Final project proposal

Sulaimon Oyeleye* Alexander Zhiliakov†

March 2020

TITLE: AUTO-ENCODER INTERPOLATION

Overview

Autoencoders are known to be a powerful framework that helps to map datapoints by encoding all of the information needed to reconstruct the datapoints to a latent code from which the data can be recovered with minimal information loss. In most cases, the latent code is of lower dimension than the data itself, which indicates that autoencoders can perform some form of dimensionality reduction. For certain architectures, the latent codes have been shown to disentangle important factors of variation in the dataset which makes such models useful for representation learning.

In some cases, autoencoders can "interpolate": By decoding the convex combination of the latent codes for two datapoints, the autoencoder can produce an output which semantically mixes characteristics from the datapoints [2]. Intuitively, a good interpolation should decode to meaningful objects, give a gradual transformation, and reflect the internal structure of the dataset. More precisely, we require that the interpolation curve (after transporting to the input space) is smooth and relatively short. However, in some cases, even for a Gaussian prior, a linear interpolation could be of poor quality, e.g. due to the so called "soap bubble effect" [1, 3].

Objective

In this project, we will study the problem of generating a meaningful interpolation from a previously trained generative model. We claim that a good interpolation should both reveal the hidden structure of the dataset, as well as be smooth and follow the true data distribution, i.e. produce realistic elements. This mix-up regularizes the neural networks to favor simple linear behavior. In order to produce curves satisfying these conditions we define a realism index [5] of a path, which takes into account both density values and differences between consecutive decoded images to ensure smoothness.

We will take a specific autoencoder which is known as *adversarially constrained autoencoder interpolation* (ACAI). We will also provide the results one gets when this autoencoder is applied to different datasets, and show that ACAI can preserve latent space interpolation across categories, even though the discriminator is never exposed to such vectors [4].

Significance

This work will provide a useful representation by showing that it elicits improved representation learning performance on downstream tasks.

References

[1] Igor Sieradzki Damian Lesniak and Igor Podolak. Distribution interpolation trade off in generative models. in proceedings of international conference on learning representations. May 2019.

*Department of Mathematics, University of Houston, Houston, Texas 77204 (soyeleye@math.uh.edu).
†Department of Mathematics, University of Houston, Houston, Texas 77204 (alex@math.uh.edu).

[2] Aurko Roy Ian Goodfellow David Berthelot, Colin Raffel. Understanding and Improving Interpolation in Autoencoders via an Adversarial Regularizer. 23 July 2018.

[3] Ferenc Husar. Gaussian distributions are soap-bubbles. 18 March 2019.

[4] Eugene Lin Sreeram Kannan Sudipto Mukherjee, Himanshu Asnani. Clustergan : Latent space clustering in generative adversarial networks. 26 Jan 2019.

[5] Igor Podolak Aleksandra Nowak Krzysztof Maziarz Łukasz Struski, Jacek Tabor. Realism index: Interpolation in generative models with arbitrary prior a preprint. September 25, 2019.