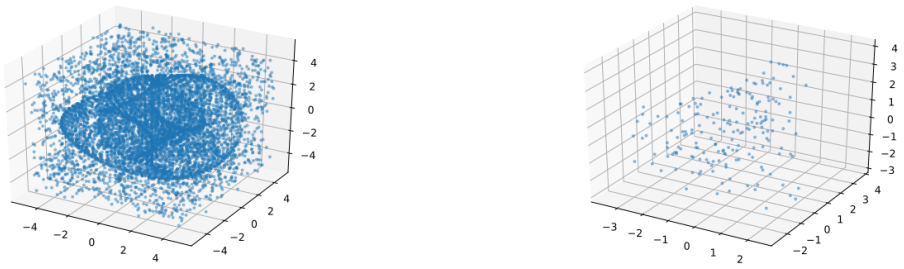


Figure 6: First two clusters of the layer 3 representations of the ReLU network.

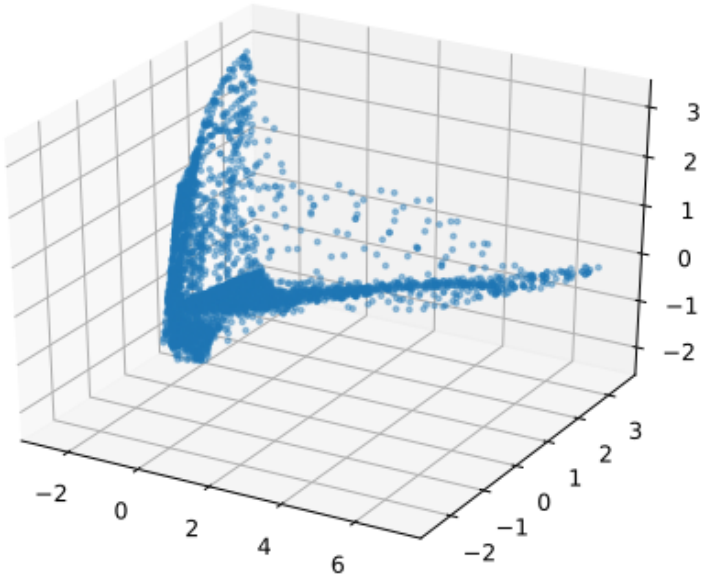
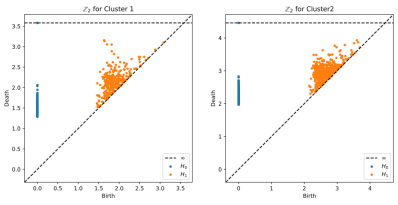
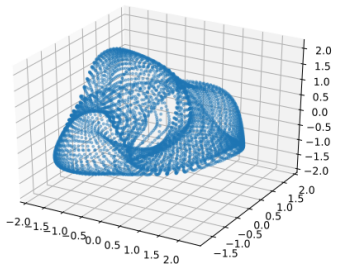


Figure 7: PCA projection of the layer 3 representation of the ReLU network.

Figure 8: First two clusters of the layer 1 representations of the Tanh network.



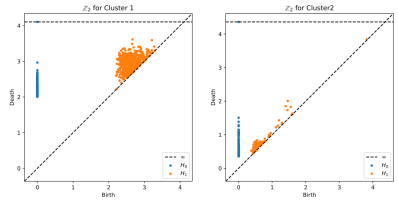
(a) Persistence diagrams of clusters.



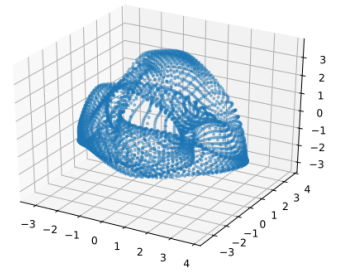
(b) PCA projection of torus.

Figure 9: Persistence diagrams of layer 1 clusters for the Tanh network and PCA projection.

Figure 10: First two clusters of the layer 2 representations of the Tanh network.



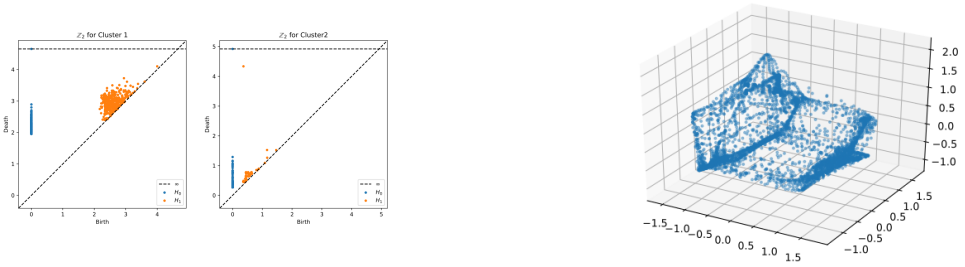
(a) Persistence diagrams of clusters.



(b) PCA projection of torus.

Figure 11: Persistence diagrams of layer 2 clusters for the Tanh network and PCA projection.

Figure 12: First two clusters of the layer 3 representations of the Tanh network.



(a) Persistence diagrams of clusters.

(b) PCA projection of torus.

Figure 13: Persistence diagrams of layer 3 clusters for the Tanh network and PCA projection.

5 Discussion and Further Work

The visual results from the experiments seem to indicate that the network is approximating homeomorphisms in early layers, before deeper layer representations obtain a topology markedly different than the original data. This could result from repeated application of non-homeomorphic layers making recovery of the underlying topology difficult. This appears to be correlated with inability of HDBSCAN to properly cluster the layer representations of deeper layers. However, we notice that while not as defined as the raw data, the persistence diagrams indicate the presence of topological features similar to the original data for the first two layers of both networks, albeit with a much shorter lifetime due to interference from the noise of the clusterings. This brings up the possibility that the network is attempting to isolate the relevant topological features for classification, and excising superfluous features. This would appear to explain the progressively sparser PCA projections obtained as the layers get deeper. The deep PCA projections seem to cluster points in more localized areas, leaving more empty portions in the space unlike the spread-out representations of earlier layers.

Furthermore, the PCA projections seem to indicate that the network approximates a continuous deformation for the first two layers of the Tanh network and first layer of the ReLU network, before the topology becomes unrecognizable. This could a result of ReLU being a surjective function, rather than bijective function like Tanh, making it harder for the layer to approximate a homeomorphism.

It is noted that nothing in this study is conclusive, and these are observations gathered from a preliminary experiment. Further work must be done in testing neural networks on more varied topologies and perhaps using alternative clustering algorithms as the noise included in these clusters appear to disrupt persistent homology.