Efficient Transfer Learning for Cardiac landmark Localization Using Rotational Entropy

Samira Masoudi[*1,2] SMasoudi@ucsd.edu
Kevin Blansit[*3] kevin.blansit@gmail.com
Naeim Bahrami[4] naeim.bahrami@gmail.com
Albert Hsiao[1,2] a3hsiao@ucsd.edu

[*] *these authors contributed equally to this work*

[1] *Halıcıoğlu Data Science Institute, University of California San Diego, San Diego, CA, USA*
[2] *Department of Radiology, University of California San Diego, San Diego, CA, USA*
[3] *Department of Biomedical informatics, University of California San Diego, San Diego, CA, USA*
[4] *GE Healthcare, Menlo Park, CA, USA*

Editors: Under Review for MIDL 2022

Abstract

Transfer learning is a common technique to address model generalization among different sources, which requires additional annotated data. Herein, we proposed a novel strategy to select new data to be annotated for transfer learning of a landmark localization model, minimizing the time and effort for annotation and thus model generalization. A CNN model was initially trained using 1.5T images to localize the apex and mitral valve on the long axis cardiac MR images. Model performance on 3T images was reported poor, necessitating transfer learning to 3T images. *Rotational entropy*, was introduced not only as a surrogate of model performance but as a metric which could be used to prioritize the most informative cases for transfer learning.

Keywords: cardiac landmark localization, rotational entropy.

1. Introduction

Cardiac MR, essential for cardiovascular function assessment, is usually acquired at 1.5T field strength. However, there is a growing interest to use 3T scanners for this purpose. Inspired by (Payer et al., 2020), a cardiac landmark localization model called LAXLoc-Net was trained to localize the apex and mitral valve in long axis MR images. The ground truth for i^{th} landmark was established by a target pseudoprobability map, P_i defined as a circular Gaussian function centered at the expert-defined ground truth coordinate vector ${X^*}_i$ and a fixed standard deviation vector Σ_i (Blansit et al., 2019).

$$P_i(X) = e^{-\frac{(X-{X^*}_i)^2}{2{\Sigma_i}^2}} \tag{1}$$

With the training-time shift, zoom, and rotational augmentation, LAXLoc-Net was optimized through minimization of the L2 loss computed between the ground truth (P_i) and predicted ($\hat{P}_i$) pseudoprobability maps. The peak point of $\hat{P}_i$ was then extracted as the predicted landmark coordinate vector $\hat{X}_i$. Having the model initially trained with 1.5T images, caused it to make erroneous predictions on the majority of the 3T images. To address the model's undertraining in case of 3T images, we adopted transfer learning which involves fine-tuning the model with more annotated images acquired at 3T. Having an efficient strategy to sample from a pool of 3T images to be included in the step 2 of training, can minimize the time and effort required for the annotation. For this, we saught to find a metric to be calculated per image, where we hypothesized that unstable response to rotational disturbances could be a sign of model's insufficiency.

2. Materials and Methods

A schematic for rotational entropy is depicted by Figure 1-A. Given the test-time augmented image I^{θ_j}, rotated by θ_j about the center of the input image I, model's prediction of psudoparability map would be $\hat{P}_i^{\theta_j}$ for landmark i. We then define the rotational entropy E_i across the rotational range of $\{\theta_1, \theta_2, ..., \theta_M\}$ as follows:

$$E_i = \frac{1}{M} \Sigma_{j=1}^{M} |\hat{P}_i - (\hat{P}_i^{\theta_j})^{-\theta_j}| \tag{2}$$

where $\hat{P}_i$ represents the i^{th} heatmap prediction from image I and $(\hat{P}_i^{\theta_j})^{-\theta_j}$ stands for re-oriented prediction of the oriented image, I^{θ_j}. The larger spatial spread of the rotational entropy indicates lower model stability. To translate rotational entropy into a single score measure, we used the spatial variance of E_i.

$$s_i = \sqrt{(\frac{M_{2,0,i}}{M_{0,0,i}} - (\frac{M_{1,0,i}}{M_{0,0,i}})^2)^2 + (\frac{M_{0,2,i}}{M_{0,0,i}} - (\frac{M_{0,1,i}}{M_{0,0,i}})^2)^2}, \quad M_{k,t,i} = \sum_k \sum_t x^k y^t E_i(x,y) \tag{3}$$

With HIPPA compliance and IRB approval, we retrospectively collected 405 cardiac MRI studies at our institution, 285 of which were collected at 1.5T and the remaining 120 images were collected at 3T. An expert radiologist located apex and mitral valve coordinates on long axis MR images to be used for ground truth.

We trained our initial LAXLoc-Net using 1.5T long axis images. A 2D U-net was modified for heatmap localization by setting the activation function at the final layer to be linear and having L2-error loss for training.

Transfer Learning was then used to adapt LAXLoc-Net from 1.5T to 3T. In an experiment, we used incremental number of 3T training cases for transfer learning. These cases were selected using 3 strategies: 1) descending s_i, 2) Ascending s_i, and 3) random (with 20 repetitions). We also used five-fold cross validation where each fold was independently trained using 60% of the 3T transfer learning images and validated on 40% of the remaining images.

3. Results and Discussion

To assess the performance of our landmark localization model, we measured the localization error in terms of the distance between the predicted $\hat{X}_i$ and ground truth X_i^*. The initial LaxLocNet errors of 9.64 and 7.18 mm for localizing the apex and mitral valve in 1.5T images grew worse in 3T images (29.79 mm ($p < 0.01$) for apex and 15.44 mm ($p < 0.01$) for mitral valve). Our analysis of localization error in this study indicated that s_i could be utilized as a surrogate for model performance. Illustrated by Figure 1-B,& 1-C, lower spatial variance in rotational entropy, s_i, was associated with low localization error for either apex and mitral valve. These observations imply the potential to use s_i as a proxy for model uncertainty. We furthur compared the results of transfer learning using 3 sampling strategies from 3T images. While the naïve strategy-3 resulted in a decremental error for apex localization, using strategy-1 required less data to achieve the same level of performance. Using strategy-2 had worse transfer effect than choosing images at random (Figure 1-D). Our experiment demonstrates the potential use of s_i as model uncertainty for efficient transfer learning.

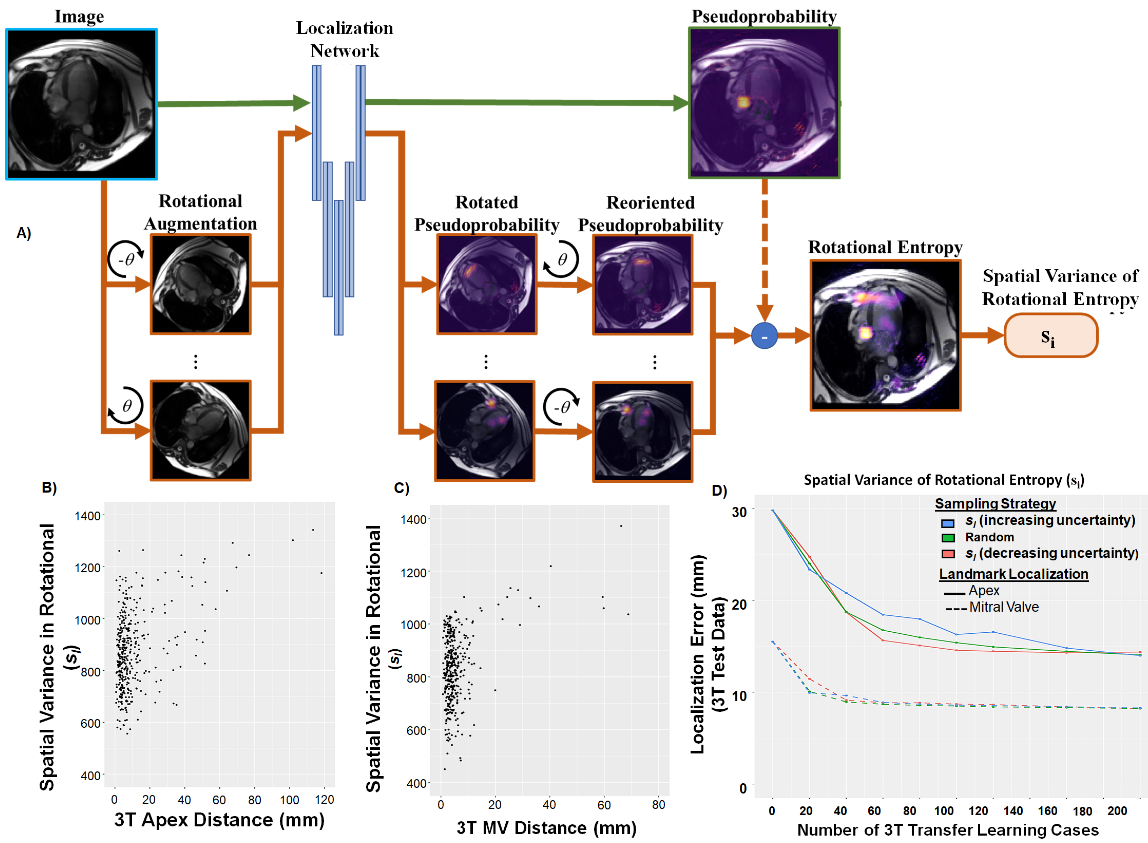

Figure 1: A) Shematic for rotational entropy, B and C) relation between s_i and localization errors, D) Comparison of 3 sampling strategies for transfer learning.

Acknowledgments

We would like to acknowledge research grant support from GE Healthcare.

References

Kevin Blansit, Tara Retson, Evan Masutani, Naeim Bahrami, and Albert Hsiao. Deep learning–based prescription of cardiac mri planes. *Radiology: Artificial Intelligence*, 1(6): e180069, 2019.

Christian Payer, Martin Urschler, Horst Bischof, and Darko Štern. Uncertainty estimation in landmark localization based on gaussian heatmaps. In *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging, and Graphs in Biomedical Image Analysis*, pages 42–51. Springer, 2020.