

# DeepDish: Multi-Object Tracking

With an off-the-shelf Raspberry Pi

Matthew Danish

Justas Brazauskas

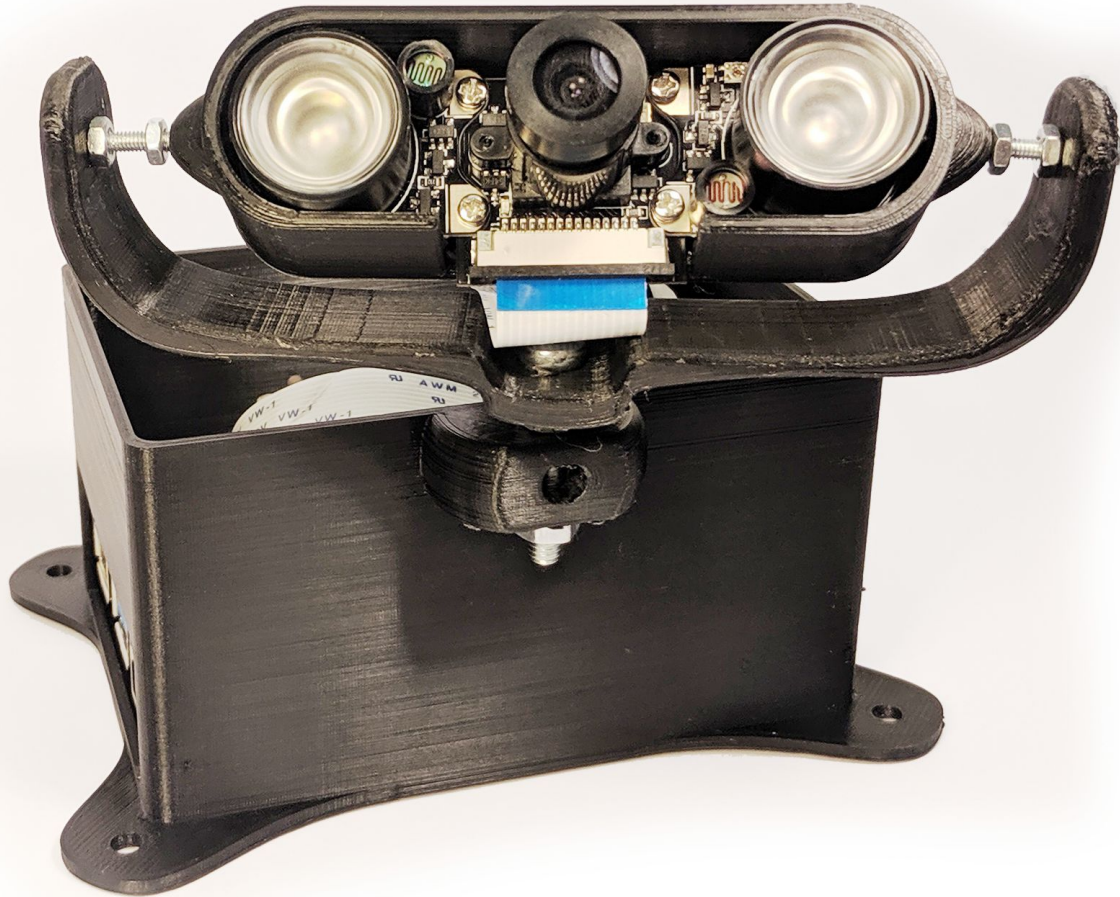
Rob Bricheno

Ian Lewis

Richard Mortier

*Cambridge University Department of Computer Science & Technology*

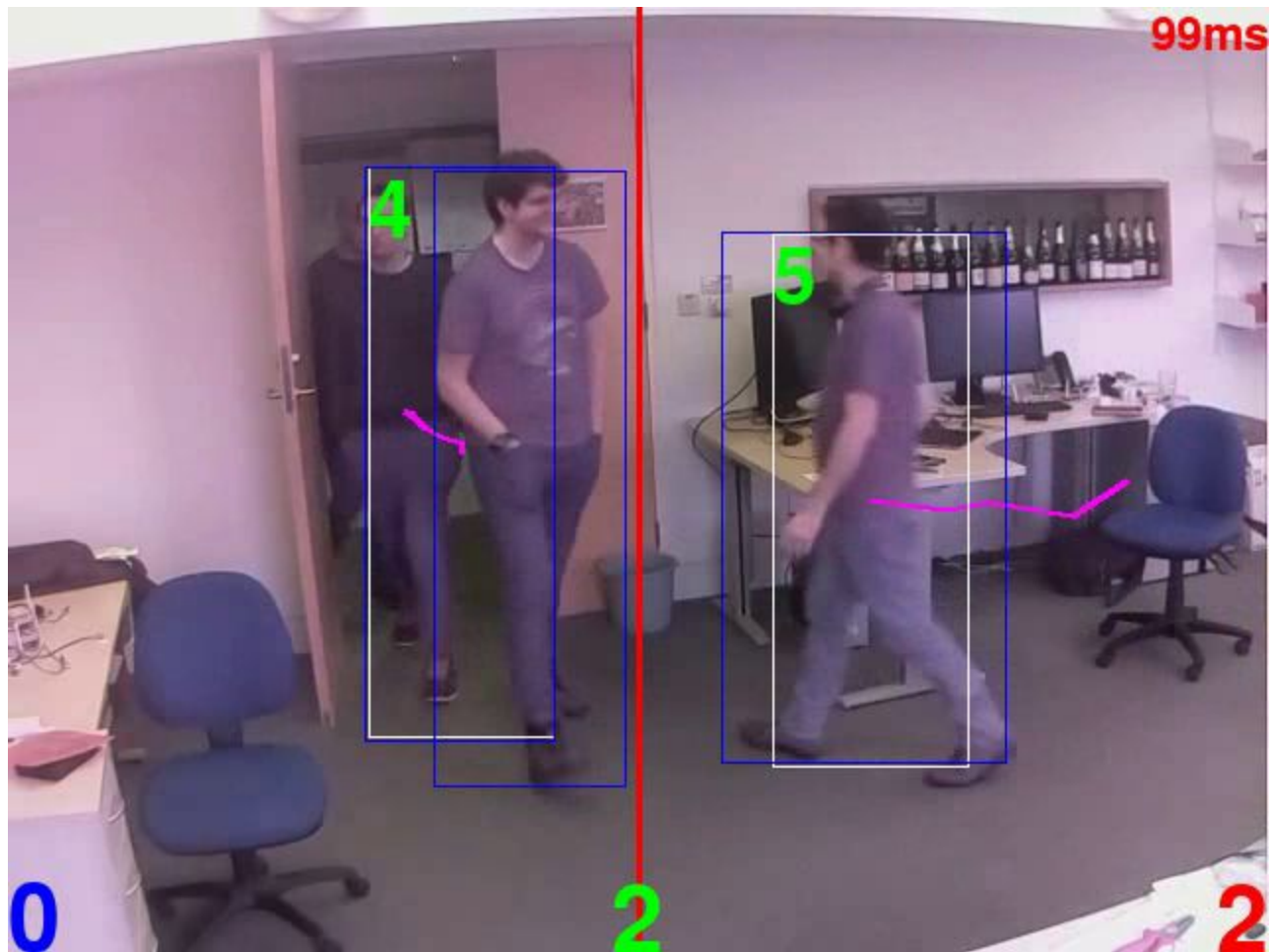
*Centre for Digitally Built Britain*



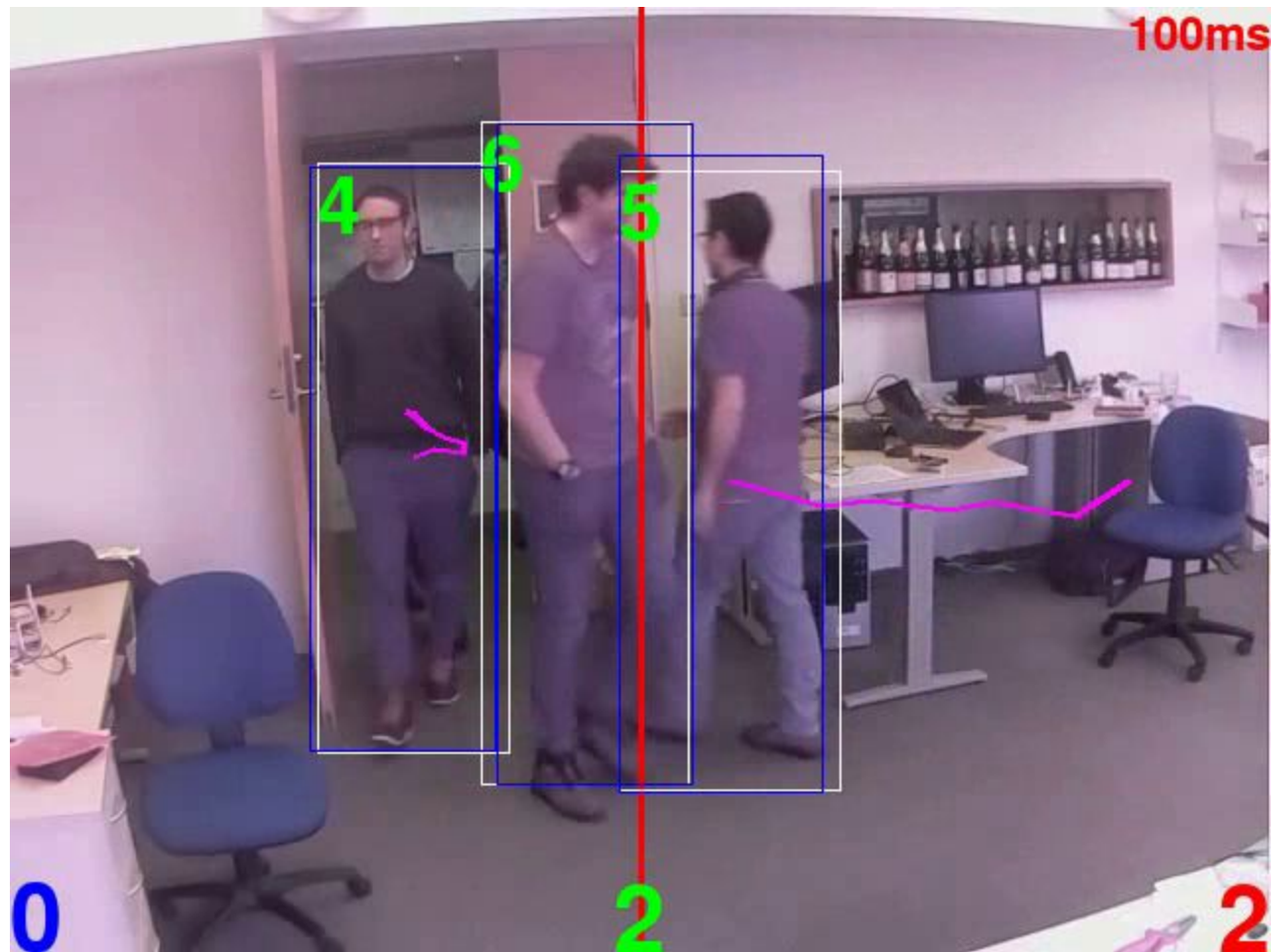
- Raspberry Pi 4B (4GB RAM)
  - Standard camera (plus infrared lights)
  - Fan SHIM cooling
  - ~£80 total
- 
- (cool 3D-printed case: extra)

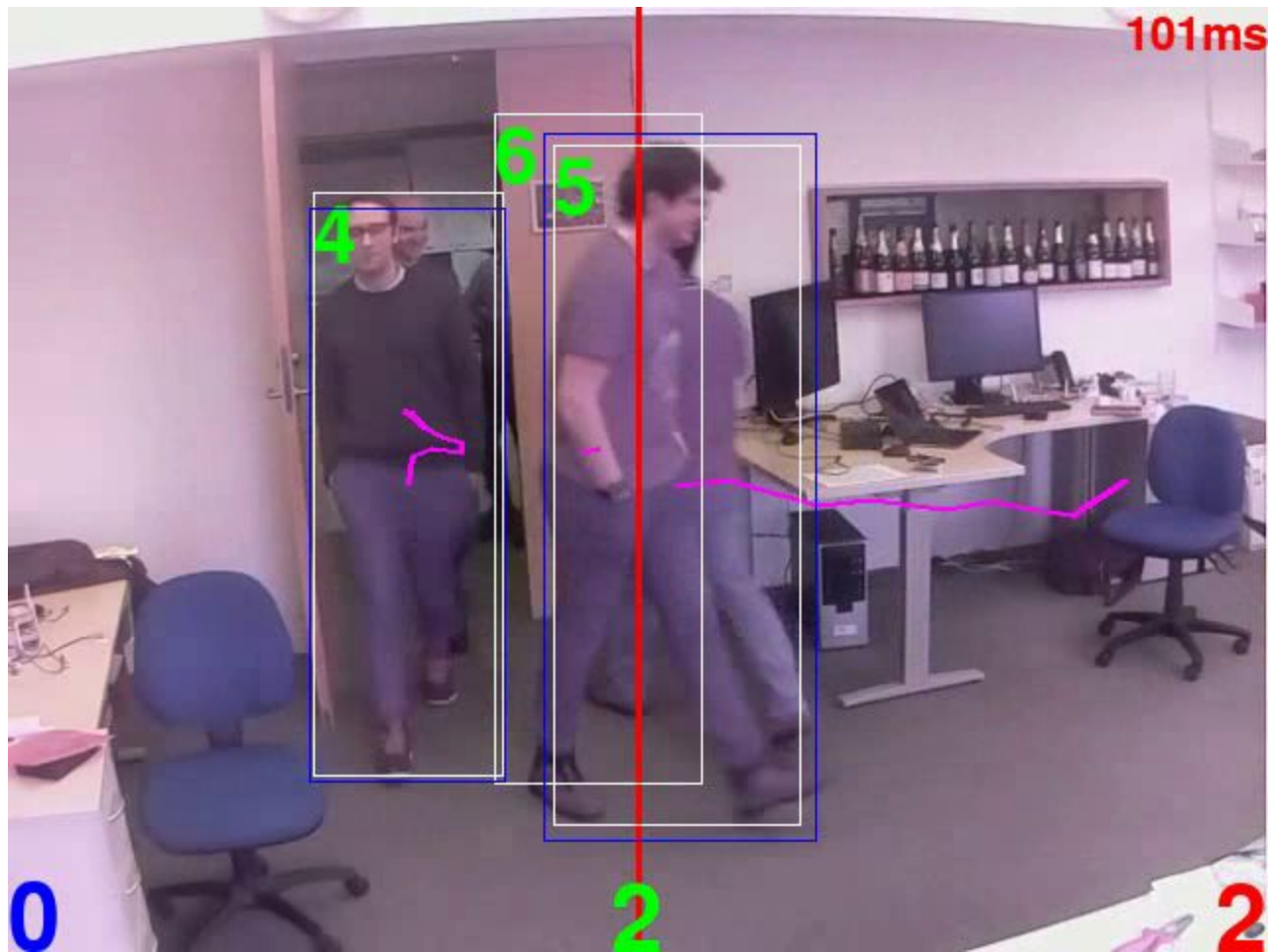
# Tracking by detection

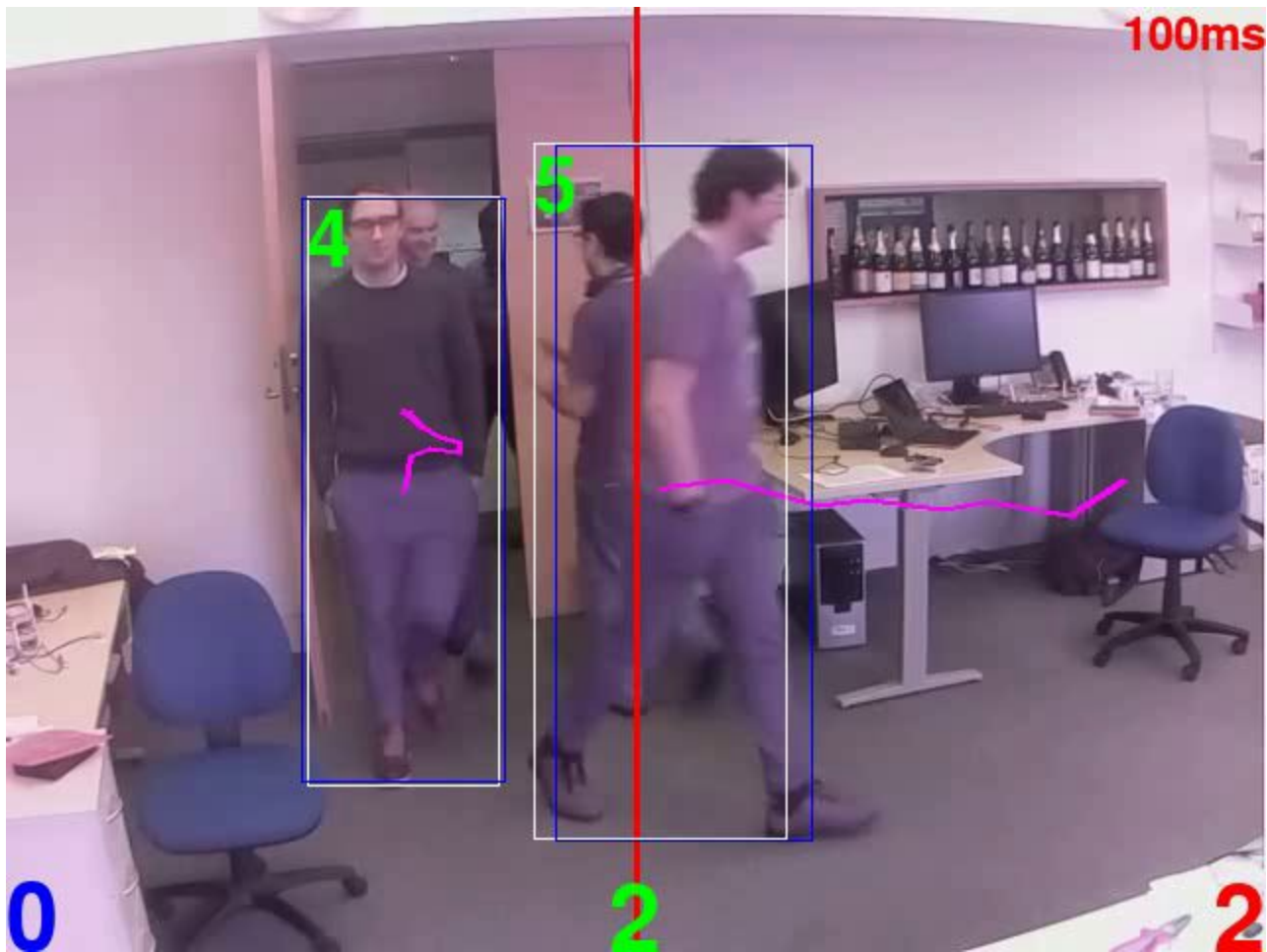
- Object detection:
  - MobileNet (v1)
  - MobileNet (v2)
  - YOLOv3
- Feature encoding: DeepSORT (Wojke 2017)
  - Trained on Motion Analysis and Re-identification Set (MARS)
  - Cosine Distance Metric training
  - Combine with Kalman Filter tracking



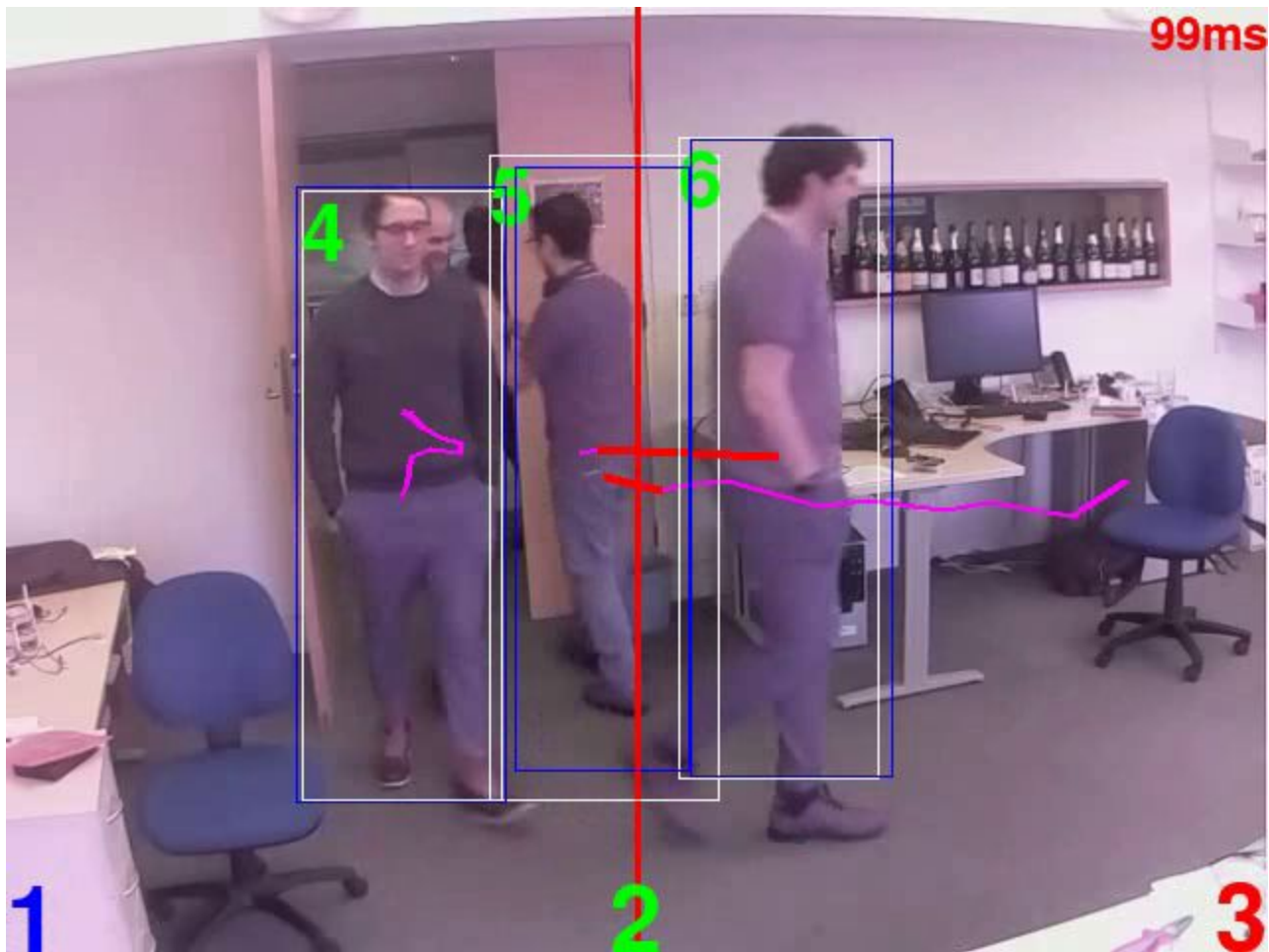












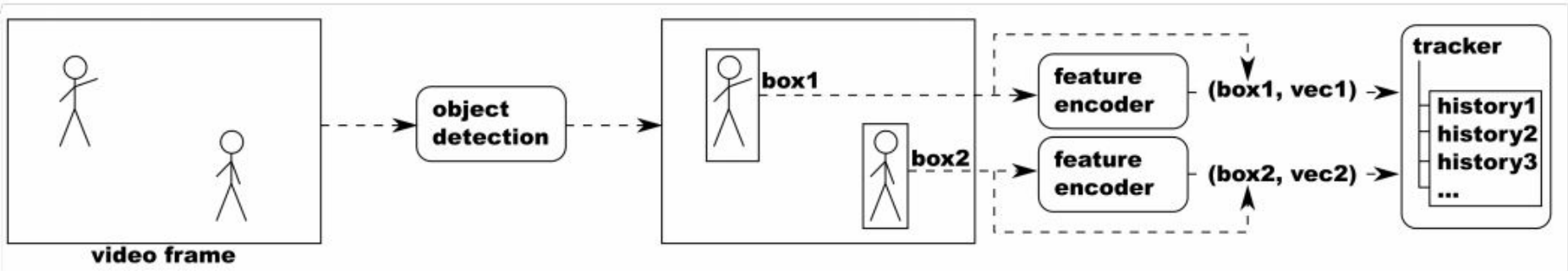


Figure 2: Tracking-by-detection pipeline.



Figure 6: Screenshot from the 'Plaza' test video.

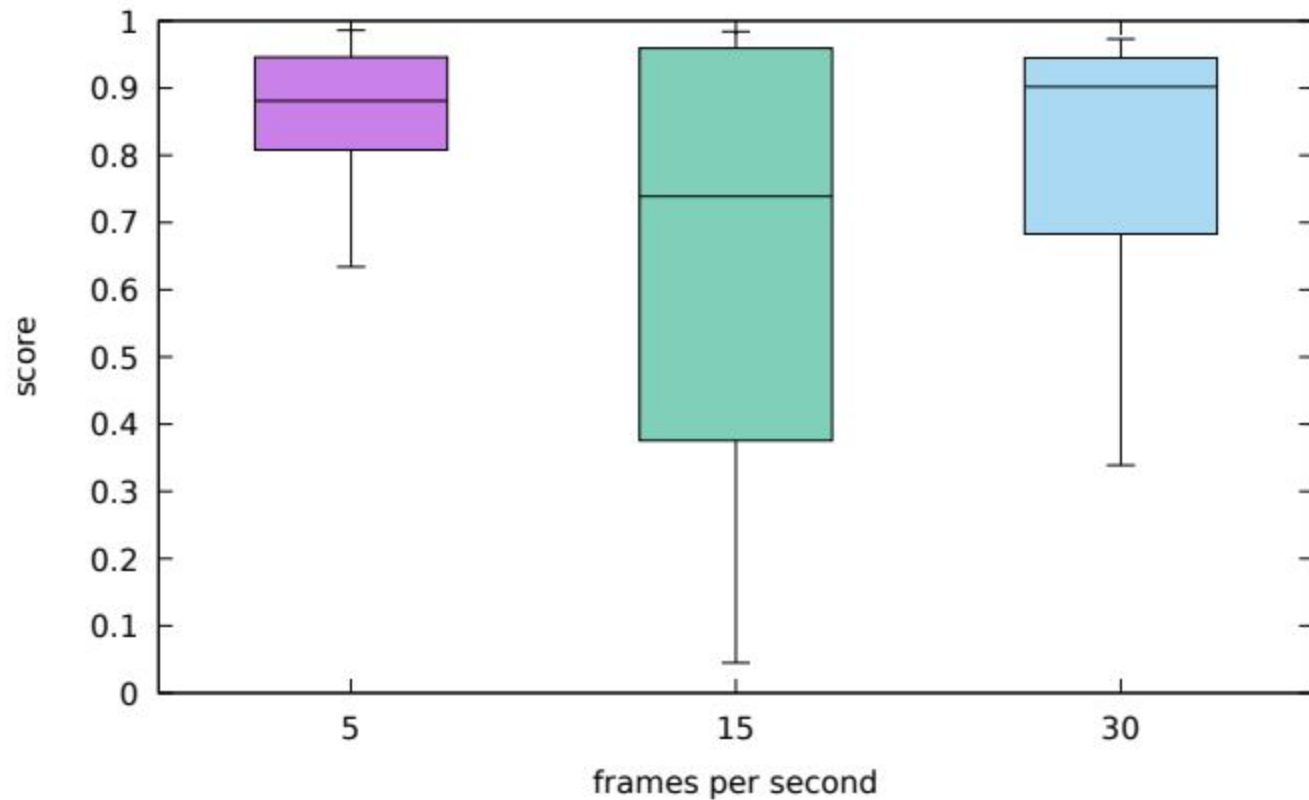
# Scoring methodology

- 'Office' video and 'Plaza' video
- Establish ground-truth by hand:
  - Count crossings in each direction, grouped into 10 second intervals
  - Subtract one direction from the other
  - Forms 'ground-truth vector'
- Testbed software performed same operation automatically
- Score is  $1 - \text{cosine distance}$  between test vector and ground-truth

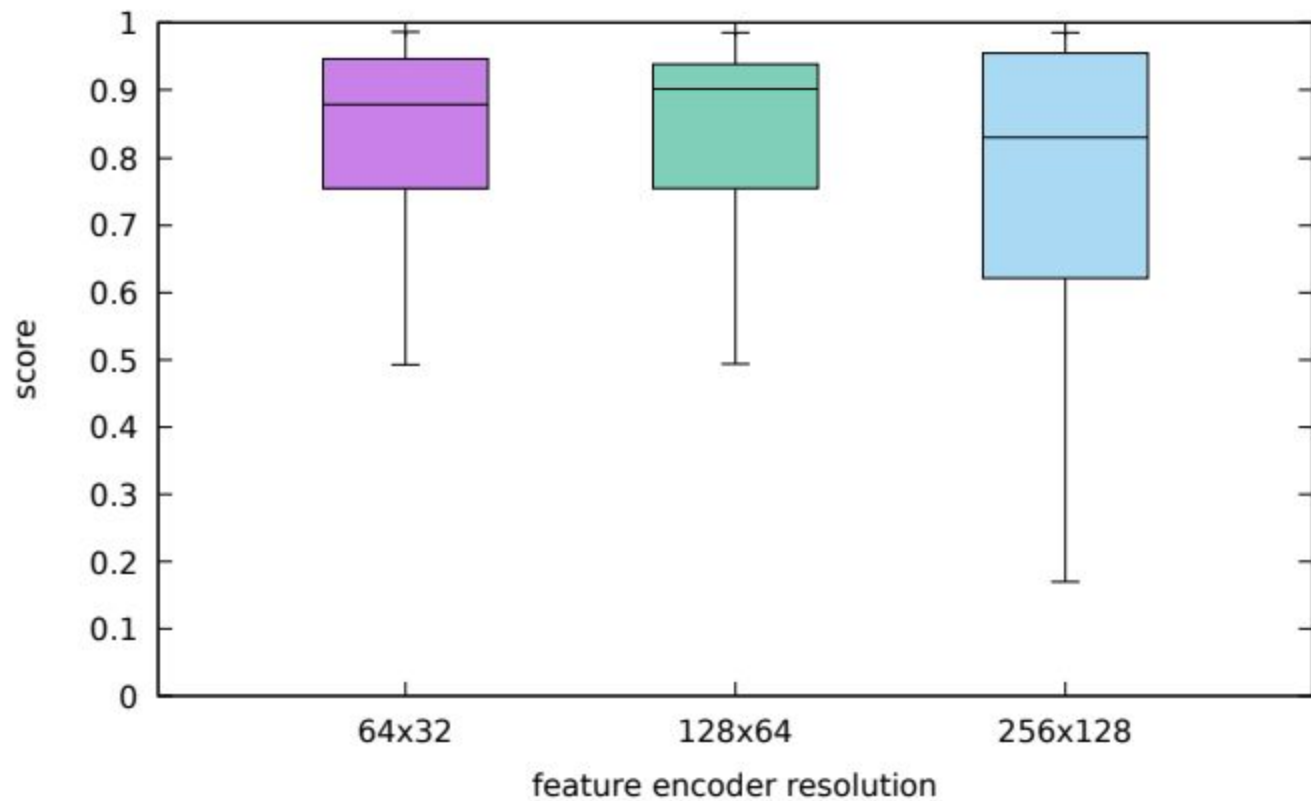
# Accuracy results

Table 2: Accuracy of counting: a selection of test results.

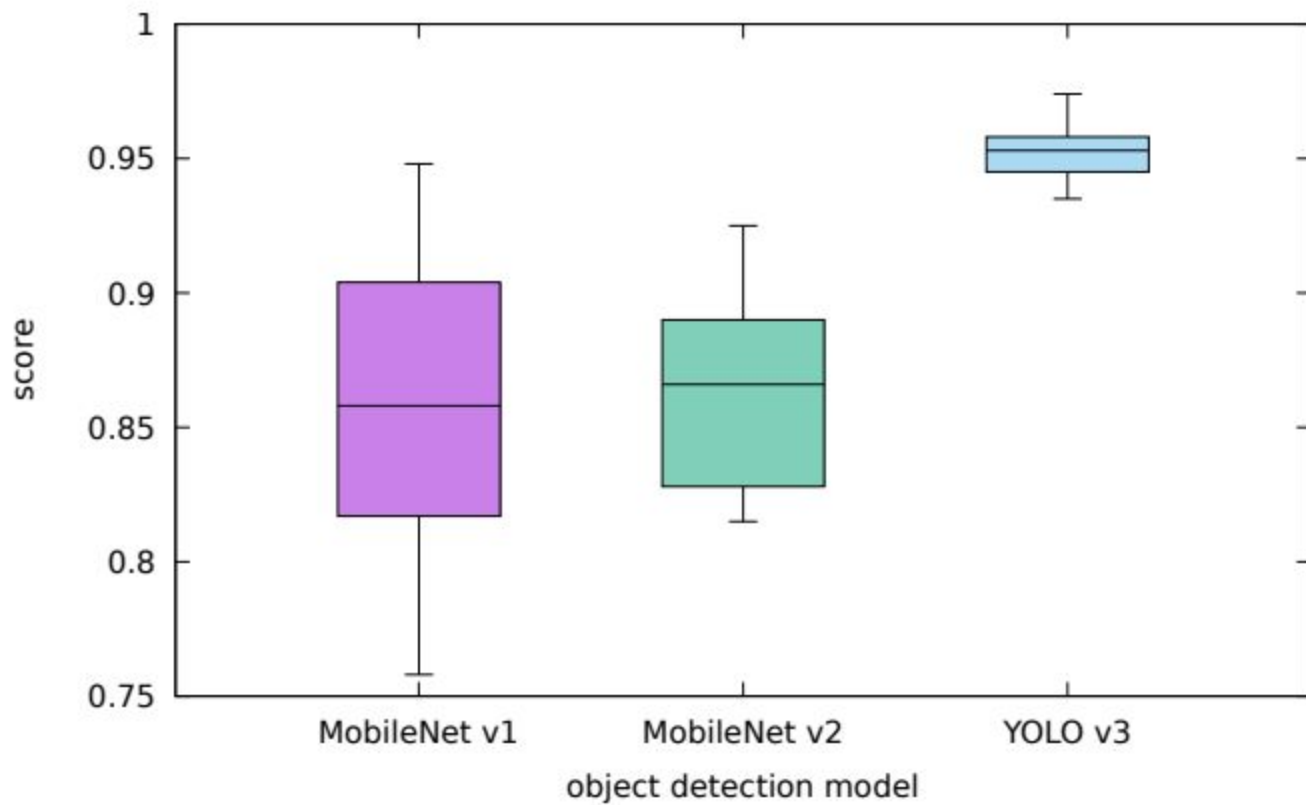
Score	Test name	Object-detector	Feature encoder	FPS	<i>max-cos-distance</i>	<i>nms-max-overlap</i>
0.968	Office	YOLO v3	128x64	5	0.3	0.6
0.948	Office	MobileNet v1	64x32	5	0.6	0.6
0.938	Office	MobileNet v2	64x32	30	0.9	0.6
0.906	Office	MobileNet v1	128x64	5	0.3	0.3
0.878	Office	MobileNet v2	64x32	5	0.9	1.0
0.794	Office	MobileNet v1	128x64	30	0.9	0.6
0.656	Office	MobileNet v2	256x128	30	0.6	0.6
0.424	Office	MobileNet v2	256x128	15	0.01	1.0
0.986	Plaza	YOLO v3	64x32	5	0.6	0.6
0.903	Plaza	MobileNet v1	128x64	15	0.9	0.6
0.880	Plaza	MobileNet v2	128x64	5	0.3	0.6
0.843	Plaza	MobileNet v1	64x32	5	0.6	0.8
0.839	Plaza	MobileNet v2	256x128	5	0.9	0.6
0.815	Plaza	MobileNet v2	64x32	5	0.6	0.3
0.713	Plaza	MobileNet v1	64x32	15	0.01	1.0
0.596	Plaza	MobileNet v2	128x64	30	0.9	1.0



**Figure 7: The effect of frame-rate on score.**

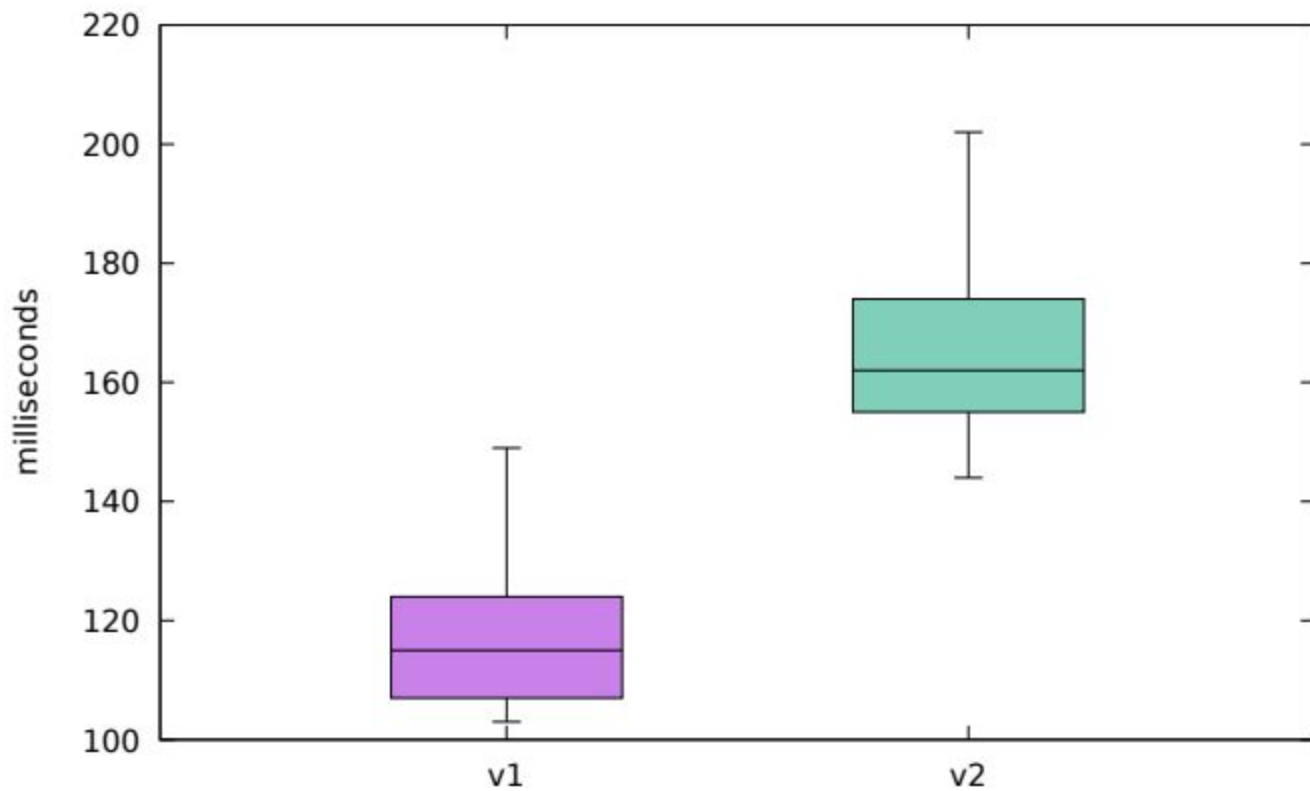


**Figure 10: The effect of feature encoder resolution on score.**

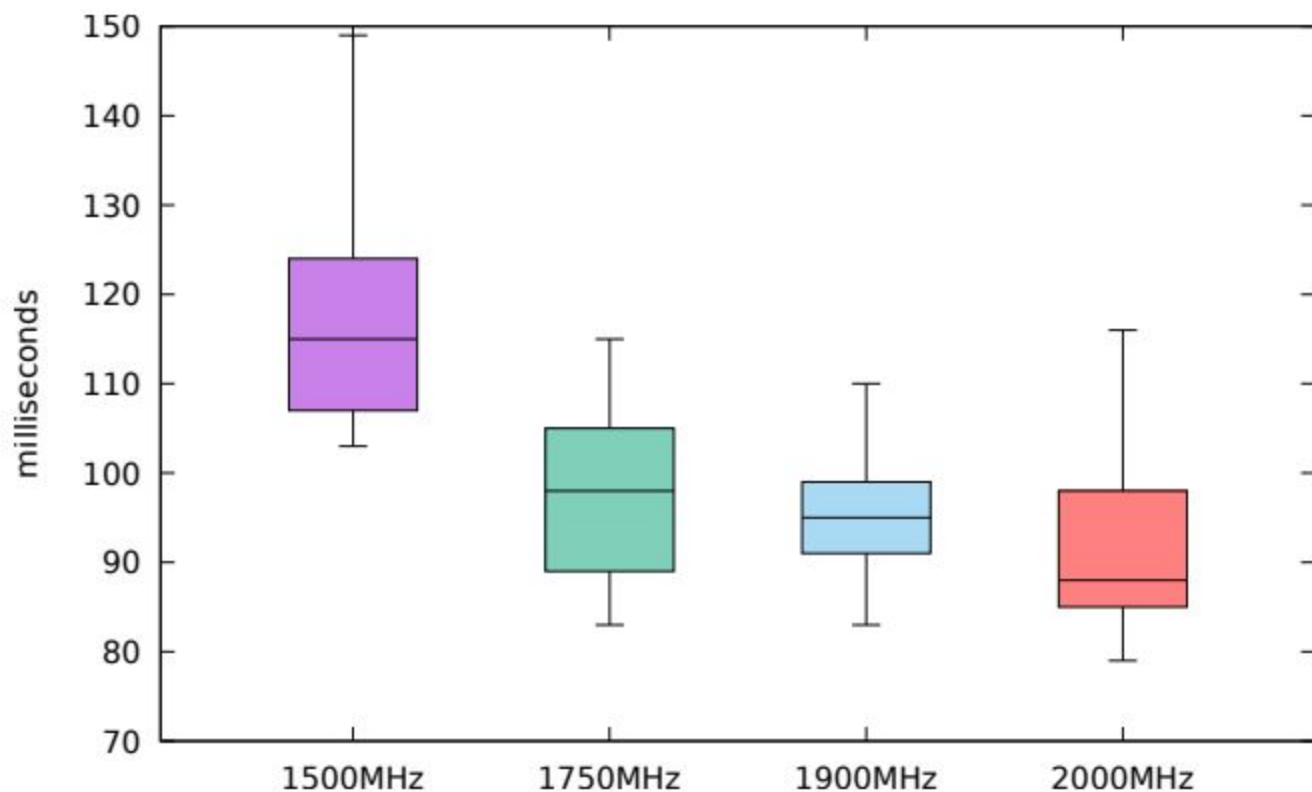


**Figure 11: The effect of model type on score.**





**Figure 4: Inference time for object detection using different versions of SSD MobileNet.**



**Figure 5: Object detection inference times for SSD MobileNet v1 with overclocking.**

**Table 1: Power draw and CPU temperature in different CPU frequencies and modes.**

CPU Clock (MHz)	Mode	Power (W)	CPU Temp. (°C)
1,500	<i>Run</i>	6.0	49
1,500	<i>Sleep</i>	3.6	35
1,500	Idle Pi	3.4	33
1,750	<i>Run</i>	7.0	51
1,750	<i>Sleep</i>	3.9	35
1,750	Idle Pi	3.6	34
1,900	<i>Run</i>	8.5	57
1,900	<i>Sleep</i>	4.2	36
1,900	Idle Pi	3.7	35

# Other issues

- Privacy
- Using 'Edge TPU' devices
- Deployment

# Acknowledgments

This research forms part of Centre for Digital Built Britain's work within the Construction Innovation Hub. The funding was provided through the Government's modern industrial strategy by Innovate UK, part of UK Research and Innovation.



Thank you!