A Wearable ECG Pipeline for Free-Living VO_2max Prediction and HRV Interpretation

Aron B. Syversen*[1], Zhi-Qiang Zhang[1], David Jayne[1], David Wong[1]

[1]University of Leeds, *scabs@leeds.ac.uk

Abstract:
Assessing cardiorespiratory fitness is a key step in evaluating patient health for surgical risk. The gold-standard method, cardiopulmonary exercise testing (CPET), measures maximal oxygen uptake (VO_2max), but it is resource-intensive, requiring specialist equipment, staff, and maximal effort from the patient. These barriers make it difficult to implement at scale, particularly in frail or preoperative populations. As a result, there is growing interest in research using wearable sensors to estimate fitness levels remotely. Processing ECG data from free-living environments poses challenges to signal quality and demands a robust extraction pipeline.

In this work, we present a pipeline developed to predict VO_2max from raw wearable ECG and accelerometer data collected during daily life, using a preoperative patient cohort. We start by evaluating four open-source signal quality index (SQI) tools on synthetic ECGs with controlled noise and compare their output to cardiologist assessments. From this, we select the SQI most suitable for heart rate extraction, though our results also highlight the inconsistencies and limitations of current SQIs for ECG monitoring.

Building on this, we applied an ECG processing pipeline to real-world data from 198 patients, collected from Leeds Teaching Hospitals NHS Trust as part of the REMOTES study. We extracted heart rate, physical activity, and heart rate variability (HRV) features, alongside demographic data and compared four machine learning models (SVR, RF, XGBoost, MLP) against multiple linear regression for VO_2max prediction. Linear regression outperformed more complex models, suggesting that in diverse clinical populations, simpler models may generalise better.

Finally, we extended the linear model to assess the contribution of HRV features. Specifically, we compare short-term HRV measures collected from clean periods of resting ECG alongside long-term HRV collected over 24 hours extrapolated from heart rate signals. We conclude that HRV measures offered moderate improvements to predictions when included, and that long-term HRV was consistently among the top predictive features, outperforming short-term HRV measures. Long-term HRV can be derived from heart rate data, making it robust to noise and suitable for real-world wearables.

This pipeline demonstrates a practical approach to remote VO_2max estimation and highlights two key takeaways: (1) signal quality assessment is a key component of wearable ECG processing but should be purpose-specific, and (2) long-term HRV features may offer more stable insight than acute measures in uncontrolled settings. These findings support further evaluation of wearable pipelines for patient monitoring in preoperative assessment.