Unsupervised Deep Learning based Techniques for MRI Segmentation

Medical Image Computing (CS 736)

Prof. Suyash P. Awate

Cheerla Vinay Kumar

Shubh Kumar

Abstract

Probabilistic atlas priors have been commonly used to derive adaptive and robust MRI segmentation algorithms. We plan to develop a strategy that combines a conventional probabilistic atlas-based segmentation with deep learning, enabling us to train an Unsupervised model.

Current Approaches

Bayesian segmentation of medical images, particularly in the context of MRI scans, is a well-studied problem. Most probabilistic models for image segmentation exploit atlas priors, and account for variations in contrast.

There are a lot of tools available in this area like FreeSurfer, SPM, FSL. While these tools achieve high robustness to changes in MRI contrast of the input scan, a significant drawback is that they are computationally demanding.
Recently, there has been a surge in the application of deep learning techniques, but a lot of these tools use Supervised learning models i.e., they need a manually segmented image dataset, which can be prohibitively expensive and should be changed when the contrast changes (happens when there is an upgrade in imaging techniques).

Bayesian Segmentation

Bayesian Segmentation relies on Baye's rule to derive the posterior probability distribution of the segmentation S (segmented labels) given the input image I (intensities of voxels).

We can estimate the labels as the mode of the posterior constructed.

$$\hat{S} = \arg\max_{S} p(S|I) = \arg\max_{S} p(I|S)p(S)$$

Both the prior and likelihood have a set of associated parameters θ_s, and θ_l, which describes attributes such as label probabilities and atlas deformation.

First, one estimates the mode of the posterior distribution for the parameters, and use the point estimates to get the final mode of the posterior, which can be computed using the EM algorithm.

Parameter Estimation

A standard approximation is to use point estimates for the parameters. We can estimate the mode of the posterior distribution for the parameters, assuming independence between parameters of the prior and likelihood.

$$(\hat{\theta}_S, \hat{\theta}_l) = \arg\max_{\theta_S, \theta_I} p(\theta_S, \theta_I | I) = \arg\max_{\theta_S, \theta_I} p(\theta_S) p(\theta_I) \sum_S p(I | S, \theta_I) p(S | \theta_S)$$

This computation often requires estimating an atlas deformation in θ_s and intensity parameters in θ_l, and is typically achieved with a combination of numerical optimization and EM algorithm.

Given the point estimates, the final segmentation is computed efficiently as:

$$\hat{S} = \arg\max_S p(S \mid \hat{\theta}_S, \hat{\theta}_I, I) = \arg\max_S p(I \mid S, \hat{\theta}_I) p(S \mid \hat{\theta}_S)$$

and is often produced directly by the same EM algorithm.

Proposed Model

The current existing models either use Supervised learning, or takes a lot of time to run. But this proposed model is fast, and is unsupervised.

We take a probabilistic atlas A as the prior distribution. The prior probability distribution is modeled using a diffeomorphic transform of the atlas, which is denoted by ϕ and parameterized by a stationary velocity field υ. Specifically, the prior probability distribution is determined by setting the parameters θ_s to be equal to the velocity field υ, such that θ_s = υ. The transformation ϕ_υ is computed as the exponential function of the velocity field, i.e., ϕ_υ = exp[υ]. Assuming independence over voxels:

$$p(S|\theta_S; A) = p(S \mid v; A) = \prod_{j \in \Omega} A(S_j, \phi_v(x_j))$$

Where S_j is the segmentation at voxel j, and x_j is its spatial location.

We also discourage strongly varying deformations by penalizing the spatial gradient $\nabla \boldsymbol{u}$_v of displacement u_v.

Proposed Model

We also discourage strongly varying deformations by penalizing the spatial gradient **∇*u*_**v of displacement u_v, where ϕ_υ = Id + u_v:

$$p(\theta_S; \lambda) = p(v; \lambda) \propto \exp[-\lambda |\nabla u_v|^2]$$

Conditioned on a segmentation, we assume that the observed intensities at different voxel locations are independent samples of Gaussian distributions:

$$p(I \mid S, \theta_I) = p(I \mid S, \mu, \sigma^2) = \prod_{j \in \Omega} \mathcal{N}(I_j; \mu_{S_j}, \sigma^2_{S_j})$$

Where N(., μ, σ^2) is the Gaussian distribution, Ij is the image intensity at voxel j, and the likelihood parameters θ_l = { μ, σ^2 } are L means μ_l, and variance σ_l^2.

Learning

To avoid computationally expensive optimization typically required for maximum a posteriori estimation, we propose to train a CNN to estimate the parameters directly from an input scan. Specifically, we design a CNN with convolutional parameters θ_c that takes as input a scan I and the probabilistic atlas A, and outputs the model parameters v, μ, σ^2 for that scan. The loss function used is:

$$-\sum_{n=1}^{N} \log p(v^n, \mu^n, [\sigma^2]^n | I^n; A, \lambda)$$

which is also equal to:

$$-\sum_{n=1}^{N} \sum_{j \in \Omega} \log \left[\sum_{l=1}^{L} \mathcal{N}\left(I_j^n; \mu_l^n, [\sigma^2] l^n\right) A(l, \phi_{v_m}(x_j)) \right] + \lambda ||\nabla {u_n}^v||^2 - K(\lambda) + \text{const}$$

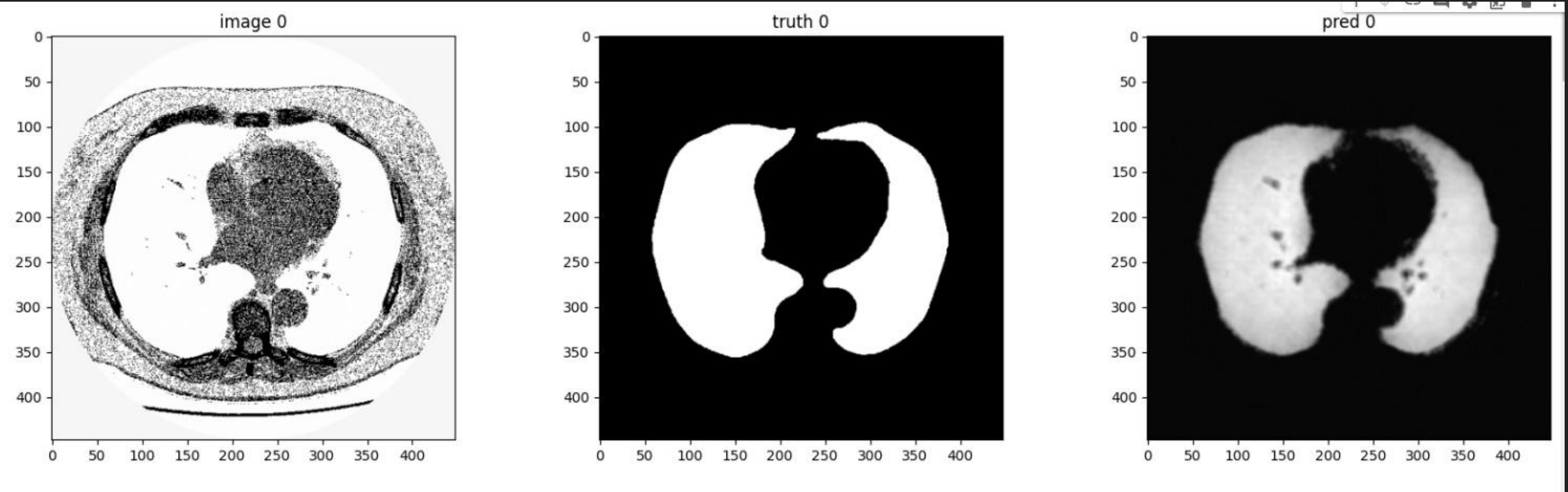
Where K(λ) is a log-partition function that depends on the hyperparameter λ.

Efficient Segmentation

Given a trained network and a new test subject, the network efficiently provides the image-specific parameter point estimates v, and θ_l via a single forward pass.
The optimal segmentation can be efficiently computed for each voxel:

$$\hat{S}_j = \arg\max_l N(I_j; \hat{\mu}l, \hat{\sigma}_l^2) A(l, \phi\hat{v}(x_j))$$

Results



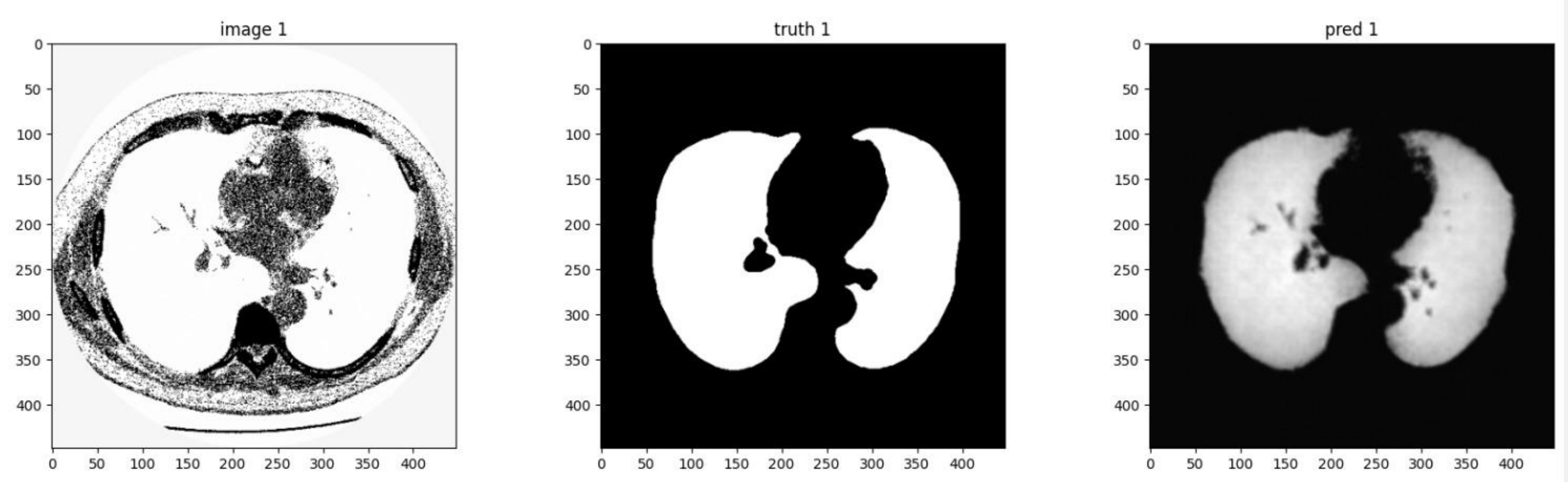


image 1
truth 1
pred 1
0
50
100
150
200
250
300
350
400

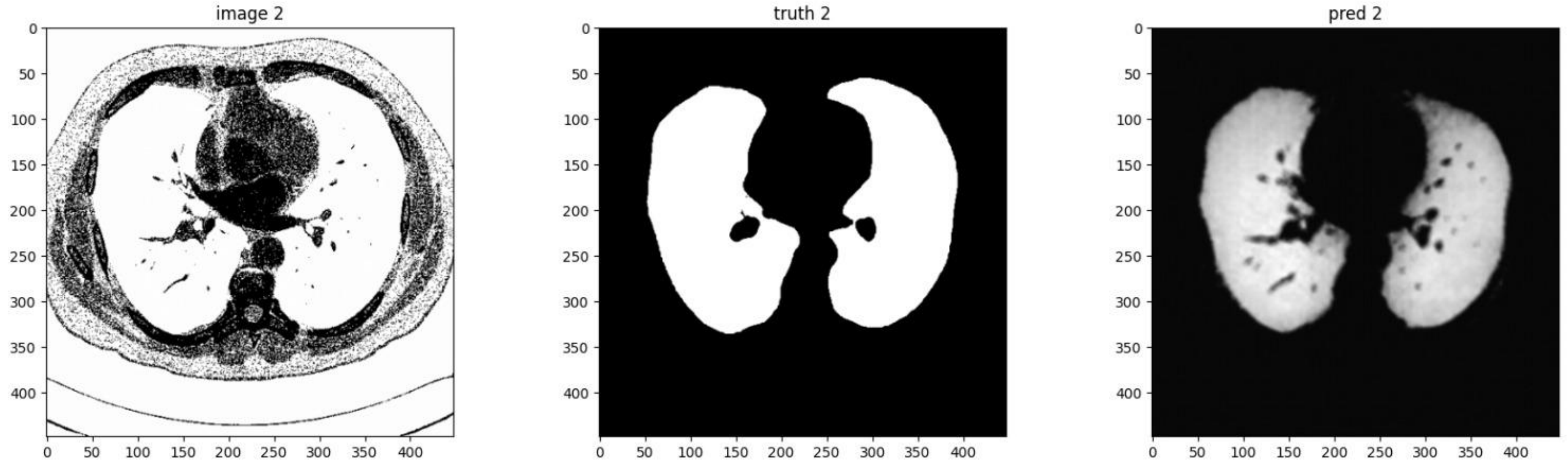


image 2
truth 2
pred 2
0
50
100
150
200
250
300
350
400

References

Unsupervised Deep Learning for Bayesian MRI Segmentation by AV Dalca, Evan Yu, Polina Golland, Bruce Fischl, Mert R.Sabuncu, Juan Eugenio Iglesias

Unsupervised Learning for Fast Probabilistic Diffeomorphic Registration by Adrian V. Dalca, Guha Balakrishnan, John Guttag, Mert R. Sabuncu

A Log-Euclidean Framework for Statistics on Diffeomorphisms by Vincent Arsigny, Olivier Commowick, Xavier Pennec & Nicholas Ayache

Thank You!

Cheerla Vinay Kumar (200050027)
Shubh Kumar (200050134)