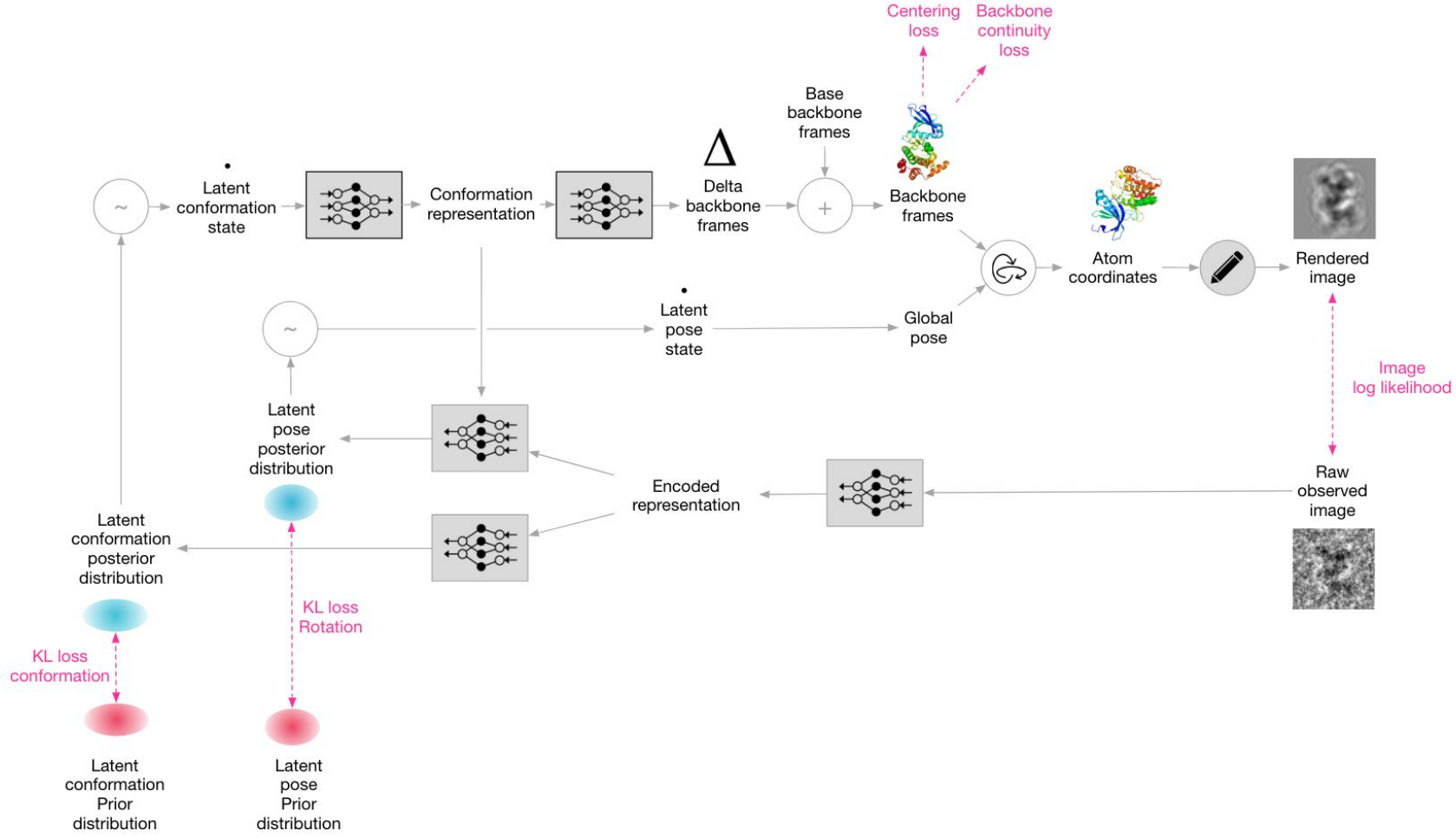
In reality, each protein appears as a distribution of structures in the dataset

Of course, no ground truth labels available

Want to learn more?

Inferring a Continuous Distribution of Atom Coordinates from Cryo-EM Images using VAEs, Rosenbuam et al (2021)

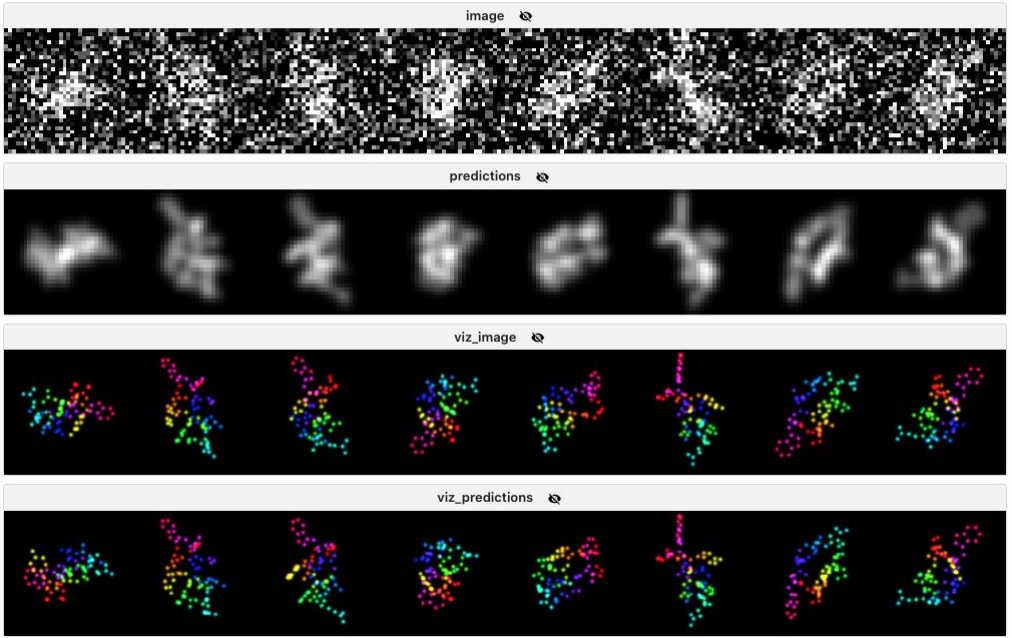



Centering loss
Backbone continuity loss
Base backbone frames
Latent conformation state
Conformation representation
Delta backbone frames
Backbone frames
Atom coordinates
Rendered image
Latent pose state
Global pose
Image log likelihood
Latent pose posterior distribution
Encoded representation
Raw observed image
Latent conformation posterior distribution
KL loss Rotation
KL loss conformation
Latent conformation Prior distribution
Latent pose Prior distribution

Results on toy data: Chignolin

Want to learn more?

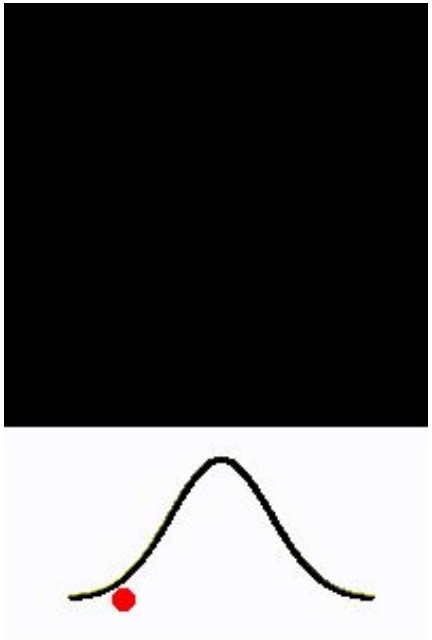
Inferring a Continuous Distribution of Atom Coordinates from Cryo-EM Images using VAEs, Rosenbuam et al (2021)



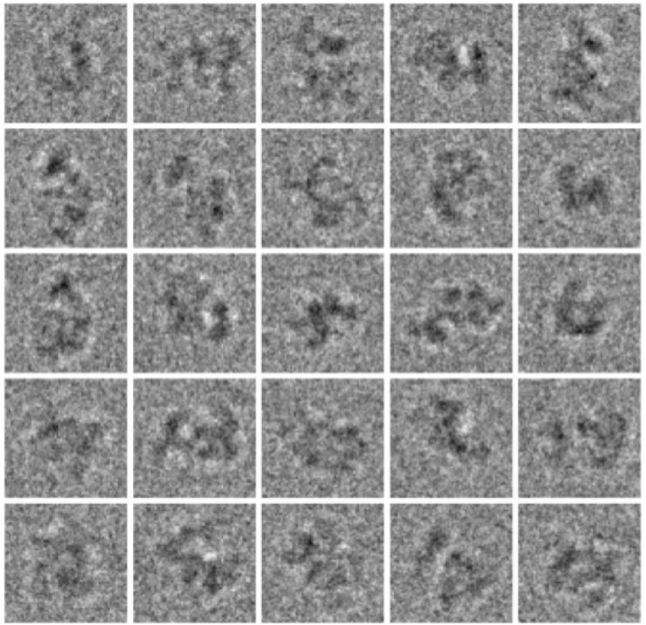
Results on toy data: Chignolin

Want to learn more?

Inferring a Continuous Distribution of Atom Coordinates from Cryo-EM Images using VAEs, Rosenbuam et al (2021)



Results on more realistic data: Aurora A Kinase



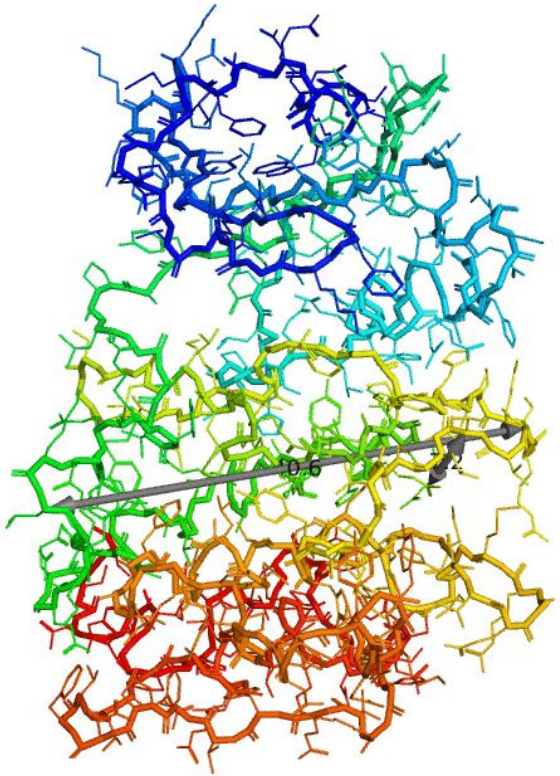
Want to learn more?

Inferring a Continuous Distribution of Atom Coordinates from Cryo-EM Images using VAEs, Rosenbuam et al (2021)

Results on more realistic data: Aurora A Kinase

Want to learn more?

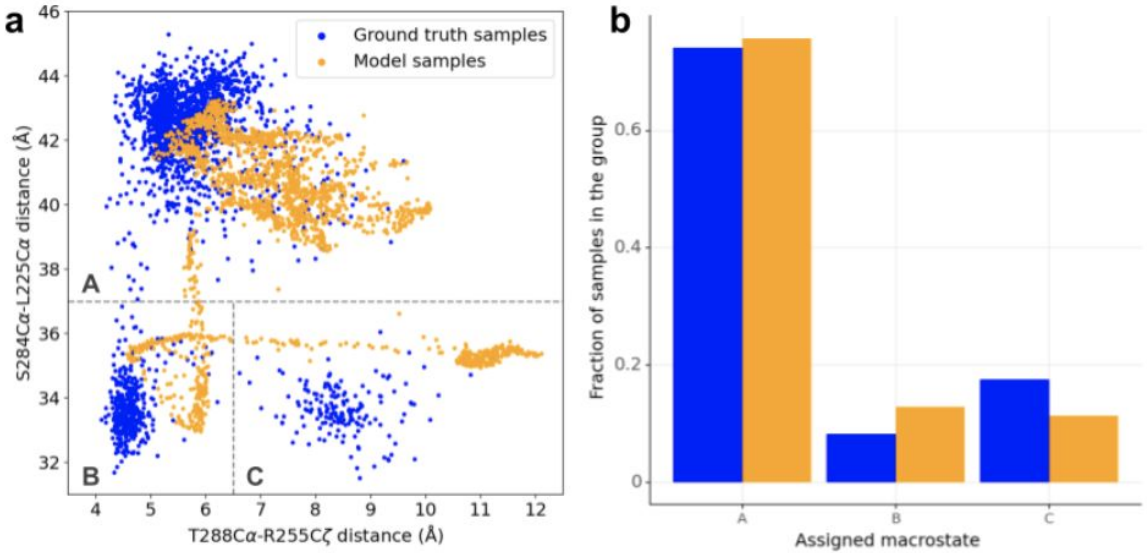
Inferring a Continuous Distribution of Atom Coordinates from Cryo-EM Images using VAEs, Rosenbuam et al (2021)



Results on more realistic data: Aurora A Kinase

Want to learn more?

Inferring a Continuous Distribution of Atom Coordinates from Cryo-EM Images using VAEs, Rosenbuam et al (2021)



Beyond likelihood-based

Discriminators / Contrastive Networks

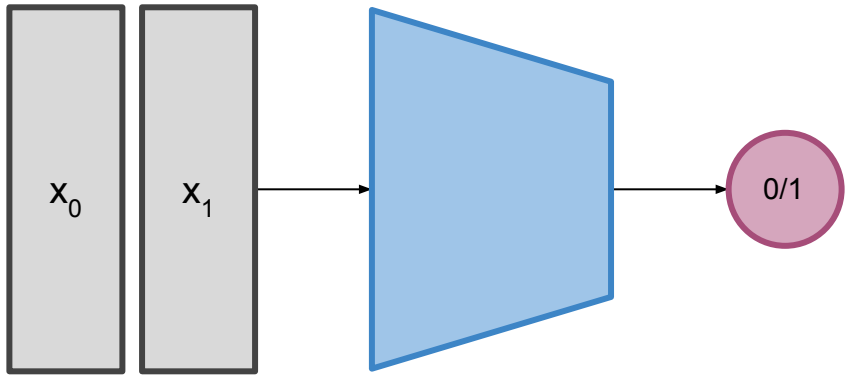


Diagram inspired by Lilian Weng, https://lilianweng.github.io

Generative adversarial networks

Want to learn more?

Generative adversarial networks. Goodfellow, et al. NeurIPS (2014)

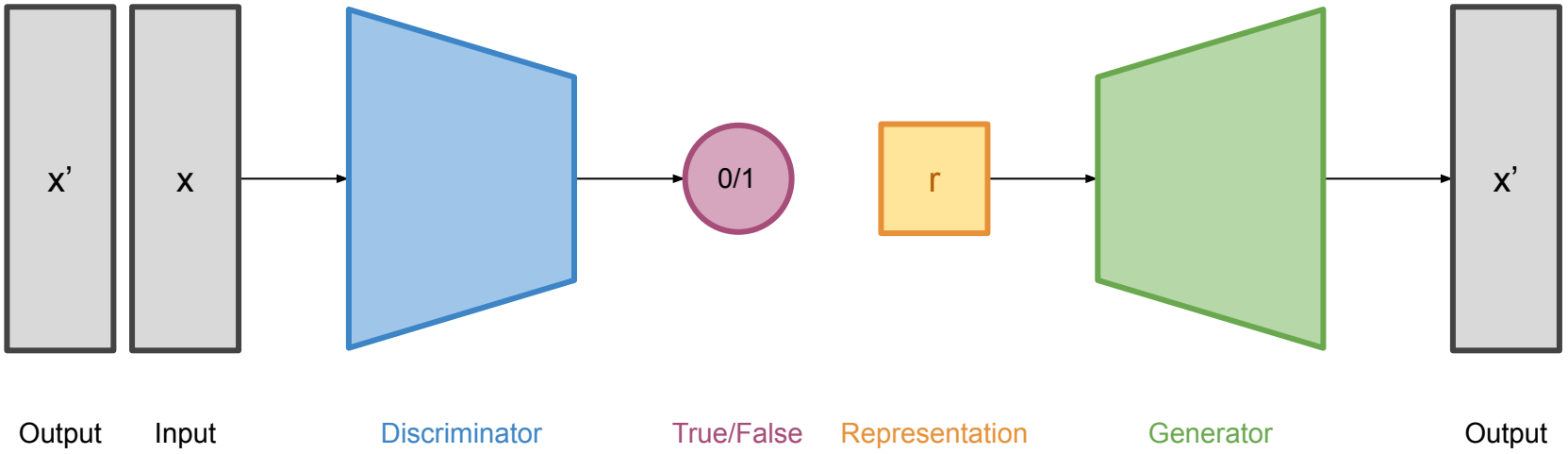


Diagram inspired by Lilian Weng, https://lilianweng.github.io

Generative adversarial networks

Want to learn more?

A Style-Based Generator for GANs, Karras et al (2018)

Large Scale GAN Training for High Fidelity Natural Image Synthesis, Brock et al (2018)

Generative adversarial networks

Want to learn more?

Generative adversarial networks. Goodfellow, et al. NeurIPS (2014)

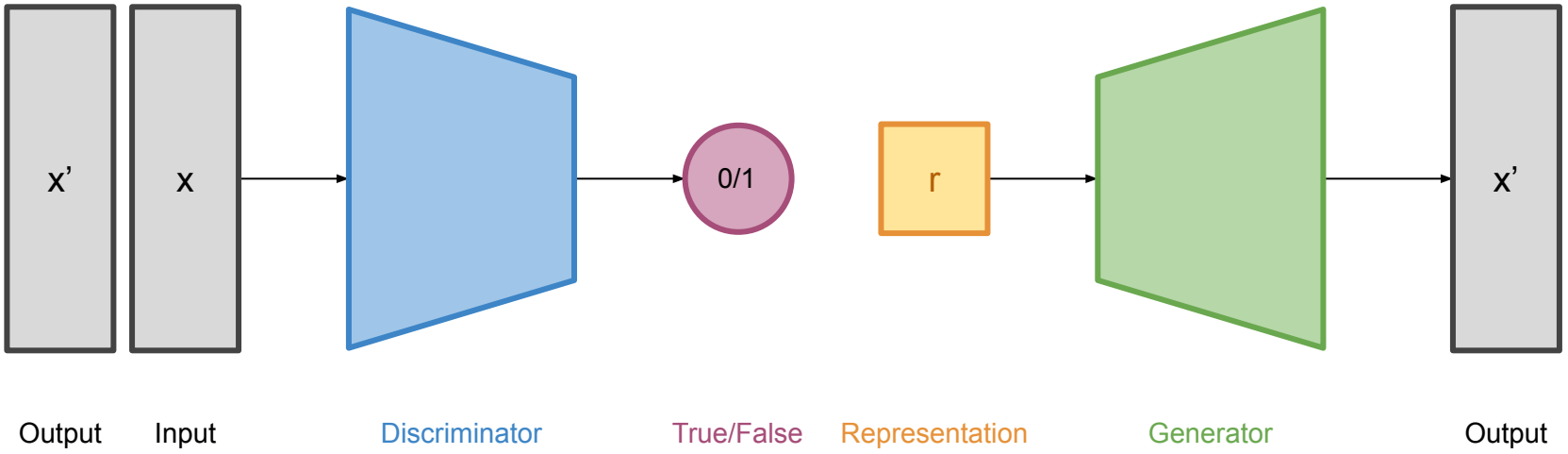
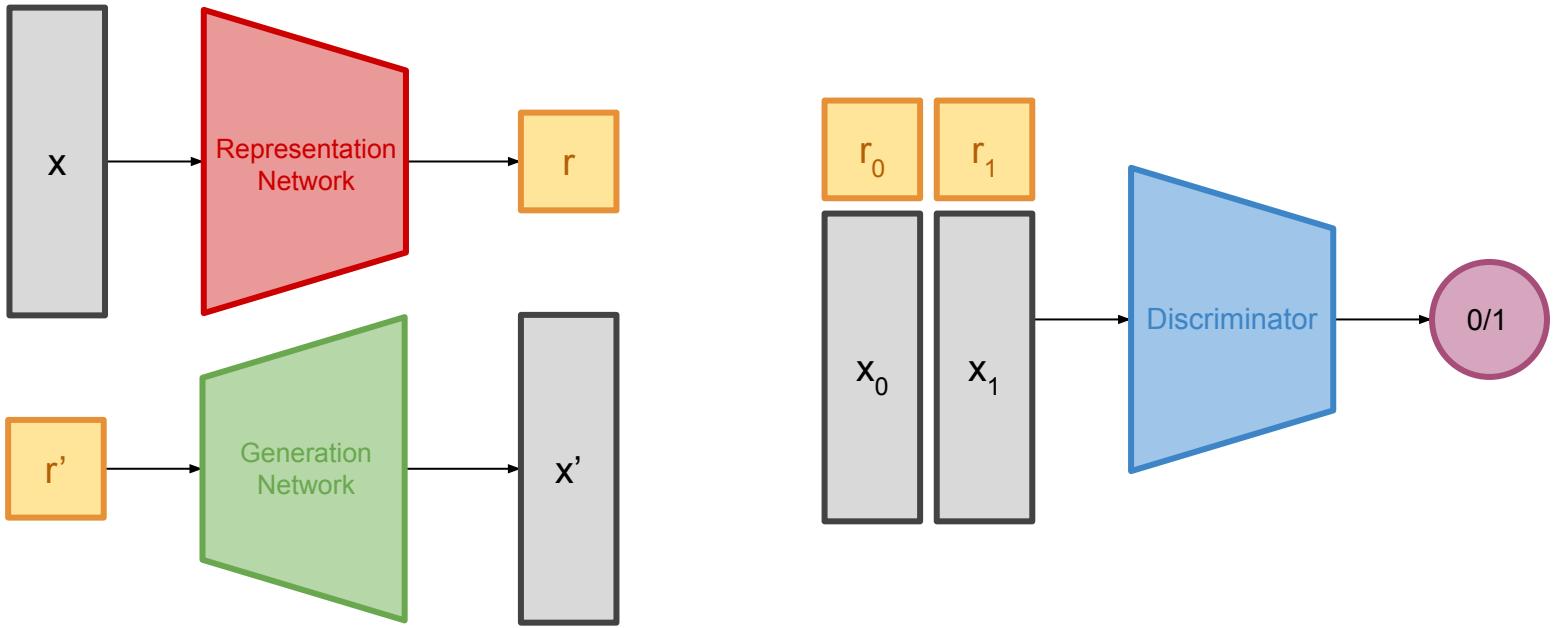


Diagram inspired by Lilian Weng, https://lilianweng.github.io

BiGAN

Want to learn more?

Adversarial Feature Learning, Donahue, et al. ICLR (2017)

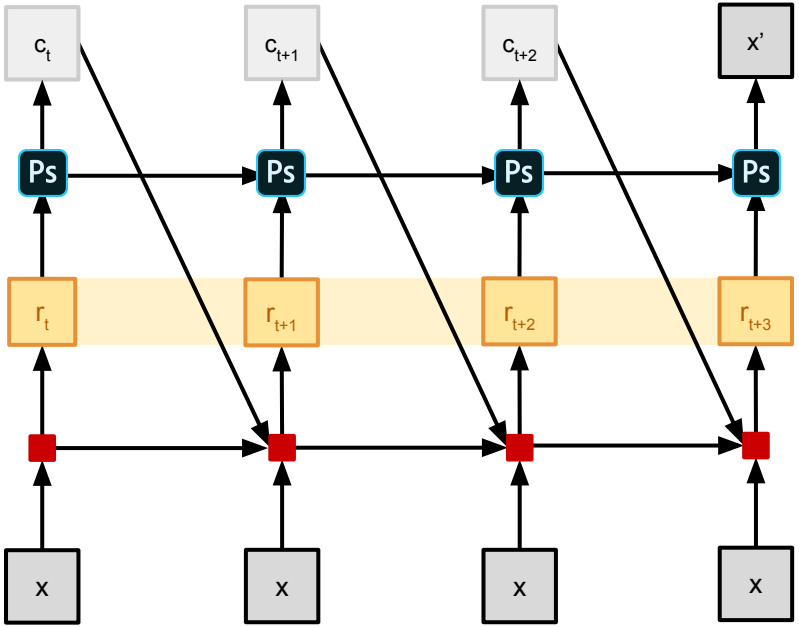


BigBiGAN

Want to learn more?

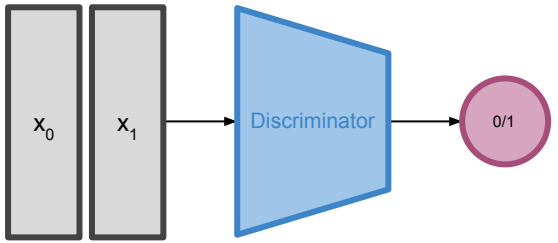
Large Scale Adversarial Representation Learning. Donahue, et al. NeurIPS (2019)

SPIRAL



Paintbrush Simulator

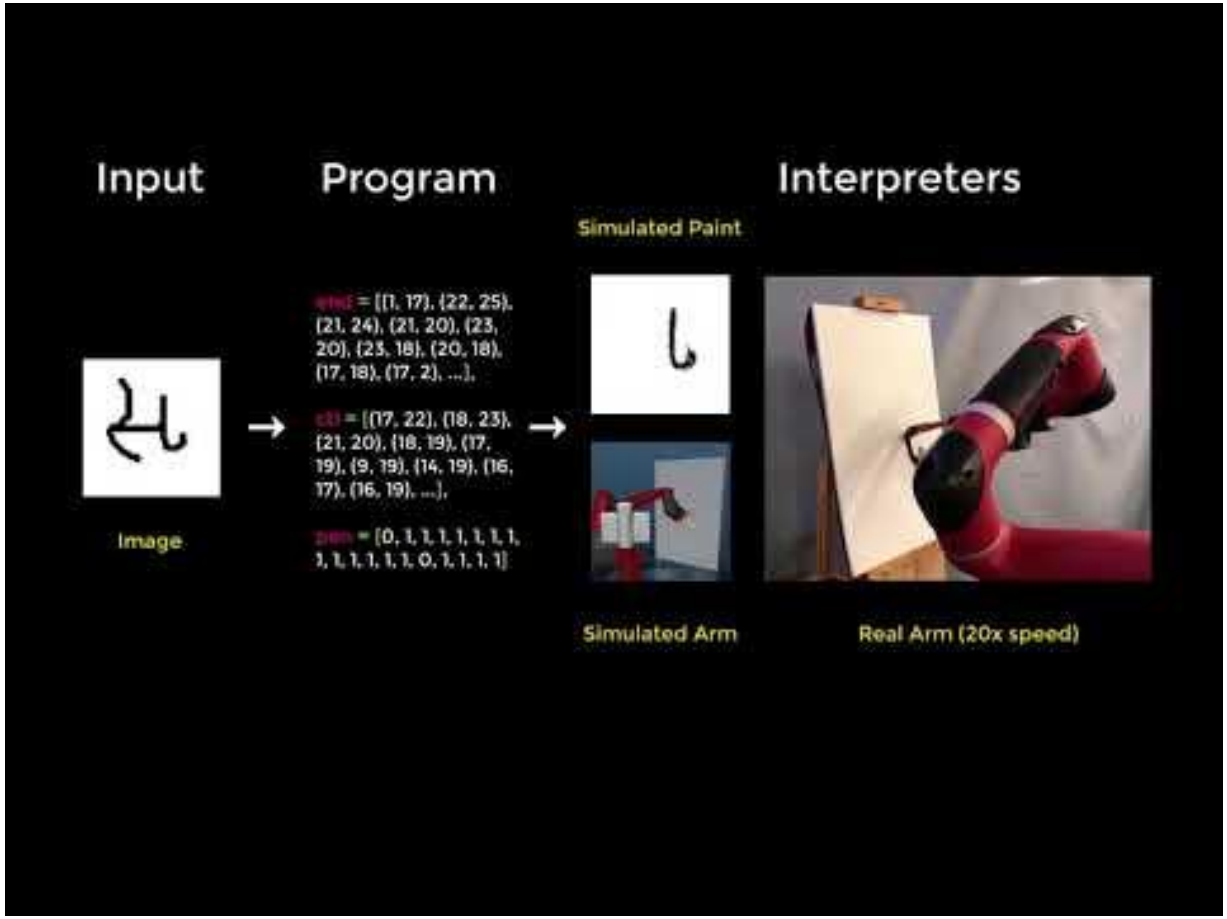
Representation RNN



Want to learn more?

Synthesizing Programs for Images using Reinforced Adversarial Learning, Ganin et al, ICML (2018)

Unsupervised Doodling and Painting with Improved SPIRAL, Mellor et al (2019)



Want to learn more?

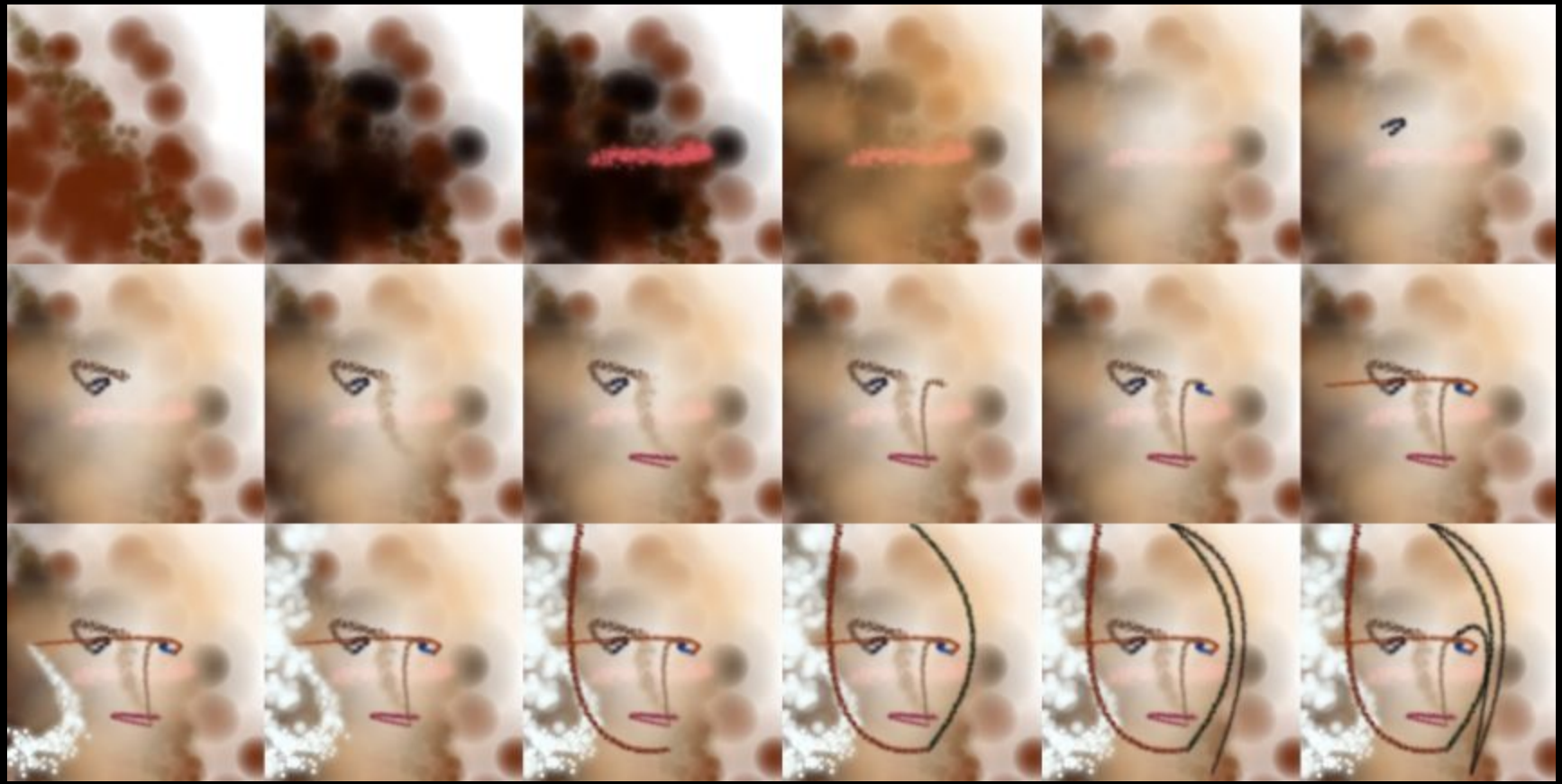
Synthesizing Programs for Images using Reinforced Adversarial Learning, Ganin et al, ICML (2018)

Unsupervised Doodling and Painting with Improved SPIRAL, Mellor et al (2019)

Want to learn more?

Synthesizing Programs for Images using Reinforced Adversarial Learning, Ganin et al, ICML (2018)

Unsupervised Doodling and Painting with Improved SPIRAL, Mellor et al (2019)

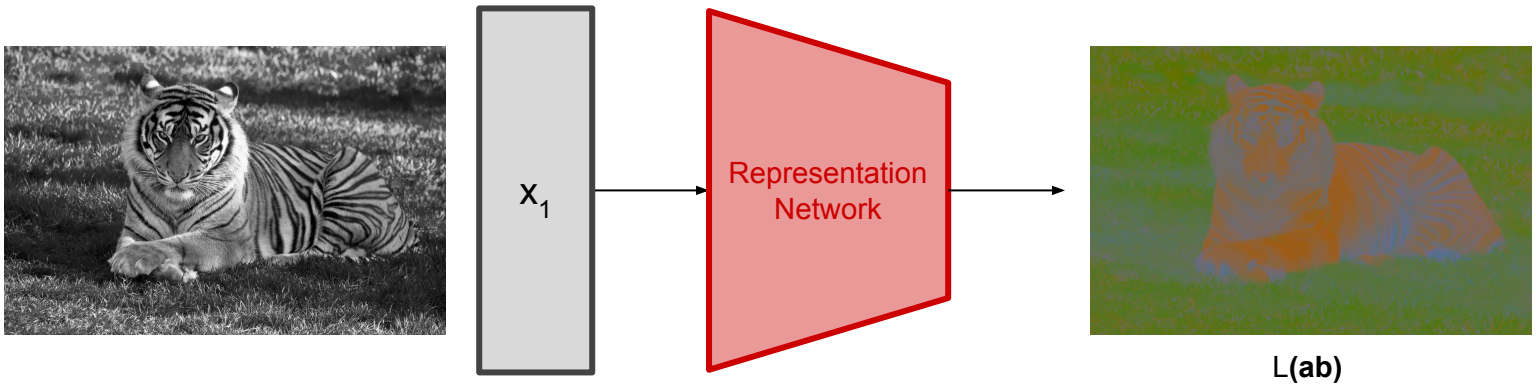


Beyond generative

Colorization

Want to learn more?

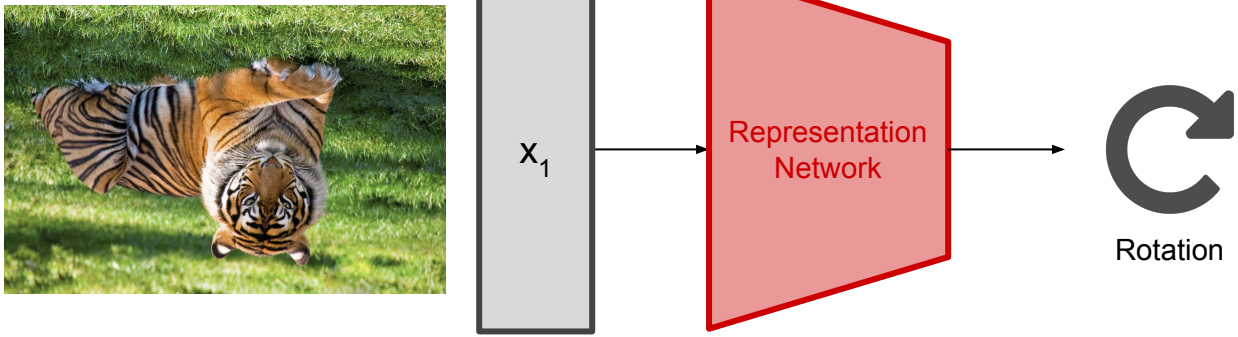
Colorization as a proxy task for visual understanding, Larsson et al, CVPR (2017)



Rotation Prediction

Want to learn more?

Unsupervised Representation Learning by Predicting Image Rotations, Gidaris et al, ICLR (2018)

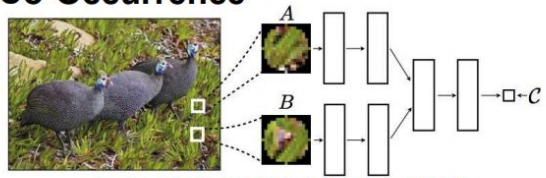


Self-supervised learning

Want to learn more?

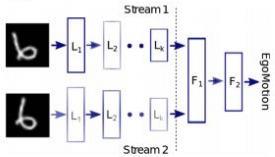
Self-Supervised Learning lecture, Andrew Zisserman, ICML (2018)

Co-Occurrence



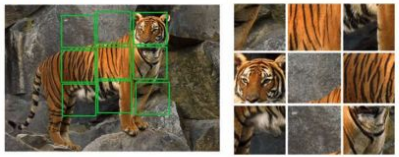
Isola *et al.* ICLR Workshop 2016.

Egomotion



Agrawal *et al.* ICCV 2015 Jayaraman *et al.* ICCV 2015

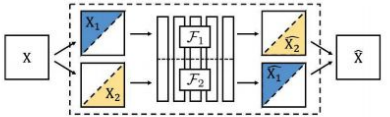
Context



Noroozi et al 2016

Pathak *et al.* CVPR 2016

Split-brain auto-encoders

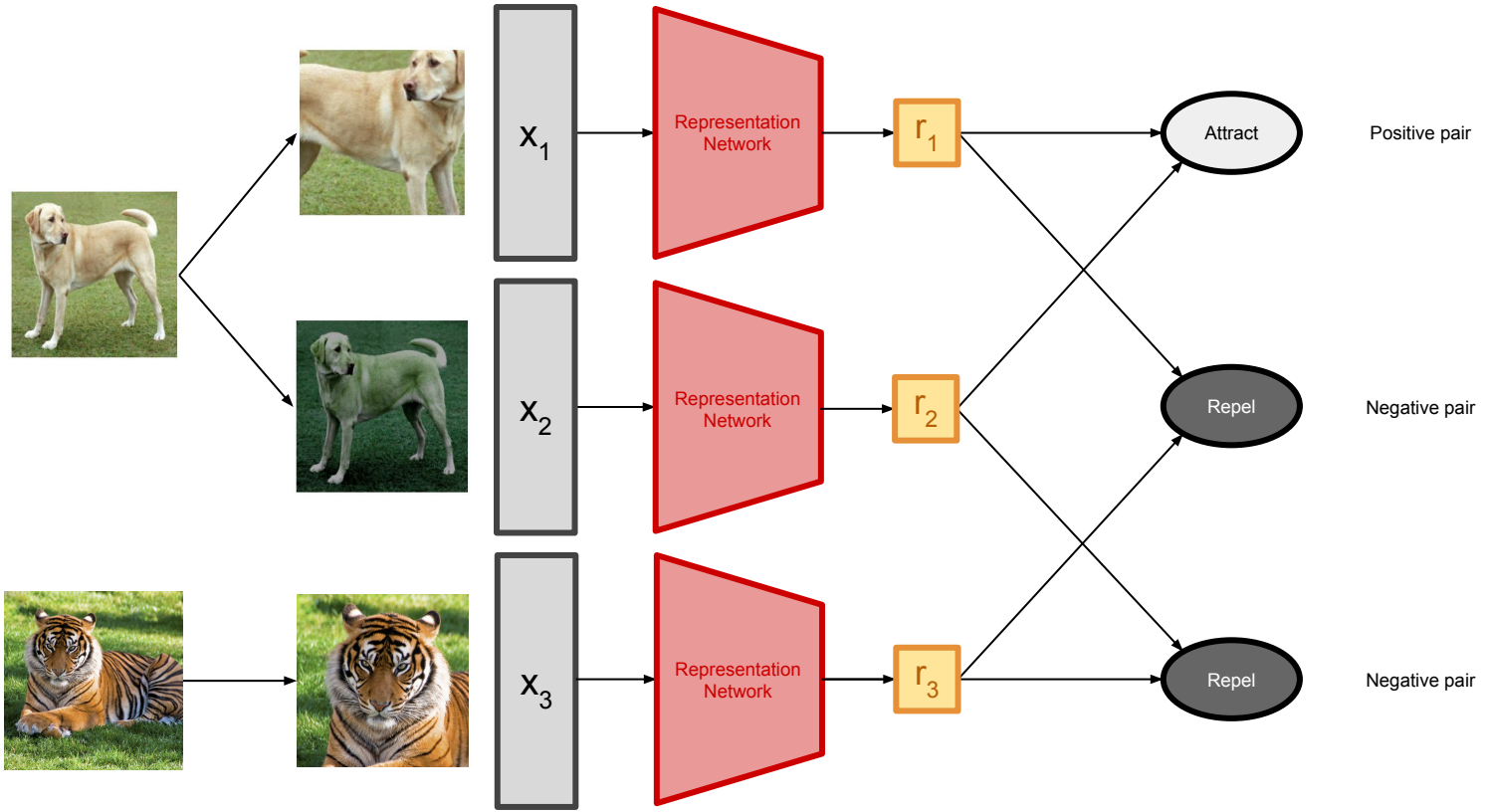


Zhang *et al.* CVPR 2017

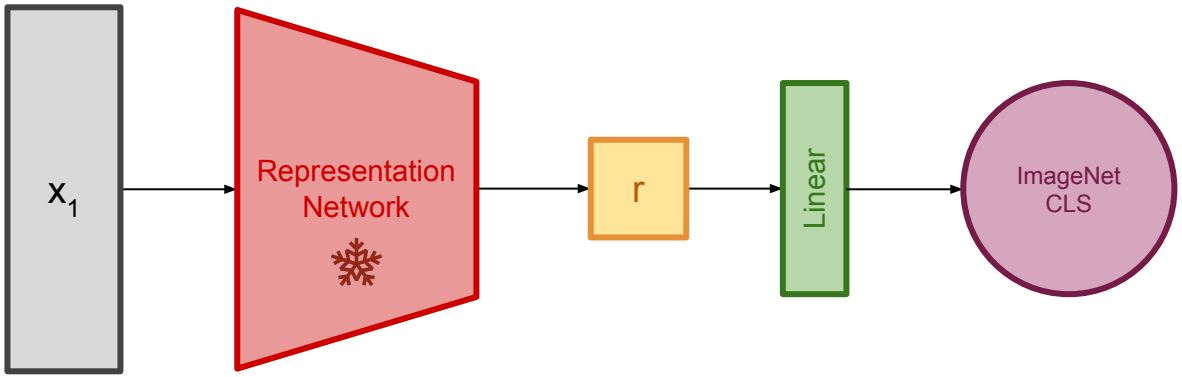
Contrastive learning

Want to learn more?

A Simple Framework for Contrastive Learning of Visual Representations, Chen et al, ICML (2020)



Evaluation: Linear separation



Contrastive learning

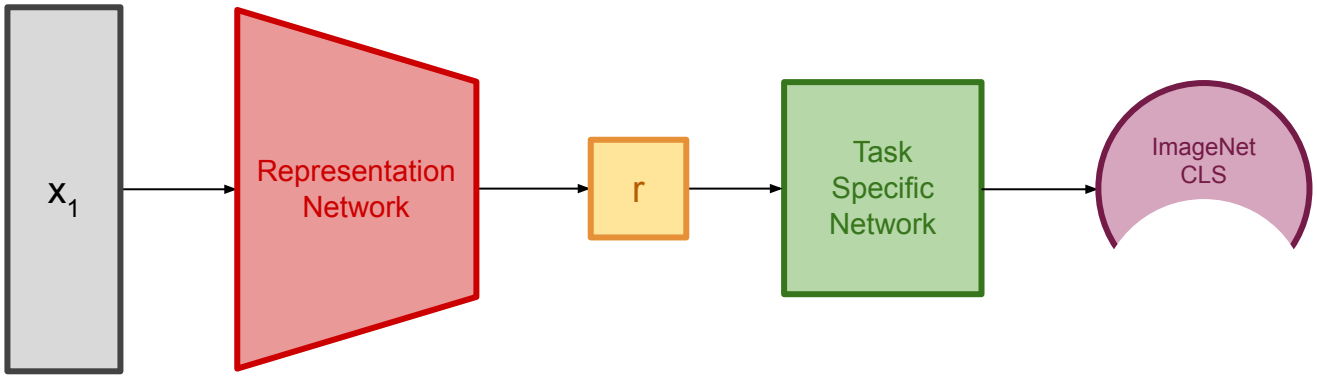
Want to learn more?

A Simple Framework for Contrastive Learning of Visual Representations, Chen et al, ICML (2020)

Method	Architecture	Param.	Top 1	Top 5
Methods using ResNet-50:				
Local Agg.	ResNet-50	24	60.2	-
MoCo	ResNet-50	24	60.6	-
PIRL	ResNet-50	24	63.6	-
CPC v2	ResNet-50	24	63.8	85.3
SimCLR (ours)	ResNet-50	24	**69.3**	**89.0**
Methods using other architectures:				
Rotation	RevNet-50 (4×)	86	55.4	-
BigBiGAN	RevNet-50 (4×)	86	61.3	81.9
AMDIM	Custom-ResNet	626	68.1	-
CMC	ResNet-50 (2×)	188	68.4	88.2
MoCo	ResNet-50 (4×)	375	68.6	-
CPC v2	ResNet-161 (∗)	305	71.5	90.1
SimCLR (ours)	ResNet-50 (2×)	94	74.2	92.0
SimCLR (ours)	ResNet-50 (4×)	375	**76.5**	**93.2**

Table 6. ImageNet accuracies of linear classifiers trained on representations learned with different self-supervised methods.

Evaluation: Data efficiency



Data efficient representation learning

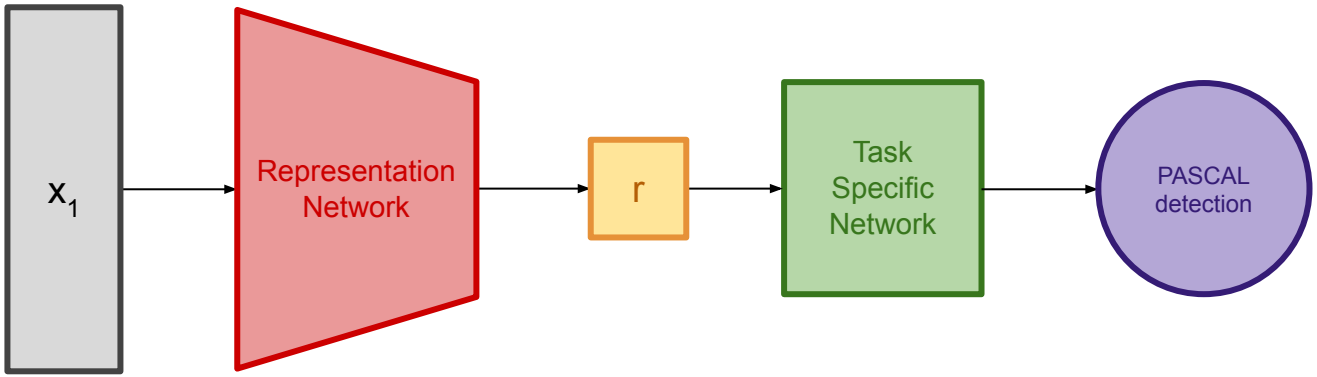
Want to learn more?

A Simple Framework for Contrastive Learning of Visual Representations, Chen et al, ICML (2020)

Method	Architecture	Label fraction 1%	10%
		Top 5	
Supervised baseline	ResNet-50	48.4	80.4
Methods using other label-propagation:			
Pseudo-label	ResNet-50	51.6	82.4
VAT+Entropy Min.	ResNet-50	47.0	83.4
UDA (w. RandAug)	ResNet-50	-	88.5
FixMatch (w. RandAug)	ResNet-50	-	89.1
S4L (Rot+VAT+En. M.)	ResNet-50 (4×)	-	91.2
Methods using representation learning only:			
InstDisc	ResNet-50	39.2	77.4
BigBiGAN	RevNet-50 (4×)	55.2	78.8
PIRL	ResNet-50	57.2	83.8
CPC v2	ResNet-161(*)	77.9	91.2
SimCLR (ours)	ResNet-50	75.5	87.8
SimCLR (ours)	ResNet-50 (2×)	83.0	91.2
SimCLR (ours)	ResNet-50 (4×)	**85.8**	**92.6**

Table 7. ImageNet accuracy of models trained with few labels.

Evaluation: Transfer learning



Transfer learning

Want to learn more?

A Simple Framework for Contrastive Learning of Visual Representations, Chen et al, ICML (2020)

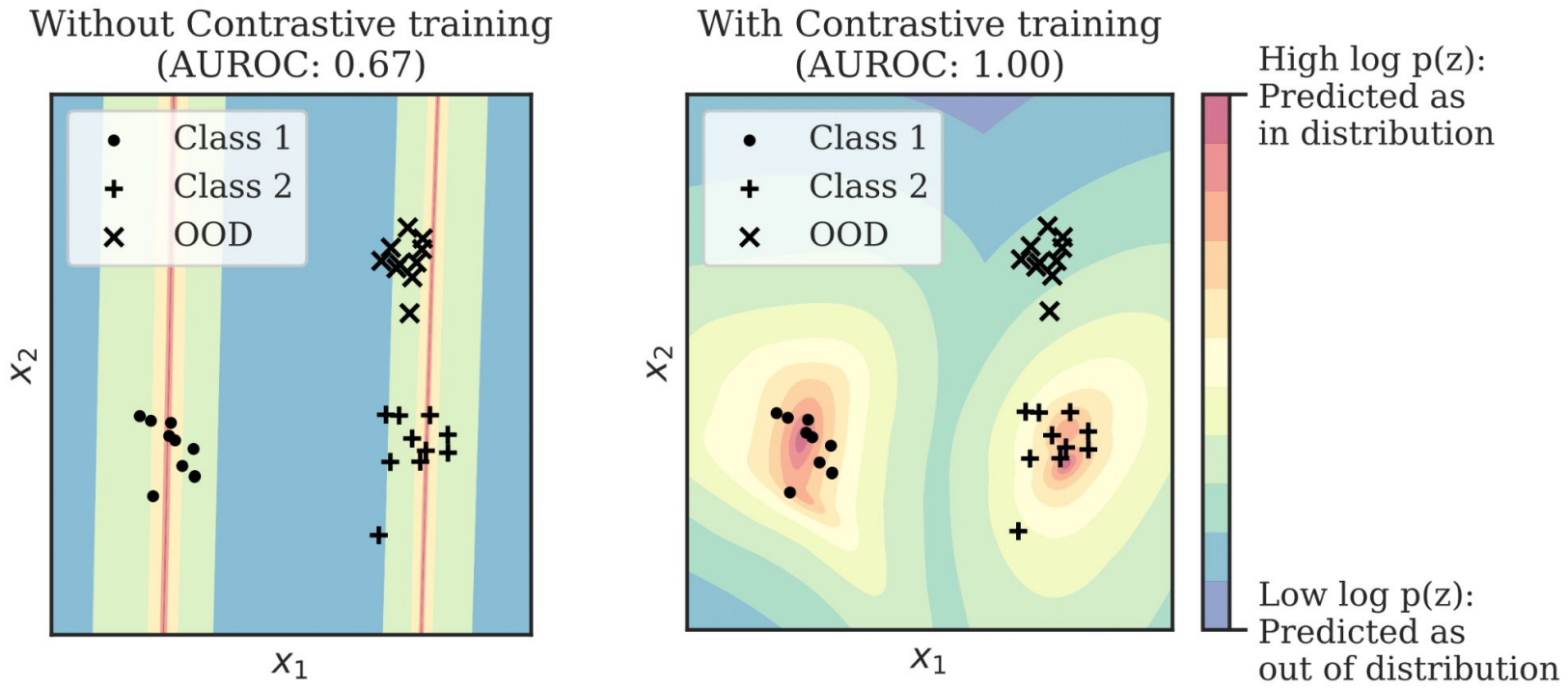
	Food	CIFAR10	CIFAR100	Birdsnap	SUN397	Cars	Aircraft	VOC2007	DTD	Pets	Caltech-101	Flowers
Linear evaluation:												
SimCLR (ours)	**76.9**	**95.3**	80.2	48.4	**65.9**	60.0	61.2	**84.2**	**78.9**	89.2	**93.9**	**95.0**
Supervised	75.2	**95.7**	**81.2**	**56.4**	64.9	**68.8**	**63.8**	83.8	**78.7**	**92.3**	**94.1**	94.2
Fine-tuned:												
SimCLR (ours)	**89.4**	**98.6**	**89.0**	**78.2**	**68.1**	**92.1**	**87.0**	**86.6**	**77.8**	92.1	**94.1**	97.6
Supervised	88.7	98.3	**88.7**	**77.8**	67.0	91.4	**88.0**	86.5	**78.8**	**93.2**	**94.2**	**98.0**
Random init	88.3	96.0	81.9	**77.0**	53.7	91.3	84.8	69.4	64.1	82.7	72.5	92.5

Table 8. Comparison of transfer learning performance of our self-supervised approach with supervised baselines across 12 natural image classification datasets, for ResNet-50 $(4\times)$ models pretrained on ImageNet. Results not significantly worse than the best ($p > 0.05$, permutation test) are shown in bold. See Appendix B.8 for experimental details and results with standard ResNet-50.

Out of distribution detection

Want to learn more?

Contrastive Training for Improved Out-of-Distribution Detection, Winkens et al (2020)

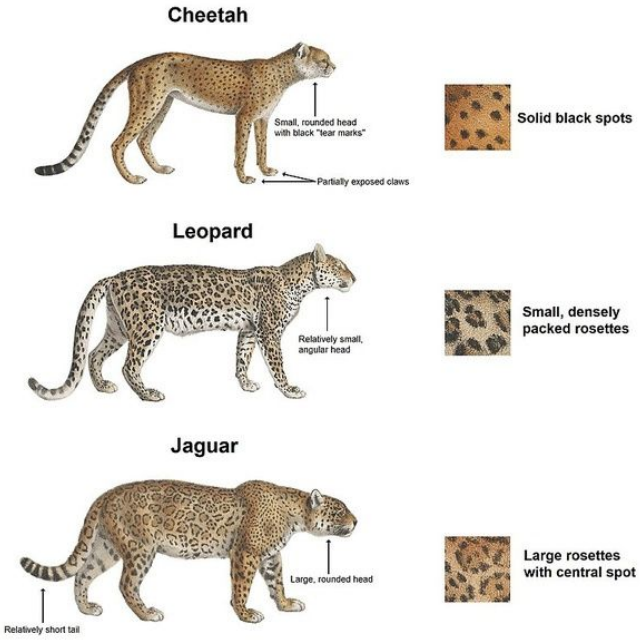


DeepMind

5 Conclusions

Only two problems that make intelligence hard

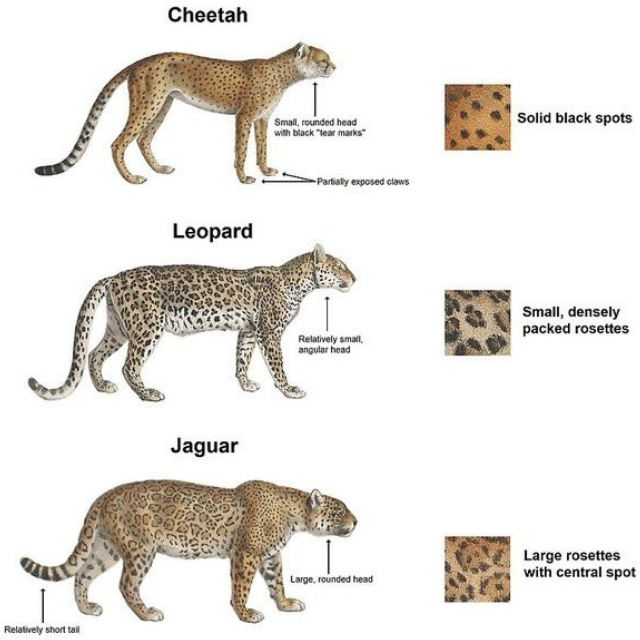
Hot take!

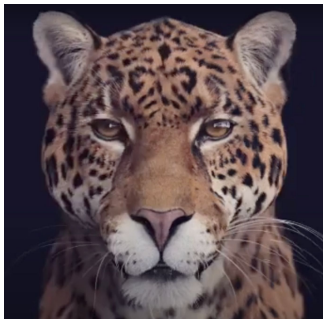


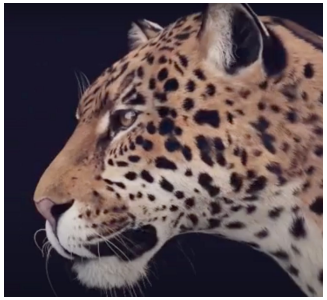
Categorisation

Only two problems that make intelligence hard

Hot take!





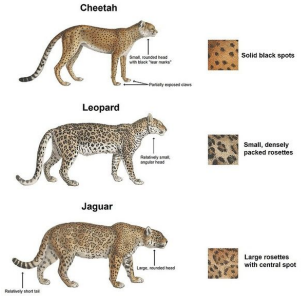


Categorisation

Partial observability

Only two problems that make intelligence hard

Hot take!



Categorisation

‘Truth’ exists only in human minds

Subjective

Prediction uncertainty

Representational form is relatively clear

Partial observability

Truth exists in reality and can be sensed (in theory)

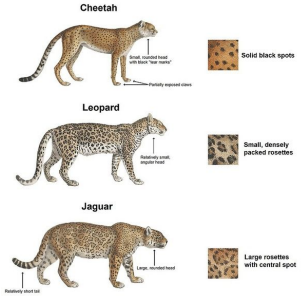
Objective

Prediction uncertainty

Representational form is mostly unclear

Only two problems that make intelligence hard

Hot take!



Pattern Recognition

Categorisation

‘Truth’ exists only in human minds

Subjective

Prediction uncertainty

Representational form is relatively clear

Information Integration

Partial observability

Truth exists in reality and can be sensed (in theory)

Objective

Prediction uncertainty

Representational form is mostly unclear

Summary

The representation learning problem is **under-specified**.

Three broad categories of approaches:

1. **Building in structure** or **inductive bias** to obtain the 'right' representations, e.g. structured autoencoders
2. **Training for proxy tasks** that can only be solved with the 'right' representations, e.g. contrastive learning
3. **Internet-scale** datasets with rich labels

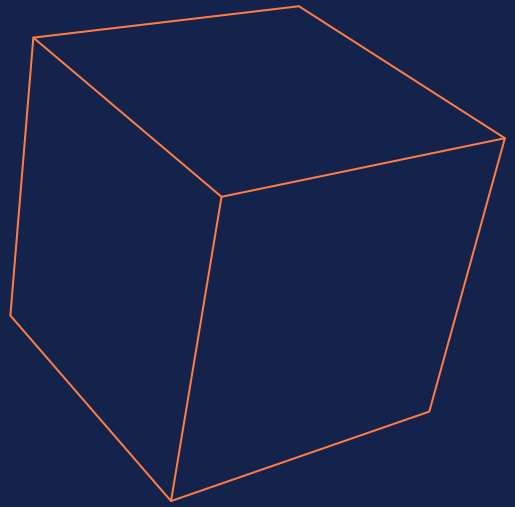
Current approaches seem to involve a trade-off between:

1. **Generality of the representation**, i.e. what range of downstream tasks the representation is good for
2. **Interpretability**, i.e. how much control we have on the representational space

General representation learning without labels is **still largely unsolved**?

But in the era of internet-scale data, is it a question still worth solving?

Stay in touch!



Twitter:
@arkitus