Predictive Encoding of Contextual Relationships using multi-layer neural networks

Project Web Page: http://www.andrew.cmu.edu/user/bojianh/

Project Description:
This project will be done in Spring 2016 under Dr. Tai Sing Lee from Center for the Neural Basis of Cognition (CNBC) in CMU. This project is meant as an extension to the previous research project conducted by Zhao [1]. In the previous project, Zhao and others developed a neural network model with a single contextual layer to achieve the effects of interpolation and extrapolation of video frames. In this project, my advisor and I will be exploring the suitability of alternative neural network models to tackle the same task and compare the output from the different networks. Additionally, we will also be conducting research to build multi-layered neural network from different neural network sub-layers in attempt to achieve more sophisticated prediction of motion in addition to simple linear translation and rotation in prediction.

The research findings from this project will further refine the approach to utilize videos as a viable data source for unsupervised training instead of labelled static images. The research findings will also provide useful insights in the overall neural network architecture in our brain regarding visual cortex and confirm the relationships between visual cortex and other sectors of the brain.

The resultant neural model can also be highly useful for the development of various new technologies such as a more efficient version of videos encoding. This technology can also be highly useful for other data sources for filling in missing sound in sound tracks or filling in blanks in text.

Project Goals
In this project, we would hope to extend the single contextual layer for prediction to multi-layer model that can predict more sophisticated motions such as accelerations in motion instead of simple linear motion and translation. This project is successful if a multi-layer neural network can be effectively trained to interpolate and extrapolate higher-level motions.

The 100% goal means that we can effectively train a multi-layer neural network consisting of convolutional neural network at lower levels and recurrent neural network (or other alternative neural network that support temporal data) at higher levels. In this goal, we will find the optimum parameters for the various models such as the number of nodes and layers of each sub-network as well as the types of training algorithm for this network.

The 75% goal means that we will be to construct a multi-layer network that can be trained on a subset of the whole dataset. This could happen if there are significant computational limitations in training the neural network that could not resolved easily given the technologies available today.

The 125% goal means that we will attempt to train this model of neural network on other types of data such as sound, text instead of just images. We will also be cross-comparing different multi-layered networks comprised of different sub-networks to figure out the best combination to solve this task.
Milestones

1st Technical Milestone: Install the various software necessary for the various types of neural network models such as convolutional neural network (CNN), recurrent neural network (RNN), long short-term memory neural network (LSTM) as well as deep boltzmann machines (DBM).

January 25th: Replicate the existing research results of the previous project using the single contextual layer network present in paper by Zhao [1].

February 8th: Compare different neural network models such as RNN, LSTM and DBM on the same prediction task of using single contextual layer and recording the results.

February 22nd: Identify new potential datasets and convert the raw data from the datasets into training and testing data that can be utilized by the neural network models.

March 14th: Design and implement multi-layered neural network architecture in Theano.

March 28th: Train the multi-layered neural network using the refined datasets and at the same time fine tuning various parameters of the neural network.

April 11th: Further improve the neural network model by refactoring the code so as to improve computational efficiency in training.

April 25th: Design and develop several multi-layered neural network architectures using different combination of sub-networks and cross compare the results from training.

Literature Search

Currently, I have read the paper by Zhao [1] which is the predecessor of this research project. Additionally, I have also read the paper by Wang [2] which discussed new approaches to train CNN using videos. However, I have yet to read papers that discusses the usage LSTM networks or DBM in training temporal data.

Resources needed

The primary software I need to conduct my study is Python Numpy and Theano libraries. These libraries are necessary to implement various neural network models and apply backpropagation algorithm to train the network. Currently, I have yet to install these software on my machine yet but these software are easily available for download. Additionally, I may also need to install other neural network software just as AlexNet for convolutional networks.

Bibliography
