Arjun Balaji - Weekly Report

Literature Survey on Loss Functions on Shape Priors:

Paper	Explanation	How it can help
Shape Description losses for medical image segmentation	Shape-based features are first computed from the probability map and the ground truth segmentation label, respectively. Then a group of losses, including volume loss, surface area loss, center localization loss, and bounding box loss, are computed using the extracted shape-based features. These losses are complementary to the common mainstream losses	This loss can help by adding global geometrical constraints to the new model, ensuring that the predicted shapes align with expected anatomical structures.
A Surprisingly Effective Perimeter-based Loss for Medical Image Segmentation	Focuses on minimizing boundary errors by using a perimeter-based loss that penalizes errors along the ground truth boundary, helping the model preserve object boundaries.	For the model, this loss could be combined with shape prior losses to improve boundary precision when integrating the shape prior. This cam be really helpful where boundary delineation is essential like in cardiac use-cases like ours.
FourierLoss: Shape-Aware Loss Function with Fourier Descriptors	Utilizes Fourier descriptors, which are mathematical representations of shapes, to calculate shape dissimilarity between predicted and ground truth segmentations. This loss encourages shape consistency.	This can help the model maintain accurate shape predictions, especially when working with shape priors. It could be applied to enforce shape consistency in regions of interest identified by the shape prior like sections of the heart.
The Benefits of Incorporating Shape Priors in Contrastive Learning [Novel Idea]	Excerpt from Paper: Introduced a novel contrastive learning framework called Shape Prototype Contrastive Learning (S-PCL). In this framework, global shape in the form of the silhouette	For 3D cardiac segmentation, shape priors (like heart contours) would help the model focus on the overall heart shape early in training, speeding up learning. The coarse-to-fine approach could

Paper	Explanation	How it can help
(ICLR 2024)	of an object, extracted using existing methods on figure-ground segregation, is used as to augment the input training image in contrastive learning. Shape prototypes, the cluster centroids of silhouette shapes of objects in the data set in the embedding space, are learned with momentum clustering and then used to organize the embedding space of the original input images. We found that deep networks learned with these shape prototype priors exhibit stronger shape representations that are more aligned with human perception. Furthermore, we found that S-PCL accelerates the learning process, particularly in the early stage of development, reminiscent of the impact of global shapes in children's lexical categorical learning.	first segment the heart's general shape, then refine details. The shape bias introduced by S-PCL would also make the model more robust to variations in patient data I think.

Survey on Potential Baseline models:

Other than nnUNet and MedSAM2

Model	Abstract	Why we need to compare
Deep Separable Spatiotemporal Learning for Fast Dynamic Cardiac MRI	Dynamic magnetic resonance imaging (MRI) plays an indispensable role in cardiac diagnosis. To enable fast imaging, the k-space data can be undersampled but the image reconstruction poses a great challenge of highdimensional processing. This challenge necessitatesextensive training data in deep learning reconstruction methods. In this work, we propose a novel and	Uses spatiotemporal priors

Model	Abstract	Why we need to compare
	efficient approach, leveraging a dimension-reduced separable learning scheme that can perform exceptionally well even with highly limited training data. We design this new approach by incorporating spatiotemporal priors into the development of a Deep Separable Spatiotemporal Learning network (DeepSSL), which unrolls an iteration process of a 2D spatiotemporal reconstruction model with both temporal low-rankness and spatial sparsity. Intermediate outputs can also be visualized to provide insights into the network behavior and enhance interpretability. Extensive results on cardiac cine datasets demonstrate that the proposed DeepSSL surpasses state-of-the-art methods both visually and quantitatively, while reducing the demand for training cases by up to 75%. Additionally, its preliminary adaptability to unseen cardiac patients has been verified through a blind reader study conducted by experienced radiologists and cardiologists. Furthermore, DeepSSL enhances the accuracy of the downstream task of cardiac segmentation and exhibits robustness in prospectively undersampled real-time cardiac MRI.	
Deep Conditional Shape Models for 3D Cardiac Segmentation (MICCAI 2023)	Delineation of anatomical structures is often the first step of many medical image analysis workflows. While convolutional neural networks achieve high performance, these do not incorporate anatomical shape information. We introduce a novel segmentation algorithm that uses Deep	Uses shape priors

Model	Abstract	Why we need to compare
	Conditional Shape models (DCSMs) as a core component. Using deep implicit shape representations, the algorithm learns a modality-agnostic shape model that can generate the signed distance functions for any anatomy of interest. To fit the generated shape to the image, the shape model is conditioned on anatomic landmarks that can be automatically detected or provided by the user. Finally, we add a modality-dependent, lightweight refinement network to capture any fine details not represented by the implicit function. The proposed DCSM framework is evaluated on the problem of cardiac left ventricle (LV) segmentation from multiple 3D modalities (contrast-enhanced CT, non-contrasted CT, 3D echocardiography-3DE). We demonstrate that the automatic DCSM outperforms the baseline for non-contrasted CT without the local refinement, and with the refinement for contrasted CT and 3DE, especially with significant improvement in the Hausdorff distance. The semi-automatic DCSM with user-input landmarks, while only trained on contrasted CT, achieves greater than 92% Dice for all modalities. Both automatic DCSM with refinement and semi-automatic DCSM achieve equivalent or better performance compared to inter-user variability for these modalities.	
TotalSegmentor	Tool for segmentation of most major anatomical structures in any CT or MR image. It was trained on a wide range of different CT and MR images	Standard out of the box tool for any organ

Model	Abstract	Why we need to compare
	(different scanners, institutions, protocols,...) and therefore should work well on most images. A large part of the training dataset can be downloaded here: CT dataset (1228 subjects) and MR dataset (298 subjects). You can also try the tool online at totalsegmentator.com or as 3D Slicer extension.	
TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation	Medical image segmentation is an essential prerequisite for developing healthcare systems, especially for disease diagnosis and treatment planning. On various medical image segmentation tasks, the ushaped architecture, also known as U-Net, has become the de-facto standard and achieved tremendous success. However, due to the intrinsic locality of convolution operations, U-Net generally demonstrates limitations in explicitly modeling long-range dependency. Transformers, designed for sequence-to-sequence prediction, have emerged as alternative architectures with innate global self-attention mechanisms, but can result in limited localization abilities due to insufficient low-level details. In this paper, we propose TransUNet, which merits both Transformers and U-Net, as a strong alternative for medical image segmentation. On one hand, the Transformer encodes tokenized image patches from a convolution neural network (CNN) feature map as the input sequence for extracting global contexts. On the other hand, the decoder upsamples the encoded features which are then combined with the high-resolution CNN feature maps	Standard baseline method for segmentation tasks.

Model	Abstract	Why we need to compare
	to enable precise localization. We argue that Transformers can serve as strong encoders for medical image segmentation tasks, with the combination of U-Net to enhance finer details by recovering localized spatial information. TransUNet achieves superior performances to various competing methods on different medical applications including multi-organ segmentation and cardiac segmentation.	