Multi-View Feature Learning

Introduction

Multi-View learning deals with the problem of learning from multiple features sets of the same data. Multi-view learning is largely being researched because data can be represented using different feature sets or views. For example, in recommender systems a user's interests can be obtained from his search queries and his online profile in social media. A multimedia can be separately described by its video and audio signals. A web page can be described by the text in the document and by the anchor text attached to hyperlinks pointing to the page.

Multi-view feature learning consists of methods to learn features of the dataset where we have access to multiple views of data. Methods developed for multi-view feature learning include techniques that involve autoencoders with a reconstruction or correlation based objectives. In this project I implemented Deep canonically correlated autoencoders(DCCAE) method which is a multi-view feature learning method to learn features of the MNIST dataset.

Deep Canonically Correlated Autoencoders (DCCAE)

Below is the Figure(Fig 1) of DCCAE method. In the below figure X and Y are two views which are inputs to two autoencoders. U and V are the two views of the learned features. Reconstructed x and reconstructed Y are the reconstructed views of X and Y.

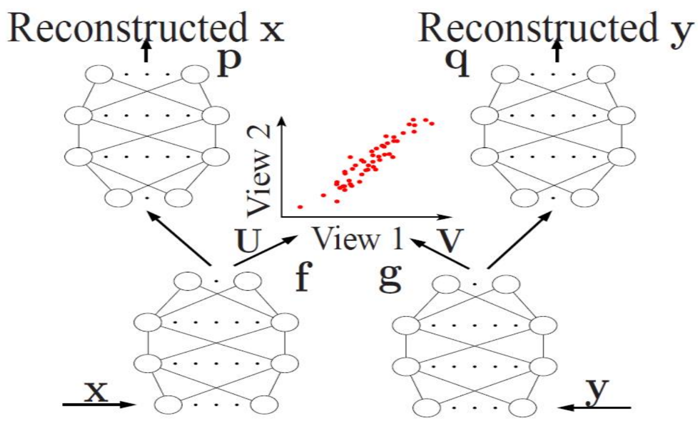


Fig 1 : Deep canonically correlated autoencoders

DCCAE optimizes the combination of canonical correlation between the learned features(U,V) and reconstruction errors of the autoencoders as mentioned in the below equation:

$$\min_{\mathbf{W_f},\mathbf{W}_g,\mathbf{W_p},\mathbf{W_q},\mathbf{U},\mathbf{V}} -\frac{1}{N}\operatorname{tr}\left(\mathbf{U}^\top \mathbf{f}(\mathbf{X})\mathbf{g}(\mathbf{Y})^\top \mathbf{V}\right) + \frac{\lambda}{N}\sum_{i=1}^{N}\left(\|\mathbf{x}_i - \mathbf{p}(\mathbf{f}(\mathbf{x}_i))\|^2 + \|\mathbf{y}_i - \mathbf{q}(\mathbf{g}(\mathbf{y}_i))\|^2\right)$$

According to the above equation DCCAE minimizes the sum of the negative correlation between learned features and mean squared error between original and reconstructed images.

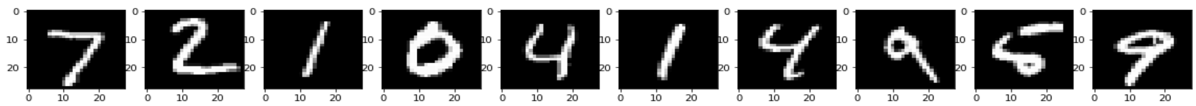
Experiment

I generated two views from MNIST dataset. First view is the MNIST digits dataset itself and the second view is noisy MNIST digits dataset which I got by adding random noise to each pixel of every image in the dataset. I implemented autoencoder using convolution neural networks using Keras package in python. For encoder, the first layer is the input layer, second is the convolution layer, third is maxpooling layer for down-sampling, fourth is the convolution layer and then fifth is the maxpooling layer. For decoder I created layers to reconstruct the image using up-sampling. Weights are adjusted for the network after checking the loss function of each image. Below is the autoencoder model:

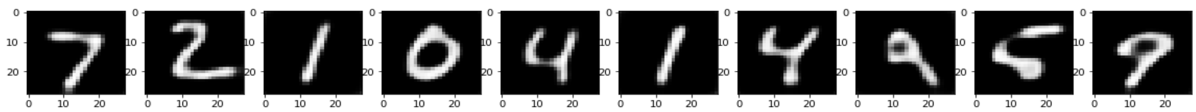
Layer (type)	Output Shape	Param #
input_18 (InputLayer)	(None, 28, 28, 1)	0
conv2d_103 (Conv2D)	(None, 28, 28, 24)	240
max_pooling2d_35 (MaxPooling	(None, 14, 14, 24)	0
conv2d_104 (Conv2D)	(None, 14, 14, 48)	10416
max_pooling2d_36 (MaxPooling	(None, 7, 7, 48)	0
conv2d_105 (Conv2D)	(None, 7, 7, 96)	41568
conv2d_106 (Conv2D)	(None, 7, 7, 96)	83040
up_sampling2d_35 (UpSampling	(None, 14, 14, 96)	0
conv2d_107 (Conv2D)	(None, 14, 14, 48)	41520
up_sampling2d_36 (UpSampling	(None, 28, 28, 48)	0
conv2d_108 (Conv2D)	(None, 28, 28, 1)	433

Results

Below are the Input images:



Below are the reconstructed images:



Input images are the MNIST digits and reconstructed images are the output of the autoencoders using convolutional neural networks. Training loss is 0.1120 which is the loss occurred after reconstructing the images for training data. Training accuracy is 0.8088 which is accuracy obtained for classifying training digits. Testing loss is 0.1140 which is loss occurred after reconstructing the images for testing data. Testing accuracy is 0.8072 which is the accuracy obtained for classifying testing digits.

References

Wang, Weiran, et al. "On deep multi-view representation learning." *International Conference on Machine Learning*. 2015.

Xu, Chang, Dacheng Tao, and Chao Xu. "A survey on multi-view learning." *arXiv preprint arXiv:1304.5634* (2013).

https://keras.io/