Critique on the Paper “Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering”

- E Santhosh Kumar (CS16B107)

Introduction:

CNNs have lead to a break-through in image, video and sound recognition tasks. They are capable of identifying local features in a translation invariant manner and further hierarchically combine these to understand more complex features. However, these data types fall in the Eucledian domain with a regular spatial structure. For instance, images are simply color values assigned over a 2 dimensional grid and similarly sound is just oscillations over a 1 dimensional time domain.

There are several cases where the data exebits lies on a irregular, non-Eucledian domain expressible through a graph. The graph consists of vertices and edges between the vertices, and each vertex may contain several features. Examples of this include banking transactions, social networks, brain mapping, etc.

In each of these cases, utilising the topology of the graph is important to understand and extract the underlying features of the data. Thus, simply embedding the data into a lower dimensional space and then use our existing techniques is not enough. As an example, this has the same effect as flattening images and then using a regular neural network on them.

The paper pushes forward the previous work of Henaff-Bruna-LeCun on extending CNNs to graphs and overcomes some of the limitations of the previous work. Chiefly, computational complexity of this model is linear with the dimensionality of the data, similar to regular CNNs.

Problem Setting:

- Given: a fixed undirected graph (composed of vertices and edges (may be weighted) between the vertices)
- Data: Input: features of each of the graph vertices
- Output: corresponding to the features of the vertices

Summary:

Due to the irregular nature of graphs, and the fact the translation in a graph is not clearly defined, there is a need to define **i) localized convolution filters for graphs**, **ii) graph coarsening and pooling** (similar in effect to the pooling layers in regular CNNs).

i) Convolution

Convolution on graphs is easier to perform in the Fourier Domain as there is no requirement for the definition of translation. An important operator in obtaining the Fourier modes of a graph is the graph **Laplacian**. It is a real symmetric, positive semidefinite matrix obtained from the edge weights of the graph, and contains information related to the topology of the graph (for instance, the eigenvalues of the Laplacian matrix can be used to find the number of connected components of the graph). The eigen-vectors of the Laplacian form the Fourier vectors (a.k.a Fourier modes). They give us the smoothest orthogonal domain on the graph domain.

The Fourier transform enables the formulation of filtering in the graph domain. By using polynomial filters defined as Kth order polynomials of the Laplacian, we are able to strictly **localize the the filter to a ball of radius K**. Furthermore, the sparse nature of the Laplacian is used to develop a recursive formulation for fast filtering.

ii) Graph Coarsening and Pooling

These operations together have the effect of pooling in regular CNNs. We cluster the graph by merging similar vertices forming meaningful neighbourhoods, and then assigning the features for these new vertices. We use the coarsening phase of the Graclus multi-level clustering, where each level produces a coarser graph corresponding to the previous layer. Moreover, this algorithm offers precise control over the size of the graph as it reduces the graph size by half at every level.

For pooling, in order to avoid the expense of maintaining a table to store the vertices matched during coarsening, we create a balanced binary tree of the vertices at each level. This allows pooling to be done similar to a 1D signal and can be parallelized.

Model Performance:

i) MNIST dataset: Given that the Eucledian grid is just a specific case of a graph, the model shows performance close to that of a regular CNN. However, it is not able to perform better than the same, with a possible explanation being the isotropic nature of spectral filters that prevent it from understanding directions (which, unlike in a generic graph, is present in a 2D grid).

ii) 20NEWS dataset: The model defeated a fully connected neural network (with more parameters), but does not outperform the Multinomial Naive-Bayes Classifier.

Limitations and Scope for Future Work:

i) The techniques seen so far are limited to cases with a single graph and cannot work on cases with varying graph sizes or even in cases where the adjacency matrix changes (keeping the number of vertices fixed). This is because the Eigen vectors of the Laplacian (the Fourier Modes) change drastically even with small changes in the graph.This implies that the model is not transferable.

ii) The isotropic nature of the model might give it an inherent disadvantage when used on certain datasets (as seen in the case of MNIST).

iii) The Laplacian is not clearly defined for directed graphs.