

Master Thesis - Ngoc Quan Pham  
Convolutional Neural Network Language Models

ABSTRACT

Convolutional Neural Networks (CNNs) have shown to yield very strong results in several Computer Vision tasks. Their application to language has received much less attention, and it has mainly focused on static classification tasks, such as sentence classification for Sentiment Analysis or relation extraction. In this work, we study the application of CNNs to language modeling, a dynamic, sequential prediction task that needs models to capture local as well as long-range dependency information. Our contribution is twofold. First, we show that CNNs achieve 11-26% better absolute performance than feed-forward neural language models, demonstrating their potential for language representation even in sequential tasks. As for recurrent models, our model outperforms RNNs but is below state of the art LSTM models. Second, we gain some understanding of the behavior of the model, showing that CNNs in language act as feature detectors at a high level of abstraction, like in Computer Vision, and that the model can profitably use information from as far as 16 words before the target.