



Hierarchical Deep Learning for Computer Vision

Matthew Zeiler
Clarifai, USA

Abstract

This lecture will cover the application of deep learning to computer vision using both unsupervised and supervised algorithms. This lecture will start by covering an unsupervised learning technique for learning features directly from images in a hierarchical fashion. This method, called deconvolutional networks, utilizes the same operations as a traditional convolutional networks and combines them with a sparse coding objective to learn without requiring labels. It becomes evident in training these models that spatial information captured in pooling layers is critical, which exposes some new techniques for pooling spatial information that prove useful for both deconvolutional and convolutional networks. Also, by leveraging deconvolutional networks as a diagnostic tool, visualizations of convolutional networks are shown which provide a better understanding of the deep learning approaches. These visualizations exposes a natural progression of complexity of features throughout the layers from simple edges and colors, to mid-level primitives, to object parts and finally complete object representations in higher layers. Leveraging these visualizations to understand the downsides of other models, we show performance gains on Imagenet. Finally, an optimization technique called ADADELTA is explained as a method to simplify the training of neural networks.

Keywords

Hierarchical Deep Learning, deconvolutional networks, ADADELTA