EE/CprE/SE 492

HAML: Heterogeneous and Accelerated Computing for Machine Learning

Semester 2 Week 7-8 Report

10/4/24 - 10/17/24

Faculty Advisor   Phillip Jones
Client   JR Spidell

Team Members:

Jonathan Tan   - Memory Affinity, Kria Board Manager

Josh Czarniak   - DPU Control Developer

Justin Wenzel   - Multi-threaded Developer

Kai Heng Gan   - Image Processing/Semantic Segmentation Developer

Santiago Campoverde   - Model Analytics

Summary for Progress These Two Weeks

During these two weeks, we worked towards "Milestone 3." That is, mainly, creating the multithreaded approach for our Blink, Eye-tracking, and semantic segmentation inference. We are also working on setting up the Vitis-AI profiling tool for performance analysis. Finally, we worked on optimizing the semantic segmentation model by modifying negative slope value of Leaky ReLu Activation function to ensure that DPU could support the converted semantic segmentation model.

These Two Weeks' Individual Contributions

- Justin
  - Began implementing eye track function in the model class format as defined in the multi-threaded application allowing the threads to interact with the model and make inferences, by calling the preprocessing, execution and postprocessing for the model.
  - Met with Josh and discussed the DPU handler code implementation
    - Discussed functions to be implemented and how the threads will interact with the DPU handler class
  - Worked with Kai about implementing semantic segmentation into the model class format allowing the threads to interact with the model and perform inferences by calling the functions implemented for it.
  - Researched into thread affinity within the pthread.h library function to isolate threads on different cores in our implementation.
    - Created a presentation to discuss thread affinity, and present our current implementation to ensure it is being implemented properly
- Jonathan
  - Debug error when using vaitrace, currently, profiling is not working due to a divide by zero error, which I believe is caused by the profiling program (vaitrace) failing to detect events.

- Things tried (since last report):
  - Learned that there are two ways of creating the DPU, the Vivado flow, and the Vitis flow. We are using the Vivado flow, and the Vitis-AI version we are using is v3.0. There is a bug with Vivado flow in Vitis-AI v3.0, according to [this thread](#).
- Moving forward:
  - Try updating to Vitis-AI v3.5.
  - If that doesn't work out, try creating the DPU using the Vitis flow (no bug as far as I know).
  - ○
- Josh
  - Met with Justin and Jonathan to discuss the DPU handler code implementation
    - Discussed the implementation of the new function and how the threads will be interacting with the DPU handler class
  - Worked on DPU handler code
- Kai
  - Resolved xmodel compilation errors
    - Re-trained the SS model with negative slope value of 0.1015625 in LeakyReLu activation function
    - Tested and validated the re-trained model and got the accuracy of 85.92% at epoch 3
  - Worked with Justin about organizing semantic segmentation function in the model class format as defined in the multi-threaded application allowing the threads to interact with the model and make inferences, by calling the preprocessing, execution and postprocessing for the model.
  - Programmed the semantic segmentation cpp in the model class format.
  - Sent the result of re-trained model to Santi for measuring the accuracy.
    - It would be beneficial to compare the accuracy reported during the training process with the accuracy obtained from Santi's script. This comparison will help determine any discrepancies or improvements in the accuracy measurements.
- Santiago
  - Created script for calculating Mean Intersection over Union for the Semantic Segmentation model.
    - Works but needs a few small changes to handle the output format.
  - Created presentation on Mean Intersection over Union with the script done so far.

Team Member	These Two Weeks' Task	Completion Date	Hours Took	These Two Weeks' Hours	Total Project Hours
	Attended meetings	NA	3	9.5	127

Name	Task	Date	Hours	Weekly Total	Cumulative Total
Justin Wenzel	Began implementing eye track function in the model class format as defined in the multi-threaded application	10/13	3		
	Thread Affinity research and presentation	10/8	2		
	DPU handler meeting with Josh	10/8	1		
	Semantic segmentation model into the model class format with Kai	10/9	0.5		
Jonathan Tan	Attended meetings	NA	6	11	148.5
	Debug error when running profiler on the board (vaitrace)	On-going	5		
Josh Czarniak	Attended meetings	NA	1	6	106
	Met with Justin and Jonathan for the DPU handler code	10/12	3		
	Worked on DPU code	10/10	2		
Kai Heng Gan	Attended meetings	NA	5	15.5	151
	Resolved xmodel compilation errors	10/15	8		
	Worked with Justin on coding the semantic segmentation model into the model class format	10/9	0.5		
	Programmed the semantic segmentation cpp in the model class format.	10/15	1		
	Sent the result of re-trained model to Santi for measuring the accuracy.	10/16	1		
Santiago Campoverde	Attended meetings	NA	2	6	94
	Created script for calculating Mean Intersection over Union for the Semantic Segmentation model.	10/14	3		
	Created presentation on Mean Intersection over Union with the script done so far.	10/15	1		

Note: 1. This is per week hours, Σ "hours taken" = "week hours". 2. Due to multiple meeting times, meetings' "completion date" are "NA".

Plans for Coming Two Weeks

Team Member	Plans for Coming Week	Planned Completion	Planned Hours Required
Justin Wenzel	Finish implementing eye tracking model to match the model class interface. Threads will be able to use the functions to interact with the model and perform inferences	10/20	4
	Test each threads functionality on the board with its ML model. Following an individual execution/testing approach ensuring models and threads perform in the application as expected.	10/24	4
	Begin Profiling critical sections, e.g. using gettimeofday() to start	10/27	3
	Build a diagram conveying the multi-threaded implementation and flow of execution of each thread	10/22	2
Jonathan Tan	Continue debug error with vaitrace	9/25	5
	Update Vitis-AI to v3.5	10/19	1
	(if vai 3.5 doesn't) Change to Vitis flow	10/24	6
Josh Czarniak	Adjust slide deck on DPU implementation based on information learned in the meetings	10/19	1
	Finalize implementation of DPU code	10/19	5
Kai Heng Gan	Continue working and testing on preprocessing cpp code that will run on the Kria KV260. Resolve the invalid output from the xmodel.	ongoing	5
	Validate the quantized Pytorch model	ongoing	5
	Continue working with Justin to implement semantic segmentation in the model class format defined in the multi-threaded application to interact with the semantic segmentation model, and perform inference with the DPU	NA	3
Santiago Campoverde	Adding additional tweaks to the segmentation testing script to handle the output properly	10/19	1
	Provide visualized interpretations of accuracy to the script	10/20	2
	Begin working on Blink Algorithm testing	10/22	2