Predicting Standardized Uptake Value of Brown Adipose Tissue from CT scans using convolutional neural networks

September 23, 2024

A Predicting the PET activity of BAT in other BAT depots

In this experiment, we trained the CNN using the original images that encompass the cervical, supraclavicular, and paraspinal regions, diverging from our initial approach of focusing solely on images cropped around the supraclavicular area, as described in Section 4 of the main paper. For ground truth segmentation, we identified brown adipose tissue (BAT) by the intersection of thresholded original CT and PET scans. Specifically, we applied Hounsfield Unit (HU) thresholds ranging from -180 to -10 for CT scans and a standardized uptake value (SUV) threshold of 1.5 for PET scans. The predicted segmentations were similarly derived by intersecting thresholded original CT scans with thresholded predicted PET scans. The predicted segmentations were similarly derived by intersecting thresholded original CT scans with the thresholded predicted PET scans. We quantified the similarity between the ground truth and predicted segmentations using the Dice score, which resulted in an average score of 0.519 across the test sets in a 5-fold cross-validation, whereas the Dice score for segmenting the supraclavicular region is 0.521, as shown in Table 1 in the main paper. These results demonstrate that CNNs are capable of predicting BAT activity across various BAT regions with consistent accuracy.

Additionally, we assessed the efficacy of the HU thresholding-based method for segmenting active brown adipose tissue (BAT) across various depots, including cervical, supraclavicular, and paraspinal regions. For this, we applied HU thresholds ranging from -180 to -10 for CT scans. It is important to note that we did not apply any standardized uptake value (SUV) threshold, as the method does not predict PET activations of BAT. The segmentations predicted by the HU thresholding-based method were compared with the ground truth using the Dice score, resulting in a notably low score of 0.098. This outcome highlights a significant disparity in the effectiveness of the HU thresholding-based method across different experiments. Specifically, the Dice score for predicting BAT activity in multiple regions is substantially lower than the score reported in Table 1 for the supraclavicular region alone in the main paper. This discrepancy arises because, in the main paper's experiments, we enhanced the accuracy of the HU thresholding method by intersecting its segmentations with manually segmented supraclavicular regions, effectively minimizing false positives. However, in our broader assessment of BAT activity across multiple regions, we did not apply manual segmentation masks to neither the CNN evaluations nor the HU thresholding method, due to the lack of manual segmentations for all examined areas.

B Figures

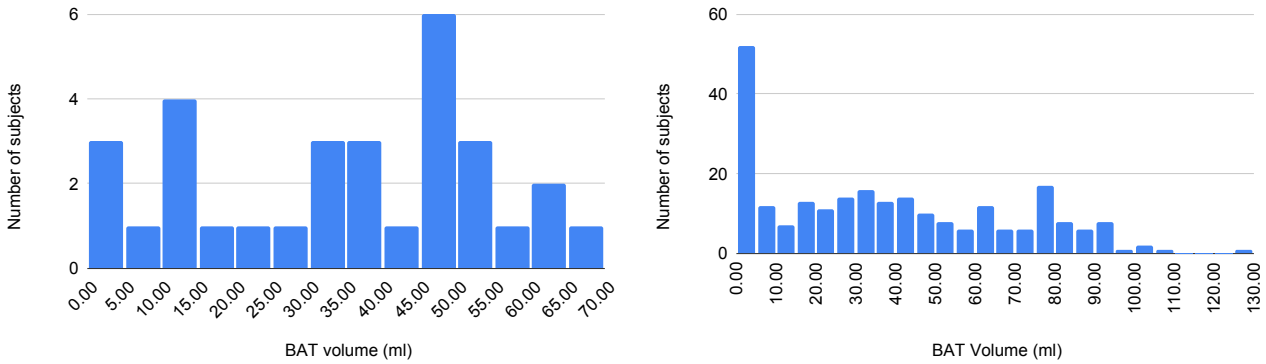

Figure 1: BAT volume histograms of (left) Basel and (right) Granada cohorts indicate that the majority of the subjects have high BAT activity measured by BAT volume due to cohort selection and cold exposure.