Variational Capsule Encoder

Harish RaviPrakash, Syed Muhammad Anwar, Ulas Bagci Department of Computer Science, University of Central Florida, Orlando, Florida Email: harishr@knights.ucf.edu

Summary

This paper proposes a novel capsule network based variational encoder architecture, called Bayesian capsules (**B-Caps**), to modulate the mean and standard deviation of the sampling distribution in the latent space. We hypothesized that this approach could learn a better representation of features in the latent space than traditional approaches.

Training B-Caps

Data-driven and random normal sampling to avoid trivial solution during back-propagation from standard normal sampling



a) Standard normal sampling b)Random normal sampling c) Data-driven sampling

Quantitative Evaluation

Comparison of B-Caps and a basic Variational autoencoder (VAE) with increasing latent dimension



Latent Space Visualization

Separation of classes using B-Caps and baseline VAE





a) Baseline VAE latent space

b)B-Caps latent space

Architecture (Encoder and Decoder)



Qualitative Evaluation

Improvement in reconstruction of MNIST digits as a function of the latent variable length (L)



A comparison between reconstruction of Fashion-MNIST images using baseline VAE and B-Caps. Left to Right: Original image, Baseline VAE reconstruction and B-Caps reconstruction for L= 10



Center for Research in Computer Vision

This project is partially supported by NIH R01-CA246704and R01-CA240639, and Florida Department of Health20K04.