Latent Space Topology of Several Machine Learning Algorithms

https://github.com/cchrisnguyen/LatentTopology

CHRIS HC NGUYEN AND JAMES M SHIHUA, THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

ABSTRACT

In this project, we study the latent spaces of different learning methods, including RPCA, VAE and BiGAN. We use the MNIST Fashion dataset.

DATASET

Fashion-MNIST is a dataset of Zalando's article images—consisting of a training set of 60,000 examples and a test set of 10,000 examples. Each example is a 28x28 grayscale image, associated with a label from 10 classes.

METHODOLOGY

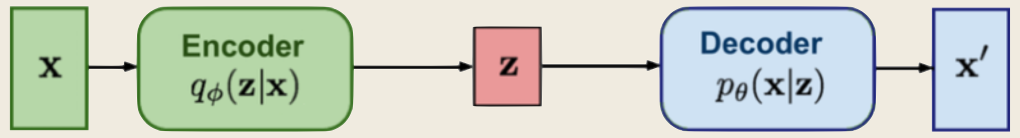
We would like to investigate the difference in the low-dimensional embedding of the original dataset, the low-rank matrix due to RPCA [1], and the latent space of a VAE and a GAN network.

We first conduct dimension reduction on the original dataset, by 11 methods: Random Projection, Truncated SVD, Linear Discriminant Analysis, Isomap, Standard LLE, Modified LLE, MDS, Random Trees, Spectral, t-SNE, and NCA.

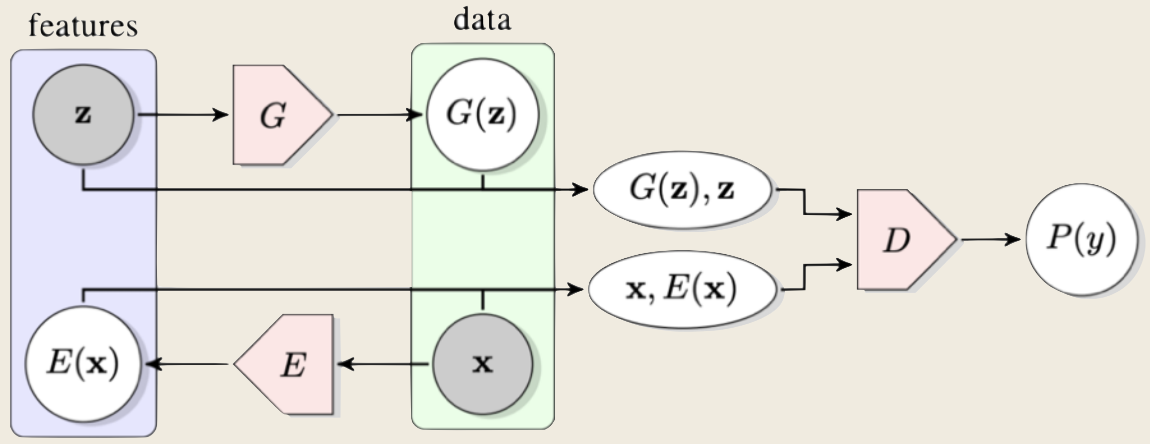
Then we perform Robust PCA to compute the rank required for the original dataset, which results in a rank of 432. This information indicates the minimum dimension required to preserve an adequate amount of information. We use the same 11 methods to study the low-rank matrix decomposed from RPCA. (Only effective ones are shown here.)

In the next step, we established two popular architectures for computer vision: VAE [2] and BiGAN [3], with the dimension of the latent space set to 433. Again, we use 11 methods to study the latent space of these two neural networks. (Only effective ones are shown here.)

VAE: maximize variational lower bound



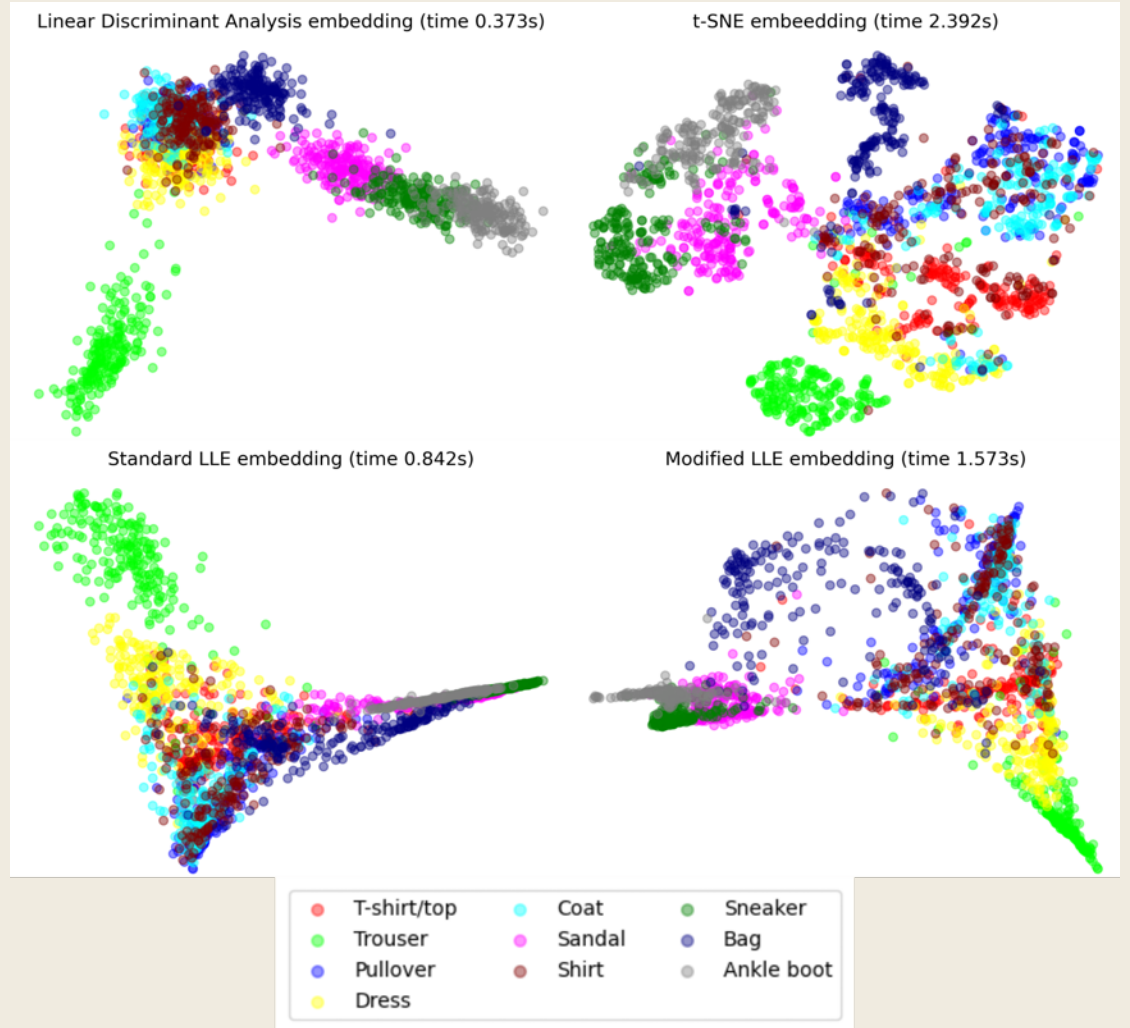
GAN: Adversarial training



[1] Candès, E. J., Li, X., Ma, Y., & Wright, J. (2011). *Robust principal component analysis?*. Journal of the ACM (JACM), 58(3), 1-37.
[2] Kingma, D. P., & Welling, M. (2013). *Auto-encoding variational bayes*. arXiv preprint arXiv:1312.6114.
[3] Zhirui Zhang, Shujie Liu, Mu Li, Ming Zhou, and Enhong Chen. 2018. *Bidirectional Generative Adversarial Networks for Neural Machine Translation*. In Proceedings of the 22nd Conference on Computational Natural Language Learning, pages 190–199, Brussels, Belgium.

RESULTS

Low Dimensional Representation of **Raw Data**



We notice that Linear Discriminant Analysis and t-SNE perform the best. Standard LLE and Modified LLE also can separate some datapoints into their respective categories. It can also be observed that trouser and dress categories have most distinct information embedded in the figures, while shirt and coat categories are relatively difficult to distinguish.

Low D Representation of Lo/Hi-Rank Matrix from **RPCA**

Low-Rank Matrix — High-Rank Matrix

RPCA converged in 34 iterations and reported a rank of 432. This rank will be later used as the dimension of the latent space of BiGAN. RPCA, reduced data dimension to slightly more than ½ of the original, their low-D topologies and information embedded remain largely the same. So does the computational time to conduct such embeddings. In addition, low-rank matrix can be viewed as the "layout" of a fashion category, while the high-rank one represents the "design".

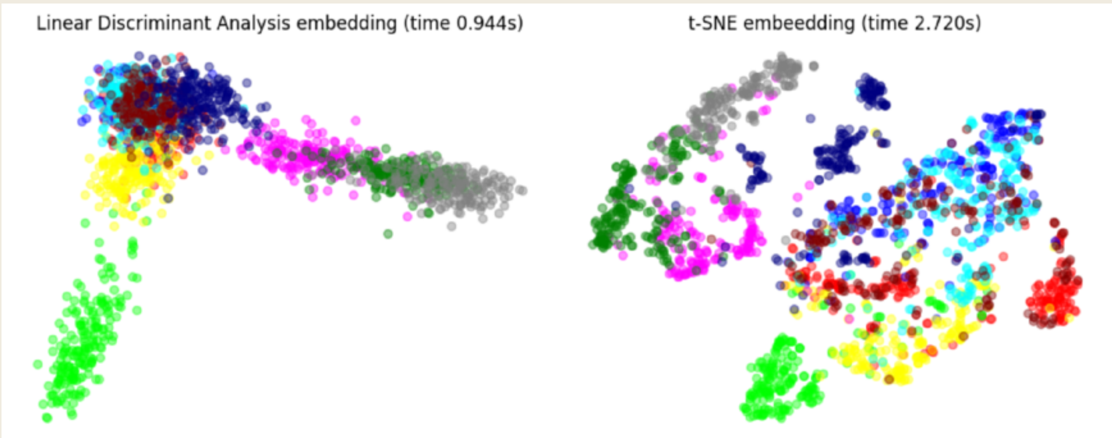
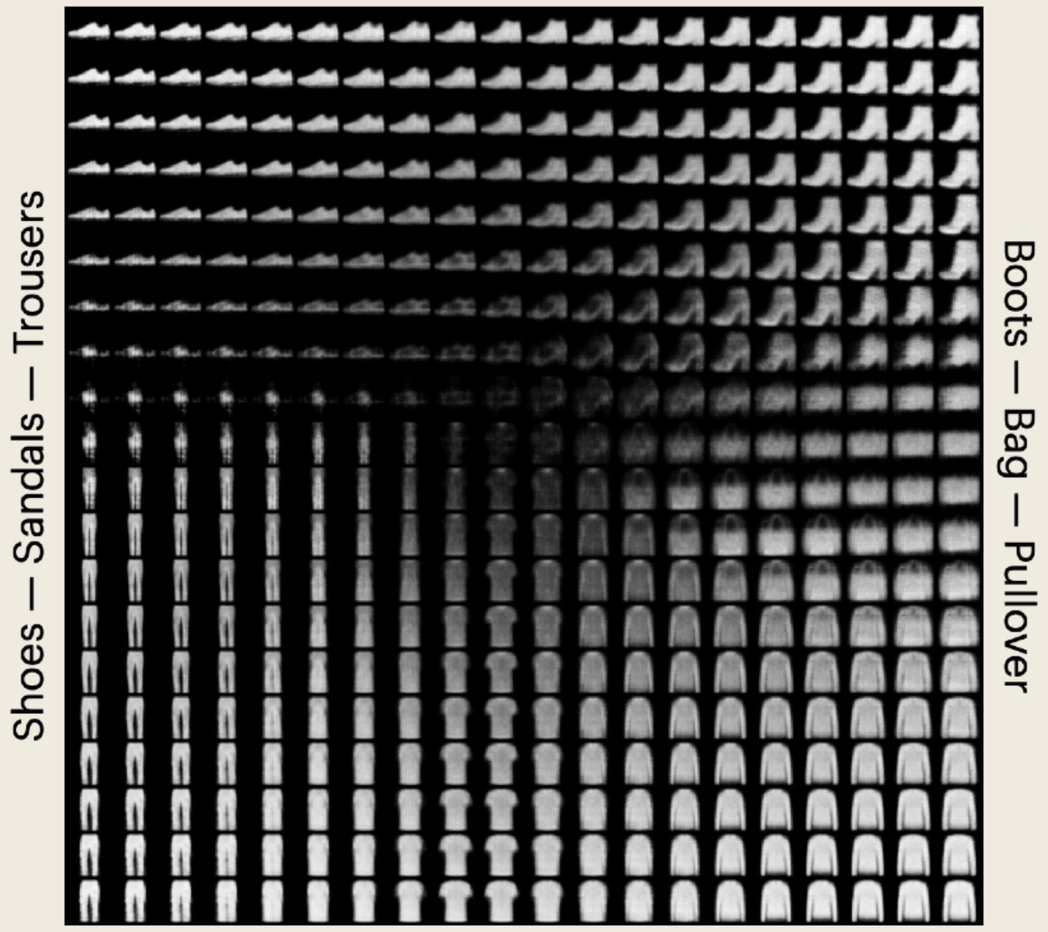


Image Generation and Latent Space of **VAE**

Shoes — Boots



Trousers — T-Shirts — Pullover

VAE itself can encode high-dimensional data into low dimensions. It has same good performance as the two non-ML algorithms. Latent space of VAE has a clear structure

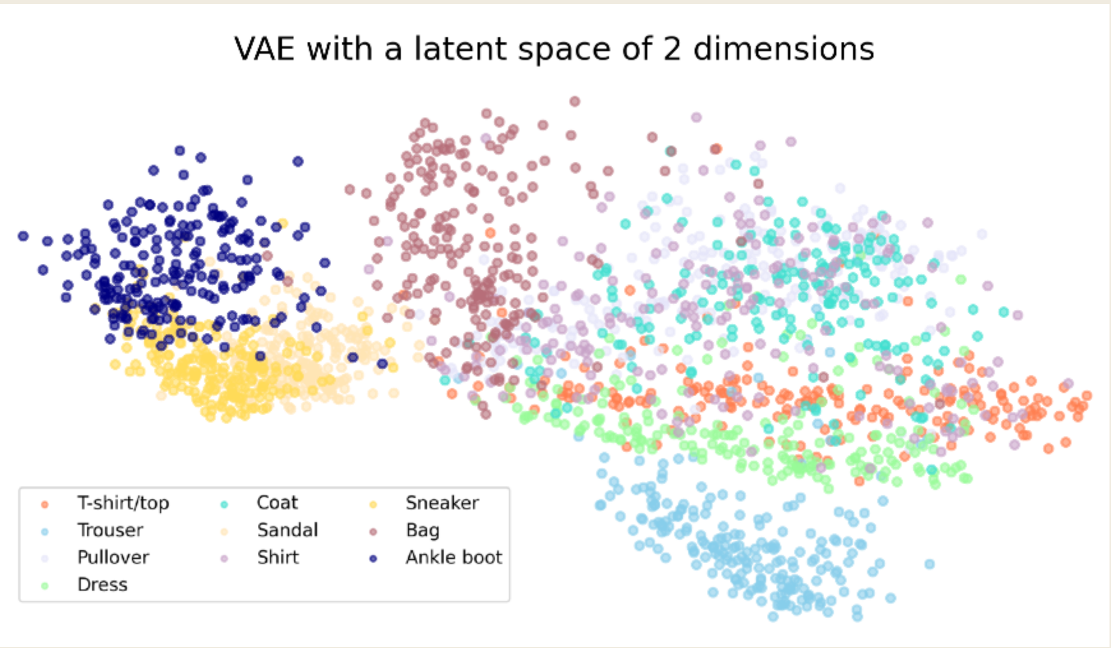


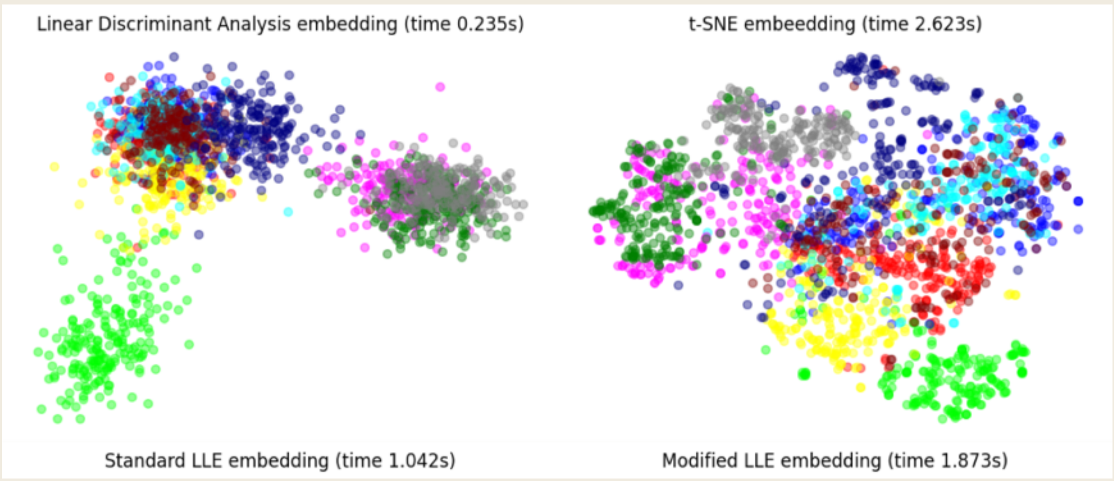
Image Generation and Latent Space of **BiGAN**

2D 1000 iters

2D 135000 iters

432D 1000 iters

432D 135000 iters



We clearly see fashion images generated by BiGAN with latent space of 432 is significantly more detail-rich and realistic. However, with only 2D latent space, we can also generate distinguishable images with enough training.

CONCLUSION

In this project, we explored the low-dimensional topology of the images in the Fashion-MNIST dataset. Linear Discriminant Analysis and t-SNE succeeded in clustering images with different categories in low-D space. We observed that trousers, dress, and pullover are among the most distinguishable. Next, we conducted RPCA on the raw image, to compute the minimum rank required to preserve necessary information. RPCA is also able to separates the "layout" of a fashionable and its "design". We use rank=432 outputted by RPCA as the dimension of the latent space of a VAE model. We see that VAE can learn the transition in pattern between categories and possesses a well-structured latent space. Similarly, we train a BiGAN model to generate fake images. However, BiGAN is good at generating images with more details, but do not have a well-organised latent space.