

# Convolutional Neural Network Application to Graph Analysis

Cameron Ibrahim

September 2017

## Abstract

I propose the use of Convolutional Neural Networks for the use of graph analysis, by making use of the ability to model a graph as a 2-d matrix, similar to the representation of an image, which CNNs are uniquely prepared to handle. This could perhaps be used to predict features of local nodes in the graph, or perhaps even to identify patterns or structures in the graph itself. This is quite an interesting proposition, as it attempts to unify two disparate topics, namely Neural Networks and Graph Analysis.

## 1 Research Context

Convolutional Neural Networks is a class of deep, feed-forward neural networks, consisting of alternating convolutional and pooling layers, before running it through a final pass of fully connected layers. CNNs were specifically designed with images in mind and work especially well on their large, matrix-like, layered data structures. Through the multitude of filters present in each convolutional layer, features of the image space may be captured by the model.

Let us now consider graphs; graphs are characterized by their nodes and edges. Each node is connected by some degree  $k$  of edges to other nodes. One way that graphs may be represented is through a matrix - each row/column identifies a node in the graph, while a non-zero value in cell indicates an edge between the nodes indicated by their row and column.

It is this that I would like to look further into - can we analyze patterns in a graph component by feeding it into a convolutional neural network.

## 2 Data/Design

The Yelp data provides ample resources for this task, as it comes fully prepared with an extensive list of users, as well as user to user friend connections. Additional data that might be useful/interesting would be to somehow gain access to the dataset of a larger social media site, such as Facebook or Twitter, to get a

more expansive set of user connections, with different possibilities for additional insight. Alternatively, we could simply write a web crawler to achieve the same purpose.

I cannot currently speak to the cleanliness of the data, as attempting to load it crashed my computer. This however, is insight on it's own: the dataset of friend connections on Yelp is quite large.

In this case, the design for the research would be to experiment with different constructions of Convolutional Neural Networks with graphs as input, while trying to ascertain a number of possible variables.

### 3 Methods

For the research itself, the star of the show is obviously the CNN - it's the main player in this endeavor. I believe we may make similar use of the matrix-like structure that graphs may be expressed in to take advantage of the feature filters of the convolutional layers.

As for the expected data, it varies depending on what one is hoping to get out of it. For example, one may want to determine which node in the graph likely has the most friends based on the graphs structure. This is comparatively easier to verify, than say, trying to figure out how many clusters are in a graph portion.

### 4 Significance

This technique, if successful, could allow Yelp to determine a great deal of information about the structure of their userbase and grant additional insights into how they experience it. Additionally, it would likely have a great deal of impact on the field of graph analysis, as well as network flow and other realated fields.

### 5 References

<https://arxiv.org/abs/1609.02907>  
<https://arxiv.org/abs/1606.09375>