ADVANCES IN
COMPUTER
VISION

Module 2: Unsupervised Representation Learning & Generative Modeling

Prof. Vincent Sitzmann

Problem Sets & Final Project Proposal

- Problem Set 5 out today! On Representation Learning & Diffusion
  - You can start today, part of it is not dependent on lectures

- Final Project Proposal due on April 15th (2 weeks from now)
  - Refer to Final Project guidelines on website

What is generative modeling?

> *Creating noise from data is easy;*
> *creating data from noise is generative modeling*
>
> \- Song et al. 2020

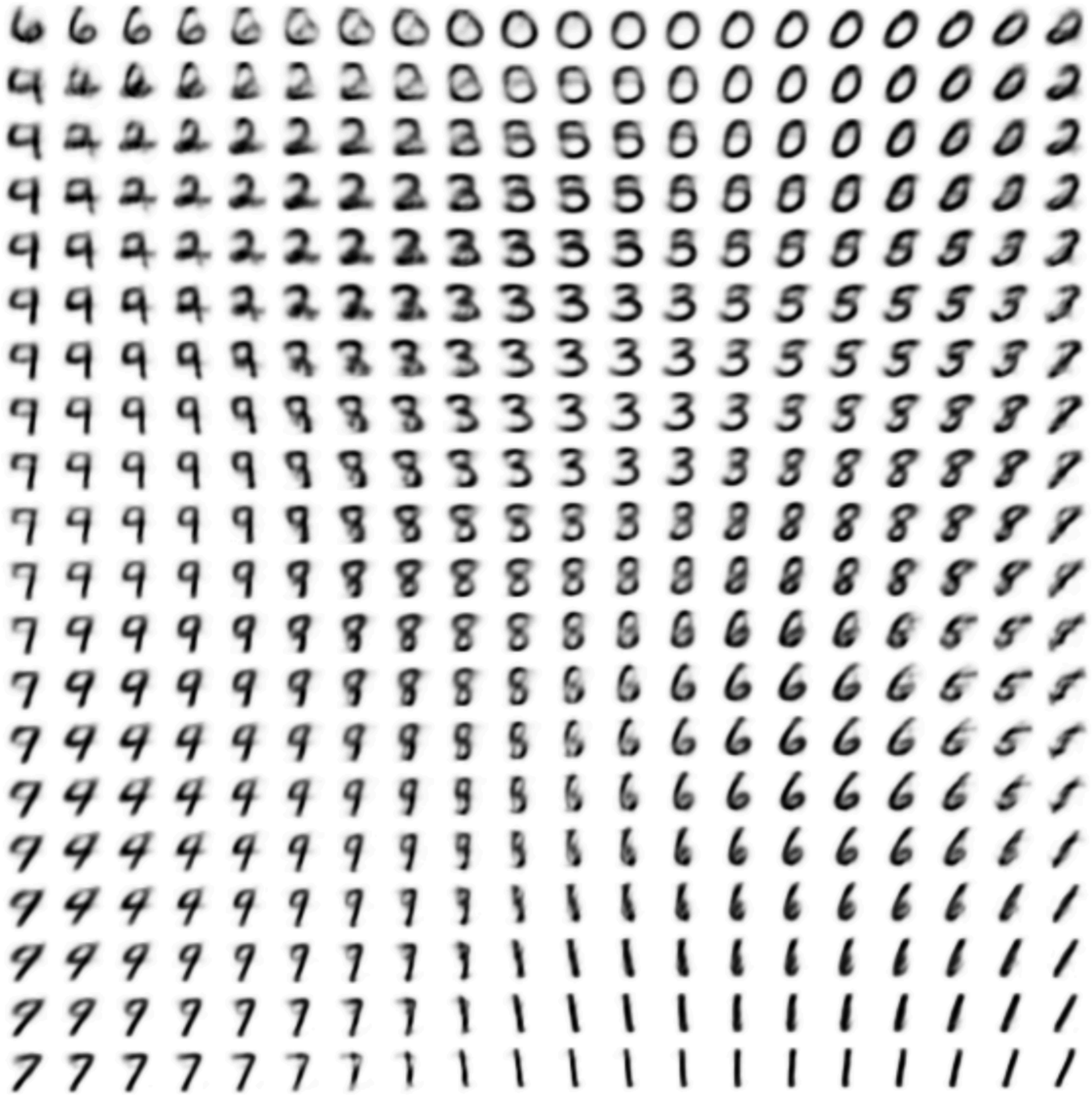
What is generative modeling?

> *Creating noise from data is easy;*
> *creating data from noise is generative modeling*
>
> \- Song et al. 2020

- Given samples from some distribution,
  learn to generate new samples
- In practice: learn to map noise to (images, audio, …)

2013: VAEs

(a) Learned Frey Face manifold

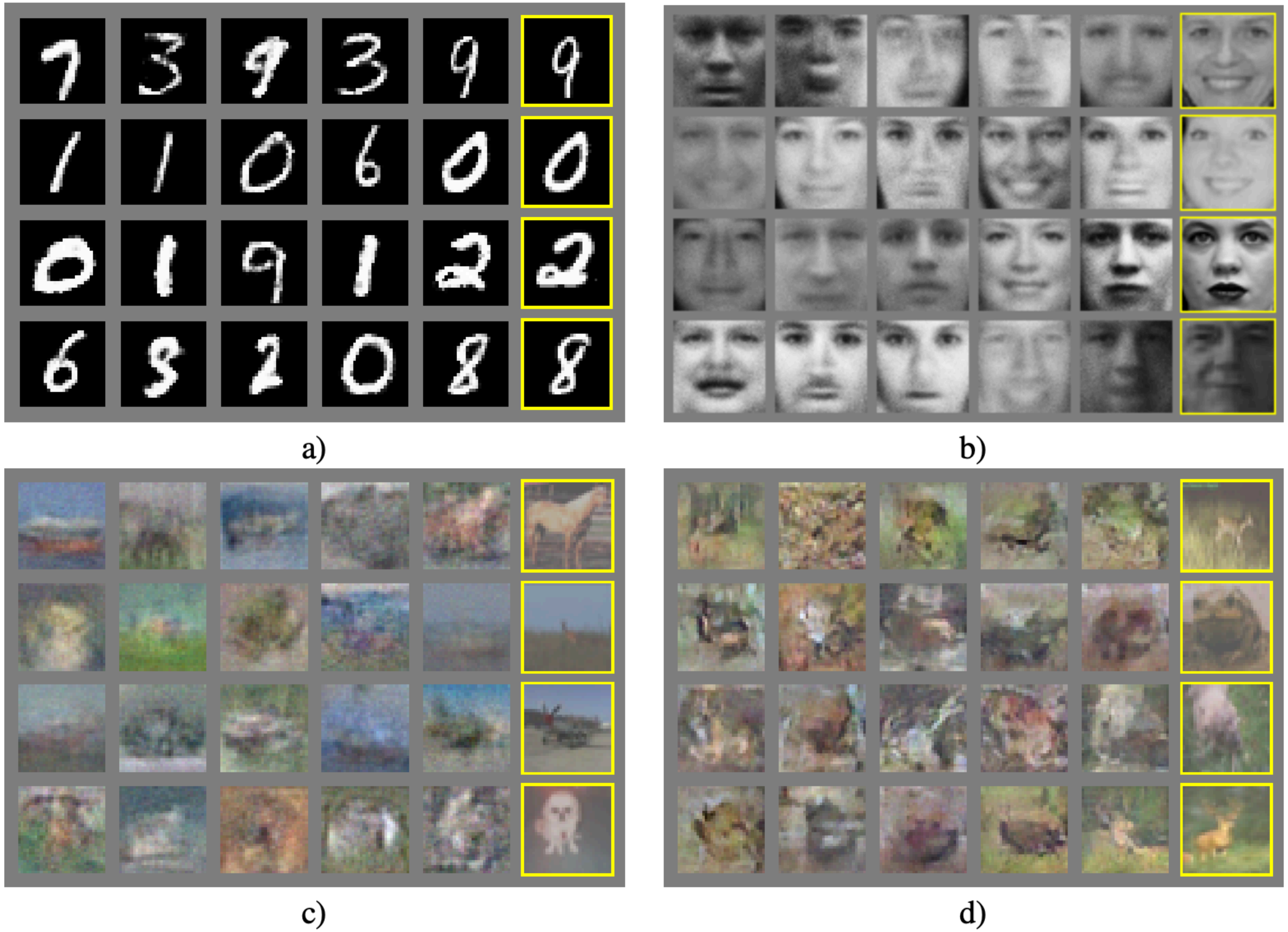


(b) Learned MNIST manifold

Video source: https://x.com/runwayml/status/1807822396415467686

Auto-encoding Variational Bayes, Kingma et al. 2013

2014: GANs



Generative Adversarial Networks, Goodfellow et al. 2014

GANs by 2019

Figure 1: Class-conditional samples generated by our model.

LARGE SCALE GAN TRAINING FOR HIGH FIDELITY NATURAL IMAGE SYNTHESIS, Brock et al. 2019

VAEs by 2020

Figure 1: 256×256-pixel samples generated by NVAE, trained on CelebA HQ [28].

NVAE: A Deep Hierarchical Variational Autoencoder, Vahdat et al. 2020

2020: Diffusion

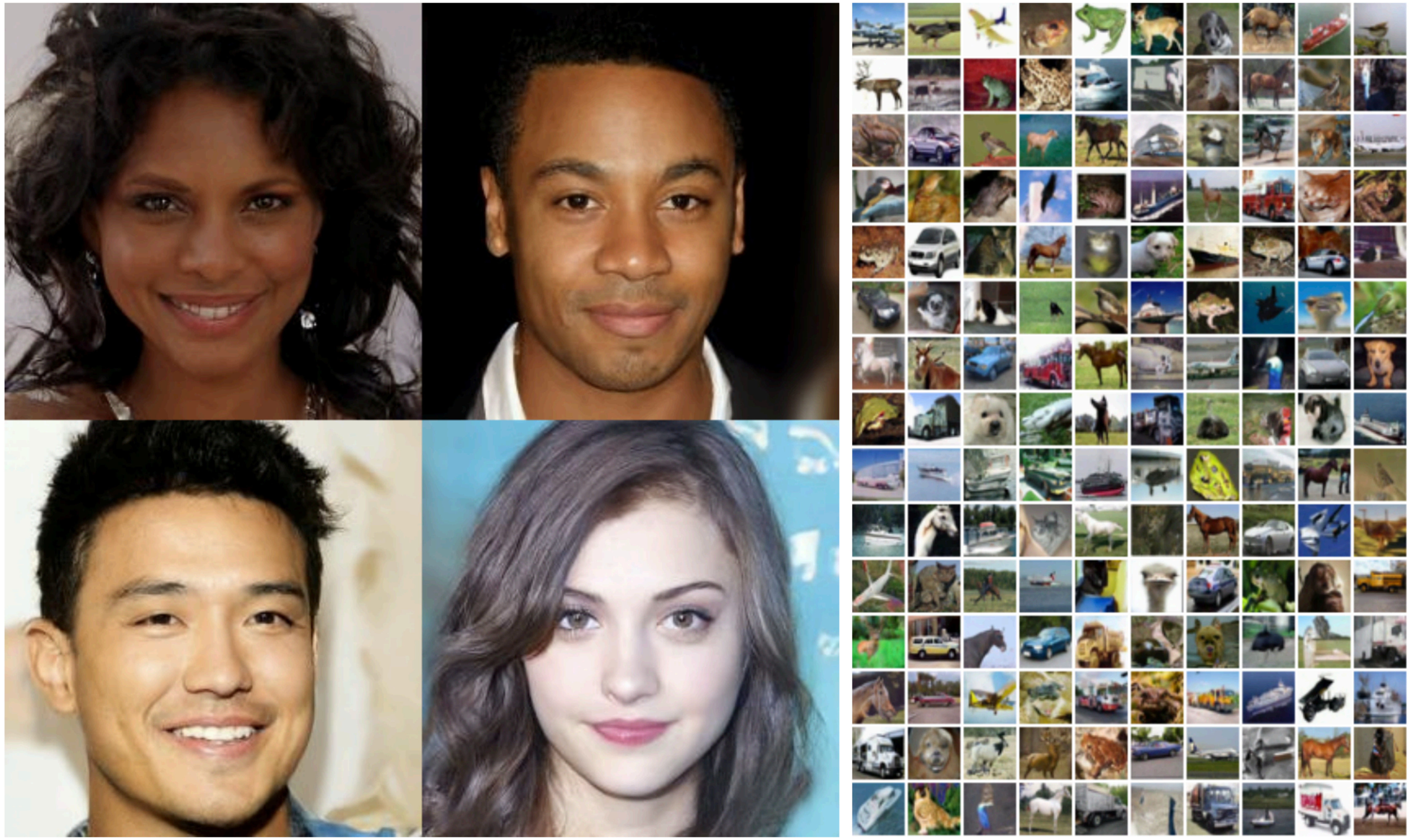


Figure 1: Generated samples on CelebA-HQ 256 × 256 (left) and unconditional CIFAR10 (right)

Deep Unsupervised Learning using Nonequilibrium Thermodynamics, Sohl-Dickstein et al. 2015
Denoising Diffusion Probabilistic Models, Ho et al. 2006

Since 2022: Video Generative Models

Video source: https://x.com/runwayml/status/1807822396415467686

Diffusion Probabilistic Modeling for Video Generation, Yang et al. 2022
Video Diffusion Models, Ho et al. 2022

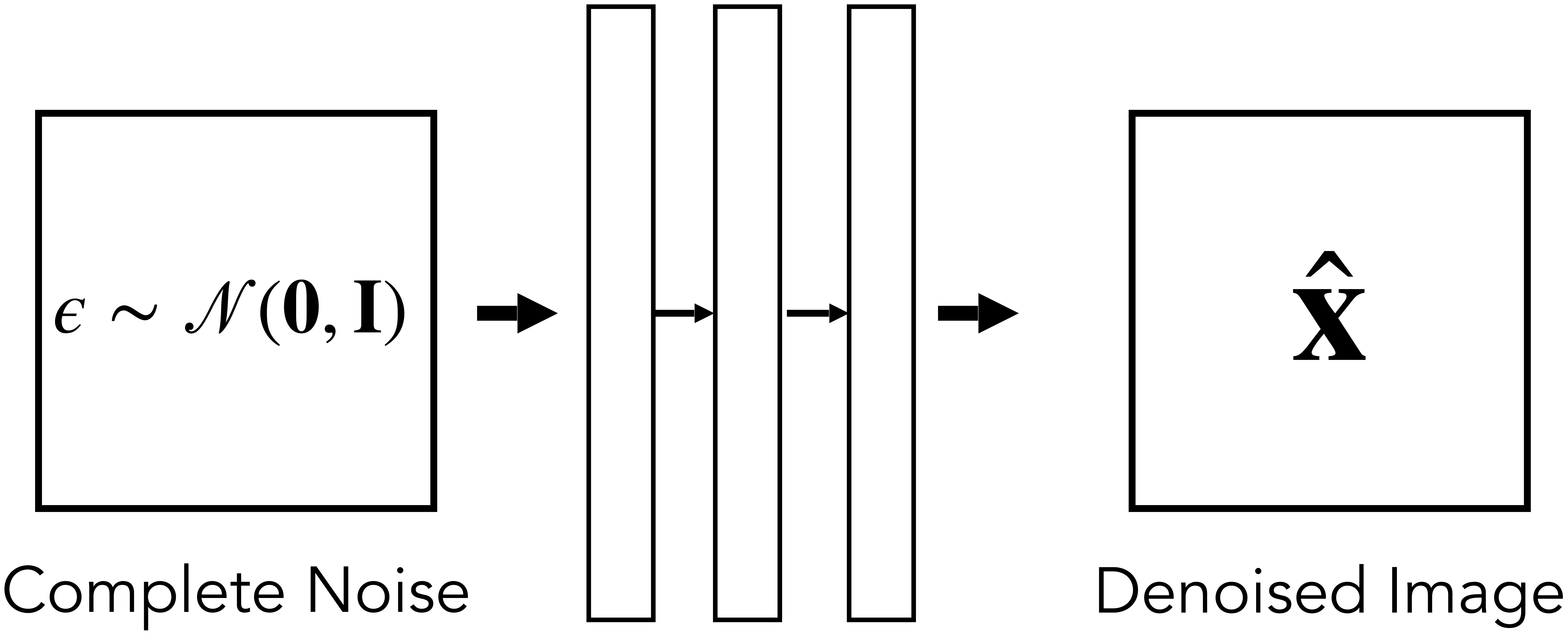
Since 2022: Video Generative Models

Video source: https://x.com/runwayml/status/1807822396415467686

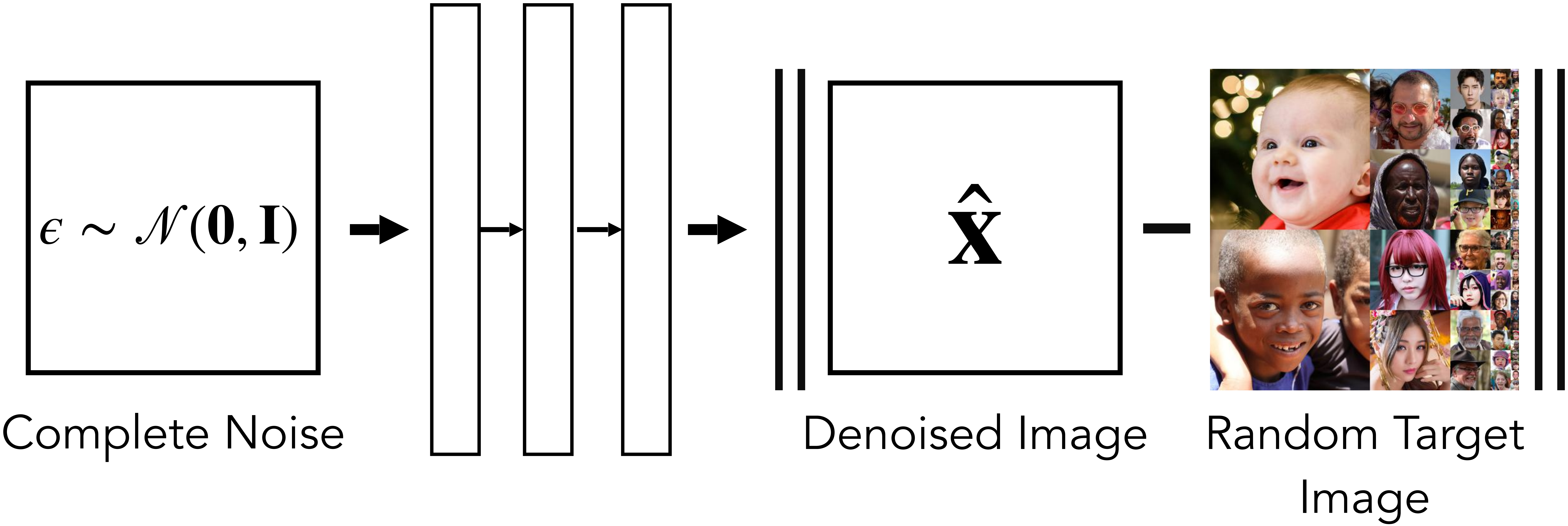
Diffusion Probabilistic Modeling for Video Generation, Yang et al. 2022

Video Diffusion Models, Ho et al. 2022

Let's be naive about it…



Let’s be naive about it…



Let's be naive about it...

$\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ → → $\left\| \hat{\mathbf{x}} - \right.$  $\left. \right\|$

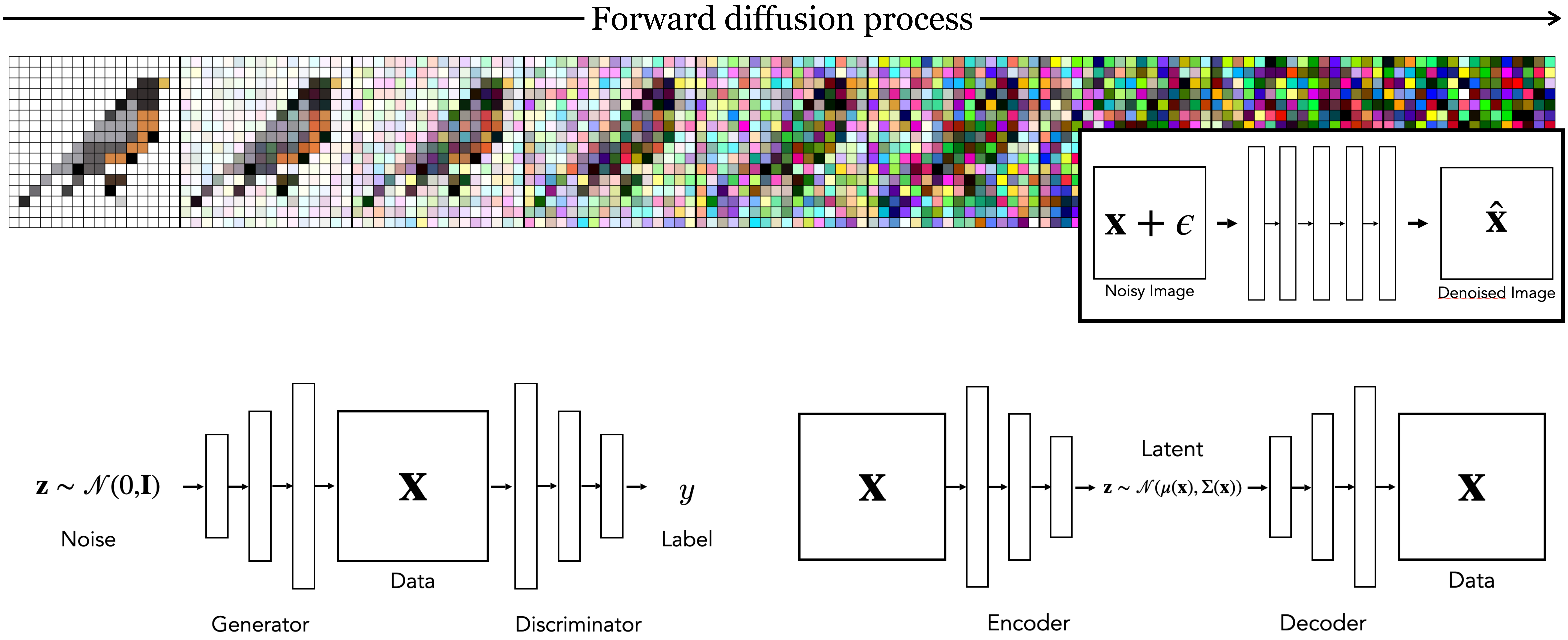
Complete Noise

Denoised Image

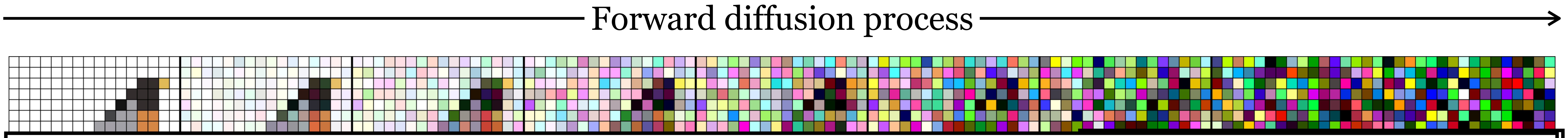
Random Target Image

What will happen?

Zeroth order: Generative Modeling is about "breaking the symmetry" of which noise to map to which image

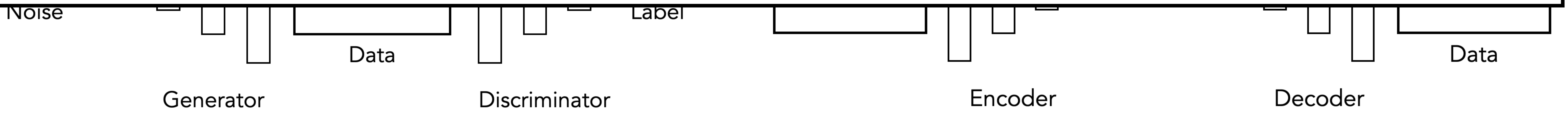


Zeroth order: Generative Modeling is about "breaking the symmetry" of which noise to map to which image



We will use this course for a deep-dive into diffusion models.

Lots of the learnings will transfer to other kinds of generative models as well!



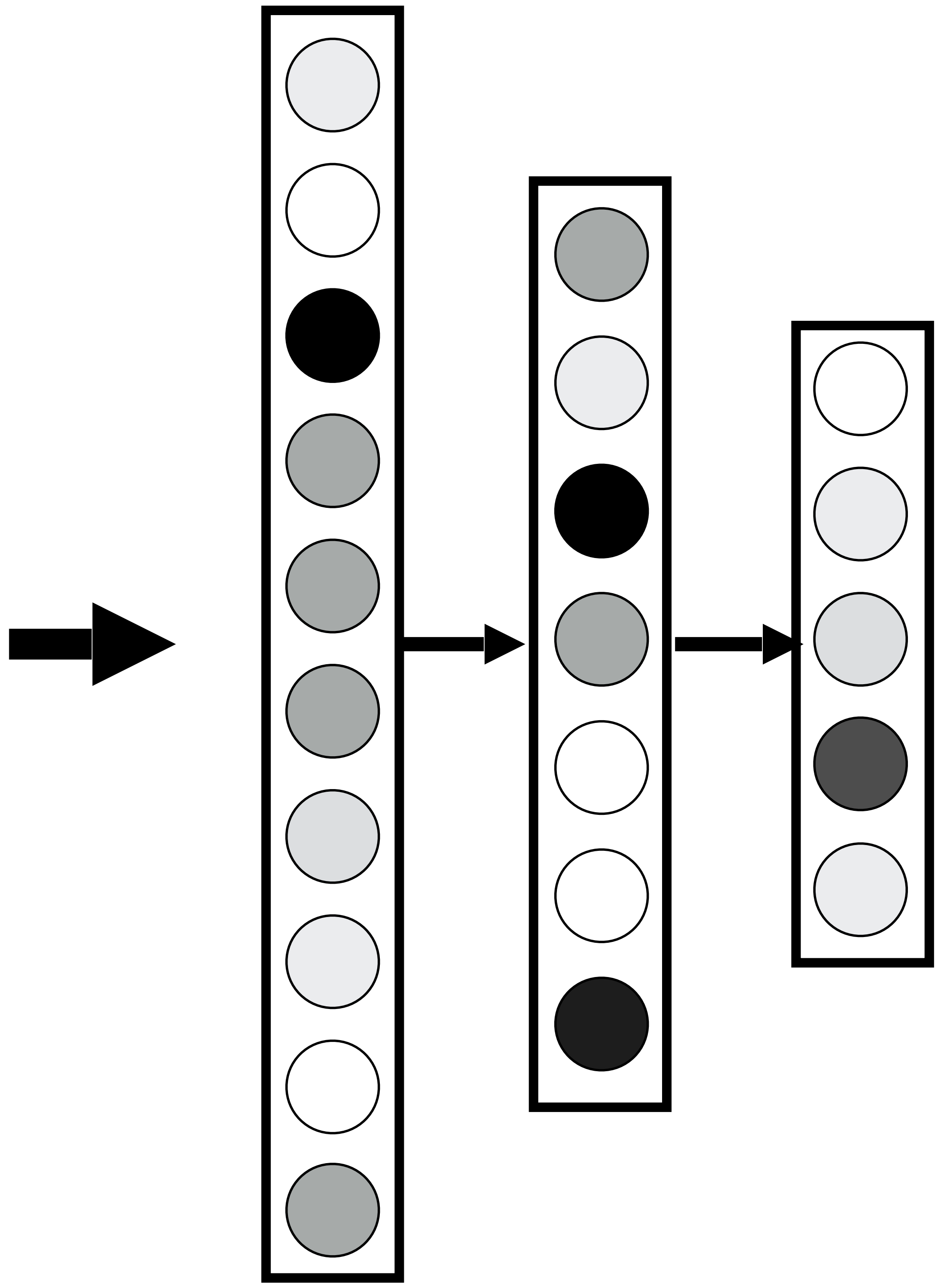
What is a *good* generative model?

- Needs to generate samples that are…
  …not close to any in the training set
  …still “in-distribution” in all important ways
- What kind of “new” images are generated by different generative models is non-obvious
- We will talk about that a bit in this module!

What is Representation Learning?

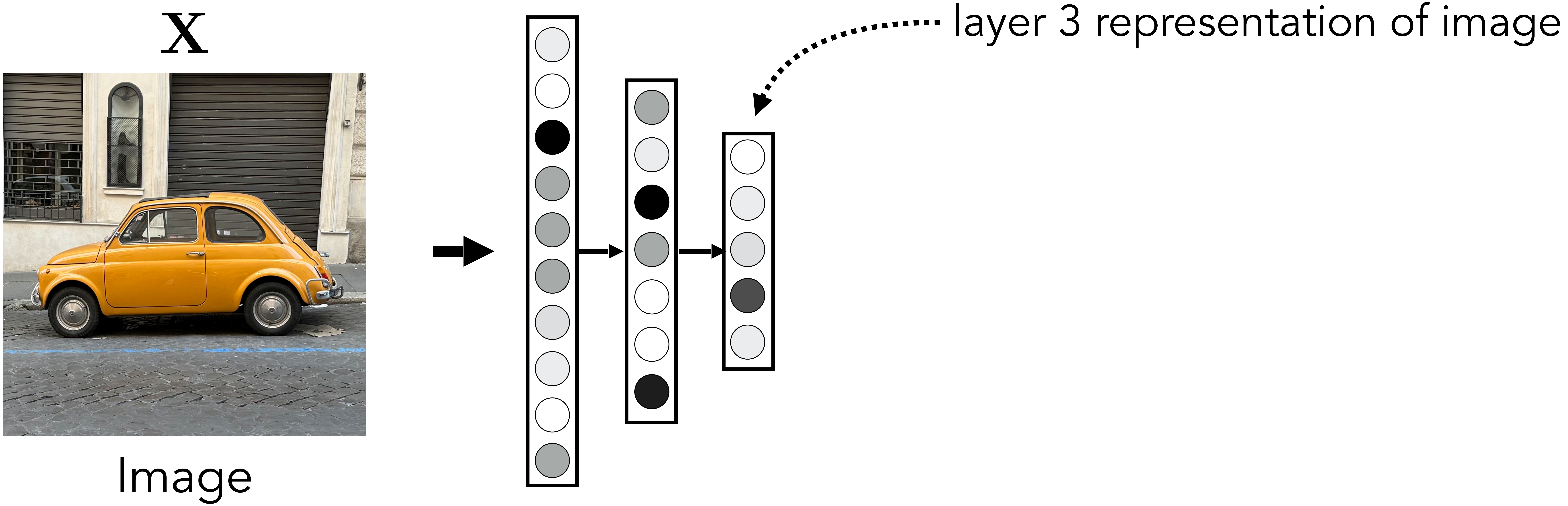
$\mathbf{x}$

Image



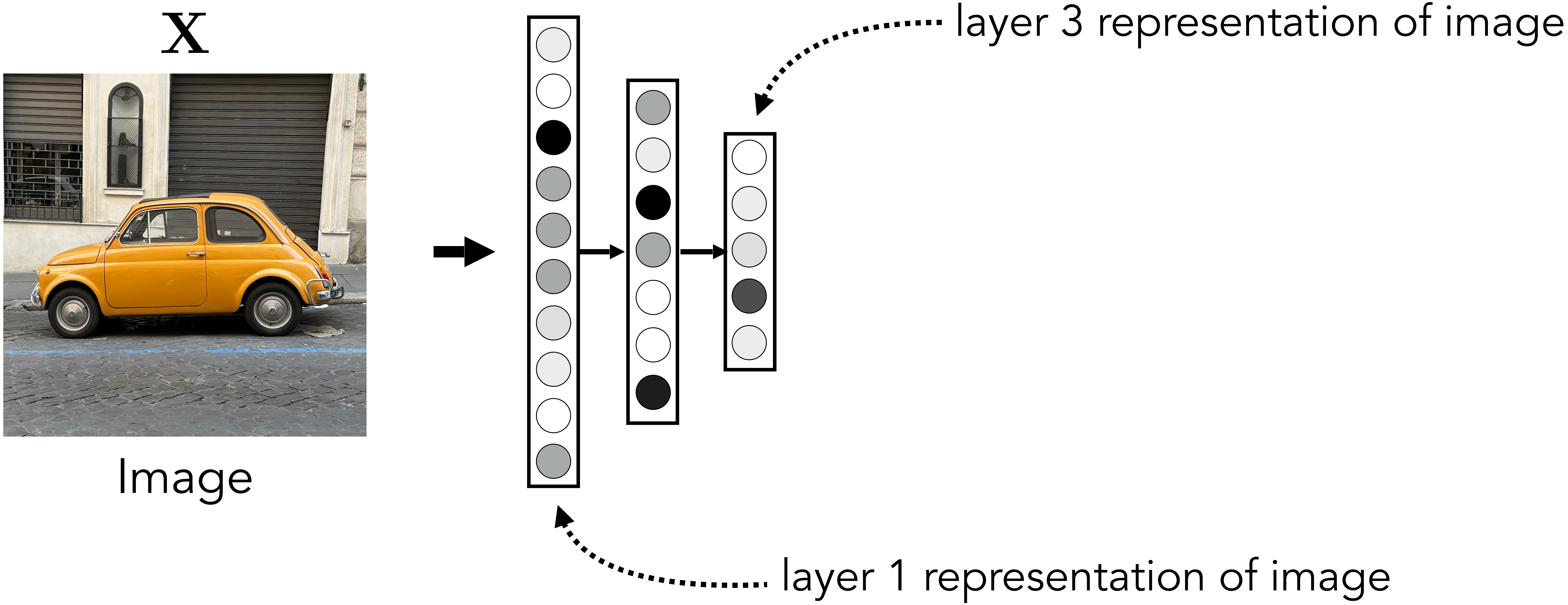
Given an (image, audio signal, …), extract a digital representation of it

What is Representation Learning?



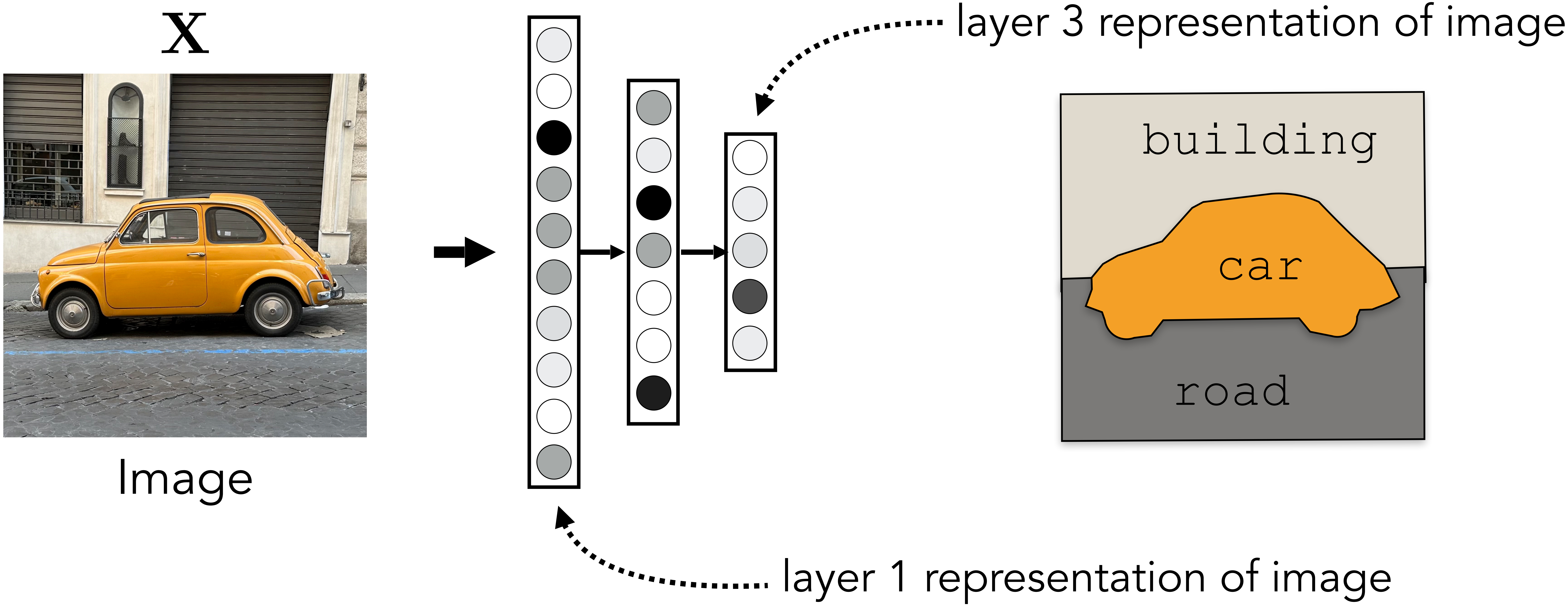
Given an (image, audio signal, …), extract a digital representation of it

What is Representation Learning?



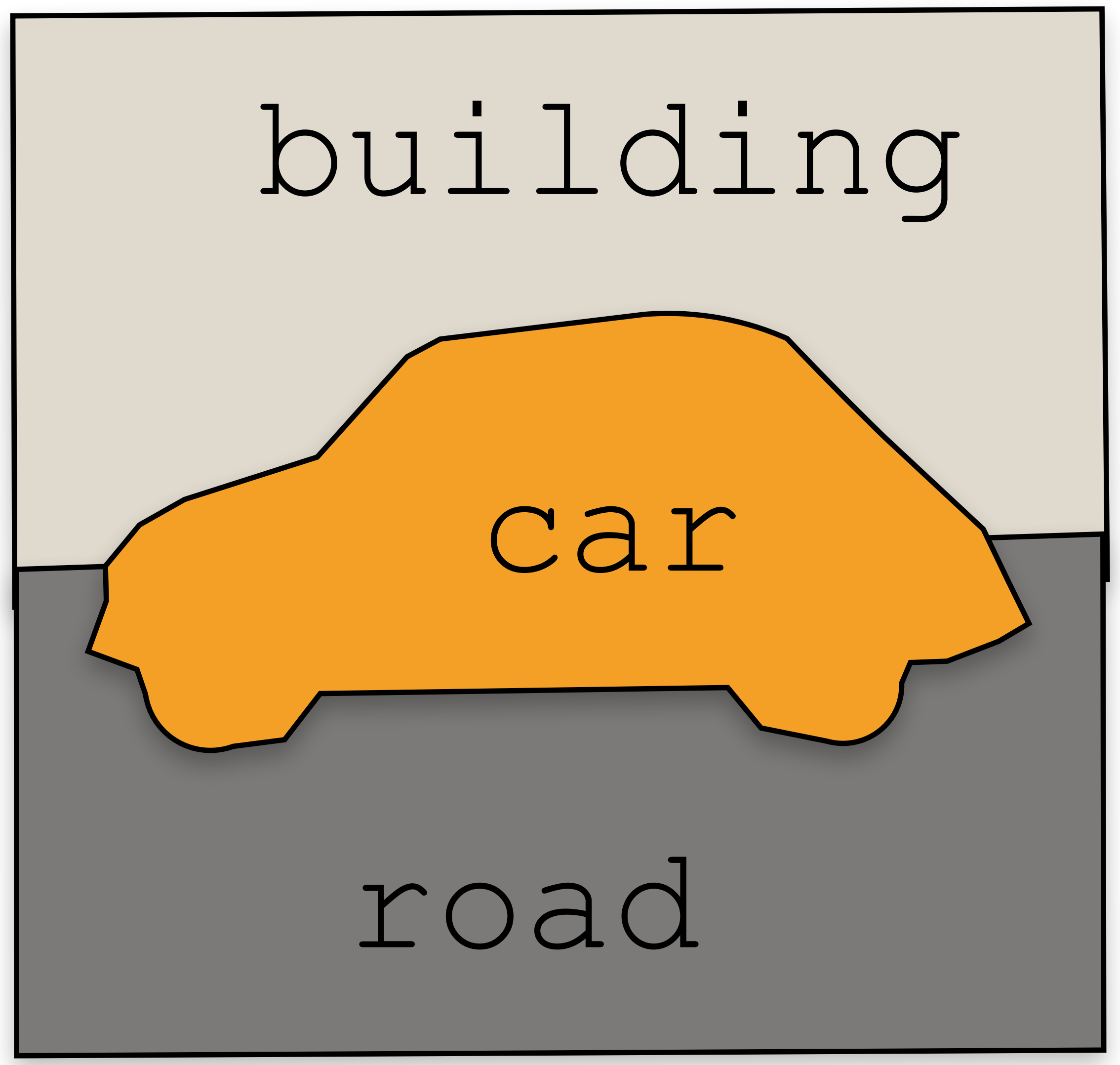
Given an (image, audio signal, …), extract a digital representation of it

What is Representation Learning?



Given an (image, audio signal, …), extract a digital representation of it

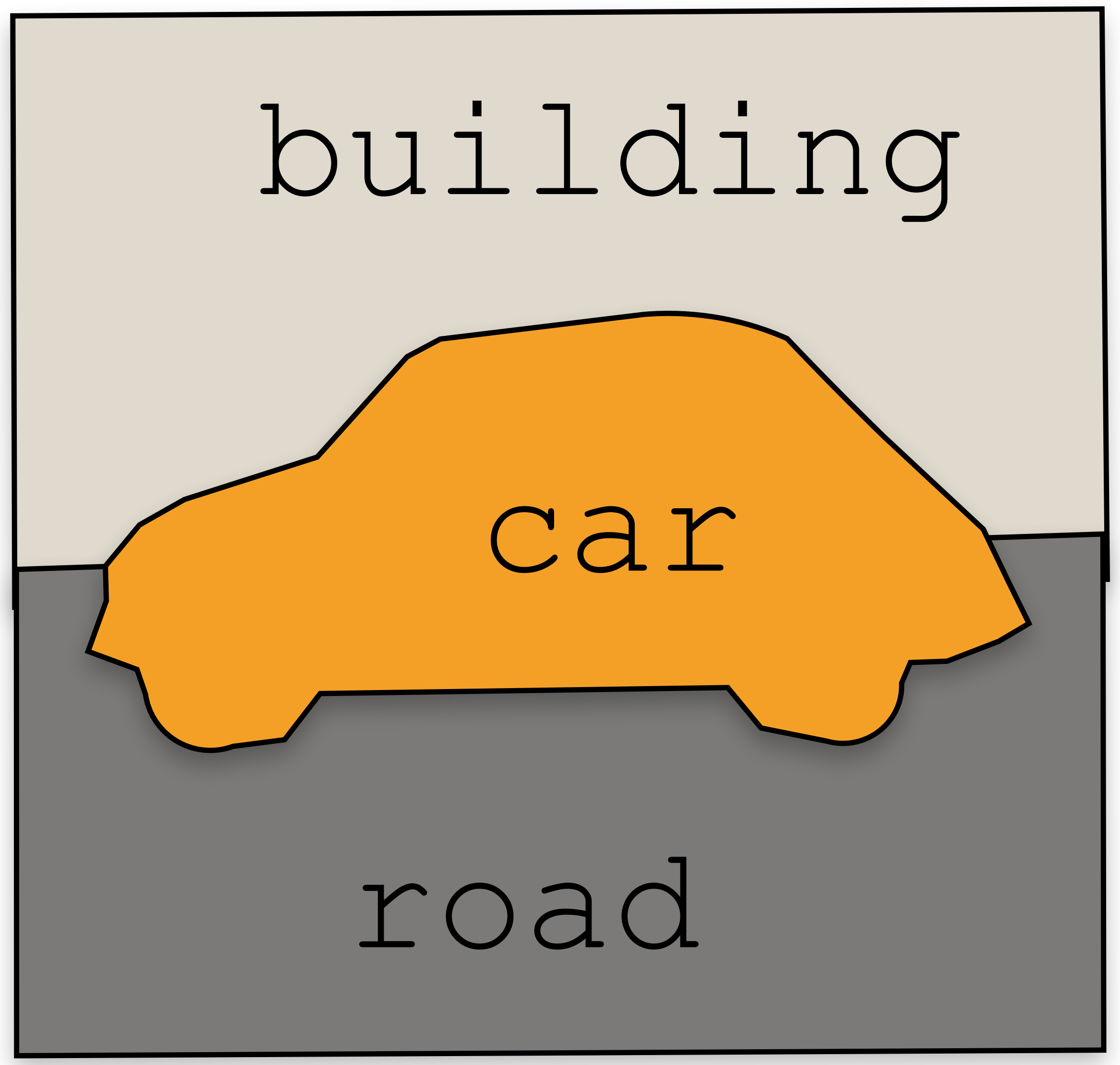
What Makes a Good Representation?



[See “Representation Learning”, Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

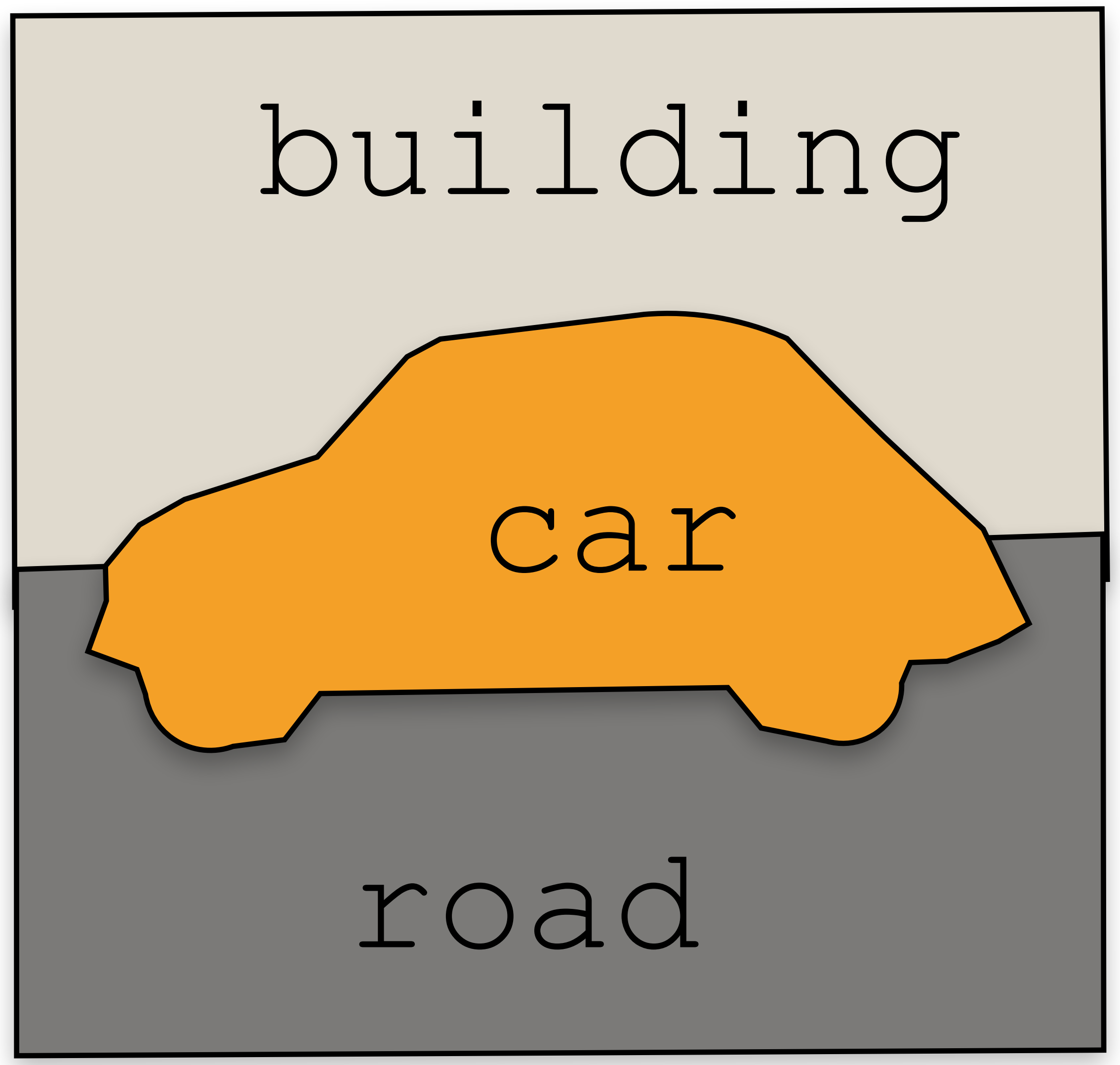


[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

1. Compact (*minimal*)



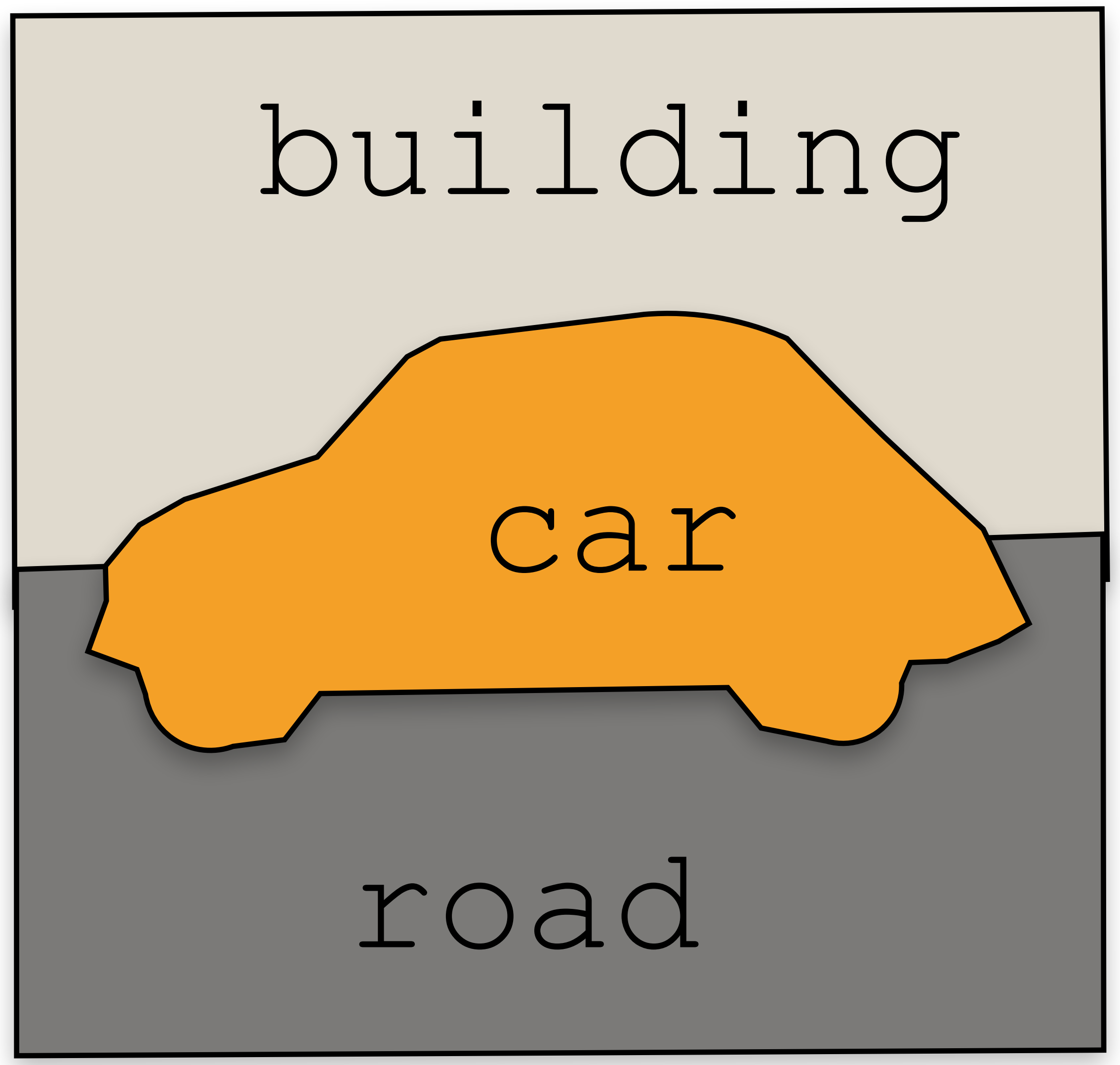
[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

1. Compact (*minimal*)

2. Explanatory (*sufficient*)

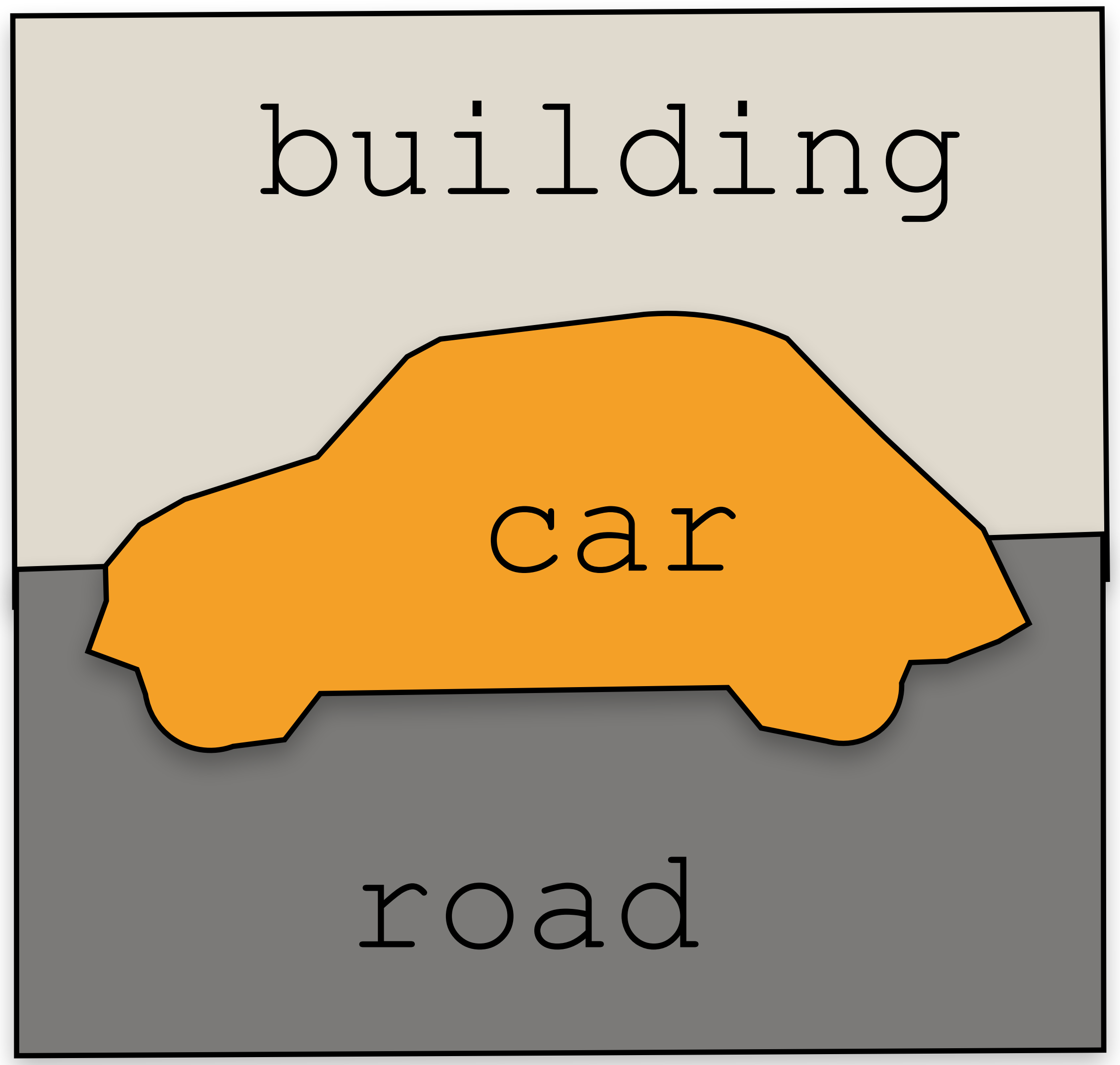


[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

1. Compact (*minimal*)
2. Explanatory (*sufficient*)
3. Disentangled (*independent factors*)

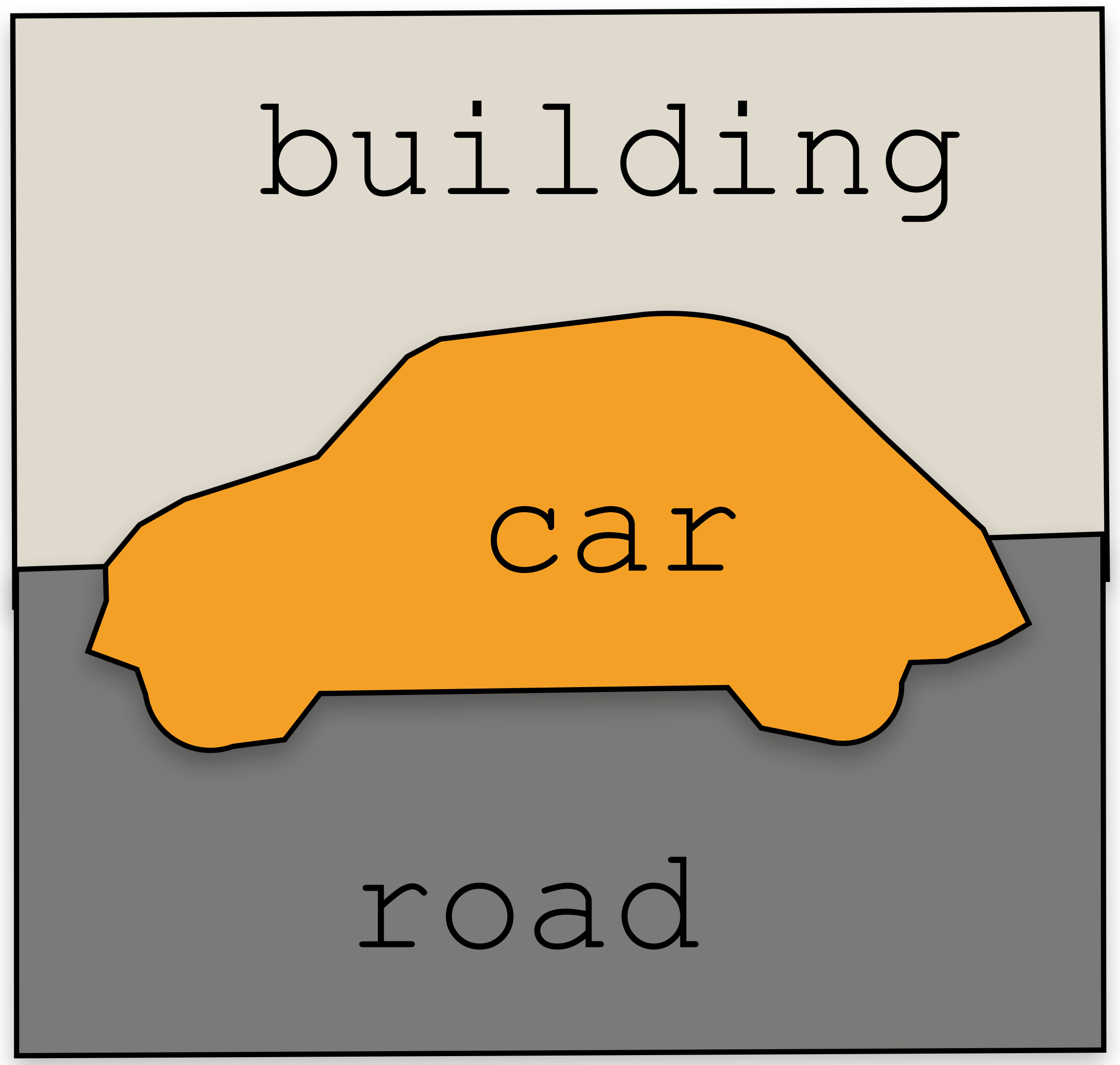


[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

1. Compact (*minimal*)
2. Explanatory (*sufficient*)
3. Disentangled (*independent factors*)
4. Equivariant (*transform predictably*)

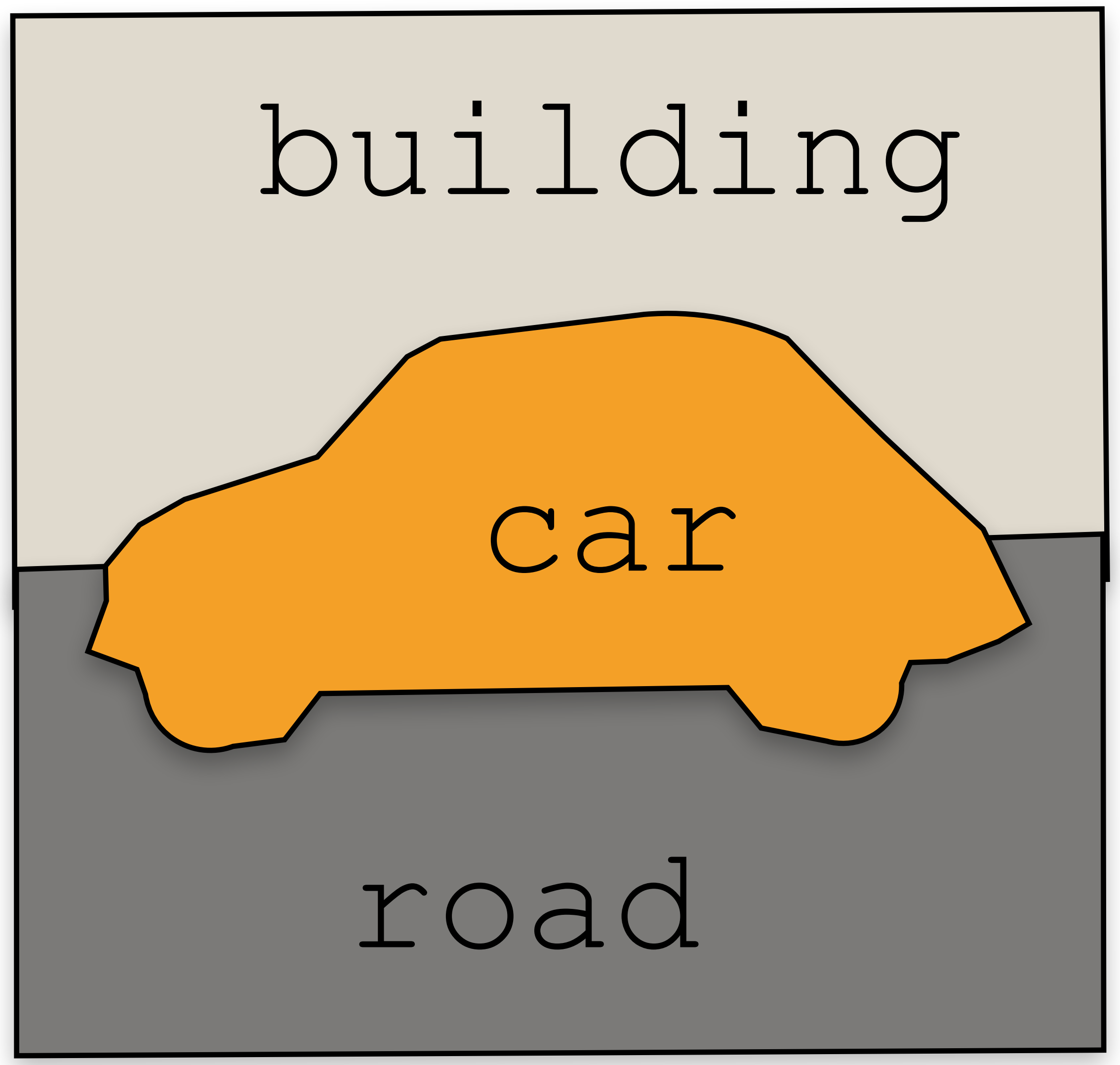


[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

1. Compact (*minimal*)
2. Explanatory (*sufficient*)
3. Disentangled (*independent factors*)
4. Equivariant (*transform predictably*)
5. Interpretable

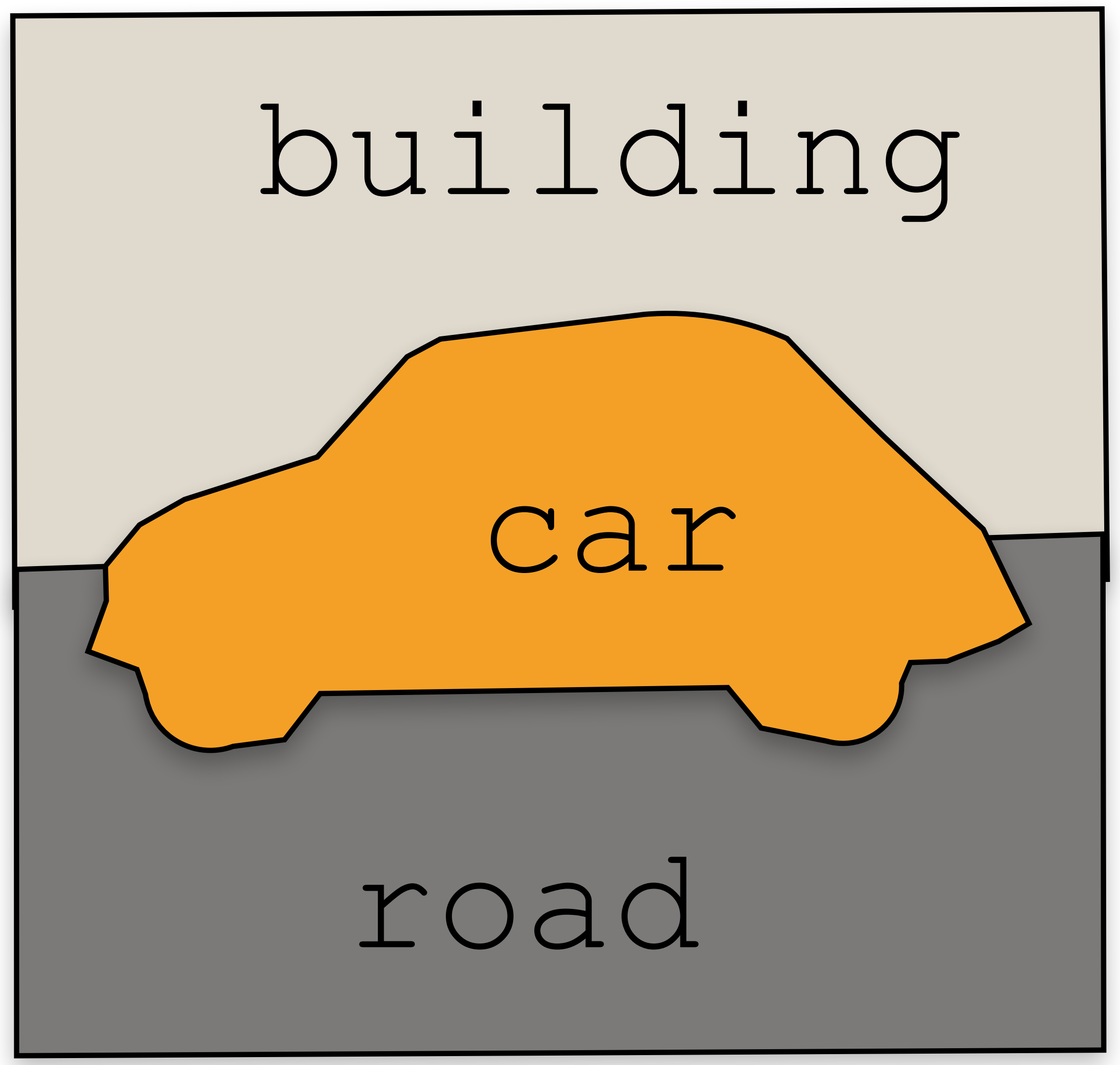


[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

1. Compact (*minimal*)
2. Explanatory (*sufficient*)
3. Disentangled (*independent factors*)
4. Equivariant (*transform predictably*)
5. Interpretable
6. *Make subsequent problem solving easy*

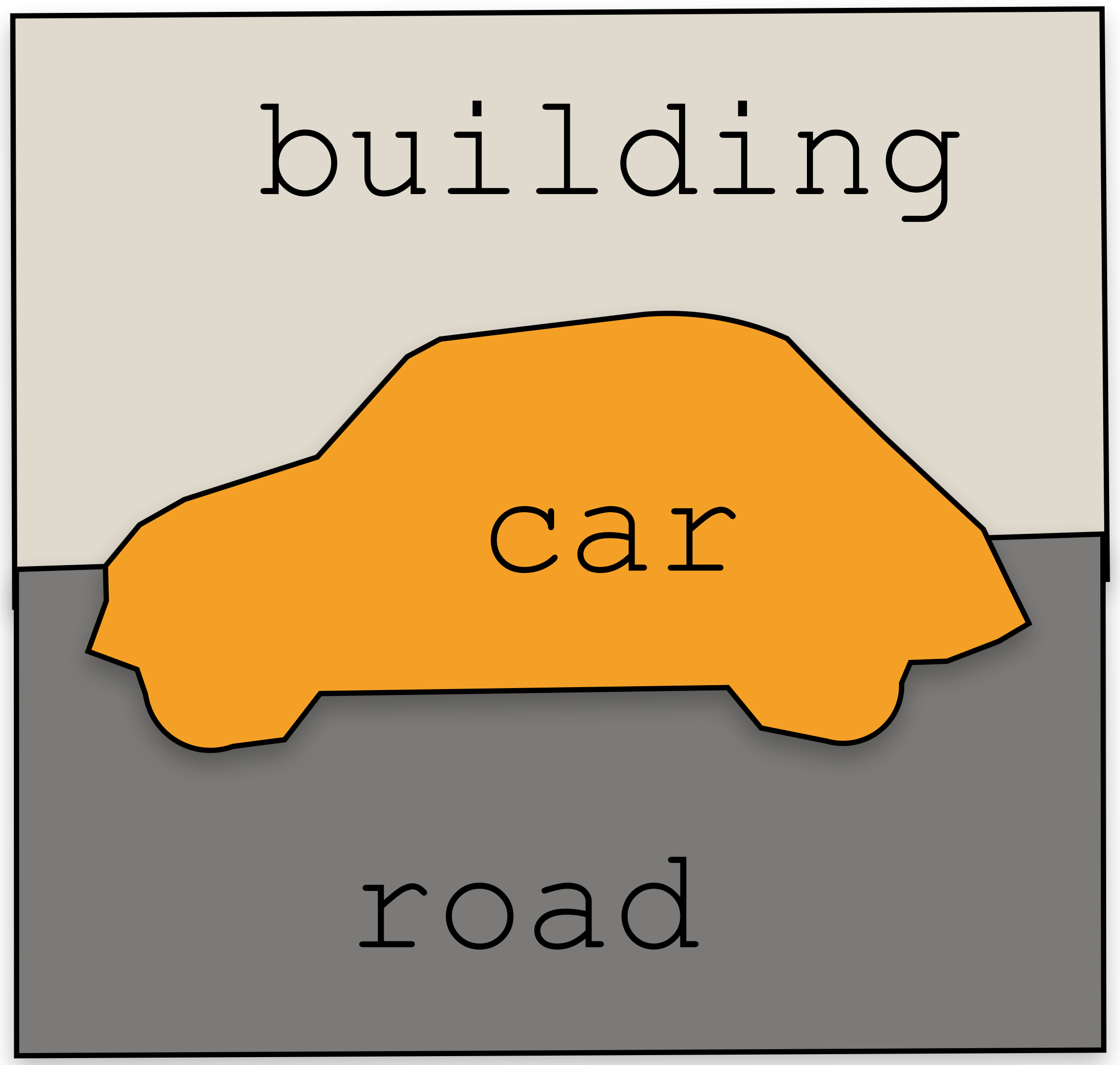


[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

1. Compact (*minimal*)
2. Explanatory (*sufficient*)
3. Disentangled (*independent factors*)
4. Equivariant (*transform predictably*)
5. Interpretable
6. *Make subsequent problem solving easy*
7. ...?



[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

This is a tall order…

If we could do this reliably, all of 3D reconstruction would be done!

[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

This is a tall order…

If we could do this reliably, all of 3D reconstruction would be done!

[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

1. Compact (*minimal*)

This is a tall order…

If we could do this reliably, all of 3D reconstruction would be done!

[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

1. Compact (*minimal*)
2. Explanatory (*sufficient*)

This is a tall order…

If we could do this reliably, all of 3D reconstruction would be done!

[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

1. Compact (*minimal*)
2. Explanatory (*sufficient*)
3. Disentangled (*independent factors*)

This is a tall order…

If we could do this reliably, all of 3D reconstruction would be done!

[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

1. Compact (*minimal*)
2. Explanatory (*sufficient*)
3. Disentangled (*independent factors*)
4. Equivariant (*transform predictably*)

This is a tall order…

If we could do this reliably, all of 3D reconstruction would be done!

[See “Representation Learning”, Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

1. Compact (*minimal*)
2. Explanatory (*sufficient*)
3. Disentangled (*independent factors*)
4. Equivariant (*transform predictably*)
5. Interpretable

This is a tall order…

If we could do this reliably, all of 3D reconstruction would be done!

[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

1. Compact (*minimal*)
2. Explanatory (*sufficient*)
3. Disentangled (*independent factors*)
4. Equivariant (*transform predictably*)
5. Interpretable

This is a tall order…

If we could do this reliably, all of 3D reconstruction would be done!

6. *Make subsequent problem solving easy*

[See "Representation Learning", Bengio 2013, for more commentary]

What Makes a Good Representation?

Good representations are:

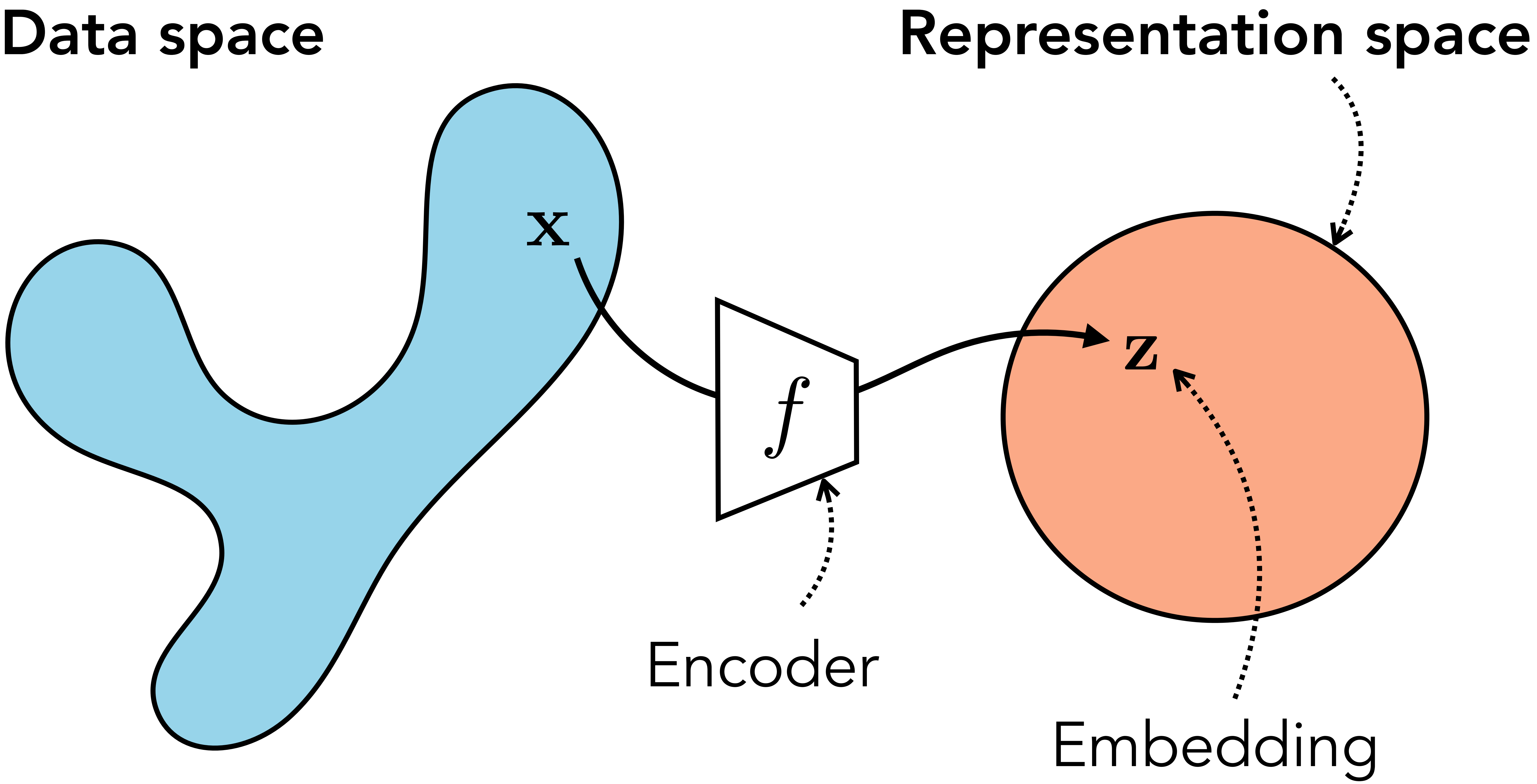
1. Compact (*minimal*)
2. Explanatory (*sufficient*)
3. Disentangled (*independent factors*)
4. Equivariant (*transform predictably*)
5. Interpretable
6. *Make subsequent problem solving easy*
7. …?

This is a tall order…

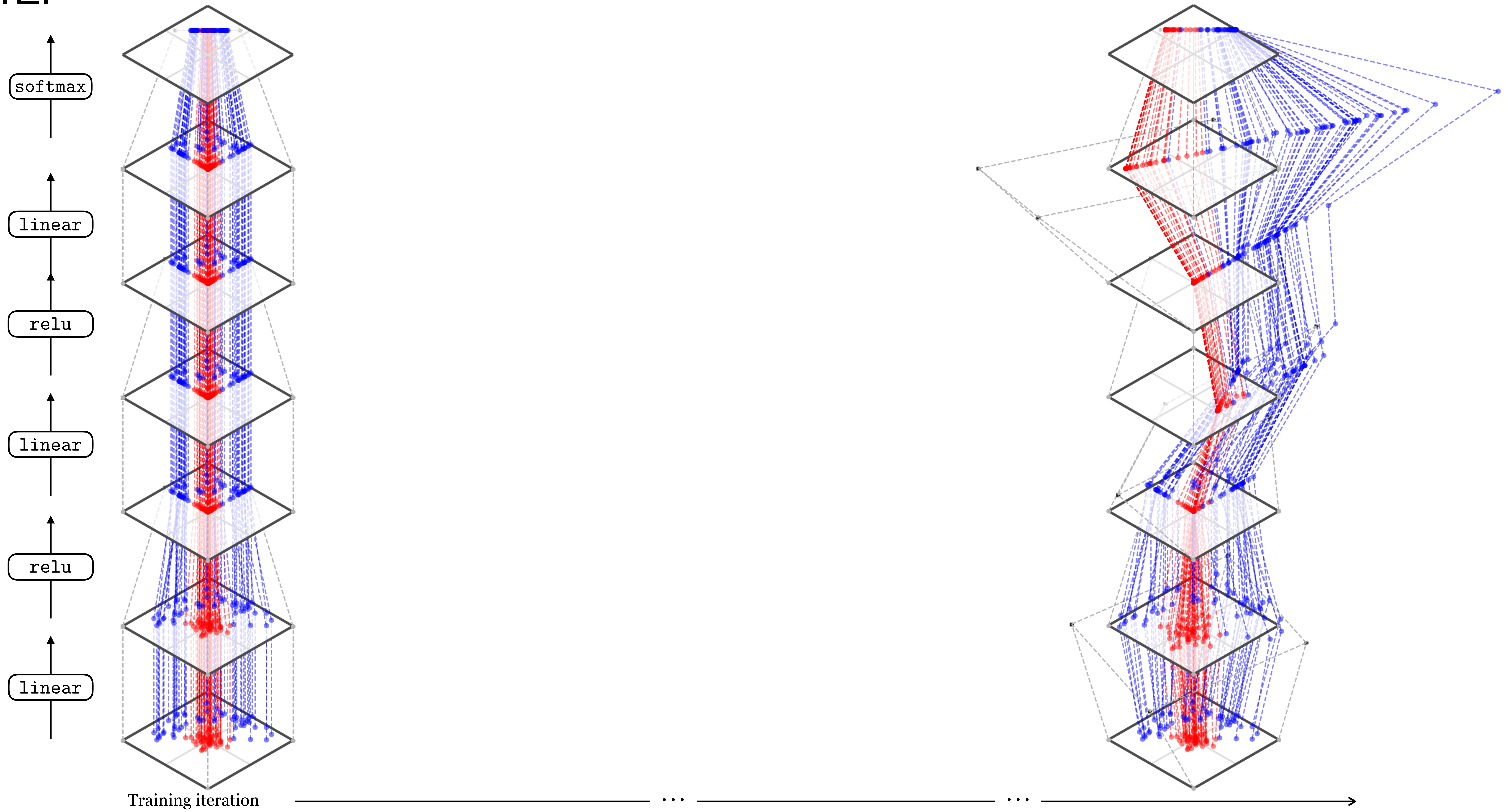
If we could do this reliably, all of 3D reconstruction would be done!

[See "Representation Learning", Bengio 2013, for more commentary]

x2vec

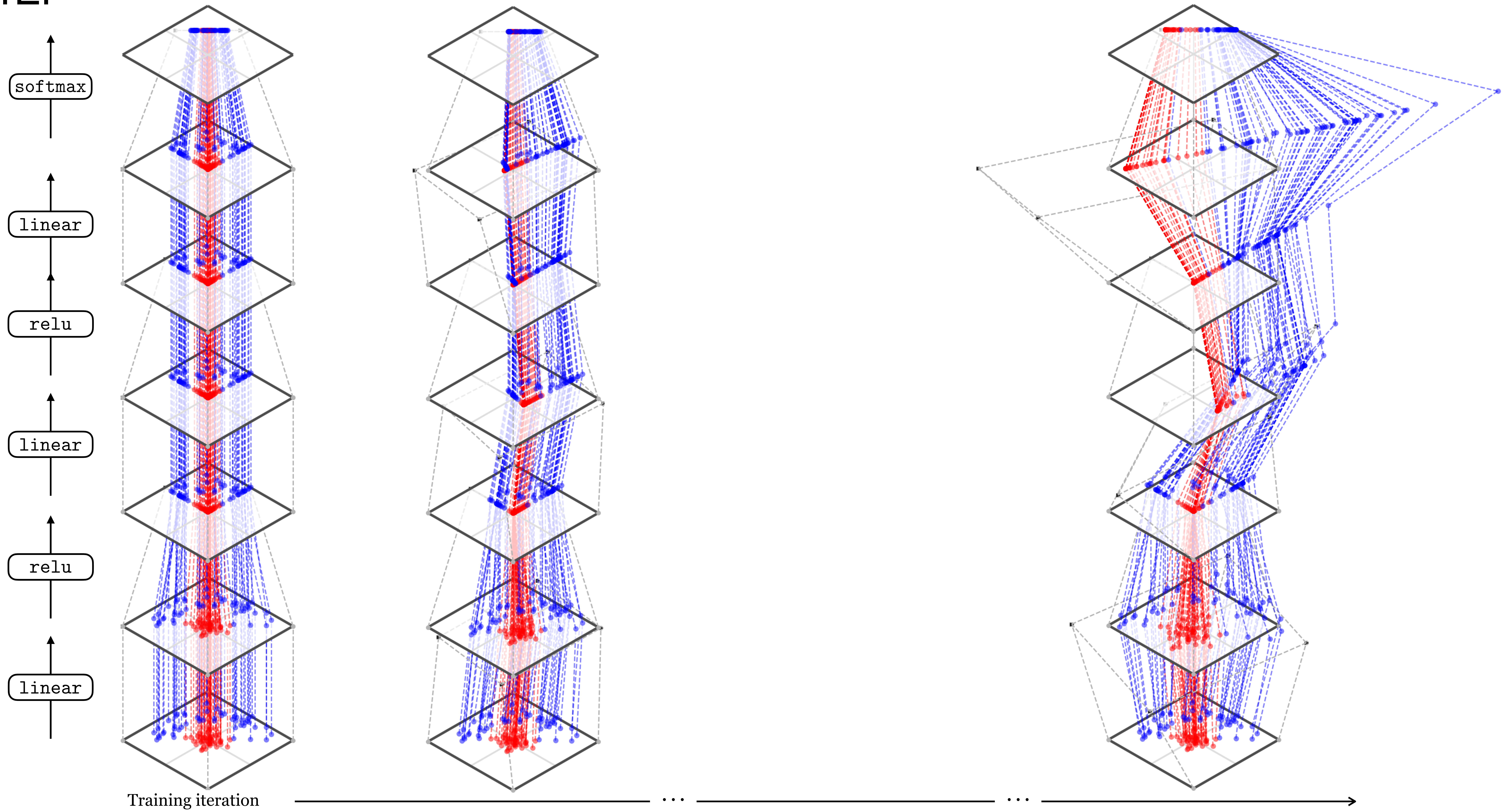


MLP



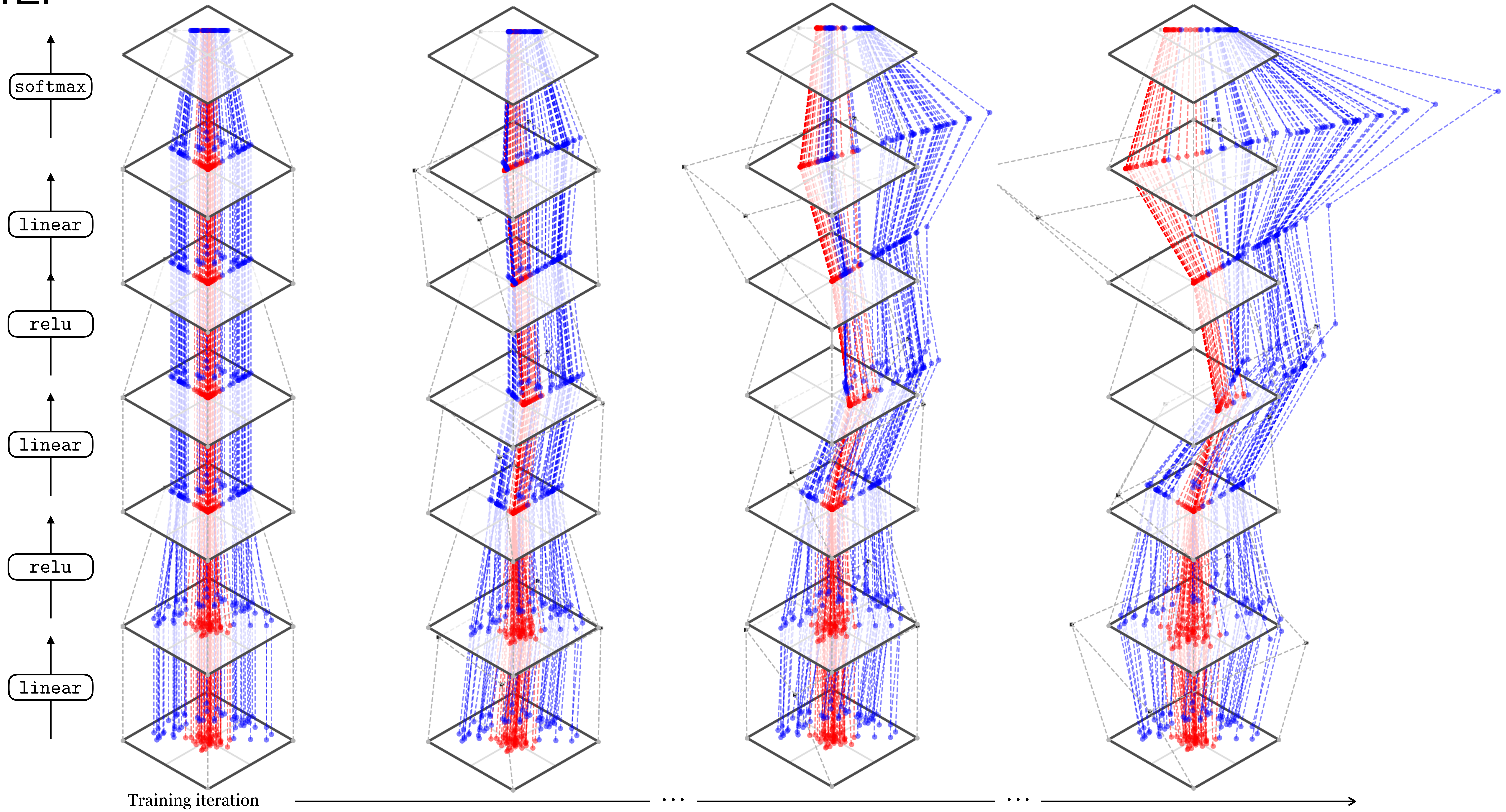
softmax
linear
relu
linear
relu
linear
Training iteration
...
...

MLP

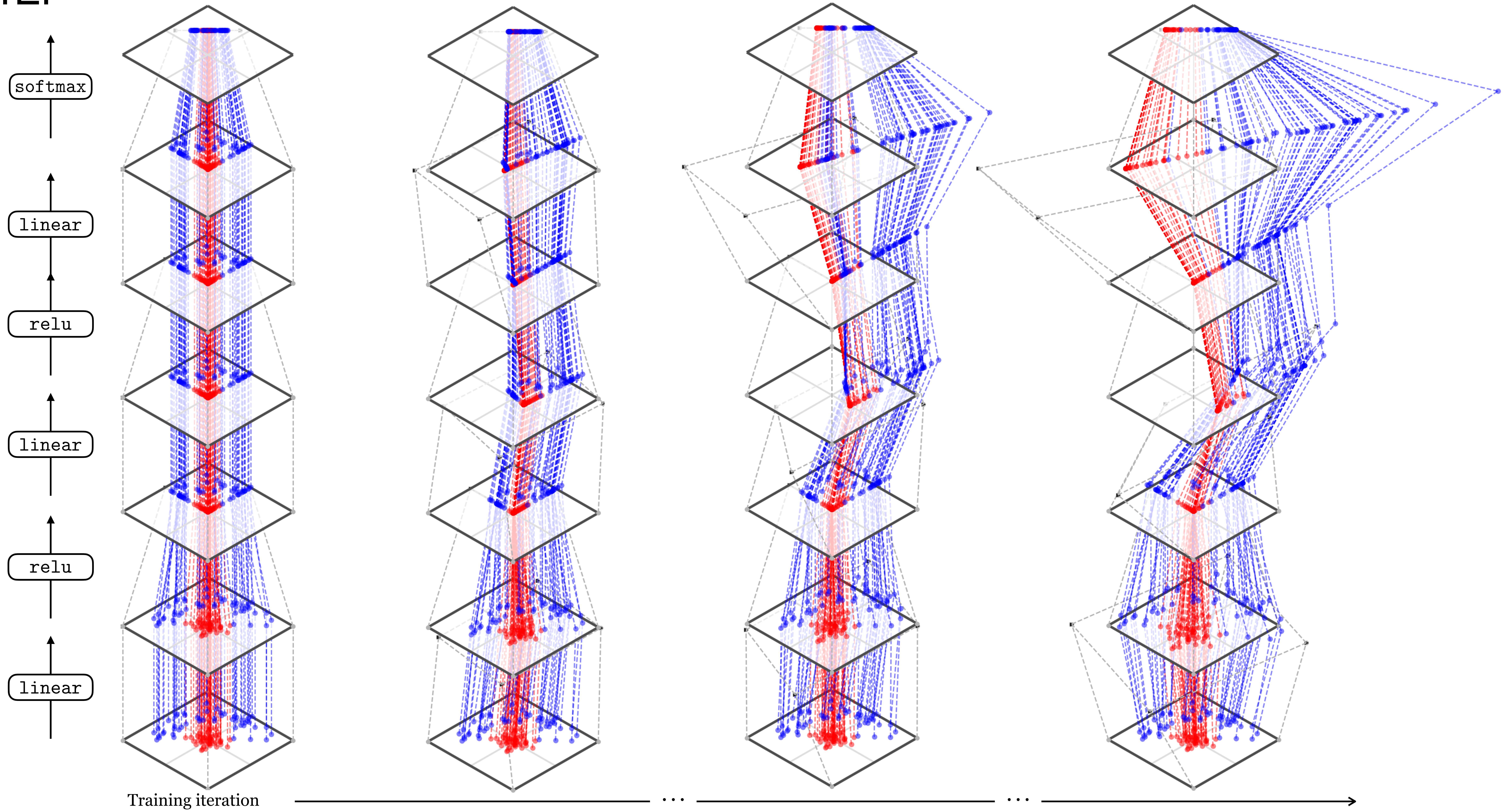


softmax
linear
relu
linear
relu
linear
Training iteration
...
...

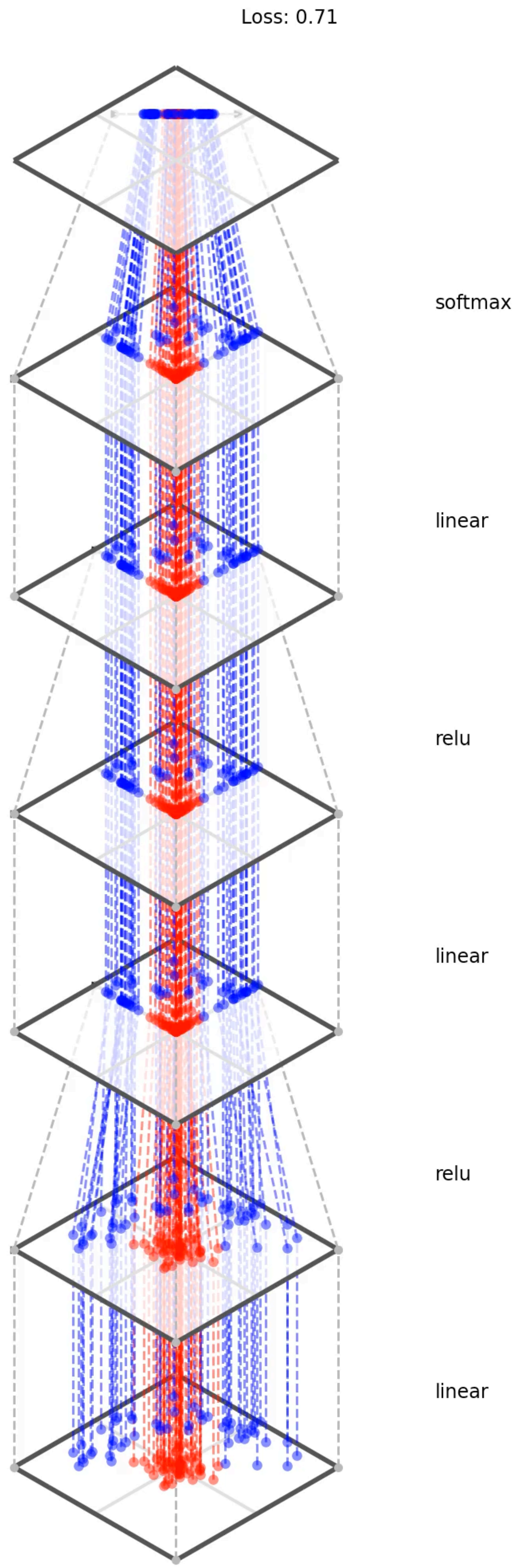
MLP


softmax
linear
relu
linear
relu
linear
Training iteration
...
...

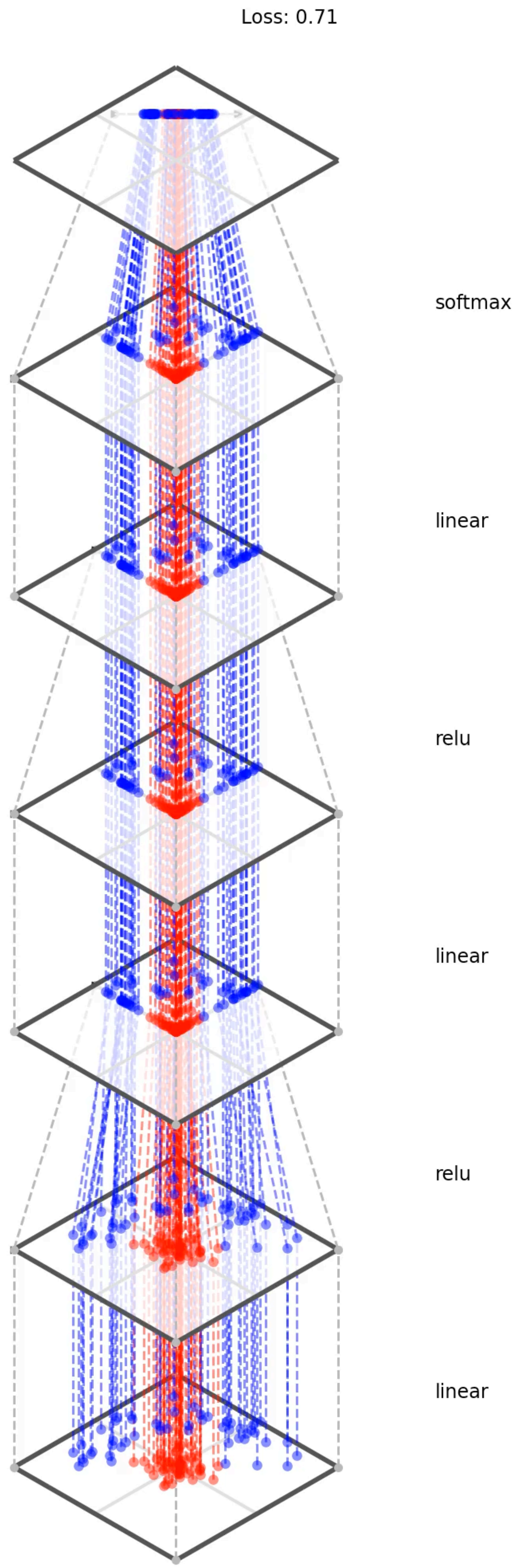
MLP


softmax
linear
relu
linear
relu
linear
Training iteration
...
...

MLP

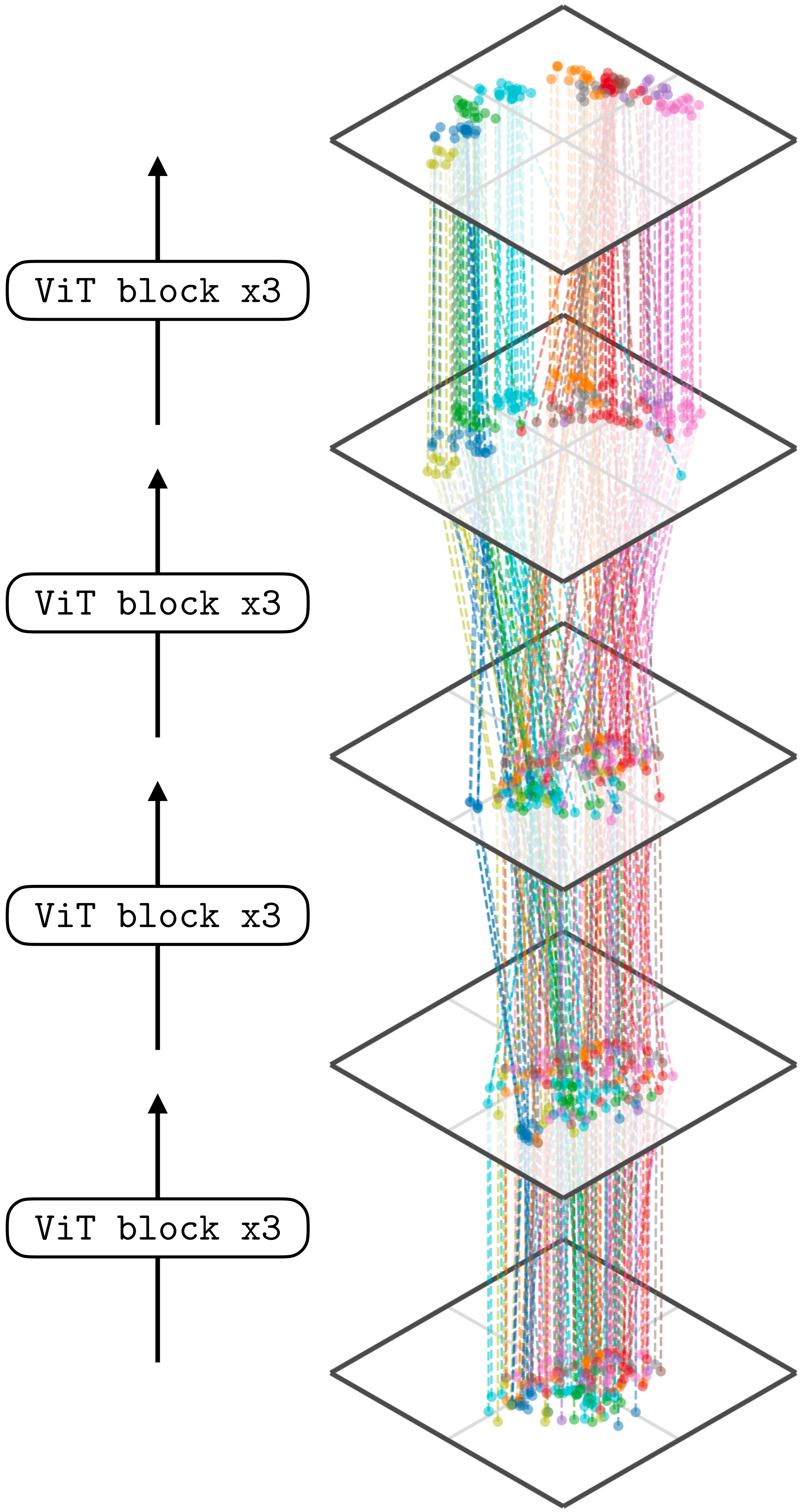


MLP



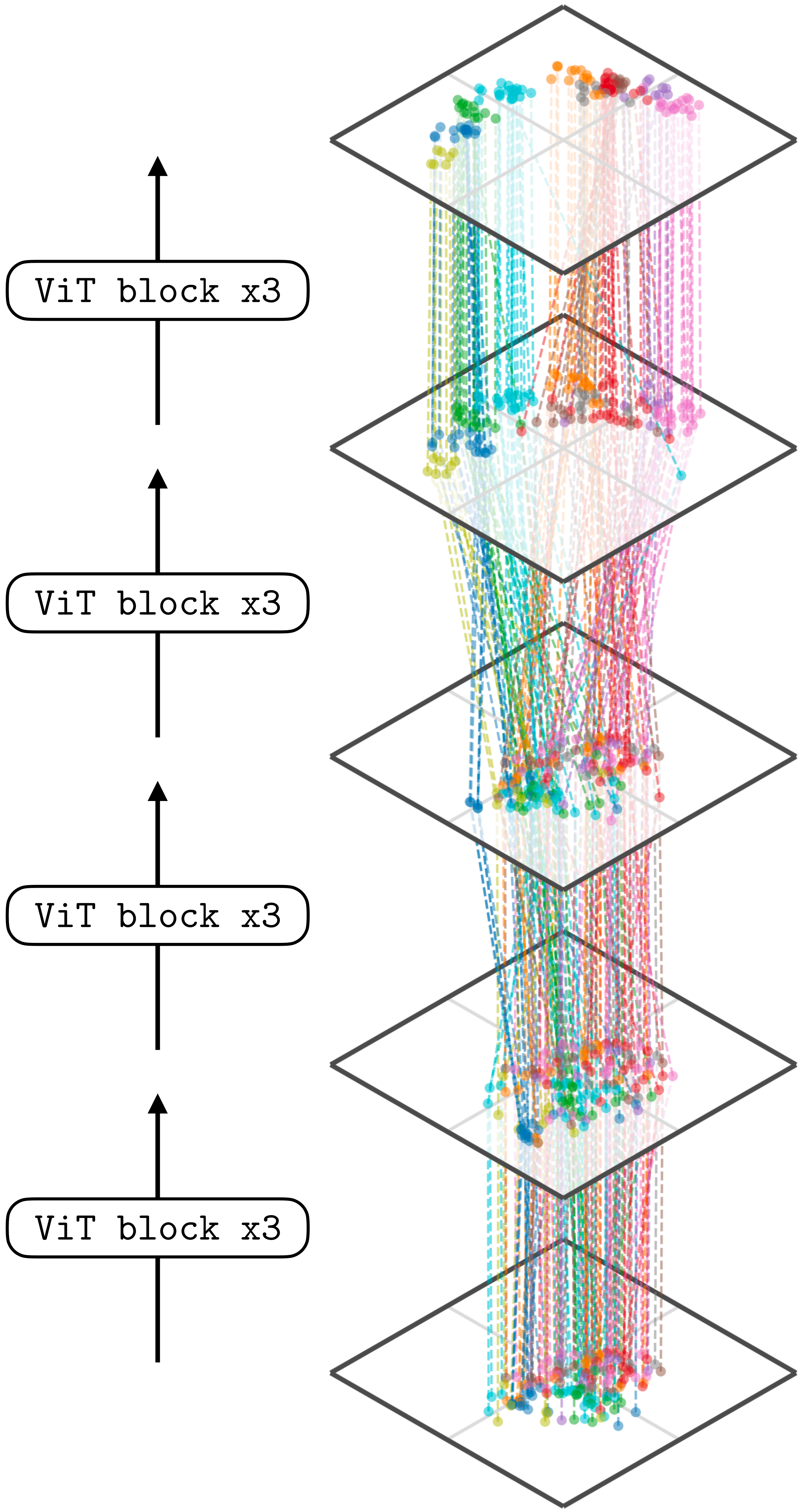
CLIP

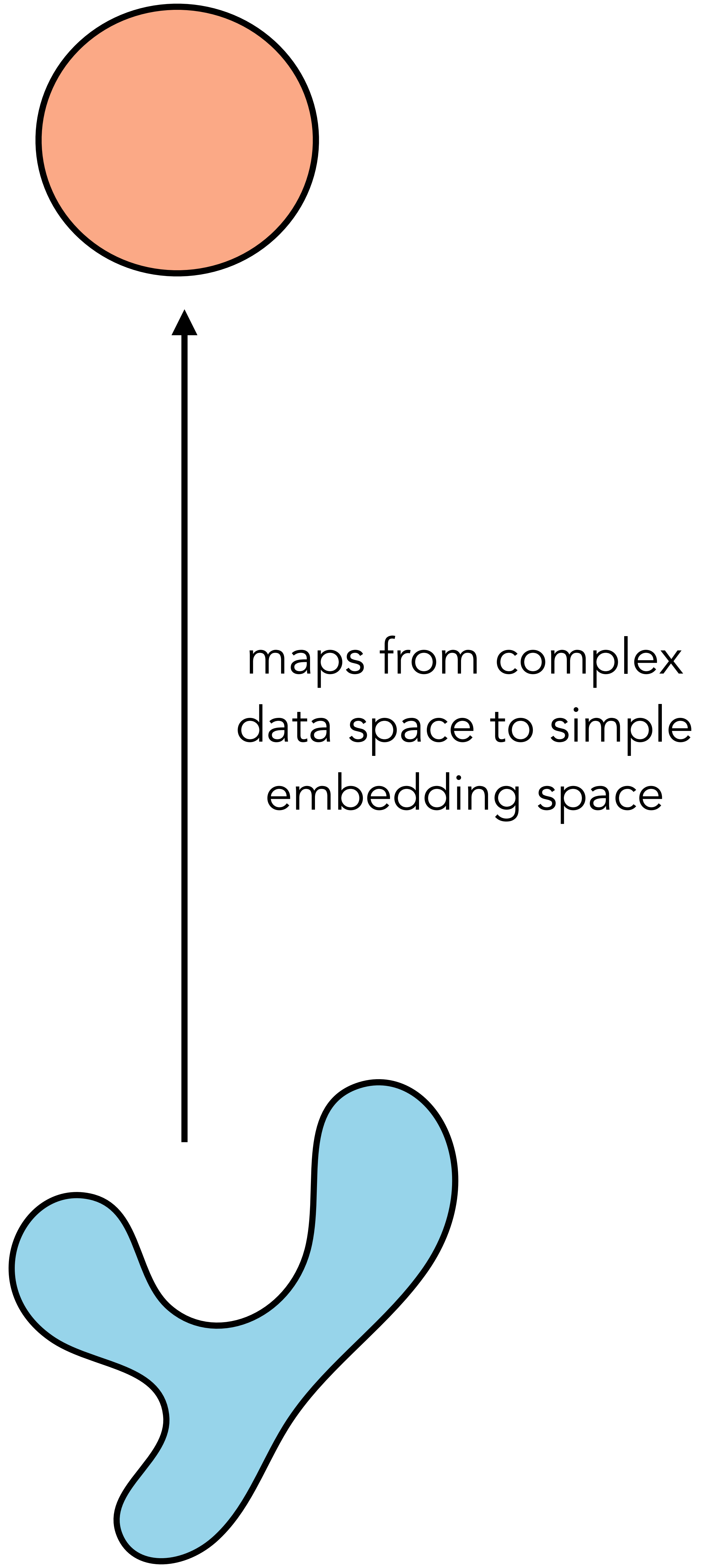
[Radford*, Kim* et al., ICML 2021]



CLIP

[Radford*, Kim* et al., ICML 2021]

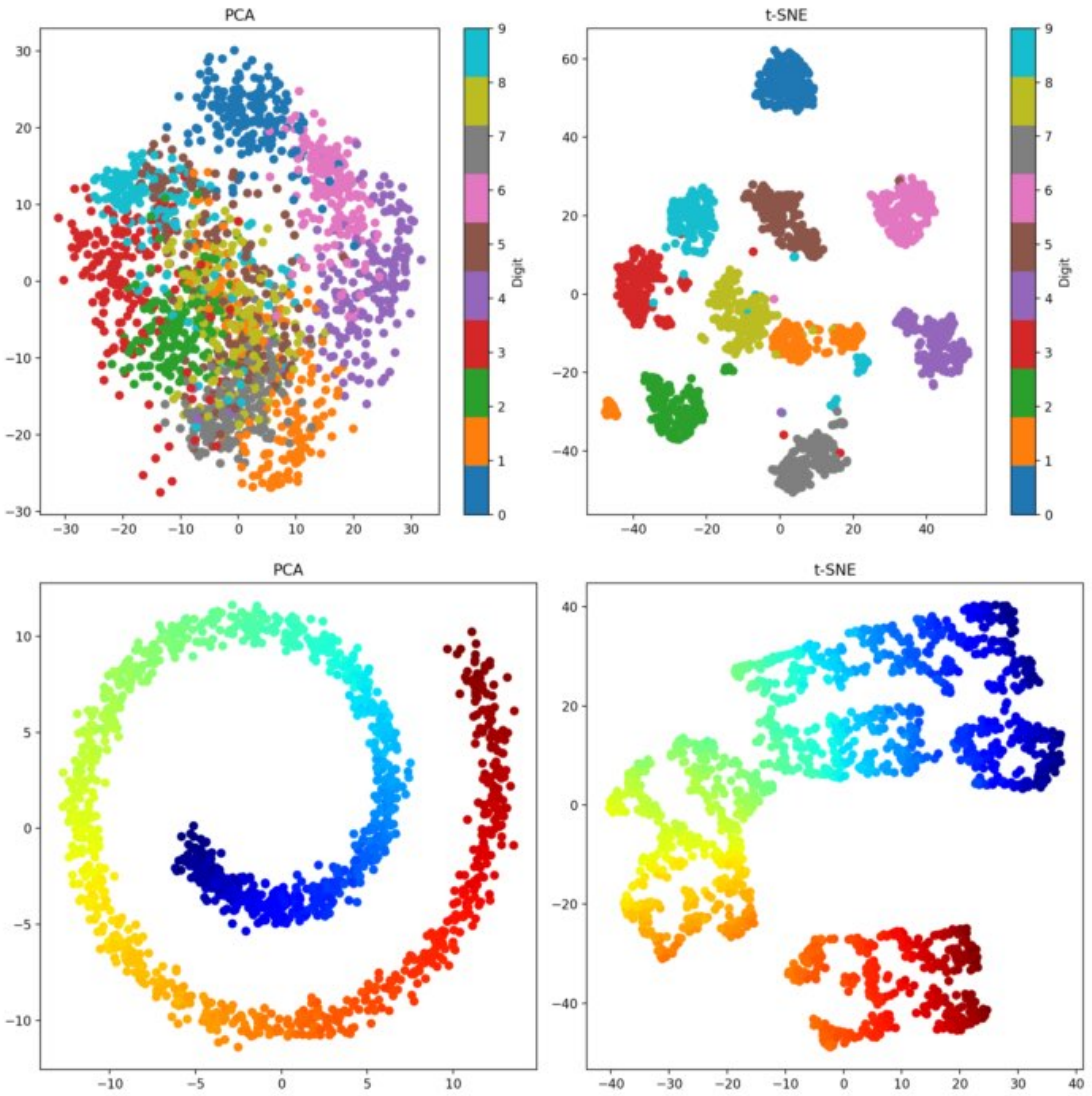


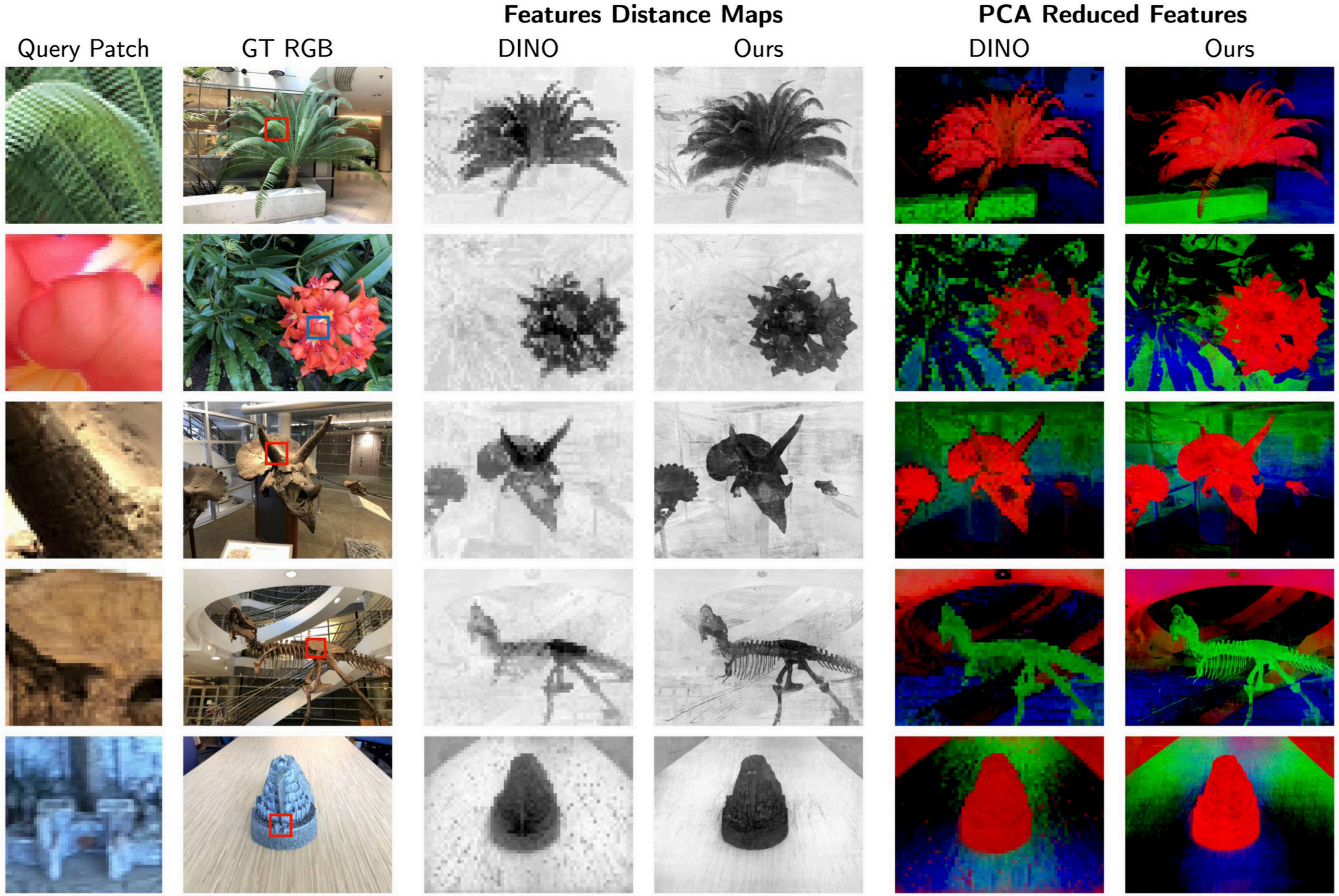


How do we test if a representation is “good”?

- Dimensionality Reduction & Visualization
- Check downstream task performance of features
  - Train small “linear probe” on top of representation
  - Fine-tune network for some downstream task

https://www.fabriziomusacchio.com/blog/2023-06-12-tsne_vs_pca/





From Neural Feature Fusion Fields: 3D Distillation of Self-Supervised 2D Image Representations

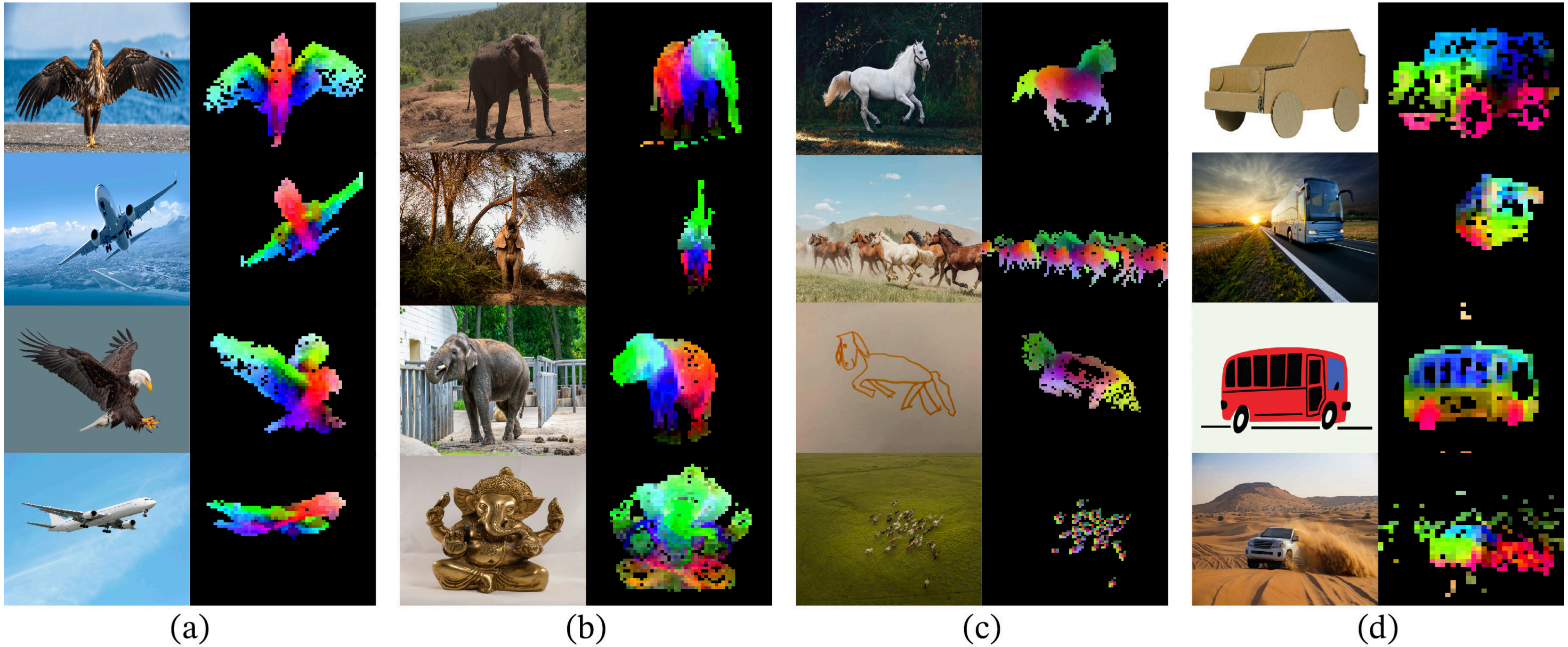


Figure 1: **Visualization of the first PCA components.** We compute a PCA between the patches of the images from the same column (a, b, c and d) and show their first 3 components. Each component is matched to a different color channel. Same parts are matched between related images despite changes of pose, style or even objects. Background is removed by thresholding the first PCA component.

From DinoV2: Learning Robust Visual Features without Supervision, Oquab et al. 2024

Table 2: **Linear and k-NN classification on ImageNet.** We report top-1 accuracy for linear and k-NN evaluations on the validation set of ImageNet for different self-supervised methods. We focus on ResNet-50 and ViT-small architectures, but also report the best results obtained across architectures. * are run by us. We run the k-NN evaluation for models with official released weights. The throughput (im/s) is calculated on a NVIDIA V100 GPU with 128 samples per forward. Parameters (M) are of the feature extractor.

Method	Arch.	Param.	im/s	Linear	k-NN
Supervised	RN50	23	1237	79.3	79.3
SCLR [11]	RN50	23	1237	69.1	60.7
MoCov2 [13]	RN50	23	1237	71.1	61.9
InfoMin [54]	RN50	23	1237	73.0	65.3
BarlowT [66]	RN50	23	1237	73.2	66.0
OBoW [21]	RN50	23	1237	73.8	61.9
BYOL [23]	RN50	23	1237	74.4	64.8
DCv2 [9]	RN50	23	1237	75.2	67.1
SwAV [9]	RN50	23	1237	**75.3**	65.7
DINO	RN50	23	1237	**75.3**	**67.5**
Supervised	ViT-S	21	1007	79.8	79.8
BYOL* [23]	ViT-S	21	1007	71.4	66.6
MoCov2* [13]	ViT-S	21	1007	72.7	64.4
SwAV* [9]	ViT-S	21	1007	73.5	66.3
DINO	ViT-S	21	1007	**77.0**	**74.5**

From DINO (Emerging Properties in Self-Supervised Vision Transformers), Caron et al. 2021

Method	Arch.	NYUd (0.330) lin. 1	NYUd (0.330) lin. 4	NYUd (0.330) DPT	KITTI (2.10) lin. 1	KITTI (2.10) lin. 4	KITTI (2.10) DPT	NYUd → SUN RGB-D (0.421) lin. 1	NYUd → SUN RGB-D (0.421) lin. 4	NYUd → SUN RGB-D (0.421) DPT
OpenCLIP	ViT-G/14	0.541	0.510	0.414	3.57	3.21	2.56	0.537	0.476	0.408
MAE	ViT-H/14	0.517	0.483	0.415	3.66	3.26	2.59	0.545	0.523	0.506
DINO	ViT-B/8	0.555	0.539	0.492	3.81	3.56	2.74	0.553	0.541	0.520
iBOT	ViT-L/16	0.417	0.387	0.358	3.31	3.07	2.55	0.447	0.435	0.426
DINOv2	ViT-S/14	0.449	0.417	0.356	3.10	2.86	2.34	0.477	0.431	0.409
	ViT-B/14	0.399	0.362	0.317	2.90	2.59	2.23	0.448	0.400	0.377
	ViT-L/14	0.384	0.333	0.293	2.78	2.50	2.14	0.429	0.396	0.360
	ViT-g/14	**0.344**	**0.298**	**0.279**	**2.62**	**2.35**	**2.11**	**0.402**	**0.362**	**0.338**

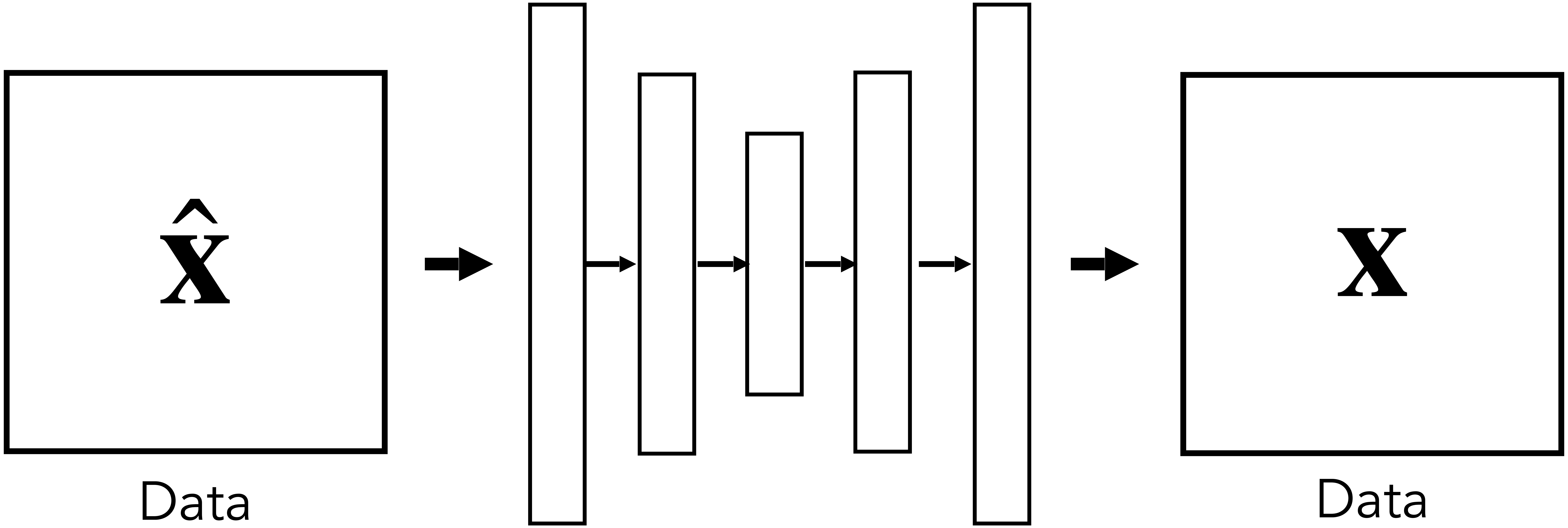
Table 11: **Depth estimation with frozen features**. We report performance when training a linear classifier on top of one (lin. 1) or four (lin. 4) transformer layers, as well, as the DPT decoder (DPT) of Ranftl et al. (2021). We report the RMSE metric on the 3 datasets. Lower is better. For reference, we report state-of-the-art results taken from Li et al. (2022b) on each benchmark on top of the Table.

From DinoV2: Learning Robust Visual Features without Supervision, Oquab et al. 2024

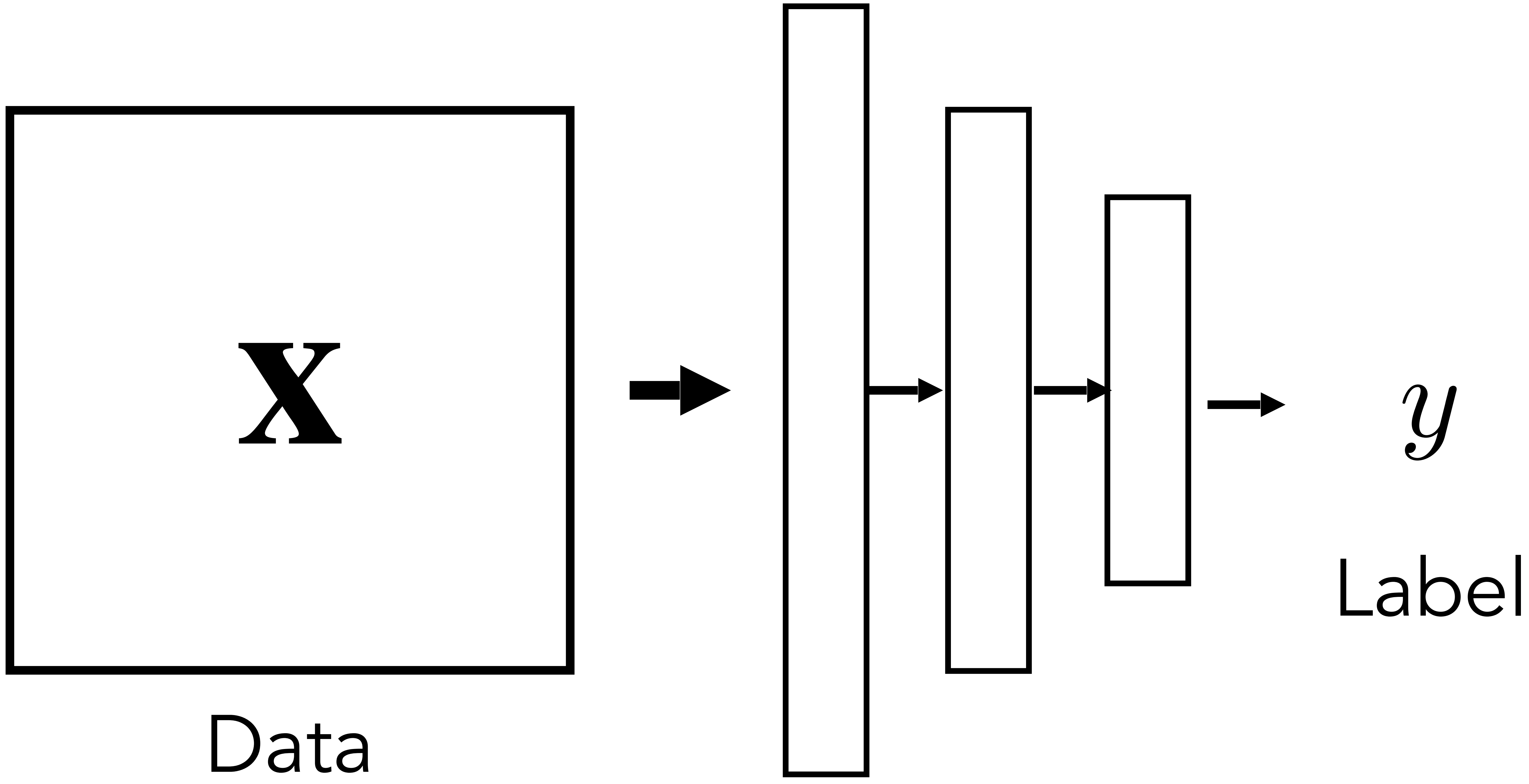
How to train them? Two basic approaches: Compression and Prediction

Learning Method	Learning Principle	Short Summary
Autoencoding	Compression	Remove redundant information
Contrastive	Compression	Achieve invariance to viewing transformations
Clustering	Compression	Quantize continuous data into discrete categories
"Prediction"	Prediction	Predict the future
Imputation	Prediction	Predict missing data
Pretext tasks	Prediction	Predict abstract properties of your data

Data compression

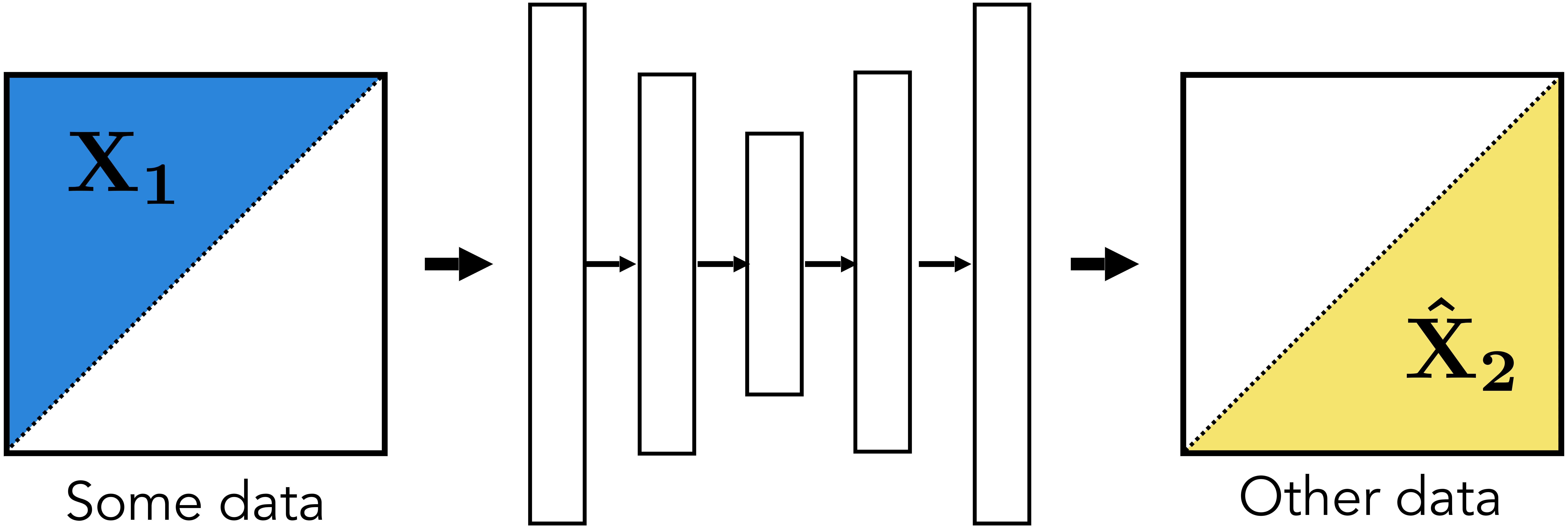


Label prediction

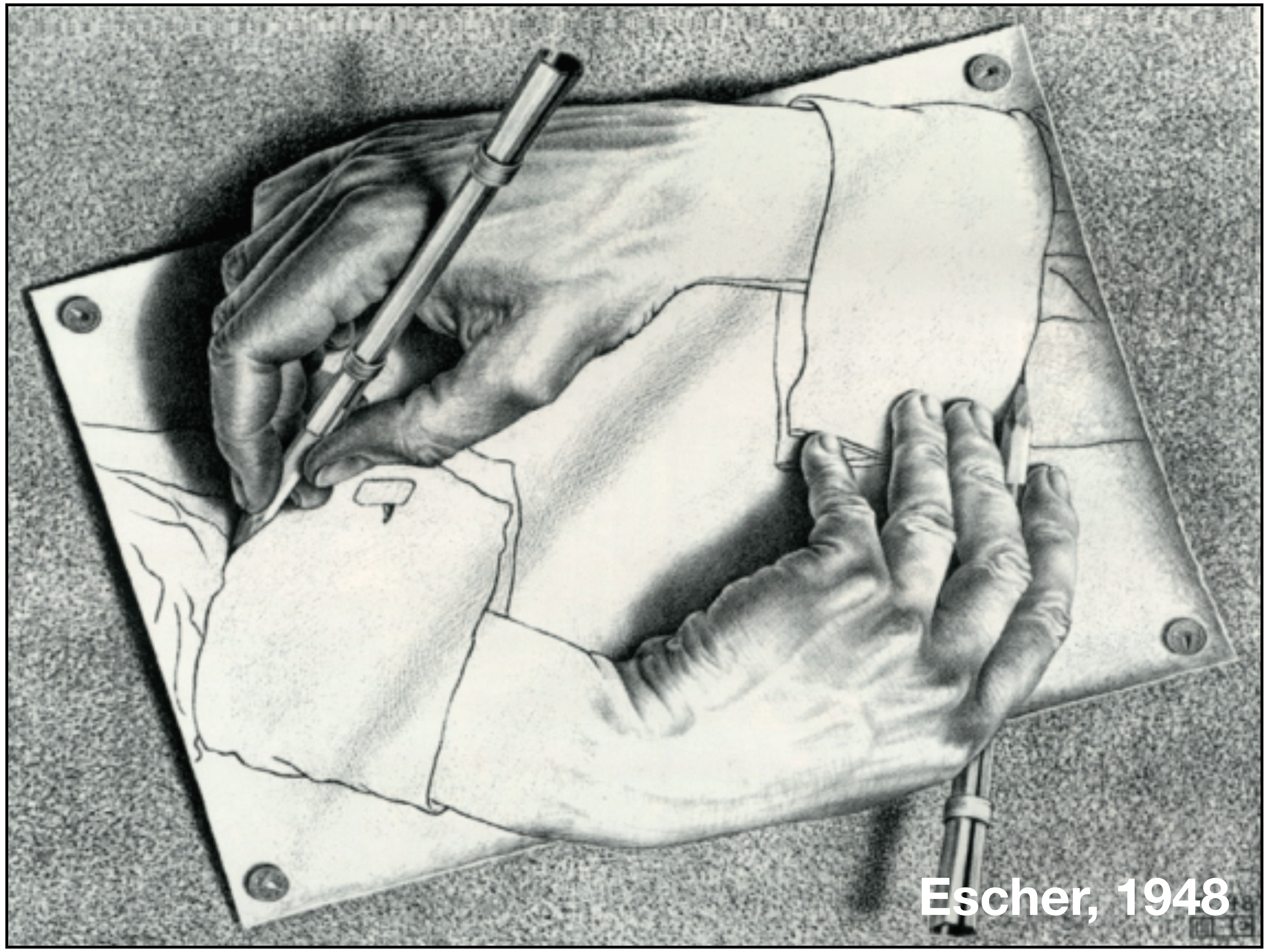


Data prediction

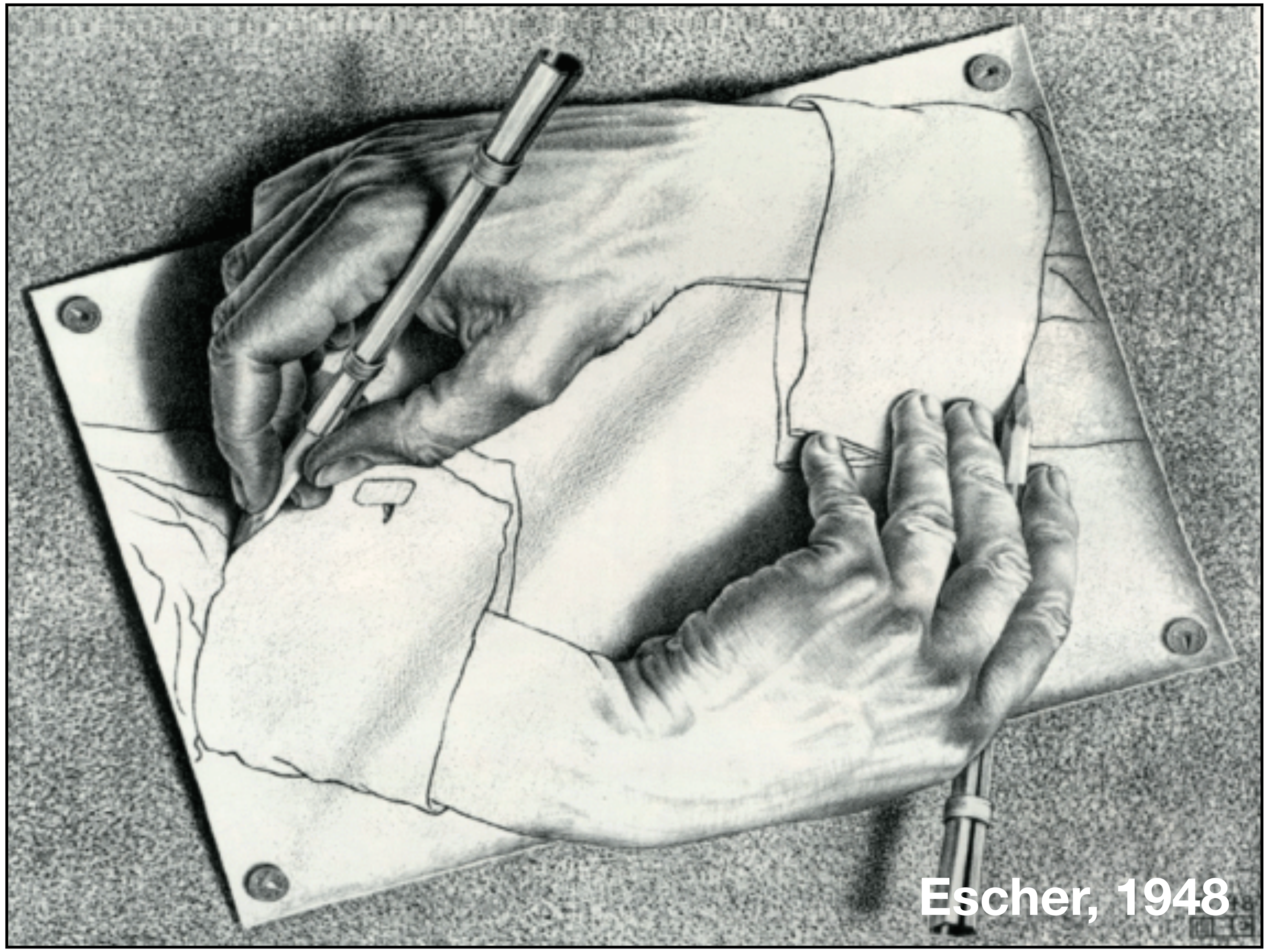
aka "self-supervised learning"



Self-supervised learning

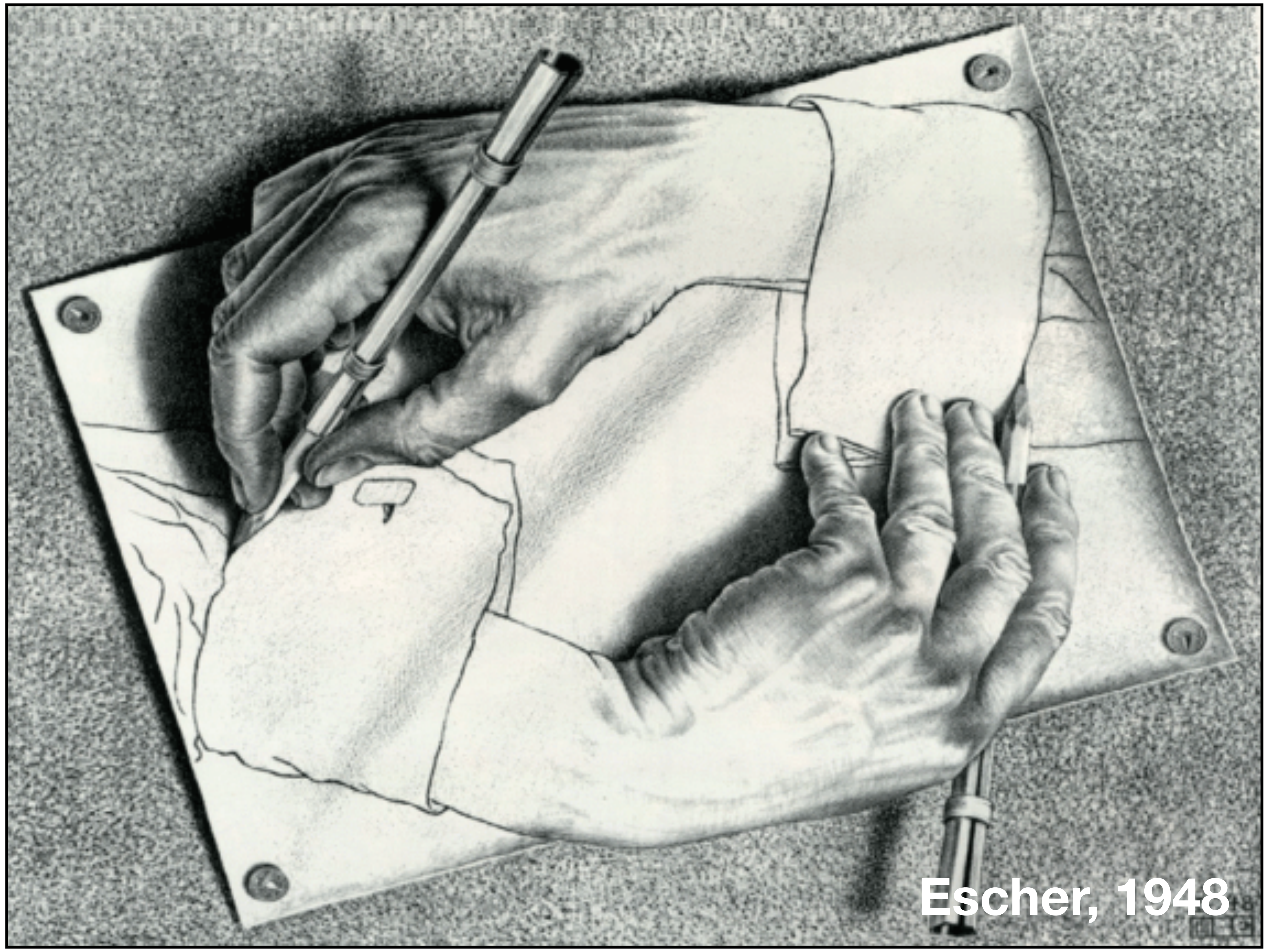


Self-supervised learning



Common trick:

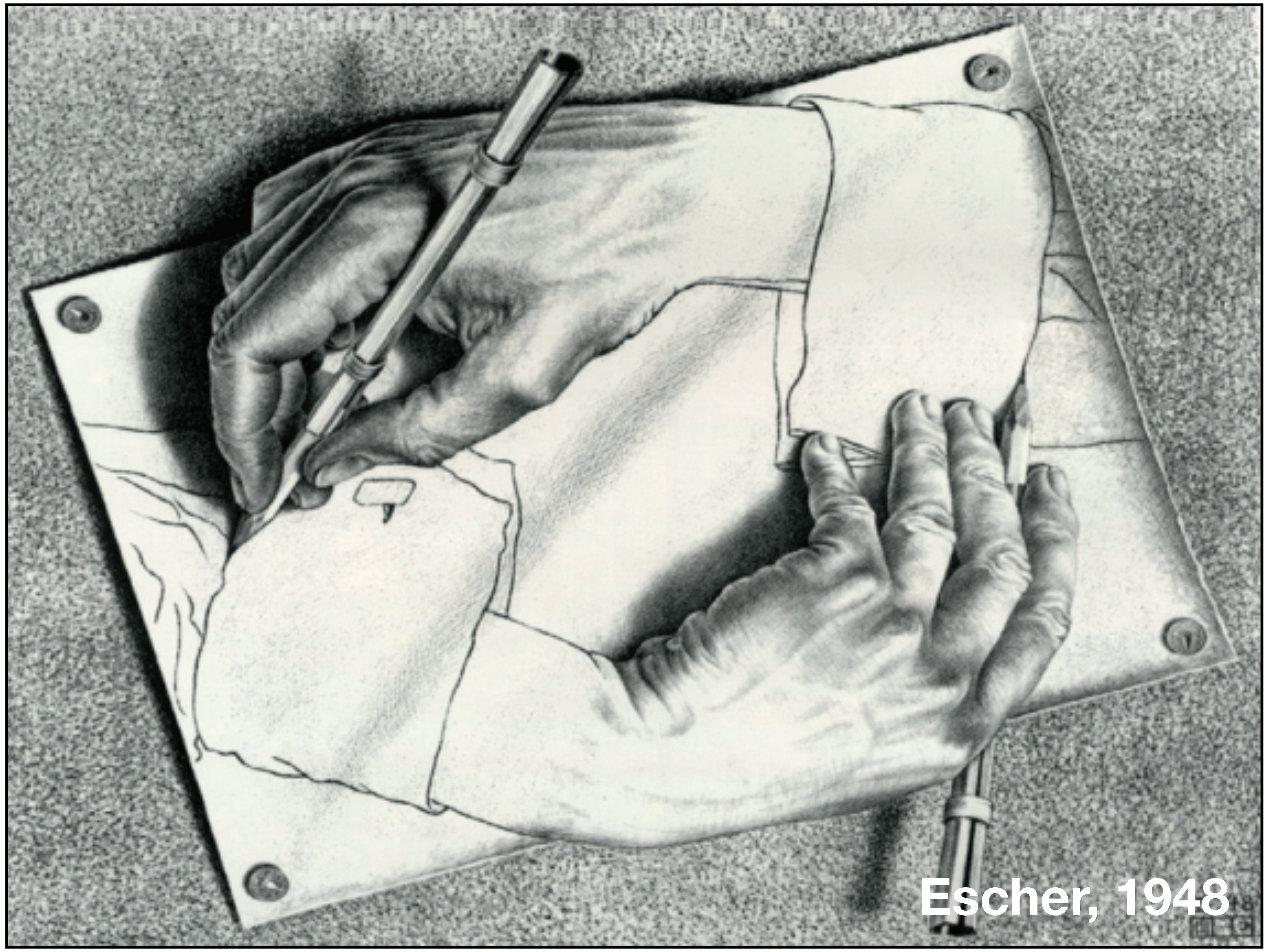
Self-supervised learning



Common trick:

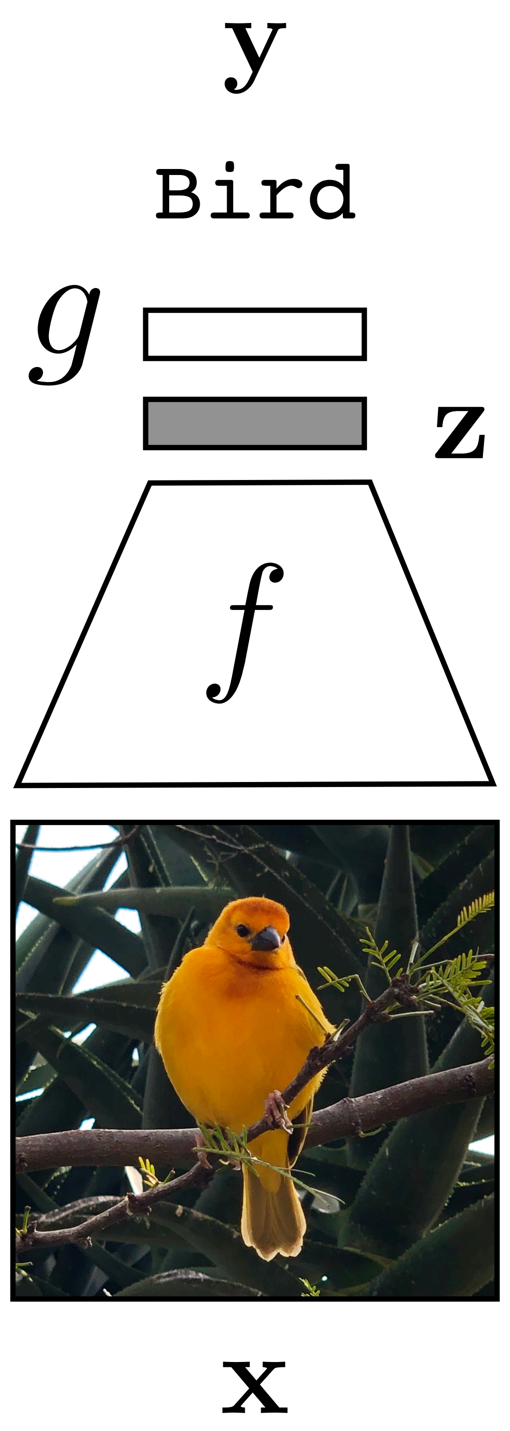
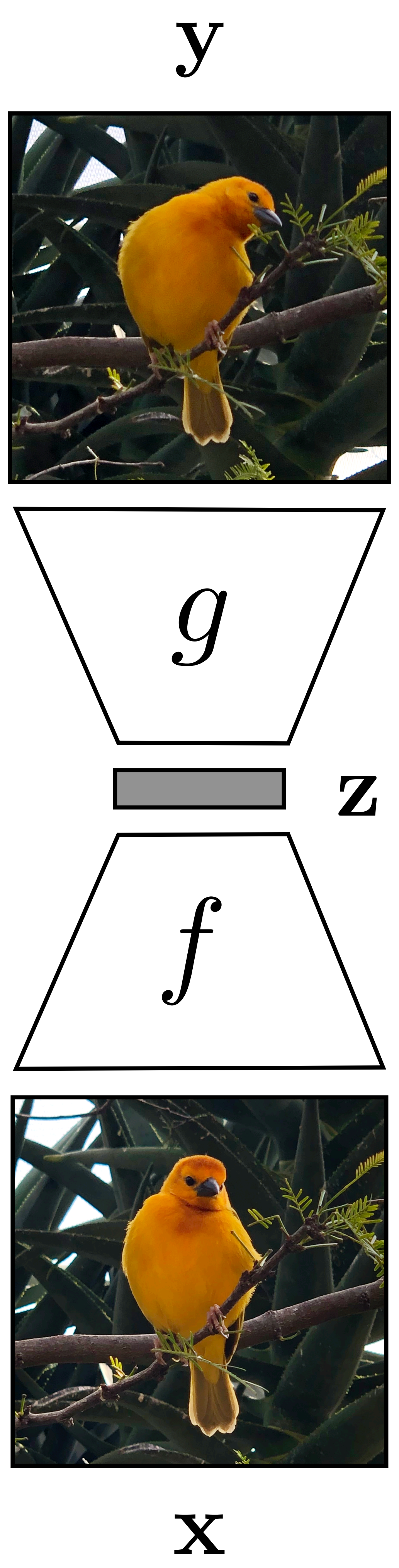
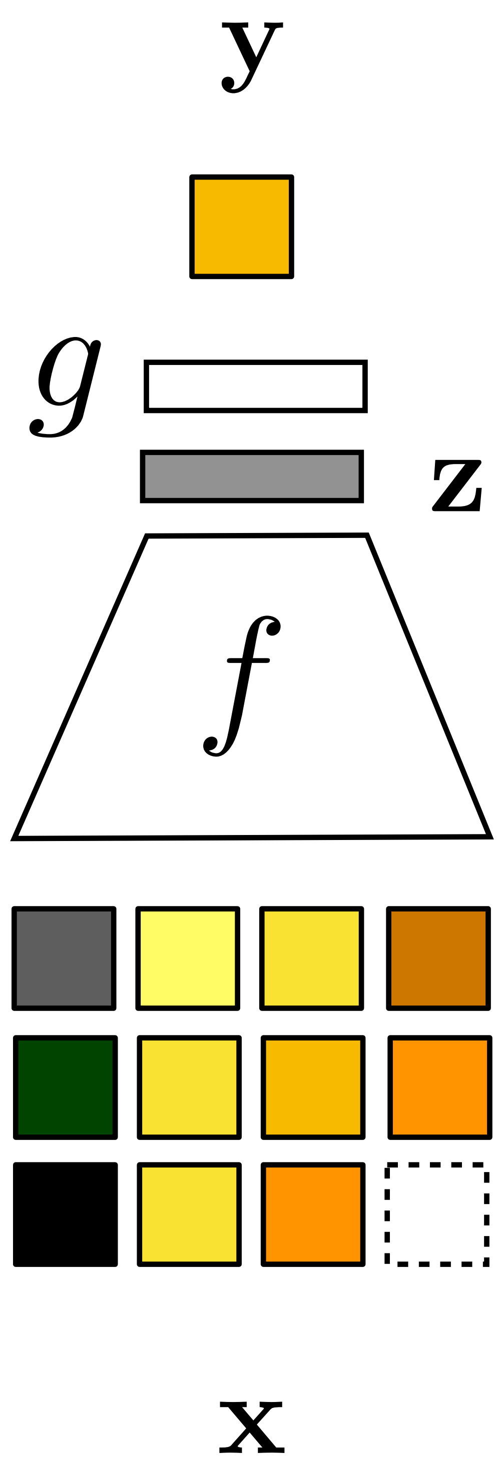
- Convert "unsupervised" problem into "supervised" empirical risk minimization

Self-supervised learning

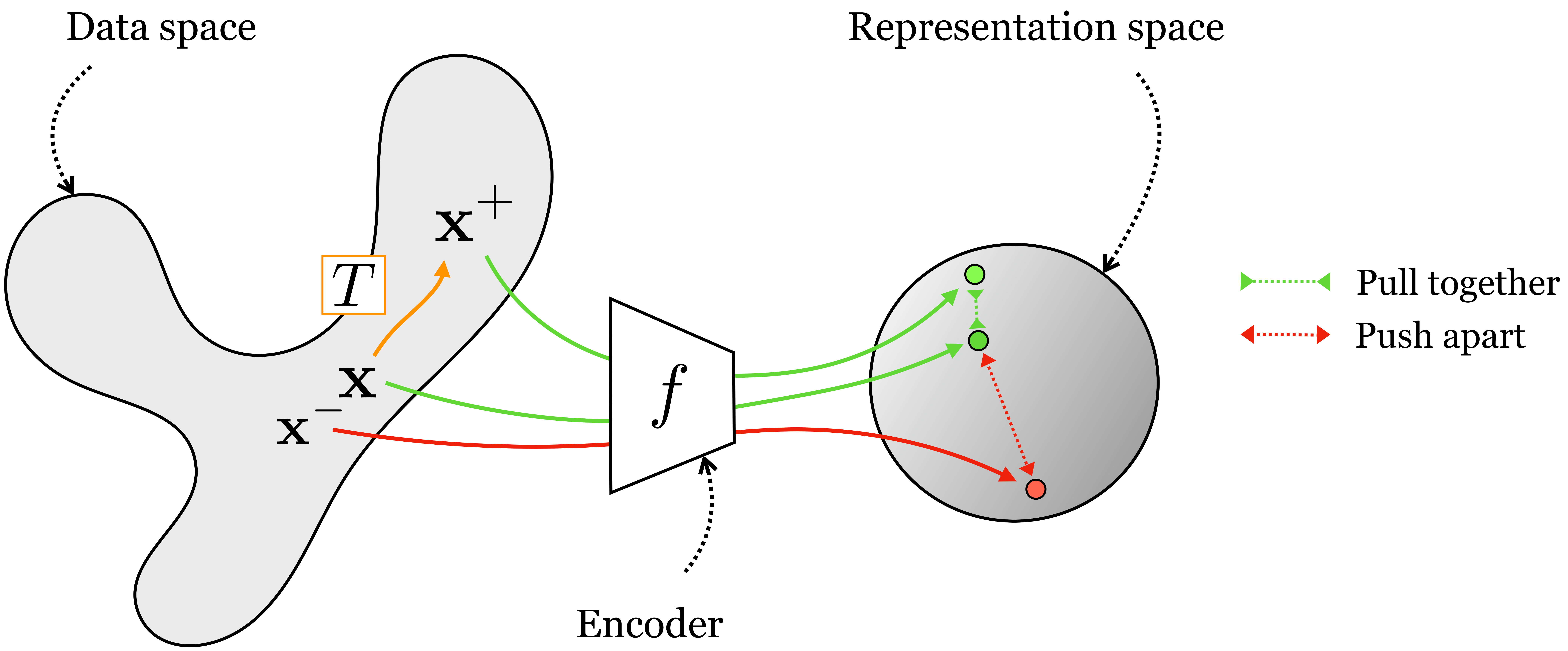


Common trick:

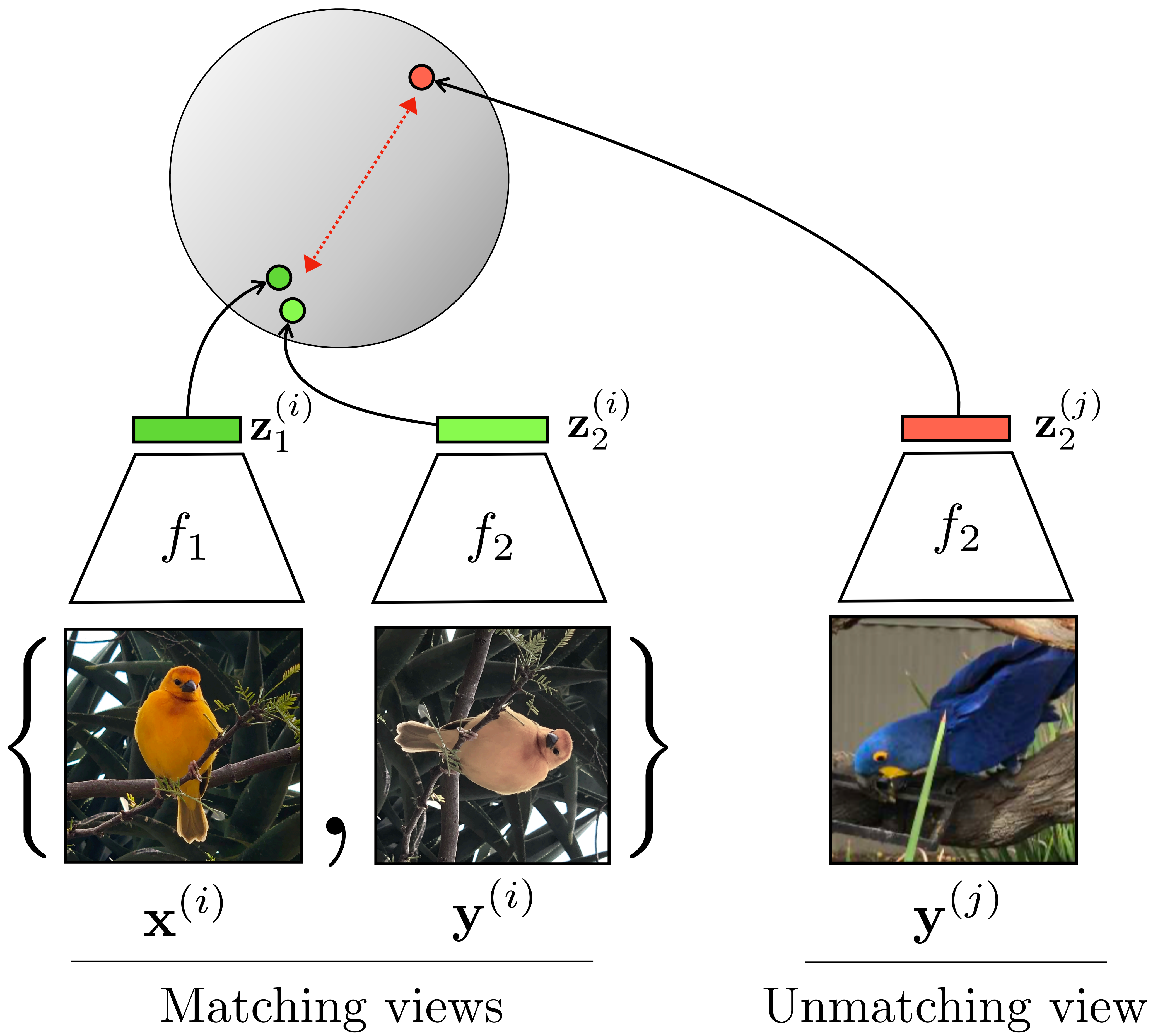
- Convert "unsupervised" problem into "supervised" empirical risk minimization
- Do so by cooking up "labels" (prediction targets) from the raw data itself — called **pretext task**

Pretext task:	Class prediction	Future frame prediction	Next pixel prediction
Model schematic:			

Contrastive Learning



Contrastive Learning — Transformations



Data

$\{\mathbf{x}^{(i)}\}_{i=1}^N, T \rightarrow$

Constrastive learning (transformations)

Objective

$\sum_{i,j} D(f(T(\mathbf{x}^{(i)})), f(\mathbf{x}^{(i)}))$

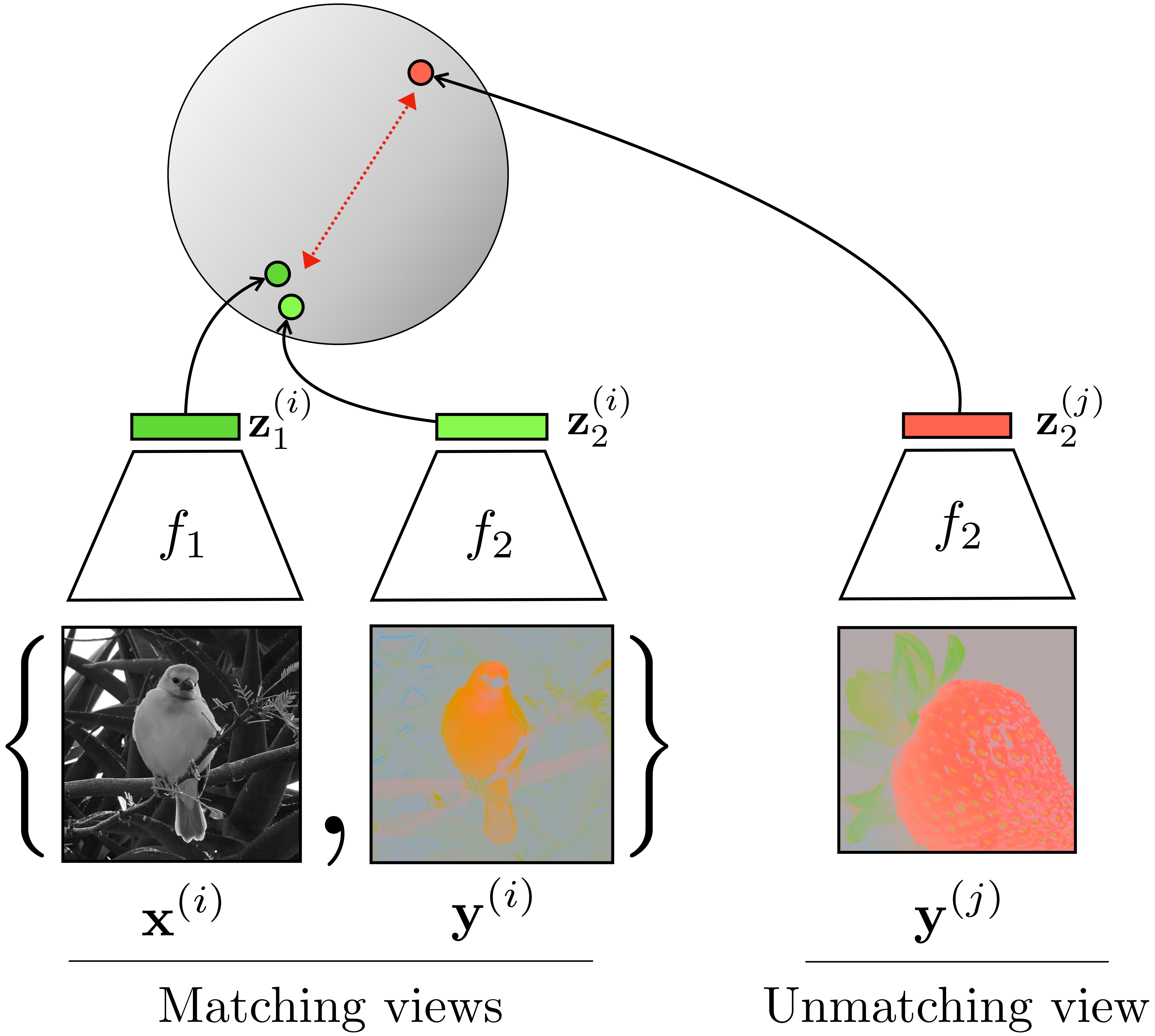
$-D(f(\mathbf{x}^{(i)}), f(\mathbf{x}^{(j)}))$

Hypothesis space

$f : \mathbb{R}^N \rightarrow \mathbb{R}^M$

$\rightarrow f$

Contrastive Learning — Co-occurrences



Data

$\{\mathbf{x}^{(i)}, \mathbf{y}^{(i)}\}_{i=1}^{N} \rightarrow$

Constrastive learning (co-occurrences)

Objective

$$\sum_{i,j} D(f_1(\mathbf{x}^{(i)}), f_2(\mathbf{y}^{(i)}))$$

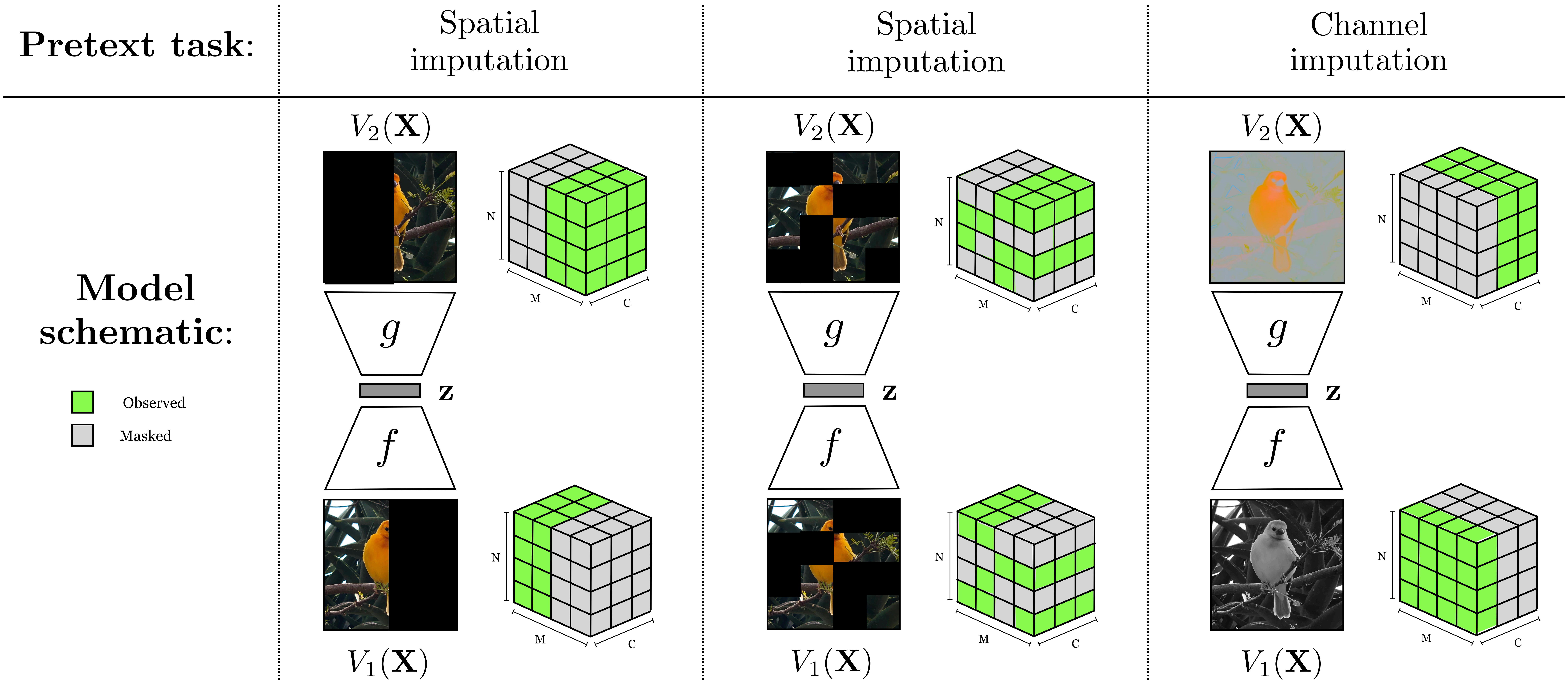
$$-D(f_1(\mathbf{x}^{(i)}), f_2(\mathbf{y}^{(j)}))$$

Hypothesis space

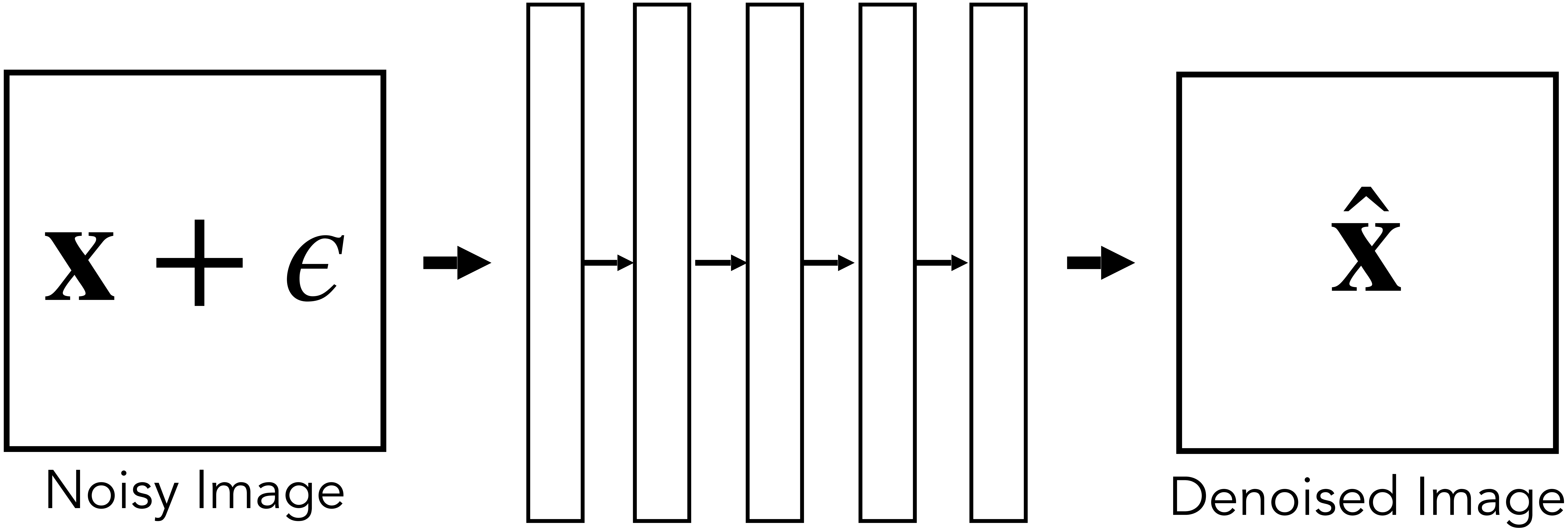
$$f_1, f_2 : \mathbb{R}^N \rightarrow \mathbb{R}^M$$

$\rightarrow f_1, f_2$

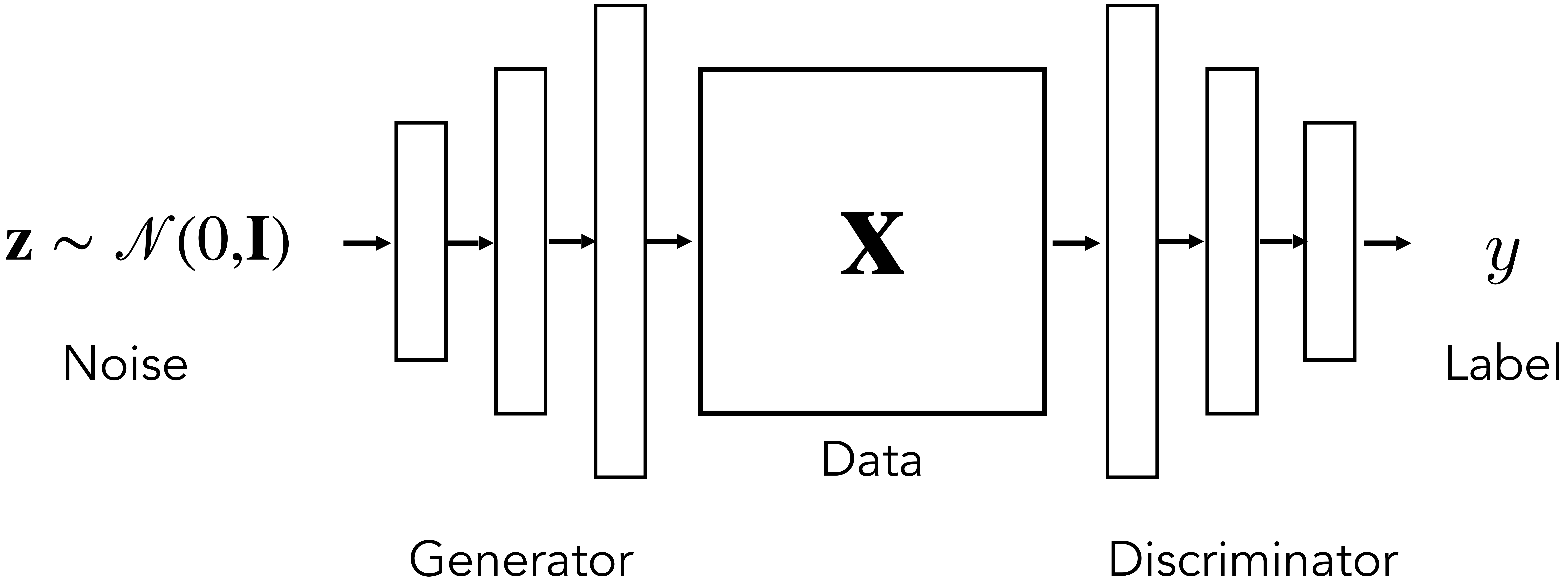
Imputation: one pretext task to rule them all?



Diffusion



GANs


$\mathbf{z} \sim \mathcal{N}(0,\mathbf{I})$
Noise
Generator
$\mathbf{X}$
Data
Discriminator
y
Label

VAEs

$\mathbf{x}$
Latent
$\mathbf{z} \sim \mathcal{N}(\mu(\mathbf{x}), \Sigma(\mathbf{x}))$
$\mathbf{x}$
Data
Encoder
Decoder

Generative Modeling - just another pretext task?

Source: https://en.wikipedia.org/wiki/Sora_%28text-to-video_model%29

Generative Modeling - just another pretext task?

Source: https://en.wikipedia.org/wiki/Sora_%28text-to-video_model%29

Generative Modeling - just another pretext task?

Video source: https://x.com/runwayml/status/1807822396415467686

Generative Modeling - just another pretext task?

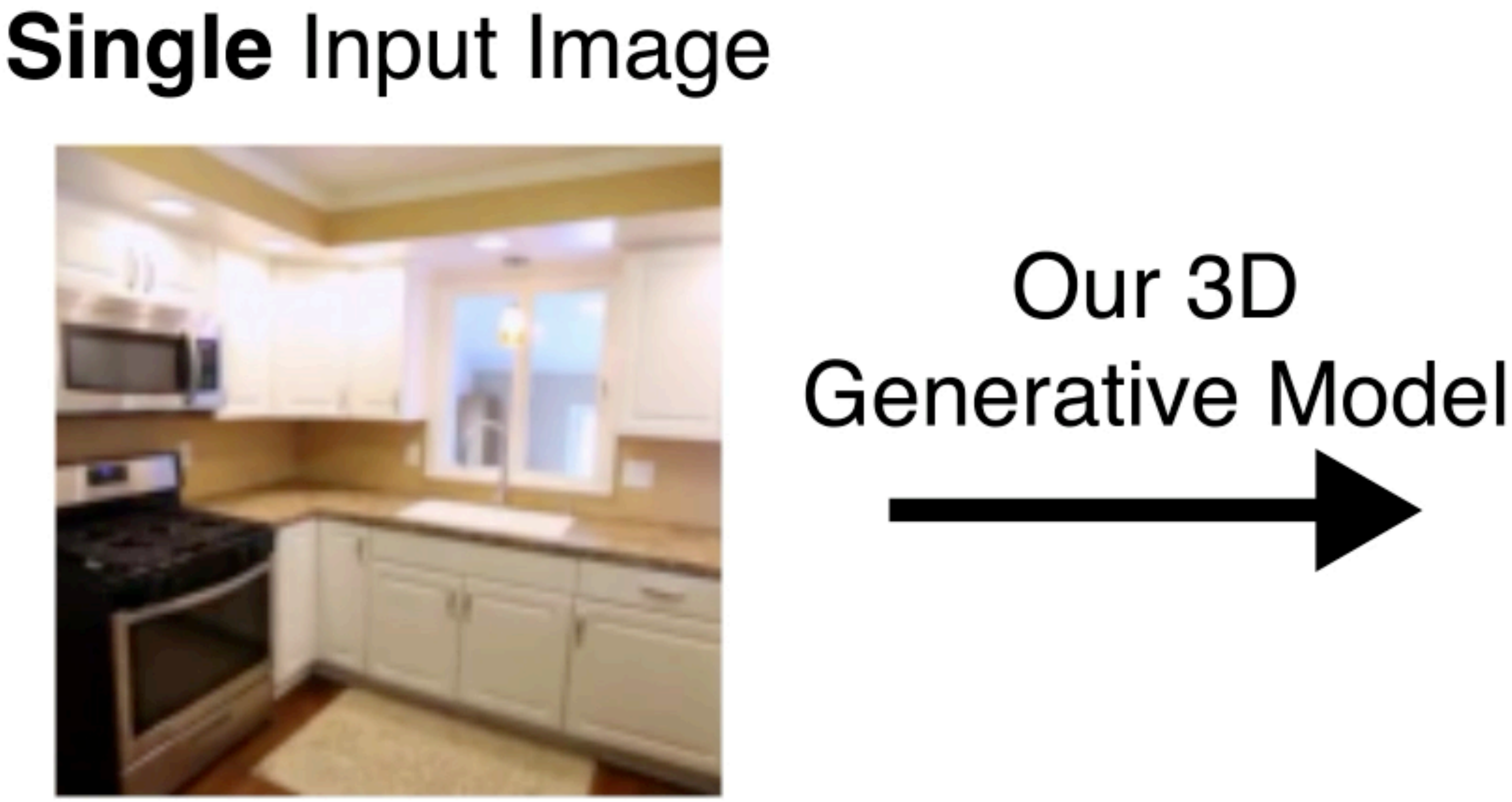
Video source: https://x.com/runwayml/status/1807822396415467686

Conditional Generative Modeling as Uncertainty-Aware Reconstruction

History-Guided Video Diffusion, Song and Chen et al. 2025

Conditional Generative Modeling as Uncertainty-Aware Reconstruction

History-Guided Video Diffusion, Song and Chen et al. 2025



Diffusion with Forward Models
Tewari et al. 2023

Single Input Image

Our 3D
Generative Model

Sampled 3D Scene 1

Sampled 3D Scene 2

Sampled 3D Scene 3

Sampled 3D Scene 4

Diffusion with Forward Models
Tewari et al. 2023

Relationship of Generative Modeling and Representation Learning

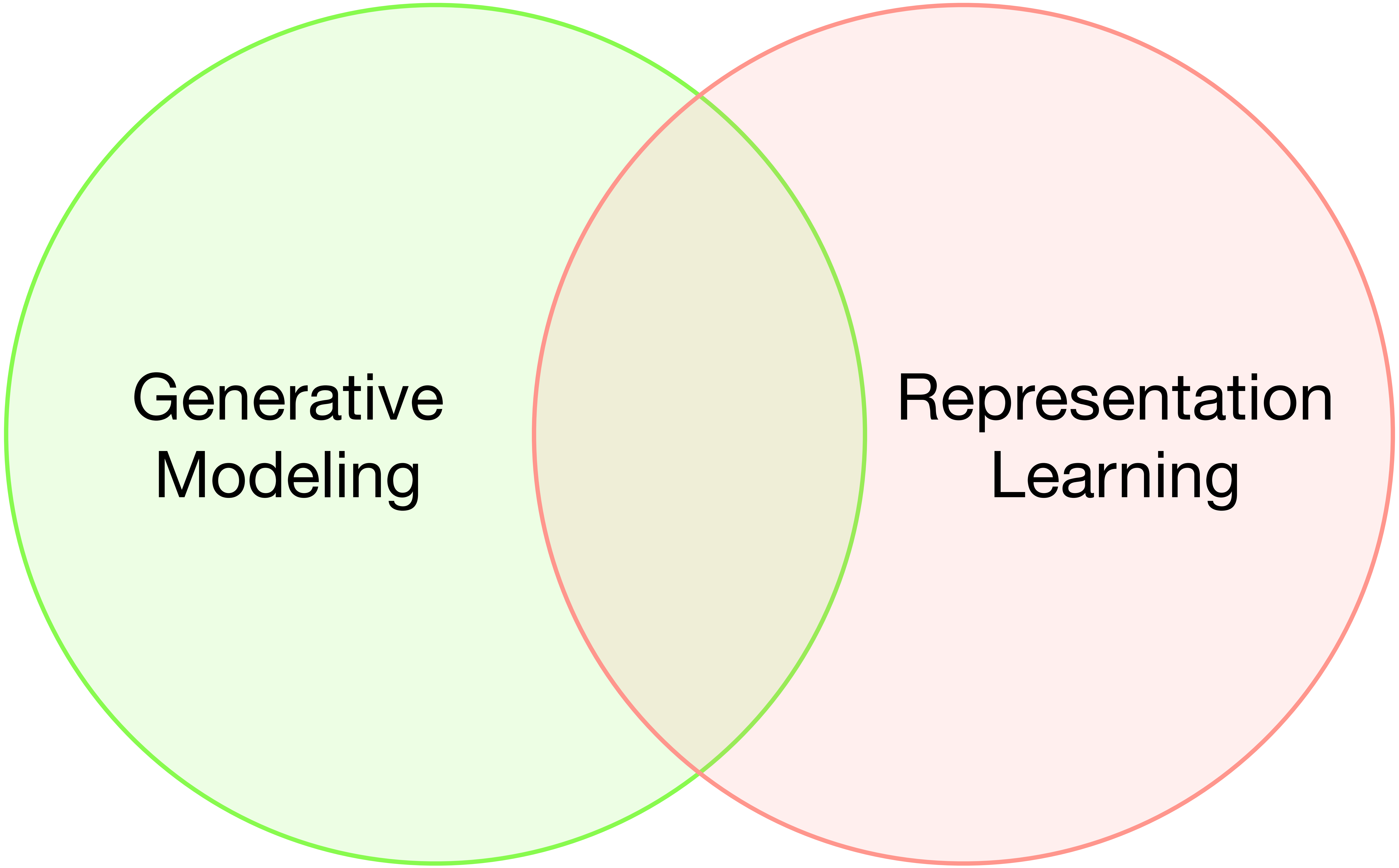
> *What I cannot create; I do not understand.*
> - Richard Feynman

Relationship of Generative Modeling and Representation Learning

> *What I cannot create; I do not understand.*
> - Richard Feynman

- It is tempting to believe that a generative model has to have a good representation if it can generate a perfect image / video / audio sample
- This is not generally true however: Consider a generative model overfit on a single image

Relationship of Generative Modeling and Representation Learning



Relationship of Generative Modeling and Representation Learning

Generative Modeling **can** be used for representation learning.

But it's not close to the best representation learning methods today (we will learn a bit about why that may be in this course).

However, lots of potential & lots of applications!

Overview of Module 2: Generative Modeling and Representation Learning

- Two lectures by **Peter Holderrieth (MIT PhD student)**
  - Mathematical Foundations of Diffusion Models
  - Guidance & case study of major diffusion model architectures for image & video generation
- Three lectures on generative modeling by Vincent:
  - Idiosyncrasies of diffusion for computer vision: spectral perspective, generalization
  - Diffusion for sequence generation
  - 3D generative models
- Three lectures on non-generative representation learning: Two “normal” lectures, one guest lecture by Dr. Mathilde Caron, author of DINO.

How to do research &
How to give talks

With slides from Vincent, Phil Isola, and Bill Freeman

~~How~~
Why to do research?

Why to do research?

- Follow your own interest.
- Be creative!
- Learn the scientific method - how to find truth through hypothesizing, predicting, experimentation.
- Discover something nobody knew before you.
- Safe space to “be bold”, i.e., tackle big problems.
- Fundamentally entrepreneurial: High “risk", high reward.

How to **recognize good** research?

Novelty

Good metrics: **surprisal** and **enabling new directions.**

What makes something surprising?

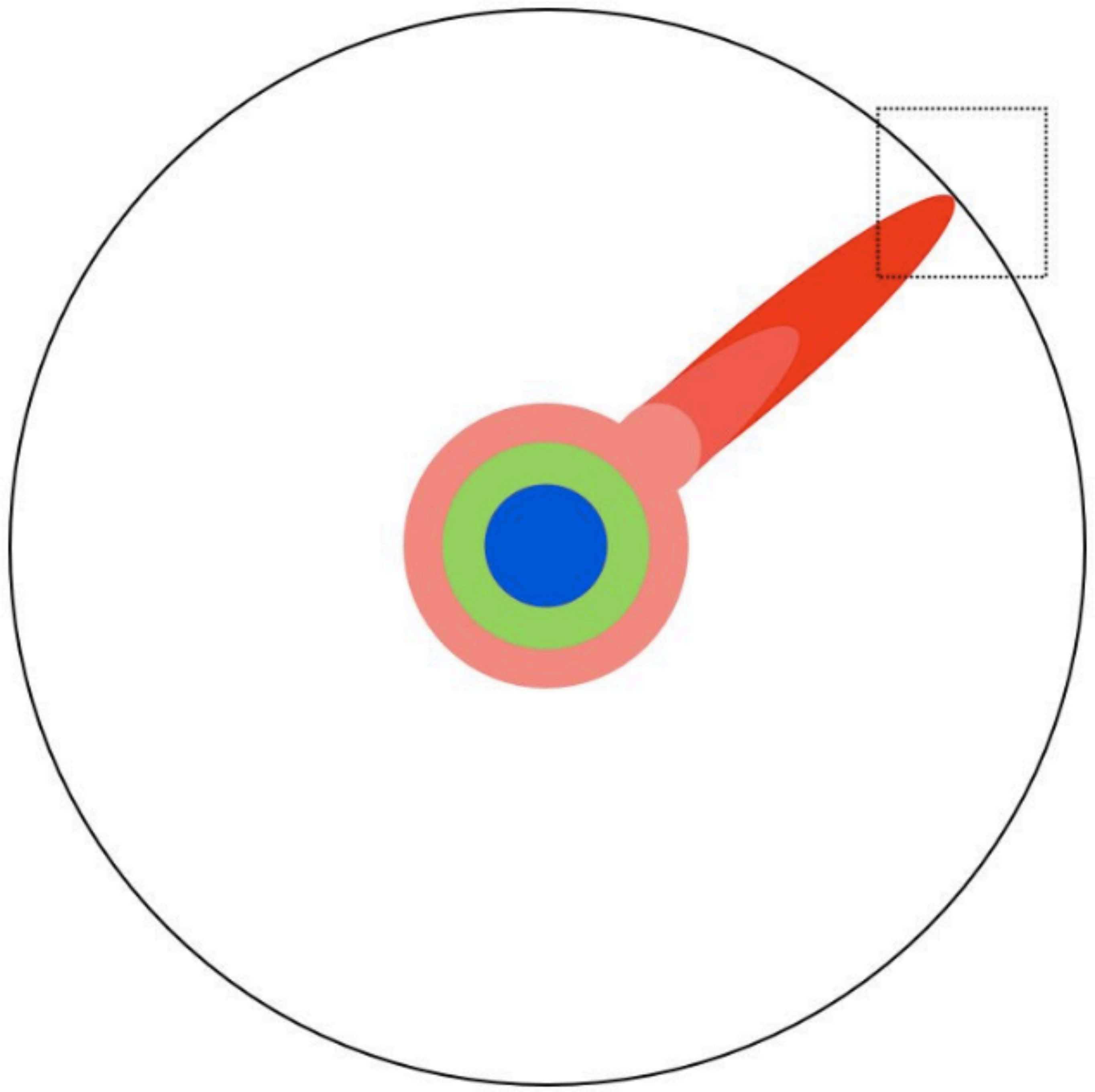
* People should *gain a lot of information from it*
  * It should be new
  * It should be understandable

Enabling new directions:

* Does this paper enable a new approach to an old problem?
* Does this paper establish a bridge between two problems, such that tools from one are now useful for the other?

Novelty

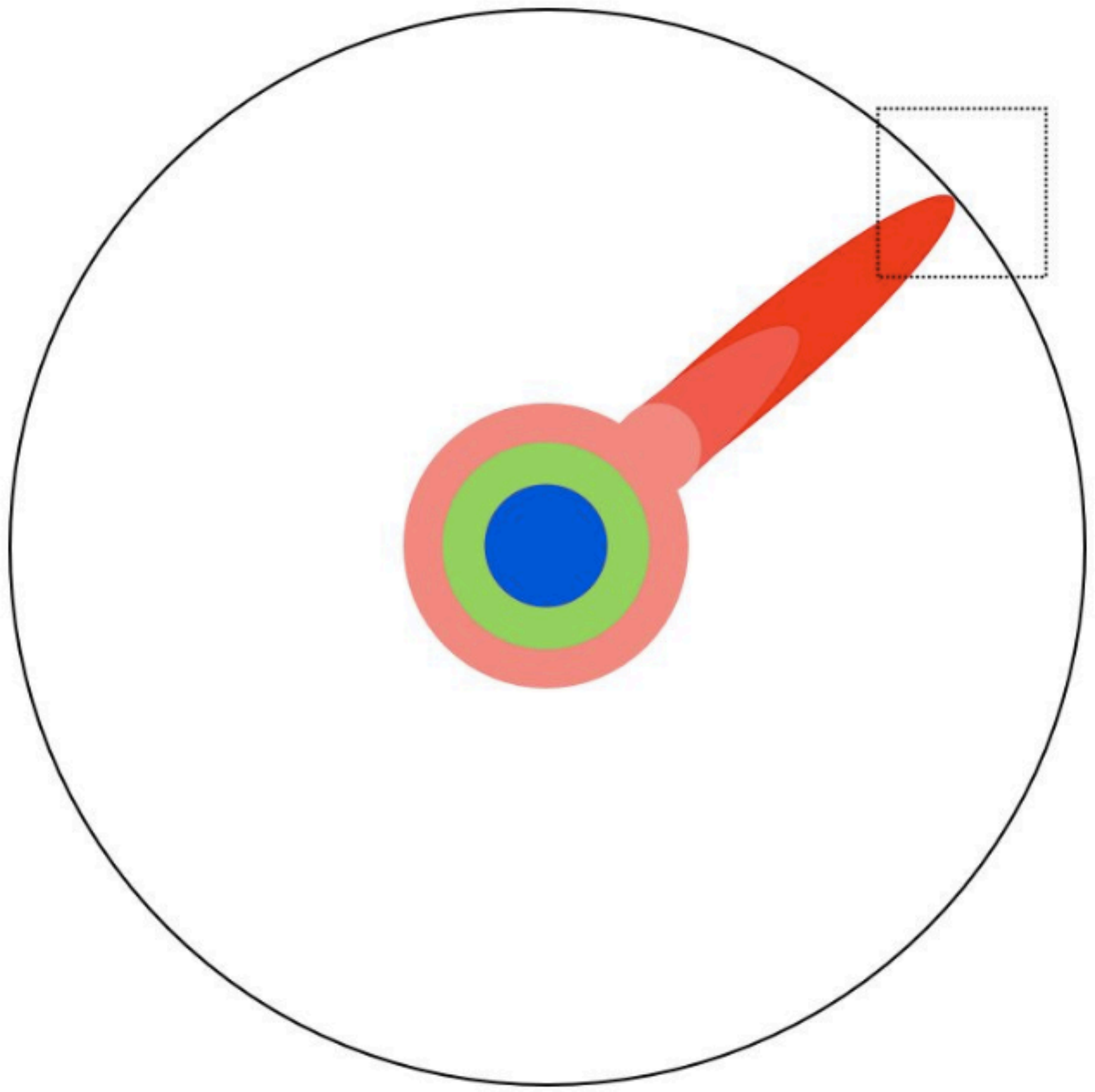
Very hard to achieve without knowing what has already been done.

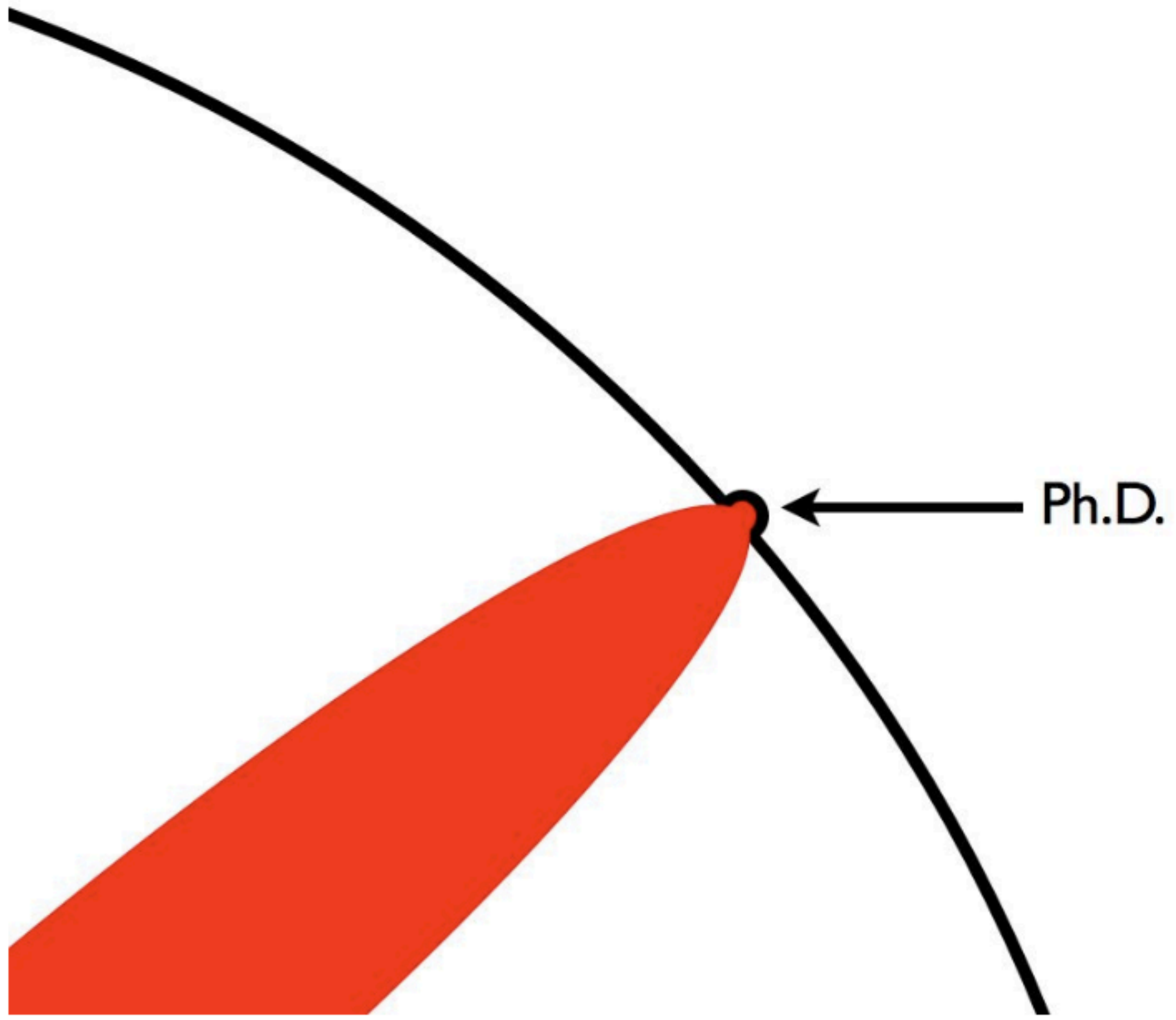


http://matt.might.net/articles/phd-school-in-pictures/

Novelty

Very hard to achieve without knowing what has already been done.





http://matt.might.net/articles/phd-school-in-pictures/

Novelty

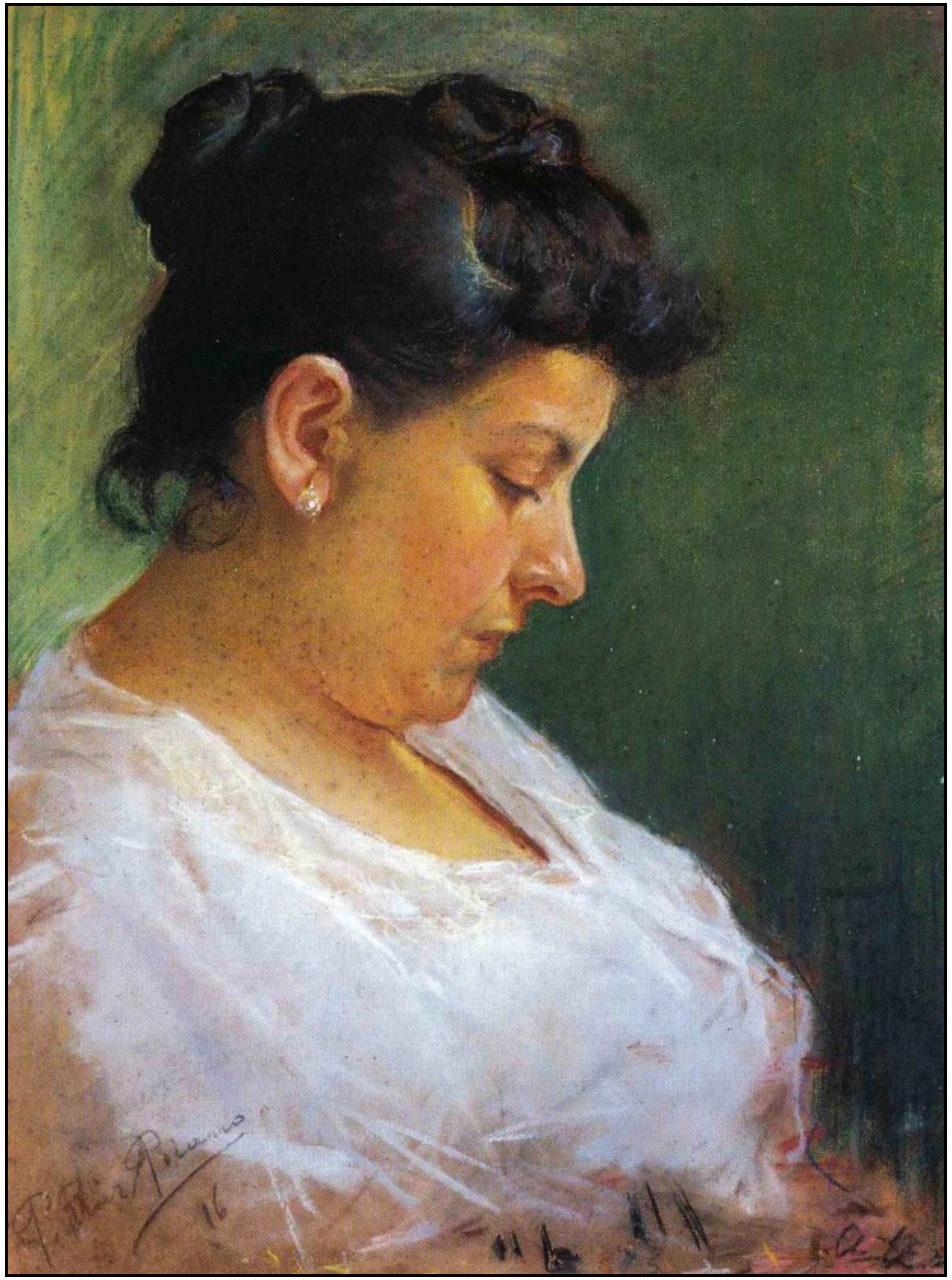
Very hard to achieve without knowing what has already been done.

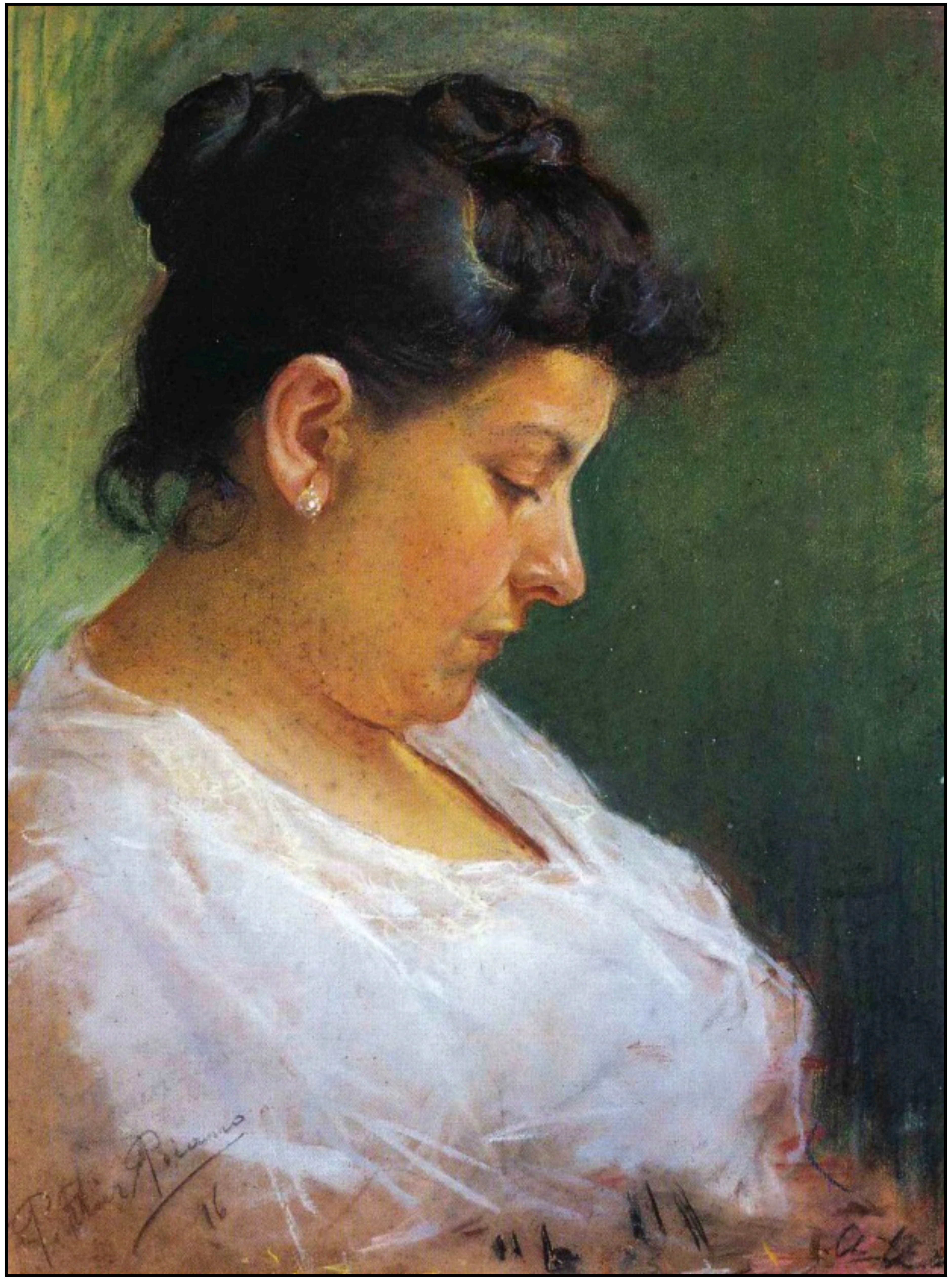
There is an exception to this: If you learn about a *problem* that has existed in a field for a long time / that is unsolved, it is **not** always useful to **know what others have tried before.**

http://matt.might.net/articles/phd-school-in-pictures/

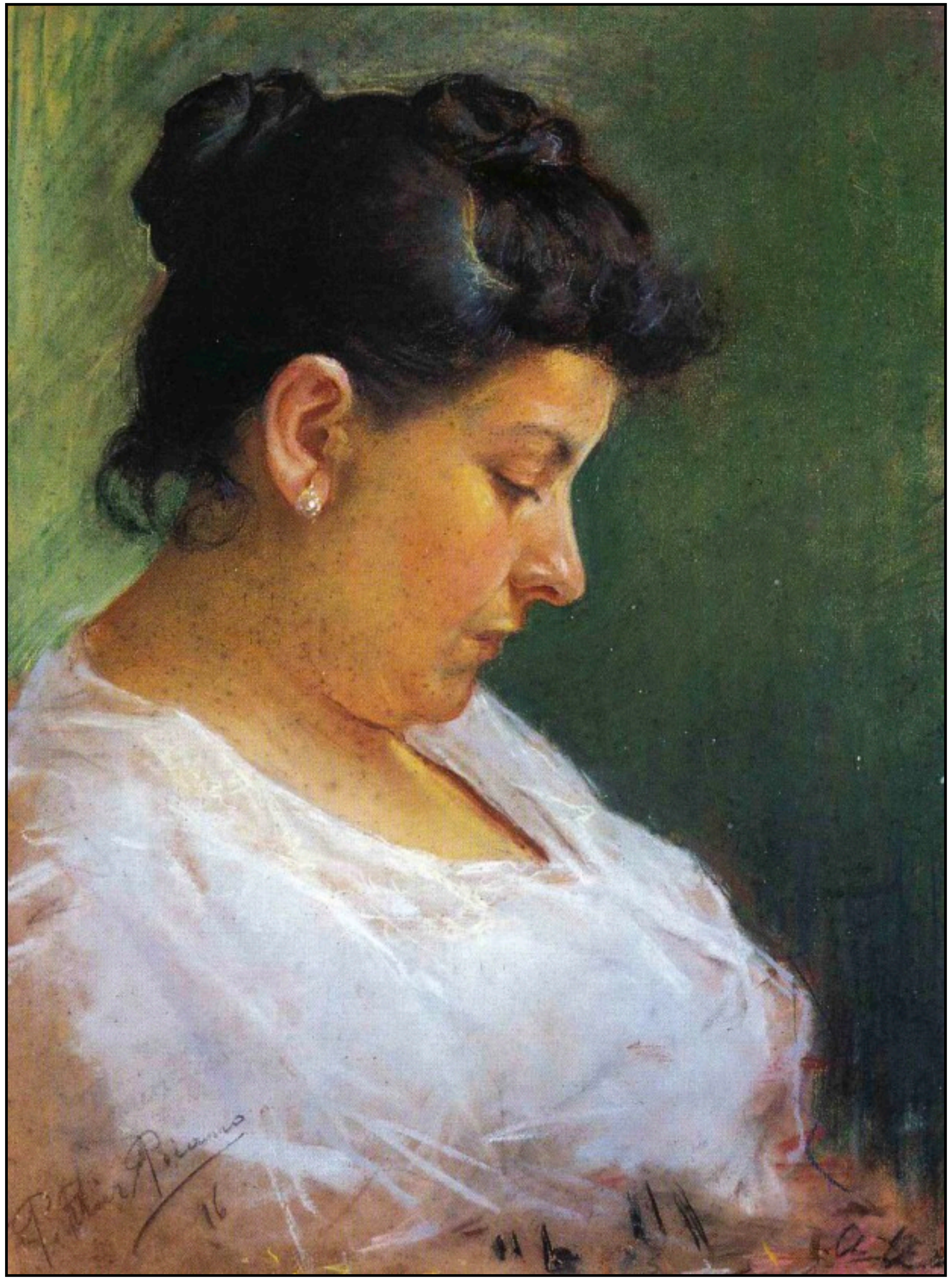
[Picasso]

[Picasso]





[Picasso]



[Picasso]

Blessing of the fool

> *In statistics, Dantzig solved two open problems in statistical theory, which he had mistaken for homework after arriving late to a lecture by Jerzy Spława-Neyman.*
>
> - Wiki on George Dantzig

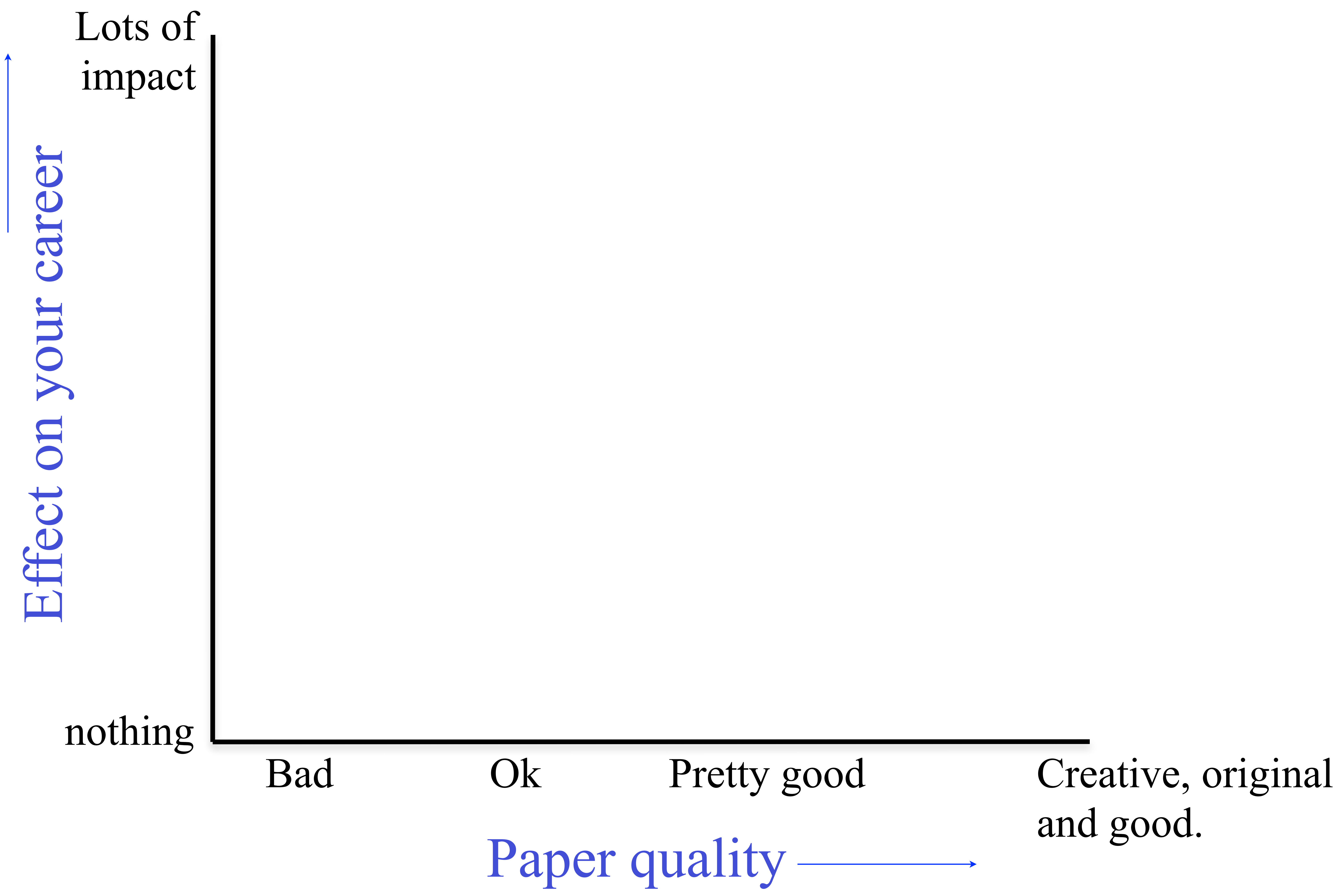
Blessing of the fool

> *In statistics, Dantzig solved two open problems in statistical theory, which he had mistaken for homework after arriving late to a lecture by Jerzy Spława-Neyman.*
>
> \- Wiki on George Dantzig

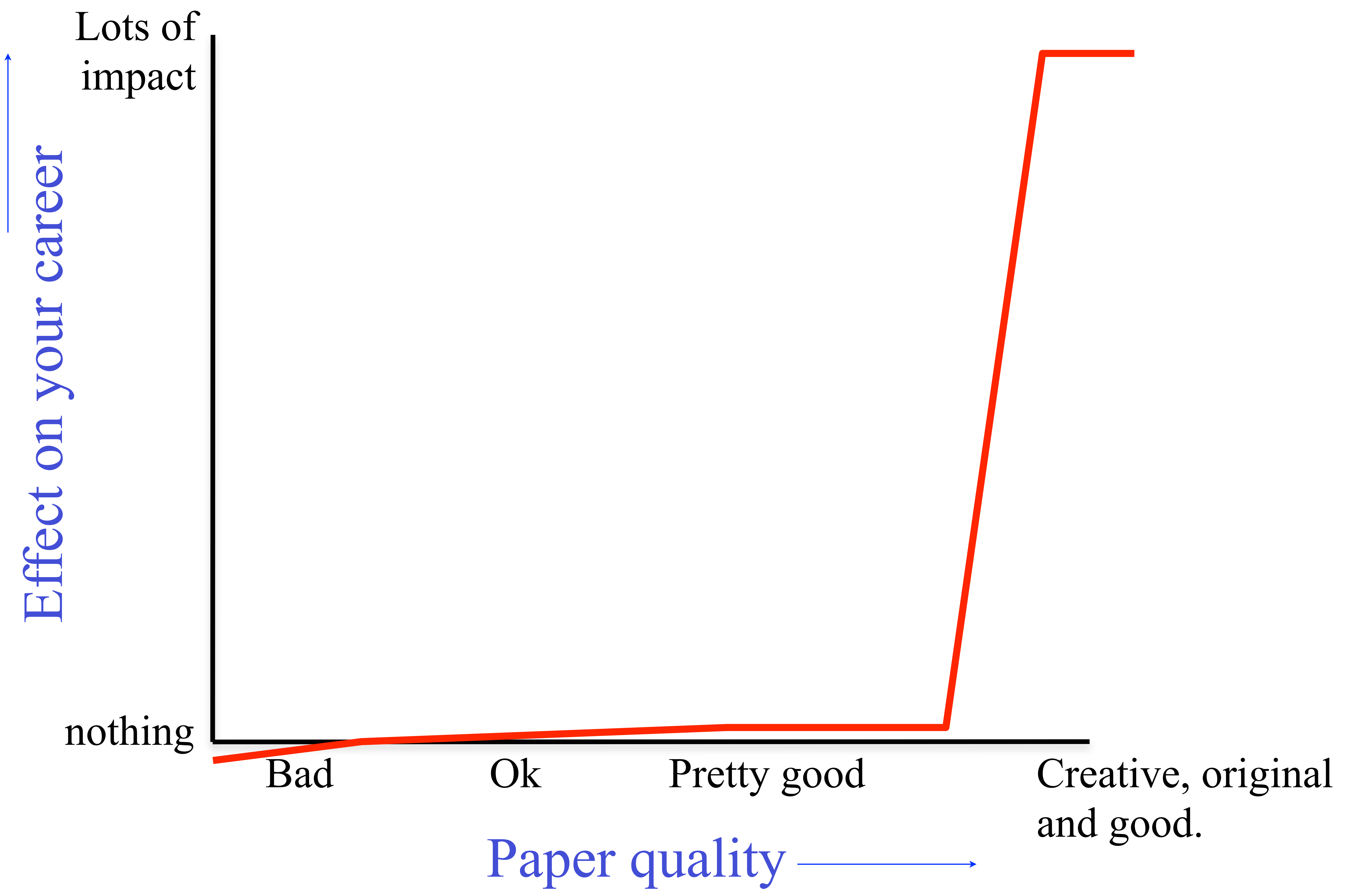
- If you learn about a problem that existed in a field for a long time but that (1) nobody knows how to solve or (2) people don't find interesting (but you disagree), then you're good!
- Knowing what others have tried before to solve it is like knowing how your peer approached a homework assignment…

How to **do good** research?

A paper's impact on your career

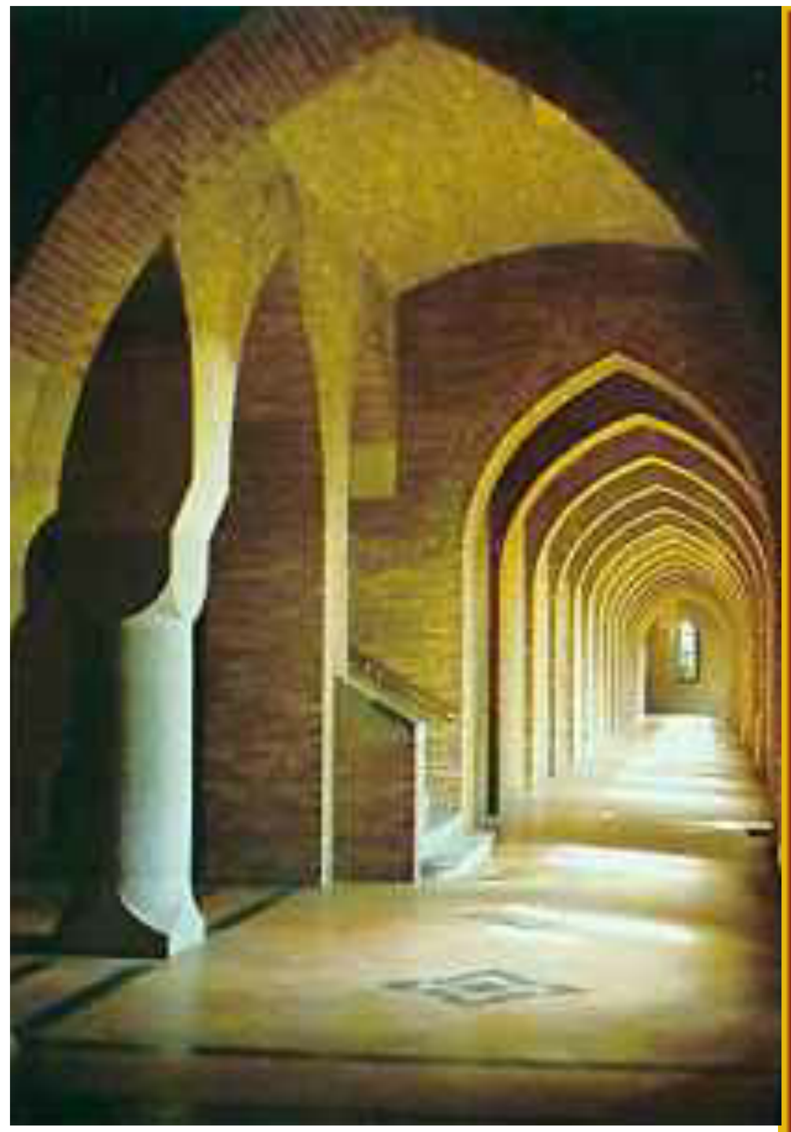


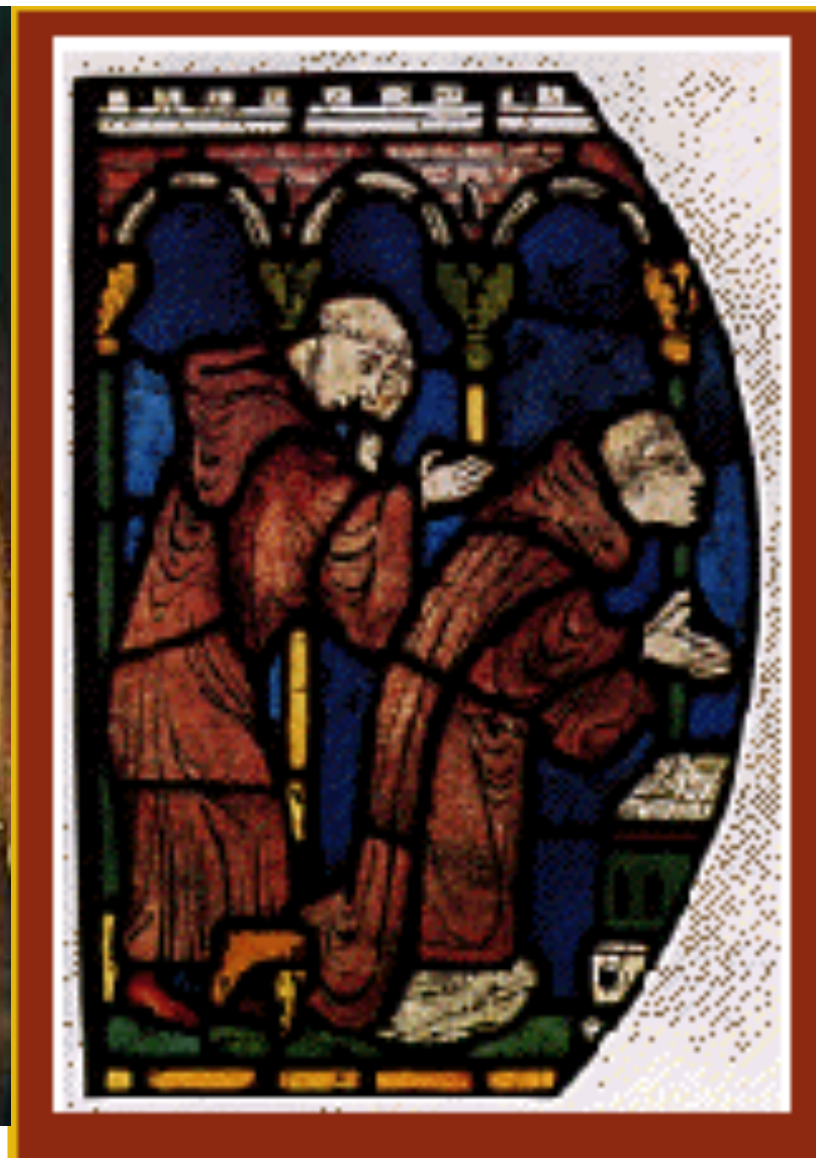
A paper’s impact on your career

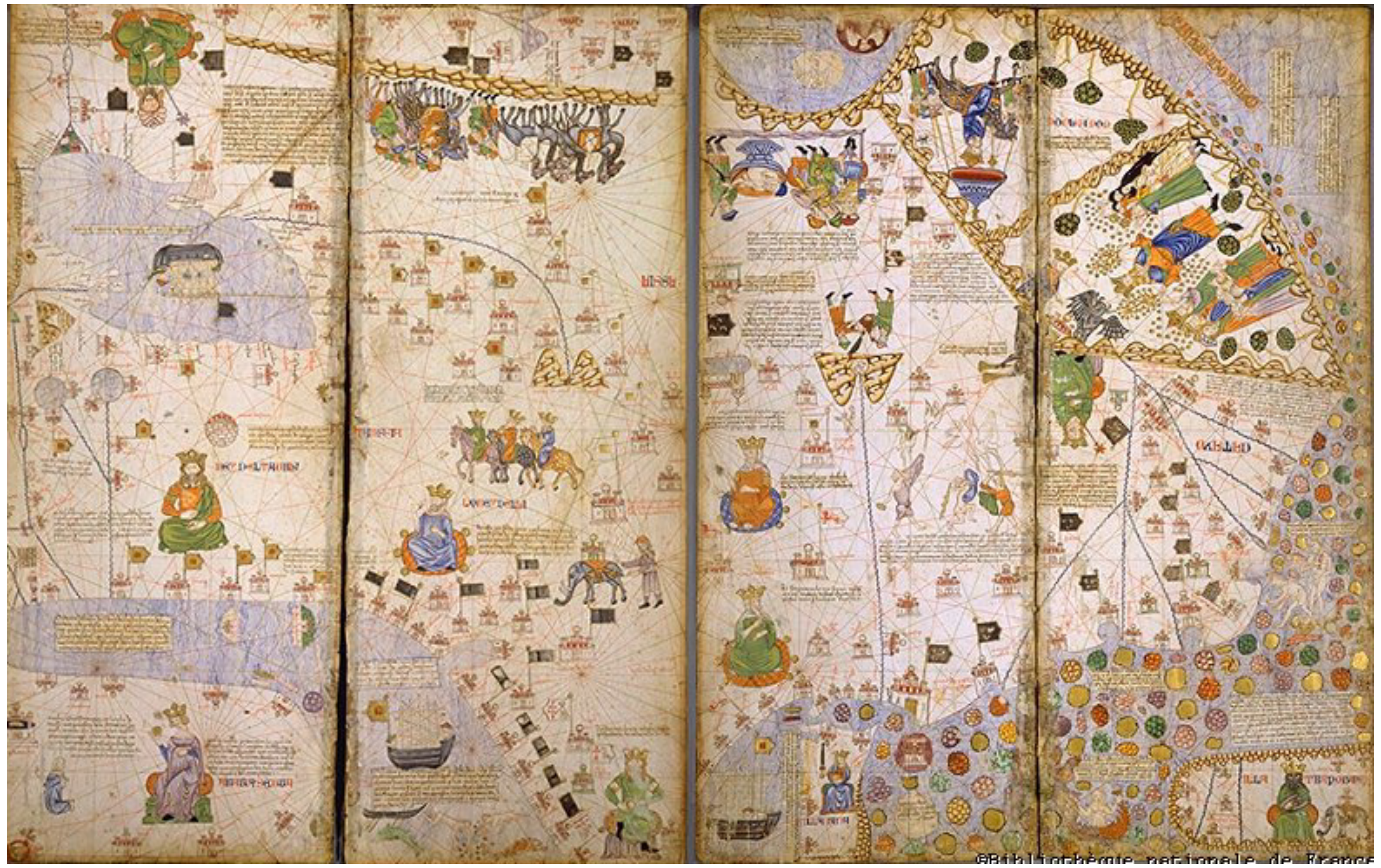


Our image of the research community

- Scholars, plenty of time on their hands, pouring over your manuscript.







The reality:
more like a large, crowded marketplace

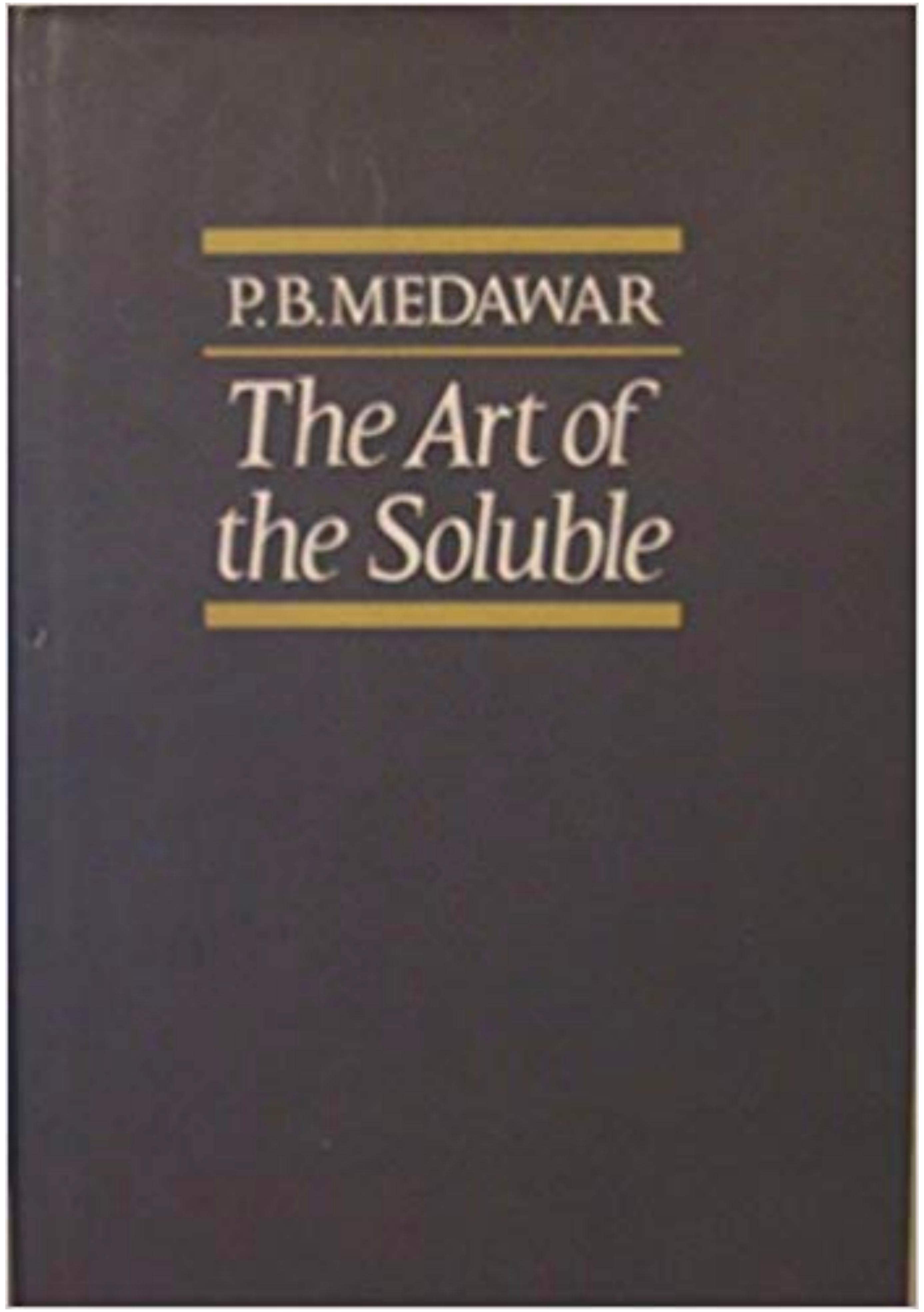
http://ducksflytogether.wordpress.com/2008/08/02/looking-back-khan-el-khalili/

Picking a topic

Picking a topic

> *Why Ghibli is still surviving until now? Because we have chosen opposite position to the trend.*
>
> *I have never thought to dominate the market or win the competition.*
>
> *It is easier if things going to just one way. Because we can go the opposite way.*
>
> *If you don't have this attitude, it would be impossible to catch up this moody trend and huge mass-consumption society.*
>
> - Hayao Miyazaki

Science is the "Art of the Soluble"



"'Good scientists study the most important problems they think they can solve. It is, after all, their professional business to solve problems not to grapple with them.' —Peter Medawar"
— Jitendra Malik

Know something no one else knows

Know something no one else knows

- It's necessary but not sufficient to master the core knowledge of your field

Know something no one else knows

- It’s necessary but not sufficient to master the core knowledge of your field
- Acquire a unique skillset

Know something no one else knows

- It's necessary but not sufficient to master the core knowledge of your field
- Acquire a unique skillset
  - Take difficult or unusual classes

Know something no one else knows

- It's necessary but not sufficient to master the core knowledge of your field

- Acquire a unique skillset
  - Take difficult or unusual classes
  - Read old papers

Know something no one else knows

- It's necessary but not sufficient to master the core knowledge of your field
- Acquire a unique skillset
  - Take difficult or unusual classes
  - Read old papers
  - Take on a complementary hobby

Know something no one else knows

- It's necessary but not sufficient to master the core knowledge of your field
- Acquire a unique skillset
  - Take difficult or unusual classes
  - Read old papers
  - Take on a complementary hobby
  - Talk to people in other fields

Know something no one else knows

- It's necessary but not sufficient to master the core knowledge of your field

- Acquire a unique skillset
  - Take difficult or unusual classes
  - Read old papers
  - Take on a complementary hobby
  - Talk to people in other fields

"My answer to "Now What" is "here is a research problem which is unusual, perhaps significant, novel, that I can pose and probably solve because of my background in physics". The situation would not be readily identified as a problem at all by those whose background seems much more relevant than my own."
— "Now What", John Hopfield

Question obvious weirdness

Question obvious weirdness

- Often, when you are new in a field, you are confronted with something that everybody accepts, but that strikes you as weird.

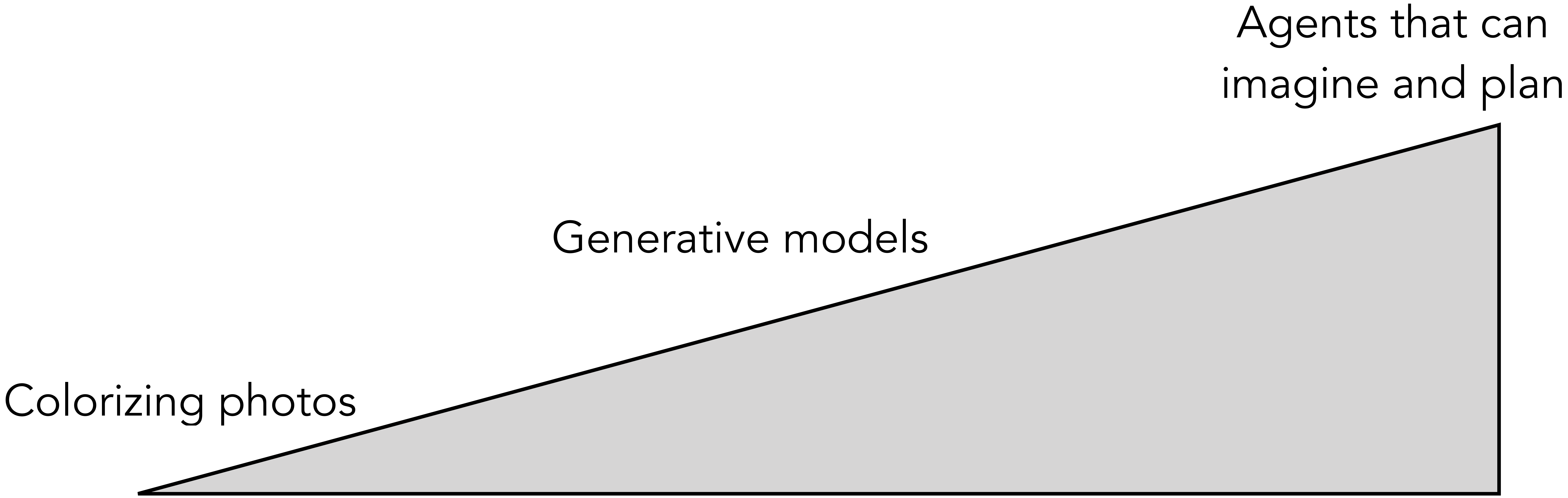
Question obvious weirdness

- Often, when you are new in a field, you are confronted with something that everybody accepts, but that strikes you as weird.
- These are often good research topics.

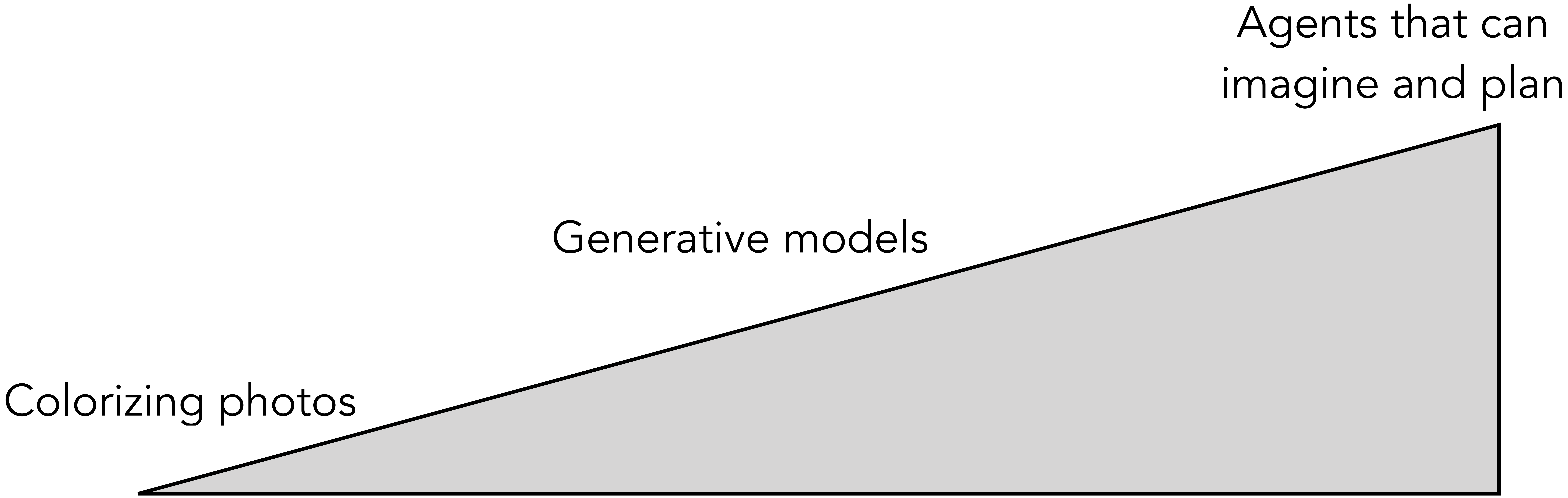
Question obvious weirdness

- Often, when you are new in a field, you are confronted with something that everybody accepts, but that strikes you as weird.
- These are often good research topics.
- Example: Structure-from-Motion…

Build a ramp


Agents that can
imagine and plan
Generative models
Colorizing photos

Build a ramp



"When you are famous it is hard to work on small problems. This is what did Shannon in. After information theory, what do you do for an encore? The great scientists often make this error. They fail to continue to plant the little acorns from which the mighty oak trees grow."
— Richard Hamming, "You and Your Research"

Think about the consequences

Why are you working on the problem you are working on?

What would happen if you were successful? Is that what you want to happen? Consider impact on science. Consider impact on society.

Get comfortable being confused

Michael Black @Michael_J_Black · 21h

Replying to @Michael_J_Black

There is an essential stage of confusion necessary for the formation of really new ideas. I call this the "high-temperature" state, where exploration happens. It can be uncomfortable for students and advisors but I encourage it.

1 2 24

Michael Black @Michael_J_Black · 21h

My role as an advisor is sometimes to raise the temperature, apparently increasing confusion by suggesting new directions. Then, at a critical point, I quench the system, dropping the temperature, and helping the student "close the deal."

1 1 25

The (abridged) Heilmeier Catechism

George H. Heilmeier, former DARPA director (1975-1977)

- Problem definition: What are you trying to do?
- How is it done today, and what are the limits of current practice?
- What is your approach and why is it better?
- Who cares? If you are successful, what difference will it make?
- What are the mid-term and final “exams” to check for success?

How to **communicate** your research?

(How to give talks)

High order bit: **prepare**

- Practice by yourself.
- Give practice versions to your friends.
- Think through your talk.
- You can write out verbatim what you want to say in the difficult parts.
- Ahead of time, visit where you'll be giving the talk and identify any issues that may come up.
- Preparation is a great cure for nervousness.

Giving good talks is **essential** to being a good scientist

- You might think: “the work itself is what really counts. Giving the talk is secondary”.
- But: you want to **inspire** people to help you work on the thing that **you believe** is the right way to go. By giving good talks, you can **change the course of science!**
- Talks are a great opportunity to offer another angle on work that helps people understand it.
- Researchers are like little startups: You have a brand and a product (your research). Talks is like doing sales: You want your product to help as many people as possible!

Your audience

Slide Credit: Bill Freeman

Your audience

- Your image of your audience:

Slide Credit: Bill Freeman

Your audience

- Your image of your audience:
  - Paying attention, listening to every word

Slide Credit: Bill Freeman

Your audience

- Your image of your audience:
  - Paying attention, listening to every word

Slide Credit: Bill Freeman

Your audience

- Your image of your audience:
  – Paying attention, listening to every word

- Your audience in reality:

Slide Credit: Bill Freeman

Your audience

- Your image of your audience:
  - Paying attention, listening to every word
- Your audience in reality:
  - Tired, hungry, not wanting to sit through yet another talk at the conference…

Slide Credit: Bill Freeman

Your audience

- Your image of your audience:
  - Paying attention, listening to every word
- Your audience in reality:
  - Tired, hungry, not wanting to sit through yet another talk at the conference…

Slide Credit: Bill Freeman

Your audience

- Your image of your audience:
  – Paying attention, listening to every word
- Your audience in reality:
  – Tired, hungry, not wanting to sit through yet another talk at the conference…

Your audience wants to be **entertained.**

Slide Credit: Bill Freeman

A tip to not be nervous that I found useful

- Get over it. They’re not there to see you, they’re there to hear the information. Just convey the information to them.
- Treat it like an opportunity for people to get to know you and what you’re excited about. Show them what you’re excited about, and they’ll like it :)

Slide Credit: Bill Freeman

A common mistake: Using your “talk voice”…

- Somehow, we are taught that giving talks is somehow a different thing than explaining something to someone. That leads to people completely changing their body language and tone of voice.
- **Don’t.** Giving a talk is essentially the same as explaining something to someone in-person, just with better preparation of the arguments and expositions :)
- Use **the same** body language and tone of voice you would use in-person!

Body Language

- If you can, walk around. Use the space. The podium is a trap!!
- Use your hands.
- Don't look at the ground or computer screen all the time. If you need to look at the text on the slide, walk to the projector and point it out to the audience as you read it out.

Add dynamics to the talk

http://www.nch.ie/dynamic/img/Mariss%20Jansons%20%20%20new.jpg

- A talk is a story: there can be different levels of excitement or tension in different parts of the talk. This makes it easier for the audience to pay attention to what you're saying. Perhaps move to another location.
- I like to find some part of the work that really grabs me, that I'm really excited about, and let that show through. (The audience loves to see you be excited. Not all the time, but when appropriate).
- People love hearing about when something confuses you - it makes you relatable. Paradoxically, stating that you *were* confused about something is *not* perceived as a "weakness" - it's perceived as confident!

Figure out how one part follows from another

Ahead of time, think through how each part motivates the next, and point that out during the talk. If one part doesn't motivate the next, consider re-ordering the talk until it has that feel.

Slide Credit: Bill Freeman


OBEY


OBEY

What the audience of a technical talk wants

To have everything follow and make sense

To learn something

To connect with the speaker, to share their excitement.

They want to watch you love something!

Alan Alda: https://www.youtube.com/watch?v=j4XgjkXDxss, and others

Slide Credit: Bill Freeman

Let the audience see your personality.

- They want to see you enjoy yourself.
- They want to see what you love about the science.
- People really respond to the human parts of a talk. Those parts help the audience with their difficult task of listening to an hour-long talk on a technical subject. What was easy, what was fun, what was hard about the work?
- **Don't be afraid to be yourself and to be quirky.**

http://is3.okcupid.com/users/112/250/11225140098321842389/mt1112532356.jpg

Slide Credit: Bill Freeman