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], Olaf Ronneberger[1]

[1]DeepMind
[2]German Cancer Research Center

*work done during an internship at DeepMind

DeepMind

dkfz.
GERMAN CANCER RESEARCH CENTER
IN THE HELMHOLTZ ASSOCIATION

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], Olaf Ronneberger[1]

[1]DeepMind
[2]German Cancer Research Center

*work done during an internship at DeepMind

Poster #127

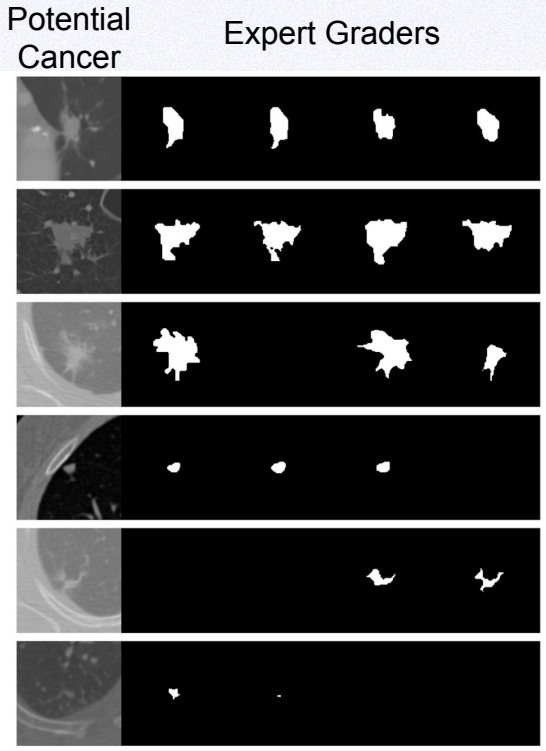
Medical Imaging Workshop Talk: Sat, Dec 8, 9:45 am

DeepMind

dkfz.
GERMAN CANCER RESEARCH CENTER
IN THE HELMHOLTZ ASSOCIATION

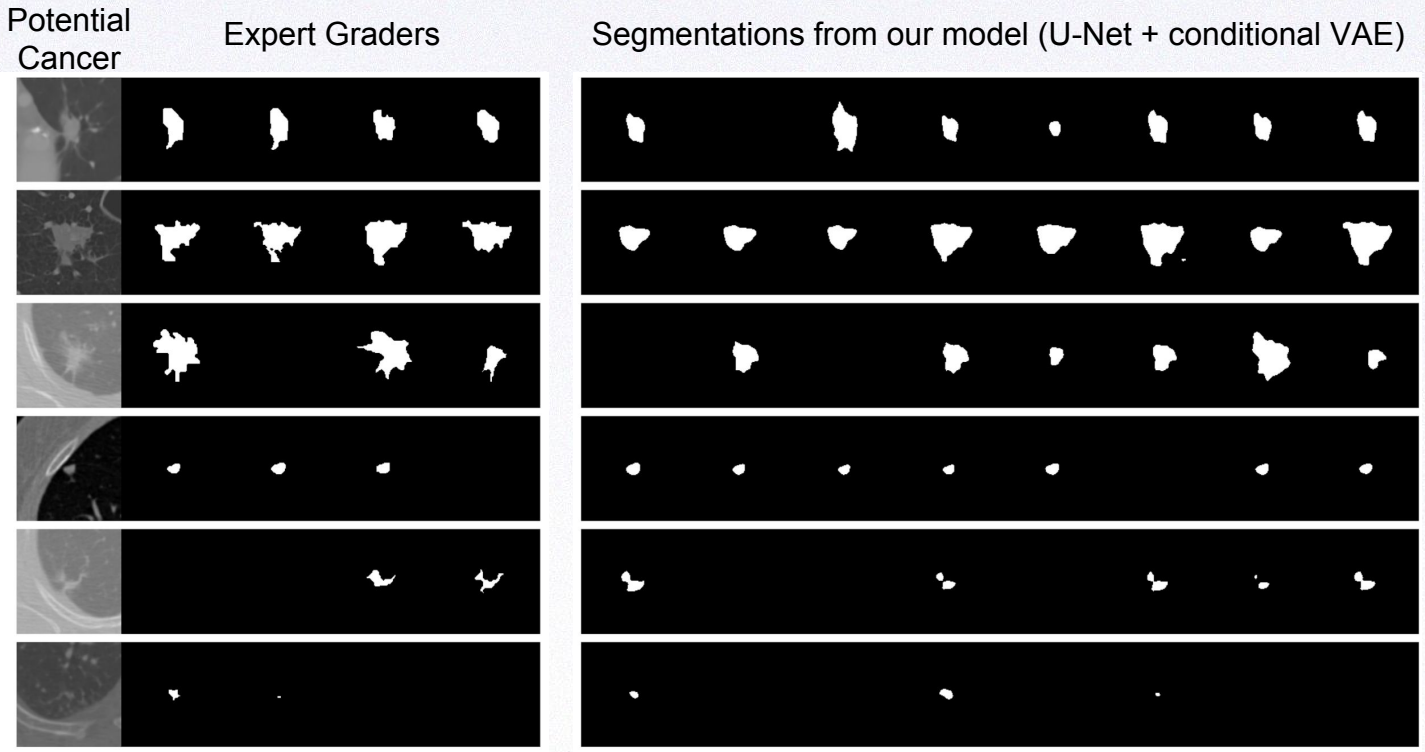
Images are often Ambiguous

Images are often Ambiguous



DeepMind

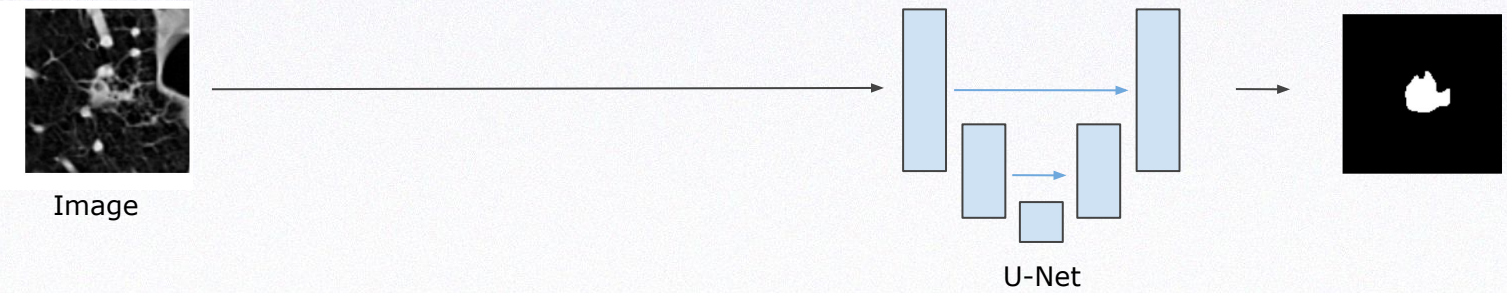
Images are often Ambiguous



DeepMind

Deterministic U-Net

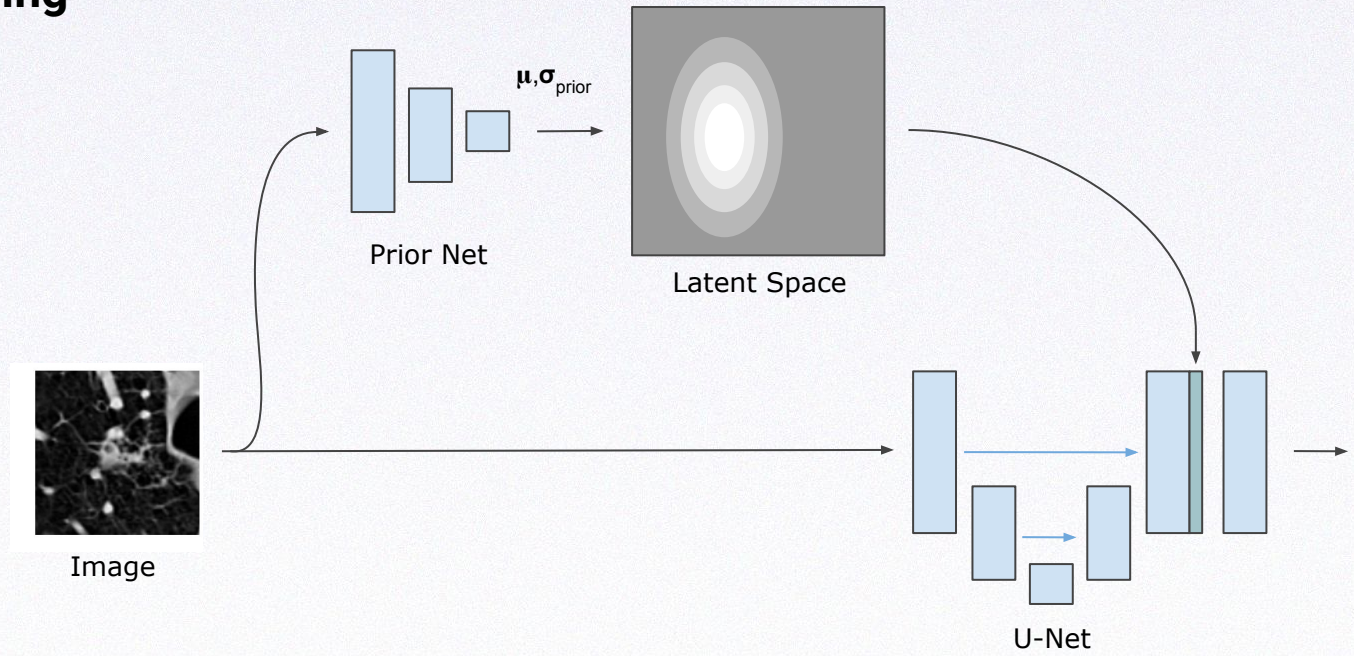
Inference



DeepMind

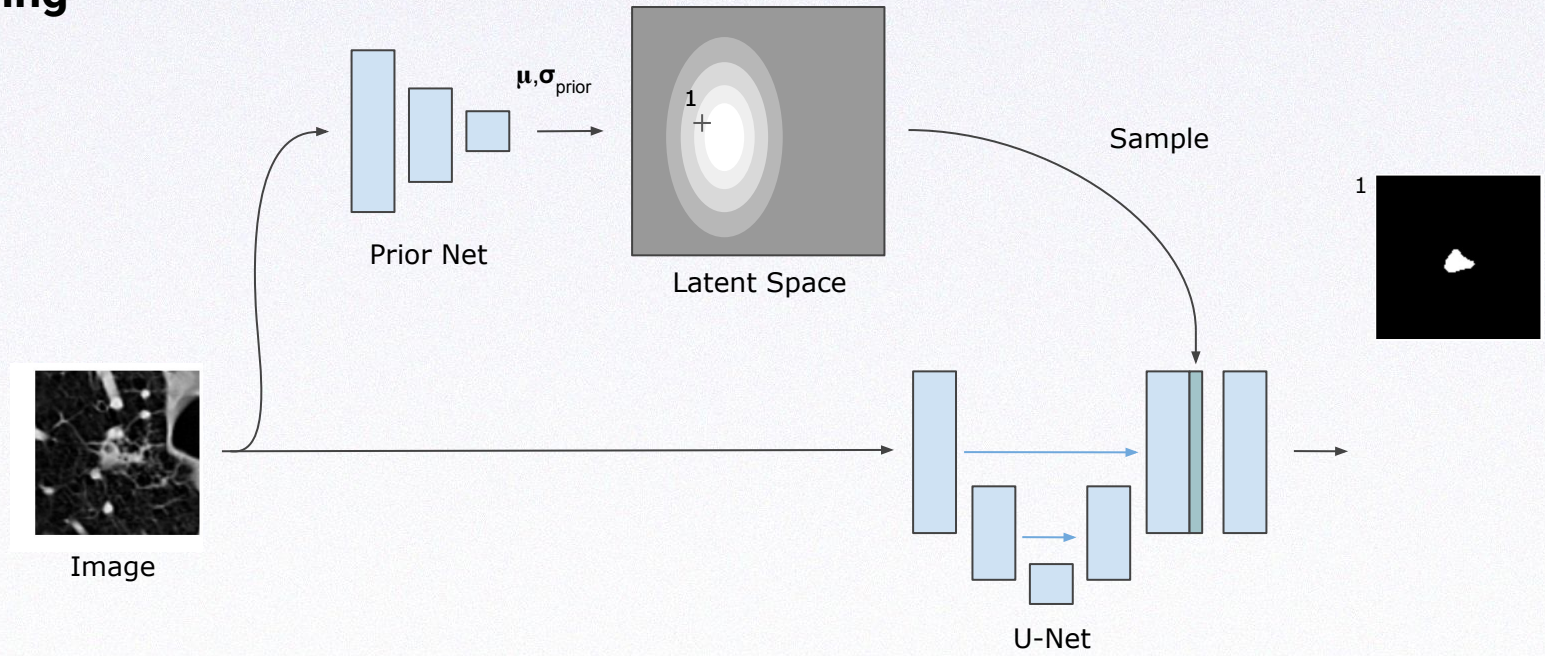
Probabilistic U-Net

Sampling



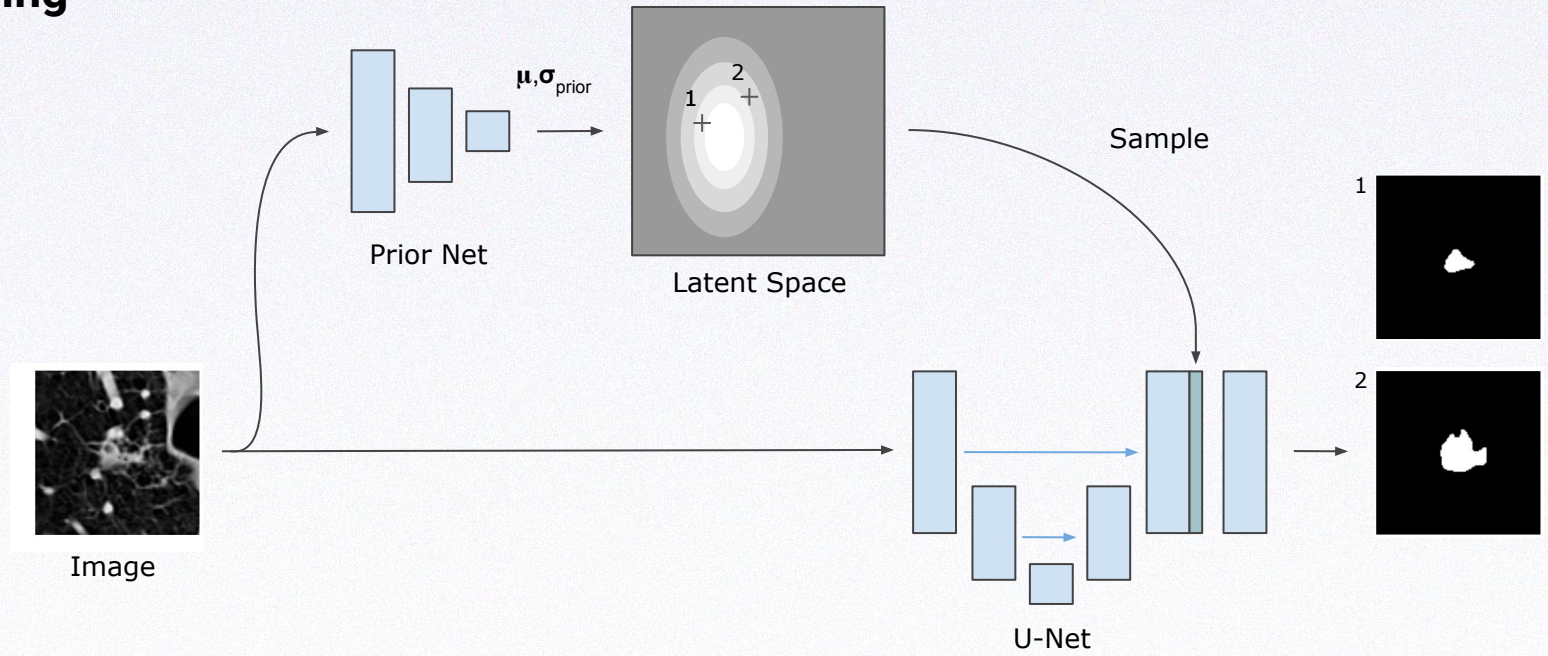
Probabilistic U-Net

Sampling



Probabilistic U-Net

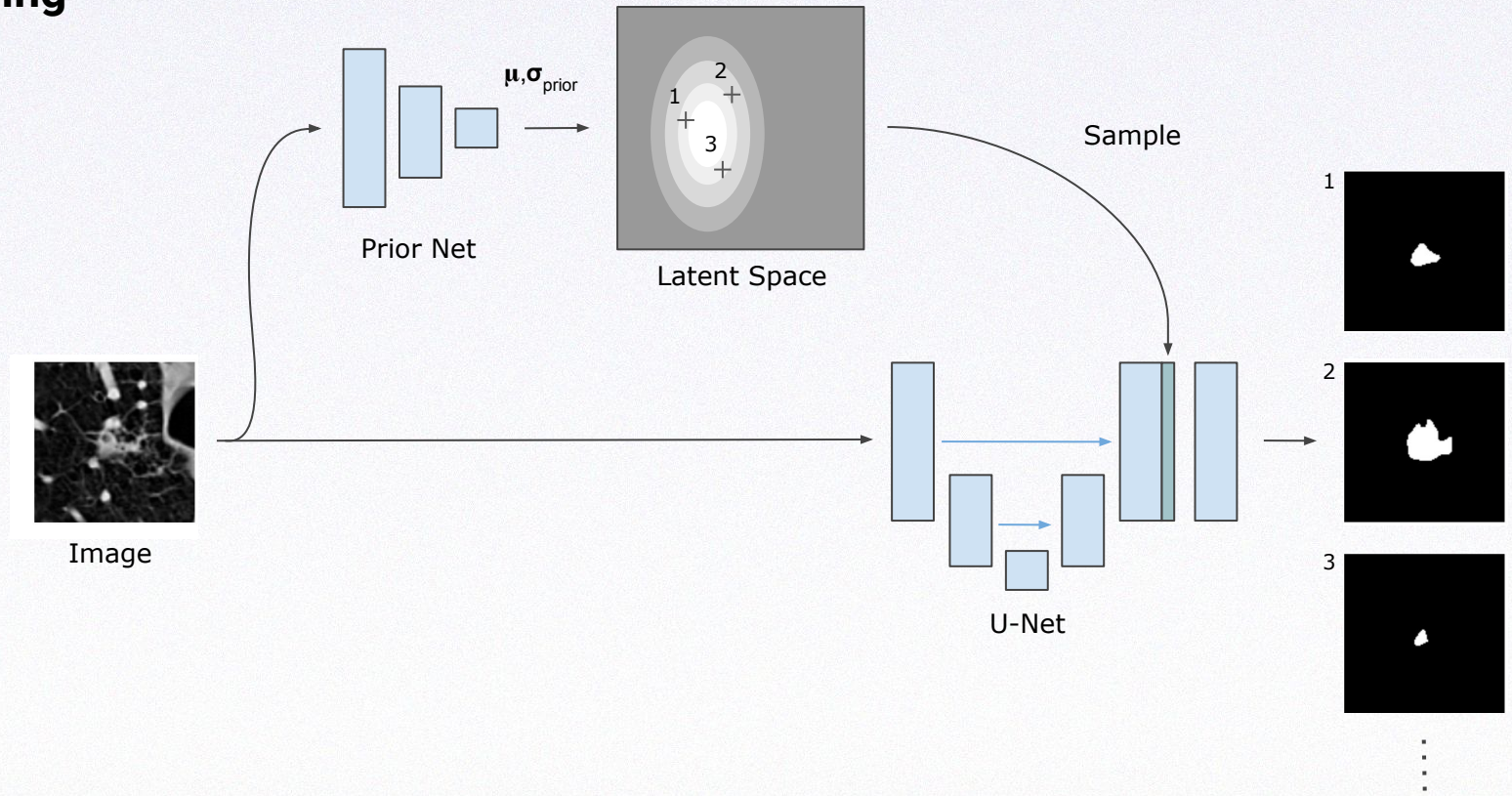
Sampling



DeepMind

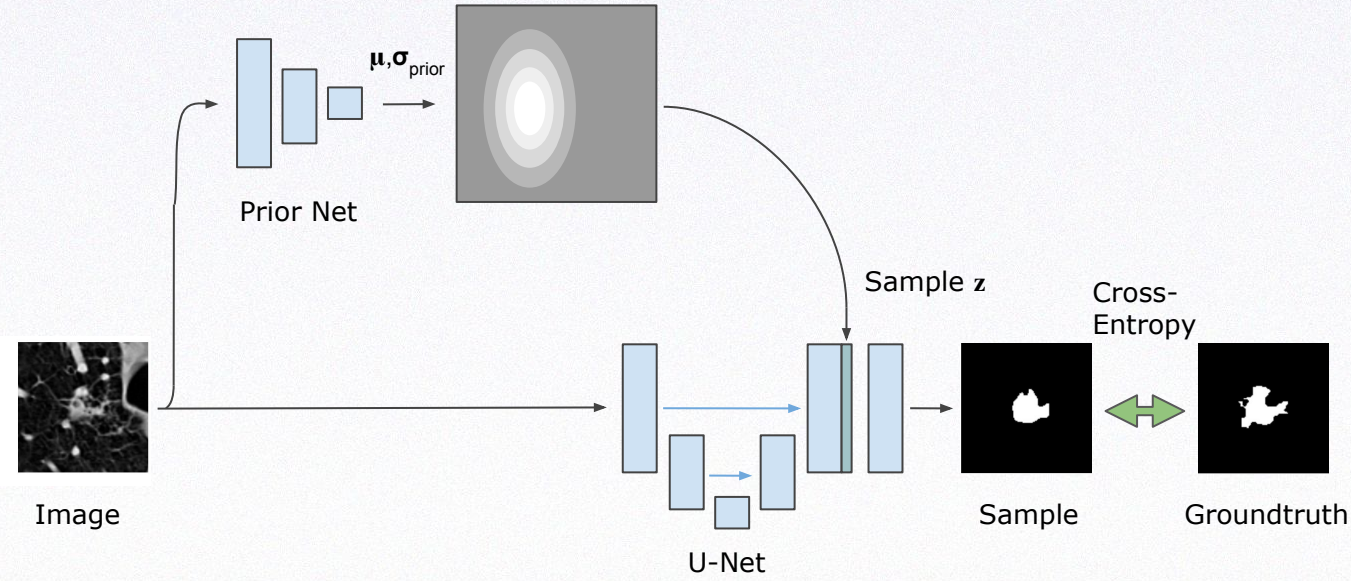
Probabilistic U-Net

Sampling



Probabilistic U-Net

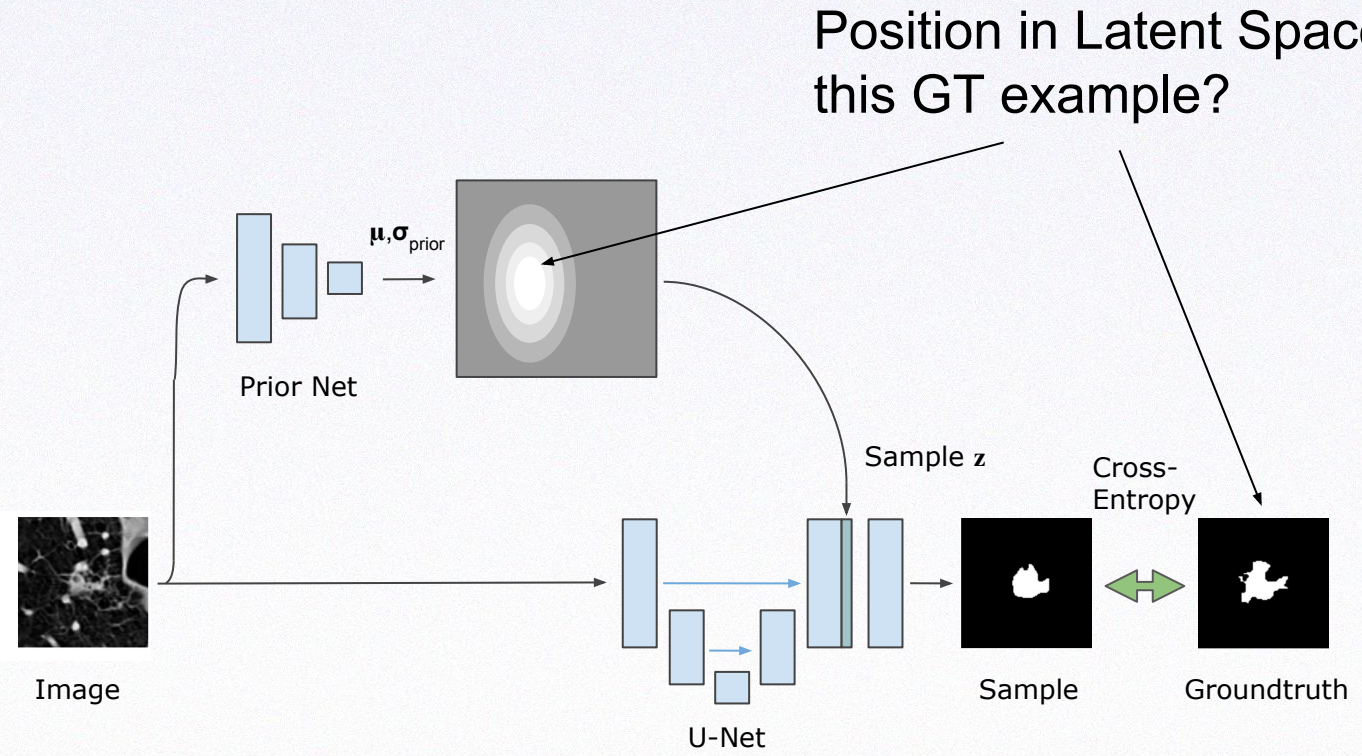
Training



DeepMind

Probabilistic U-Net

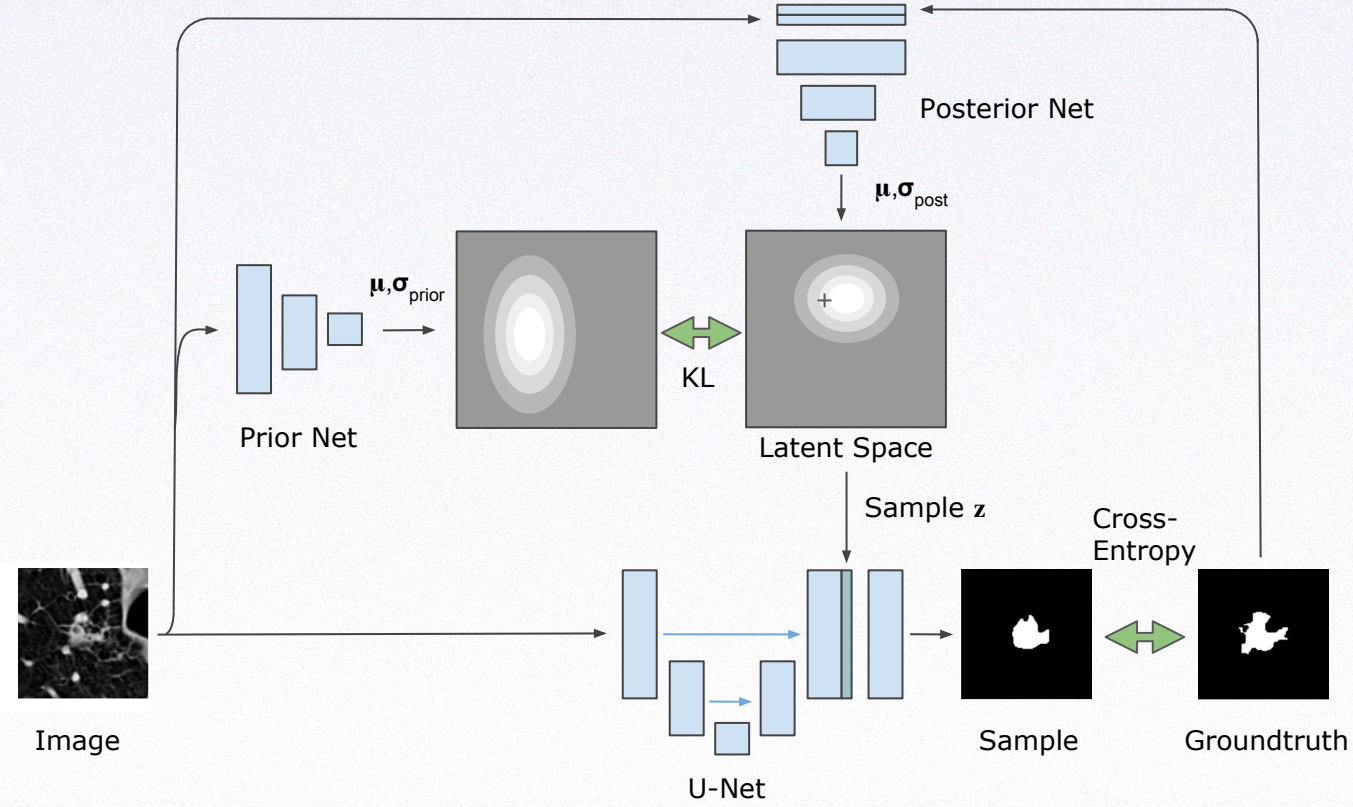
Training



DeepMind

Probabilistic U-Net

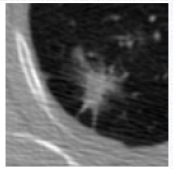
Training



Latent Space Analysis

Probabilistic U-Net

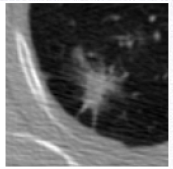
Image

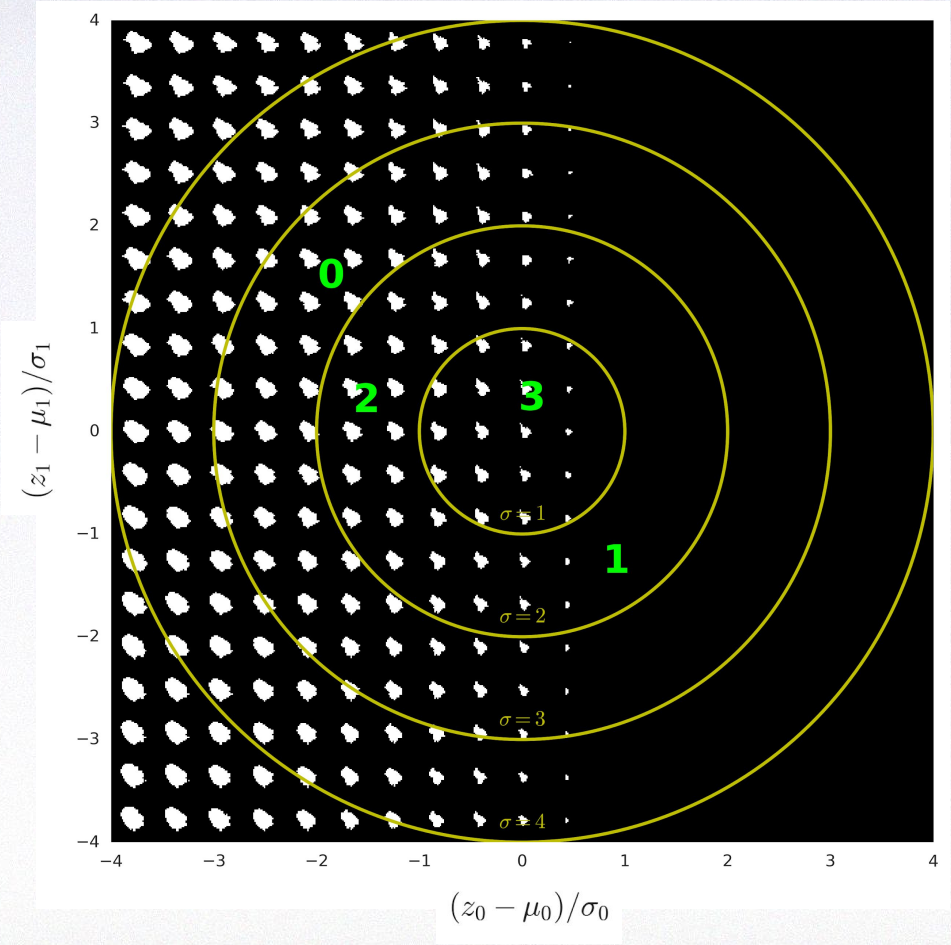


Latent Space Analysis

Probabilistic U-Net

Image



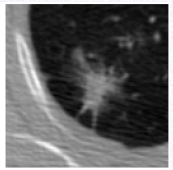


DeepMind

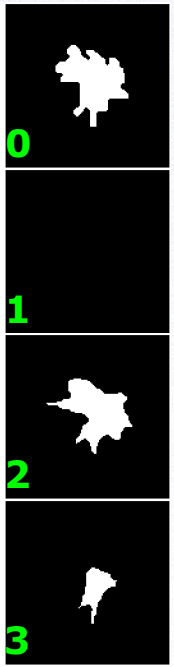
Latent Space Analysis

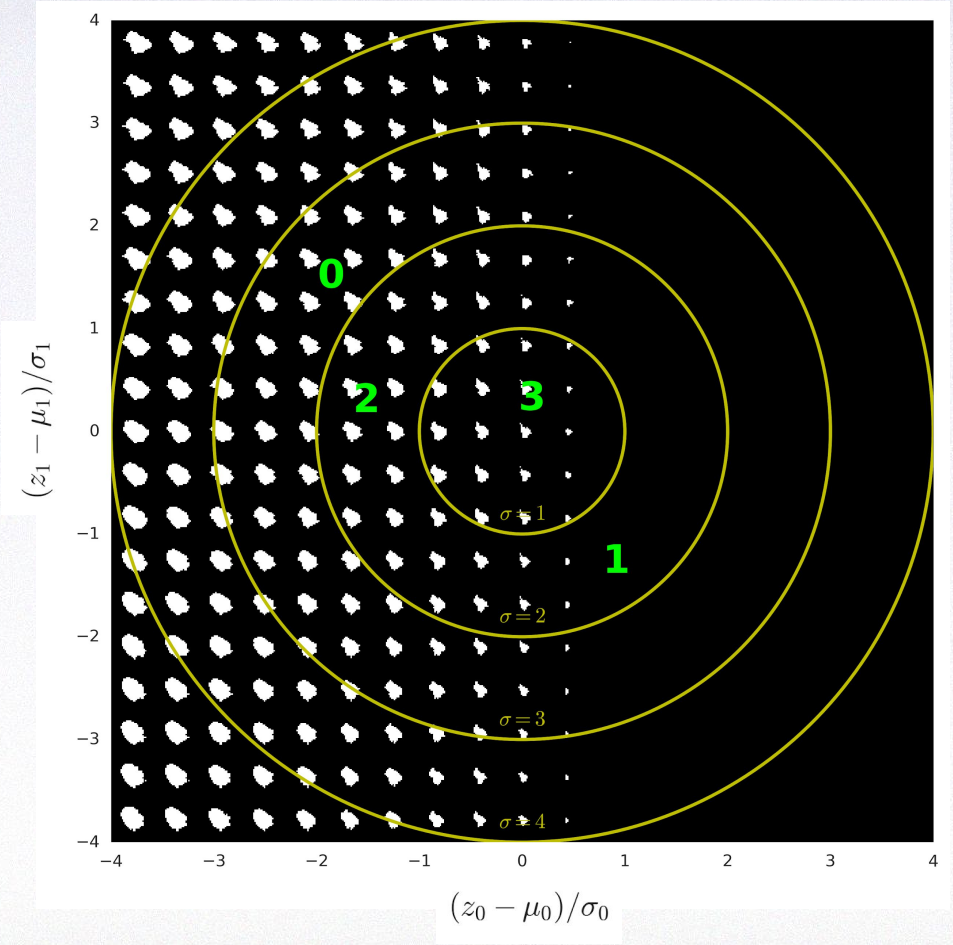
Probabilistic U-Net

Image



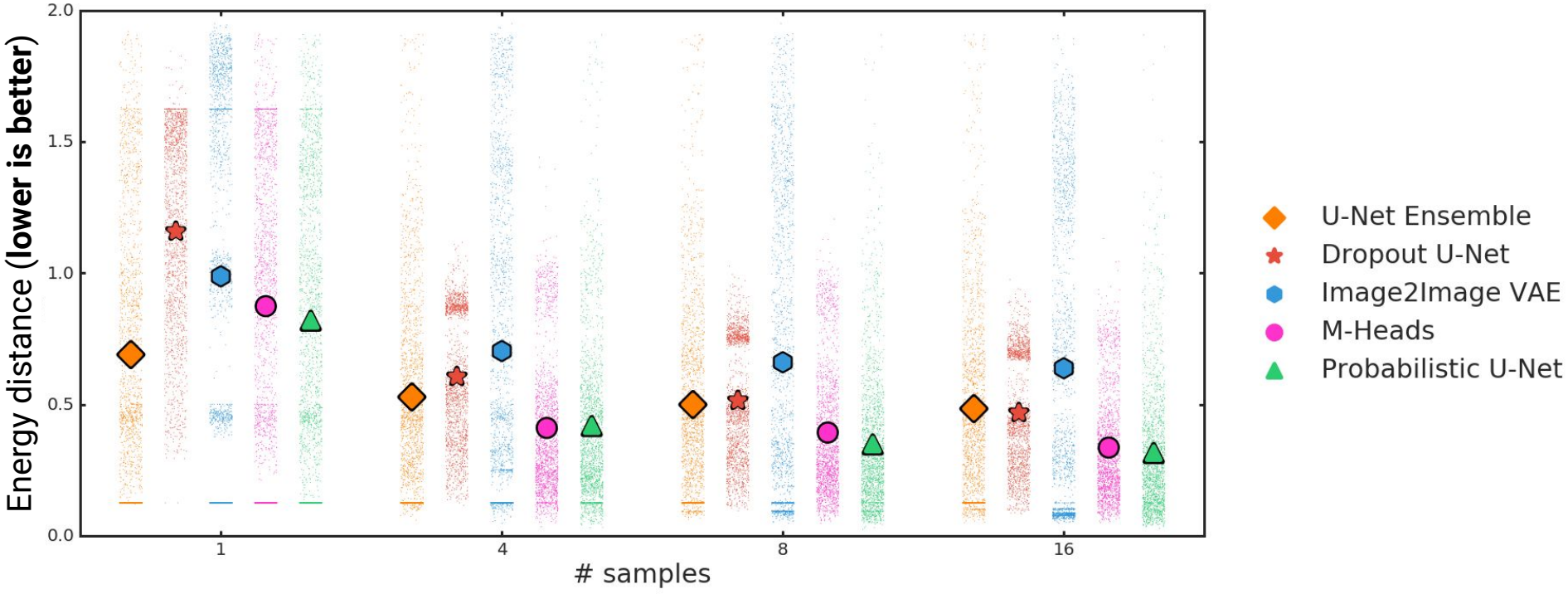
Graders





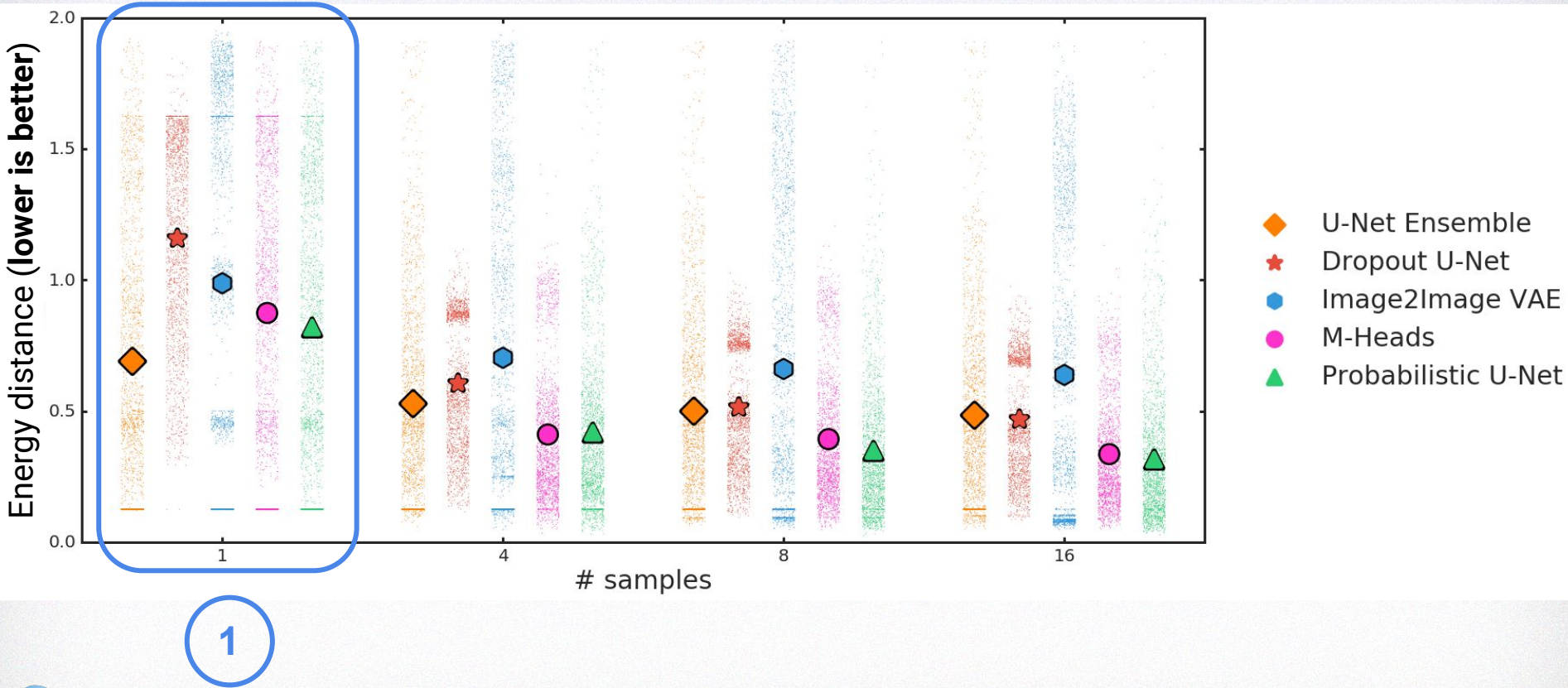
DeepMind

Lung Abnormalities Segmentation: Quantitative Results



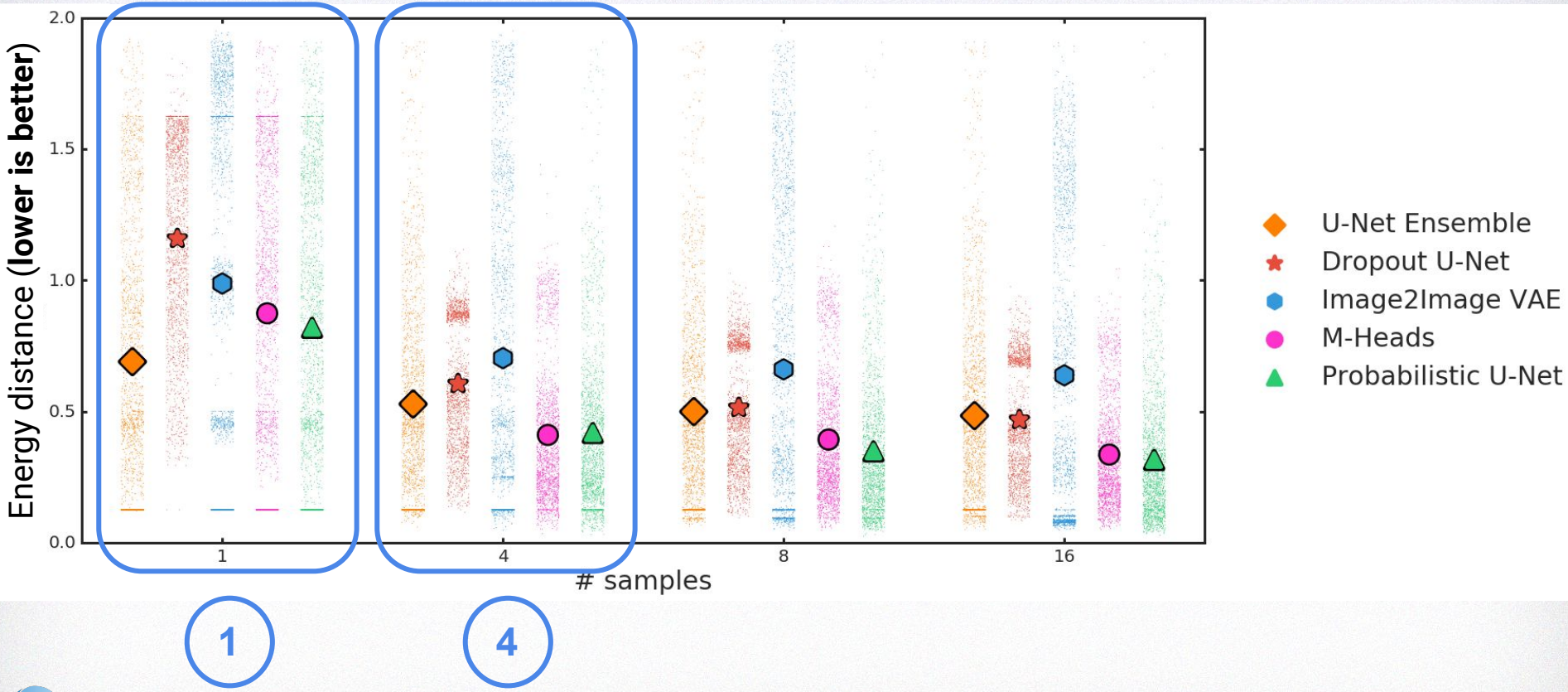
DeepMind

Lung Abnormalities Segmentation: Quantitative Results



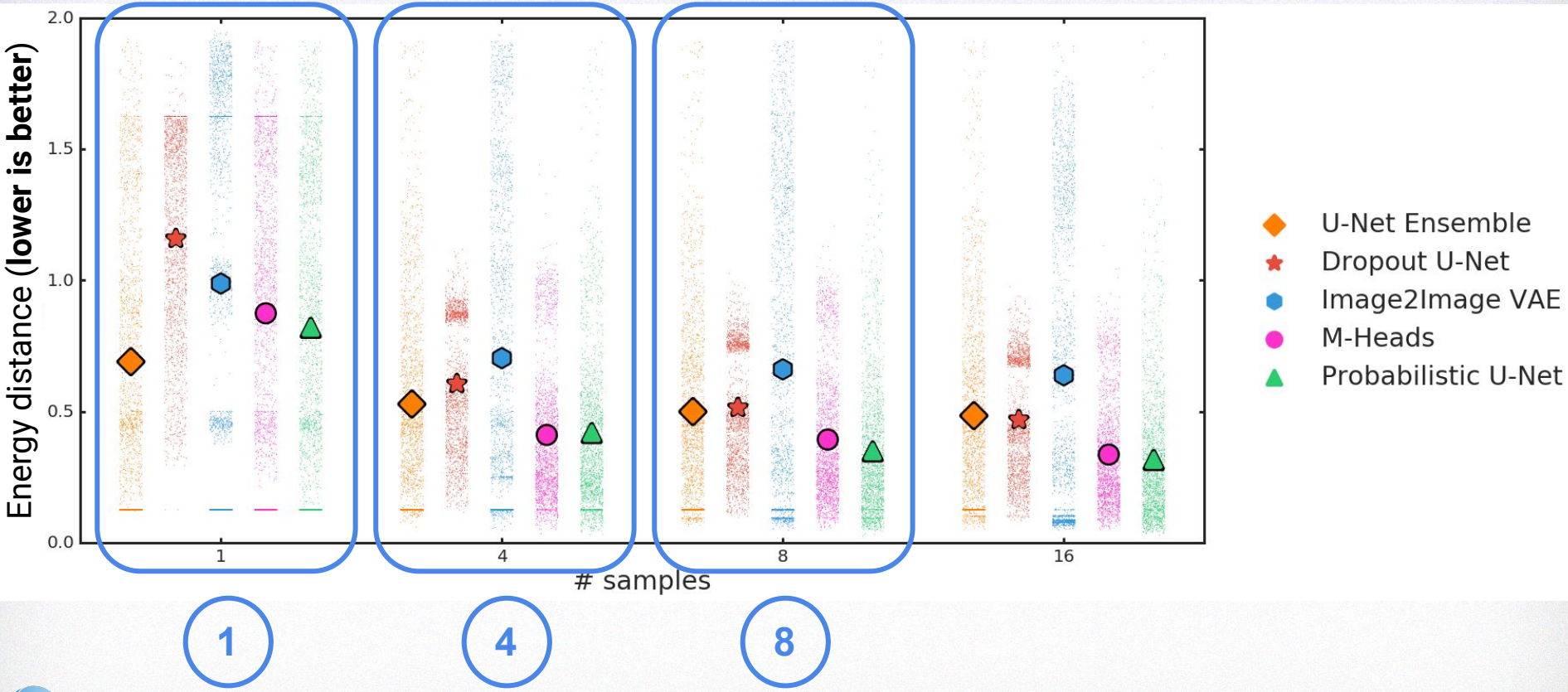
DeepMind

Lung Abnormalities Segmentation: Quantitative Results



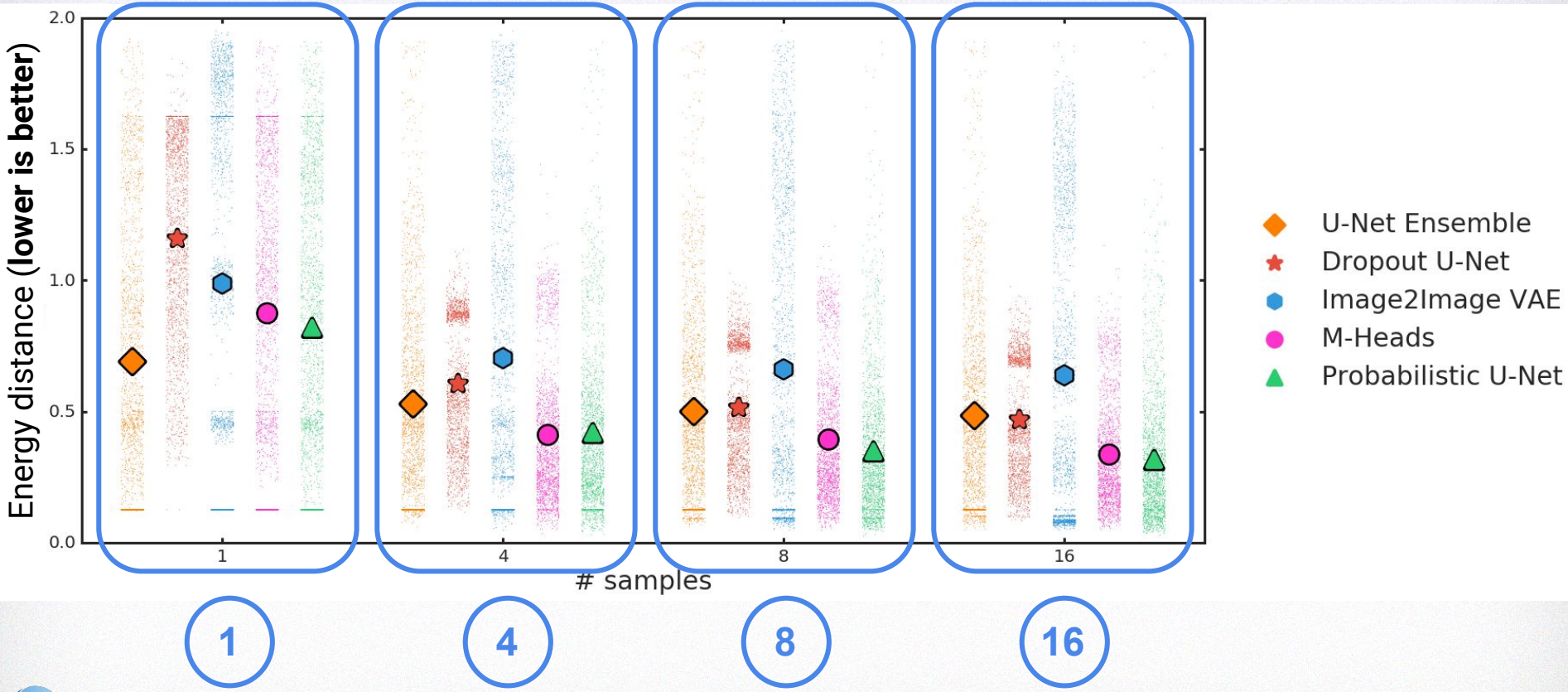
DeepMind

Lung Abnormalities Segmentation: Quantitative Results



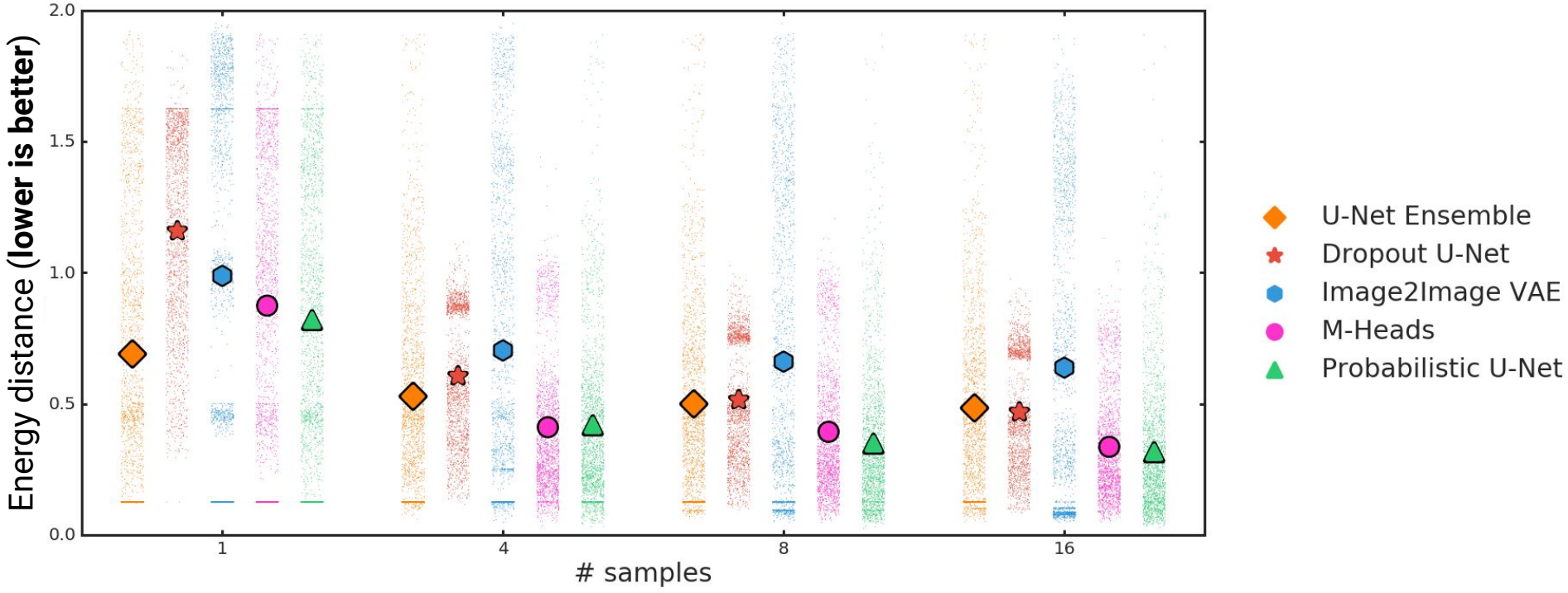
DeepMind

Lung Abnormalities Segmentation: Quantitative Results



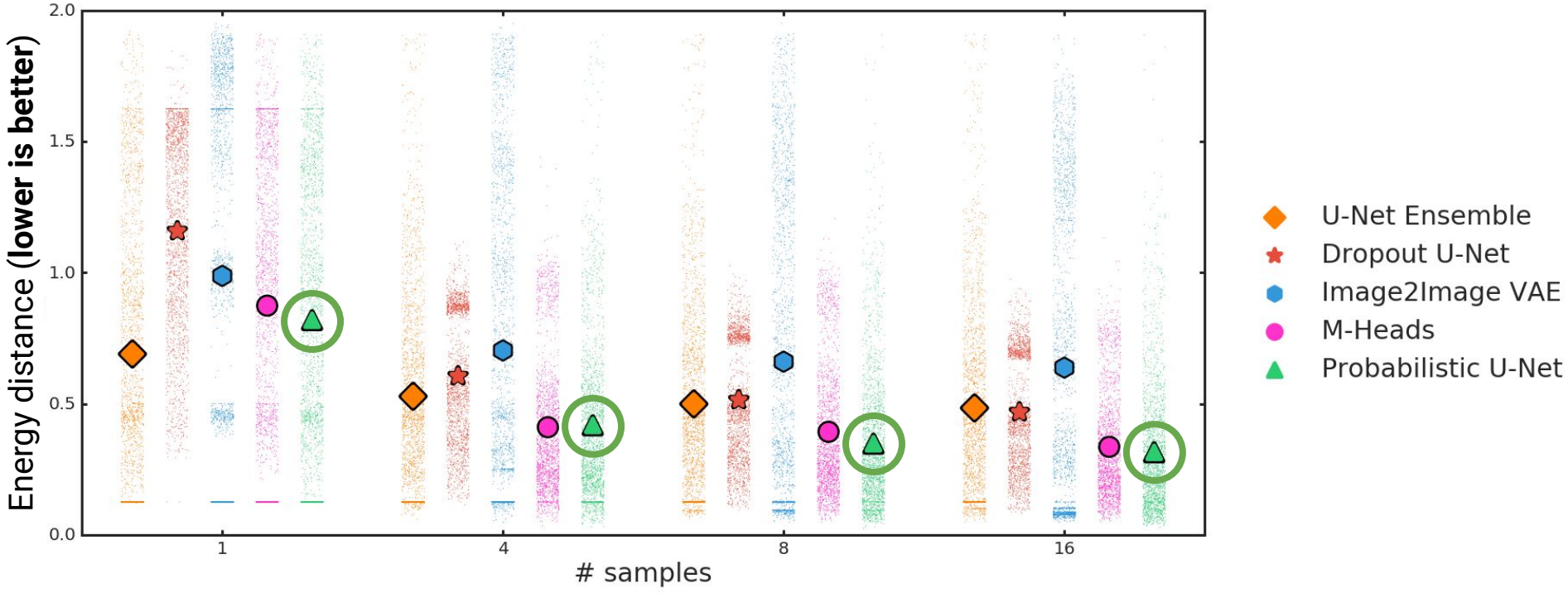
DeepMind

Lung Abnormalities Segmentation: Quantitative Results



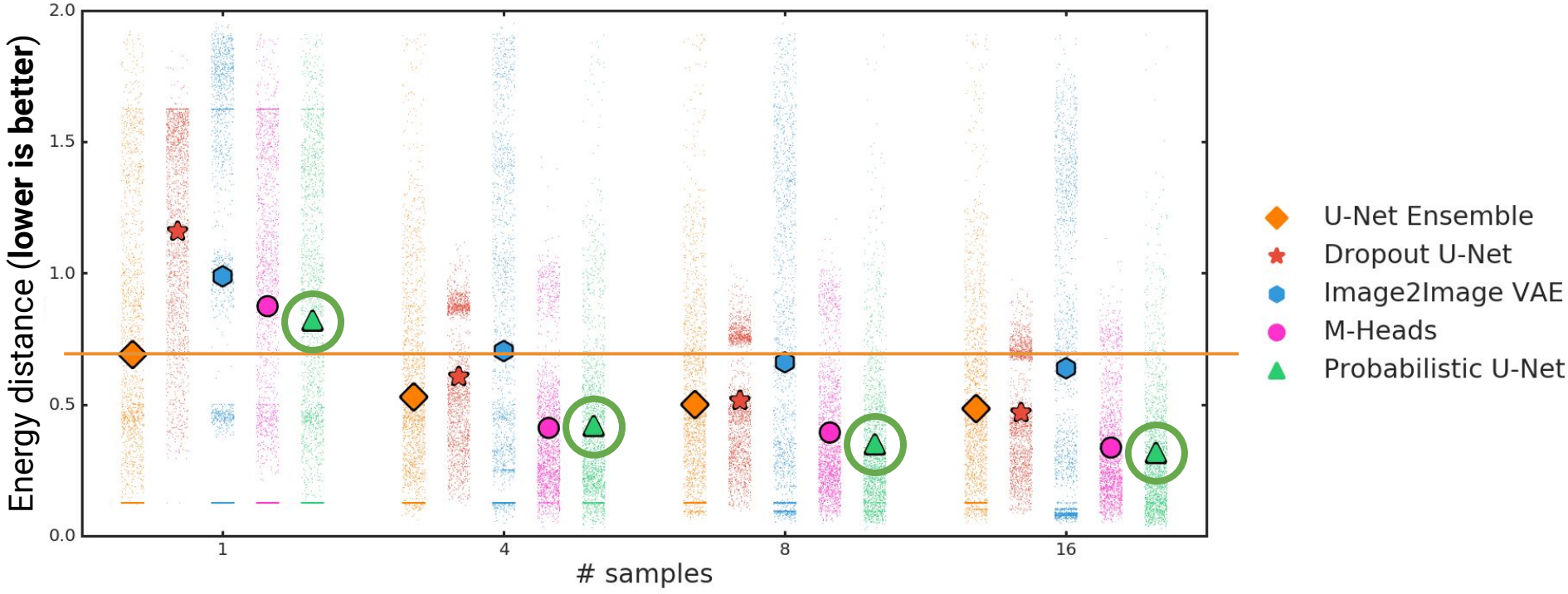
DeepMind

Lung Abnormalities Segmentation: Quantitative Results



DeepMind

Lung Abnormalities Segmentation: Quantitative Results

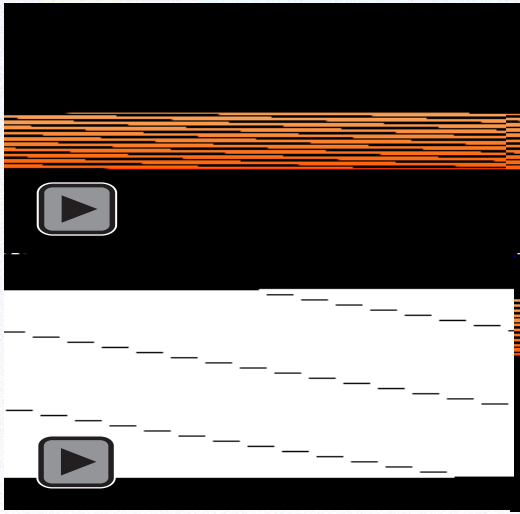


DeepMind

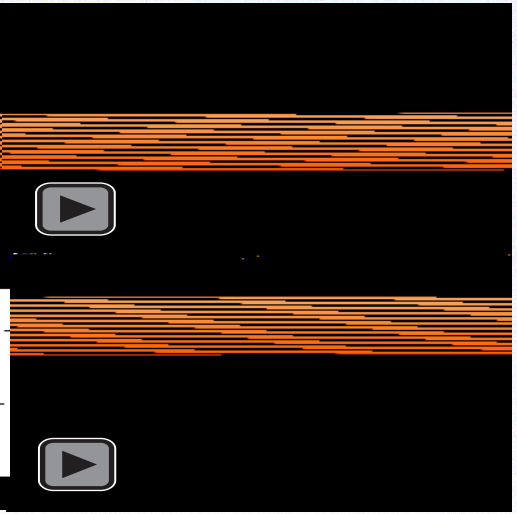
Cityscapes segmentation: Qualitative Results

Input Image

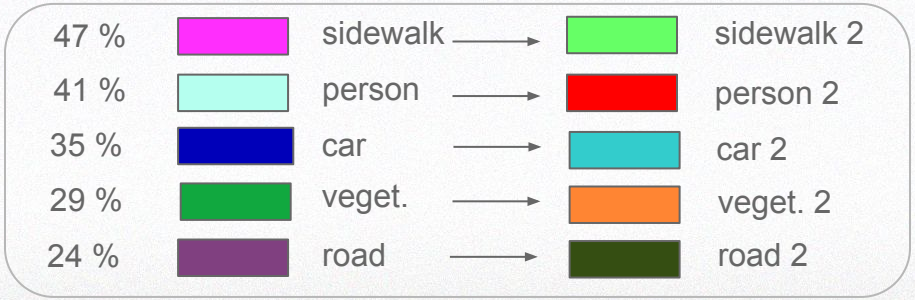
Ground-truth Grader



Samples (Probabilistic U-Net)



stochastic flips:

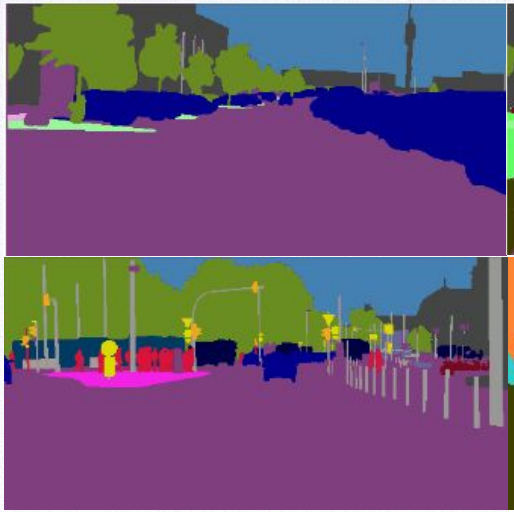


DeepMind

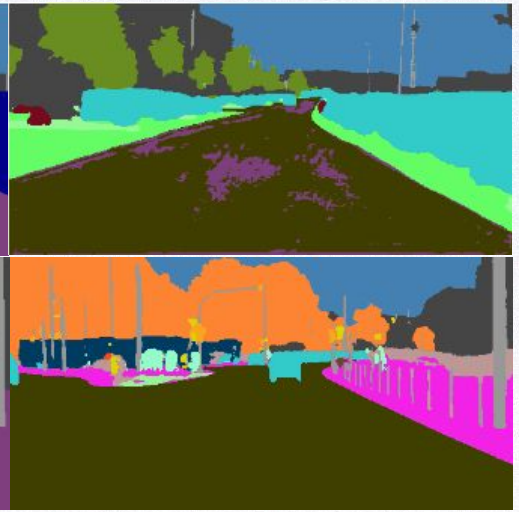
Cityscapes segmentation: Qualitative Results

Input Image

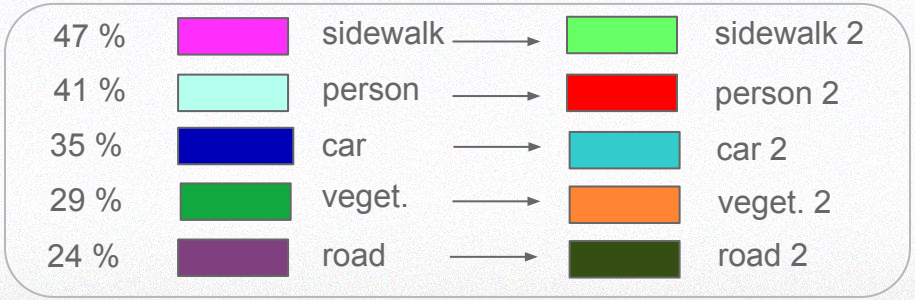
Ground-truth Grader



Samples (Probabilistic U-Net)

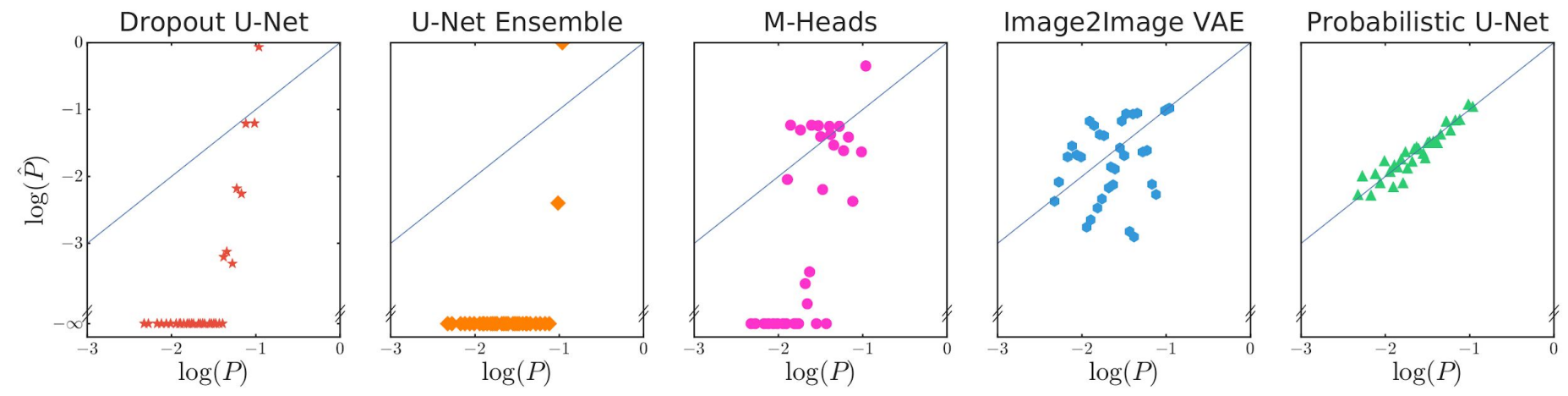


stochastic flips:



DeepMind

Cityscapes segmentation: Quantitative Results



DeepMind

Conclusions

- Learn conditional probability over segmentation maps
- Each sample is a valid & consistent segmentation
- The likelihoods are well calibrated
- Works on large-scale, real-world data
- Can also be trained with a uni-modal GT
- Can be used to asses annotations under the model

code: github.com/SimonKohl/probabilistic_unet

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], Olaf Ronneberger[1]

[1]DeepMind
[2]German Cancer Research Center

*work done during an internship at DeepMind

Poster #127

Medical Imaging Workshop Talk: Sat, Dec 8, 9:45 am

DeepMind

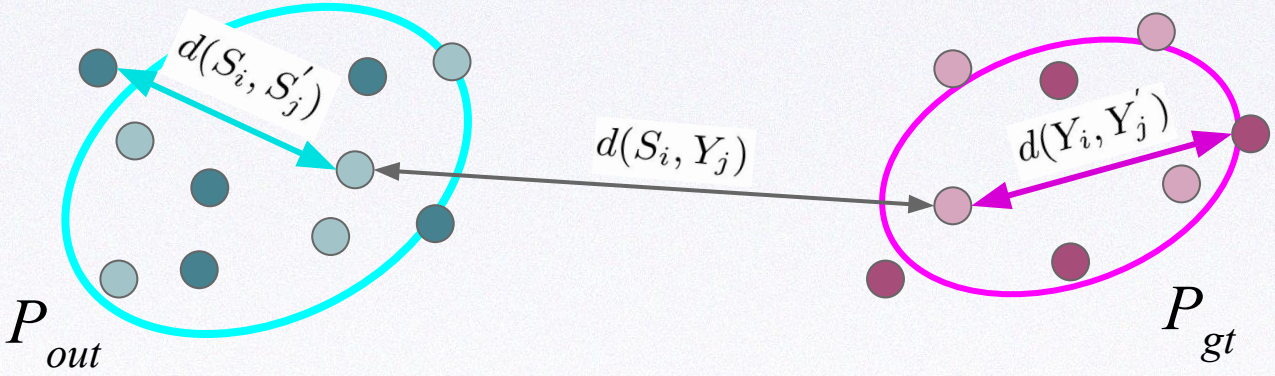
dkfz.
GERMAN CANCER RESEARCH CENTER
IN THE HELMHOLTZ ASSOCIATION

Probabilistic Segmentation: Clinical Use-Cases

- Best-fit could be picked by clinician and adjusted if necessary.
- Hypotheses could be propagated into next diagnostic pipeline steps.
- Hypotheses could inform actions to resolve ambiguities.

Evaluation Metric for Quantitative Comparison

We use the Energy Distance[1] statistic (aka MMD):

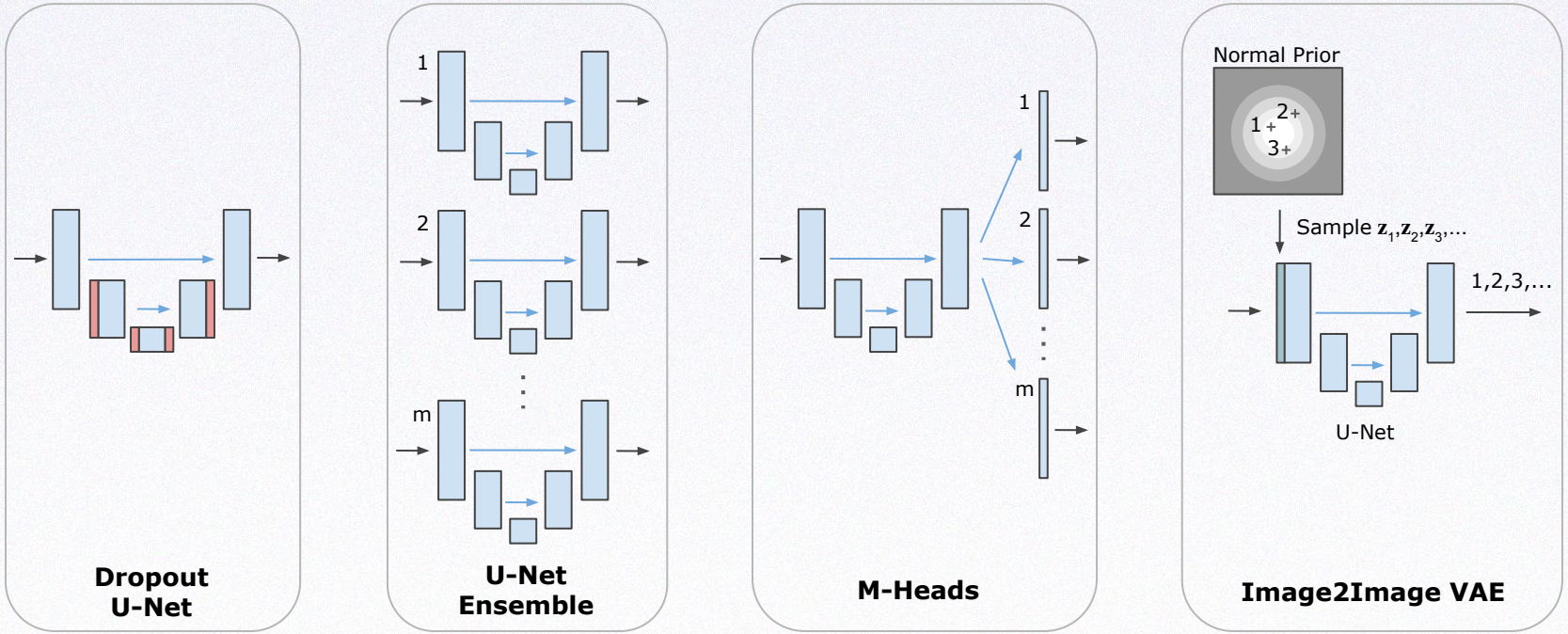


$$\hat{D}^2_{\text{GED}}(P_{\text{gt}}, P_{\text{out}}) = \frac{2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} d(S_i, Y_j) - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} d(S_i, S_j') - \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m} d(Y_i, Y_j')$$

where d(x,y) = 1 - IoU(x,y) and

[1] Székely, G.J., Rizzo, M.L.: Energy statistics: A class of statistics based on distances. Journal of statistical planning and inference 143(8) (2013) 1249–1272

Baselines


Dropout U-Net
1
2
m
U-Net Ensemble
1
2
m
M-Heads
Normal Prior
1
2
3
Sample $z_1, z_2, z_3, \ldots$
1,2,3,...
U-Net
Image2Image VAE