Self-supervised Learning for Human Activity Recognition Using 700,000 Person-days of Wearable Data

Hang Yuan*[1,2], Shing Chan*[1,2], Andrew P. Creagh[4], Catherine Tong[3], David A. Clifton[4], and Aiden Doherty[1,2]

[1]Nuffield Department of Population Health, University of Oxford
[2]Big Data Institute, University of Oxford
[3]Department of Engineering Science, University of Oxford
[4]Department of Computer Science, University of Oxford

Current human activity recognition (HAR) models typically rely on manual feature engineering partly due to the limited size of existing labelled datasets. This small data issue caps the effectiveness of data-hungry deep learning methods. In general, obtaining *labelled* data is labour intensive, especially for HAR data, because one would need to annotate the corresponding video stream for the ground truth. On the other hand, collecting large-scale *unlabelled* HAR data is highly feasible, as evidenced by projects such as the UK Biobank [1] and NHANES [2]. This prompts the use of self-supervised learning (SSL) methods to leverage unlabelled data in a similar manner as language models and vision models. Recent studies explored the utility of SSL for HAR [3, 4], but these still relied on small-scale laboratory-style datasets. Hence the full potential of SSL-HAR remains unknown.

We thus investigate how learning three simple self-supervised learning tasks independently and jointly could facilitate HAR across various environments using the UK Biobank dataset, which contains terabytes of wearable sensor data collected in the real world.

Our main contributions are as follows: (1) We show for the first time that multi-task SSL on a large free-living accelerometry can train a HAR model that generalises well across seven external datasets that differ in activity classes, devices, populations and recording environments. The relative F1 improvement against the baseline model on the benchmark datasets ranged between 2.5% (Capture24) and 100.0% (ADL) with a median of 18.4%. (2) In contrast to previous works, we provide a more realistic evaluation of the utility of SSL-HAR by factoring in common issues seen in the practical use cases of pre-trained models such as domain shift and task shift. (3) Our publicly available models will enable the research community to build high-performing activity recognition models even in a resource-restricted environment.

References

[1] Aiden Doherty et al. "Large scale population assessment of physical activity using wrist worn accelerometers: the UK biobank study". In: *PloS one* 12.2 (2017).

[2] Britni R Belcher et al. "US Population-referenced Percentiles for Wrist-Worn Accelerometer-derived Activity." In: *Medicine and Science in Sports and Exercise* (2021).

[3] Aaqib Saeed, Tanir Ozcelebi, and Johan Lukkien. "Multi-task self-supervised learning for human activity detection". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3.2 (2019), pp. 1–30.

[4] Chi Ian Tang et al. "Exploring Contrastive Learning in Human Activity Recognition for Healthcare". In: *arXiv preprint arXiv:2011.11542* (2020).