A Probabilistic U-Net for Segmentation of Ambiguous Images

Simon A. A. Kohl[1*,2], **Bernardino Romera-Paredes**[1], **Clemens Meyer**[1], **Jeffrey De Fauw**[1], **Joseph R. Ledsam**[1], **Klaus H. Maier-Hein**[2], **S. M. Ali Eslami**[1], **Danilo Jimenez Rezende**[1], and **Olaf Ronneberger**[1]

[1]DeepMind, London, UK
[2]Division of Medical Image Computing, German Cancer Research Center, Heidelberg, Germany
`{simon.kohl,k.maier-hein}@dkfz.de`
`{brp,meyerc,defauw,jledsam,aeslami,danilor,olafr}@google.com`

NeurIPS 2018

Overview

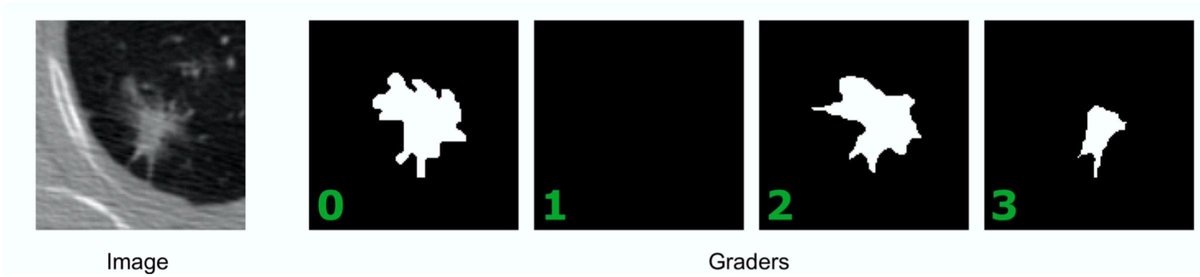
1. Introduction
2. Preliminary
   1. Semantic Segmentation
   2. Uncertainty Estimation
   3. Variational Inference
3. Probabilistic U-Net
4. Experiments

Overview

1. Introduction
2. Preliminary
    1. Semantic Segmentation
    2. Uncertainty Estimation
    3. Variational Inference
3. Probabilistic U-Net
4. Experiments

Introduction

- Inherent ambiguities in medical image analysis
  - Annotation：



  - Deterministic：eliminate disagreements，provide the most likely hypothesis
  - Generative：Distribution over segmentation given an image

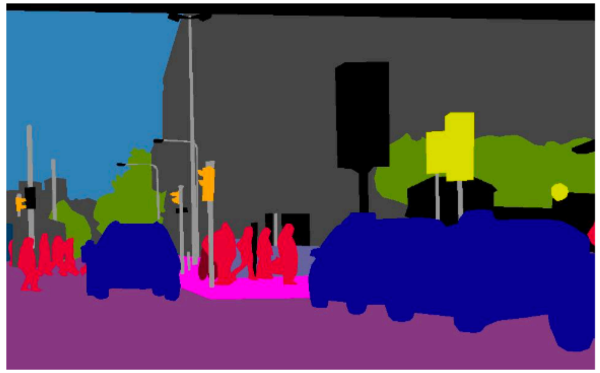
Overview

1. Introduction
2. Preliminary
    1. Semantic Segmentation
    2. Uncertainty Estimation
    3. Variational Inference
3. Probabilistic U-Net
4. Experiments

Semantic Segmentation

- Per-pixel classification
- Context

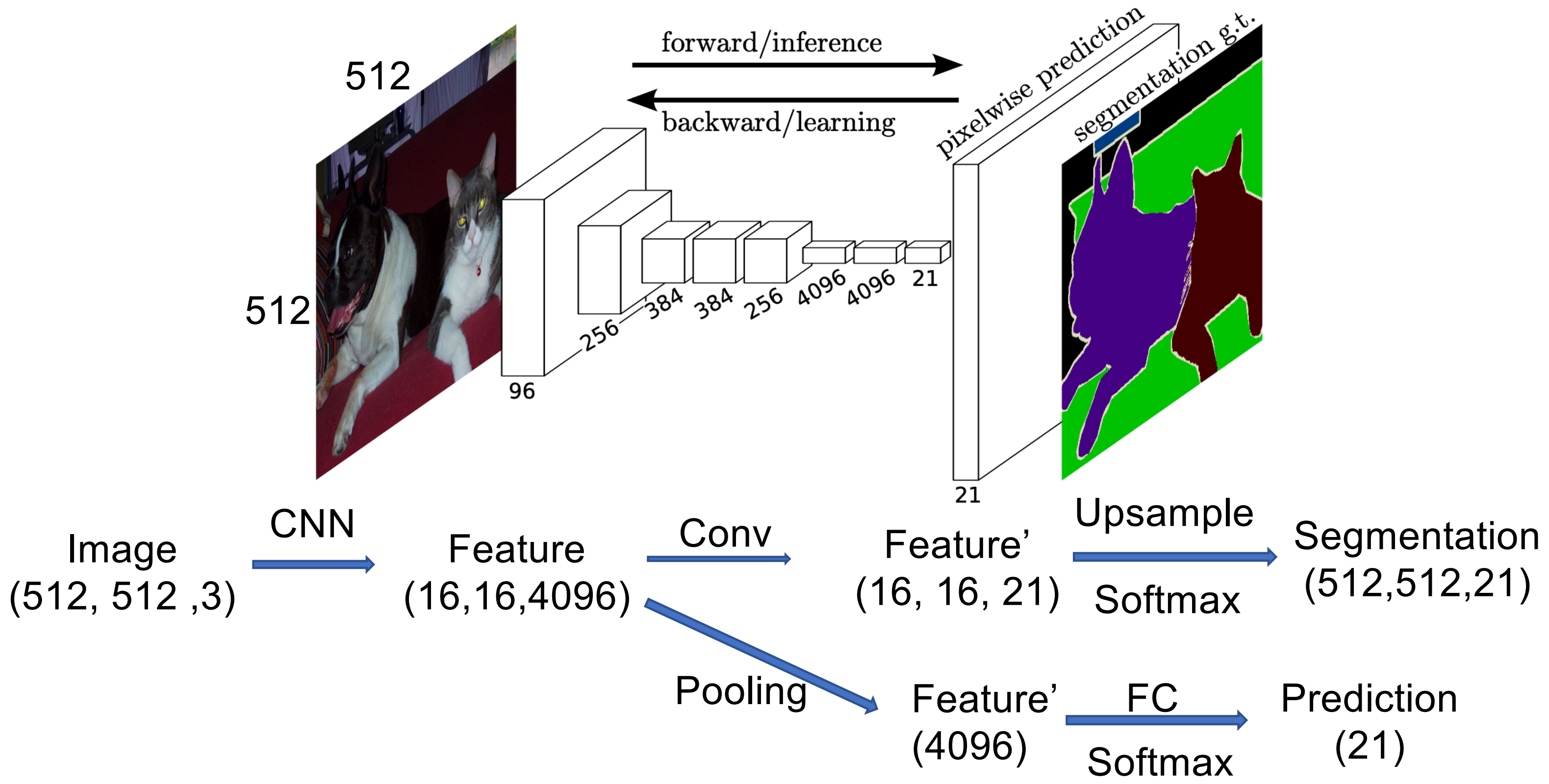
(a) image



(b) semantic segmentation

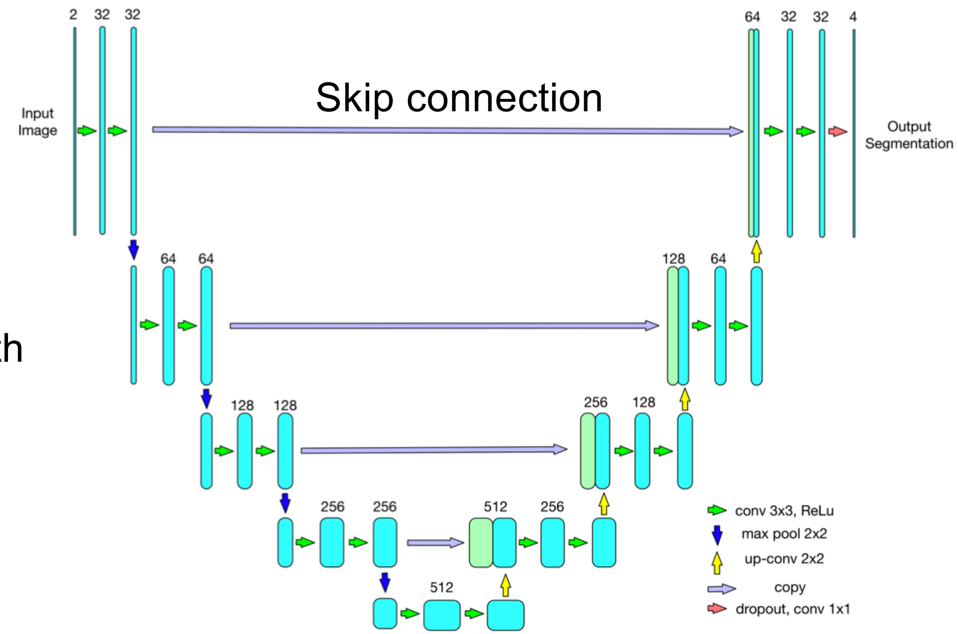
Semantic Segmentation

- Fully Convolutional Networks



Semantic Segmentation

- U-Net



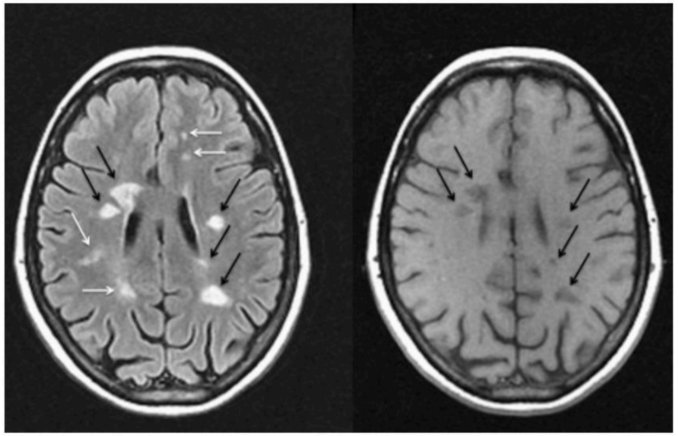
Contracting path
Context

Expanding path
Precise localization

Overview

1. Introduction
2. Preliminary
    1. Semantic Segmentation
    2. Uncertainty Estimation
    3. Variational Inference
3. Probabilistic U-Net
4. Experiments

Uncertainty



Different MRI contrasts (T2/T1). Source: http://www.msdiscovery.org. 2019.

Medical Analysis

Autonomous Driving

高风险的应用，当模型不确定性高时，需要人的介入

Uncertainty in Classification

- Discriminative Model
  - LDA, SVM, CNN…

model → Cat: 2.5 / Dog:0.8 → softmax → Cat: 0.85 / Dog:0.15　　不是概率！！！

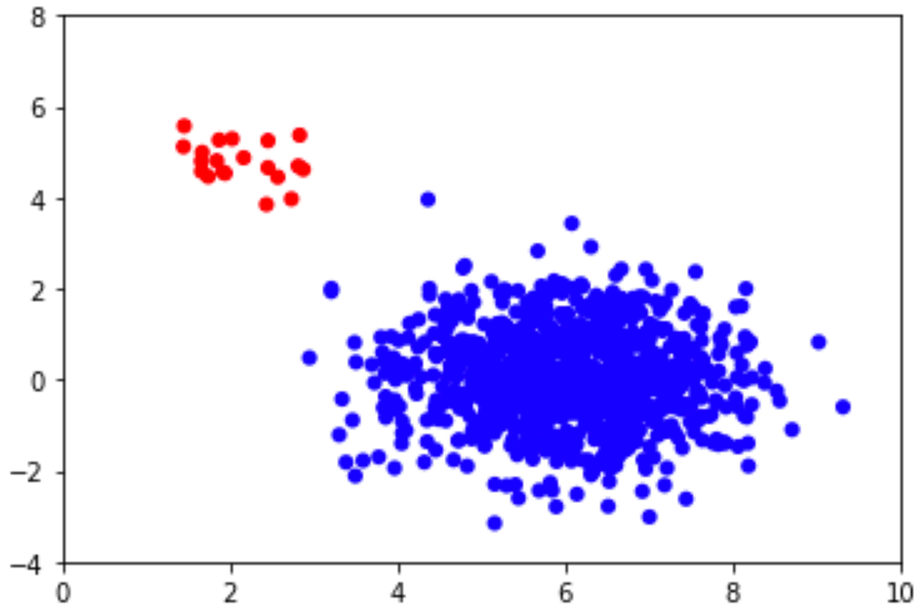
model → Confidence score，90%

- Generative Model
  - 联合概率分布：　$p(x,y) = p(x|y)p(y)$
  - 贝叶斯公式：

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)}$$

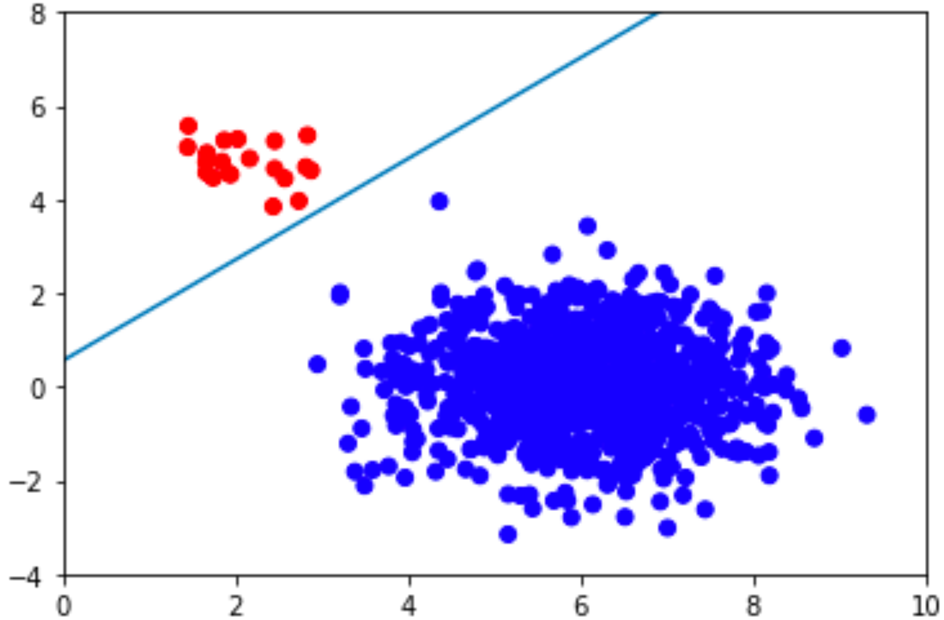
Uncertainty in Classification

- Generate Observations
  - $\mu_R = (2,5), \mu_B = (6,0), \sigma_R = (0.5,0.5), \sigma_B = (1,1)$
  - $R \sim N(\mu_R, \sigma_R^2), B \sim N(\mu_B, \sigma_B^2)$
  - $\mu = \{\mu_R, \mu_B\}, \sigma = \{\sigma_R, \sigma_B\}$
  - $c_i \sim Categorical(0.02, 0.98)$
  - $x_i | c_i, \mu, \sigma \sim N(c_i^T \mu, c_i^T \sigma)$



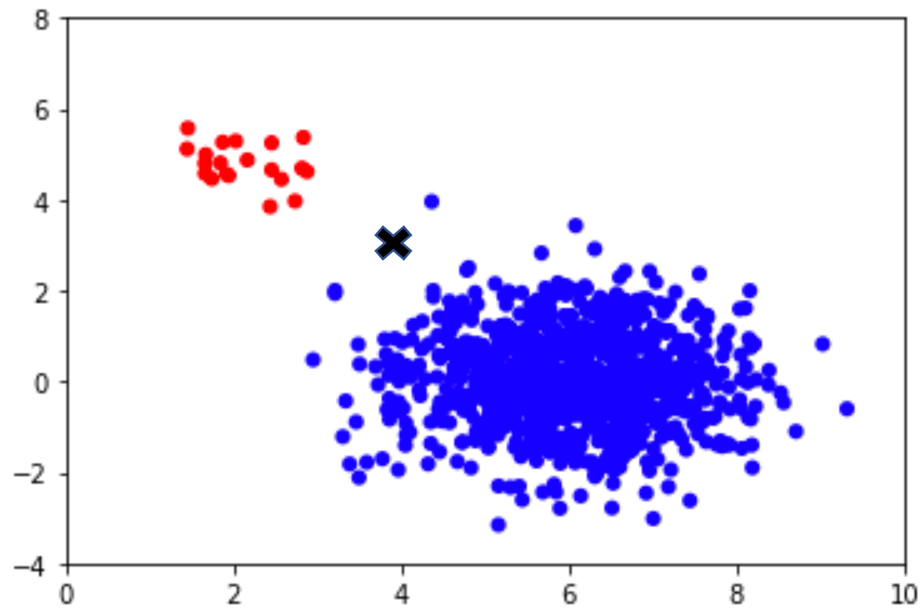
Uncertainty in Classification

- Discriminative

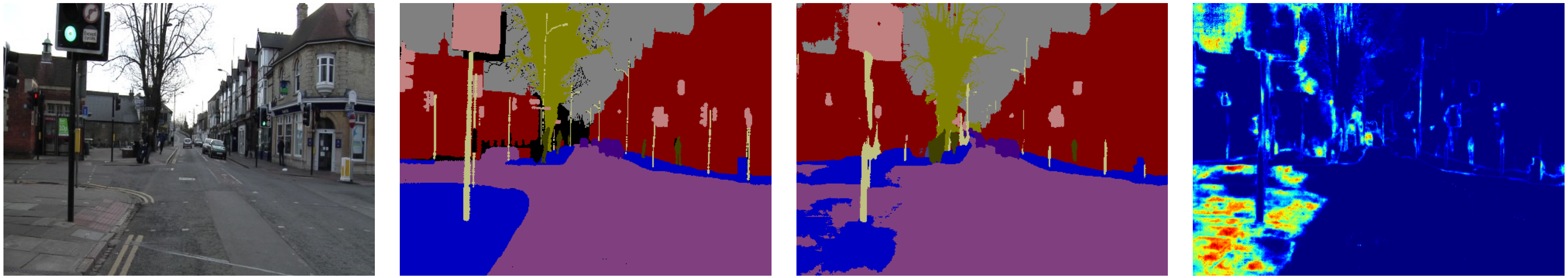


Uncertainty in Classification

- MLE
  - $\hat{u}_R = (2.12, 4.78), \hat{\sigma}_R = (0.46, 0.43)$
  - $\hat{u}_B = (6.03, 0.03), \hat{\sigma}_B = (1.00, 1.00)$
  - $p(R) = 0.02, p(B) = 0.98$
- $x = (4,3)$
- $p(R|x) = \frac{p(x|R)p(R)}{p(x)} = 0.0165$
- $p(B|x) = \frac{p(x|B)p(B)}{p(x)} = 0.9835$

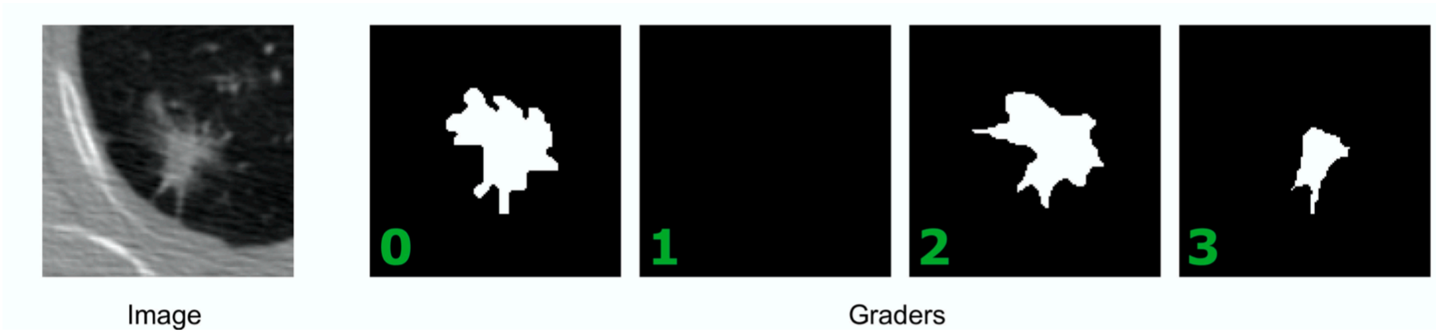


Uncertainty in Segmentation



Input image | Ground truth | Segmentation | Uncertainty Map

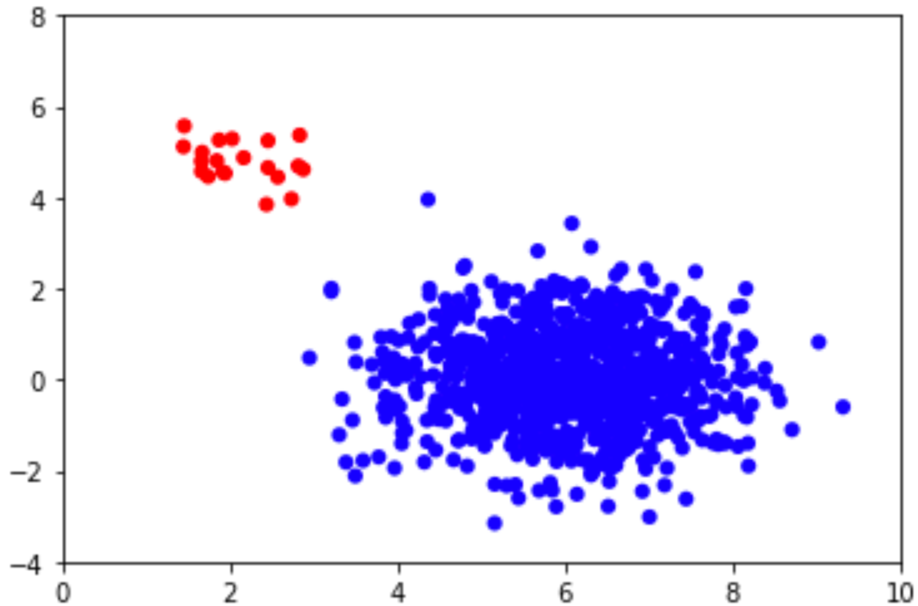
Providing only pixel-wise probabilities ignores all co-variances between the pixels



Overview

Uncertainty in Classification

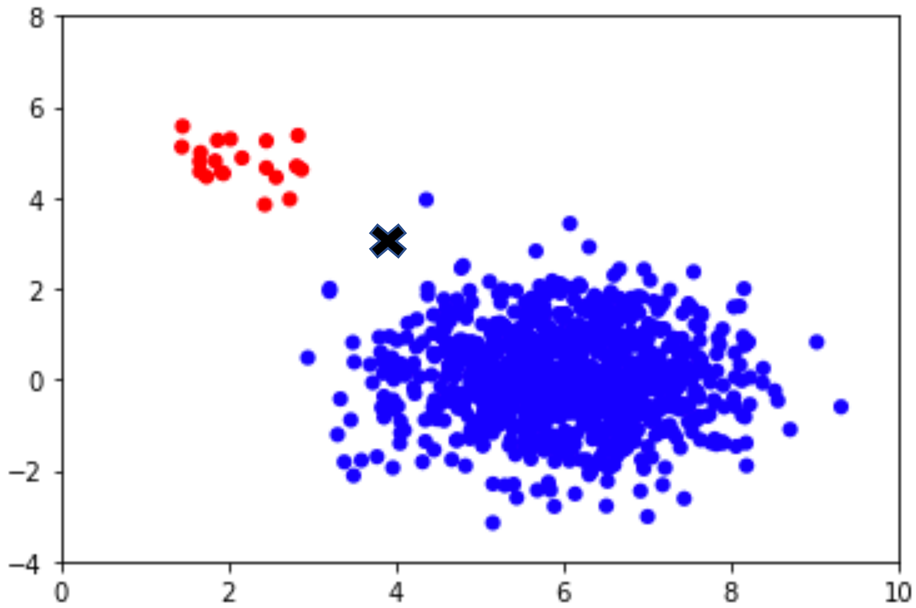
- Generate Observations
  - $\mu_R = (2,5), \mu_B = (6,0), \sigma_R = (0.5,0.5), \sigma_B = (1,1)$
  - $R \sim N(\mu_R, \sigma_R^2), B \sim N(\mu_B, \sigma_B^2)$
  - $\mu = \{\mu_R , \mu_B \}, \sigma = \{\sigma_R , \sigma_B\}$
  - $c_i \sim Categorical(0.02, 0.98)$
  - $x_i | c_i, \mu, \sigma \sim N(c_i^T \mu, c_i^T \sigma)$



1. $z^{(i)} \sim p\ (z)$
2. $x^{(i)} \sim p\ (x|z^{(i)})$

Uncertainty in Classification

- MLE
  - $\hat{u}_R = (2.12, 4.78), \hat{\sigma}_R = (0.46, 0.43)$
  - $\hat{u}_B = (6.03, 0.03), \hat{\sigma}_B = (1.00, 1.00)$
  - $p(R) = 0.02, p(B) = 0.98$
- $x = (4,3)$
- $p(R|x) = \frac{p(x|R)p(R)}{p(x)} = 0.0165$
- $p(B|x) = \frac{p(x|B)p(B)}{p(x)} = 0.9835$



- $p(z|x)$

Variational Inference

- Data Generation
  1. $z^{(i)} \sim p\ (z)$, continuous random variable
  2. $x^{(i)} \sim p\ (x|z^{(i)})$
- Posterior
  - $p(z|x)$
- Bayes theorem
  - $p(z|x) = \frac{p(z,x)}{p(x)}$
  - $p(x) = \int p(z,x)\mathrm{d}z$

1. Posterior is intractable
   - EM algorithm cannot be used
2. Large dataset
   - MCMC sampling is slow

Variational Inference

- Approximation
  - Candidate distribution space: $\mathfrak{D}$
  - $q^* (z) = \arg \min_{q(z) \in \mathfrak{D}} KL(q(z)||p(z|x))$
- The Evidence Lower Bound
  - $KL(q(z)||p(z|x)) = E[\log q(z)] - E[\log p(z|x)]$
    $= E[\log q(z)] - E[\log p(z,x)] + \log p(x)$
  - $\log p(x) = KL(q(z)||p(z|x)) + (E[\log p(z,x)] - E[\log q(z)])$

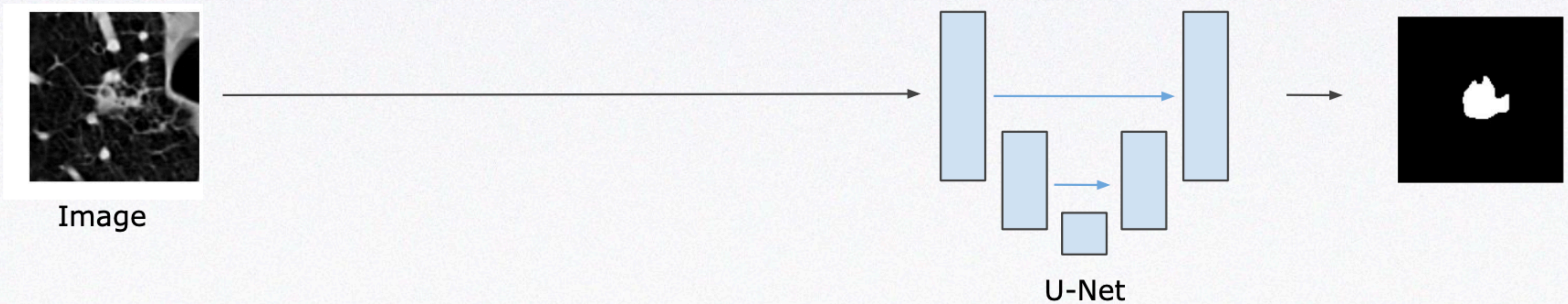
    ELBO
  - Minimize KL = Maximize ELBO
  - $ELBO(q) = E[\log p(z)] + E[\log p(x|z)] - E[\log q(z)]$
    $= E[\log p(x|z)] - KL(q(z)||p(z))$

Overview

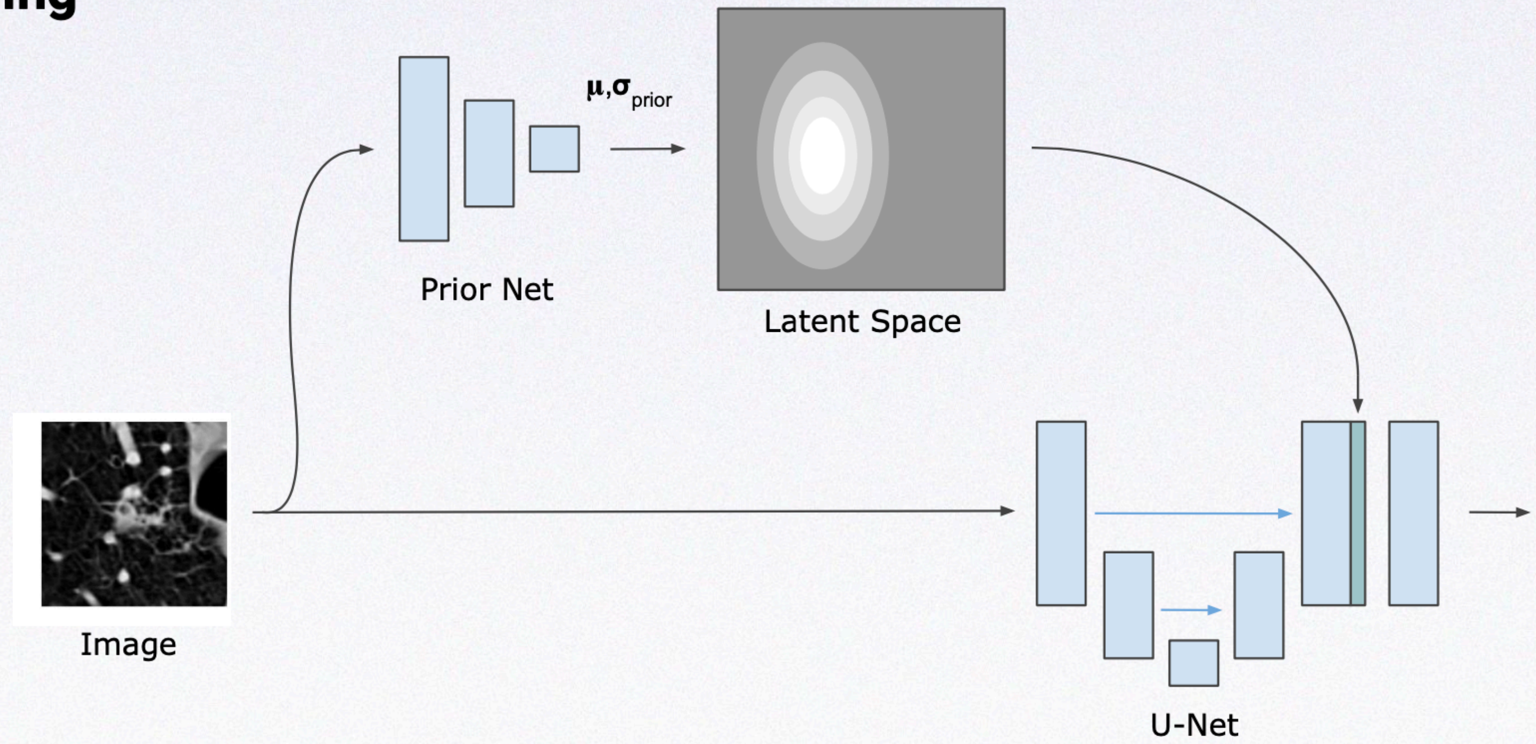
1. Introduction
2. Preliminary
    1. Semantic Segmentation
    2. Uncertainty Estimation
    3. Variational Inference
3. Probabilistic U-Net
4. Experiments

Deterministic U-Net

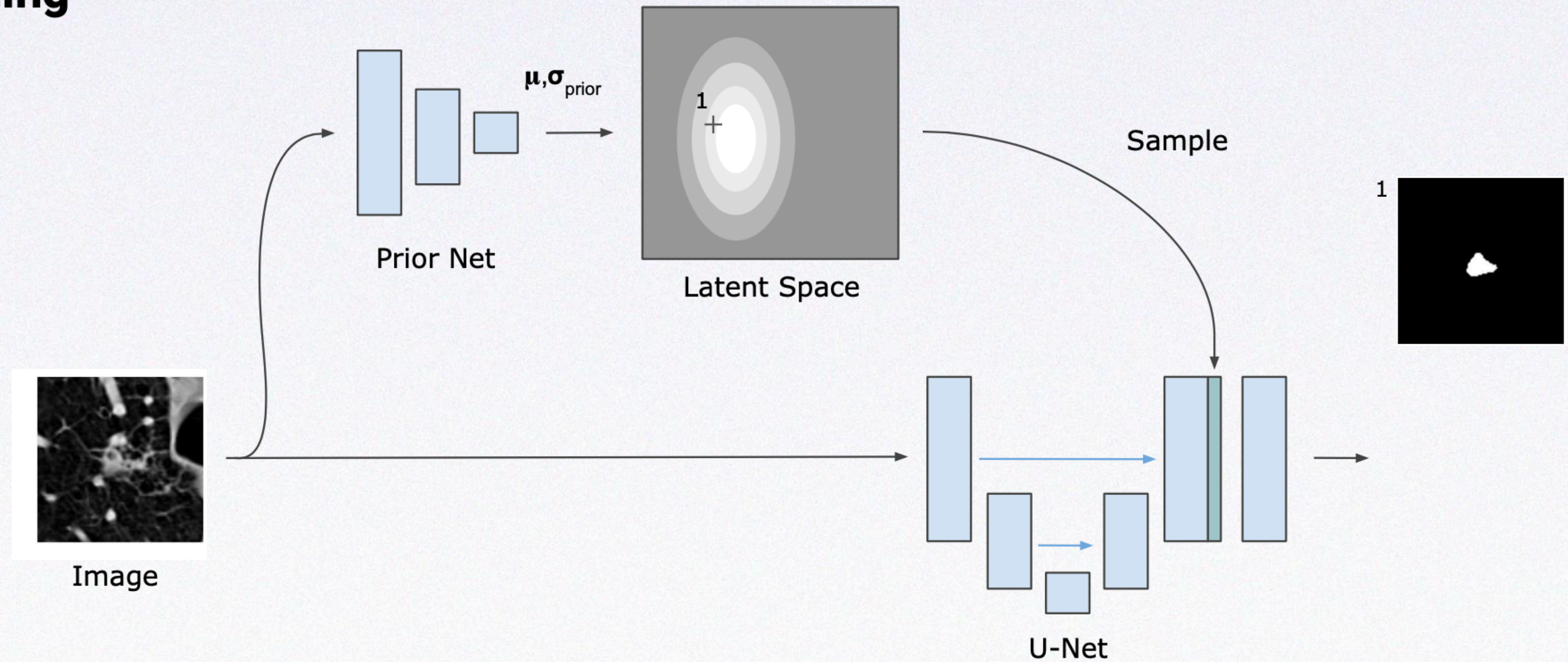
Inference



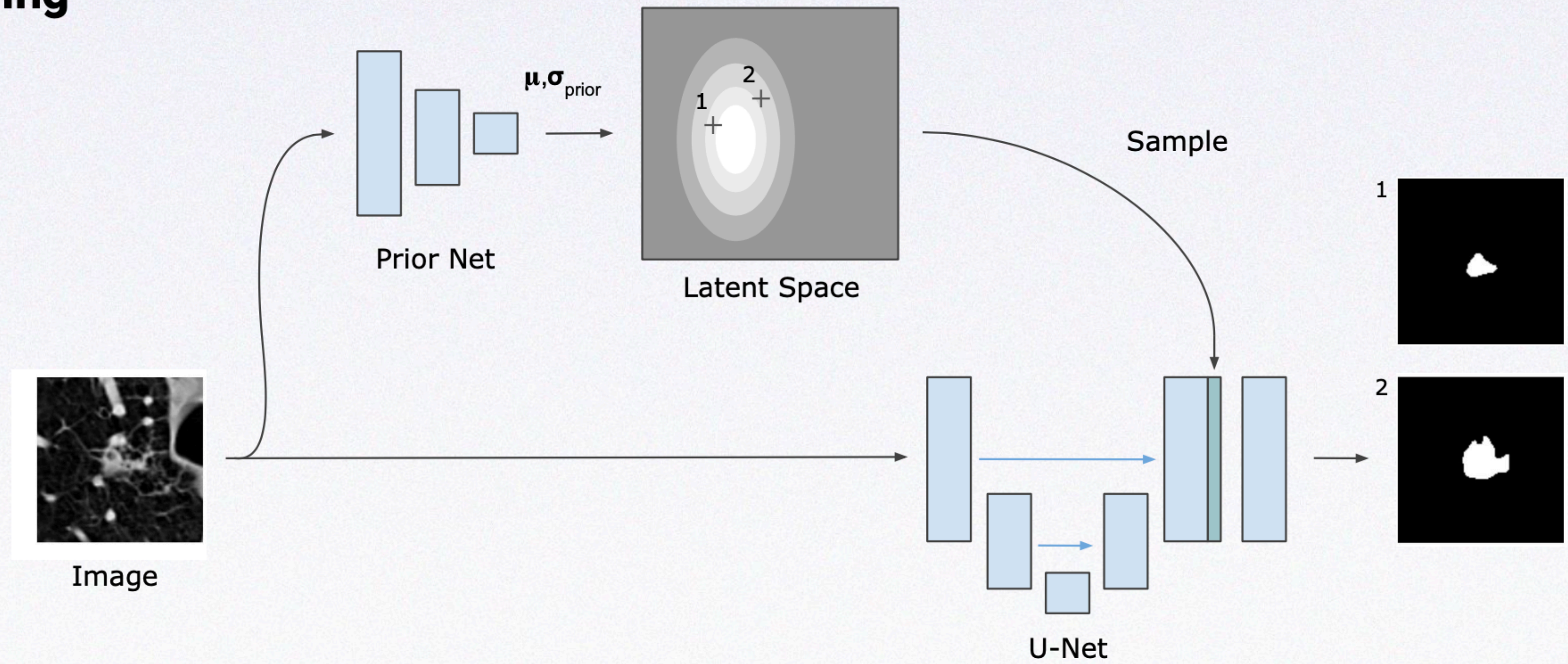
Probabilistic U-Net
Sampling



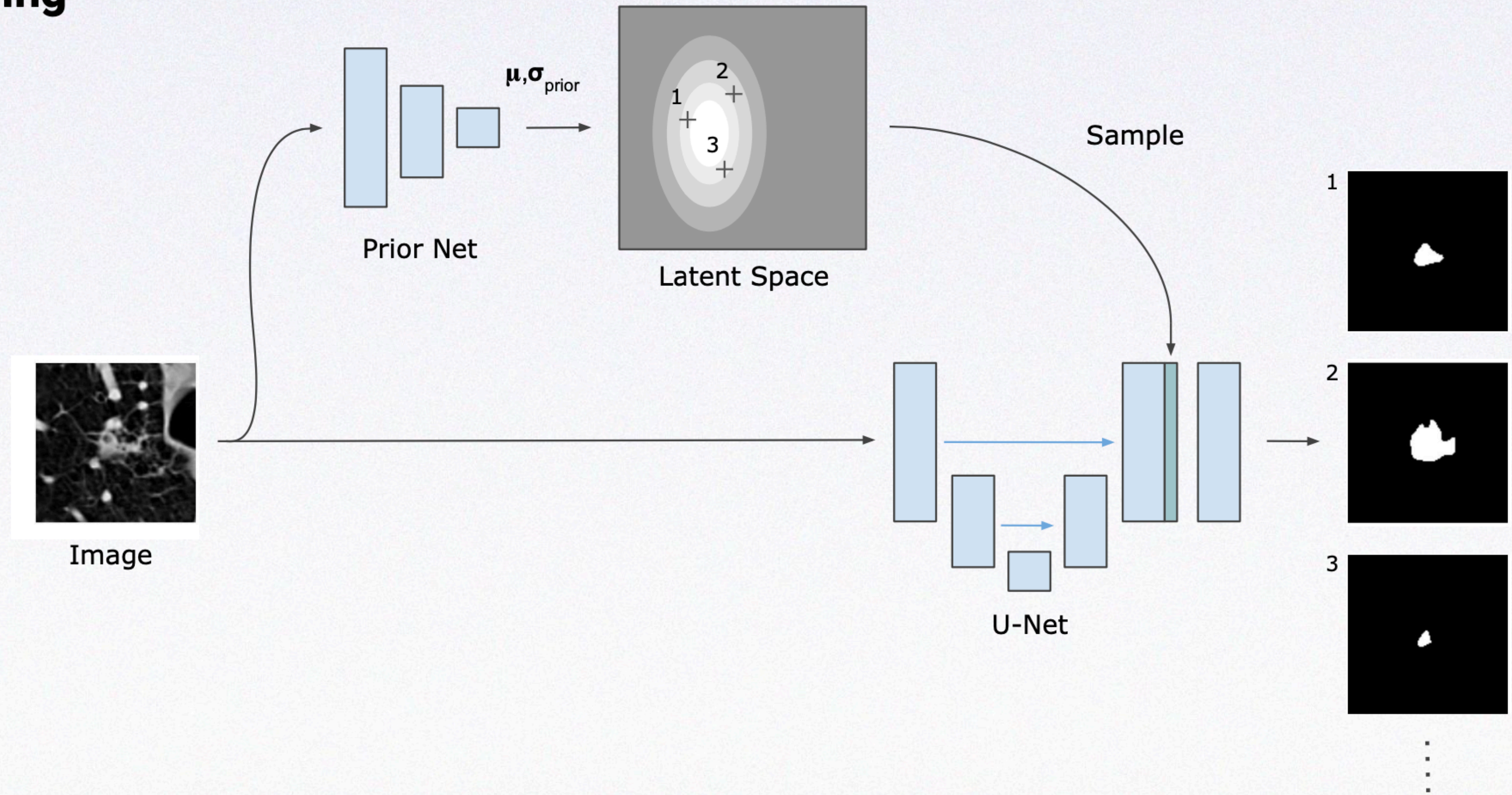
Probabilistic U-Net
Sampling



Probabilistic U-Net
Sampling



Probabilistic U-Net
Sampling

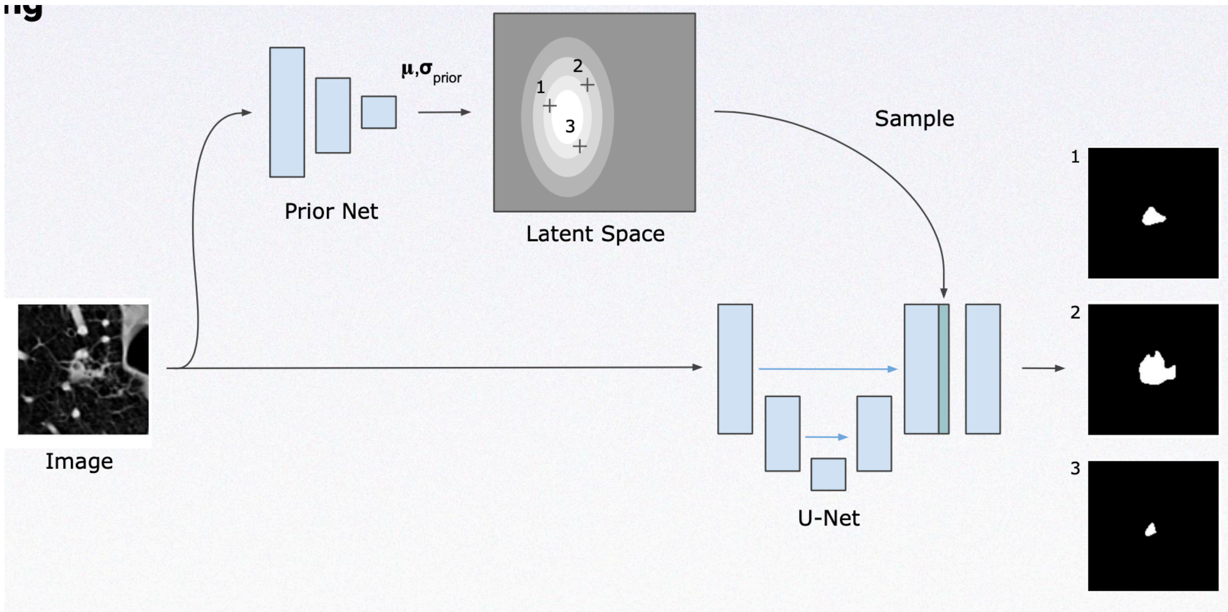


Probabilistic U-Net Sampling

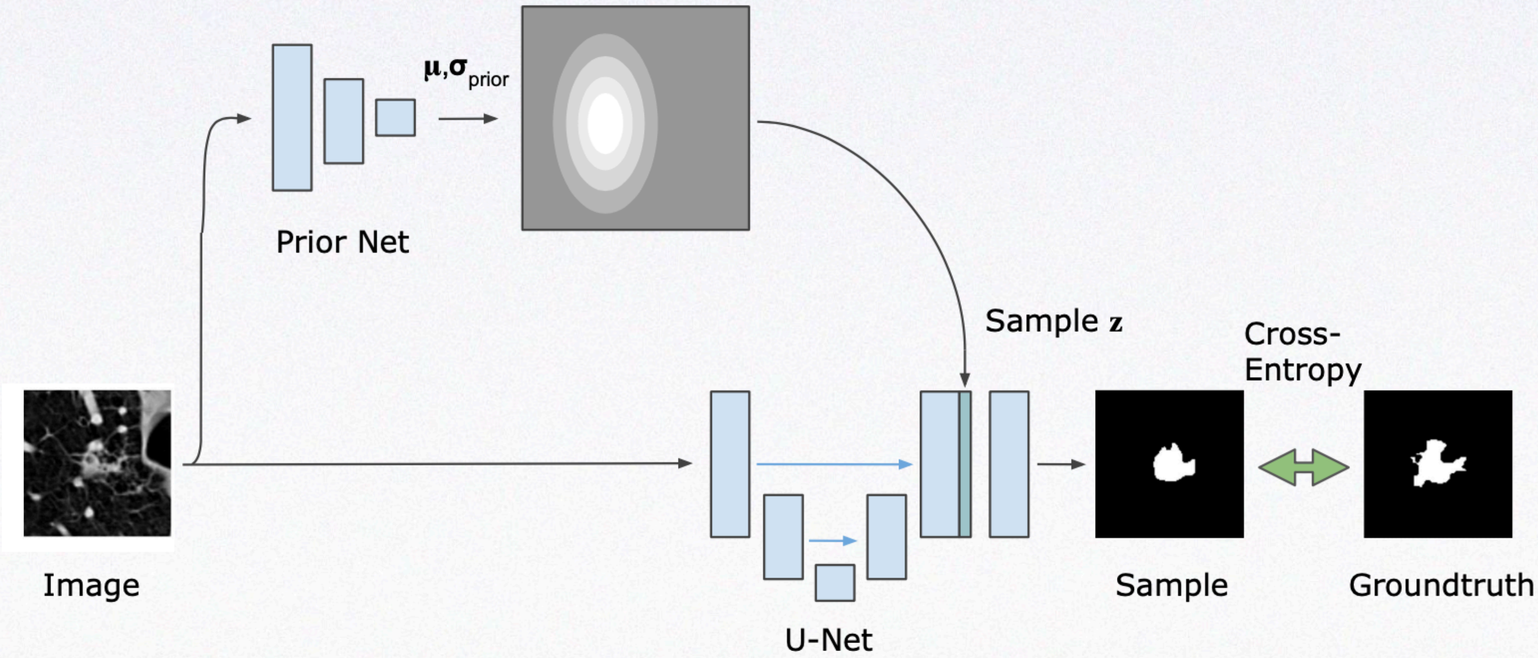
- $z^{(i)} \sim p(z|x) = N\big(\mu(x; w), \sigma(x; w)\big)$
- $y^{(i)} = f_{comb}\big(f_{unet}(x; \theta), z^{(i)}; \phi\big)$

Data Generation

1. $z^{(i)} \sim p\,(z)$, continuous random variable
2. $x^{(i)} \sim p\,(x|z^{(i)})$



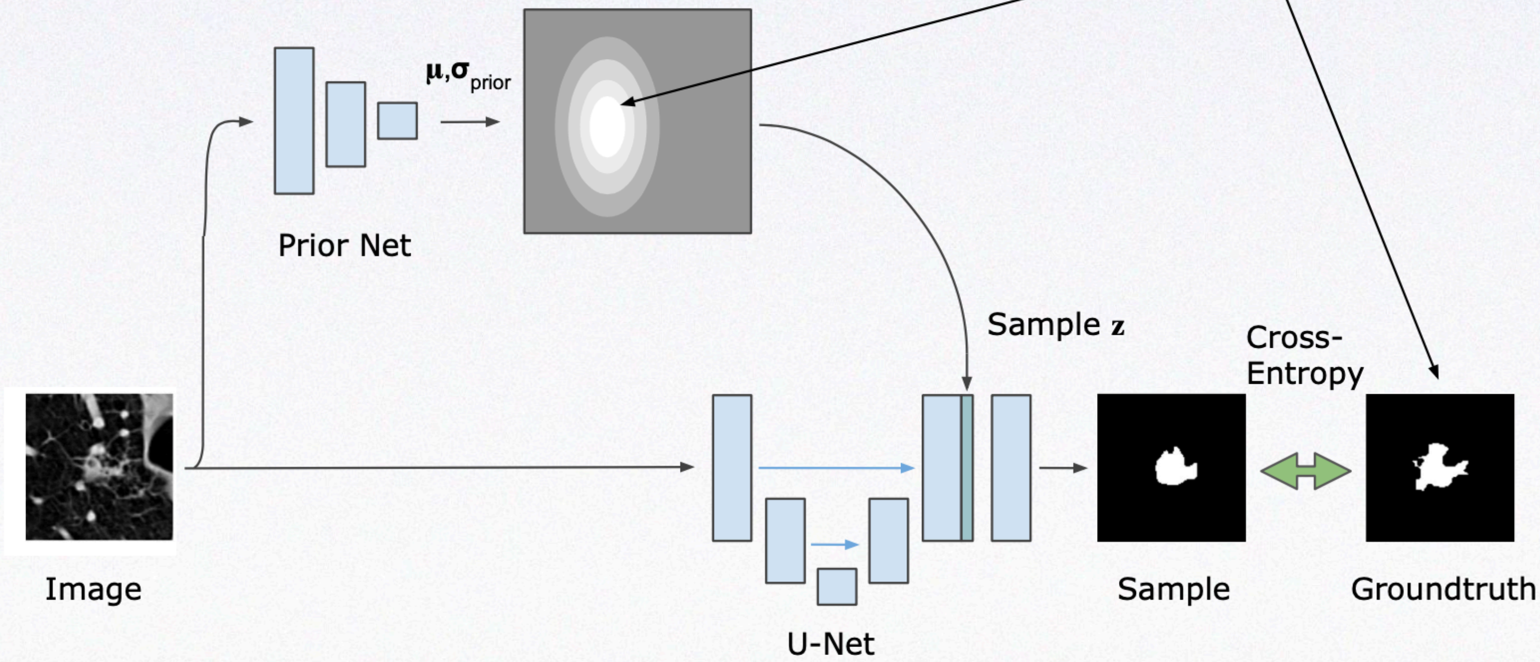
Probabilistic U-Net
Training



DeepMind

Probabilistic U-Net
Training

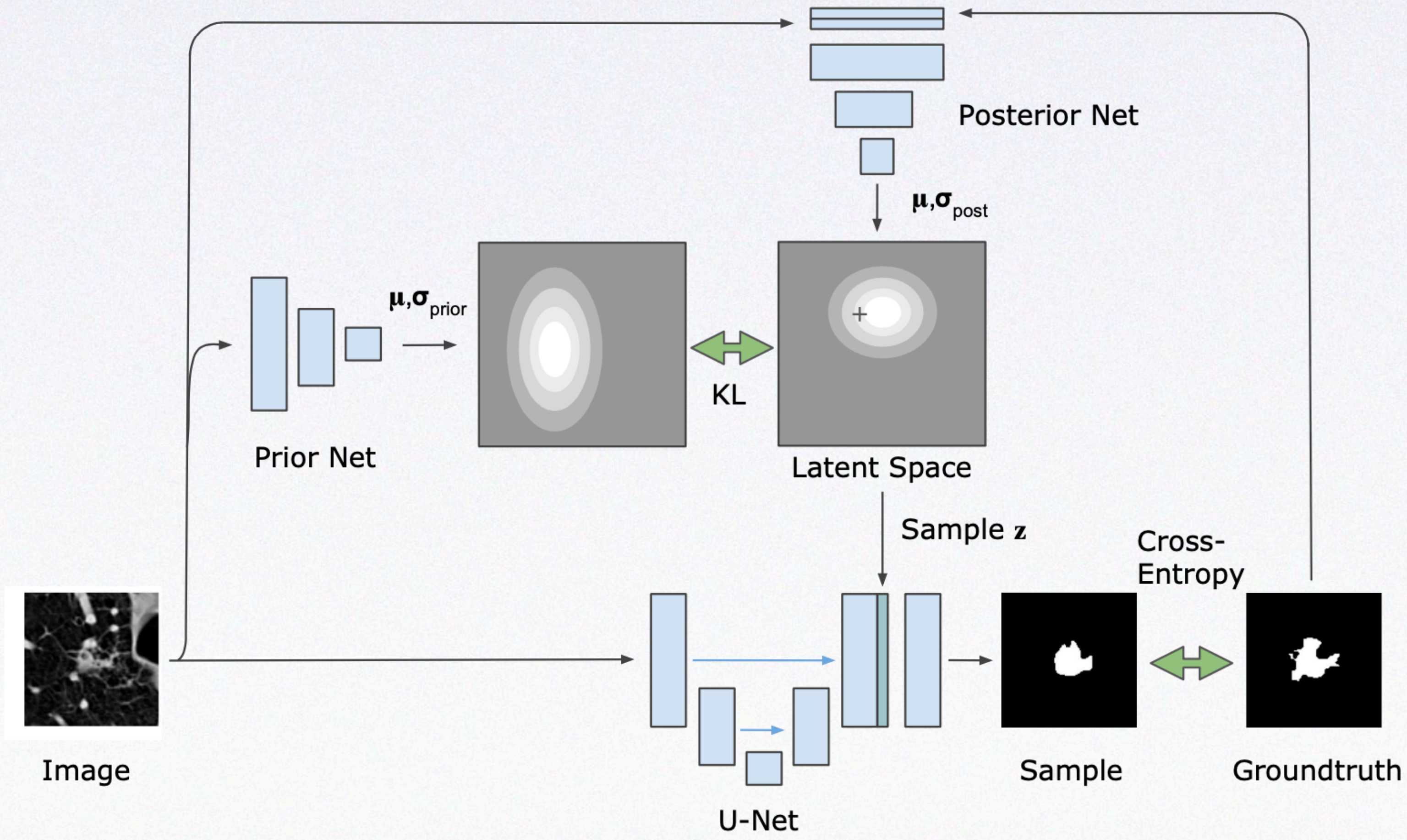
Position in Latent Space for this GT example?



DeepMind

Probabilistic U-Net

Training

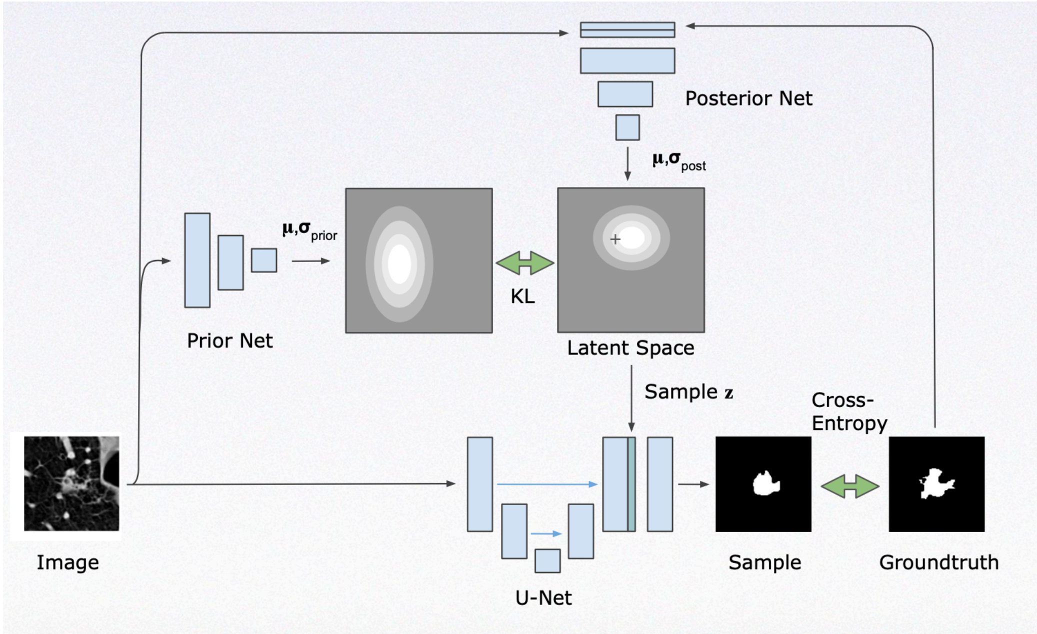


DeepMind

Probabilistic U-Net

- $ELBO(q) = E[\log p(z)] + E[\log p(x|z)] - E[\log q(z)]$
  $= E[\log p(x|z)] - KL(q(z)||p(z))$

- $L(x, y) = -E[\log p(y|z, x)] + KL(q(z|x, y)\|p(z|x))$



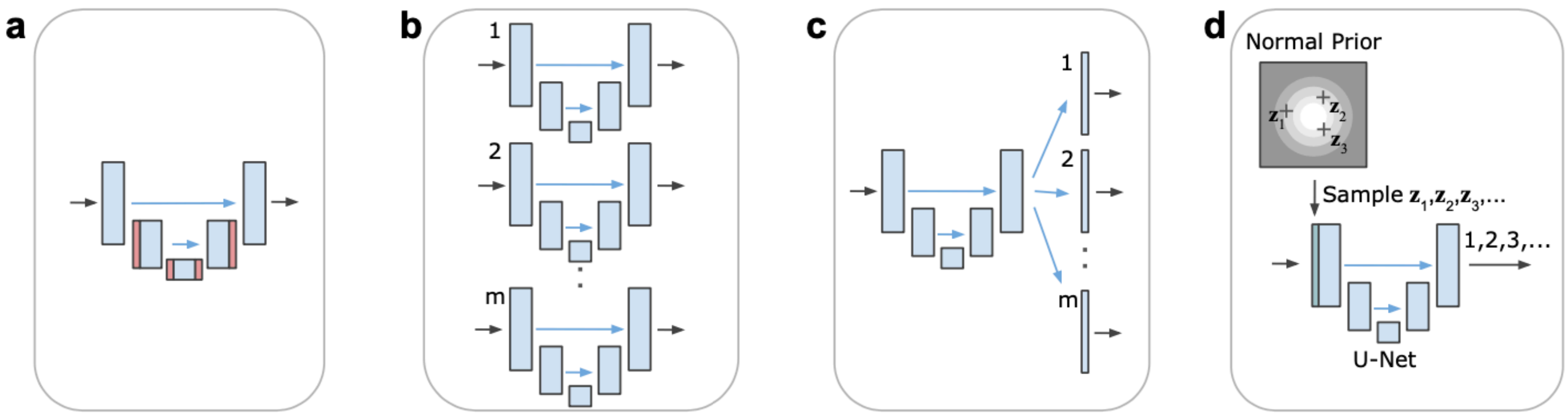
Overview

1. Introduction
2. Preliminary
   1. Semantic Segmentation
   2. Uncertainty Estimation
   3. Variational Inference
3. Probabilistic U-Net
4. Experiments

Datasets

- Lung abnormalities segmentation
- Cityscapes semantic segmentation
  - ‘sidewalk’ to ‘sidewalk 2’ with a probability of 8/17
  - ‘person’ to ‘person 2’ with a probability of 7/17
  - ‘car’ to ‘car 2’ with 6/17
  - ‘vegetation’ to ‘vegetation 2’ with 5/17
  - ‘road’ to ‘road 2’ with probability 4/17

Baseline



Dropout-Unet

Ensemble

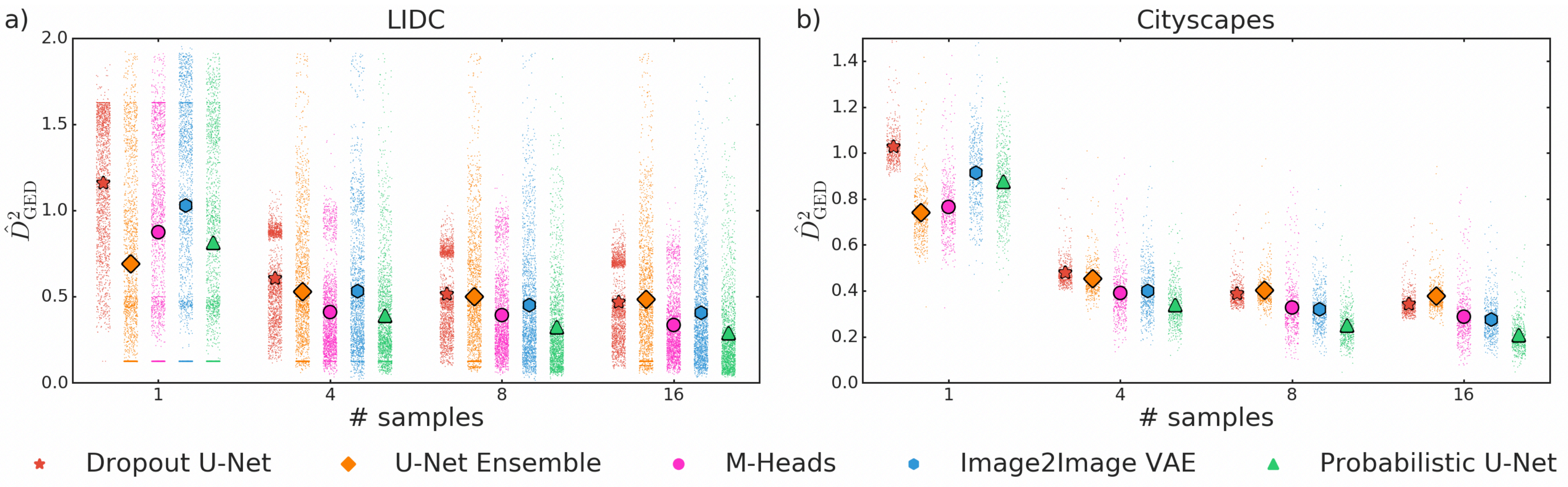
M-Head

Image2Image VAE

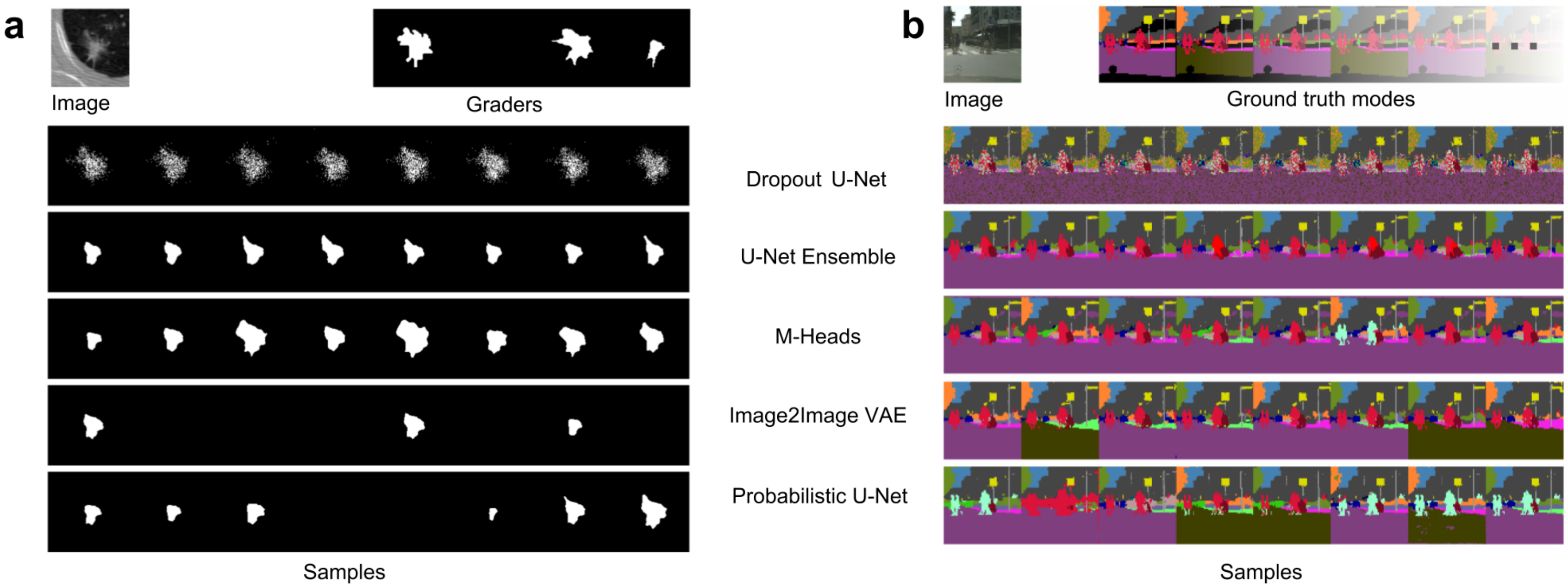
Experiments

- generalized energy distance

$$D^2_{\mathrm{GED}}(P_{\mathrm{gt}}, P_{\mathrm{out}}) = 2\mathbb{E}\Big[d(S,Y)\Big] - \mathbb{E}\Big[d(S,S')\Big] - \mathbb{E}\Big[d(Y,Y')\Big]$$

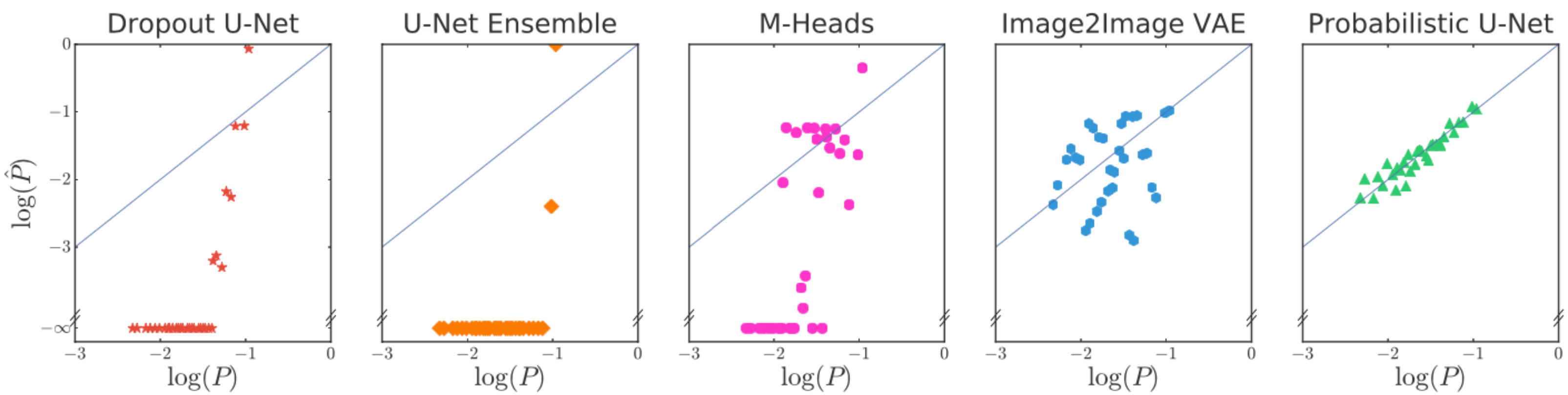


Experiments



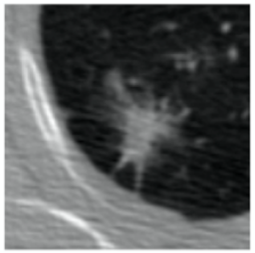
Experiments

- Cityscapes
  - Flip 5 classes -> 32 modes



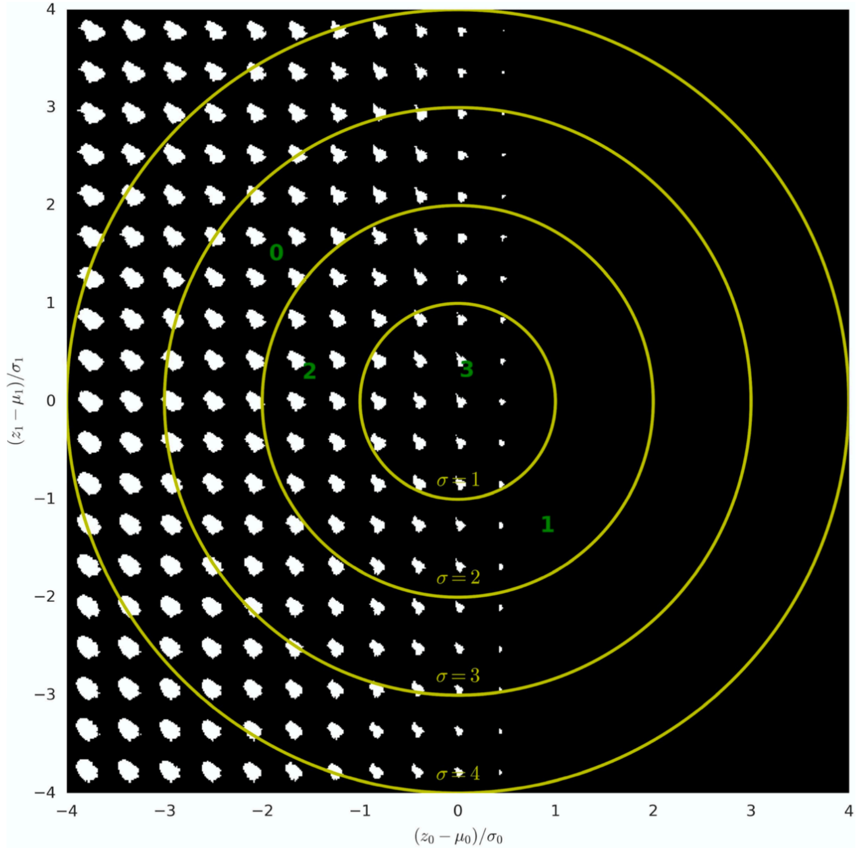
Car, road
Car, road2
Car2, road
Car2, road2

Experiments



Image

Graders



Experiments

Input image