Evaluating Contrastive Learning on Wearable Timeseries for Downstream Classification Outcomes

Kevalee Shah, Dimitris Spathis, Chi Ian Tang and Cecilia Mascolo

UNIVERSITY OF CAMBRIDGE

NEURAL INFORMATION PROCESSING SYSTEMS

Applying SimCLR and BYOL methodologies to high-dimensional health signals shows promising results for task-agnostic downstream classification

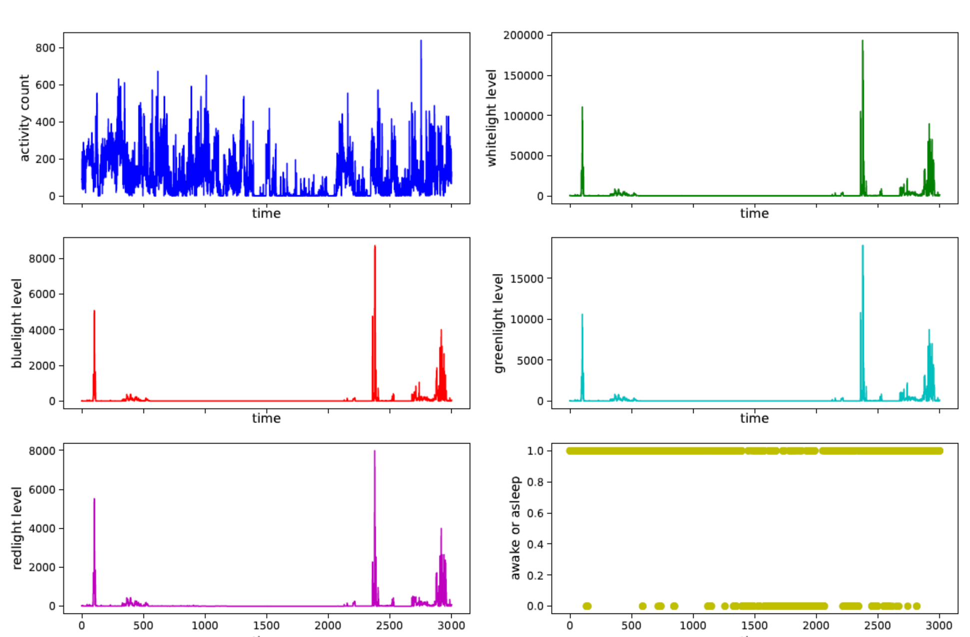
The current bottleneck for supervised learning on data from wearables is the lack of labelled datasets. Features extracted from self-supervised methods can be leveraged in downstream disease classification tasks. Through this we can see how user-generated data can be used to predict user-specific diseases and conditions, which would aid healthcare through early diagnosis. Therefore, there is a clear motivation to use self-supervised learning to extract features from unlabelled datasets which can be leveraged in downstream classification tasks. Early diagnosis of conditions can lead to a better understanding of the prognosis and can lessen the burden of the healthcare system.

Main contributions:

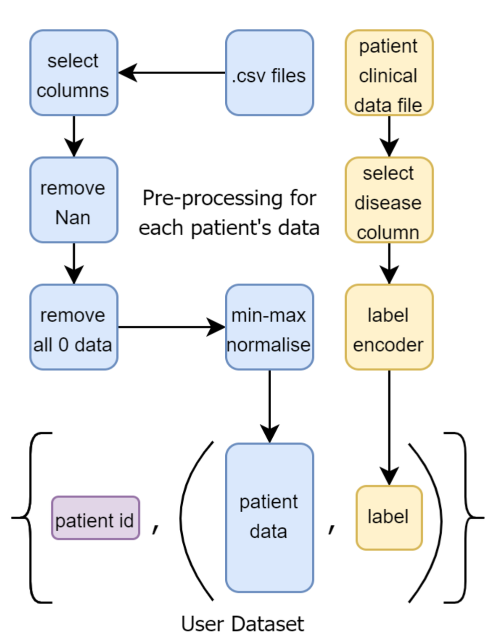
- We adapt and apply two self-supervised methods, SimCLR and BYOL, to learn features from a wearable activity dataset and use downstream classification tasks of different medical conditions to evaluate the quality of learned representations.
- We show that SimCLR outperforms the adversarial method and a fully-supervised method in the majority of evaluation tasks, and self-supervised methods always outperform the fully-supervised methods
- This work provides a comprehensive benchmark for contrastive methods applied to the wearable time-series domain, showing the promise of task-agnostic representations for downstream clinical outcomes.

Dataset and Pre-processing

The selected dataset is the Hispanic Community Health Study (HCHS) from the National Sleep Research Resource (NSSR). 1887 people with Latino origin between the ages of 18 and 75 had their activity data measured using a Philip's Actiwatch Spectrum wristwatch for 7 consecutive days. The time-series for each participant was sampled every 30 seconds, and metrics such as mean activity count, sleep or awake state and light levels were measured. Clinical annotations were provided that denote the insomniac, sleep apnoeic, diabetic, hypertension and metabolic syndrome status of each participant.

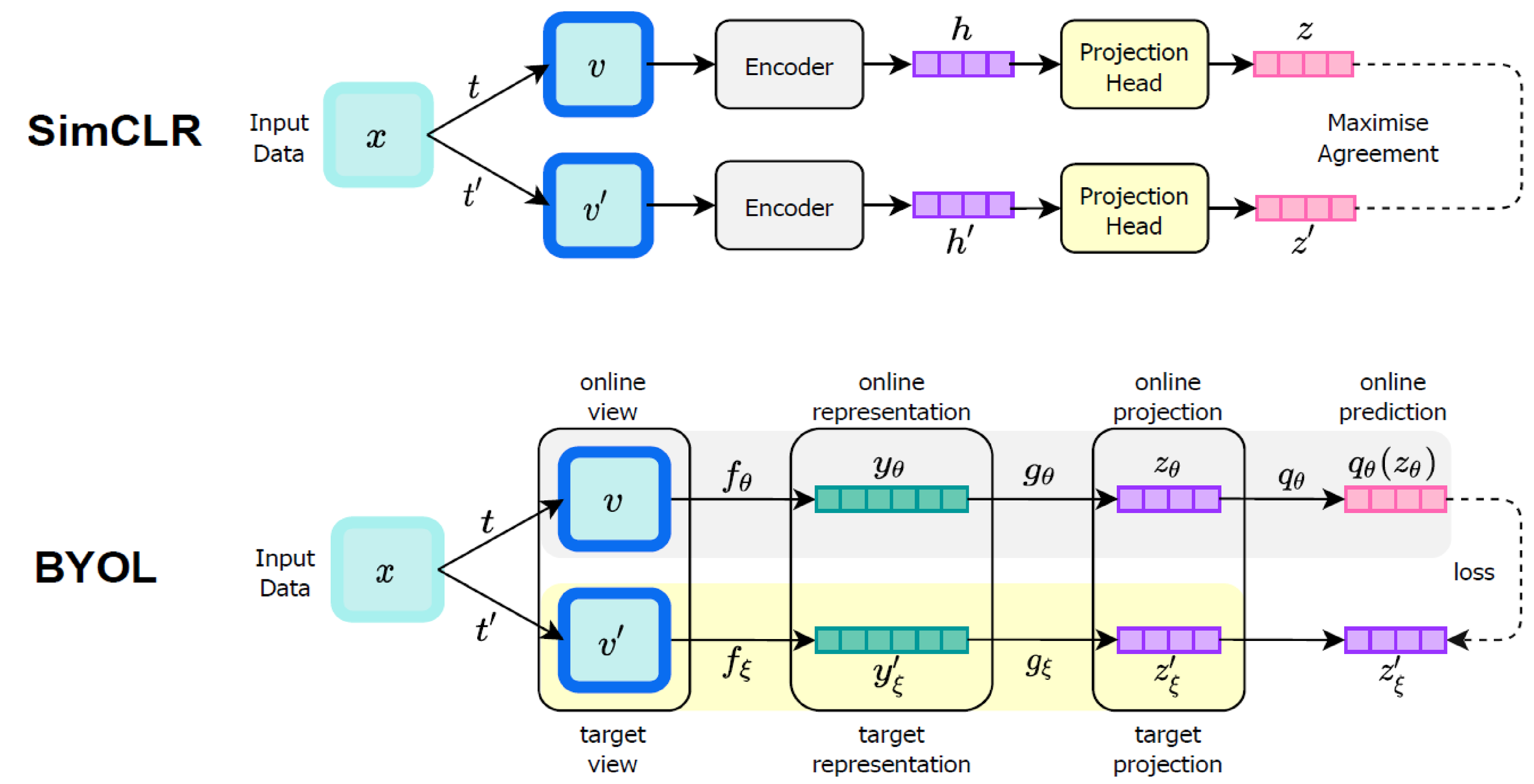


Plots showing the activity level, light levels and sleep or wake status of Participant 5270581 from the HCHS dataset



Pipeline for data pre-processing

SimCLR and BYOL Methodology



Components of the SimCLR and BYOL Frameworks

SimCLR

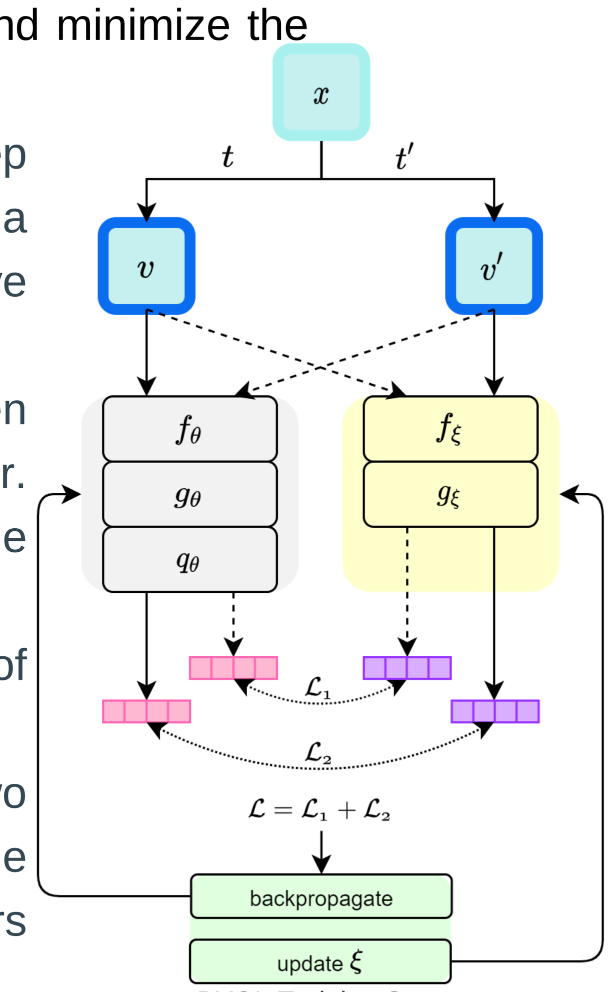
The SimCLR framework, originally proposed for image data, is able to efficiently learn useful representations without requiring specialised architectures or a memory bank. The framework is made up of 4 major components:

- **Data Augmentation:** Transformations such as Gaussian noise, scaling, negation and time reversal are used for time-series data are used to transform input data into correlated views.
- **Base Encoder:** The base encoder takes the augmented data as input and returns embeddings h_i. We use an encoder that consists of three 1D convolution layers with 10% dropout.
- **Projection Head:** A multilayer perceptron head, with three hidden layers, is used to map the representations to get projections.
- **Contrastive Loss Function:** NT-Xent loss with a LARS optimiser and a cosine decay schedule is used.

BYOL

This is a self-supervised framework that uses contrastive learning *without* negative samples. We use two copies of an encoder network, called the online and target networks, to obtain representation pairs, and minimize the contrastive loss between them. Details:

- **Transformations:** The augmentation of data is the first step of the BYOL algorithm, and this is done by applying a sequence of transformations to the input data, to achieve two different views.
- **Encoder and Projector:** We train an autoencoder, and then use the encoder part of the network as the BYOL encoder. An MLP head is attached to the encoder, which outputs the projection representation.
- **Predictor:** This is a linear layer that maps the projection of the online network to the projection of the target.
- **Training Step:** Each input data is transformed to get two views and the loss is calculated between the online prediction and the target projection. The online parameters are adjusted.



BYOL Training Step

Results

Evaluation Protocol

We use an 80%, 10%, 10% split for train, validation and test sets of the labelled HCHS dataset. For the self-supervised training of the contrastive models, we use the train set without using the labels. We report the mean F_1-scores with 95% confidence intervals of 10 runs, with the resulting learned representations evaluated on 5 downstream linear classification tasks, utilising a logistic regression classifier.

	HCHS									
Method	Sleep Apnea		Diabetes		Insomnia		Hypertension		Metabolic Syndrome	
	F_1-macro	F_1-micro	F_1-macro	F_1-micro	F_1-macro	F_1-micro	F_1-macro	F_1-micro	F_1-macro	F_1-micro
SimCLR	$\mathbf{50.8 \pm 1}$	$\mathbf{91.6 \pm 2}$	40.0 ± 2	48.8 ± 1	32.0 ± 4	60.0 ± 4	$\mathbf{47.6 \pm 3}$	75.1 ± 1	$\mathbf{52.7 \pm 4}$	$\mathbf{66.8 \pm 3}$
BYOL	48.0 ± 1	91.4 ± 2	23 ± 2	$\mathbf{50 \pm 3}$	25 ± 1	$\mathbf{61 \pm 3}$	43 ± 1	$\mathbf{76 \pm 2}$	42 ± 1	65 ± 2
Supervised	47.7 ± 0.4	91.2 ± 1.5	45.2	41.0	50.7	40.1	43.1	75.9 ± 2.2	39.7 ± 0.9	65.8 ± 2.6
day2vec	43.6	-	**45.8**	42.5	**55.7**	41.4	44.1	-	-	-

The best results achieved for each of the 5 downstream tasks with SimCLR and BYOL, with day2vec and task-specific CNN results shown for comparison. The mean and 95% confidence intervals of 10 runs is reported, apart from the results taken from *day2vec*, which reports just the mean.

Findings

- SimCLR and BYOL outperform the other techniques in the majority of the downstream classification tasks carried out, with the exception of the F_1-macro for the classification of diabetes. They outperform the fully supervised method and *day2vec* method for sleep apnea and metabolic syndrome.
- The classification of metabolic syndrome is a novel downstream evaluation task which has not been studied previously as far as we are aware with promising results seen especially with the SimCLR method for its classification.
- With hypertension, SimCLR can be observed to be performing the best when looking at the macro metric and BYOL when considering the micro metric.
- The classification task of both insomnia and diabetes shows the common trend that BYOL performs significantly better than the task specific CNN and *day2vec* in the micro metric, but for the macro metric the *day2vec* method works better.
- These results, and other related work, show that there is definite meaning extracted from the contrastive learning of time-series data, and these representations can be used to boost the performance of linear classification when compared to other fully supervised methods.

References

Grill, Jean-Bastien, et al. "Bootstrap your own latent: A new approach to self-supervised learning." *arXiv preprint arXiv:2006.07733* (2020).

Aggarwal, Karan, et al. "Adversarial unsupervised representation learning for activity time-series." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33. No. 01. 2019.

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

Chen, Ting, et al. "A simple framework for contrastive learning of visual representations." *International conference on machine learning*. PMLR, 2020.