

Project Proposal Report: CNN Accelerator Using U250 FPGA

1. Student Information and Team-Up Information

This is a single-person project. The student responsible for this project is Zhenyu Huang (zhenyu_huang@brown.edu).

2. Project Problem Scope

The project involves designing and optimizing a convolutional neural network (CNN) accelerator on a Xilinx U250 FPGA platform. Specifically, the tasks include:

- **Version 0:** Establishing a baseline for the CNN accelerator without modifications to the kernel code (`cnn.cpp`).
- **Version 1:** Implementing optimizations such as computation unrolling and pipelining in the kernel code to enhance performance.

The primary focus is on improving the dataflow of CNN computations and applying code optimization techniques within the given constraints.

3. Expected Delivery

The project's evaluation will include:

- **Correctness:** Both `v0` and `v1` should produce correct results on the FPGA at 200 MHz.
- **Performance:** The optimized `v1` version should achieve a performance that is at least 400 times faster than the baseline `v0`.
- **Reports:** Submission of resource utilization, timing analysis, and performance metrics for both `v0` and `v1` designs. This includes:
 - Kernel execution time
 - Throughput in GOPS
 - Resource utilization (LUT, BRAM, DSP, etc.)
 - Frequency and worst negative slack (WNS)

4. Background and Related Work

- **Tools Used:**
 - **Vivado HLS** for high-level synthesis and hardware optimization.
 - **Vivado GUI** for generating resource reports and visualizing chip layouts.
 - **Provided Tools:** The project source code includes `host.cpp` for host-side operations and `cnn.cpp` for HLS kernel operations.

5. Supplement