Implementing a U-Net Architecture in MagmaDNN

Students: Chow Tsz Ching, Edward Karak, Spencer Smith
Chinese Univ. of Hong Kong, Baruch College at City Univ. of New York, Univ. of North Texas
Mentors: Dr. Kwai Wong, Dr. Stan Tomov

Abstract
MagmaDNN is an open-source deep-learning library written in C++. It is based on Magma, a linear algebra package and is designed to handle supervised problems. MagmaDNN is unique in that it is tailored for parallel computing and, consequently, supercomputing applications.

A U-Net is a convolutional neural network developed originally for biomedical image segmentation to detect tumors. It can be defined in terms of down-sampling and up-sampling layers. Our U-Net implementation is called semantic segmentation. It aims to learn the classifications of individual pixels in an image.

Introduction
• A U-Net is a convolutional neural network developed originally for biomedical image segmentation to detect tumors. It can be defined in terms of down-sampling and up-sampling layers. Our U-Net implementation is called semantic segmentation. It aims to learn the individual pixels in an image.

• Similar to PyTorch and Keras, MagmaDNN is a machine learning package. It is still in the development phase, so it is very limited in its scope. Take the loss function as an example, MagmaDNN only supports categorical cross-entropy loss and MSE. MagmaDNN can only do classification but not segmentation and hence the Output Layer of a neural network must be a flattened two-dimensional tensor or else errors will occur. Therefore, the main task of our research is to implement segmentation in MagmaDNN.

Methodology

Data
We will train our neural network with the Oxford-IIIT Pet Dataset. The dataset contains around 7000 cats and dogs classified by breed. Each image has its corresponding masking. The masking is in trimap format. Trimap is in three colors only: 1 for foreground, 2 for background and 3 for not classified.

Data Extraction
MagmaDNN was only able to input MNIST, CIFAR10, CIFAR100 data and one-hot encoded ground-truth data. Moreover, MagmaDNN is only able do classification instead of image segmentation. Therefore, we need to use a custom dataset for the training and testing for the U-Net. After, we can integrate OpenCV with MagmaDNN so it can input data from ImageNet and Oxford-IIIT Pet Dataset.

Model
We have adopted convolution transpose instead of using bilinear interpolation for upsampling in the U-Net. Convolution transpose is better than bilinear interpolation because convolution transpose will learn when it is training. However, up-sampling using bilinear interpolation will consume less resources. Still, we still adopt the convolution transpose because it overperform bilinear interpolation theoretically.

Results

Next Steps
• Finish the implementation of U-Net
• HDF5 implementation in MagmaDNN
• ResU-Net implementation

Acknowledgements/References
This project was sponsored by the National Science Foundation through Research Experience for Undergraduates (REU) award, with additional support from the Joint Institute for Computational Sciences at University of Tennessee Knoxville. This project used allocations from the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation. In addition, the computing work was also performed on technical workstations donated by the BP High Performance Computing Team.