Final project for EC720: **ICPR 2010 Contest on Semantic Description of Human Activities**

the data base and the task is provided at:

**[http://cvrc.ece.utexas.edu/SDHA2010/Human\_Interaction.html#Data](http://cvrc.ece.utexas.edu/SDHA2010/Human_Interaction.html" \l "Data)**

The challenge require us to recognize ongoing human activities from continuous videos.

The thesis [**Large-scale Video Classification with Convolutional Neural Networks**](https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/42455.pdf)written by Andrej Karpathy provided me the inspiration of how to design a CNN network to do the behavior of a frame by aquire the information of not only one frame but also frames in a time span.

The challenging part of this project is:

1. How to design a CNN network to receive a converge and effective result. Intuitivly, we may think that frames in time span would provide a more confident predict result. Thus, I want to make a comparation between single frame CNN and sequence frame CNN.

2. How to do the pre processing of the video using motion detection method. I will try both continuous frames subtraction, background subtraction, and no preprocessing video as the input to find if the motion detection method could provide a better result.