Generative Models part II

Cian M. Scannell, PhD
Assistant Professor
Medical Image Analysis
c.m.scannell@tue.nl

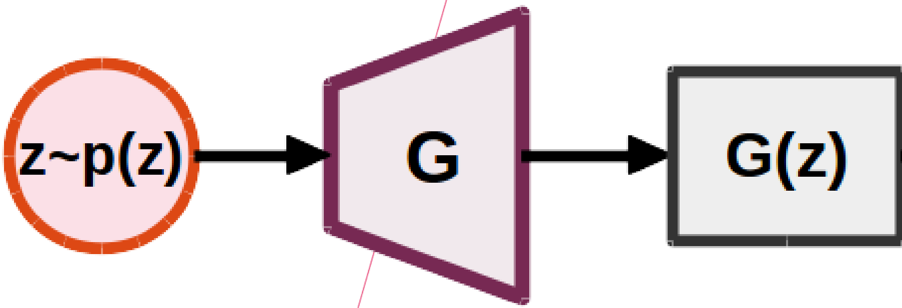
Topics:
- Limitations and extensions to GANs
- Conditional generation
- Image-to-image translation
- Fréchet inception distance
- Normalizing flow
- Diffusion models

Intended learning outcomes

The student can:

- Understand the benefits of GANs in the context of the limitations of VAEs
- Formulate the GAN training process and loss function
- Classify applications of GANs
- Give examples of issues with GANs
- Motivate solutions to these issues
- Understand the benefits of normalising flows in the context of the limitations of VAEs
- Describe the origin of the name "normalising flows"
- Summarise the normalising flow algorithm
- Relate diffusion models to VAEs

Recap VAEs



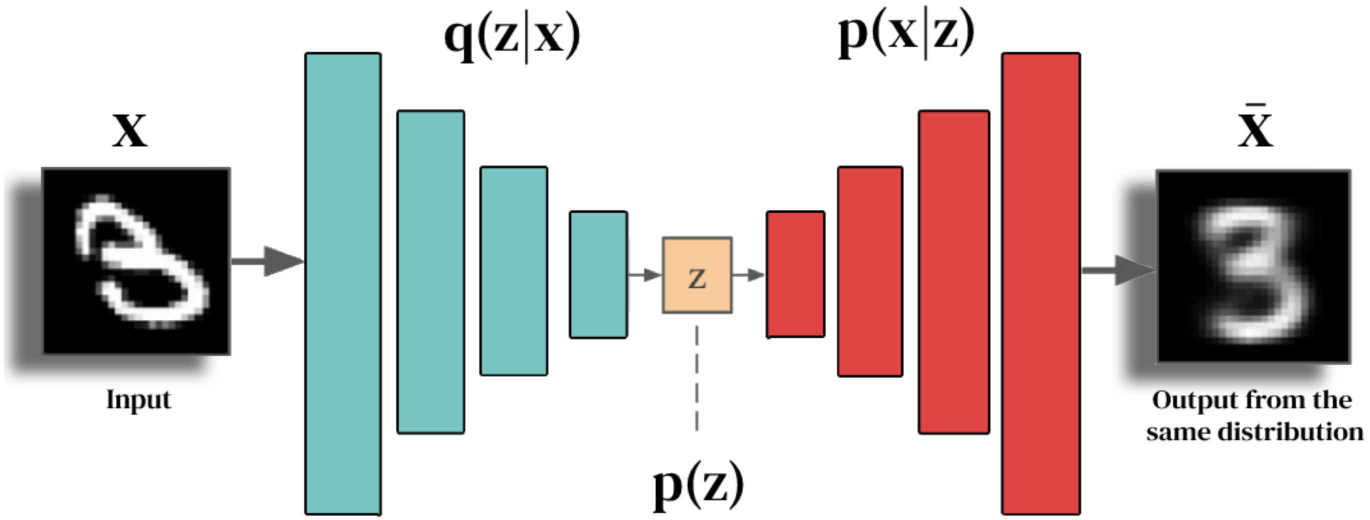
- A VAE is a latent variable model
  - Sample $z \sim p(z)$
  - Generate $x \sim p(x|z)$
- A VAE is an *explicit* likelihood model
  - It tries to explicitly compute (approximate) the density $p(x)$
  - But we need to do a lot of work to get around the lack of an analytical solution and the curse of dimensionality

Limitation:

- Intractable likelihood - > Maximise ELBO rather than likelihood directly

Limitations:

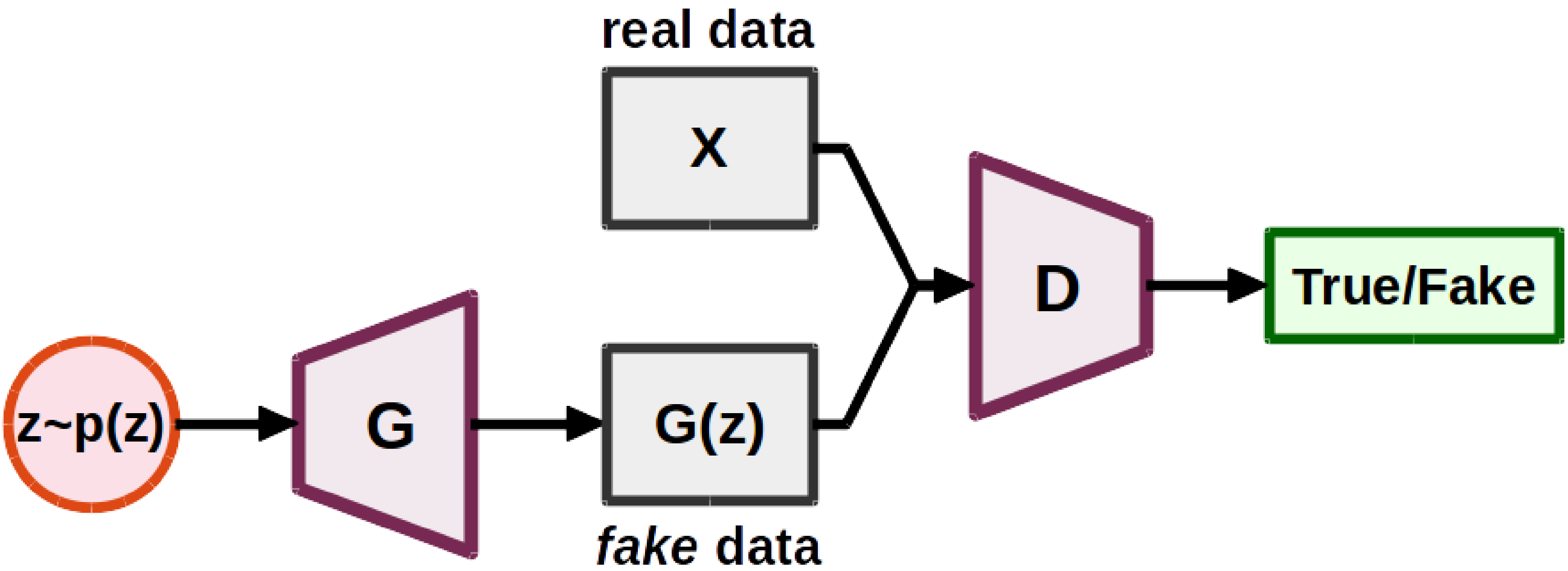
- VAEs may not generate samples that are as high-quality or realistic, especially for images.



- Classes may overlap in latent space. (Latent space is entangled).
- Reconstruction losses (MSE, MAE) may not high quality image generation

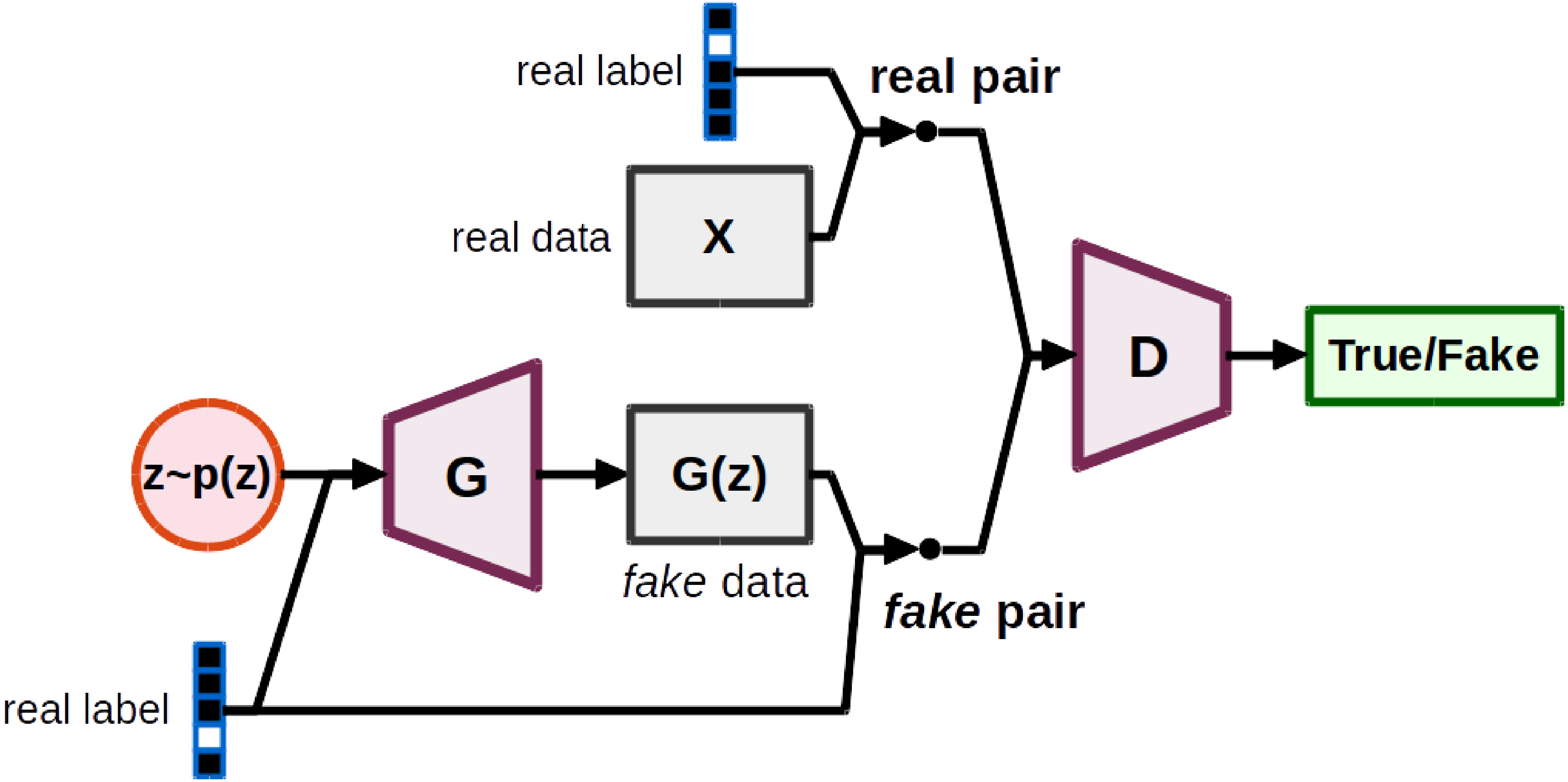
GAN:

GANs are implicit likelihood models



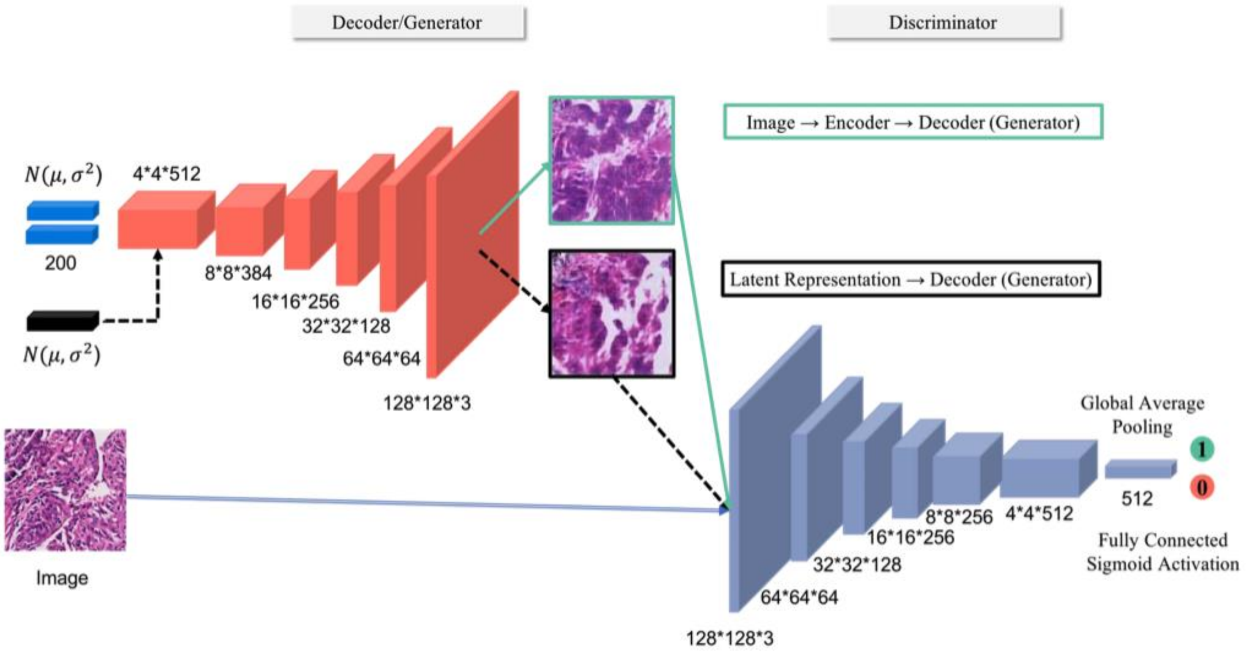
Conditional generation

Conditional GAN (cGAN):

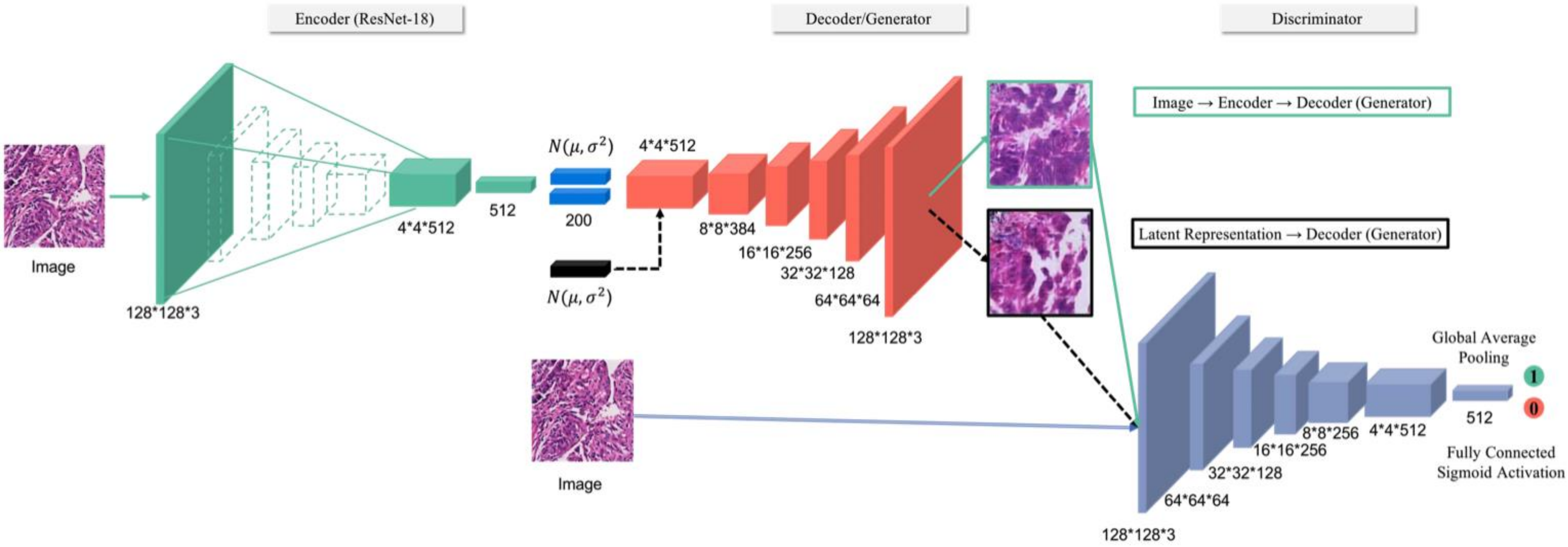


$D_\alpha(x|y)$ and $G_\beta(z|y)$ instead of $D_\alpha(x)$ and $G_\beta(z)$.

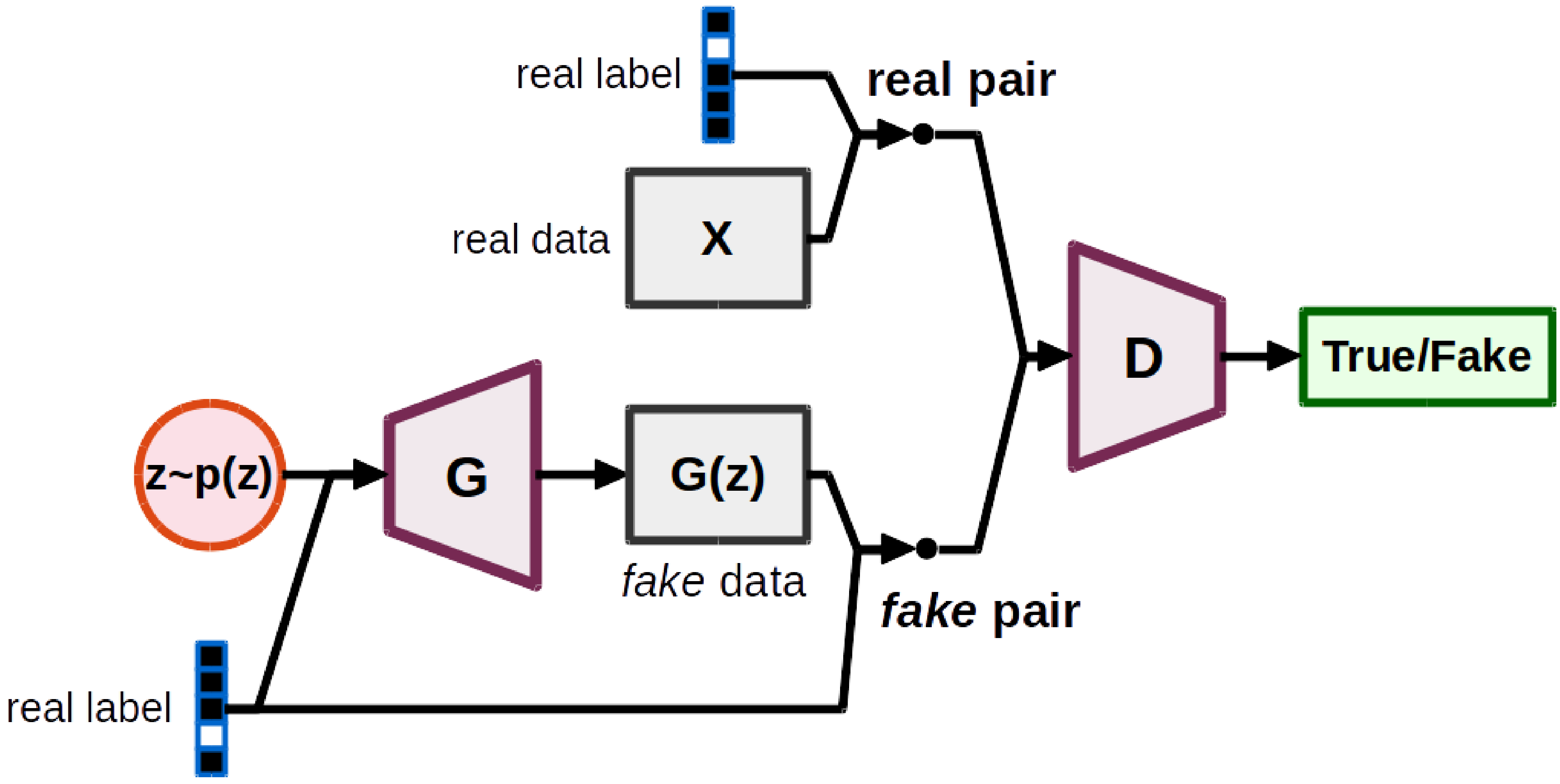
Hybrid models:



Hybrid models – VAE-GAN:



Conditional GAN (cGAN):

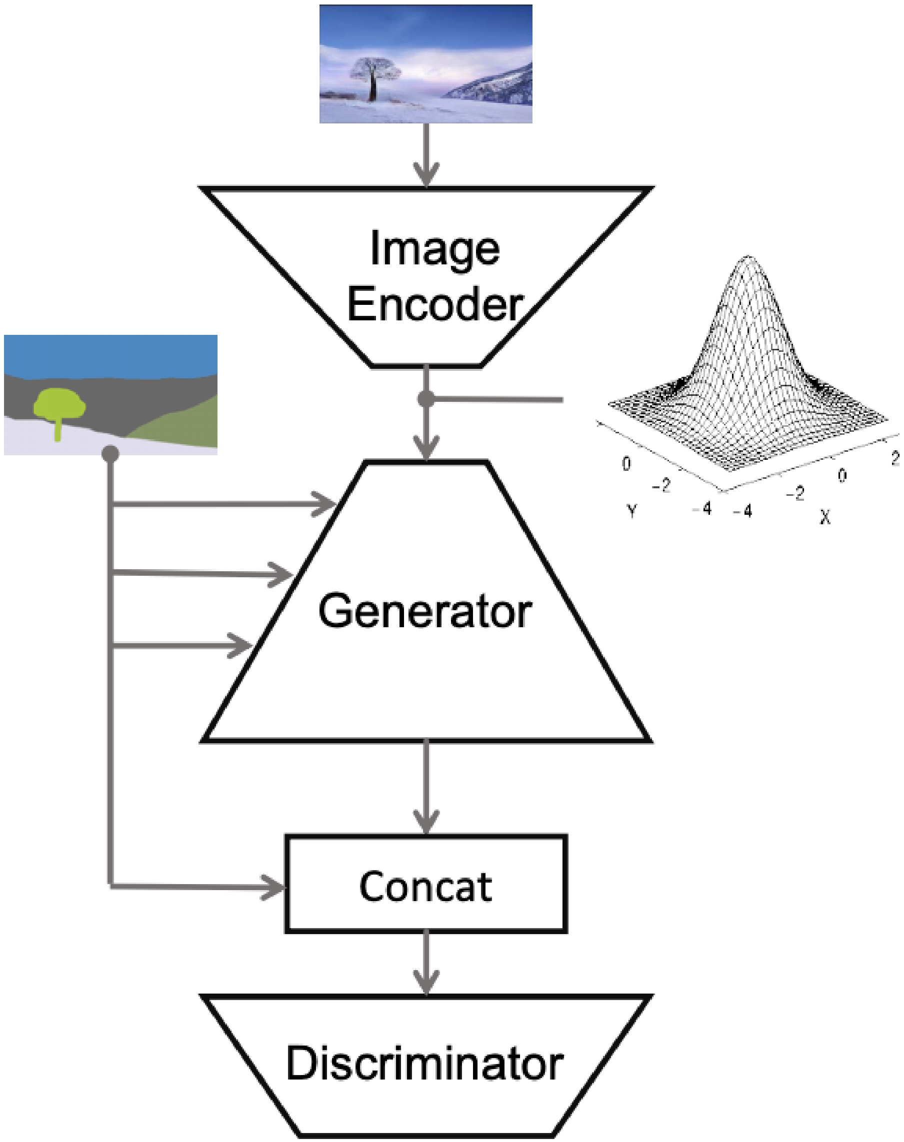


$D_\alpha(x|y)$ and $G_\beta(z|y)$ instead of $D_\alpha(x)$ and $G_\beta(z)$.

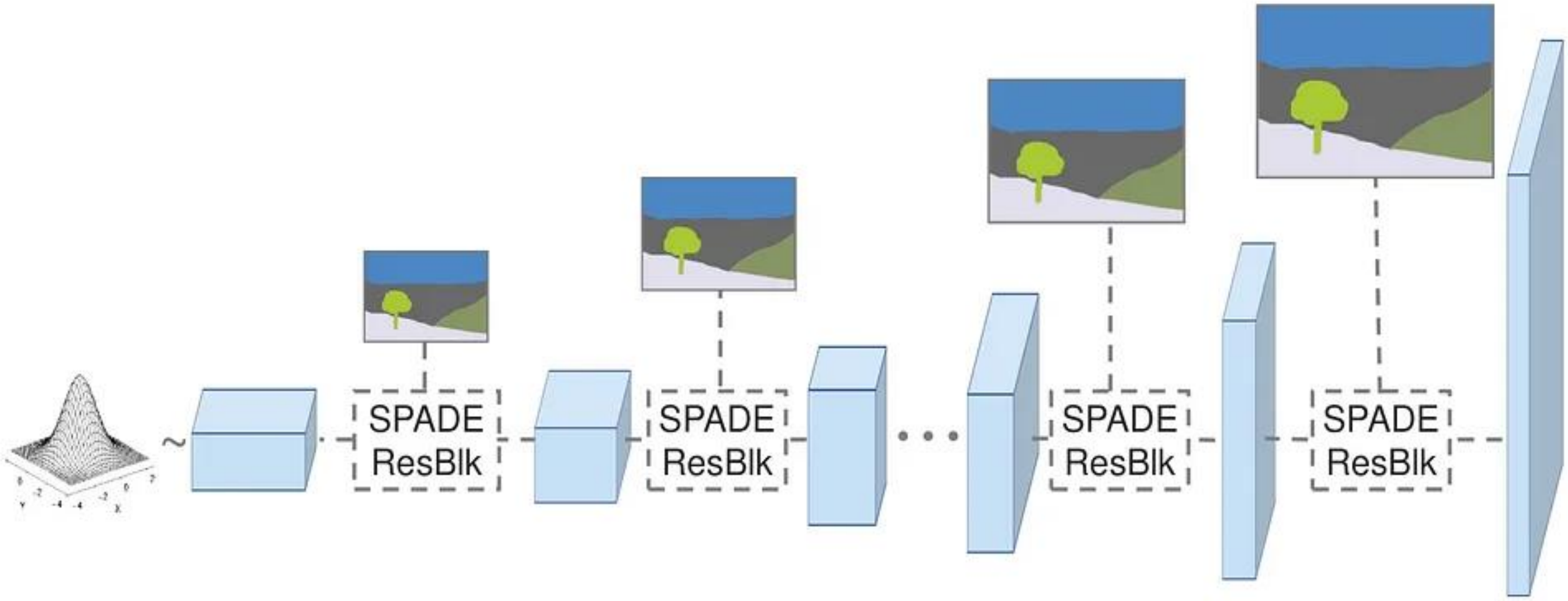
Condition on segmentations

SPADE-GAN

Spatial adaptive (de)normalisation (SPADE):



SPADE:



SPADE:

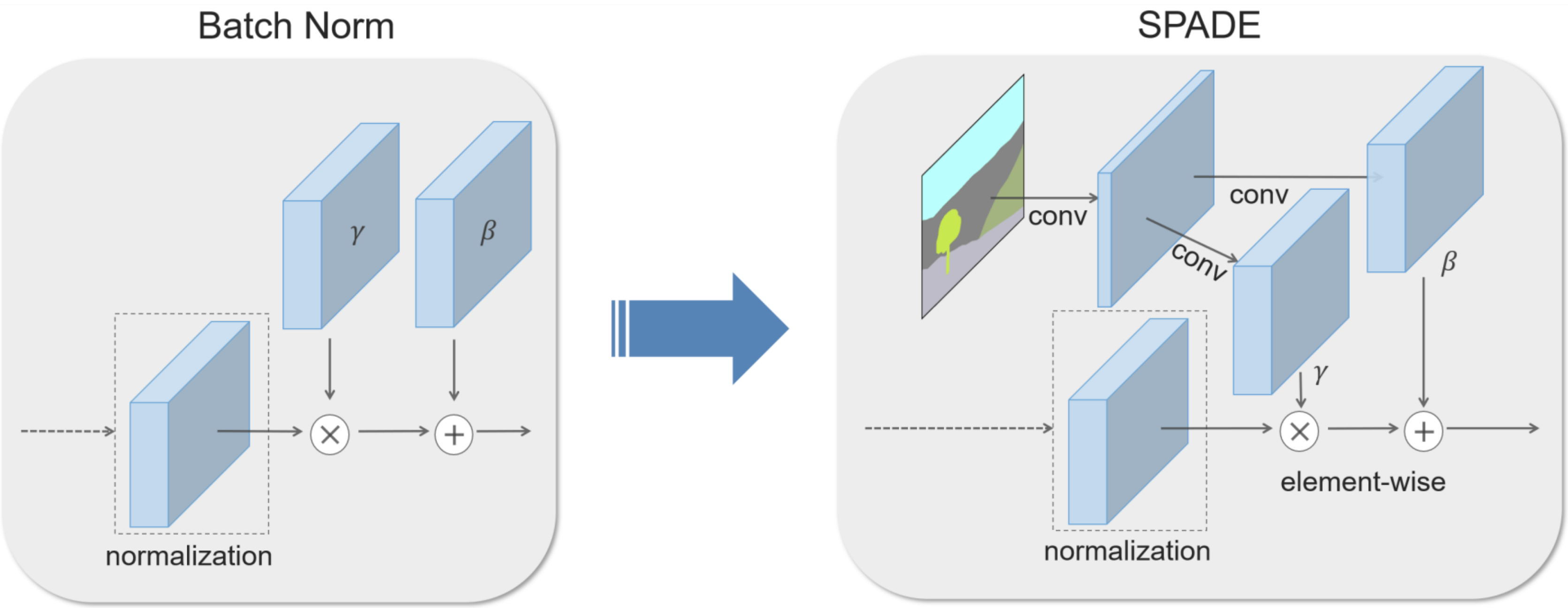
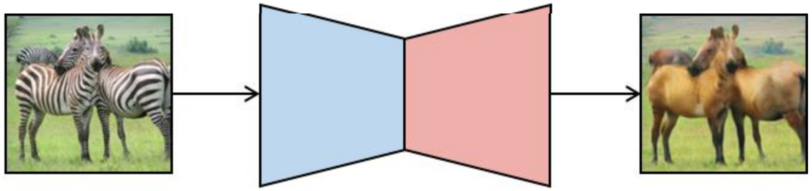
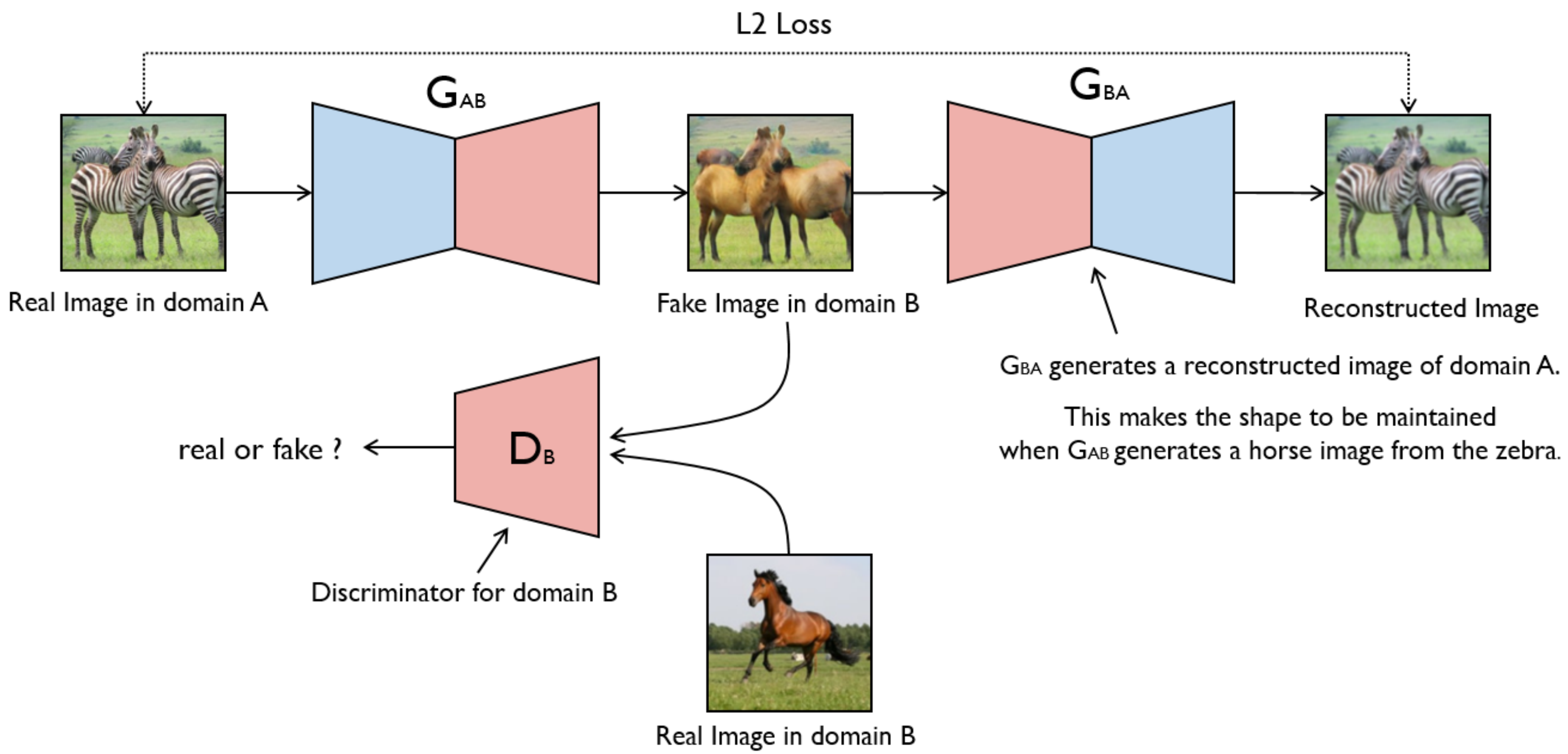

Batch Norm
γ
β
normalization
SPADE
conv
conv
conv
β
γ
element-wise
normalization

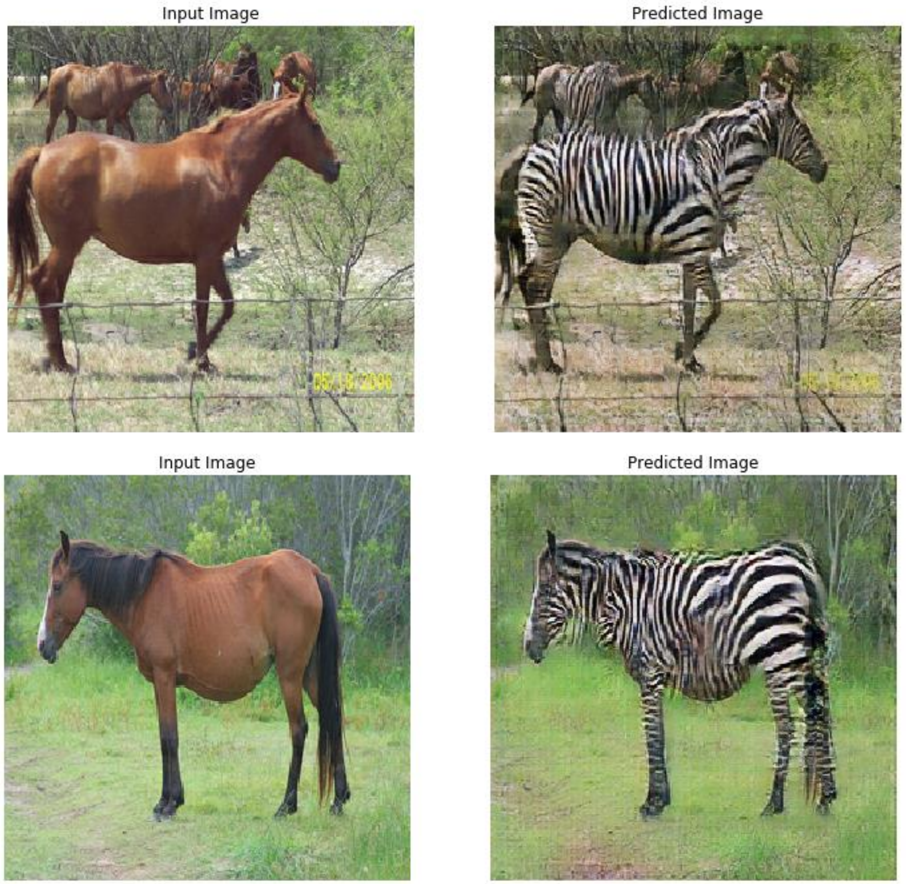
Image to image translation:



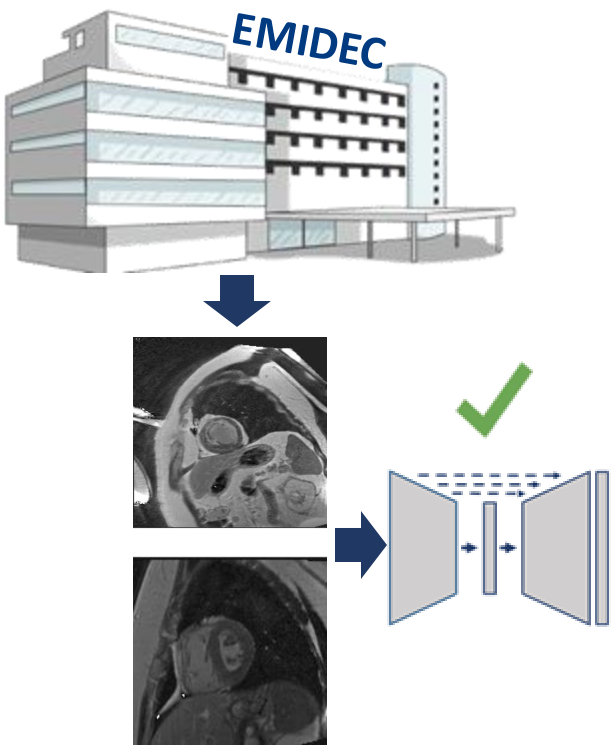
CycleGAN:

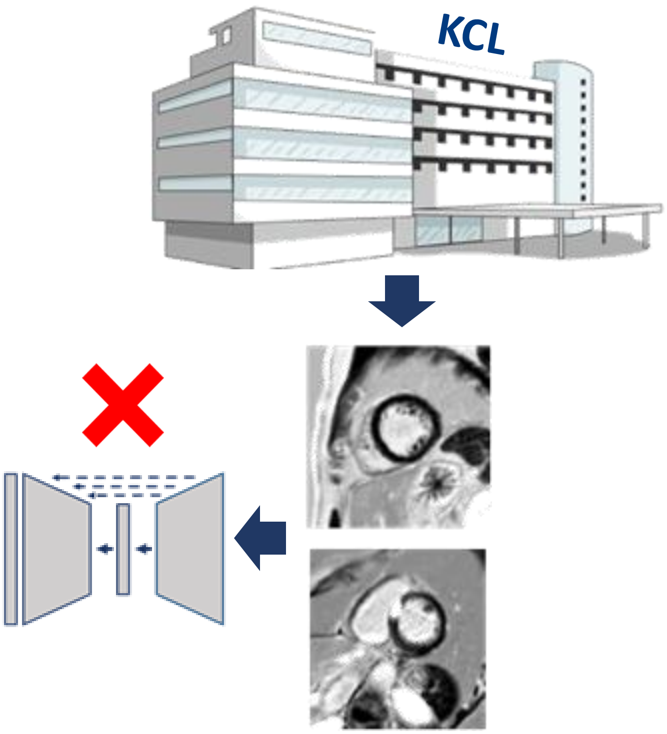

L2 Loss
GAB
GBA
Real Image in domain A
Fake Image in domain B
Reconstructed Image
GBA generates a reconstructed image of domain A.
This makes the shape to be maintained
when GAB generates a horse image from the zebra.
real or fake ?
DB
Discriminator for domain B
Real Image in domain B

CycleGAN examples:

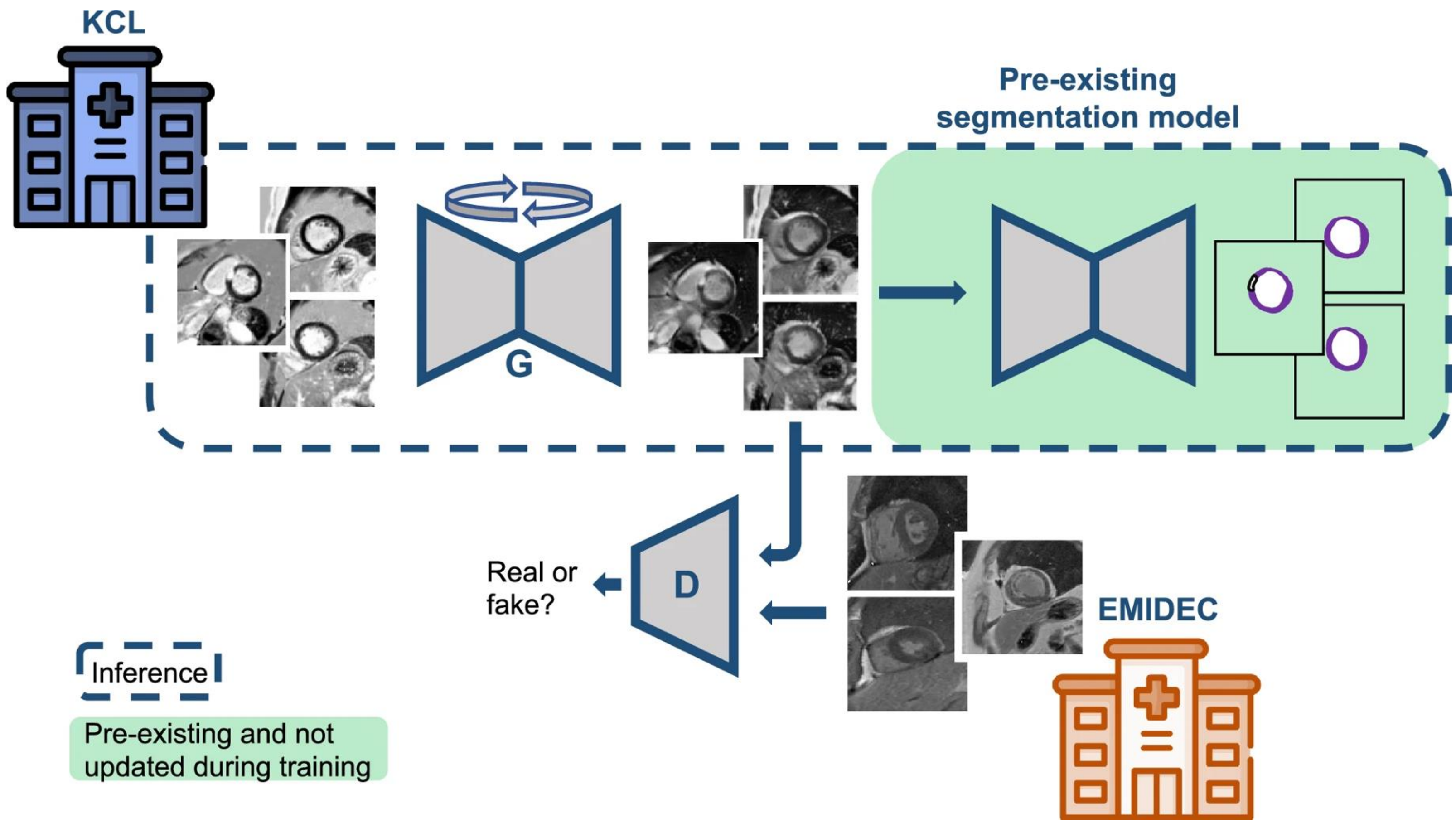


CycleGAN examples:



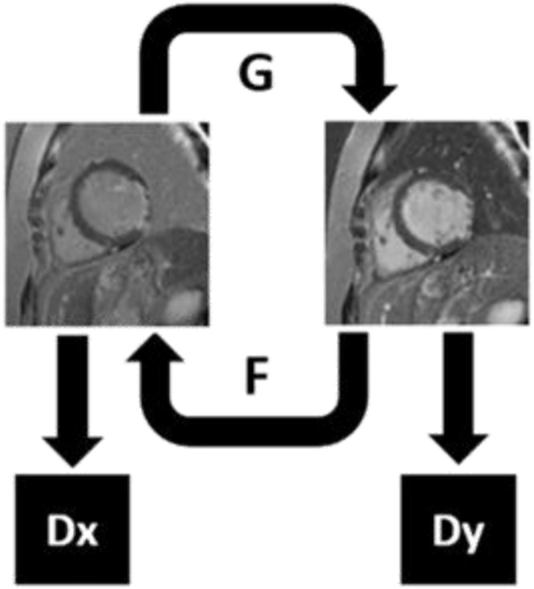


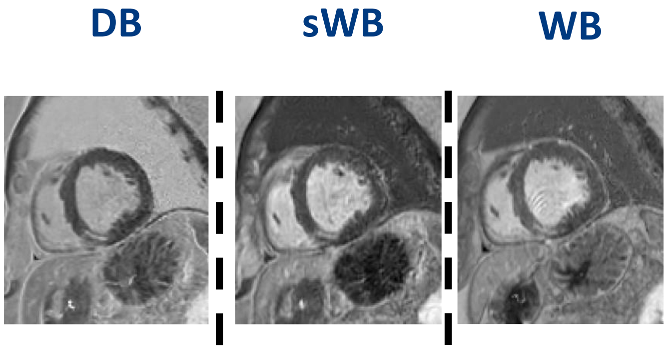
CycleGAN examples:



R Crawley, …, CM Scannell *European Radiology Experimental* 2024

CycleGAN example:





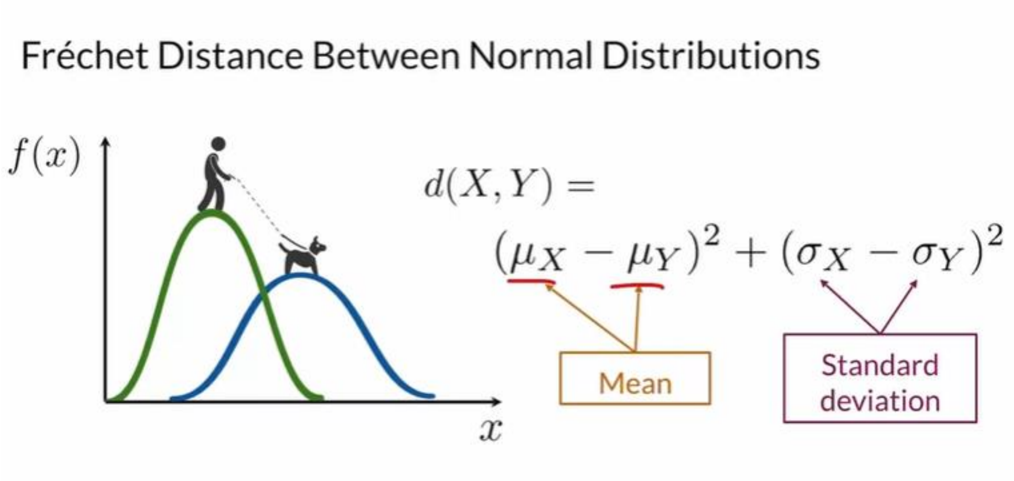
T Jaspers, …, CM Scannell *Investigative Radiology* 2024

Evaluation

Evaluating synthetic data is very hard and still requires a lot of research!!

(especially for medical domains)

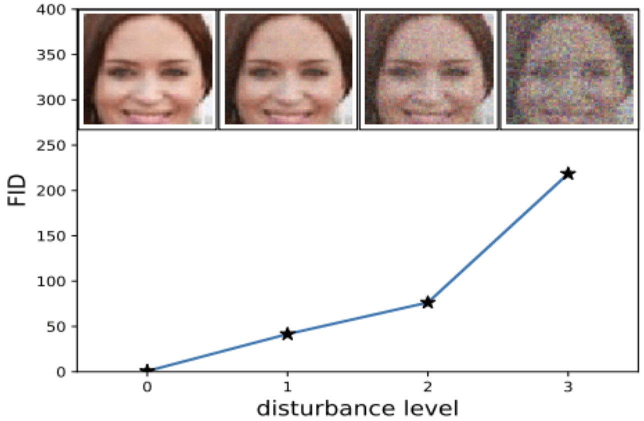
Fréchet inception distance (FID):



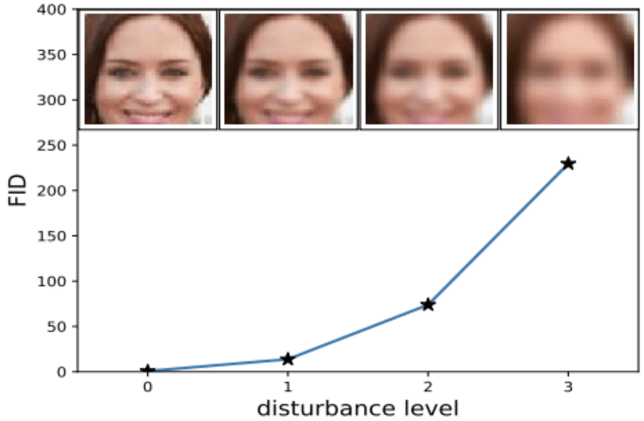
We use the activations from the pre-trained Inception V3 model to summarise each image and compute the distance between these features

$$FID = \|\mu_r - \mu_g\|^2 + T_r(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{1/2})$$

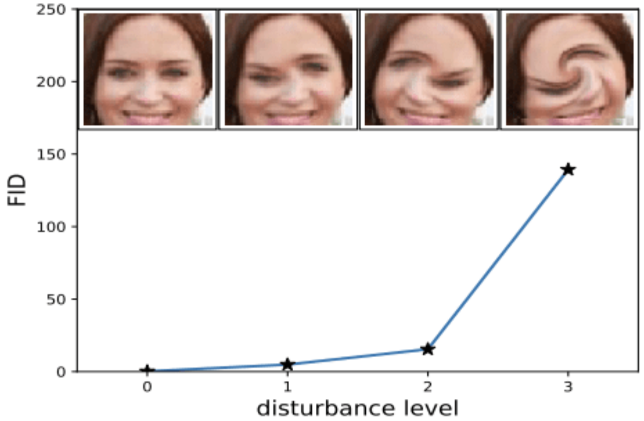
FID example:



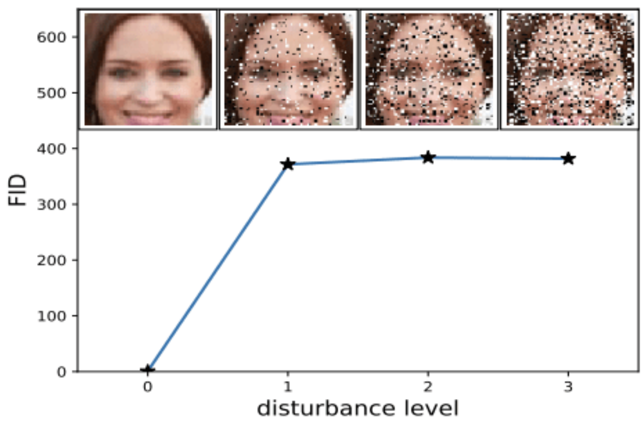
400
350
300
250
200
150
100
50
0
FID
0
1
2
3
disturbance level



400
350
300
250
200
150
100
50
0
FID
0
1
2
3
disturbance level



250
200
150
100
50
0
FID
0
1
2
3
disturbance level



600
500
400
300
200
100
0
FID
0
1
2
3
disturbance level

Evaluation

- Talk to professionals
- Downstream tasks

Normalising flows

Motivation:

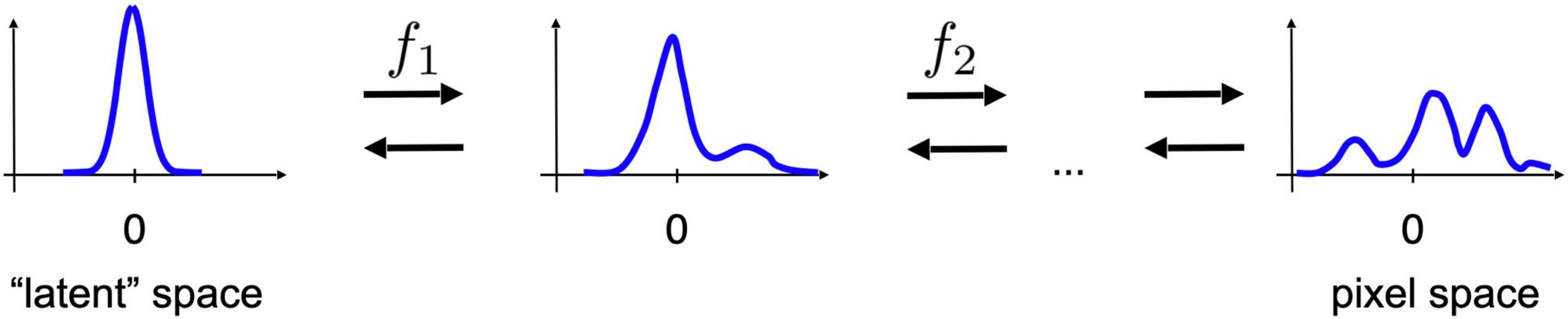
- Can we design a latent variable model with a tractable likelihood?
- Ideally:
    - Easy to evaluate
    - Easy to sample from

  (like a Gaussian)

Normalising flows

Normalising flows - main idea:

- Learn to map simple distributions to complex ones with an invertible mapping
- In particular, map Z to X with a deterministic, invertible function f_θ such that $X = f_\theta(Z)$ and $Z = f_\theta^{-1}(X)$



f_1
f_2
...
0
0
0
"latent" space
pixel space

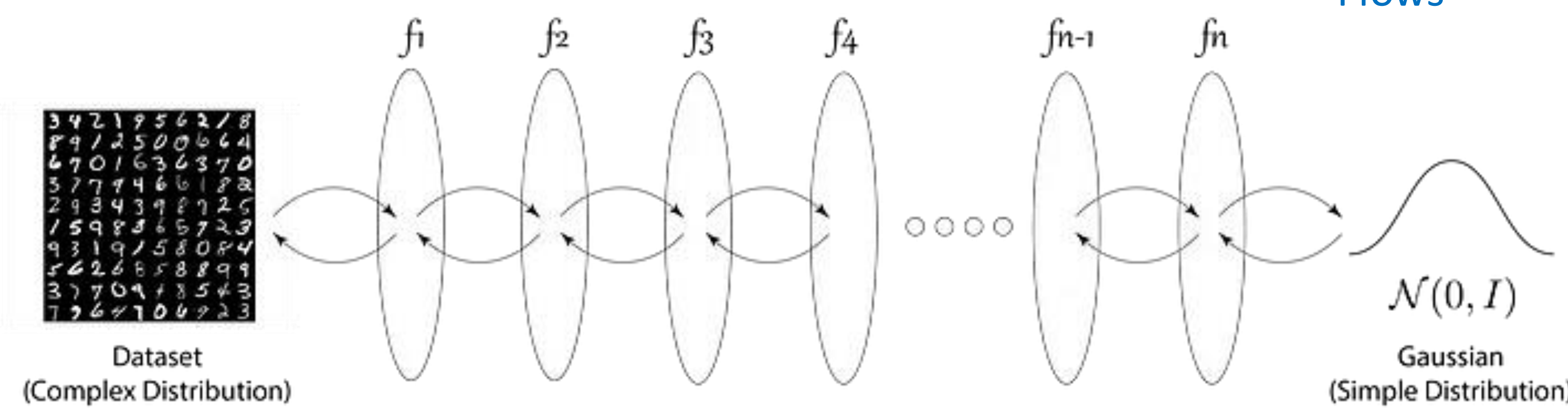
Normalising flows:

- If $z \sim p(z)$ and $f : Z \to X$
- Then $x \sim p(x) = p(z = f^{-1}(x)) \boxed{\left| det\left(\frac{\partial f^{-1}(x)}{\partial x}\right) \right|}$

Normalising

- To model complex high dimensional distributions, a sequence of these invertible transformations is applied.

Flows

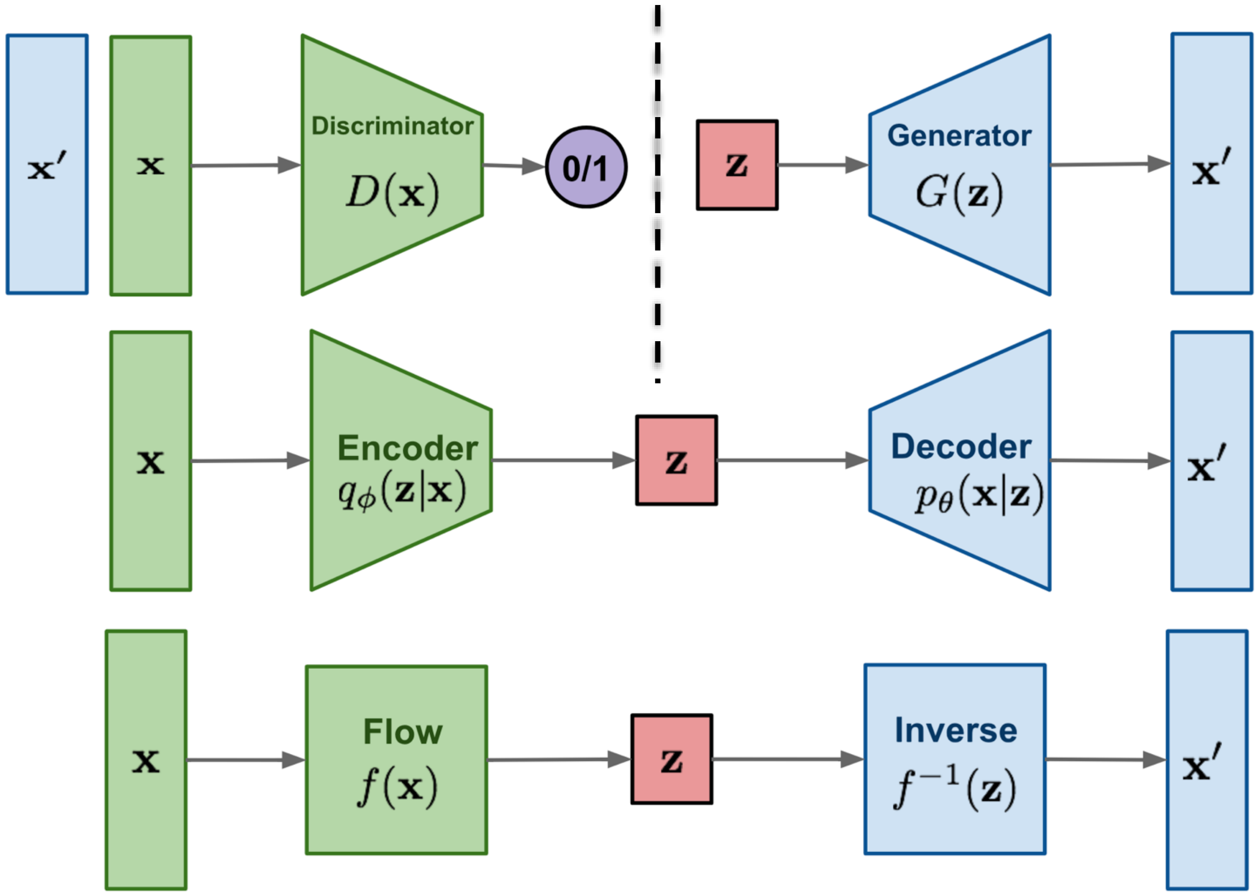


Summary:

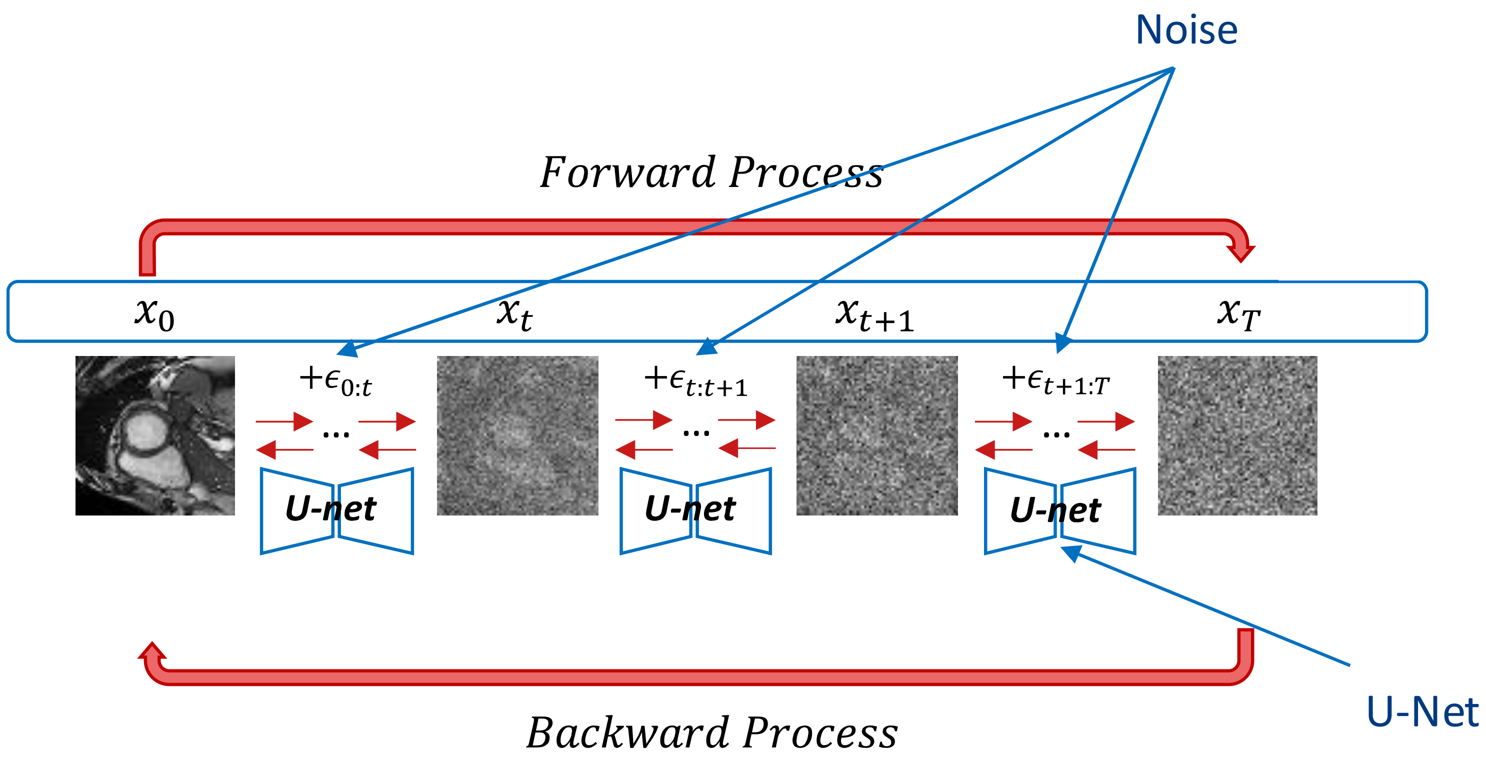
GAN: minimax the classification error loss.

VAE: maximize ELBO.

Flow-based generative models: minimize the negative log-likelihood

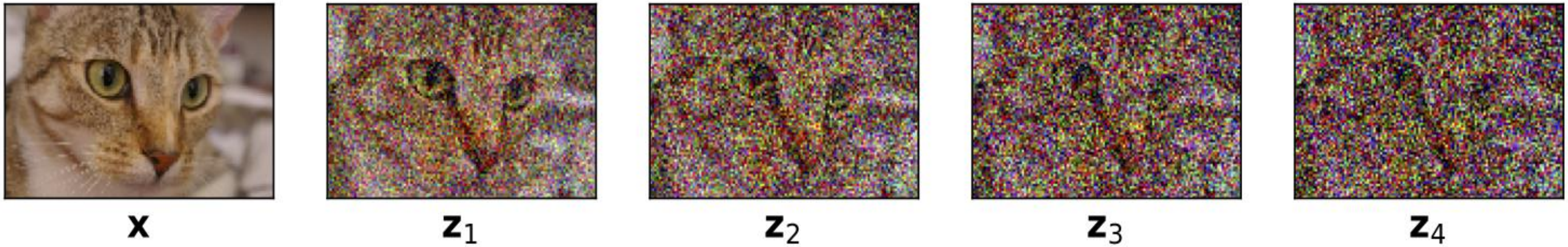


Denoising diffusion probabilistic model (DDPM):

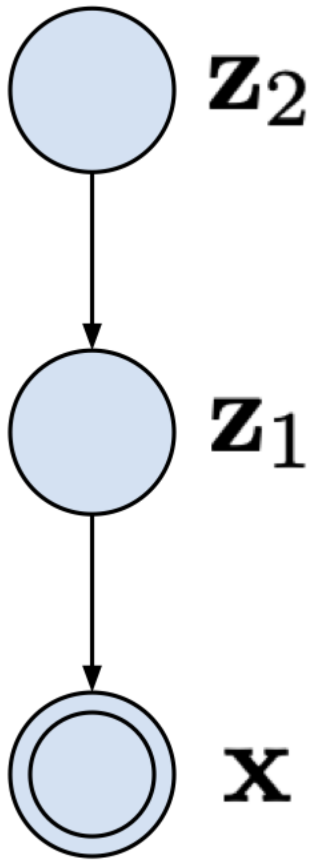


Diffusion models:

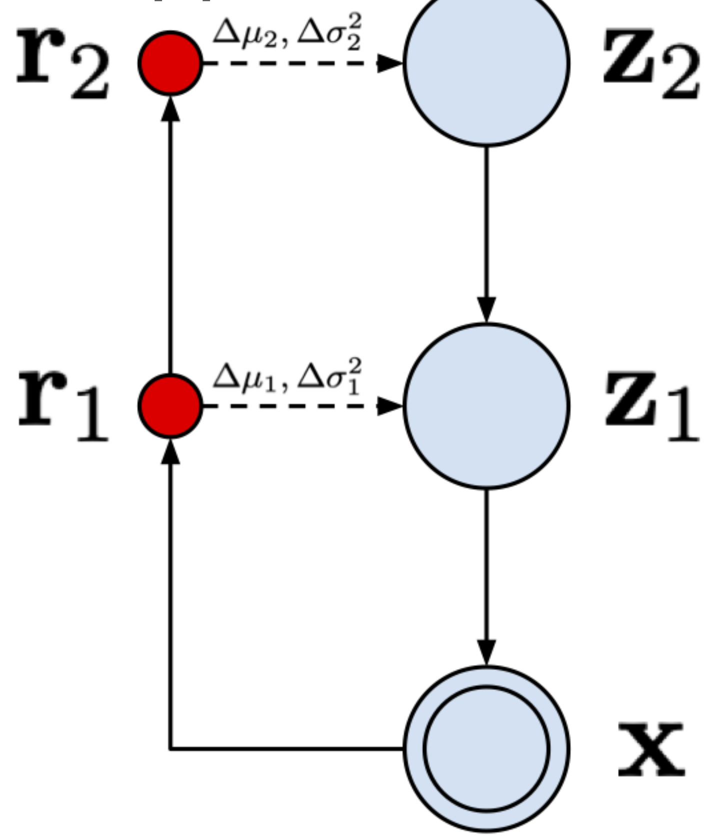
A diffusion model is a hierarchical VAE with the bottom-up path defined by a diffusion process and the top-down path parametrised by neural networks (reversed diffusion).



Hierarchical VAE:



Top-down hierarchical VAE with bottom-up path:



Generative Models applications

Cian M. Scannell, PhD
Assistant Professor
Medical Image Analysis
c.m.scannell@tue.nl

Topics:

- Application of generative modelling in medical image analysis
- How to generated cardiac MR images with variations
- Evaluation of synthetic data
- Usability of synthetic data
- Use cases of synthetic data

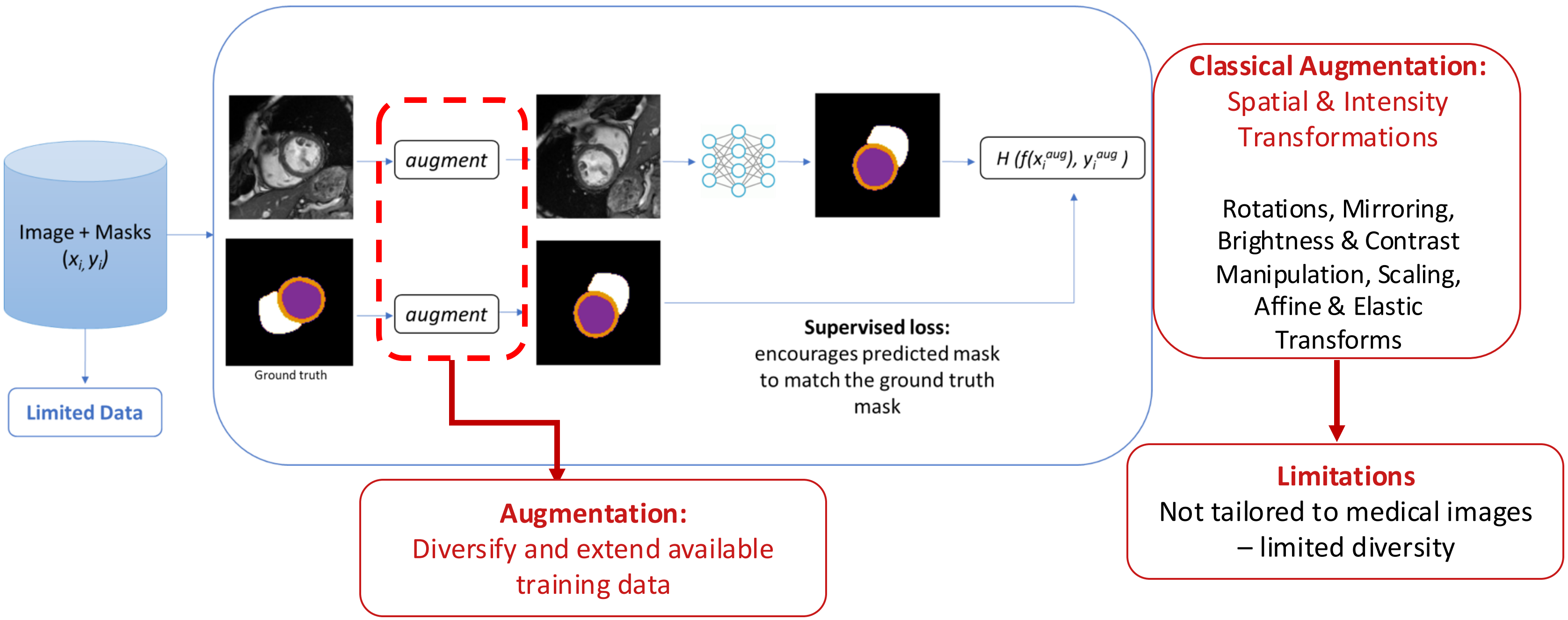
Learning objectives

The student can:

- Understand some of the applications of the synthetic data
- Learn about types of generative models
- Use generative models to enrich data
- Evaluate the quality of the synthetic data
- Investigate the usability of the synthetic data for DL training

Improving model performance in the presence of limited data

Using generated images as a data augmentation strategy to increase data diversity



Courtesy of Yasmina Al Khalil

IMAG/e

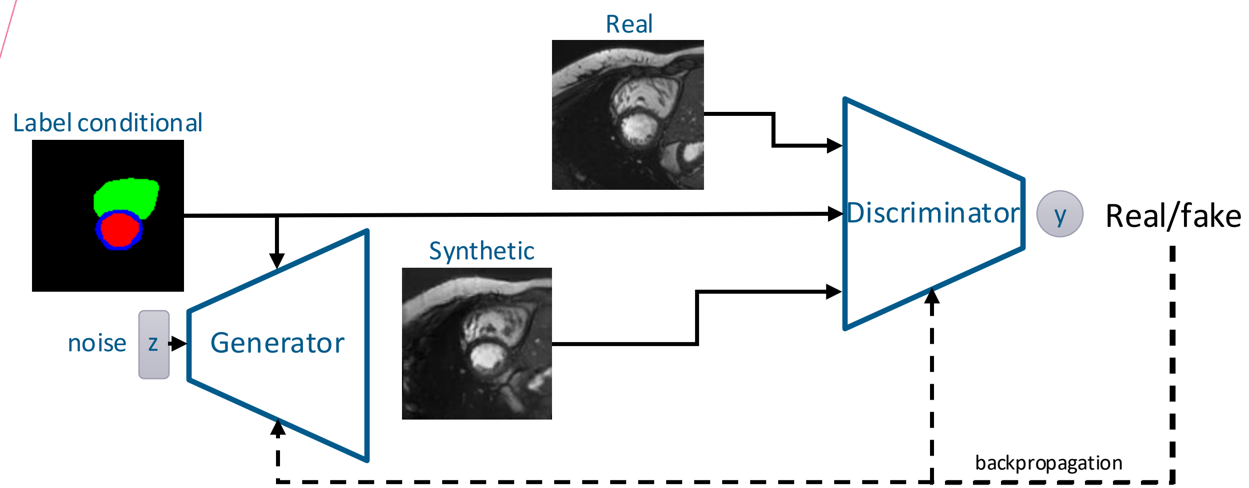
TU/e Technische Universiteit Eindhoven University of Technology

Can we use synthetic images to improve the performance of segmentation models?

Let’s first understand how to synthesize images with variations…

The conditional **generator** translates input labels into synthesized images to trick the discriminator

The conditional **discriminator** tries to identify real images from synthesized ones created by the generator

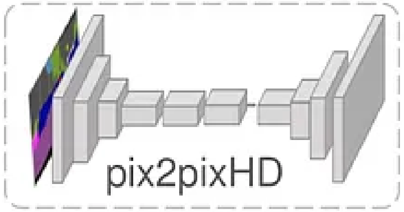


Why condition on labels?

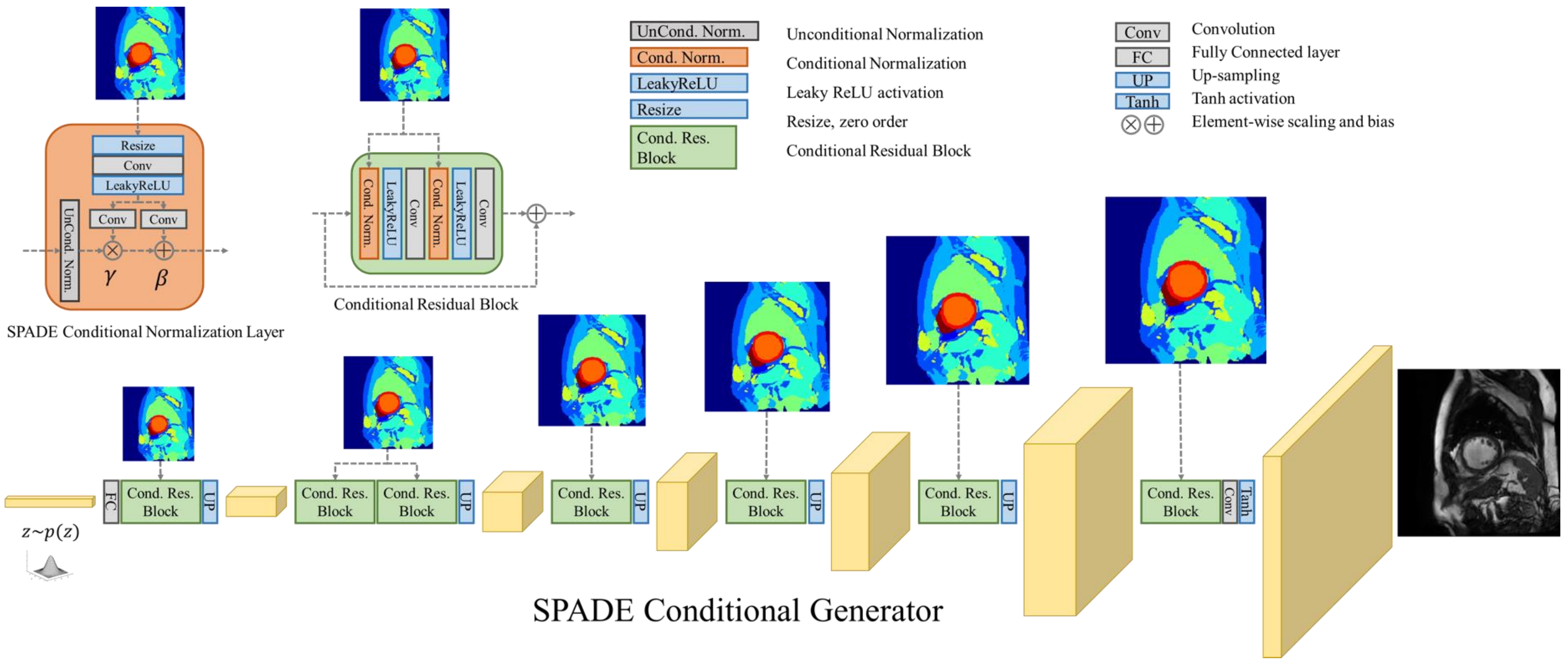
How to generated cardiac MR images using conditional GANs

Generator architecture

Previous work

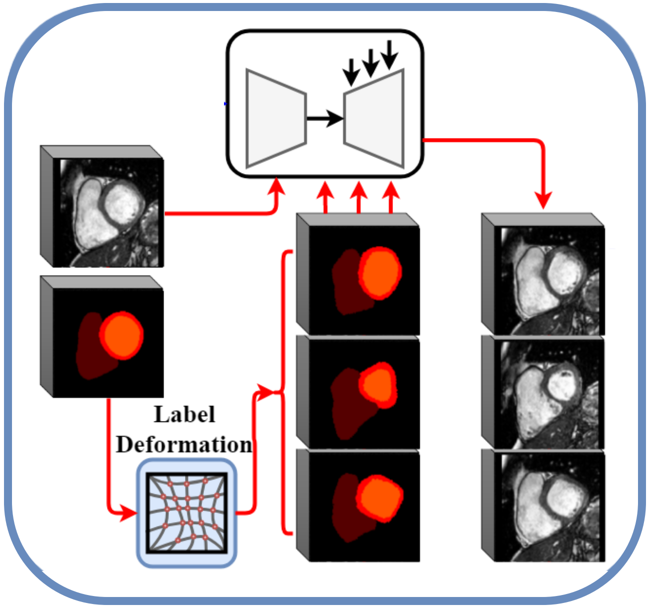


SPADE normalization layers to preserve the anatomical information of the labels

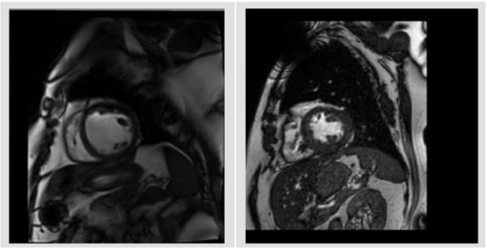


Important notes on generating diverse examples

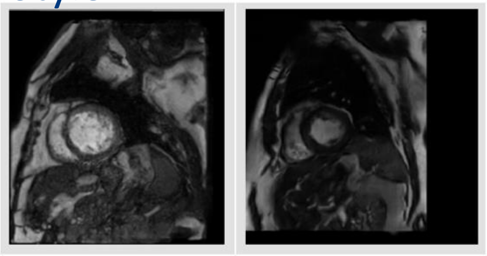
- Include contrast/style variation
- Include label deformation to provide anatomical variation in the generated data



Style 1



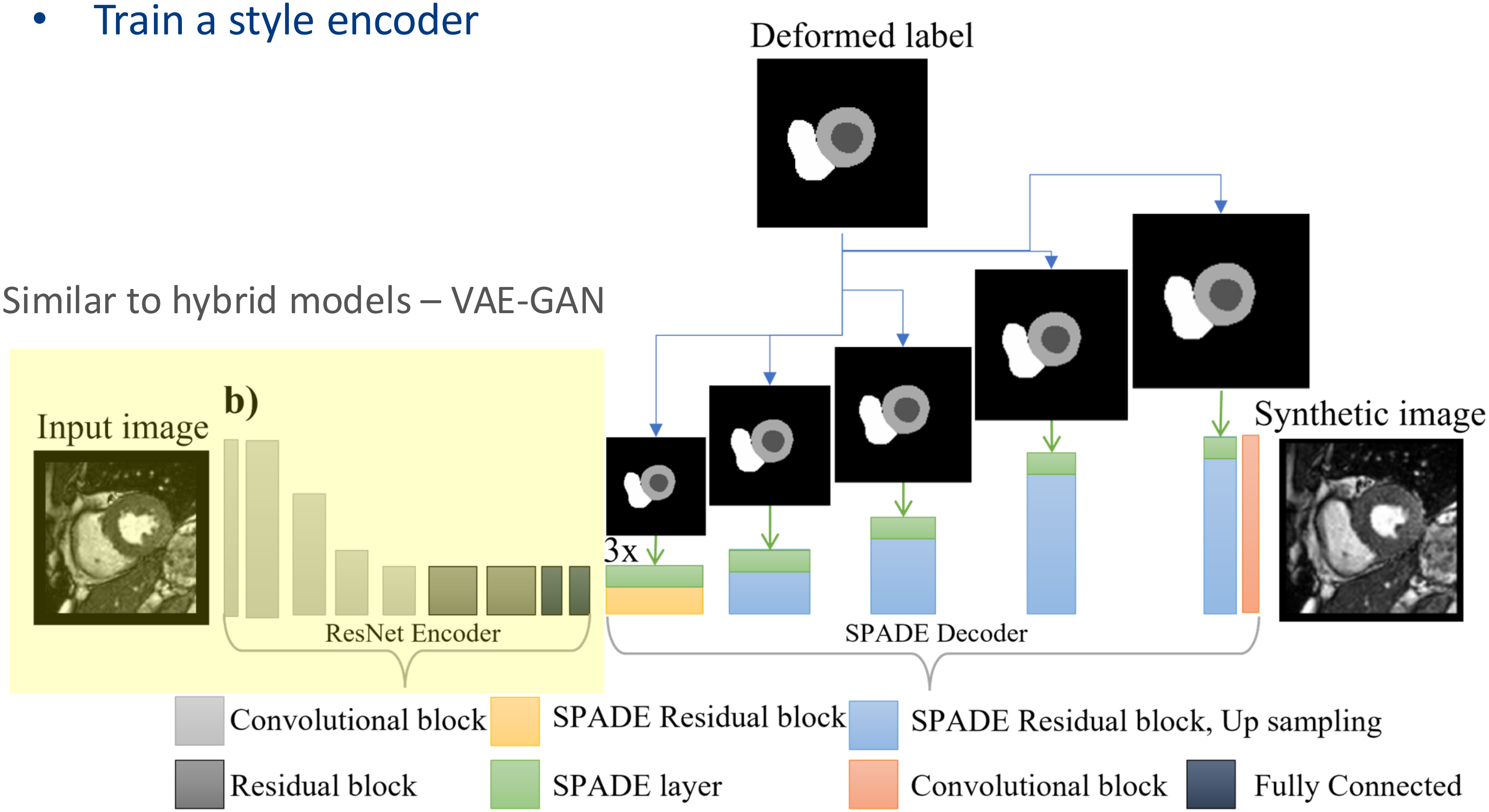
Style 2



How to control the style of the image?

- Train a style encoder

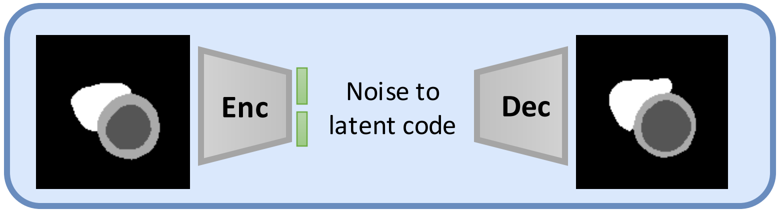
Similar to hybrid models – VAE-GAN



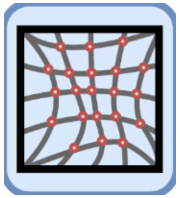
How to deform labels

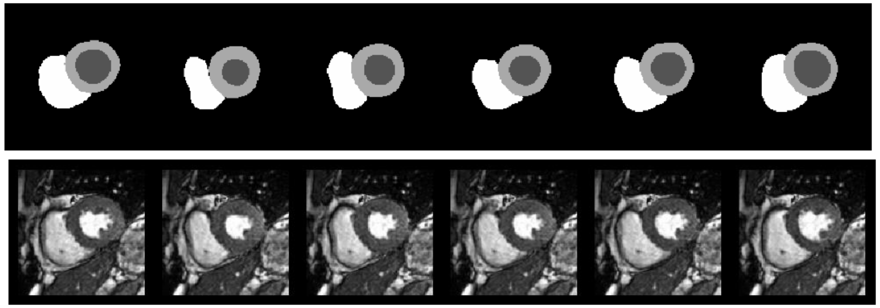
- Elastic deformation and morphological operations
- VAE to learn anatomical deformation

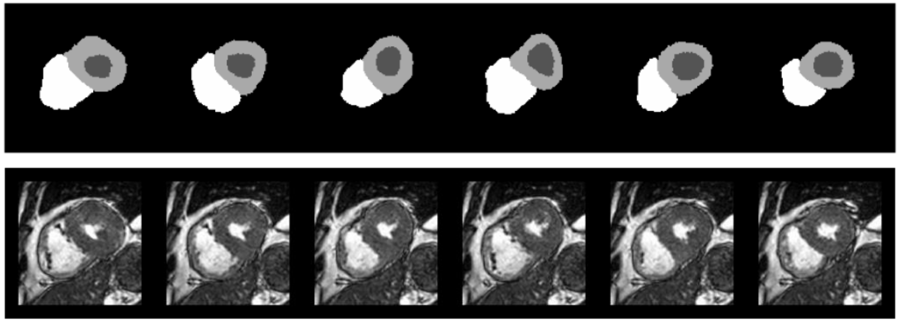
VAE based deformation



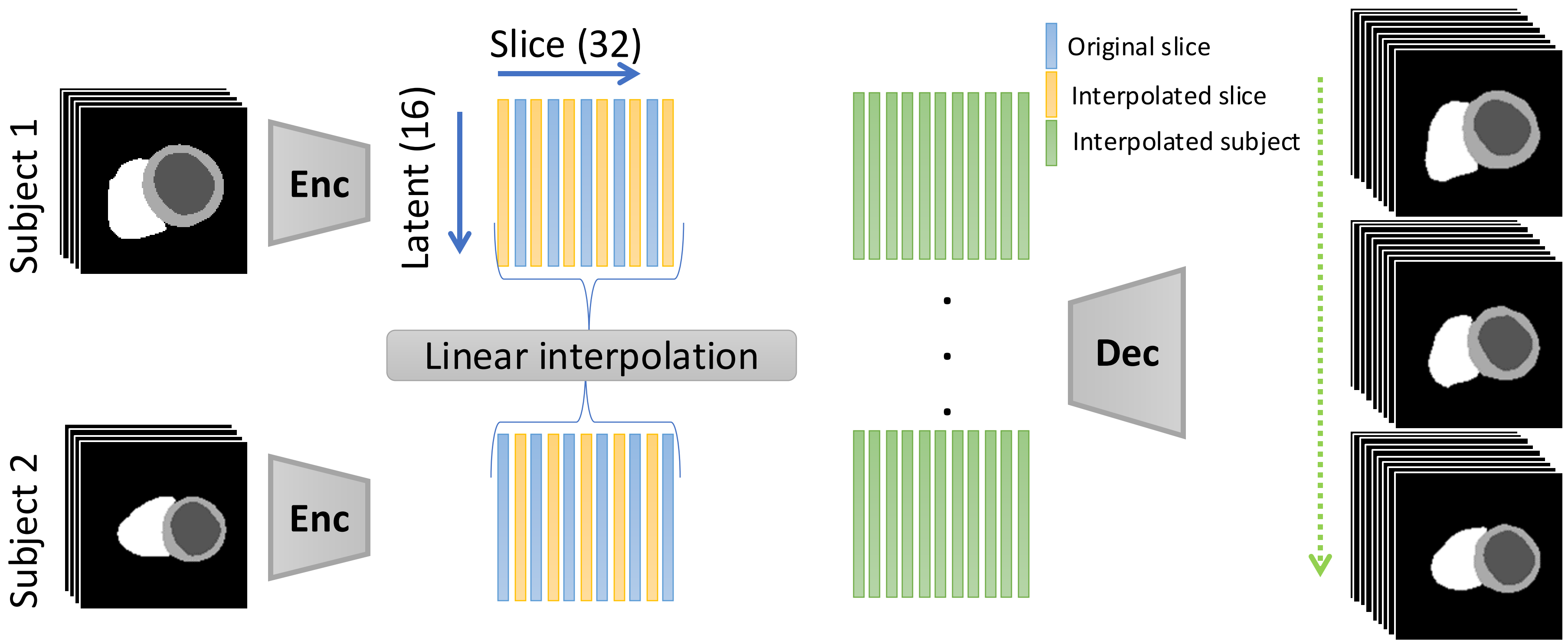
Random elastic deformation



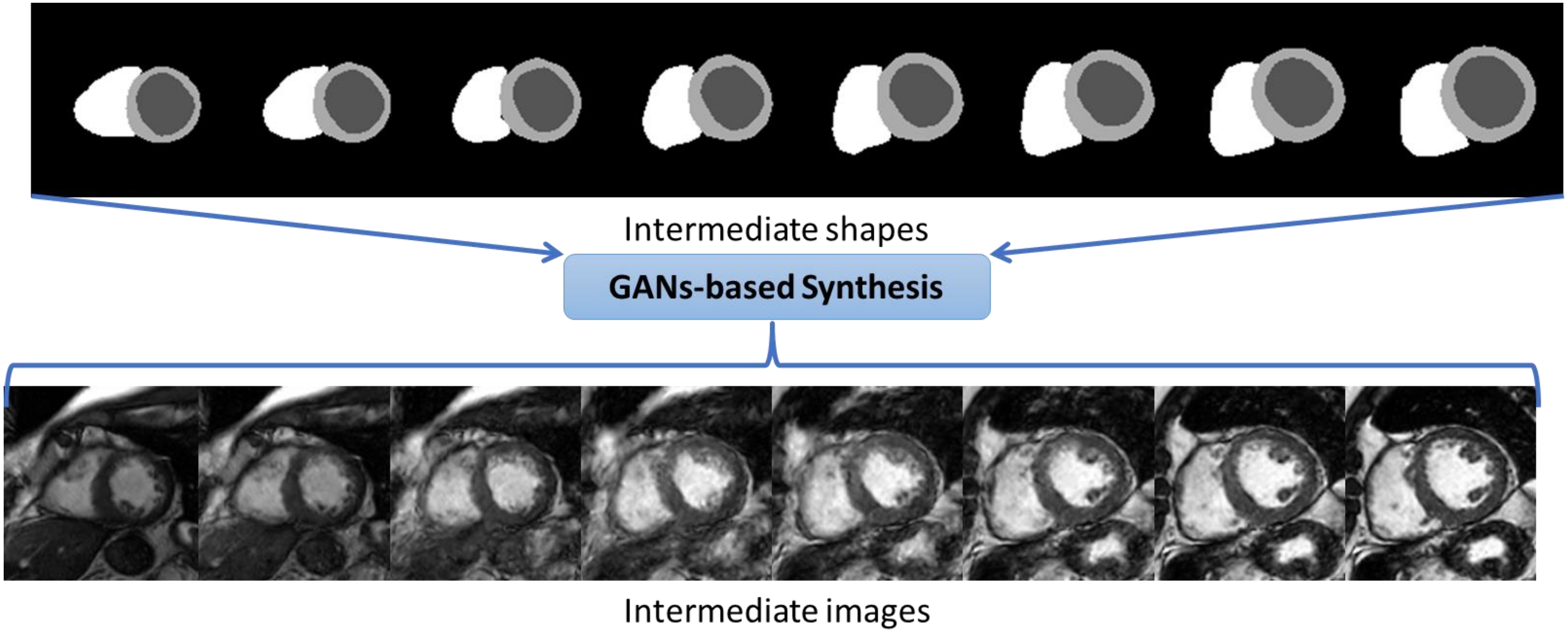




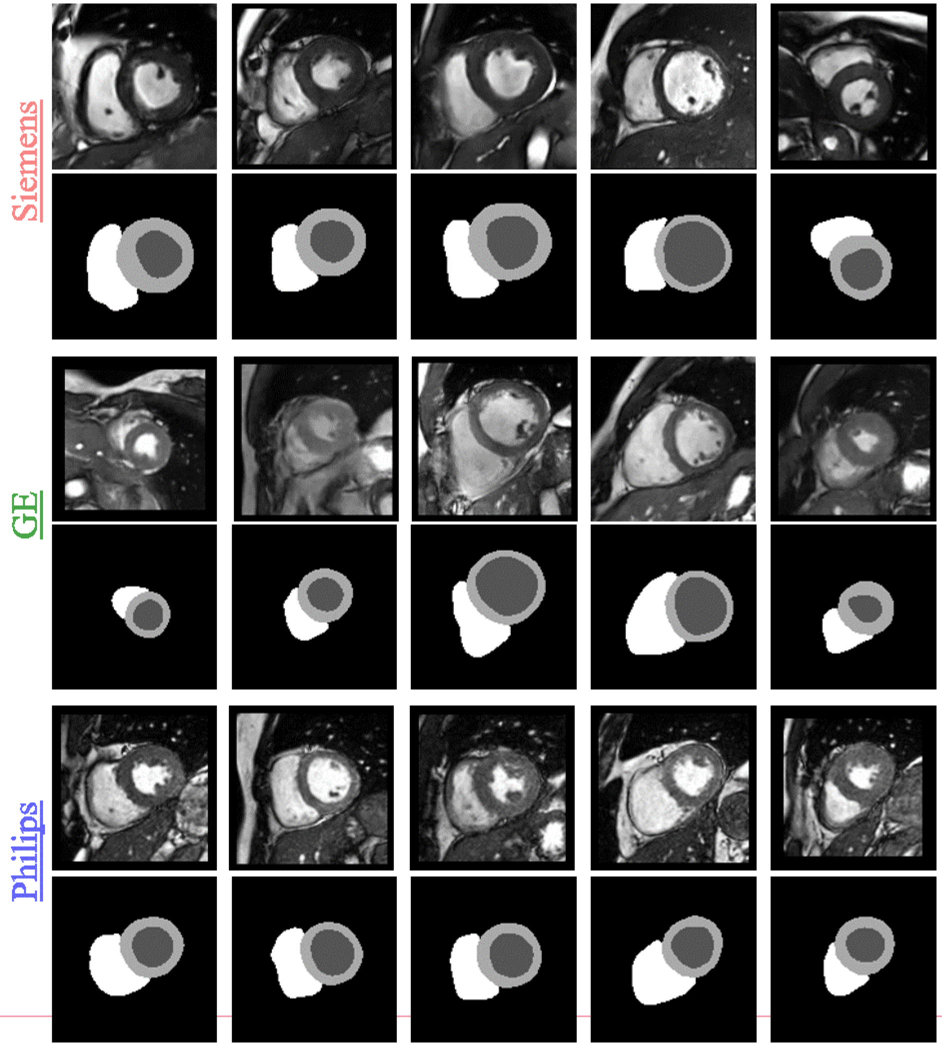
VAE to generate images with heart geometry in-between two subjects



VAE to generate images with heart geometry in-between two subjects

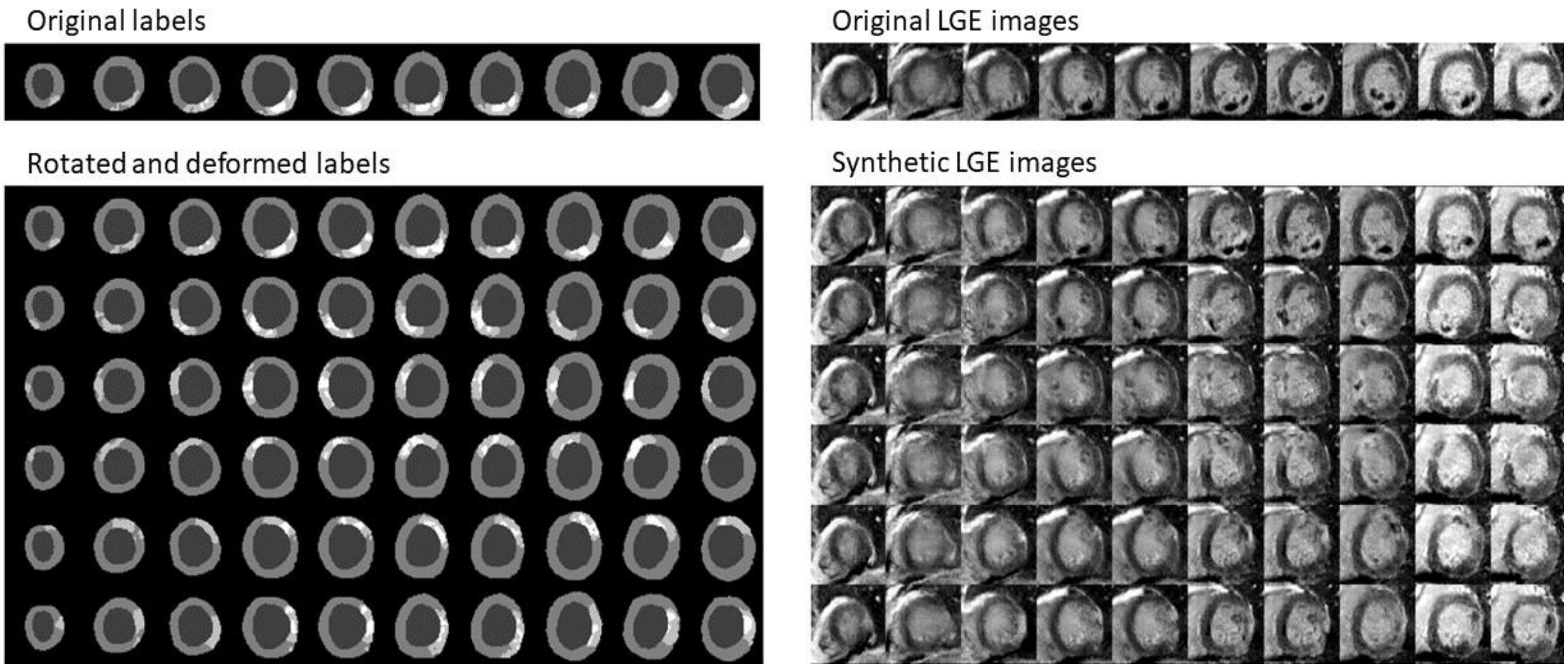


- Synthesizing data for each scanner vendor



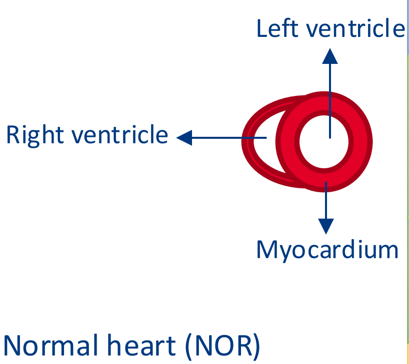
Myocardial scar synthesis

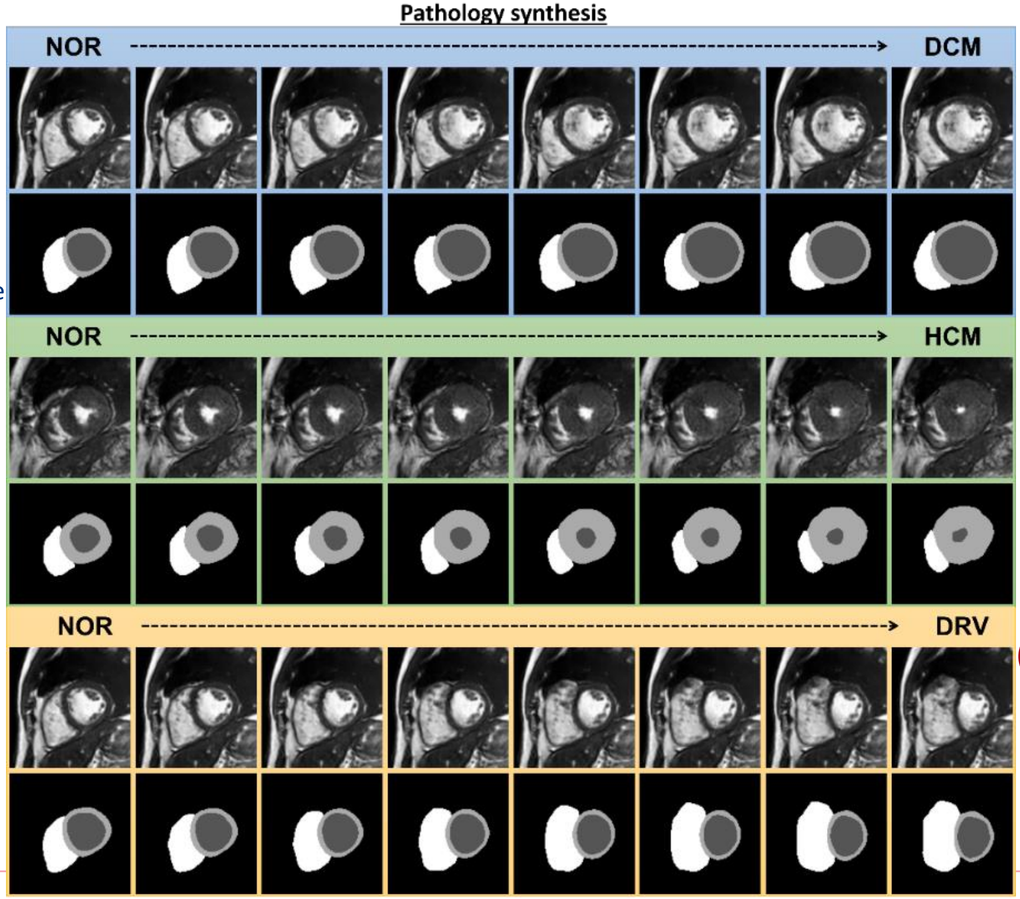
- *Augmentation with synthetic data* ***improves myocardial scar quantification*** *in late-gadolinium enhancement (LGE)*

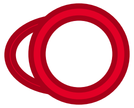


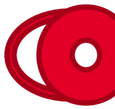
Targeted synthesis to increase pathological cases

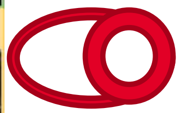
- *Subjects with plausible anatomy and heart disease characteristics are synthesized in pathology synthesis*



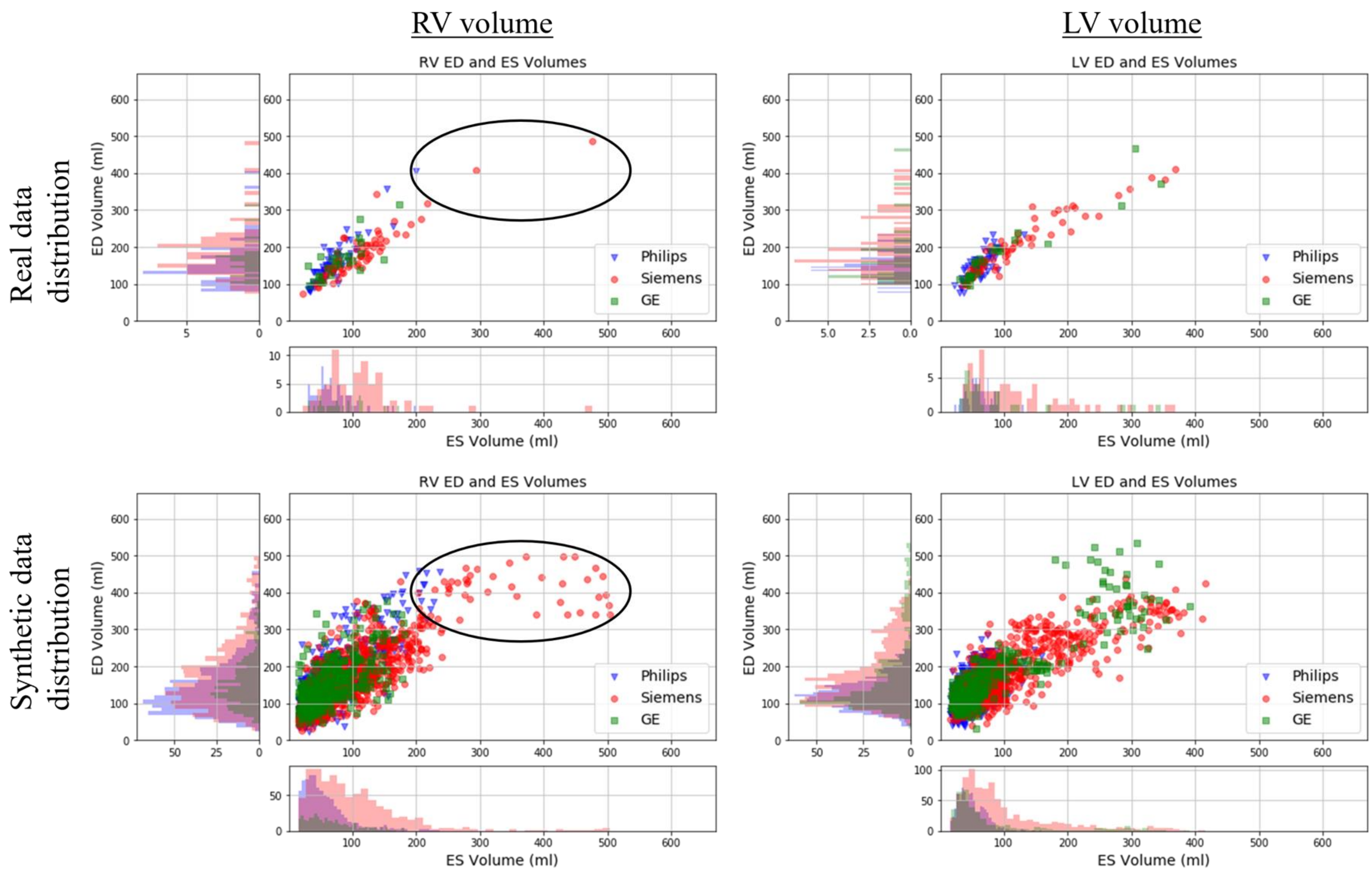








Targeted synthesis to increase the outlier cases



Conclusions

- Image synthesis can provide solutions to
  - Generate substantial number of images with variations in anatomy and contrast
  - Tackle medical data scarcity for deep-learning applications
  - Boost the performance of segmentation models
  - Improve the robustness of models to handle subjects with pathology

Questions?