Cuff-Less and Calibration Free Blood Pressure Estimation Using the Pulse Transit Time Method

Daniel Tabatabai, ECE 699-002, Learning from Data Project

Abstract—Cuff-less blood pressure estimation using the pulse transit time method (PTT) has shown promise in the literature. Although PTT shows correlation to blood pressure, the method requires calibration for individual patients. A recent paper has proposed a cuff-less blood pressure estimation technique based on PTT that does not require calibration. This project attempts to replicate there results using support vector machines. Results from this study show consistency with the paper mentioned above.

I. INTRODUCTION

According to the United States Centers for Disease Control, about 70 million American adults (29%) have high blood pressure. Of that 29 percent, only half of them have their condition under control. Another one out of three Americans are pre-hypertensive. Furthermore, high blood pressure costs the nation $46 billion each year [1]. Non-invasive cuff-based blood pressure measurement methods are the most commonly used methods to date. This is usually done using a device called a sphygmomanometer. Although providing adequate accuracy, the cuff-based approach has a number of disadvantages. For example, continuous blood pressure monitoring is not possible since a pause of 1-2 minutes between measurements is required to avoid errors. In addition, the cuff itself would be a nuisance in daily life.

Since fluctuation in blood pressure can be dependent on food intake, activity level, and stress levels, continuous blood pressure monitoring can help to understand and prevent hazardous triggers and better individualize patient treatment. With the advancement of new wearable bio-sensors along with an increase in the development of phone applications for monitoring biological signals such as heart rate and blood oxygen saturation, cuff-less blood pressure monitoring could be achievable if algorithms were developed to take advantage of these new sensors. Methods of cuff-less blood pressure measurements have been attempted using the photo-plethysmograph (PPG) waveform generated from a pulse oximeter. One method in particular that has shown promise uses the pulse wave velocity from heart beats to estimate blood pressure [2]. The main disadvantage of this technique is that it requires individual patient calibration that has to be updated periodically [3].

One paper in particular attempts to eliminate the need for calibration while using techniques similar to the pulse wave velocity. This is done by extracting features from ECG and PPG waveforms and applying machine learning algorithms to estimate blood pressure [3]. This project attempts to replicate the methods and results described in [3].

II. RELATED WORK

The authors in [3] provide a method for estimating blood pressure by using features from ECG and PPG waveforms. The methodology is described as follows:

- Collect data with a sufficient sample size
- Preprocess the data to eliminate high frequency noise, as well as eliminating invalid signals
- Extract features from the ECG and PPG signals
- Partition the data into training, validation, and testing sub-sets
- Train the regression models with machine learning algorithms
- Perform model evaluation

A. Database

The authors used the Multi-parameter Intelligent Monitoring in Intensive Care (MIMIC) II online waveform database [4] provided by PhysioNet.org for their data set. The data set was reposted on the UCI database website and this project makes use of that data.

B. Preprocessing

The authors preprocess the data to eliminate the unusable signals and prepare for feature extraction. The steps are described as follows:

Step I: Smoothing all the signals with an averaging filter.
Step II: Removing signal blocks with irregular and unacceptable human blood pressure values.
Step III: Removing signal blocks with unacceptable heart rates.
Step IV: Removing signal blocks with severe discontinuities, which were not resolved with the help of the smoothing filter in step I.
Step V: Calculate the autocorrelation of the PPG signal, which indicates the degree of similarity between successive pulses, and removing blocks with high alteration.
C. Feature Extraction
For this project, the same features as in [3] were used. This will be discussed further in the methods section.

D. Partitioning
In [3] the data is partitioned into 60% training, 20% validation, and 20% testing. For this project the data was partitioned into 80% training and 20% testing. The training set was used along with cross validation.

E. Machine Learning
In [3], three approaches were attempted and the best performing model was selected. Regularized Linear Regression (RLR), Artificial Neural Networks (ANN), and Kernelized Support Vector Machines (SVM) were attempted. The SVM algorithm was found to outperform the other two. In this project the SVM algorithm was attempted and compare to the results in [3].

F. Model Evaluation
In [3], the Mean Absolute Error (MAE) and Standard Deviation (STD) of estimation errors were used for the model evaluation. Calculations were performed on the test data. The same criteria was used in this project and compared to the results in [3].

III. BACKGROUND
The main premise behind the method used in this project is that when the heart beats, a pressure pulse is produced that propagates through the arteries of the circulatory system. The velocity of the pressure pulse, called the pulse wave velocity (PWV), is related to the elasticity of the artery walls by

$$PWV = \sqrt{\frac{E}{2\rho}}$$  \hspace{1cm} (1)

where \( R \) is the inner radius of vessels, \( p \) represents the blood density, \( t \) is the vessel thickness and \( E \) is Young’s modulus, which is related to the vessels elasticity. For an elastic vessel, the relation between the blood pressure and \( E \) is given by:

$$E = E_0 e^{(p-p_0)}$$  \hspace{1cm} (2)

where \( P \) is the pressure in the arteries and \( E_0 \) and \( P_0 \) are constants. One method of calculating PWV is the pulse transit time (PTT). PTT is defined as the time it takes a pulse from the heart to propagate to the peripherals. The PWV can be estimated by measuring the distance \( d \) and calculating

$$PWV = \frac{d}{PTT}$$  \hspace{1cm} (3)

(Eq 3) implies that PWV is highly dependent on an individual’s height. In addition, from (Eq 1 and 2) PWM is dependent on the elasticity of the arteries, which becomes less elastic with age. Therefore, PWV is dependent on age. PTT can be estimated by measuring the time interval between the R-peak of the ECG signal and various points on the PPG signal.

IV. METHODS
For this project the data from the UCI machine learning repository was used [3]. According to the website the data has been preprocessed and cleaned. Therefore, the preprocessing step was skipped in this work. The following steps were performed.

A. Features
The features used for this project were the same as listed in [3]. The following features were extracted from the data:

1) Heart Rate:
The heart rate was calculated by measuring the time between the peaks of the ECG signals. The heart rate was also calculated by measuring the time interval between the systolic peaks of the PPG signals. To insure that the features were valid, the heart rates of the two signals were compared and if the difference between the two were more than 10 beats per minute apart from one another the record was deemed invalid and removed from the data set.

2) PTT:
PTT is calculated by measuring the time interval between the ECG R-peak and three points on the PPG waveform. The three points were PPG maximum peak (PPGp), PPG minimum (PPGm), and the point of maximum slope (PPGd).

3) PPG features:
To eliminate the need for calibration, features from the PPG waveform were used. These include:

- Augmentation Index (AI): The augmentation index is a measure of the wave reflection of the arteries. It is calculated as the ratio between the diastolic peak and the systolic peak.

$$AI = \frac{\text{diastolic peak}}{\text{systolic peak}}$$  \hspace{1cm} (4)

- Large Artery Stiffness Index (LASI): The LASI is a measure of the stiffness of the arteries. It is measure as the time interval between the systolic peak and the diastolic peak.

- Inflection Point Area ratio (IPA): The areas S1, S2, S3, and S4 were used in replace of the IPA. S1 is the area between the minimum point and the point of maximum slope. S2 is the area between the point of maximum slope and the systolic peak. S3 is the area between the systolic peak and the diastolic peak. S4 is the area between the diastolic peak and the next minimum point.

V. FEATURE EXTRACTION
The first step to extract features was to find the ECG R-peaks. The method to do this was as follows: Step 1: High pass filter the signal to eliminate low frequency variations. Step 2: Determine the maximum value of the signal. This was
done by finding the first 10 maximum samples and taking the median value. Step 3: use the MATLAB function ‘findpeak’ with a threshold of 80% of the maximum value.

The next step was to find the peaks in the PPG waveform. To find the systolic peak the following approach was used:
Step 1: Calculate the PPG first derivative.
Step 2: Find the zero crossing of the first derivative where it crosses from positive to negative and where the PPG signal is greater than 40% of the maximum value.

The next step is to find the point of minimum value. This is accomplished by finding the minimum value between systolic peaks. The point of maximum slope was found by finding the peaks in the first derivative. This was accomplished by finding the zero crossing points in the second derivative. The diastolic peak was found by the maximum points in the second derivative, selecting the second highest peak, and then finding the next minimum point. This was accomplished by finding the zero crossings third derivatives. (Fig 1 and Fig 2) show the PPG signal, first, and second derivatives with the points of interest shown.

The same process was performed on the blood pressure waveforms to extract the systolic and diastolic peaks. Once the R-peak of the ECG signal and the peaks of the PPG signal were found, the blocks were validated by ensuring that every each block had all the signals found, that the time intervals were within one standard deviation of the mean, and that the PPG heart rate and the ECG heart rate were within 10 beats per second of one another. If the criteria were not met, then that record was thrown out.

VI. MACHINE LEARNING
Support vector machines were used to train and predict systolic and diastolic blood pressure. The LIBSVM library was used to implement the SVM algorithms [5]. The epsilon-SVR with the radial basis kernel was used for regression. Cross validation was performed to optimize the epsilon and gamma parameters. Once the parameters were determined, training was performed using the “best fit” parameters.

VII. MODEL EVALUATION
The mean absolute error and the error standard deviation were used to evaluate the model performance. Both were calculated from the testing data set. Results were compared to the results from [3].

VIII. RESULTS

<table>
<thead>
<tr>
<th></th>
<th>SBP (mmHg)</th>
<th>DBP (mmHg)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAE</td>
<td>STD</td>
</tr>
<tr>
<td>My Results</td>
<td>12.818</td>
<td>16.6761</td>
</tr>
<tr>
<td>Their Results</td>
<td>12.38</td>
<td>16.17</td>
</tr>
</tbody>
</table>

Table 1: Comparison of mean absolute error and standard deviation for systolic and diastolic blood pressure.
Results show that systolic MAE is the same for both cases. The diastolic MAE was greater in this project. This may be because of bad performance for the diastolic feature extraction. Figure 3 shows the distribution of the systolic blood pressure. Figure 4 shows the distribution of the diastolic blood pressure. Figure 5 and Figure 6 show the histogram of the systolic and diastolic prediction error, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Systolic</th>
<th>Diastolic</th>
</tr>
</thead>
<tbody>
<tr>
<td>My Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>≥ 5 mmHg</td>
<td>81.8%</td>
</tr>
<tr>
<td></td>
<td>≥ 10 mmHg</td>
<td>90.3%</td>
</tr>
<tr>
<td></td>
<td>≥ 15 mmHg</td>
<td>93.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Their Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>51.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>78.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>93.6%</td>
</tr>
<tr>
<td>Systolic</td>
<td></td>
<td>28.8%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>51.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>69.5%</td>
</tr>
</tbody>
</table>

Table 2: Comparison of percentage of predictions outside a set range.
IX. CONCLUSION

The results discussed above show that cuff-less blood pressure estimation is possible and that results can be repeated. Since blood pressure is dependent on age and height having those feature could potentially improve the results. There was a lot of time dedicated to feature extraction. Future work could include attempting automatic feature extraction with some unsupervised learning technique. Furthermore, many records were discarded because of invalid features and signals, but there is no evidence in this work that the invalid records do not have useful information, especially if the event is seen in multiple channels. Future work could focus on learning if signals have artifacts or are actual features.

REFERENCES


