Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

Statistical approach based on Deep Neural Networks for Oscillometric Blood Pressure Estimation

Soojeong Lee, Joon-Hyuk Chang

School of Electronic Engineering, Hanyang University 222 Wangsimni-ro, Seongdong, Seoul 133-791, Korea

Tel.: + 82 2 2220 0355; fax: +82 2 2281 9912; jchang@hanyang.ac.kr

Abstract

Oscillometric blood pressure (BP) devices currently estimate a single point but do not identify fluctuations in BP or distinguish them from variations in response to physiological properties. In this paper, to analyze BP normality based on oscillometric measurements, we use statistical approaches including kurtosis, skewness, Kolmogorov-Smirnov, and correlation tests. Then, to mitigate uncertainties, we use a deep neural network (DNN) to determine the confidence limits (CLs) of BP measurements based on their normality. The proposed DNN regression model decreases the standard deviation of error (SDE) of the mean error (ME) and the mean absolute error (MAE) and reduces the uncertainty of the CLs and SDEs of the proposed technique. We validate the normality of the distribution of the BP estimation distribution which fits the Gaussian distribution very well. We use a rank test in the

© 2018 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

DNN regression model to demonstrate the independence of the artificial SBP and DBP estimations. First, we perform statistical tests to verify the normality of the BP measurements for individual subjects. The proposed methodology provides accurate BP estimations and reduces the uncertainties associated with the CLs and SDEs based on the DNN regression estimator.

Key words: Blood pressure, oscillometric measurement, statistical analysis, normality, confidence interval, deep belief networks.

---

1 Introduction

Blood pressure (BP), one of the most important vital signals, is a key consideration when making decisions about the cardiovascular activity of patients. In 2012, 18 million people died from cardiovascular diseases, which represents 31% of all global deaths [1]. The oscillometric method has recently gained popularity and is used in automatic BP measurement devices that are now readily available in the marketplace. Although vendors of oscillometric BP devices seldom disclose their algorithms, the maximum amplitude algorithm (MAA) is referred to in the literature as the one most commonly used for estimating arterial blood pressure (ABP) by the oscillometric method. The mean arterial BP is easily estimated via the cuff pressure whereby the oscillation amplitude becomes maximal based on the oscillometric waveform envelope when the cuff pressure is equal to the mean arterial BP. The compliance of the arterial wall is then maximized, so the ABP volume change with respect to the maximum

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

change in the arterial pressure and the amplitude of the oscillation signal in the cuff is also at maximum. Even though the ABP estimation relies on the oscillometric method, it is subject to various errors because the BP signal is characterized by continuous fluctuations. Specifically, the systolic blood pressure (SBP) and diastolic blood pressure (DBP) are subject to significant and continuous change over time [13] in response to physiological oscillations due to factors such as food intake, emotional status, level of exercise, and disease states, such that they can shift up to 20 mmHg within a few heartbeats [14]. However, the SBP and DBP are generally estimated at two random instances in time with respect to the systolic and diastolic ratios (SBPR and DBPR) in the oscillometric waveform envelope [13], which offers no information on the significant BP fluctuations. There is no established criterion for determining accurate SBP and DBP or SBPR and DBPR [8].

A recent journal editorial asserted that there is a clear recognition of biological variability [15]. This means that BP estimation using available automated monitors is subject to two sources of uncertainty the first being a measurement system inaccuracy and the second physiological variability. The first source of uncertainty is introduced by BP measurement error based on the ANSI/AAMI SP 10 standard [12], in which physiological oscillations in BP measurements are not perceived by most clinicians. Thus, oscillometric BP devices commonly provide a single-point estimate with no confidence limits (CLs) and fail to consider fluctuations in BP or to distinguish them from vari-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

ations due to physiological properties [15]. When CL estimates are available in these BP monitors, broad CL estimates can be utilized to recommend repeat analysis of BP measurements. Without a CL estimate, it is difficult to make any meaningful determination based on BP estimates. Although it is very important to determine CLs during BP estimation, there have been no significant studies to determine the CLs for oscillometric BP measurements [16] because many BP measurements are required to determine CLs using the student's t-distribution [17]. However, it is not practical to acquire a large number of measurements for each subject using oscillometric BP monitors because repeatable conditions for reproducible measurements cannot be guaranteed [4]. In this situation, it is necessary to estimate CLs using fewer BP measurements. Therefore, the bootstrap technique is employed to obtain CL estimates from BP measurements with small sample sizes [4]. Unfortunately, this approach assumes that BP based on the oscillometric measurement for individual subjects is an independent and identical distribution (iid), which supposes that the populations for which BP are measured are normally distributed. The assumption of normality is especially critical when determining CLs for BP measurement for individual subjects. The normality assumption also plays an important role in the reliability of estimations and statistical tests [38]. Thus, this assumption should be taken seriously because it is impossible to estimate BP accurately when this assumption does not hold. For this reason, the assumption of normality must be verified with respect to BP measurements for each subject.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

In this investigation, our research objective is to use statistical approaches including kurtosis, skewness, Kolmogorov-Smirnov (KS), and rank tests to examine the normality of BP values based on oscillometric measurements. We also estimate the SBP, DBP, and CLs of these values based on the normality for individual subject using a deep neural network [19]. Deep learning is a part of machine learning technique that models high-level abstractions in data by utilizing multi-layered architecture with complex nonlinear transformations [34]. However, our BP measurements for individual subjects were drawn from only a small sample size due to limited measurements available for individual subjects, which is a disadvantage when using deep neural networks (DNNs) that work best with big data. To address this problem, we first generated artificial features from the original feature data using a parametric bootstrap technique for the SBP and DBP estimates. This approach can efficiently represent the highly nonlinear relationship between the feature vectors obtained from oscillometric signals and those from target BPs. Then, we utilize the bootstrap technique again to determine the CLs based on the estimated target BP for the DNNs. Next, we statistically analyze the normality of the BP measurements. We believe this to be one of the first studies using statistical analysis for individual BP estimation.

This paper has the following contributions:

- We propose a novel approach for estimating accurate BP parameters that mitigates uncertainties (physiological variability), such as the CLs and stan-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

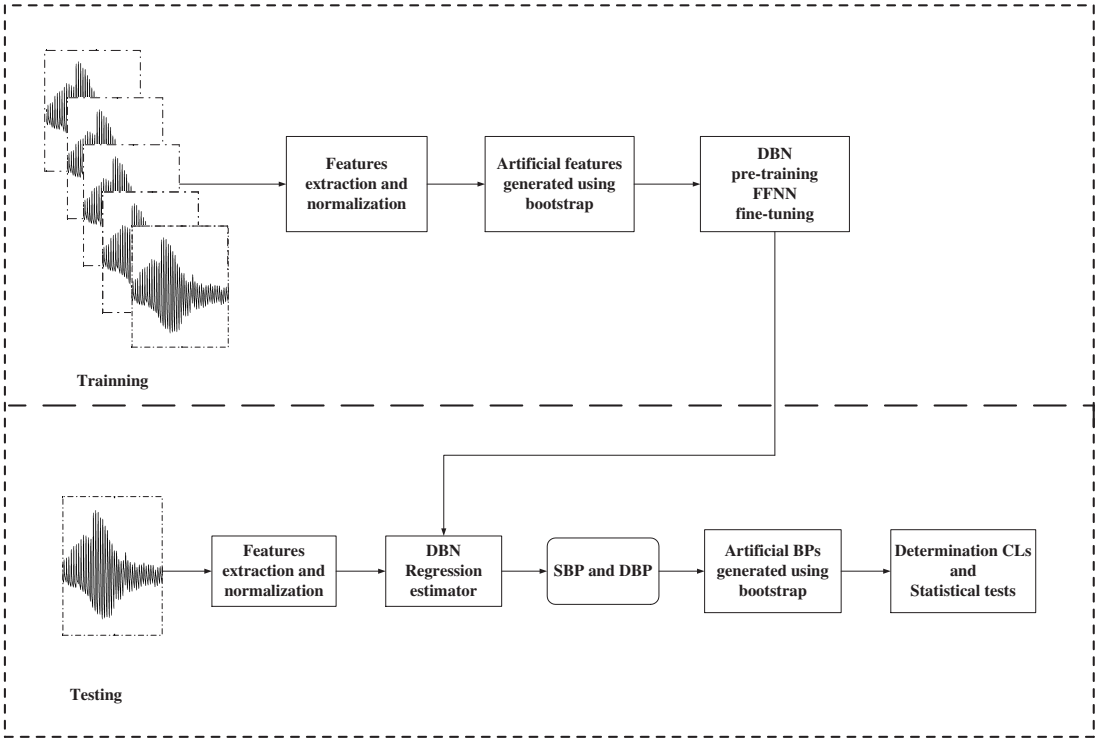

Fig. 1. A block diagram of the proposed methodology.

dard deviation of errors (SDEs), based on the DNN regression model. To do so, we use a small sample of oscillometric BP measurements and bootstrap techniques for individual subjects.

- First, we perform the kurtosis and skewness tests to verify the normality of the BP measurements for individual subjects.
- Next, we conduct KS tests to validate the distribution of the artificial BP measurements that fit the Gaussian distribution very well.
- We then use a rank test to analyze the independence between the artificial SBP and DBP estimations.

Fig. 1 shows a block diagram of the proposed methodology and we describe the procedure below. This is a two-step methodology. The upper path of the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

block diagram in the figure show the first step in which the proposed algorithm is trained and the lower path of the block diagram illustrates the testing step. We combine these two paths to estimate SBP and DBP. After bio-signal processing the oscillometric BP measurements, we obtain the features from the oscillometric waveform signals and envelopes. Next, we generate the artificial features from the original features and evaluate the normality of the distribution of all the features. We then build the DNN via the two training processes and perform pre-training and fine-tuning using a back-propagation algorithm [19]. In the second step, we use the unseen original feature set to evaluate the proposed methodology. Using the DNN regression models, we estimate the target BP values (SBP and DBP) for the individual subjects. Subsequently, to determine the CLs of the BPs, we generate artificial BPs using target BP values generated using the parametric bootstrap scheme. Finally, we fully analyze the normality of the artificial BPs for individual subjects.

This paper is organized as follows. In the next section, we describe the experimental data set and BP measurement methodology. In Section 3, we briefly describe the oscillometric blood pressure estimation using the DBN regression model. We describe the statistical analysis of the BP estimation and verify the normality of BP values based on oscillometric measurements in Section 4. In Section 5, we discuss the results and draw our conclusions.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

2 Data set

With reference to the measurement protocol of the institutional research ethics board, this work has been approved by a local research ethics committee and all test subjects provided informed consent prior to BP measurement. Biosign Technologies Inc. (Toronto, Ontario, Canada) provided the oscillometric BP measurements for this study. The experimental BP measurements were originally recorded from 85 healthy subjects with no history of cardiovascular disease, aged 12 to 80 years, of which 37 were females and 48 were males. We obtained five sets of oscillometric BP measurements from each volunteer (duration range of a single measurement: 31-95 sec, median duration: 55 sec) using a wrist-worn blood pressure device with a sampling rate of 100 Hz, which satisfies the recommendations of the ANSI/AAMI SP 10 standard [12], [16]. We averaged the readings made by two independent specialists (nurses) to yield one SBP and one DBP reading [4]. The BP measurement set contained comparatively stable observer (nurse) readings with a maximum difference between the results of the two observers being no more than 2 mmHg, which is well within the recommendations of the ANSI/AAMI SP 10 standard protocol (a mean difference less than 5 mmHg). Our BP measurements involved an oscillometric BP recording guided by the two trained nurses following one minute of rest. We repeated this step four more times, thereby generating five measurements. During data collection, each subject sat comfortably upright on a chair and the BP monitor cuff was strapped to the subject's left

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

wrist. Then the subject's arm was raised to the heart level at which time the auscultatory cuff, which served as the reference device, was fastened to the upper left arm, also at heart level [6]. The upper cuff around the arm was then inflated to occlude the brachial artery. As the cuff was allowed to deflate, turbulent blood flow generated Korotkoff sounds that could be easily heard using a stethoscope placed next to the upper cuff. The first Korotkoff sound (K1) was measured in mmHg by a manometer connected to the upper cuff and this reading was used to estimate SBP. The fifth sound (K5) was used to estimate DBP [24].

Due to the difficulty in occluding brachial arteries by upper arm sphygmomanometers, obtaining concurrent brachial and wrist measurements was not possible. Therefore, roughly 1.5 min after each pulse wave signal was acquired by the BP automated device, the two nurses utilized a classic upper arm sphygmomanometer to simultaneously record the subject's systolic blood pressure (SBP_1 and SBP_2) and diastolic blood pressure (DBP_1 and DBP_2). In this way, we obtained the mean values of the simultaneous readings as the reference BPs (SBP and DBP). Readings acquired by the first nurse were given the subscript 1 and those by the second nurse were readings are denoted as $[SBP_{1,j}, SBP_{2,j}, |j = 1, ..., 5]$ and $[DBP_{1,j}, DBP_{2,j}, |j = 1, ..., 5]$ for each subject, respectively.

With respect to time, each of the five classic upper arm sphygmomanometer measurements acquired by the first and second nurses $[(SBP_{1,j}, SBP_{2,j}), (DBP_{1,j}, DBP_{2,j}), |j =$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

$1, ..., 5]$ corresponded to a delay of roughly 1.5 min for each of the five pulse-wave bio-signals recorded by the automated oscillometric blood pressure monitor $[\text{UFIT}_j | j = 1, ..., 5]$ for each subject. We minimized this delay time between the classic upper arm and wrist measurements to the degree possible to minimize the natural blood pressure variability over time, while allowing enough time for the system to settle after the occlusion of the arteries during measurements. The 1.5 minute delay time between measurements is a compromise to minimize the method errors [12]. More details regarding this issue can be found in [26] for the interested readers.

With in the above conditions, we conducted training and test experiments to evaluate the proposed DNN regression estimator. In the proposed scenario, we sequentially separated the BP measurements of the subjects into a training set (300 measurements obtained from 60 subjects with five measurements) and testing set (125 measurements obtained from 25 subjects with five measurements). Then, we repeated this procedure to ensure that individual subjects were included only once in the testing step, so the proposed test scenario may be viewed to examine the generalization ability of the proposed BP measurement approach in new subjects. Five BP measurements (i.e., five samples for each feature) in an individual subject constitute only a small amount of input data in the training step. Thus, we also utilized the artificial features generated from the original feature vectors. In the test step, we used the unseen original features to verify the proposed algorithm. As mentioned above, we ob-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

tained our feature vectors from the oscillometric waveforms (OMWs), which we utilized to create artificial feature vectors based on the parameter bootstrap technique [22]. Thus, we obtained artificial feature vectors with training samples (ranging from $B = 5 - 1,000$) for each subject, such that each feature had training samples (ranging from $B = 5 - 1,000$). That is, we obtained 5,000 input vectors = 50 subject $\times$ B, for $B = 100$ for each feature in the training step.

3 Deep learning-based regression estimator

3.1 Features acquired from oscillometric wave signals

We extracted informative features from the oscillometric waveform signals [8] and envelopes after processing the signals of the oscillometric BP measurements to estimate the reference systolic blood pressure (RSBP) and reference diastolic blood pressure (RDBP). We then utilized these estimates as target BPs in the proposed DBN regression techniques. As the first feature, we used the mean arterial pressure (MAP:1), estimated using the maximum amplitude (MA) of the oscillometric waveform envelope. We then used (MA:2), the area under the envelope (AE:3), and the asymmetry ratio (AR:4) of the oscillometric waveform envelope (=LMAP/LE). We obtained this ratio by determining the length of the maximum amplitude's position (LMAP:5) and the length of the envelope (LE:6) acquired from the oscillometric wave signals because these

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

features are representative of the physiological characteristics of BP of individual subjects. We used features such as the amplitude, (σ_1:7), and (σ_2:8) values acquired from the asymmetrical Gaussian curve function on the oscillometric waveform envelope. We included in the features the averages of the oscillometry signal heart rate (9) [26], subject age (10), and subject gender (11) [8] were included in the features, where the number denotes the feature number. In all, we used 11 feature vectors to estimate the target BPs (RSBP and RDBP). We used a general normalization technique to restrict the ranges of the allowed feature values to between a minima and maxima of the predefined ranges [8].

3.2 Artificial features obtained using bootstrap

Since we deal with only five BP measurements for each individual subject, we created artificial features using the bootstrap technique as in [4], [22], which is a computer-intensive technique for improving the accuracy of estimates based on a small number of measurements for situations in which conventional methods are not valid for use in enhancing the accuracy [4], [22]. In this work, we assume $\mathbf{X} = [x_1, ..., x_N]$ to be a random sample of the distribution $\mathbb{F}$ with unknown parameters $[\theta, \sigma^2]$. Thus, we can approximate $\mathbb{F}$ by $\widehat{\mathbb{F}}(\widehat{\theta}, \widehat{\sigma}^2|\mathbf{X})$, where we define the mean and variance by $\mathbb{E}(\theta|\mathbf{X}) \simeq \widehat{\theta} = \bar{x} = \frac{1}{N}\sum_{n=1}^{N} x_i$ and $\mathbb{E}(\sigma^2|\mathbf{X}) \simeq \widehat{\sigma}^2 = [\frac{1}{N-1}\sum_{n=1}^{N}(x_n - \bar{x})^2]$, respectively, where $\widehat{\mathbb{F}} \simeq \mathcal{N}(\widehat{\theta}, \widehat{\sigma}^2)$ is approximated by a Gaussian distribution, which is called the parametric bootstrap technique [22]. We utilized this parametric way to create the artifi-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

cial features $\widehat{\mathbf{\Xi}}^* = [\theta_1^*, \theta_2^*, ..., \theta_B^*]$, which we acquired from the original feature vectors $\mathbf{X}$, and which have *priori* information such as the mean and variance for each subject, whereas the nonparametric bootstrap technique estimates $\mathbb{F}$ without making any *priori* assumptions.

3.3 KS test and its analysis for artificial feature

To use the artificial feature distribution with confidence, we fully evaluated the normality of each. First, we conducted the KS test to validate the normality of each artificial feature distribution. For this test, we suppose that $\mathbb{F}^*$ is a distribution of an artificial feature vector $[\theta_1^*, \theta_2^*, ..., \theta_B^*]$, which we obtain by utilizing an unknown sample distribution $\mathbb{F}$, where B is the number of replication. We thus can describe the result of the KS test as a probability that measures the correspondence between the artificial feature distribution and the hypothesis [16]. Based on the test results, we clearly confirmed the distribution of the artificial feature to be a Gaussian distribution as shown in Fig. 2, which shows histogram plots of the distribution of the artificial features that fit the Gaussian distribution. As the number of B replication increases, the distribution more closely adheres to a Gaussian distribution [4]. In this work, the null hypothesis supposes that the obtained distribution of the artificial features has an approximately Gaussian distribution, whereas the alternative hypothesis supposes that the distribution of the artificial features does not approximate a Gaussian distribution [29].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

- $\mathbb{H}_0$: The distributions of the artificial features and normal distributed empirical measurements are normal.
- $\mathbb{H}_1$: There is a large difference between the distribution of the artificial features and that of the normally distributed empirical measurements.

If the fitting test results of all the artificial features were 0, as represented in Table 1, we could not reject the null hypothesis at the $\alpha(= 0.05)$ significance level. Moreover, we note that all p values of the KS test are greater than the required significance level. Additionally, if the KS values are greater than the critical values (CV), the null hypothesis will be rejected. Therefore, we would necessarily accept the null hypothesis that the distributions of the artificial features follow Gaussian distributions [40].

3.4 Deep neural networks

We can see that the deep learning estimator is essentially based on the distributed representation, which implies that we can describe the obtained data by the interactions of various components at different levels [34]. We organized our deep learning-based regression estimator in two training stages, with greedy layer-wise unsupervised pre-training at each layer to preserve information from the input features and supervised back-propagation for fine-tuning the deep networks with respect to the target BPs. The pre-training phase consists of DNNs, which comprise a top-down graph calculated se-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

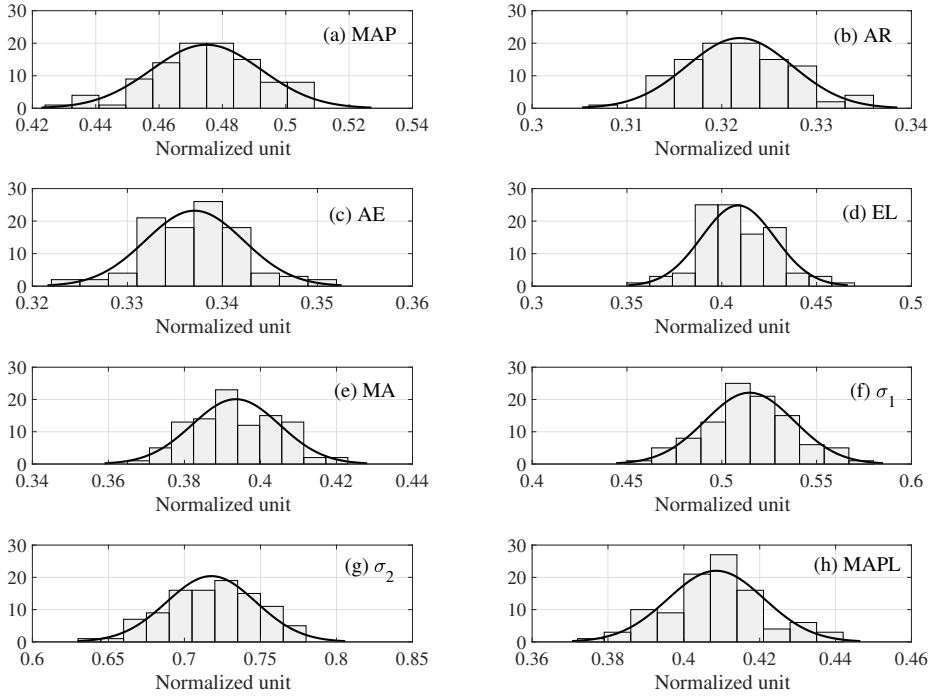

Fig. 2. The plots of distribution of artificial features based on the parametric bootstrap approach with replication numbers (B=100). Note that these artificial figures are examples obtained from one subject with 5 samples for 8 features [16].

quentially from the top layer to the input data. The DNN is a probabilistic generative model that is configured by multiple-hidden layers of stochastic latent variables. The top two layers are undirected connections, whereas the other hidden layers have a top-down acyclic graph wherein the units of the bottom layer represents visible units (which is an input feature vector), which are subsequently connected to two layers known as the restricted Boltzmann machine (RBM). The RBM is a basic component of the DNN that contains multiple layers of hidden units. The DNN can thus be given as $\mathbb{P}(\mathbf{X}^*, \mathbf{h}^1, \mathbf{h}^2, ..., \mathbf{h}^l) = \mathbb{P}(\mathbf{X}^*|\mathbf{h}^1)\mathbb{P}(\mathbf{h}^1|\mathbf{h}^2)\cdots\mathbb{P}(\mathbf{h}^{l-2}|\mathbf{h}^{l-1})P(\mathbf{h}^{l-1}, \mathbf{h}^l)$, where the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

Table 1

The hypothesis test values of pseudo features based on the parametric bootstrap approach with replication numbers (B=100), where h is the hypothesis test result, p denotes the p-value, ks is the statistic value of the Kolmogorov-Smirnov test, and cv denotes the critical value.

Feature/Test values	h	p	ks	cv
MAP	0	0.900	0.056	0.134
AR	0	0.904	0.055	0.134
AE	0	0.885	0.057	0.134
EL	0	0.716	0.068	0.134
MA	0	0.599	0.075	0.134
σ_1	0	0.996	0.040	0.134
σ_2	0	0.949	0.051	0.134
MAPL	0	0.969	0.048	0.134

probabilities of the condition layers $\mathbb{P}(\mathbf{h}^i|\mathbf{h}^{i+1})$, $\mathbf{h}^i$ represent the hidden units at layer i, and $\mathbf{X}^*$ denotes the resampled input data. We can define a joint probability with input data $\mathbf{X}^*$ and hidden layer activations $\mathbf{h}$ as follows:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

$$\mathbb{P}(\mathbf{X}^*, \mathbf{h}) = \frac{1}{Z} \exp^{-\mathbf{h}'\mathbf{W}\mathbf{X}^* - \mathbf{b}^{\mathrm{T}}\mathbf{X}^* - \mathbf{c}^{\mathrm{T}}\mathbf{h}} \tag{1}$$

where Z denotes the normalization factor for the joint probability, $\mathbf{b}$ is the bias vector for the visible units, $\mathbf{c}$ denotes the bias vector for the hidden units, and $\mathbf{W}$ represents the weight matrix for the layer. Thus, we an define the conditional probability of one layer, given the other by $\mathbb{P}(\mathbf{h}^i|\mathbf{h}^{i+1}) = \mathrm{sigm}\left(-\mathbf{b} - \sum_{k=1}^{n^{i+1}} \mathbf{W}\mathbf{h}_k^{i+1}\right)$, where $\mathrm{sigm}(x)$ is $1/(1+\exp(-x))$, $\mathbf{b} = [b_j^i]$ and denotes the bias vector for unit j of layer i, and $\mathbf{W} = [W_{kj}^i]$ is the weight matrix for layer i. In our case, the RBM comprises a two-layer, bipartite, generative undirected model with a set of binary hidden units $\mathbf{h}$ and a set of visible units $\mathbf{X}^*$. Here, a weight matrix $\mathbf{W}$ connects the visible and hidden units [33]. We also use a logistic function to activate the hidden units and the type of output as a linear value.

To connect a Gaussian visible layer with a binary hidden layer, we use the Gaussian-Bernoulli RBM [19] in this work, because we assume the artificial input feature vectors to follow an asymptotic Gaussian distribution. Then, we stack multiple Bernoulli-Bernoulli RBMs behind the first Gaussian-Bernoulli RBM [19], as shown in Fig. 3. Then, to minimize the negative log probability of the training vectors, we perform contrastive divergence to train the first Gaussian-Bernoulli RBM [19], [34] as an unsupervised learning algorithm. Next, we train the second Bernoulli-Bernoulli RBM by using the hidden layer of the first Gaussian-Bernoulli RBM as the second RBMs visible layer [19].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

In the training step of this work, we initialize the weights and biases by pre-training [34], and using them as an effective starting point for fine-tuning by back-propagation [19] based on feed-forward neural networks (FFNN) [11]. We use the minimum mean square error (MMSE) criterion [35] as a cost function based on a mini-batch-scaled conjugate gradient optimizer [31], for the estimated and reference BPs, as follows:

$$\mathbf{\Omega}(\mathbf{W},\mathbf{b}) = \frac{1}{N}\sum_{n=1}^{N}\sum_{d=1}^{D}\left(\widehat{\mathbf{Y}}_n^{*d}(\mathbf{W},\mathbf{b}) - \mathbf{Y}_n^d\right)^2 \tag{2}$$

where $\mathbf{\Omega}$ is the cost function; $\widehat{\mathbf{Y}}_n^{*d}(\mathbf{W},\mathbf{b})$ and $\mathbf{Y}_n^d$ denote the d^{th} estimated and reference BP vectors at the sample index n, respectively; N and D represent the mini-batch size and the feature vector's size, respectively; and (W^i, b^i) represents the weight bias parameters learned at the i^{th} layer. Then, we can iteratively update the estimated weights and bias, as follows:

$$\left(\mathbf{W}_{n+1}^i, \mathbf{b}_{n+1}^i\right) = -\epsilon\frac{\partial\mathbf{\Omega}}{\partial\left(\mathbf{W}_n^i, \mathbf{b}_n^i\right)} + \eta\left(\mathbf{W}_n^i, \mathbf{b}_n^i\right), 1 \le i \le L+1 \tag{3}$$

where ϵ is the learning rate, η denotes the momentum parameter, L represents the number of hidden layers, and $L+1$ is the output layer.

To summarize, we develop DNN regression model to learn the complex mapping function between the artificial feature vectors and the target BPs. The objetive to automatically learn this complicated relationship to estimate target BPs (RSBP and RDBP) from the artificial feature vectors, given a sufficient number of training samples, using the parametric bootstrap technique [22] in the training step. Finally, we estimate the SBP and DBP in our estimation

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

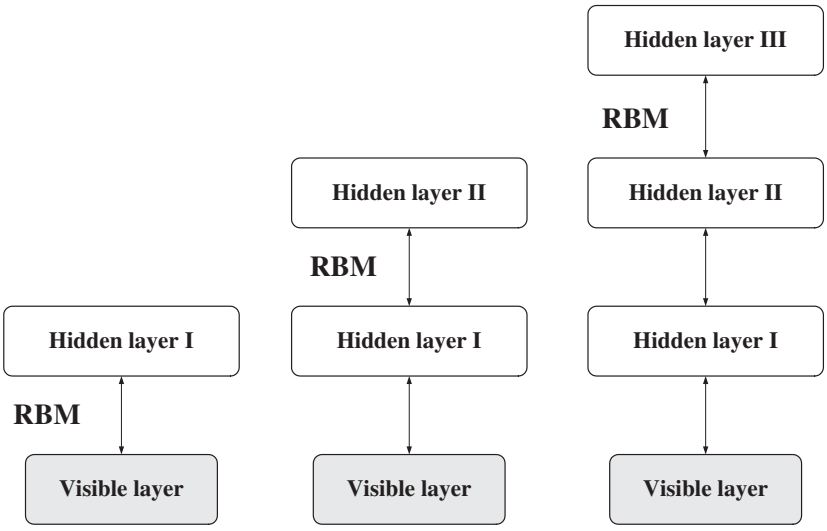

Fig. 3. Concept of the DNN, which is the generative model with several layers of latent variables while there are no intra-layer connection.

stage using the deep learning based regression model obtained in the training step.

4 Statistical analysis for BP estimation

4.1 CL estimation

Our goal in estimating the CLs is to use the bootstrap technique to identify the uncertainty (physiological variability) and to provide the CLs of five BP estimations obtained from the deep learning regression model for individual subjects. First, we describe the bootstrap principles of the nonparametric and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

parametric approaches. The basic idea is to generate many pseudo-BP estimates by resampling real BP estimates, $\mathbf{X} = [x_1, ..., x_n]$, utilizing n independent measurements from an unknown probability distribution $\mathbb{F}$ to generate a CL for $\widehat{\mu}(\mathbf{X})$.

In addition, we assume $\mathbf{X}$ to be a random sample of the distribution $\mathbb{F}$ with unknown parameter $[\mu, \sigma^2]$. As suh, we sample the pseudo sample $\mathbf{X}^*$ from $\widehat{\mathbb{F}}(\widehat{\mu}, \widehat{\sigma}^2|\mathbf{X})$ using the Monte-Carlo approach, whereby $[\widehat{\mu}, \widehat{\sigma}^2]$ is generally the maximum likelihood estimate from $\mathbf{X} = [x_1, ..., x_n]$. Therefore, when $B \rightarrow \infty$, we can approximate a Gaussian distribution as given by $\widehat{\mathbb{F}}(\widehat{\mu}^*, \widehat{\sigma}^{2*}|\mathbf{X}^*) \cong \mathcal{N}(\mu, \sigma^2)$.

In our approach, we an also obtain the CLs using the parametric bootstrap technique, where $\widehat{\mu}$ and $\widehat{\sigma}$ represent the estimated parameter variables acquired from the distribution, which was approximated using the BP target results in the DBN regression model. First, we can obtain the estimated BP ($\mathbf{Y}^*$) matrix, as shown in line 1 of Algorithm 1, which contains the the estimated SBP and DBP values, for each subject with five measurements, as shown in line 2. Then, we acquire two matrices as given by the following:

$$\mathbf{M}^{\mathrm{S}*}(i \mid \widehat{\mathbf{Y}}_i^{\mathrm{S}*}) = \left[[\mathbf{s}_{i,b}^j]^T, ..., [\mathbf{s}_{i,b}^j]^T\right] \tag{4}$$

$$\mathbf{M}^{\mathrm{D}*}(i \mid \widehat{\mathbf{Y}}_i^{\mathrm{D}*}) = \left[[\mathbf{d}_{i,b}^j]^T, ..., [\mathbf{d}_{i,b}^j]^T\right] \tag{5}$$

where superscripts S and D indicate SBP and DBP, respectively, and $*$ denotes the resampled data from the parametric bootstrap. Next, we vertically

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

calculate each column, as in Algorithm 1, to obtain the mean of each. We then perform ascending sorts as shown in lines 11-12. The sorted values are given as $\widehat{\Xi}^{\mathrm{S}*} = [\widehat{\theta}_1^{\mathrm{S}*}, \widehat{\theta}_2^{\mathrm{S}*}, \cdots, \widehat{\theta}_B^{\mathrm{S}*}]$, supposing $\widehat{\theta}_\alpha^{\mathrm{S}*}$ is the 100α th percentile of B bootstrap replications $[\widehat{\theta}_1^{\mathrm{S}*}, \widehat{\theta}_2^{\mathrm{S}*}, \cdots, \widehat{\theta}_B^{\mathrm{S}*}]$. We obtain the percentile interval $\widehat{\theta}_{\mathrm{lower}}^{\mathrm{S}*}, \widehat{\theta}_{\mathrm{upper}}^{\mathrm{S}*}$ of the intended coverage $1-2\cdot\alpha$, from this parametric percentiles bootstrap technique, as follows:

$$[\widehat{\theta}_{\mathrm{lower}}^{\mathrm{S}*}, \widehat{\theta}_{\mathrm{upper}}^{\mathrm{S}*}] = [\widehat{\theta}_{\alpha}^{\mathrm{S}*}, \widehat{\theta}_{1-\alpha}^{\mathrm{S}*}]. \tag{6}$$

Algorithm 1 CL estimation with bootstrap based on DNN model

1: **procedure** CL $(\widehat{\mathbf{Y}}^*)$

2: **for** $i \leftarrow 1, N$ **do** $\widehat{\mathbf{Y}}_i^{\mathrm{S}} = \{\widehat{s}_1, ..., \widehat{s}_n\}$ and $\widehat{\mathbf{Y}}_i^{\mathrm{D}} = \{\widehat{d}_1, ..., \widehat{d}_n\}$

3: $\widehat{\mu}_i^{\mathrm{S}} = \frac{1}{n}\sum_{j=1}^{n} \widehat{s}_j$ and $\widehat{\mu}_i^{\mathrm{D}} = \frac{1}{n}\sum_{j=1}^{n} \widehat{d}_j$

4: $\mathbb{E}(\widehat{\mu}_i^{\mathrm{S}}, \widehat{\sigma}_i^{\mathrm{S}} \mid \widehat{\mathbf{Y}}_i^{\mathrm{S}})$ and $\mathbb{E}(\widehat{\mu}_i^{\mathrm{D}}, \widehat{\sigma}_i^{\mathrm{D}} \mid \widehat{\mathbf{Y}}_i^{\mathrm{D}})$

5: $\mathbf{M}^{\mathrm{S}*}(i \mid \widehat{\mathbf{Y}}_i^{\mathrm{S}}) = \widehat{\mu}_i^{\mathrm{S}} + \widehat{\sigma}_i^{\mathrm{S}} \times \mathbf{RANDN}(\mathrm{n}, \mathrm{B})$

6: $\mathbf{M}^{\mathrm{D}*}(i \mid \widehat{\mathbf{Y}}_i^{\mathrm{D}}) = \widehat{\mu}_i^{\mathrm{D}} + \widehat{\sigma}_i^{\mathrm{D}} \times \mathbf{RANDN}(\mathrm{n}, \mathrm{B})$

7: **for** $b \leftarrow 1, B$ **do**

8: $\widehat{\theta}^{\mathrm{S}*}(i, b \mid \widehat{\mathbf{Y}}_i^{\mathrm{S}}) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{s}_{i,b}^{j}$

9: $\widehat{\theta}^{\mathrm{D}*}(i, b \mid \widehat{\mathbf{Y}}_i^{\mathrm{D}}) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{d}_{i,b}^{j}$

10: **end for**

11: $\widehat{\Xi}_i^{\mathrm{S}*} = \mathbf{SORT}\left[\widehat{\theta}^{\mathrm{S}*}(i, b \mid \widehat{\mathbf{Y}}_i^{\mathrm{S}})\right]$

12: $\widehat{\Xi}_i^{\mathrm{D}*} = \mathbf{SORT}\left[\widehat{\theta}^{\mathrm{D}*}(i, b \mid \widehat{\mathbf{Y}}_i^{\mathrm{D}})\right]$

13: **end for**

14: **end procedure**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

where N is the number of subjects, and n and B denote the number of measurements and replications, respectively, for individual subjects. It is difficult to obtain many BP measurements due to the cost and range of the physiological properties of individual subjects [4]. Even when cost is not considered, the experimental conditions might prevent reproducible BP measurements. Therefore, we estimate CLs using the bootstrap method based on the results of the DNN regression model.

4.2 Computing and testing for kurtosis and skewness

To verify the normality of the artificial BP measurements for individual subjects, we conducted kurtosis and skewness tests [41], and determined whether the distribution of the artificial BP measurements that resample a normal curve is an approximately Gaussian distribution. We computed the artificial sample mean and variance as follows: $\widehat{\mu}^* = \frac{1}{B}\sum_{b=1}^{B}\widehat{\theta}_b^{\mathrm{S}*}$ and $\widehat{\sigma}^{*2} = \mathrm{var}(\widehat{\boldsymbol{\Xi}}^{\mathrm{S}*})$, where B denotes the number of the artificial BP measurements for each subject and $\widehat{\boldsymbol{\Xi}}_{\mathrm{S}*} = [\widehat{\theta}_1^{\mathrm{S}*}, \widehat{\theta}_2^{\mathrm{S}*}, \cdots, \widehat{\theta}_B^{\mathrm{S}*}]$. We used Eqs. (7) and (8) to determine the kurtosis and the standard error of the kurtosis.

$$\mathrm{kurt}(\widehat{\boldsymbol{\Xi}}_{\mathrm{S}*}) = \left[\frac{\mathbb{E}(\widehat{\boldsymbol{\Xi}}_{\mathrm{S}*}^4) - 4\widehat{\mu}^*\mathbb{E}(\widehat{\boldsymbol{\Xi}}_{\mathrm{S}*}^3) + 6\widehat{\mu}^{*2}\widehat{\sigma}^{*2} + 3\widehat{\mu}^{*4}}{\widehat{\sigma}^{*4}}\right] \tag{7}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

$$\mathrm{se}_{\mathrm{kurt}(\widehat{\Xi}_{S*})} = \left(\frac{24n(n-1)^2}{(n-2)(n-3)(n+5)(n+3)} \right)^{\frac{1}{2}} \tag{8}$$

We used Eq. (9) to determine the skewness and the standard error of the skewness as follows:

$$\mathrm{skew}(\widehat{\Xi}_{S*}) = \left[\frac{\mathbb{E}(\widehat{\Xi}^3_{S*}) - 3\widehat{\mu}^* \mathbb{E}(\widehat{\Xi}^2_{S*}) + 2\widehat{\mu}^{*3}}{\widehat{\sigma}^{*3}} \right] \tag{9}$$

$$\mathrm{se}_{\mathrm{skew}(\widehat{\Xi}_{S*})} = \left(\frac{6n(n-1)}{(n-2)(n+1)(n+3)} \right)^{\frac{1}{2}} \tag{10}$$

Therefore, we evaluate the normality based on the z-scores for kurtosis and skewness, as given by the following, respectively:

$$z_{\mathrm{kurt}} = \frac{\mathrm{kurt}(\widehat{\Xi}_{S*}) - 0}{\mathrm{se}_{\mathrm{kurt}(\widehat{\Xi}_{S*})}} \tag{11}$$

$$z_{\mathrm{skew}} = \frac{\mathrm{skew}(\widehat{\Xi}_{S*}) - 0}{\mathrm{se}_{\mathrm{skew}(\widehat{\Xi}_{S*})}} \tag{12}$$

where we set α to 0.05, and then z-scores to verify that we have an approximately Gaussian distribution. Based on this outcome, we considered further joint testing for kurtosis and skewness. With normality, the null hypotheses for these situations are expressed as follows:

$$\mathbb{H}_0 : \mathrm{kurt}(\widehat{\Xi}_{S*}) = 3 \quad \text{and} \quad \mathrm{skew}(\widehat{\Xi}_{S*}) = 0. \tag{13}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

4.3 Normality test using KS

The KS one-sample test is a technique for evaluating the correspondence between two sets of values [40]. The null hypothesis states that the artificial BP measurements have approximately normal distributions. The alternative hypothesis is that the artificial BP measurement distributions are not approximately normal. We set the frequency level at $\alpha(= 0.05)$. That is, there is a 95% chance that any observed measurement statistical difference will be real and not due to chance. To conduct the KS test, we began with a decision about the relative empirical distribution $\widehat{\mathbb{F}}(\widehat{\mathbf{\Xi}}_{\mathrm{S}*})$, based on the observed artificial measurements. This test can be used to find a two-tailed probability p to determine if the artificial BP measurements are statistically similar or different by utilizing the point at which these two distributions exhibit the greatest divergence [41]. A p value smaller than the level of significance represents an artificial BP measurement distribution that is not adequately Gaussian. A p value greater than the level of significance is associated with the null hypothesis, meaning that the artificial BP measurement distribution is sufficiently Gaussian.

4.3.1 Theoretical background for KS test

The KS test is based on the empirical distribution function [40] and uses a cumulative distribution function (CDF) $\mathbb{F}(x) = \mathbb{P}(X_1 \leq x)$ (CDF) of the true

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

underlying distribution of the data. Here, to simplify the following equations, we assumed that $x \equiv \widehat{\theta}^{\mathrm{S}*}$, $X_1 \equiv \widehat{\theta}_1^{\mathrm{S}*}$ and $\mathbf{X} \equiv \widehat{\mathbf{\Xi}}_{\mathrm{S}*}$. We can define an empirical CDF in the following:

$$\mathbb{F}_B^*(x) = \mathbb{P}_B(\mathbf{X} \leq x) = B^{-1} \sum_{b=1}^{B} \mathbf{I}_{[-\infty,x]}(\mathbf{X}_b \leq x) \tag{14}$$

where $\mathbf{I}_{[-\infty,x]}(\mathbf{X}_b \leq x)$ represents an indicator equal to 1 if $(\mathbf{X}_b \leq x)$ and equality to 0 otherwise, which determines the proportion of sample points below level x. The law of large numbers is described such as

$$B^{-1} \sum_{b=1}^{B} \mathbf{I}_{[-\infty,x]}(\mathbf{X}_b \leq x) \longrightarrow \mathbb{E}\Big[\mathbf{I}_{[-\infty,x]}(\mathbf{X}_b \leq x)\Big] = \mathbb{P}(\mathbf{X} \leq x) = \mathbb{F}(x) \tag{15}$$

The proportion of the BP sample in the set $[-\infty, x]$ approximates the probability of this set.

Theorem 1: If $\mathbb{F}(x)$ continues then the distribution of this supremum does not depend on the unknown distribution $\mathbb{P}$ in the BP sample:

$$\| \mathbb{F}_B^*(x) - \mathbb{F}(x) \|_\infty \longrightarrow 0 \tag{16}$$

Therefore, this approximation holds uniformly over all $x \in \mathbb{R}$, where $\| \cdot \|_\infty$ is $\sup_{x \in \mathbb{R}} |\cdot|$. The result of Eq. (16) proves that the distribution $\mathbb{F}_B^*$ converges to that of $\mathbb{F}$ in the sense that the sup-norm of the difference has almost surely 0 probability. More details regarding theorem 1 are given in the appendix.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

4.4 Independence test based on rank

Next, we assume the B bivariate observations $\widehat{\boldsymbol{\Xi}}^{\mathrm{S}*} = [\widehat{\theta}_1^{\mathrm{S}*}, \widehat{\theta}_2^{\mathrm{S}*}, \cdots, \widehat{\theta}_B^{\mathrm{S}*}]$ and $\widehat{\boldsymbol{\Xi}}^{\mathrm{D}*} = [\widehat{\theta}_1^{\mathrm{D}*}, \widehat{\theta}_2^{\mathrm{D}*}, \cdots, \widehat{\theta}_B^{\mathrm{D}*}]$ for each individual subject (i) to be random resamples from the BP estimation results generated by the proposed approach, as shown in the algorithm. That is, the member of the $[\widehat{\boldsymbol{\Xi}}^{\mathrm{S}*}, \widehat{\boldsymbol{\Xi}}^{\mathrm{D}*}]$ set were mutually independent and identically distributed in the bivariate population. We suppose that $\mathbb{F}_{\widehat{\boldsymbol{\Xi}}^{\mathrm{S}*}, \widehat{\boldsymbol{\Xi}}^{\mathrm{D}*}}$ are joint distribution functions for the bivariate population of the $[\widehat{\boldsymbol{\Xi}}^{\mathrm{S}*}, \widehat{\boldsymbol{\Xi}}^{\mathrm{D}*}]$ set. Unlike other approaches, the Spearman rank-order correlation method deals with the relationship between two populations [40]. That is, this approach addresses how one population changes with respect to another. We obtained the $\mathbb{Z}$-value of a correlation coefficient (CORR) r and the CORR r for a large sample using the Spearman rank-order test [40].

As shown below, the null hypothesis indicates that there is no correlation (independence) between the artificial SBP and DBP measurements.

$$\mathbb{H}_0 : \left[\mathbb{F}_{\widehat{\boldsymbol{\Xi}}^{\mathrm{S}*}, \widehat{\boldsymbol{\Xi}}^{\mathrm{D}*}}(\widehat{\zeta}, \widehat{\eta}) \equiv \mathbb{F}_{\widehat{\boldsymbol{\Xi}}^{\mathrm{S}*}}(\widehat{\zeta}) \mathbb{F}_{\widehat{\boldsymbol{\Xi}}^{\mathrm{D}*}}(\widehat{\eta}), \quad \forall (\widehat{\zeta}, \widehat{\eta})\right]. \tag{17}$$

The level is the frequency set at $\alpha(= 0.05)$, so there is a 95% confidence that the statistical difference of any observed BP measurement will be real and not due to chance. The asymptotic approximation is based on the normality of r. To apply this theory, we must know the expected value and variance of r when the null hypothesis of independence is true. Thus, the expected value

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

and variance of r are given by $\mathbb{E}(r) = 0$ and $\mathrm{var(r)} = 1/\mathrm{B} - 1$ under $\mathbb{H}_0$, respectively. We verify these equations in the following calculations.

5 Experimental Results and Comparison

5.1 Evaluation protocols

In this section, we evaluate the statistical analysis results of the proposed approach for individual BP measurements by determining the ME and SDE between the estimated BP and the reference auscultatory measurement computed according to the recommendations of the AAMI standard protocol [12]. We also confirm the normality of the individual BP measurements based on the kurtosis, skewness, KS, and rank test results [40], [41] and we estimate the CLs of the BP measurement, based on the normality of the BP measurement for individual subjects.

We evaluate the BP measurement device based on the AAMI protocol to ensure that the ME is less than ± 5 mmHg with an SDE no greater than 8 mmHg [12]. Lower ME and SDE values indicate better overall performance. However, the SDE is of greater importance than the ME because BP devices can be very inaccurate, so the BP measurement can have a low ME and high SDE, which can be either positive or negative. Based on the British hypertension (BHS) protocol, we confirmed the percentages of the mean absolute

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

errors (MAEs) for three groups, $\leq$ 5 mmHg, $\leq$ 10 mmHg, and $\leq$ 15 mmHg, for all measurements (425 measurements). The BHS protocol graded system assigns devices to one of four classes, A through D. If 60% of a device's error measurements are within 5 mmHg, 85% of its error measurements are within 10 mmHg, and 95% of its error measurements fall within 15 mmHg [36], the device is classified as grade A. The BHS system has less stringent criteria for grades B and C and assigns a grade of D if the device performs worse than the criteria for grade C. We computed the ME and SDE values between the estimated BPs $(\widehat{t}_i, i = 1, ..., n)$ and the reference measurements $(t_i, i = 1, ..., n)$ according to the AAMI protocol. We calculated the MEs of the proposed method by $(e_i = t_i - \widehat{t}_i)$ and the ME and MAE were given by $(n^{-1} \sum_{i=1}^{n} e_i), (n^{-1} \sum_{i=1}^{n} \mid e_i \mid)$. We readily computed the SDE values of ME and MAE as well.

5.2 Statistical analysis

Table 2 shows a statistical information summary for the 85 subjects in this study and Table 3 shows the configuration and parameter settings of our deep learning technique. In the tables, the number associated with eah feature vectors represents a feature such as the MAP or AE, and we obtained the number of samples for each artificial feature (100) using bootstrap techniques.

Based on this configuration, we compared the MEs of the SBP and DBP

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

Table 2

The statistical information about eighty five subjects.

Statistical information	Value
Age (Male)	12 to 80
Age (Female)	17 to 65
Arm size	25 (cm) to 42 (cm)
Wrist size	13.5 (cm) to 23 (cm)
Deflation rate	3.0 (mmHg/s)
Male	48 of 85 (56.5%)
Female	37 of 85 (43.5%)

values acquired using the proposed DNN regression estimator with those when using the MAA [2], FFNN [11], and support vector regression (SVR) [18] models, as shown in Table 4. We determined the SDEs obtained by our DNN regression estimator were observed to be 6.30 and 5.45 mmHg for the SBP and DBP, respectively. The SDEs determined by the proposed DNN regression estimator were better by 1.18 and 1.35 mmHg for the SBP and DBP values, respectively, compared with those of the conventional FFNN, as shown in Table 4. We confirmed differences of 0.9 and 0.73 mmHg in the SDE values for the SBP and DBP, respectively, between the proposed DBN regression estimator and the SVR estimator. In addition, we used an SVR model with a linear epsilon insensitive cost function [18] to evaluate the performance of the proposed DNN regression estimator. We also constructed Bland-Altman

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

Table 3

Parameter setting [33], [19], [34] of the DBN based regression model, where 11 denotes the number of input units (namely, the input vector's dimension) and 2 is the number of output units (namely, the target vector's [SBP and DBP] dimensions).

Number of the hidden unit in three layers:	$[(11,(32),(32),(32),4)]$
Number of feature vector $\mathbf{X}$	11
Number of target vector $\mathbf{Y}$)	2
Number of sample over each pseudo feature	100
Number of sample over each original feature	5
Number of hidden layers	3
Number of hidden unit on the layers	16 to 256
Number of ensemble	50
Learning rate for weight	0.001
Learning rate for biases of visible units	0.01
Learning rate for biases of hidden units	0.01
Momentum rate	0.9
Activation type	sigmoid type function
Maximum epoch in the pre-training	200
Maximum epoch in the fine-tuning	200
Initial weights and biases	randomly between (-1, 1)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

plots to compare the performance of the proposed DBN regression model with that of the auscultatory nurse measurements (425 measurements), as shown in Fig. 4. The results of the BHS protocol [36] grading indicate that the DNN regression estimator provided more accurate BP estimates compared with those of the MAA, FFNN, and SVR techniques. Based on the BHS protocol system, the proposed methodologies obtained grades of A for both the SBP and DBP values. The results of the proposed DNN regression estimator were 69.18 % ($\leq$ 5 mmHg), 88.71 % ($\leq$ 10 mmHg), and 97.18 % ($\leq$ 15 mmHg) for the SBP in the given experimental test scenario and 76.24 % ($\leq$ 5 mmHg), 93.17 % ($\leq$ 10 mmHg), and 98.12 % ($\leq$ 15 mmHg) for the DBP in the experimental test scenario, as presented in Table 5. Therefore, the probabilities of the BHS criteria for the proposed methodologies are superior to those of the conventional techniques, as shown in Table 5.

To summarize, we found the performance of the proposed DBN regression estimator to be superior to that of conventional algorithms, as indicated in Tables 4 and 5. The proposed DNN regression estimator earned an overall grade of A according to the BHS grading system. In addition, we assessed the performance of the DNN regression estimator using the Bland-Altman plots shown in Fig. 4. These plots indicate that the BP estimates acquired by the proposed DNN regression estimator agree very well with the reference BPs (SBP and DBP). The bounds of agreement (see bold horizontal lines in Fig. 4) we used were (ME $\pm$ 2$\times$ SDE) for both plots, and most of the black asterisks

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

fall within these bounds of agreement. The biases (see the horizontal center lines) of the two plots were small ($\leq \pm$ 0.5 mmHg). Furthermore, the results shown in Fig. 4 accord with those in the last column in Table 5. Based on the overall performance evaluation results, we can conclude that the proposed DNN regression mitigates the ME variance and improves the performance confidence level.

As mentioned above, we compared the CLs results for SBP and DBP estimated by the proposed methodologies with those estimated by conventional methods (PMAE [4], MAA using ST [4], and the "Guide to the expression of uncertainty in measurements (GUM) [42]" as shown in Table 6. Specifically, to obtain the CLs, we used the bootstrap values based on the results of the proposed DBN regression estimator for SBP and DBP. Note that the parametric bootstrap technique to determine the CLs of the auscultatory BP measurements based on the results of the DNN regression approach. One goal of this paper was to determine the CLs for SBP and DBP estimates when only five BP measurements were available. The CLs of the SBP and DBP obtained by the proposed method were tighter than those acquired from the conventional methods. This was due to the decrease in the standard deviation because the results of the proposed method were improved using the parameter bootstrap technique in the algorithm. Although the CLs results obtained by the proposed method were wider than those obtained by PMAE, we noted differences of 2.0 mmHg and 1.6 mmHg in the SDEs of the CLs for the SBP

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

and DBP, respectively, between the proposed DNN regression estimator and the PMAE estimator. Thus, we can argue that the uncertainties of the SDEs were reduced when using the DNN regression estimator.

To evaluate normality with respect to the individual BP measurements, we statistically analyzed the kurtosis, skewness, KS, correlation test results with their associated standard deviations (std) using the artificial BP measurements based on the results of the DNN regression estimator. Kurtosis is a measure of a population that identifies how flat or peaked it is in terms of a Gaussian distribution. The kurtosis for a Gaussian distribution is 3. Therefore, a kurtosis value greater than 3 indicates a heavy-tailed distribution and a kurtosis value of less than 3 indicate a light-tailed distribution. As shown in Table 7 and Fig. 5, the kurtosis results for the SBP and DBP estimations were 2.99 and 3.01, respectively, which means that the distributions of the artificial BP measurements were normal. The skewness of a population can be represented as measure of its horizontal symmetry with respect to a Gaussian distribution. The skewness of a Gaussian distribution is 0 and symmetric data have a skewness of almost 0, with negative values skewed to the left and positive values skewed to the right. We confirmed that the symmetric distributions of the artificial BP measurements were -0.01 and 0.01, respectively, for the SBP and DBP estimations, which indicates that they were very close to 0 for the number of bootstrap replications (B=1000), as shown Fig. 5 and Table 7.

We performed a KS test to validate the normality of each distribution and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

these fit the Gaussian distribution [29] very well. We supposed that we obtained $\mathbb{F}_*$, the distribution of the artificial BP measurements, using an unknown sample distribution $\mathbb{F}$. Thus, we represent the test result as a probability that measures the correspondence between the distribution of the artificial BP measurement and the hypothesis [16]. As the number of B bootstrap replications became large, the distribution of the artificial BP measurement was close to a Gaussian distribution. We confirmed the distribution of the artificial BP measurement to be a Gaussian distribution based on the test results, which represent the cumulative distribution function (CDF) of the artificial BP measurements versus the theoretical CDF of a Gaussian distribution, as shown in Fig 6. Moreover, we confirmed that the hypothesis results for the mean artificial BP measurements were 0 for both the SBP and DBP estimations. These results indicate that we did not reject the null hypothesis at the 0.05 significance level. In addition, we could not reject the null hypothesis because the $ks(= 0.02)$ values were less than the critical values $cv(= 0.04)$. Also, our $p(= 0.78)$ and $(= 0.79)$ KS test values were greater than the $\alpha(= 0.05)$ significance level for SBP and DBP, as shown in Table 7. Therefore, we can accept the normality of the distribution of the artificial BP measurements for individual subjects. In addition, we obtained CORR $r_s(= 0.01)$ and $r_d(= 0.01)$ values based on the asymptotic normality between the artificial SBP and DBP measurements and acquired variances of 0.0009 and 0.0009 for r_s and r_d using the independence based on rank as described the last column in Table 7, respectively. Therefore, these results are clearly close to the expected values and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

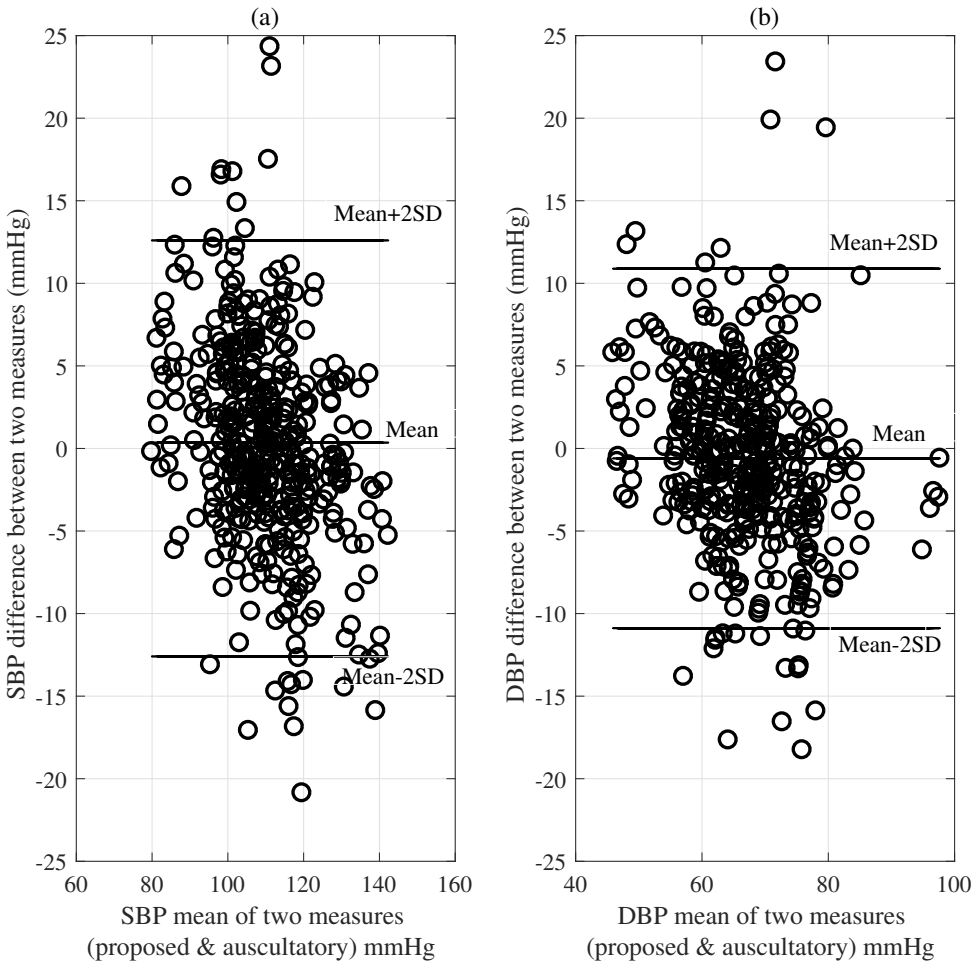

Fig. 4. The Bland-Altman plots comparing the performance between the proposed DNN based regression estimator and the auscultatory nurse measurements [12]. (a) Bland-Altman plot for the SBP. (b) Bland-Altman plot for the DBP.

variances. Thus, we did not reject $\mathbb{H}_0$ (independence) at the $\alpha = 0.05$ level.

6 Conclusion

The main contribution of this paper was our verification of the the normality of the BP measurements, based on only five samples, using various statistical

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

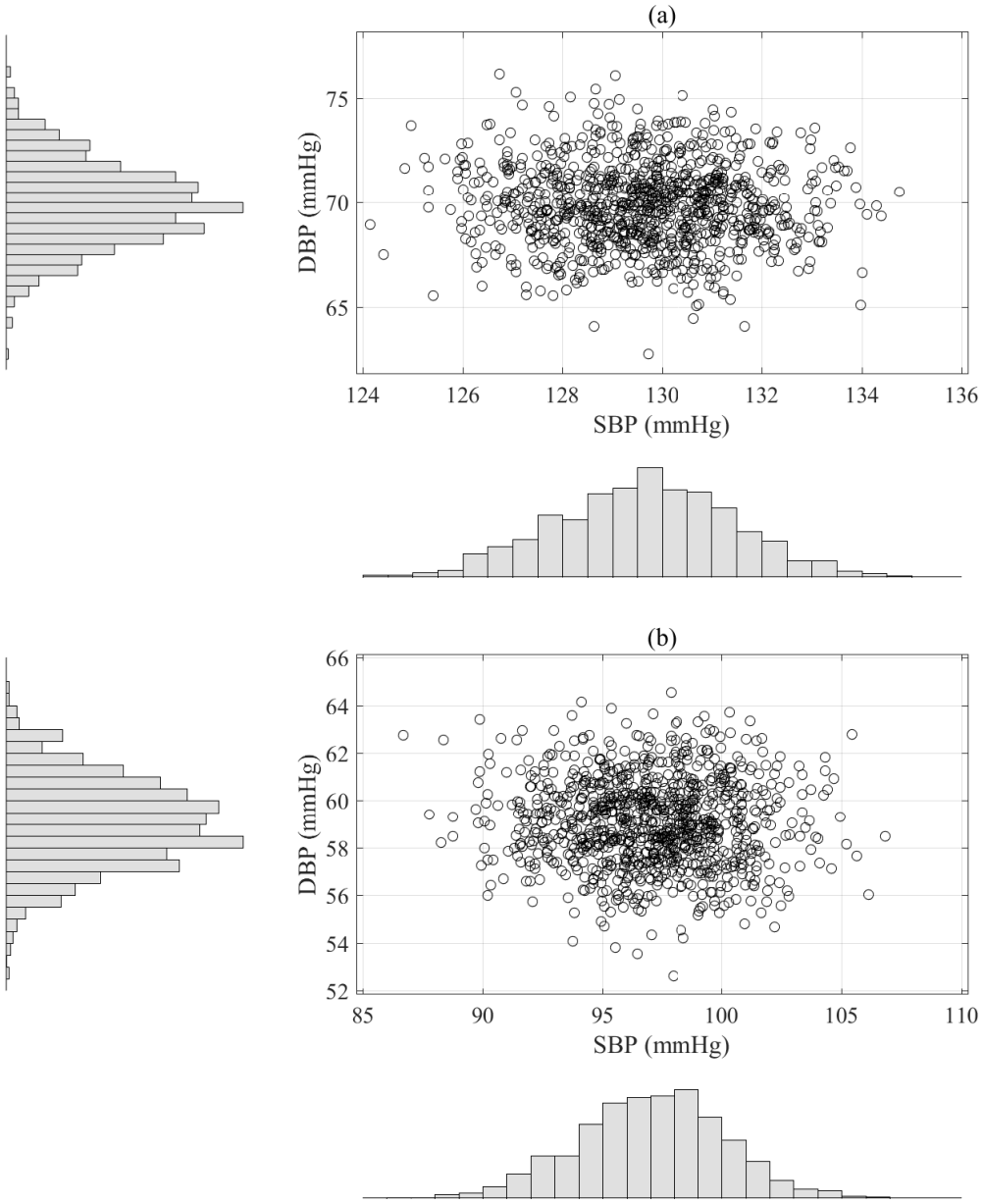

Fig. 5. Scatter histograms of BP estimation based on the parametric bootstrap approach with replication numbers (B=1000) using the results of the DNN estimator, where the upper plot (a) is an example obtained from one subject and the bottom plot (b) is another example acquired from different subject.

methodologies for individual subjects. We clearly determined the independence of the artificial SBP and DBP estimations using the DNN regression

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

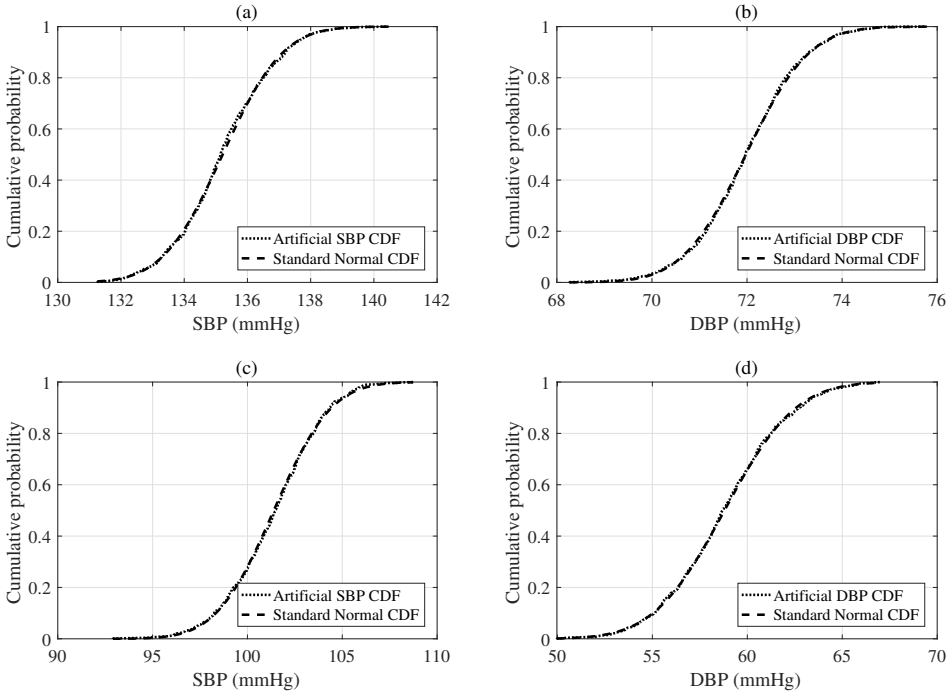

Fig. 6. Cumulative distribution functions (CDFs) of selected artificial BP estimations obtained from the parametric bootstrap approach with replication numbers (B=100) based on the results of the DNN regression model. Note that the upper plot (a) and (b) are the examples obtained from one subject with 5 BP estimations and the bottom plot (c) and (d) are the examples acquired from another subject with 5 BP estimations.

model based on the distribution-free test of rank. The proposed methodology also provides accurate BP estimations and reduces uncertainties such as the CLs and SDEs, as determined by the DNN regression estimator. In future work, we will pursue clinical testing on a wide range of new subject populations based on the ANSI/AAMI SP 10 and BHS protocols.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

Table 4

ME and SDE relative to the reference auscultatory method obtained with the conventional MAA, FFNN, SVR[18], and DNN regression model, where the results are the average values for our test data.

mmHg	MAA		FFNN		SVR		DBN	
Test	SBP	DBP	SBP	DBP	SBP	DBP	SBP	DBP
ME	0.07	-0.89	0.25	-0.22	-0.51	0.17	0.36	-0.61
SDE	9.28	7.76	7.48	6.80	7.20	6.18	6.30	5.45

Table 5

Grading of the proposed algorithm based on the BHS standard using the results of MAA, FFNN, SVR [18], and DNN on $(5 \times 85 = 425)$ measurements.

	SBP			DBP			Standard (SBP/DBP)
Tests	Absolute difference (%)			Absolute difference (%)			BHS
	≤5 mmHg	≤10 mmHg	≤15 mmHg	≤5 mmHg	≤10 mmHg	≤15 mmHg	Grade
MAA	47.06	85.88	96.47	56.47	88.24	97.65	C/B
FFNN	53.88	85.65	95.53	66.12	94.12	98.82	B/A
SVR	62.59	86.12	95.53	74.12	93.65	96.94	A/A
DNN	69.18	88.71	97.18	76.24	93.17	98.12	A/A

Conflict of interest statement

None declared.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

Table 6

Summary of the CI of the SBP and DBP for nurse measurements, the proposed DNN regression model and the conventional methods, n (= 85) is the number of subjects and SDE is a standard deviation of error; L and U are the lower and upper limits, respectively.

BP (mmHg)	SBP (SDE) 95%CI	DBP (SDE) 95%CI	SBP L (SDE)	SBP U (SDE)	DBP L (SDE)	DBP U (SDE)
MAA_{ST}	13.2 (8.0)	9.4 (5.8)	106.7 (14.3)	120.2 (16.5)	62.4 (10.4)	71.7 (11.0)
MAA_{GUM}	13.9 (7.9)	10.0 (5.4)	106.4 (14.3)	120.5 (16.4)	62.0 (10.4)	72.1 (10.9)
$PMAE_{NPB}$	2.8 (3.3)	1.7 (2.4)	112.4 (13.9)	115.7 (14.1)	66.7 (10.5)	68.2 (9.9)
DNN_{Boot}	5.5 (1.3)	4.2 (0.8)	107.4 (12.7)	113.0 (12.6)	64.5 (8.3)	68.6 (8.4)

Table 7

Statistical results such as skewness, kurtosis, KS, Spearman's correlation between the SBP and DBP estimations, where the results are the average values for our test data.

Tests	KS test				Normality test		
α(= 0.05)	h (std)	p (std)	ks (std)	cv (std)	kurtosis (std)	skewness (std)	corr (std)
SBP	0 (0)	0.78 (0.2)	0.02 (0.00)	0.04 (0.00)	2.99 (0.14)	-0.01 (0.08)	0.01 (0.03)
DBP	0 (0)	0.79 (0.2)	0.02 (0.01)	0.04 (0.00)	3.01 (0.16)	-0.01 (0.07)	0.01 (0.03)

Acknowledgments

This work was supported by

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

References

[1] $www.who.int/cardiovascular_diseases/en/$, World Health Organization (WHO), (2013).

[2] G. Drzewiecki, R. Hood, H. Apple, Theory of the oscillometric maximum and the systolic and diastolic ratios, ANN. Biomed. Eng. 22 (1) (1994) 88-96.

[3] P. D. Baker, D. R. Westenskow, K. Kuck, Theoretical analysis of non-invasive oscillometric maximum amplitude algorithm for estimating mean blood pressure, Med. Biol. Eng. Comput. 35 (3) (1997) 271-278.

[4] S. Lee, M. Bolic M, V. Groza, H. Dajani, S. Rajan, Confidence interval estimation for oscillometric blood pressure measurements using bootstrap approach, IEEE Trans. Instumen. Meas. 60 (10) (2011) 3405-3415.

[5] L. A. Geddes, M. Voelz, C. Combs, D. Reiner, C. F. Babbs, Characterization of the oscillometric method for measuring indirect blood pressure, ANN. Biomed. Eng. 10 (6) (1982) 271-280.

[6] S. Chen, "Improving algorithms for oscillometric blood pressure estimation by suppressing breathing effects," M.S. thesis, Univ. Ottawa, Ottawa, ON, Canada, 2010.

[7] S. Lee, J.-H. Chang, S. W. Nam, C. Lim, S. Rajan, H. Dajani, and V. Groza, "Oscillometric blood pressure estimation based on maximum amplitude algorithm employing Gaussin mixture regression," *IEEE Trans. Instumen. Meas.*, vol. 62, no. 12, pp. 3387-3389, Dec. 2013.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

[8] S. Lee *et al.*, "Improved Gaussian mixture regression based on pseudo feature generation using bootstrap in blood pressure measurement," *IEEE Trans. Ind. Informat.*, 12 (6) (2016) 2269-2280.

[9] R. Raamat *et al.*, "Errors of oscillometric blood pressure measurement as predicted by simulation," *Blood Press. Monit.*, vol. 16. no. 5, pp 73-76, Jun. 2011.

[10] J. Liu *et al.*, "Error mechanisms of the oscillometric fixed-ratio blood pressure measurement method," *Ann. Biomed. Eng.*, vol. 41, no. 3, pp. 587-597, Mar. 2013.

[11] M. Forouzanfar, H. Dajani, V. Groza, M. Bolic, S. Rajan, Feature-based neural network approach for oscillometric blood pressure estimation, IEEE Trans. Instumen. Meas. 60 (8) (2011) 2786-2796.

[12] *Assiociation for the advancement of medical instrumentation (AAMI), American national standard manual*, electronic or automated sphygmonanometers, 2003; AASI/AAMI SP 10:2002.

[13] K. Soueidan *et al.*, "Augmented blood pressure measurement through the noninvasive estimation of physiological arterial pressure variability," *Physiol. Meas.*, vol. 33, no. 6, pp. 881-899, May 2012.

[14] S. Hansen and M. Staber, Oscillometric blood pressure measurement used for calibration of the arterial tonometry method contributes significantly to error, Eur. J. Anaesthesiol. 23 (2006) 781-787.

[15] D. W. Jones and J. E. Hall, Hypertension-pathways to success, Hypertens.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

Editorial. 51 (2008) 1249-1251.

[16] S. Lee, S. Rajan, C. H. Park, J.-H. Chang, H. Dajani, and V. Groza, "Estimated confidence interval from single blood pressure measurement based on algorithm fusion," *Comput. Biol. Med.*, vol. 62, pp. 154-163, Jul. 2015.

[17] L. R. Krakoff, Confidence limits for interpretation of home blood pressure recordings, Blood Press. Monit. 14 (4) (2009). 172-177.

[18] A. Rakotomamonjy, "Analysis of SVM regression bound for variable ranking," *Neurocomputing*, vol 70, pp 1489-1491, Mar. 2007.

[19] G. Hinton, S. Osindero, and Y. W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527-1554, Jul. 2006.

[20] L. Breiman, "Bagging predictors," *Mach. Learn.* , vol. 24. no. 2, pp. 123-140, 1996.

[21] Y. Freund and R. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119-139, 1997.

[22] B. Efron, R. Tibshirani, Bootstrap methods for standard errors, confidence interval, and other measures of statistical accuracy, Statistical science 1 (1) (1986) 54-77.

[23] D. P. Landau and K. Binder, A guide to Monte Carlo simulations in statistical physics Cambridge university press, Cambridge, 2014.

[24] M. R. Neuman, "Measurement of blood pressure," *IEEE Pulse*, vol. 2, no. 2, pp. 39-43, Mar./Apr. 2011.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

[25] P. Buhlmann and B. Yu, "Analyzing bagging," *ANN. STAT.*, vol. 30, no. 4, pp. 927-961, 2002.

[26] S. Ahmad, M. Bolic, H. Dajani, V. Groza, I. Batkin, S. Rajan , Measurement of heart rate variability using an oscillometric blood pressure monitor, IEEE Trans. Instumen. Meas. 59 (10) (2010) 2575-2590.

[27] M. Forouzanfar, H. Dajani, V. Groza, M. Bolic, S. Rajan, and I. Batkin"Oscillometric blood pressure estimation: past, present, and future," *IEEE Rev. Biomed. Eng.*, vol. 8, pp. 44-62, 2015.

[28] H. C. Lin, "Specialised non-invasive blood pressure measurement algorithm," M. E. Thesis, Auckland Univ. Tec., Auckland, New Zealand, 2007.

[29] A. Ghasemi and S. Zahediasl, Normality tests for statistical analysis: a guide for non-statisticians Int. J. Endocrinol. Metab. 10 (2012) 486-489.

[30] C. M. Bishop, *Pattern recognition and machine learning*, NY, USA: Springer, 2006.

[31] M. Moller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks*, vol. 6, no. 4, pp. 525-533, 1993.

[32] S. Theodoridis, Machine learning, London, UK: Academic press, 2015.

[33] X. L. Zhang and J. Wu, "Deep belief networks based voice activity detection, " *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 4, pp. 697-710, Apr. 2013.

[34] Y. Bengio, "Learning deep architectures for AI," *Foundat. Trends in Mach. Learn.*, vol. 2, no. 1, pp. 1-127, 2009.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

[35] Y. Xu *et al.*, "A regression approach to speech enhancement based on deep neural networks," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 23. no. 1, pp. 7-19, Jan. 2015.

[36] E. O'Brien, R. Asmar, L. Beilin, Y. Imai, J.-M. Mallion, G. Mancia, T. Mengden, M. Myers, P. Padfield, P. Palatini, G. Parati, T. Pickering, J. Redon, J. Staessen, G. Stergiou, P. Verdecchia, European society of hypertension recommendations for conventional, ambulatory and home blood pressure measurement, J. of Hypertension, 21 (5) (2003) 821-848.

[37] BIPM, IEC, IFCC, ISO, IUPAC, OIML, Guide to the expression of uncertainty in measurement, (1993).

[38] A. F. Galvao, G. M.-Rojas, and W. S.-Escudero, Test for skeness and kurtosis in the one-way error component model, J. of Multivariate Analysis. 122 (2013) 35-52.

[39] A. M. Zoubir, B. Boashash, The Bootstrap and its application in signal processing, IEEE Signal Process. Mag. 15 (1) (1998) 56-76.

[40] M. Hollander and D. A. Wolfe, Nonparametric statistical methods, NewYork: Willey (1999).

[41] G. W. Corder and D. I. Foreman, Nonparametric statistics for non-statisticians, New Jersey: Willey (2009).

[42] "Guide to the expression of uncertainty in measurement, "BIPM, IEC, IFCC, ISO, IUPAC, and OIML, 1993.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

Appendix

Proof 1: We can define the inverse of the CDF F given by

$$\mathbb{F}^{-1}(y) = \lim\Big[x : \mathbb{F}(x) \geq y\Big]. \tag{18}$$

Then, building the change of variable $y = \mathbb{F}(x)$ or $x = \mathbb{F}^{-1}(y)$, we define such that

$$\mathbb{P}\Big(\parallel \mathbb{F}^{*}_{B}(x) - \mathbb{F}(x) \parallel_{\infty} \leq \epsilon\Big) = \mathbb{P}\Big(\parallel \mathbb{F}^{*}_{B}(\mathbb{F}^{-1}(y)) - y \parallel_{0 \leq y \leq 1} \leq \epsilon\Big) \tag{19}$$

where ϵ denotes a small value. Based on the definition of the empirical CDF F, we can define as follows:

$$\mathbb{F}^{*}_{B}(\mathbb{F}^{-1}(y)) = B^{-1}\sum_{b=1}^{B}\mathbf{I}_{[-\infty,x]}(\mathbf{X}^{*}_{b} \leq \mathbb{F}^{-1}(y)) = B^{-1}\sum_{b=1}^{B}\mathbf{I}_{[-\infty,x]}(\mathbb{F}(\mathbf{X}^{*}_{i}) \leq y) \tag{20}$$

Thus,

$$\mathbb{P}\Big(\parallel \mathbb{F}^{*}_{B}(\mathbb{F}^{-1}(y)) - \mathbb{F}(x) \parallel_{0 \leq y \leq 1} \leq \epsilon\Big) = \mathbb{P}\Big(\parallel B^{-1}\sum_{b=1}^{B}\mathbf{I}_{[-\infty,x]}(\mathbb{F}(\mathbf{X}^{*}_{b}) \leq y) - y \parallel_{0 \leq y \leq 1} \leq \epsilon\Big) \tag{21}$$

where the distribution of $\mathbb{F}(\mathbf{X}^{*}_{b})$ denotes uniform on the interval $[0, 1]$ because the CDF of $\mathbb{F}(X_1)$ is given by

$$\mathbb{P}\Big(\mathbb{F}(X_1) \leq \quad\Big) = \mathbb{P}\Big(X_1 \leq \mathbb{F}^{-1}(\epsilon)\Big) = \mathbb{F}\Big(\mathbb{F}^{-1}(\epsilon)\Big) = \epsilon. \tag{22}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2018 doi:10.20944/preprints201805.0276.v1

Therefore, the random variables $\mathbf{U}_i = \mathbb{F}(\mathbf{X}_b^*)$, $\forall b \leq B$ are independent and have uniform distribution on $[0, 1]$, we thus confirmed that

$$\mathbb{P}\Big(\parallel \mathbb{F}_B^*(x) - \mathbb{F}(x) \parallel_\infty \leq \epsilon\Big) = \mathbb{P}\Big(\parallel B^{-1} \sum_{b=1}^{B} \mathbf{I}_{[-\infty,x]}(\mathbf{U}_i) \leq y) - y \parallel_{0 \leq y \leq 1} \leq \epsilon\Big) \quad (23)$$

That is clearly independent of $\mathbb{F}$.

$$\sqrt{B}\Big[\mathbb{F}_B^*(x) - \mathbb{F}(x)\Big] \longrightarrow \mathcal{N}\Big(\widehat{\mu}, \mathbb{F}(x)(1 - \mathbb{F}(x))\Big) \quad (24)$$

because $\mathbb{F}(x)(1 - \mathbb{F}(x))$ denotes the variance of $\mathbf{I}(X_1 \leq x)$.

$$\parallel \mathbb{F}_B^*(x) - \mathbb{F}(x) \parallel_{x \in \mathbb{R}} \longrightarrow \mathcal{N}\Big(\widehat{\mu}, \widehat{\sigma}^2)\Big) \quad (25)$$

Therefore, it is closed to the normal distribution.