Purpose: To detect images that do not belong to the training distribution of a deep learning-based segmentation model (out-of-distribution images).

Methods: Our liver dataset contained 3,234 abdominal computed tomography (CT) scans from 456 patients windowed with level 50 and width 350. All axial slices were mapped to 512x512 PNG images, resulting in 154,945 images. 1,000 randomly selected images (split patient-wise) were set aside for testing (in-distribution). An additional 1,000 (250 each) images of the brain, neck, lung, and cervix structures were used for testing (out-of-distribution). A StyleGAN2-ADA architecture, the state-of-the-art generative model for high resolution images, was trained to model the distribution of the liver images. Images were reconstructed using backpropagation on the model's input latent space. Reconstructions were evaluated using Wasserstein distance (WD). Out-of-distribution detection was evaluated with area under the receiver operating characteristic curve (AUC).

Results: Our model could distinguish between liver and non-liver CT with 95.59 AUC. The livers with high WDs contained anomalous objects, such as: metal artifacts, large tumors, and high amounts of contrast.

Conclusion: We defined an out-of-distribution detection paradigm by using a generative adversarial network to model in-distribution images. Using this paradigm, we were able to distinguish between liver and non-liver images.