Generative Modelling in Machine Learning for Medical Imaging

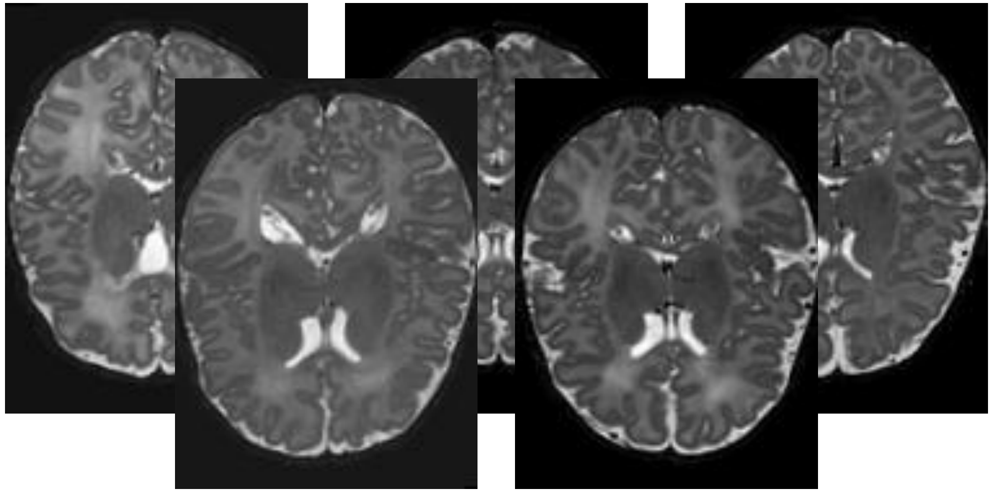
Benjamin Hou | BioMedical Image Analysis
Imperial College London

BioMedIA

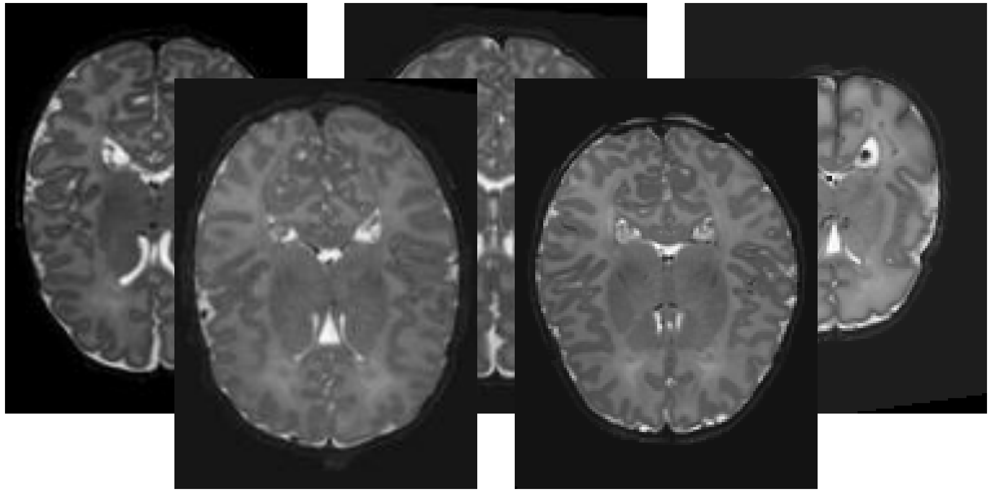
Overview

- Part 1
  - What are Generative Models?
  - Generative Adversarial Models (GANs)
  - Variational Auto-Encoders (VAEs) and Conditional VAEs (C-VAEs)
  - Normalizing Flow Models
- Part 2
  - 3D Fetal Skull Reconstruction from 2DUS via Deep Conditional Generative Networks
  - Conditional Image Generation of Missing Data with Deep Mental Maps

What are Generative Models?



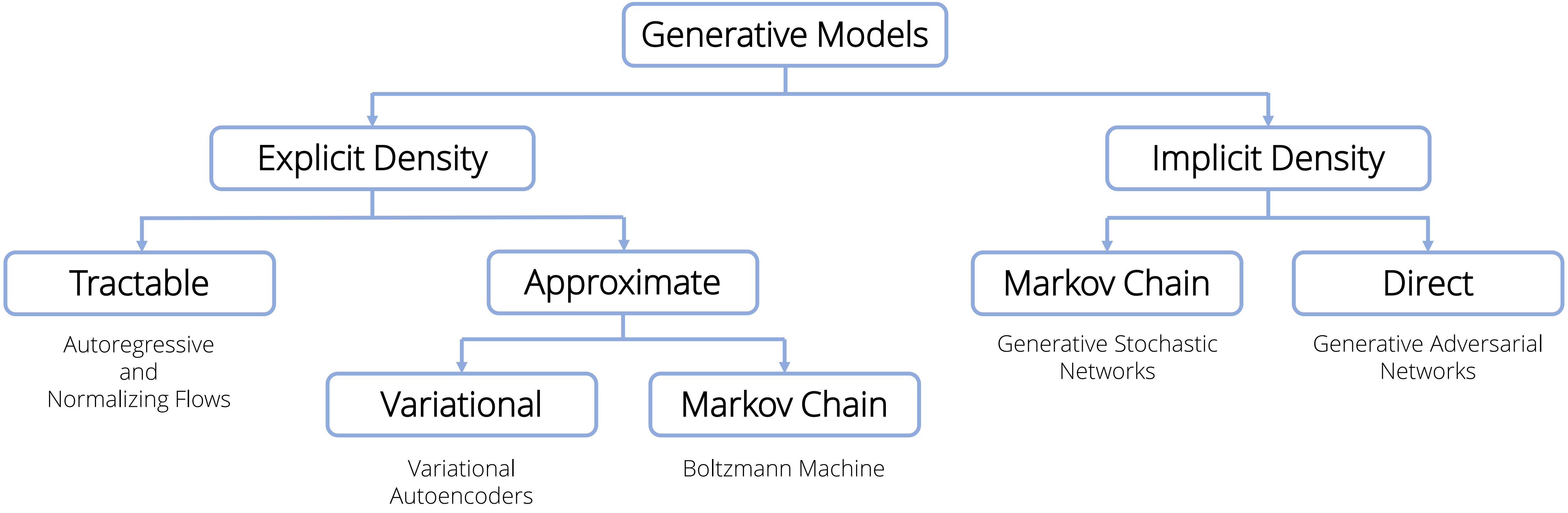
Training Images $\sim p_{data}(x)$



Sample Images $\sim p_{model}(x)$

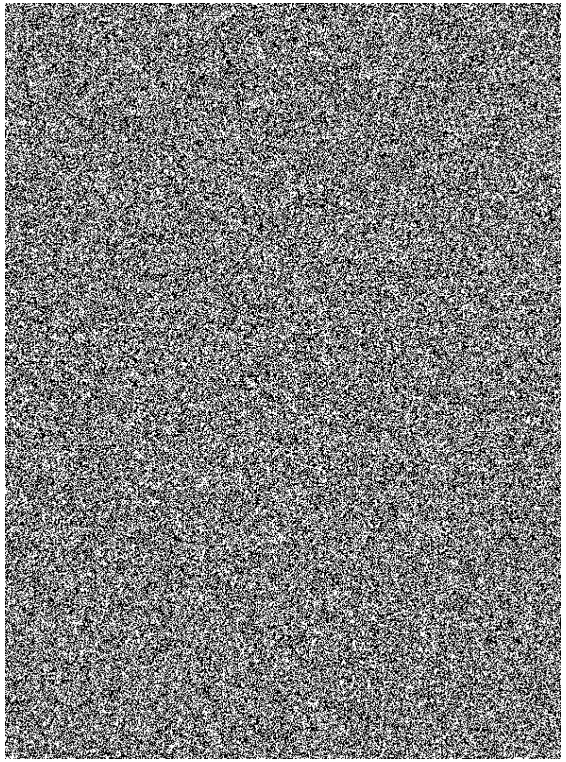
$$p_{data}(x) \approx p_{model}(x)$$

What are Generative Models?



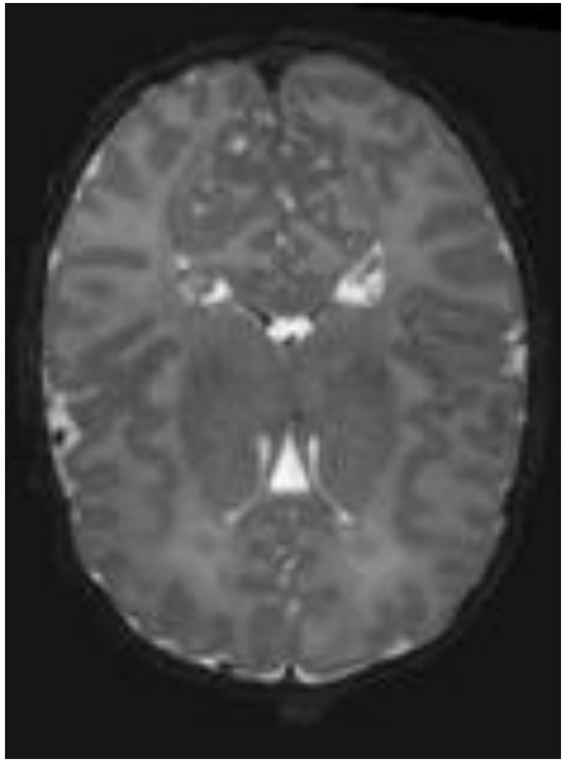
*Goodfellow et al. - NIPS 2016 Tutorial: Generative Adversarial Networks

What are Generative Models?



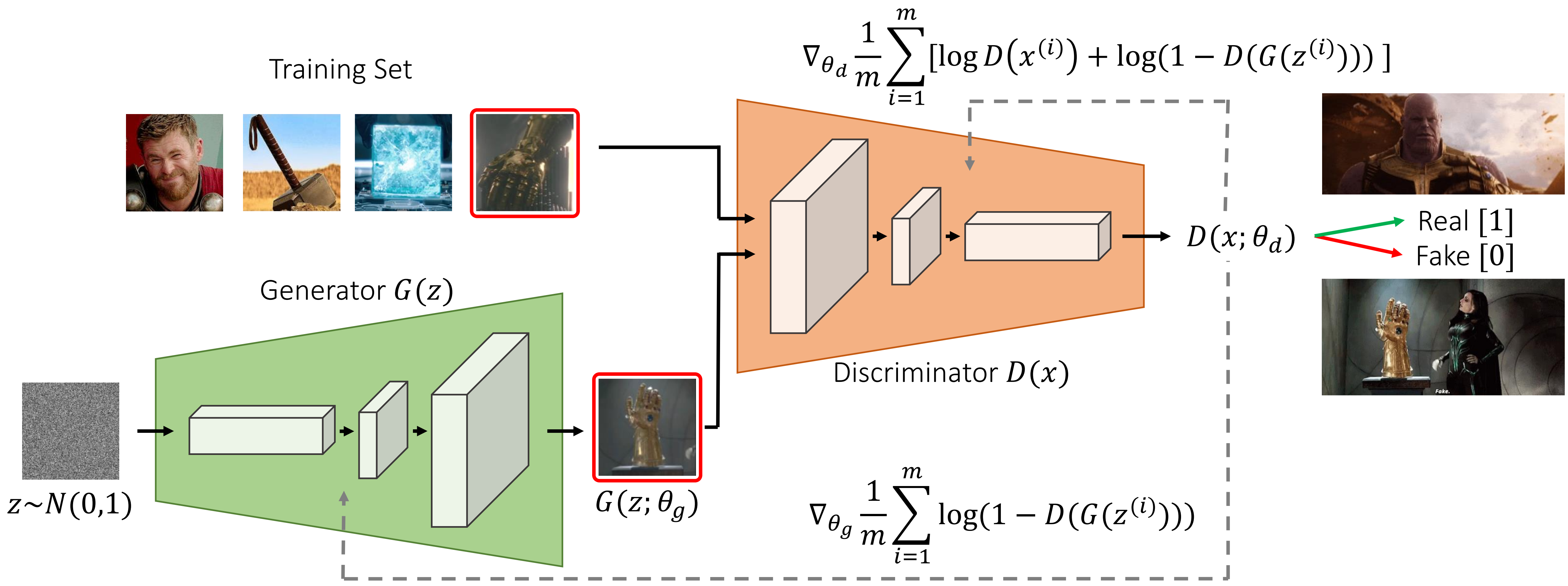
$z \sim \mathcal{N}(0,1)$

$x = f(z;\theta)$



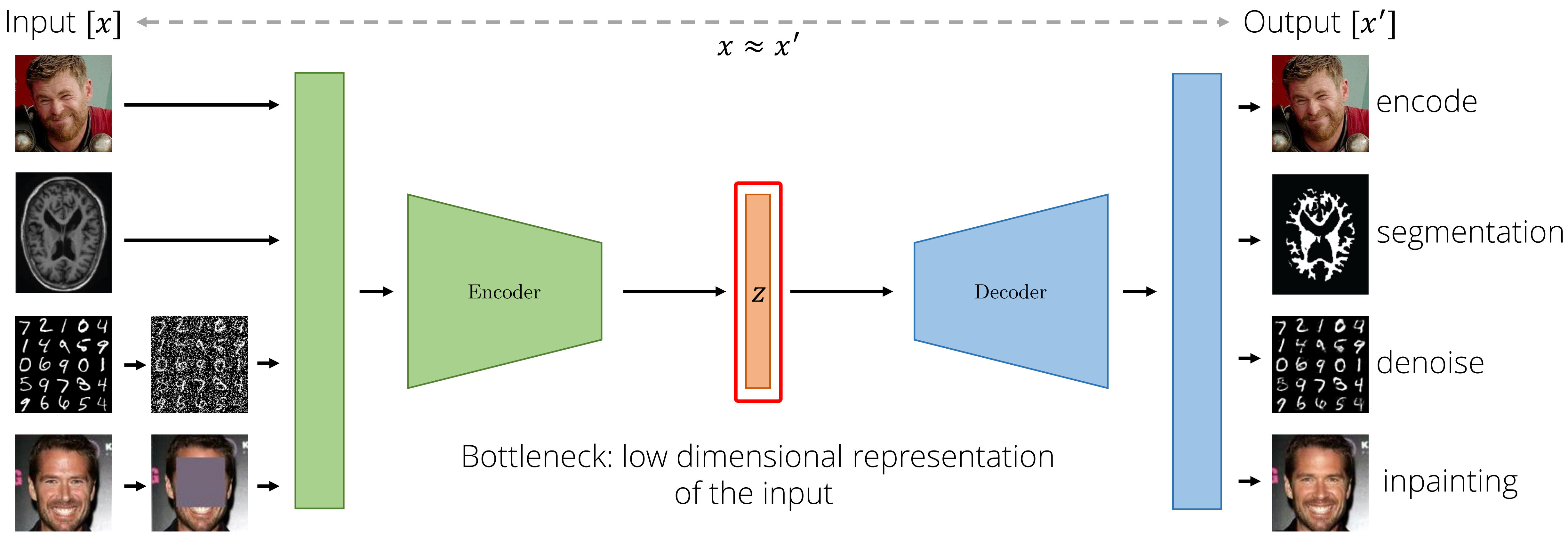
$x \sim \mathcal{N}(f(z;\theta), \sigma^2 I)$

Generative Adversarial Networks (GANs)



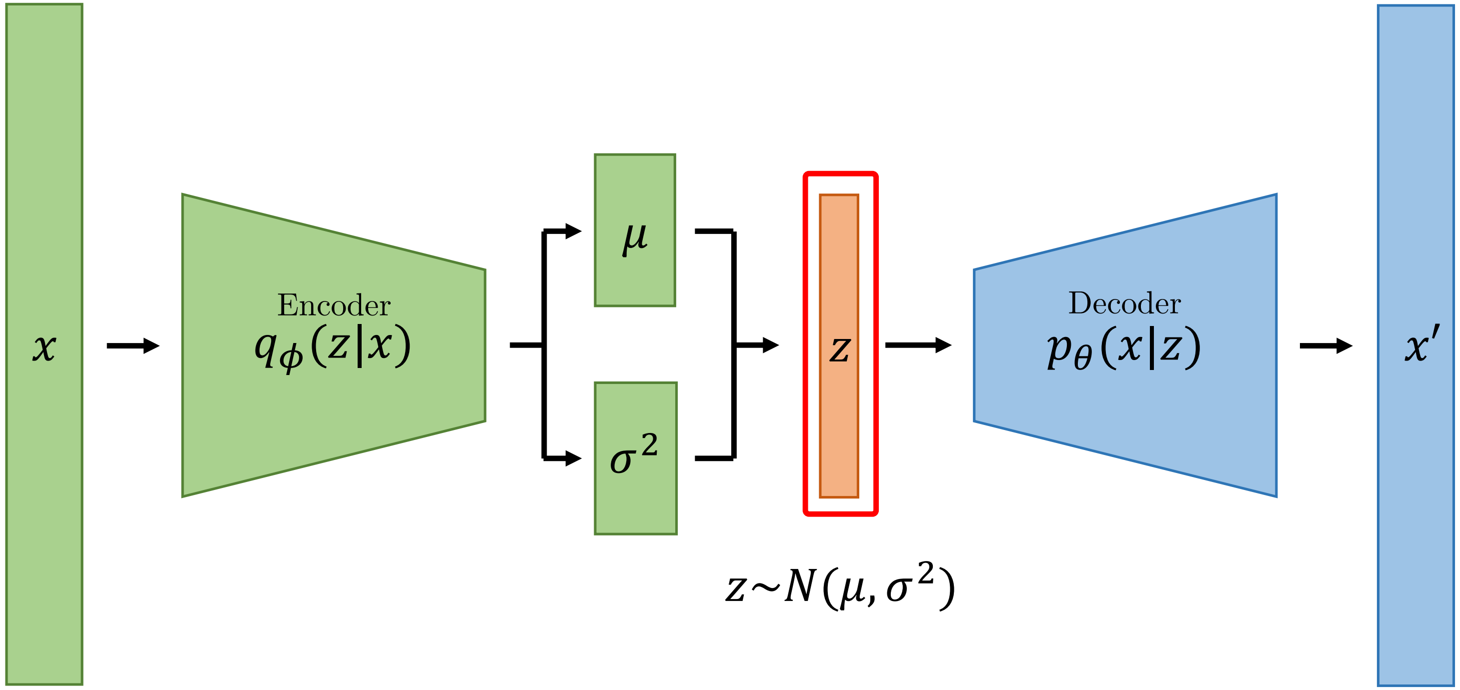
$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z)))]$$

Auto-Encoders



$$Loss = \|x - x'\|_2^2$$

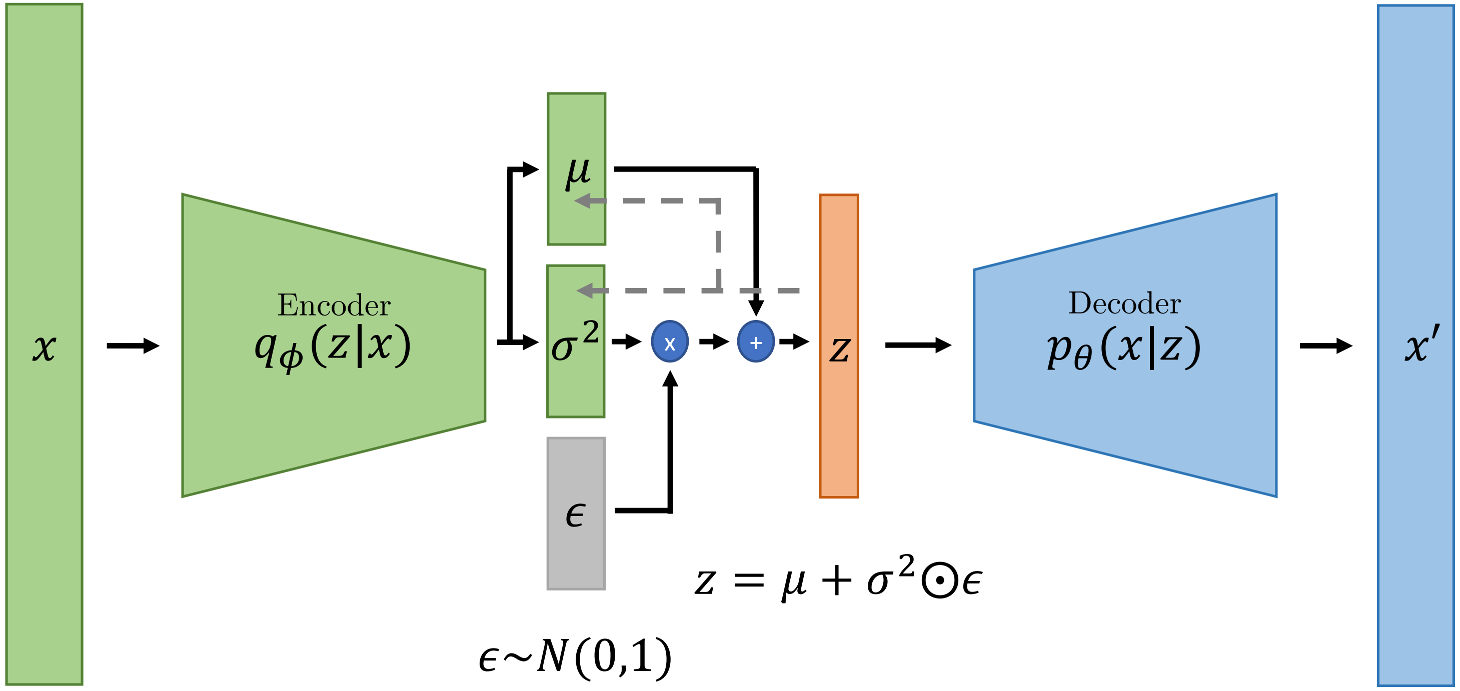
Variational Auto-Encoders



Reconstruction Loss — KL Divergence

$$L(x, z; \theta, \phi) = \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] - KL[q_\phi(z|x) || p(z)]$$

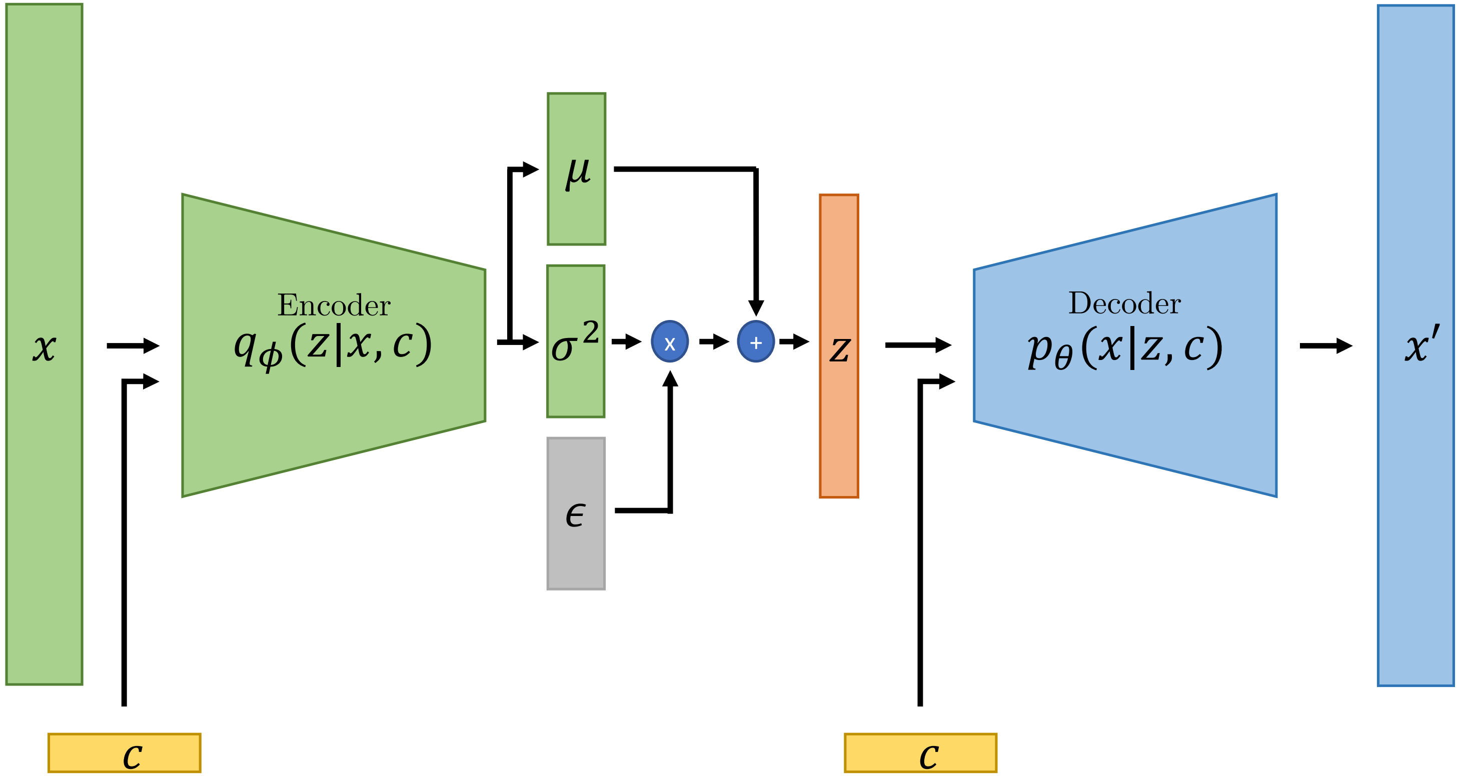
Variational Auto-Encoders



Reconstruction Loss KL Divergence

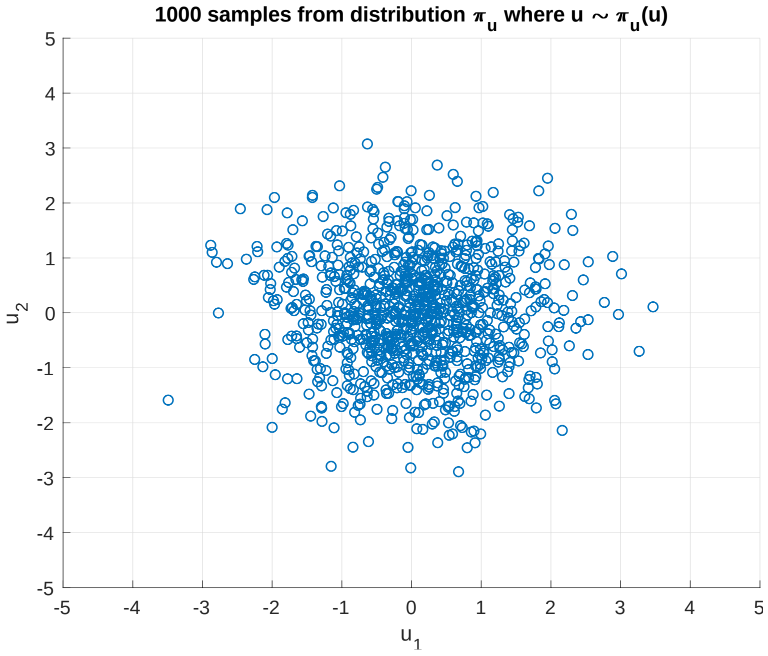
$$L(x, z; \theta, \phi) = \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] - KL[q_\phi(z|x)||p(z)]$$

Conditional Variational Auto-Encoders



$$L(x,z,c;\theta,\phi) = \mathbb{E}_{q_\phi(z|x,c)}[\log p_\theta(x|z,c)] - KL[q_\phi(z|x,c)||p(z)]$$

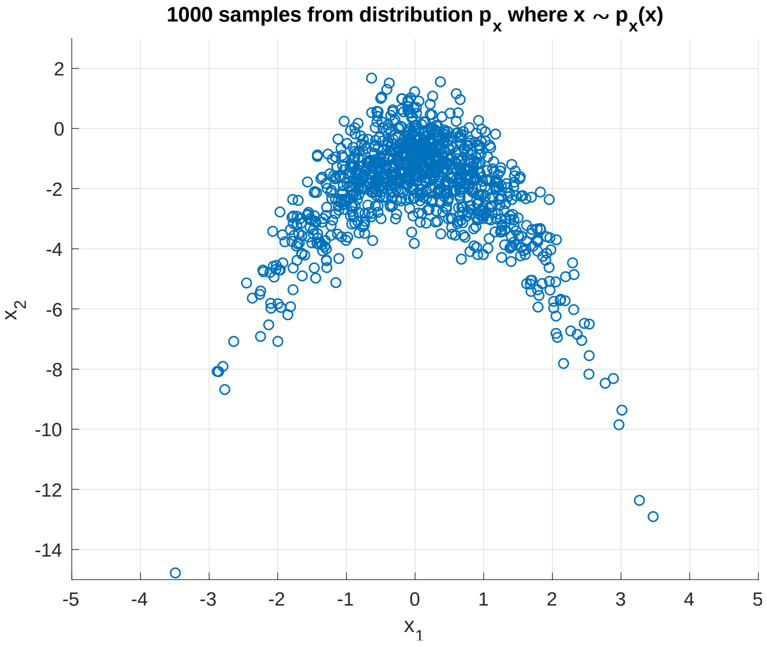
Normalizing Flows



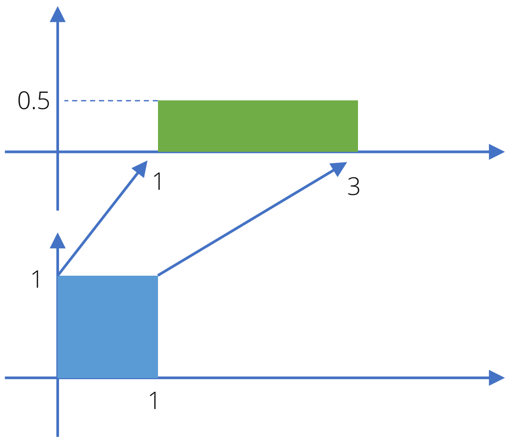
$$x = f(u) = \begin{bmatrix} u_1 \\ u_2 - u_1^2 - 1 \end{bmatrix}$$

f must be invertible

$$u = f^{-1}(x) = \begin{bmatrix} x_1 \\ x_2 + x_1^2 + 1 \end{bmatrix}$$



Normalizing Flows



$$\int \pi_u(u)\partial u = \int p_x(x)\partial x$$

Total probability must be preserved (i.e. =1)

$$p_x(x) = \pi_u(u)\left|\frac{\partial u}{\partial x}\right|$$

Recall:
$x = f(u)$
and
$u = f^{-1}(x)$

$$p_x(x) = \pi_u(f^{-1}(x))\left|\frac{\partial}{\partial x} f^{-1}(x)\right|$$

Mapped probability density

$$p_x(x) = \pi_u(f^{-1}(x))|\det J(f^{-1}(x))|$$

$$\log p_x(x) = \log \pi_u(f^{-1}(x)) + \log|\det J(f^{-1}(x))|$$

Normalizing Flows

$$\log p_x(x) = \log \pi_u(f^{-1}(x)) + \log|\det J(f^{-1}(x))|$$

We estimate this

We know this
P.D.F. of Normal Distribution

By learning this

We can then use MLE:

$$\theta = \arg\max_\theta \log p_x(x;\theta)$$

Where θ are the weights of f

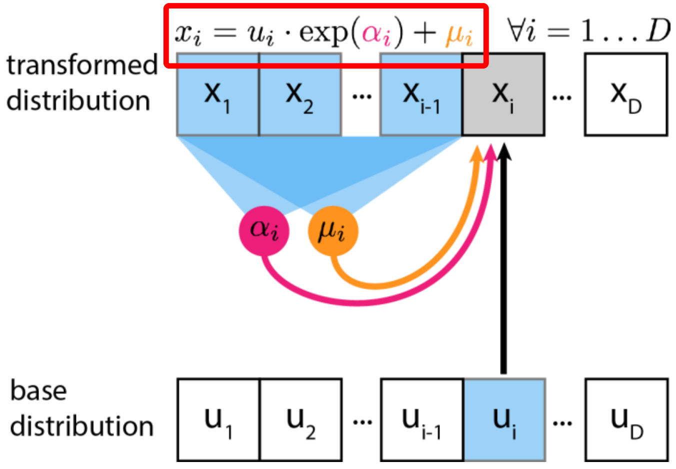
How to we construct f so that it's invertible?

Normalizing Flows

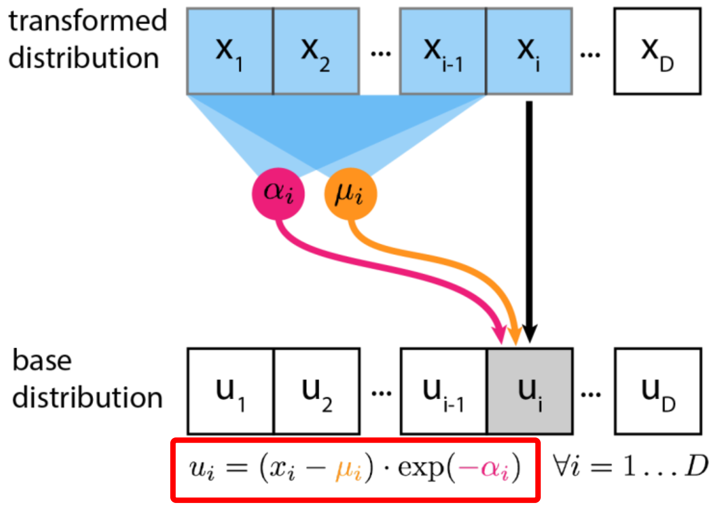
- 2014 – Dinh et al. – NICE: Non-linear Independent Components Estimation
- 2015 – Germain et al. – MADE: Masked Autoencoder for Distribution Estimation
- 2016 – van den Oord et al. – Pixel Recurrent Neural Networks
- 2016 – van den Oord et al. – WaveNet: A Generative Model for Raw Audio
- 2016 – Kingma et al. - Improving Variational Inference with Inverse Autoregressive Flow
- 2016 – Dinh et al. – Density estimation using Real NVP
- 2017 – Papamakarios et al. – Masked Autoregressive Flow
- 2017 – van den Oord et al. - Parallel WaveNet: Fast High-Fidelity Speech Synthesis
- 2018 – Kingma et al. – Glow: Generative Flow with Invertible 1x1 Convolutions

Normalizing Flows

Masked Autoregressive Flow Forward



Masked Autoregressive Flow Inverse



Neural Networks

$$\mu_i = \psi_{\mu_i}(x_{1:i-1})$$

$$\alpha_i = \psi_{\alpha_i}(x_{1:i-1})$$

$$x_i = f(u_i) = u_i \cdot \exp(\alpha_i) + \mu_i$$

$$u_i = f^{-1}(x_i) = (x_i - \mu_i) \cdot \exp(-\alpha_i)$$

*Jiang et al. – Normalizing Flows Tutorial, Part 2: Modern Normalizing Flows

Normalizing Flows

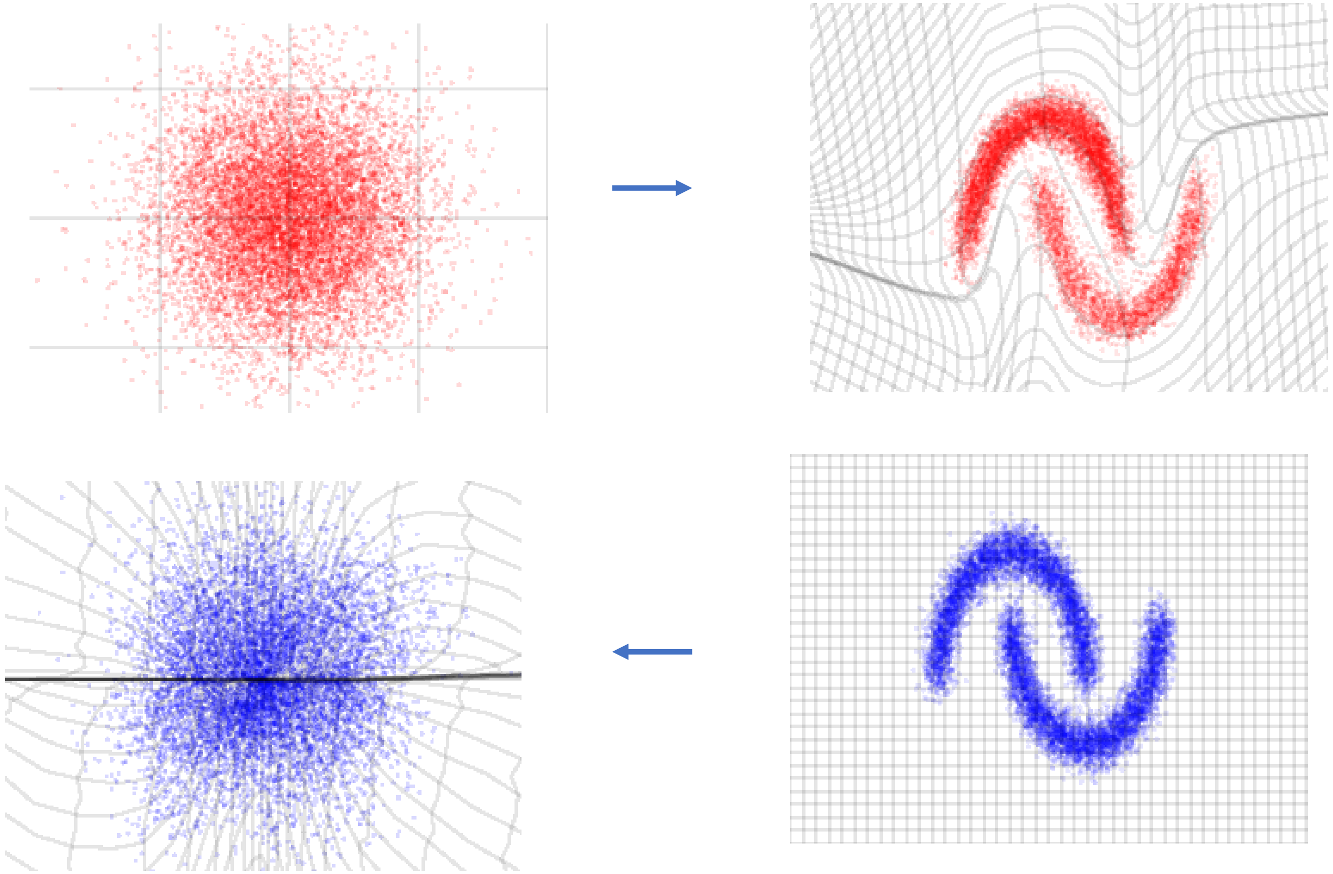
Latent Distribution π_u | Data Distribution p_x

Forward

$$u \sim \pi_u(u)$$
$$x = f(u)$$

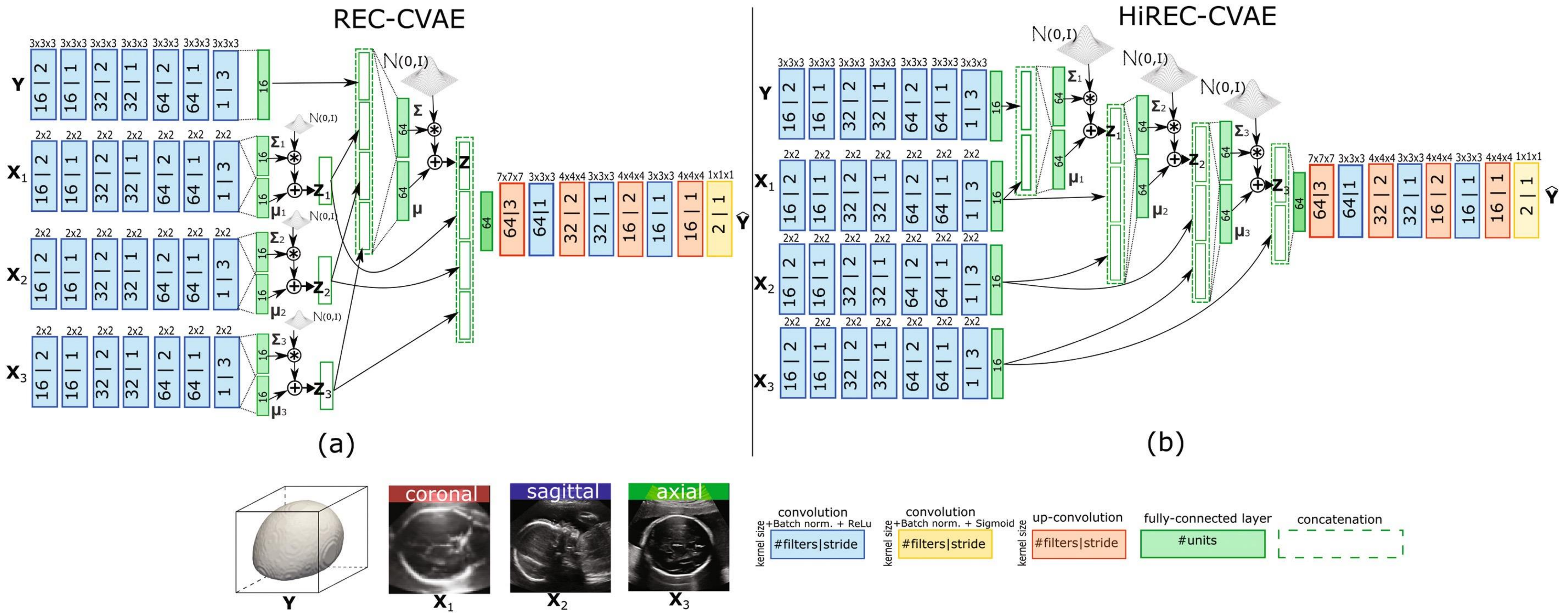
Inverse

$$x \sim p_x(x)$$
$$u = f^{-1}(x)$$



*Dinh et al. – Density Estimation Using RealNVP

3D Fetal Skull Reconstruction from 2DUS via Deep Conditional Generative Networks



*J. J. Cerrolaza, Y. Li, C. Biffi, J. Matthew, M. Sinclair, A. Gomez, B. Kainz, D. Rueckert

3D Fetal Skull Reconstruction from 2DUS via Deep Conditional Generative Networks

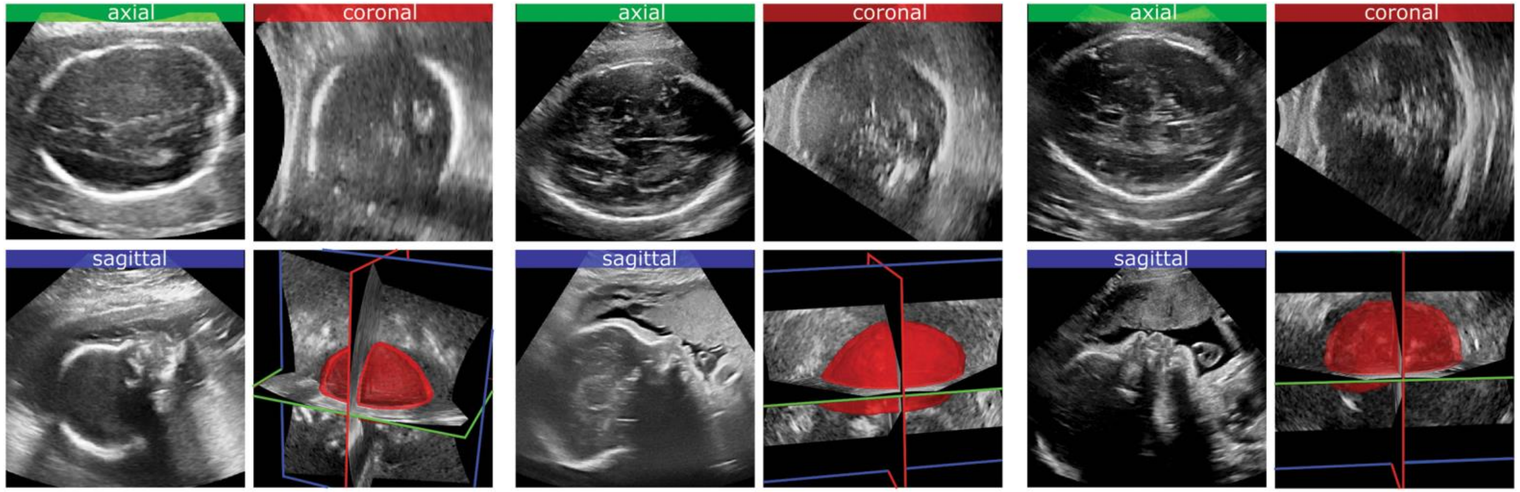
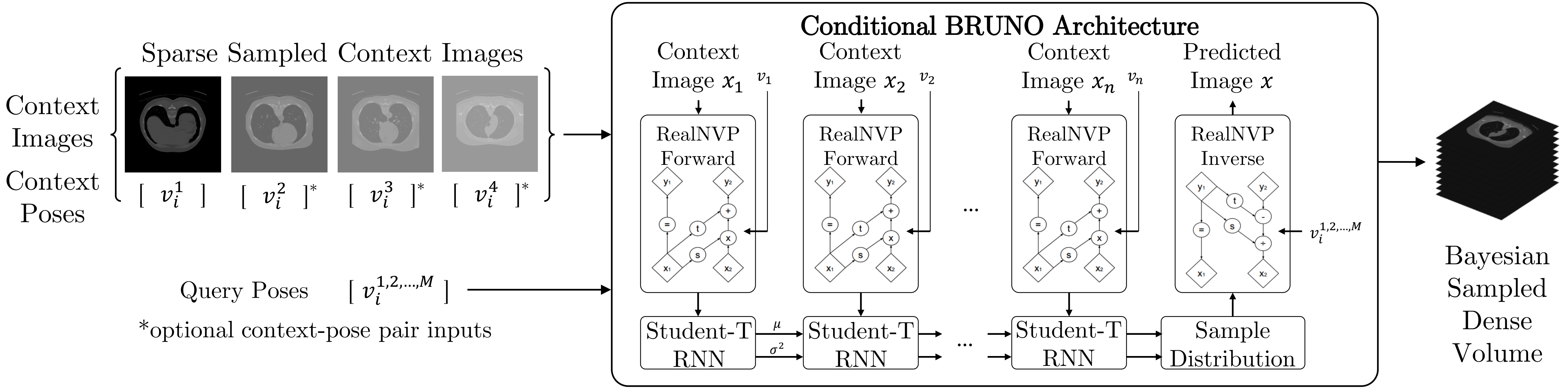


Table 1. The table presents the average and standard deviation for the Dice's coefficient (DC), sensitivity (SEN.), and precision (PPV) of the reconstruction of the fetal skull, and the effect of using three, two, or one standard US views as predictors.

	axial + sagittal + coronal			axial + sagittal			axial		
	DC	SEN.	PPV	DC	SEN.	PPV	DC	SEN.	PPV
REC-CVAE	0.91 ± 0.02	0.91 ± 0.05	0.91 ± 0.06	0.86 ± 0.05	0.88 ± 0.13	0.87 ± 0.09	0.83 ± 0.06	0.86 ± 0.15	0.84 ± 0.13
HiREC-CVAE	0.91 ± 0.04	0.89 ± 0.06	0.93 ± 0.06	0.89 ± 0.05	0.90 ± 0.10	0.91 ± 0.08	0.86 ± 0.05	0.86 ± 0.11	0.90 ± 0.08
TL-net	0.89 ± 0.03	0.93 ± 0.05	0.86 ± 0.07	0.89 ± 0.05	0.90 ± 0.11	0.90 ± 0.06	0.85 ± 0.04	0.88 ± 0.05	0.80 ± 0.09

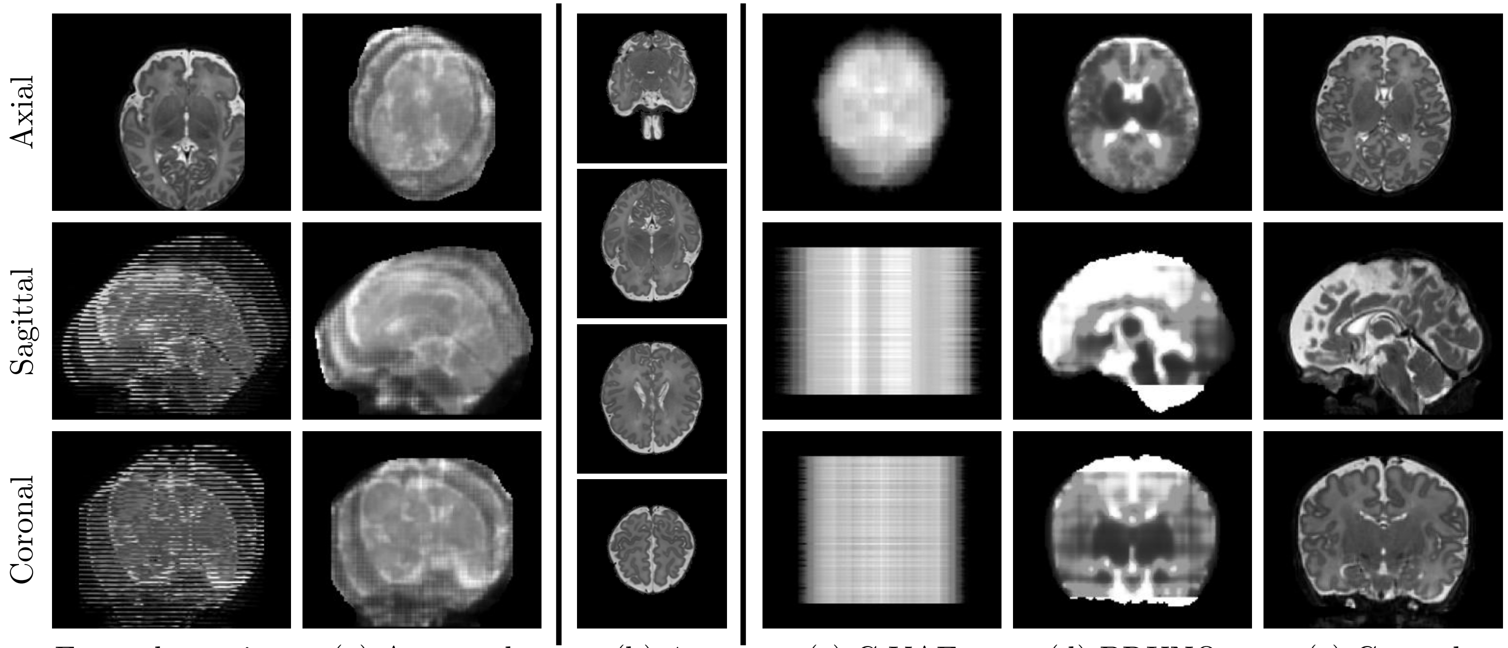
*J. J. Cerrolaza, Y. Li, C. Biffi, J. Matthew, M. Sinclair, A. Gomez, B. Kainz, D. Rueckert

Conditional Image Generation of Missing Data with Deep Mental Maps



*B. Hou, A. Vlontzos, A. Alansary, D. Rueckert, B. Kainz

Conditional Image Generation of Missing Data with Deep Mental Maps



Example motion Corrupted stack | (a) Averaged Template | (b) 4 Context Slices | (c) C-VAE Predicted Template | (d) BRUNO Predicted Template | (e) Ground Truth

*B. Hou, A. Vlontzos, A. Alansary, D. Rueckert, B. Kainz

THANK YOU