UNIVERSITY of HOUSTON
Department of Computer Science

Community Analysis of Deep Networks

Behrang Mehrparvar (behrang@cs.uh.edu)
Ricardo Vilalta (r.vilalta.us@ieee.org)

PATTERN ANALYSIS LAB
finding meaningful patterns in scientific data

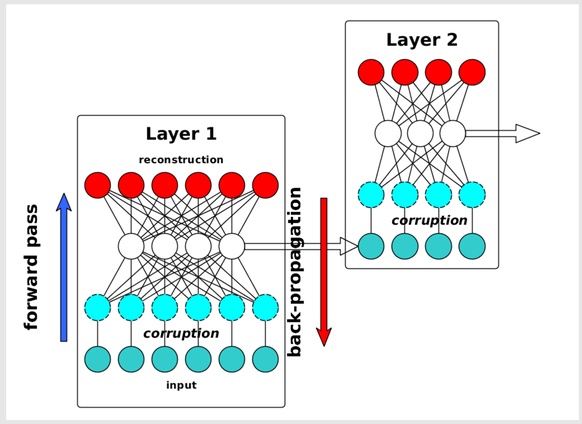
Introduction

The core motivation for this research is based on the collaboration between hidden units in each layer of the deep network.

Definition: Nodes *X* and *Y* are defined as **collaborating** if they have high mutual information.

Definition: A set of nodes is defined as a **community** if the pairwise collaboration within the set is greater than the pairwise collaboration of each of nodes with nodes outside the community.

In order to provide community analysis of deep networks, we formulate leading questions of the research. For each leading research question, we provide a motivating reason for the question, potential outcomes of answering the question, initial hypotheses, basic definitions, corresponding experiments, results, final analysis and potential theories.



Stacked Denoising Auto-encoder

Preliminary Experiments

In order to provide initial observations to develop hypotheses, we conducted two sets of experiments:

- Real-world data
- Toy example

For real-world experiments we used three datasets:

- **MNIST** handwritten digits
- **MNIST_rot** rotated handwritten digits
- **CIFAT-10** object recognition dataset

Representations are learned using a 6 layer **stacked denoising auto-encoder** [1]. Comparing different **community detection algorithms** [2] we chose fastgreedy, leading eigenvector and multilevel optimization.

We measured modularity and number of communities in each epoch. **Modularity** [3] is a measure of how disjoint a community structure is.

For toy examples, we trained a neural network with one hidden layer to approximate a function with 2, 3 and 4 hidden units.

Research Direction

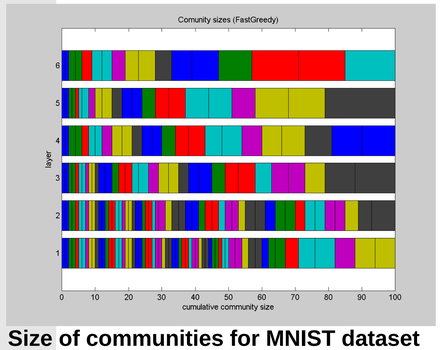
1 Existence
Do units in hidden layers form collaborating communities?
Experiment – significance of modularity

2 Necessity
Are collaborating communities necessary for representing the input or are they redundant?
Experiment – feature selection among communities followed by performance evaluation

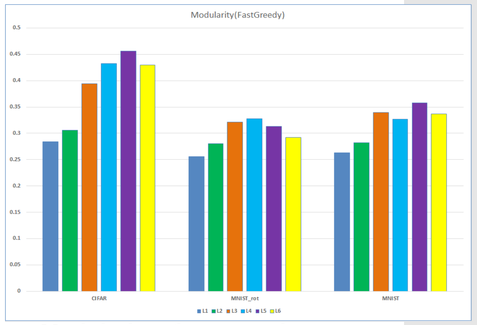
3 Specialization
Is each community of collaborating units specialized in a specific task?
Experiment – performance evaluation on predicting known factors of variation of synthetic data

4 Abstraction
Is the structure of communities in lower-level related to the formation of higher-level features?
Experiment – matching membership in communities and involvement in abstraction

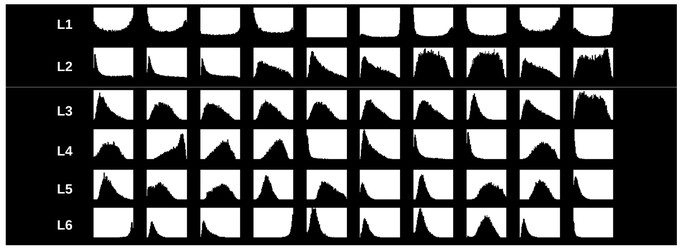
Real-world Dataset Results



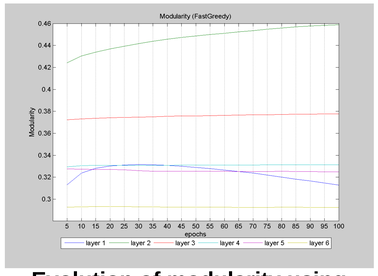
Size of communities for MNIST dataset



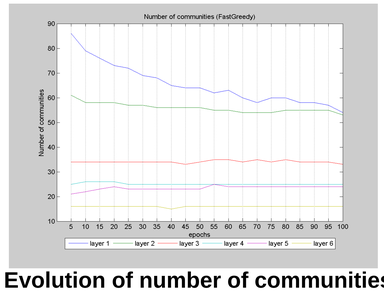
Modularity of community structures (best performance settings)



Sample distributions of activations along layers



Evolution of modularity using fastgreedy on MNIST

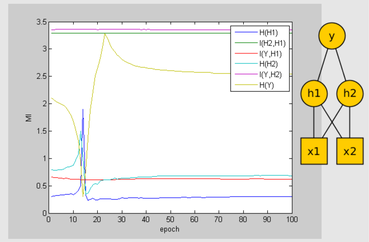


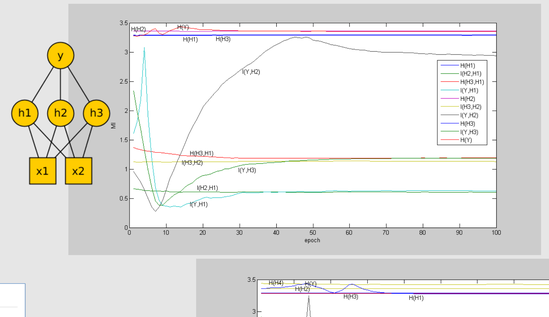
Evolution of number of communities using fastgreedy on MNIST

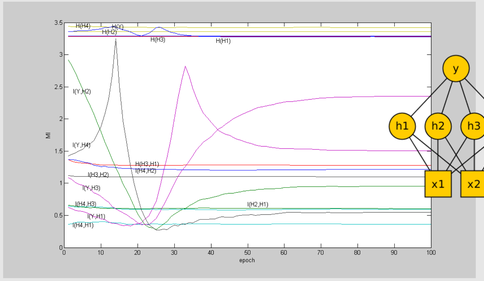
Toy Example Results

Observations:

- Adding noise to the output results in a fluctuation
- After enough epochs entropy and mutual information will converge to a stable value







Conclusion

Althouh deep networks have provided significant performance boost in different AI tasks, however further understanding of details of the behavior of the learning algorithms is required. Sporadic research papers such as [4] and [5] have been devoted to understanding different aspects of deep learning algorithms. In the proposed research similarly we plan to provide insight about deep networks in the light of community analysis.

Future Works

Future works of the proposed research can be listed as follows:

- **Overlapping** community analysis of deep networks
- Studying the **evolution** of communities in relation with the gradient descent
- Effect of **extension** of hidden layers on existing collaborating communities
- Relation between **over-fitting** and detected communities
- Developing **new learning algorithms** based on communities
- Choosing **number of hidden layers** and **hidden units**
- A **measure of goodness** of representation

References

[1] Vincent, Pascal, et al. "**Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion**." The Journal of Machine Learning Research 11 (2010): 3371-3408.

[2] Fortunato, Santo. "**Community detection in graphs**." Physics Reports 486.3 (2010): 75-174.

[3] Clauset, Aaron, Mark EJ Newman, and Cristopher Moore. "**Finding community structure in very large networks**." Physical review E 70.6 (2004): 066111.

[4] Montavon, Grégoire, Mikio L. Braun, and Klaus-Robert Müller. "**Kernel analysis of deep networks**." The Journal of Machine Learning Research 12 (2011): 2563-2581.

[5] Goodfellow, Ian, et al. "**Measuring invariances in deep networks**." Advances in neural information processing systems. 2009.