MagmaDNN and Applications in Materials Science

Daniel Nichols University of Tennessee
Sedrick Keh Hong Kong University of Science and Technology
Kam Fai Chan Chinese University of Hong Kong

Mentor: Dr. Stan Tomov

Presentation Structure

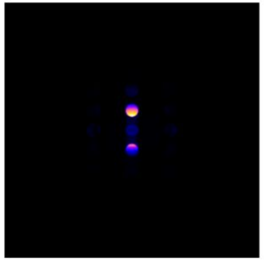
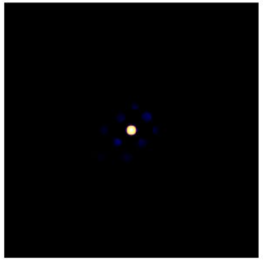
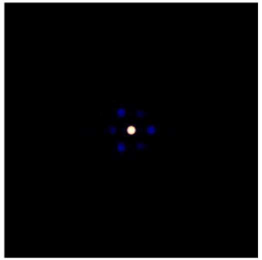
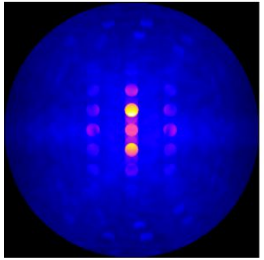
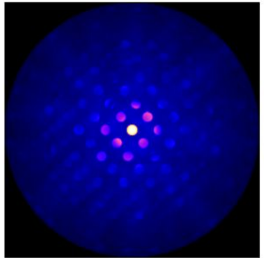
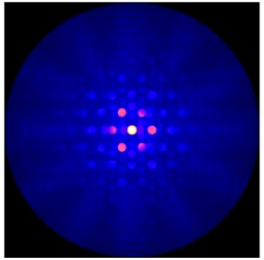
1. Lattice structure classification
2. Energy computation in Ising model
3. MagmaDNN development

Lattice Structure Classification

Lattice Structure Classification

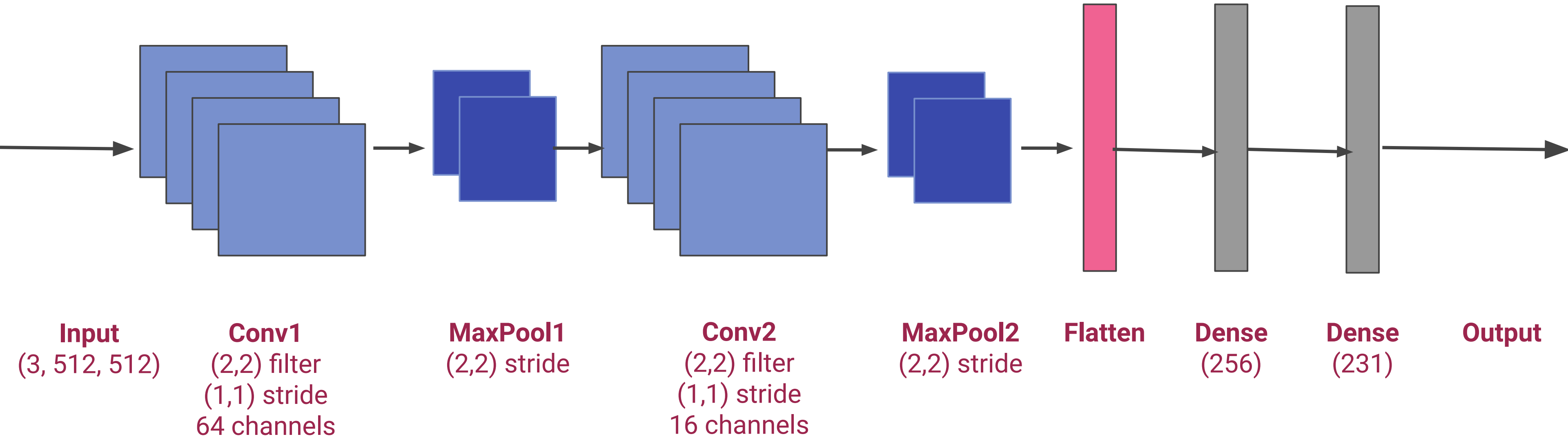
Task description:

- Given (**3** x **512** x **512**) image, we want to classify electron diffraction patterns into 230 distinct crystallographic space groups
- Each channel corresponds to a different material projection
- Also given other features (lattice angles, electron beam energy, d-spacing, etc.)
- Current accuracy is around **0.40**

Lattice Structure Classification

Baseline model:



Lattice Structure Classification

Parameters:

- **Adam** optimizer
- **0.001** learning rate
- Batch size **8**
- **5** epochs

Problems with baseline model:

- Memory limit
- Low accuracy
- Slow training

Evaluation:

- **Accuracy**: 0.11
- **Time**: Start to finish is around 3 minutes

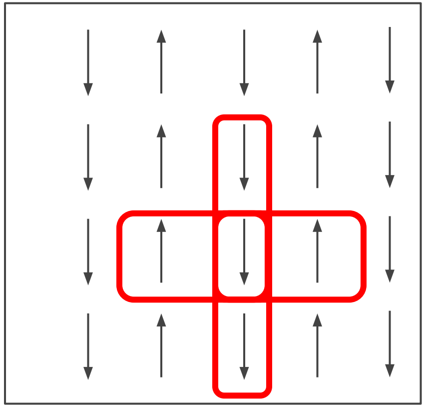
Lattice Structure Classification

Solutions (to do):

- MagmaDNN (try ResNet 50)
- Parallelism

Problems related to Ising Model

Energy in Ising Model



Neighbour spins

Specific heat, magnetic susceptibility and other quantities are related to (the expected value of) energy

$$H = -\sum_{<i,j>\text{neighbour}} J_{ij}\sigma_i\sigma_j - \mu\sum h_i\sigma_i$$

Energy in Ising Model

Problem:

- Want to simulate real-life alloy with Ising model
- Need to compute energy for lots of Ising configurations
- Current method is slow (with large number of configurations)
- Want a fast method

Solution:

- Use machine learning based model
- Use MagmaDNN for optimal performance
- Parallelism?

Energy in Ising Model Progress

- Done:
    - Generated data
    - Have a simple model prototype in tensorflow/keras
- To do:
    - Check if overfitted
    - Implement in MagmaDNN
    - Construct a more general model (currently 2D 8x8 grid with periodic boundary and uniform interaction)
    - Transfer to 3D model

Towards 3D lattice structure

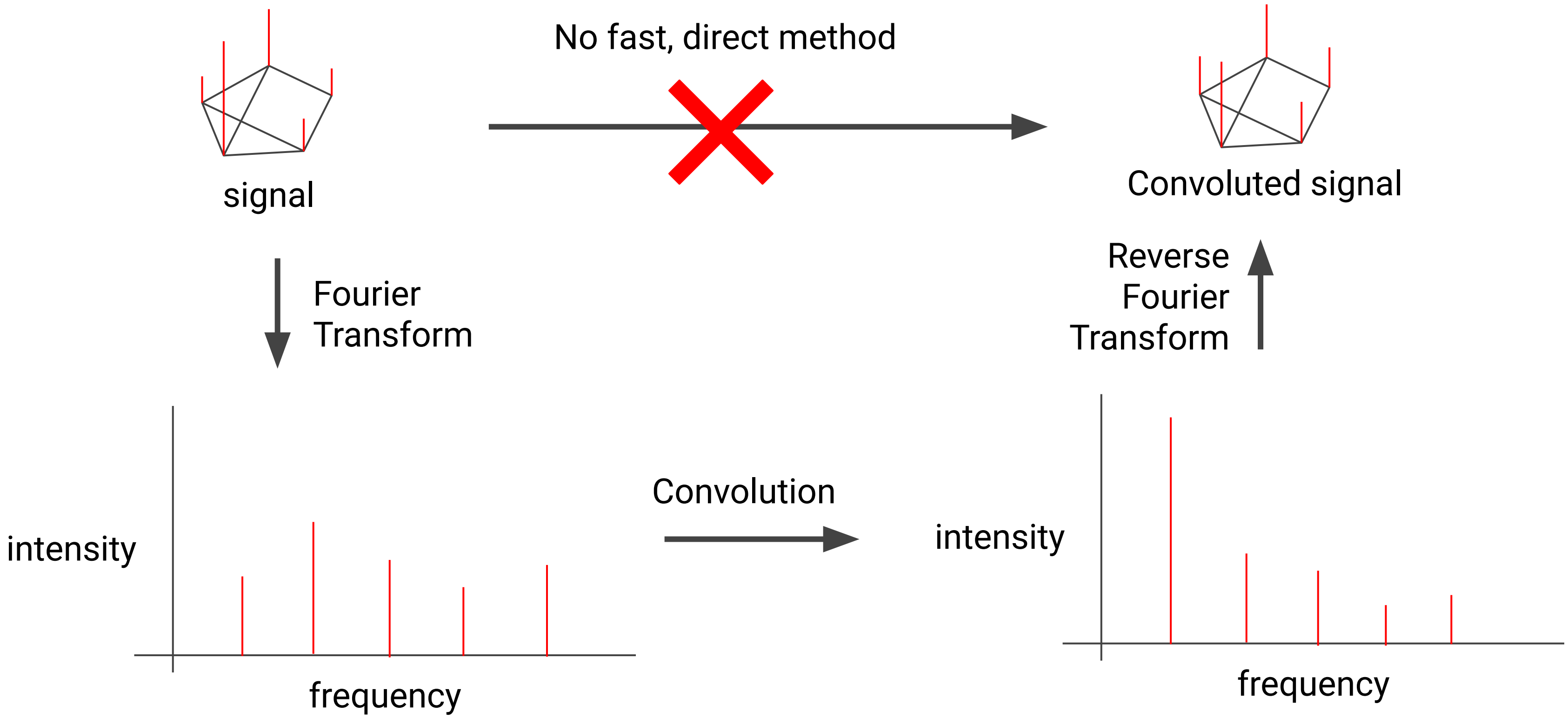
Problem:

- Need to do CNN on 3D lattice structure
- No Euclidean structure (grid-like)
- No natural concept of “neighbourhood”
- Want high flexibility on defining lattice structure
- Want fast method to do convolution

Solution:

- Use Graph Convolution Network (GCN)
- Implement in MagmaDNN

Graph Convolutional Network



Graph Convolutional Network Progress

- GCN already exists in some extension library for PyTorch[1], but there are
  - More than 23 schemes for GCN
  - More than 14 schemes for graph coarsening on GCN
- To do:
  - Test different schemes
  - Implement in MagmaDNN

[1] M. Fey. PyTorch Geometric Documentation. https://pytorch-geometric.readthedocs.io/en/latest/modules/nn.html

MagmaDNN Core Development

MagmaDNN Core Problems and Proposed Solutions

Problems:

- Optimal training speed (fast, fast, fast…)
- Lattice convolutions
- Optimized distributed training
- Measure performance on a state of the art model
- Keep MagmaDNN up to date and bug free so that it can support a range of projects

Solutions:

- Lattice Convolutions
- Compute graph optimizations
- Explore distributed training techniques using Nvidia NCCL and provide a modular workflow for training in parallel
- ResNet 50 test
- Conduct heavy use of MagmaDNN (on microscopy and monte carlo problems) to expose bugs and speed issues

Lattice Convolutions Progress

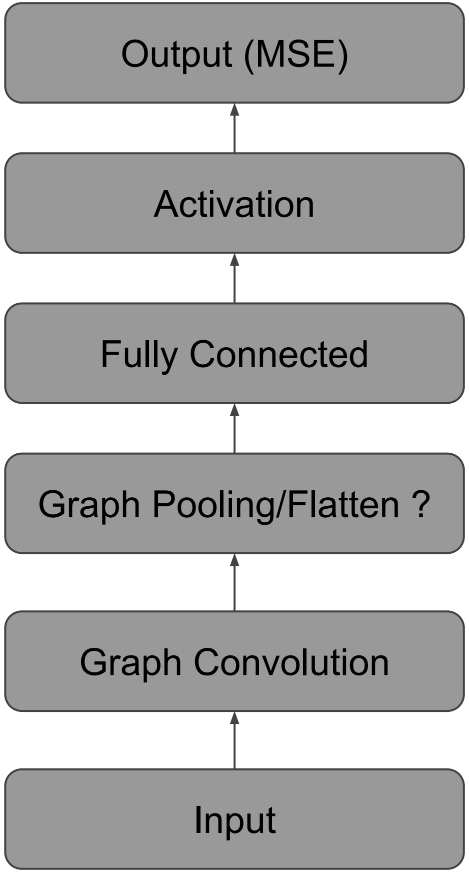
Graph convolutions can be performed using

$$H^{(l+1)} = \sigma\left(\widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right)$$

where H(L) is the output of layer L [1].

Computational benefits

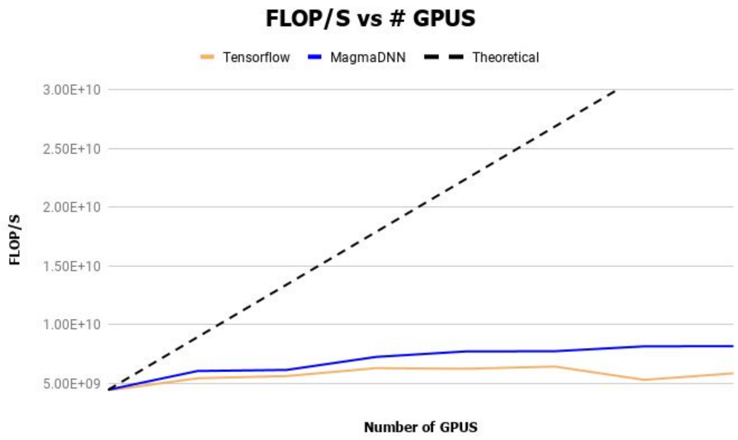
- 1 gemm and 1 spmm
- Magma and CuSparse support spmm

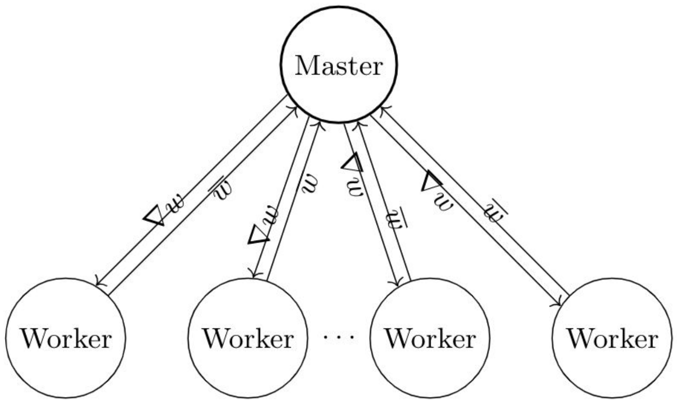


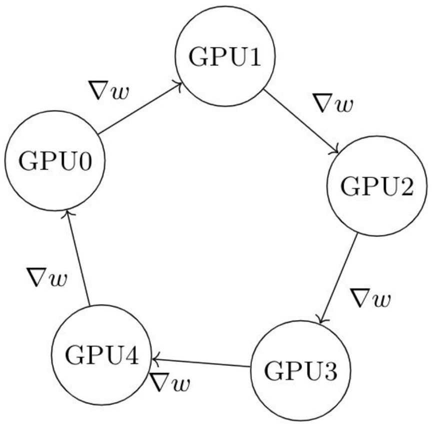
[1] T. N. Kipf and M. Welling, “Semi-Supervised Classification with Graph Convolutional Networks,” *arXiv:1609.02907 [cs, stat]*, Sep. 2016.

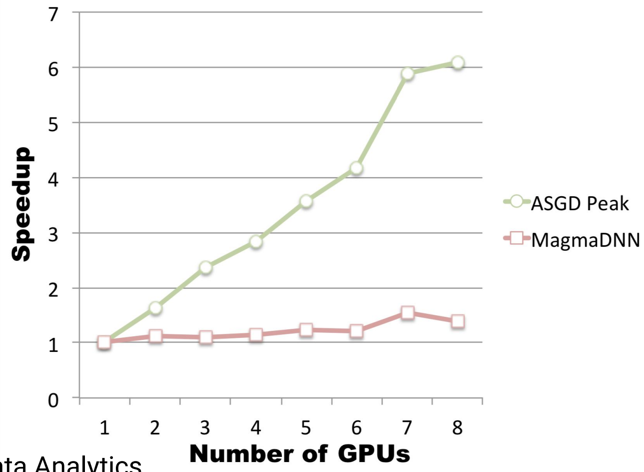
Distributed Training Progress

- Better speedup than tensorflow (1.7) distributed training
- Simple AllReduce method [1]
- Ring AllReduce
- Port from OpenMPI to NCCL





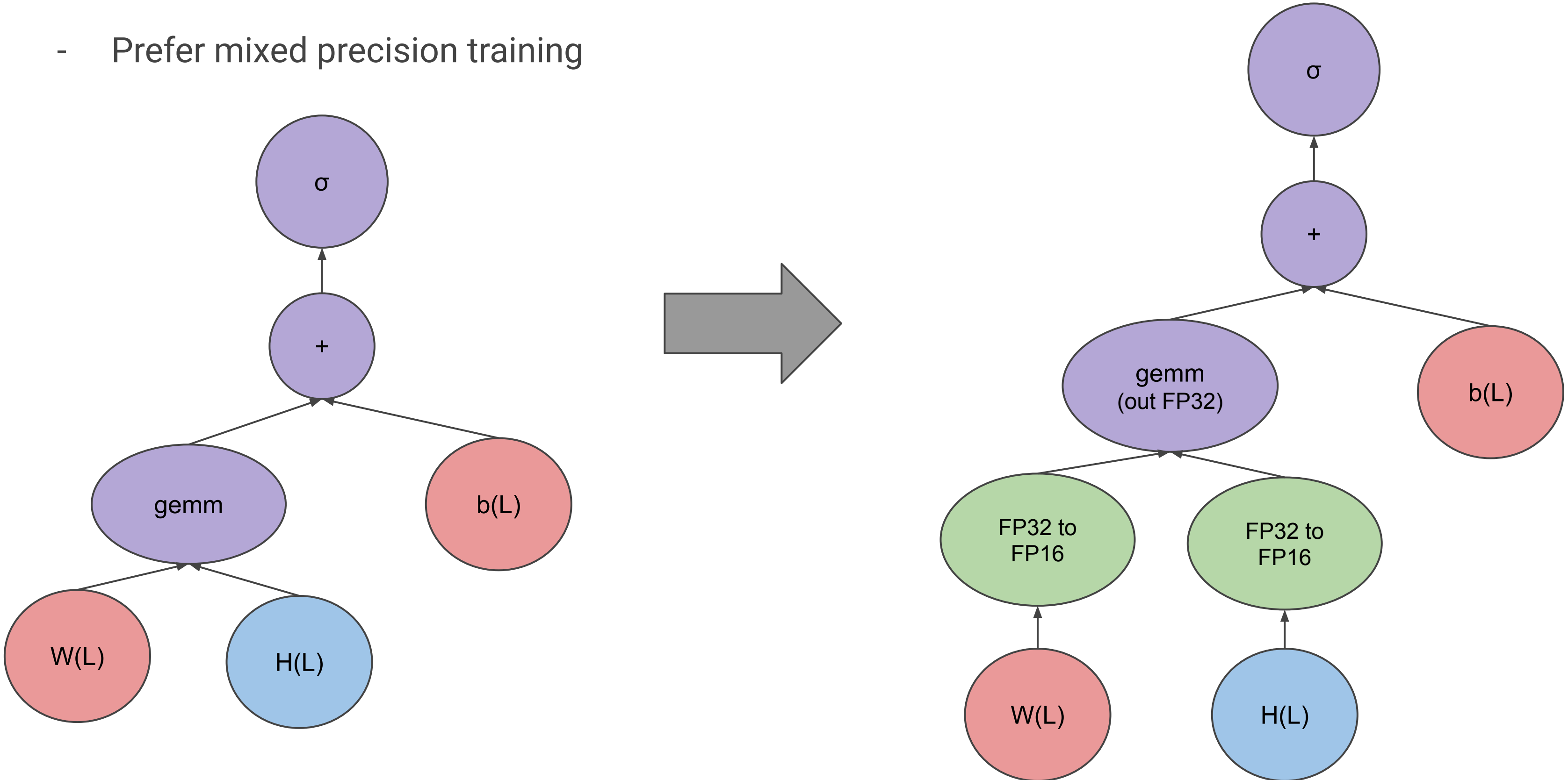




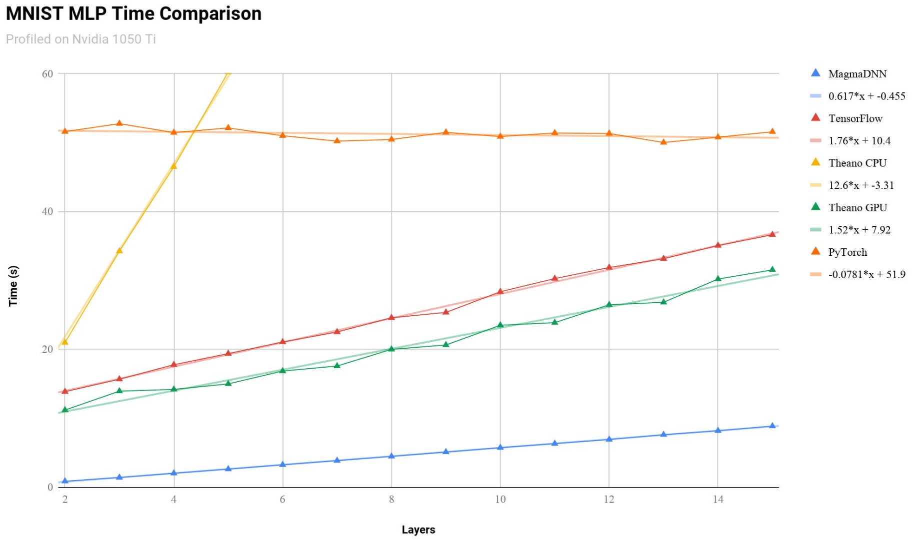
[1] Nichols N. Tomov Betancourt S. Tomov Wong Dongarra, “MagmaDNN: Towards High-Performance Data Analytics and Machine Learning for Data-Driven Scientific Computing” . ISC, June 2019.

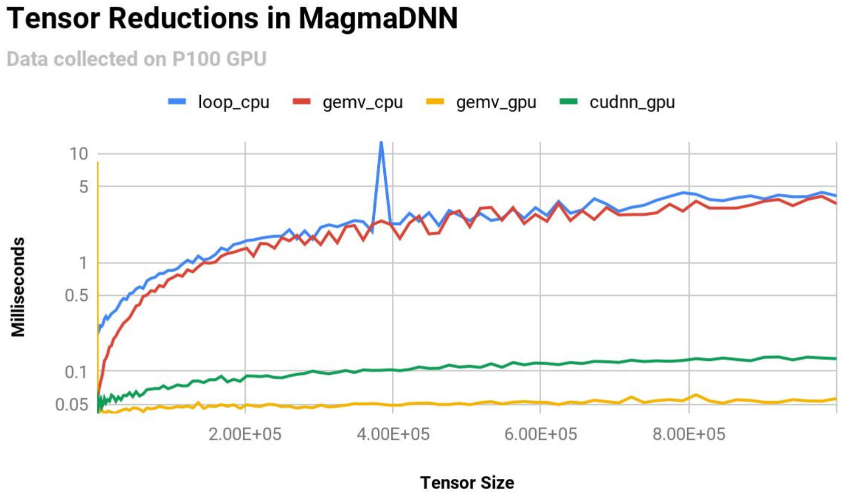
Compute Graph Optimizations Progress

- Prefer mixed precision training



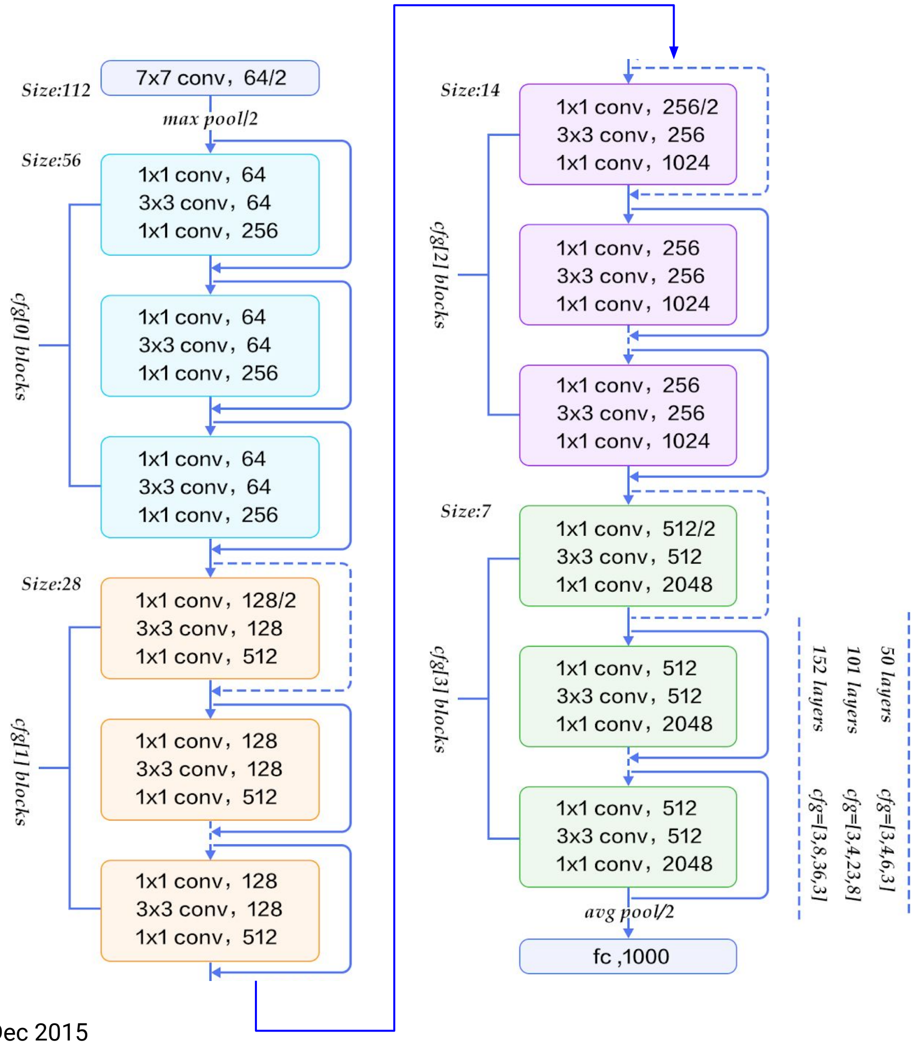
MagmaDNN speed progress





ResNet50[1] Progress

- Convolutions
- Careful memory management
- Model Parallelism
- Todo
    - Residual Layers
    - Custom gradients with residual layers



[1] He Zhang Ren Sun, “Deep Residual Learning for Image Recognition” arxiv:1512.03385, Dec 2015

Logo Contest

MagmaDNN needs a logo.

Magma Logo

Current “Ideas”