BE275-PROJECT PROPOSAL: SHIFT-ROBUST LOSS FUNCTION FOR RPPG

Ananya Deepak Deoghare
UID: 005627628
ananyadd1607@g.ucla.edu

Kimaya Kulkarni
UID: 805528337
kimayak@g.ucla.edu

ABSTRACT

Asynchronized data can't be used to train models as state of the art methods expect inputs and targets to align perfectly. In case of datasets where the shift in data exists, manual alignment of data is required. Recent methods focus more on architectural changes which may not be practical for all applications. Hence we propose a novel shift robust loss function that enables the network to learn from misaligned input data and ground truth data.

1 INTRODUCTION AND MOTIVATION

Contactless HR estimation methods have been developed with the use of Computer vision or deep learning algorithms. Remote PPG (rPPG) works by looking at slight colour changes in mainly the facial region.

Recent work has improved upon this method by incorporating face tracking, skin segmentation, color space transformation, signal decomposition and filtering steps. However, these techniques are expensive to implement and adapt to various lighting conditions, motion and noise. Additionally, these deep learning models are trained on ground truth data collected by contact devices. Existing methods on vital signs detection have focused on obtaining state-of-the-art performance, with little practical consideration of the input data for the network. One such consideration is the misalignment of the input and ground truth data. For example, in rPPG, we estimate the heart rate using videos of the subject's face, but the ground truth data is collected using finger oximeter. The rPPG signal measured from the relevant areas of the face is slightly shifted from the PPG signal detected by pulse oximeter, resulting in misaligned video frames and ground truth data.

2 DATASET

UBFC-rPPG (stands for Univ. Bourgogne Franche-Comté Remote PhotoPlethysmoGraphy) comprises two datasets which are focused specifically for rPPG analysis. The UBFC-RPPG database was created using a custom C++ application for video acquisition with a simple low cost webcam (Logitech C920 HD Pro) at 30fps with a resolution of 640x480 in uncompressed 8-bit RGB format. A CMS50E transmissive pulse oximeter was used to obtain the ground truth PPG data consisting of the PPG waveform as well as the PPG heart rates. During the recording, the subject sits in front of the camera (about 1m away from the camera) with his/her face visible. All experiments are conducted indoors with a varying amount of sunlight and indoor illumination.

3 APPROACH

Apart from asynchronized data, another reason for the shift is the physiological delay in remote PPG signal measured from the face as compared to the reference PPG signal measured from finger oximeter. The pulse transit time from finger to face leads to the shift between the PPG signal extracted from the finger oximeter and camera.

We will use the below 2 models and try making modifications to the same, Since they have established a ground level/ base level which will help us start the project.

- **DeepPhys** is the first end-to-end system for video-based measurement of Heart Rate and Breathing Rate using CNN. The network learns spatial masks that are shared between the models and features, important for recovering Blood Volume Pulse and respiration signals.
- **Physnet** (another model) is an end-to-end spatio-temporal network for rPPG signal measurement from raw facial videos. The steps of projecting RGB into color subspace with stronger representation capacity and getting rid of the irrelevant information to achieve the target signal are combined into one step to obtain the final rPPG signal.

The concepts that we learnt in class will be very helpful for us to move forward in this project. Some of the concepts which we will use are:

- K fold Cross Validation
- Pearson Correlation (Covariance)
- PCA
- Mean Squared Error Metrics
- Mean Absolute Error Metrics

We use K fold Cross Validation and not bootstrapping for the videos dataset, even though we might get better results with Bootstrapping. Because, Cross Validation methods overestimate the error unlike Bootstrapping where the error is underestimated.

We want to create a loss function which can learn from misaligned signal data, since current state-of-the-art methods require perfect alignment. The ground truth signal from the pulse oximeter has a considerable lag behind the estimated signal from the relevant areas of the face. To account for this, we want to add a shift in our input and target data. We can then train our model on various loss functions.

4 More about the Project

We believe the project should not take more than a month. Below is the timeline for the same:

- Week 1: Gathering the dataset and code base
- Week 2: Try and run DeepPhys and PhysNet on UBFC Dataset
- Week 3: Come up with 2-3 Loss functions which we can test
  Implementing PCA on the dataset, and passing that into PhysNet and DeepPhys
- Week 4: Add K fold Cross Validation to the loss functions.
- Week 5: Document the results

We are aware the project may not be very easy and there might be a few hiccups in implementing the loss functions and getting everything ready but we are cautiously optimistic about the same!

5 Final Goal

We are hoping to be able to work on modifying and generating atleast 2-3 loss functions to understand and improve on the present Loss functions used in rPPG. We are also hoping to understand how much error does asynchronized data cause, and if the data should be synchronized or not.