

Non-Invasive Blood Pressure Estimation from Photoplethysmography Signals using Artificial Neural Networks

Nicolas Bersano

Allm Inc.

Email: n.bersano@allm.net

Horacio Sanson

Allm Inc.

Email: horacio@allm.net

Abstract—This work focuses on the potential of artificial neural networks to classify biological signals in a healthcare setting, specifically in the estimation of blood pressure from photoplethysmography signal readings obtained via medical devices. This signal is known to have valuable cardiovascular information and has been related to heart rate and blood pressure pulsewave. Among the literature there have been attempts to correlate this signal directly to a single blood pressure value and/or classify it into one of the blood pressure clinical states (e.g. Hypotension, Normal, Pre Hypertension, Stage 1 Hypertension, Stage 2 Hypertension). We propose models based on artificial neural networks that achieve similar performance to those in previous works, without needing engineered nor demographic features. These models are capable of learning how to extract descriptive features from only the raw photoplethysmography signals, and use them for classification into a blood pressure class. Test results are promising and validate the usefulness of artificial neural network architectures for this task.

I. INTRODUCTION

Continuous monitoring of blood pressure in a non-invasive and portable way has been a hot topic of research due to its application in early detection and monitoring of hypertension, a major cardiovascular risk.

The photoplethysmogram (PPG) signal is easy to obtain from any tissue (e.g. skin) by using a simple light emitter and photodetector pair. This signal can be used to measure blood volume changes [1] and has many clinical applications such as the measurement of heart rate and oxygen saturation. Studies have found that the PPG's temporal, spatial and frequency features provide valuable cardiovascular information [1]–[9] and it has also been related to pulse and blood pressure (BP) pulsewave [10].

Initial attempts to estimate systolic and diastolic blood pressure from PPG signals concentrate on feature engineering from the signal [11]–[14] and then applying regression methods [15], support vector machines [11] or neural networks [16] with the engineered features as input and the estimated blood pressure as output.

On the other hand, deep neural networks have been studied for time series classification tasks. In [17] the authors apply a multi-task learning method using a recurrent neural network for time series classification tasks. Authors in [18] evaluate Long Short-Term Memory networks for multi-label classification of clinical signals into one or more diagnosis labels.

One of the benefits of deep neural networks for the classification of time series is their ability to perform perception tasks [19]. This allows these networks to learn by themselves the relevant features of the input required to perform the task at hand (e.g. classification), thus removing the need for complex feature engineering [20].

In this work, we first formulate the problem of estimating systolic and diastolic blood pressure from PPG signals as one of multi-class classification and evaluate three different neural network models performing the classification task in terms of accuracy and absolute error. Also, comparison with previous work using linear regression, support vector machines and shallow neural networks is presented.

II. RELATED WORK

Visvanathan et. al. [11] use linear regression and support vector machines to classify each pulse in a PPG signal into one of six classes shown in Table I. For each pulse they extract fourteen time domain features shown in Table II.

TABLE I
BLOOD PRESSURE CLASSES

BP Class	Systolic	Diastolic
Hypotension	< 90	< 60
Normal	90 – 119	60 – 79
Prehypertension	120 – 139	80 – 89
Stage 1 Hypertension	140 – 159	90 – 99
Stage 2 Hypertension	160 – 179	100 – 109
Hypertensive Emergency	≥ 180	≥ 110

They report accuracies between 80% and 87% for both systolic and diastolic blood pressure classification using only the extracted time domain features and 100% accuracy for systolic blood pressure when using SVM with additional demographic features such as height, weight and age. Given the small sample size we are inclined to believe the impressive results of SVM are due to overfitting on the demographic features.

The same authors further improve their results by adding additional frequency domain features of the PPG [12]. To reduce processing complexity, they apply the maximal information coefficient (MIC) in order to find and select the most

TABLE II
VISVANATHAN ET AL. FEATURES

Feature	Name	Description
Avn	Valley amplitude	Value of PPG signal at the valley pre systolic phase.
Asn	Systolic peak amplitude	Maximum value of PPG during systolic phase.
Adn	Dicrotic notch amplitude	Value of PPG signal on the dicrotic notch.
S_{area}	Systolic area	Area under the curve of the PPG signal between the valley and the dicrotic notch.
D_{area}	Dicrotic notch area	Area under the curve of the PPG signal between the dicrotic notch and the next valley.
T_{area}	Total area	Systolic area plus dicrotic notch area.
R_{ratio}	Ratio of area	Dicrotic notch area divided by systolic area.
$Pk_{interval}$	Peak interval	Time between systolic peaks of two consecutive pulses.
Pl_{height}	Pulse height	Difference between systolic peak amplitude and valley amplitude.
$Pl_{interval}$	Pulse interval	Time between pre systolic valleys of two consecutive pulses.
C_t	Crest time	Time difference between the systolic peak and the pre systolic valley.
D_t	Delta time	Time difference between dicrotic notch and systolic peak.
A_{index}	Augmentation index	$(Adn - Avn)/(Asn - Avn)$
R_{index}	Reflection index	$1 - A_{index}$

relevant features for blood pressure estimation. They find that frequency features play a more important role than temporal features for BP classification of PPG signals.

Ding et al. [14] propose a new feature: Photoplethysmogram Intensity Ratio (PIR) that they show is highly correlated with arterial diameter variation and mean blood pressure. The same authors then developed a linear model that allows them to obtain blood pressure measurements from the proposed PIR value [15].

The above methods work under the assumption that there is a linear relation between the input features and the blood pressure. The work by Kurylyak et. al. [16], proposes to use a feed forward neural network to perform regression on the input features. They feed the network input layer with 21 PPG morphologic features (Table III) proposed by Linder et. al. [13] and at the output layer they obtain two estimated values for systolic and diastolic BP respectively. The results are reported in terms of error rates: 3.80 ± 3.46 mmHg for systolic and 2.21 ± 2.09 mmHg for diastolic BP that are within the accepted maximal errors of 5 ± 8 mmHg set by the Association for the Advancement of Medical Instrumentation (AAMI).

TABLE III
KURYLYAK ET AL. FEATURES

Feature	Description
CT	The time difference between two consecutive systolic peaks.
SUT	The time interval between the onset time and the systolic peak time.
DT	The time interval between the next pulse's onset time and the current pulse's systolic time.
$DW10\%$	Diastolic width at 10% pulse max height.
$DW10\% + SW10\%$	Diastolic width and systolic width sum at 10% pulse max height.
$DW10\%/SW10\%$	Diastolic width to systolic width ratio at 10% pulse max height.
$DW25\%$	Diastolic width at 25% pulse max height.
$DW25\% + SW25\%$	Diastolic width and systolic width sum at 25% pulse max height.
$DW25\%/SW25\%$	Diastolic width to systolic width ratio at 25% pulse max height.
$DW33\%$	Diastolic width at 33% pulse max height.
$DW33\% + SW33\%$	Diastolic width and systolic width sum at 33% pulse max height.
$DW33\%/SW33\%$	Diastolic width to systolic width ratio at 33% pulse max height.
$DW50\%$	Diastolic width at 50% pulse max height.
$DW50\% + SW50\%$	Diastolic width and systolic width sum at 50% pulse max height.
$DW50\%/SW50\%$	Diastolic width to systolic width ratio at 50% pulse max height.
$DW66\%$	Diastolic width at 66% pulse max height.
$DW66\% + SW66\%$	Diastolic width and systolic width sum at 66% pulse max height.
$DW66\%/SW66\%$	Diastolic width to systolic width ratio at 66% pulse max height.
$DW75\%$	Diastolic width at 75% pulse max height.
$DW75\% + SW75\%$	Diastolic width and systolic width sum at 75% pulse max height.
$DW75\%/SW75\%$	Diastolic width to systolic width ratio at 75% pulse max height.

III. MATERIALS AND METHODS

A. Problem Definition

Similar to Visvanathan et. al. [11], [12] we define the problem of blood pressure estimation from PPG signals as one of time series classification, where the input are features extracted from PPG pulses and the output is one of l defined classes or bins, where each bin corresponds to a range of blood pressure. We use four bin sizes: 1, 3, 5 and 10 as shown in Table IV.

TABLE IV
OUTPUT BIN SIZES AND RANGES

Bin Size	BP Ranges (mmHg)
1	60, 61, 62, 63, ..., 200
3	60 - 62, 63 - 65, 66 - 68, ..., 199 - 201
5	60 - 64, 65 - 69, 70 - 74, ..., 195 - 199
10	60 - 69, 70 - 79, 80 - 89, ..., 190 - 199

B. Data Set

We obtained PPG signals from 52 healthy people with age ranges between 25 and 60, male and female. The signals were obtained using a Nonin Xpod pulse oximeter with a 8000AA articulated finger clip [21] for a duration of 60 seconds, while taking measurements of diastolic and systolic blood pressure before and after capturing the PPG signal.

Each PPG signal was segmented in pulses using a thresholding algorithm proposed by Zong et al. [22] and the same time domain features proposed by Visvanathan [11] and Kurylyak [16] were extracted from each one. This resulted in a set of 8135 pulse features.

The set of 8135 pulse features was then split into training, test and validation sets of 5694, 1220 and 1221 pulse features, respectively. We ensured that no single pulse belongs to more than one set.

C. Softmax Regression

Our first experiment was replication of the work by Visvanathan et al. [11] using a simple softmax regression model. We performed the experiments with the same fourteen time domain features (Table II) as input and each bin size (Table IV) as output. The same regression model was trained independently for systolic and diastolic BP inference.

In Table V the column *14 features* shows the accuracy and absolute errors of the validation set for both diastolic and systolic BP. We can see that softmax regression using only the time domain features results in absolute errors beyond the maximum accepted by the AAMI. It is important to note that the error values are in bin size units. To obtain the error in *mmHg* units it must be multiplied by the bin size. For example, an error of 1.0 ± 2.0 becomes 1.0 ± 2.0 mmHg for a bin size of 1 but becomes 10.0 ± 20.0 mmHg for a bin size of 10.

In the next experiment, we reduced the number of features used in training and inference by removing redundant features. For example, features that are measured from the same morphological part of the PPG wave, and features that are computed as some combination of the other features. After this feature reduction we ended with nine out of the fourteen features, namely: *Avn*, *Asn*, *Adn*, *Rration*, *Pk_{interval}*, *Pl_{height}*, *Pl_{interval}*, *C_t*, and *D_t*.

Table V shows the results for the experiment using nine features under the column *9 features*. We can see that there is almost no significant change in the results when using fourteen or nine features.

Finally, we added demographic features such as age, height, weight, and gender to the softmax input combined with the other fourteen and nine features used before. The results are presented in the same Table V under the columns *14 features + demographics* and *9 features + demographics*. We can see the accuracy is now much better and the error is within the limits of the AAMI for bin sizes 1 and 3 for both diastolic and systolic blood pressure.

Given the low number of subjects, and therefore the low distribution of demographic features; we are inclined to believe

that the softmax regression is inferring the blood pressure classes based mostly in the demographic features, giving less or no weight to the features extracted from the PPG. This is true for our own data set and the data set used by the original authors, which contained only 17 subjects with demographic features.

D. Support Vector Machine

We repeated the same experiments described by Visvanathan et al. [11] using a support vector machine (SVM) instead of softmax regression. Table VI shows the results of these experiments.

In Table VI we can see that without using demographic features, SVM already has good results in terms of accuracy. This is attributed to the ability of SVM to model non-linearities using different kernels, such as a Gaussian radial basis function for example. On the other hand, the absolute error is high for all bin sizes, with only the systolic case with bin size 1 being within the accepted max error value set by the AAMI.

When demographic features are factored in the input we see further improvements in accuracy, jumping from 86% to 98% and from 87% to 99% accuracy for diastolic and systolic blood pressure, respectively, with a bin size of 10. Similar improvements can be observed on all bin sizes. More importantly, we can observe a reduction in the absolute error to within the limits of the AAMI for all bin sizes for both diastolic and systolic blood pressures.

Finally, when reducing the number of PPG features used in the input from fourteen to nine we see a slight improvement in accuracy and absolute error on most cases.

E. Artificial Neural Network (ANN)

We implemented the same feed-forward back propagation ANN with *N* inputs and two outputs as proposed in [16], in order to simultaneously estimate Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP). The number of inputs *N* depends on the features from Table III. We performed experiments using two different two-hidden-layer ANNs with 35:20 neurons and 50:30 neurons. That is, one network with 35 neurons in the first hidden layer and 20 on the second, and another network with 50 neurons in the first hidden layer and 30 on the second. The experiments were performed varying the input layer size to accept different features for a total of 8 different experiments:

- ANN 35:20 with all 21 features of Table III as input.
- ANN 50:30 with all 21 features of Table III as input.
- ANN 35:20 with only 9 features: CT, SUT, DT, and the six DW/ST ratios.
- ANN 50:30 with only 9 features: CT, SUT, DT, and the six DW/ST ratios.
- ANN 35:20 with all 21 features plus demographic features (age, gender, height, weight).
- ANN 50:30 with all 21 features plus demographic features (age, gender, height, weight).
- ANN 35:20 with only 9 features plus demographic features (age, gender, height, weight).

TABLE V
SOFTMAX REGRESSION RESULTS (VALIDATION SET).

Systolic Blood Pressure								
Model	14 features		9 features		14 features + demographics		9 features + demographics	
Bin Size	Acc.	Error	Acc.	Error	Acc.	Error	Acc.	Error
1	58%	6.6 ± 11.6	57%	7.2 ± 12.8	86%	1.3 ± 4.9	84%	1.5 ± 4.9
3	46%	3.3 ± 4.4	44%	3.6 ± 4.7	72%	1.0 ± 1.9	69%	1.1 ± 2.0
5	52%	2.1 ± 2.9	49%	2.2 ± 3.0	73%	0.6 ± 1.2	72%	0.7 ± 1.4
10	54%	1.1 ± 1.5	56%	1.0 ± 1.5	78%	0.3 ± 0.7	75%	0.4 ± 0.7
Diastolic Blood Pressure								
Model	14 features		9 features		14 features + demographics		9 features + demographics	
Bin Size	Acc.	Error	Acc.	Error	Acc.	Error	Acc.	Error
1	48%	4.6 ± 7.2	47%	4.9 ± 7.4	83%	1.0 ± 2.8	78%	1.4 ± 3.5
3	41%	1.9 ± 2.6	40%	2.0 ± 2.6	69%	0.7 ± 1.2	71%	0.7 ± 1.3
5	50%	1.2 ± 1.6	47%	1.2 ± 1.6	74%	0.4 ± 0.9	72%	0.5 ± 0.9
10	49%	0.7 ± 0.9	45%	0.8 ± 0.9	72%	0.3 ± 0.5	68%	0.3 ± 0.5

TABLE VI
SVM RESULTS (VALIDATION SET).

Systolic Blood Pressure								
Model	14 features		9 features		14 features + demographics		9 features + demographics	
Bin Size	Acc.	Error	Acc.	Error	Acc.	Error	Acc.	Error
1	86%	2.3 ± 7.6	81%	3.0 ± 8.8	96.8%	1.2 ± 1.3	98.5%	0.04 ± 0.42
3	85%	0.8 ± 2.5	84%	0.9 ± 2.7	99.2%	0.01 ± 0.13	98.9%	0.02 ± 0.34
5	86%	0.6 ± 1.7	81%	0.8 ± 2.0	99.2%	0.02 ± 0.22	99.9%	0.005 ± 0.17
10	87%	0.3 ± 0.9	86%	0.3 ± 0.9	99.6%	0.005 ± 0.08	99.4%	0.006 ± 0.09
Diastolic Blood Pressure								
Model	14 features		9 features		14 features + demographics		9 features + demographics	
Bin Size	Acc.	Error	Acc.	Error	Acc.	Error	Acc.	Error
1	83%	1.6 ± 4.8	84%	1.4 ± 4.1	96.4%	0.17 ± 0.98	98.7%	0.08 ± 0.7
3	84%	0.5 ± 1.6	85%	0.5 ± 1.7	97.1%	0.06 ± 0.36	96.8%	0.06 ± 0.37
5	85%	0.3 ± 1.0	81%	0.4 ± 1.2	97.5%	0.04 ± 0.3	99.0%	0.02 ± 0.17
10	86%	0.2 ± 0.6	85%	0.2 ± 0.6	98.1%	0.02 ± 0.16	97.1%	0.03 ± 0.17

- ANN 50:30 with only 9 features plus demographic features (age, gender, height, weight).

Table VII shows the absolute errors for each of the experiments performed¹. The errors we see are in *mmHg* and we can observe that using only the PPG pulse features as input results in errors higher than SVM approach. When demographics are added to the input, the error decreases considerably to the point of being within the AAMI requirements. We can see a small improvement when using 21 features instead of 9 when no demographics are considered, and no significant difference between the number of PPG pulse features when demographics are considered. This further increases our belief that these models are learning to estimate BP from demographic data alone, giving little or no weight to the PPG pulse features. Finally, the size of the hidden layers has little impact on the error results.

¹The experiment reports only absolute errors as it does not perform a classification task, the output values are estimates values for SBP and DBP in *mmHg* units.

F. Fully Connected Network

Two fully connected networks (FCN) with a softmax output layer were implemented. The first network had three hidden layers with eight neurons with ReLU activations each and the second had six hidden layers with eight neurons with ReLU activations each.

The input layer for both networks was configured with four PPG pulse features (see list below) and four optional demographic features: age, weight, gender and height. We selected the features that related most with the physiological inner workings of the cardiac cycle:

- *PIR*: Photoplethysmogram Intensity Ratio.
- *R_{ratio}*: Ratio between systolic and diastolic area (Table II).
- *Pl_{interval}*: Pulse interval (Table II).
- *Pk_{interval}*: Peak interval (Table II).

The softmax output layer has as many output nodes as bins defined in Table IV. This layer performs classification of the input features into one of the BP bins. We implemented and trained two versions of each network, one for SBP and another

TABLE VII
ANN ABSOLUTE ERRORS IN MMHG (VALIDATION SET).

	No Demographics				Demographics			
Inputs	21 Features		9 Features		21 Features		9 Features	
Model	ANN 35:20	ANN 50:30	ANN 35:20	ANN 50:30	ANN 35:20	ANN 50:30	ANN 35:20	ANN 50:30
SBP	13.30 ± 9.41	12.71 ± 9.22	13.02 ± 9.64	12.66 ± 9.75	2.20 ± 2.60	2.01 ± 2.25	1.68 ± 1.75	1.33 ± 1.67
DBP	6.63 ± 6.22	6.27 ± 6.12	7.69 ± 6.98	6.94 ± 6.64	1.45 ± 1.42	1.54 ± 1.57	1.60 ± 1.53	1.72 ± 1.63

for DBP for a total of 32 different FCNs.

Table VIII lists the accuracy and absolute errors for the FC experiments. We get worse performance when compared to the results obtained by using a simple softmax regression (Section III-C) with no demographic features used in the input. When demographic features are added to the input, we obtain a significant improvement in accuracy and error values, better performance than simple softmax regression with demographic inputs and results on par with SVM (Section III-D).

G. Convolutional Neural Network

In previous sections we implemented and evaluated models for BP estimation from PPG pulses that required engineered features as input. Unfortunately, in practice it is often not feasible to reliably extract the required features from the pulses due to a number of factors [23]. Morphological variations between subjects in addition to noise, make it difficult to extract the pulses themselves from the raw PPG signal and further features from those pulses.

Convolutional neural networks (CNNs) have shown to be good at extracting features from images for a number of different image processing tasks, including classification [24], [25]. Building on these results we implemented a CNN on top of the FCN with softmax output. The idea is that the CNN should learn to extract relevant features from the raw PPG input signal that can then be passed as input to a FCN with a softmax output layer for BP classification. The architecture of this network can be seen in Table IX.

The input layer has N inputs where N is the size of a pulse extracted from the raw PPG signal using the algorithm proposed by Zong et al. [22]. During scanning of the PPG signal, each data point $P[i]$ in the pulse P is standardized using Equation 1.

$$SP[i] = \begin{cases} \frac{P[i] - \text{avg}(P)}{\max(P) - \min(P)}, & \text{if } P[i] \leq \text{avg}(P) \\ \frac{P[i]}{\max(P) - \min(P)}, & \text{otherwise} \end{cases} \quad (1)$$

After the pulse is standardized it is zero padded if the length of the data is less than the maximum N in the set of pulses.

The convolutional part of the network is composed of four convolutional layers with ReLU activations with max pooling layers in between. Each max pooling layer has a kernel size of 2 and stride size of 2. These pooling layer settings have

the effect of reducing the dimension of the output of each convolution by half.

After the convolutional part of the network, we added two fully connected layers, also with ReLU activations, similar to the FCN described in Section III-F. Finally, the last layer is a softmax layer with l outputs that correspond to the bins for each bin size as described in Table IV.

We build two identical CNNs for DBP and SBP estimation and trained them using batch sizes of 25 and 50. Table X shows the accuracy and absolute error values observed when estimating both SBP and DBP with the trained CNNs. At first, we can observe that the accuracies are slightly better than those of SVM without demographic inputs while, on the other hand, the absolute error values are worse.

It is important to remark that the CNN is learning the features to perform classification by itself, thus completely removing the need for complex feature engineering and heavy pre-processing of the PPG signal. This alone overweights the higher error values obtained.

H. Convolutional-Recurrent Neural Network

We build two identical networks, one for DBP and other for SBP estimation, made from a CNN, FCN, and Recurrent Neural Network (RNN) blocks one after another. The CNN architecture is the same used for our CNN experiments, described in Section III-G. The idea behind this is letting the CNN [24], [25] learn how to extract descriptive features along the signal. The output of the convolutional network is passed through a final average pooling layer, averaging all the outputs into only one array, which is fed as input to a three layered FCN block, the first two layers with a ReLU activation and with the last one being a sigmoid activation. These FCN layers apply an expansion and a bottleneck to surface more descriptive features obtained by the CNN. Finally, the output of the FCN is fed as input to a RNN block with 24 LSTM hidden cells, the output layer being a softmax layer.

As the CNN block will learn and extract features along the BP pulse, these features will be enhanced by the FCN block, the RNN block will try to learn a pattern in the sequence of features to finally classify them into one of l outputs that correspond to the bins for each bin size as described in Table IV.

It must be noted that the variables such as kernel size for each CNN layer, output size for each FCN layer and size of the RNN, were selected taking into account the length of the size of the input layer, and they can be changed depending on the sampling rate of the signal. Preliminary results presented

TABLE VIII
FC RESULTS (VALIDATION SET).

Systolic Blood Pressure								
Model	5 Layer Network				8 Layer Network			
Input	PPG Features		PPG + Demographic Features		PPG Features		PPG + Demographic Features	
Bin Size	Acc.	Error	Acc.	Error	Acc.	Error	Acc.	Error
1	43.24%	11.08 ± 14.25	88.70%	1.04 ± 3.98	38.49%	12.97 ± 16.35	92.55%	0.35 ± 1.73
3	50.61%	2.88 ± 4.41	92.96%	0.15 ± 0.68	44.06%	4.16 ± 5.56	85.75%	0.26 ± 0.80
5	50.94%	1.98 ± 2.70	99.67%	0.005 ± 0.09	51.35%	2.14 ± 3.04	93.45%	0.13 ± 0.62
10	59.21%	0.90 ± 1.40	93.53%	0.08 ± 0.32	44.72%	1.30 ± 1.53	98.85%	0.013 ± 0.13
Diastolic Blood Pressure								
Model	5 Layer Network				8 Layer Network			
Input	PPG Features		PPG + Demographic Features		PPG Features		PPG + Demographic Features	
Bin Size	Acc.	Error	Acc.	Error	Acc.	Error	Acc.	Error
1	47.6%	5.04 ± 7.90	96.97%	0.13 ± 0.83	45.62%	5.85 ± 8.68	92.87%	0.36 ± 1.45
3	49.0%	1.64 ± 2.50	94.59%	0.13 ± 0.64	22.28%	2.58 ± 2.78	94.76%	0.09 ± 0.44
5	51.1%	1.23 ± 1.72	91.48%	0.13 ± 0.46	55.93%	1.17 ± 1.80	94.18%	0.09 ± 0.41
10	60.0%	0.57 ± 0.86	95.17%	0.05 ± 0.22	60.60%	0.56 ± 0.86	96.64%	0.04 ± 0.20

TABLE IX
CNN ARCHITECTURE

Layer	Kernel shape	Output size	Activation
<i>Conv</i> ₁	[1, 10]	25	ReLU
<i>Conv</i> ₂	[1, 10]	50	ReLU
<i>Conv</i> ₃	[1, 10]	100	ReLU
<i>Conv</i> ₄	[1, 10]	100	ReLU

TABLE X
CNN RESULTS (VALIDATION SET).

Batch Size	25				50			
Type	SBP		DBP		SBP		DBP	
Bin Size	Acc.	Error	Acc.	Error	Acc.	Error	Acc.	Error
1	89.19%	1.83 ± 7.20	90.50%	0.79 ± 3.40	88.94%	1.83 ± 6.89	89.93%	0.97 ± 3.86
3	80.10%	5.73 ± 15.55	87.39%	1.15 ± 5.28	89.52%	0.58 ± 2.27	87.71%	0.87 ± 4.27
5	90.90%	0.33 ± 1.42	88.53%	0.59 ± 2.77	88.12%	0.81 ± 3.60	88.79%	0.40 ± 1.92
10	86.40%	0.88 ± 3.35	86.73%	0.77 ± 7.72	89.43%	0.24 ± 0.87	90.17%	0.26 ± 1.25

in Table XI shows a slight decrease in performance if we consider the accuracy of the classification made, but for most of the cases, the variance on the error of the prediction is significantly lower which is very promising.

IV. CONCLUSIONS AND REMARKS

Through a series of empirical experiments we have evaluated the possibility of estimating blood pressure from raw PPG signals using Neural Networks (NNs). To do this we formulated the problem of blood pressure estimation as one of classification of PPG signal pulse features, or PPG signal pulses in the case of CNNs. We found in our preliminary results that NNs, such as CNNs and CNN-RNNs, are capable of estimating SBP and DBP with good accuracy and low error, with the CNN-RNN results having less error variance but slightly losing accuracy. Using the raw PPG as input, these

results fall within the guidelines proposed by the AAMI for DBP estimation, while the performance for SBP estimation falls slightly short from these guidelines. Even though they do not reach the same level of accuracy and error of SVM, please note that they achieve this performance without requiring complex feature engineering, PPG signal pre-processing, or additional input features (e.g. demographics). This makes CNN solutions more attractive for real scenarios where such tasks are unfeasible or difficult to perform. As future work, we are expanding our data set, with both synthetic and real data. The authors envision a future where wearable devices become ubiquitous and continuous measurement of vital signals such as PPG and ECG becomes commonplace. Thus generating large amounts of vital sign data that can then be used to train deeper NNs, which could better generalize the classification of SBP and DBP from vital signals.

TABLE XI
CONVOLUTIONAL-RECURRENT NN. RESULTS (VALIDATION SET).

Type	SBP		DBP	
Bin Size	Acc.	Error	Acc.	Error
1	84.26%	2.56 ± 8.41	87.38%	1.08 ± 3.90
3	85.73%	0.82 ± 2.74	84.18%	0.47 ± 1.43
5	84.83%	0.60 ± 1.88	87.21%	0.27 ± 0.88
10	83.36%	0.36 ± 0.99	91.23%	0.12 ± 0.45

REFERENCES

- [1] J. Allen, "Photoplethysmography and its Application in Clinical Physiological Measurement," *Physiological Measurement*, no. 3, p. R1.
- [2] L. Wang, E. Pickwell-MacPherson, Y. P. Liang, and Y. T. Zhang, "Noninvasive Cardiac Output Estimation Using a Novel Photoplethysmogram Index," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Sept 2009, pp. 1746–1749.
- [3] S.-H. Liu, K.-M. Chang, and T.-H. Fu, "Heart Rate Extraction from Photoplethysmogram on Fuzzy Logic Discriminator," *Eng. Appl. Artif. Intell.*, no. 6, pp. 968–977, Sep.
- [4] K. Chellappan, E. Zahedi, and M. Ali, *Effects of Physical Exercise on the Photoplethysmogram Waveform*, 2007.
- [5] K. Takazawa, N. Tanaka, M. Fujita, O. Matsuoka, T. Saiki, M. Aikawa, S. Tamura, and C. Ibukiyama, "Assessment of Vasocative Agents and Vascular Aging by the Second Derivative of Photoplethysmogram Waveform," *Hypertension*, vol. 32, no. 2, pp. 365–370, 1998.
- [6] J. M. Padilla, E. J. Berjano, J. Saiz, L. Facila, P. Diaz, and S. Merce, "Assessment of Relationships Between Blood Pressure, Pulse Wave Velocity and Digital Volume Pulse," in *2006 Computers in Cardiology*, Sept 2006, pp. 893–896.
- [7] U. Rubins, A. Grabovskis, J. Grube, and I. Kukulis, *Photoplethysmography Analysis of Artery Properties in Patients with Cardiovascular Diseases*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 319–322.
- [8] R. K. J. R. S.C. MILLASSEAU and P. CHOWIENCZYK, "Determination of Age-related Increases in Large Artery Stiffness by Digital Pulse Contour Analysis," *Clinical Science*, vol. 103, no. 4, pp. 371–377, 2002.
- [9] S. R. Alty, N. Angarita-Jaimes, S. C. Millasseau, and P. J. Chowienczyk, "Predicting Arterial Stiffness from the Digital Volume Pulse Waveform," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 12, pp. 2268–2275, Dec 2007.
- [10] S. C. Millasseau, F. G. Guigui, R. P. Kelly, K. Prasad, J. R. Cockcroft, J. M. Ritter, and P. J. Chowienczyk, "Noninvasive Assessment of the Digital Volume Pulse," *Hypertension*, no. 6, pp. 952–956.
- [11] A. Visvanathan, A. Sinha, and A. Pal, "Estimation of Blood Pressure Levels from Reflective Photoplethysmography Using Smart Phones," in *Bioinformatics and Bioengineering (BIBE), 2013 IEEE 13th International Conference on*, Nov 2013, pp. 1–5.
- [12] A. Visvanathan, R. Banerjee, A. Dutta Choudhury, A. Sinha, and S. Kundu, "Smart Phone Based Blood Pressure Indicator," pp. 19–24.
- [13] S. P. Linder, S. M. Wendelken, E. Wei, and S. P. McGrath, "Using The Morphology of Photoplethysmogram Peaks to Detect Changes in Posture," *Journal of Clinical Monitoring and Computing*, no. 3, pp. 151–158.
- [14] X. R. Ding and Y. T. Zhang, "Photoplethysmogram Intensity Ratio: A Potential Indicator for Improving the Accuracy of PTT-based Cuffless Blood Pressure Estimation," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2015, pp. 398–401.
- [15] X. R. Ding, Y. T. Zhang, and H. K. Tsang, "A New Modeling Methodology for Continuous Cuffless Blood Pressure Measurement," pp. 264–267, Feb 2016.
- [16] Y. Kurylyak, F. Lamomaca, and D. Grimaldi, "A Neural Network-based Method for Continuous Blood Pressure Estimation from a PPG Signal," pp. 280–283, May 2013.
- [17] M. Hüsken and P. Stagge, "Recurrent Neural Networks for Time Series Classification," *Neurocomputing*, pp. 223–235.
- [18] Z. C. Lipton, D. C. Kale, C. Elkan, and R. C. Wetzell, "Learning to Diagnose with LSTM Recurrent Neural Networks," *CoRR*.
- [19] L. Oakden-Rayner, "The End of Human Doctors - Understanding Medicine," <https://lukeoakdenrayner.wordpress.com/2017/04/24/the-end-of-human-doctors-understanding-medicine/>, Apr. 2017, Accessed: 2017-09-21.
- [20] Z. Wang, W. Yan, and T. Oates, "Time Series Classification from Scratch with Deep Neural Networks: A Strong Baseline," in *IJCNN*. IEEE, pp. 1578–1585.
- [21] Nonin, "Nonin Xpod Pulse Oximeter." [Online]. Available: <http://www.nonin.com/Xpod>
- [22] W. Zong, T. Heldt, G. B. Moody, and R. G. Mark, "An Open-source Algorithm to Detect Onset of Arterial Blood Pressure Pulses," in *Computers in Cardiology*, 2003, Sept 2003, pp. 259–262.
- [23] B. Hu, Y. Chen, and E. J. Keogh, "Time Series Classification under More Realistic Assumptions," in *SDM*. SIAM, pp. 578–586.
- [24] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based Learning Applied to Document Recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov 1998.
- [25] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation Applied to Handwritten Zip Code Recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989. [Online]. Available: <http://dx.doi.org/10.1162/neco.1989.1.4.541>

Nicolas Bersano Born in Santiago, Chile, 1984. Bachelor of Science in Engineering, computer sciences and telecommunications, Diego Portales University, Santiago, Chile in 2009, Computer Sciences and telecommunications Engineer, Diego Portales University, Santiago, Chile in 2009, Magister of Sciences, computer science and telecommunications, Diego Portales University Santiago, Chile in 2011.

Worked as Research and Development Engineer at Delphos mining planning laboratory, Universidad de Chile, Lecturer at Diego Portales university, Development Engineer at JATCH Technology. Currently working as Research Engineer at Allm inc. Shibuya Tokyo. His published works Estimating the size of peer-to-peer networks using Lamberts W function (undergraduate thesis work), Metrics and models for social networks (Magister thesis work).

Horacio Sanson Obtained his Masters Degree at the Global Information and Telecommunication Institute of Waseda University in Tokyo, Japan, in 2005, where he researched distributed localization on emergency wireless networks. From there, he worked on several engineering positions until he became lead research engineer at Allm Inc. on 2014, where he spends his time applying engineering and machine learning to solve practical medical and health-care issues.