

# Large Scale Representation Learning In-the-wild

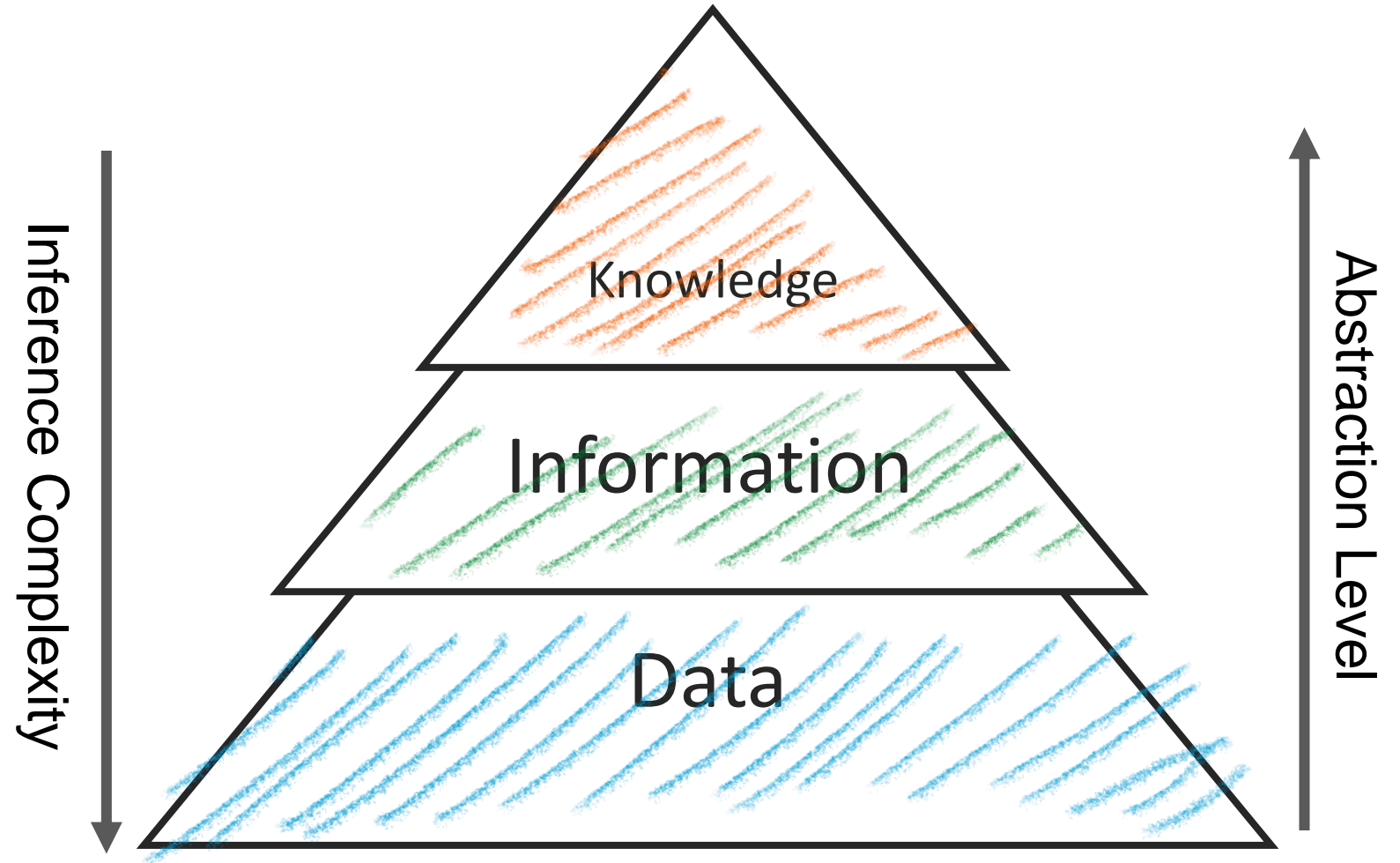
Hadi Abdi Khojasteh

TELIGHT, Czech Republic  
hadi.abdikhojasteh@telight.eu

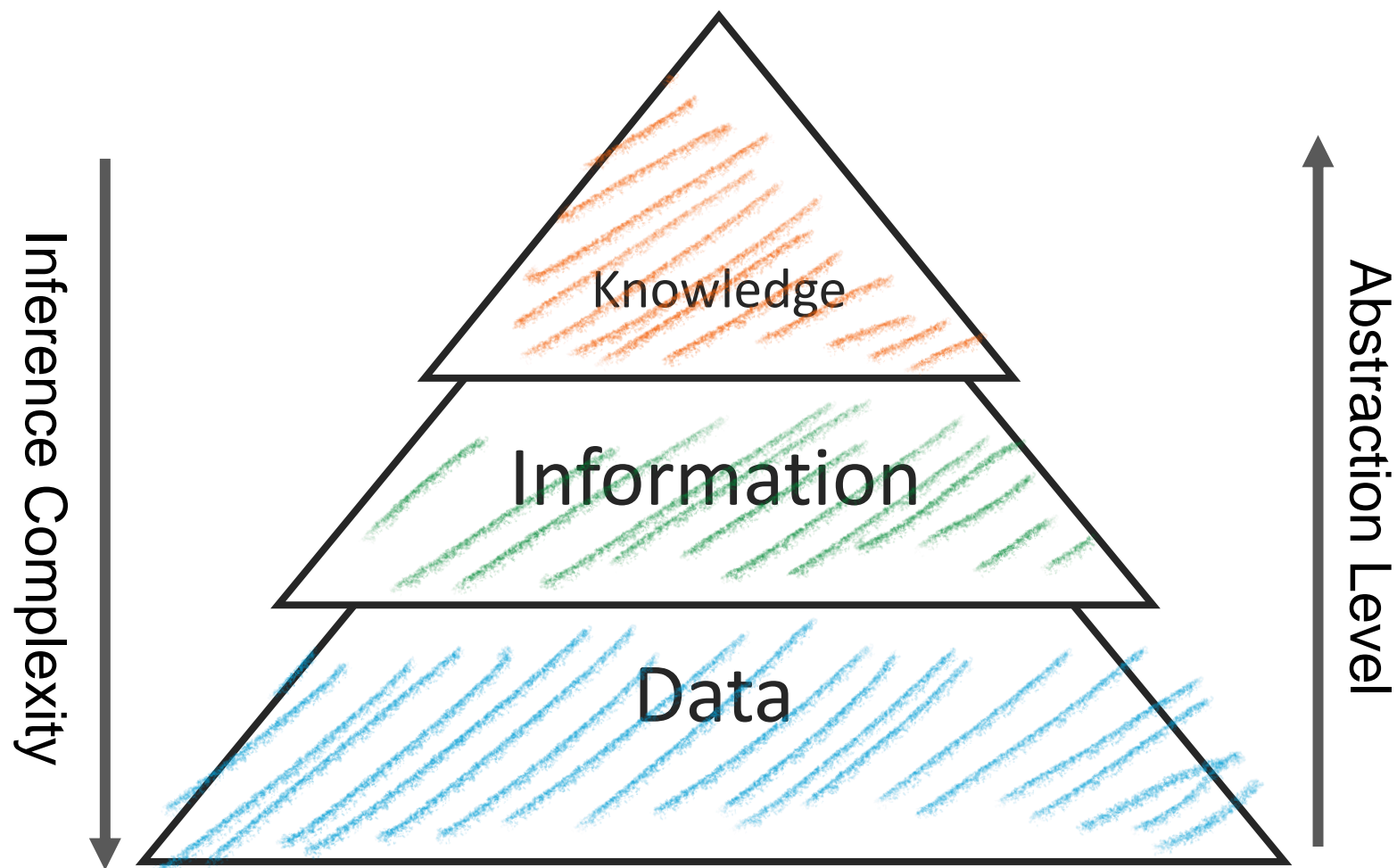
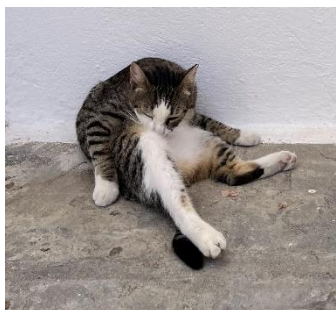


August 13th, 2022

# Definition

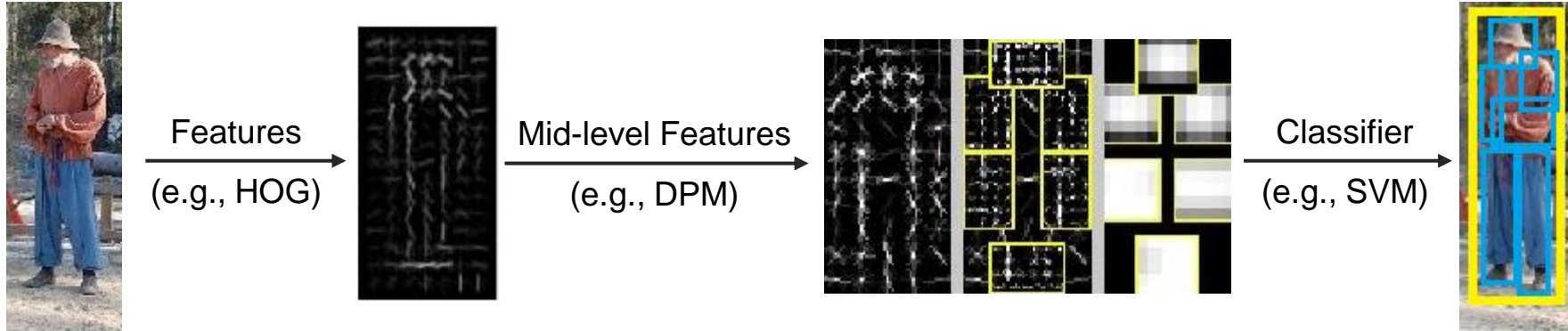


# Definition



# Deep Learning: End-to-end approach

General  
Computer  
Vision



## Learn More:

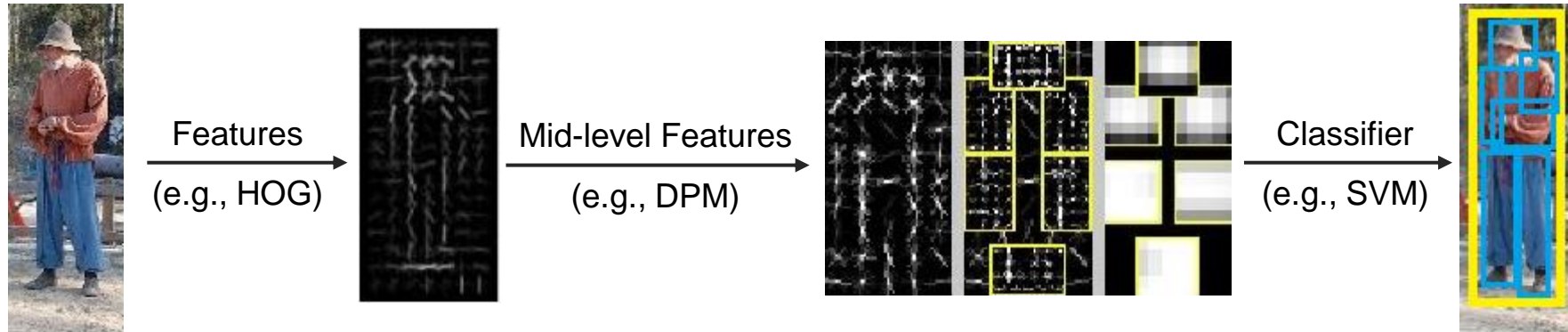
Object Detection with Discriminatively Trained  
Part-Based Models, Felzenszwalb et al (2010)



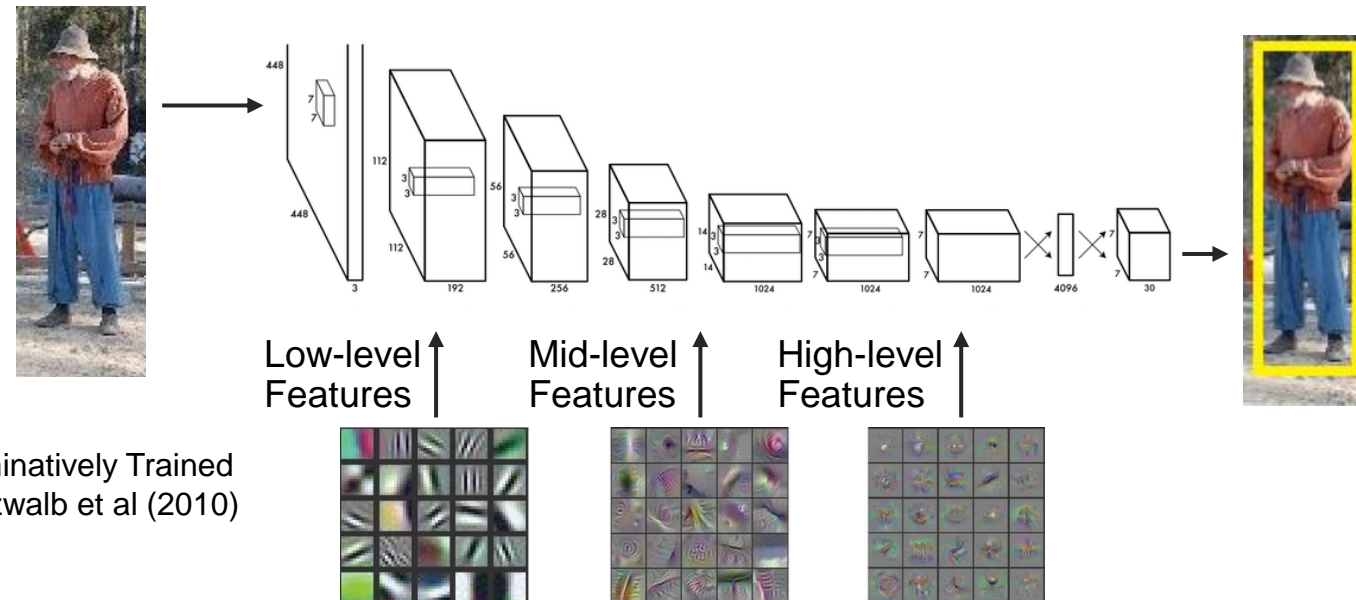
August 13th, 2022

# Deep Learning: End-to-end approach

# General Computer Vision



# Deep Learning



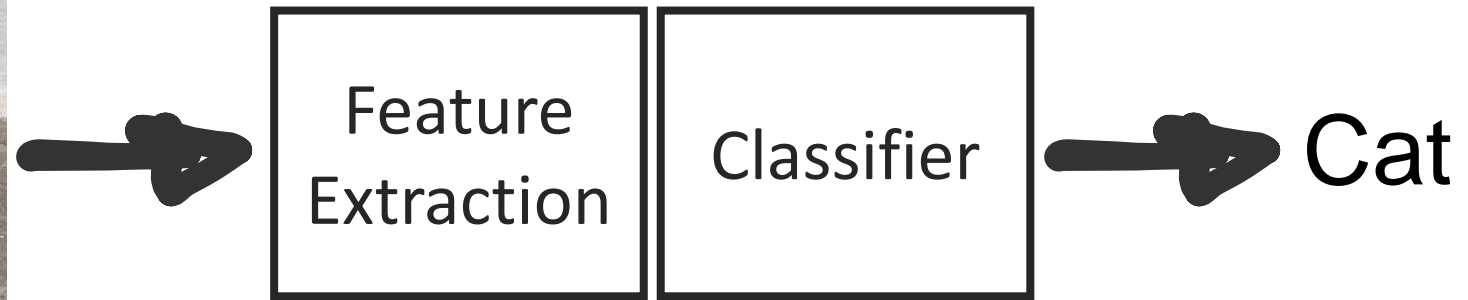
## Learn More:

## Object Detection with Discriminatively Trained Part-Based Models, Felzenszwalb et al (2010)

# Supervised Representation Learning



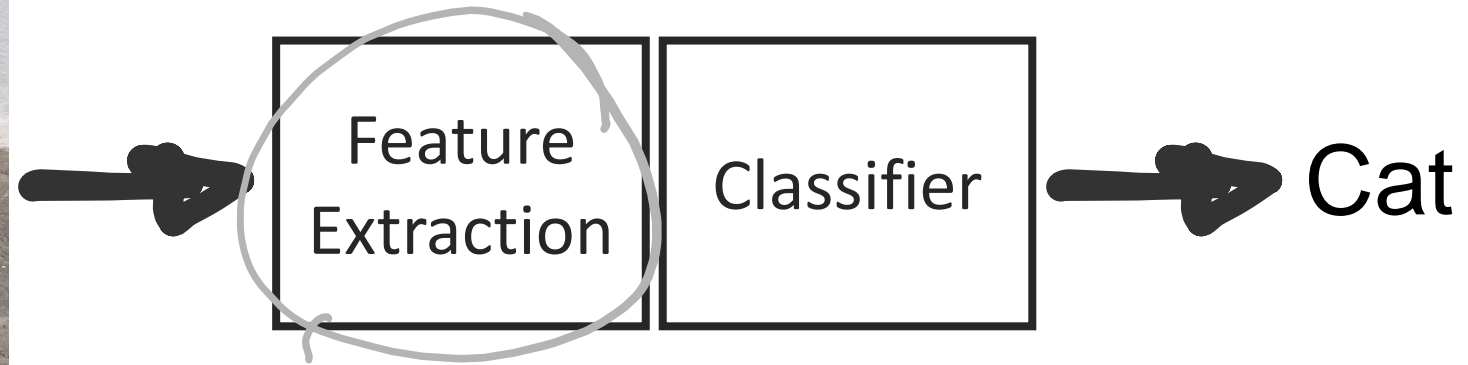
Repurpose



# Supervised Representation Learning



Repurpose



Object Detection

Semantic Segmentation

Visual Question Answering

...



# Supervised Representation Learning



Repurpose



Object Detection

Semantic Segmentation

Visual Question Answering

...

Expensive

Time-consuming

Prone to error



# Why unsupervised/self-supervised learning?

- Nature does not use supervised learning most of the time
- Taking advantage of huge unlabeled data
- Answer new questions

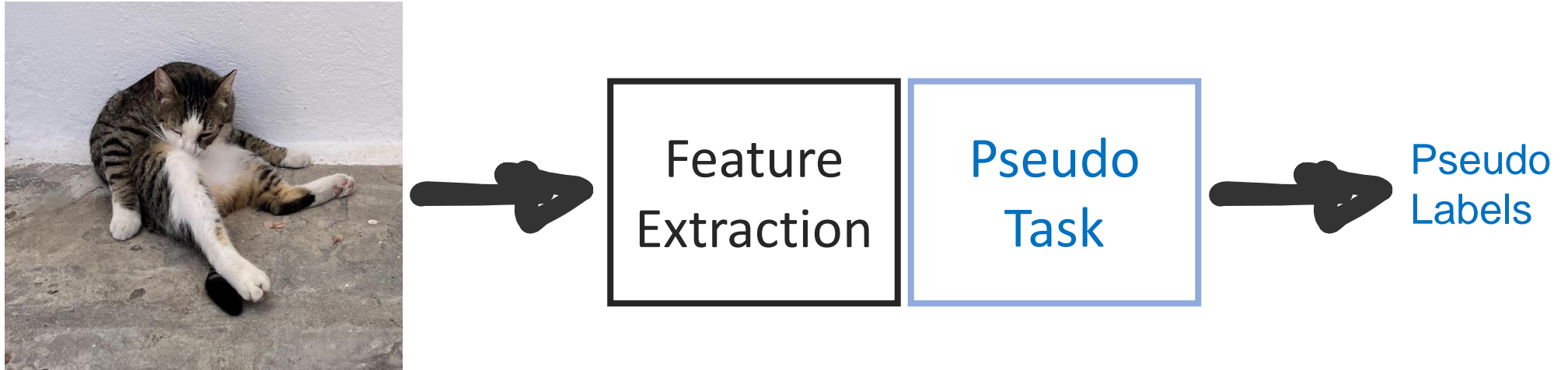
# Learning how the world ticks

- Objective is to learn distribution that data comes from

Supervised:  $p(y|x)$   
Unsupervised:  $p(x)$

- So long as our machine learning models “cheat” by relying only on surface statistical regularities

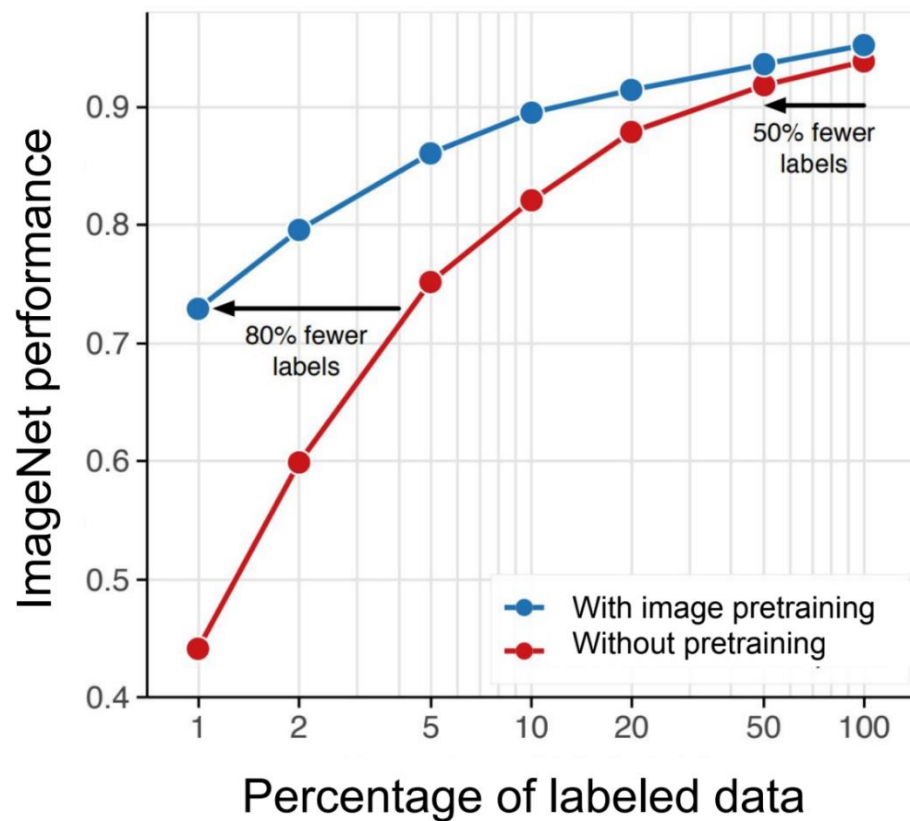
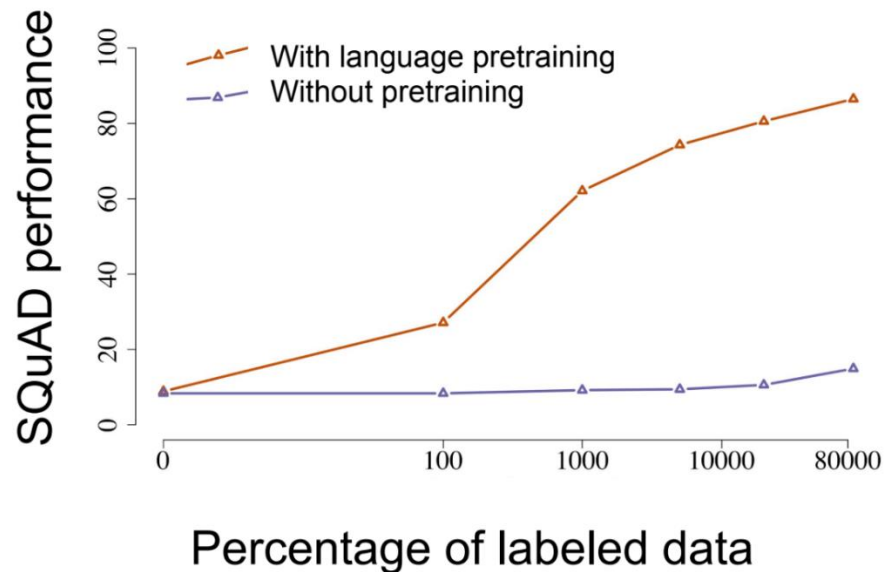
# Self-supervised Representation Learning



**Self-supervised:** pseudo labels

- Low cost
- More scalable
- Flexible

# Recent Progress

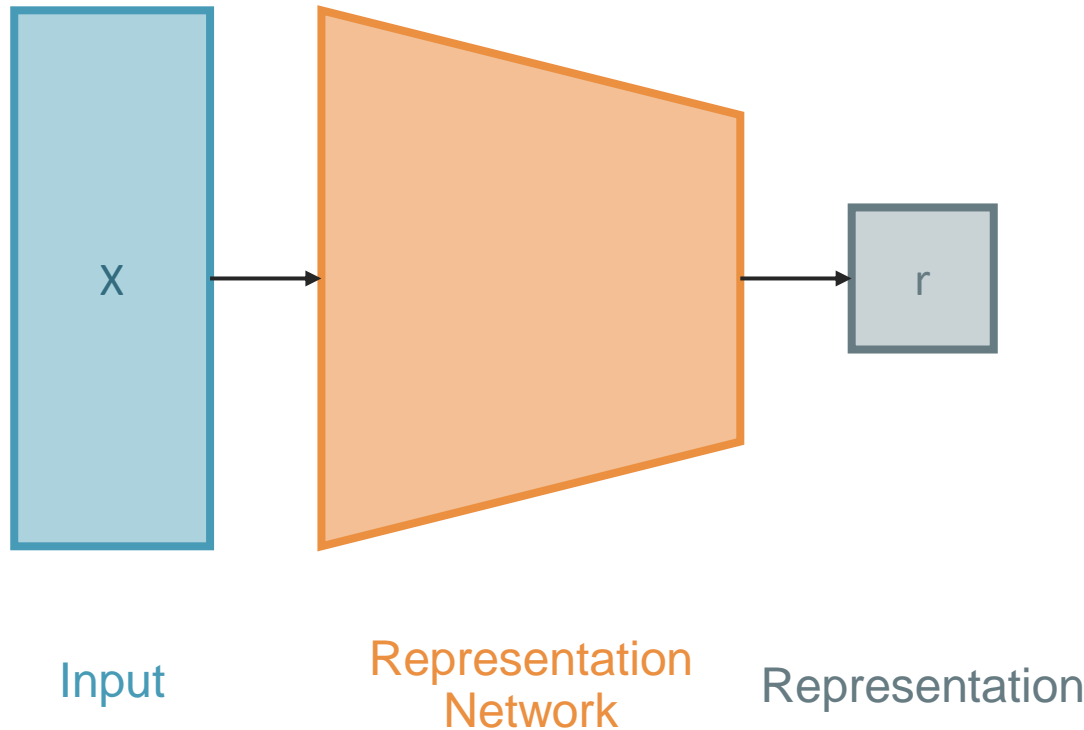


## Learn More:

Learning and Evaluating General Linguistic Intelligence, Yogatama et al

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

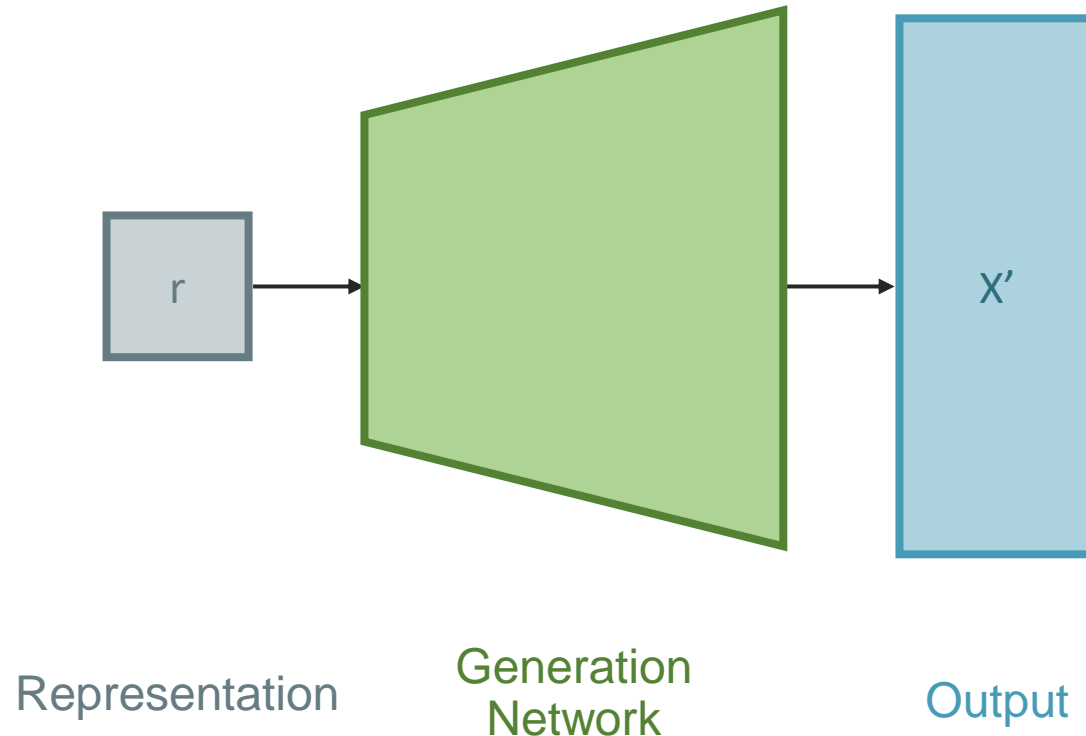
# (Representation / Encoder / Inference) Networks



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

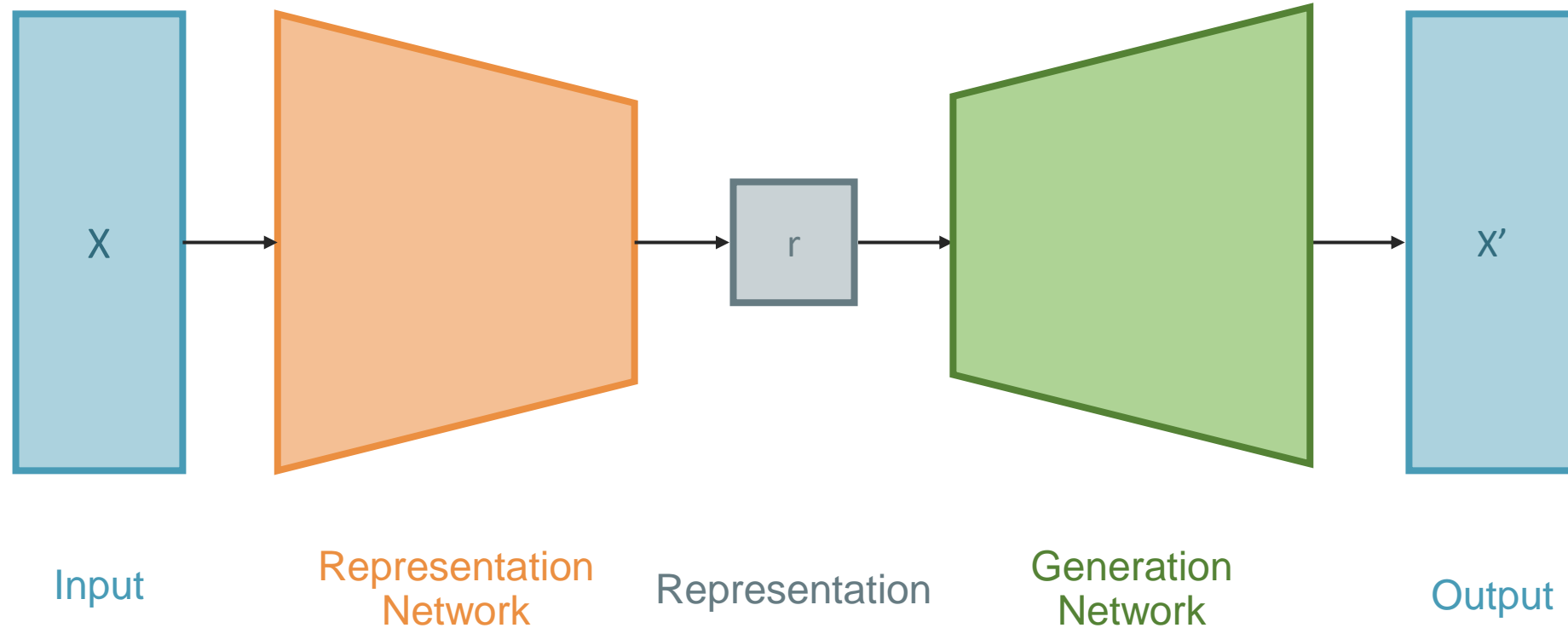
- Multi-layer perceptron
- ConvNet
- Transformer
- Recurrent neural net

# (Generation / Generator / Decoder) Networks



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

# Autoencoders



## Learn More:

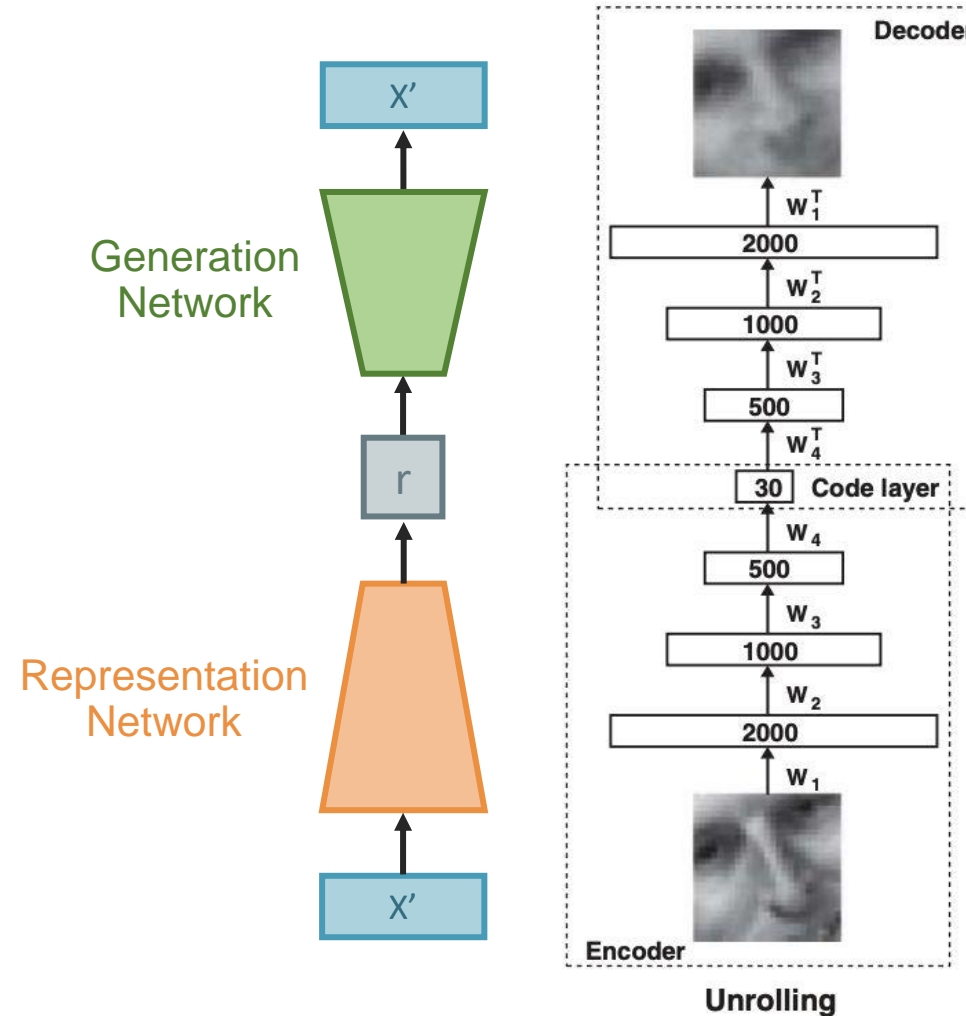
Auto-Encoding Variational Bayes, Kingma et al, ICLR (2014)

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



# Autoencoders: What are they for?

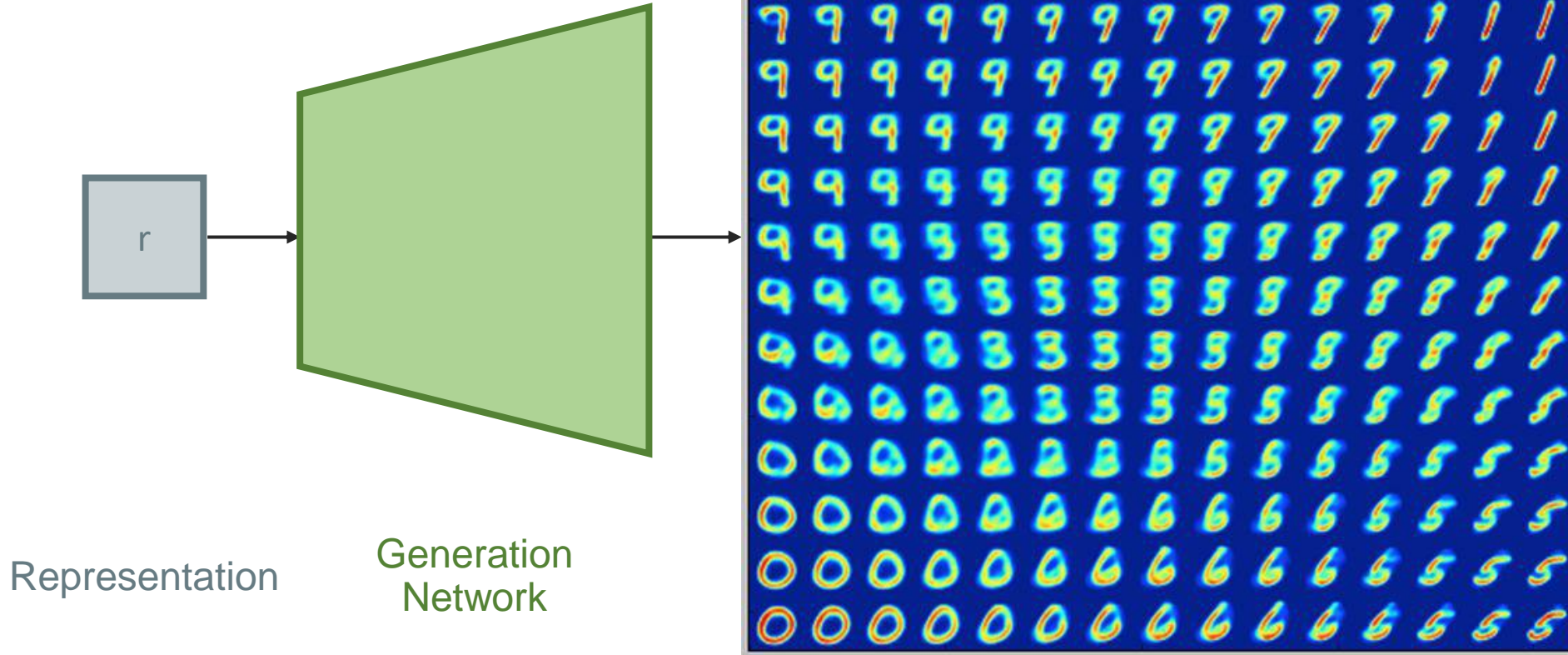
- Density estimation
- Dimensionality reduction
- Image generation
- Denoising
- **Representation learning**



## Learn More:

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

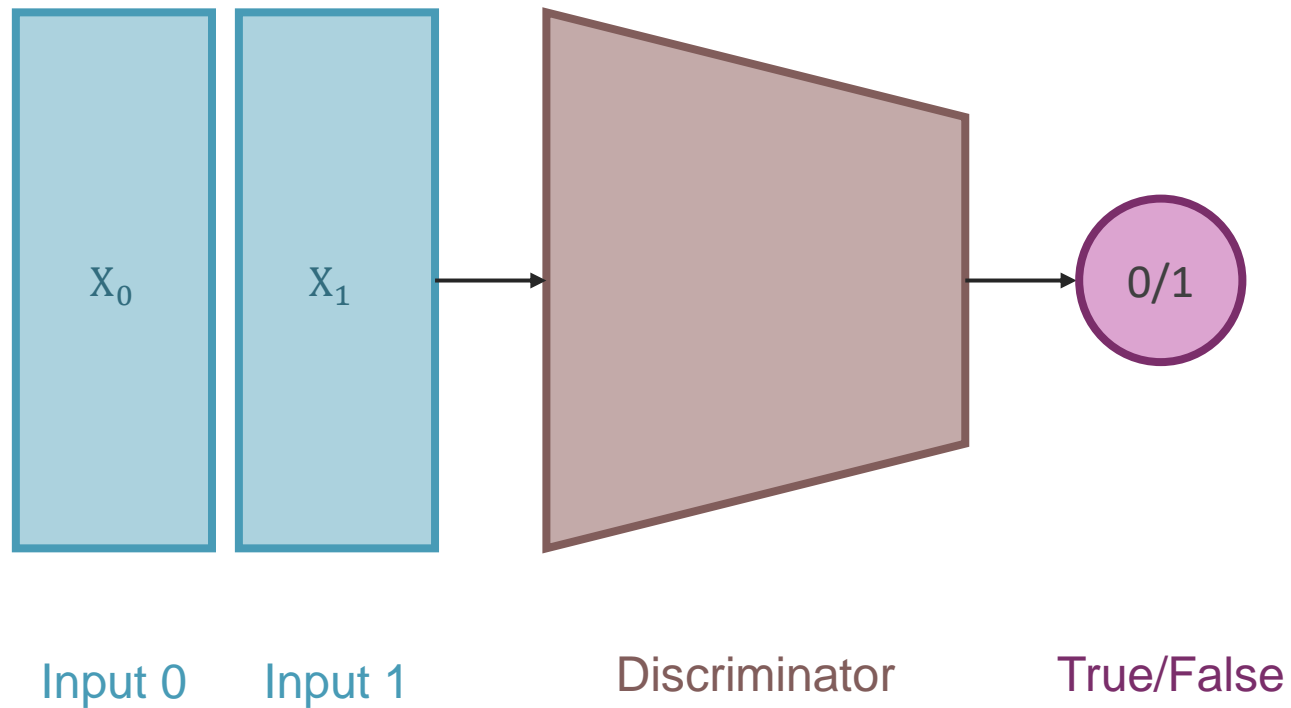
# Autoencoders



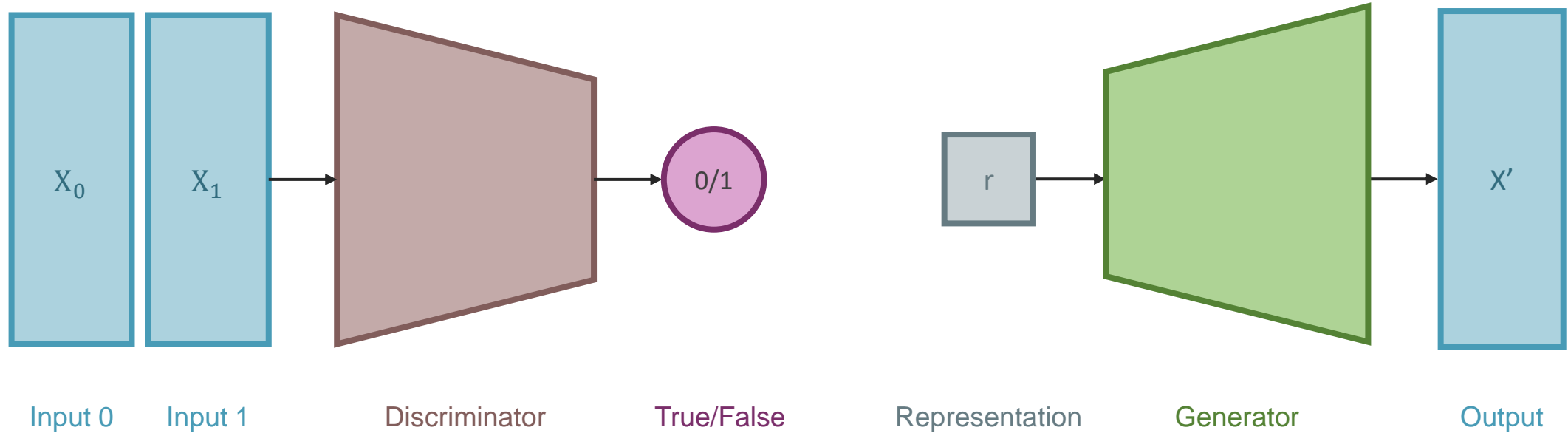
## Learn More:

Building Autoencoders in Keras, Chollet (2016)

# Discriminators / Contrastive Networks



# Generative adversarial networks



## Learn More:

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



# Generative adversarial networks

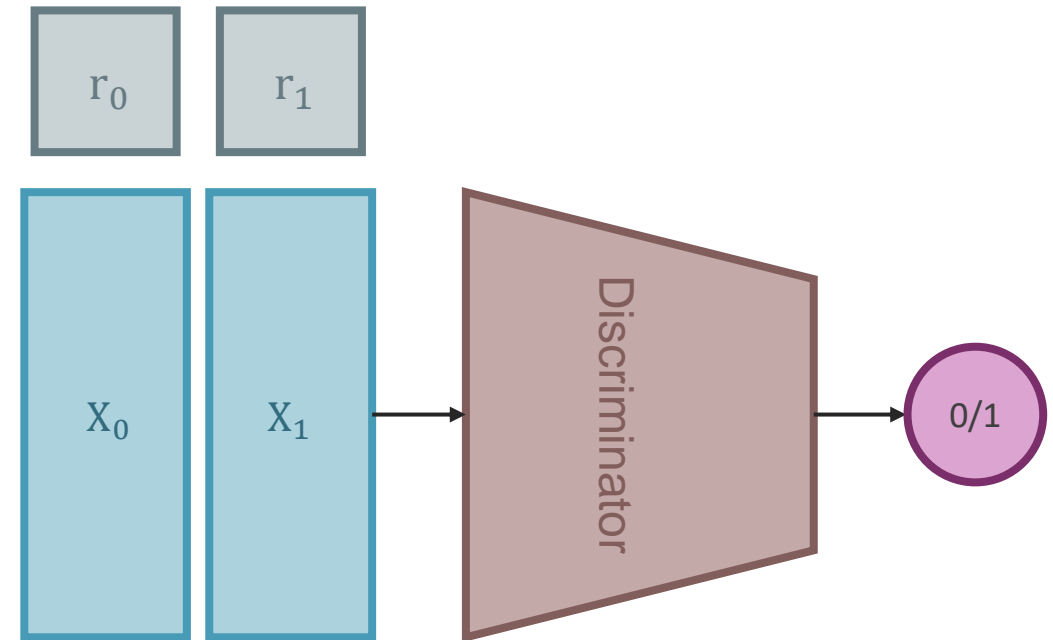
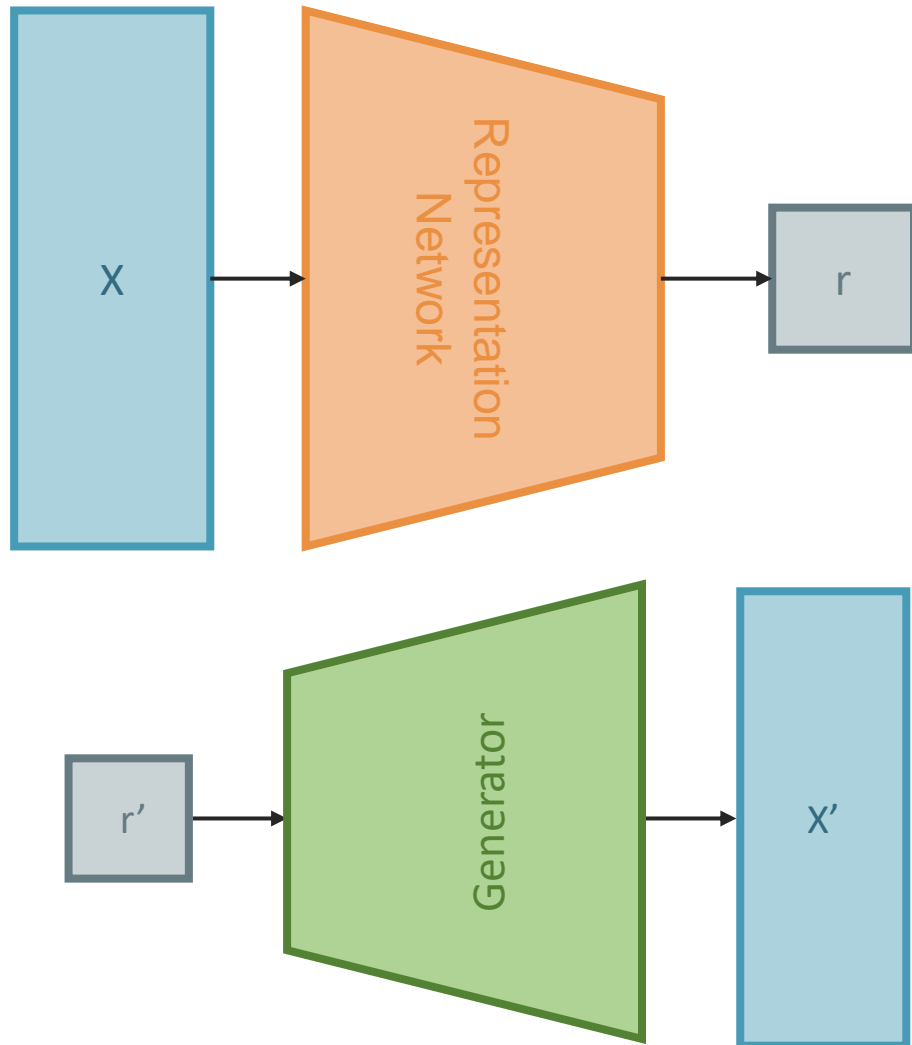


## Learn More:

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

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

# BiGAN



**Learn More:**

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



# BigBiGAN

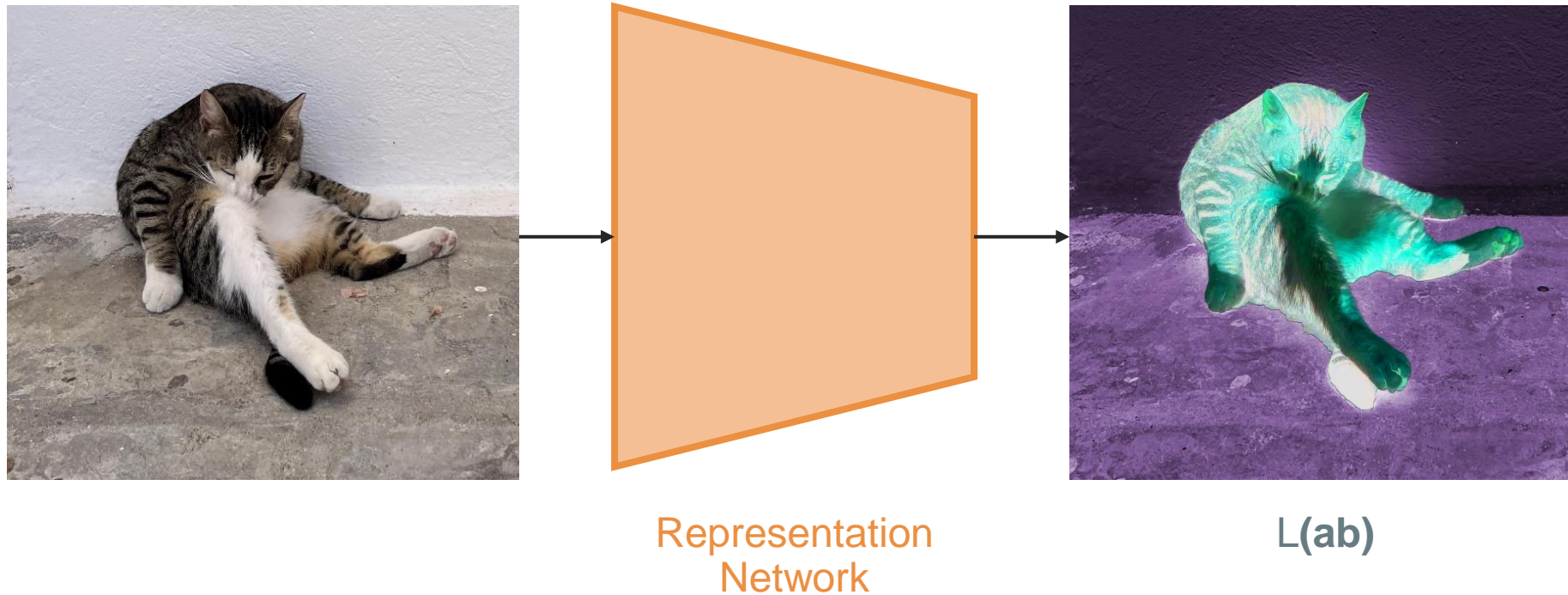


## Learn More:

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



# Generative Adversarial Networks: Colorization



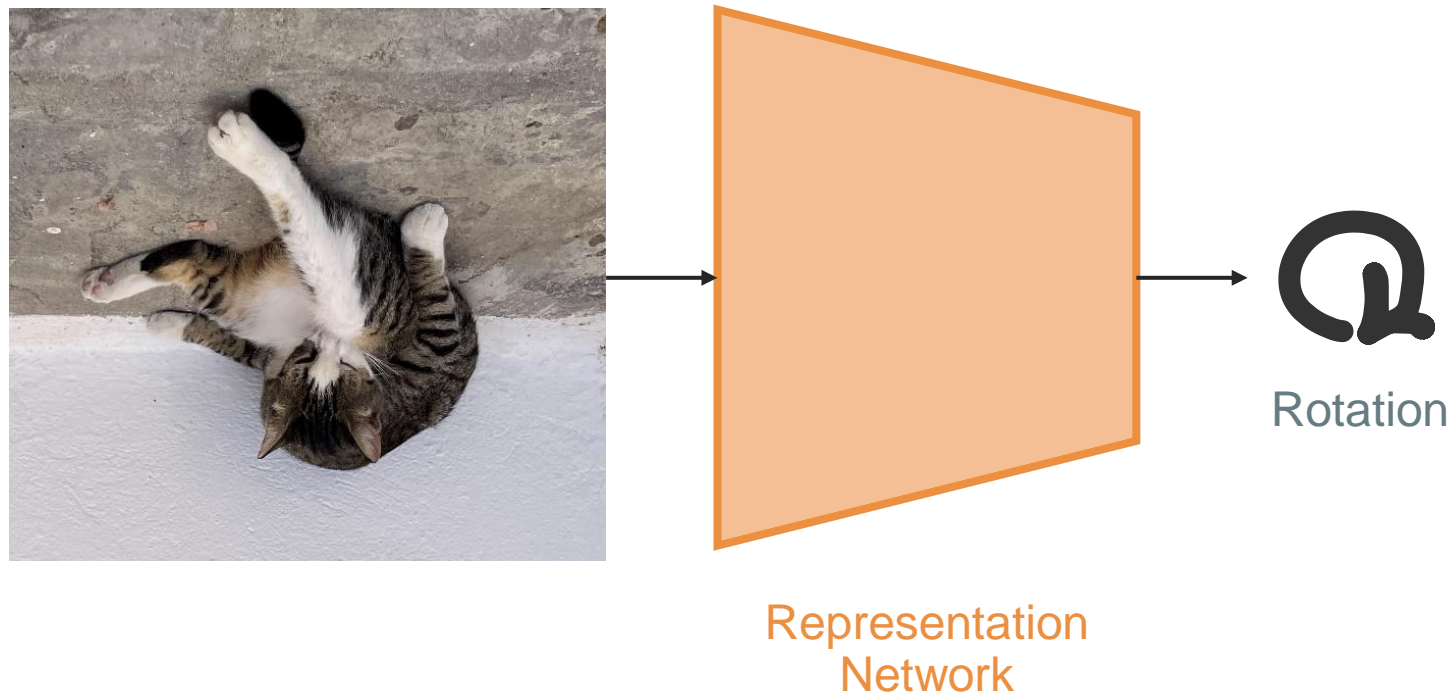
## Learn More:

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



August 13th, 2022

# Generative Adversarial Networks: Rotation Prediction



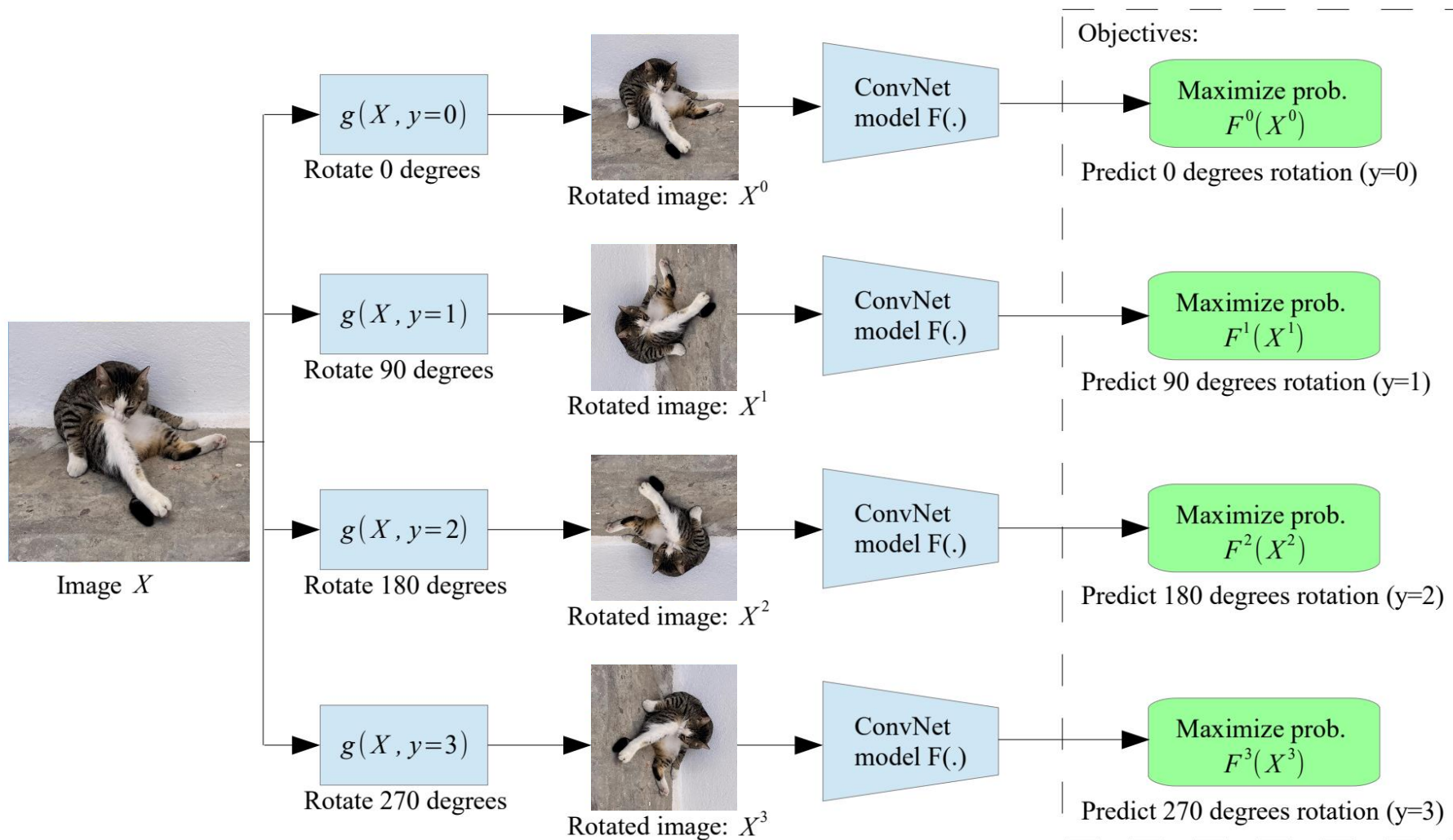
## Learn More:

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



August 13th, 2022

# Generative Adversarial Networks: Rotation Prediction

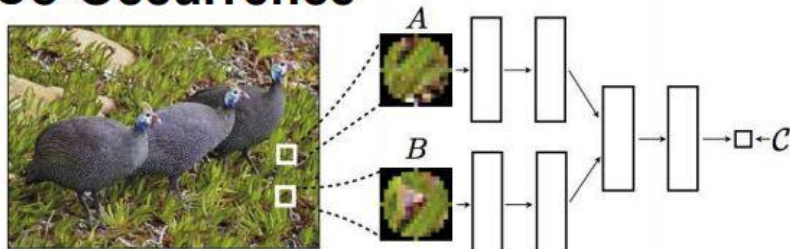


## Learn More:

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

# Self-supervised learning

## Co-Occurrence



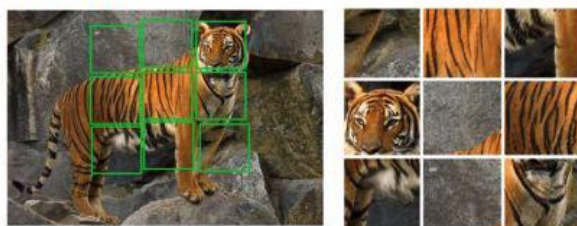
Isola *et al.* ICLR Workshop 2016.

## Egomotion



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

## Context

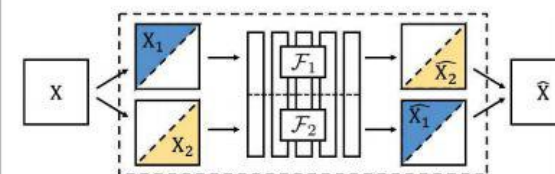


Noroozi *et al.* 2016



Pathak *et al.* CVPR 2016

## Split-brain auto-encoders



Zhang *et al.* CVPR 2017

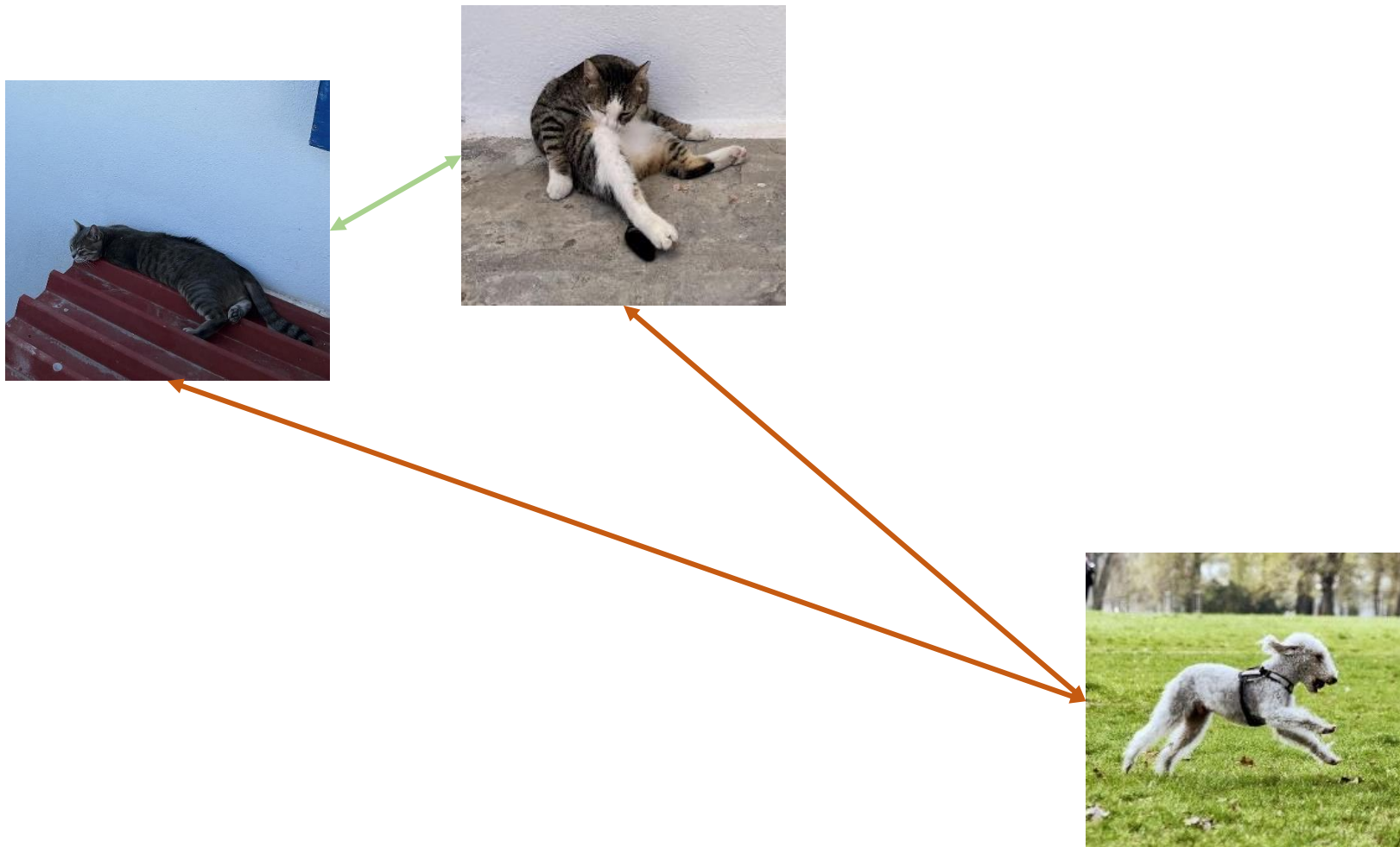


**Learn More:**

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



# Contrastive learning



# Contrastive learning

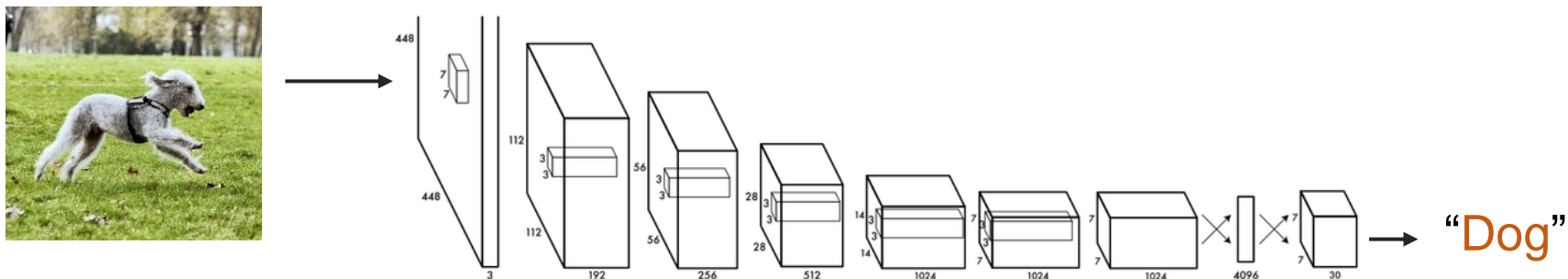
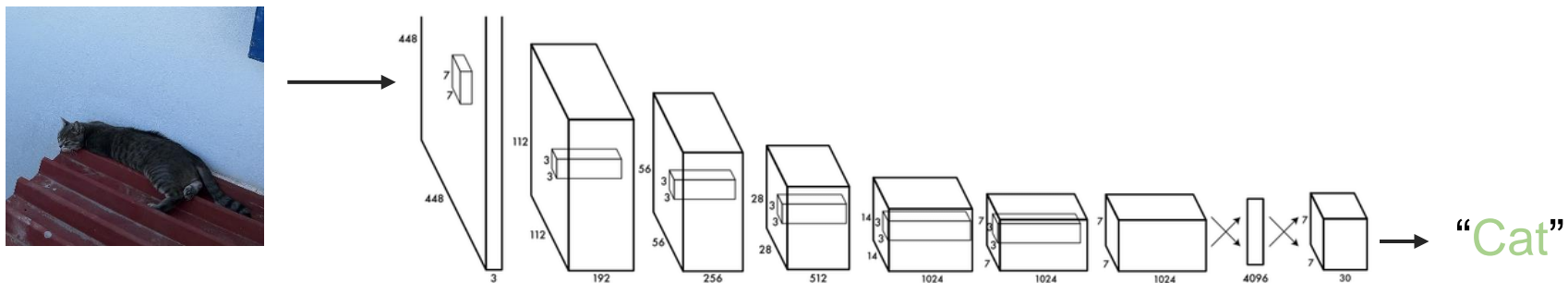
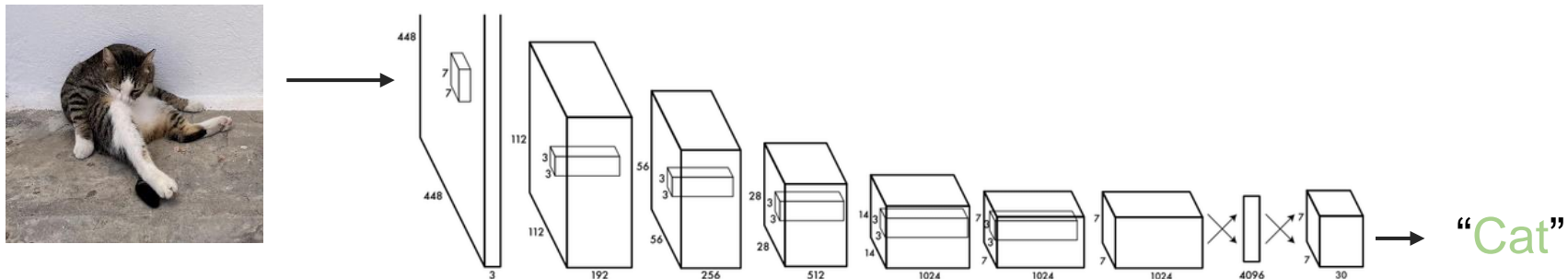
What data points are **similar**?



What data points are **dissimilar**?



# Supervised Representation Learning





# Contrastive learning

What data points are **similar**?



What data points are **dissimilar**? —————> Random Selection



# Contrastive learning

What data points are **similar**? —————> Multiple Views



What data points are **dissimilar**? —————> Random Selection



# How to obtain different views?



## Learn More:

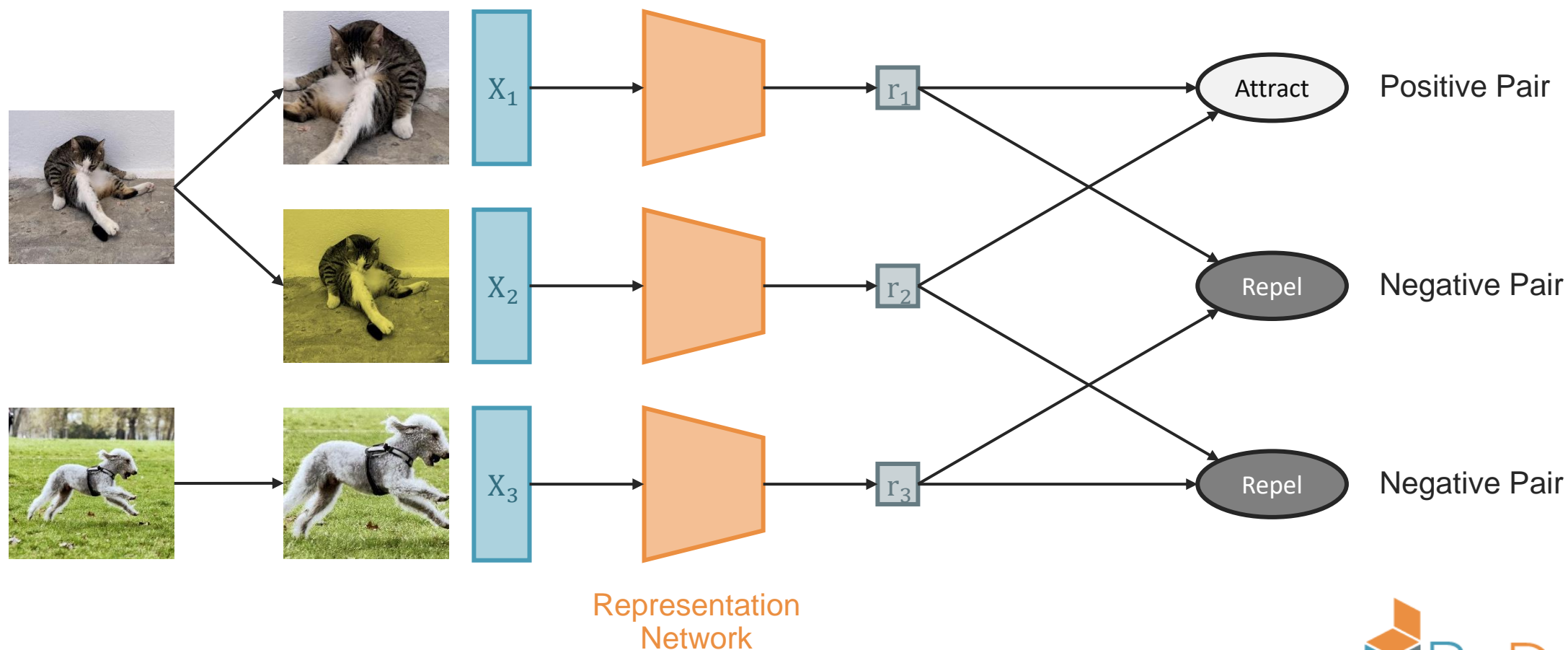
Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles, Noroozi et al, ICCV (2017)



## Learn More:

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

# Contrastive learning

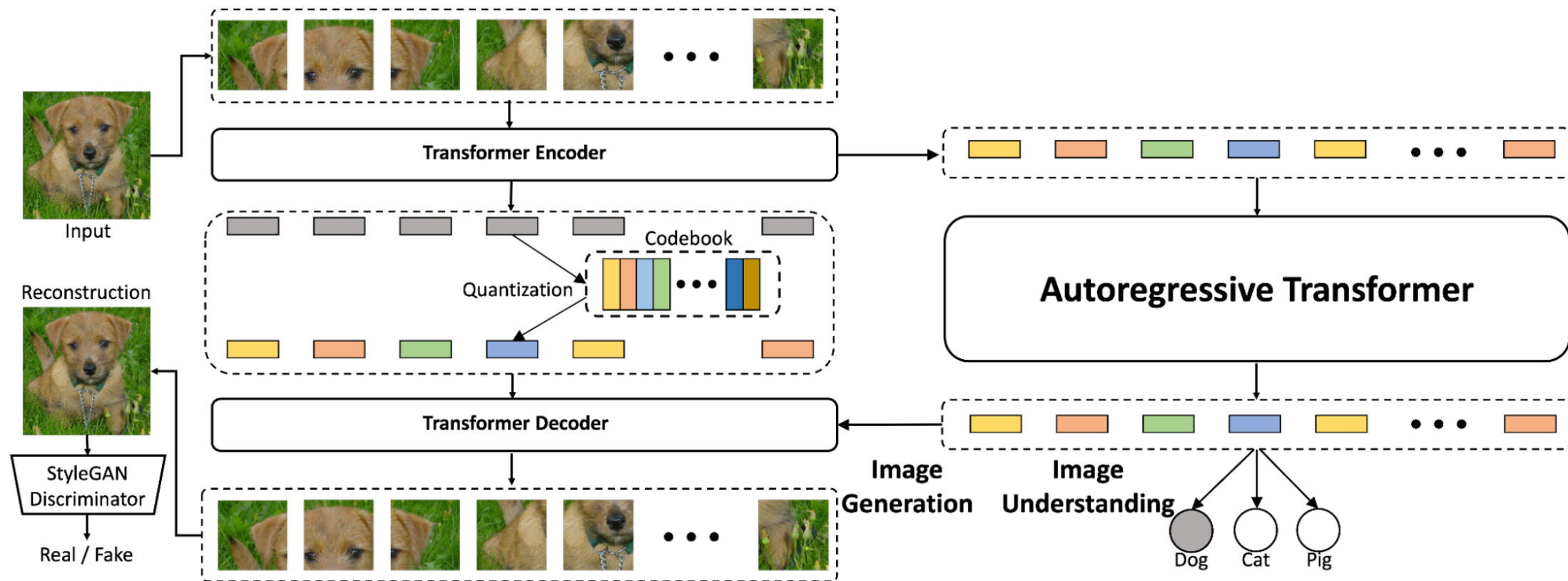


## Learn More:

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



# Vector-Quantized Image Modeling with Improved VQGAN *New*



Stage 1: Image Quantization

Stage 2: Vector-quantized Image Modeling

**Learn More:**

Vector-quantized Image Modeling with Improved VQGAN, Yu et al, ICLR (2022)



\* this photo was taken in Lindos on my last trip

Materials have  
been shared!



<https://github.com/hkhojasteh/PyData-Yerevan>



@hadiakhojasteh



hadiabdikhojasteh



August 13th, 2022