GAN-Driven Anomaly Detection for Active Learning in Medical Imaging Segmentation

Automatic image segmentation is crucial for automation in cancer detection, image-guided treatment, and response assessment. MD Anderson has several ongoing initiatives to develop and clinically deploy automated segmentation of diseased and normal tissues to support tumor measurement initiatives, clinical trials in image-guided interventions, and image quality assurance and quality control. To achieve these goals, extensive research has been performed to develop image segmentation models using deep learning. A critical barrier to the deployment of these models is the need to detect "anomalous" cases where the segmentation model may fail, even after significant external validation has been performed.

Aim: *Develop interpretable, deep learning-based anomaly detectors to improve the robustness of segmentation models*. I will build a generative adversarial network (GAN)-based anomaly detector that will allow for the safe integration of automatic segmentation models into a clinical workflow. The detector will flag images that a segmentation model will not perform well on so that they can be manually segmented (Figure 1). These manually segmented images could then be added to the training set of the automatic segmentation model, consequently diversifying both the dataset and the model's proficiencies.

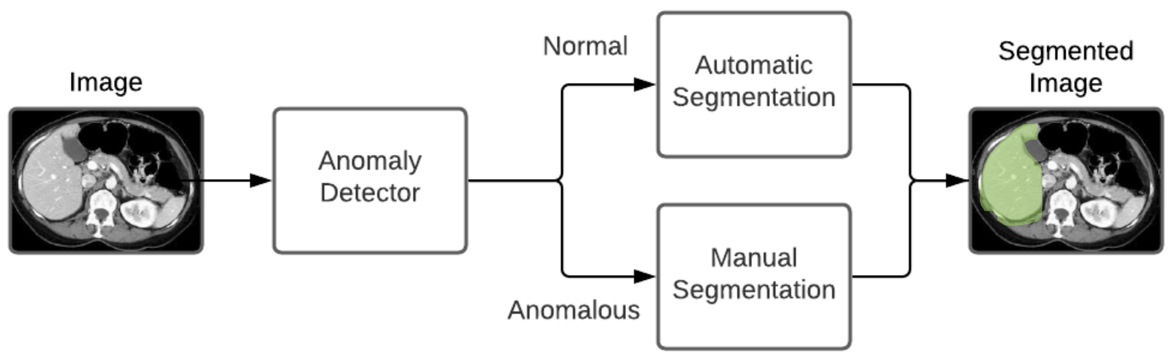

Figure 1: An active learning paradigm for the integration of automatic segmentation models into a clinical workflow.

Significance

Deep learning algorithms are the state-of-the-art automatic segmentation models for medical imaging, with research spanning many anatomical regions and imaging modalities[1]. Although these models perform exceptionally well on new data that is similar to the data that they were trained on, they do not generalize well to novel structures (Figure 2). Due to this poor generalization capability, the success of many deep learning models today hinges on training on large amounts of labeled data. For most medical imaging applications, this is infeasible, due to both the cost of labeling and protecting patient privacy. I propose anomaly detectors as a solution. An anomaly detector could flag images that a deployed deep learning model will not perform well on. It could also detect the images in an unlabeled dataset that contain the most unique features, thereby prioritizing which images should be labeled manually.

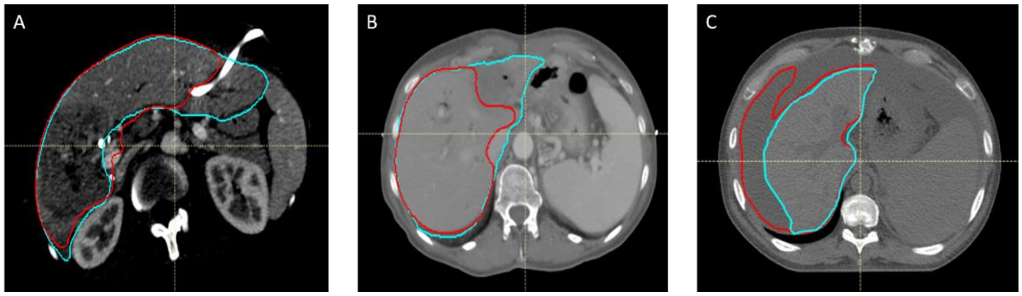

Figure 2: Instances where a trained liver segmentation model failed due to abnormalities not present in the training dataset: (A) a stent, (B) a tumor, and (C) fluid buildup. Red is the automated segmentation and light blue is the manual contour. The model had a Dice coefficient over 0.96 on 50 unseen computed tomography (CT) scans. Thank you to Brian Anderson for this image.

Innovation

Using GANs to detect anomalies has recently become an active area of research. Schlegl et al. were the first to use a GAN for anomaly detection. They used it to locate retinal fluid in OCT images[2,3]. Their inaugural work has since been applied to various medical domains and applications[4-12]. In every paper, GANs were used to flag diseases (tumors, hemorrhages, etc.) I will be the first to use GAN-based anomaly detectors for active learning, a field of machine learning where a model may query a user if it is uncertain about a given input. Current GAN methods on medical imaging modalities generate slices (or patches) of medical images independently from one another. In contrast, I plan to exploit the intricacies of medical imaging modalities by incorporating the locational relationship between slices.

Methods

I will build an anomaly detector that will function as an active labeler of an automatic segmentation model. If the detector finds an image to be within its training distribution, the image will be automatically segmented. Otherwise, the image will be manually segmented (Figure 1).

To model the training distribution, I will use a semi-supervised StyleGAN[14] due to its performance on high-resolution images. I plan to incorporate the 3D locational information into the GAN by experimenting with both conditioning the GAN on the relative slice location and by making the GAN 3D. I will then train an encoder to map images to the generator's latent space. With this encoder, I will then reconstruct images using the generator.

Once trained, the model can then be used to assign an anomaly score. This score is twofold: (1) a reconstruction score, and (2) a feature matching score, as in Salimans et al.[13] If the score is over a specified threshold, the image will be classified as anomalous. The full anomaly detection scheme is presented in Figure 3. The first experiment will be trained on 96 CT scans and tested on 50 unseen CT scans.

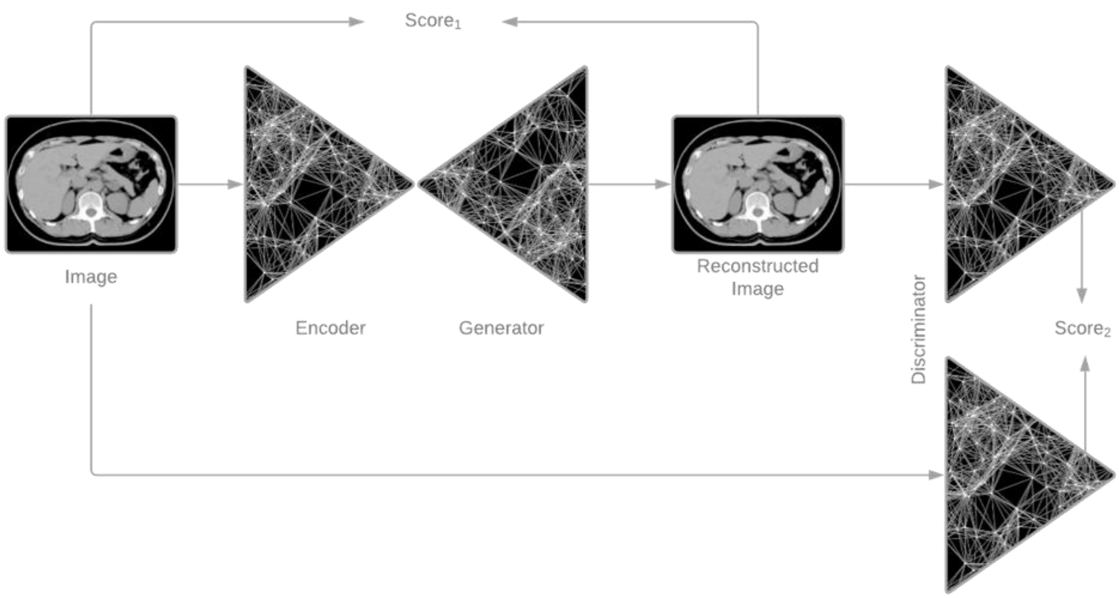

Figure 3: The anomaly detection scheme.

References

1. Cardenas, C. E., Yang, J., Anderson, B. M., Court, L. E., & Brock, K. B. (2019). Advances in Auto-Segmentation. Seminars in Radiation Oncology, 185-197.
2. Schlegl, T., Seeböck, P., Waldstein, S. M., Schmidt-Erfurth, U., & Langs, G. (2017). Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery. *Information Processing in Medical Imaging*. Boone, North Carolina.
3. Schlegl, T., Seeböck, P., Waldstein, S. M., Langs, G., & Schmidt-Erfurth, U. (2019). f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks. Medical Image Analysis, 30-44.
4. Han, C., Rundo, L., Murao, K., Nogucki, T., Shimahara, Y., Milacski, Z. A., . . . Satoh, S. (2020). MADGAN: unsupervised Medical Anomaly Detection GAN using multiple adjacent brain MRI slice reconstruction.
5. Zhou, K., Gao, S., Cheng, J., Gu, Z., Fu, H., & Tu, Z. (2020). Sparse-GAN: Sparsity-constrained generative adversarial network for anomaly detection in retinal OCT Image. International Symposium on Biomedical Imaging. Iowa City, Iowa.
6. Zhang, C., Wang, Y., Zhao, X., Guo, Y., Xie, G., Lv, C., & Lv, B. (2020). Memory-Augmented Anomaly Generative Adversarial Network for Retinal OCT Images Screening. International Symposium on Biomedical Imaging (pp. 1971-1974). Iowa City, Iowa: IEEE.
7. Wang, N., Chen, C., Xie, Y., & Ma, L. (2020). Brain Tumor Anomaly Detection via Latent Regularized Adversarial Network. CoRR abs/2007.04734.
8. Fijioka, T., Kubota, K., Mori, M., Kikuchi, Y., Katsuta, L., Kimura, M., . . . Tateishi, U. (2020). Efficient Anomaly Detection with Generative Adversarial Network for Breast Ultrasound Imaging. Diagnostics (Basel).
9. Han, C., Rundo, L., Murao, K., Ádám Milacski, Z., Umemoto, K., Sala, E., Nakayama, H., & Satoh, S. GAN-based Multiple Adjacent Brain MRI Slice Reconstruction for Unsupervised Alzheimer's Disease Diagnosis. Lecture Notes in Bioinformatics.
10. Singanamalli, A., Mitra, J., Wallace, K., Venugopal, P., Smith, S., Mo, L., Leung, L.Y., Morrison, J., Rasmussen, T., & Marinelli, L. (2020) Blood flow anomaly detection via generative adversarial networks: a preliminary study. Proc. SPIE 11315, Medical Imaging 2020: Image Guided Procedures, Robotic Interventions, and Modeling. Houston, Texas.
11. Budianto, T., Nakai, T., Imoto, K., Takimoto, T., & Haruki, K. (2020) Dual-encoder Bidirectional Generative Adversarial Networks for Anomaly Detection. CoRR abs/2012.11834.

12. Benson, S., & Beets-Tan, R. (2020) GAN-based anomaly detection in multi-modal MRI images. bioRxiv. Cold Spring Harbor Laboratory.
13. Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, Al, & Chen, Xi. (2016) Improved Techniques for Training GANs. Neural Information Processing Systems. Barcelona, Spain.
14. Nie, W., Karras, T., Garg, A., Debnath, S., Patney, A., Patel, A. B., & Anandkumar, A. (2020). Semi-Supervised StyleGAN for Disentanglement Learning. International Conference on Machine Learning. Vienna, Austria.