Calibration of A Commercial Wearable Device with Research-Grade Physiological Signals

HINF5300 Term Project

Tinashe M. Tapera

Introduction

This course project aims to develop a calibration algorithm to improve the accuracy of heart rate data collected by a Garmin wearable using clinical research-grade heart rate data in a driving simulator study.

Wearable devices have become ubiquitous and effective tools for sensing digital biomarkers of health (e.g. physical activity, heart rate, respiration, location, etc.) and using these biomarkers to predict (and improve) health outcomes in at-risk groups and the general population (Coughlin and Stewart 2016; Knight et al. 2021; Burnham et al. 2018). As more devices become commercially available, research has scrutinised the accuracy and validity of these devices for measuring digitial biomarkers in controlled lab settings. A systematic review by Fuller et. al (2020) examined 158 publications that compared data collected on multiple Apple, Garmin, Samsung, Fitbit, Misfit, Polar, Xiaomi, Withings, and Mio wearables, to controlled in-lab data collection. The review found that while no one device demonstrated more systemic bias in error estimates, there were some interesting inaccuracies in device data reporting. When compared to manually counting steps, studies with numerous commercial wearable devices *underestimated* measurement error by 9%. Compared to electrocardiography, pulse oximetry, and the clinical Polar chest band device for measuring heart rate, 43.5% of studies with wrist worn devices were ±3% measurement error (Fuller et al. 2020). These inaccuracies suggest the need for calibration of commercially available wearables to ensure research conducted using these devices is valid and accurate.

Previous studies have documented the need for calibration of wearable devices. Freedson et. al (2012) and Bassett et. al (2012) both outline strategies for calibrating sensors for research "in the wild", including physical unit calibration (e.g. a vigorous shake for an accelerometer), converting signals to other measures for comparison, statistical aggregation, and various machine learning approaches. In line with these suggestions, a recent study on in-home ageing demonstrated that researchers were able to use a medical grade wearable to calibrate the quantification of in-home physical activity by passive infrared motion sensors. This lead to improved accuracy of the motion sensor in identifying and quantifying physical activity (Schütz et al. 2021). In a further example, in the development and deployment of a wearable phonocardiogram (PCG, wearable stethoscope), the proposed PCG device is calibrated using an embedded force sensor (Shyam Kumar, Ramasamy, and Varadan 2022).

Resources: The Dataset

This proposal will make use of the In-Vehicle Drunk Driving Detection (DRIVE) project dataset from the Bosch Internet of Things (IoT) Lab (Martin, n.d.), in collaboration with the UbiWell Lab at Northeastern University; the University of Berne Institute of Forensic Medicine; Center for Digital Health Interventions (CDHI) at ETH Zurich; and the University of St. Gallen. Data availability is provided by the UbiWell Lab. This project was initially proposed to "build a reliable in-vehicle drunk driving detection system". In one study, participants completed a driving simulator task in-lab, once while sober, once after consuming enough alcohol for blood alcohol content (BAC) to be over the legal limit for driving while intoxicated, and once more after their BAC had returned below the legal limit. Importantly for this project, participants were wearing a number of passive sensors throughout this study, including a Garmin wrist wearable and a clinical-grade electrocardiogram (ECG).

Proposed Analysis

It is expected that compared to the ECG, the Garmin may under-perform in a machine learning prediction task using this dataset. The proposed plan is to develop a baseline machine learning pipeline using the Garmin data, as a proxy for current research methods for activity recognition using commercial devices. Then, the signal from the Garmin can be compared to the clinical grade ECG signals. The Garmin signal can then be augmented by calibrating it to the ECG, similar to a pre-processing step. Once this is done, the machine learning pipeline re-run and evaluated using this calibrated signal. The expected deliverable for this plan should be ready at the end of the fall semester 2022. For future work, the calibration algorithm can be packaged into a shareable module that can be evaluated on other machine learning tasks using the Garmin. Additionally, more complex models such as Transfer Learning could be investigated.

References

Bassett, David R., Ann V. Rowlands, and Stewart G. Trost. 2012. "CALIBRATION AND VALIDATION OF WEARABLE MONITORS." *Medicine and Science in Sports and Exercise* 44 (1 Suppl 1): S32–38. https://doi.org/10.1249/MSS.0b013e3182399cf7.

Burnham, Jason P, Chenyang Lu, Lauren H Yaeger, Thomas C Bailey, and Marin H Kollef. 2018. "Using Wearable Technology to Predict Health Outcomes: A Literature Review." *Journal of the American Medical Informatics Association* 25 (9): 1221–27. https://doi.org/10.1093/jamia/ocy082.

Coughlin, Steven S., and Jessica Stewart. 2016. "Use of Consumer Wearable Devices to Promote Physical Activity: A Review of Health Intervention Studies." *Journal of Environment and Health Sciences* 2 (6): 10.15436/2378–6841.16.1123. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5395205/.

Freedson, Patty, Heather R. Bowles, Richard Troiano, and William Haskell. 2012. "ASSESSMENT OF PHYSICAL ACTIVITY USING WEARABLE MONITORS: RECOMMENDATIONS FOR MONITOR CALIBRATION AND USE IN THE FIELD." *Medicine and Science in Sports and Exercise* 44 (1 Suppl 1): S1–4. https://doi.org/10.1249/MSS.0b013e3182399b7e.

Fuller, Daniel, Emily Colwell, Jonathan Low, Kassia Orychock, Melissa Ann Tobin, Bo Simango, Richard Buote, et al. 2020. "Reliability and Validity of Commercially Available Wearable Devices for Measuring Steps, Energy Expenditure, and Heart Rate: Systematic Review." *JMIR mHealth and uHealth* 8 (9): e18694. https://doi.org/10.2196/18694.

Knight, Stephen R., Nathan Ng, Athanasios Tsanas, Kenneth Mclean, Claudia Pagliari, and Ewen M. Harrison. 2021. "Mobile Devices and Wearable Technology for Measuring Patient Outcomes After Surgery: A Systematic Review." *Npj Digital Medicine* 4 (1): 1–14. https://doi.org/10.1038/s41746-021-00525-1.

Martin. n.d. "DRIVE: In-Vehicle Drunk Driving Detection." https://www.iot-lab.ch/projects-connectedmobility/drive/.

Schütz, Narayan, Hugo Saner, Angela Botros, Philipp Buluschek, Prabitha Urwyler, René M. Müri, and Tobias Nef. 2021. "Wearable Based Calibration of Contactless in-Home Motion Sensors for Physical Activity Monitoring in Community-Dwelling Older Adults." *Frontiers in Digital Health* 2. https://www.frontiersin.org/articles/10.3389/fdgth.2020.566595.

Shyam Kumar, Prashanth, Mouli Ramasamy, and Vijay K. Varadan. 2022. "Evaluation of Signal Quality from a Wearable Phonocardiogram (PCG) Device and Personalized Calibration." *Electronics* 11 (17): 2655. https://doi.org/10.3390/electronics11172655.