

Towards Deep Generative Models in Game Development

Jorge del Val Research Engineer / SEED (Electronic Arts)



SEED // SEARCH FOR EXTRAORDINARY EXPERIENCES DIVISIO





SEED







Agenda

- Motivation and fundamentals 1.
- Variational autoencoders (VAE) 2.
- Generative adversarial networks (GAN) 3.
- Conditional generative models 4.
- Some applications to game development 5.



In a sentence...

Models that generate or remix stuff



In a better sentence...

Models that learn the data probability distribution and are able to sample from it



But... why in games?



Nysdattystant		ant rickory	Had.	Hittest	teres topont and tarve risks must		-7 A	Deplicate	Append
stek Presection Subtast		Strength Hultiplier I	trab (0.0000)	fillayer there	Takis bath good throughton man		BIN COMM FR	Contraction of	Del Other
tine Inter PolyNexkID		Innail d Pressore I		tiese leath tress	Carry	Terrate Insertifie	· 20 21	Gulete	Del Al
Test New Occurrent		Constantfilt Tilt Branh		Gradient Della	Second Stars Incoders	Greats Inserbia	theore, and deput by	59.01	
Render Past	Ent Dupth May			Whereaster Street in	Parate 1 1 1			Manga.	
And in case of the local division of the loc	re View Hight Relation	A aga break the Local House	SwitchColor	Labora March	Trability Country Country Country		allvantes 3.011 Mil	Remieste	
	Back Left Lapytours			Caler Hash	Constant and Sharest Correct Correct Correct		weBoints 1.276 Mil	Project	
Part Dia Langua Das Ed. 7	2400115	POLYGON ACTIONS	STREET, STREET	the terms in the			nu Kis	Entract	
	Add To Curve	Deflate	Q*icsh					Secondary	et - 11
©	Revel	boort Nanamesh	Scale	The second se				Contract Contract	the second second
FISH WY & FILL	Bridge	Invert Paint	Sphurizu						
	Ereate	Innert Palyloops	5 pas	CONTRACTOR OF A					
ALL DESCRIPTION OF	Delete	Inet	Spin Edge#						
The Street Sectors	Do Nothing	Mashi	Spit.					Recentruct	t subday
TTT H Tiles a Abgelietath	Equation	Meah To Drush	Transpose						
Rev V. Tiles 1 Abduttle 1	tatesale	Have	Université		and the second second			Divide	2010
tarface Holl	The Faces	Polyproup	2Nodelar Nodifiers					CTARE	A PA
Intel Georgiess & Rive Dise 1	TARGET		LENGO GITTINI AKA					Dynamik to	ded in
	A Share Party	Facing Front All	beland					Edgel rep-	
100 A	MI Polygona	Facing Front Island	Polygraup All					Lesate	
	WI Quade	Hat & Polygramp	Polygroup Ronder	A CONTRACT OF	A CONTRACTOR OF			ShadonBox	
are Off Hele Off Deta	All Triangles	Flot Border	Polygroup Innor	California (California)				tlayPolish	
Deectoral HeplayLast	Indexd	Flat Incor	Polygroup Island		ACCESSION OF A DESCRIPTION OF A DESCRIPR			Dynaticsh	
- 🗇	Behind & Polygroup	Flot Island	Polyloog					Homesher Top	
Malphi Brooth Malife	Brush Radius	tofrant	Polyloop to Hat.					Position	c+3 (9
	Turved Island	Intront & Polygroup	Polyloog & Pelygroup					Sice	
Carlos Ca	Street and a sure of				1000			Meditatege	
C Bloober CloRact Terretue	Man Territi Stee	Align Quarter Step	Align Third Step	Street, Street					
	Nign Half Step	Allign Full Step	No Alignment					ferentiesh	the second s
a a a Son Lana	HODIFICAS	Sec. P. Co. Constants	No. of Concession, Name	Contraction of the local division of the loc	A Street Street			Lightbers A	
Allers Effort I. North CO	Dise title Daily	Nate: Sides By Brush	Stag Size 6.1						have:
By Dabagenages B Drest Dect	Incrutal Alteration	Weak Attraction	No Attraction	ALC: NO				Array Heak	
	Disable Transle Seas	Couble Trie	right Smap		ALC: NOT A REAL PROVIDED IN				
Prok Jaw_DiStanter	Because extended visio	trable esti	maled trap						
🗢 🥏 🦉 🕾 😽	of the second	and the second se	2 COLORADO						
c Laparde Firmder, Welst Spray Brudlet					Contraction of the Party of the				
N Argd					A CONTRACTOR OF A CONTRACTOR OFTA CONTRACTOR O			Copy	
🔵 🦣 🧠 Oler S					A CONTRACT OF A				
Stradye theits there.				and the second division of the second divisio					
Sault Rain Oproch Contract		/ Dekich To filer in	the second second second	Hole Hat Hat		Draw Her	white site		



GAME DEVELOPERS CONFERENCE MARCH 18–22, 2019 | #GDC19

(Thanks A. Opara ☺)

Duplicabe	Append
	lesert
Delete	Del Other Del Al
Split	
Nariga.	
Rements	
Project	
Extract	
Succestry	teres interes
	Higher Kerl
Heconetruc	t intelle
	R Te Gin
	a state of the sta
Olvide	Sere Della
Dynamic to	
Edgel rep	
Line aller	
ShadowBox	
tinytoist	
Dyn attesh	
Remesher	
Hodify Top	dama.
Position	
Sice	
Vestatore	it e
ferentesh	
	rrays Preset
Open	
Arrey Heat	
	Abairtaban Abbirtartart Estimation

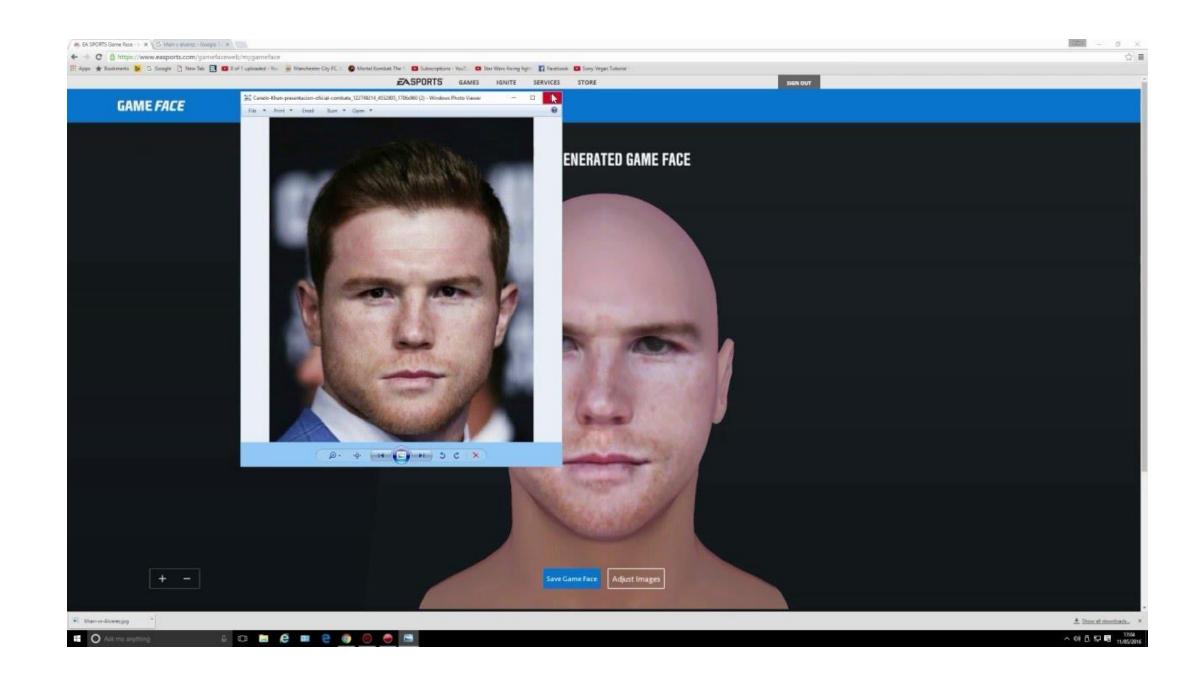






Photo Wake-Up: 3D Character Animation from a Single Photo. Weng et al. 2018



Which is real?





A Style-Based Generator Architecture for Generative Adversarial Networks. Karras et al. 2018 (NVIDIA)





Which is real?









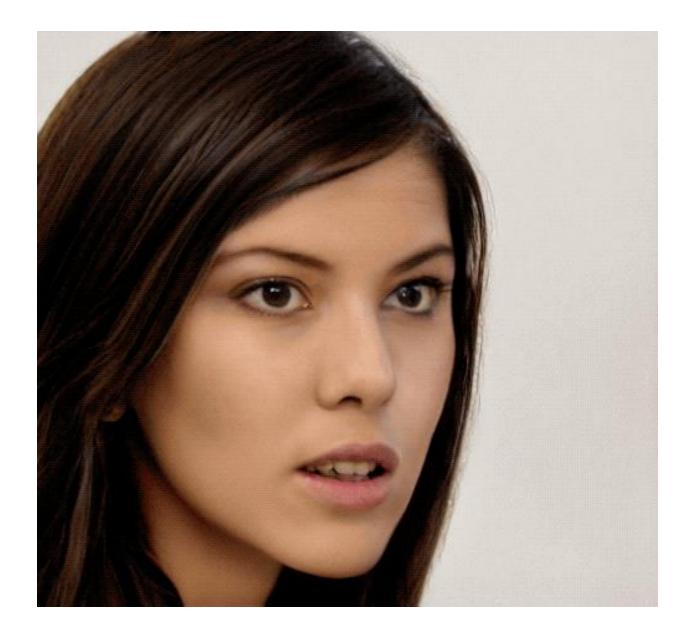
A Style-Based Generator Architecture for Generative Adversarial Networks. Karras et al. 2018 (NVIDIA)











https://thispersondoesnotexist.com

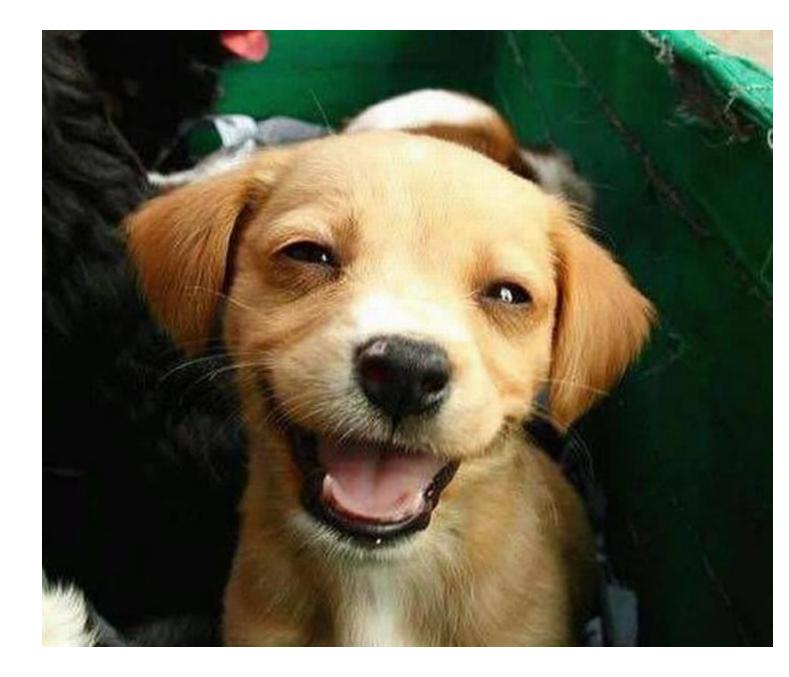
(A Style-Based Generator Architecture for Generative Adversarial Networks. Karras et al. 2018)





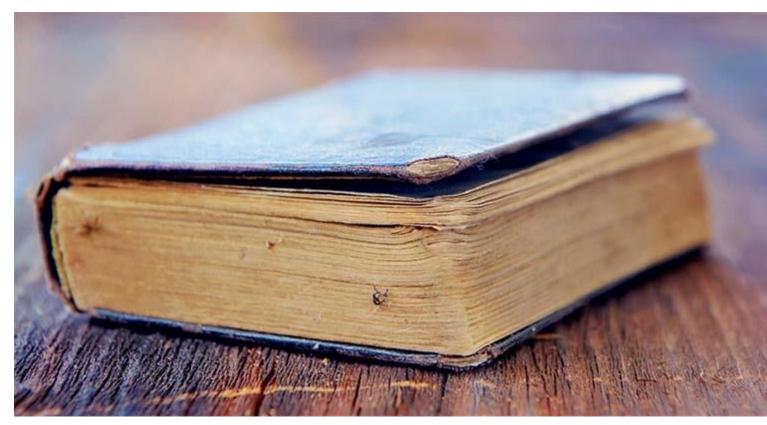
So what do they actually do?







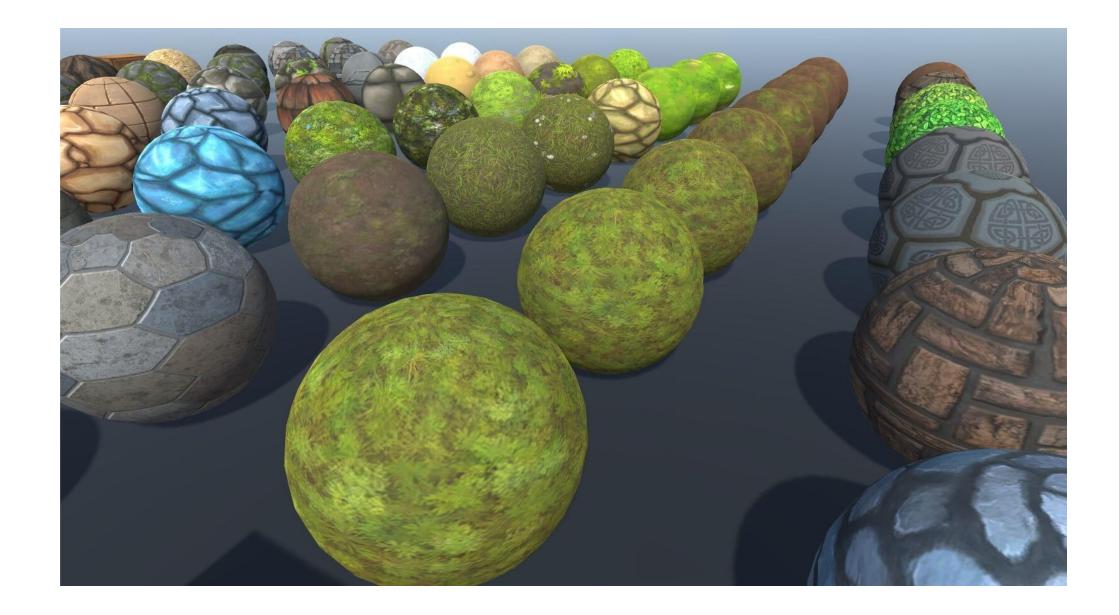






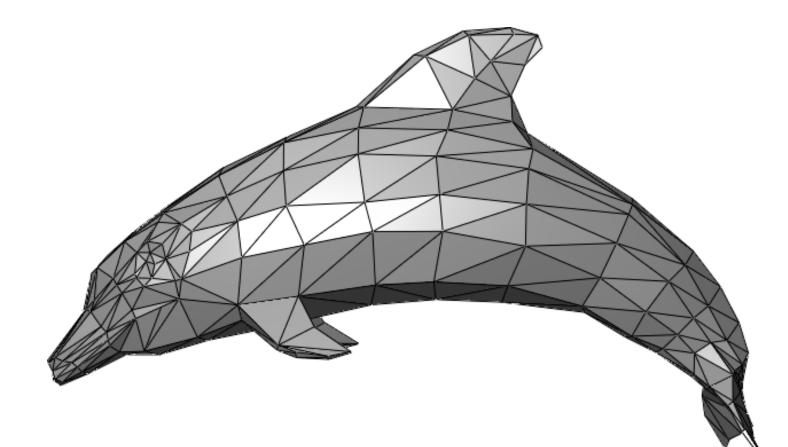






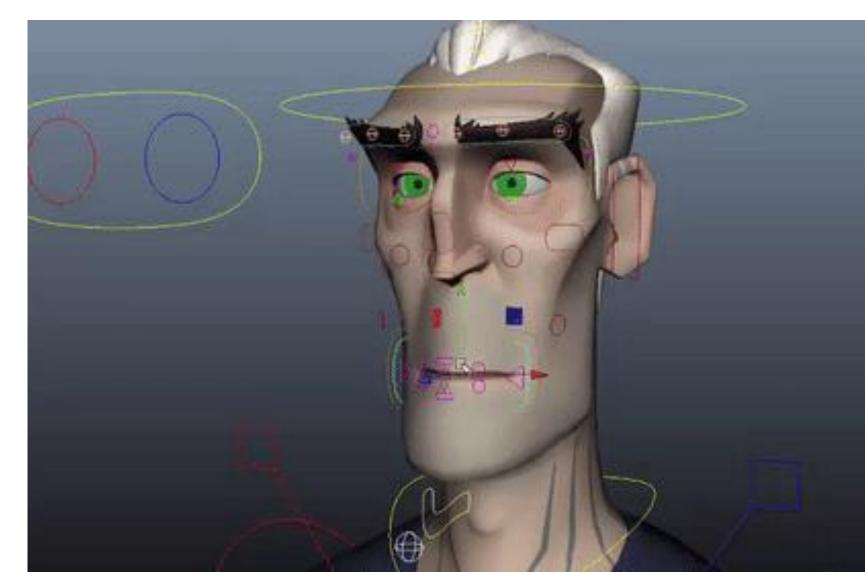
















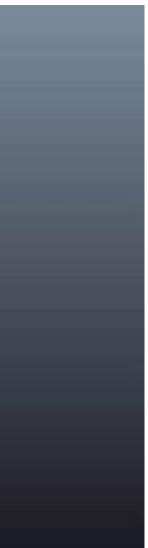
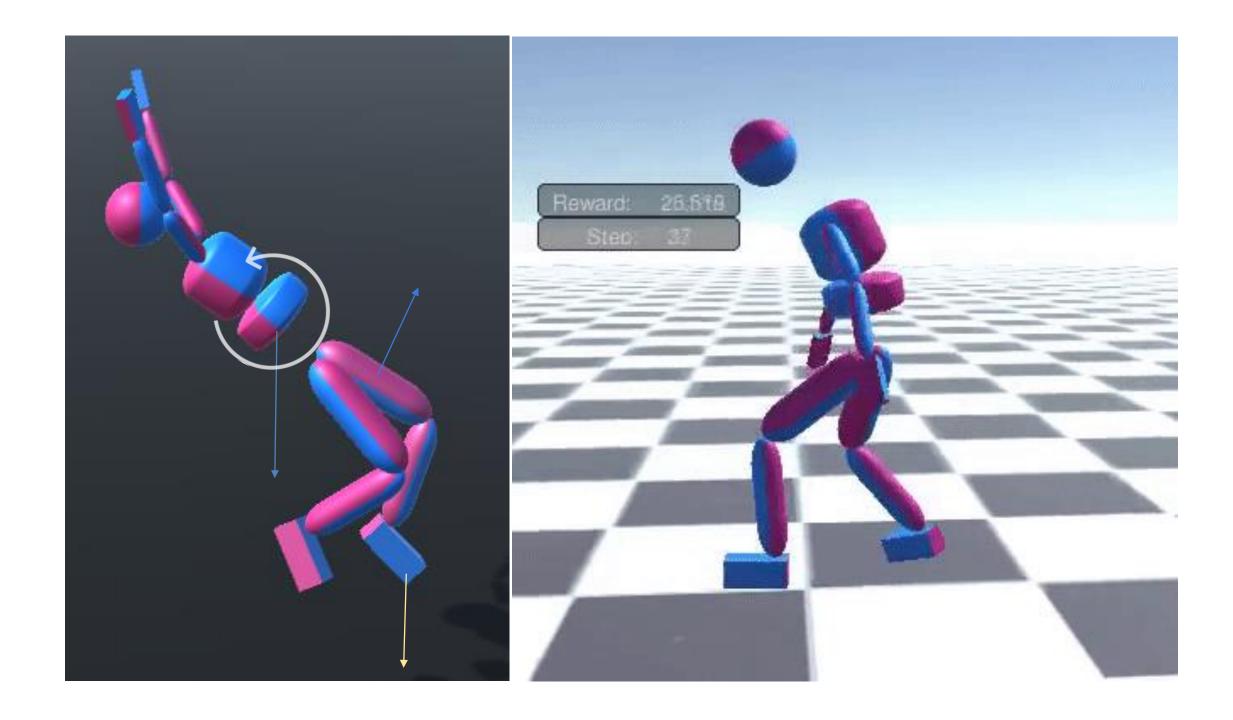


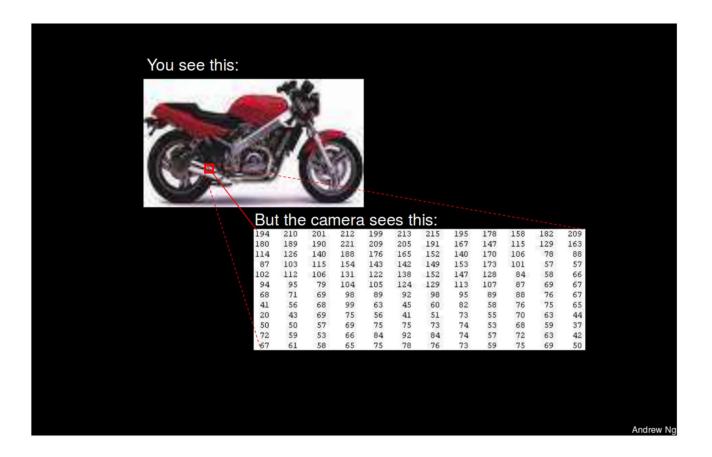
Image credit Animation Mentor







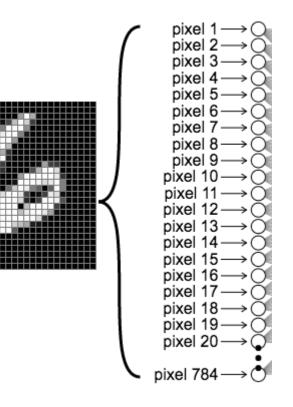
In the end... It's all numbers



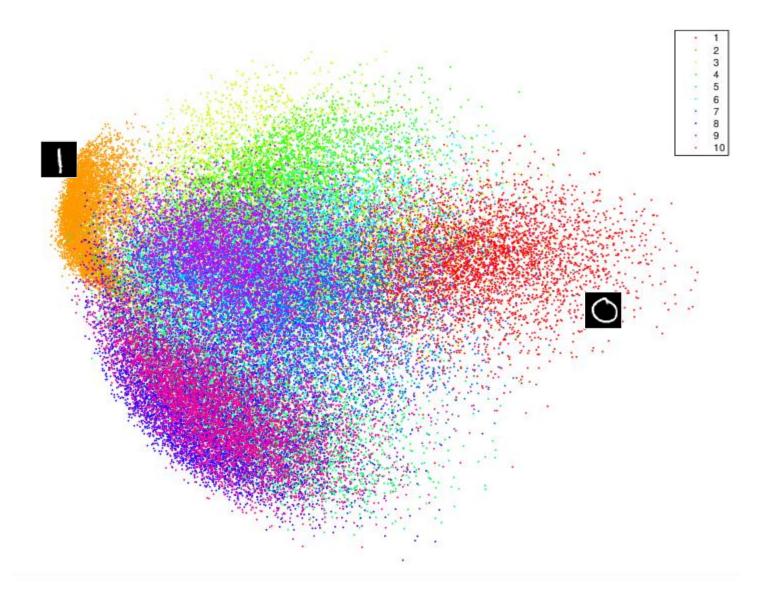
... in particular, *M*-dimensional vectors in $\mathcal{X} \subset \mathbb{R}^M$.







Data is far from random

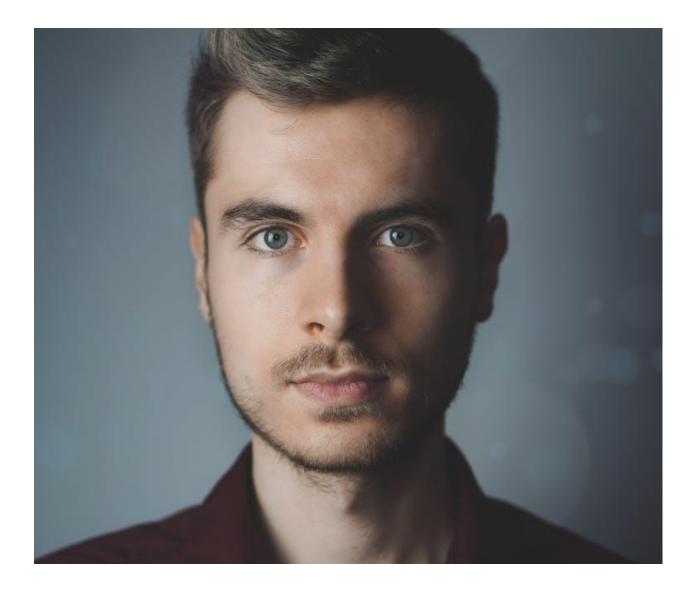






							,		1
		Z	2	3	2	2	2		æ
	3	3	3	3	S	3	3		3
	r	4	Ч	Ч	<i>4</i>	Ч	4	4	4
	5	5	5	5	ς	5	১		5
	6	6	6	6	4	6	6		6
1	7	7	1	7	7	7	7		7
1.00	К	8	8	8	8	8	8		ଷ୍ଠ
1	9	9	ら	9	9	9	9		9

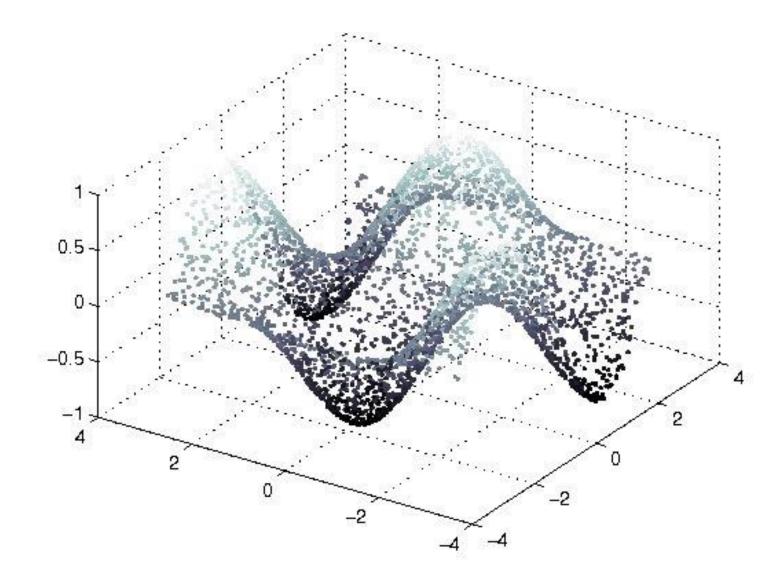
Do we need *M* pixels to represent a face?





M=1000000 pixels!

Data is not really M-dimensional

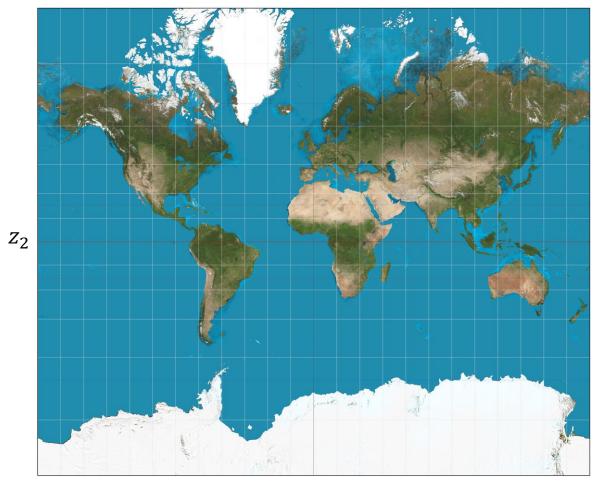


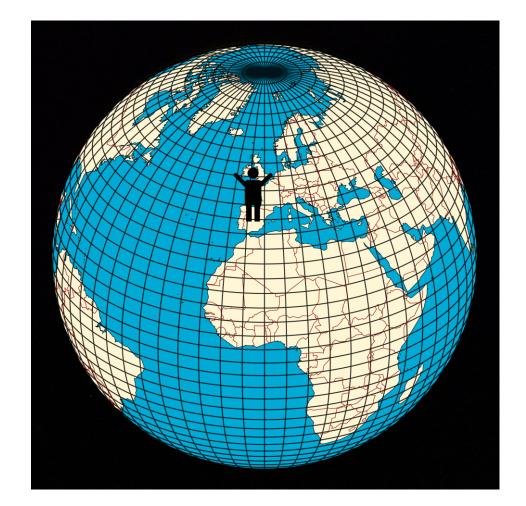




It rather lays on a lower dimensional *manifold*!

Manifold?





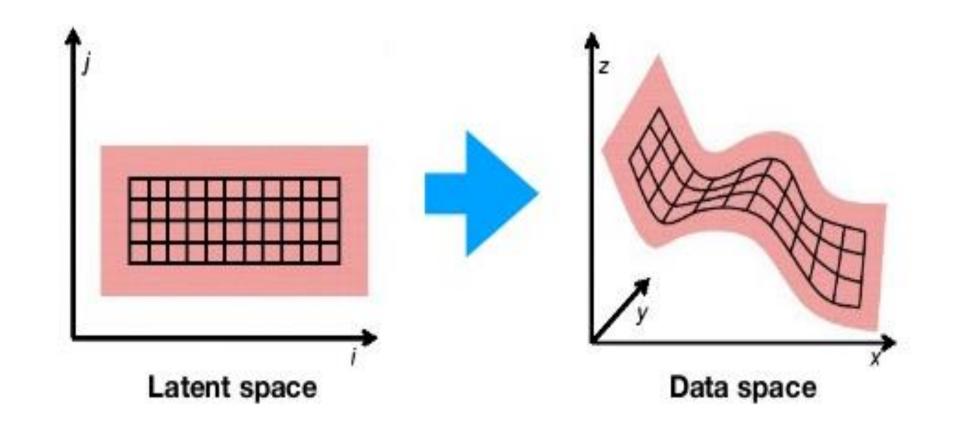
 Z_1

$$z = (z_1, z_2)$$



 $x = (x_1, x_2, x_3)$

Latent dimensions of data



***Spoiler**: Generative models learn both: the intrinsic geometry and the probability distribution!





Images credit Ward A.D. et al 2007

Ð Ð OВ Ð Ð E D

 Z_1



 Z_1

Auto-Encoding Variational Bayes. Kingma et al. 2013

 Z_2



n D s

 Z_2

A walk through the latent space



A Style-Based Generator Architecture for Generative Adversarial Networks. Karras et al. 2018 (NVIDIA)





And how do they work?



Random variables and generative modelling

For us, each datapoint x_i is just a *realization* of an underlying random variable.

$$\mathbf{x} \sim p(x)$$

- Unsupervised learning is the field which attempts to infer properties of x from samples.
- Generative modelling is a subset of unsupervised learning which attempts to <u>approximate</u> **x** as some parametrized combination of "simple" random variables which you can sample.

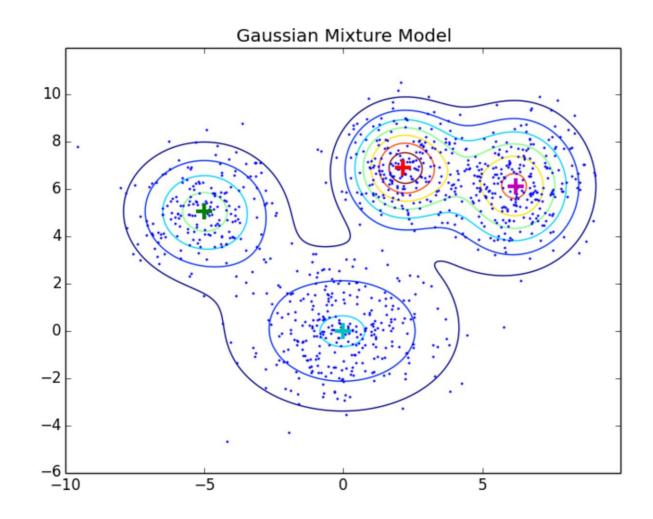
$$\mathbf{x} \approx f_{\theta}(\mathbf{z_1}, \mathbf{z_2}, \dots, \mathbf{z_K}) \triangleq \hat{\mathbf{x}}$$



GVWE DEVELODERS (COVERSION MARCH 18–22, 2019 | #GDC19

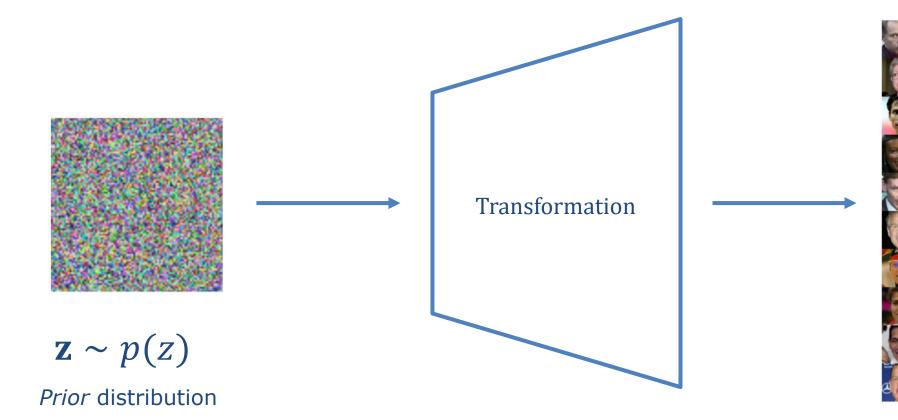
Example: Gaussian Mixtures

Here every \mathbf{z}_i is normal (gaussian) $\mathcal{N}(\mu_i, \Sigma_i)$ and the combination $f_{\theta}(\cdot)$ is a *mixture*.





Latent variable models





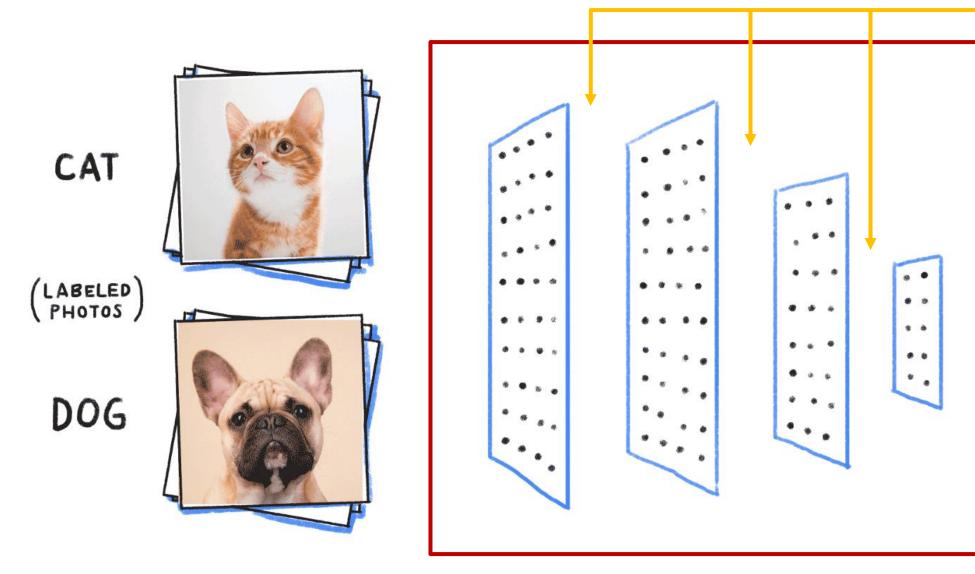




Â

Architectures: neural networks

GDC



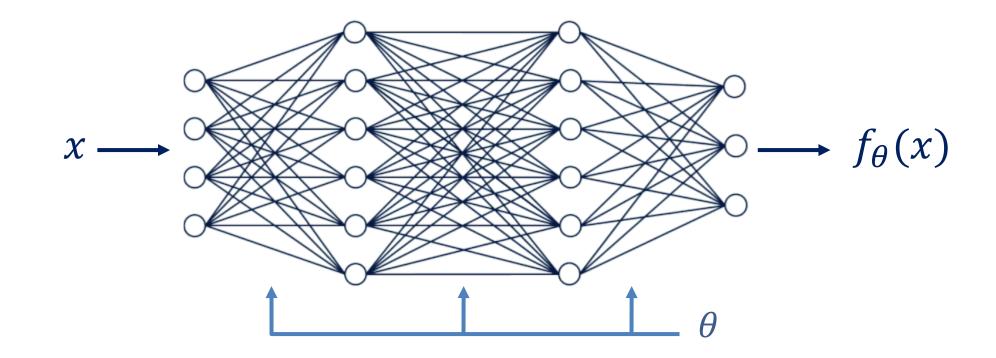
 f_{θ}



→ weights

OUTPUT

Architectures: neural networks



Neural networks can approximate any function f(x) to any precision! *

*G. Cybenko 1989, K. Hornik 1991, Z. Lu et al 2017, B. Hanin 2017





Image credit Sydney Firmin

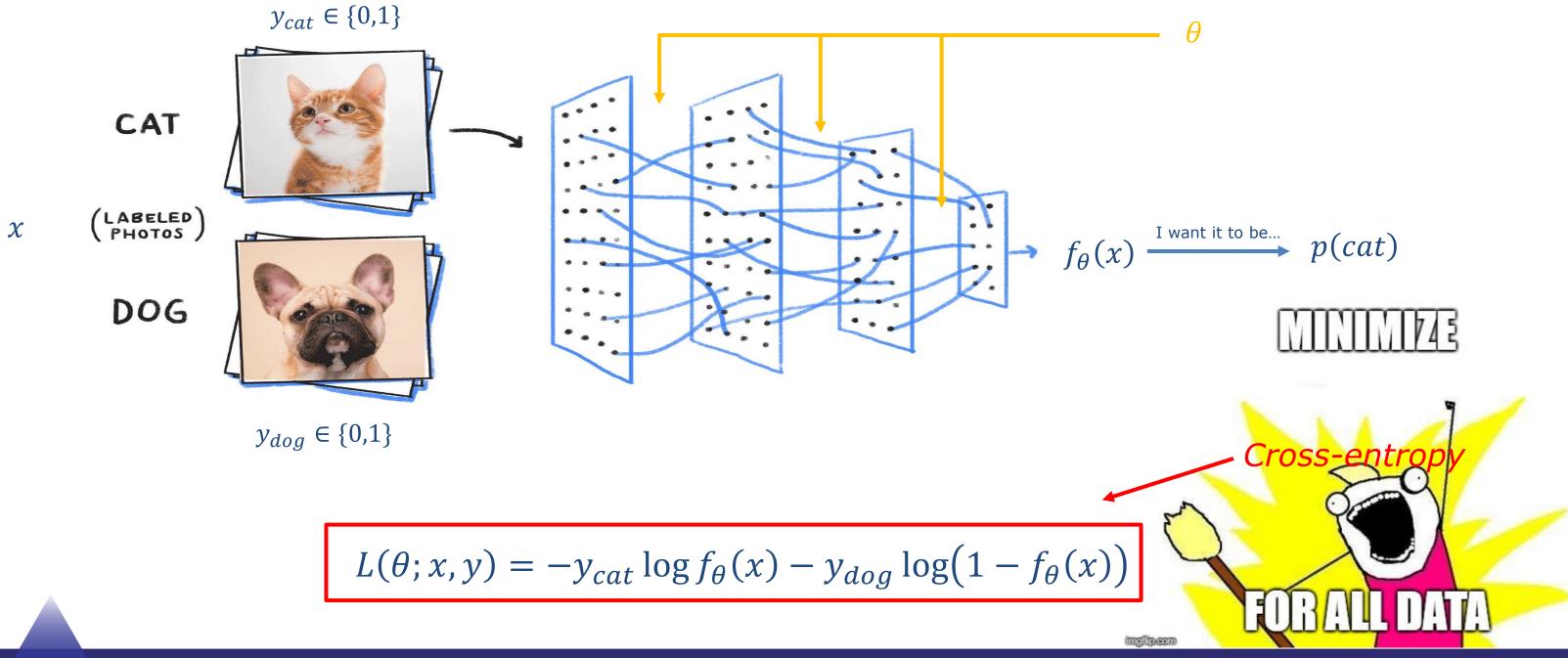
But to find the right θ (training)?

You optimize some loss (error) function!

 $\min_{\theta} L(\theta; data)$



E.g. Classifier







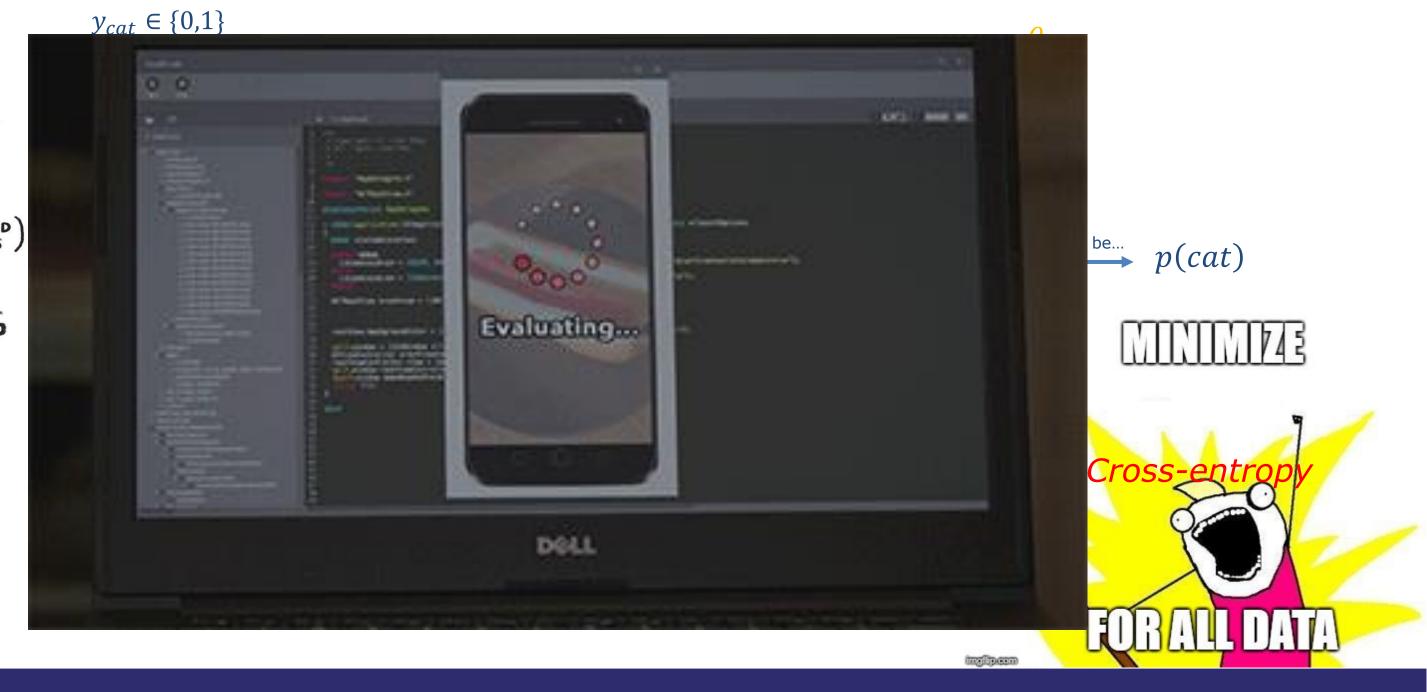
E.g. Classifier

CAT

(LABELED PHOTOS)

X

```
DOG
```

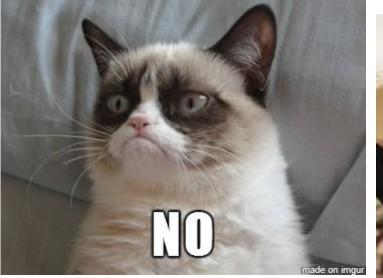


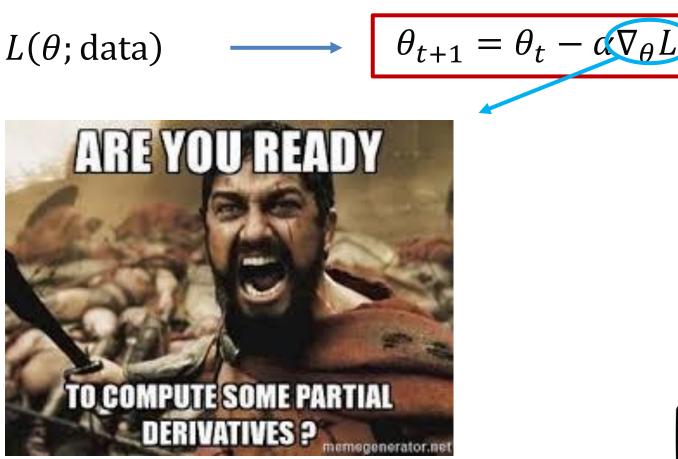


But to find the right θ (training)?

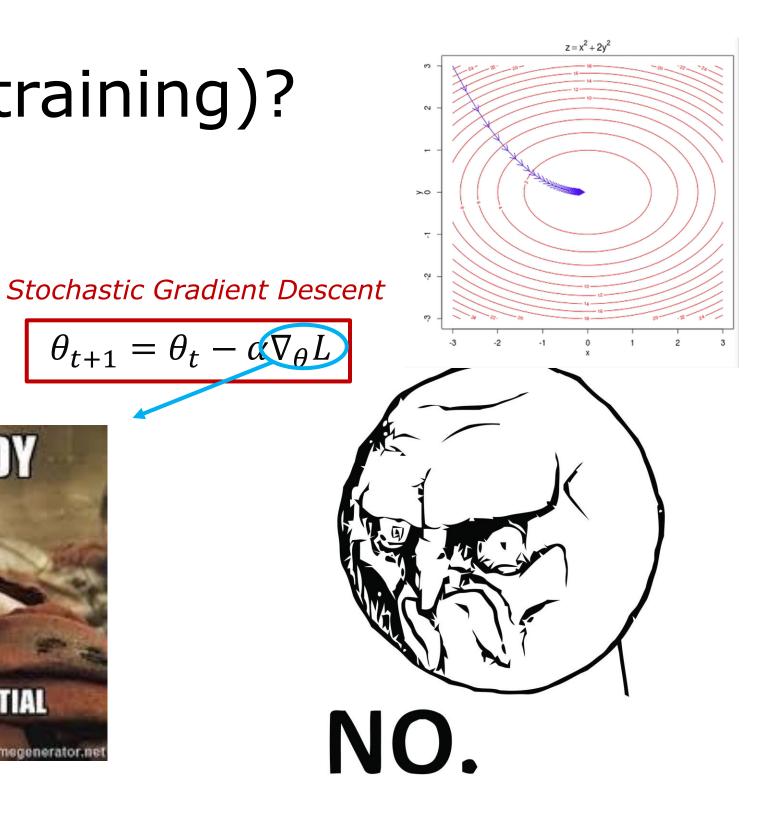
You optimize some loss (error) function!

 $\min_{\boldsymbol{\theta}} L(\boldsymbol{\theta}; data)$



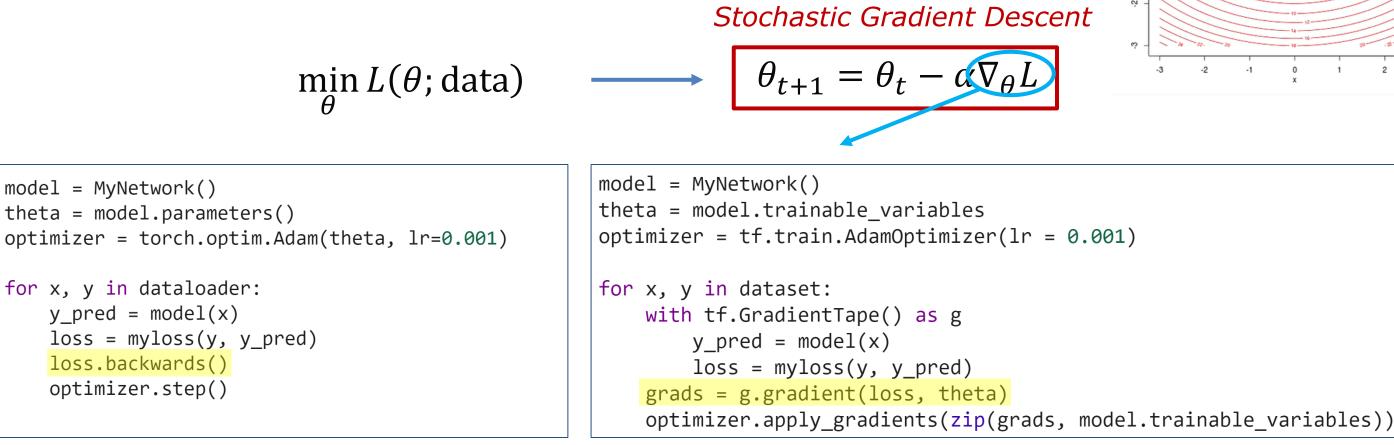






But to find the right θ (training)?

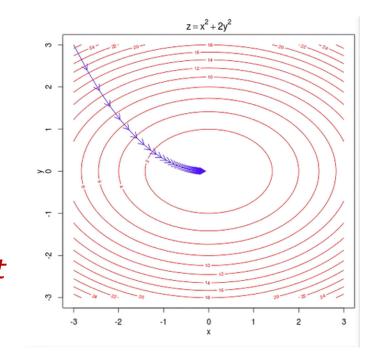
You optimize some loss (error) function!

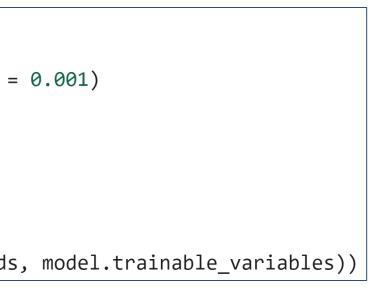


O PyTorch

TensorFlow

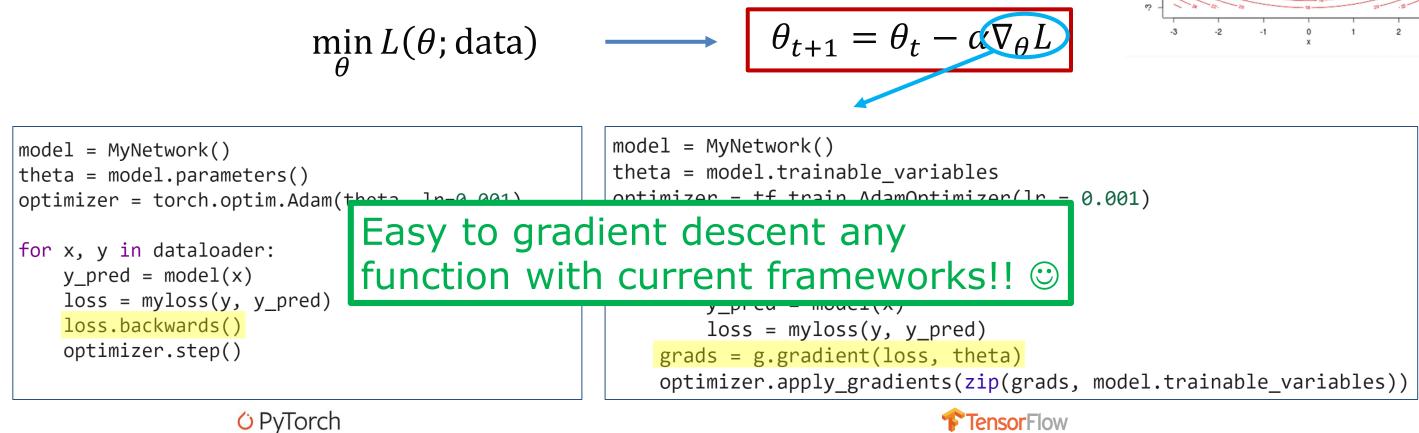




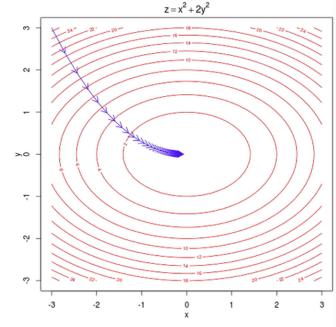


But to find the right θ (training)?

You optimize some loss (error) function!

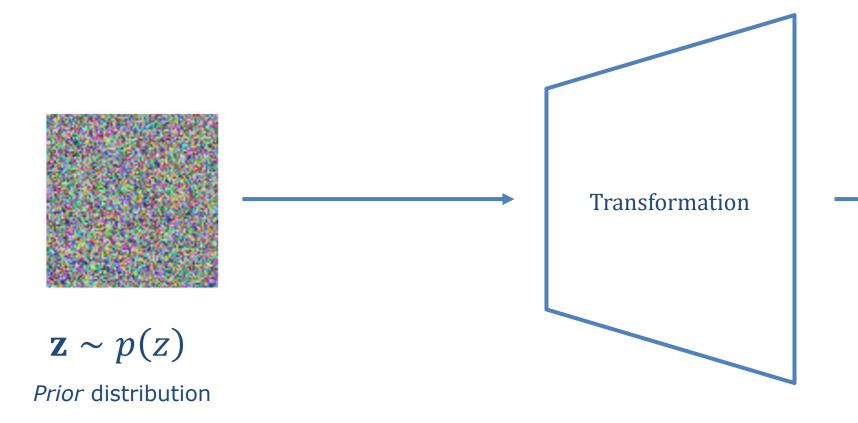






Stochastic Gradient Descent

Latent variable models

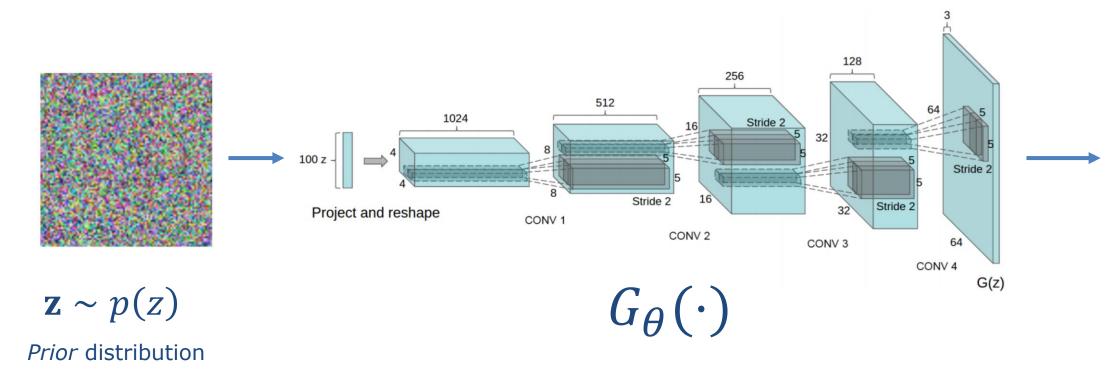






Â

Deep latent variable models



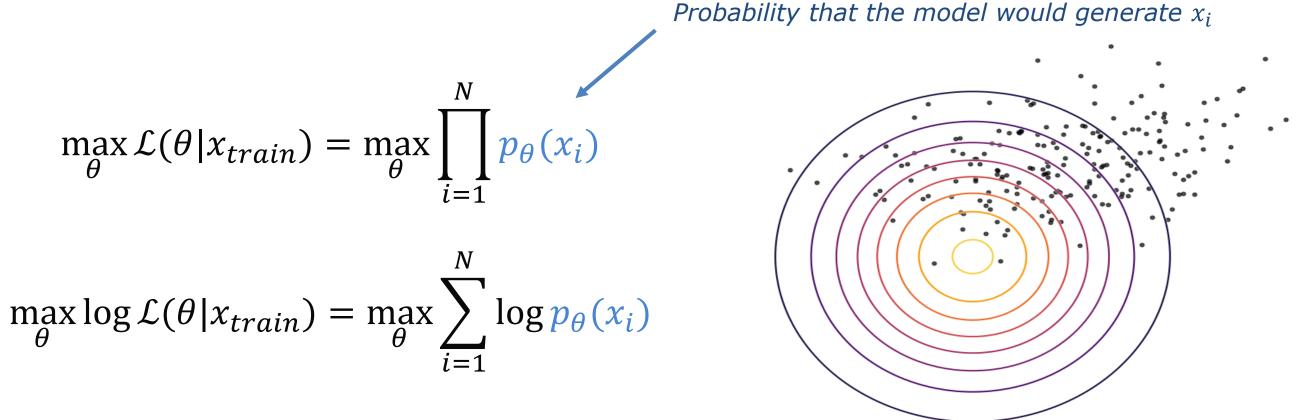




$\hat{\mathbf{x}} = G_{\theta}(\mathbf{z})$

Training

We want to approximate x as $\hat{\mathbf{x}} = G_{\theta}(\mathbf{z})$. How do we find the optimal θ ? Maximize the likelihood of the data!



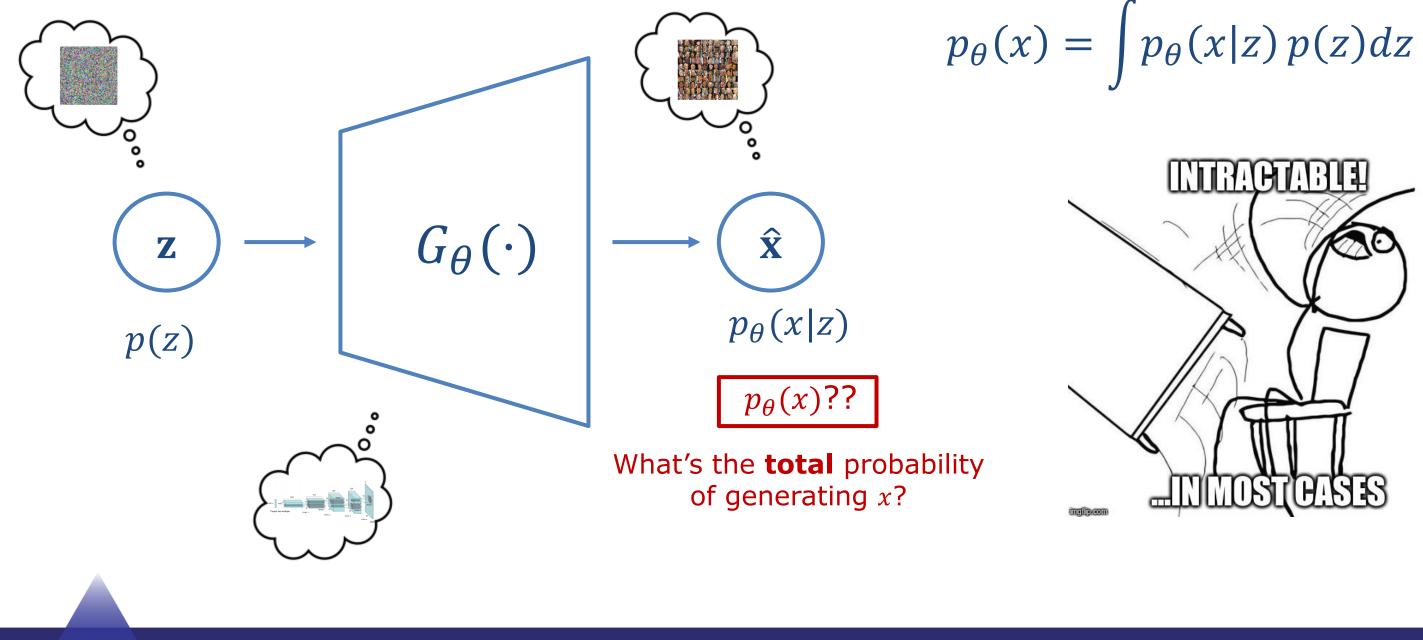
But... we need $p_{\theta}(x)$ explicitly!



Image credit: Colin Raffel

$(\mathbf{C}(\mathbf{O}))$ MARCH 18-22, 2019 | #GDC19

What about deep latent variable models?

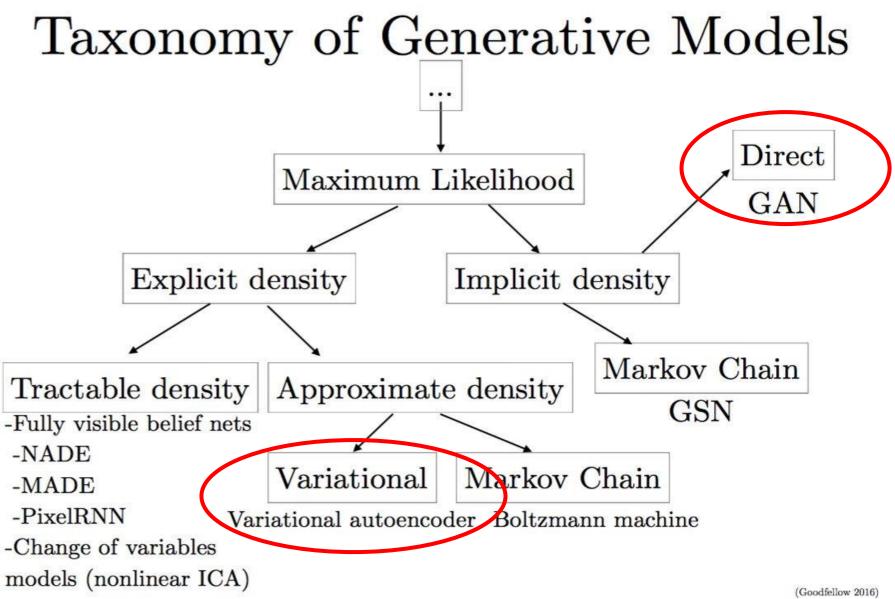




Different models – different methods

- 1. We have $p_{\theta}(\hat{x})$ explicitly: maximize the likelihood.
- 2. $p_{\theta}(\hat{x})$ is intractable: we can approximate it instead
 - Markov Chain Monte Carlo (MCMC) methods lacksquare
 - Variational methods (e.g. Variational Autoencoders)
- 3. We don't need $p_{\theta}(\hat{x})$; it's implicit.
 - Adversarial methods (e.g. GANs)



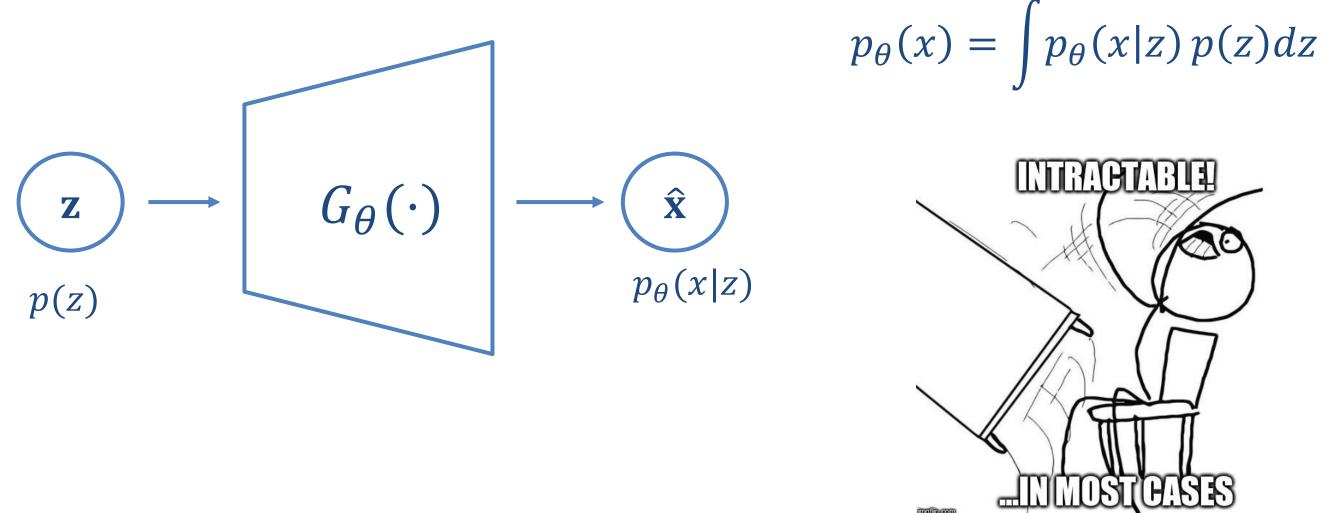






Goodfellow. 2016

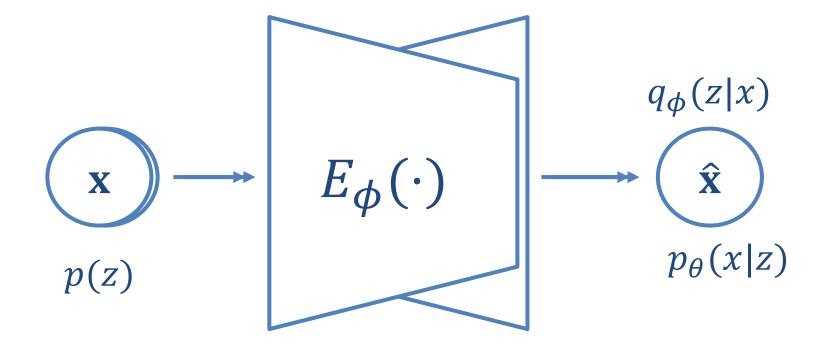
Variational autoencoder (VAE)







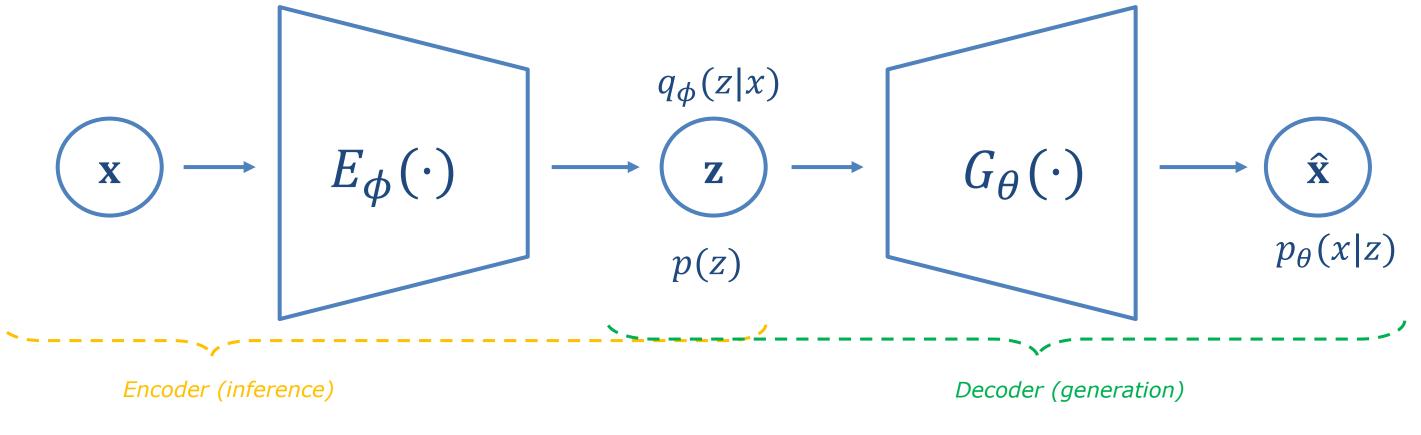
Variational autoencoder (VAE)







Variational autoencoder (VAE)



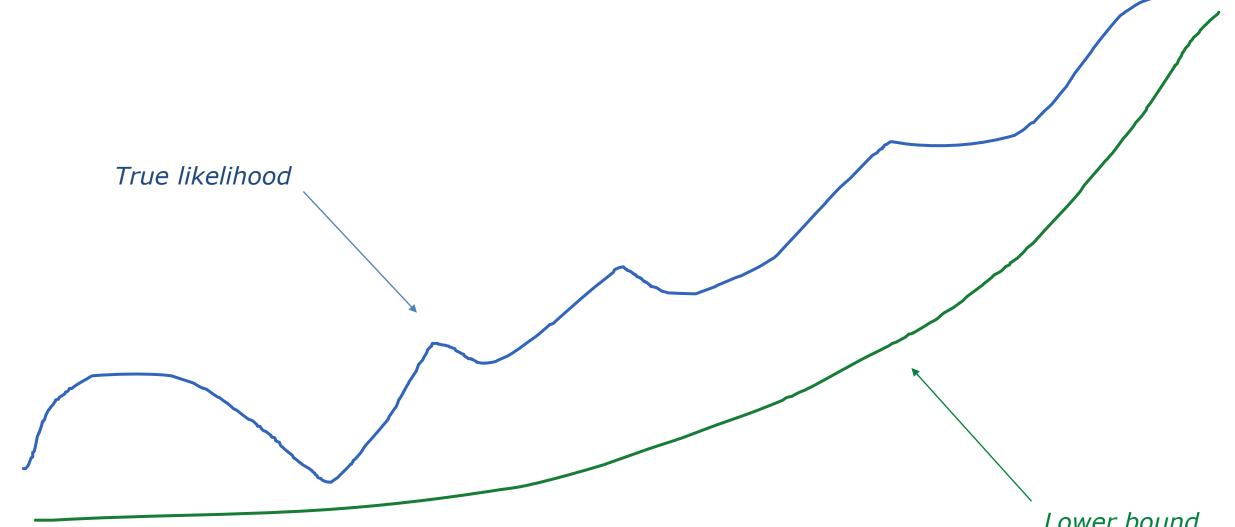
$$\log p_{\theta}(x) \ge \mathbb{E}_{z \sim q_{\phi}(z|x)}[\log p_{\theta}(x|z)] - \mathrm{KL}[q_{\phi}(z|z)]$$





 $|x) \parallel p(z)$

Maximize this instead!







GAME DEVELOPERS CONFERENCE MARCH 18-22, 2019 | #GDC19

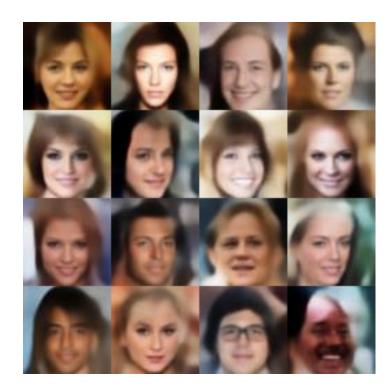
Iterations



Variational autoencoders

Pros:

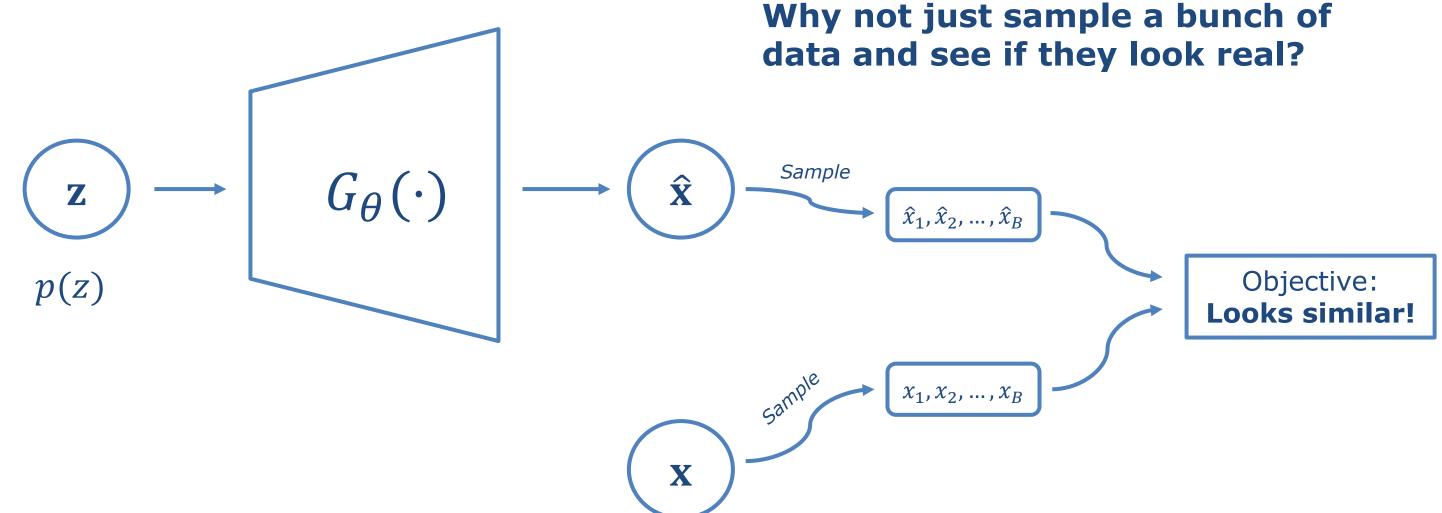
- Efficient inference for free!
 - Great tool for modelling the hidden structure of data. 0
- Stable to train.
- Good theoretical ground. Cons:
- Not very good samples.







Generative adversarial networks (GANs)





But... how do we measure similarity between groups of samples?



How to measure similarity of samples

One solution: train a <u>classifier</u> $D_{\phi}(x)$ to discriminate!

- If the classifier can not tell if a sample is real or fake, both distributions are close.
- Trained with the standard *cross-entropy loss*: lacksquare

$$\max_{\phi} L_d(\phi) = \max_{\phi} \left(\mathbb{E}_{x_r \sim p_{real}} \log \left(D_{\phi}(x_r) \right) + \mathbb{E}_{x_f \sim p_{fake}} \log \left(1 - \frac{1}{2} \right) \right)$$

It can be shown that the *optimal* classifier performance $L_d(\phi^*)$ is related to the *closeness* between both distributions (JS divergence).



 $D_{\phi}(x_f)))$

The GAN game

We want to minimize "closeness" between the generated and real samples, as measured by the discriminator loss:

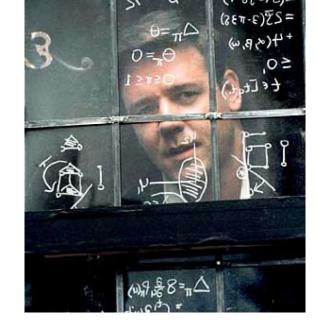


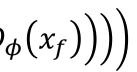
$$= \min_{\theta} \left(\max_{\phi} \left(\mathbb{E}_{x_r \sim p_{real}} \log \left(D_{\phi}(x_r) \right) + \mathbb{E}_{x_f \sim p_{fake}} \log \left(1 - D_{\phi} \right) \right) \right)$$

It's formally a two-player minimax game!!

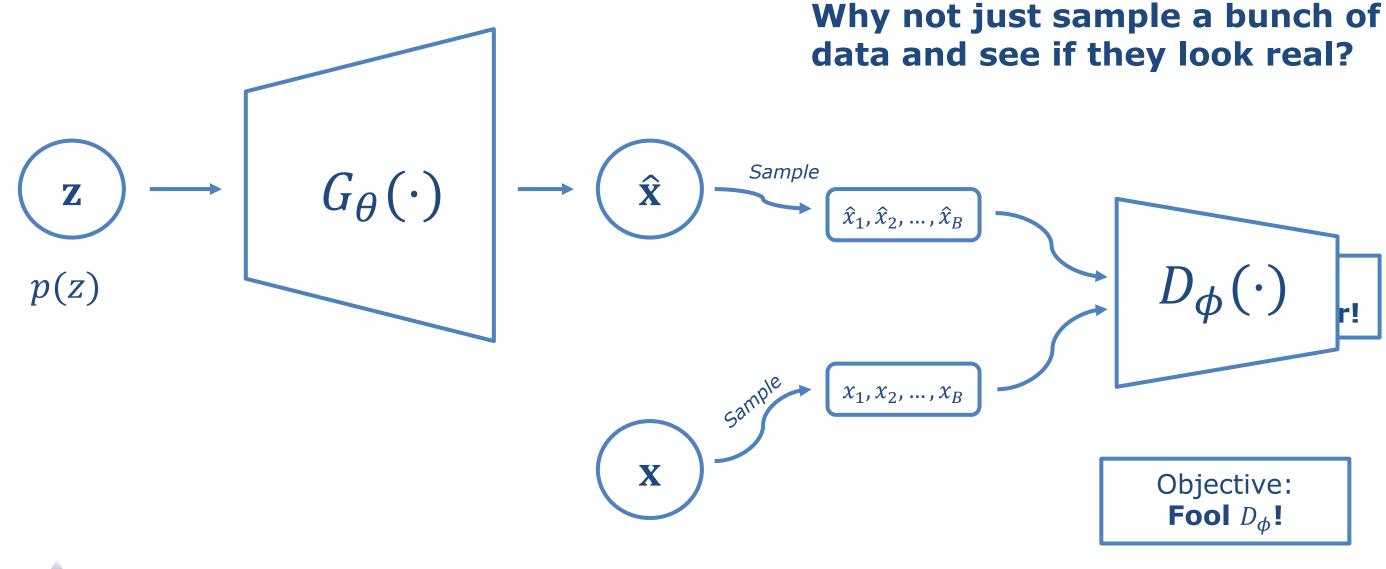
min "closeness"



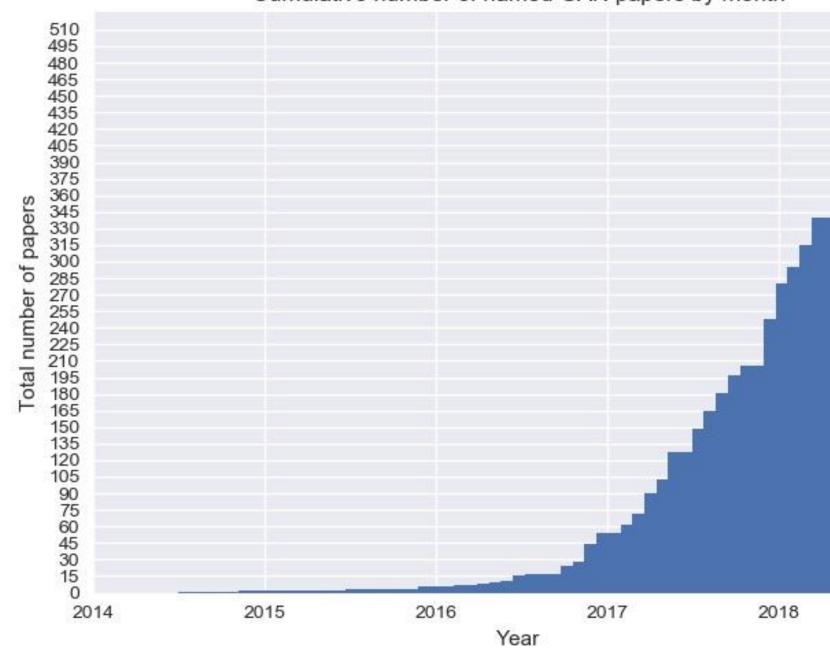




Generative adversarial networks



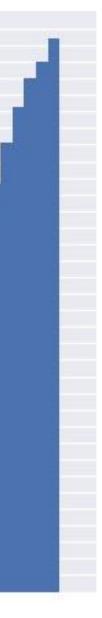




Cumulative number of named GAN papers by month







GANs

- Pros:
 - Awesome samples
- Cons:
 - Unstable training
 - No explicit probability density
 - No direct inference

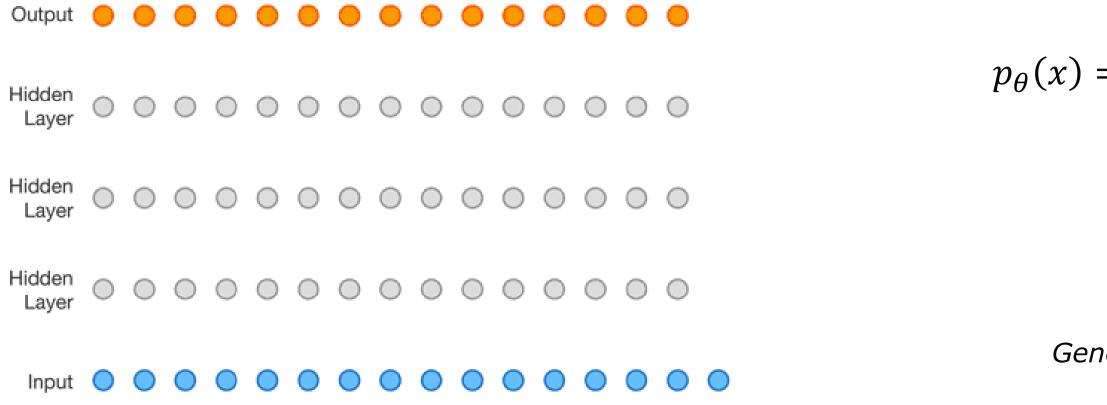






Large Scale GAN Training for High Fidelity Natural Image Synthesis. Brock et al. 2018

Bonus: autoregressive methods



Wavenet: A Generative Model for Raw Audio. Van den Oord et al. 2016.



$p_{\theta}(x) = \prod_{t=1}^{T} p_{\theta}(x_t | x_1, \dots, x_{t-1})$

Generate little by little!

OK!





OK! I can generate stuff.



OK! I can generate stuff.

But how do I influence what I generate?



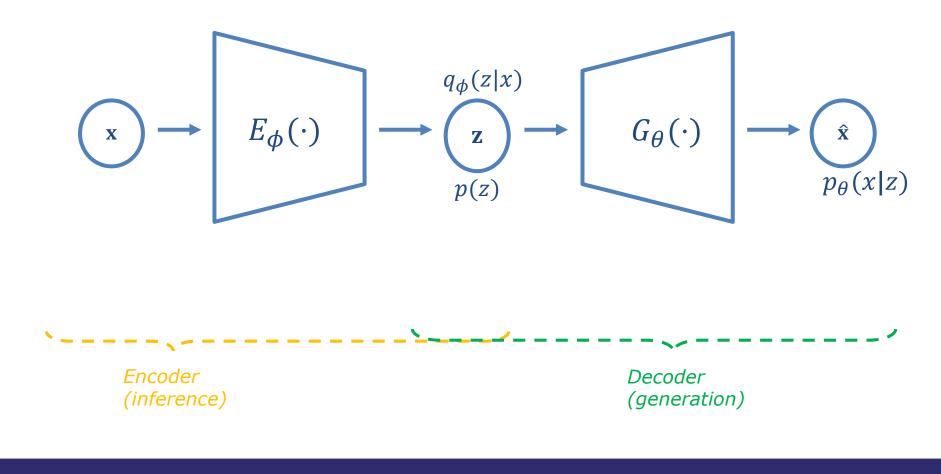
OK! I can generate stuff.

How do I remix existing stuff??



What if I have information \mathbf{c} to *condition* the generation/inference, e.g., class labels?

Just introduce them in the networks!

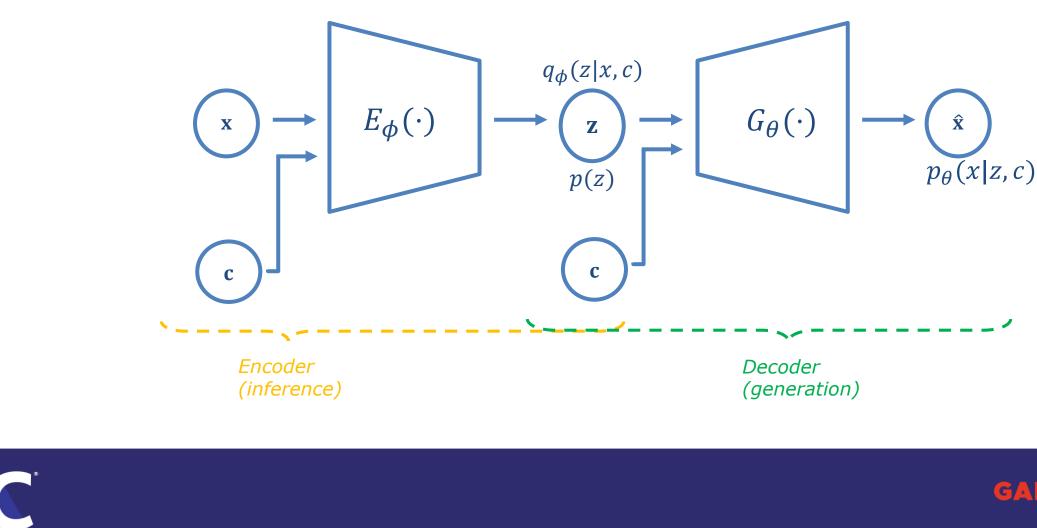




Variational Autoencoder

What if I have information \mathbf{c} to *condition* the generation/inference, e.g., class labels?

Just introduce them in the networks!

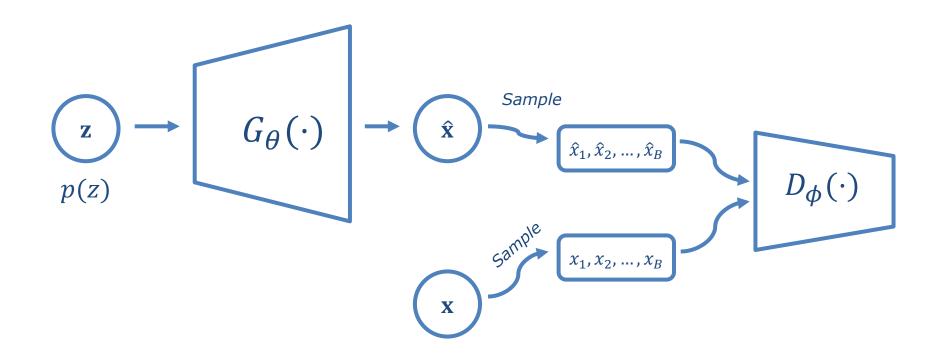


Conditional Variational Autoencoder

(q_____)= D=V=L0D=235 ((0)V==22=2(0= MARCH 18–22, 2019 | #GDC19

What if I have information c to *condition* the generation/inference, e.g., class labels?

Just introduce them in the networks!



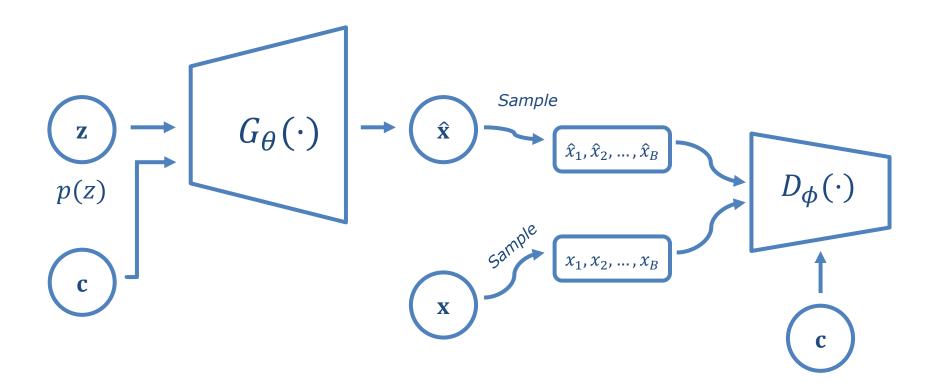


GAN

GAME DEVELOPERS MARCH 18–22, 2019 | #GDC19

What if I have information c to *condition* the generation/inference, e.g., class labels?

Just introduce them in the networks!

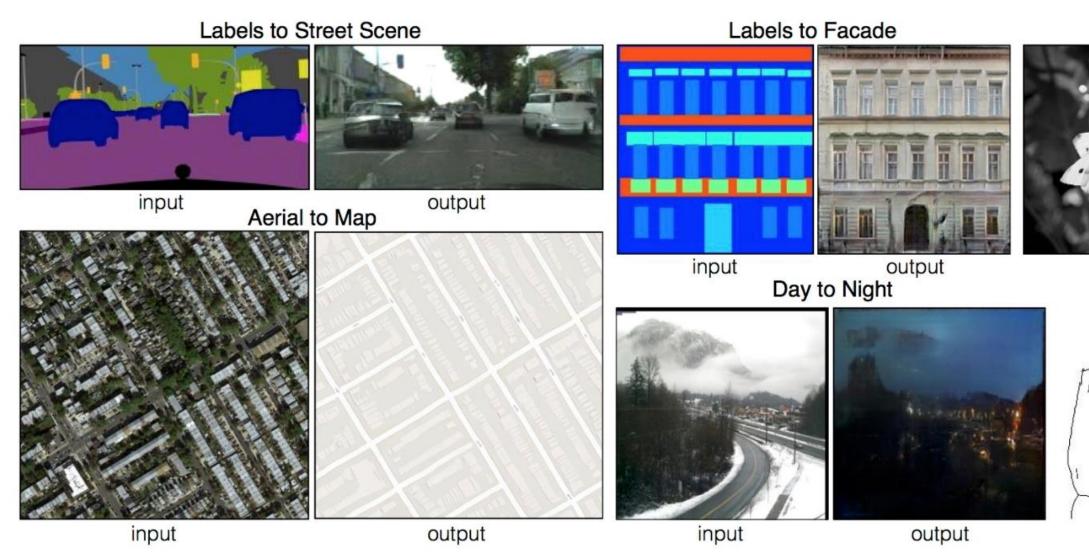




Conditional GAN

GAME DEVELOPERS (C(O))MARCH 18-22, 2019 | #GDC19

Conditional GMs are very important!



Pix2Pix: Image-to-Image Translation with Conditional Adversarial Networks. Isola et al.



BW to Color



input

output Edges to Photo



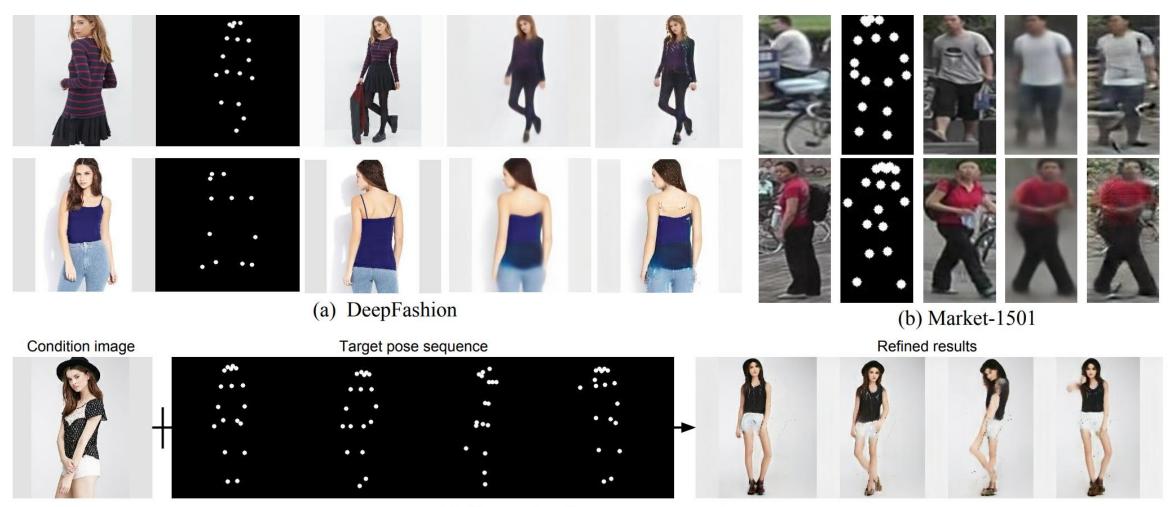


input

output

Conditional GMs are very important!

GDC



(c) Generating from a sequence of poses

Pose Guided Person Image Generation. Ma et al. 2017.





Some applications to game dev so far?



Generation of terrain



Interactive Example-Based Terrain Authoring with Conditional Generative Adversarial Networks. Guérin et al. 2017.



3D Content Generation





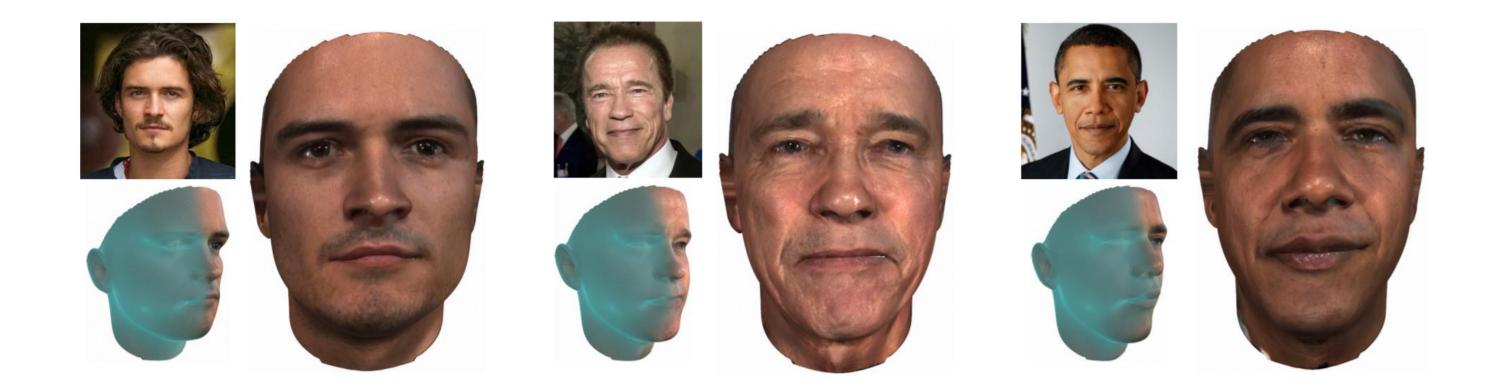
Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling. Wu et al. 2016

DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation. Park et al. 2019.





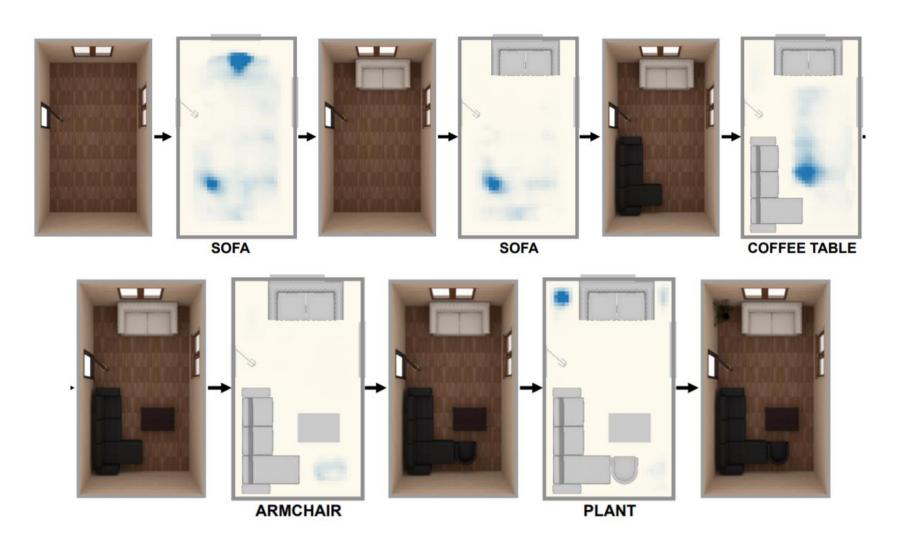
Face generation



GANFIT: Generative Adversarial Network Fitting for High Fidelity 3D Face Reconstruction. Gecer et al. 2019 (FaceSoft.io)



Procedural placement



Deep Convolutional Priors for Indoor Scene Synthesis. Wang et al. 2018.



Generation of behaviour policies



Variational Discriminator Bottleneck: Improving Imitation Learning, Inverse RL, and GANs by Constraining Information Flow. Peng et al. 2018.





Robustness

Generation of behaviour policies

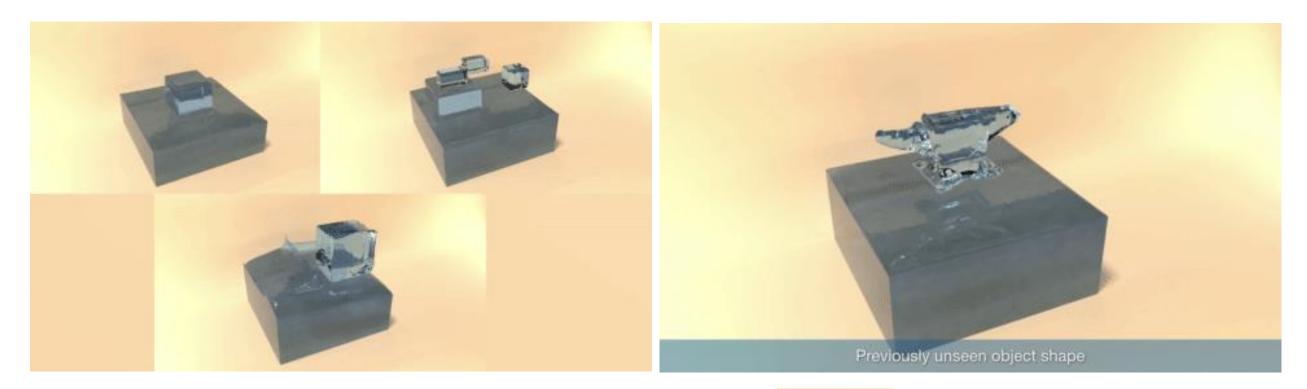


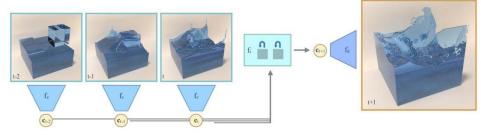
Imitation Learning with Concurrent Actions in 3D Games. Harmer et al. 2018 (SEED)





Learn and accelerate physics





Latent-space Physics: Towards Learning the Temporal Evolution of Fluid Flow. Wiewel et al. 2018





Thanks for inspiration and insightful conversations ③

Anastasia Opara Camilo Gordillo Colin Barré-Brisebois Hector Anadón León Henrik Johansson Jack Harmer Joakim Bergdahl Johan Andersson Kristoffer Sjöö Ken Brown

Linus Gisslen Martin Singh-Blom Mattias Teye Mark Cerny Mónica Villanueva Aylagas Magnus Nordin **Olivier Pomarez** Paul Greveson Roy Harvey Tomasz Stachowiak







SEED // SEARCH FOR EXTRAORDINARY EXPERIENCES DIVISION

STOCKHOLM – LOS ANGELES – MONTRÉAL – REMOTE

SEED.EA.COM

WE'RE HIRING!

Jorge del Val Santos

jdelvalsantos@ea.com

