CS81 Project Proposal

Alan Zhao, Gabriel Meyer-Lee, and Harsha Uppili

1 Introduction

The great breakthrough that deep learning has provided over past AI algorithms was to remove the necessity of feature engineering systems designed by humans and to allow the AI system to learn complex representations itself from the raw data. This breakthrough has led to a new level of sophistication in AI systems and massive increases in the performances of neural networks trained for supervised learning tasks. However, deep neural networks have introduced a new set of topological variations and hyperparameters that must be designed by humans based on vague guidelines and experimentation. Designing the architectures and tuning the hyperparameters of deep neural networks by hand remain tedious and potentially incorporates human biases. One approach to overcoming this shortfall is to utilize Evolutionary Algorithms to evolve neural networks. The core task for our research will be the development of a framework capable of using Evolutionary Algorithms to make these design decisions with minimum human involvement and input. Specifically, we will be using an evolutionary framework to create deep neural networks in Keras and will evaluate them on both image recognition and language modeling benchmarks. Our evolutionary framework will be a novel algorithm that we will develop, inspired by by the CoDeepNEAT framework [4], Hierarchical SANE [5], ESP [2], and CoSyNE [3].

2 Hypothesis

Deep learning traditionally focuses on programming a deep neural network. We believe that evolutionary algorithms can be developed to use and generate architectures of powerful deep neural networks. We intend for these algorithms to have the ability to incorporate the many types of layers of Keras. We expect our evolutionary algorithm to successfully evolve deep neural networks capable of solving the tasks we train them on.

On the other hand, the two most established types of deep neural networks are the Convolutional Neural Network (CNN) and the Recurrent Neural Network (RNN). We believe that having a evolutionary framework can help decide which type (or types) of deep neural networks should be used for a specific task, and could potentially perform better than a manually-built CNN or RNN architecture.

3 Implementation and Experiments

The starting point for the evolutionary algorithm development will be the DEvol framework[1] developed by Joe Davison. The DEvol framework utilizes genetic programming to architect deep neural networks based on t a limited number of hyperparameters of thee Keras library. Each model within DEvol is encoded as a fixed-width genome. During evolution, genetic crossover takes place over two generated models to get a resulting model architecture. We performed a preliminary experiment to evolve a CNN to classify MNIST hand-written digits images. We observed evolution patterns across generations, and the resulting accuracy was on par with hand-designed CNNs.

Our evolutionary framework will extend on this to create a system largely inspired by CoDeepNEAT, but also drawing from the neural network coevolution work of Mikkulainen. Our goal is to create a system capable of evolving significantly different types of structures of neural networks for different tasks. Experimentation with these evolutionary algorithms and design of our implemented system will be led by Gabriel. We plan to extend the hyperparameters with which evolution can evolve, especially the number and type of layers.

We are going to test the evolutionary framework on benchmark image recognition datasets such as CIFAR10 and CIFAR100. The CIFAR datasets are labeled images and are widely used in computer vision benchmarks. Each image has the dimensions 32x32 and 3 RGB channels. CIFAR10 has 10 categories while CIFAR100 has 100. Alan will lead the experiments of image recognition.

We also plan on designing evolutionary algorithms to evolve neural networks trained on language modeling datasets such as the Penn Tree Bank. Specifically, we will seek to evolve Recurrent Neural Networks that fit probabilistic models which assign probabilities to sentences by predicting future words based on the previous words. The Penn Tree Bank is a popular benchmark for measuring the quality of these RNNs. Harsha will lead language modeling experimentation.

4 Analysis

We will analyze the performances based on the results of both our image recognition and language modeling tasks. As our technique is based on an evolutionary algorithm, we will consider the performance of our evolutionary framework to be the performance of the best network evolved by that framework for each of our evolutionary tasks. The evolutionary framework can be compared against hand designed networks to determine its effectiveness.

References

[1] Devol - deep neural network evolution. https://github.com/joeddav/devol.

[2] Faustino Gomez and Risto Miikkulainen. Incremental evolution of complex general behavior. *Adaptive Behavior*, 5(3-4):317–342, 1997.

[3] Faustino Gomez, Juergen Schmidhuber, and Risto Miikkulainen. Accelerated neural evolution through cooperatively coevolved synapses. *Journal of Machine Learning Research*, pages 937–965, 2008.

[4] R. Miikkulainen, J. Liang, E. Meyerson, A. Rawal, D. Fink, O. Francon, B. Raju, H. Shahrzad, A. Navruzyan, N. Duffy, and B. Hodjat. Evolving deep neural networks. *ArXiv e-prints*, March 2017.

[5] D. E. Moriarty and R. Miikkulainen. Hierarchical evolution of neural networks. In *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98TH8360)*, pages 428–433, May 1998.