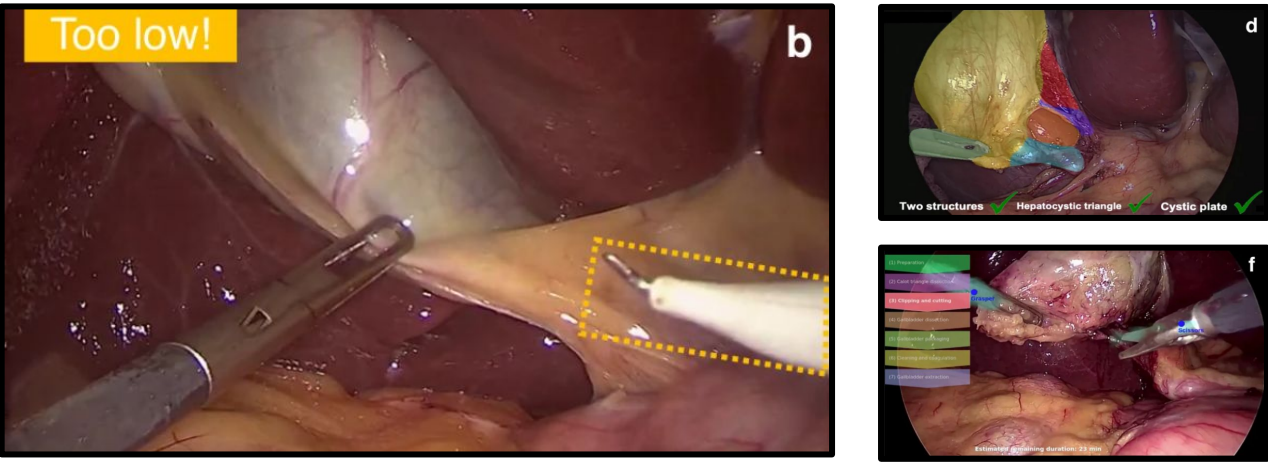


Evaluating Edge Computing vs. Cloud Computing in the Operating Room

Nana Ama Gyamfi-Kordie¹, Harkirat Singh², Chandra Shekhar Pandey³, Fatih Berkay Sarpkaya³, Fraida Fund³, Shivendra Panwar³
Democracy Prep Charter High School ¹ Francis Lewis High School² New York University Tandon School of Engineering ³

Introduction

Real-time AI assistance in operating room can give surgeon helpful feedback (images: [1]).



But, the response time must be very small!

Edge vs. Cloud Computing

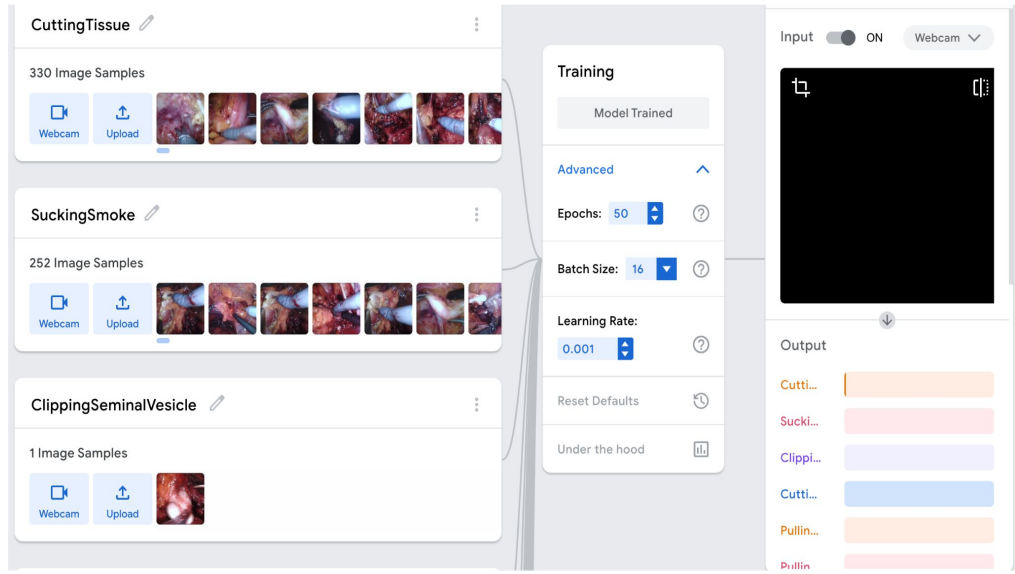
Key decision: where is the computation placed?

Type of Computation	Location	Trade Offs
Edge Computing	Computations are done locally	Little to no network delays but less powerful
Cloud Computing	Computations are done remotely	Powerful but prone to network delays

Goal: Compare the response time of Edge computing v.s. Cloud computing for real-time assistance in the operating room using computer vision

Methodology


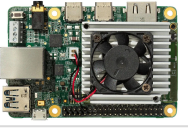
Dataset [2]: SARAS-MESAD Multi-domain Endoscopic Surgeon Action Detection, 8k images, 21 action classes.



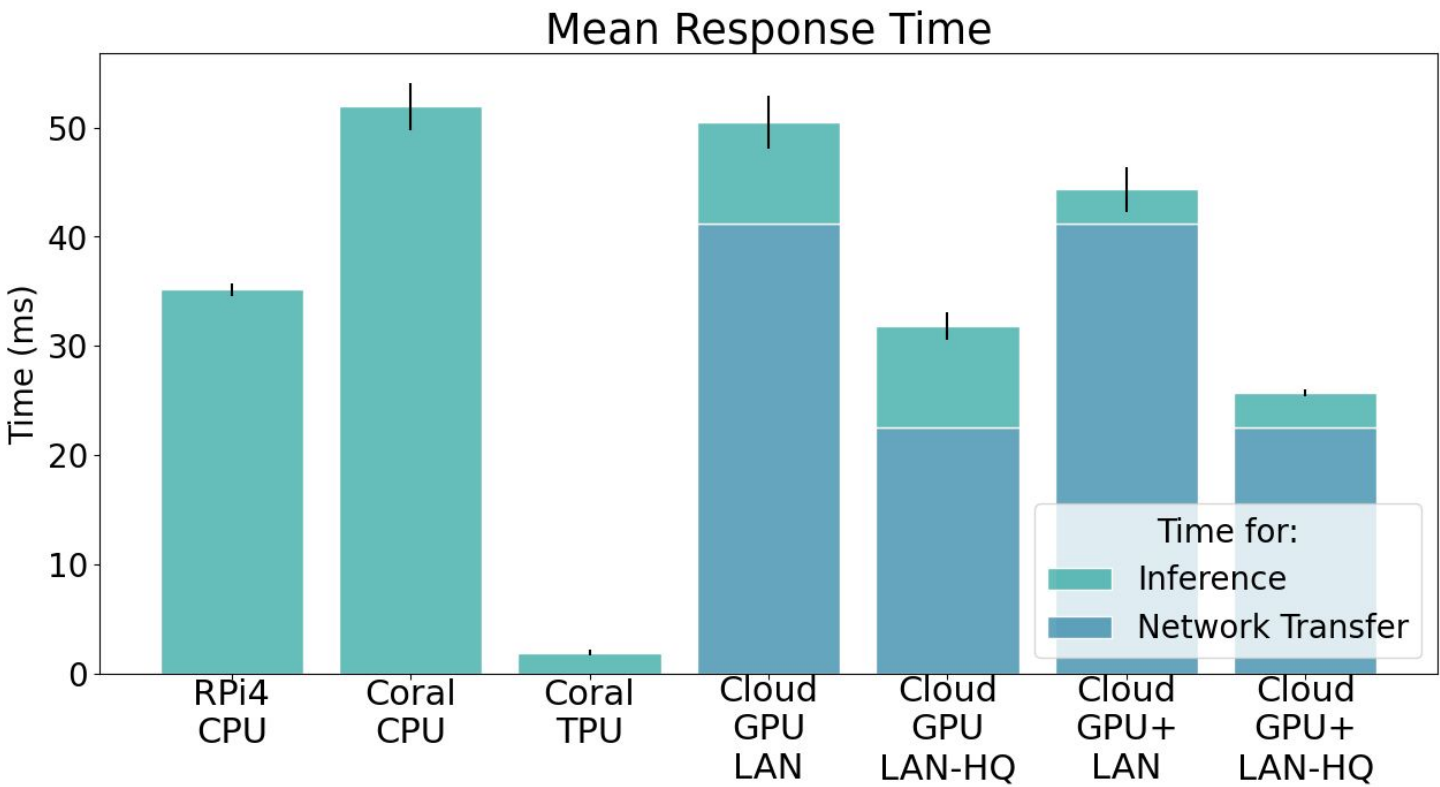
Model training on Google's Teachable Machine platform.

Network scenarios: LAN (100 Mbps, 20 ms RTT, 3 ms jitter), high-quality LAN (1 Gbps, 10 ms RTT)

Inference devices:

Raspberry Pi 4	CPU	
Coral Dev Board	CPU, TPU	
Cloud server	GPU (RTX6000), GPU + optimization	

Results



GPU/TPU can both be fast for inference, but for cloud computing we must also consider the network delay as part of response time.

Works Cited

[1] Mascagni, P., Alapatt, D., Sestini, L. et al. Computer vision in surgery: from potential to clinical value. npj Digit. Med. 5, 163 (2022). <https://doi.org/10.1038/s41746-022-00707-5>

[2] Bawa, Vivek Singh, et al. "The SARAS Endoscopic Surgeon Action Detection (ESAD) dataset: Challenges and methods." arXiv preprint arXiv:2104.03178 (2021).

Supported By

