Variational Interpretable Deep Canonical Correlation Analysis

Lin Qiu[1] **Vernon M. Chinchilli**[2] **Lin Lin**[3]
[1] Genentech AI [2] The Pennsylvania State University [3] Duke University
lin.qiu.stats@gmail.com vchinchi@psu.edu l.lin@duke.edu

ICLR 2022

Abstract

The main idea of canonical correlation analysis (CCA) is to map different views onto a common latent space with maximum correlation. We propose a deep interpretable variational canonical correlation analysis (DICCA) for multi-view learning. The developed model extends the existing latent variable model for linear CCA to nonlinear models through the use of deep generative networks. DICCA is designed to disentangle both the shared and view-specific variations for multi-view data. To further make the model more interpretable, we place a sparsity-inducing prior on the latent weight with a structured variational autoencoder that is comprised of view-specific generators. Empirical results on real-world datasets show that our method is competitive across domains.

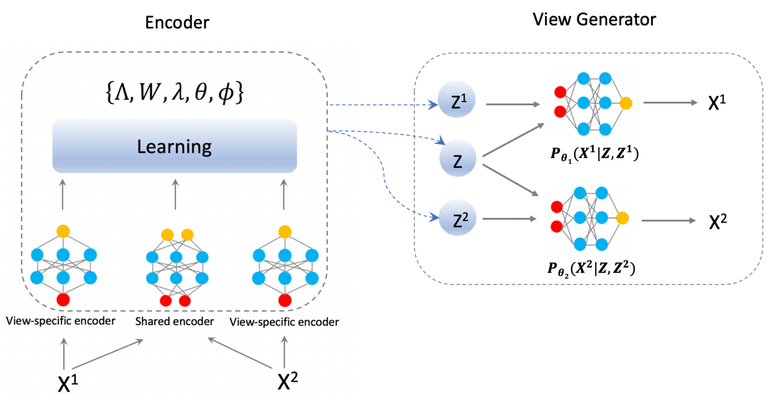
Model

$$\mathbf{X}^m \sim \mathcal{N}(f^{(m)}_{\theta_m}(\mathbf{\Lambda}^m\mathbf{Z} + \mathbf{W}^m\mathbf{Z}^m), \mathbf{\Psi}^m),$$

$$\mathbf{Z} \sim \mathcal{N}(\mathbf{0}_K, \mathbf{I}_K),$$
$$\mathbf{Z}^m \sim \mathcal{N}(\mathbf{0}_K, \mathbf{I}_K).$$

$$\gamma^2_{mj} \sim \text{Gamma}\left(\frac{d_m+1}{2}, \frac{\lambda^2}{2}\right),$$
$$\mathbf{\Lambda}^{(m)}_{\cdot,j}, \mathbf{W}^{(m)}_{\cdot,j} \sim \mathcal{N}(\mathbf{0}, \gamma^2_{mj}\mathbf{I}),$$

Model architecture



Collapsed objective function

We construct a collapsed variational objective function by marginalizing the γ^2_{mj}

$$\begin{aligned}
&\log p_\theta(\mathbf{x}) = \log \int p(\mathbf{x}|\mathbf{z}, \mathbf{z}^1, ..., \mathbf{z}^m, \mathcal{W}, \Lambda, \theta)p(\mathbf{z}) \\
&\times \prod_{m=1}^{M} p(\mathbf{z}^m)p(\mathcal{W}|\gamma^2) \times p(\Lambda|\gamma^2)p(\gamma^2)p(\theta)\, d\gamma^2\, dz ... dz^m \\
&\geq \sum_{m=1}^{M} E_{q_\phi(\mathbf{z}|\mathbf{x}^m), q_\phi(\mathbf{z}^m|\mathbf{x}^m)} \left[\log p_\theta(x^m|\mathbf{z}, \mathbf{z}^m, \mathcal{W}, \Lambda, \theta_m)\right] \\
&- D_{KL}(q_\phi(\mathbf{z}|\mathbf{x}^1, ..., \mathbf{x}^m)||p(\mathbf{z})) \\
&- \sum_{m=1}^{M} D_{KL}(q_\phi(\mathbf{z}^m|\mathbf{x}^m)||p(\mathbf{z}^m)) \\
&+ \sum_{m=1}^{M} \log p(\theta_m) - \lambda \sum_{m,j} ||\mathbf{\Lambda}^{(m)}_{\cdot,j}||_2 - \lambda \sum_{m,j} ||\mathbf{W}^{(m)}_{\cdot,j}||_2 \\
&= \mathcal{L}(\phi, \theta, \mathcal{W}, \Lambda).
\end{aligned}$$

Noisy MNIST Dataset Disentanglement learning

Two-view noisy MNIST datasets are widely used for testing multi-view models. This dataset is generated from the MNIST dataset, the first view of the dataset is generated by rotating each image at angles randomly sampled from uniform distribution, while the second view is from randomly sampled images with the same identity to the first view but not necessarily the same image.

Table 1: Reconstruction comparison on noisy two-view MNIST

Method	View 1 MSE (STD)	View 2 MSE (STD)
oi-VAE	0.059 (0.009)	0.172 (0.009)
DPCCA	0.052 (0.012)	0.134 (0.003)
VCCA	0.023 (0.011)	0.088 (0.0042)
VCCA-p	0.024 (0.011)	0.084 (0.005)
DICCA (Ours)	**0.016 (0.005)**	**0.080 (0.005)**

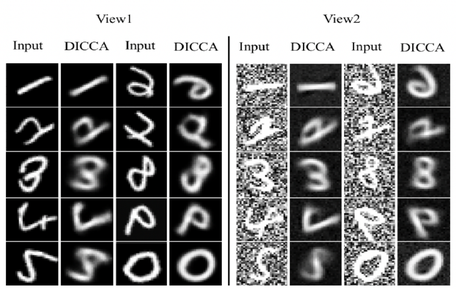
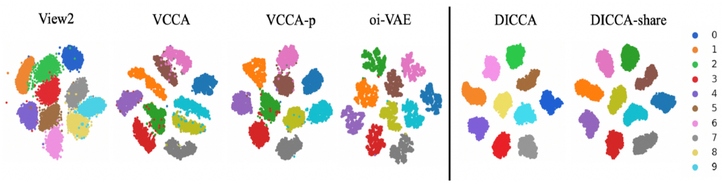
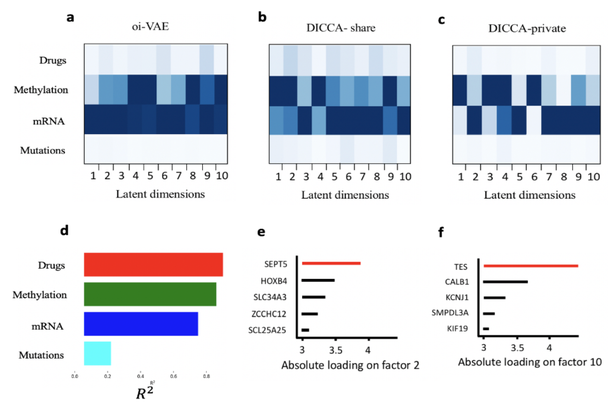


Figure 2: Reconstruction of images from the noisy MNIST test set by DICCA.



Genetic Data

This dataset contains 200 patients with chronic lymphocytic leukaemia (CLL}. This data combines vivo drug response measurements with somatic mutation status, transcriptome profiling and DNA methylation assays. Thus, there are four measurements on the same patients.



Group dependency relationship: each latent dimension of z influences only a sparse subset of the observational groups. We can view the observational groups associated with a specific latent dimension.
Latent dimension interpretation and biomarker discovery: Based on the top weights in mRNA data, factor 2 is aligned with SEPT5 which is a member of the septin gene family of nucleotide binding proteins. Disruption of septin function can disturb cytokinesis and result in large multinucleate or polyploid cells. Factor 10 is aligned with TES which maps to a common fragile site on chromosome 7q31.2 designated FRA7G. TES is a negative regulator of cell growth and may act as a tumour suppressor gene that is inactivated primarily by transcriptional silencing resulting from CpG island methylation.