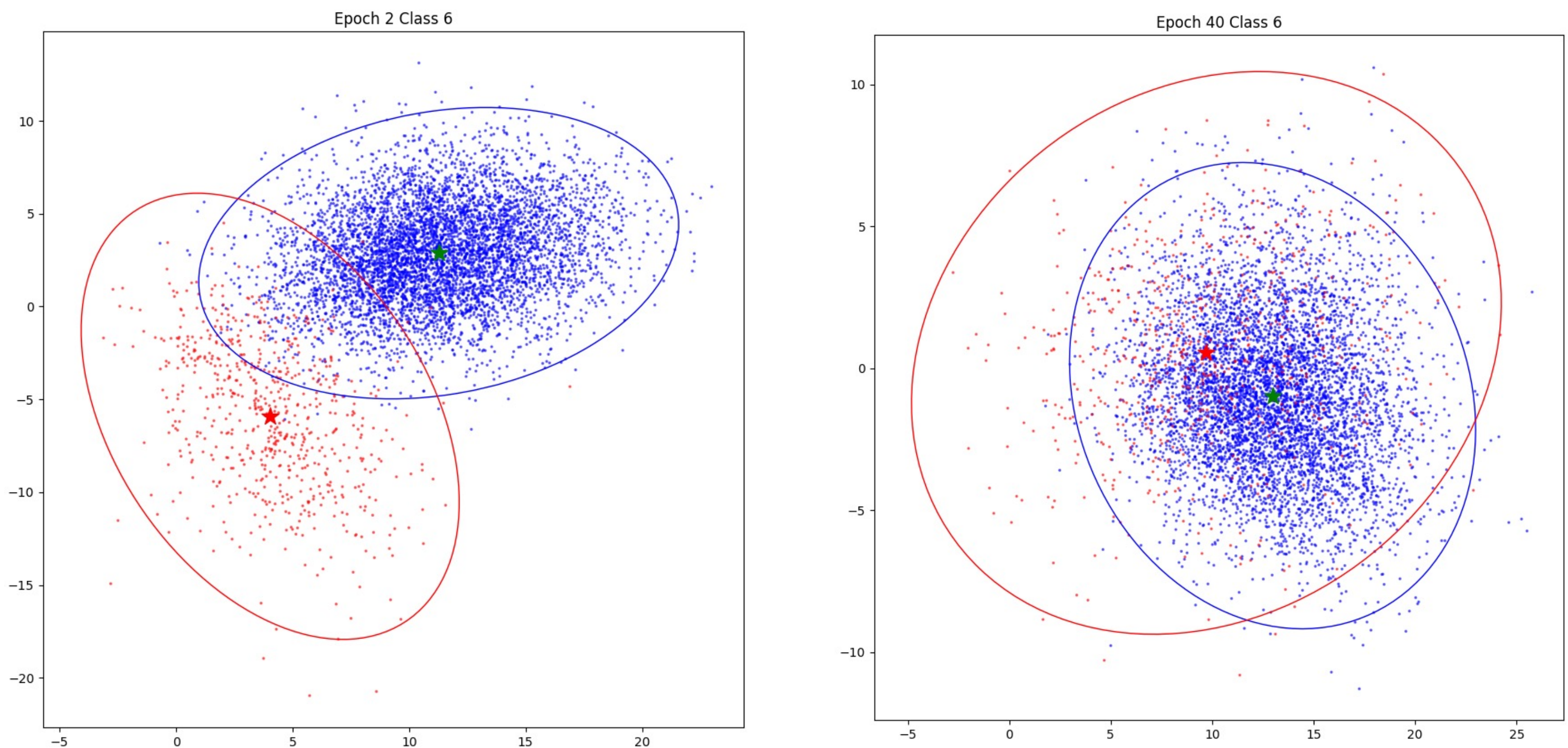


Subfigure 1: Features with no adaptation in Epoch 2 and Epoch 40.



Subfigure 2: Features with adaptation by CS divergence in Epoch 2 and Epoch 40.

Figure: Visualization of features of data samples from MNIST (source: blue) to USPS task (target: red). PCA is performed for dimension reduction. Distributions with no adaption (top) and adaptation by CS divergence (bottom) are shown. The visualization shows that even after adaption, the gap between two distributions are still no sufficient small.