Hardware choice

Proposal Document

Scenario 1: Manufacturing

Client Requirements and Potential Hardware Solution

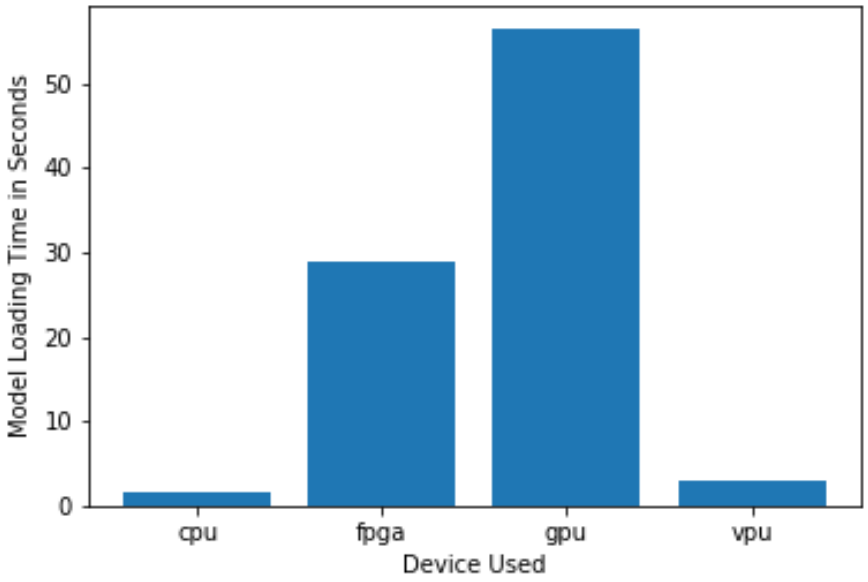
Which hardware might be most appropriate for this scenario?
FPGA - ***Intel® Arria® 10 GX1150 FPGA***

Requirement Observed	How does the chosen hardware meet this requirement?
Industrial grade system that last for at least 5-10 years.	FPGAs have a long lifespan and made to last long in industrial conditions, typically at least 10 years. The ambient operating temperature is in the range 5ºC - 60º, which is suited for factory line needs.
System should have ability to be repurposed for other issues.	FPGAs are field-programmable, they can be re-programmed to adapt to new, evolving, and custom networks. They can also be optimised for different deep learning tasks.
Image processing task of the system needs to be completed 5 times per seconds.	FPGAs can excute neural networks with high performance and very little latency.

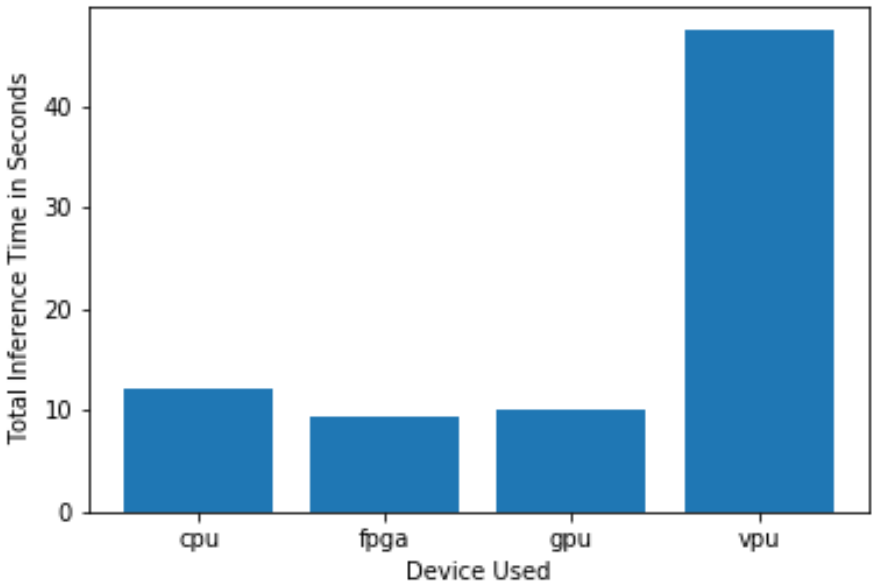
Queue Monitoring Requirements

Maximum number of people in the queue	*5*
Model precision chosen	*FP16*

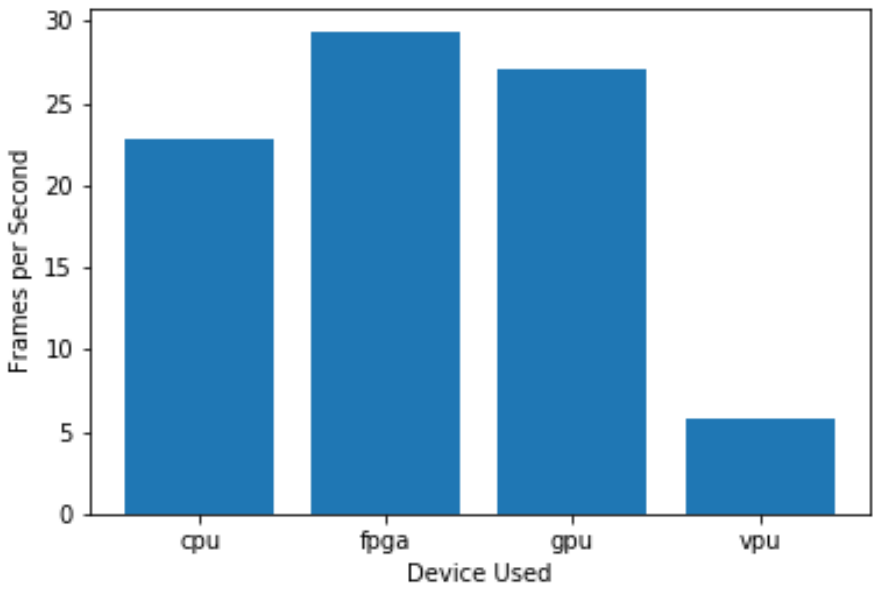
Test Results



Model Load Time



Inference Time



FPS

Final Hardware Recommendation

Final Hardware Recommendation
Considering the requirements like flexibility, long term solution and elevate image processing capabilities, FPGA s are the best hardware for this scenario. This is also proved by the test result: FPGA performed better in terms on inferencing time and number of processed frames per second.

Scenario 2: Retail

Client Requirements and Potential Hardware Solution

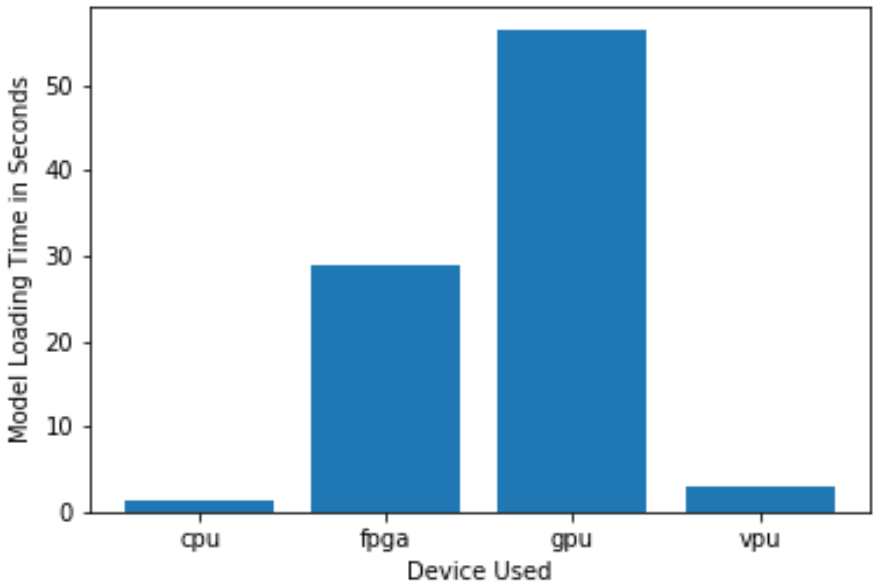
Which hardware might be most appropriate for this scenario?
IGPU *- Integrated Graphic Processor Unit*

Requirement Observed	How does the chosen hardware meet this requirement?
No budget for additional hardware.	As the IGPU is already present in the checkout computers there won't be need for additional hardware.
Save on power consumption	*An Intel i7 has a typical TDP of 95W. Moreover, unused sections in a GPU can be powered down to reduce power consumption.*
Need of an Edge AI System	Using an integrated GPU, the system does not need to be have an Internet connection, and the model loading and inference can all be done in the edge itself.

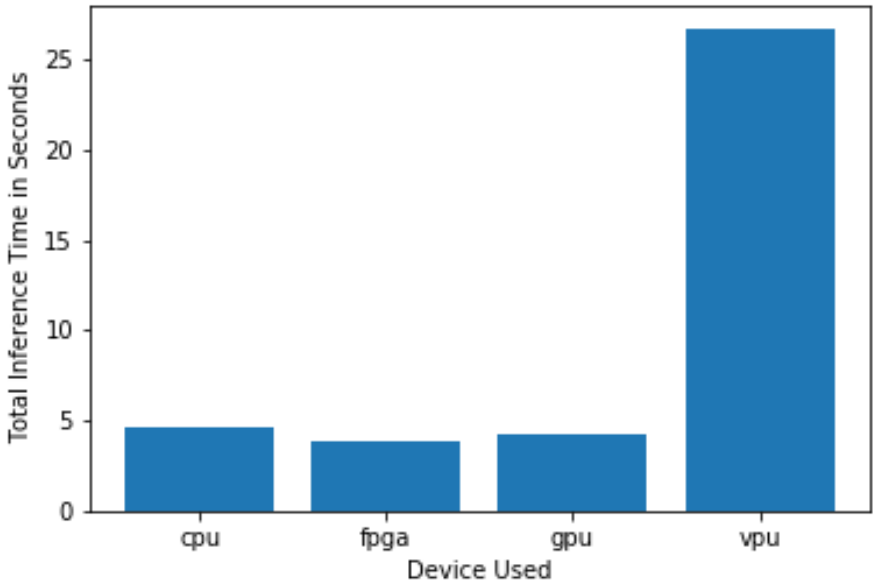
Queue Monitoring Requirements

Maximum number of people in the queue	*2*
Model precision chosen	*FP16*

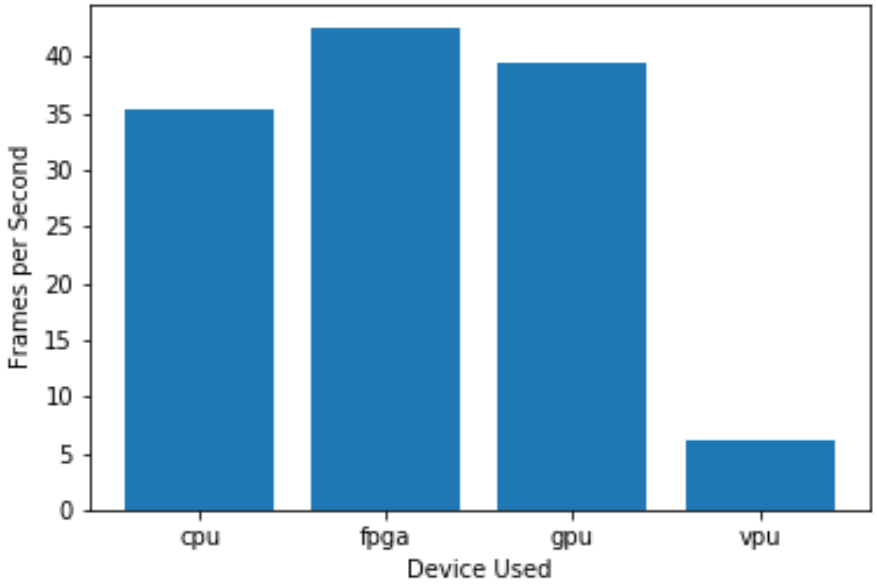
Test Results



Model Load Time



Inference Time



FPS

Final Hardware Recommendation

Final Hardware Recommendation
The use of an integrated graphic unit address the problems of space and budget. Moreover, as shown by the test results, an integrated GPU has a good inference time which would not impact the normal operations of the checkout computers.

Scenario 3: Transportation

Client Requirements and Potential Hardware Solution

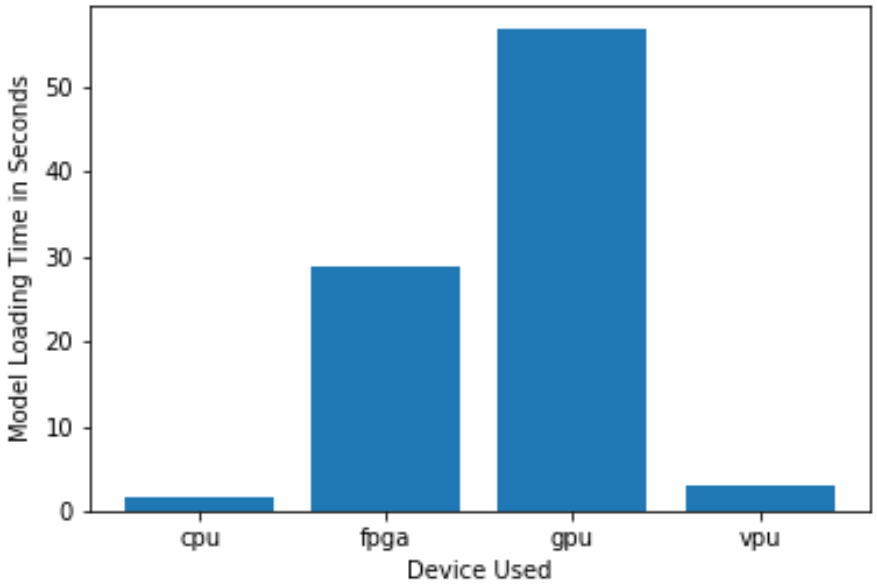
Which hardware might be most appropriate for this scenario?
VPU - ***Intel® Neural Compute Stick 2***

Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
Maximum budget of $300 per machine.	NCS2 cost only less than $100
Current machines have no additional processing power available.	NCS2 requires only a USB3 port. The inference load can be done exclusively on the NCS2 without additional CPU load.
Save on power requirements	NCS2 have a power consumption of only 1W
System need to process stream from 7 CCTV cameras	NCS2 can excite 4 trillion of operations per seconds and support multiple video stream per device.

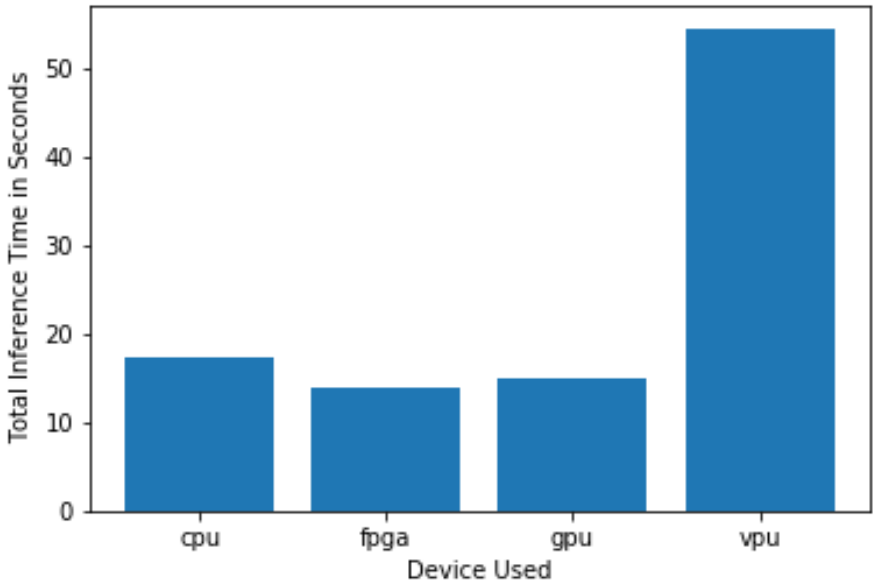
Queue Monitoring Requirements

Maximum number of people in the queue	*7*
Model precision chosen	*FP16*

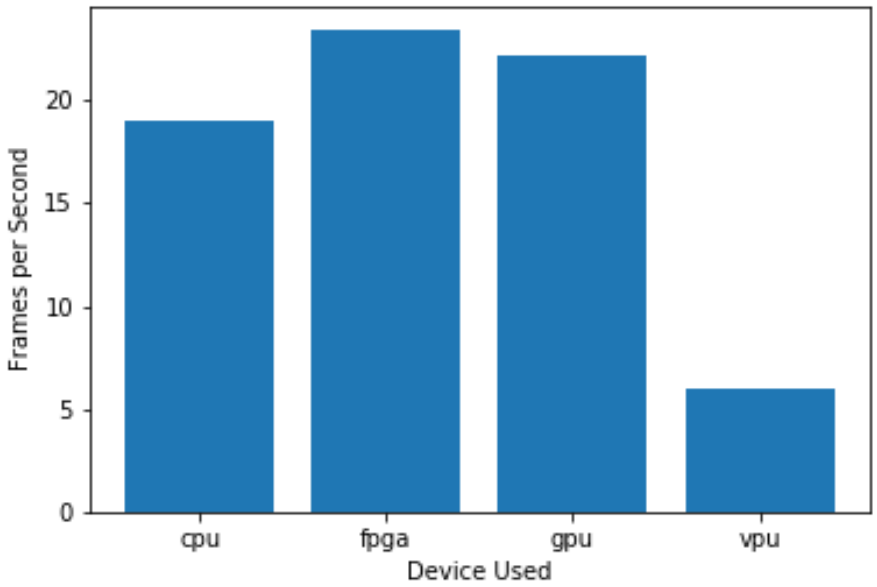
Test Results



Model Load Time



Inference Time



FPS

Final Hardware Recommendation

Final Hardware Recommendation
The test results show that an FPGA device would be a better choice. However, taking into account the budget and saving power requirements, a NCS2 is a more suitable choice.