

CDA5125 Programming Assignment No. 5: CUDA C Deep Neural Network Code (optional)

(Due: April 22)

Purpose:

- Practice GPU programming with CUDA C

Statement of work:

This assignment is optional.

This is a group assignment. Each group can have 2 people. In this assignment, you will develop CUDA C Deep Neural Network program based on the deep neural network code that you developed in Assignment 1.

Due dates:

The assignment is due on April 22, 11:59pm. Put all related source code, the makefile, and a README file in a tar file and submit the tar file. In the README file, you must describe (1) how to compile and run the program, (2) whether and how the GPU program achieves high accuracy (GPU code should be correct to get any point in this assignment), and (3) report the speedups over the baseline code. What is described in the README file must be repeatable with your submitted files.

Grading:

1. Submission has all components (all related source code, makefile, README file); the executable can be successfully produced with a 'make' command in the directory; a deep neural network for handwriting digit recognition with the MNIST dataset is built; the GPU program must generate correct results (3 points).
2. The README file describes the information as required and the program is correct (1 points).
3. The GPU program is at least 1.5 times faster than the baseline code (2 points).