Progress Report

Kuei-Yueh Ko

Current Progress

- Read the code of CellCnn and had discussed the code with Scott
- Try to implement neural network from scratch
- Finding papers about graph convolutional network.

Bruna et al 2013
Spectral Networks and Locally Connected Networks on Graphs

Spectral Networks and Deep Locally Connected Networks on Graphs

Joan Bruna
New York University
bruna@cims.nyu.edu

Wojciech Zaremba
New York University
woj.zaremba@gmail.com

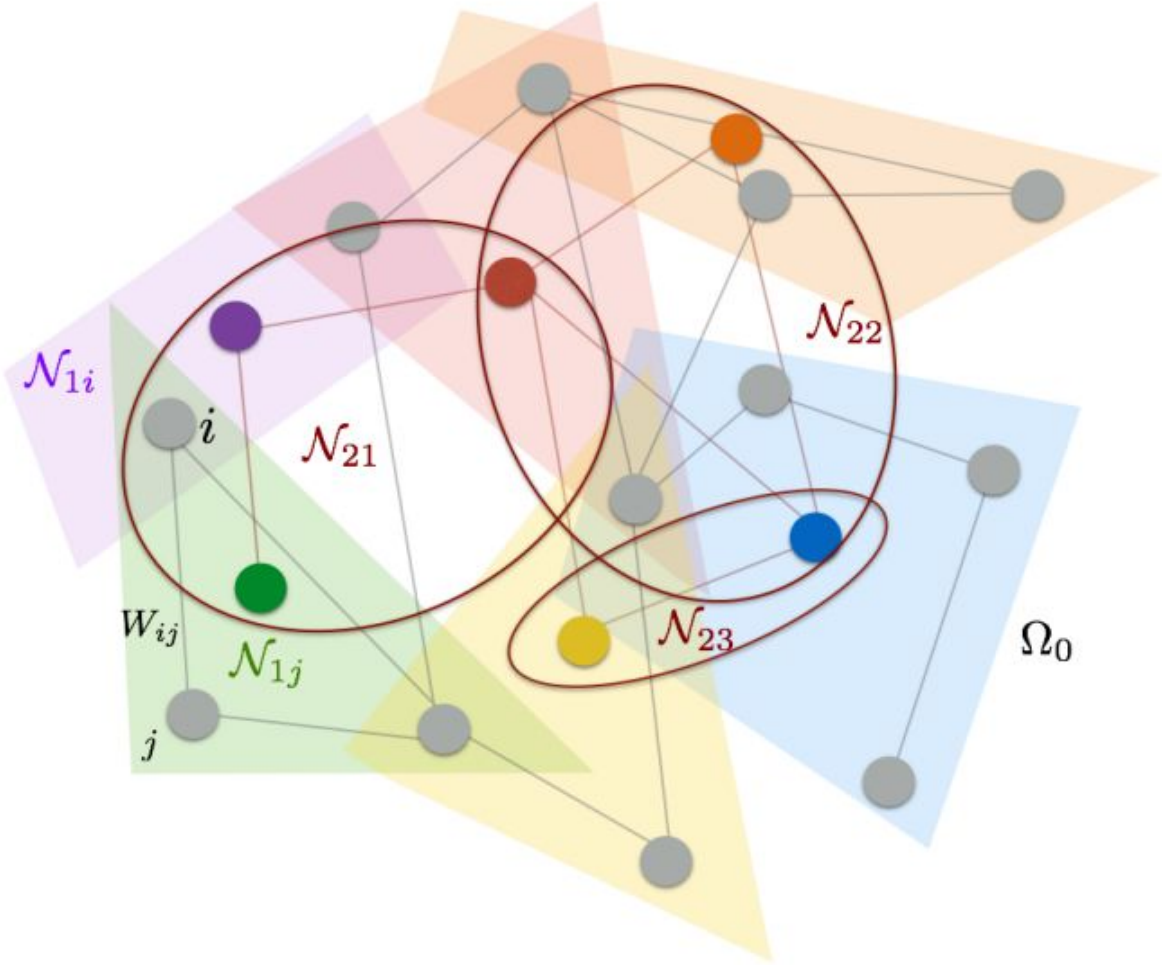
Arthur Szlam
The City College of New York
aszlam@ccny.cuny.edu

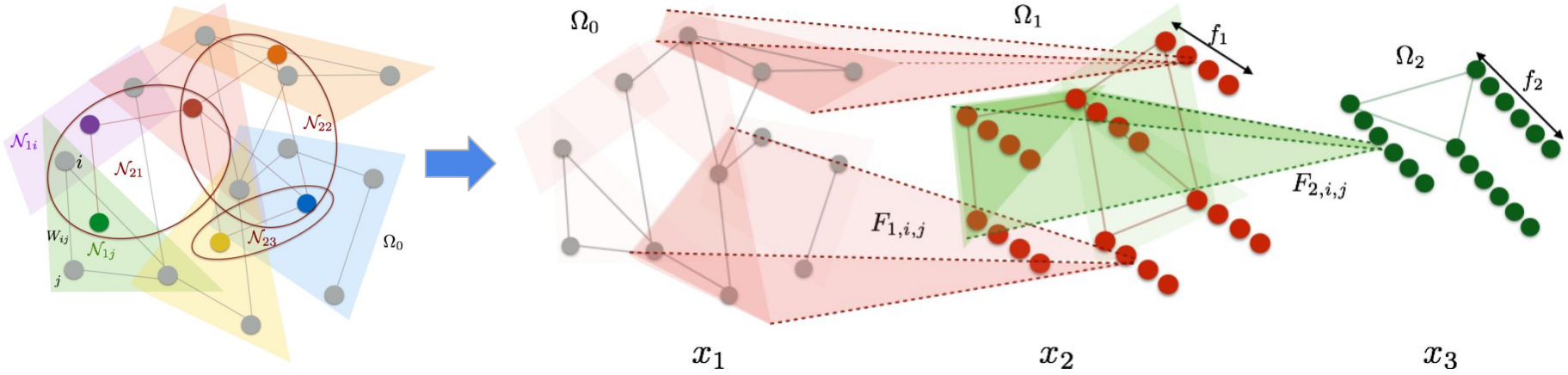
Yann LeCun
New York University
yann@cs.nyu.edu

Abstract

Convolutional Neural Networks are extremely efficient architectures in image and audio recognition tasks, thanks to their ability to exploit the local translational invariance of signal classes over their domain. In this paper we consider possible generalizations of CNNs to signals defined on more general domains without the action of a translation group. In particular, we propose two constructions, one based upon a hierarchical clustering of the domain, and another based on the spectrum of the graph Laplacian. We show through experiments that for low-dimensional graphs it is possible to learn convolutional layers with a number of parameters independent of the input size, resulting in efficient deep architectures.

https://arxiv.org/abs/1312.6203


$\mathcal{N}_{1i}$
i
$\mathcal{N}_{21}$
$\mathcal{N}_{22}$
W_{ij}
$\mathcal{N}_{1j}$
j
$\mathcal{N}_{23}$
Ω_0


$\mathcal{N}_{1i}$
i
$\mathcal{N}_{21}$
$\mathcal{N}_{22}$
W_{ij}
$\mathcal{N}_{1j}$
$\mathcal{N}_{23}$
j
Ω_0
Ω_0
Ω_1
f_1
Ω_2
f_2
$F_{2,i,j}$
$F_{1,i,j}$
x_1
x_2
x_3

Henaff et al., 2015
Deep Convolutional Networks on Graph-Structured Data

Deep Convolutional Networks on Graph-Structured Data

Mikael Henaff
Courant Institute of Mathematical Sciences
New York University
mbh305@nyu.edu

Joan Bruna
University of California, Berkeley
joan.bruna@berkeley.edu

Yann LeCun
Courant Institute of Mathematical Sciences
New York University
yann@cs.nyu.edu

Abstract

Deep Learning's recent successes have mostly relied on Convolutional Networks, which exploit fundamental statistical properties of images, sounds and video data: the local stationarity and multi-scale compositional structure, that allows expressing long range interactions in terms of shorter, localized interactions. However, there exist other important examples, such as text documents or bioinformatic data, that may lack some or all of these strong statistical regularities.

In this paper we consider the general question of how to construct deep architectures with small learning complexity on general non-Euclidean domains, which are typically unknown and need to be estimated from the data. In particular, we develop an extension of Spectral Networks which incorporates a Graph Estimation procedure, that we test on large-scale classification problems, matching or improving over Dropout Networks with far less parameters to estimate.

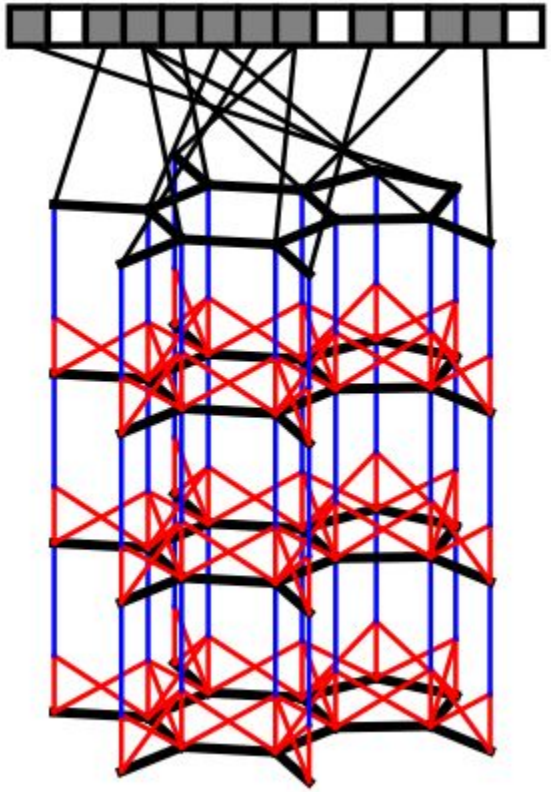
Duvenaud et al., 2015
Convolutional Networks on Graphs
for Learning Molecular Fingerprints

Convolutional Networks on Graphs for Learning Molecular Fingerprints

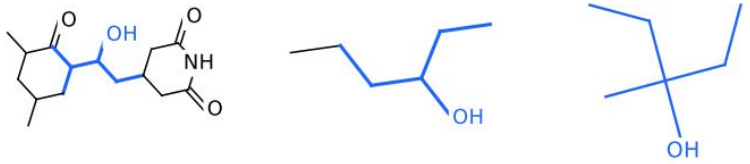
David Duvenaud[†], **Dougal Maclaurin**[†], **Jorge Aguilera-Iparraguirre**
Rafael Gómez-Bombarelli, Timothy Hirzel, Alán Aspuru-Guzik, Ryan P. Adams
Harvard University

Abstract

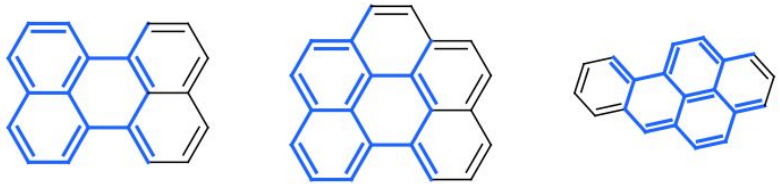
We introduce a convolutional neural network that operates directly on graphs. These networks allow end-to-end learning of prediction pipelines whose inputs are graphs of arbitrary size and shape. The architecture we present generalizes standard molecular feature extraction methods based on circular fingerprints. We show that these data-driven features are more interpretable, and have better predictive performance on a variety of tasks.



Fragments most activated by pro-solubility feature



Fragments most activated by anti-solubility feature



Fragments most activated by toxicity feature on SR-MMP dataset

Fragments most activated by toxicity feature on NR-AHR dataset

Li et al., 2016
Gated Graph Sequence Neural Networks

GATED GRAPH SEQUENCE NEURAL NETWORKS

Yujia Li* **& Richard Zemel**
Department of Computer Science, University of Toronto
Toronto, Canada
{yujiali,zemel}@cs.toronto.edu

Marc Brockschmidt & Daniel Tarlow
Microsoft Research
Cambridge, UK
{mabrocks,dtarlow}@microsoft.com

ABSTRACT

Graph-structured data appears frequently in domains including chemistry, natural language semantics, social networks, and knowledge bases. In this work, we study feature learning techniques for graph-structured inputs. Our starting point is previous work on Graph Neural Networks (Scarselli et al., 2009), which we modify to use gated recurrent units and modern optimization techniques and then extend to output sequences. The result is a flexible and broadly useful class of neural network models that has favorable inductive biases relative to purely sequence-based models (e.g., LSTMs) when the problem is graph-structured. We demonstrate the capabilities on some simple AI (bAbI) and graph algorithm learning tasks. We then show it achieves state-of-the-art performance on a problem from program verification, in which subgraphs need to be described as abstract data structures.

Defferrard et al., 2016
Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering

Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering

Michaël Defferrard **Xavier Bresson** **Pierre Vandergheynst**

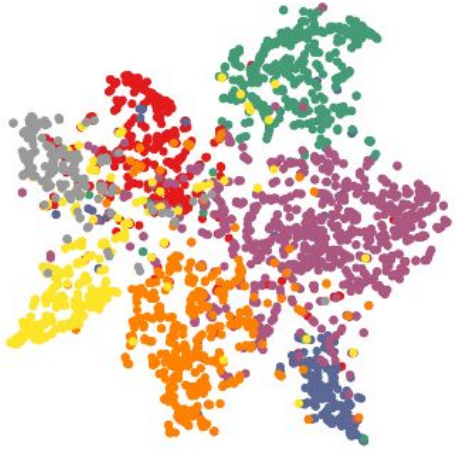
EPFL, Lausanne, Switzerland
{michael.defferrard,xavier.bresson,pierre.vandergheynst}@epfl.ch

Abstract

In this work, we are interested in generalizing convolutional neural networks (CNNs) from low-dimensional regular grids, where image, video and speech are represented, to high-dimensional irregular domains, such as social networks, brain connectomes or words' embedding, represented by graphs. We present a formulation of CNNs in the context of spectral graph theory, which provides the necessary mathematical background and efficient numerical schemes to design fast localized convolutional filters on graphs. Importantly, the proposed technique offers the same linear computational complexity and constant learning complexity as classical CNNs, while being universal to any graph structure. Experiments on MNIST and 20NEWS demonstrate the ability of this novel deep learning system to learn local, stationary, and compositional features on graphs.

Kipf & Welling, 2017

Semi-Supervised Classification with Graph Convolutional Networks



(b) Hidden layer activations

SEMI-SUPERVISED CLASSIFICATION WITH GRAPH CONVOLUTIONAL NETWORKS

Thomas N. Kipf
University of Amsterdam
T.N.Kipf@uva.nl

Max Welling
University of Amsterdam
Canadian Institute for Advanced Research (CIFAR)
M.Welling@uva.nl

ABSTRACT

We present a scalable approach for semi-supervised learning on graph-structured data that is based on an efficient variant of convolutional neural networks which operate directly on graphs. We motivate the choice of our convolutional architecture via a localized first-order approximation of spectral graph convolutions. Our model scales linearly in the number of graph edges and learns hidden layer representations that encode both local graph structure and features of nodes. In a number of experiments on citation networks and on a knowledge graph dataset we demonstrate that our approach outperforms related methods by a significant margin.

After Discussion 180213

To-Do List

1. **Learn Keras**
   a. **BOOK: Deep Learning with Python**
   b. **Stop implement from scratch**

2. **Try to implement convolution**
   a. **Image convolution**
   b. **Graph convolution**

Image Convolution

- **Get the idea of convolution and pooling -> implement in Python**
- **Try to perform the operation on some simple image**
  - **(ex: edge detection)**
- **Signal Convolution (Optional)**

Graph Convolution

Paper

Kipf & Welling, 2017

Semi-Supervised Classification with Graph Convolutional Networks

Try to understand what is happening here and see if you could implement it in numpy

2 Fast Approximate Convolutions on Graphs

In this section, we provide theoretical motivation for a specific graph-based neural network model $f(X, A)$ that we will use in the rest of this paper. We consider a multi-layer Graph Convolutional Network (GCN) with the following layer-wise propagation rule:

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) . \tag{2}$$

Here, $\tilde{A} = A + I_N$ is the adjacency matrix of the undirected graph $\mathcal{G}$ with added self-connections. I_N is the identity matrix, $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$ and $W^{(l)}$ is a layer-specific trainable weight matrix. $\sigma(\cdot)$ denotes an activation function, such as the $\mathrm{ReLU}(\cdot) = \max(0, \cdot)$. $H^{(l)} \in \mathbb{R}^{N \times D}$ is the matrix of activations in the l^{th} layer; $H^{(0)} = X$. In the following, we show that the form of this propagation rule can be motivated[1] via a first-order approximation of localized spectral filters on graphs (Hammond et al., 2011; Defferrard et al., 2016).

Graph Convolution

Paper

Kipf & Welling, 2017

Semi-Supervised Classification with Graph Convolutional Networks

Note:

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) . \qquad (2)$$

- Graph-based neural netwtork model f(X, A)
- Undirected graph G -> adjacency matrix A
- Ā = A+I (add self-connections)
- D_{ii} = sum_j Ā_{ij}
- W^{(l)} is a layer-specific trainable weight matrix
- \sigma: activation function

- Graph Laplacian
  - L = I - D^{-1/2} A D^{-1/2}

Graph Convolution

Paper

Kipf & Welling, 2017

Semi-Supervised Classification with Graph Convolutional Networks

Try to work on some simple graphs

Linear graph
Cycle
Complete graph
Star
Erdos-Renyi
Scale free