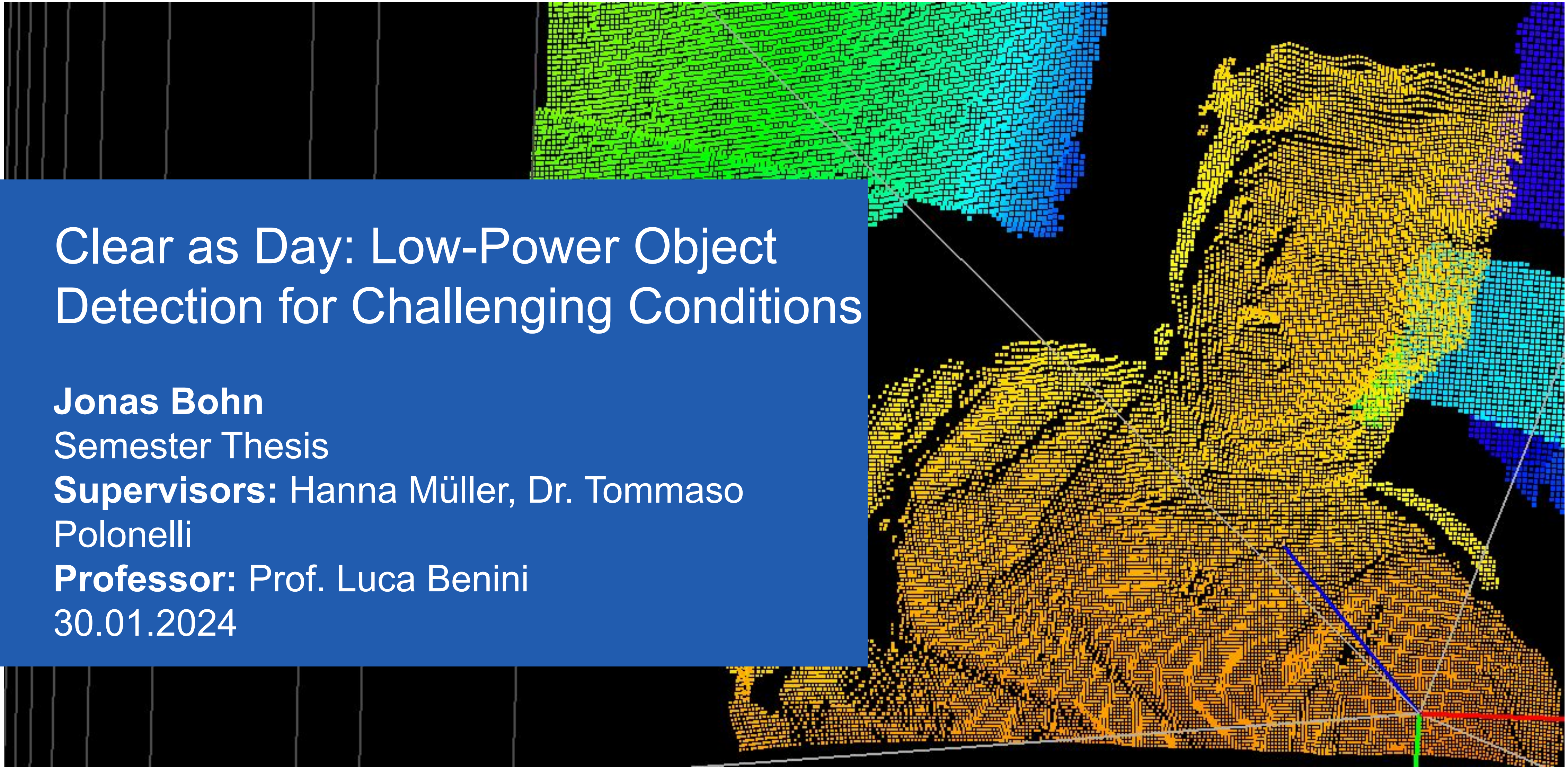


Clear as Day: Low-Power Object Detection for Challenging Conditions

Jonas Bohn
Semester Thesis
Supervisors: Hanna Müller, Dr. Tommaso Polonelli
Professor: Prof. Luca Benini
30.01.2024

Clear as Night

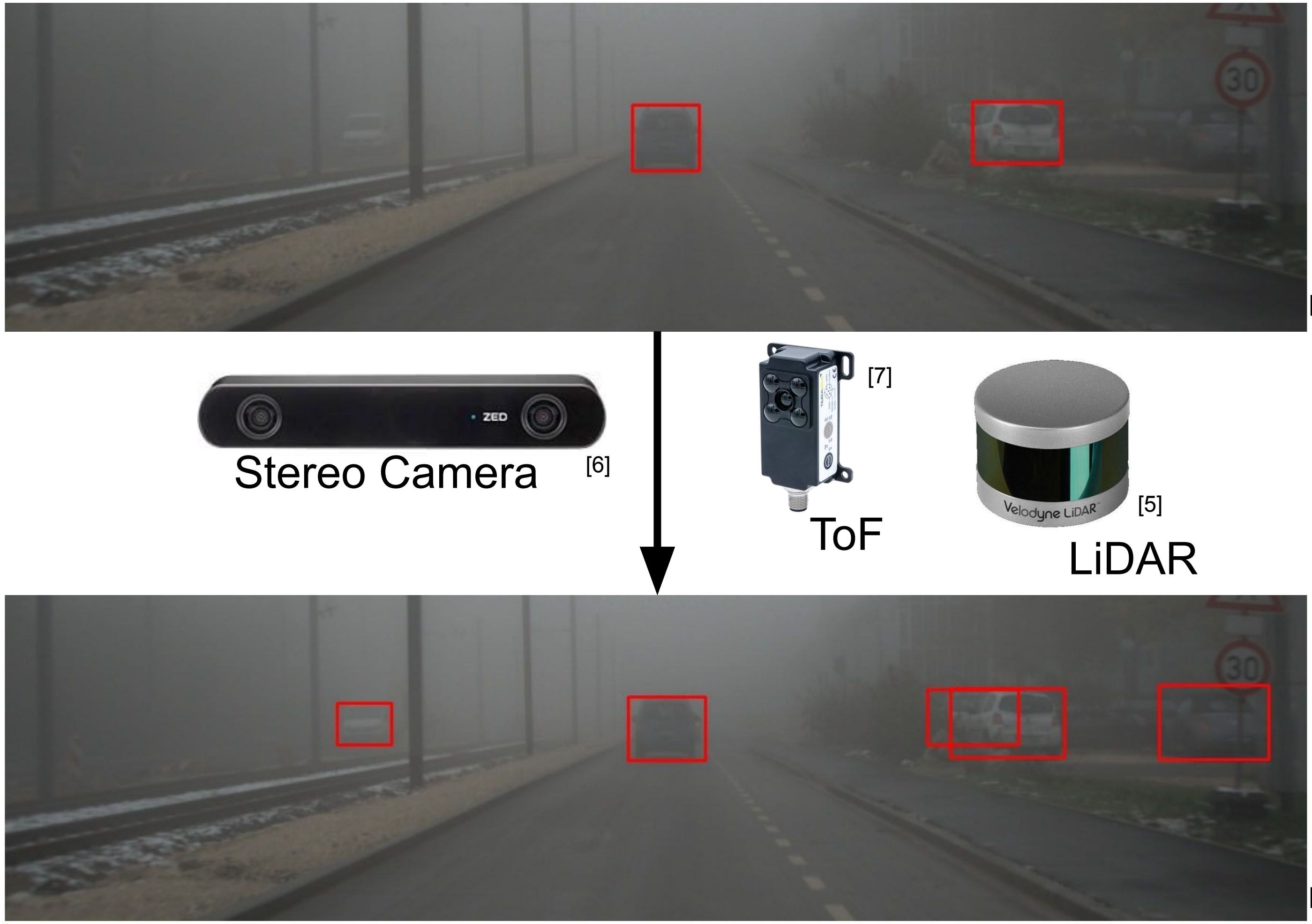
- Challenges Going to Bed:
    - Restricted View
    - Fast Perception
    - Limitations in Computation and Power
- Challenges in Mobile Robotics:
    - Restricted View
    - Fast Perception
    - Limitations in Computation and Power

[1,2]

[3,4]

What is Currently Done?

- Object Detection with CNN
- Sensor Fusion
- Low Resolution ToF [8, 10]

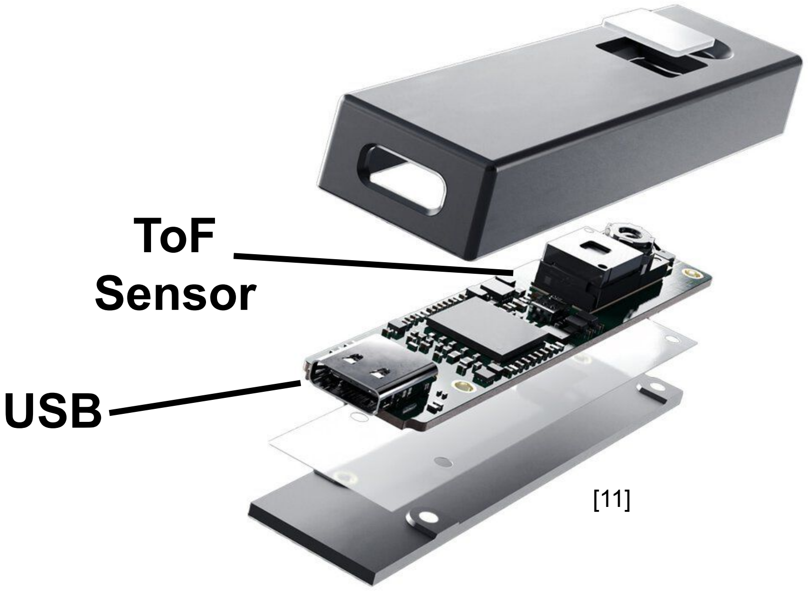


[9]

[9]

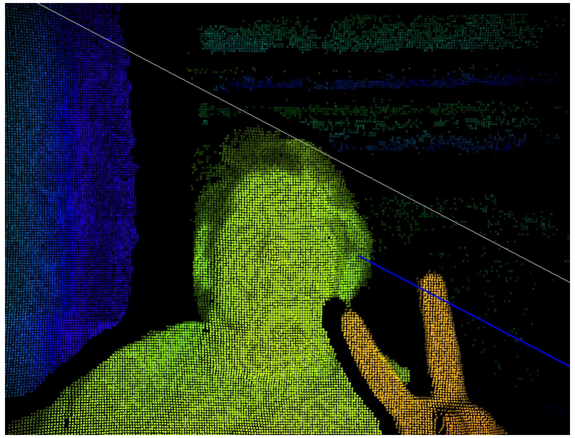
Flexx2 3D Camera

- ToF Image Sensor
- 224 x 172 Pixel
- IR Image & 3D Point Cloud
- USB Interface
- Depth from 0.1 - 4 m
- Up to 60 FPS with 9 Predefined Modes
- 13 g
- Average 300 mW



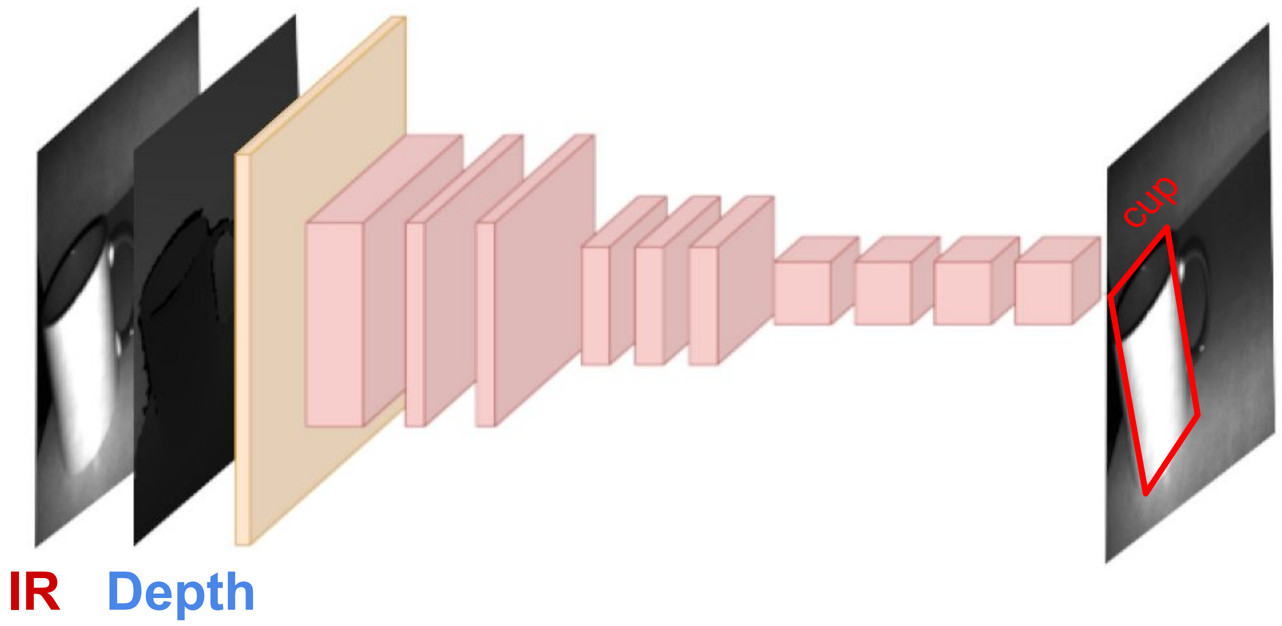
Infrared Image

3D Point Cloud



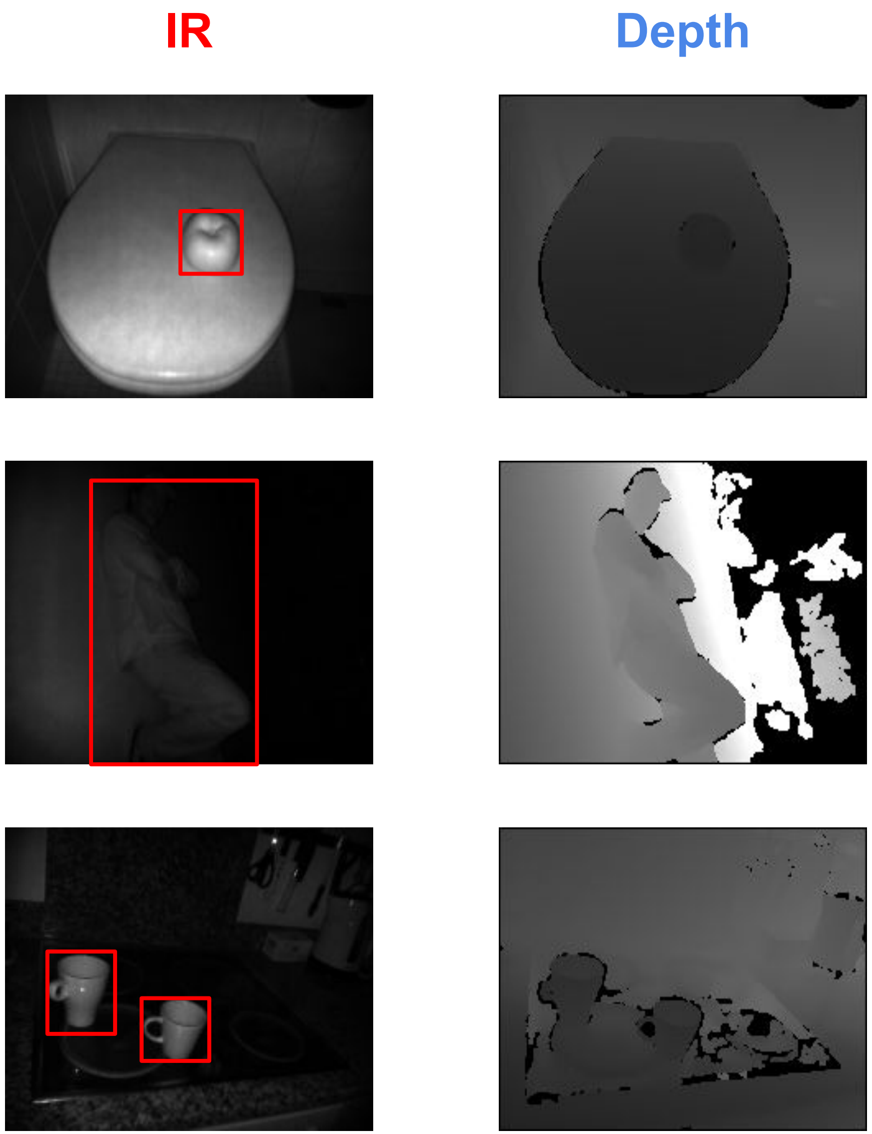
Object Detection for Challenging Conditions

- Ultra-Small, Ultra-Fast Network
- Impact of Sensor Fusion
- Deployment Ready Model



Dataset

- 1486 Image Pairs
- 3 Classes (Apple, Cup, Person)
- Bounding Box Labels



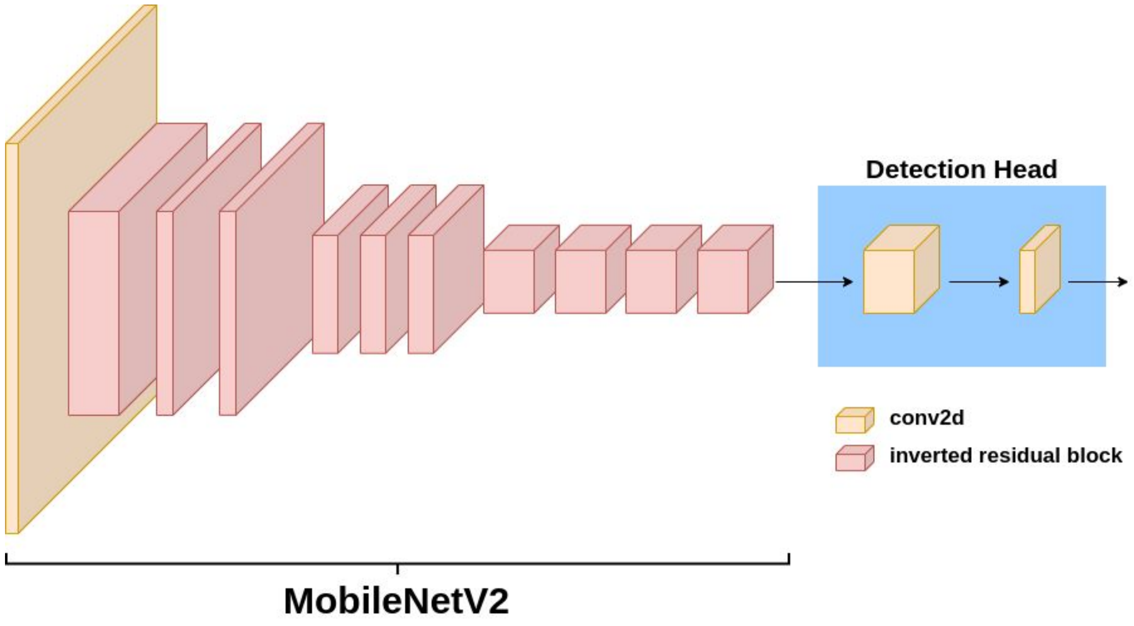
Fast Object Detection Models

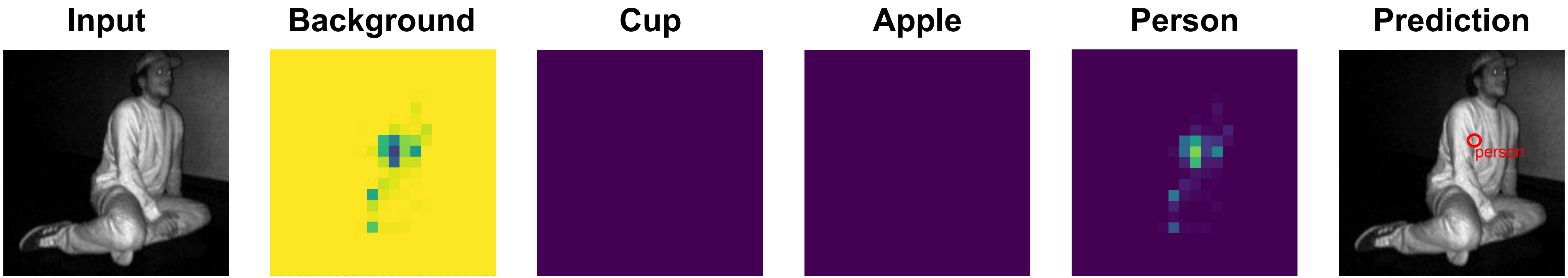
- YOLO-based Networks
- Smallest Network ~ 400K Parameters
- FOMO
- 20K Parameters in Demo
- 30x Faster than YOLOv5 [17]

Model	Input Size	#Params
YOLOv8n [12]	640 x 640	3.2M
YOLO-X Nano [13]	416 x 416	0.91M
Tiny YOLO-Lite [14]	416 x 416	0.6 M
TinyissimoYOLO [15]	88 x 88	0.42 M
FOMO [17]	192 x 192 [16]	**0.02M** [16]

Faster Objects, More Objects (FOMO)

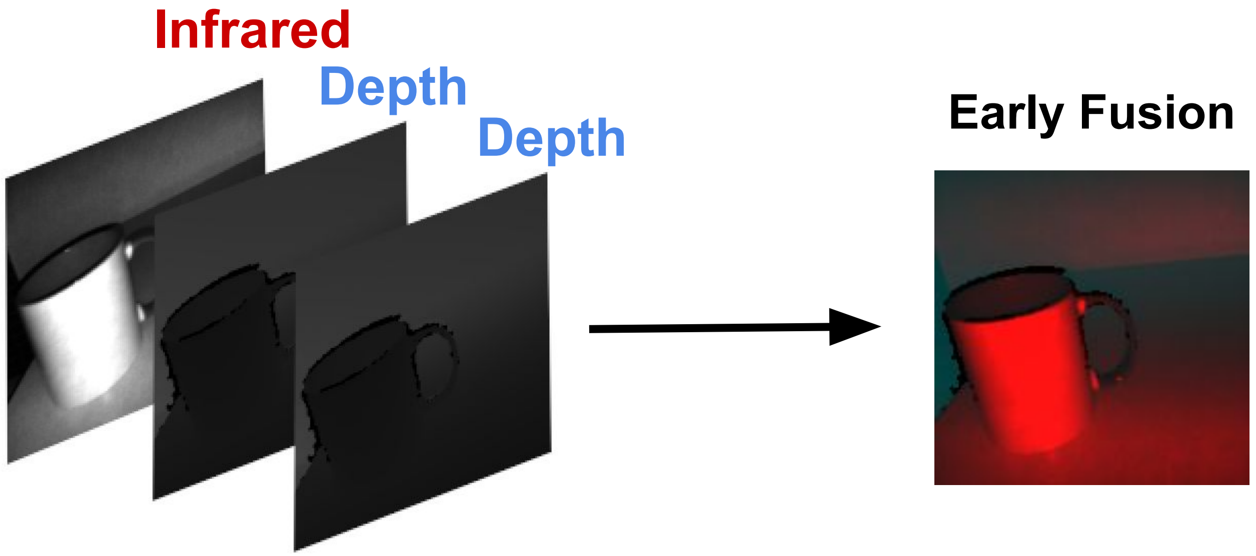
- MobileNetV2 Backbone
- Outputs Class Heatmaps
- Centroid Predictions





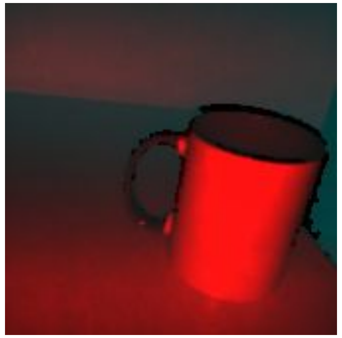
Fusion Experiments

- Three Fusion Strategies
- Early Fusion

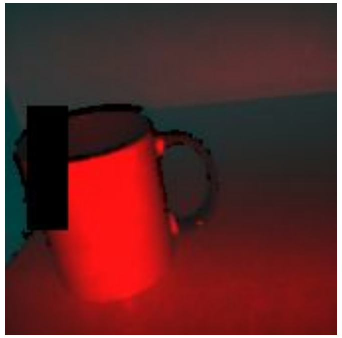


- 80 / 20 Train-Test Split
- Data Augmentation
- F1 Score Evaluation

Flip



Cut Out



Channel Dropout

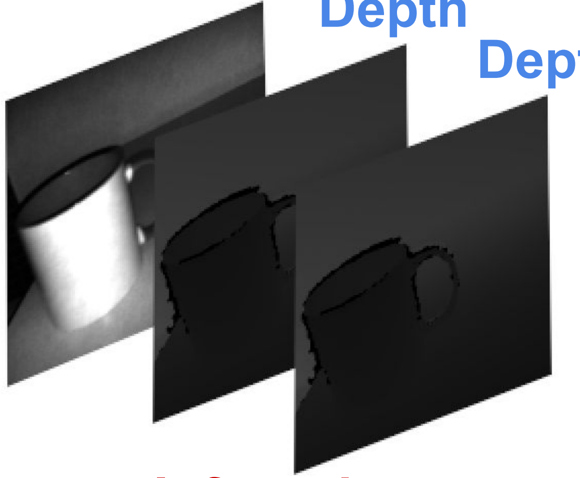
Quantitative Results

Infrared

F1-Score: 0.86

Params: 17’916

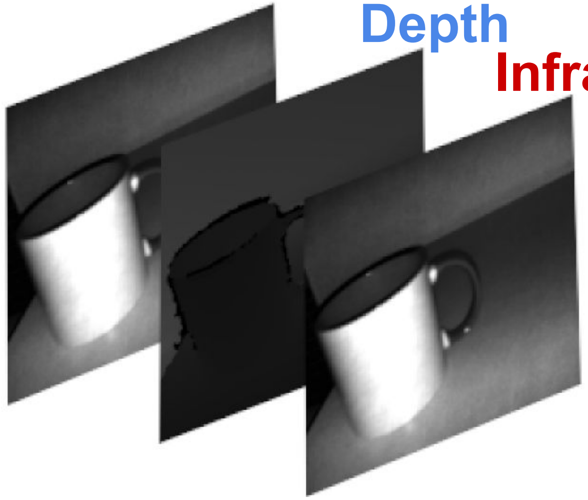
Infrared Depth Depth



F1-Score: 0.82

Params: 18’204

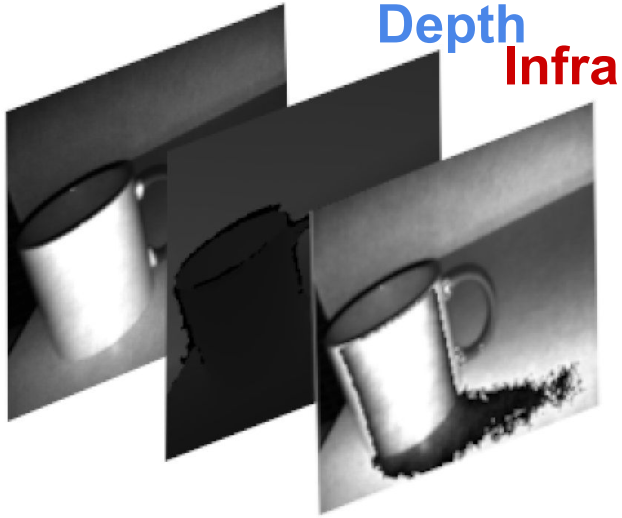
Infrared Depth Infrared



F1-Score: 0.84

Params: 18’204

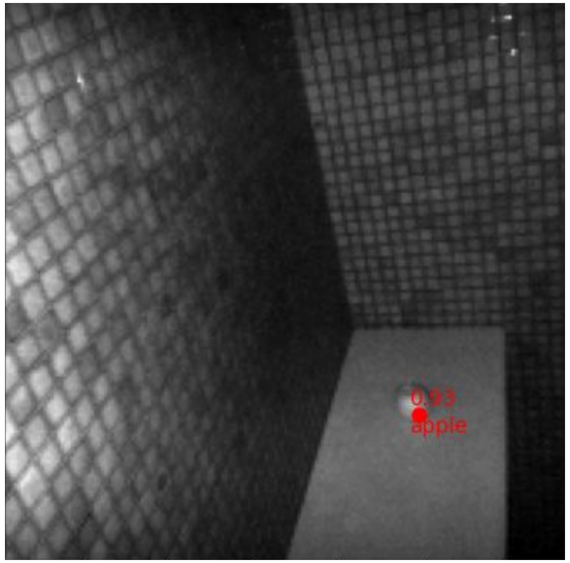
Infrared Depth Infrared + Depth

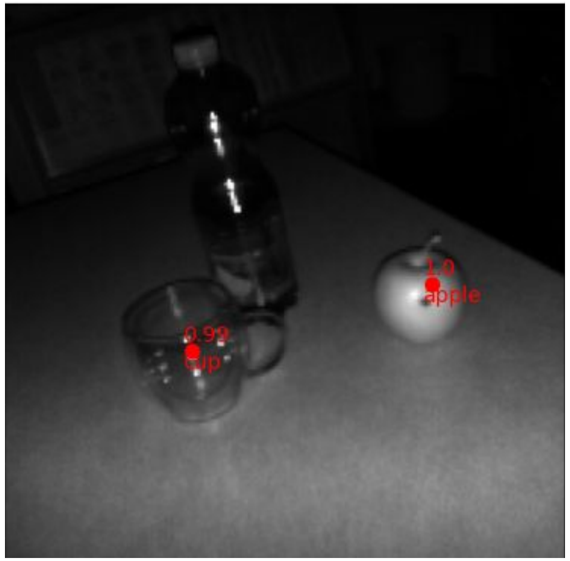


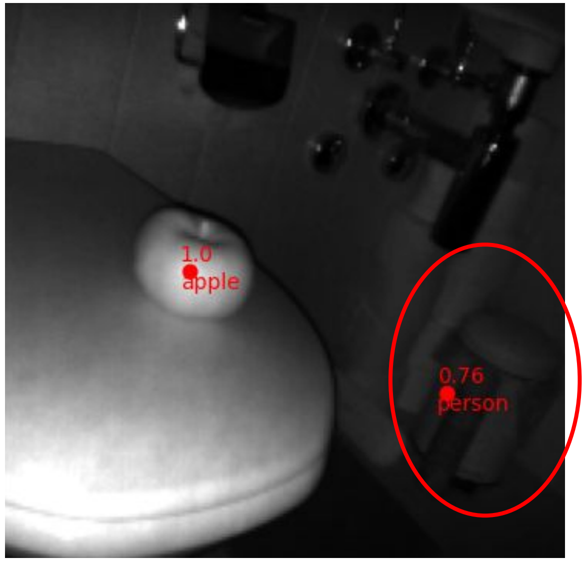
F1-Score: 0.86

Params: 18’204

Qualitative Results



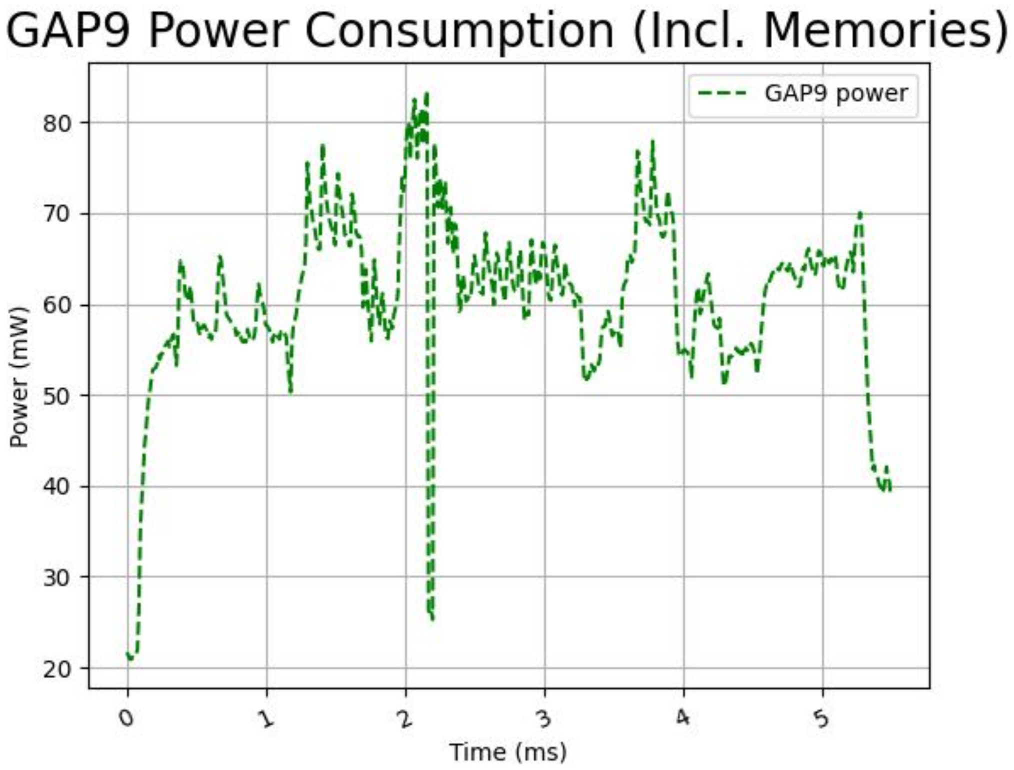




Performance Measurements on GAP9

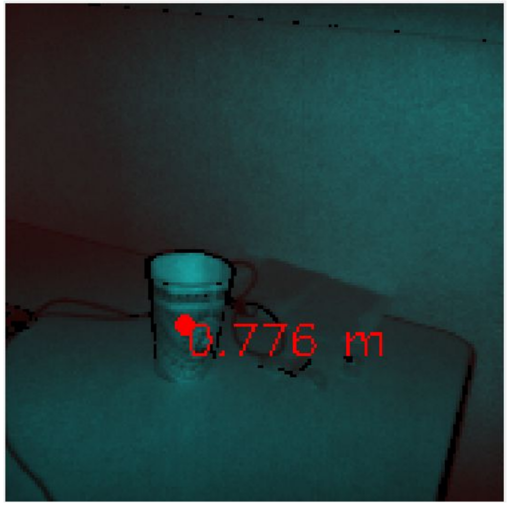
- Quantized Model with GAPFlow
- 10 Core RISC-V
- 370 MHz Internal Clock

- Average Power: 60 mW
- ~ 5 ms Processing
- 18 Total MOps
- 2 Total MCycles



Clear as Day!

- 60 FPS
- Detection in the Dark
- Distance to Object

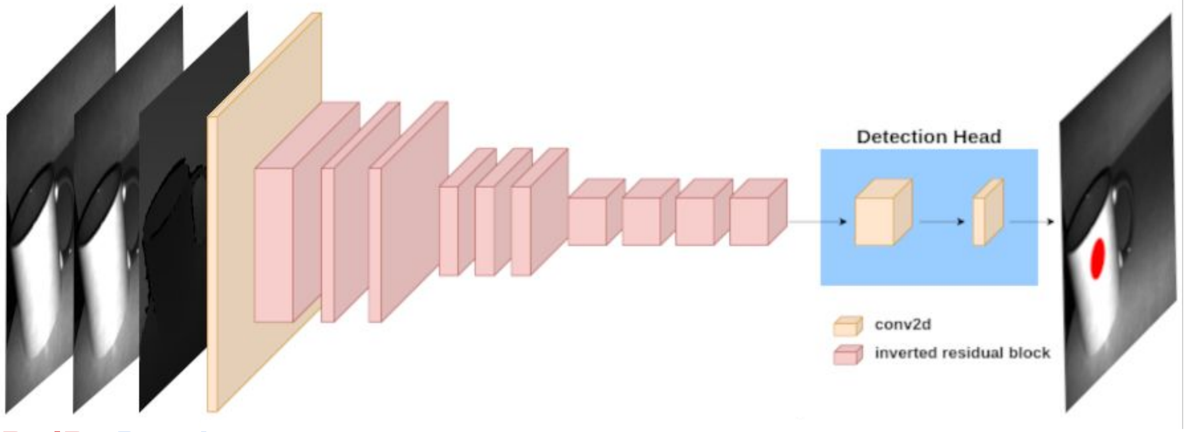


IPhone Camera

Flexx2 + FOMO

Conclusion + Future Work

- Dataset for Detection with Infrared and Depth
- Sensor Fusion for Object Detection
- Fast Detection in the Dark

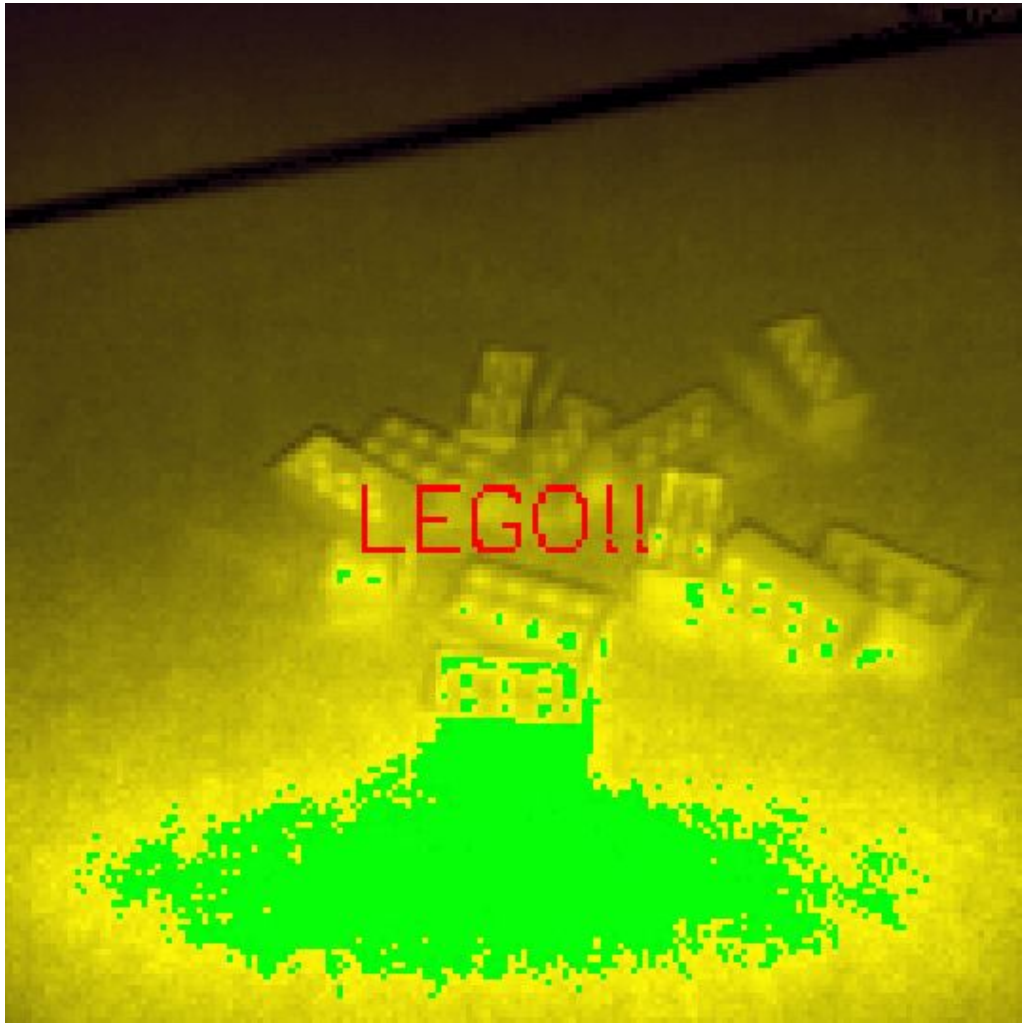


IR IR Depth

FOMO

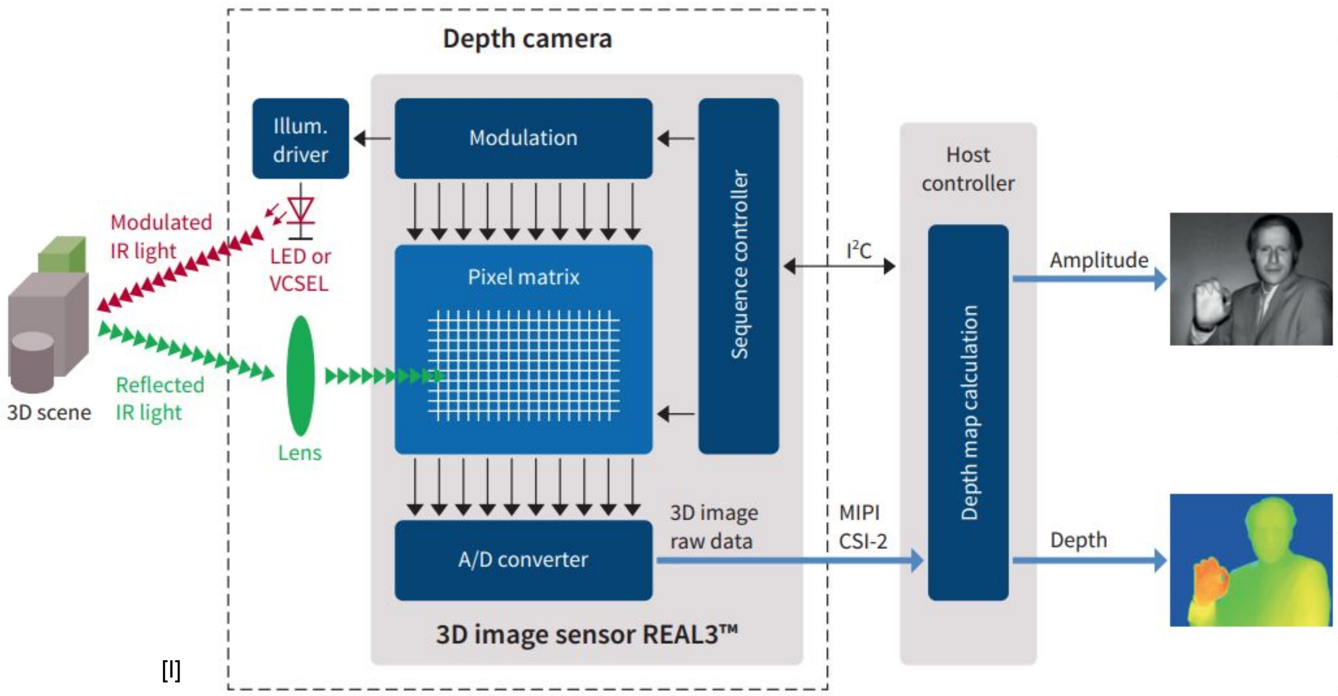
- Deploy and Test Network on GAP9
- Use Depth for Obstacle Avoidance or Tracking
- Improve FOMO for More Stable Predictions

Watch Out!



Flexx2 3D Camera Bonus Information

Time-of-Flight principle and block diagram



[I]

2. flexx2 Specifications

Parameter	flexx2
Dimensions	71.9 mm x 19.2 mm x 10.6 mm
Time of Flight (ToF) Sensor	IRS2381 Infineon® REAL3™ 3D Image Sensor IC based on pmd technology
Resolution	224 x 172 Depth Pixels 38,000 Depth Points
Weight, without USB Cable	13g
Measurement range	0.1 – 4 m
Framerate	Up to 60fps (3D frames); 9 pre-defined operation modes
Power consumption	USB3.0 compliant, average 300mW for full system
Illumination	940 nm, VCSEL, Laser Class 1
Software	Royale SDK (C/C++ based, supports Matlab, OpenCV, ROS 1/2)
Viewing angle (H x V)	56° x 44°
Interface	USB3.0 (data & power)
Depth resolution	<= 1% of distance, all modes
Sunlight Tolerance	At 100K Lumens (Full Sunlight), Loses ~10% max range vs. Indoors
Operating System	Windows 10, Linux/ARM

[I] Data Brief REAL3, Infineon
https://www.infineon.com/dgdl/Infineon-REAL3+Image+Sensor+Family-PB-v01_00-EN.PDF?fileId=5546d462518ffd850151a0afc2302a58

References

[1] LEGO Brick, https://www.lego.com/en-ch/product/8-stud-storage-brick-red-5006867
[2] Adobe Stock Image, https://stock.adobe.com/de/search/images?k=bedroom+night
[3] IStock Image, https://www.istockphoto.com/de/fotos/street-lights-foggy-misty-night-lamp-post-lanterns-deserted-road
[4] Wall-E. Directed by Andrew Stanton,, Pixar, 2008.
[5] BusinessWire, Velodyne LiDAR
[6] Reichelt, ZED 2 Stereo Kamera
[7] Terabee, TOF sensor
[8] Yijin Li et al., DELTAR: Depth Estimation from a Light-weight ToF Sensor And RGB Imag, ECCV 2022
[9] M. Bijelic. Seeing Through Fog Without Seeing Fog: Deep Multimodal Sensor Fusion in Unseen Adverse Weather, 2020.
[10] C. Brander et al., Improving Data-Scarce Image Classification Through Multimodal Synthetic Data Pretraining, IEEE Sensors, 2023.
[11] pmdtechnologies, flexx2 camera, https://3d.pmdtec.com/en/3d-cameras/flexx2/, 2024.
[12] Ultralytics, YOLOv8, https://github.com/ultralytics/ultralytics, 2024
[13] Zheng Ge et al., “YOLOX: Exceeding YOLO Series 2021”, 2021
[14] Shiqi Chen et al., “Learning Slimming SAR Ship Object Detector Through Network Pruning and Knowledge Distillation”, 2021
[15] J. Moosmann, M. Giordano, C. Vogt, M. Magno, “TinyissimoYOLO: A Quantized, Low-Memory Footprint, TinyML Object Detection Network for Low Power Microcontrollers”, 2023
[16] L. Boyle, N. Baumann, S. Heo and M. Magno, "Enhancing Lightweight Neural Networks for Small Object Detection in IoT Applications," *2023*
[17] EdgeImpulse, FOMO: Object detection for constrained devices, 2023

Jonas Bohn
Semester Thesis
bohnj@student.ethz.ch

ETH Zürich
Center for Project-Based Learning
ETF F109
Sternwartstrasse 7
8092 Zürich

www.pbl.ee.ethz.ch