EE/CprE/SE 491

HAML: Heterogeneous Computing for Machine Learning Algorithms

Week 7 Report

2/27/24 - 3/4/24

Faculty Advisor   : Phillip Jones

Client   : JR Spidell

_____

Team Members:

Jonathan Tan   - DPU Management, Kria Board Manager

Josh Czarniak   - Pupil Center Location Algorithm

Justin Wenzel   - Blink/No-Blink Algorithm, Meeting Leader of the Week

Kai Heng Gan   - OpenCV

Santiago Campoverde   - Data Profile/Model Analytics

Summary for Progress This Week

This week's objectives were to continue working with our advisor, client, and ETG on setting up a workstation. We decided that the original set up with RHEL8 will not work for our purposes because we have Docker images and file types that isn't directly supported by the OS. Hence, we decided to go with a Windows workstation. To better understand how our environment will look like while waiting for ETG to set up our workstation, Justin prototyped the environment on his personal PC.

Moreover, after meeting with previous team member (Rudy), our team got a better understanding of how the project works and got a better understanding of how each of our tasks should be completed. Each of us did research to further our understanding in our respective areas.

This Week's Individual Contributions

- Justin
  - Set up Docker image/container environment to better the groups program and work development.
    - Docker is normally a terminal based application.
    - Using X11 forwarding protocol with X server software to display graphical applications on the host machine through the X server the host machine is running.
    - Packing containers into Docker images for easy compatibility and usability between team.
  - Researched bounce diagrams, code mappings, and stack diagrams to outline our program structure and layout flow of libraries/utilities through the software and hardware components.
- Jonathan
  - Continue to communicate with ETG to get workstation set up in Coover 1301.
  - Read documentations on VART and Vitis AI to better understand our program flow, and while doing that I came up with this plan to get progress started:
    - As a proof of concept, we should create a naïve solution to put everything (pre-processing, blink detection, eye-tracking detection, profiling) into one program flow:

      ```
      main() --> preprocessing() --> blink() --> track() --> analysis()
      ```

    - The idea is to, **by the end of the week of 4/8**, our team should be able to pass **one** frame through the entire flow. I discussed the plan and our team agreed during team meeting on 3/4/2024 that this will help us move forward towards our milestone.
    - Doing so will allow each team member to identify areas of confusion in our respective task; and allow the team to identify overheads in the flow (the entire system).
  - Because there're multiple models that shares the DPU resource, I looked into how the DPU could be shared, I found that the part of the models that run the model on the DPU is:

    ```
    auto runner = vart::Runner::create_runner(subgraph, "run");
    auto job_data = runner->execute_async(inputs, outputs);
    runner->wait(job_data.first, -1);
    ```

    Knowing this, one of my straightforward idea is to create a function, let's call it `use_dpu(args)` in this example. This function will be the only code that can access the DPU directly, any algorithms and requires the DPU will call this function instead, and all the resource management and sharing will be done in this function.
  - Reorganized files on the Kria board, there was a lot of temporary files and backup files, I cleaned it up and pushed the code onto GitHub so team members have easy access to it.
  - (On-going) Do testing on how to set up coding environment in Vitis, the previous team coded directly on the board using Vim. If I could get a Vitis development environment set up, we will have access to a debugger, which I believe will greatly benefit our workflow.
- Josh
  - Researched further on pupil tracking
    - Looked up online and from previous teams code how it can be done
  - Studied previous team's code

- What type of model that they trained as well as how they trained that model
- They used a TensorFlow model and turned it into an xmodel using XIR
  - o Worked on multi-threading slides and presentation for the team
    - Multi-threading in C++
    - Went over join methods similar to that of pthreads.
- Kai
  - o Researched on the segmentation technique on video/image processing
  - o Discussed about image processing with the client
    - Understand the ML algorithm on segmentation mechanism
- Santiago
  - o Researched Vitis AI Profiler VAITRACE commands
  - o Researched Vitis AI Performance Profiling and Methods using GMIO with AIE
  - o Read through Yocto tracing and profiling documentation

Team Member	This Week's Task	Completion Date	Hours Took	This Week's Hours	Total Project Hours
Justin Wenzel	Attended meetings	NA	3	9	30
	Set up Docker environment involving an interactive desktop environment.	3/2	5		
	Researched into bounce diagrams/code mapping for VART.	2/26	1		
Jonathan Tan	Attended meetings	NA	3	12	34
	(On going) Communication w/ ETG and setting up env on workstation.	2/26	~0		
	Research and preparing for proposal	3/3	2		
	Look into ways to share DPU resource	3/3	1		
	Mess with Kria board to gain familiarity with it.	3/3	6		
Josh Czarniak	Attended meetings	NA	3	10	37
	Looked at previous team's code	3/3	4		
	Work on multi-threading slides	3/2	3		
Kai Heng Gan	Attended meetings (Client, Team, Advisor, and 1:1 meetings)	NA	3.5	7.5	31.5
	Researched on the segmentation technique on video/image processing	Ongoing	4		
Santiago Campoverde	Attended meetings	NA	3	8	30
	Researched Vitis AI Profiler VAITRACE commands	2/28	1		

	Researched Vitis AI Performance Profiling and Methods using GMIO with AIE	2/29	2		
	Read through Yocto tracing and profiling documentation	2/29	2		

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 Week

Team Member	Plans for Coming Week	Planned Completion	Planned Hours Required
Justin Wenzel	Complete Docker setup by adding multiple tools to assist in development and ability to access host machine devices through container.	3/8	4
	Begin implementing bounce diagrams, code mappings, and stack diagrams of program structure.	3/10	2
	Lead meetings and create meeting agenda for client and team meetings.	3/8	1
Jonathan Tan	Learn how to compile h5/pb model files into xmodel. (side quest, low priority)	3/29	5
	Assist Justin in figuring out the best way to set up our team's development environment	3/18	1
	Make presentation about my findings on sharing DPU	3/15	2
	Continue messing with Kria board to understand how to set up Vitis development environment for the board	3/18	3
Josh Czarniak	Understand more about machine learning	3/8	5
	Continue work on multi-threading slides	3/6	3
	Understand more of the previous Teams code	3/7	3
Kai Heng Gan	Continue research on the image segmentation technique. Run some test on the image segmentation ML algorithm that discussed with the client.	3/16	5
	Work on taking some images into image pre-processing process and analyze the result	3/16	4
Santiago Campoverde	Getting a Linux environment to test profiling and tracing environments.	8/1	4
	Deep dive into GMIO ports.	7/1	2
	Find if previous build of the environment has necessary profiling tools.	7/1	2