👍 or 👎: Simple quality assessment tools for physiological signals

OHBM 2024
JUNE 23~27, SEOUL, KOREA

Roza G. Bayrak[1], Richard Song[2], Rithwik Guntaka[2], Ruogi Yang[1], Catie Chang[1]

[1]Computer Science, Vanderbilt University, [2]Vanderbilt Brain Institute, Nashville, TN,

NEURDY lab
VANDERBILT UNIVERSITY

HIGHLIGHTS

- A simple classification pipeline is developed to assess the quality of raw physiological measures, currently focusing on respiration and cardiac waveforms.
- The proof-of-concept tool is able to classify, the quality of respiration data, with 83.35 ± 1.01% and cardiac data with 88.49 ± 1.42% accuracy.

INTRODUCTION

Background: While traditionally regarded as noise, systemic physiological processes are frequently shown to be linked with cognitive processes and may contribute valuable information to fMRI studies [1,2,3] . Recognizing this, neuroimaging research increasingly draws upon concurrent recordings of peripheral physiology to enhance fMRI analysis.

Motivation: However, usefulness of physiological data is contingent upon the quality of the recordings as well as expertise in data handling. Quality assessment is not only a tedious process, but the assessments vary significantly between raters.

Gap: While there are manual and template-based tools assessing peak detection quality (physiopy's *peakdet*, PhysIO, etc.), and automated exclusion criteria based on statistical summary metrics, currently there are no automated approaches that can provide a quick quality check.

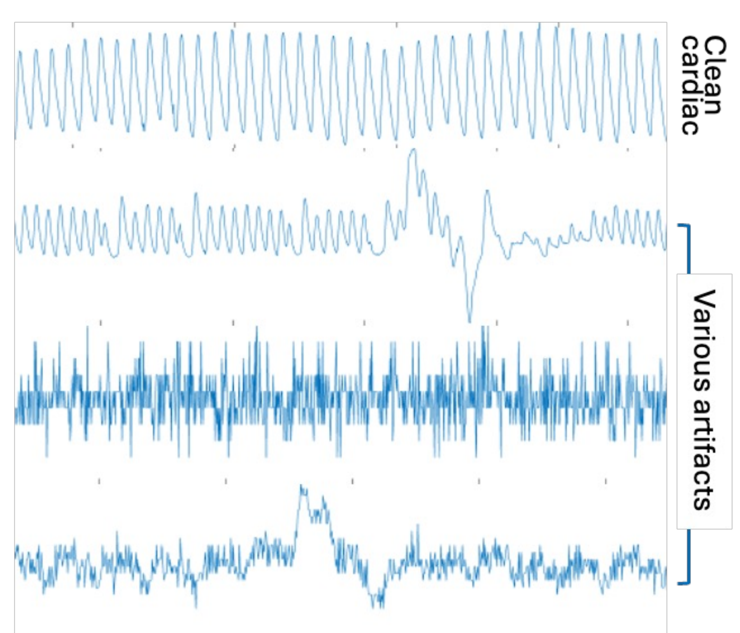

Fig. 1 Various cardiac waveforms, some containing artifacts.

Proposed: A simple deep-learning powered tool to assess the quality of peripheral physiological recordings and a complementary interactive GUI for human-in-the-loop assessments.

RESULTS AND DISCUSSIONS

- Here, we provide a simple thumbs up / thumbs down tool that can save several hours of manually vetting physiological recordings.
- To the best of our knowledge, this is the first attempt at developing a DL-based physiological QA method.
- The models were able to classify the quality of respiration data with 83.35 ± 1.01% and cardiac data with 88.49 ± 1.42% accuracy.
- ML tool aims to ensure data integrity, while it could be utilized to describe data quality issues and suggest steps for fixing the data, thereby promising to improve the accuracy and reliability of downstream research.
- The rater annotations were not consistent, we speculate the current accuracy ceiling is due to this issue.

FUTURE DIRECTIONS

We envision that in future iterations, the tool can be further developed by adding new modules to:
- generate text-based reports detailing specific reasons for why the quality check for a given recording has failed and whether it is fixable (e.g., if a recording is partly usable, or if a simple interpolation algorithm could fix the problem),
- provide suggestions for fixing the data, (i.e. HeartPy clipping spline interpolation, and
- apply the suggested fix and return the corrected data (See Fig 3.).

Our GUI is currently developed in MATLAB, we will shift the efforts to developing a python tool.

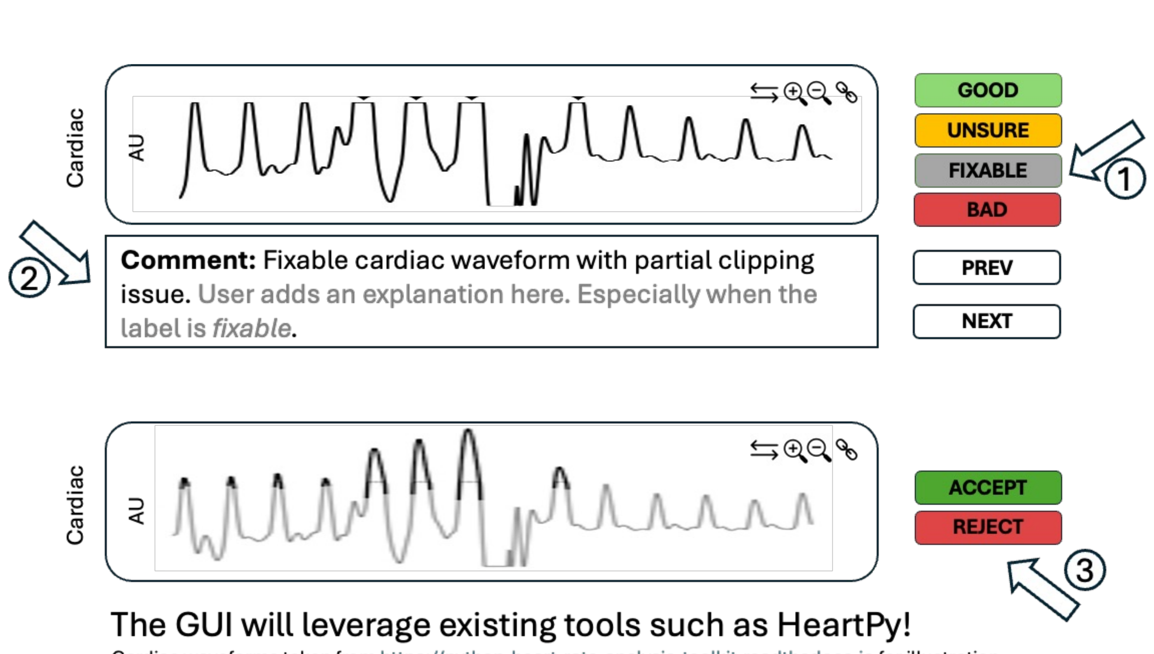

Fig. 3 A possible artifact correction scenario. The annotator:
(1) chooses a label, i.e. Fixable,
(2) provides a text-based description of the artifact and the potential fix,
(3) accepts or rejects the suggested fix returned by the intreactive platform.

METHODS

Quality Assessment

Option to manually and/or *automatically* QA raw physiological recordings

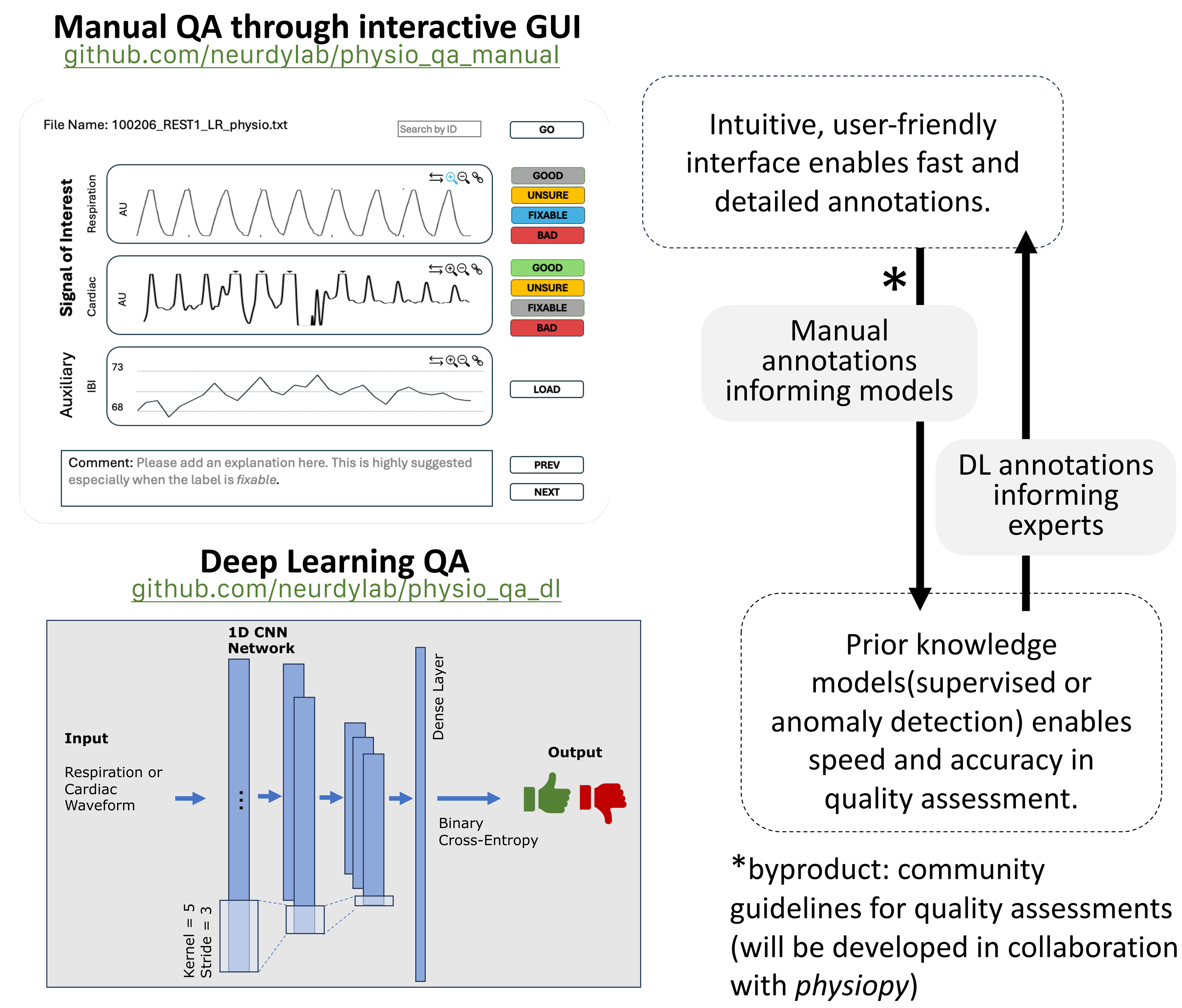

Fig. 2 Quality assessment tools and the iterative pipeline.

Deep Learning (DL) Models:

Data and preparation:
- Cardiac and respiratory belt recording from the HCP1200 resting-state dataset are used. Models are trained only with edge-cases (good/bad).
- Originally waveforms sampled at 400 Hz, downsampled by a factor of 4, temporally normalized to zero mean and unit variance.

Model Training:
- The network is composed of stacked 1D CNNs (Convolutional Neural Networks) with decreasing feature map sizes at each layer (Fig. 2, bottom)
- Learning rate of 0.0001, batch size of 2, employing Adam as the optimizer and binary cross-entropy as the loss function.

Model Testing and Validation:
- The dataset is divided into training and test sets, employing rotating partitions in a 5-fold cross-validation framework.
- To prevent overfitting, an early stopping criterion is applied based on the performance on the validation set.

The DL tool:
(1) takes in raw recordings,
(2) writes out the annotations in a .csv format.

Interactive GUI:
Manual annotations are cumbersome. However, manually assessing the quality of labels is vital in this initial stage to ensure the accuracy of our models. So, to enable our supervised neural networks, we:
- developed an in-house annotation tool that is a matlab-based GUI that enables fast annotations of physiological signals,
- labeled the HCP cohort using this GUI.

The manual tool:
(1) takes in raw recordings,
(2) plots full length raw time series,
(3) provides the rater (annotater) with visual information for quality inspection and annotation (Fig 2, top),
(4) writes out the annotations in a .csv format.

REFERENCES

[1] Shokri-Kojori et al. 2018, PMID: 29955858
[2] Mather & Thayer 2018, PMID: 29333483
[3] Yuan et al. 2013, PMID: 23631982

SCAN ME

ACKNOWLEDGEMENTS

This work was supported by NIH grant R01 MH125931 (C.C.).

For more info please contact: roza.g.bayrak@vanderbilt.edu
NEURDY lab git repo for code and more: github.com/neurdylab