ColloSSL: Collaborative Self-Supervised Learning for Human Activity Recognition

ECE 69500 IoT: Paper Review | April 27th, 2023 | **Reviewer:** Yuan-Yao Lou

1 Introduction & Motivation

The authors propose a new technique called Collaborative Self-Supervised Learning (ColloSSL) to address the need for large-scale labeled sensor datasets in Human Activity Recognition (HAR) models. In terms of the motivation, the author address the limitations of current contrastive learning techniques in the Time-Synchronous Multi-Device System (TSMDS) problem setting, where multiple intertial measurement unit (IMU)-enabled devices are collecting time-synchronized sensor data while the user is performing a physical activity.

2 Problem

01. Limitations of current contrastive learning techniques in capturing appropriate positive and negative samples.
02. Need for large-scale labeled sensor datasets.
03. Difficulty in extending conventional self-supervised learning algorithms to a multi-device setting.

3 Solution

ColloSSL is an extension of unsupervised and semi-supervised learning techniques that can learn good features from the data without requiring any labels. The paper presents three technical innovations to extend conventional self-supervised learning algorithms to a multi-device setting: a Device Selection approach, a Contrastive Sampling algorithm, and a loss function called Multi-view Contrastive Loss.

4 Experimental Results

The experimental results on three multi-device datasets show that ColloSSL outperforms both fully-supervised and semi-supervised learning techniques in majority of the experiment settings, resulting in an absolute increase of up to 7.9% in F1 score compared to the best performing baselines.

5 Pros and Cons

ColloSSL outperforms both fully-supervised and semi-supervised learning techniques in a majority of the experiment settings. Additionally, ColloSSL outperforms the fully-supervised methods in a low-data regime, by just using one-tenth of the available labeled data in the best case.

However, the paper primarily focuses on applying ColloSSL to Human Activity Recognition (HAR) with motion data, although the paper mentions that the Time-Synchronous Multi-Device System (TSMDS) setting is common to other sensor modalities such as audio and vision. The technical solutions proposed in ColloSSL, such as device selection, contrastive sampling, and group contrastive loss, need to be redesigned to reflect the characteristics of sensory signals, user behavior, and environments in other settings. In future work, the authors plan to explore technical solutions to extend ColloSSL to audio- and vision- based TSMDS settings.

6 Questions

01. The paper does not address ethical or privacy concerns related to large-scale labeled sensor datasets. Given this concern, how to mitigate or alleviate this issue?
02. How to extend ColloSSL to audio- and vision- based TSMDS settings? What performance might be impacted?