

ARGUS: Wi-Fi Telemetry Compression for Efficient Person Identification

Compact CSI statistics. Selective patch retention. Fewer classifier FLOPs under the same fixed protocols.

Nayan Sanjay Bhatia¹ Pranay Kocheta² Yuhan Li² Katia Obraczka¹

¹Department of Computer Science and Engineering, University of California, Santa Cruz ²Independent Researcher, USA
nbhatia3@ucsc.edu · pranayko021@gmail.com · yuhanli27t@gmail.com · katia@soe.ucsc.edu

BayLearn 2026

Santa Clara University
15 October 2026

CONTRIBUTION 1 / STATGRAM REPRESENTATION

A CSI window becomes a compact map of temporal statistics per frequency bin (20 rows $\times$ 64 bins per view). It replaces raw packet sequences as classifier input and is compared against dimension-matched pooling and PCA with the same classifiers.

CONTRIBUTION 2 / SELECTIVE PATCH RETENTION

A fixed subset of statgram patches is chosen offline and the classifier is retrained on it. Transformer arithmetic falls by 44.8% (EHealth) and 31.6% (WiMANS) at $-0.12/-0.73$ and $+1.20$ ($+0.24$ recipe-fixed) accuracy points.

CONTRIBUTION 3 / CONTROLLED EVALUATION

15 classifier families $\times$ 7 view combinations on fixed closed-set splits; occlusion vs. attention vs. fixed vs. random rankings; token deletion vs. retraining. Attention is neither required nor a reliably better ranking signal.

01 What is a statgram? What is a patch?

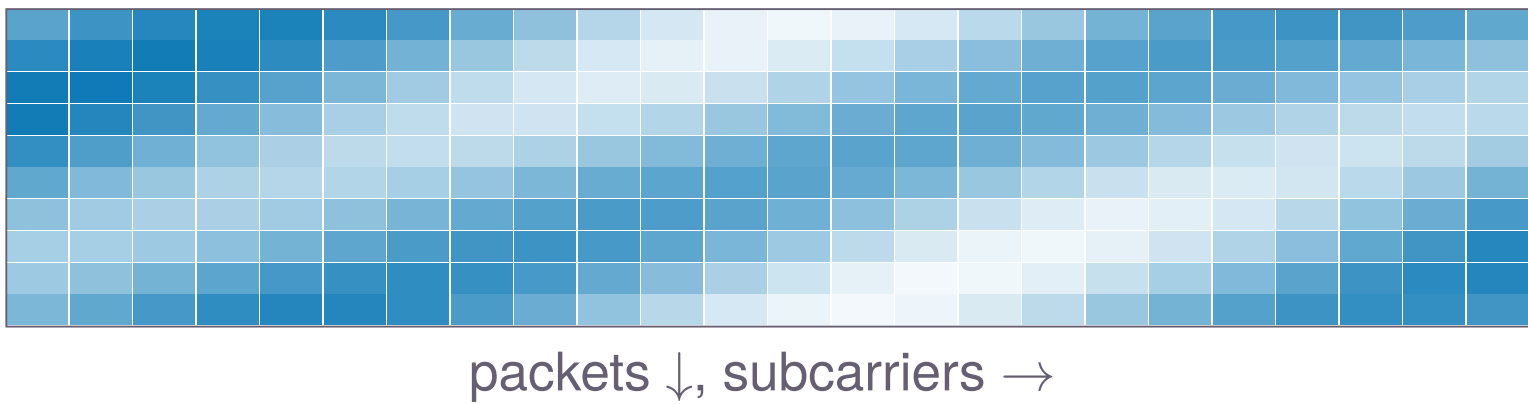
Window to views. A window holds complex CSI $H \in \mathbb{C}^{T \times K}$ (T packets, K subcarriers). Each view is $X = \frac{X - \text{median}(X)}{\max(\text{IQR}(X), \epsilon)}$, normalised per window with $\rho(X) = \frac{X - \text{median}(X)}{\max(\text{IQR}(X), \epsilon)}$. amplitude $A = g_W(\rho(\log(1 + |H|)))$; phase $P = g_W(\rho(\Phi))$ after unwrapping across subcarriers and removing a per-packet linear trend; phase-delta D is the temporal difference of P . g_W averages adjacent subcarriers into W bins (EHealth: 256 $\rightarrow$ 64).

Statistic rows. For every view and bin, 18 statistics summarise the T packets: mean, SD, min, max, last-first; low/mid/high FFT band energies; p10, p25, p50, p75, p90; IQR, RMS, zero-crossing rate; autocorrelation at two lags; plus two reserved rows. A 50×256 complex window becomes a $3 \times 20 \times 64$ real map (APD).

Patches and tokens. The map is cut into 10×16 patches: 8 per window in EHealth, 15 (4×10) in WiMANS. Each patch is flattened and projected to a 192-d token; a four-block causal Transformer with a CLS token classifies the sequence. *Retention* keeps a fixed subset of patch IDs chosen offline; the dropped tokens are never computed.

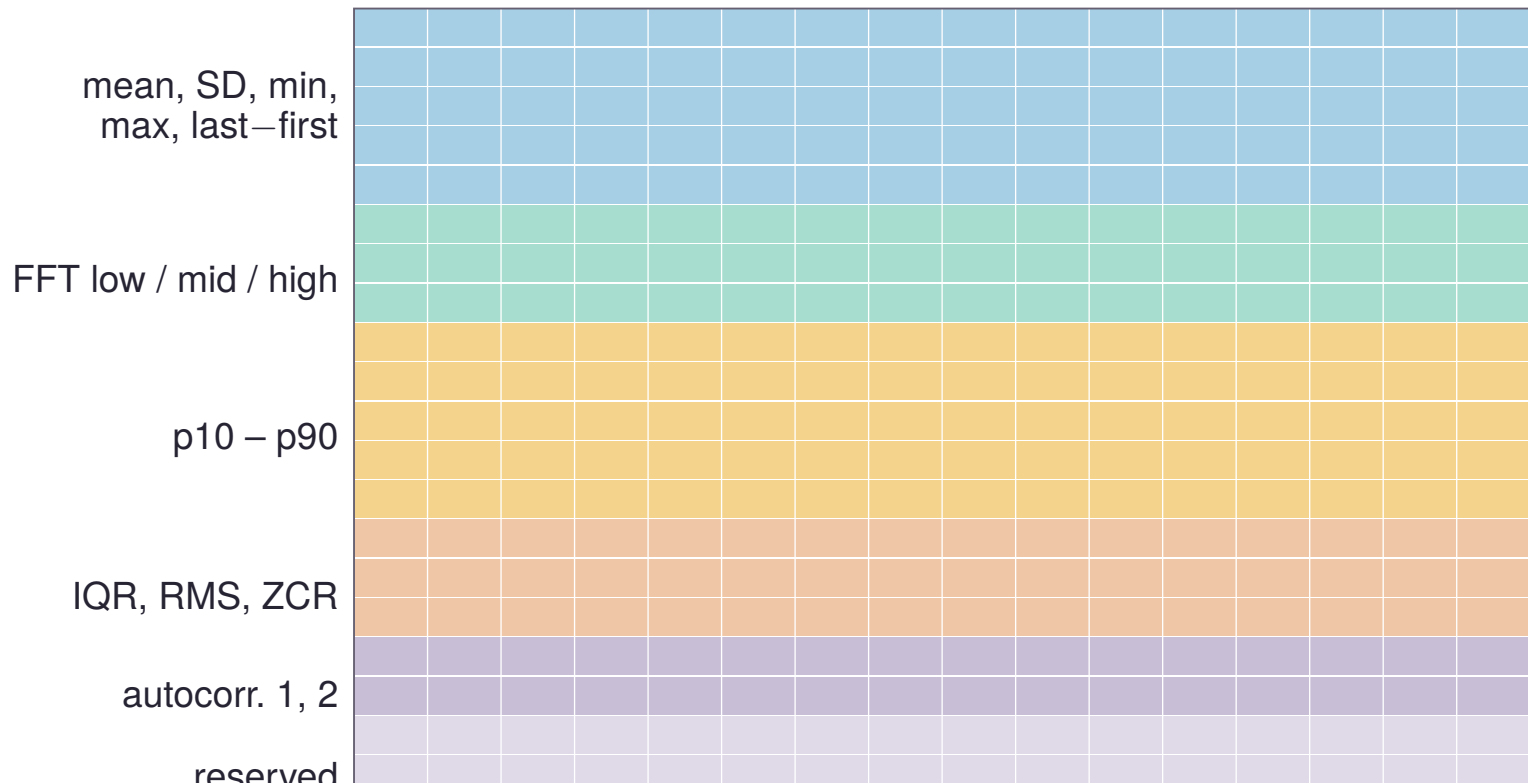
Window $\rightarrow$ statgram $\rightarrow$ patches (one view)

(a) Window: $T = 50$ packets $\times K = 256$ subcarriers



18 statistics per bin

(b) Statgram: 20 statistic rows $\times$ 64 bins



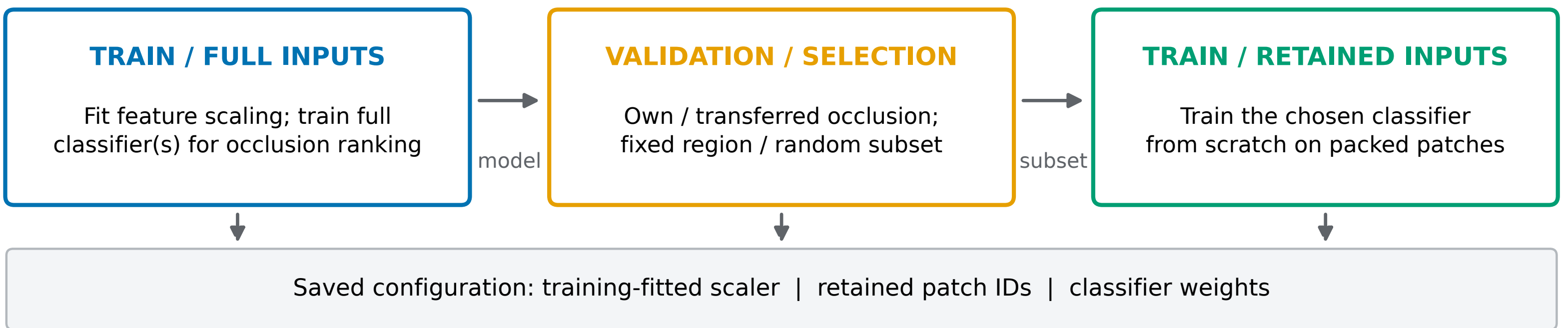
64 frequency bins $\rightarrow$

(c) 2×4 patches $\rightarrow$ 8 tokens; fixed-lower keeps 4-7

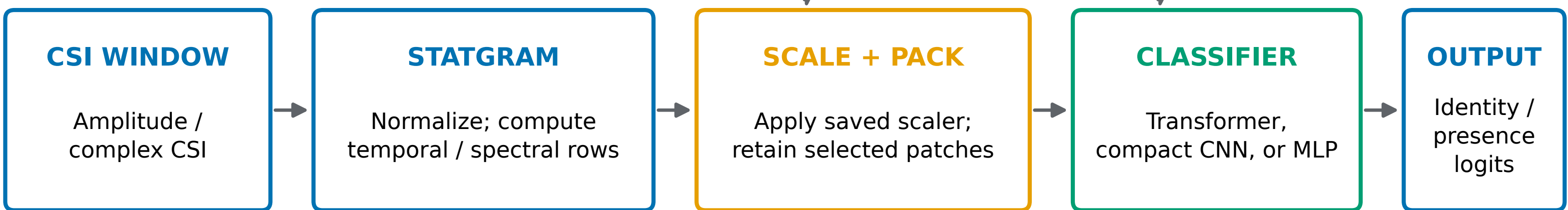


02 ARGUS: train, select patches offline, retrain, infer

(a) Offline preparation



(b) Inference



Solid arrows: data flow. Dashed: configuration fixed before test inference.

Offline: fit the scaler, pick a fixed patch subset on training/validation data, retrain on the packed patches. Inference never runs the full ranking model.

03 Two closed-set tasks, fixed splits

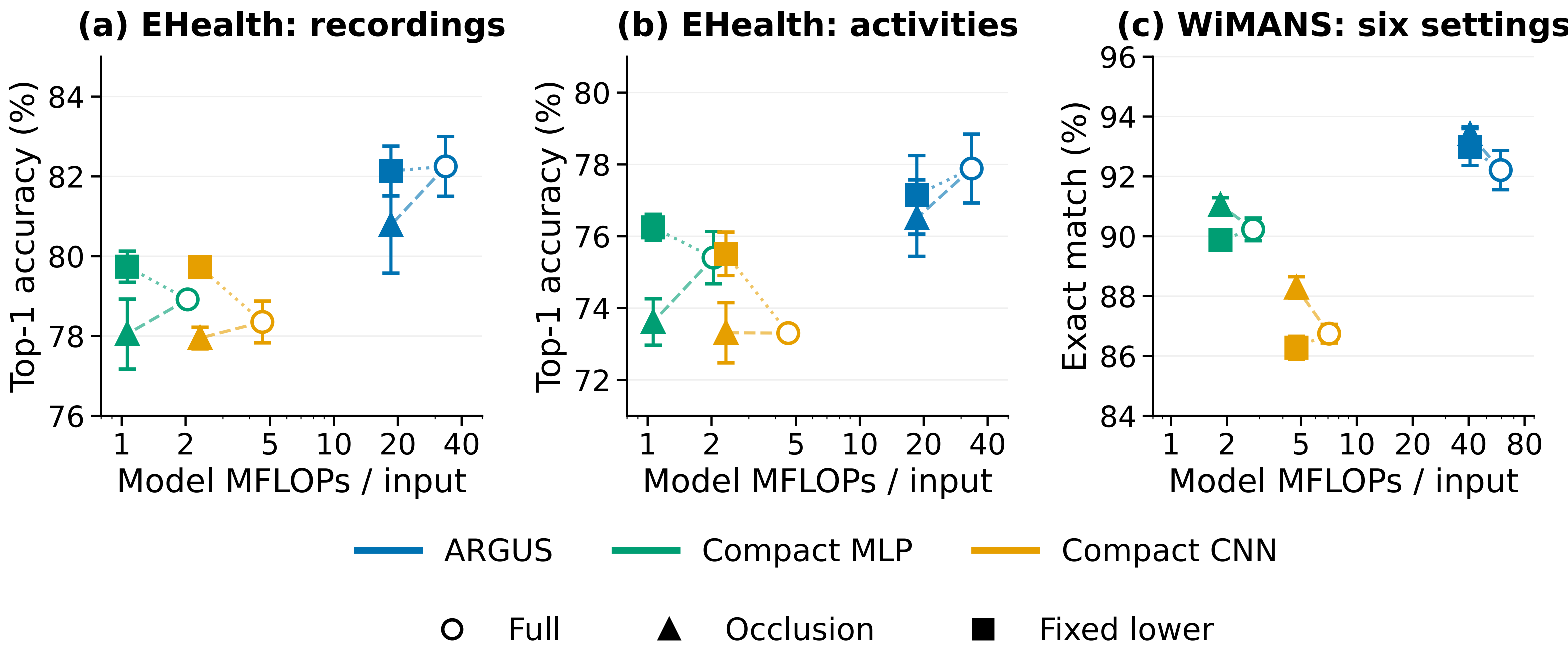
	EHealth	WiMANS
Task	Identify 1 of 154 enrolled people	Six-user presence vector; 0-5 users per recording
Data	154 people $\times$ 17 recordings; 500 rows $\times$ 256 subcarriers	11,286 recordings; 3 rooms $\times$ 2 bands; 9 links $\times$ 30 subcarriers
Input	50-row window, stride 25 (19 per recording); APD statgram	3,000-row recording; amplitude statgram over 9 links
Held out	Recording holdout (11/3/3 per person) and activity holdout; 8,778 test windows each	20% of recordings per setting; six settings trained and scored separately
Metric	Window Top-1 (recording Top-1: mean of 19 win-dow logits)	Exact match of all six outputs at probability > 0.5
Budget / SD	4 of 8 patches; mean $\pm$ SD over 5 seeds	10 of 15 patches; settings averaged within seed, then 5 seeds

08 Raw-input baselines on the same EHealth windows

Input / classifier	Window Top-1	Recording Top-1	MFLOPs/window
Raw amplitude / THAT	81.36 $\pm$ 0.84	85.50 $\pm$ 1.86	149.08
Raw amplitude / CNN-BiLSTM	85.82 $\pm$ 0.75	89.39 $\pm$ 0.88	295.30
Statgram A / CNN1D	80.68 $\pm$ 0.18	86.54 $\pm$ 0.64	16.71
Scaled APD / full ARGUS	82.25 $\pm$ 0.75	87.79 $\pm$ 0.50	33.63
Scaled APD / 4-patch ARGUS	82.13 $\pm$ 0.62	86.88 $\pm$ 1.12	18.57

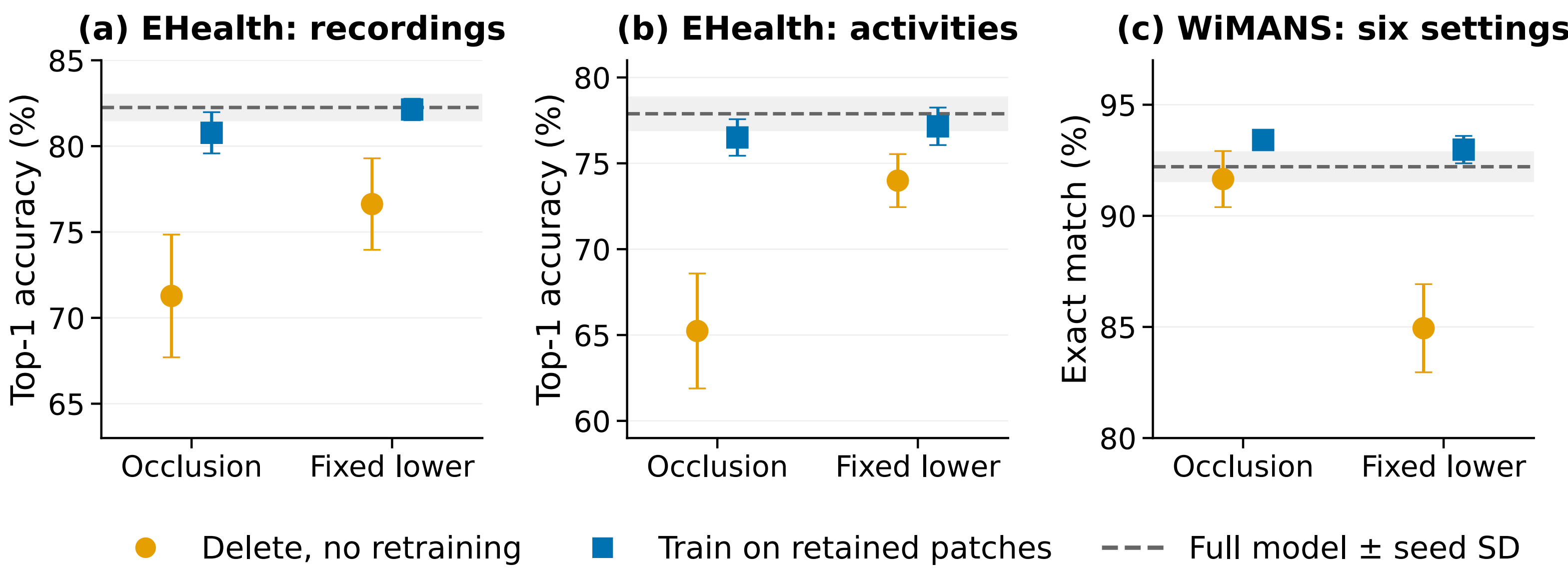
Pipelines differ in input, architecture, and recipe: an accuracy-computation trade-off, not an isolated representation effect.

04 Accuracy against classifier FLOPs



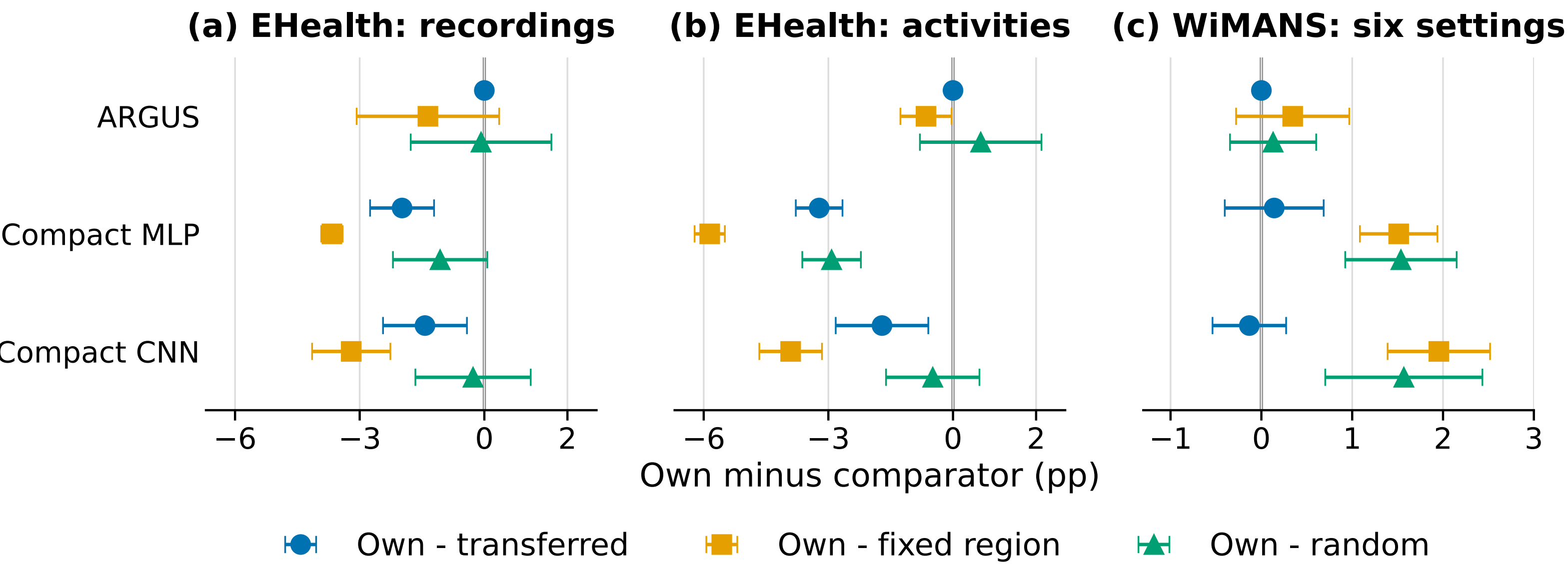
Open: full input. Filled: 4/8 or 10/15 patches, retrained from scratch, recipes tuned separately. **EHealth** prefers the fixed lower region ($-0.12, -0.73$ points). **WiMANS** prefers occlusion: $+1.20$, but $+0.24$ once the full model's recipe is held fixed.

05 Deleting tokens vs. training on the retained patches



Same patches, two procedures. Deleting tokens from the unchanged full model loses up to 12.7 points; retraining lands within 1.5 points of, or above, the full model.

06 Which patches to keep? No selector is universally best



ARGUS, full-model recipe held fixed

Selector	EH rec. (%)	WiMANS (%)
Full input	82.25 $\pm$.75	92.21 $\pm$.65
Occlusion	80.78 $\pm$ 1.20	92.45 $\pm$.11
Fixed lower	82.13 $\pm$.62	92.10 $\pm$.61
Random	80.86 $\pm$.77	92.32 $\pm$.38

Each classifier's own occlusion ranking minus a transferred Transformer ranking, the fixed region, or random subsets; bars are seed SD. **EHealth**: the fixed region wins for every classifier. **WiMANS**: own beats random by about 1.5 points for the compact models; own vs. transferred changes sign. Occlusion ranking is not attention ranking.

07 Controls: representation and scaling, not attention

Dimension-matched inputs, same 256-unit MLP

Input (320-d / 720-d)	EH rec.	EH act.	WiMANS
Statgram	77.6	72.6	87.4
Mean pooling	72.9	66.0	81.1
Pooling + PCA	76.0	70.7	74.7

Training-fitted standardisation, EHealth APD

Classifier	Original	Scaled	MFLOPs
MLP	0.65	76.2	2.04
CNN1D	79.0	81.8	18.34
ARGUS	80.4	81.7	33.63

Statgrams beat dimension-matched pooling and PCA with the MLP on both datasets (ridge disagrees on EHealth). Training-fitted standardisation rescues the EHealth MLP, so the unscaled failure is not evidence that attention is needed.

Scope. Within-collection, within-setting evidence: not independent-session, device, or day splits; not cross-room; no open-set rejection. FLOPs count classifier matrix arithmetic only; seed SDs are initialisation variation, not confidence intervals; fixed-region controls are exploratory.