Skoltech
Skolkovo Institute of Science and Technology

MIDL
Montréal 2020

Bayesian generative models for knowledge transfer in MRI semantic segmentation problems

Anna Kuzina and Evgenii Egorov, Evgeny Burnaev
{a.kuzina, e.egorov, e.burnaev}@skoltech.ru

Motivation: Semantic Segmentation of MRI

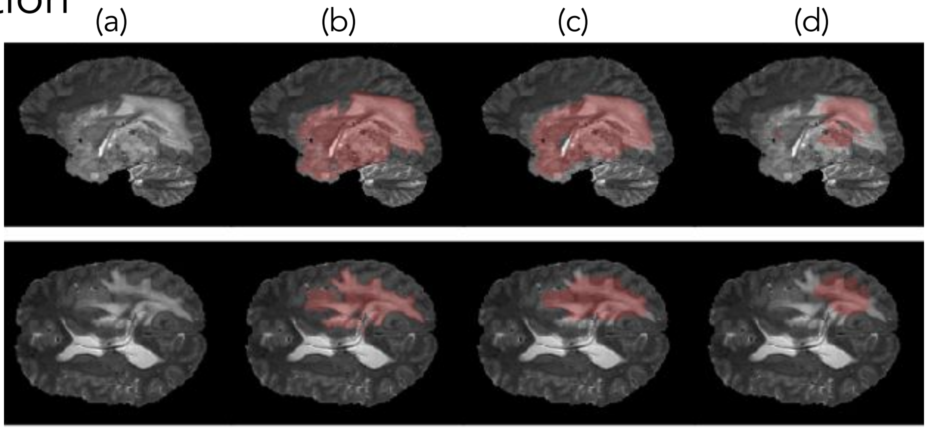
Applications in medicine

- Tumors (e.g. brain, liver) analysis and monitoring
- Multiple sclerosis plaques detection
- White matter hyperintesities detection
- etc.

Challenges:

- Expensive annotation
- Privacy concerns

→ Small sample sizes



Sample from BRATS18 dataset: (a) MRI with brain tumor; (b) ground truth segmentation; (c) prediction of the proposed model; (d) prediction of the fine-tuned model; (c,d) trained on 5 samples from BRATS18

Problem: Bad Performance on Small Datasets

How to train deep networks on *small* datasets with *high* dimensional medical objects

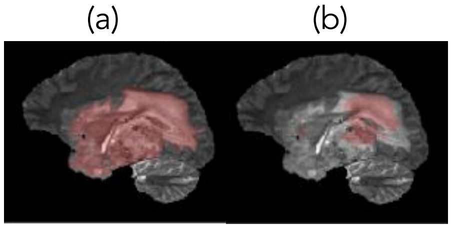
3D convolutional filters trained on large enough dataset

Important Observations

- Trained 3D convolutional filters have structure
- Bad performance of transfer learning due to disease specificity

Proposed Solution

- Train prior on “good” convolutional kernels
- Use Deep Weight Prior (DWP*) to transfer knowledge to small datasets



Predictions:
(a) of the proposed model;
(b) of the fine-tuned model;

* Atanov, A., Ashukha, A., Struminsky, K., Vetrov, D., & Welling, M. (2018). The deep weight prior.

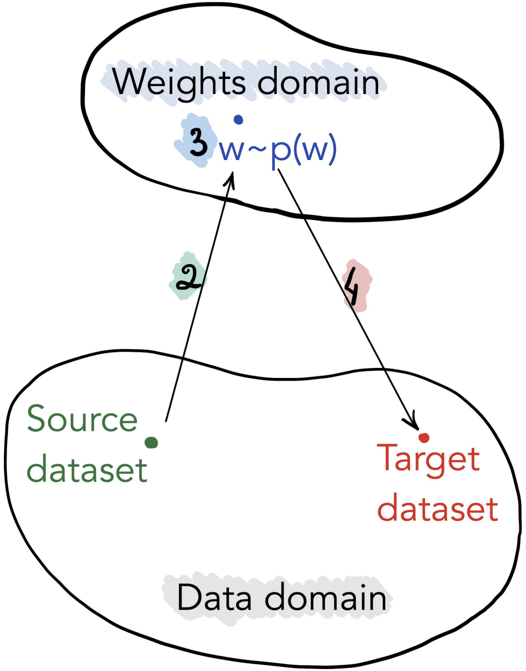
Method: Deep Weight Prior*

Main idea

Perform variational inference with implicit prior p(w) (VAE), trained on convolutional filters

Algorithm

1. Train network on the bootstrapped source dataset
2. Collect learned filters
3. Train implicit prior distribution (VAE)
4. Use trained prior for variational inference on the target dataset



* Atanov, A., Ashukha, A., Struminsky, K., Vetrov, D., & Welling, M. (2018). The deep weight prior.

Setup: Experiments on BRATS18 and MS

- Train Unet models on the full MS dataset

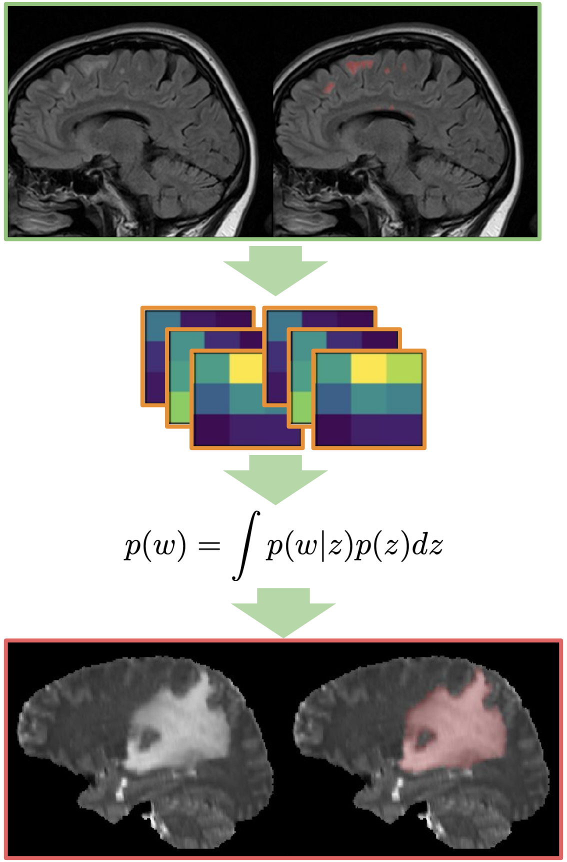
  Use bootstrapped sample from the initial dataset

- Collect filters from the trained models

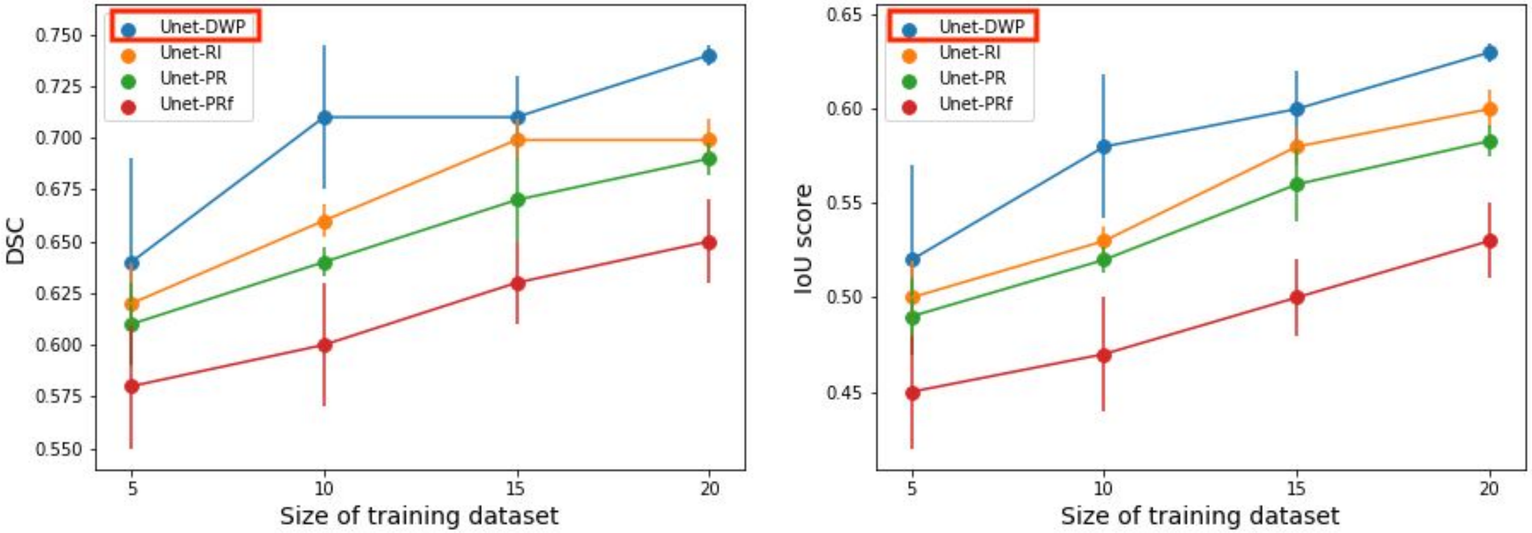
  Use cycling learning rate to expand set of learned filters

- Train VAE for each Unet block to learn $p(w)$.

- Do variational inference with implicit prior $p(w)$ on subset of BRATS18 dataset (5-20 images)



Results: Brain Tumor Segmentation Task



Unet-RI: without transfer learning
Unet-PR: fine-tuning of the whole network
Unet-PRf: fine-tuning of the input and output blocks only

Skoltech

Skolkovo Institute of Science and Technology

MIDL

Montréal 2020

Bayesian generative models for knowledge transfer in MRI semantic segmentation problems

Anna Kuzina and Evgenii Egorov, Evgeny Burnaev

[Full paper]

[Github]