Omdena HeartKinetics Project

Abstract. This document discusses a model designed as an early diagnosis tool for heart failure based on patients' biometric characteristics and transformed input signals representing the mechanical activity of the heart. This presents the methodology used for data preparation, data splitting, and the model architecture. It also introduces a segment discriminator technique based on autocorrelation to handle irregular signal amplitudes. The text concludes with a summary of the results, future work suggestions, and concluding remarks.

Introduction

The goal of the model is to estimate whether a patient presents a heart failure condition by analyzing their biometric characteristics and transformed input signals. These signals capture the mechanical activity of the heart at the surface of the human chest. To achieve this, this document explores the data preparation process and addresses the challenges posed by the original data format.

Methodology

Data Preparation

The original raw data consists of subject study records, each containing biometric data like age, weight, height, and gender, and five signal recordings or channels for linear energy, linear power, rotational energy, rotational power, and respiratory activity. To improve exploratory and training tasks, the data was reformatted and normalized into three entities: subjects, records, and signals, as shown in Figure [Data Schema]. This normalization reduces redundancies and provides semantic separation of the data sources.
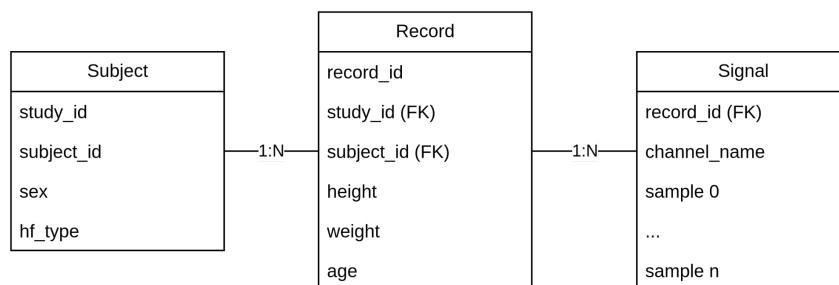

Figure [Data Schema]. Data entities

Data Splits

Since one subject may have multiple associated records, it is crucial to prevent data leakages by carefully assigning each subject to only one of the dataset while performing data splits. This guarantees that no patient information is present in both the train and test or validation datasets. To ensure proper and standardized data splitting for model training and

evaluation we introduced the DataManager class, check [notebook] for implementation details.

Model

Pool of Predictions

Instead of inputting the entire signal into the model, we split it into consecutive segments of a predefined duration, such as 10 seconds, and train a model using them. These segments may or may not have time overlaps between contiguous segments. The advantage of this approach is twofold: first, it works as a data augmentation technique by generating multiple data points for the record, and second, it mimics the human focus mechanism used when checking this type of signal, where the attention span is limited to a couple of seconds at a time.
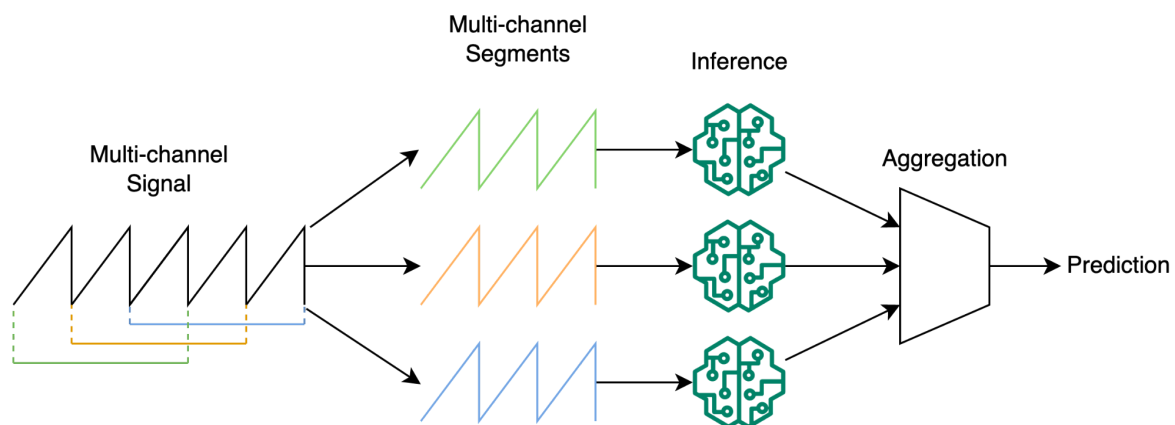

Figure [Pool Architecture] Multi segment pool approach diagram. Where one channel signal is split in a pool of three contiguous segments.

This process also helps us face the heterogeneity of the signal durations. The pool of prediction scores can be aggregated in several ways. In our experiments, we tried using the maximum aggregator, which takes the maximum score of the pool of predictions as the output. Generally, this technique seems to increase the F1 and F2 scores when comparing it with the pure-model inference.

It is worth noticing that while the trained model inputs multi-channel segments, the inference pipeline takes a full multi-channel signal and builds the previously described pool of predictions.

Model Architecture

During our experimentation, we tried several network architectures. The one that seems to have smoother learning curves and better results is represented in the figure architecture. This has one independent CNN backbone for each of the five record channels (side-to-side dotted boxes labeled as sequential 12 to 16 in the figure architecture). Each backbone is a stack of two (CONV 1D + Max pooling) blocks with 32 filters each, which should help to identify useful features from the channel signals.

The five backbone outputs are combined with the biometric patient characteristics and the signal features obtained from the segment feature engineering stage to be the input of a shallow head with a 32-neuron dense layer (see sequential 17 in the figure architecture). To increase generalization and avoid overfitting, this head includes a Dropout layer with a dropout rate of 0.5.
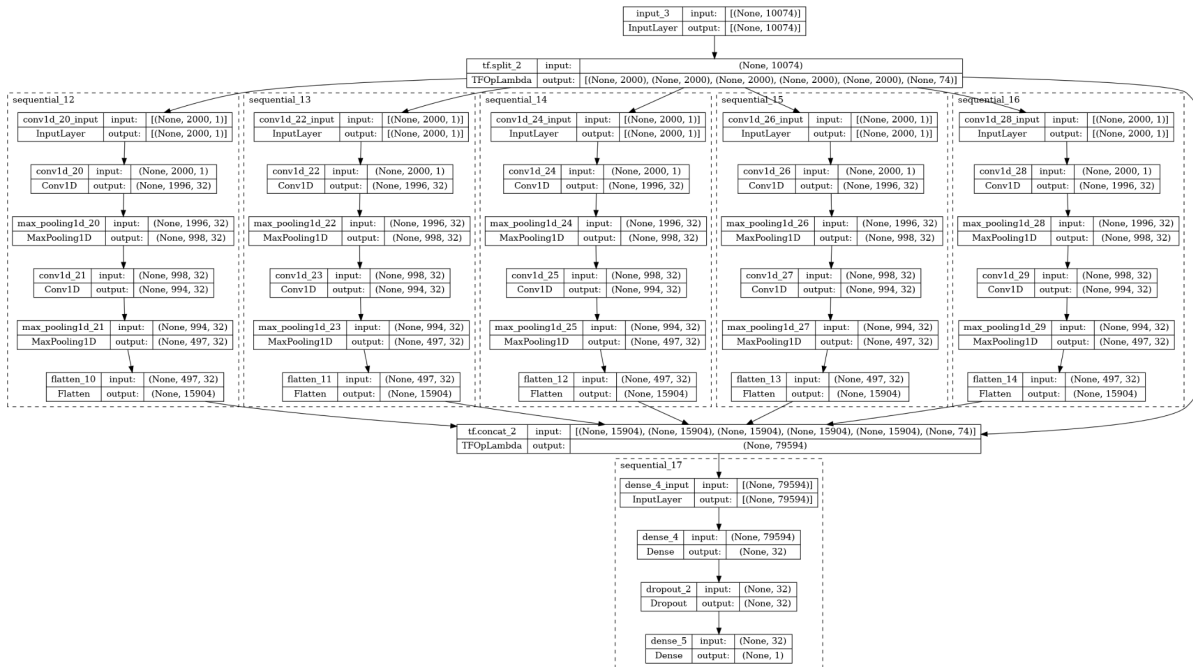

Figure [Architecture]. Five-backbone model architecture for 10 seconds duration segments with sample rate 200 Hz and 14 features engineered from signals per channel[1].

To get more familiar with the input structure, let's consider the input size of the architecture in the Figure [Architecture] as an example. It is for 10 seconds segments at a sample rate of 200 Hz, which represents 2000 inputs per channel, for a total of 10,000 inputs associated with the convolutional backbone of the network. In this case, 14 features were engineered from each channel segment, adding 70 extra inputs. Finally, the biometric characteristic inputs are 4 (age, weight, height, and sex), making a total of 10,074 model inputs.

It is worth noticing that the model does not use data with an unknown heart failure type for training. Similarly, to exclude the effect of different devices, we excluded the Frontiers study samples as they were not collected with the KINO.md device.

Segment Discriminator

Due to different reasons such as instrumentation noise or involuntary subject movements, the signal amplitude may become irregular or present spikes, see Figure [Segments comparison]. To tackle this, we have proposed a segment discrimination technique based on autocorrelation. Our main assumption is that, in the absence of external influences (i.e., ideal conditions), the underlying phenomena should produce highly periodic signals, as seen in heartbeats and regular breathing.

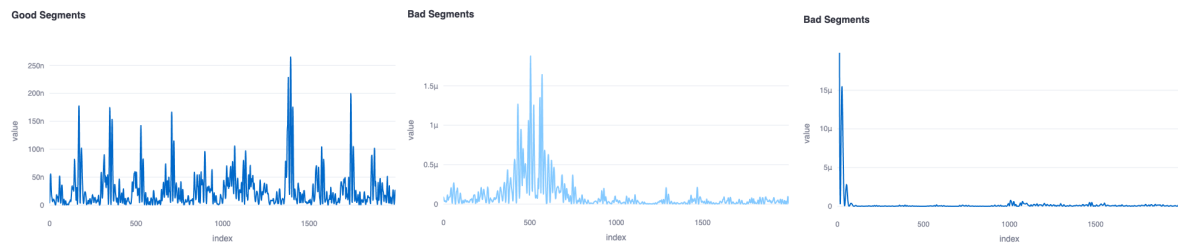[1] Other architectures we tried include a single backbone in a multidimensional CNN [check notebook].

Figure [Segments comparison]. Three segments of the same signal (HAP-PWV_AR_1_20220216-104303_MA5xMxRwYhEESb) are automatically discriminated by the autocorrelation technique.

To illustrate the segment discrimination working principle, consider the figure [Autocorrelations], where you will see different segments and their autocorrelation plots and scores. Note that the more regular the signal, the larger the area between the autocorrelation's maximum and minimum envelopes[2].

In order to establish whether a segment is regular or not, we set a threshold for the discriminator. The higher the threshold, the higher the rejection rate of "regular" segments. The selection of this limit was done manually after some trial and error. However, it can be included as part of the hyperparameter fine-tuning.

_____
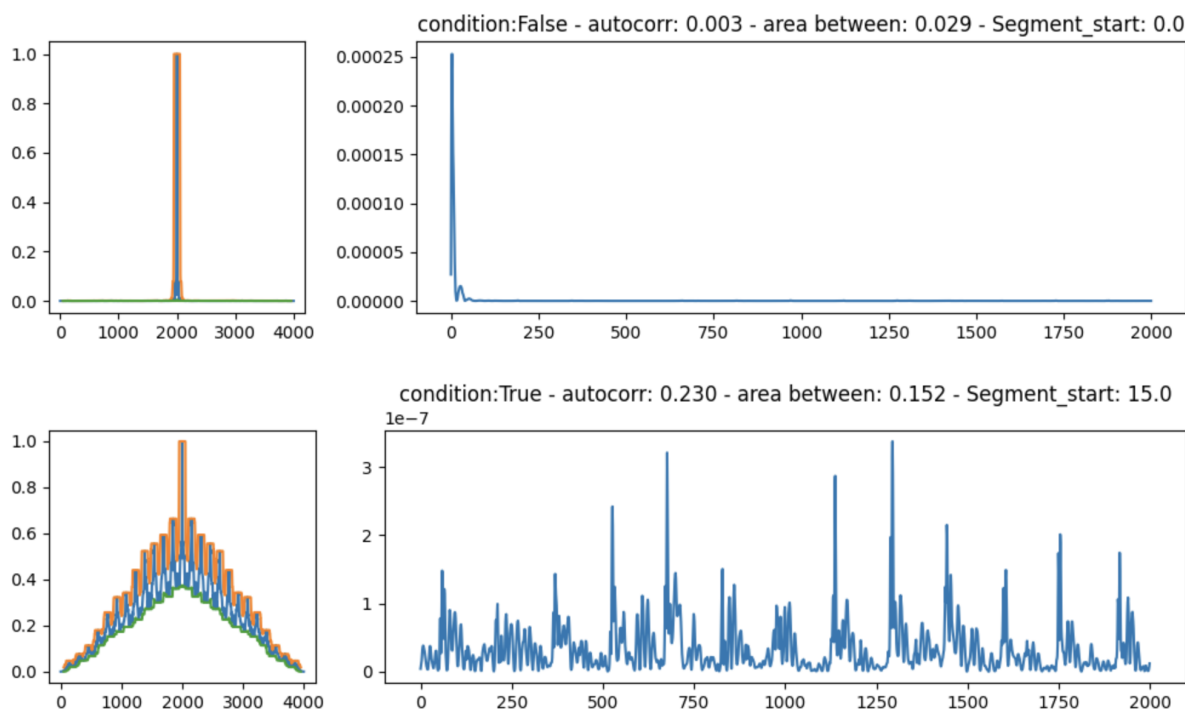
[2] Note this area is divided by twice the length of the signal. To get the discrimination score (check autocorrelation segment checker.ipynb for more details about its implementation).

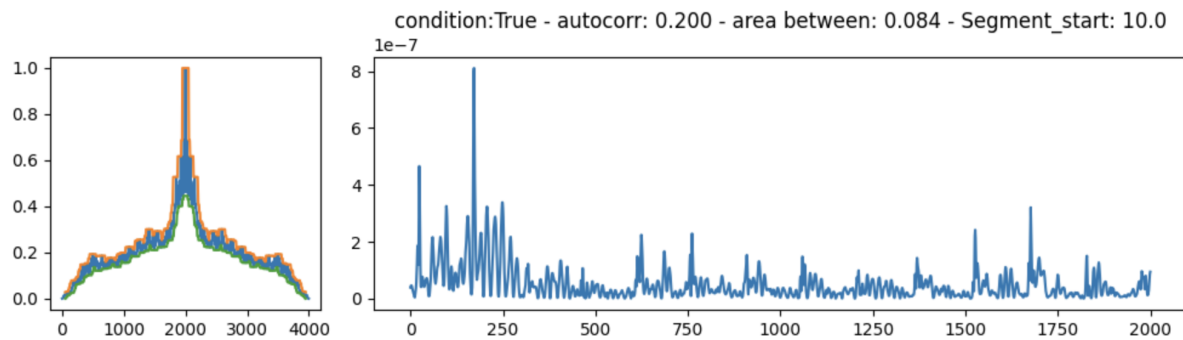condition:True - autocorr: 0.200 - area between: 0.084 - Segment_start: 10.0

Figure [Autocorrelations]. Three signal segments (right column) and their corresponding autocorrelations and min (green) - max envelopes (orange) in the left column.

Result Summary

The results of the model's performance are summarized.
Concerning the multi-headed (5 backbone) model, we validated the model via 5-fold cross validation, and the following mean metrics were obtained (on pooled segment performance):

- F1-score = 0.81 (0.12)
- Recall = 0.81 (0.15)
- Precision = 0.83 (0.09)
- AUC = 0.91 (0.04)

Note: Metrics are reported as mean (standard deviation)

Future Work

Identifying the contribution of each of the features in a model is an insightful task, in this case we are interested in identifying the contribution of not just the patient biometrics but also the device derived signals. In order to identify, we propose to use a variation of the permutation feature importance [ref-permutation-feature-importance] technique but modified for. Instead of considering each of the readings in a channel as a feature we consider an entire channel as a feature. Once the model is fitted we can estimate the importance of a given channel as the decrease in a model score when this channel readings are shuffled within patients in our dataset. Notice this breaks the relationship between the feature and the target, thus the drop in the model score is indicative of how much the model depends on the feature.

Final Remarks

The methodology used for data preparation, data splitting, and the model architecture seems to be promising. The segment discriminator technique, based on autocorrelation, was introduced to handle irregular signals, and it is model-agnostic.

In general, the performance of the model seems to improve when using the pool strategy. Note that the pool strategy only makes sense when the signal's information is taken into account, showing some evidence that the device-related information might enhance the performance of this type of model.

We hope the ideas discussed here are valuable for our partners' future research and prototyping. Overall, this project has the potential to significantly improve the early detection and treatment of heart failure.