Weak Label Learning for Wearable Cardio-respiratory Fitness Prediction

Yu Wu, Dimitris Spathis, Ignacio Perez-Pozuelo, Tomas I. Gonzales, Soren Brage, Nicholas J. Wareham, Cecilia Mascolo

University of Cambridge

Introduction. Wearable devices and smartphone applications for tracking physical activities and other health behaviors have enabled population-level health inferences studies, including sleep stage classification, stress detection, and fitness prediction [1]. However, compared with gold-standard annotations labeled by time-consuming and costly procedures, these silver-standard tests still contain various levels of estimation noise which may reduce model inference accuracy. Therefore, effectively training on labels with different levels of noise is challenging. This abstract presents our recent work on weak label learning by developing a transfer learning framework and an end-to-end (E2E) semi-supervised method. We validate our models on a challenging real-world health application: predicting lab-measured maximal oxygen consumption (VO_2max) through free-living wearables on two independent cohorts [2]. We use submaximal VO_2max tests (N=12,318) as the silver-standard (weak labels) and a smaller cohort that performed a maximum test to exhaustion (N=191) as the gold-standard for VO_2max and cardiorespiratory fitness (CRF)

Transfer learning. We first developed a transfer learning framework to leverage large weakly-labeled data and fine-tune it with gold-standard data. The pre-training model was trained using physiological signals derived from wearable sensors with weak labels. Then, we froze the backbone network during fine-tuning with a set of parameter sharing techniques. Our framework achieved the best performance of corr = 0.72±0.07 and RMSE = 5.085±0.584, outperforming baselines by 33%. Most notably, compared with baseline methods that struggled with extreme label shifts between high-quality and weak labels, we showed that our transfer learning network addressed the distribution shifts problem.

Semi-supervised learning. Second, we presented a versatile semi-supervised learning (SSL) framework for learning both weak labels and gold-standard labels. Our SSL network was trained in an E2E fashion and combined data augmentations with a joint loss function containing weak and high-quality labels. This SSL model demonstrated that learning labels with different noise level together also benefits the VO_2max prediction and gain similar performance to the transfer learning method (corr = 0.7±0.07 and RMSE = 5.662±0.789).

Conclusion. Our work examines the effectiveness of deep learning approaches for learning labels with varying degrees of noise by presenting a general transfer learning framework and an end-to-end semi-supervised learning method for the VO_2max prediction task. We demonstrate that through transfer learning and SSL, we can successfully employ noisy labels to enhance CRF prediction, which has implications for more accurate population-scale health monitoring with implications for overall health and mortality.

References

1. Hicks, J. *et al.* Best practices for analyzing large-scale health data from wearables and smartphone apps. *npj Digital Medicine* **2** (Dec. 2019).
2. Lindsay, T. *et al.* Descriptive epidemiology of physical activity energy expenditure in UK adults (The Fenland study). *International Journal of Behavioral Nutrition and Physical Activity* **16** (Dec. 2019).