European Partnership on Metrology
Call 2022 – Digital Transformation, Health, Integrated European Metrology, Normative and Research Potential

Selected Research Topic number: **SRT-H11**
Version: 1.0

Title: Uncertainty quantification for machine learning models applied to photoplethysmography signals

Abstract

Photoplethysmogram (PPG) signals are easy to collect non-invasively using cheap devices and are used in the clinic and in wearable devices for home monitoring. It is recognised that PPG signals contain a wealth of valuable physiological information for monitoring or diagnosing a range of health conditions. Machine learning (ML) is applied to PPG signals but there is a lack of work on trustworthiness, which is crucial in a medical context. By developing methods to quantify both the data and model uncertainty for ML applied to PPG signals, this project aims to generate reference datasets to benchmark such models and to identify models with high accuracy and low uncertainty thus providing trustworthy models that are ripe for implementation.

Keywords

Photoplethysmography; supervised machine learning; deep learning; uncertainty quantification; aleatoric (data) uncertainty; epistemic (model) uncertainty; benchmarking

Background to the Metrological Challenges

Photoplethysmography (PPG) is an optical technique that makes measurements at the surface of the skin to detect volumetric changes in peripheral blood circulation. PPG signals, read using a pulse oximeter, are routinely used in clinical settings for determining heart rate and blood oxygen saturation. It is now recognised that PPG signals contain valuable information on the cardiovascular, respiratory and autonomic nervous systems which are not yet routinely exploited.

In addition, PPG signals are collected by many wearable devices and have the potential to be used in a variety of continuous home monitoring applications such as blood pressure (to detect hypertension), continuous blood glucose monitoring and for the detection of peripheral artery disease. The data is easy to obtain non-invasively and PPG devices are inexpensive.

Similarly, detection of atrial fibrillation, monitoring stress, and detection of sleep apnoea, are all common conditions which can be detected from PPG signals and are easily treatable, but which often go undetected.

Currently issues with PPG signals include signal quality, motion artefacts and inaccuracies due to different skin tones. In addition, it is known that there are sex differences in ECG parameters, so signal-based diagnosis should be sex-specific, although many machine learning models currently do not take this into account.

Supervised machine learning, a subcategory of machine learning and artificial intelligence that uses labelled datasets to train algorithms, classify data, or accurately predict outcomes, can be used on PPG data to get information on regression problems (such as determining blood pressure or arterial stiffness) and classification problems (including diagnosis of heart arrhythmias, and peripheral artery disease).

In addition, however, when applying machine learning in medicine, confidence in the model and its predictions is crucial since mistakes due to false negatives can be harmful to patients and potentially fatal and false positives can lead to patient anxiety and unnecessary treatment.

Uncertainty arises from two sources, namely the model (epistemic) and the data (aleatoric), and an understanding of both is required. It is also important to validate the uncertainties in a variety of contexts to ensure that they are reliable and consistent measures of uncertainty. Machine learning is a developing field with a variety of different methods proposed and the best approach, particularly in the context of PPG signals, is not known which limits its usefulness. By providing uncertainties for machine learning models, both on the

PPG data and the model itself this project aims to transform the analysis of PPG signals from academic study to trustworthy algorithms which are ripe for implementation and exploitation by European industry.

Objectives

Proposers should address the objectives stated below, which are based on the PRT submissions. Proposers may identify amendments to the objectives or choose to address a subset of them in order to maximise the overall impact, or address budgetary or scientific / technical constraints, but the reasons for this should be clearly stated in the protocol.

The overall objective is to provide trustworthy machine learning models for analysing photoplethysmography signals in a medical context by developing methods for the quantification of uncertainty in supervised machine learning and deep learning models applied to photoplethysmography signals and generating reference datasets to benchmark those models.

The specific objectives are:

1. To develop methods for quantifying the uncertainty in at least 3 existing supervised machine learning and deep learning models for classification and regression problems using photoplethysmography data, considering the effects of both aleatoric (data) uncertainty and epistemic (model) uncertainty on model predictions.
2. To generate at least 5 measurement problems and their corresponding datasets, using real and/or synthetic photoplethysmography data, that can be used to benchmark accuracy and uncertainty of supervised machine learning and deep learning models. In addition, to make those reference problems and datasets widely available to the medical device and digital health communities.
3. To validate the uncertainties obtained for existing machine learning and deep learning models of Objective 1 and to compare the accuracy and uncertainty of at least 3 existing machine learning and deep learning models in order to identify models and methods, which have high accuracy and low uncertainty for a wide range of tasks.
4. To engage with the medical device, digital and health communities to (a) promote and enable the uptake and implementation of the methods for uncertainty quantification developed during the project, (b) support the adoption of the benchmarking problems and datasets by providing guidelines for their use, and (c) develop a framework for independently reviewing machine learning models proposed by industry to assist them in getting regulatory approval.
5. To facilitate the take up of the technology and measurement infrastructure developed in the project by the measurement supply chain (NMIs, DIs, medical device calibration services,), standards developing organisations (IEC, ISO,), and end users (clinical practitioners, digital experts within the health communities, manufacturers of medical and healthcare products).

These objectives will require large-scale approaches that are beyond the capabilities of single National Metrology Institutes and Designated Institutes, and it is expected that multidisciplinary teams will be required. To enhance the impact of the research, the involvement of the appropriate user community such as medical practitioners, medical (academic) hospitals and industry is strongly recommended, both prior to and during methodology development.

Proposers should establish the current state of the art and explain how their proposed project goes beyond this.

EURAMET expects the average EU Contribution for the selected JRPs in this TP to be 1.9 M€ and has defined an upper limit of 2.6 M€ for this project.

EURAMET also expects the EU Contribution to the external funded beneficiaries to not exceed 35 % of the total EU Contribution across all selected projects in this TP.

Any industrial beneficiaries that will receive significant benefit from the results of the proposed project are expected to be beneficiaries without receiving funding or associated partners.

Potential Impact

Proposals must demonstrate adequate and appropriate participation/links to the ‘end user’ community, describing how the project partners will engage with relevant communities during the project to facilitate knowledge transfer and accelerate the uptake of project outputs. Evidence of support from the ‘end user’ community (e.g. letters of support) is also encouraged.

You should detail how your JRP results are going to:

- Address the SRT objectives and deliver solutions to the documented needs,
- Feed into the development of urgent documentary standards through appropriate standards bodies,
- Transfer knowledge to the medical sector.

You should detail other impacts of your proposed JRP as specified in the document “Guide 4: Writing Joint Research Projects (JRPs)”

You should also detail how your approach to realising the objectives will further the aim of the Partnership to develop a coherent approach at the European level in the field of metrology and include the best available contributions from across the metrology community. Specifically, the opportunities for:

- improvement of the efficiency of use of available resources to better meet metrological needs and to assure the traceability of national standards
- the metrology capacity of EURAMET Member States whose metrology programmes are at an early stage of development to be increased
- organisations other than NMIs and DIs to be involved in the work.

Time-scale

The project should be of up to 3 years duration.