RUSSIAN ACADEMY OF SCIENCES
INSTITUTE FOR INFORMATION TRANSMISSION PROBLEMS
(Kharkevich Institute)

MIPT MOSCOW INSTITUTE OF PHYSICS AND TECHNOLOGY

Adapting Probabilistic U-Net for Midline Shift Detection

Alim Bukharaev, Maxim Pisov, Mikhail Belyaev

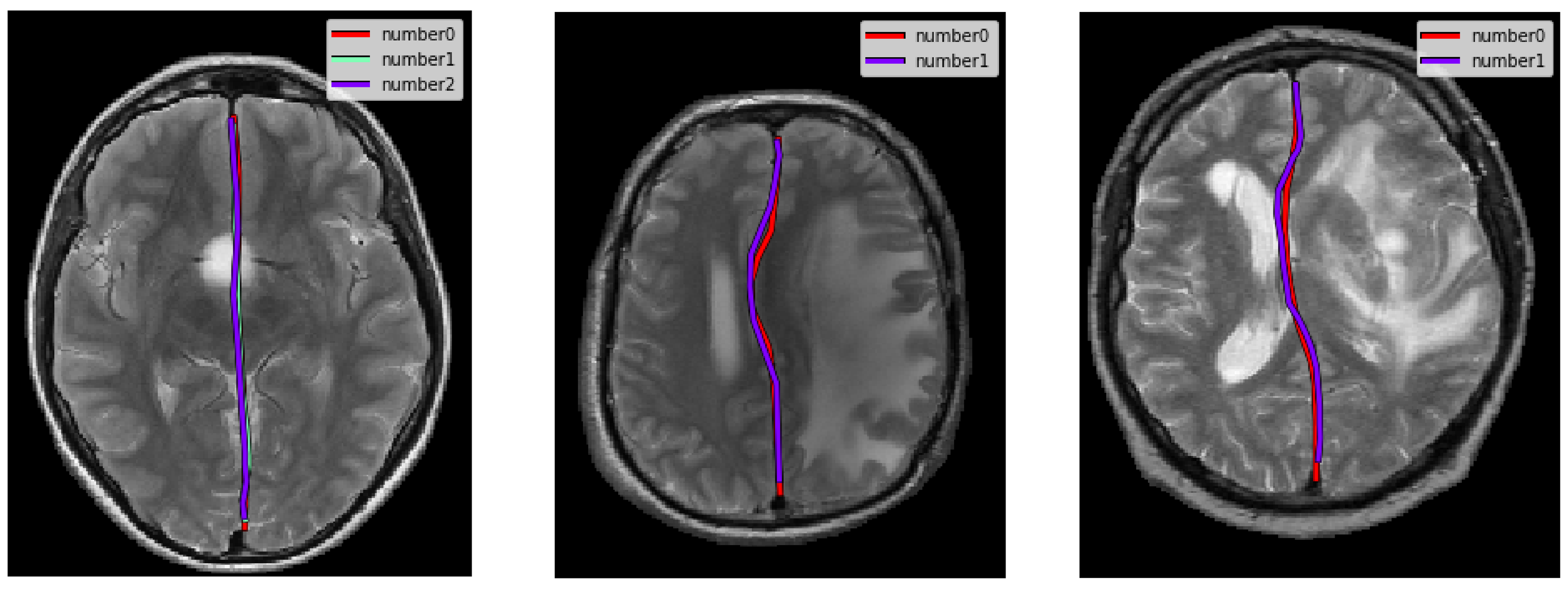
bukharaev.an@phystech.edu

Objective

In many supervised machine learning tasks researches have no access to ground truth data, but rather to a number of annotations that are merely expert *opinions*. This is especially true for problems with potentially high inter-expert variability such as segmentation.
Currently the state-of-the-art approach for segmentation with ambiguous annotations is Probabilistic U-Net, proposed in [2]. Inspired by variational autoencoders [1], the authors proposed a similar generative neural network, that predicts a ***distribution*** of segmentations for each input image. In this article we adapt the approach from [2] for the task of brain midline detection, and compare its performance with a simple U-Net [3] with the same architecture. The ability to generate such distributions would be useful in various medical image segmentation tasks, so it would be beneficial to know whether [2] is capable of that.

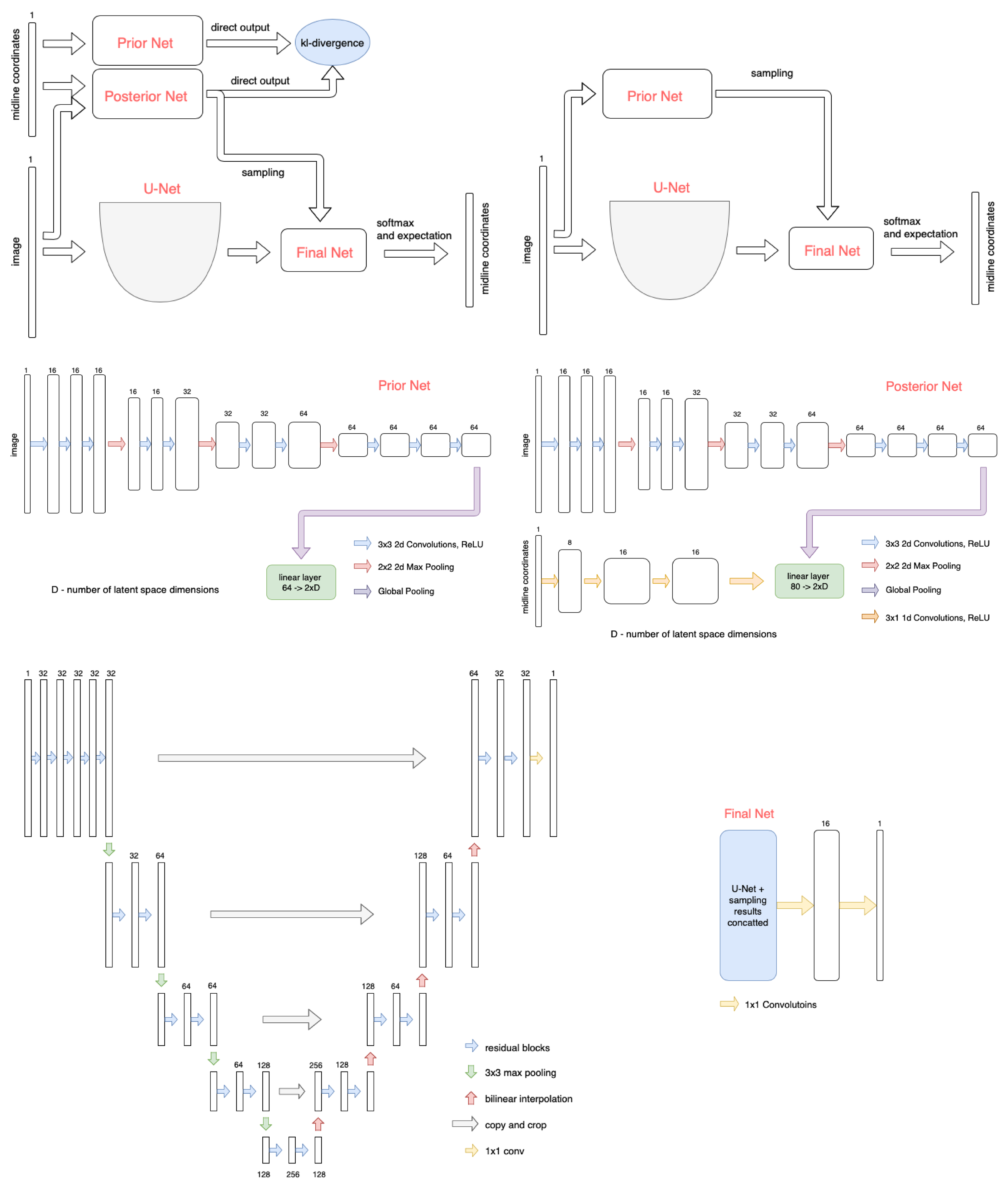
Data

The dataset consists of 352 MRI series that belong to patients with brain tumors. The dataset was labeled by an experienced radiologist (E1) and three specialists with limited experience in neuroradiology (E2-E4). The series contain axial slices with various voxel spacial sizes ranging from $0.2 \times 0.2 \times 1$mm to $1 \times 1 \times 5$mm, and modalities: T1 (25%), T2 (68%) and FLAIR (7%). The images were collected using GE and Siemens scanners.



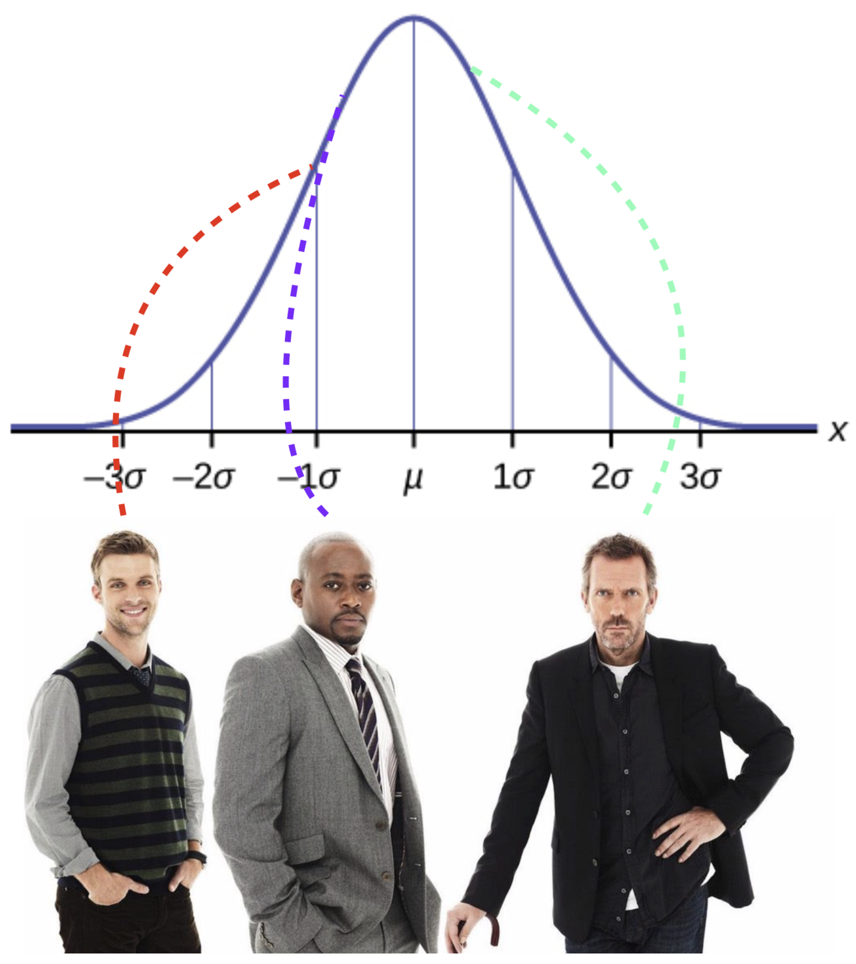
Methods

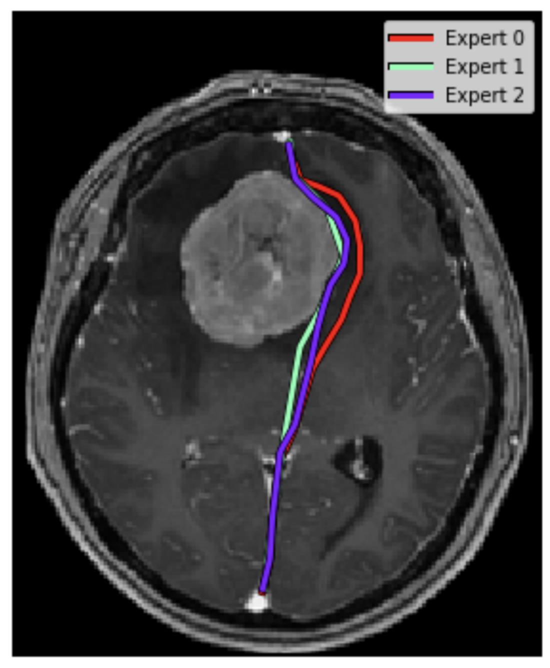
The architecture used in the article is shown in the pictures below. From top to bottom: training and inference steps; prior and posterior nets; the u-net and final net.



Desired Results

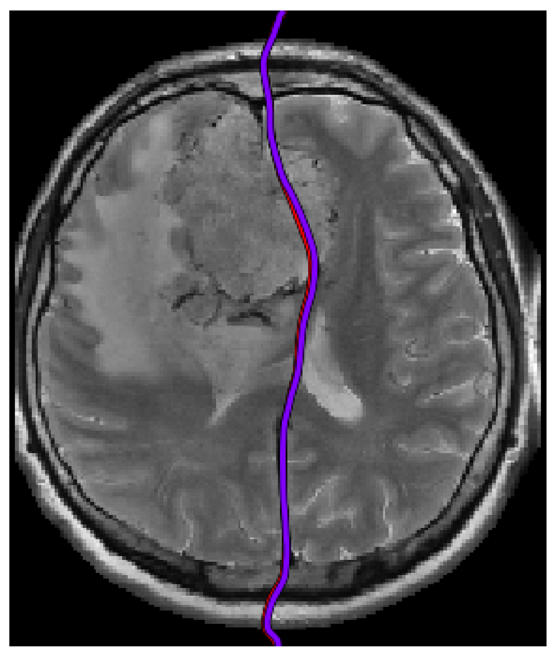
We expected the model to learn a distribution of experts' opinions upon any given brain image such that we could sample from it n times and get n independent curves (opinions).





Actual Results

However, the model failed in generating a representative distribution. μ and σ were very close to zero for any input image. In the example on the right the curves are 20σ away from each other. Moreover, no performance enhancement was achieved (in comparison with the plain u-net, see the table below). 5-fold cross-validation was applied. As for the metrics, we used mean absolute error (MAE), root-mean-square error (RMSE) and maximal error (MAX):



$$MAE(\hat{x}, x) = \frac{1}{I} \sum_{y \in I} |\hat{x}_y - x_y|$$

$$RMSE(\hat{x}, x) = \sqrt{\frac{1}{I} \sum_{y \in I} |\hat{x}_y - x_y|^2}$$

$$MAX(\hat{x}, x) = \max_{y \in I} |\hat{x}_y - x_y|$$

Table 1: Midline estimation metrics (± std) calculated on 5-fold cross-validation.

	MAE	RMSE	MAX
Plain U-Net	0.87 ± 0.68	1.24 ± 0.95	5.49 ± 3.45
Probabilistic U-Net	0.87 ± 0.61	1.25 ± 0.88	5.48± 3.37

Conclusion

The reason why our implementation of the variational u-net performed so poorly remains unclear to us. It may be a relatively low variabilty of the input data to blame. However, it is also possible that the real cause lies in the Probabilistic U-Net itself. In recent years, some researches doubted the usage of KL-divergence metric in the training process of variational autoencoders (for instance, see https://towardsdatascience.com/with-great-power-comes-poor-latent-codes-representation-learning-in-vaes-pt-2-57403690e92b). So it might be fine-tuning of the variational u-net method researches should focus on in order to make it applicable to a wider variety of image segmentation problems.

References

[1] D. P. Kingma and M. Welling. Auto-encoding variational bayes. 2013. URL http://arxiv.org/abs/1312.6114. cite arxiv:1312.6114.

[2] S. A. A. Kohl, B. Romera-Paredes, C. Meyer, J. D. Fauw, J. R. Ledsam, K. H. Maier-Hein, S. M. A. Eslami, D. J. Rezende, and O. Ronneberger. A probabilistic u-net for segmentation of ambiguous images. *CoRR*, abs/1806.05034, 2018. URL http://arxiv.org/abs/1806.05034.

[3] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015. URL http://arxiv.org/abs/1505.04597.