Israel Jesus Santos Filho
Kaust-ID 179023

Exploration of Deep Learning regression capabilities for Diastolic and Systolic Blood Pressure Estimation (DBP and SBP) through ECG (electrocardiogram) and PPG (photoplethysmogram)

Introduction:

Many discoveries and applications have been explored through the combination of novel machine learning algorithms for solving healthcare problems. Nowadays, there is a lot of interest in developing machine learning models, and engineering systems, that could use biological signals to some applications: self-monitoring, pre-diagnosis and helping decision making for some physicians. Some of those signals are oxygen saturation, heart rate, respiratory rate, and systolic and diastolic blood pressure (SBP and DBP, respectively). Due to the success of deep learning the research community and practitioners have been trying to apply those models to solve estimation/regression problems for the biomarkers mentioned before.

This work aims to use two different biological signals, ECG and PPG respectively, through a subset of a public dataset, PULSE-DB. Through the replication of [1] we will explore ways of using deep learning algorithms for solving the estimation of two quantities, SBP and DBP, from latent representations of the input signals, ECG and PPG, that will be learned from two separate models: one model for the SBP and another model for the DBP. It is well known that deep learning algorithms are good for learning features from structured signal manifolds and as both signals, ECG and PPG, have a singular signature, aka their own structure, this information motivates the use of those machines to solve this estimation problem.

Methodology:

- Implementation of the model proposed by [1].
- Computing the metric on the training and test dataset.
- Network configuration:
    - Model follows the figure 2.
    - Optimizer Adam for both models.
    - Cost function: L1Loss function.
    - Learning rate: 1e-7.
- Dataset Description:
    - A subset of PULSE-DB dataset containing both training and test set specifically crafted for model evaluation.

- We are using the ECG and PPG waveforms, both sampled from 125 Hz and 10 seconds length.
- The labels are also provided as features from the dataset, in our case the labels are the systolic and diastolic blood pressures.
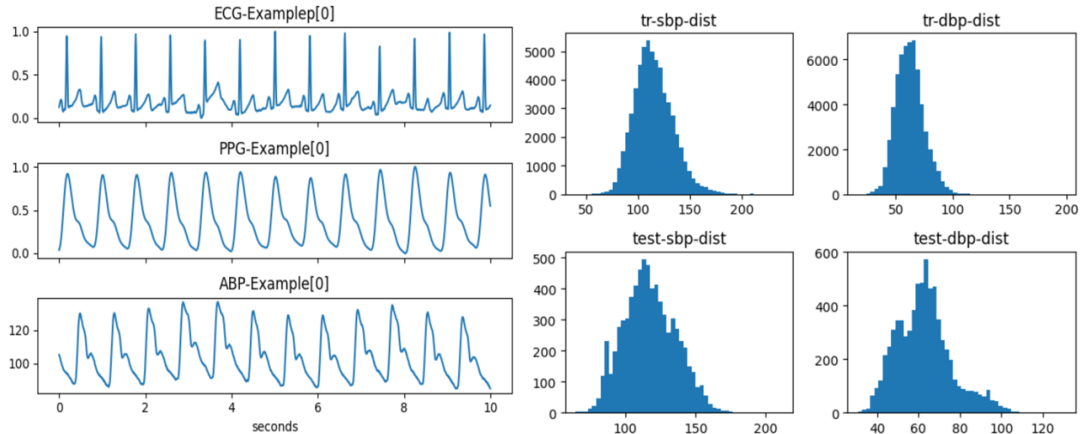

Figure 1 Left - example of the signals in the subset. Right, distribution of the labels for both training and test set (source, the author)
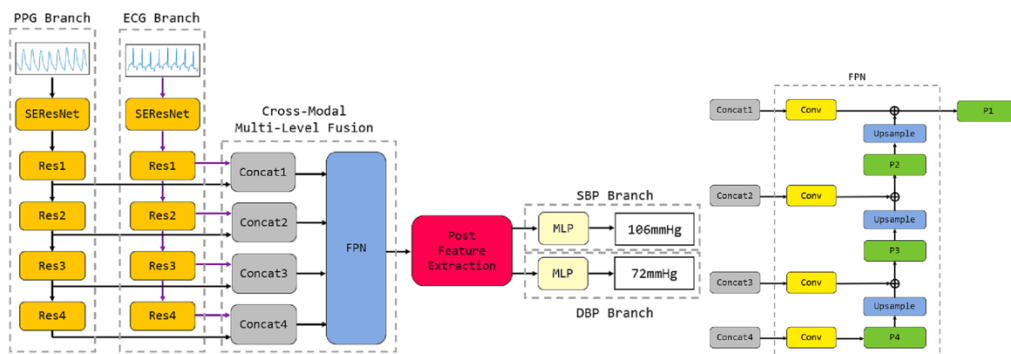

Figure 2 The model implemented. It is noteworthy that SEResnet stands for squeeze and excitation network[3] and FPN, feature pyramid networks[2] (source, reference 1)

Results and Discussion:

The development followed the one proposed in [1]. It is well known that deep learning models are well suited to unfold and figured out inner structure [1, 4] and the combination of different modalities, signals sources, have been used to enhance the final representation of the features used to solve the task at hand. It is also clear that multi modal deep learning is a trend [5, 6, 7] and has been applied to many problems, therefore it is natural to follow the trend with the hope that the representations of biomedical signals would benefit from this kind of approaches.

In the replication, it was detected an unnatural behavior associated with the prediction of only one quantity of interest, the **sbp**. It was found that the definition of their loss function was not so clear, which ended up making the model to cancel the influence of the DBP in the learning process.

$$Loss = \frac{1}{n}\sum_{i}^{n}\left[\left(SBP_{pi} - SBP_{ti}\right)^2 + corrcoef*\left(DBP_{pi} - DBP_{ti}\right)^2\right]$$

$$corrcoef = \frac{\left(SBP_{pi} - SBp_{ti}\right)^2}{\left(DBP_{pi} - DBP_{ti}\right)^2}$$

Motivated by this situation I changed, instead of learning minimizing the mean squared error, I decided to minimize the L1 error and to create a different model for each signal of interest. Both models followed the same structure pointed out in the figure 2 and the evaluation metric is the mean absolute error and the root mean squared error for both quantities.
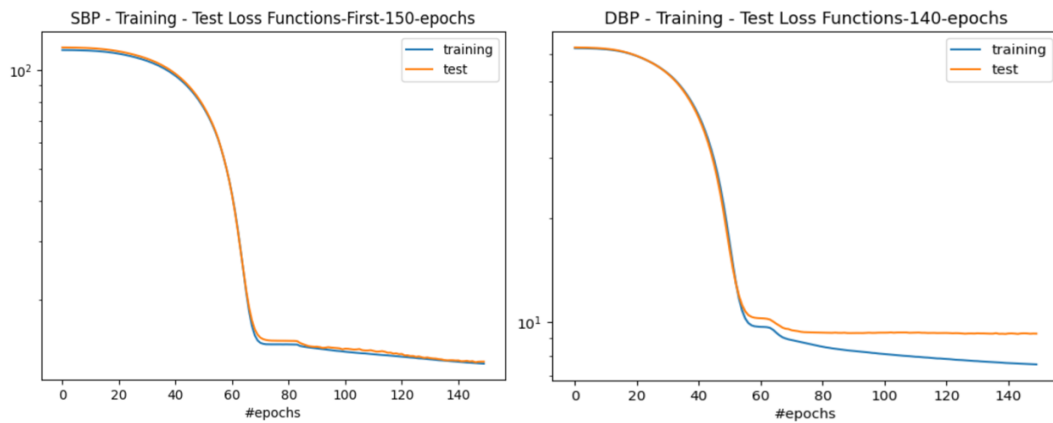

Figure 3 Left training and test loss for the sbp model, Right training and test loss for the dbp model.
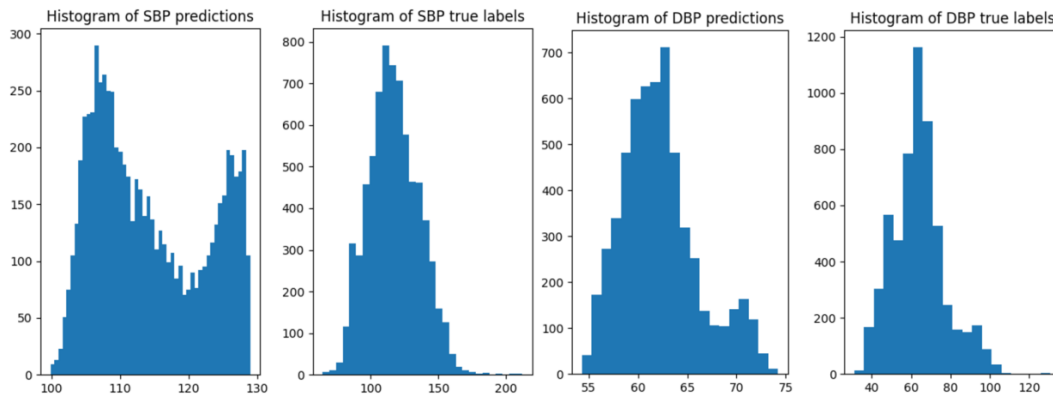

Figure 2 Left predicted sbp values from the model vs the true distribution of sbp labels. Right, predicted values of dbp from the model vs the true dbp labels.

By the end, achieved:
- DBP model:
    - DBP RMSE on test set: 11.8361
    - DBP MAE on test set: 9.2952
- SBP model:
    - SBP RMSE on test set: 16.2643
    - SBP MAE on test set: 12.9888

Conclusion:

The utilization of deep learning models for learning representations has been causing a lot of impacts in the scientific community, showing how much we have to explore in terms of research in order improve and further run those learned representations. The biomedical signal processing community is working on how to learn best ways of fusion and mix information from different bio signals components and how to get the most of those representations to build better models. I was not able to completely replicate the paper, but as we can see, from the predicted distributions of both models, we are able to capture some useful relationship (specially from the DBP).

This project aimed to develop a model that would learn jointly representations of different bio signals and use this representation to solve a regression problem. In this project we were able to apply deep modal fusion in order to estimate the systolic and diastolic blood pressure, which has shown to be a task not easy and that requires more knowledge about the signals in order to improve the estimation.

References:

[1] Long, W., & Wang, X. (2023). BPNet: A multi-modal fusion neural network for blood pressure estimation using ECG and PPG. In Biomedical Signal Processing and Control (Vol. 86, p. 105287). Elsevier BV. https://doi.org/10.1016/j.bspc.2023.105287

[2] T.Y. Lin, P. Dollar, ´ R. Girshick, et al. Feature pyramid networks for object detection, in: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2117-2125.

[3] Jie Hu, Li Shen, Samuel Albanie, Gang Sun, Enhua Wu: "Squeeze-and-Excitation Networks", 2017; <a href='http://arxiv.org/abs/1709.01507'>arXiv:1709.01507

[4] Y. Bengio, A. Courville and P. Vincent, "Representation Learning: A Review and New Perspectives," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798-1828, Aug. 2013, doi: 10.1109/TPAMI.2013.50.

[5] Li, J., Hong, D., Gao, L., Yao, J., Zheng, K., Zhang, B., & Chanussot, J. (2022). Deep Learning in Multimodal Remote Sensing Data Fusion: A Comprehensive Review (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2205.01380

[6] Ortega, J. D. S., Senoussaoui, M., Granger, E., Pedersoli, M., Cardinal, P., & Koerich, A. L. (2019). Multimodal Fusion with Deep Neural Networks for Audio-Video Emotion Recognition (Version 1). arXiv. https://doi.org/10.48550/ARXIV.1907.03196

[7] Akkus, C., Chu, L., Djakovic, V., Jauch-Walser, S., Koch, P., Loss, G., Marquardt, C., Moldovan, M., Sauter, N., Schneider, M., Schulte, R., Urbanczyk, K., Goschenhofer, J., Heumann, C., Hvingelby, R., Schalk, D., & Aßenmacher, M. (2023). Multimodal Deep Learning (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2301.04856