The Latent space divided the different numbers into patches. This is expected as the embedding of two different images of a "4" should be very similar while the embedding of a "4" and a "7" should be different. This is especially the case as we hope to represent causal elements in our embedding.