

# Medical Image Generation and Analysis using Bayesian Generative Models

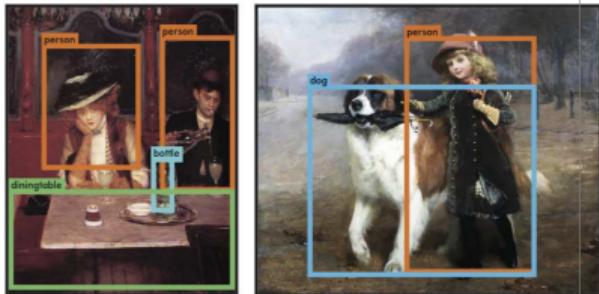
Răzvan V. Marinescu

Massachusetts Institute of Technology



# Machine Learning algorithms have achieved impressive milestones

Object detection (YOLO)



Text-to-Image Generation (DALL-E)



prompt: "an armchair in the shape of an avocado"

Image Generation (StyleGAN2)

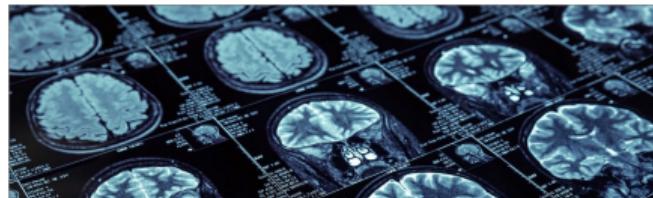


Text generation (GPT-3)

Title: United Methodists Agree to Historic Split  
Subtitle: Those who oppose gay marriage will form their own denomination  
Article: After two days of intense debate, the United Methodist Church has agreed to a historic split - one that is expected to end in the creation of a new denomination, one that will be "theologically and socially conservative," according to The Washington Post. The majority of delegates attending the church's annual General Conference in May voted to strengthen a ban on the ordination of LGBTQ clergy and to write new rules that will "discipline" clergy who officiate at same-sex weddings. But those who opposed these measures have a new plan: They say they will form a separate denomination by 2020, calling their church the Christian Methodist denomination.

# Maching Learning holds great promise for improving healthcare

## Diagnose with unprecedented accuracy



## Augment doctors



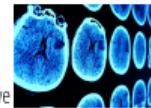
### Top 12 Ways Artificial Intelligence Will Impact Healthcare

Artificial intelligence is poised to become a transformational force in healthcare. How will providers and patients benefit from the impact of AI-driven tools?



### How Artificial Intelligence Improves Medical Imaging in Hospitals

Deep learning software, such as artificial intelligence, can improve



## Improve patient healthcare and save lifes



However, such milestones have not been translated to medical applications

## Prediction of clinical variables not always working

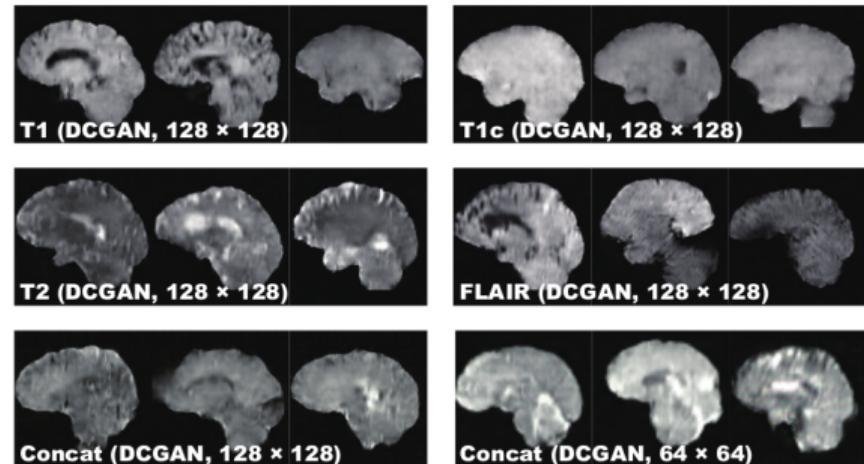
No algorithm/33 could predict cognitive scores in Alzheimer's (TADPOLE Challenge, Marinescu 2020)



- Forecasts were very good for clinical diagnosis and ventricle volume -- on the other hand, predicting ADAS turned out to be very difficult -- no team was able to regenerate forecasts that were significantly better than random guessing

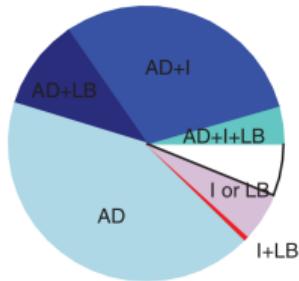
## Generated images are crude, not high-resolution, mostly 2D

Brain MRI generation (Han, 2018)



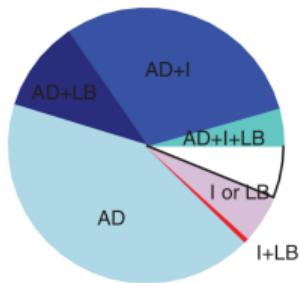
## Lack of good labels

- Alzheimer's diagnosis accuracy just 42%

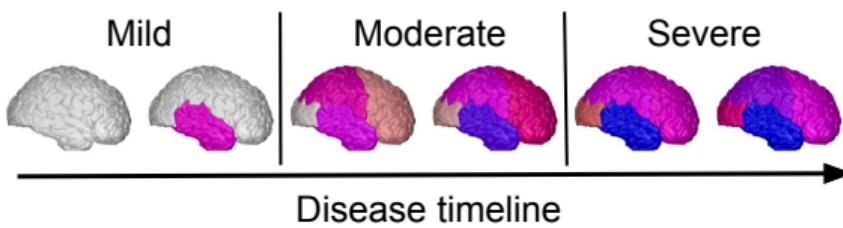


## Lack of good labels

- ▶ Alzheimer's diagnosis accuracy just 42%



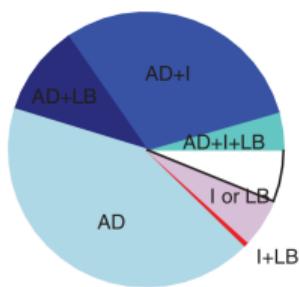
- ▶ Labels are categorical instead of continuous



# Why are Machine Learning models not working on medical applications?

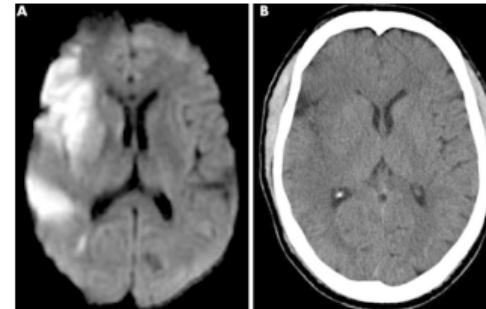
## Lack of good labels

- ▶ Alzheimer's diagnosis accuracy just 42%

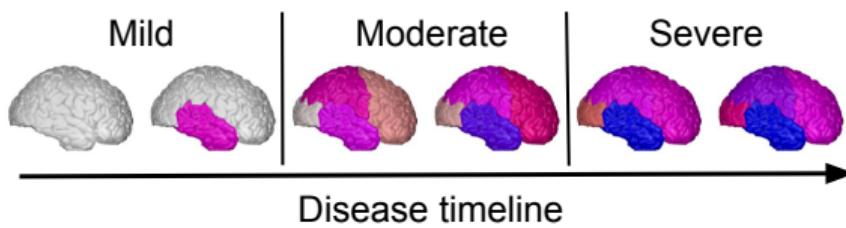


## Lack of good input data/signal

- ▶ Limited contrast



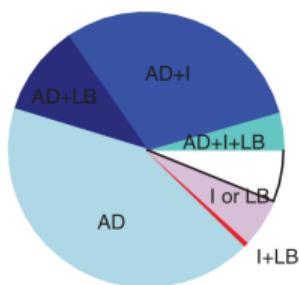
- ▶ Labels are categorical instead of continuous



# Why are Machine Learning models not working on medical applications?

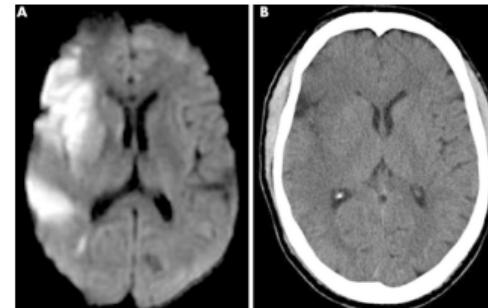
## Lack of good labels

- Alzheimer's diagnosis accuracy just 42%



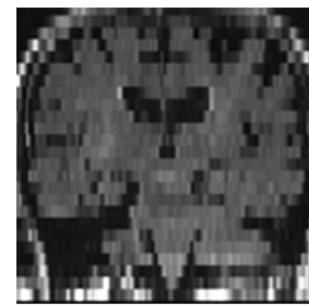
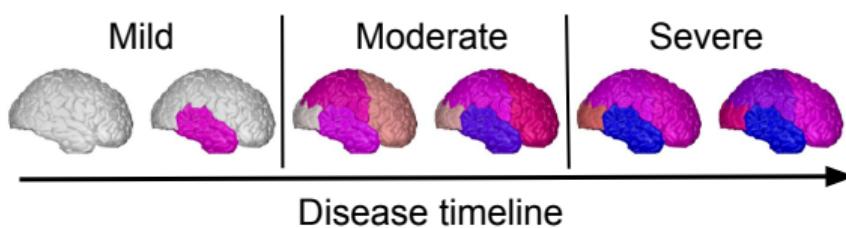
## Lack of good input data/signal

- Limited contrast



- Low-resolution

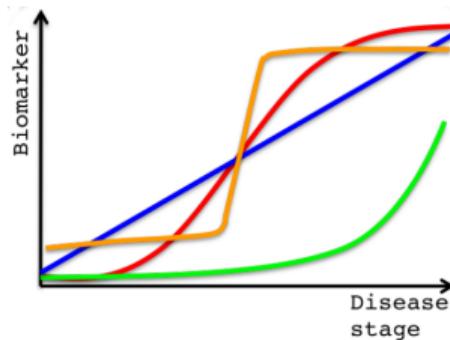
- Labels are categorical instead of continuous



What can we do?

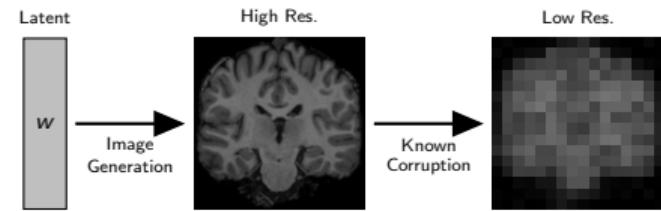
## Lack of good labels

Solution: Unsupervised Learning of Continuous Dynamics  
= Disease Progression Modelling



## Lack of good input data/signal

Solution: Image Reconstruction using Deep Generative Models

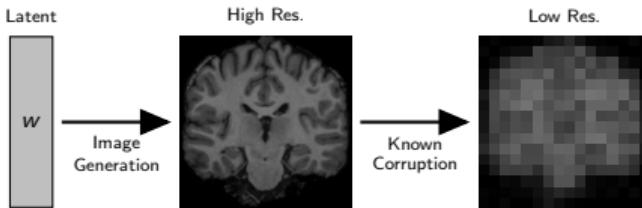


## 1. Disease progression modelling of Alzheimer's disease

### 1.1 Towards unsupervised clustering of biomarker trajectories



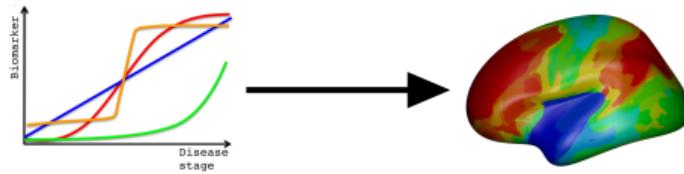
## 2. Image Reconstruction using Deep Generative Models



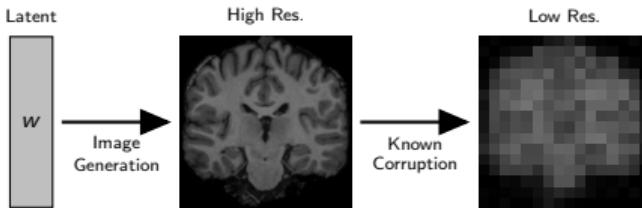
## 3. Future work

## 1. Disease progression modelling of Alzheimer's disease

### 1.1 Towards unsupervised clustering of biomarker trajectories



## 2. Image Reconstruction using Deep Generative Models



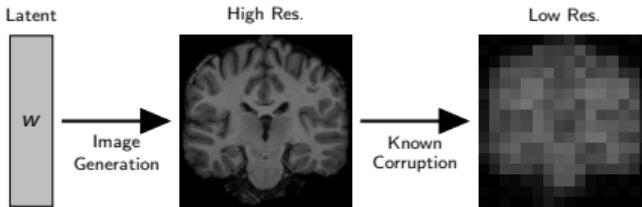
## 3. Future work

## 1. Disease progression modelling of Alzheimer's disease

### 1.1 Towards unsupervised clustering of biomarker trajectories



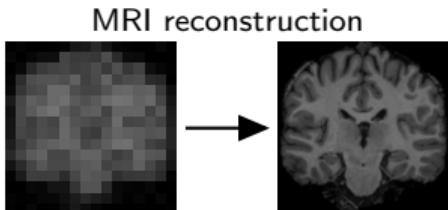
## 2. Image Reconstruction using Deep Generative Models



## 3. Future work

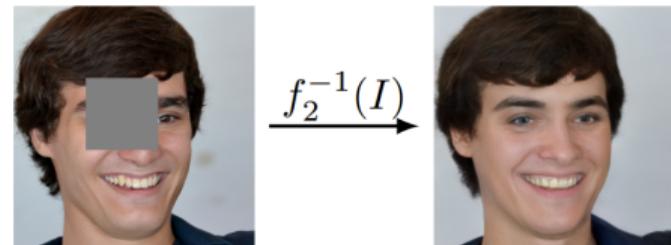
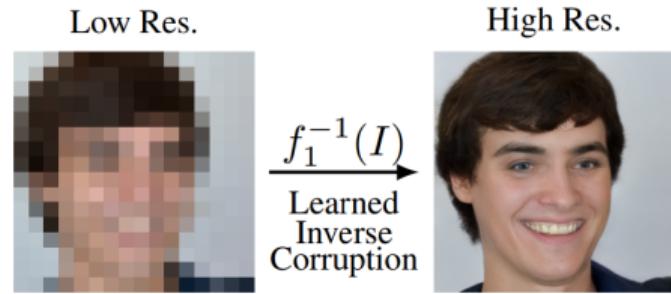
Aim: image reconstruction using **\*pre-trained\*** generator models

- Adapt the state-of-the-art StyleGAN2 for medical image reconstruction



## Current image reconstruction methods have several limitations

- ▶ Require large computational resources and data
- ▶ Are specific to particular corruption tasks
- ▶ Cannot deal with distribution shifts:
  - ▶ in inputs: e.g. older populations
  - ▶ in corruption type: e.g. change in blur kernel



## Limitation 1: State-of-the-art DL methods have large computational requirements

- Requirements = Computation Time + Advanced Hardware + Large Datasets
- Most computation now runs on clouds

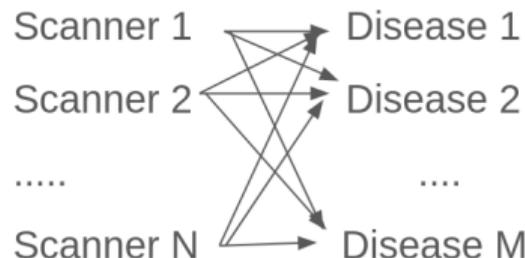


- Currently few labs/companies have the resources to train state-of-the-art models
  - StyleGAN2: 9 days on 4 GPUs
  - GPT-3: 355 years on single GPU
- Solutions moving forward:
  - Adapting previously-trained models
  - Combine smaller models into larger ones

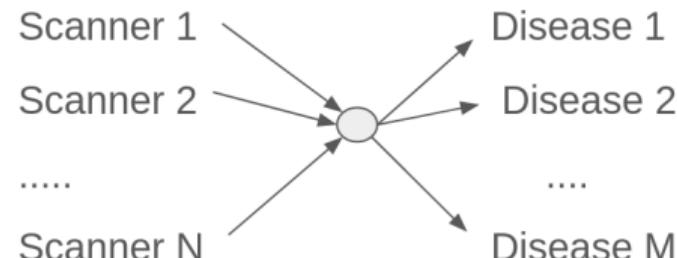
## Limitation 2: Distribution shifts require model re-training

- ▶ Distribution shifts happen all the time:
  - ▶ Changes in hospital scanners, protocols, software upgrades
  - ▶ Can be continuous: population getting older due to better healthcare
- ▶ Shifts can result in combinatorial effects in number of re-training instances!
- ▶ Compositionality is one potential solution

Without compositionality: **N x M**

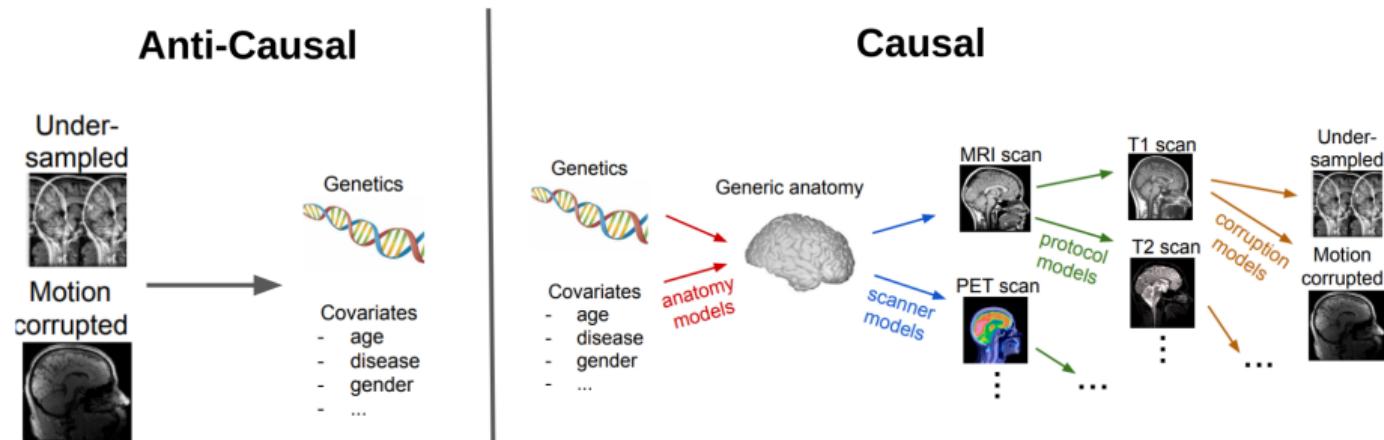


With compositionality: **N + M**



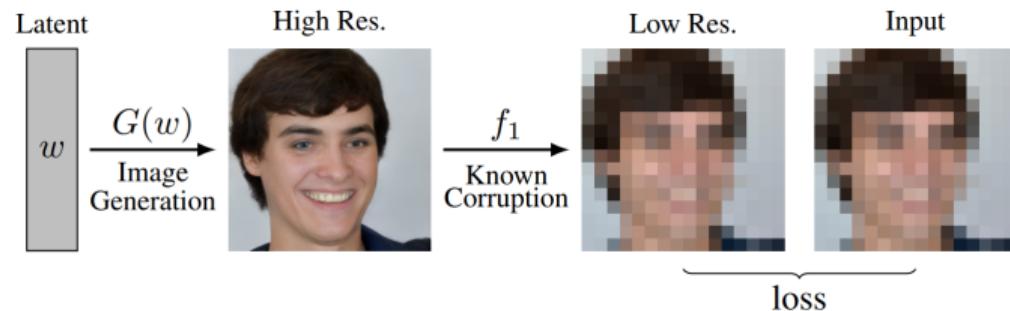
## Limitation 3: Models are anti-causal

- ▶ Existing model don't follow the data-generation process
  - ▶ Discriminative modelling easier than generative
- ▶ Causal modelling is the **right solution** to deal with distribution shifts



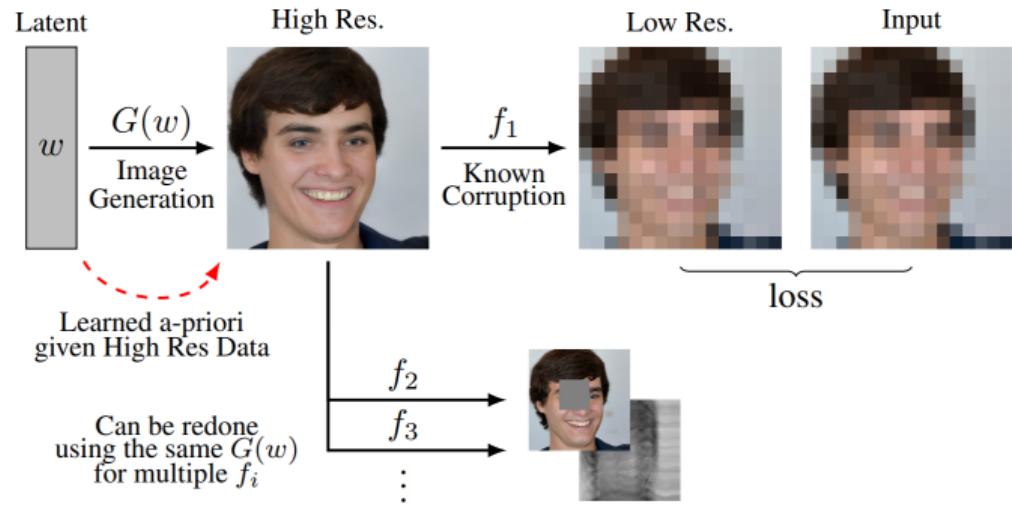
Method: We perform image reconstruction by combining two models

1. a pre-trained generator  $G$  (StyleGAN2)
2. a known forward corruption model  $f_1$



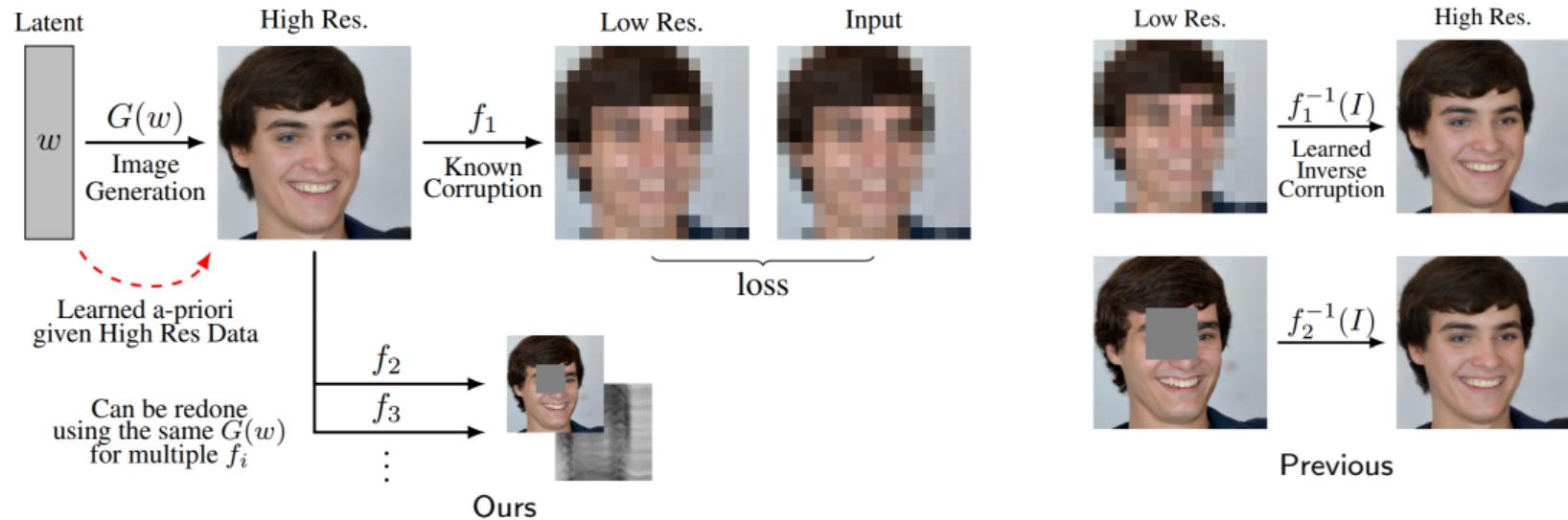
Method: We perform image reconstruction by combining two models

1. a pre-trained generator  $G$  (StyleGAN2)
2. a known forward corruption model  $f_1$



Method: We perform image reconstruction by combining two models

1. a pre-trained generator  $G$  (StyleGAN2)
2. a known forward corruption model  $f_1$



Reconstructed image is given by computing the Bayesian maximum a-posteriori (MAP) estimate

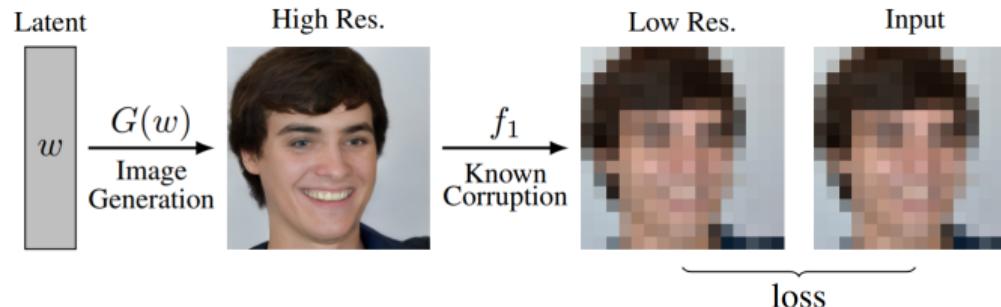
- We optimise:

$$w^* = \arg \max_w p(w)p(I|w)$$

- For uninformative prior  $p(w)$  and Gaussian noise model (pixelwise independent), we get:

$$w^* = \arg \min_w \|I - f \circ G(w)\|_2^2$$

- This can be optimised with SGD
- Once we get  $w^*$ , the the reconstructed image is  $G(w^*)$



## Good reconstructions require further modifications

- We started from the original StyleGAN2 inversion

- ▶ We started from the original StyleGAN2 inversion
- ▶ Yet the reconstruction was not good → required several changes

$$w^*, \eta^* = \arg \min_{w, \eta} \|\phi(I) - \phi \circ f \circ G(w, \eta)\|_2^2$$

- ▶ We started from the original StyleGAN2 inversion
- ▶ Yet the reconstruction was not good → required several changes
  - ▶ remove noise layers

$$w^* = \arg \min_w \|\phi(I) - \phi \circ f \circ G(w)\|_2^2$$

- We started from the original StyleGAN2 inversion
- Yet the reconstruction was not good → required several changes
  - optimize latents at all resolutions

$$\mathbf{w} = w_1, \dots, w_L$$

$$\mathbf{w}^* = \arg \min_{\mathbf{w}} \|\phi(I) - \phi \circ f \circ G(\mathbf{w})\|_2^2$$

- ▶ We started from the original StyleGAN2 inversion
- ▶ Yet the reconstruction was not good → required several changes
  - ▶ add pixelwise loss

$$\mathbf{w}^* = \arg \min_{\mathbf{w}} \|\phi(I) - \phi \circ f \circ G(\mathbf{w})\|_2^2 + \|I - f \circ G(\mathbf{w})\|_2^2$$

- We started from the original StyleGAN2 inversion
- Yet the reconstruction was not good → required several changes
  - gaussian prior on latents

$$\begin{aligned}\mathbf{w}^* = \arg \min_{\mathbf{w}} & ||\phi(I) - \phi \circ f \circ G(\mathbf{w})||_2^2 + ||I - f \circ G(\mathbf{w})||_2^2 + \\ & + \sum_i \left( \frac{w_i - \mu}{\sigma_i} \right)^2\end{aligned}$$

- We started from the original StyleGAN2 inversion
- Yet the reconstruction was not good → required several changes
  - force latents to be colinear

$$\begin{aligned}\mathbf{w}^* = \arg \min_{\mathbf{w}} & ||\phi(I) - \phi \circ f \circ G(\mathbf{w})||_2^2 + ||I - f \circ G(\mathbf{w})||_2^2 + \\ & + \sum_i \left( \frac{\mathbf{w}_i - \mu}{\sigma_i} \right)^2 - \sum_{i,j} \frac{\mathbf{w}_i \mathbf{w}_j^T}{|\mathbf{w}_i| |\mathbf{w}_j|}\end{aligned}$$

- We started from the original StyleGAN2 inversion

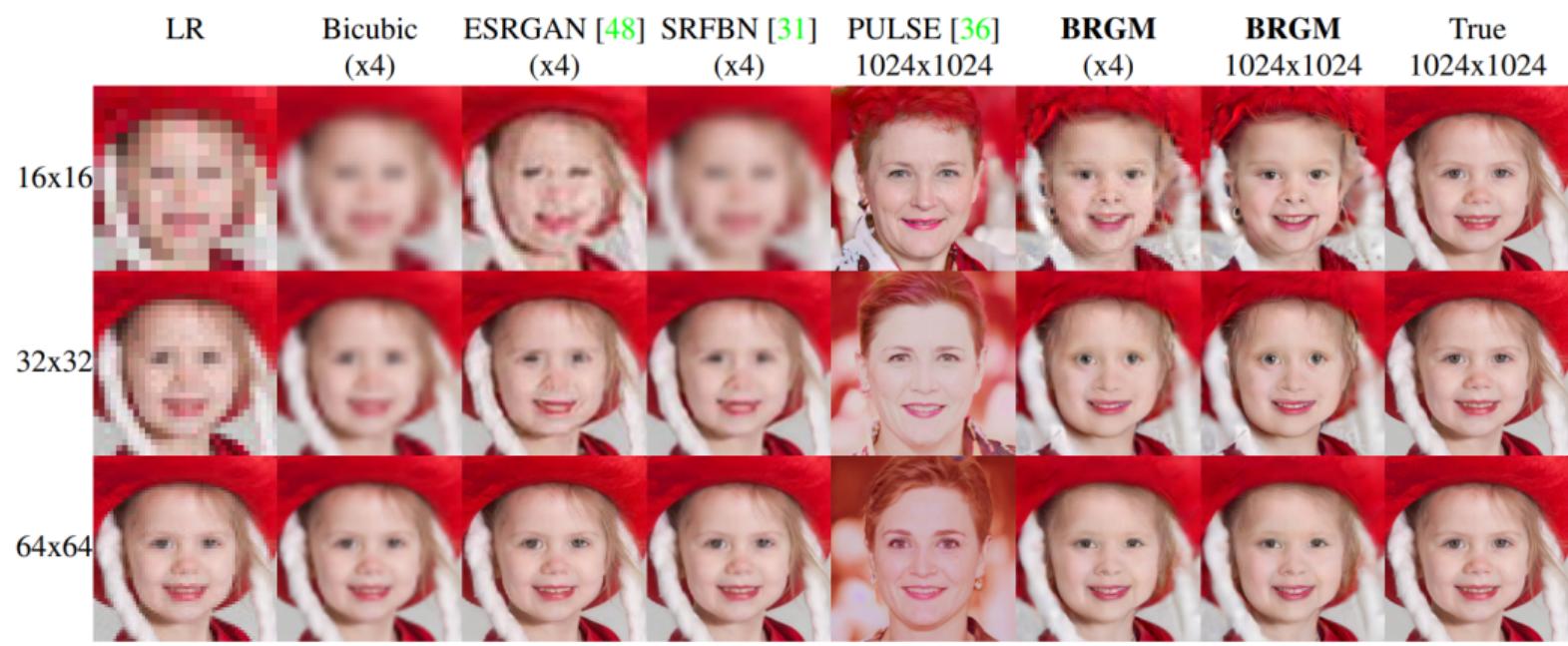
- Yet the reconstruction was not good → required several changes

- Analytically expressed the full likelihood (Marinescu et al, 2021)

$$\begin{aligned} \mathbf{w}^* = \arg \min_{\mathbf{w}} & ||\phi(I) - \phi \circ f \circ G(\mathbf{w})||_2^2 + ||I - f \circ G(\mathbf{w})||_2^2 + \\ & + \sum_i \left( \frac{\mathbf{w}_i - \mu}{\sigma_i} \right)^2 - \sum_{i,j} \frac{\mathbf{w}_i \mathbf{w}_j^T}{|\mathbf{w}_i| |\mathbf{w}_j|} \end{aligned}$$

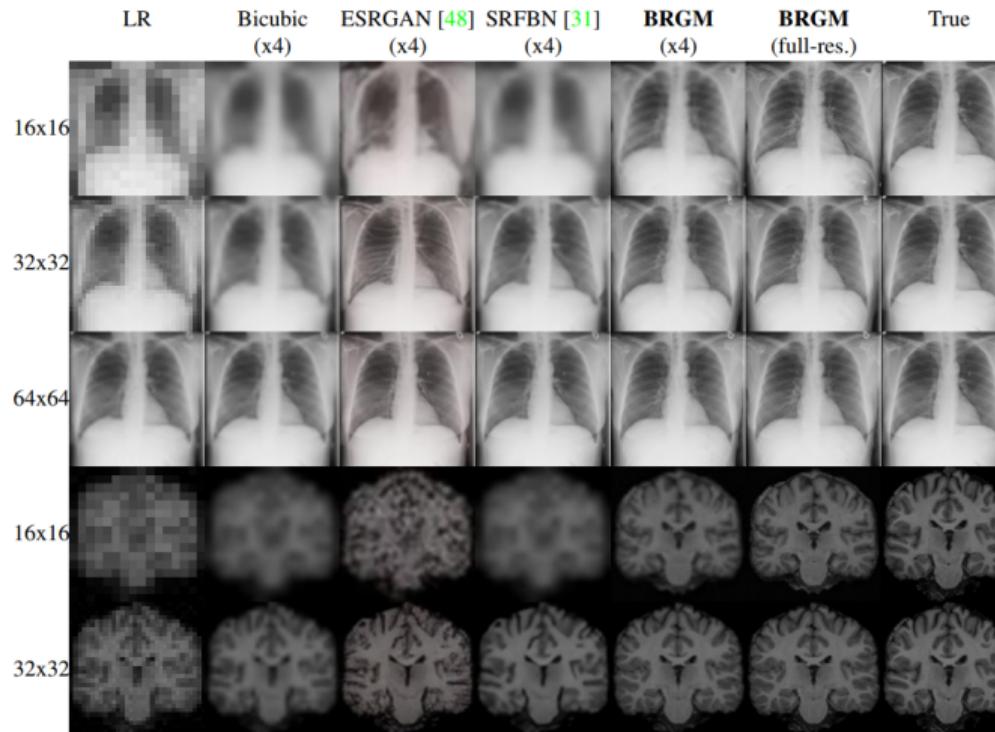
## Results on super-resolution using the FFHQ dataset

- We achieve state-of-the-art (SOTA) results on small inputs resolutions 16x16
- On larger resolutions (>32x32), we achieve very good results, albeit not SOTA



## Similar results on super-resolution for medical datasets

- We achieve state-of-the-art (SOTA) results on small inputs resolutions 16x16
- On larger resolutions (>32x32), we achieve very good results, albeit not SOTA



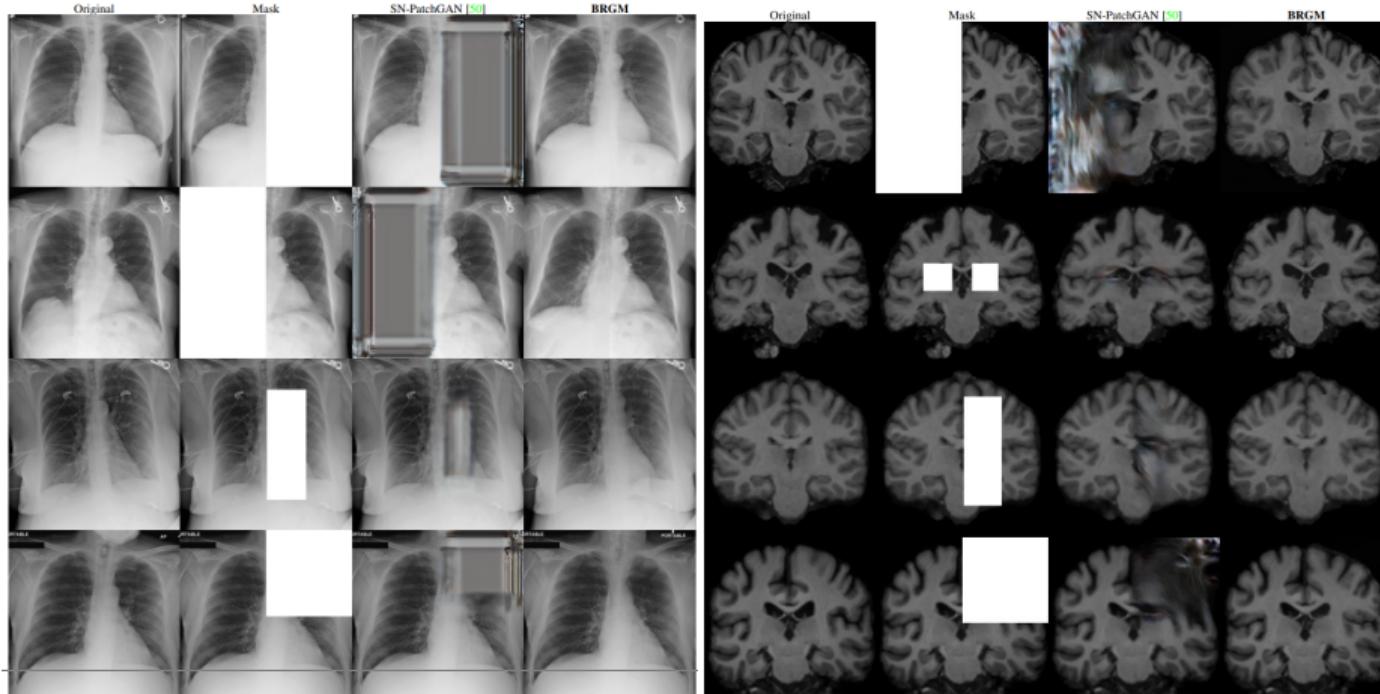
## Inpainting also achieves state-of-the-art results

- Best previous method (SN-PatchGAN, CVPR 2019) does not work for large masks
- Our method can “hypothesize” missing structure



## Inpainting also achieves state-of-the-art results

- Best previous method (SN-PatchGAN, CVPR 2019) does not work for large masks
- Our method can “hypothesize” missing structure



# Results confirmed through quantitative evaluation

- ▶ Three different datasets, at different resolutions
- ▶ Human study with 20 raters

## Super-resolution

Dataset	BRGM	PULSE [36]	ESRGAN [48]	SRFBN [31]
FFHQ 16 <sup>2</sup>	<b>0.24</b> /25.66	0.29/27.14	0.35/29.32	0.33/ <b>22.07</b>
FFHQ 32 <sup>2</sup>	0.30/18.93	0.48/42.97	0.29/23.02	<b>0.23</b> / <b>12.73</b>
FFHQ 64 <sup>2</sup>	0.36/16.07	0.53/41.31	0.26/18.37	<b>0.23</b> / <b>9.40</b>
FFHQ 128 <sup>2</sup>	0.34/15.84	0.57/34.89	0.15/15.84	<b>0.09</b> / <b>7.55</b>
X-ray 16 <sup>2</sup>	<b>0.18</b> / <b>11.61</b>	-	0.32/14.67	0.37/12.28
X-ray 32 <sup>2</sup>	0.23/10.47	-	0.32/12.56	<b>0.21</b> / <b>6.84</b>
X-ray 64 <sup>2</sup>	0.31/10.58	-	0.30/8.67	<b>0.22</b> / <b>5.32</b>
X-ray 128 <sup>2</sup>	0.27/10.53	-	0.20/7.19	<b>0.07</b> / <b>4.33</b>
Brains 16 <sup>2</sup>	<b>0.12</b> / <b>12.42</b>	-	0.34/22.81	0.33/12.57
Brains 32 <sup>2</sup>	<b>0.17</b> /11.08	-	0.31/14.16	0.18/ <b>6.80</b>

## Inpainting

Dataset	BRGM				SN-PatchGAN [50]			
	LPIPS	RMSE	PSNR	SSIM	LPIPS	RMSE	PSNR	SSIM
FFHQ	<b>0.19</b>	<b>24.28</b>	<b>21.33</b>	<b>0.84</b>	0.24	30.75	19.67	0.82
X-ray	<b>0.13</b>	<b>13.55</b>	<b>27.47</b>	<b>0.91</b>	0.20	27.80	22.02	0.86
Brains	<b>0.09</b>	<b>8.65</b>	<b>30.94</b>	<b>0.88</b>	0.22	24.74	21.47	0.75

## Human evaluation

Dataset	BRGM	PULSE [36]	ESRGAN [48]	SRFBN [31]
FFHQ 16 <sup>2</sup>	<b>0.42</b>	0.32	0.11	0.15
FFHQ 32 <sup>2</sup>	0.39	0.02	0.12	<b>0.47</b>
FFHQ 64 <sup>2</sup>	0.14	0.08	0.32	<b>0.45</b>
FFHQ 128 <sup>2</sup>	0.14	0.10	<b>0.39</b>	0.38

## Our method also has limitations

- ▶ It can fail for images that are too dissimilar to the training ones
  - ▶ Because generator cannot extrapolate easily



- ▶ Can be inconsistent with the input image



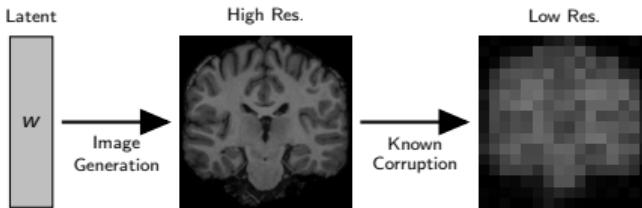
- ▶ Proposed a method for image reconstruction using pre-trained deep generative models
- ▶ Solution is given by the Bayesian MAP estimate
- ▶ State-of-the-art results on super-resolution and inpainting

## 1. Disease progression modelling of Alzheimer's disease

### 1.1 Towards unsupervised clustering of biomarker trajectories



## 2. Image Reconstruction using Deep Generative Models



## 3. Future work

## Accurate diagnosis and prognosis through AI

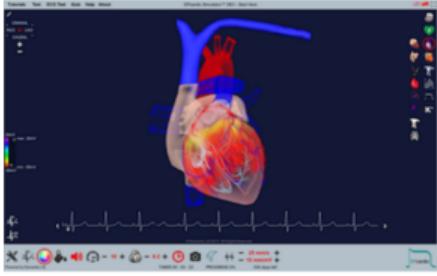


## AI to augment humans

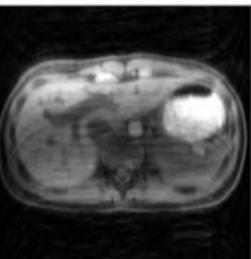
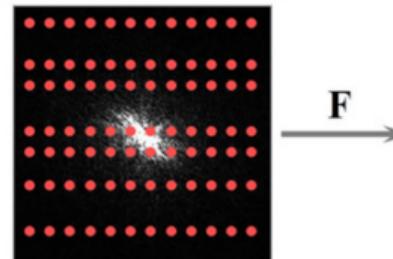


## Future work

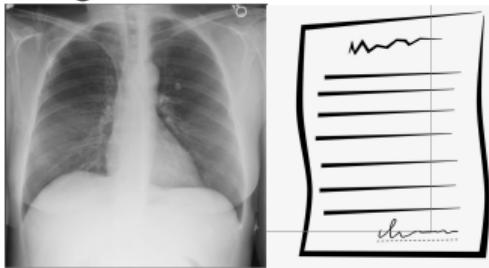
### Biological simulators



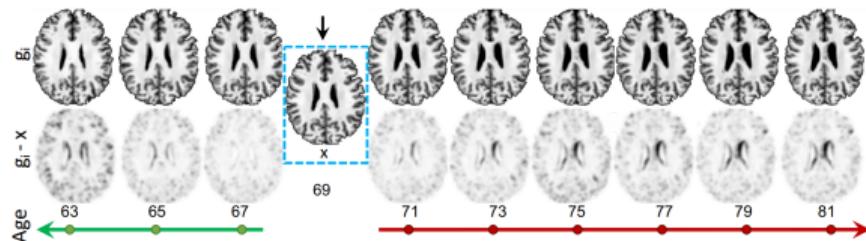
### Better and faster reconstruction of medical images Undersampled k-space



### Multimodal modelling images + text + structural data



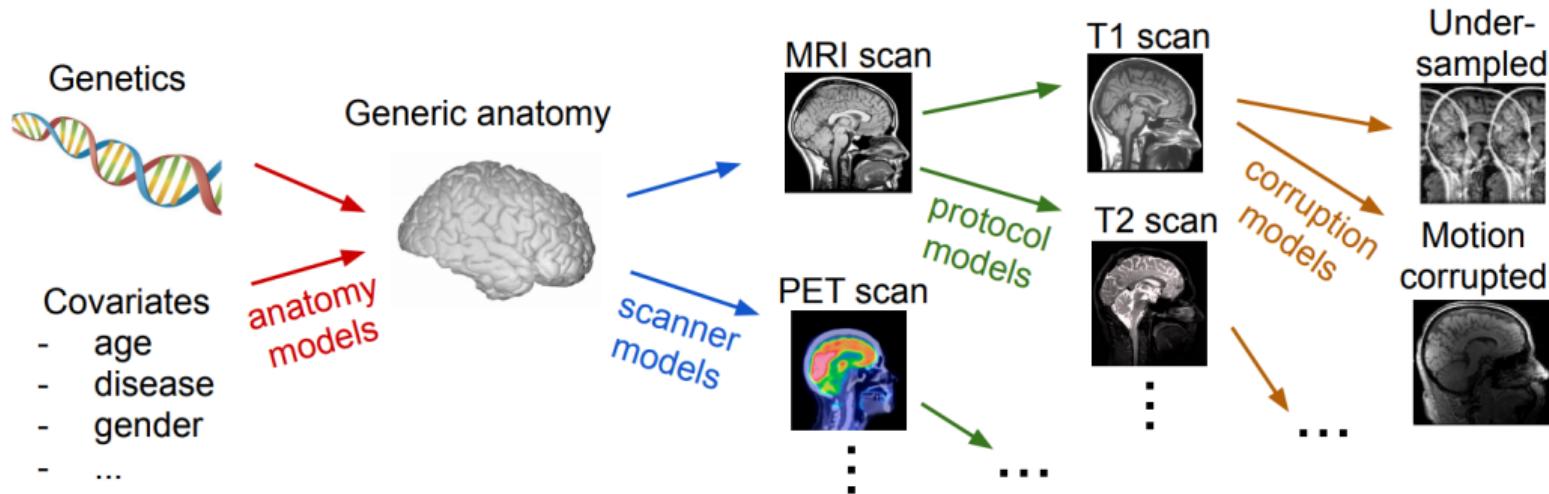
### Disease Progression Modelling



## Future work: Brain tissue and anatomy simulator

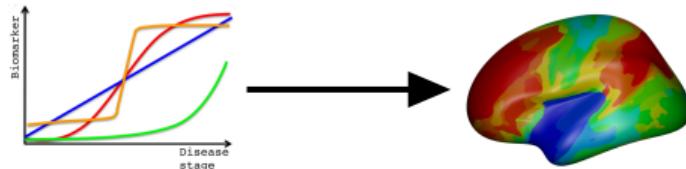
Simulator for brain anatomy from genetics:

- ▶ Using deep generative models
- ▶ Accounting for distributions shifts
- ▶ Following causal principles



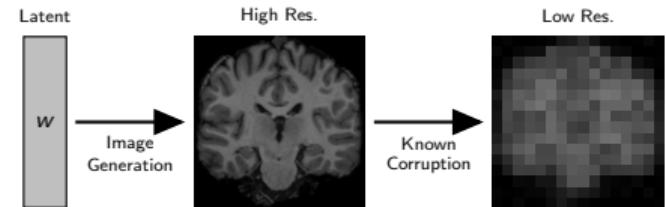
## Problem: Lack of good labels

Solution: Unsupervised Learning through Disease Progression Modelling



## Problem: Lack of good input data

Solution: Image Reconstruction using Deep Generative Models



Long-term vision

## Accurate diagnosis and prognosis through AI



## AI to augment humans

