Person Detection with TinyML on the Raspberry Pi 5

// Edge AI · Real-Time Inference · Privacy-Friendly Detection

Author: Murat Aslan Type: Bachelor's Thesis Field: Computer Science

GRADE: 1.0

01 PROBLEM STATEMENT

This thesis investigates whether reliable, privacy-friendly person detection can run locally on a resource-constrained Raspberry Pi 5 in warehouse and warehouse-like environments. The main challenge is balancing detection quality with real-time performance while handling distant people, occlusion, motion blur, changing lighting, and person-like background objects that can trigger false alarms.

02 APPROACH & METHOD

A complete camera-based edge-AI pipeline was developed and fine-tuned on a domain-specific dataset of 2,226 images containing 7,183 annotations, then exported as INT8-quantized TensorFlow Lite models for CPU-based on-device inference. Models with 320, 480, and 640-pixel inputs were evaluated through reproducible offline benchmarks and live RTSP camera tests. The system also included resource monitoring and a Telegram-based human-in-the-loop feedback mechanism for collecting false-positive examples.

Python · YOLOv8n · Ultralytics · OpenCV · Roboflow · TensorFlow Lite · Raspberry Pi 5

03 KEY RESULTS

0.88	0.77	0.85	0.58
PRECISION	RECALL	MAP@0.5	MAP@0.5:0.95

- Evaluation used 63 videos (44 with people, 19 without), ~21 minutes total runtime.
- v7/640 fully detected all expected people in 32 of 44 person-containing videos; 9 partially, 3 not detected.
- A three-camera RTSP live test ran at approximately 4-6 FPS per stream.

MODEL	320PX	480PX	640PX
v7 (fine-tuned, INT8)	**26.84 FPS**	**11.33 FPS**	**6.29 FPS**
Pre-trained INT8	**28.88 FPS**	—	**8.14 FPS**
PyTorch (640px)	—	—	**3.54 FPS**
Float32 TFLite (640px)	—	—	**2.90 FPS**

04 CONCLUSION

The work demonstrates that practical, fully local person detection is feasible on the Raspberry Pi 5 without cloud inference. Domain-specific fine-tuning improved detection of difficult, distant people, while INT8 quantization enabled usable edge performance. Remaining limitations are mainly false positives from clothing and shelves, reducible through hard-negative mining and temporal filtering.