Partner Report—Neural Architecture Search

Overview

Our group has made substantial progress over the last few weeks towards the goal of understanding & evaluating Differentiable Architecture Search (DARTS) on scientific (image-like) datasets. We have:

1. Created a Docker container and instructions for easily running DARTS on any CUDA-based infrastructure
2. Refactored DARTS code to accept arbitrary inputs, a variety of datasets, and save more information during training (e.g. architecture weights)
3. Set up reference “hand-designed” models (e.g. ResNet) for benchmarking
4. Run multi-day training & evaluation experiments on MNIST, FashionMNIST, CIFAR-10, and Graphene Kirigami
5. Collected results and drawn conclusions for next steps

In this report we will expand on the current status of our work and detail the future steps we plan to take. Going forward, we will seek out more challenging scientific datasets to experiment with.

Current stage

Our preliminary results indicate that DARTS achieves equal performance (0.9225 R^2) when compared to a residual-based network of comparable number of parameters (0.92 R^2) on the fine-grid Graphene Kirigami dataset. However, as expected, DARTS requires significantly more computational resources. We expected this result before we started running DARTS, as when we fit a simple random forest model, we were able to achieve strong results on the (smaller) coarse-grid dataset; additionally, all of our exploratory analysis and modeling efforts (without DARTS) consistently performed near the state-of-the-art. The conclusion we draw from this is that the Graphene dataset admits many relatively simple models that perform well.

We believe that this dataset is just too simple for us to be able to fully evaluate, and gain the benefits from, DARTS. In fact, when we investigate the learned cell, we note that the bulk of the operations are pooling or skip-connections, which have no trainable parameters! This is consistent with the fact that the graphene data is too easy.

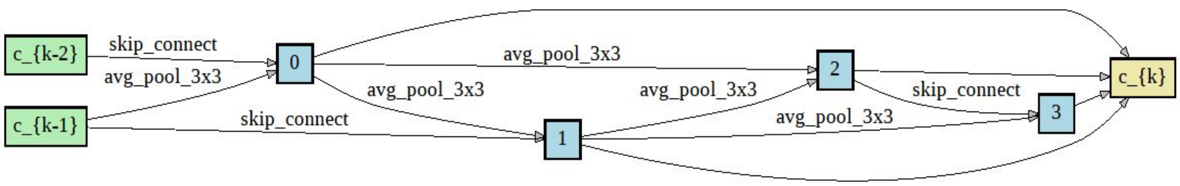

Figure 1: Learned Graphene cell (normal)

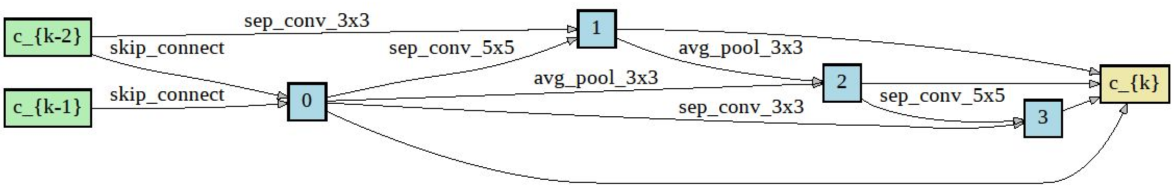

Figure 2: Learned Graphene cell (reduction)

Current Plans

We are interested in exploring more challenging scientific datasets as we have already realized that the Graphene Kirigami dataset may be too simple for full power of the DARTS method to be realized. We are in very good shape for running the full pipeline of DARTS, as well as hand-designed architectures, and should be able to get results quickly.

Moving forward, we will look into other datasets to compare/contrast DARTS with hand-designed architectures. Specifically, we have already thought of the following datasets:

- PLAsTiCC Astronomical Classification (https://www.kaggle.com/c/PLAsTiCC-2018)
- Galaxy Zoo (https://data.galaxyzoo.org/)
- Music Genre Classification (https://github.com/mlachmish/MusicGenreClassification)
- Structural Optimization (https://arxiv.org/pdf/1909.04240.pdf)
- Potentially some radiology/healthcare imaging datasets

We will use the results from above to offer advice and analysis related to architectures discovered, and architecture search in general, for scientific datasets. We also plan to evaluate the *transferability* of learned architectures (e.g. from Galaxy Zoo to Structural Optimization, etc.) and to study the degree of sparsity (in architecture space) induced by DARTS.

As always, we welcome comments, questions, and any other feedback.