Choose the Right Hardware

Proposal Template

Scenario 1: Manufacturing

Client Requirements and Potential Hardware Solution

Look through the scenario and find any relevant client requirements. Then, suggest a potential hardware type and explain how this hardware would satisfy each of the requirements.

Which hardware might be most appropriate for this scenario? (CPU / IGPU / VPU / FPGA)
FPGA

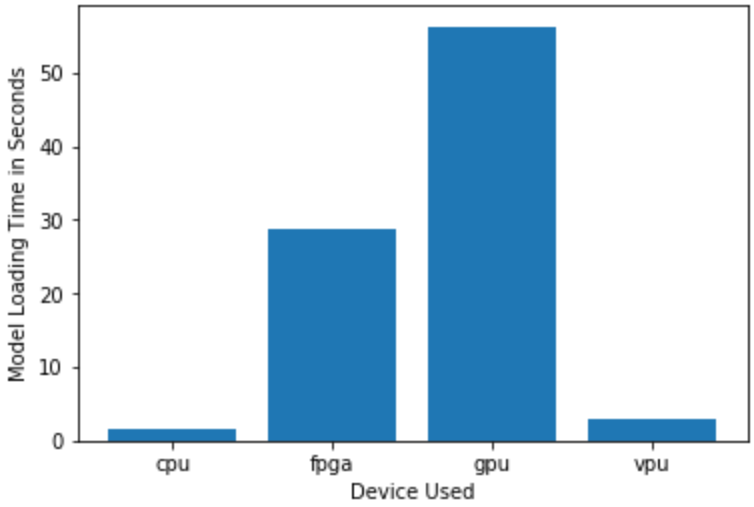
Requirement Observed (Include at least two.)	**How does the chosen hardware meet this requirement?**
Example requirement: The client requires a tiny device to be connected to their CPU—and their budget is only about $100 for each device.	*Example explanation:* VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
Image processing task to be completed five times per second.	*Inference time is important for this job and cpu, gpu, and fpga will likely be the main choice*
to be able to repurpose the system to address a second issue.	*This one I think is beyond the scope of the study with People queues, but I think fpga being programmable is also able to handle this task.*
Each camera records video at 30-35 FPS (Frames Per Second)	*How much a camera can record means how much data we have for inference*

Queue Monitoring Requirements

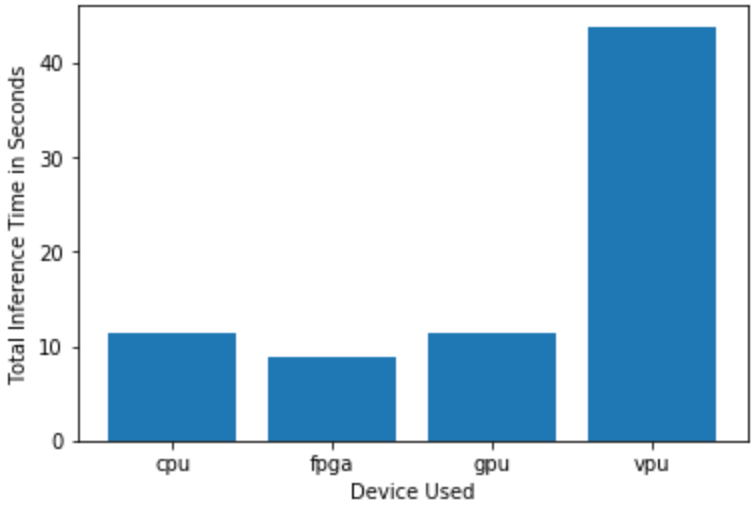
Maximum number of people in the queue	*2*
Model precision chosen (FP32, FP16, or Int8)	*FP16*

Test Results

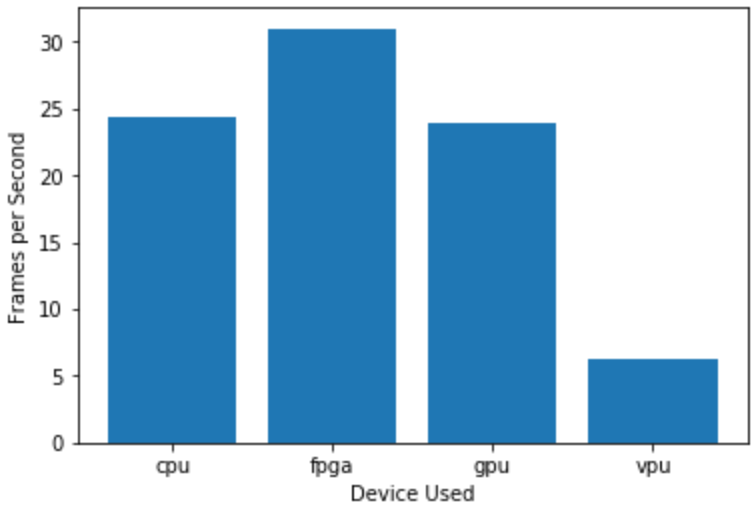
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



Model Load Time



Inference Time



FPS

Final Hardware Recommendation

Now synthesize your points from above and provide a brief write-up describing why the chosen hardware is the best choice for this scenario. Be sure to discuss the client's requirements, the test results, and how these relate to one another (e.g., perhaps one of the devices performed better than the rest, but does not meet one of the client's requirements).

Write-up: Final Hardware Recommendation
FPGA is what I recommend as the task has specific window time , and that requirement of 30-35 fps from an existing video capturing device is given. Repurposing the system is also a requirement. CPU and GPU are also candidates but false short on utilizing the FPS the camera can handle based on the graphs.

Scenario 2: Retail

Client Requirements and Potential Hardware Solution

Look through the scenario and find any relevant client requirements. Then, suggest a potential hardware type and explain how this hardware would satisfy each of the requirements.

Which hardware might be most appropriate for this scenario? (CPU / IGPU / VPU / FPGA)
CPU

Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
Example requirement: The client requires a tiny device to be connected to their CPU—and their budget is only about $100 for each device.	*Example explanation:* VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
Mr. Lin does not have much money to invest in additional hardware, and also would like to save as much as possible on his electric bill.	*It was stated that there are lightly loaded i7 machines on each queue, we can use those.*
Most of the store's checkout counters already have a modern computer, each of which has an Intel i7 core processor. Currently these processors are only used to carry out some minimal tasks that are not computationally expensive.	*Again, we can use the machines already existing on this scenario. I think the scenario doesn't require a top of the line machine that has the fastest inference as we are counting people on a queue in a retail store. It was given on the report that customers usually wait 230-400 seconds which I think would best describe that this task allows use of CPU.*

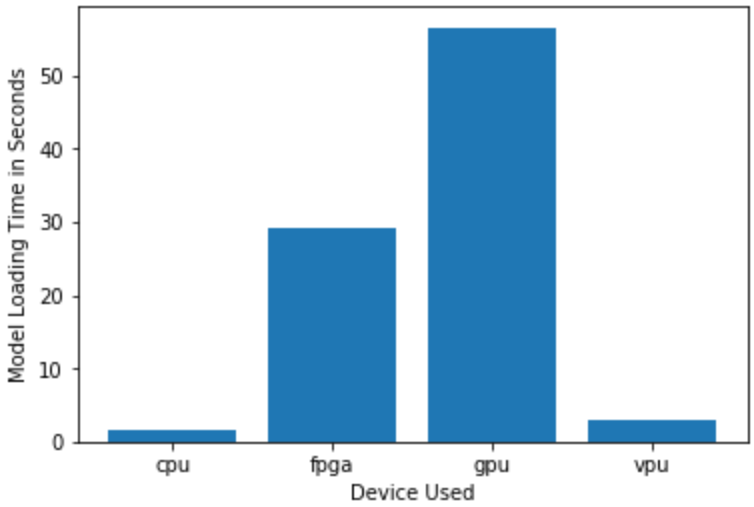
The total number of people in the checkout queue ranges from an average of 2 per queue (during normal daily hours) to 5 per queue (during rush hours).	*This is something a CPU solution can handle.*

Queue Monitoring Requirements

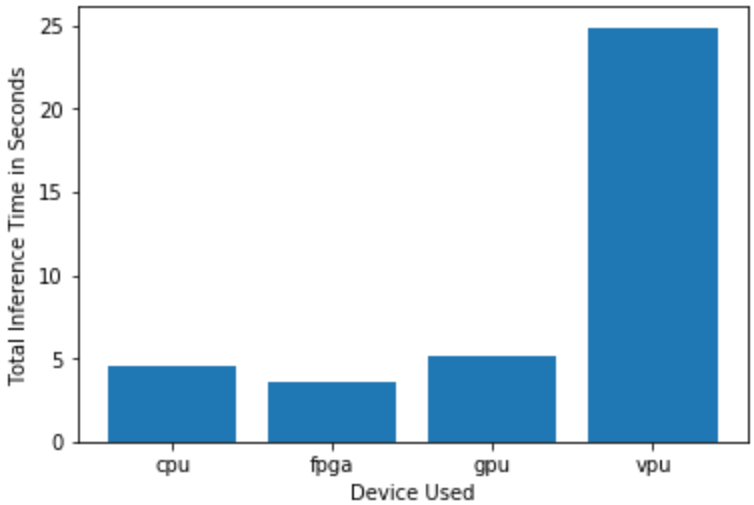
Maximum number of people in the queue	*2(normal), 5(rush)*
Model precision chosen (FP32, FP16, or Int8)	*FP32*

Test Results

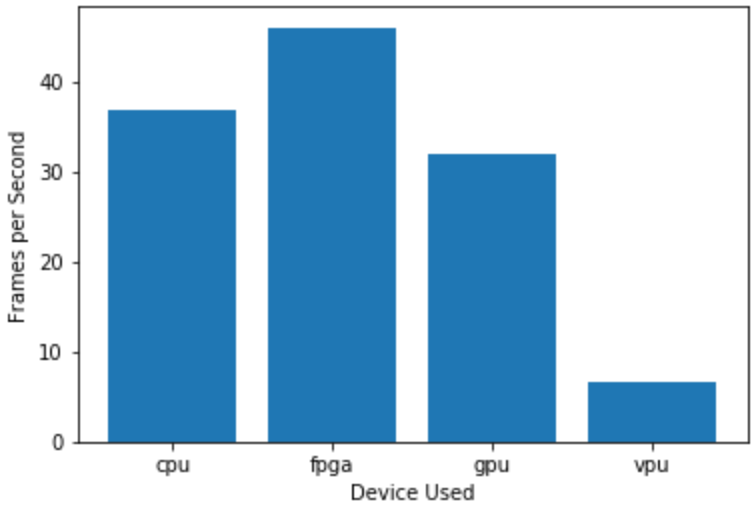
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



Model Load Time



Inference Time



FPS

Final Hardware Recommendation

Now synthesize your points from above and provide a brief write-up describing why the chosen hardware is the best choice for this scenario. Be sure to discuss the client's requirements, the test results, and how these relate to one another (e.g., perhaps one of the devices performed better than the rest, but does not meet one of the client's requirements).

Write-up: Final Hardware Recommendation
CPU is what I recommend as it's already there, and that the requirement doesn't require that we look into solutions such as FPGA. Moving forward, if load on the machines increase, we can switch to say an Integrated GPU if the machines have one.

Scenario 3: Transportation

Client Requirements and Potential Hardware Solution

Look through the scenario and find any relevant client requirements. Then, suggest a potential hardware type and explain how this hardware would satisfy each of the requirements.

Which hardware might be most appropriate for this scenario? (CPU / IGPU / VPU / FPGA)
VPU

Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?

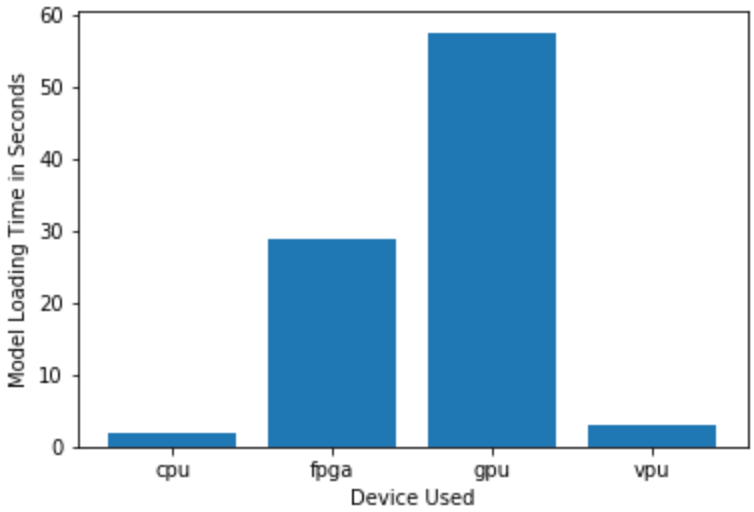
Example requirement: The client requires a tiny device to be connected to their CPU—and their budget is only about $100 for each device.	*Example explanation:* VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
They monitor the entire situation with 7 CCTV cameras on the platform.These are connected to closed All-In-One PCs that are located in a nearby security booth. The CPUs in these machines are currently being used to process and view CCTV footage for security purposes and no significant additional processing power is available to run inference.	*GIven this circumstance and also the budget stated, I think we can look into VPUs as solution.*
Ms. Leah's budget allows for a maximum of $300 per machine, and she would like to save as much as possible both on hardware and future power requirements.	*Again, this limits us to i think the cheapest solution which is VPUs.*
[TODO: Type your answer here]	*[TODO: Type your answer here]*

Queue Monitoring Requirements

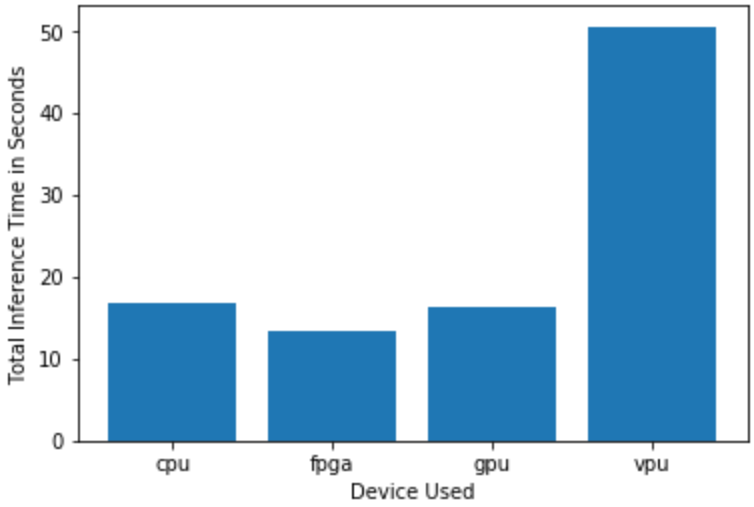
Maximum number of people in the queue	*7(Non-peak hours), 15*
Model precision chosen (FP32, FP16, or Int8)	*No strict requirement to speed and generally VPUs allow FP32*

Test Results

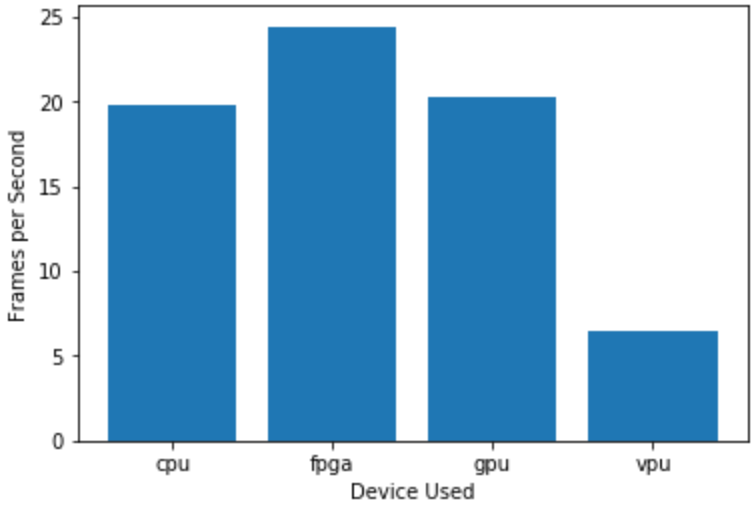
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



Model Load Time



Inference Time



FPS

Final Hardware Recommendation

Now synthesize your points from above and provide a brief write-up describing why the chosen hardware is the best choice for this scenario. Be sure to discuss the client's requirements, the test results, and how these relate to one another (e.g., perhaps one of the devices performed better than the rest, but does not meet one of the client's requirements).

Write-up: Final Hardware Recommendation
VPU is what I would like to recommend for this job. There are no available processing power from the existing device on the given scenario, sticking VPUs to the machine can help. With the given budget limitation, I think we cannot take it that we can buy new machines for the task, and thus we can proceed with VPUs,