

Article

Uncertainty in Blood Pressure Measurement Estimated Using Ensemble-Based Recursive Methodology

Soojeong Lee ^{1,*}, Hilmi R Dajani ², Sreeraman Rajan ³, Gangseong Lee ⁴ and Voicu Z Groza ²

¹ Department of Computer Engineering, Sejong University, 209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Korea

² School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N6N5, Canada; hdajani@uottawa.ca (H.R.D.); Groza@eecs.uOttawa.ca (V.Z.G.)

³ Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S5B6, Canada; sreeramanr@sce.carleton.ca

⁴ Ingenium College, Kwangwoon University, 20 Kwangwoon-ro, Nowon-gu, Seoul 01897, Korea; gslee0115@gmail.com

* Correspondence: leesoo86@sejong.ac.kr; Tel.: +82-2-6935-2480

Received: 18 February 2020; Accepted: 27 March 2020; Published: 8 April 2020



Abstract: Automated oscillometric blood pressure monitors are commonly used to measure blood pressure for many patients at home, office, and medical centers, and they have been actively studied recently. These devices usually provide a single blood pressure point and they are not able to indicate the uncertainty of the measured quantity. We propose a new technique using an ensemble-based recursive methodology to measure uncertainty for oscillometric blood pressure measurements. There are three stages we consider: the first stage is pre-learning to initialize good parameters using the bagging technique. In the second stage, we fine-tune the parameters using the ensemble-based recursive methodology that is used to accurately estimate blood pressure and then measure the uncertainty for the systolic blood pressure and diastolic blood pressure in the third stage.

Keywords: uncertainty; confidence interval; oscillometry blood pressure measurement; deep neural network; ensemble method

1. Introduction

Blood pressure (BP) always fluctuates due to factors such as stress, exercise, disease, and inherent physiological oscillations [1]. However, the physiological variance of BP, which can rise up to 20 mmHg, has been neglected so far [2]. Even though physiological uncertainty is larger than the margin of error of standard BP measurement protocols. The issue of accuracy, precision and uncertainty in the measurement of physiological parameters has been of constant concern for practitioners [3]. Even though the standard for the expression of uncertainty in measurement [4] states that it is applicable to a broad span of fields, in practice it has been applied only to measurements determined on the basis of a series of observations obtained under repeatable conditions, a situation which is rarely reproducible in physiological measurements. Recently cuff-less BP devices using the photoplethysmograph (PPG) sensors have recently been used to measure BP [5–7]. Kachuee et al. [5] proposed a method to estimate BP based on the PPG for continuous health-care monitoring. A continuous BP estimation method using machine learning was introduced by Chen et al. [6]. Tjahjadi et al. [7] a novel classification method for BP based on PPG signal using long short-term memory (LSTM). However, automated oscillometric blood pressure measurement methods [8,9] are more commonly and popularly used to measure BP for many patients at the home, office, and medical centers. These devices usually offer a single BP value. However, these devices are unlikely to provide values that are superior to those that may be obtained by repeated BP measurements [10]. That is

because individual BP measurements are inherently subject to sources of uncertainty sources that cause deviations of the measured value (i.e., estimate) from the true BP value (i.e., reference BP) [11]. The sources of uncertainty can be separated into random errors and systematic errors [11], which will be dealt with in more detail in the following section. If BP measurements are simultaneously influenced by many sources of uncertainty, the distribution function of these measurements converges towards Gaussian distribution as the number of uncertainties increases, regardless of the distribution function of parameters representing the source of uncertainty [11]. Few researchers have tried to study uncertainty in physiological measurements [12,13] and there have been no attempts to include the characterization of the quality of the acquired signal and its compatibility with the employed estimation algorithms in a global figure of merit of confidence in the measurement accuracy. Thus, the confidence interval (CI) should be provided in such a way as to assess and express uncertainty in BP measurements where CI provides an estimated range of BP values, which possibly includes significant unknown sources [11]. Based on some aggregated statistics, wide CIs can provide alerts to patients, medical staff, and families. Thus, measuring the CI of a blood pressure measurement is very important, but unfortunately, there has been very little study done to determine the CI for an oscillometric BP measurement. Indeed, in order to estimate the CI for each patient, we need many BP measurements. However, it is very difficult to measure BP multiple times for each patient using an oscillometry BP device since repeatable circumstances for reproducible BP measurements cannot be guaranteed [14]. For this reason, it is necessary to calculate CI using a small number of measurements, and as a result, the bootstrap technique was proposed to obtain CI estimates from BP measurements using a small sample size [14].

Soueidan et al. also proposed a new method to augment the noninvasive measurement by providing the mean systolic blood pressure (SBP) and diastolic blood pressure (DBP) with CIs [15]. However, these methods did not satisfy the allowable bias specified by the standard protocol [16]. To address this issue, Lee et al. recently introduced the deep neural network (DNN) estimator [17] to estimate BP measurements. However, this method has many random and initialized parameters in the training procedure, such as weights and bias, which can cause unstable estimates such as large standard deviations of errors. To address the above challenge, Lee et al. also provided a method to obtain accurate BP estimates using the DNN ensemble estimator [18]. Here, we introduce a novel methodology using an ensemble-based recursive methodology (EBRM) to measure uncertainty for oscillometric BP measurements. There are three stages to the methodology: the first is pre-learning to initialize good parameters using the bagging technique [19], where the sample number of artificial features is augmented using the bagging technique to obtain effective ensemble parameters by training each estimator. After that, in the second stage, we fine-tune the parameters using the EBRM that is used to accurately estimate BPs and then measure the uncertainty for the SBP and DBP in the third stage. As far as we know, the proposed method is one of the first studies using the EBRM based on the DNN model to measure the uncertainty for the SBP and DBP. This paper is an expanded version of the paper [14,18] with the following contributions:

- The proposed methodology can measure uncertainty such as CIs, the standard deviation of error, bias, standard uncertainty, and expanded uncertainty for the SBP and DBP.
- We provide the standard uncertainty u , the combined uncertainty u_c , and the expanded uncertainty U and all are computed based on the approaches detailed in GUM [4] using the bias and standard error for artificial features for the SBP and DBP.
- The previous estimated SBP and DBP are also initialized as another input matrix for the EBRM with the DNN model. This is a novel method as the EBRM is different from the conventional AdaBoost technique.
- We execute Lilliefors test to validate that the distribution of the artificial BP features approaches the Gaussian distribution and to identify similarities between the actual data and the artificial data.

The introduced methodology is expressed as follows. First, we acquire the features from the oscillometric waveform (OMW) signals and envelopes after pre-signal processing on the BP signals.

We then create the artificial features from the original features and evaluate the normality of the distribution of all the features. We then constitute the proposed EBRM through the pre-training and a fine-tuning [17]. Based on the EBRM, we estimate the BP values (SBP and DBP) for the individual subjects. Subsequently, to measure the CIs, standard deviation of error, bias, standard uncertainty, and expanded uncertainty for the SBP and DBP. We then confirm the normality of the artificial BPs for individual subjects.

2. Methods

2.1. BP Measurement and Protocol

The study was approved by the institutional ethics committee, and each test subject provided informed consent. The BP data were measured from 85 people who do not have cardiovascular disease, aged 12 to 80 years with 48 men and 37 women. Particularly, the mean age of 85 subjects was 40.4, the standard deviation was 15.2 and there were six people under 19 in all subjects. A wrist-mounted blood pressure device was used to obtain five sets of oscillometric BP measurements using a piezoelectric sensor embedded in a pressure cuff from each subject following the American National Standards Institute (ANSI)/Association for the Advancement of Medical Instrumentation (AAMI) protocol [16,20]. The average value measured by the two trained observers was used as the reference value for SBP and DBP [14]. This process was repeated four more times to generate five sets of BP data for each subject, with a one-minute break between each BP measurement. Each subject comfortably sat on a chair during the measurement, with a BP cuff wrapped around the subject's left wrist and comfortably raised her or his arm, which was raised to heart level. For the reference measurement, an auscultatory BP cuff was worn at the top of the left arm to match the height of the heart. Notably, two nurses applied a blood pressure cuff to the patient's upper arm. Then, the pressure was increased on the cuff to block the brachial artery. Hence, the blood flow generated a Korotkoff signal (KS) that was heard with the aid of a stethoscope. The first KS, which was determined in units of mmHg by a manometer connected to the upper cuff, was used to predict the SBP, while the fifth KS was used to predict the DBP. However, it was impossible to measure the blood pressure of the arm and wrist at the same time because of the occlusion problem of brachial arteries by arm sphygmomanometers. Hence, almost 1.5 min after each pulse wave signal acquired by the wrist BP monitor, two trained nurses concurrently recorded (SBP₁) and (SBP₂) and (DBP₁) and (DBP₂) using a upper arm sphygmomanometer. The first and second nurses concurrently obtained the BP readings. Hence, the results of five sets were given by (SBP_{1i}, SBP_{2i}, |i = 1, ..., 5) and (DBP_{1i}, DBP_{2i}, |i = 1, ..., 5) for each subject, respectively, where subscript 1 and 2 denoted the first and second nurses. Therefore, each of five classic arm sphygmomanometer measurements obtained by the first and second nurses as [(SBP_{1i}, SBP_{2i}), (DBP_{1i}, DBP_{2i}), |i = 1, ..., 5] corresponded with an interval of 1.5 min to each of the five pulse wave signal obtained by the automated oscillometric BP device (OBPD), as (OBPD_i |i = 1, ..., 5) for each subject as shown in Table 1. This interval between traditional arm and wrist measurements was not only as short as possible to minimize natural BP variability over time, but could be long enough for the system to stabilize after the occlusion of arteries during measurements. The approximately 1.5 min interval between arm and wrist measurements was selected by compromise to minimize method errors [12].

Table 1. The process of blood pressure (BP) measurements for one subject.

Sequence (i)	Primary Nurse	Second Nurse	Average	OBPD
1	SBP ₁₁ (=90) and DBP ₁₁ (=60)	SBP ₂₁ (=90) and DBP ₂₁ (=60)	SBP ₁ (=90) and DBP ₁ (=60)	OBPD ₁
2	SBP ₁₂ (=96) and DBP ₁₂ (=64)	SBP ₂₂ (=94) and DBP ₂₂ (=64)	SBP ₂ (=95) and DBP ₂ (=64)	OBPD ₂
3	SBP ₁₃ (=100) and DBP ₁₃ (=64)	SBP ₂₃ (=98) and DBP ₂₃ (=64)	SBP ₃ (=99) and DBP ₃ (=64)	OBPD ₃
4	SBP ₁₄ (=96) and DBP ₁₄ (=70)	SBP ₂₄ (=96) and DBP ₂₄ (=70)	SBP ₄ (=96) and DBP ₄ (=70)	OBPD ₄
5	SBP ₁₅ (=98) and DBP ₁₅ (=66)	SBP ₂₅ (=96) and DBP ₂₅ (=66)	SBP ₅ (=97) and DBP ₅ (=66)	OBPD ₅

2.2. Features Obtained from Oscillometric Signals and Artificial Data Obtained Using Bootstrap Technique

In order to estimate the reference BP value, we removed outliers using a signal processing technique and the effective features of the oscillometric waveform (OMW) signals were extracted [21]. Because the five BP data for individual volunteers represent small amounts of input data for the training process, we used the bootstrap method [22] to increase the amount of data for each volunteer, where this data is referred to artificial data or features in this study. The artificial input data were generated using the bootstrap technique [14,22] to improve estimation accuracy with traditional approaches when the datasets are limited. More details regarding these features can be found in [21].

2.3. Lilliefors Test for Artificial Data

The normality assumption is the key to a majority of standard steps [23]. We thus verify the normality of the artificial feature. The Lilliefors test is executed to evaluate the normality of each artificial feature as well as to correct the Kolmogorov–Smirnov test of goodness of fit for small values at the tails of probability distributions [24]. Here, we assume that \mathbb{D}^* is a probability distribution for an artificial feature $(\mu_1^*, \dots, \mu_B^*)$, where B denotes the size of replication. We measure the homogeneity between the Gaussian distribution hypothesis and the distribution of artificial features [20]. If the Lilliefors test returns a decision value for the null hypothesis then the artificial feature comes from a normal distribution family, against the alternative in which it does not come from such a distribution. The result h denotes one if the test rejects the null hypothesis at the 5% significance level, and zero otherwise. Thus, we accept the null hypothesis as shown in Table 2. It is also noted that all p values in the Lilliefors test are greater than α ($=0.05$) and that null hypothesis is rejected if the Lilliefors test value k is larger than the critical values c . Therefore, we can accept the null hypothesis that the distribution of artificial features converges to the normal distribution [25]. We examine the consistency and convergence for the artificial data [22]. Therefore, we verify that our artificial data are suitable for actual data convergence for sample means based on the theorem [26] in that if $\mathbb{E}(x^2) < \infty$, then $\|\mathbb{D}^*\{\sqrt{n}(\hat{\mu}^* - \hat{\mu}) \leq x\} - \mathbb{D}\{\sqrt{n}(\hat{\mu} - \mu) \leq x\}\|_{\infty} \rightarrow 0$, where \mathbf{X} is the actual feature and $\|\cdot\|_{\infty}$ is $\sup_{x \in \mathbb{R}} |\cdot|$, respectively. It is found that the distribution of $\sqrt{n}(\hat{\mu}^* - \hat{\mu})$ approximates to $\sqrt{n}(\hat{\mu} - \mu)$ [26], where $\beta(\mu) = \mathbb{E}(\hat{\mu}(\mathbf{X}) - \mu)$ denotes a bias and μ is the original feature. When the bias approaches zero, estimates are considered to be unbiased and we can easily compute uncertainties such as the bias and standard error for the artificial features as shown below:

$$\begin{aligned} \beta(\hat{\mu}^*(\cdot)) &= \frac{1}{N} \sum_{b=1}^N \hat{\mu}_b^* - \mathbb{E}(\mu|\mathbf{X}) \\ &\cong \mathbb{E}(\hat{\mu}^*(\mathbf{X}^*) - \hat{\mu}(\mathbf{X})) \end{aligned} \quad (1)$$

$$\hat{\sigma}^* = \sqrt{\left(\frac{1}{N-1} \sum_{b=1}^N (\hat{\mu}_b^* - \hat{\mu}^*(\cdot))^2 \right)}, \quad (2)$$

where $\hat{\mu}^*(\cdot)$ denotes $N^{-1} \sum_{b=1}^N \hat{\mu}_b^*$.

Table 2. The p are obtained from the artificial features utilizing the bootstrap technique $N (=100)$, where N is a number of replication and k , c , and h are obtained from the Lilliefors-test function [24], where TSBP* and TDBP* denote target artificial SBP and DBP values, respectively.

Features/Parameters	p	k	c	h
TSBP*	0.500	0.060	0.089	0
TDBP*	0.470	0.090	0.089	0
MAP	0.500	0.050	0.089	0
AR	0.500	0.049	0.089	0
AE	0.420	0.063	0.089	0
EL	0.368	0.065	0.089	0
MA	0.352	0.065	0.089	0
STD ₁	0.059	0.371	0.089	0
STD ₂	0.063	0.500	0.089	0
MAPL	0.485	0.061	0.089	0

3. Ensemble-Based Recursive Methodology (EBRM) for Measured BP

3.1. EBRM with DNN Regression

Here, we present the EBRM based on the bagging [19] and AdaBoost [27] techniques used with the DNN model [18]. There are three parts: the first part is pre-learning to initialize good parameters using the bagging technique, and the second part is to fine-tune the parameters using the recursive AdaBoost technique. Our input features are given as $(\mathbf{V} \quad \mathbf{T})$, where \mathbf{V} and \mathbf{T} denote the input matrix and the output matrix, respectively. Then, the mean μ and standard error σ are computed for each feature vector. The bootstrap method is then used as a generator to build the distribution for artificial feature, as described in Section 2. Indeed, the Gaussian–Bernoulli restricted Boltzmann machine (GBRBM) is used to connect between a Gaussian input layer and a binary hidden layer because the distribution of artificial feature approaches a normal distribution [17]. The GBRBM [28] is utilized to mitigate the local minima problem and the issue of overfitting through the choice of initial parameters. However, random initialization parameters such as training data sets, weights, and biases at the pre-learning step are sources of uncertainty. Hence, ensemble parameters are used to solve random initialization parameter problems, and they are used in fine-tuned steps. The parameter matrix $\mathbf{W} \in \mathbb{R}^{m \times n}$ is built based on the given input features matrix \mathbf{V} . We implement the bagging technique in order to minimize the standard deviation of the error of the difference between the input data and reconstructed data (i.e., estimated input data) using the Kullback-Leibler divergence [28]. Therefore, the ensemble parameters can be given as $\tilde{\mathbf{W}}_{\varphi} = \sum_{k=1}^K \mathbf{W}_k(\cdot)$, where K denotes the number of ensemble.

On the other hand, we acquire each artificial feature after the adaptive adjustment of the distribution of training data sets using the EBRM in the second step, as presented in Algorithm 1. A weight vector $w_m^{(1)}$ is initialized to be used in the second step. We then create a different training set for each estimator according to the weighted sample, M from the sequence of L with $M \leq L$, where L denotes the number of artificial samples ($L = B \times S$). Note that, in $L = B \times S$, B and S denote number of replication and the number of subjects, respectively. Note that I in Algorithm 1 denotes the number of features, as shown in lines 4–8. In detail, the artificial samples are obtained from different distributions, which are updated repeatedly through relative errors and estimated BPs to be used in the next estimator. The weight of each instance is updated based on an error. In other words, it is more likely that an instance with a large error in the previous distribution exists in the next distribution as shown $(\mathbf{V}^* | w_m^{(k)})$ in line 7. Estimate BPs (SBP and DBP) are also initialized as another input matrix $\hat{\mathbf{T}}_m^{(1)}$ for training EBRM with the DNN model. The estimate BPs are thus concatenated to an artificial block matrix as $[\tilde{\mathbf{V}}^* \quad \hat{\mathbf{T}}_m^{(k)}]$ and are updated recursively as shown in algorithm line 8, where $\hat{\mathbf{T}}_m^{(k)}$ denotes the previous BPs' estimator. This is a novel method as the EBRM is different from the conventional AdaBoost approach [27].

Algorithm 1

```

procedure EBRM( $\mathbf{V}, \mathbf{T}$ )
2:   for  $k \leftarrow 1, K$  do  $w_m^{(1)}$  and  $\hat{\mathbf{T}}_m^{(1)}$ 
   for  $i \leftarrow 1, I$  do
4:     for  $b \leftarrow 1, B$  do  $V_{i,b}^* = (v_1^*, v_2^*, \dots, v_N^*)$  and  $T_{j,b}^* = (t_1^*, t_2^*, \dots, t_N^*)$ ,
        $\bar{V}_{i,b}^* = \frac{1}{N} \sum_{n=1}^N v_n^*$  and  $\bar{T}_{j,b}^* = \frac{1}{N} \sum_{n=1}^N t_n^*$ ,
6:      $\mathbf{V}_i^* = (\bar{V}_{i,1}^*, \bar{V}_{i,2}^*, \dots, \bar{V}_{i,B}^*)$  and  $\mathbf{T}_j^* = (\bar{T}_{j,1}^*, \bar{T}_{j,2}^*, \dots, \bar{T}_{j,B}^*)$ 
        $\tilde{\mathbf{V}}^* = (\mathbf{V}^* | w_m^{(k)})$ 
8:      $\mathbf{U}^* = [\tilde{\mathbf{V}}^* \quad \hat{\mathbf{T}}_m^{(k)}]$ 
   end for
10:  end for
  call learning: back-propagation  $\{\hat{f}^{*k}(\mathbf{U}_m^*, \mathbf{T}_m^*)\}$ ,
12:  output:  $\hat{\mathbf{T}}_m^*, \forall m = 1$  to  $M$ 
   $\varepsilon_{\max} = \max_{m=1, \dots, M} [\hat{\mathbf{T}}_m^* - \mathbf{T}_m^*]^2$ 
14:   $\varepsilon_m = \frac{[\hat{\mathbf{T}}_m^* - \mathbf{T}_m^*]^2}{\varepsilon_{\max}}$ 
   $\bar{\varepsilon} = \sum_{m=1}^M \varepsilon_m w_m^{(k)}$ 
16:   $\beta_k = \frac{1}{1 - \bar{\varepsilon}}$ 
   $w_m^{(k+1)} = w_m^{(k)} \beta_k^{(1 - \varepsilon_m)}$ 
18:   $w_m^{(k+1)} = \frac{w_m^{(k+1)}}{\sum_m w_m^{(k+1)}}$ 
  end for
20: end procedure

```

Back-propagation with scaled conjugate gradients is then executed to optimize the parameters as shown in line 11. In turn, we repeatedly calculate the error between the hypothesis ($\hat{\mathbf{T}}^*$) and reference (\mathbf{T}^*) until the minimum value is reached as expressed in line 12. Here, the estimated ($\hat{\mathbf{T}}^*$) BPs are used as an input feature to train the recursive ensemble estimator. If the error is smaller, the artificial block matrix for the next estimator in EBRM is closer to the reference BPs (\mathbf{T}^*). The mean error is computed as shown in lines 13–15, and the weight parameter β_k is represented as in line 16. Finally, we update the weight vector for instances and normalize them as shown in steps 17–18. If the error values for each instance are very small, the weight parameters also have a small value, respectively, and if the error values for each instance are large in the current iteration, β_k also have large values. The output of the EBRM is given as follows: $\hat{\mathbf{T}}^*(x) = \inf \{t \in \mathbf{T} : \sum_{k: \hat{t}_k \leq t} \log(1/\beta_k) \geq 1/2 \sum_k \log(1/\beta_k)\}$. Here, each k estimator predicts $\hat{t}_k, \forall k = 1, \dots, K$. If β_k is all equal, it will be equivalent to the median. We add up the log until we reach the smallest k so that the inequality can be satisfied [27]. The EBRM with DNN model is utilized as a stable estimator given as $\hat{f}_\varphi^*(\cdot) = \mathcal{D}_{NN}((\mathbf{U}^*, \mathbf{T}^*))(\cdot) : \mathbb{R}^I \rightarrow \mathbb{R}^J$. In this study, $\mathcal{D}_{NN}(\cdot)$ is the DNN model which is used as a weak learner.

3.2. DNN Model [28]

Our deep learning model consisting of two hidden layers is a probability-generating structure [28]. The deep learning model is given by $P(\mathbf{v}^*, \mathbf{h}^1, \mathbf{h}^2, \dots, \mathbf{h}^l) = P(\mathbf{v}^* | \mathbf{h}^1)P(\mathbf{h}^1 | \mathbf{h}^2) \dots P(\mathbf{h}^{l-2} | \mathbf{h}^{l-1})P(\mathbf{h}^{l-1}, \mathbf{h}^l)$, where in $P(\mathbf{h}^i | \mathbf{h}^{i+1})$, \mathbf{h}^i denotes the hidden units at layer i , and \mathbf{v}^* is the re-sampled input data (artificial data). We can write a probability as $P(\mathbf{v}^*, \mathbf{h}) = \frac{1}{U} \exp^{-\mathbf{h}' \mathbf{W} \mathbf{v}^* - \mathbf{c}' \mathbf{v}^* - \mathbf{b}' \mathbf{h}}$, where U denotes the normalized constant value, \mathbf{c} and \mathbf{b} are the bias for the input data and the bias for the hidden units, respectively, and \mathbf{W} denotes the weight parameter. We thus rewrite the conditional layer as $P(\mathbf{h}^i | \mathbf{h}^{i+1}) = \text{sgm}\left(-\mathbf{c} - \sum_{k=1}^{n^{i+1}} \mathbf{W} \mathbf{h}_k^{i+1}\right)$, where $\text{sgm}(x)$ denotes a sigmoid function. Looking at the data learning process, we use pre-learning to make the initial values for weights and biases efficiently, and use it as an effective point to fine-tune [29]. Thus, the cost function [17,30] can be obtained as

$$\mathbb{L}(\tilde{\mathbf{W}}_\varphi, \mathbf{c}) = \frac{1}{C} \sum_{c=1}^C \sum_{d=1}^D \left(\hat{\mathbf{T}}_c^{*d}(\tilde{\mathbf{W}}, \mathbf{c}) - \mathbf{T}_c^{*d} \right)^2 \quad (3)$$

where \mathbf{T}_c^{*d} represents the d^{th} target BP data (referred to as TSBP* and TDBP*) at the sample index c ; D and C denote the size of data and the size of batch, respectively. Next, we iteratively renew the parameters as

$$\left(\widetilde{\mathbf{W}}_{\varphi^{(c+1)}}^i, \mathbf{c}_{c+1}^i\right) = -\zeta \frac{\partial \mathbb{L}}{\partial (\widetilde{\mathbf{W}}_{\varphi^{(c)}}^i, \mathbf{c}_c^i)} + \eta (\widetilde{\mathbf{W}}_{\varphi^{(c)}}^i, \mathbf{b}_c^i), 1 \leq i \leq K + 1, \quad (4)$$

where ζ , η , K and $K + 1$ denote a learning rate, a momentum parameter, the number of hidden layers and the output layer, respectively.

4. Uncertainty Estimation

4.1. Measurement Uncertainty

The quantity that we intend to measure is called measurand [11]. The aim of a measurement is to acquire the true value of the measurand. We do not know exactly how close the measured BP value is to the true BP value. Therefore, our estimates always have some uncertainty associated with it. The difference between the measured BP value and the true BP value being measured is called an error. The errors of BP can be thought of as consisting of two parts: BP random errors and BP systematic errors. We do not know BP error since true BP value is not known. Therefore, it is impossible to indeed use the quality characteristics of the BP measurement results. The quality and accuracy of the BP measurement results are characterized by uncertainty of BP measurement, defining the interval around the measured BP value, where true BP value exists with some probability. Uncertainty of BP measurement U itself is half-width of that interval and is always positive [11]. This uncertainty is considered to be as the estimate that is the highest probable absolute difference between the measured BP value and the true BP value.

Random errors result in differences among repeated BP measurement results. However, the more repeated BP measurements are made, the less likely the mean value is to be affected by the random errors. Thus, the influence of random errors can be reduced by increasing the number of repetitions in BP measurements. On the other hand, systematic errors cause deviations in the same direction by the same size of all BP measurements in the series. Increasing the number of iterations does not reduce the influence regarding system errors such as bias [11]. One of the most common ways to improve the measurement reliability is to make the same amount of repeated BP measurements. We can perform basic statistical calculations to increase the number of information obtained from BP measurements by taking multiple readings. The arithmetic mean value can be used as an estimate for the true BP value. Since we have different results when making repeated BP measurements, we seek to know the width of the range of BP measurements. The spread of results informs us regarding the uncertainty of a BP measurement and standard deviation where the standard deviation is the basis of defining standard uncertainty denoted by u . The standard uncertainty is calculated as $u = \hat{\sigma} / \sqrt{n}$. In general, if an uncertainty estimate is acquired from the standard deviation of the repeated BP measurement results, it is referred to as Type A uncertainty estimate. All uncertainty estimates acquired without repeated BP measurement are called type B uncertainty estimates using assumed probability distributions, where the assumption can be made through either experience or information [11]. As mentioned above, since we do not know the true BP value, we need to know the reference value (BP obtained by the trained observers) for improving accuracy. This performance characteristics can be quantitatively expressed. Bias is the difference between the measured BP value obtained from multiple repeated BP measurements with the same sample and the reference value which is the considered the quantitative expression of the true value. These two values are combined into BP measurement uncertainty estimates and are considered to be a quantitative representation of accuracy.

We offer a combined measurement uncertainty for BPs through the bias, standard deviation etc. for the BP measurements. The combined standard uncertainty $u_c = \sqrt{u_\alpha^2 + u_\beta^2 + u_\gamma^2}$, where u_α is the

standard uncertainty as a random error, u_β is the bias as a system error, and u_γ is the maximum permissible error (against a mercury sphygmomanometer as a system error about ± 1 mmHg [31]) which are the sources of uncertainty. We also provide a CI for expressing and evaluating uncertainty. The CI is a type of interval estimate, calculated from the statistics of the BP measurements, which might include the true BP value of an unknown population parameter. The CI is used as expanded uncertainty, i.e., $U = K \times u_c$, so that the CI of the measurand is acquired as $\bar{x} \pm U$ [4], where \bar{x} denotes the mean of the measurand. If the measurand distribution close to converges toward a Gaussian distribution, the arithmetic mean value is given by \bar{x} , and the standard uncertainty is provided by standard deviation σ of this distribution. If $K = 2$, then the CI is $\bar{x} \pm 2\sigma$, and the level of confidence climbs up to 95%.

4.2. CI Estimation Using the Bootstrap

The basic concept of this method is that it can utilize the uncertainty ranges of each BP measurement value to compute the maximum and minimum values for the CI. Thus, we provide the CI of the five BP estimates for each patient obtained from the EBRM algorithm, and explain the bootstrap principle of parameter estimation approach. The idea is to resample blood pressure hypotheses to produce many artificial blood pressure hypotheses, $\hat{\mathbf{T}}^* = (\hat{t}_1^*, \dots, \hat{t}_n^*)$, based on n estimates obtained from an unknown distribution $\mathbb{D}(\mu, \sigma)$ to compute a CI for $\hat{\mu}(\mathbf{T}^*)$. Here, $[\hat{\mu}, \hat{\sigma}]$ denotes the maximum likelihood estimate obtained using $\hat{\mathbf{T}} = (\hat{t}_1, \dots, \hat{t}_n)$. Thus, when $n \rightarrow \infty$, we obtain a normal distribution given as $\hat{\mathbb{D}}(\hat{\mu}^*, \hat{\sigma}^* | \hat{\mathbf{T}}^*) \cong \mathcal{N}(\mu, \sigma)$. In our work, we measure the CIs utilizing the bootstrap technique [14,22] which can be obtained using the BP estimates of the EBRM. We then obtain a matrix as follows:

$$\mathbf{M}^*(i | \hat{\mathbf{T}}_i^{S*}) = \begin{bmatrix} t_{1,1}^{*i} & \cdots & t_{1,B}^{*i} \\ \vdots & \ddots & \vdots \\ t_{n,1}^{*i} & \cdots & t_{n,B}^{*i} \end{bmatrix}, \quad (5)$$

where Equation (5) is acquired as $\hat{\mu}_i^* + \hat{\sigma}_i^* \times \text{RANDN}(n, B)$, we then vertically compute each column to obtain the average of each column as $\hat{\mu}_b^{S*} = 1/n \sum_{j=1}^n t_{j,b}^{*i}$, where, s denotes SBP, and $*$ indicates the resampled data obtained from the bootstrap technique. We, hence, do ascending sorts and the sorted BP estimate are given by $\hat{\mathbf{E}}^{S*} = (\hat{\mu}_1^{S*}, \hat{\mu}_2^{S*}, \dots, \hat{\mu}_B^{S*})$, assuming $\hat{\mu}_\alpha^{S*}$ is the 100α -th percentile of B bootstrap replications $(\hat{\mu}_1^{S*}, \hat{\mu}_2^{S*}, \dots, \hat{\mu}_B^{S*})$. We acquire the CI as $\hat{\mu}_{\text{lower}}^{S*}, \hat{\mu}_{\text{upper}}^{S*}$ of the $1 - 2 \cdot \alpha$, from this bootstrap technique as $(\hat{\mu}_\alpha^{S*}, \hat{\mu}_{1-\alpha}^{S*})$. Similar process is used to estimate the CI for DBP.

4.3. CI Estimation with the Monte Carlo Technique

In this work, we assume the expected parameters $\mathbb{E}(\hat{\mu}_i^*, \hat{\sigma}_i^* | \hat{\mathbf{T}}_i^*)$ using a random variable based on the Monte Carlo technique [32]. Hence, we can create $\{\hat{\mathbf{T}}_1^*, \dots, \hat{\mathbf{T}}_n^*\}$ to be independent and identically distributed (IID) random variables from the distribution of $\hat{\mathbf{T}}^*$ using the estimated BP results using the EBRM and acquired their mean and standard deviation. Hence, we express a variance of $\hat{\mu}_i^*$ as $\mathbb{E}[(\hat{\mu}_i^* - \mu)^2] = \frac{\sigma^{*2}}{n}$. The Monte-Carlo technique is a way to achieve an approximate error using sample values as $\frac{\sigma^{*2}}{n}$. However, we do not know σ^2 exactly. Hence, we commonly use estimates of σ^2 given as

$$\hat{\sigma}^{*2} = \frac{1}{n-1} \sum_{i=1}^n (\hat{\mathbf{T}}_i^* - \hat{\mu}_i^*)^2. \quad (6)$$

A standard Gaussian distribution with a zero mean and unit variance has a probability density function as shown below:

$$\rho(z) = \frac{1}{\sqrt{2\pi}} \exp^{-\frac{1}{2}z^2}. \quad (7)$$

Hence, a cumulative distribution function is calculated as

$$\Phi(\alpha) = \int_{-\infty}^{\alpha} \rho(z) dz, \quad -\infty < z < \infty. \quad (8)$$

Based on [32], if $\hat{\mathbf{T}}_i^*$ is an IID with mean μ^* and variance $\sigma^{*2} > 0$, then for all $z \in \mathbb{R}$

$$\mathbb{P}\left(\sqrt{n} \frac{\hat{\mu}_i^* - \mu}{\hat{\sigma}^*} \leq z\right) \rightarrow \Phi(z). \quad (9)$$

We then acquire the CI based on the theorem in [32] for μ but it requires that we know σ^* . $\mathbb{P}(|\hat{\sigma}^* - \hat{\sigma}| > \epsilon)$ converges to zero for any $\epsilon > 0$. Hence, we can substitute s for $\hat{\sigma}$ where s is $\hat{\sigma}^*$. We then rewrite Equation (9) as

$$\begin{aligned} \mathbb{P}\left(|\hat{\mu}_i^* - \mu| \geq \frac{\lambda s}{\sqrt{n}}\right) &= \mathbb{P}\left(\sqrt{n} \frac{\hat{\mu}_i^* - \mu}{s} \leq -\lambda\right) + \mathbb{P}\left(\sqrt{n} \frac{\hat{\mu}_i^* - \mu}{s} \geq \lambda\right) \\ &\rightarrow \Phi(-\lambda) + (1 - \Phi(-\lambda)) = 2\Phi(-\lambda). \end{aligned} \quad (10)$$

We can find a 5 % chance of non-coverage for a 95 % CI, and hence set $2\Phi(-\lambda) = 0.05$, where $\lambda > 0$. Therefore, we can denote as $\lambda = -\Phi^{-1}(0.025) = \Phi^{-1}(0.975)$ as

$$\hat{\mu}_i^* - \Phi^{-1}(0.975) \frac{s}{\sqrt{n}} \leq \mu \leq \hat{\mu}_i^* + \Phi^{-1}(0.975) \frac{s}{\sqrt{n}}, \quad (11)$$

where $\Phi^{-1}(0.975)$ is 1.96. Thus, we acquire the CIs based on the estimated BP results using the EBRM.

5. Experimental Results

The BP measurements of volunteers were sequentially separated into training data and test data. Five sets of 300 BP measurements each of which taken from 60 volunteers were used as training data, and 125 measurements from 25 volunteers were used as validation data. This procedure was repeated so that each volunteer was included only once in the test process. In Table 3, an example result is shown in order to differentiate the artificial features from the original features. We summarized the parameters for the EBRM algorithm as in Table 4. As required by the ANSI/AAMI BP measurement protocol [16], the EBRM algorithm was evaluated to verify that the mean error (ME) is less than ± 5 mmHg and that the standard deviation of error (SDE) is less than 8 mmHg, as shown in Table 5. Moreover, in accordance with the British Hypertension Society (BHS) protocol [1], the EBRM results were compared to the conventional methods, and the mean absolute error was evaluated for three groups of less than 5 mmHg, less than 10 mmHg and less than 15 mmHg, respectively (Table 5). If 60% of the mean absolute error of a BP measurement method is within 5 mmHg, 85% within 10 mmHg and 95% within 15 mmHg, the method is classified as class A. In Table 5, we provided the results of EBRM with respect to the uncertainty. Table 6 and Figure 1 show the results of statistical analysis through the Lilliefors function to evaluate the normality in terms of the artificial BP estimates for each subject.

Table 3. An exemplary result (one subject) is represented to verify between the artificial features and original features for consistency and convergence [22].

Features/Parameters	$\hat{\mu}$	$\hat{\mu}^*$	CI_L^*	CI_U^*	$\hat{\sigma}$	$\hat{\sigma}^*$	$\beta(\hat{\mu}^*(\cdot))$	u	U
TSBP	93.20	93.35	91.56	95.84	2.39	1.03	0.152	0.103	± 0.367
TDBP	59.80	59.90	57.69	61.93	2.17	0.95	0.101	0.095	± 0.278
MAP	0.311	0.312	0.253	0.365	0.057	0.027	0.001	0.0027	± 0.0057
AR	0.494	0.496	0.451	0.533	0.045	0.020	0.002	0.002	± 0.0052
AE	0.065	0.066	0.057	0.077	0.012	0.005	0.001	0.0005	± 0.0022
EL	0.236	0.236	0.231	0.242	0.006	0.002	0.000	0.0002	± 0.0005
MA	0.166	0.165	0.141	0.194	0.026	0.011	-0.001	0.001	± 0.002
STD ₁	0.150	0.151	0.101	0.204	0.054	0.025	0.001	0.003	± 0.03
STD ₂	0.184	0.183	0.133	0.228	0.047	0.022	-0.001	0.002	± 0.005
MAPL	0.391	0.390	0.360	0.416	0.031	0.013	-0.001	0.001	± 0.003

Table 4. Summarized parameters [17,28] of the ensemble-based recursive methodology (EBRM) algorithm, where 12 denotes the dimension of input vector, 2 is the number of output units (namely, the target vector [SBP and DBP] dimensions), and 32 is the number of hidden unit.

Number of the Units:	[(12,(32),(32), (32), 2)]
Dimension of feature	12
Dimension of target	2
Number of hidden layers	3
Number of hidden unit on the layers	32
Number of sample over original feature	5
Number of sample over each artificial feature	100
Number of epoch in the pre-training	10 to 50
Number of epoch in the fine-tuning	10 to 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	logistic function
Initial weights and biases	randomly between (-1, 1)

Table 5. Evaluating of the EBRM algorithm through the British Hypertension Society (BHS) and Association for the Advancement of Medical Instrumentation (AAMI) protocols utilizing the results of maximum amplitude algorithm (MAA), neural network (NN), support vector regression (SVR) [33], deep neural network (DNN), DNN _{ϕ} and EBRM on ($5 \times 85 = 425$) measurements, where DNN _{ϕ} denotes the Adaboost with DNN model.

Methods	SBP			DBP			SBP/DBP BHS Grade	SBP AAMI ME(SDE)	DBP ME(SDE)
	Mean Absolute Difference (%)			Mean Absolute Difference (%)					
	≤ 5 mmHg	≤ 10 mmHg	≤ 15 mmHg	≤ 5 mmHg	≤ 10 mmHg	≤ 15 mmHg			
MAA	47.06	85.88	96.47	56.47	88.24	97.65	C/B	0.07 (9.28)	-0.89 (7.76)
NN	53.88	85.65	95.53	66.12	94.12	98.82	B/A	0.25 (7.48)	-0.22 (6.80)
SVR	62.59	86.12	95.53	74.12	93.65	96.94	A/A	0.10 (7.15)	-0.34 (6.45)
DNN	69.18	88.71	95.53	76.24	93.17	98.12	A/A	0.02 (6.44)	0.11 (5.24)
DNN _{ϕ}	71.06	90.82	95.53	81.18	96.24	99.29	A/A	-0.05 (5.72)	0.05 (4.70)
EBRM	73.65	93.88	96.94	83.06	97.17	99.76	A/A	0.02 (5.50)	0.18 (4.59)

Table 6. Comparison of CIs between the proposed EBRM and conventional methods, where $n (=85)$ denotes the number of subject, L and U denote the lower and upper limits, respectively.

BP (mmHg) $n (=85)$	SBP (SDE) 95%CI	DBP (SDE) 95%CI	SBP L (SDE)	SBP U (SDE)	DBP L (SDE)	DBP U (SDE)
MAA _{ST} [14]	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} [14]	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} [14]	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)
DNN _{φ(BOOT)}	4.8 (1.5)	4.2 (0.9)	107.3 (12.7)	112.1 (12.8)	65.1 (8.2)	69.3 (8.8)
EBRM _{BOOT}	3.1 (2.9)	3.2 (2.7)	107.9 (13.9)	111.0 (13.4)	65.5 (9.4)	68.7 (9.0)
EBRM _{MC}	1.4 (0.4)	1.2 (0.4)	107.8 (12.8)	109.2 (13.4)	65.0 (9.2)	66.3 (9.4)
EBRM _U	6.6 (2.7)	6.8 (3.3)	105.7 (12.8)	112.3 (13.4)	63.8 (9.3)	70.6 (9.3)

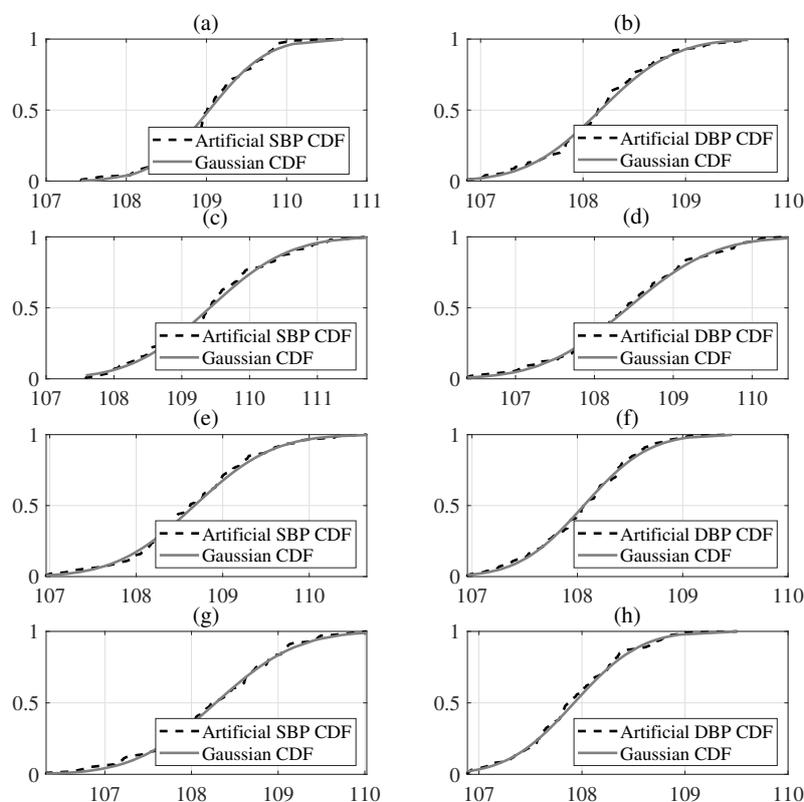


Figure 1. These figures denote the cumulative distribution function (CDF) of the selected artificial BP estimates from the parameter bootstrap approach with ($N = 100$) replicas based on the EBRM results, where x -axis denotes mmHg and y -axis denotes cumulative probability. Note that the plots (a) and (b) are the examples acquired from 5th subject, the plots (c) and (d) are the examples acquired from 6th subject, the plots (e) and (f) are the examples acquired from 7th subject, and the plots (g) and (h) are the examples acquired from 8th subject.

6. Discussion

We confirmed that the uncertainty such as the bias for artificial data was extremely small, and that the standard error $\hat{\sigma}^*$, obtained from bootstrap technique, was smaller than $\hat{\sigma}$; hence, the bootstrap technique could be used as an efficient method to increase the number of samples of features. As a result, artificial features could be identified as being very close to actual features, and in addition, we found that the CI of artificial features subsumes the CI of all artificial and true features. We provided the standard uncertainty $u = \hat{\sigma}^* / \sqrt{B(=100)}$, the combined uncertainty $u_c = \sqrt{\beta^2 + u^2}$, and the

expanded uncertainty $U = ku_c$ and all were easily calculated based on the approaches detailed in GUM [4] using the bias and standard error for artificial features (Table 3). The u and U in terms of the artificial features corresponded to very small values, which implies that there is little difference between artificial data and actual data.

Based on the BHS protocol [1] and as shown in Table 5, the mean absolute errors of the EBRM were 73.65% (≤ 5 mmHg), 93.88 % (≤ 10 mmHg), and 96.94 % (≤ 15 mmHg), respectively, for SBP and 83.06% (≤ 5 mmHg), 97.17 % (≤ 10 mmHg), and 99.76 % (≤ 15 mmHg), respectively, for DBP. Therefore, the EBRM obtained Class A designation for the measurement of SBP and DBP. We found that EBRM estimates are more accurate than BP estimates obtained using conventional methods. The accuracy of estimates obtained using EBRM was calculated by comparing them with estimates obtained using the stethoscope method in accordance with the ANSI/AAMI protocol [16]. The SDE of a device is generally considered more significant because unless there is a systematic bias, then MEs would be small even in the presence of large positive and negative errors. The proposed method met the ANSI/AAMI criteria and provided more accurate BP estimates compared with the conventional methods, as shown in Table 5. The SDE acquired using the EBRM was found to be 5.50 mmHg and 4.59 mmHg for the SBP and DBP, respectively. These results indicate superior performance compared to conventional algorithms. Therefore, we can conclude that the EBRM decreases the uncertainty related to measures such as the SDE of ME and increases the performance reliability. Moreover, the CIs for SBP and DBP obtained with EBRM_{BOOT} were smaller than CIs obtained using the conventional methods, as shown in Table 6. We confirmed the difference between 1.7 mmHg and 1.0 mmHg of the EBRM and the conventional method [14] in the SDEs of CI for SBP and DBP although the CI results acquired from the EBRM are larger than those acquired from the DNN $_{\phi}$. The CIs in terms of the artificial BP values corresponded to very small values, which indicates that through the EBRM we successfully decreased the uncertainty with respect to a random error such as SDE and systematic error such as bias. We also provided the expanded uncertainty for the EBRM_U where EBRM_U was computed using the bias and standard error for artificial BP values according to the GUM recommendations [4] as shown in Table 6. We also obtained the CIs EBRM_{MC} from the Monte Carlo technique based on the results of the EBRM as shown in Table 6. Specifically, the results of the CIs were smaller than those of the CIs obtained from the conventional methods.

The Lilliefors test is executed to verify the normality of each distribution, indicating that these distributions are very similar to Gaussian distributions [23]. We thus conducted the Lilliefors, correlation, kurtosis, and skewness tests based on the results of the EBRM in order to verify the normality for the individual BP measurements. As the number of B ($=100$) bootstrap replicas increased, the distribution of artificial BP measurements converged towards the Gaussian distribution. Hence, it was confirmed that the distribution of artificial BP measurements approaches the Gaussian distribution, as shown in Figure 1. The hypothesis for the average artificial BP measurement confirmed that h ($=0.07$) and h ($=0.05$) were almost zero for SBP and DBP estimates, which means that the null hypothesis was accepted at the 5% significance level. In addition, the null hypothesis could not be rejected because the k ($=0.02$) value was less than the threshold cv ($=0.29$) for SBP and DBP. Because our p ($=0.33$) and p ($=0.36$) test values were greater than 5% ($=0.05$) significant level for SBP and DBP, we affirmed the normality of the distribution of artificial BP measurements for each subject. In addition, kurtosis is also a measure of the population that determines how flat or peaked the probability distribution is compared to the Gaussian distribution, where the kurtosis of Gaussian distribution is 3. Thus, if the kurtosis value is greater than 3, distribution is heavy-tailed, and if the kurtosis value less than 3 the distribution is light-tailed. The kurtosis values for the artificial SBP and DBP measurements were 2.97 and 3.00, respectively, which indicates that the distributions of the artificial BP measurements were almost Gaussian. The skewness of the population is a measure of the horizontal symmetry for the distribution. Here, negative values show that the distribution is skewed to the left and positive values to the right whereas skewness of the Gaussian distribution is zero. We found that the symmetric distributions of the artificial BP data were 0.02 and -0.01 , respectively,

for the SBP and DBP, which means that they were nearly converging on the Gaussian distribution, as shown in Table 7.

Table 7. Summary of the Lilliefors and Normality tests for SBP and DBP (85 subjects).

Tests	Lilliefors Test				Normality Test		
	α (=0.05)	h (std)	p (std)	k (std)	cv (std)	kurtosis (std)	skewness (std)
SBP	0.07 (0.27)	0.33 (0.18)	0.02 (0.005)	0.29 (0.00)	2.97 (0.17)	0.02 (0.08)	
DBP	0.05 (0.21)	0.36 (0.16)	0.02 (0.005)	0.29 (0.00)	3.00 (0.18)	−0.01 (0.08)	

Although we experimented with the BP data of 85 subjects based on the ANSI/AAMI protocol [31], this study was limited due to the small number of samples with relatively a small amount of participants. However, this limitation was resolved using artificial samples. Second, simultaneous measurements at the brachial and wrist locations were impossible because of the occlusion of the brachial artery by the upper arm cuff. Therefore, there was approximately 1.5 minutes' separation between the BP measurements obtained using the automatic wrist BP monitor, and the reference measurements received simultaneously by the two trained nurses using the classic upper arm sphygmomanometer. This delay contributes to the measurement error because of the natural variation in SBP and DBP over even short periods. Third, although there are other measures of uncertainty for BP estimation, we focused on the random error (SDE) and systematic errors (i.e., ME, bias, and calibration errors) because we could only evaluate these errors in this study. We represented that CIs were short, but did not compensate for errors arising from the current setting of the experiment using oscillometric BP measurement.

7. Conclusions

In this paper, we propose a novel method using EBRM to measure uncertainty such as CIs, the standard deviation of error, bias, standard uncertainty, and expanded uncertainty for the SBP and DBP. We verify that the distribution of the artificial BP data is close to the Gaussian distribution, and identify similarities between the real data and the artificial data. The Lilliefors test is performed to investigate the normality of the artificial BP measurements for each subject. The main contribution of this work is that the accuracy and the stability of the blood pressure estimates are improved using the EBRM algorithm. We will carry out additional non-normality testing with a new subject population in the near future.

Author Contributions: Data curation, V.Z.G.; formal analysis, S.L.; funding acquisition, G.L.; investigation, S.L.; methodology, S.L.; resources, S.L.; writing—original draft, S.L.; writing—review and editing, H.R.D. and S.R. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation (NRF) of Korea (2016R1D1A1B03932925).

Conflicts of Interest: The authors declare no conflict of interest.

References

- O'Brien, E.; Petrie, J.; Littler, W.A.; De Swiet, M.; Padfield, P.L.; Altman, D.; Bland, M.; Coats, A.; Atkins, N. The British hypertension society protocol for the evaluation of blood pressure measuring devices. *J. Hypertens.* **1993**, *11*, S43–S63.
- Hansen, S.; Staber, M. Oscillometric blood pressure measurement used for calibration of the arterial tonometry method contributes significantly to error. *Eur. J. Anaesthesiol.* **2006**, *23*, 781–787. [[CrossRef](#)] [[PubMed](#)]
- Dieterle, T.; Battagay, E.; Bucheli, B.; Martina, B. Accuracy and 'range of uncertainty' of oscillometric blood pressure monitors around the upper arm and the wrist. *Blood Press Monit.* **1998**, *3*, 339–346. [[PubMed](#)]
- BIPM; IEC; IFCC; ILAC; ISO; IUPAC; IUPAP; OIML. *Guide to the Expression of Uncertainty in Measurement (GUM)*; International Standards Organization: Geneva, Switzerland, 1995.

5. Kachuee, M.; Kiani, M.M.; Mohammadzade, H.; Shabany, M. Cuffless blood pressure estimation algorithms for continuous health-care monitoring. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 859–869. [[CrossRef](#)]
6. Chen, S.; Ji, Z.; Wu, H.; Xu, Y. A non-invasive continuous blood pressure estimation approach based on machine learning. *Sensors* **2019**, *19*, 2585. [[CrossRef](#)]
7. Tjahjadi, H.; Ramli, K.; Murfi, H. Noninvasive classification of blood pressure based on photoplethysmography signals using bidirectional long short-term memory and time-frequency analysis. *IEEE Access* **2020**, *8*, 20735–20748. [[CrossRef](#)]
8. Lee, S.; Lee, G.; Jeon, G. Statistical approaches based on deep learning regression for verification of normality of blood pressure estimates. *Sensors* **2019**, *19*, 2137. [[CrossRef](#)]
9. Lee, S.; Chang, J.-H. Dempster–Shafer Fusion Based on a Deep Boltzmann Machine for Blood Pressure Estimation. *Appl. Sci.* **2019**, *9*, 96. [[CrossRef](#)]
10. Ferrero, A.; Salicone, S. Measurement uncertainty. *IEEE Instrum. Meas. Mag.* **2006**, *9*, 44–51. [[CrossRef](#)]
11. Estimation of Measurement Uncertainty in Chemical Analysis. Available online: <https://sisu.ut.ee/measurement/uncertainty> (accessed on 10 January 2020).
12. Karagöz, İ.; Cecelioğlu, S. The analysis of different approaches related to the measurement of uncertainty in biomedical calibration. *Gazi Univ. J. Sci.* **2007**, *20*, 61–67.
13. Parvis, M.; Vallan, A. Medical measurements and uncertainties. *IEEE Instrum. Meas. Mag.* **2002**, *5*, 12–17. [[CrossRef](#)]
14. Lee, S.; Bolic, M.; Groza, V.; Dajani, H.; Rajan, S. Confidence interval estimation for oscillometric blood pressure measurements using bootstrap approach. *IEEE Trans. Instrum. Meas.* **2011**, *60*, 3405–3415. [[CrossRef](#)]
15. Soueidan, K.; Chen, S.; Dajani, H.; Bolic, M.; Groza, V. Augmented blood pressure measurement through the noninvasive estimation of physiological arterial pressure variability. *Physiol. Meas.* **2012**, *33*, 881–899. [[CrossRef](#)] [[PubMed](#)]
16. AAMI (Association for the Advancement of Medical Instrumentation); ANSI (American National Standards Institute). *Manual, Electronic or Automated Sphygmomanometers*; AAMI: Arlington, VA, USA, 2003; 1–70.
17. Lee, S.; Chang, J.-H. Oscillometric Blood pressure estimation based on deep learning. *IEEE Trans. Ind. Informat.* **2017**, *13*, 461–472. [[CrossRef](#)]
18. Lee, S.; Chang, J.-H. Deep belief networks ensemble for blood pressure estimation. *IEEE Access* **2017**, *5*, 9962–9972. [[CrossRef](#)]
19. Buhlmann, P.; Yu, B. Analyzing bagging. *Ann. Stat.* **2002**, *30*, 927–961. [[CrossRef](#)]
20. Lee, S.; Rajan, S.; Park, C.H.; Chang, J.H.; Dajani, H.; Groza, V. Estimated confidence interval from single blood pressure measurement based on algorithm fusion. *Comput. Biol. Med.* **2015**, *62*, 154–163. [[CrossRef](#)]
21. Lee, S.; Park, C.H.; Chang, J.H. Improved Gaussian mixture regression based on pseudo feature generation using bootstrap in blood pressure measurement. *IEEE Trans. Ind. Informat.* **2016**, *2*, 2269–2280. [[CrossRef](#)]
22. Efron, B.; Tibshirani, R. Bootstrap methods for standard errors, confidence interval, and other measures of statistical accuracy. *Stat. Sci.* **1986**, *1*, 54–77. [[CrossRef](#)]
23. Abdi, H.; Molin, P. Lilliefors/Van Soest’s Test of Normality. Available online: <https://www.utdallas.edu/~herve/Abdi-Lillie2007-pretty.pdf> (accessed on 10 January 2020).
24. Dallal, G.E. An analytic approximation to the distribution of Lilliefors’s test statistic for normality. *Ame. Stat.* **1986**, *40*, 294–296.
25. Hollander, M.; Wolfe, D.A. *Nonparametric Statistical Methods*; Wiley: New York, NY, USA, 1999.
26. Singh, K. On the asymptotic accuracy of Efron’s bootstrap. *Ann. Stat.* **1981**, *9*, 1187–1195. [[CrossRef](#)]
27. Freund, Y.; Schapire, R.E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [[CrossRef](#)]
28. Hinton, G.; Osindero, S.; Teh, Y.W. A fast learning algorithm for deep belief nets. *Neural Comput.* **2006**, *18*, 1527–1554. [[CrossRef](#)] [[PubMed](#)]
29. Bengio, Y. Learning deep architectures for AI. *Found. Trends Mach. Learn.* **2009**, *2*, 1–127. [[CrossRef](#)]
30. Moller, M. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Netw.* **1993**, *6*, 525–533. [[CrossRef](#)]
31. Stergiou, G.S.; Alpert, B.; Mieke, S.; Asmar, R.; Atkins, N.; Eckert, S.; Frick, G.; Friedman, B.; Graßl, T.; Ichikawa, T.; et al. A Universal standard for the validation of blood pressure measuring devices. *Hypertension* **2018**, *71*, 368–374. [[CrossRef](#)]

32. Owen, A.B. Monte Carlo Theory, Methods and Examples. Available online: <https://statweb.stanford.edu/~owen/mc/> (accessed on 10 January 2020).
33. Rakotomamonjy, A. Analysis of SVM regression bound for variable ranking. *Neurocomputing* **2007**, *70*, 1489–1491. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).