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

M. STAUFFER[1,2], A. PATEL[2-4], B. M. ANDERSON[1], B. RIGAUD[1], B. REBER[1], W. NIE[4], E. KOAY[5], and K. BROCK[1]

1. Imaging Physics Department, The University of Texas MD Anderson Cancer Center, Houston, Texas, United States
2. Computer Science Department, Rice University, Houston, Texas, United States
3. Electrical and Computer Engineering Department, Rice University, Houston, Texas, United States
4. Neuroscience Department, Baylor College of Medicine, Houston, Texas, United States
5. Gastrointestinal Radiation Oncology Department, The University of Texas MD Anderson Cancer Center, Houston, Texas, United States

INTRODUCTION

- Automatic image segmentation is crucial for automation in cancer detection, image-guided treatment, and response assessment.
- 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.

SIGNIFICANCE

- Challenges with deep learning-based segmentation models:
  - They require large, diverse, and labelled datasets which are difficult to collect in the medical field.
  - They do not generalize well to unseen anatomies (Figure 1).
- Our methods can be used to:
  - Detect images on which a trained segmentation model will likely fail.
  - Diversify medical imaging datasets.
  - Prioritize images for manual labeling.

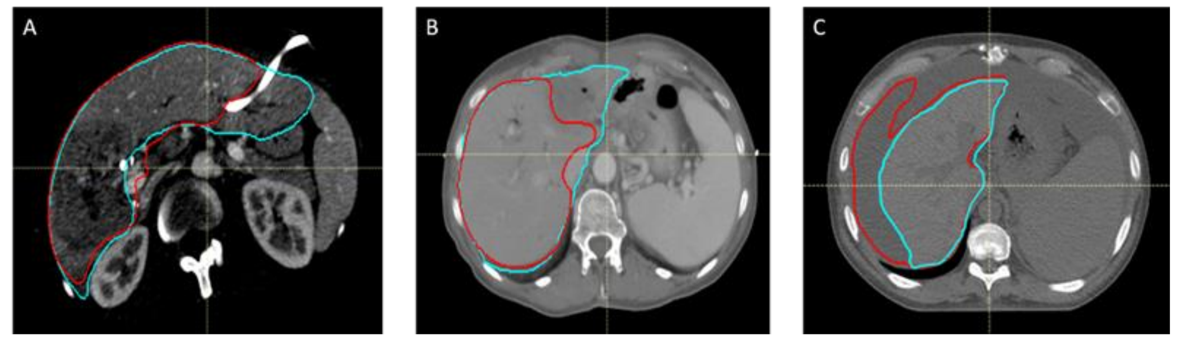

Figure 1: Instances where a trained liver segmentation model failed due to abnormalities not present in the training dataset: (A) a stent, (B) a tumor, (3) ascites. Red is the automated segmentation and light blue is the manual contour. The model had an average Dice coefficient of 0.96 on 50 unseen CT scans.

APPROACH

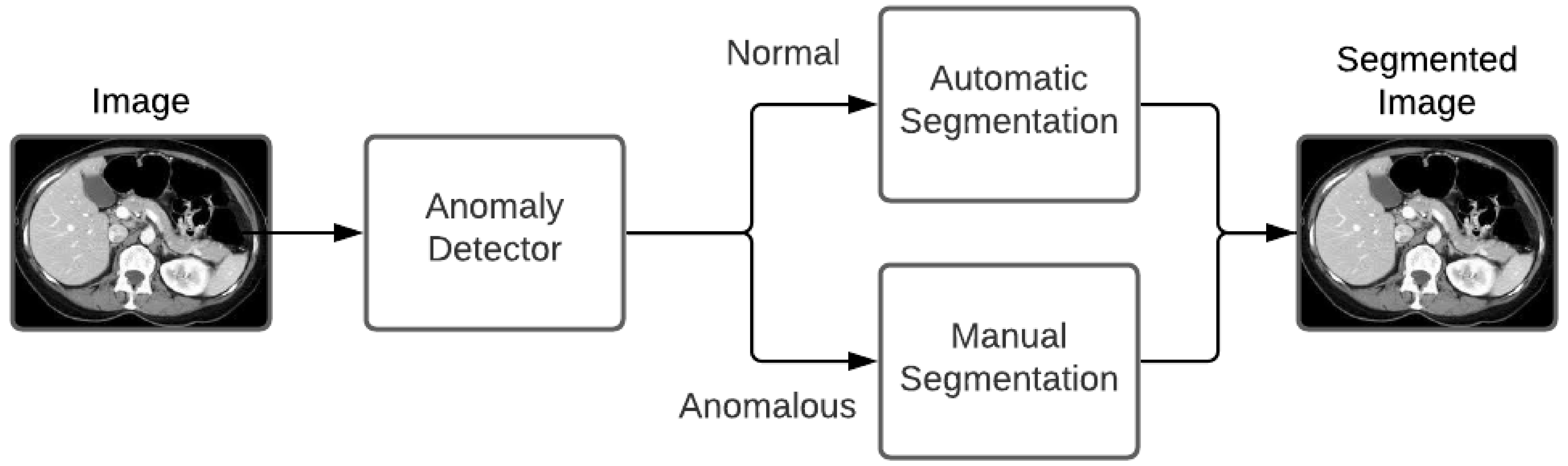

Figure 2: An active learning paradigm for the integration of automatic segmentation models into a clinical workflow. An image will first go through the anomaly detector. The detector will flag images that a trained automatic segmentation model will likely fail on. If the automatic model is likely to fail on a given image, the image will be segmented manually. Otherwise, it will be segmented automatically. The automatic segmentation model has already been trained, but it can be updated with images and their manual segmentations that were deemed to be "anomalous."

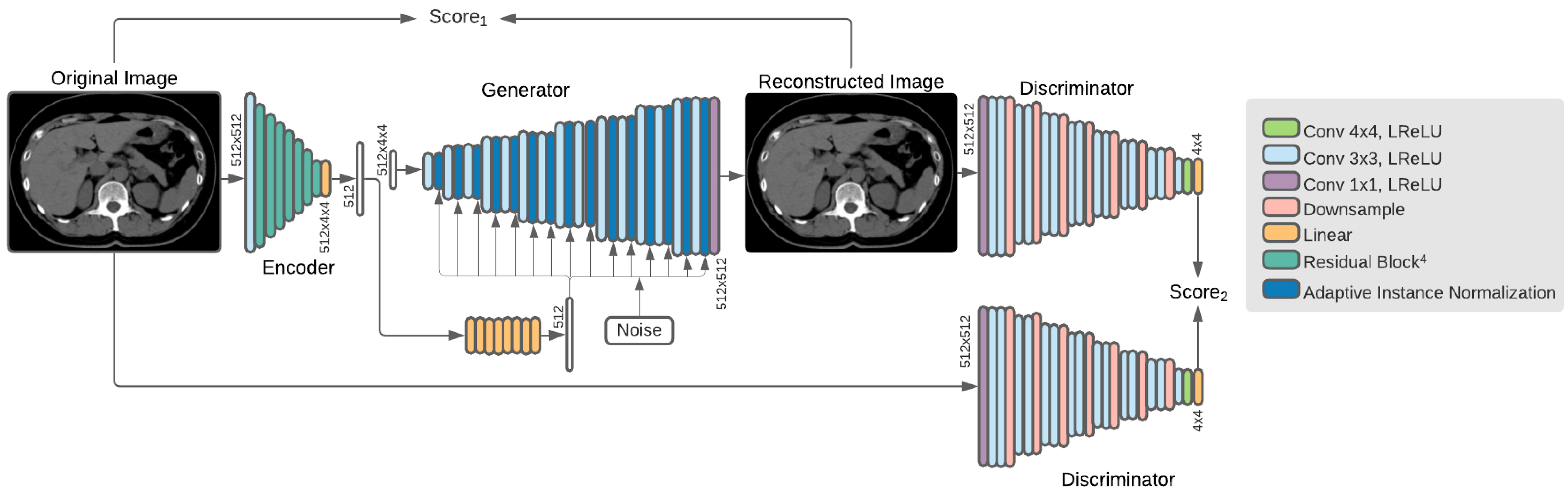

Figure 3: The anomaly detection scheme. The scheme consists of two neural network models, an encoder and a generative-adversarial network (GAN). In particular, we use a StyleGAN2[1] due to its performance on high-resolution images. We plan to incorporate the 3D locational information into the GAN by experimenting with both conditioning the GAN on the relative slice location and by making the GAN 3D. The image is first encoded into a vector space using the encoder. It is then reconstructed using the generator from the GAN. We use the generator instead of a standard decoder in an autoencoder architecture due to its superior generative capabilities. The first score is the reconstruction error between the original and reconstructed images. Then both the original and reconstructed images are put through the discriminator and their internal representations (the penultimate layer of the network) are extracted and compared for the second anomaly score, as in Salimans et al.[2] If the combination of the first and second anomaly scores is over a specified threshold, the image is classified as anomalous. We will train the detection model on 96 non-contrast enhanced abdominal CT images. We also have 50 "normal" and 3 "anomalous" (from Figure 1) unseen CT scans available for evaluation.

INNOVATION

- GAN-based anomaly detectors were introduced by Schlegl et al.[3,4]
  - They have since been used to identify diseases.[5-6]
- To the best of our knowledge, we are the first to use an anomaly detector for active learning.
- Current GAN methods on medical images generate slices (or patches of slices).
  - We plan to exploit the intricacies of medical images by incorporating the locational relationship between the slices.

INITIAL RESULTS

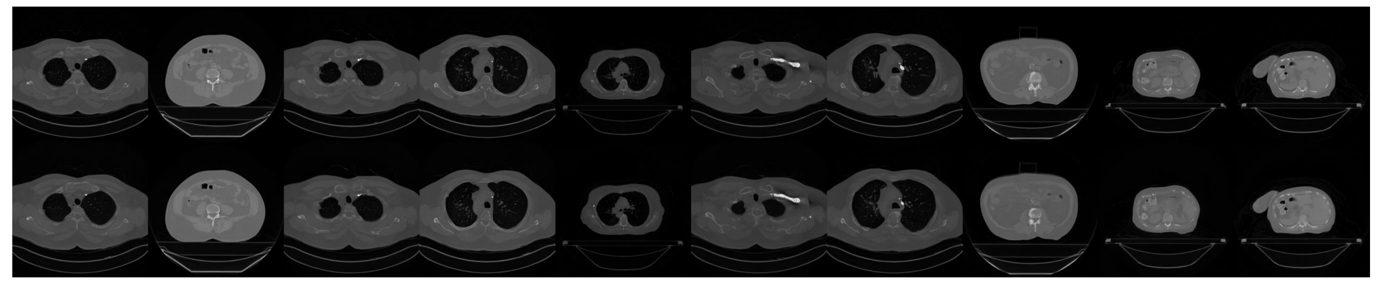

Figure 4: The top row contains original images. The bottom row contains their reconstructions.

ACKNOWLEDGEMENTS

Research supported by the Tumor Measurement Initiative at MD Anderson Cancer Center.

REFERENCES

1. Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2019) Analyzing and Improving the Image Quality of StyleGAN. arXiv preprint arXiv:1912.04958.
2. Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, Al, & Chen, Xi. (2016) Improved Techniques for Training GANs. Neural Information Processing Systems. Barcelona, Spain.
3. 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.
4. 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.
5. 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.
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.

mstauffer@mdanderson.org