Group Supervised Learning: Extending Self-Supervised Learning to Multi-Device Settings

Yash Jain*[1]. Chi Ian Tang*[2]. Chulhong Min[3]. Fahim Kawsar[3]. Akhil Mathur[3].

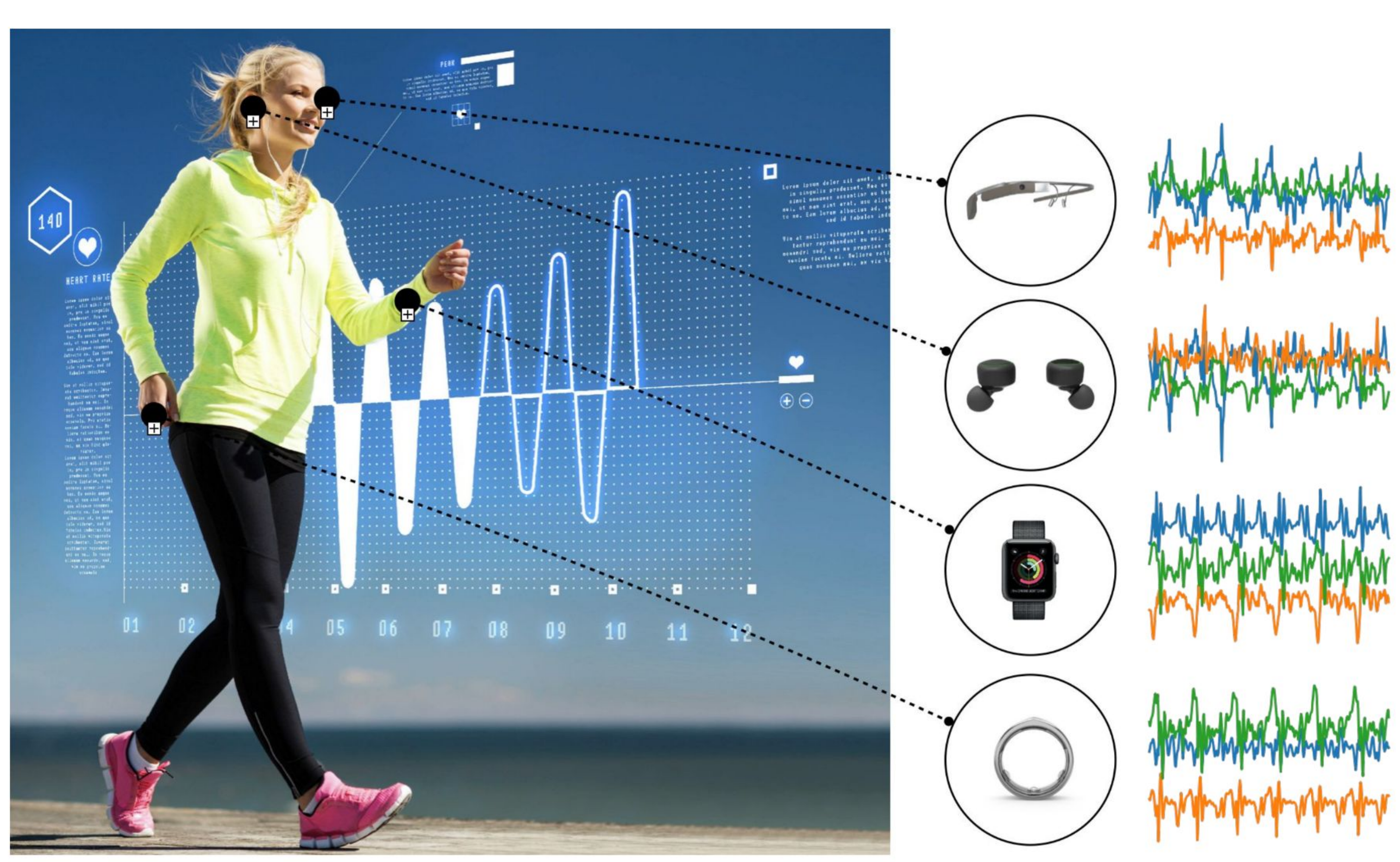
*[1]Indian Institute of Technology, Bombay. [2]University of Cambridge. [3]Nokia Bell Labs, Cambridge. *Equal Contribution*

HIGHLIGHTS

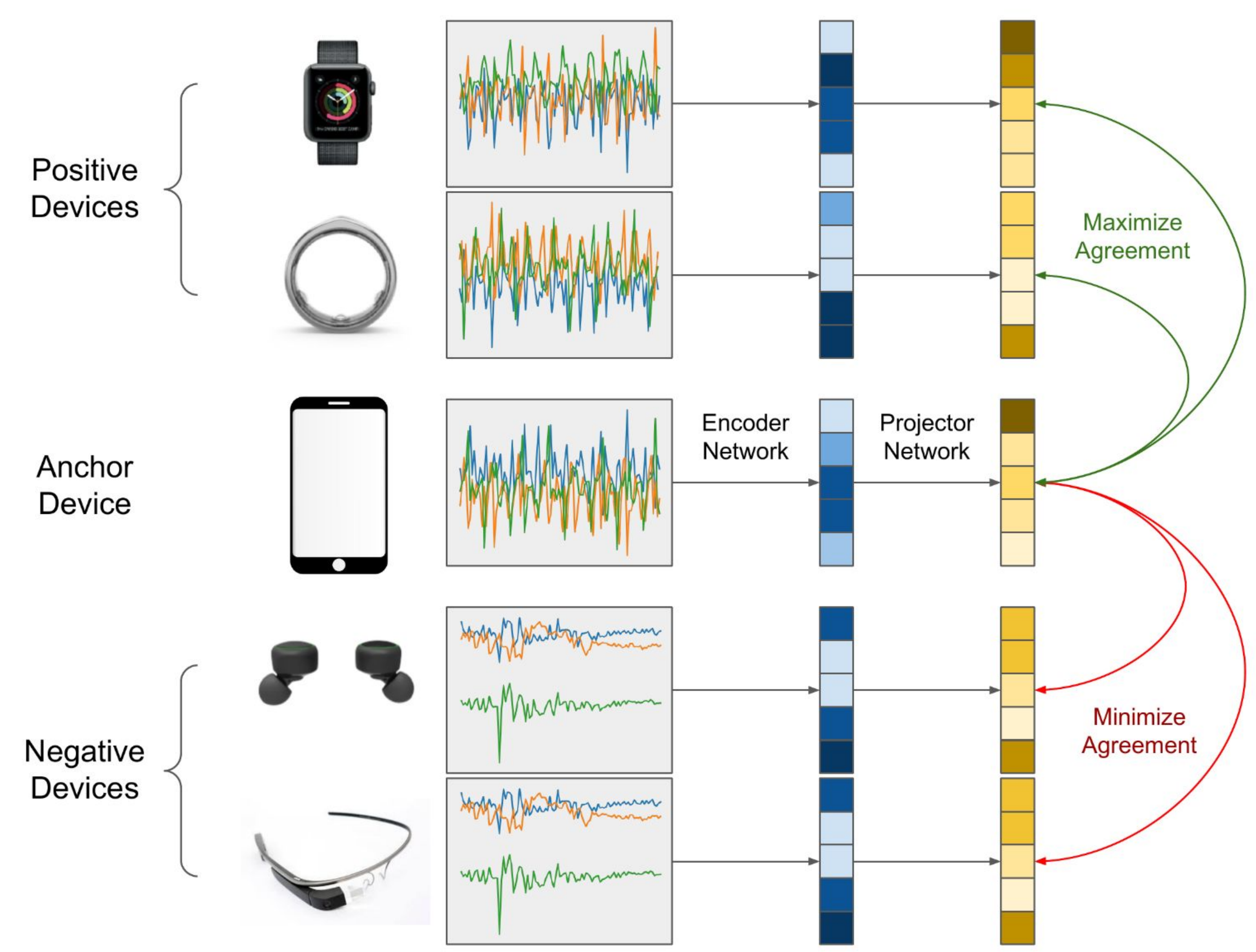
1. Proposed a new problem setting, TSMDS, which exists in many domains in the real world but has not been thoroughly explored yet.
2. A novel framework, GSL, addressing the TSMDS problem, utilizing the principles of contrastive learning in a group setting.
3. Early results demonstrate that GSL outperforms supervised and semi-supervised training baselines proposed in the HAR literature by as high as 0.15 in F-1 score.

PROBLEM DEFINITION: TSMDS

- **Time-Synchronous Multi-Device System**
- **Given**: Time-aligned unlabeled data samples from K devices including an anchor device
- **Goal**: Leverage the time-aligned, unlabeled multi-device datasets to learn a feature extractor that can generate effective feature representations for anchor device



METHODOLOGY: Group-supervised learning (GSL)



- **Group-supervised learning.** A contrastive self-supervised learning framework which extends contrastive learning to a setting with groups of time-aligned devices
- **Key intuition.** Take the time-aligned samples from devices similar to anchor device , and pull them closer to it in the embedding space while pushing samples from dissimilar devices away
- **Group Supervised Contrastive Loss.** We train the model using a novel loss function called Group Supervised Contrastive Loss, which is an extension of the standard contrastive loss function but compatible with multiple positive and negative samples

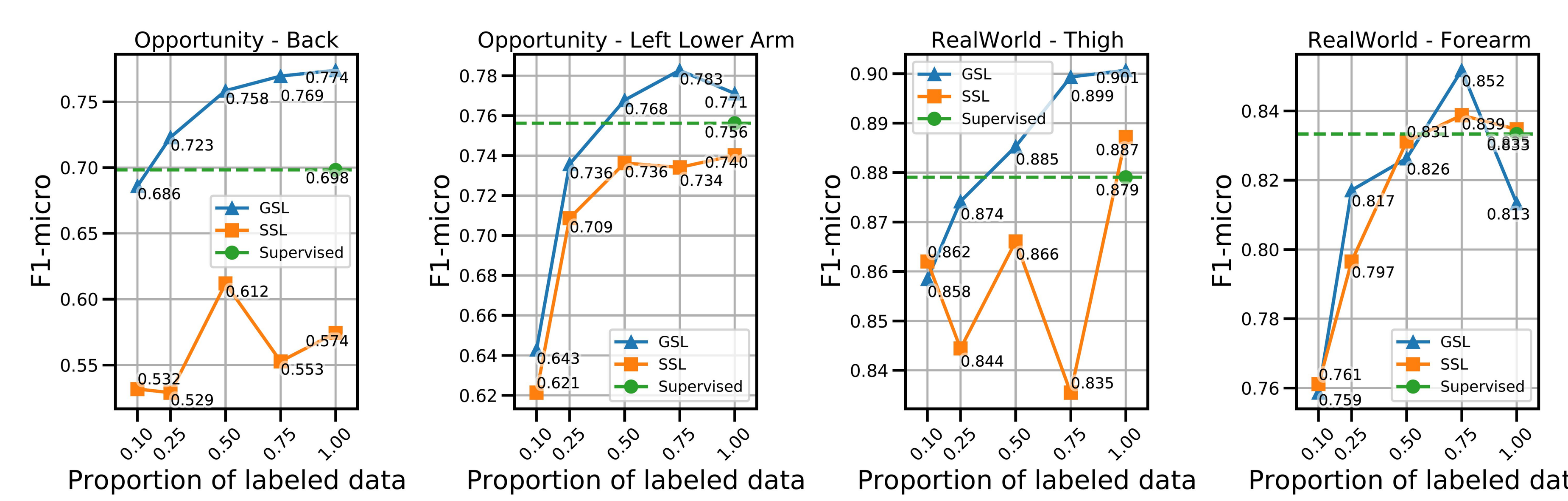
$$\mathcal{L}_{GSL} = \frac{\sum_{i=0}^{|D^+|} \exp\left(\mathrm{sim}\left(z^0, z_i^+\right)/\tau\right)}{\left(\begin{array}{l}\sum_{i=0}^{|D^+|} \exp\left(\mathrm{sim}\left(z^0, z_i^+\right)/\tau\right)\\ \quad + \sum_{j=0}^{|D^-|} \exp\left(\mathrm{sim}\left(z^0, z_j^-\right)/\tau\right)\end{array}\right)}$$

RESULTS

Method Proportion of data	GSL $\leq$ 75%	SSL $\leq$ 75%	Supervised 100%
OPP - Back	**0.769**	0.612	0.698
OPP - Left Lower Arm	**0.783**	0.736	0.756
OPP - Left Shoe	**0.732**	0.706	0.700
OPP - Right Shoe	0.722	**0.735**	0.726
OPP - Right Upper Arm	**0.831**	0.599	0.681
RW - Chest	**0.906**	0.788	0.899
RW - Forearm	**0.852**	0.839	0.833
RW - Head	**0.834**	**0.834**	0.788
RW - Shin	**0.891**	0.886	0.885
RW - Thigh	**0.899**	0.866	0.879
RW - Upper Arm	**0.876**	0.862	0.857
RW - Waist	**0.916**	0.808	0.887

Table 1. Comparison of classification performance (F1-micro scores) between *GSL* and other training pipelines on two HAR datasets (OPP - Opportunity, RW - RealWorld).

- **Datasets:** RealWorld, Opportunity
- **Baselines:** Fully-supervised training, SimCLR Contrastive Training for HAR (SSL) [1]
- **Takeaway**: GSL outperformed the other baselines in the vast majority of cases, with a performance gain compared to the second-best pipeline as high as 0.15 in F1-score.



[1] Tang, C. I., Perez-Pozuelo, I., Spathis, D., & Mascolo, C. (2020). Exploring Contrastive Learning in Human Activity Recognition for Healthcare. arXiv preprint arXiv:2011.11542.