

# Parsing Natural Scenes and Natural Language with Recursive Neural Networks

Richard Socher  
Cliff Chiung-Yu Lin  
Andrew Y. Ng  
Christopher D. Manning  
Computer Science Department, Stanford University

## Introduction

This paper uses a Recurrent Neural Networks to find a recursive structure is commonly found in the inputs of different modalities such as natural scene images or natural language sentences. This is the first deep learning work which learns full scene segmentation, annotation and classification. The same algorithm can be used both to provide a competitive syntactic parser for natural language sentences from the Penn Treebank and to outperform alternative approaches for semantic scene segmentation, annotation and classification.

This particular approach for NLP is different in that it handles variable sized sentences in a natural way and captures the recursive nature of natural language. Furthermore, it jointly learns parsing decisions, categories for each phrase and phrase feature embeddings which capture the semantics of their constituents

## Input Representation

The images are divided in segments by scenes as described by a paper by Gould et al., (2009)  
Then the feature representations are found by training the RNN using a the sigmoid as the activation function.

For sentences the bag of words model is used where we map the words to a vector made from a fixed size vocabulary.