Generative Modeling: GANs

Akash Srivastava

neuromatch academy

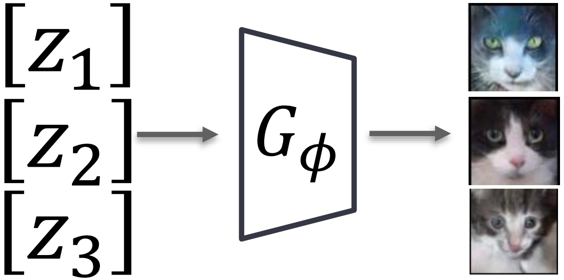
Conditional Generative Models

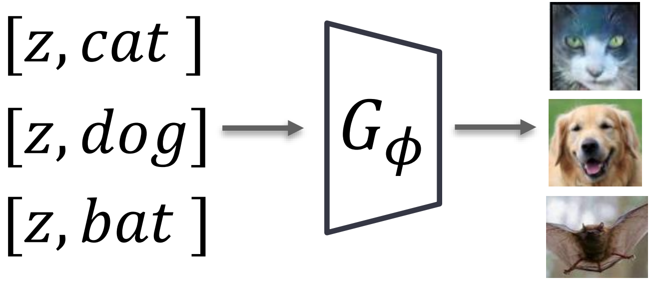
Lecture 8

neuromatch academy

Conditional Generative Model

- So far, we have seen that given a random sample Z, a trained GAN generates a random sample from the model distribution.
- This is great! But how do we control this generation process?
    - For Example,

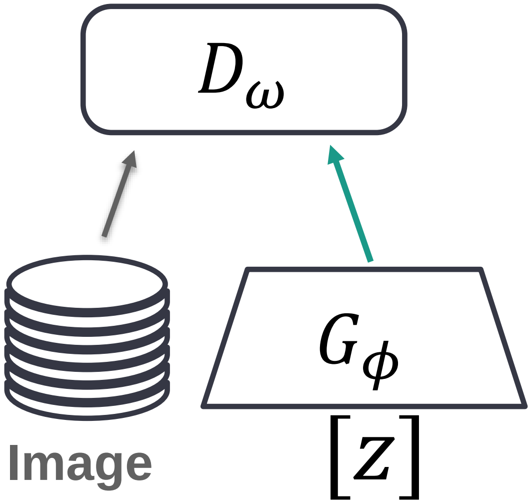




GAN vs Conditional GAN (Mirza and Osindero, 2014)

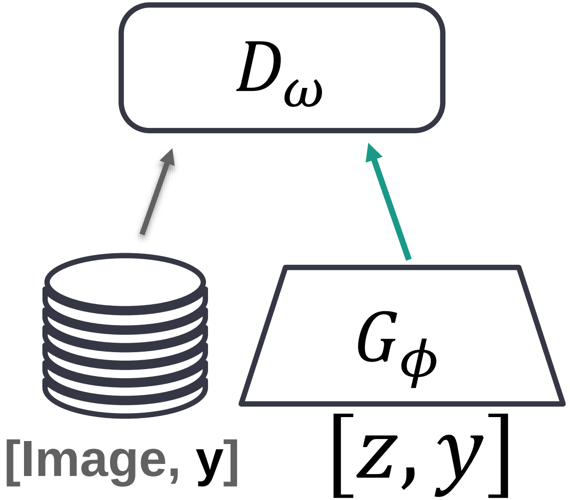
GAN

$$\min_{\phi}\max_{\omega}\mathbb{E}_{x\sim p_x}[\sigma(D_\omega\boxed{(x)})] + \mathbb{E}_{z\sim\mathcal{N}(0,I_K)}[1-\sigma(D_\omega(G_\phi\boxed{(z)}))]$$



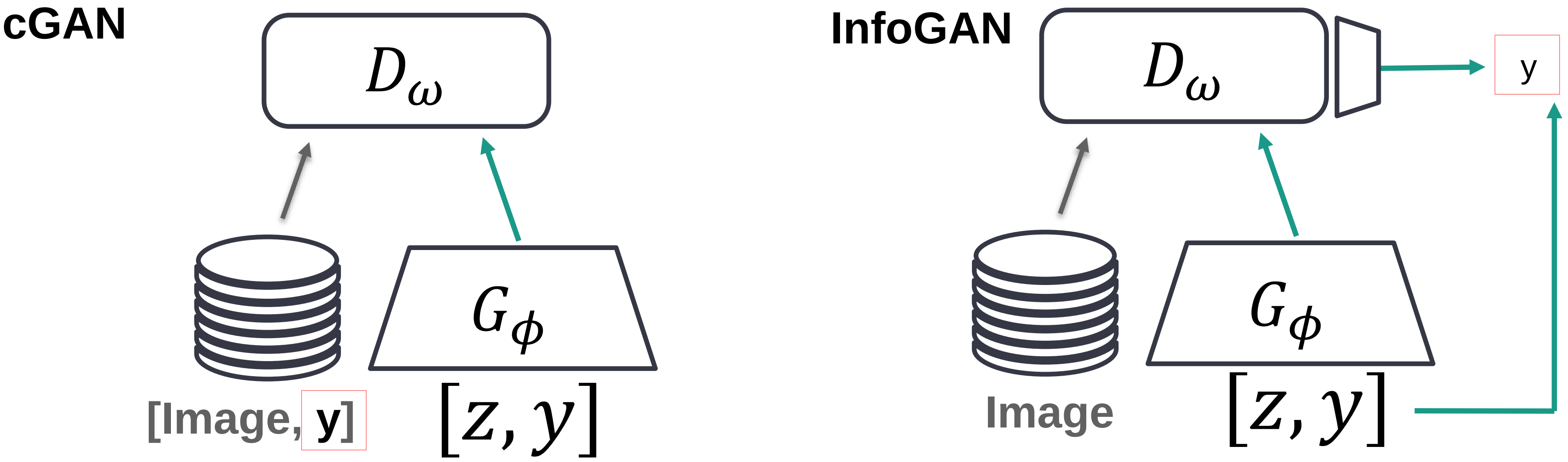
cGAN

$$\min_{\phi}\max_{\omega}\mathbb{E}_{x\sim p_x}[\sigma(D_\omega\boxed{(x|y)})] + \mathbb{E}_{z\sim\mathcal{N}(0,I_K)}[1-\sigma(D_\omega(G_\phi\boxed{(z|y)}))]$$



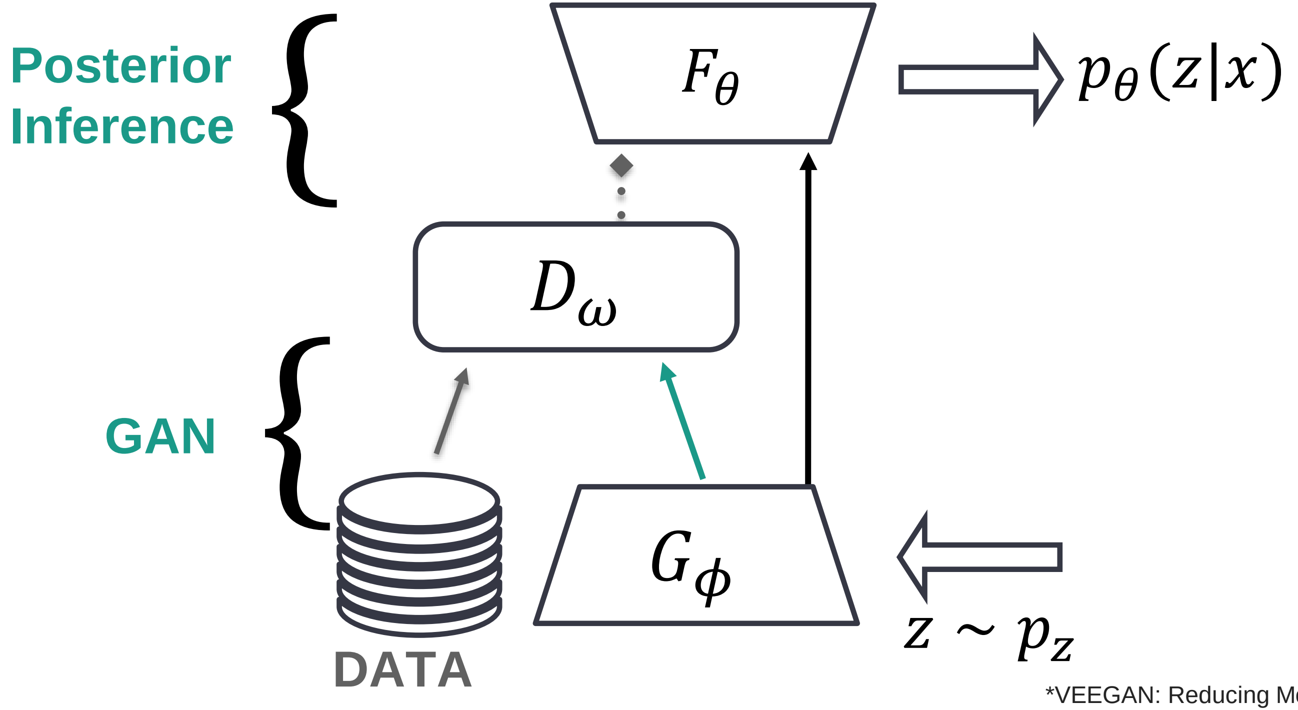
cGAN vs InfoGAN

cGAN allows conditioning on a label. In contrast, InfoGAN can be used to condition on a latent random variable using mutual information maximisation.



*InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets

GANs as upside-down VAEs

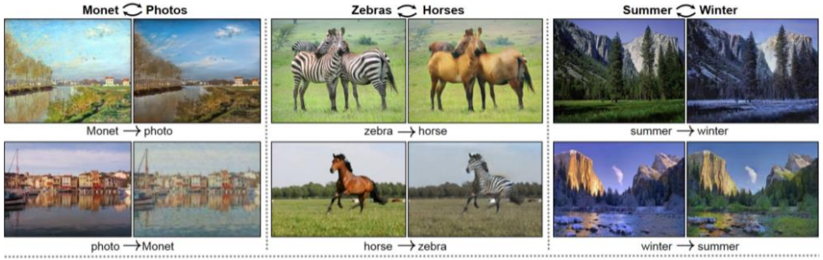


- Using conditioning in a specific way, allows us to write down GANs as upside-down VAEs that can do posterior inference!

*VEEGAN: Reducing Mode Collapse in GANs using Implicit Variational Learning.

Conditional Generative Model

- Conditioning in GANs comes in many different flavours.
    - AAE, ALI, BiGAN, BigGAN, BigBiGAN, … this is an endless list of methods
- Even methods such as Disco GAN and Cycle GAN (and all the newer avatars) use conditioning to achieve style transfer.



*Source: https://junyanz.github.io/CycleGAN/

Ethical Issues

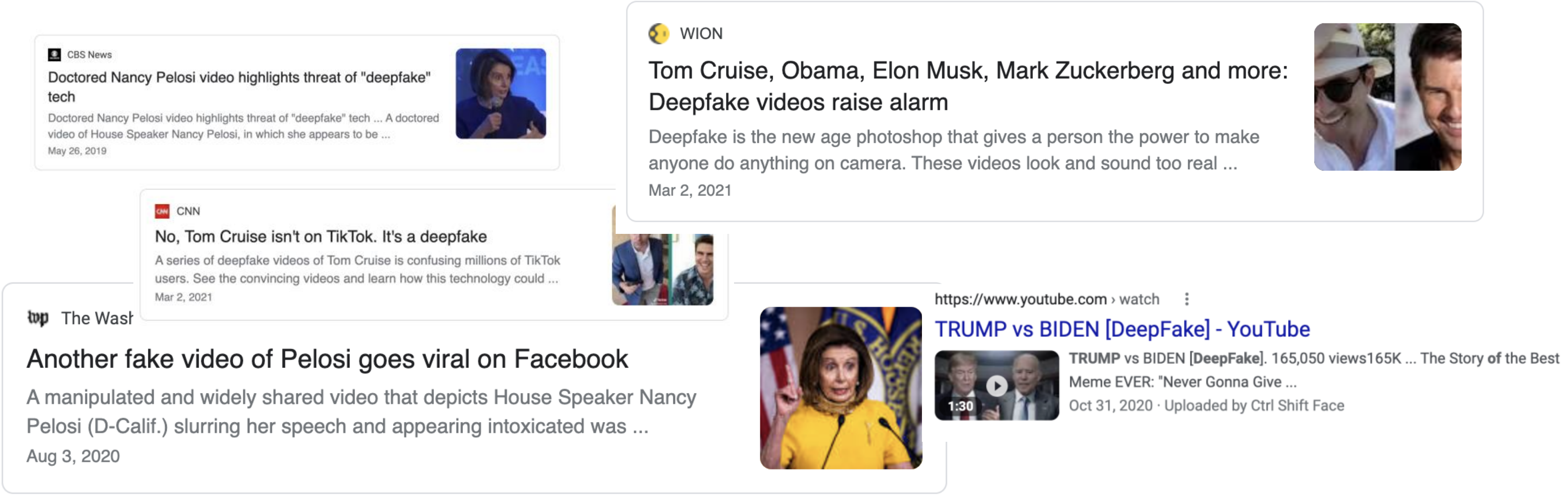
Lecture 9

neuromatch academy

Ethical Issues

- Deepfakes.
- Energy consumption.

Deep Fakes


CBS News
Doctored Nancy Pelosi video highlights threat of "deepfake" tech
Doctored Nancy Pelosi video highlights threat of "deepfake" tech ... A doctored video of House Speaker Nancy Pelosi, in which she appears to be ...
May 26, 2019
WION
Tom Cruise, Obama, Elon Musk, Mark Zuckerberg and more: Deepfake videos raise alarm
Deepfake is the new age photoshop that gives a person the power to make anyone do anything on camera. These videos look and sound too real ...
Mar 2, 2021
CNN
No, Tom Cruise isn't on TikTok. It's a deepfake
A series of deepfake videos of Tom Cruise is confusing millions of TikTok users. See the convincing videos and learn how this technology could ...
Mar 2, 2021
The Wash
Another fake video of Pelosi goes viral on Facebook
A manipulated and widely shared video that depicts House Speaker Nancy Pelosi (D-Calif.) slurring her speech and appearing intoxicated was ...
Aug 3, 2020
https://www.youtube.com › watch
TRUMP vs BIDEN [DeepFake] - YouTube
1:30
TRUMP vs BIDEN [DeepFake]. 165,050 views165K ... The Story of the Best Meme EVER: "Never Gonna Give ...
Oct 31, 2020 · Uploaded by Ctrl Shift Face

Generative Models Power Deep Fakes

The drastic improvement in high resolution image generation has unfortunately allowed some bad actors to use this technology for malicious intent.

Tools	Links	Key Features
Faceswap	https://github.com/deepfakes/faceswap	- Using two encoder-decoder pairs. - Parameters of the encoder are shared.
Faceswap-GAN	https://github.com/shaoanlu/faceswap-GAN	Adversarial loss and perceptual loss (VGGface) are added to an auto-encoder architecture.
Few-Shot Face Translation	https://github.com/shaoanlu/fewshot-face-translation-GAN	- Use a pre-trained face recognition model to extract latent embeddings for GAN processing. - Incorporate semantic priors obtained by modules from FUNIT [42] and SPADE [43].
DeepFaceLab	https://github.com/iperov/DeepFaceLab	- Expand from the Faceswap method with new models, e.g. H64, H128, LIAEF128, SAE [44]. - Support multiple face extraction modes, e.g. S3FD, MTCNN, dlib, or manual [44].
DFaker	https://github.com/dfaker/df	- DSSIM loss function [45] is used to reconstruct face. - Implemented based on Keras library.
DeepFake_tf	https://github.com/StromWine/DeepFake_tf	Similar to DFaker but implemented based on tensorflow.
AvatarMe	https://github.com/lattas/AvatarMe	- Reconstruct 3D faces from arbitrary "in-the-wild" images. - Can reconstruct authentic 4K by 6K-resolution 3D faces from a single low-resolution image [46].
MarioNETte	https://hyperconnect.github.io/MarioNETte	- A few-shot face reenactment framework that preserves the target identity. - No additional fine-tuning phase is needed for identity adaptation [47].
DiscoFaceGAN	https://github.com/microsoft/DiscoFaceGAN	- Generate face images of virtual people with independent latent variables of identity, expression, pose, and illumination. - Embed 3D priors into adversarial learning [48].
StyleRig	https://gvv.mpi-inf.mpg.de/projects/StyleRig	- Create portrait images of faces with a rig-like control over a pretrained and fixed StyleGAN via 3D morphable face models. - Self-supervised without manual annotations [49].

*Deep Learning for Deepfakes Creation and Detection: A Survey

Energy Consumption

- Training deep generative models require sophisticated hardware such as GPUs and TPUs.
- Do you know, roughly how much energy these specialized processors consume?

  The short answer is a lot! We will explore this further in the tutorial session.

There is a significant emphasis on reducing **inference time power consumption** for all sorts of deep learning methods as that allows us to run ML on small **edge devices.** But there is significantly less awareness about the **environmental impact that training large models** create.

Recap and Advanced Topics

Lecture 10

neuromatch academy

Generative Modelling

Given: $X = \{x_i | \forall j, x_j \in \mathbb{R}^D\}_{i=1}^N$ where $x \sim p_x$ is the unknown true data distribution.

Goal: Estimate p_x using only the samples in X.

Lets define a class of probability or statistical models: p_θ

and cast our estimation problem as an optimisation task of finding θ^* such that,

$$\theta^* = \arg\min_\theta D[p_x \parallel p_\theta]$$

some measure of distance D between p_x and p_θ.

Quantifying Discrepancy Between Probability Distributions

Using the ratio of the two densities:

$$D_f[P_x \parallel P_\theta] = \mathbb{E}_{x \sim p_\theta}[f(\frac{p_x}{p_\theta})]$$

Using the sup of the difference of expectations under the two distributions:

$$IPM_{\mathcal{F}}[P_x \parallel P_\theta] = \sup_{f \in \mathcal{F}} \mathbb{E}_{x \sim p_x}[f(x)] - \mathbb{E}_{x \sim p_\theta}[f(x)]$$

GAN Recipe

f-GAN

Estimate Density Ratio

Stable

Plug the estimator in an f-divergence and train the generator.

Un-Stable

Examples:
1. GAN
2. f-GAN
3. cGAN
4. VEEGAN

IPM-GAN

Maximise/Estimate IPM

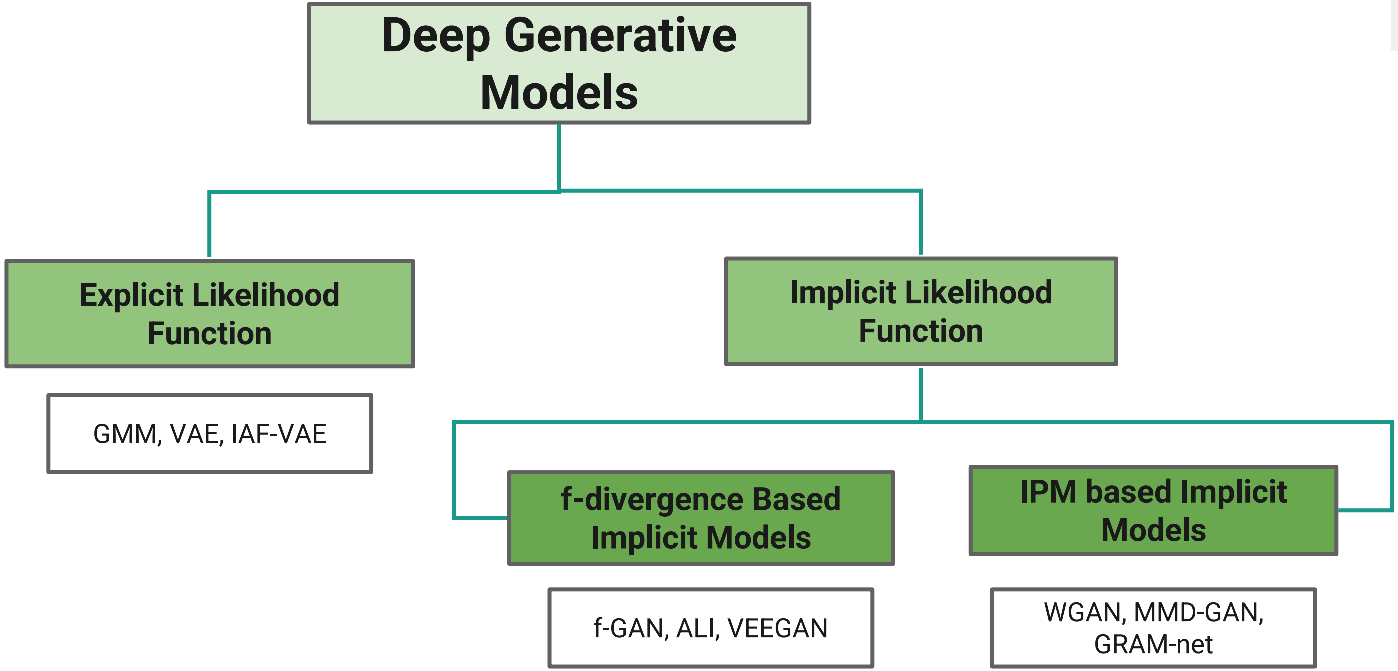
Un-Stable

Minimize the IPM to train the generator.

Stable

Examples:
1. MMD-GAN
2. WGAN
3. WGAN-GP
4. GRAM-net

Landscape of DGMs we covered!



(Too Many) Methods for Unsupervised Representation Learning

Variational Autoencoders

Auto-encoding variational bayes
DP Kingma, M Welling - arXiv preprint arXiv:1312.6114, 2013 - arxiv.org
How can we perform efficient inference and learning in directed probabilistic models, in the presence of continuous latent variables with intractable posterior distributions, and large datasets? We introduce a stochastic variational inference and learning algorithm that scales …
Cited by 13162 Related articles All 28 versions

Generative Adversarial Networks

Generative adversarial networks
IJ Goodfellow, J Pouget-Abadie, M Mirza, B Xu… - arXiv preprint arXiv …, 2014 - arxiv.org
We propose a new framework for estimating generative models via an adversarial process, in which we simultaneously train two models: a generative model G that captures the data distribution, and a discriminative model D that estimates the probability that a sample came …
Cited by 28510 Related articles All 55 versions

FLOW-based Models

Variational inference with **normalizing flows**
D Rezende, S Mohamed - International Conference on …, 2015 - proceedings.mlr.press
Cited by 1380 Related articles All 6 versions

Improving variational inference with inverse autoregressive flow
DP Kingma, T Salimans, R Jozefowicz, X Chen… - arXiv preprint arXiv …,
Cited by 970 Related articles

Glow: **Generative flow** with invertible 1x1 convolutions
DP Kingma, P Dhariwal - arXiv preprint arXiv:1807.03039, 2018 - arxiv.org
… a **generative flow** coined Glow, with various new elements as described in Section 3. In Section 5, we compare our **model** quantitatively with previous **flows**, and in Section 6, we study the qualitative aspects of our **model** on high … 2 Background: **Flow-based Generative** Models …
Cited by 886 Related articles All 6 versions

Contrastive Learning

A simple framework for **contrastive learning** of visual representations
T Chen, S Kornblith, M Norouzi… - … on machine **learning**, 2020 - proceedings.mlr.press

Exploring Simple Siamese Representation Learning
Siamese networks have become a common structure in various recent models for unsupervised visual representation learning. These models maximize the similarity between two augmentations of one image, subject to certain conditions for avoiding collapsing

Supervised contrastive learning
Cited by 101 Related articles All 6 versions

Clue: They all involve Density Ratios!

neuromatch academy