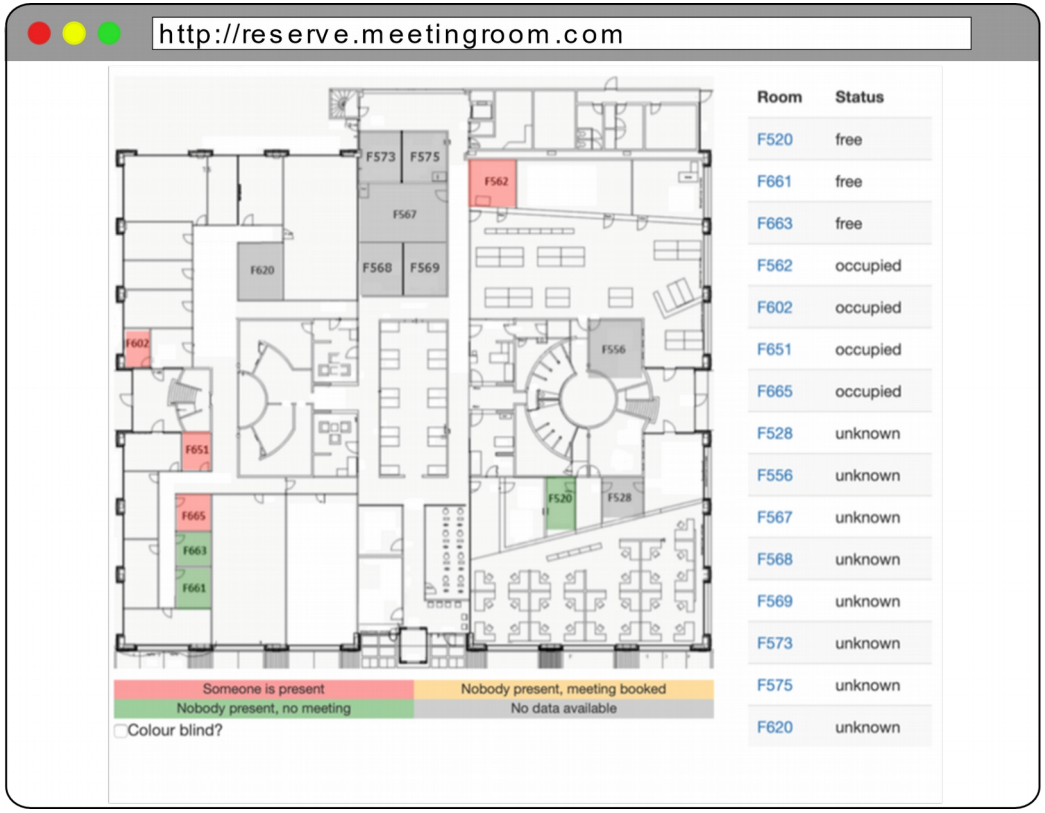
Anyone here? Smart embedded low-resolution omni-directional video sensor to measure room occupancy

Timothy Callemein, Kristof Van Beeck, Toon Goedemé

KU LEUVEN eavise

Current solution - PIR sensors



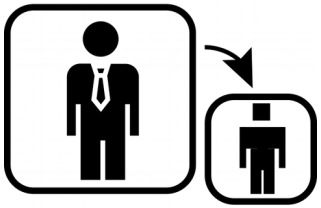
- ✓ No privacy issues
- ✓ Easy to install
- ✓ Cheap

- ✗ Requires a level of movement
- ✗ Binary output (YES / NO)
- ✗ No people count

Research Overview

Camera

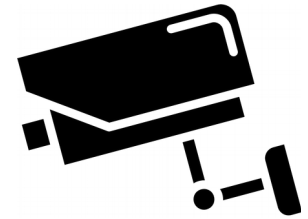
Performance

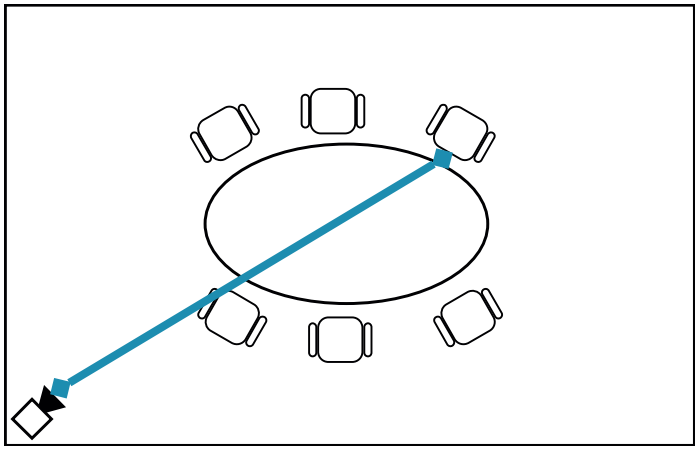


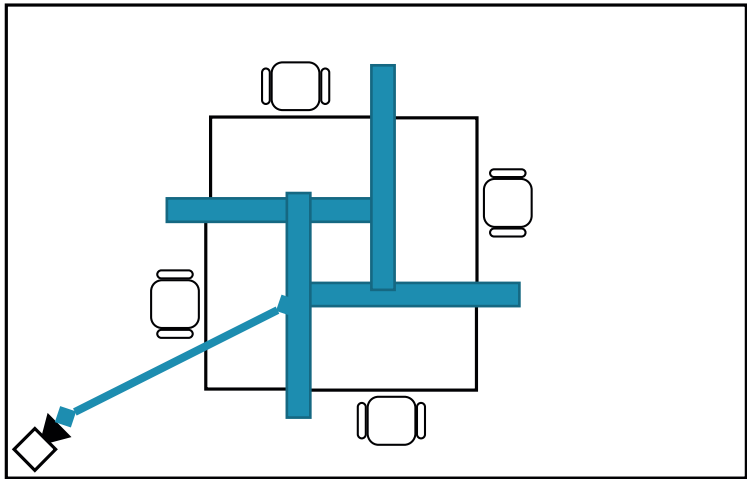
High Low

Privacy

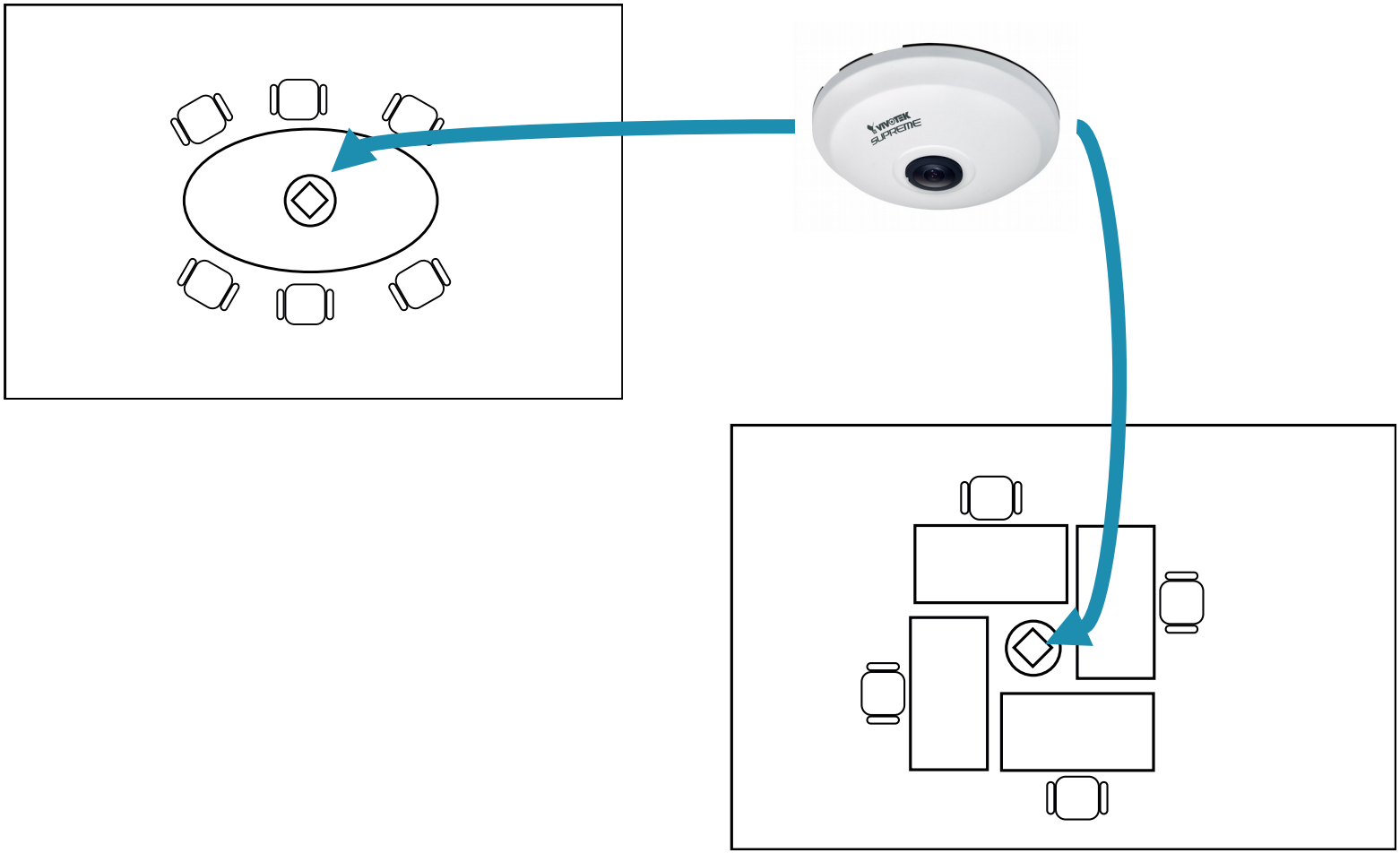
Camera



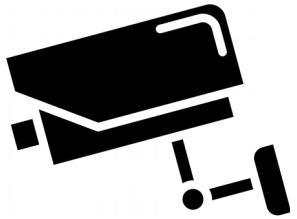


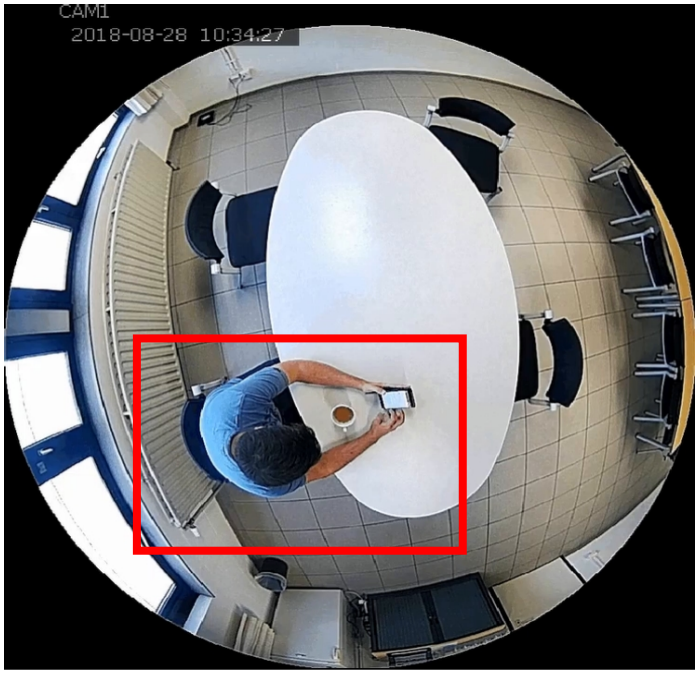


Camera



Camera



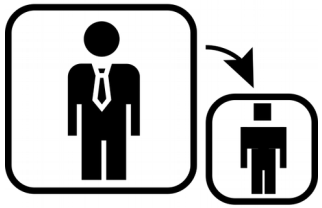


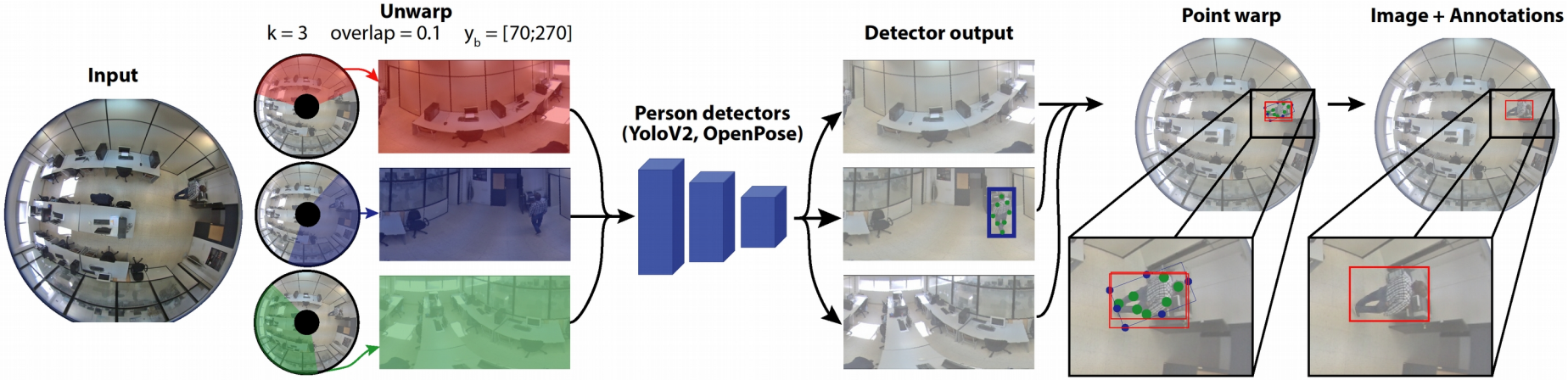
- ✓ Large field-of-view
- ✓ Single camera

- ✕ Image distortion
- ✕ Limited number of annotated datasets

Can we count people using these images?

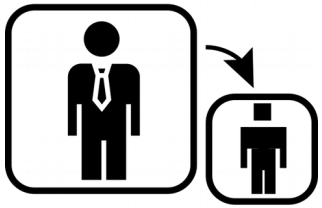
Approach – Generating data





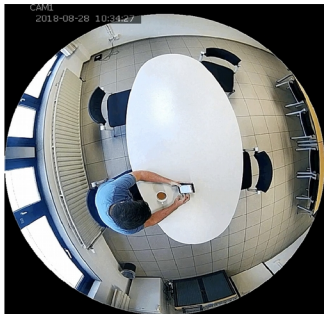
Zhe Cao et al. "OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields*" 2018 (CVPR)*
*Redmond et al. "*YOLO9000: Better, Faster, Stronger*" 2017 (CVPR)*

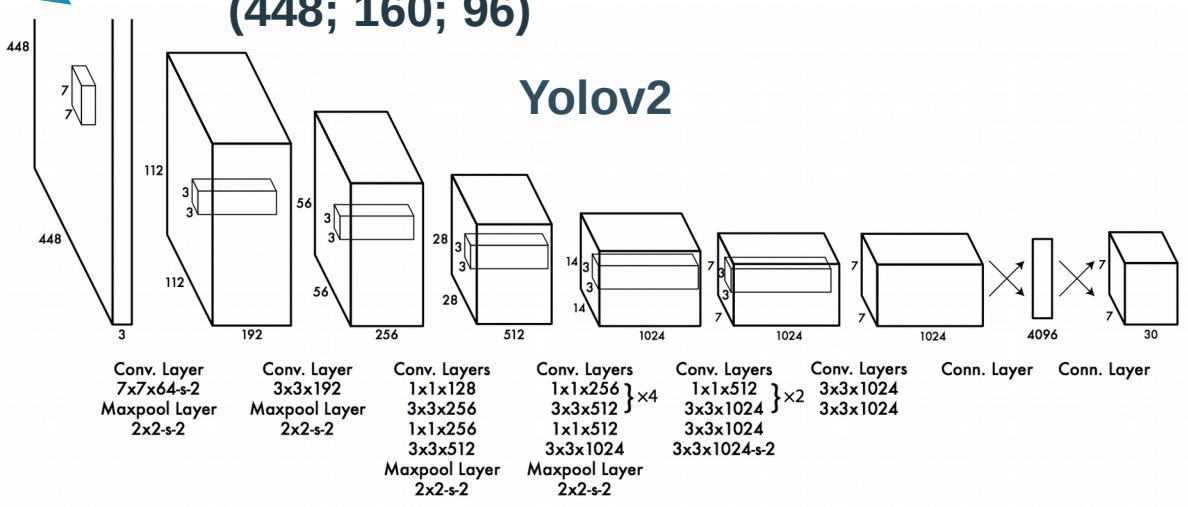
Approach – Lowering network resolution

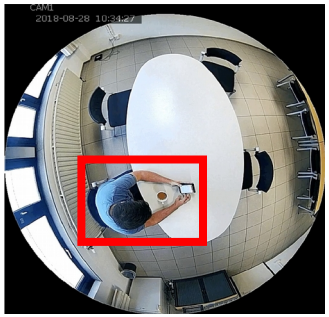


Decrease network resolution

(448; 160; 96)

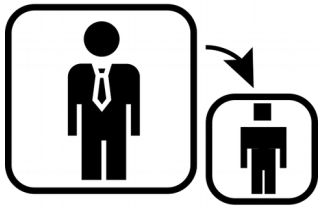






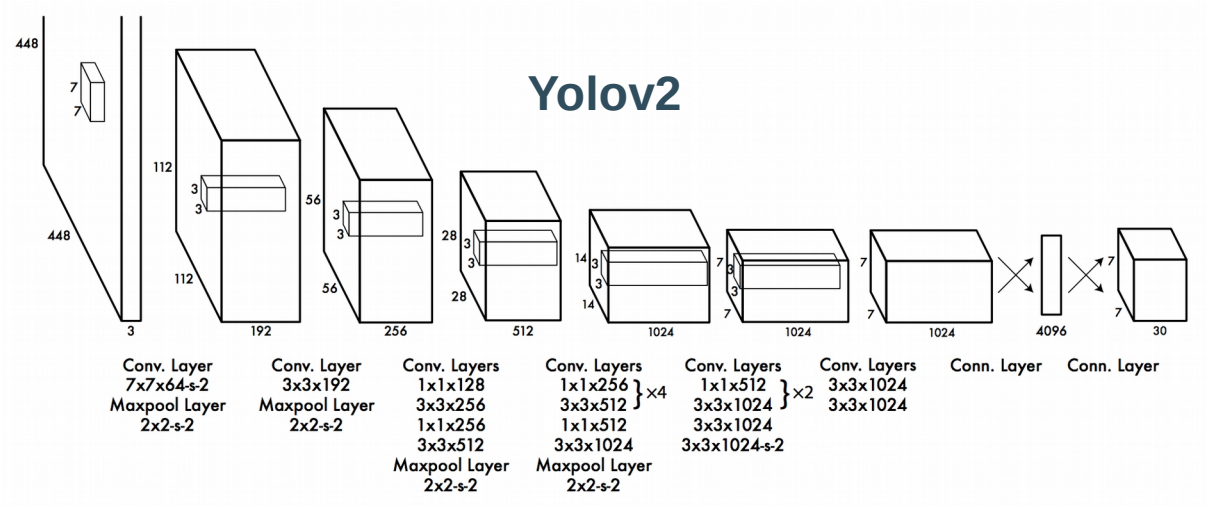
*Redmond et al. "*YOLO9000: Better, Faster, Stronger*" 2017 (CVPR)*

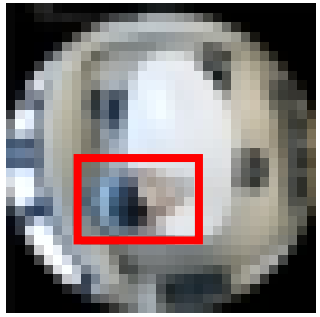
Approach – Lowering image resolution



Decrease image resolution

(64; 448; 32)





*Redmond et al. "*YOLO9000: Better, Faster, Stronger*" 2017 (CVPR)*

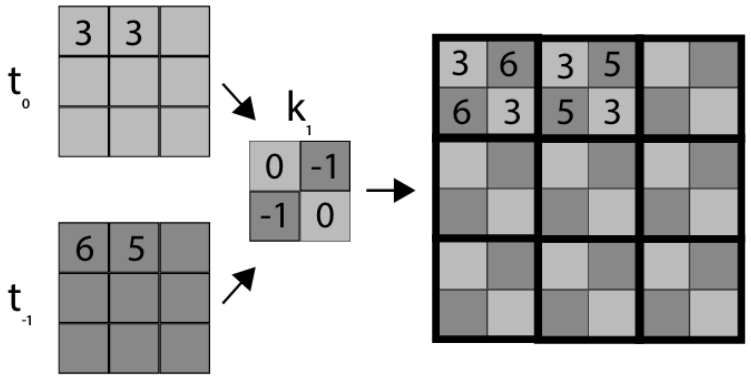
Approach – Using temporal data

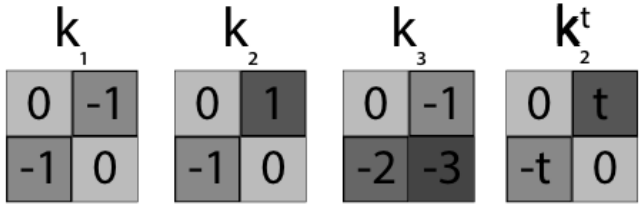
People move

vs.

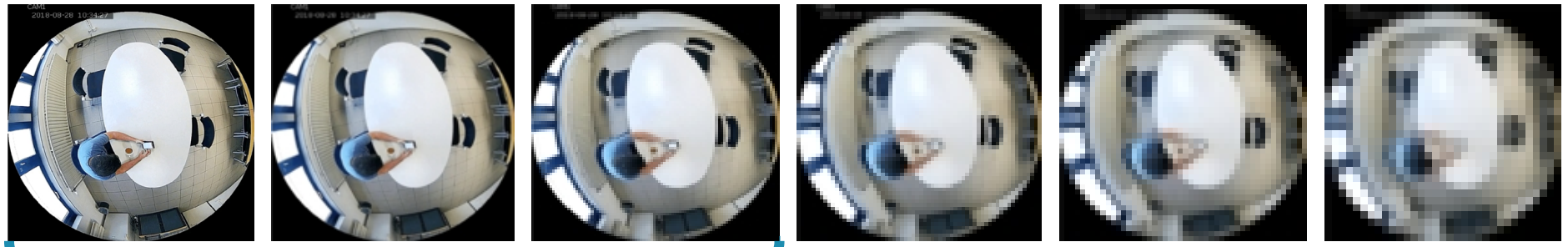
Static background

Can we improve our system by integrating temporal information?





Performance

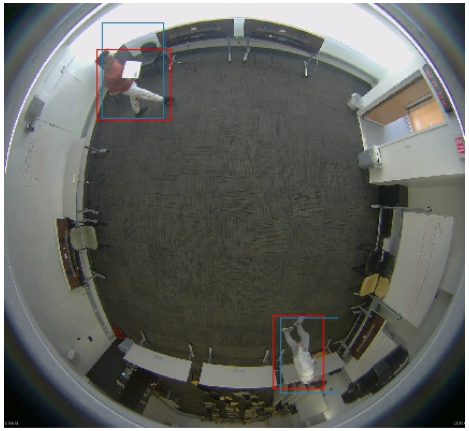


Device	Resolution	Seconds per Frame
Raspberry Pi 2	448	18.60
	160	3.60
	96	2.17
Raspberry Pi 3B	448	16.60
	160	2.96
	96	1.83
Raspberry Pi 3B+	448	11.72
	160	2.07
	96	1.30

Results

Test on model trained with generated labels

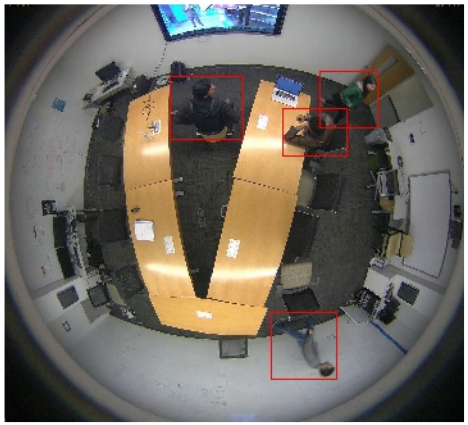
- Different network resolutions
- Different image resolutions



MIRROR A

PIROPO

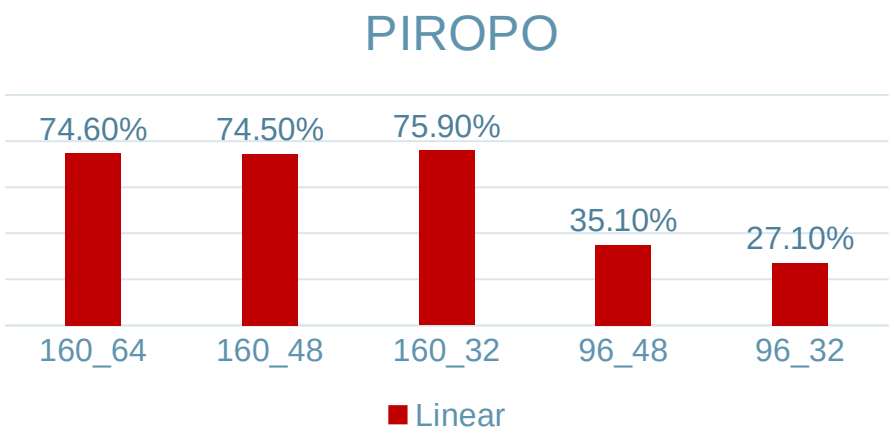
PRIVATE OFFICE

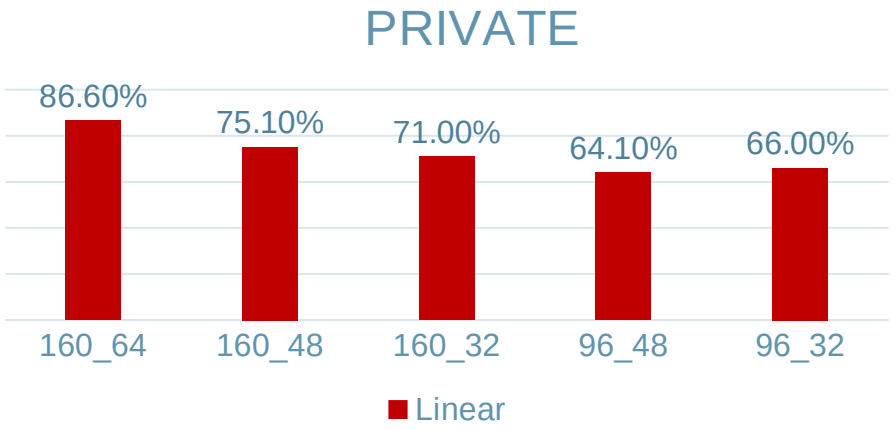


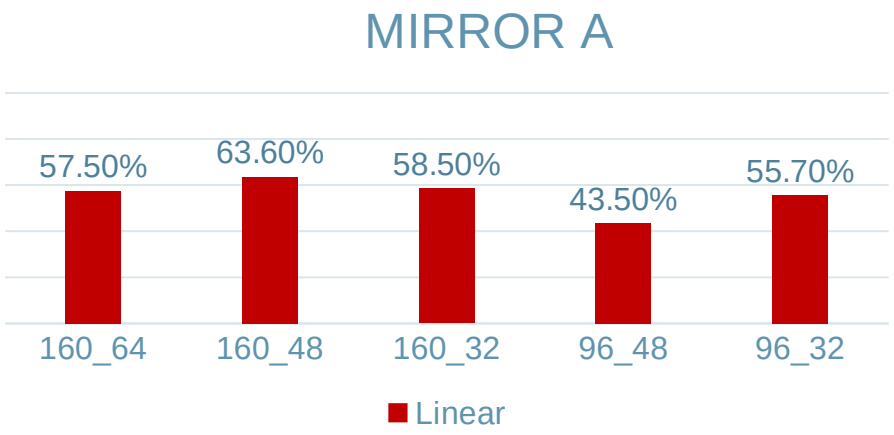
MIRROR B

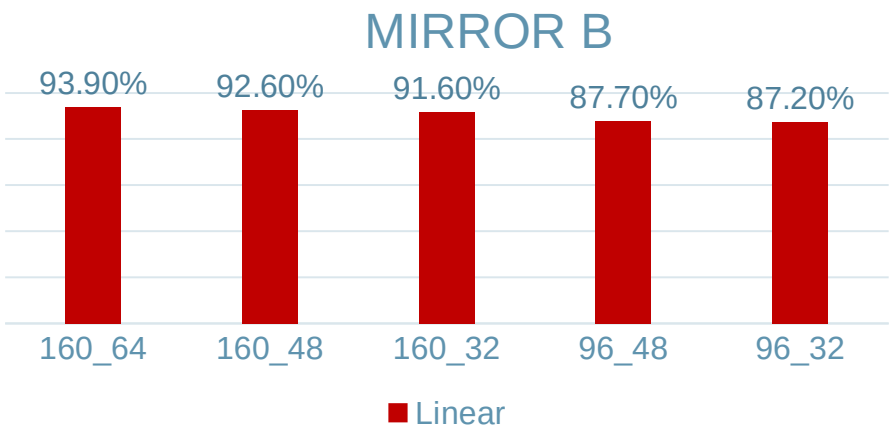
PIROPO *- https://www.gti.ssr.upm.es/research/gti-data/databases*
MirrorChallenge *- https://www.hcd.icat.vt.edu/mirrorworlds/challenge/index.html*

Results

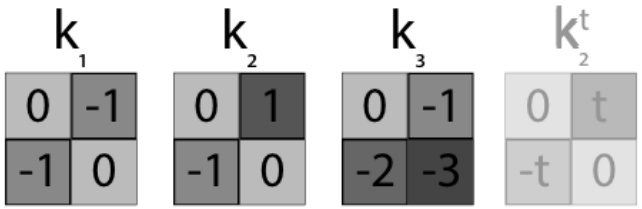

PIROPO
74.60%
74.50%
75.90%
35.10%
27.10%
160_64
160_48
160_32
96_48
96_32
Linear


PRIVATE
86.60%
75.10%
71.00%
64.10%
66.00%
160_64
160_48
160_32
96_48
96_32
Linear


MIRROR A
57.50%
63.60%
58.50%
43.50%
55.70%
160_64
160_48
160_32
96_48
96_32
Linear

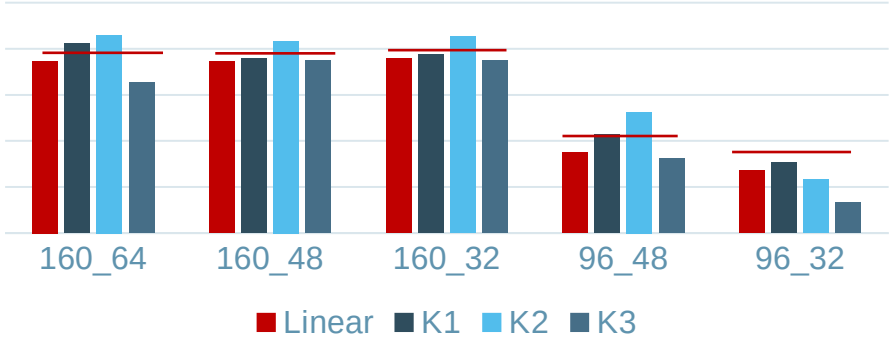

MIRROR B
93.90%
92.60%
91.60%
87.70%
87.20%
160_64
160_48
160_32
96_48
96_32
Linear

Results



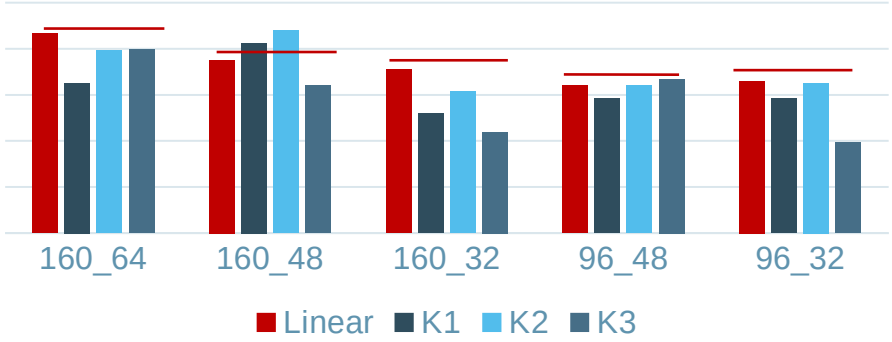
PIROPO

High movement level



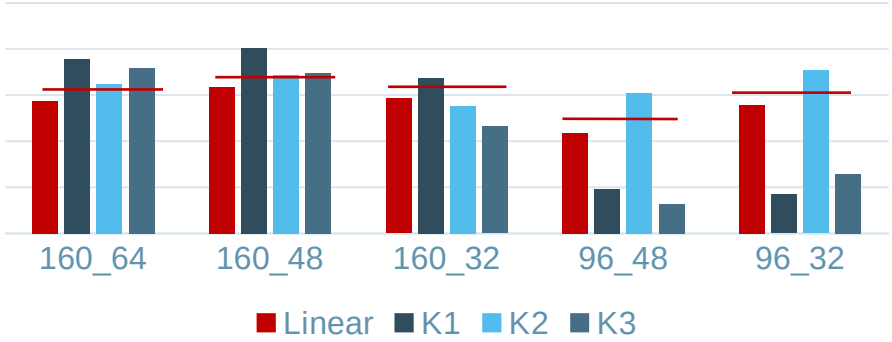
PRIVATE

Low movement level



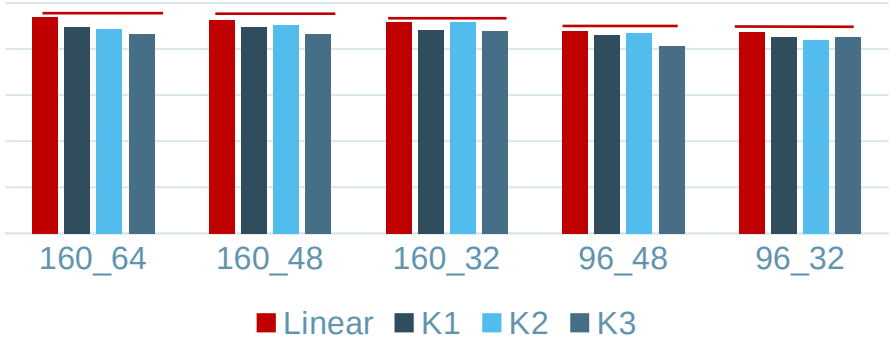
MIRROR A

Medium movement level

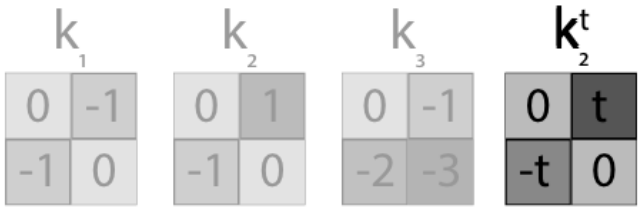


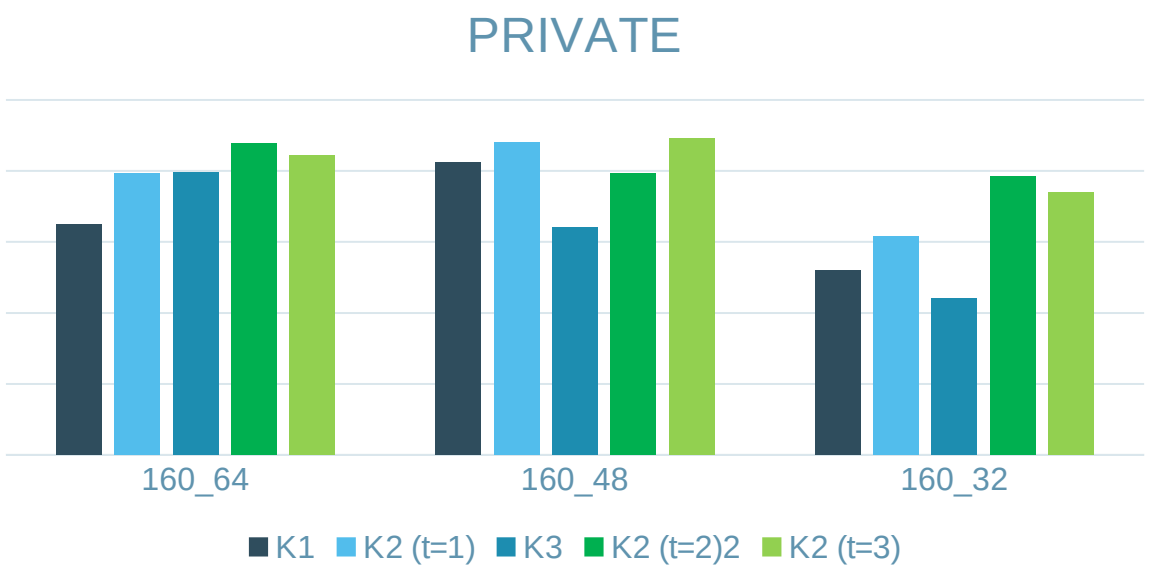
MIRROR B

Medium movement level



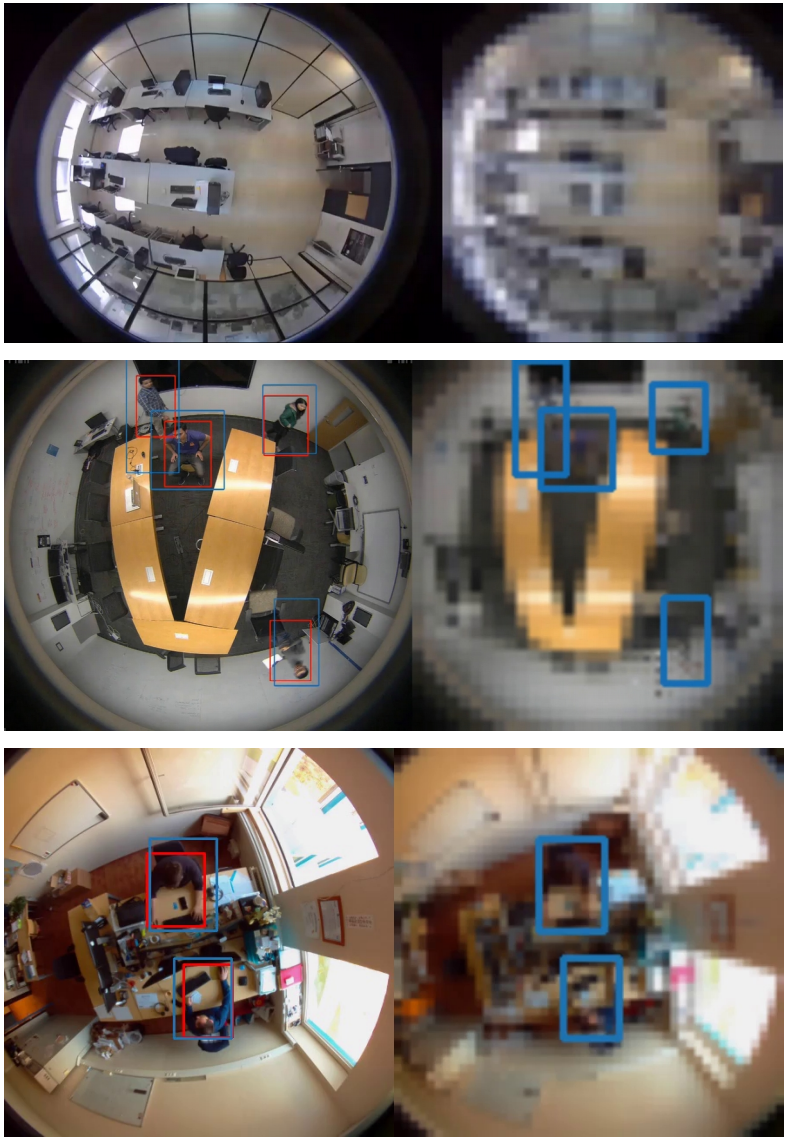
Results


k1
0 -1
-1 0
k2
0 1
-1 0
k3
0 -1
-2 -3
k2t
0 t
-t 0


PRIVATE
160_64
160_48
160_32
K1
K2 (t=1)
K3
K2 (t=2)2
K2 (t=3)

Conclusion

- Training possible with automatically generated annotations
- Can run on embedded hardware
- Already good performance
- Image resolution of 32 pixels



Future Work

- Improving label generation
- Influence of large room changes?
- What when the room gets bigger?

Thank you for your attention!
Questions?

timothy.callemein@kuleuven.be
phd.callemein.be – www.eavise.be