Scope of Work—Neural Architecture Search

Background Information

Great progress has been made to automate the process of training neural networks given an architecture. But architectures have historically been selected by human experts in a laborious and error prone process. NAS is a technique for systematically searching a space of candidate network architectures to identify a good one.

Problem statement

We began with a literature review on NAS to educate ourselves on recent progress in the field. We chose to focus on DARTS because it is one of two methods (along with ENAS) that have achieved the best results to date, and is easier to implement and understand (ENAS requires training a separate RNN controller network).

We next run DARTS on a collection of common and scientific datasets. We will compare the results obtained with DARTS to those with state-of-the-art architectures e.g. ResNet, GoogLeNet, VGG, etc. We will run DARTS on as large and diverse a collection of scientific datasets as we can. Ultimately we hope to gain insights on the computational load and prediction performance (and limitations) of DARTS when applied to scientific datasets. We also hope to develop some generally applicable guidelines to practitioners who may wish to try using DARTS on a scientific dataset for which no hand-tuned architectures are known to work well. We plan to write up our findings in a blog post for Towards Data Science.

Resources available

Compute resources: For model training, currently we use the free GPUs on Google Colab or Kaggle, as well as an in-house GPU machine. We will also have 8 AWS p2.xlarge GPU instances for the later half of the semester, so we can train multiple models in parallel.

Data resources: The original project documentation recommended four kinds of scientific dataset: graphene stretching properties, head CT scans, astronomical object classification, and weather classification. Until the mid-term milestone, we have been focusing on the graphene data, and have managed to run both DARTS and hand-designed ResNet on it. We have also downloaded and explored the CT scan data[1]; however, only 500 scans are available for download while the original paper uses 23, 000 scans. The weather dataset is not made available by the author[2]. We plan to explore the astronomical data after the mid-term milestone.

---

[1] via http://headctstudy.qure.ai/dataset

[2] https://github.com/ZebaKhanam91/SP-Weather/issues/1

Expected deliverables

1. An improved DARTS implementation that works on all kinds of datasets
2. A reproducible training & evaluation pipeline with Nvidia-Docker containers
3. Evaluation on scientific datasets:
    - Graphene Kirigami (https://arxiv.org/abs/1808.06111)
    - 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, or other datasets
4. Random search and hand-designed nets for comparison
5. Observations of patterns in discovered cells (including sparsity, distributions of edge weights) and network depths that perform well
6. Advice for practitioners who wish to use DARTS on novel scientific datasets
7. Study the properties of learned architecture transfer to other datasets

Main Repo: https://github.com/capstone2019-neuralsearch/AC297r_2019_NAS
Project Tracking: https://github.com/capstone2019-neuralsearch/AC297r_2019_NAS/projects/1
DARTS Fork: https://github.com/capstone2019-neuralsearch/darts

Timeline

Date	Milestone
10/15 - 10/22	Collect new datasets
10/23 - 11/01	Run experiments (DARTS, hand-designed, random search)
11/02 - 11/10	Visualize and analyze results
11/10 - 11/30	Evaluate learned architecture transferability, compile results into recommendations, use findings to suggest future direction