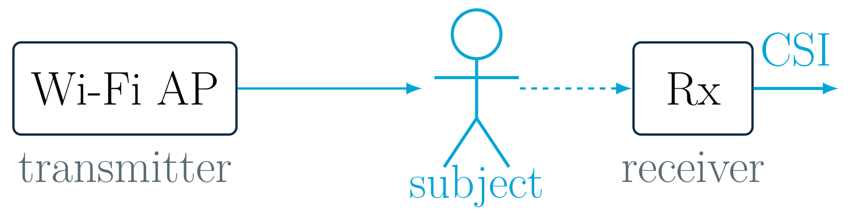
Position Robustness in WiFi-Based Human Activity Recognition

A Controlled Comparison of Five Cross-Position Strategies

Jokūbas Viršilas · Supervisors: Arash Asadi, Fabian Portner · Resp. Prof. Hayley Hung · TU Delft, EEMCS · J.Virsilas@student.tudelft.nl

1 Why this matters

To decode a WiFi signal, the receiver estimates how the indoor environment altered it on the way from transmitter to receiver — the **Channel State Information (CSI)**. The CSI is the sum of every propagation path, and a moving person perturbs some of those paths, so the CSI also carries cues about the person's **activity** — enough to recognise it without a camera or a worn sensor.



The catch: the CSI reflects the *full* radio geometry, not the activity alone. Standing elsewhere perturbs a *different* set of paths, so the *same* activity looks different across positions — at a position the model never trained on, it is recognised **about a fifth less accurately**. Surviving that shift is the problem we study.

2 Research questions

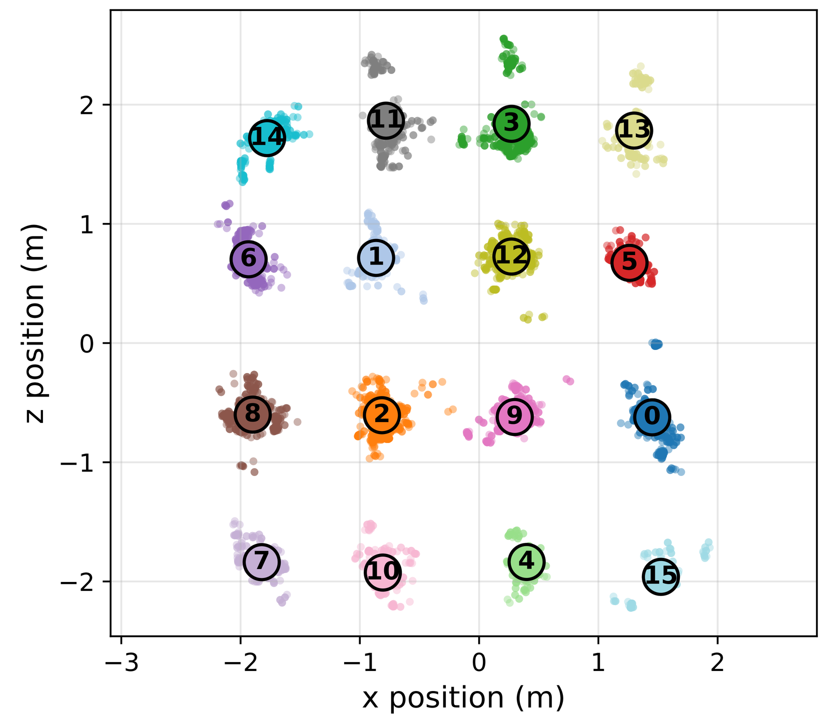
Which design choices make WiFi activity recognition robust to a change in the person's position — and does explicit domain adaptation help?

RQ1 How well do the selected robustness strategies recognise activities at *held-out subject positions* compared with no-invariance supervised learning?

RQ2 How *position-invariant* is each strategy's representation (Proxy-A-distance), and does invariance predict *cross-position* accuracy?

3 Setup: one protocol

- 1 subject, **8 stationary activities** (chance 12.5%), 3 receivers.
- **Leave-one-position-out** (16 folds), recording-disjoint splits.



The 16 standing positions (x,z); one held out per fold.

4 Five methods, three families

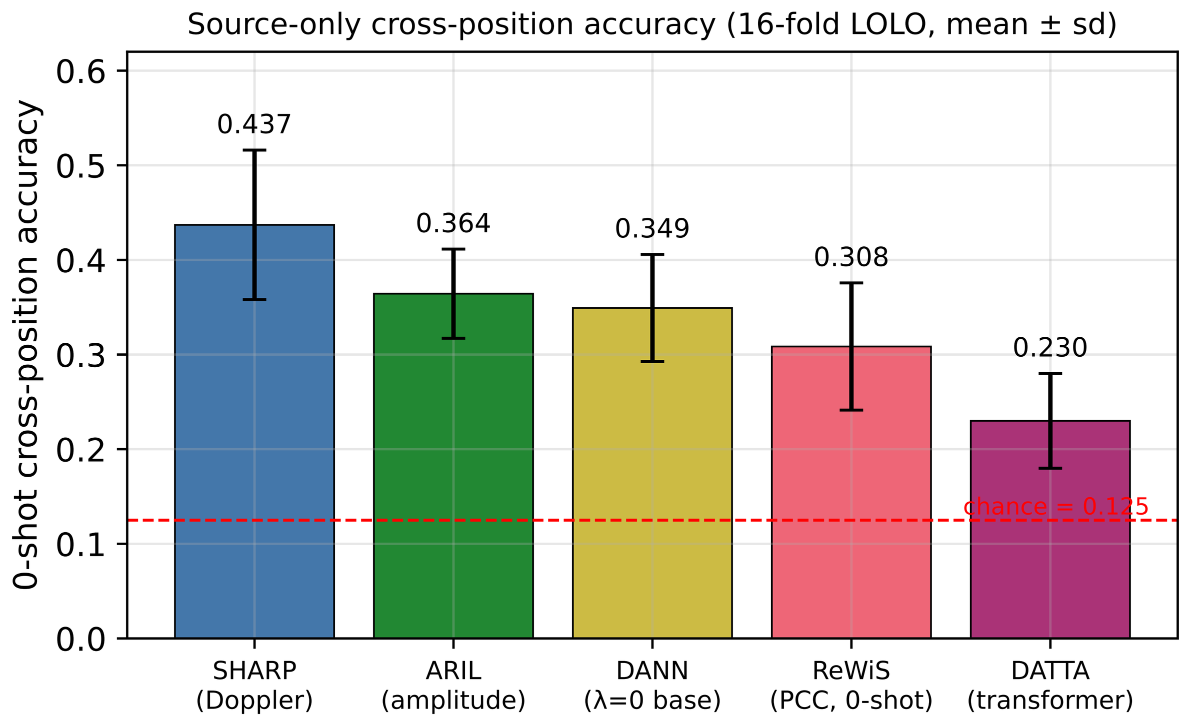
Method	Family / idea
ARIL	baseline — no invariance
SHARP	signal-aware: Doppler
DANN	adversarial domain confusion
ReWiS	few-shot prototypes (metric)
DATTA	test-time adaptation

One representative per robustness family, plus a no-invariance baseline (ARIL).

5 Which beats the baseline?

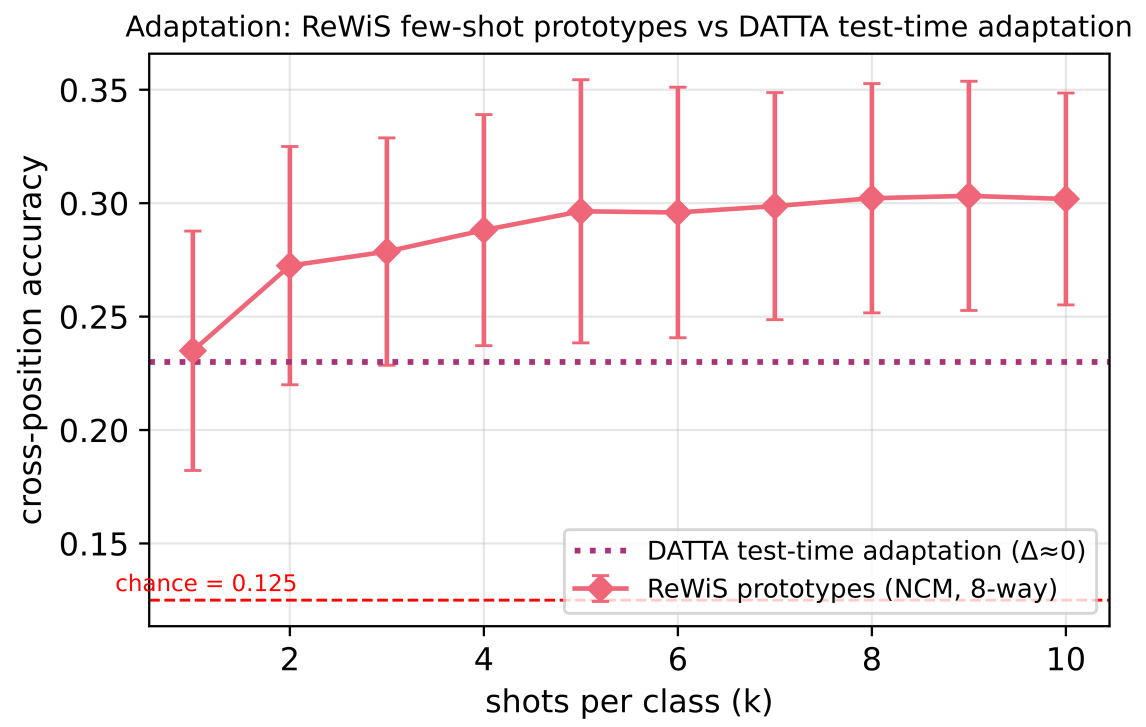
A: only the signal-aware representation.

16-fold **LOLO** (leave-one-location-out): each position is held out once for testing; chance 12.5%.



6 Does adaptation help?

A: no — adaptation gives little or no gain.

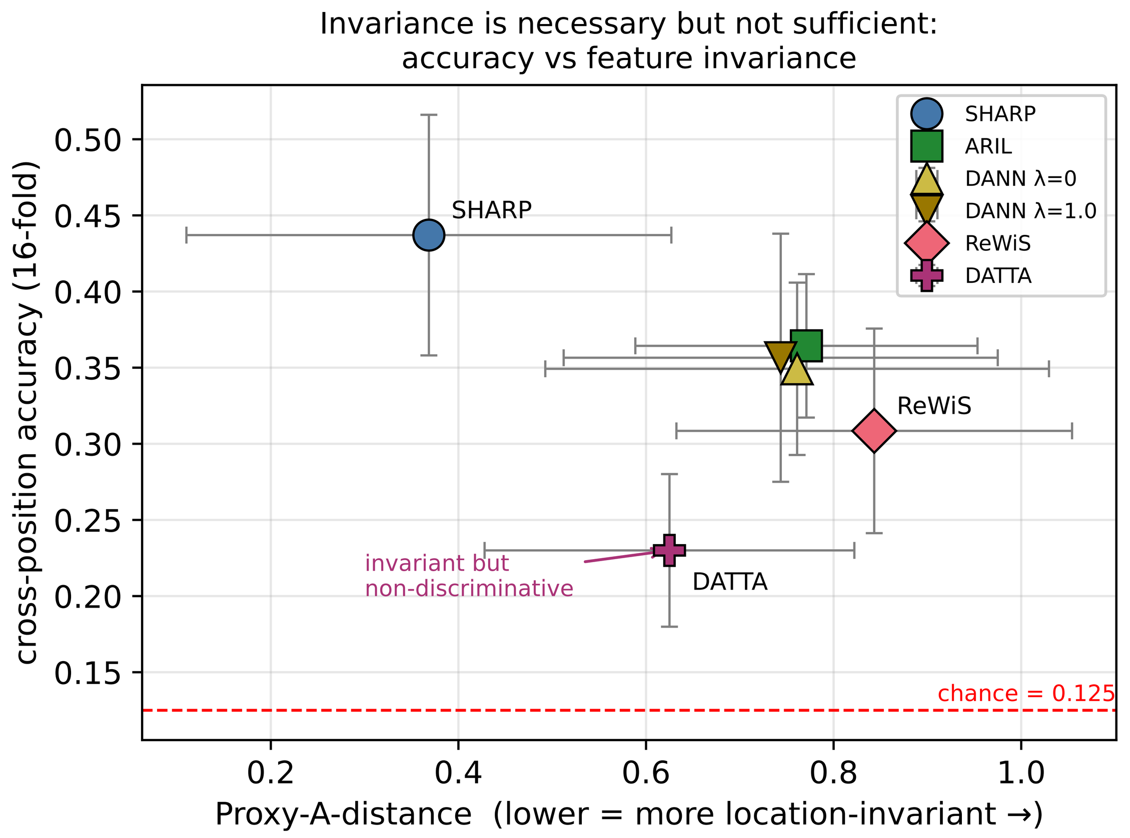


DANN is not on this plot — it trains for *invariance* (gradient-reversal, no target labels) rather than using target shots, and its λ-sweep likewise gives no significant gain (+0.7pp, n.s.).

7 Does invariance predict transfer?

A: necessary, but not sufficient.

Proxy-A-distance: lower = more position-invariant.



8 Takeaway

Across five methods under one protocol, what makes recognition survive a change of position is the signal *representation* — not an added adaptation step.

And invariance is not enough on its own: the features must stay *discriminative*, or a model can be position-invariant yet still unable to tell the activities apart, as DATTA shows.

Limitations. A single-subject study on one internal dataset and protocol, with only stationary activities (no locomotion). Some methods are adapted rather than faithful re-implementations — notably SHARP, which uses a Doppler representation. The dataset is not shared, but the full hardware, protocol and seeds are reported in the paper.