Maria Eugenia Nystrom
STAT 517: Statistical Learning and Predictive Modeling
Professor Stephen Lee
Literature Review

Applications of Convolutional Neural Networks

A Convolutional Neural Network (CNN) is a feed forward neural network that has many applications in analyzing images. Like a neural network, a Convolutional Neural Network is a computer system modeled on the human brain and nervous system. The main difference between neural networks and Convolutional Neural Networks is that CNN makes the assumption that the inputs are images. This allows for the architecture of a CNN to be different which makes it easier to implement.

Convolutional Neural Networks were first discovered when scientists were looking at cells in the visual cortex. The scientist recognized that the cells in the visual cortex are very sensitive. This proposed the idea of neocognitron. Neocognitron is a hierarchical multilayered neural network. It is trained with backpropagation algorithm that made it possible to recognize pattern images from raw pixels. Because of its excellent abilities for pattern recognition, neocognitron served as a inspiration for CNN. Unfortunately, there was a lack of large training data and it was not successful at the time. Disadvantages because of this lead to the failure to recognize video pattern images.

Since the development of AlexNet, Convolutional Neural Networks have been able to gain a lot of recognition for their strong ability to analyze images. These days Convolutional Neural Networks have many applications. These applications include: face recognition, scene labeling, image classification, action recognition, human pose estimation, and document analysis.

CCNs consist of trainable multistage architecture with each stage consisting many layers. In order to fully understand how CNN's work it is important to understand how the layers of CNN work. The layers include: the convolutional layer, non-linearity layer, pooling layer, and the fully connected layer. Together

these layers help determine important features. The convolutional layer is the first layer to extract images from the input image. It takes in two inputs the image matrix, and the kernel. It is a linear operation. The convolutional layer also preserves the relationship of pixels. After the convolutional layer is the nonlinearity layer. The nonlinearity layer (also known as ReLu) aims to replace all the negative values in the feature map with zero. This layer introduces nonlinearity which is important for multi-layer networks. The pooling layer reduces the spatial size, amount of parameters, and computation within the network. It is one way to manage overfitting. The last layer is the fully connected layer of the network. This layer creates global semantic information by taking the output neurons from the previous layers and connecting them to the output neurons in the current layer. It connects them based on which features most correlate to a certain class.