



# **Enable Spatial Understanding for Embedded/Edge Devices with DepthAI**

Erik Kokalj

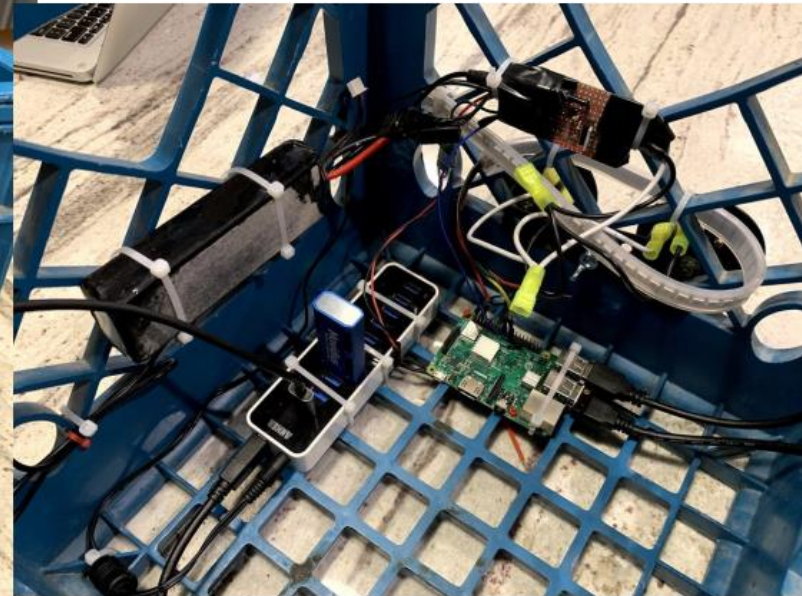
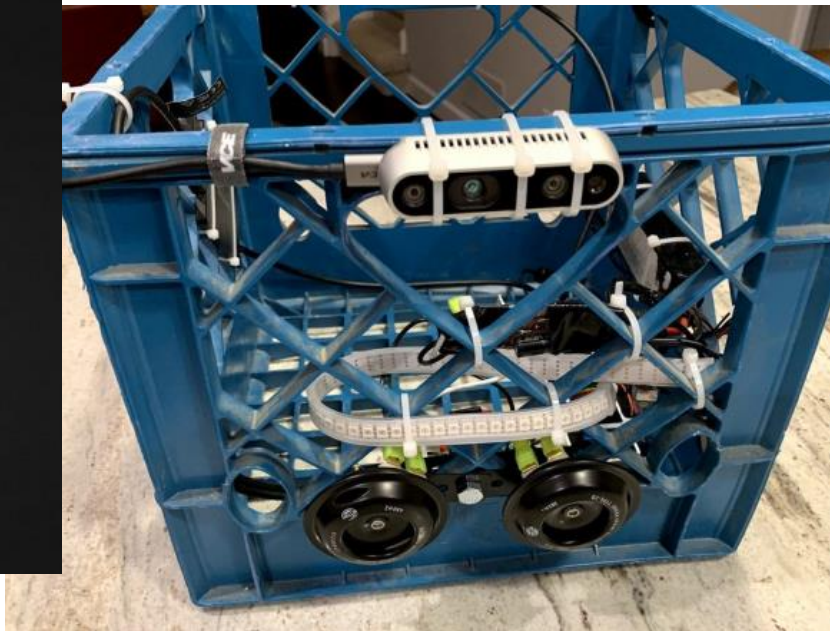
Director of Applications Engineering

Luxonis

# How it Started: Smart Cam for Saving Cyclist Lives



- Smart camera to detect vehicles, their speed and trajectory in an embedded, low power, high performance (FPS, resolution) solution.
- Required: AI, CV, stereo camera pair, high resolution color camera
- No such platform available – until we engineered **DepthAI** and **OAK cameras**



# Our Competition and How We Differentiate



**RealSense** - Stereo depth perception only



**Say goodbye to Intel's RealSense tech by remembering its incredible demos**

*The cameras had some great demos over the years, but Intel is "winding down" the business*

By [Mitchell Clark](#) | Aug 17, 2021, 5:55pm EDT

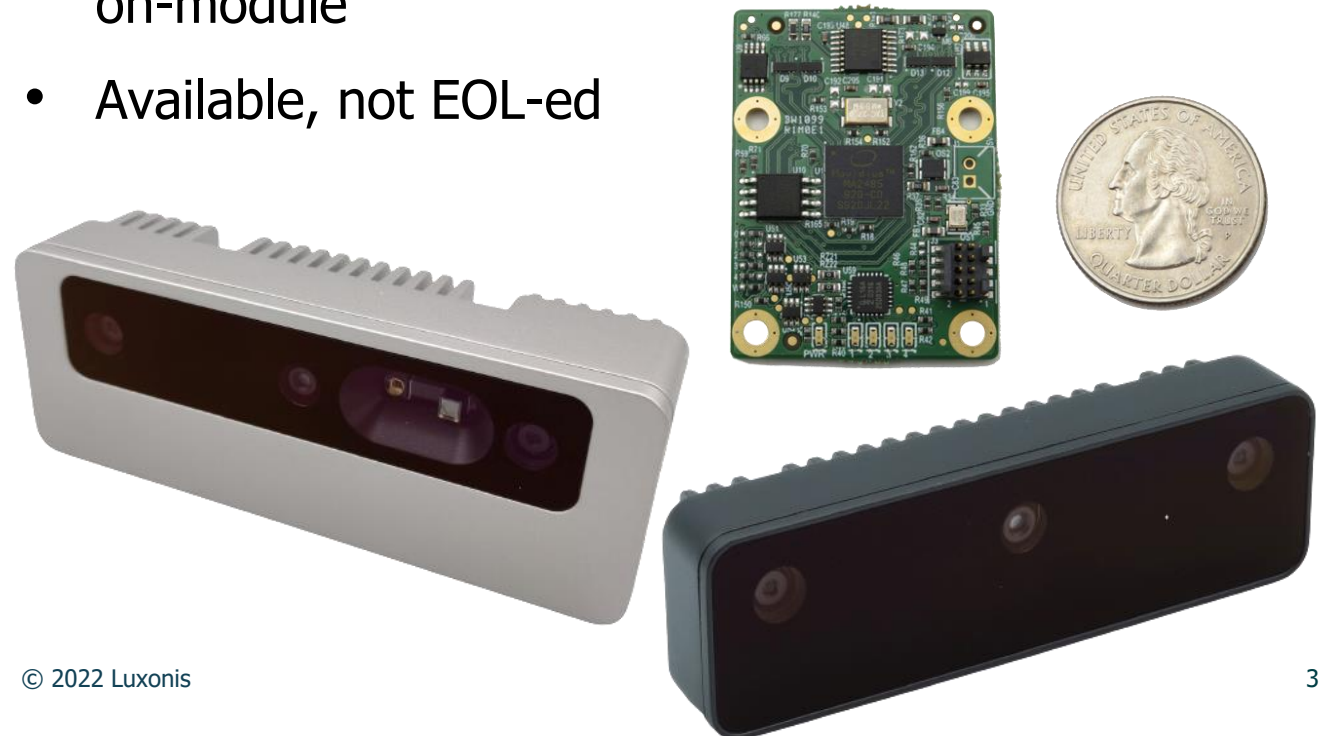
NEWS ROBOTICS

**Intel Will Keep Selling RealSense Stereo Cameras >**  
Intel's lidar, face recognition, and tracking cameras will be discontinued, but stereo cameras are safe, for now

BY EVAN ACKERMAN | 19 AUG 2021 | 3 MIN READ |

## OAK cameras:

- Onboard features: AI, CV, stereo, video encoding, object and feature tracking, object localization, Python scripting, etc.
- Integrate into your own products with system-on-module
- Available, not EOL-ed

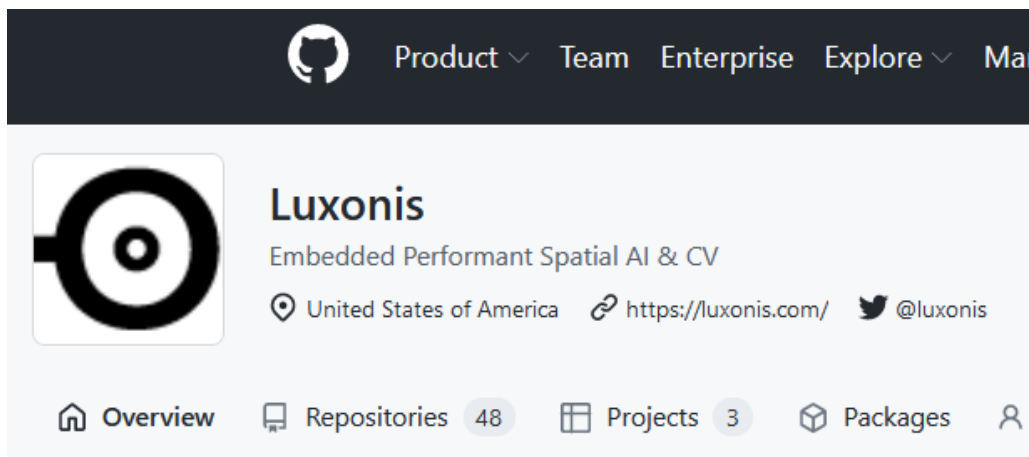


# Software — DepthAI Platform



Software side of our Spatial AI stack:

- **Firmware:** Running on Myriad X VPU, embedded system
- **API:** communicates with our OAK cameras, C++/Python/Java
- **Apps:** Open-source apps like face recognition, license plate recognition, gaze estimation, road segmentation, social distancing, and 50+ more
- **SDK:** Ease of use, abstraction of the API
- **AI training:** Notebooks for training, conversion and deployment of models
- **AI model zoo:** Models prepared to run directly on the cameras
- **Software integrations:** ROS/ROS2, Android, Unity, Foxglove, etc.

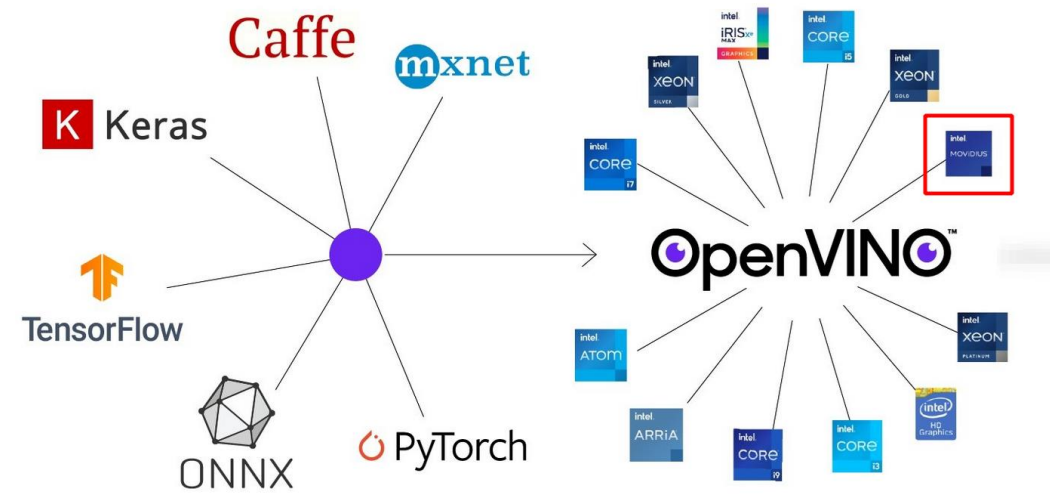




# Hardware Capabilities



- **Myriad X** vision processing unit (VPU)
- **AI:** 1.4 TOPS for AI, run any AI model
- **CV:** warp/dewarp, resize, crop, edge detection, feature tracking, etc.
- **Encoding:** H.264/H.265, MJPEG - 4k / 30 FPS
- **Stereo depth:** with filtering, post-processing, and RGB-depth alignment
- **Color camera:** Up to 60 FPS, up to 20MP
- **Object tracking:** 2D/3D
- **Low power:** below 5 W at maximum workload



**Myriad™ X Specs**  
Visual Processing Unit

Built-in to every OAK module

Compute Capacity:	4 Trillion Ops/sec
Vector Processors:	16 SHAVEs
Vision Accelerators:	20+
Memory Bandwidth:	450 GB/sec

**Also Features**

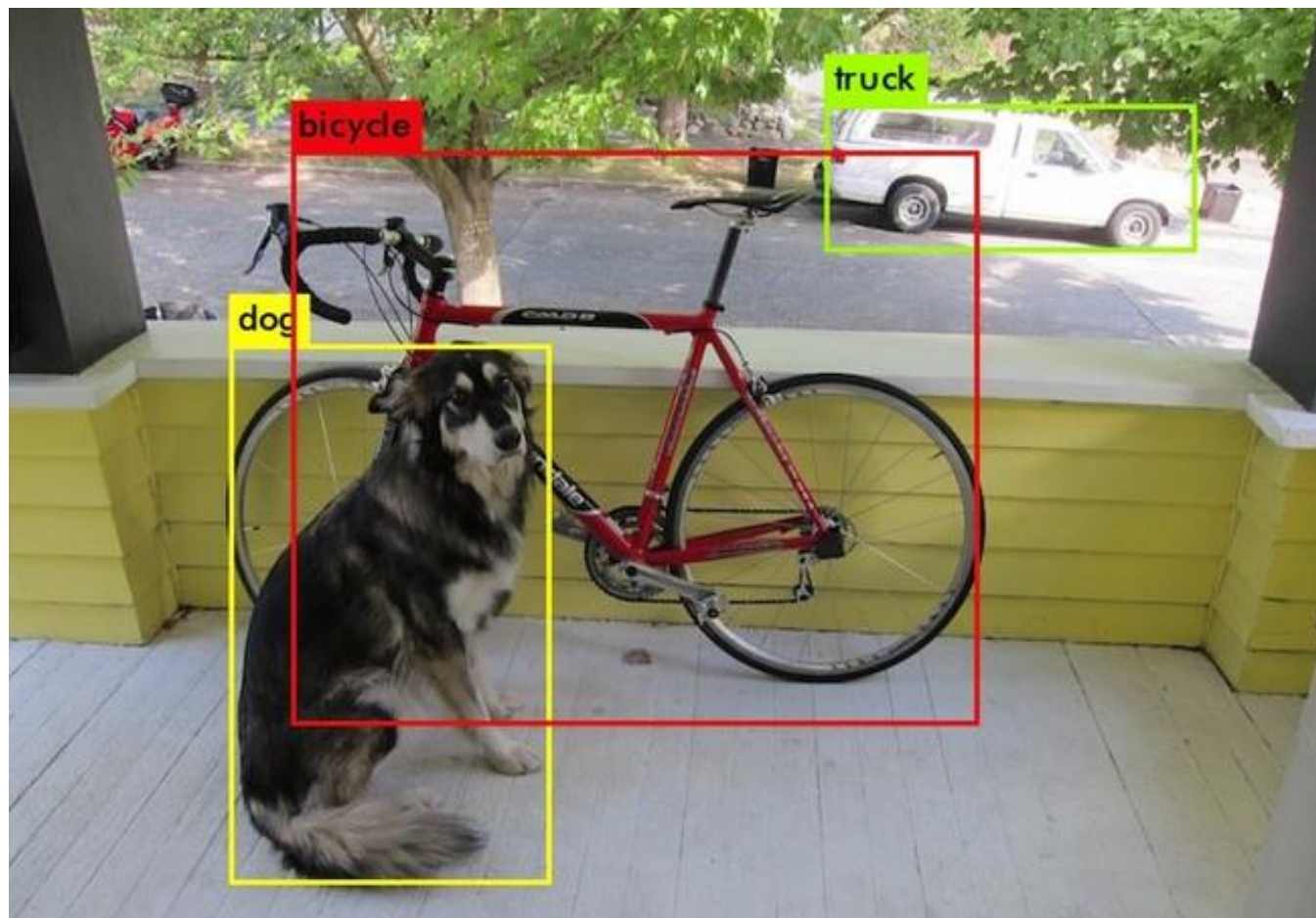
2x Neural Compute Engine (1.4 TOPS)

# Object Detection



On-device decoding:

- YOLO (3,4,5)
- MobileNet-SSD



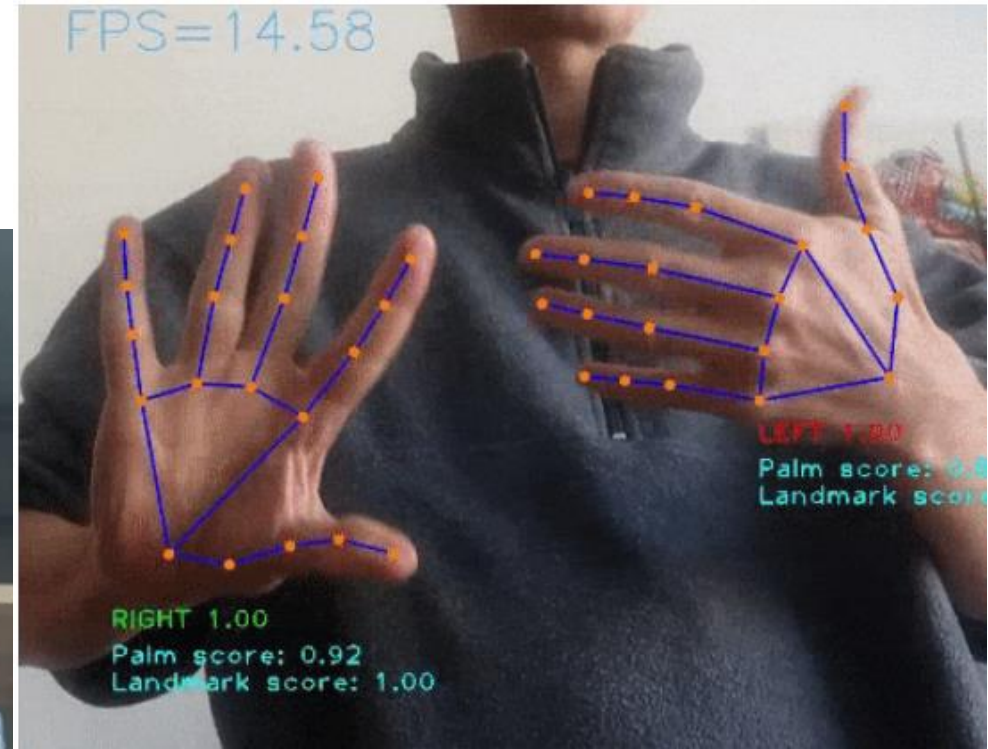
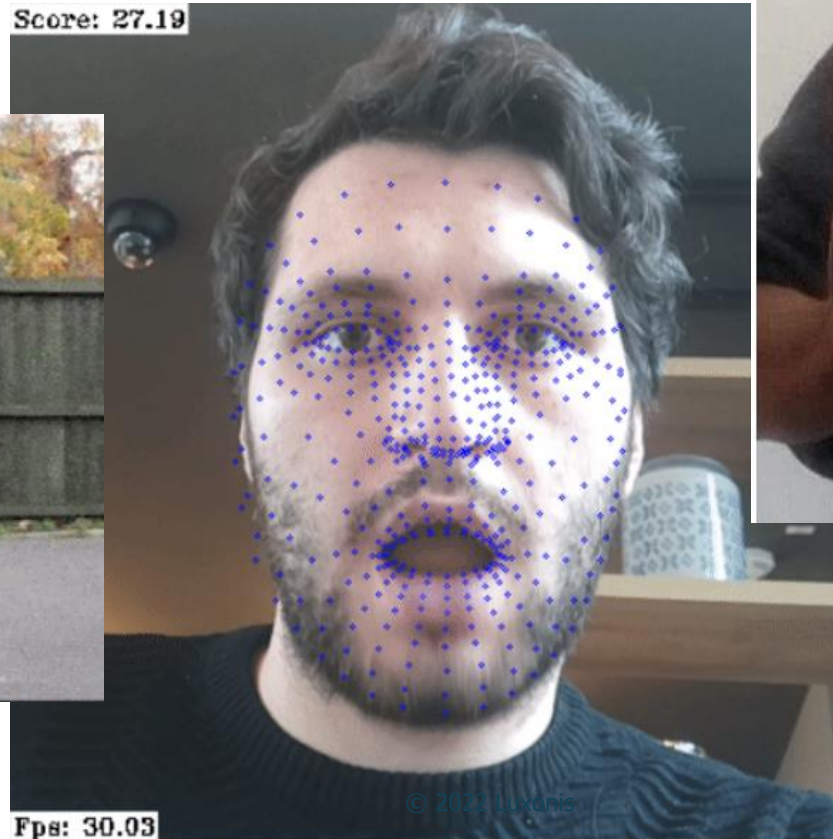


# Landmarks Detection



Examples of landmark detection:

- Pose estimation
- Hand tracking
- Facial landmarks

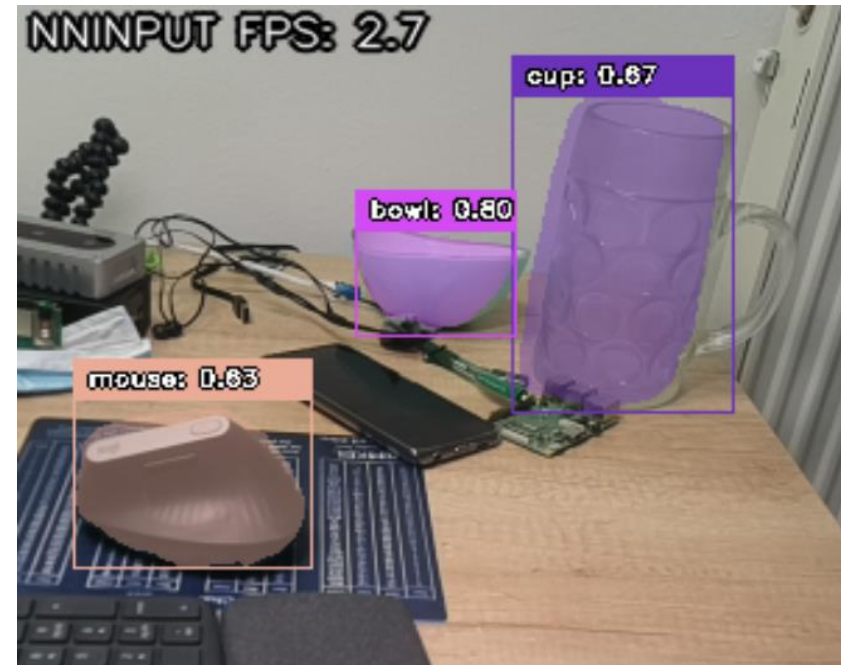
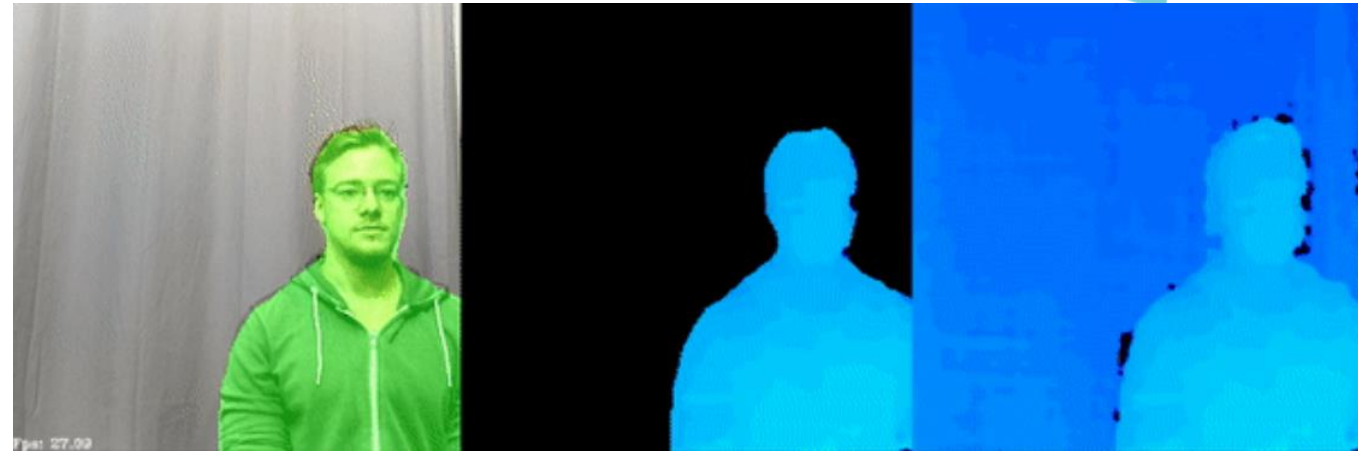


# Semantic Segmentation



Examples of semantic segmentation:

- Person segmentation
- Road segmentation
- Multi-class segmentation



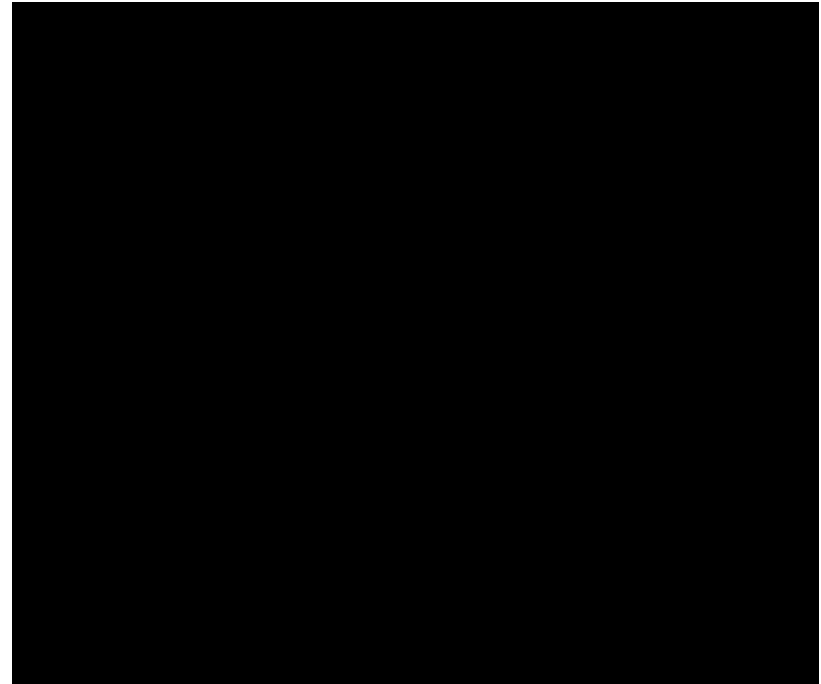
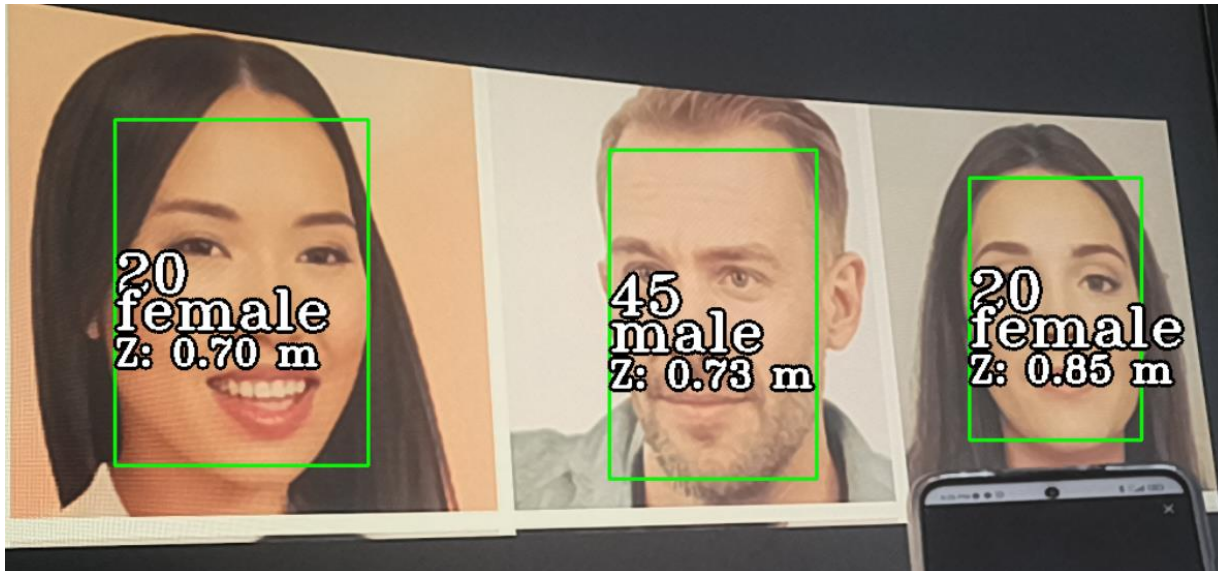


# Object Recognition



Examples of object recognition:

- Face recognition
- Age/gender, emotions recognition
- Optical Character Recognition
- License plate recognition



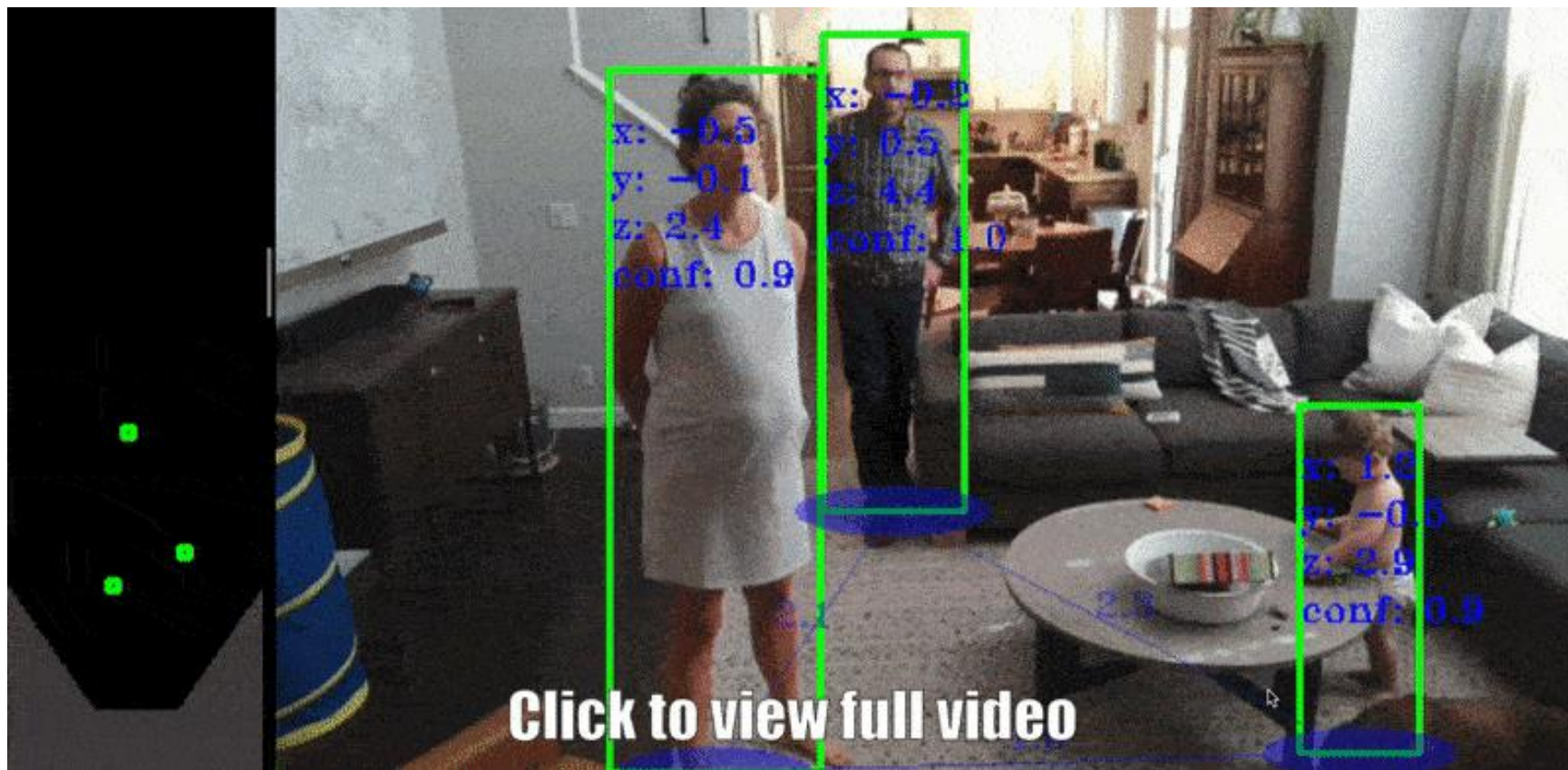
# Spatial Object Localization



Fuses existing 2D object detection models with depth perception

Used to detect:

- Objects 3D location
- Distances between objects
- Objects displacement over time (speed estimation)



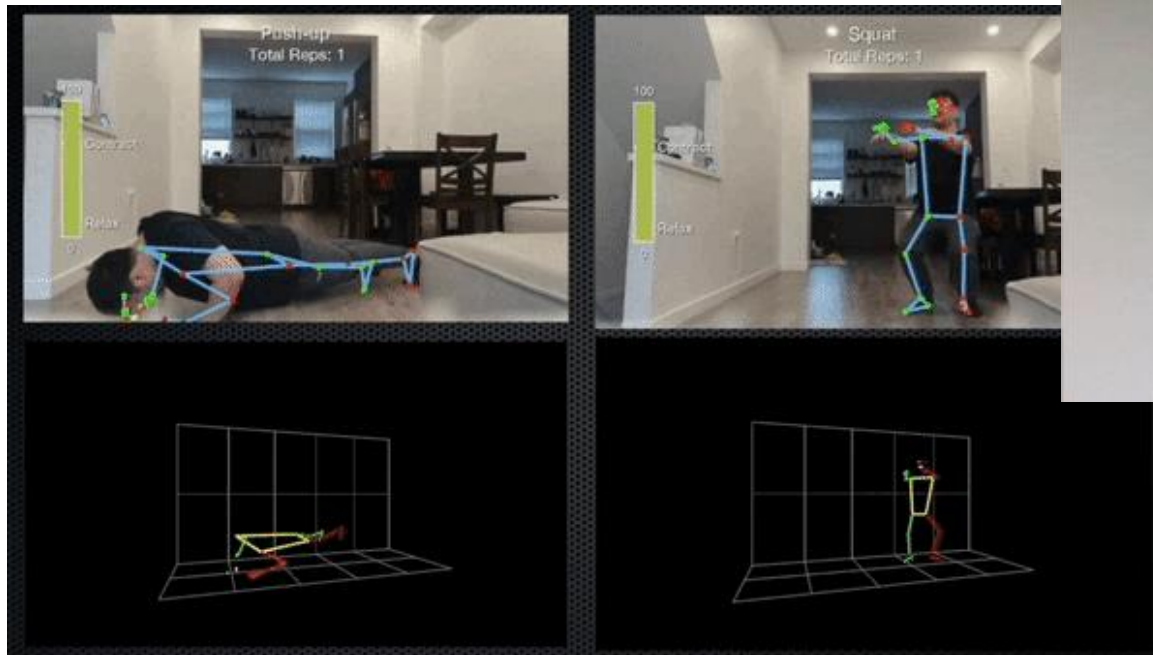


# Spatial Landmarks Localization



Fuses existing 2D landmarks detection models with depth perception. Used to detect:

- Landmarks 3D location
- Distances between landmarks
- Landmarks displacement over time (speed estimation)

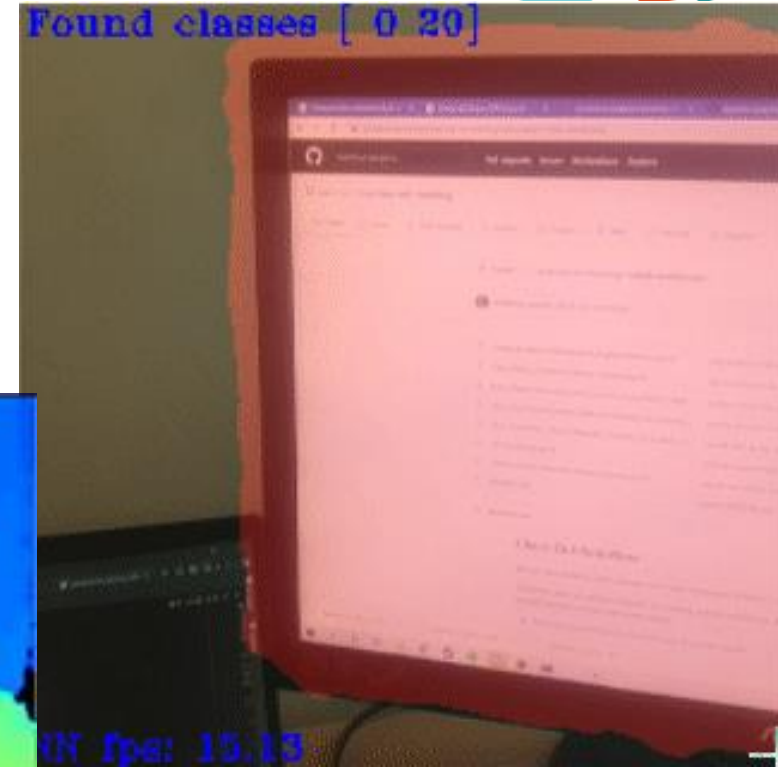
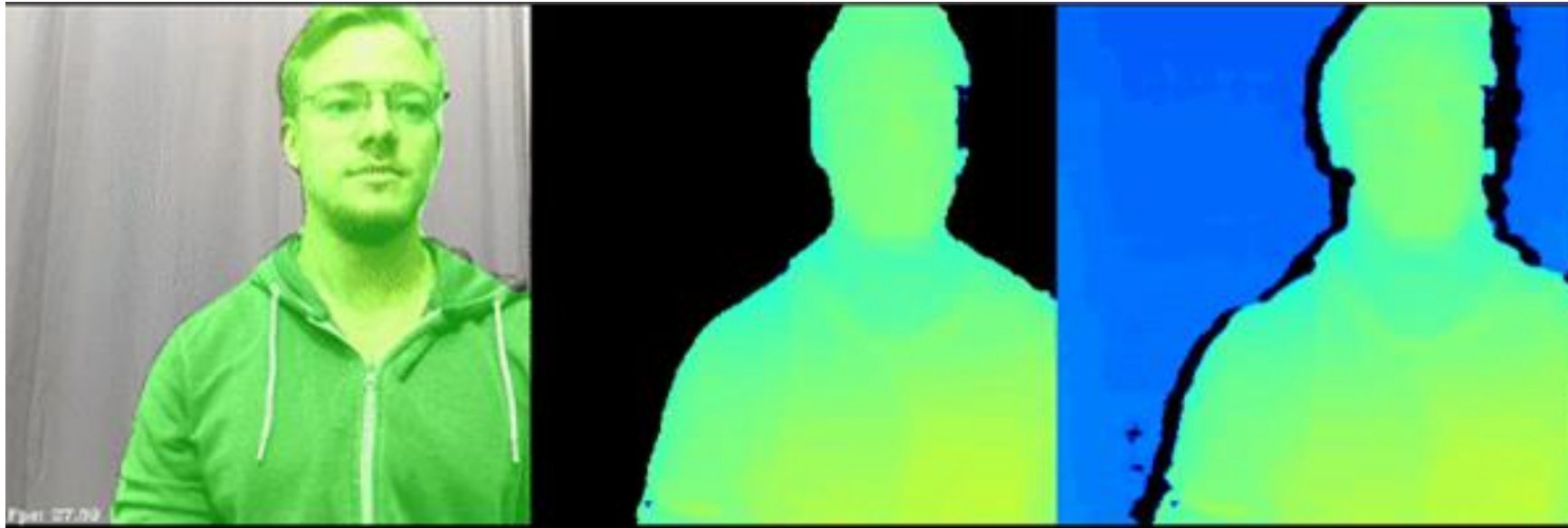




# Semantic Depth



- Fuses existing semantic segmentation models with depth perception
- Provides class for each depth point
- Extremely valuable in robotics navigation

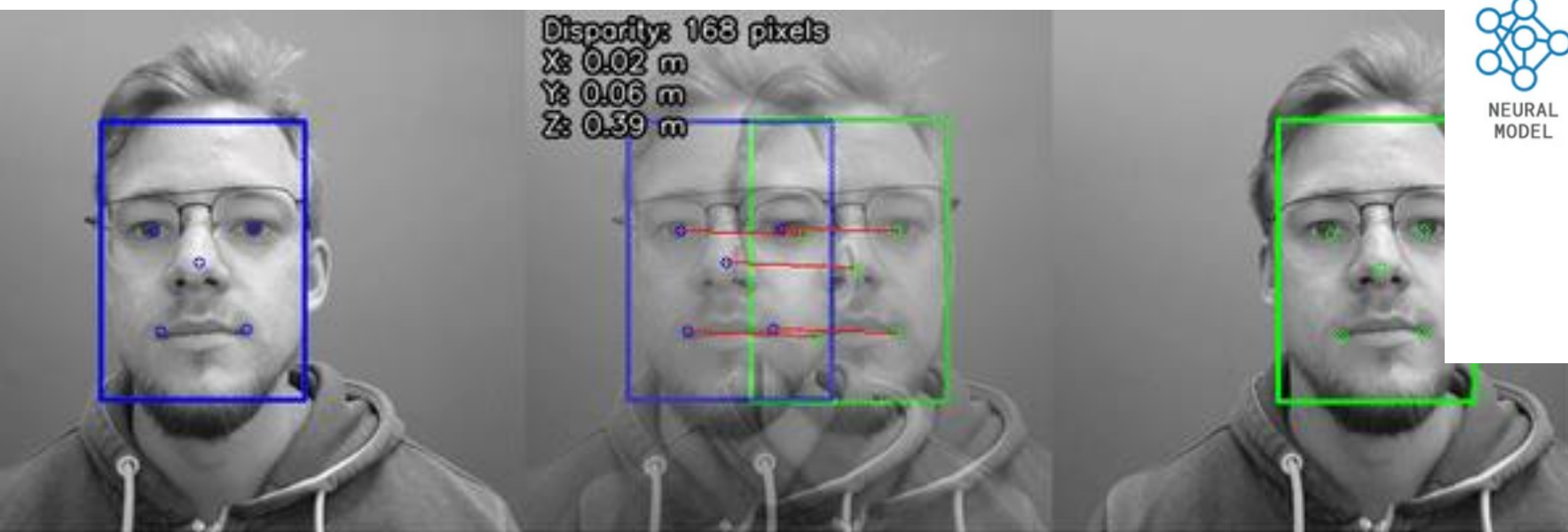
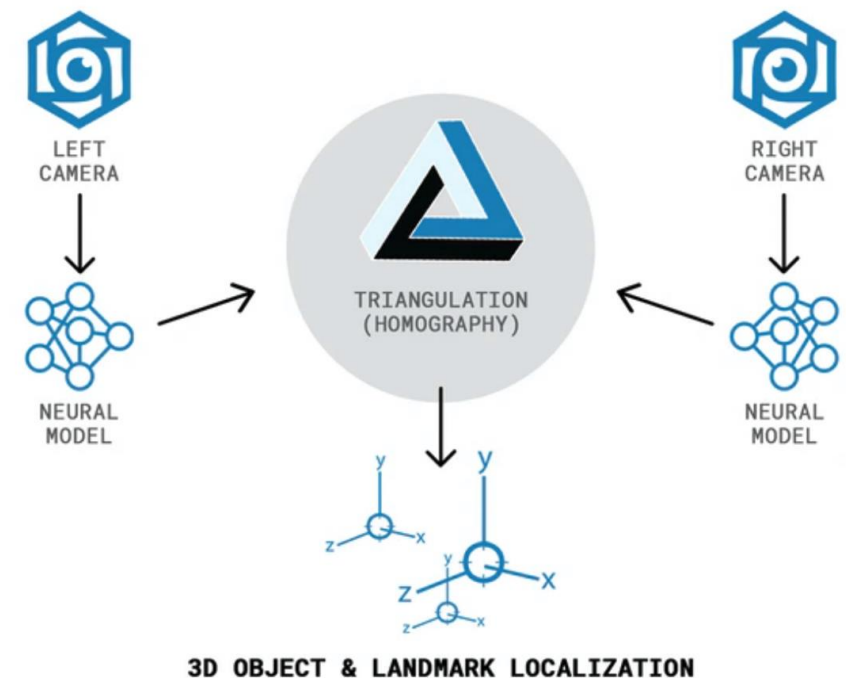


# Stereo Neural Inference



- Provides spatial coordinates of features
- Valuable for features that wouldn't have valid depth:
  - Holes: Screw holes
  - Light sources: Vehicle headlights
  - Objects with repetitive texture: Wire mesh fence
  - Thin objects: Tree branches

## STEREO NEURAL INFERENCE



# Open-source Hardware — Different Camera Models



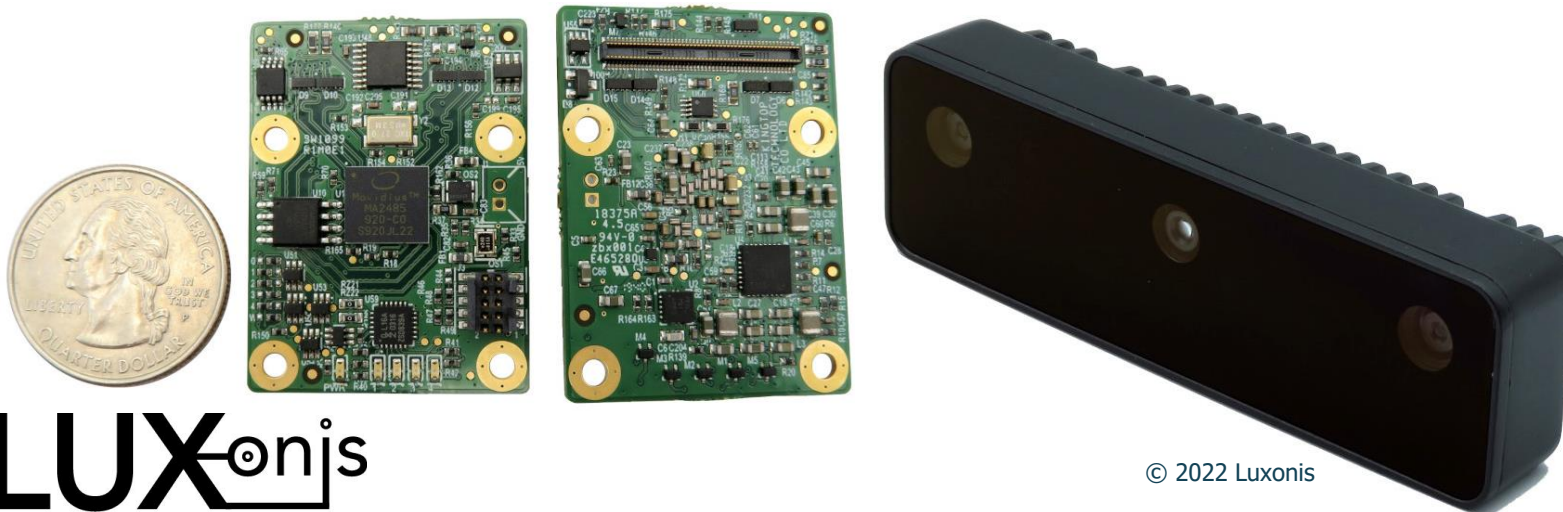


# Conclusion



## Key points

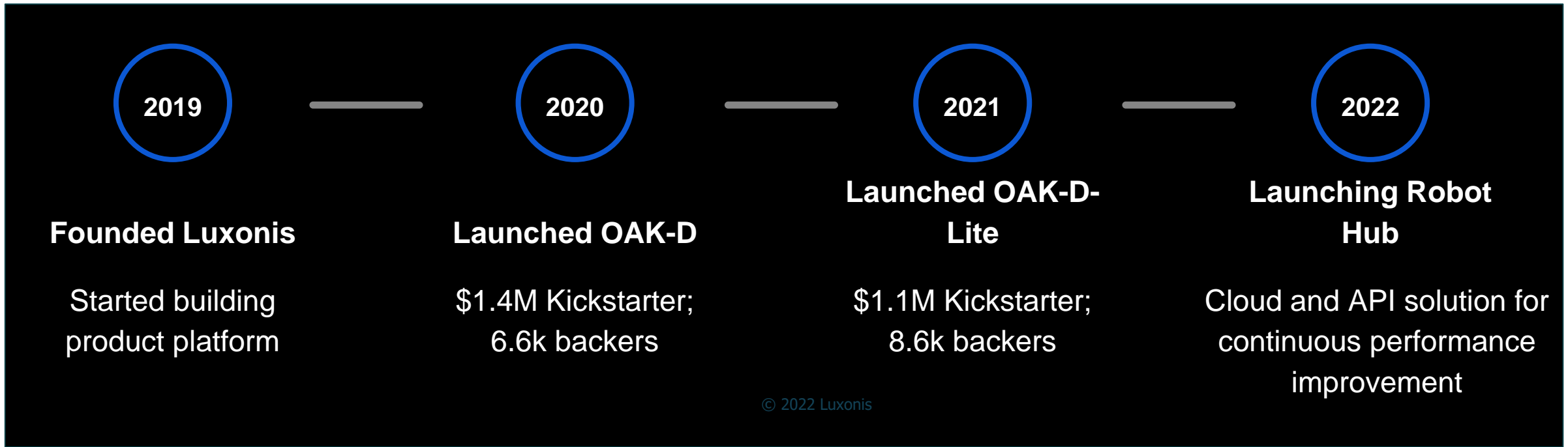
1. Run **any AI model** on the device itself, **fuse** results **with depth perception**
2. **Spatial AI** is the key to effective **robotic vision**
3. **OAK cameras** are performant, low-power and leverage the power of Spatial AI and CV on the edge
4. **Embed** Spatial AI technology into your own products **with System-on-Module (SoM)**



# About Luxonis



- **Business model:** Sell cameras / hardware
- **Team:** 30 employees, 25 engineers
- **Future:** Next generation of VPU in Q4 this year, 10x AI performance, onboard quad-core ARM
- **Goal:** Improve the engineering efficiency of the world





Bicycle safety device

<https://www.edge-ai-vision.com/2021/11/noroc-demonstration-of-from-behind-collision-detection-for-bicycle-riders/>

DepthAI documentation

<https://docs.luxonis.com/en/latest/>

Company web page

<https://www.luxonis.com/>

## 2022 Embedded Vision Summit

➔ Visit us in booth **#619!**