The "Networks"

Learning how to build artificial neural networks from real dynamic networks

Sikun Lin

University of California, Santa Barbara

1 Introduction

It is the complexity of automatic thinking and reasoning that distinguishes human beings from machines. There is still a very large gap between current machine learning ability and the general artificial intelligence, although researchers have made profound advances recently. I've been continuously fascinated by the idea of creating a machine with true intelligence. My aim is to acquire knowledge from human intelligence, and then apply it to build smarter machines. It is not to say the best way towards a general intelligence is through imitating nature, but surely there are more inspirations to find.

My research goal is designing efficient schemes for NAS (Neural Architecture Search) using the knowledge of dynamic networks. NAS is useful in reducing human intervention during the model designing process and a crucial step towards automated machine learning and general AI. To achieve this goal, I am researching on modeling real-world dynamic networks (discussed in section 2) and will apply the insight to NAS. Inside the supergraph search space that I will explore, as discussed in 3.1, NAS not only helps with the problem of multi-task learning but also low-shot learning (section 3.2). My previous experience in computer vision, as mentioned in section 3.3, also helps me establish a strong base for exploring this goal in the deep learning context and testing new models with familiar tasks.

2 Analyzing and modeling dynamic networks

Humans are capable of carrying out complex tasks and thinking thanks to our intricate brain networks, which haven't unraveled their mystery to us despite the fast development in neuroscience. As the abstract essence of human intelligence is dynamic networks, I regard 'understand how the real-world dynamic networks work' as the first part of my research.

Currently, I am working on modeling the heterogeneity of brain MRI data through optimization approaches. By estimating the inverse covariance matrix of fMRI data under network constraints, we find pathways of brain activation when subjects are doing different tasks. We are also investigating the relationship between structural and functional networks, with the former indicating the physical neuronal wiring between brain regions and the latter showing activity patterns between brain regions under different conditions. Finding this relationship can be useful for predicting brain activities based on specific structural connectivity. Taking sample variations into account, it is more reasonable to jointly estimate multiple models for heterogeneous groups, instead of estimating a single model for the whole population. We are also solving this through optimization approaches by formulating an objective function with added group assignment penalties. The central theme of this project is to get a better understanding of '*how* intelligence arises' by studying brain activation patterns and how it handles different tasks within or across brain regions. We can convert these activation patterns into representations for each task, and they will provide valuable insights into creating models that can manage multiple tasks, and eventually contribute to building general purpose AI.

Besides developing convex optimization based techniques, exploring GDL (Geometric Deep Learning) [1] can also help with graph analysis. As an exemplar of GDL, GCN (Graph Convolutional Networks) builds upon spectral graph theory and shows its usefulness especially in medicine and relational analysis. Its robustness in learning graph embeddings is well demonstrated in [2]. I see GCN (and Graph Neural Networks in general) as a neat combination of graph theory and deep learning. I will do more research on it as there is much more can be learned, including studying evolving networks on different stages and finding robust network representations, both of which are important for dynamic network modeling and explanation.

3 Towards Artificial Neural Networks

After getting an understanding of real-world networks, it is time for creating artificial ones. My goal is to use the knowledge of brain network and possibly social network mechanisms as the guide for devising 'brains' and actions of artificial agents. NAS (section 3.1) will be my main research focus: it can utilize the knowledge from dynamic network modeling and benefit low-shot learning (section 3.2) as well. My previous projects in computer vision (section 3.3) also got me prepared for handling artificial neural networks.

3.1 NAS: Neural Architecture Search

Deep Neural Networks save researchers from laborious work of designing features by hand, but put them into a new situation which is no better, and comes with even more drudgery: carefully crafting network architectures and tuning hyperparameters, which takes even more time and computational resources. Yet one network with a certain architecture and hyperparameters is often task-specific. This kind of task oriented manual model design has low efficiency and will hardly lead us to the true machine intelligence. Creating better NAS algorithms will not only alleviate repeated human labor but also discover artificial models with higher level intelligence.

During my experience of using deep neural networks, I found choosing hyperparameters to be a difficult and sometimes unreasonable task. Then I started to think about incremental models with neurons and layers added and controlled by some search strategies. I then experimented with only a few neurons using evolutionary algorithm during my computational neuroscience class and established automatic model search as my research interest. That is how I encountered, read about, and committed myself to NAS, a much more formal and better-defined version of my research interest. It targets at searching an optimal network architecture for a task inside a particular search space with a set of searching strategies [3]. It is a nascent topic and has a large room for improvement:

- Improvement in search space: to constrain the computation time into a realistic scale, current researchers make particularly strict requirements about NAS search space, allowing only a few predefined cell types, and their connection ordering must obey particular rules. These cell types and connecting rules exploit empirical knowledge of existing successful networks, yet they restrict too much from the exploration to unknown and maybe better possibilities. I encountered an idea this summer, which in my opinion can provide a more flexible search space and is worth more explorations: One-Shot Architecture Search [3]. It treats all task-specific models as subgraphs of a supergraph (the one-shot model), and share weights between different architectures based on whether they have common edges in the supergraph. Although in reality, its realizations are still limited in predefined manners, I can see from here the potentials of resembling brain network pathways and will work on it as my future research project. Together with the idea of PathNet [4] and Network Morphism [5], I consider the search space to be a fully connected graph with different pathways activated for various tasks, just like brain networks, and the search for these pathways can be monitored by hierarchical meta-controllers.

- Improvement in search strategy and performance estimation: These two are critical issues that determine whether NAS is possible for general researchers. Although being a promising way to reduce manual labor, NAS algorithm can be extremely resource consuming, converging using up to thousands of GPU days just for searching a single architecture in a restricted search space. The bottleneck lies in training and evaluating individual network choices during each round. As for search strategy, RL (reinforcement learning) and evolutionary algorithms take longer time than Bayesian optimization and gradient-based methods. I used gradient-based DARTS [6] in searching a suitable fashion pattern classification model during the internship at Markable.ai this summer. The model has a DenseNet [7] backbone, and we were attaching an additional head before the final classification layers for improving the accuracy. The search converges within one GPU day, and the resulting head not only outperforms our original DenseNet block design, but also has fewer parameters. Thus, NAS can give us more efficient model architectures than the hand-designed ones. But the DARTS I used only performs search in a highly constrained layer type and connection settings partly due to the nature of its gradient-based searching strategy, and I would like to further explore NAS using RL or evolutionary methods, which has more flexibility to find better models. Therefore, together with exploring the supergraph idea mentioned above, I will research how to reduce the performance evaluation time efficiently (approximation through the learning curve, etc.) to make NAS practical.

Ideally, the learned network should be sustainable, in the sense of being able to reuse its architecture and parameters, can be easily transferred to a large variety of tasks, can be fine-tuned jointly by wider users, and being capable of mastering new skills through active learning.

3.2 Low-shot learning

When it comes to building human-like intelligence, it is hard not to mention one-shot and low-shot learning, namely seeing one or few examples to learn a new concept. Through comparison, through focusing attention on the most essential part, through learning by analogy, human beings can master the skill of one-shot learning. On the contrary, deep neural networks require a large amount of data to carry out a specific task. Currently, researchers attempt to alleviate this problem in several directions: metric learning which essentially implements the similarity comparisons and clustering, meta-learning (so-called 'learning to learn') which learns better initializations, optimizers, etc. for a faster model convergence across different tasks, external attached memory modules that can be written and read by the network, and several generative models for one-shot generalization. I worked on low-shot learning in two projects, both using metric learning techniques based on object similarity: the first one is a working paper on one-shot one-class learning using similarity networks, which solves the class granularity problem of newly encountered classes given a few positive examples. Query images will be classified as similar to the examples or not according to positive examples category level (ex. Dog, Animal, Living-things are three different levels). The second project was at Markable.ai, and we were using metric learning to get a better classification result for clothes attributes, such as sleeve types and dress types, with only a few labeled data. Apart from metric learning, I will use attention or memory mechanisms for the low-shot setting, especially combining them as sub-modules of a greater supergraph as mentioned in section 3.1. The parameter sharing scheme in the supergraph will also transfer knowledge between tasks, which may help warm start a new task with only a few data.

3.3 Experience in Computer Vision

Machine learning attracted me through its application in computer vision. During my master's degree, I implemented gesture recognition on Google glass for smoother AR file transfer experience, resulting in a paper [8]. I also worked on an AR platform with real-time image recognition and tracking, as written in paper [9]. My Masters thesis [10] works on detecting personalized attention for different user preferences using deep learning. Apart from the low-shot learning projects mentioned in section 3.2, I also worked on weakly-supervised fashion instance segmentation this summer at Markable.ai, which helped improve the final product with respect to color prediction and item search. As the research progresses, I appreciate both traditional computer vision and recent advances in deep learning, and now want to find even better ways of designing the models through the understanding of networks.

Computer vision experience gives me a base in deep learning, whereas my current and future research focus will be more on dynamic network modeling and apply that knowledge to NAS. I want to design algorithms for machines to be able to learn better, not only restricted to vision-specific tasks. As for testing, improving, and demonstrating the algorithm's ability, I will make use of my familiarity with vision tasks and benchmark models.

4 Conclusions

As a summary, my research will focus on developing efficient algorithms to better model real-world dynamic networks, especially brain networks, and seeking inspiration on how to build artificial network models that are more human-like. The target model will be capable of handling multiple tasks and unfamiliar scenarios. By utilizing the knowledge of real dynamic networks, NAS can be useful for finding the optimal model structure for multiple tasks automatically. This will result in fewer parameters needed for all the tasks combined, as well as transferring the performance gain from one task to another, and handle newly encountered concepts with less effort than usually required.

References

[1] Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

[2] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

[3] Thomas Elsken, Jan Hendrik Metzen, and Frank Hutter. Neural architecture search: A survey. *arXiv preprint arXiv:1808.05377*, 2018.

[4] Chrisantha Fernando, Dylan Banarse, Charles Blundell, Yori Zwols, David Ha, Andrei A Rusu, Alexander Pritzel, and Daan Wierstra. Pathnet: Evolution channels gradient descent in super neural networks. *arXiv preprint arXiv:1701.08734*, 2017.

[5] Tao Wei, Changhu Wang, Yong Rui, and Chang Wen Chen. Network morphism. In *International Conference on Machine Learning*, pages 564–572, 2016.

[6] Hanxiao Liu, Karen Simonyan, and Yiming Yang. Darts: Differentiable architecture search. *arXiv preprint arXiv:1806.09055*, 2018.

[7] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. *CVPR*, 1(2):3, 2017.

[8] Sikun Lin, Hao Fei Cheng, Weikai Li, Zhanpeng Huang, Pan Hui, and Christoph Peylo. Ubii: Physical world interaction through augmented reality. *IEEE Trans. Mob. Comput.*, 16(3):872–885, 2017.

[9] Wenxiao Zhang, Sikun Lin, Farshid Hassani Bijarbooneh, Hao Fei Cheng, and Pan Hui. Cloudar: A cloud-based framework for mobile augmented reality. In *Proceedings of the on Thematic Workshops of ACM Multimedia 2017*, pages 194–200. ACM, 2017.

[10] Sikun Lin and Pan Hui. Where's your focus: Personalized attention. *arXiv preprint arXiv:1802.07931*, 2018.