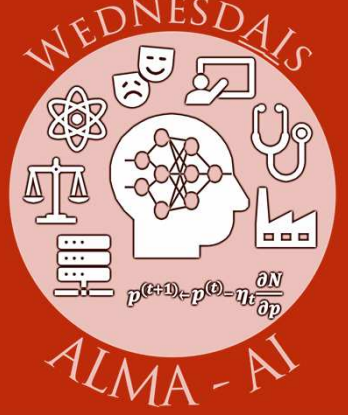


Unsupervised learning: AIs don't need no education?

Samuele Salti

Dipartimento di Informatica – Scienza e Ingegneria (DISI)

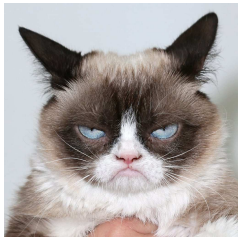
Let’s leave the AIs alone…

Supervised dataset (i.e. labeled)
Example: cats versus dogs

Cat

Dog

Cat



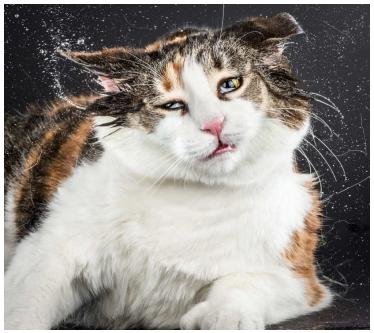
Cat

Dog

Unsupervised dataset (i.e. no labels)

?

?



?

?

?

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Traditional use cases of unsupervised learning

Clustering

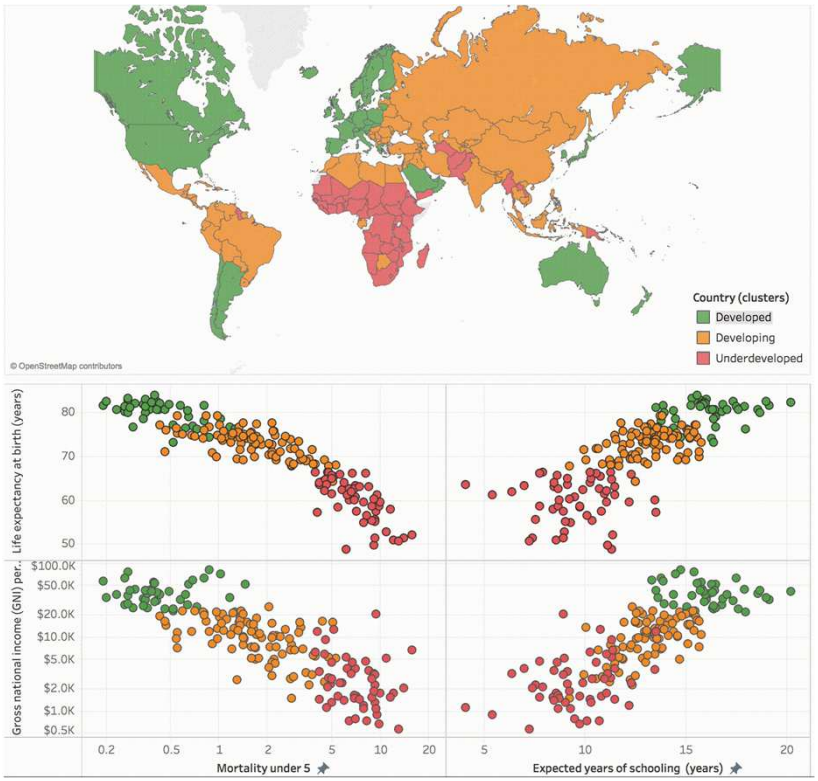


Image from:https://www.tableau.com/about/blog/2016/7/uncover-patterns-your-data-tableau-10s-clustering-feature-56373

Dimensionality reduction

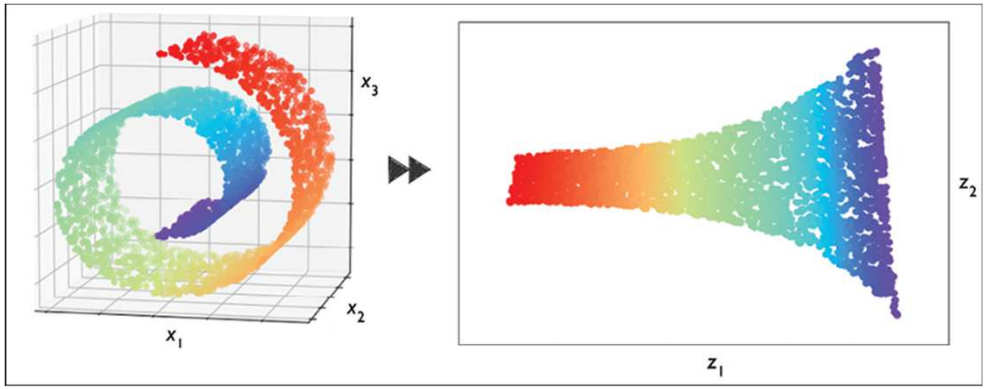


Image from:https://subscription.packtpub.com/book/data/9781789955750/1/ch01lvl1sec03/the-three-different-types-of-machine-learning

- Association (market/basket analysis)
- Probability density estimation
- …

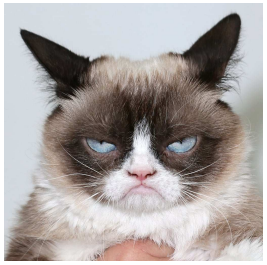
ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Can unlabeled data help to solve better/faster the supervised problems?

Manifold assumption: real data is high dimensional, but lie close to a low–dimensional manifold

Example: images

https://onlineimagetools.com/generate-random-image



Example: text

http://randomtextgenerator.com/

Style too own civil out along. Perfectly offending attempted add arranging age gentleman concluded. Get who uncommonly our expression ten increasing considered occasional travelling.

Up maids me an ample stood given. Certainty say suffering his him collected intention promotion. Hill sold ham men made lose case.

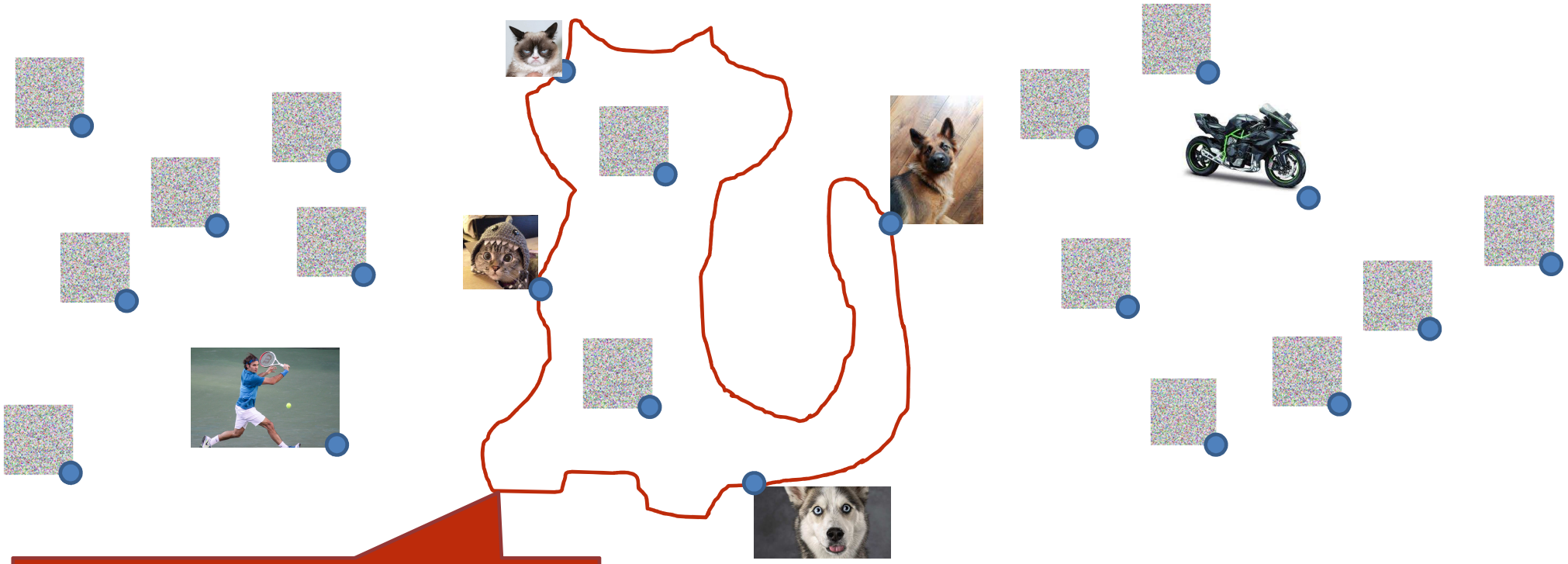
In the common parlance, randomness is the apparent lack of pattern or predictability in events.[1][2] A random sequence of events, symbols or steps often has no order and does not follow an intelligible pattern or combination.

Natural-language generation (NLG) is a software process that transforms structured data into natural language. It can be used to produce long form…

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Yes, they ~~cat~~ can

Effective low-dimensional representations (i.e. the manifold) can be learned from unsupervised data

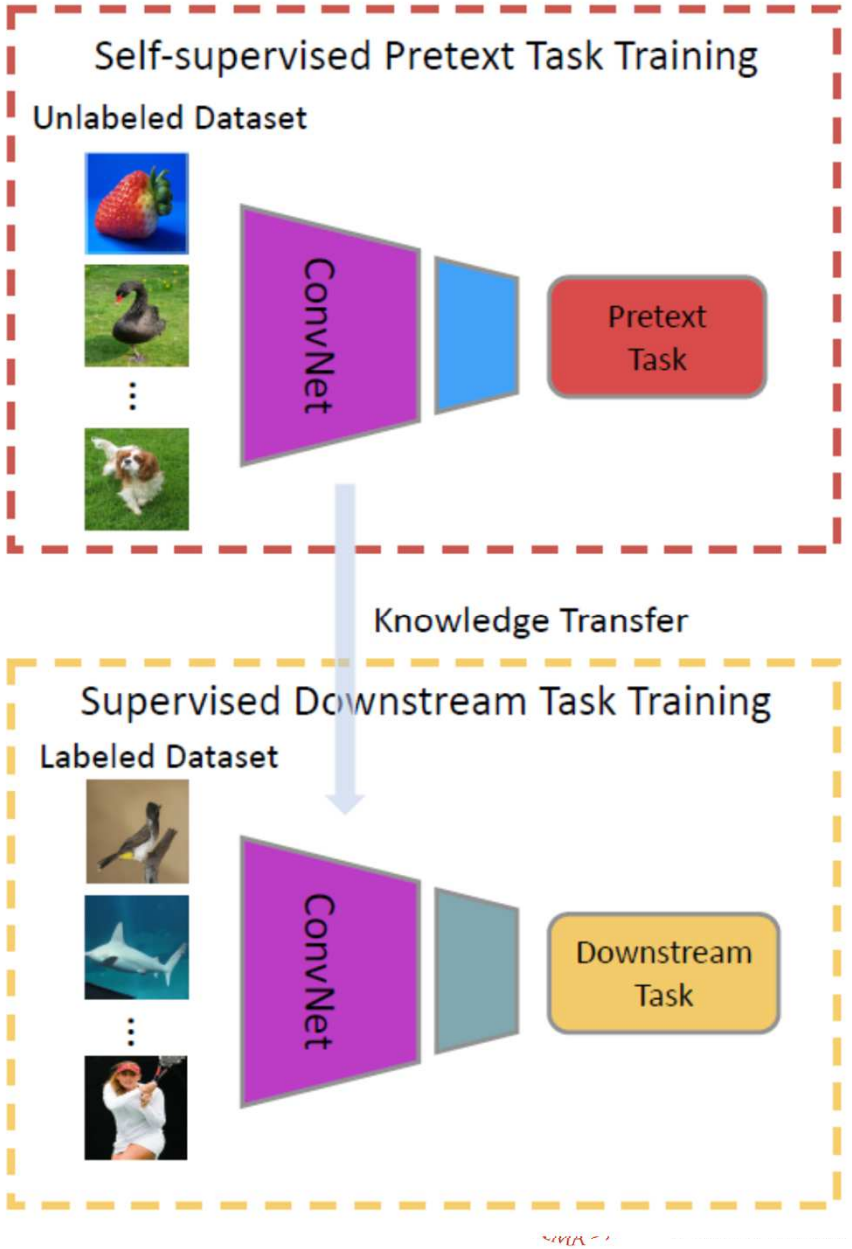


Unsupervised data helps to identify/learn the manifold

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Self-supervised learning

- When used to learn effective representations to bootstrap/improve supervised learning, unsupervised learning is (often) referred to as **self-supervised learning.**
- "Self-supervised learning is a subset of unsupervised learning methods [...] in which [neural networks] **are explicitly trained with automatically generated labels (pseudo-labels)**."



Longlong Jing and Yingli Tian, "Self-supervised Visual Feature Learning with Deep Neural Networks: A Survey", https://arxiv.org/abs/1902.06162

Let’s go back in the classroom…

Semi-supervised dataset

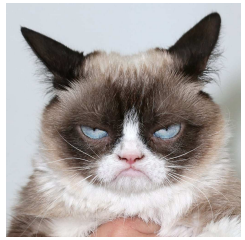
Cat

Cat

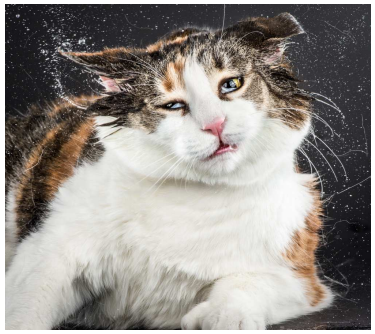
Dog

Dog

?



Cat



?

?

?

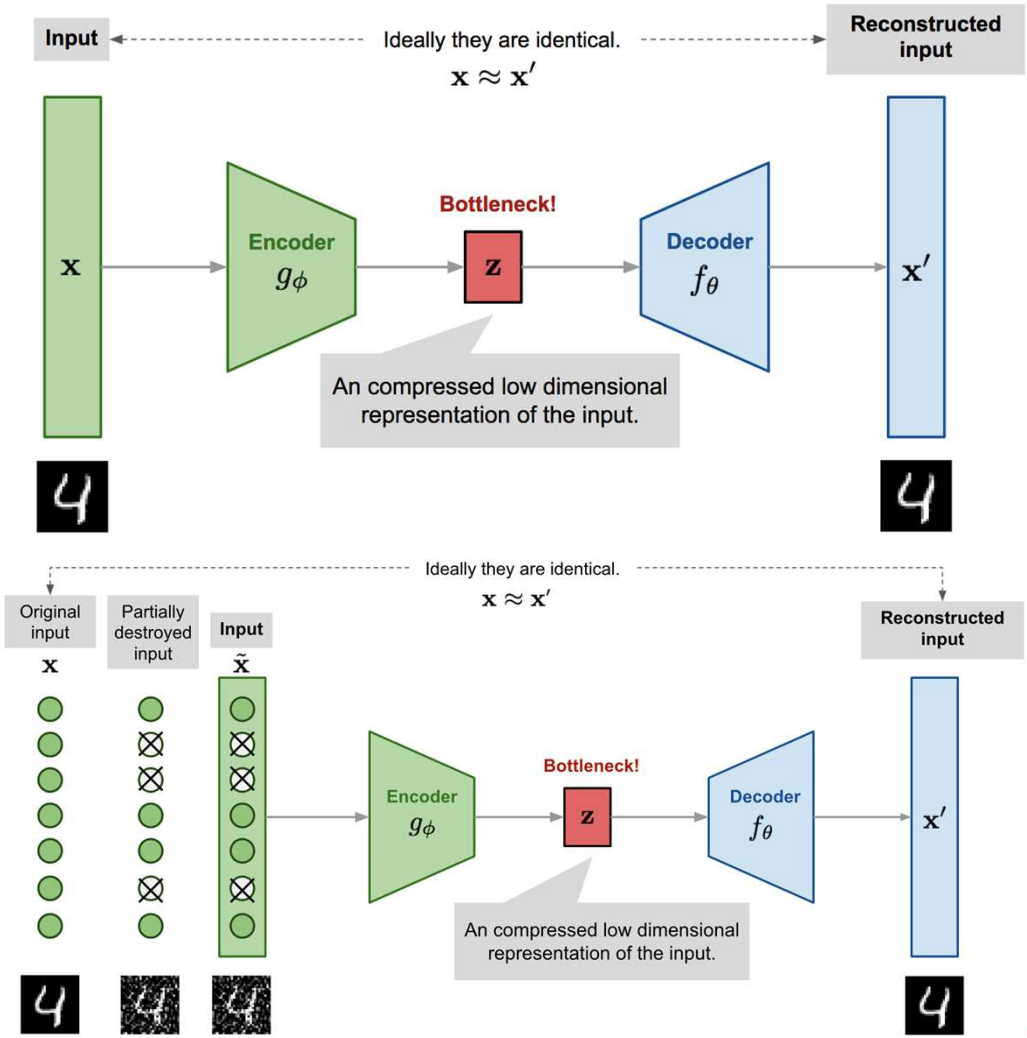
?

Auto-encoders

The simplest pretext task is the **autoencoder**: reconstruct the input data.

We usually categorize them according to how the task is made difficult

- **Autoencoder :** an information bottleneck whose dimensionality is smaller than the input data (auto-encoder)
- **Denoising auto-econder**: the bottleneck is smaller, and the input is distorted
- **Sparse auto-encoder**: the bottleneck is larger than the input, but we require sparse activations

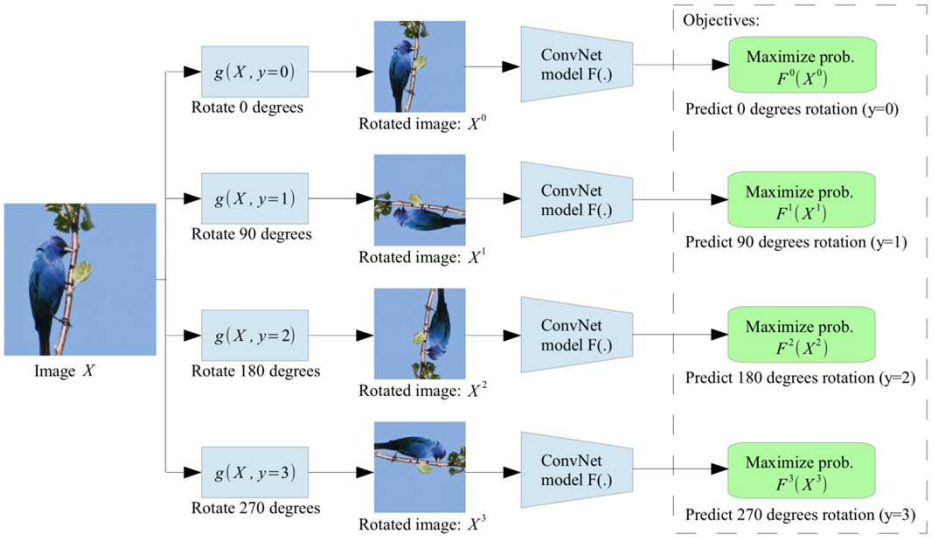


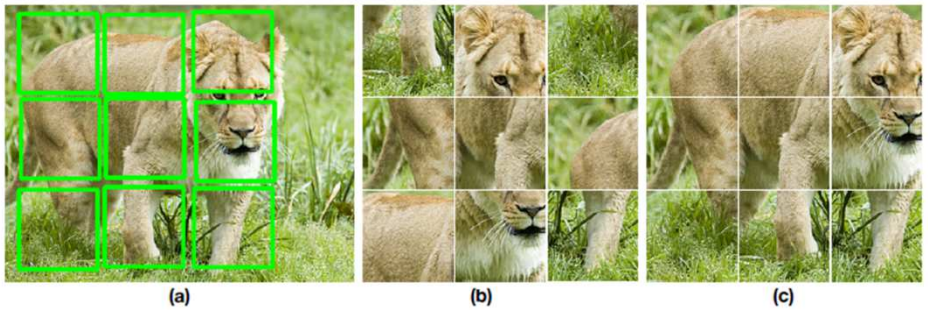
Images from:https://lilianweng.github.io/lil-log/2018/08/12/from-autoencoder-to-beta-vae.html#denoising-autoencoder

ALMA - AI

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Self-supervision from (spatial) context





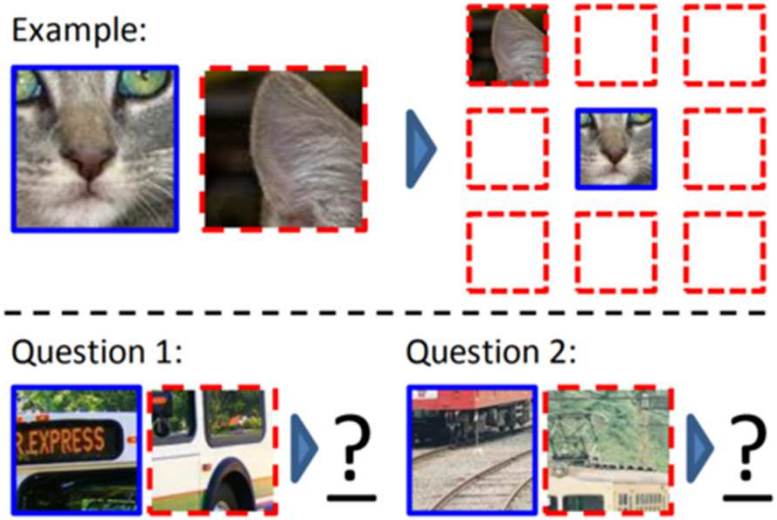


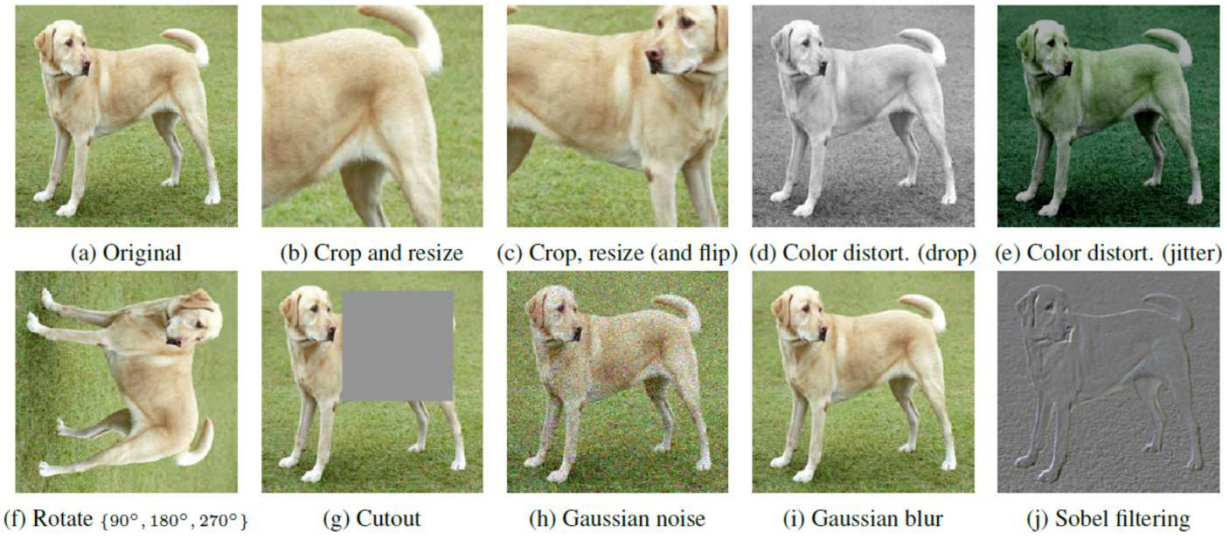
Figure 1. Our task for learning patch representations involves randomly sampling a patch (blue) and then one of eight possible neighbors (red). Can you guess the spatial configuration for the two pairs of patches? Note that the task is much easier once you have recognized the object!

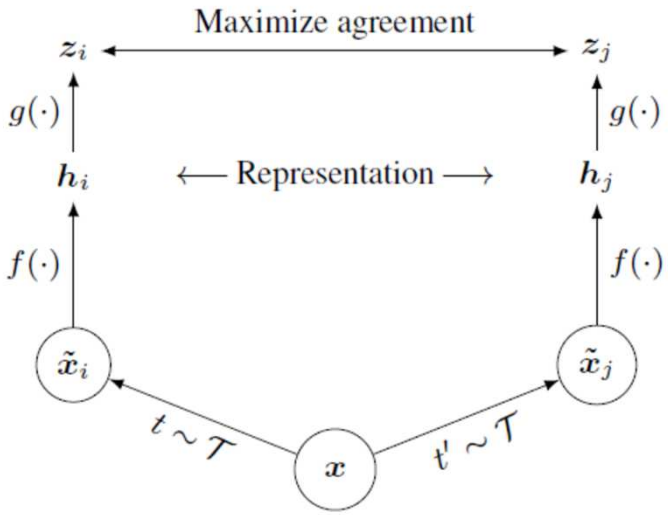
Answer key: Q1: Bottom right Q2: Top center

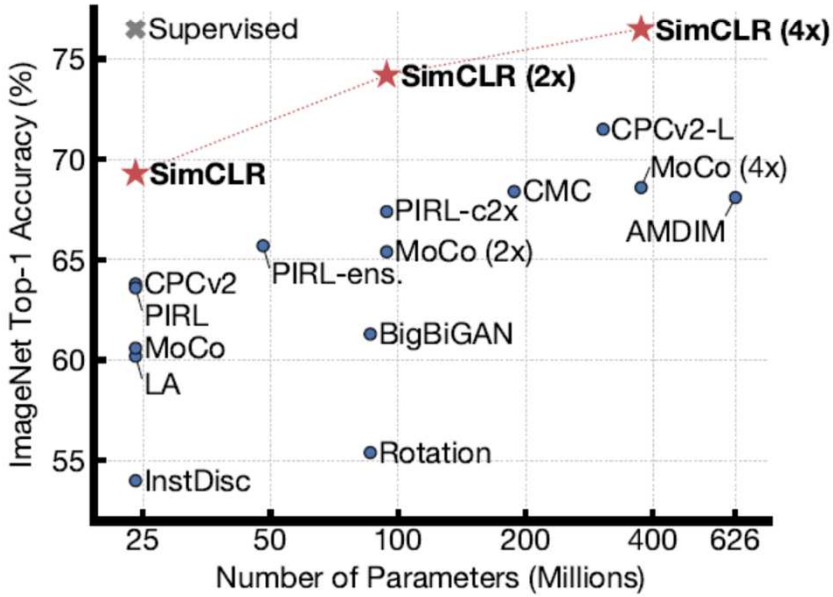
S. Gidaris, P. Singh, and N. Komodakis, "Unsupervised representation learning by predicting image rotations" in ICLR 2018.
Doersch et al. "Unsupervised Visual Representation Learning by Context Prediction", ICCV 2015
M. Noroozi and P. Favaro, "Unsupervised learning of visual representions by solving jigsaw puzzles", ECCV 2016.

Self-supervision by contrastive learning

State-of-the-art self-supervised methods **are closing the gap** with respect to the supervised couterpart in some tasks.



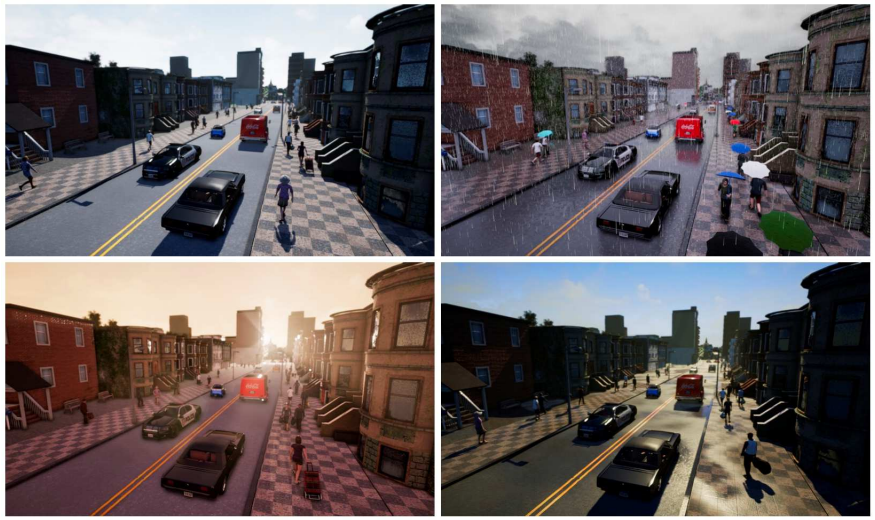


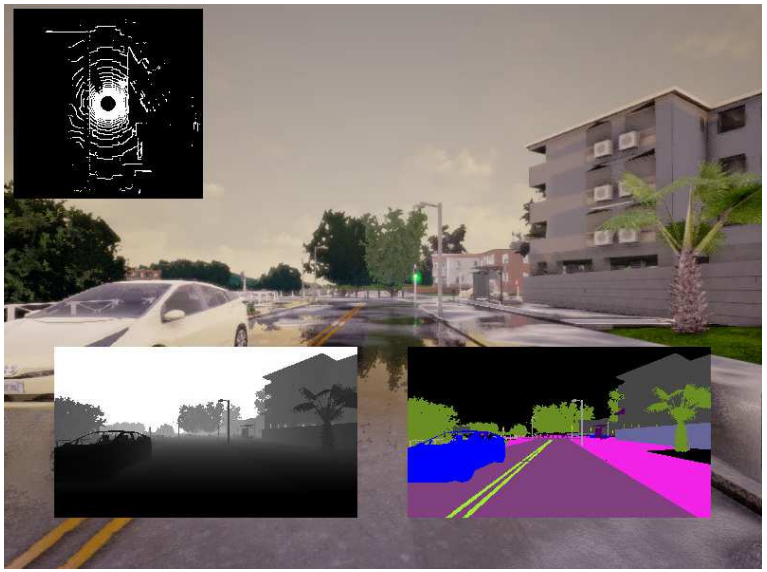


Ting Chen et al. , "A Simple Framework for Contrastive Learning of Visual Representations", https://arxiv.org/abs/2002.05709

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Domain adaptation

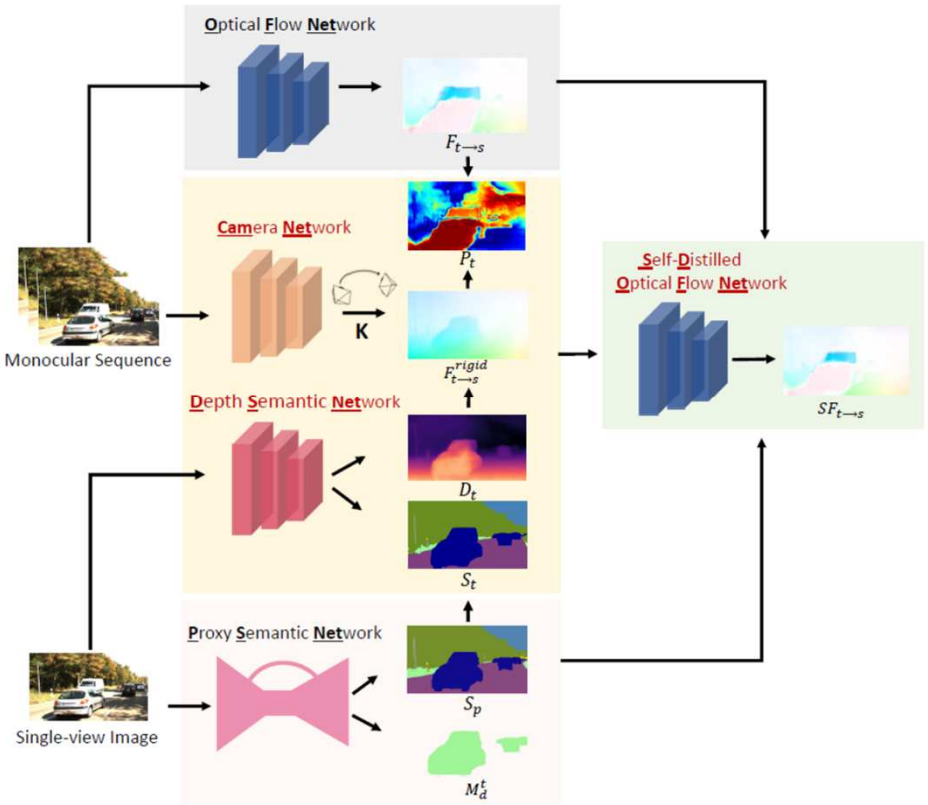


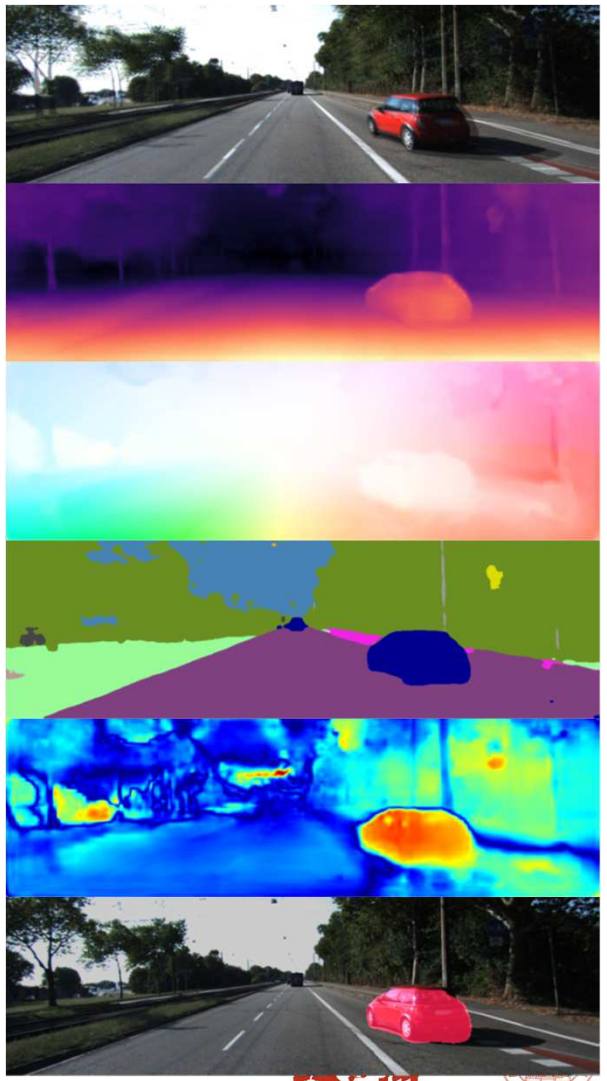


AI

A personal example: «Real-time comprehensive scene understanding» (CVPR 2020)

Learning of multiple related tasks by **self-supervision** and **proxy labels**





Tosi et al. , “Distilled Semantics for Comprehensive Scene Understanding from Videos”, http://arxiv.org/abs/2003.14030

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Samuele Salti

Dipartimento di Informatica – Scienza e Ingegneria (DISI)

samuele.salti@unibo.it

www.unibo.it