Leveraging Agonistic-Antagonistic Coactivation in Single-Grid HDsEMG for Hand Gesture Recognition

Firas Darwish[1*], Dhiyaa Al Jorf[2*], Costanza Armanini[3], Eion Tyacke[4], Farah E. Shamout[3]

[1] University of Oxford, UK [2] ETH Zurich, Switzerland [3] New York University Abu Dhabi, UAE
[4] New York University, USA [*] Equal contribution

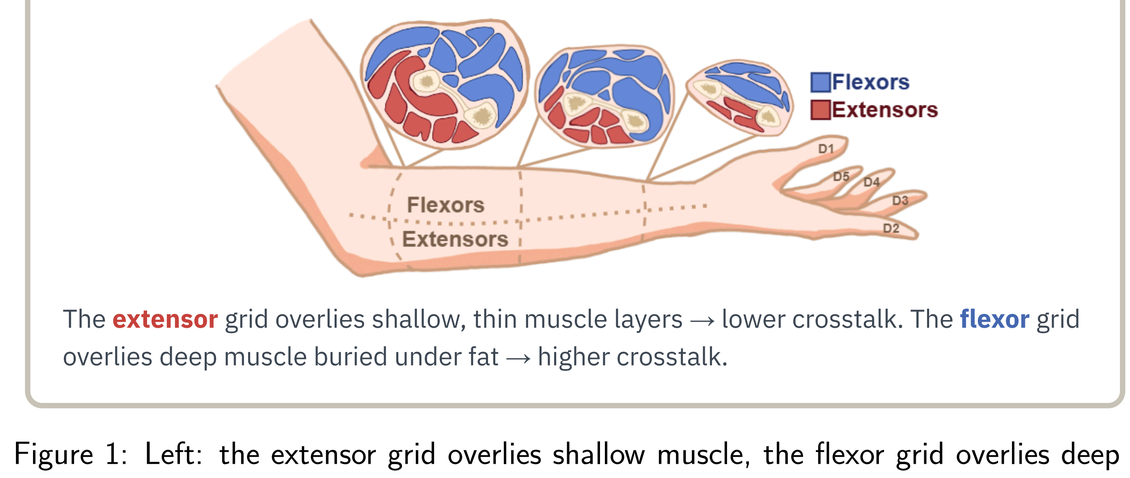

Figure 1: Left: the extensor grid overlies shallow muscle, the flexor grid overlies deep muscle under fat. Right: during a pinch, the flexor (agonist) drives the motion while the extensor (antagonist) coactivates in parallel – each grid recording a trace of the other.

Introduction

- Surface EMG powers prosthetic control, but top gesture recognizers need **dense, dual-grid electrode arrays** that are expensive, bulky, and slow to run.
- Muscles work in **antagonist pairs**: when one contracts, its partner coactivates to stabilize the joint, so every grid picks up a faint echo of the *other* muscle's activity (Fig. 1).
- If that echo carries enough signal, a **single grid** might be all a classifier needs – and fusing both grids more tightly might reveal nothing extra. We test both.

Methods

- **Data** [1]: **20** participants, **16** single-degree-of-freedom gestures, two 8×8 HDsEMG grids over the flexor and extensor compartments, 5 repetitions each.
- **Two evaluation settings**: a *generalized* model pools all 20 subjects for training and testing, while *subject-specific* models retrain the same architecture separately for each person (Fig. 2).
- **Architectures**: Single-Stream (one grid), Joint Fusion (both grids, late merge), and Slow Joint Fusion (both grids, merged early to probe for spatial structure).
- **Explainability**: GradCAM [2] on the Joint Fusion model shows which grid the network leans on, gesture by gesture.

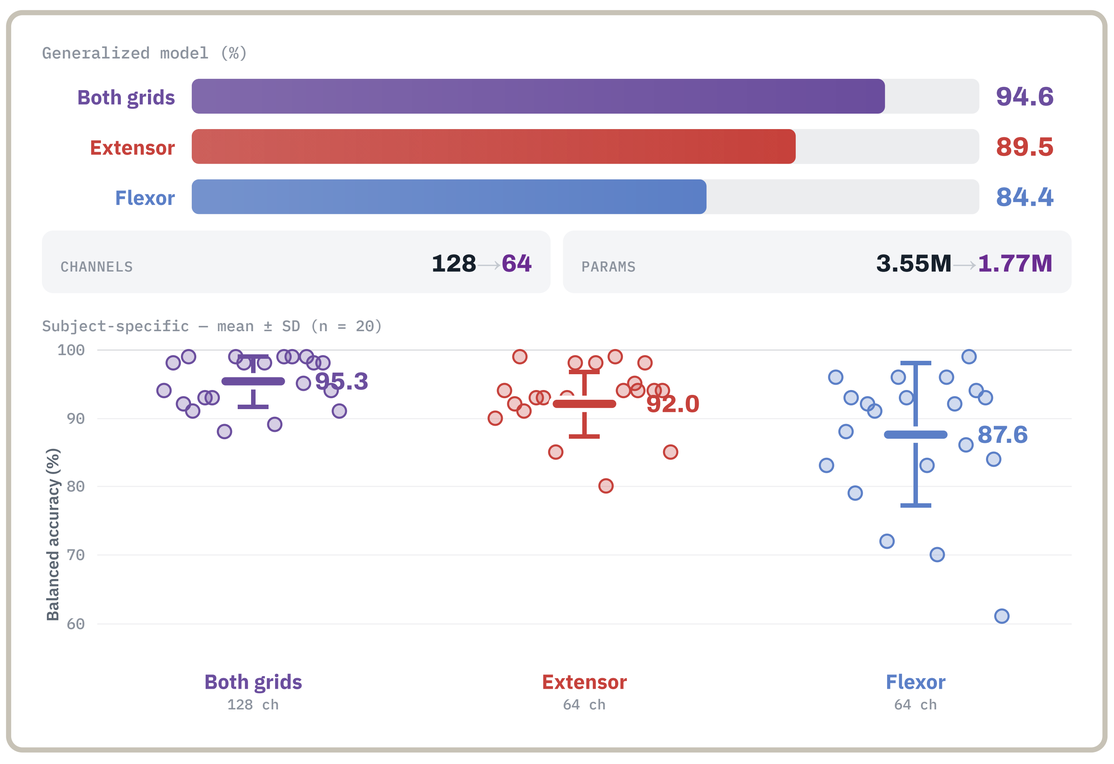

Figure 2: Dropping the flexor grid barely moves the needle; dropping the extensor grid does.

Results & Discussion

- Keeping only the **extensor** grid costs just **5 points** of balanced accuracy (**94.6%**→**89.5%**) while cutting channels and trainable parameters **in half** (Fig. 2).
- Keeping only the **flexor** grid instead is a far worse trade: accuracy drops to **84.4%** on average, and swings wildly from subject to subject (**61%–98.5%**).
- **Fusing both grids** more tightly buys nothing – Slow Joint Fusion lands at **94.0%**, statistically indistinguishable from simply concatenating the grids late.
- **GradCAM agrees**: across the joint model, the **extensor** grid drives the decision for **12 of 16** gestures, the **flexor** grid for only **10** (Fig. 3); the extensor's edge is statistically significant ($t = 3.18$, $p < 0.01$).
- Anatomy explains the gap: the **extensor** sits in thin muscle close to the skin, while the **flexor** is buried under extra muscle and fat, adding crosstalk before the signal reaches the electrodes.

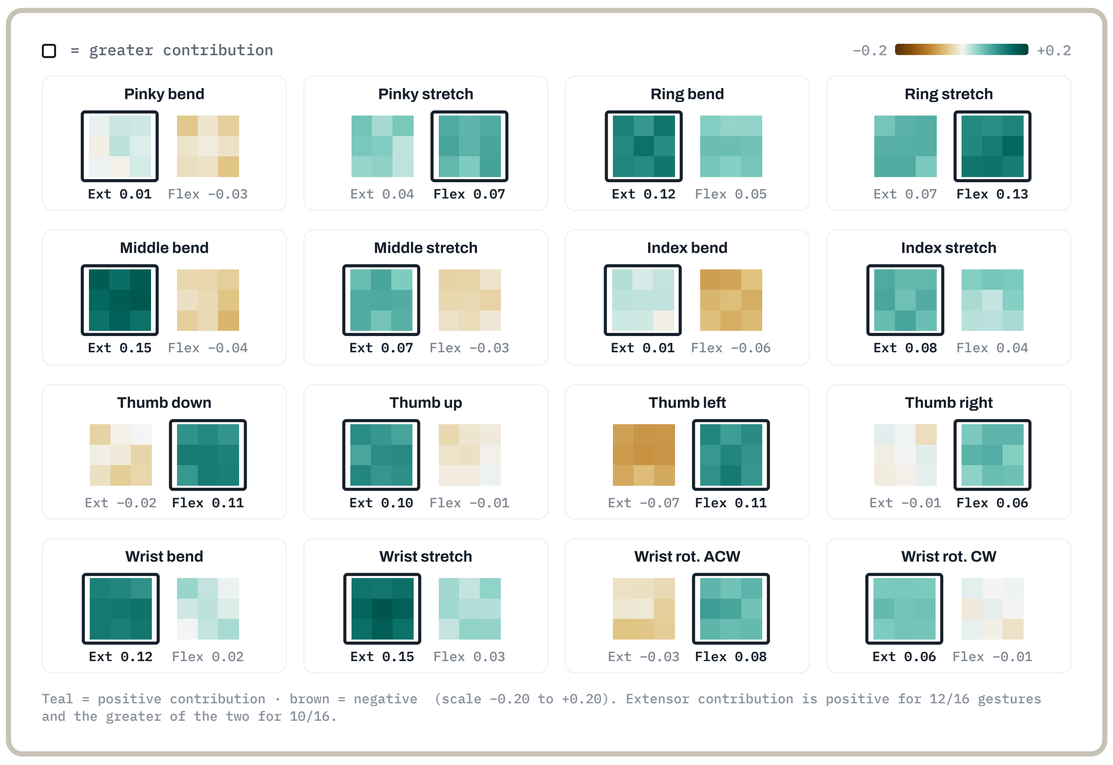

Figure 3: **Teal** marks where a grid pushes the model toward the correct gesture; brown marks where it pushes away.

References

[1] N. Malesevic *et al.*, "A database of high-density surface electromyogram signals comprising 65 isometric hand gestures," *Scientific Data*, vol. 8, no. 1, p. 63, 2021.
[2] R. R. Selvaraju *et al.*, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE ICCV*, 2017, pp. 618–626.
[3] J. P. M. Mogk and P. J. Keir, "Crosstalk in surface electromyography of the proximal forearm during gripping tasks," *J. Electromyography and Kinesiology*, vol. 13, no. 1, pp. 63–71, 2003.

Supported by the NYUAD Center for Artificial Intelligence and Robotics (Tamkeen, award CG010) and the Meem Foundation; computation on the NYUAD HPC cluster.
Correspondence: ca3072@nyu.edu

Scan for full paper

Clinical AI Website