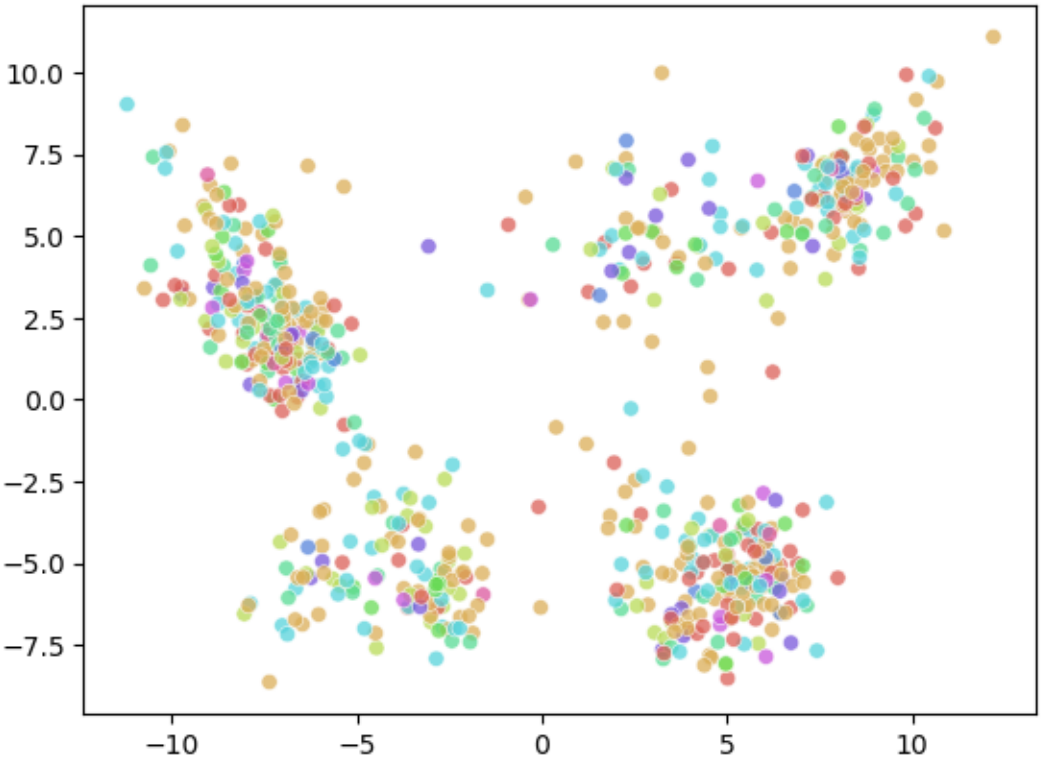

Figure 1: PCA on Dataset

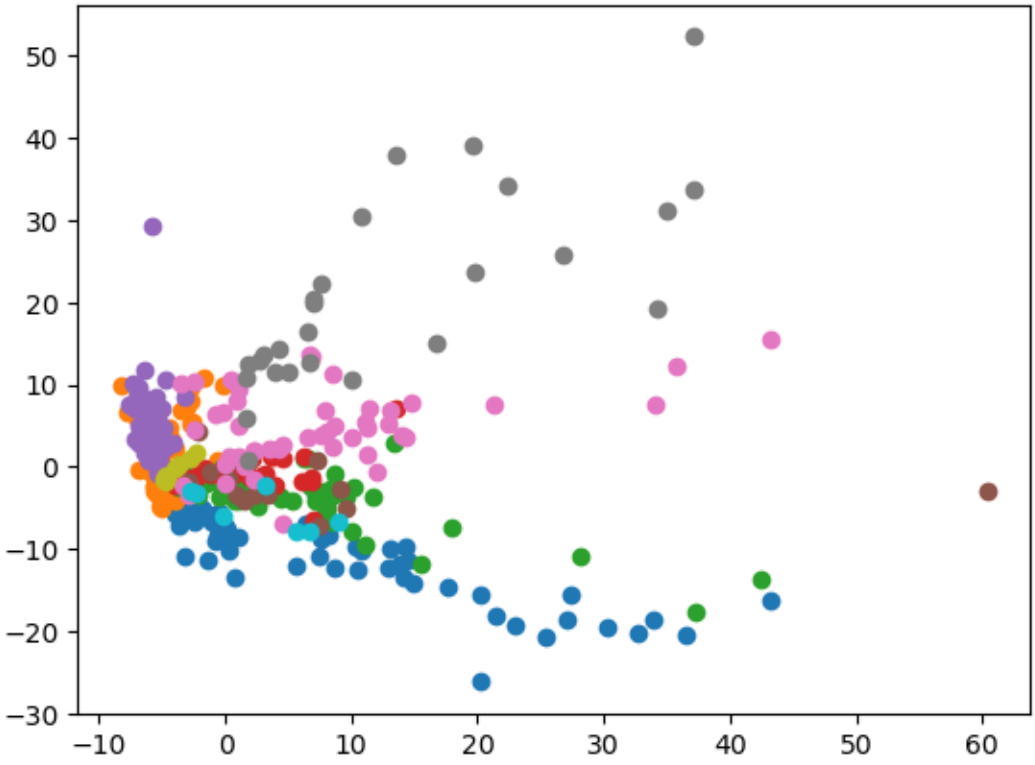

Figure 2: PCA on Encoded Train Set

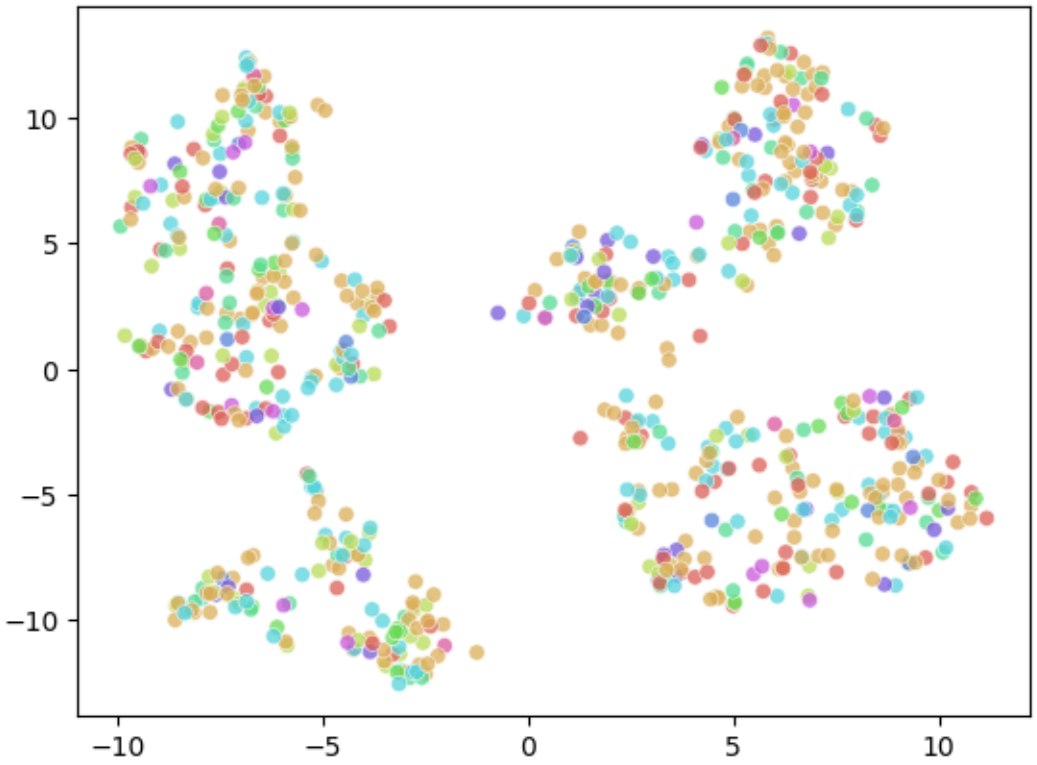

Figure 3: TSNE on Dataset

Three analyses were performed on the given dataset of cells and cell types: Principal Component Analysis (PCA) on the entire dataset prior to encoding (Figure 1), PCA on the train dataset post encoding (Figure 2), and t-distributed Stochastic Neighbor Embedding (t-SNE) on the entire dataset (Figure 3). The results indicated that the second analysis was the most accurate in clustering the different cell types together. Specifically, when observing the distribution of the color-coded cell types in Figures 1 and 3, the different cell types tended to be evenly spread throughout the image. However, in Figure 2, the different cell types were observed to be more closely clustered together, though not entirely. Although there was still considerable variation in the clusters, particularly among the grey, green, pink and blue cells, the cell types were still more discernible to the human eye compared to the other figures, in which each cell type was dispersed uniformly with other cell types.