



## Final Project

Shenghou Ma

All Sections

All projects require Verilog implementation for the module and testbench. In your report, you will need to show:

- (1) analysis, block diagrams and test methodology for the design
- (2) comparison of potential designs/architecture and rationale why a particular architecture is chosen
- (2) ASIC synthesis results (power, delay and area) with FreePDK45 library
- (3) other info required by each individual project.

All of the project require you implement some datapath element of an ANN accelerator, you can disregard the control part of your circuit and focus only on the data path. However, in your testbench, you do need to write behavior code for the controller to drive your datapath to accept data from outside to achieve what the goal of the project.

Write a six page report (in **IEEE conference format**

(<https://www.ieee.org/conferences/publishing/templates.html>), including references, if necessary, you can have more pages, but keep it less than 10 pages) and prepare a 30 minute presentation about your project. The final deliverable are the pdf report, presentation slides and all Verilog, necessary data and scripts files.

Due date: the presentation is due May 12 (4:10PM - 6:50PM), and the report and all other collateral files due that night.

Marwan:

implement a scalable unsigned fixed-point multiply-accumulate unit that supports  $16b + 8b \times 8b$ ,  $16b + 4b \times 4b$  and  $16b + 2b \times 2b$ ; also implement a testbench for your MAC unit.

Nithyaa:

Google found a better activation function called **Swish** (<https://arxiv.org/pdf/1710.05941.pdf>), based on what you learned from the paper, implement it using the best approach for 5-bit input and 8-bit output,  $x$  limited to  $[-5, 5]$ , and experiment with the various values of  $\beta$  and report how the choice of parameter affects the delay/area/speed of your circuit.

Chandler:

Implement a 6x6 systolic array based 2D convolution neural network for 3x3 kernel size using the row stationary form (see Eyeriss paper). You can use 8-bit for the weight and input activation and 24-bit accumulator for the output. Only the PE array itself needs to be implemented (i.e. control logic can be left out), but you do need to create a testbench to reformat the data to drive the PE array.

Varsha:

Implement a 5x5 sparse matrix matrix multiplication accelerator. Only the PE array itself needs to be implemented (i.e. control logic can be left out), but you do need to create a testbench to reformat the data to drive the PE array.

Please let me know if you have any questions.

Please note that we intentionally left open some details of the project so that you have some flexibility. Please document your assumptions in the final presentation & report. If you are unsure about some unspecified details, you are free to ask but as this is a graduate level course, you should spend some time research about the problem before asking.

Thanks.

Minux

