










Article

Prediction of Blood Pressure after Induction of Anesthesia Using Deep Learning: A Feasibility Study

Young-Seob Jeong ^{1,†}, Ah Reum Kang ^{1,‡}, Woohyun Jung ², So Jeong Lee ²,
Seunghyeon Lee ², Misoon Lee ², Yang Hoon Chung ², Bon Sung Koo ² and
Sang Hyun Kim ^{2,*}

¹ SCH Media Labs, Soonchunhyang University, Asan 31538, Korea; bytecell@sch.ac.kr (Y.-S.J.); armk@arkang.net (A.R.K.)

² Department of Anesthesiology and Pain Medicine, Soonchunhyang University Bucheon Hospital, Soonchunhyang University College of Medicine, Bucheon 420767, Korea; mulungtomato@naver.com (W.J.); sojung228@naver.com (S.J.L.); t0470@schmc.ac.kr (S.L.); misoonlee@schmc.ac.kr (M.L.); drcyh79@schmc.ac.kr (Y.H.C.); kbs0803@schmc.ac.kr (B.S.K.)

* Correspondence: skim@schmc.ac.kr; Tel.: +82-32-621-5328; Fax: +82-32-621-5322

† Current address: SCH Media Labs, Soonchunhyang University, 22 Soonchunhyang-ro, Sinchang-myeon, Asan 31538, Korea.

‡ These authors contributed equally to this work.

Received: 9 October 2019; Accepted: 22 November 2019; Published: 27 November 2019



Abstract: Anesthesia induction is associated with frequent blood pressure fluctuation such as hypotension and hypertension. If it is possible to precisely predict blood pressure a few minutes ahead, anesthesiologists can proactively give anesthetic management before patients develop hemodynamic problem. The objective of this study is to develop a real-time model for predicting 3-min-ahead blood pressure from the start of anesthesia induction to surgical incision. We used only vital signs and anesthesia-related data obtained during anesthesia-induction phase and designed a bidirectional recurrent neural network followed by fully connected layers. We conducted experiments on our collected data of 102 patients, and obtained mean absolute errors between 8.2 mmHg and 11.1 mmHg and standard deviation between 8.7 mmHg and 12.7 mmHg. The average elapsed time for prediction of a batch of 100 unseen data was about 26.56 milliseconds. We believe that this study shows feasibility of real-time prediction of future blood pressures, and the performance will be improved by collecting more data and finding better model structures.

Keywords: blood pressure prediction; deep learning; recurrent neural network; real-time sequence prediction

1. Introduction

General anesthesia for surgery can be divided into three phases (i.e., induction, maintenance and emergence from anesthesia). Especially in the anesthesia-induction phase, the changes in blood pressure is rapid and can range variously from hypotension to hypertension. This usually caused not only by the administration of intravenous anesthetic agents (propofol and remifentanyl), volatile anesthetic agents, neuromuscular blocking agents but also by airway manipulation to intubate patient's trachea for the mechanical ventilation. It has been reported that hypotension, even a short duration of mean arterial pressure less than 55 mmHg, is associated with acute kidney injury and myocardial injury [1]. On the other hand, hypertension, if left untreated, increases risks of bleeding, cerebrovascular events, and myocardial infarction postoperatively [2]. If blood pressure can be accurately predicted, anesthesiologist may proactively search for possible causes to prevent severe hemodynamic changes.

This may enable early intervention such as adjustment of anesthetic agents, fluids, and vasoactive drugs. Thus, patient may not experience harmful consequences caused by hypotension or hypertension as such events can be prevented ahead. However, intraoperative accurate blood pressure predictions is nearly impossible because of complex mechanisms that cause blood pressure changes, and at least similar predictions require much experience and knowledge of factors that affect blood pressure fluctuations during surgery.

The volume of modern anesthesia data has increased with the use of electronic medical records. It is difficult for anesthesiologist to use these data to judge patients hemodynamic status during operation, so it will be helpful if there is a tool to support clinical decision making for anesthesiologist based on these data. Machine learning technique is one of the tools for supporting clinical decision making, as it is known to be effective in learning arbitrary patterns of data. There have been few studies on hypotension prediction using machine learning models. One study used various machine learning models based on existing information in the electronic health records to predict hypotension within 10 min of anesthesia induction [3]. It should be noted that it does not predict when exactly the hypotension event will occur, but predicts whether the event will occur within 10 min; so it is not applicable to real-time service because the event may occur right after 1–2 s. There was another study that predicts the probability of developing hypotension 15 min before its actual occurrence by applying a machine learning model to the waveform of invasive arterial blood pressure [4]. Although it predicts the potential hypotension event before its actual occurrence, it is not suitable to real-time prediction because it assumes that the hypotension event does not occur again within 20 min. It is obvious that we do not know when future events will occur, so it is impossible to use it for real-time prediction based on such assumption. As far as we know, most previous studies basically aimed at a classification problem (i.e., predict the potential event), but there was no study for a real-time regression problem (i.e., predict the actual blood pressure).

Machine learning models have been widely adopted for the regression problem. Support vector regression (SVR) [5], which is based on the same principles of the support vector machine (SVM) [6], has been used to predict various real-numbers (e.g., stock price, demand/supply of pulpwood) [7,8]. Random forest (RF) [9] is one of ensemble learning methods. It is known to keep low bias of decision trees and avoid overfitting by controlled variance. The RF can be used for either classification or regression [10]. These traditional machine learning models have shown quite successful results, but they suffer from a common limitation; they strongly depend on a hand-crafted feature set that requires much effort of experts.

Deep learning is one of solutions for the limitation as it automatically extracts arbitrary patterns (i.e., features) beneath the observed data. The deep learning is rooted from artificial neural network (ANN) [11], and it can be used for the regression problem by adopting a particular loss function (e.g., mean squared error). The deep learning is theoretically capable of modeling all non-linear patterns by stacking many layers. As it is discovered that stacking too many layers might cause worse outcome (e.g., low accuracy, high error) due to gradient vanishing [12], there have been several approaches to effectively building deeper layers: rectified linear unit (ReLU) [13], residual connection [14], shortcut connection [15], Inception module [16], and pretraining concept [17,18]. Thanks to these studies, convolutional neural network (CNN), which mainly consists of convolutional layers and pooling layers, was widely used for detection and recognition problem [19–21]. The convolutional layer extracts latent local features, and the pooling layer picks the most meaningful feature among the extracted local features; the CNN effectively captures local patterns and makes a decision by summarizing the most meaningful patterns. On the other hand, recurrent neural network (RNN) [22] allows a layer to have a recursive connection to itself, so that the RNN effectively captures sequential patterns by memorizing previous inputs. Such property makes the RNN to be widely used for machine translation [23,24] and real-time prediction for sequences [25,26].

In this paper, we aim at real-time prediction of blood pressure between the induction of anesthesia and the beginning of operation. This is basically a problem of real-time regression for blood pressure.

Please note that we do not predict the current blood pressure, but the blood pressure of future (e.g., three minutes later). We adopt the RNN to capture arbitrary features from the sequential vital signs, and it makes prediction based on the features. As far as we know, this is the first study of applying the RNN to the real-time prediction of future blood pressure. We believe that this might be helpful for preventing some patients from falling into a critical condition.

This paper is structured as follows. Section 2 describes the characteristics of target data (e.g., vital signs) and how we preprocess the data. It also provides details of our proposed approach as well as the definition of input and output. Section 3 demonstrates the performance of our approach by experimental results, and Section 4 interprets some sample results and discuss about additional experiments with different settings. Finally, Section 5 summarizes and concludes this paper.

2. Materials and Methods

This paper aims at solving a new problem that predicts future blood pressures in real time. We basically follow Data Science (DS) methodology from problem to approach. As mentioned so far, real-time prediction of blood pressures will help to prevent patients from falling into critical condition. Practically, we follow Cross Industry Standard Process for Data Mining (CRISP-DM) methodology that is an iterative process of several steps such as business understanding, data understanding, data preparation, modeling, evaluation, and deployment. We collect and examine the data of vital signs, and preprocess the data to feed them to train our proposed model as shown in Figure 1. The model is designed to incorporate underlying sequential patterns of the vital signs, and evaluated by averaged absolute errors of 10-fold cross-validation. At the running phase, future blood pressures will be predicted given the vital signs for previous few minutes (e.g., 3 min). As this paper is a feasibility study, it is not ready for deployment; it must be carefully deployed because this is a life and death situation. We will keep collecting more data and improving the model for deployment.

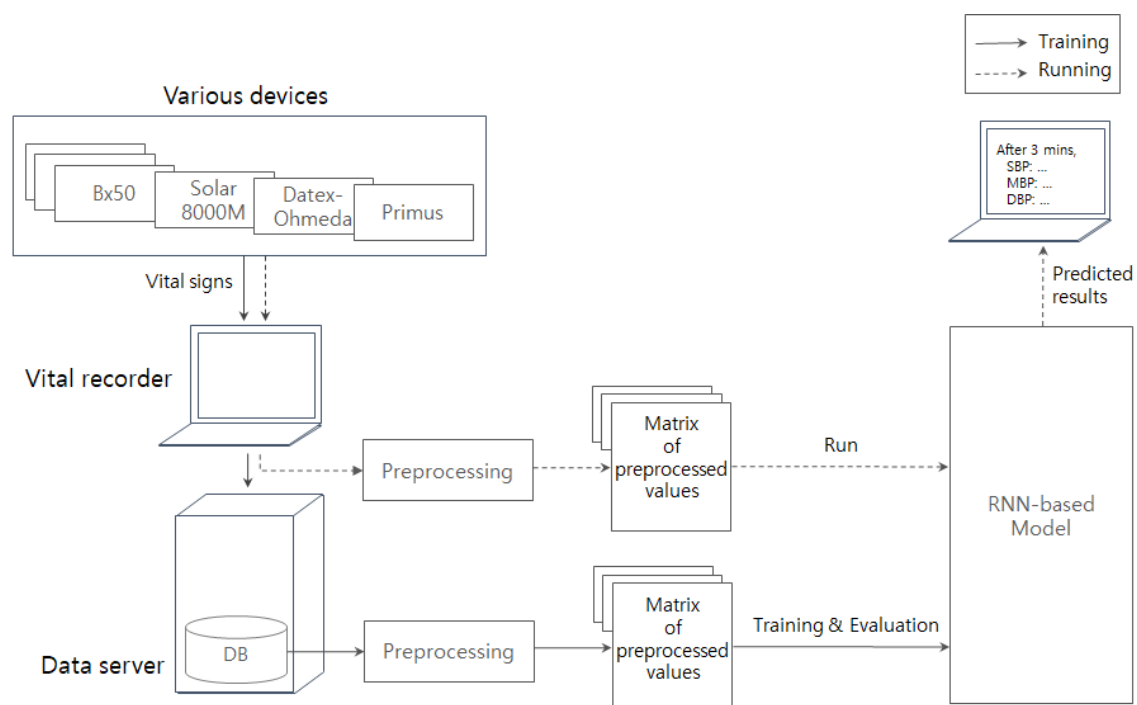


Figure 1. Entire process from data acquisition to output visualization, where the solid line represents the training phase and the dotted line corresponds to the running phase.

This retrospective study was approved by the institutional review board of Soonchunhyang University Bucheon hospital (approval No. 2019-08-016). We collect data from three operation rooms of Soonchunhyang University Bucheon Hospital, where the operations are performed between 29 October 2018 and 18 January 2019. The data are obtained from various devices using Vital Recorder: B \times 50 (patient monitor), Solar 8000M (patient monitor), Datex–Ohmeda (anesthesia machine), Primus (anesthesia machine), BIS (brain monitor) and Orchestra (infusion pump), which results in a K -dimensional real-valued vector. As the vector has few missing values, we employ two strategies: (1) replacing the missing values with the mean of surrounding values, and (2) replacing the missing values with the lastly observed previous value. We apply the first strategy to the vital signs obtainable from the BIS (e.g., signal quality index (SQI)), and make use of the second strategy to other values. In this paper, the vector dimension K is 27, and the detail of the vital signs is described in Table 1. Each dimension of the vector has a distinct sampling rate; for example, BIS/SQI and BIS/BIS are collected every second, whereas TV and MV are collected every six seconds. To address this issue, we assume that all dimensions have the common sampling rate (i.e., three seconds). For example, the blood pressure values (e.g., mean blood pressure (MBP), systolic blood pressure (SBP), diastolic blood pressure (DBP)) are obtained every 1~3 min (mostly every minute), so these values are assumed to be fixed until their new values come in. That is, if the MBP value is sampled every minute, then the MBP values for every 20 timesteps will be the same.

For each r -th surgery operation, we collect K -dimensional vectors for T_r seconds, where $1 \leq r \leq R$ and R denotes the number of operations. As we assume that all vital signs are sampled every three seconds, the total data becomes $R \times K \times (T_r/3)$ tensor (i.e., R sequence of $K \times (T_r/3)$ matrices). Please note that the T_r for different operation will be different because different operations probably have different operation time. For our collected data, the number of patients (i.e., the number of operations) R is 102. The statistics of the collected data are summarized in Table 2. Figure 2 depicts a sample sequence of the collected data. Please note that the three blood pressure values are fixed for 20 timesteps (i.e., one minute) while other values change.

Time	HR	ETCO2	NMT_TOF_CNT	...	REMIFENTANIL_VOL	NIBP_SBP	NIBP_MBP	NIBP_DBP
00:00:00	95	0.2	0		0	156	110	68
00:00:03	94	0.2	0		0	156	110	68
...
00:00:57	92	2.8	0		0.236	156	110	68
00:01:00	94	2.9	0		0.438	137	96	66
...

Figure 2. Sample sequence of real-valued vectors.

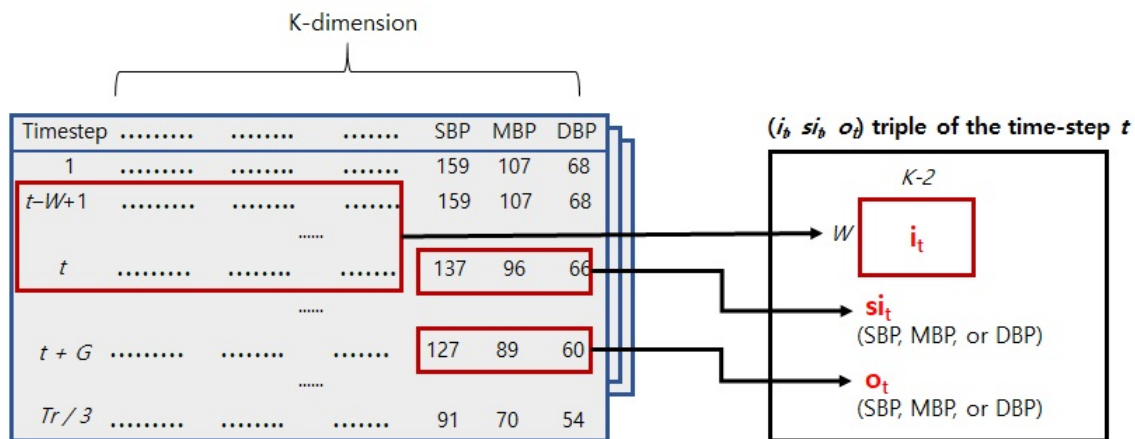
Table 1. Details of collected data.

Name	Description
HR	Heart-rate (number of heartbeats per minute)
ETCO2	End tidal carbon dioxide (carbon dioxide concentration at the end of exhalation)
NIBP_SBP	Non-invasive systolic blood pressure (systolic blood pressure measured by sphygmomanometer)
NIBP_MBP	Non-invasive mean blood pressure (mean blood pressure measured by sphygmomanometer)
NIBP_DBP	Non-invasive diastolic blood pressure (diastolic blood pressure measured by sphygmomanometer)
NMT_TOF_CNT	Neuromuscular transmission train-of-four count (number of single twitches caused by consecutive four electrical stimuli)
RR_TOTAL	Total respiratory rate (number of breaths per minute including mechanical and spontaneous breath)
TV	Tidal volume (volume in single breath)
MV	Minute ventilation (volume of air moved into and out of the lungs in a minute)
PIP	Peak inspiratory pressure (highest airway pressure during inspiration)
MAWP	Mean airway pressure (average airway pressure during inspiration)
PEEP_TOTAL	Total Positive End-Expiratory Pressure (PEEP) (alveolar pressure above atmospheric pressure that exists at the end of expiration)
BIS/SEF	Spectral edge frequency (frequency at which 95% of the total power lies below it)
BIS/SQI	Signal quality index (signal quality of the electroencephalography signal between 0 and 100%)
BIS/EMG	Electromyogram power (a logarithmic scale of total power in the 70–110 Hz range, averaged over the preceding 10 s. It has a minimal value of 25 dB, and in the awake patients it is 40–60 dB)
BIS/BIS	BIS value (value provided by the bispectral index algorithm (range 0–100))
BIS/TOTPOW	Total power (absolute total power in the frequency range (0.5–30 Hz))
PROPOFOL_RATE	Propofol flow rate (speed of propofol administration (mL per hour))
PROPOFOL_CP	Propofol plasma concentration (propofol concentration in the blood (microgram per mL))
PROPOFOL_CE	Propofol effect-site concentration (propofol concentration in the brain (microgram per mL))
PROPOFOL_CT	Propofol target concentration (set concentration of propofol to target brain concentration (microgram per mL))
PROPOFOL_VOL	Propofol infused volume (volume of propofol infused (mL))
REMIFENTANIL_RATE	Remifentanil flow rate (speed of remifentanil administration (mL per hour))
REMIFENTANIL_CP	Remifentanil plasma concentration (remifentanil concentration in the blood (nanogram per mL))
REMIFENTANIL_CE	Remifentanil effect-site concentration (remifentanil concentration in the brain (nanogram per mL))
REMIFENTANIL_CT	Remifentanil target concentration (set concentration of remifentanil to target brain concentration (nanogram per mL))
REMIFENTANIL_VOL	Remifentanil infused volume (volume of remifentanil infused (mL))

Table 2. Statistics of the collected data.

Item	Value
Sex ratio (female:male)	57:45
Age (mean \pm sd)	52.11 \pm 13.13 years
Weight (mean \pm sd)	67.66 \pm 12.96 kg
Height (mean \pm sd)	162.42 \pm 9.51 cm
American Society of Anesthesiologist (ASA) physical status ratio (1:2:3)	42:47:13

We transform the sequence of $K \times (T_r/3)$ matrices into a shape for the real-time sequential prediction of future blood pressure as follows. First, for each t -th timestep where $1 \leq t \leq T_r/3$, we define the sequence of vital signs excluding blood pressures for previous W timesteps as an input i_t ; in other words, the input i_t is a $(K-3) \times W$ matrix for the timesteps between $[t-W+1, t]$. We also add the timestep t into the i_t , so finally the i_t becomes a $(K-2) \times W$ matrix. Second, we define a normalized blood pressure at the t -th timestep as a supplementary input si_t . If the blood pressure value is 125, then it is divided by 250 to be normalized (e.g., 125/250). We take only the latest observed blood pressure, but not the blood pressure values for W timesteps because the inconsistent sampling rate (e.g., every minute) of the blood pressure may harm the results of the RNN. Third, we define the blood pressure value of the timestep $t+G$ as an output o_t . It is important that the output o_t is not the blood pressure at the timestep t , but the future blood pressure at the timestep $t+G$. Through the steps above, for each timestep t , we generate a triple of the input i_t , the supplementary input si_t , and the output o_t . Assuming that $t = 100$, $W = 60$, and $G = 20$. The input i_{100} will be a $(K-2) \times 60$ matrix and the si_{100} will be a real-number of the normalized blood pressure at the 100-th timestep. The output o_{100} will be a blood pressure value at the 120-th timestep. This can be interpreted that it predicts the blood pressure of one minute later (i.e., after 20 timesteps) given the lastly observed vital signs for three minutes (i.e., 60 timesteps). As we generate the triple (i_t, si_t, o_t) for every timestep, the total number of triples for the r -th operation will be $T_r/3 - W - G + 1$. We conduct the above transformation process to the three blood pressures (e.g., MBP, DBP, and SBP) independently, and got the triples for each of them. The transformation process is summarized in Figure 3. In short, the input consists of W vital signs i_t and a current blood pressure si_t , while the output is a future blood pressure o_t after G timesteps.

**Figure 3.** Data preparation for the prediction of future blood pressure in real time.

We observe that different operations exhibit different sequential patterns (e.g., different aspect of heart-rate changes). To incorporate such diversity of sequential patterns, we design an RNN model followed by fully connected layers as shown in Figure 4. Given the input i_t for r -th operation, the W vital vectors are sequentially injected to the RNN. Please note that our RNN has bidirectional and hierarchical structure. The bidirectional RNN consists of a forward RNN and a backward RNN, where the forward RNN and the backward RNN can capture forward patterns and backward patterns,

respectively. There might be sequential patterns of a forward direction and a backward direction, so we take the bidirectional RNN to incorporate such patterns.

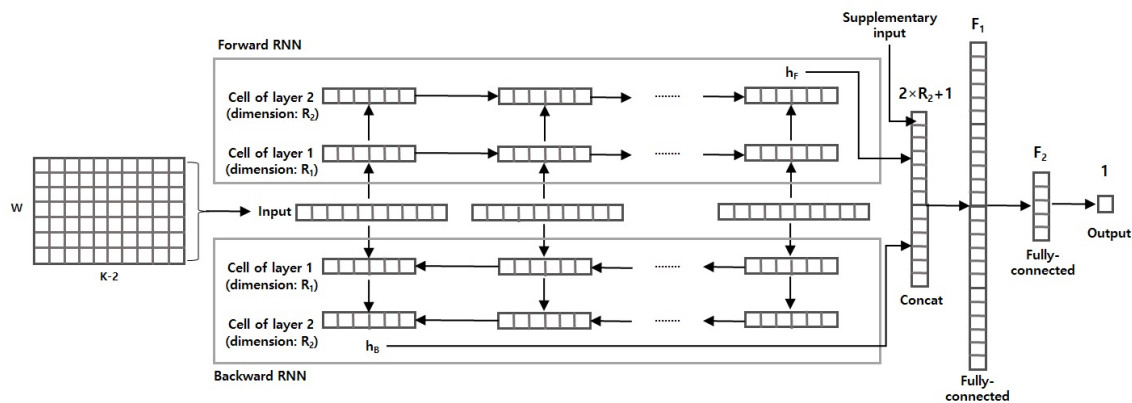


Figure 4. Structure of the proposed model.

Meanwhile, both forward and backward RNN are hierarchical as they have two stacked layers. The first RNN layer may capture sequential correlations between different vital signs (e.g., a propofol rate and the heart-rate), and the second RNN layer catches high-level sequential correlations between the correlations found at the first layer. Thanks to the bidirectional and hierarchical structure, the RNN will memorize high-level sequential patterns in both directions. The forward RNN yields a R_2 -dimensional summary vector h_F , and the backward RNN also gives a R_2 -dimensional summary h_B . These two summary vectors are then concatenated and the supplementary input si_t comes into the vector, resulting in a $2 \times R_2 + 1$ -dimensional vector. The concatenated vector is passed to the fully connected layers that are supposed to find some correlations between the h_F and h_B . For example, when the RNN layers may capture ‘increasing trend of heart rates’ and ‘fluctuating ETCO2’ patterns, the fully connected layers may find how positive or negative correlation they have. Finally, given the F_2 -dimensional vector generated by the second fully connected layer, the output layer predicts the future blood pressure.

For the cell of the RNN layers, we adopt the Gated Recurrent Unit (GRU) [27] that is one of the most widely used RNN cells. The most important aspect of the RNN cell is that it remembers previously observed information. Although the RNN cell must be capable of preserving every previous information theoretically, it loses long-term information practically. The GRU is one of solutions to settle such issue by two types of gates (i.e., an update gate and a reset gate). These two types of gates help to preserve important long-term information while discarding unnecessary information. Thanks to the GRU cells, the bidirectional RNN layers give two vectors (e.g., h_F and h_B) that capture important sequential patterns in both directions. For the two fully connected layers, we adopt the rectified linear unit (ReLU) [13] as an activation function as it is known to prevent from the gradient vanishing problem. For the output layer, we take mean squared error (MSE), which is widely used for regression, as a loss function.

3. Results

We set $W = 60$ and $G = 60$, which implies that we predict the blood pressure of three minutes later, given the observed vital signs for latest three minutes. Please note that we use the vital data only obtained between the induction of anesthesia and the beginning of operation; we do not employ any other information (e.g., age, sex, base blood pressures, ASA). The total number of transformed data is 26,887. Each dimension of the transformed data is normalized except for the timestep value. The normalization, of course, is done with only training data. We take 10-fold cross-validation and compute mean absolute error (MAE). We conduct three independent experiments: SBP prediction, MBP prediction, and DBP prediction. All experiments are performed using a computer with eight

Central Processing Units (CPU) of i7-7700 3.6 GHz and two NVIDIA GeForce 1080 Ti. The proposed model is implemented with Python3 language with Google TensorFlow packages.

The training recipe and parameter setting are as follows. The dimensions of RNN layers R_1 and R_2 are equally 15, and the dimensions of the fully connected (FC) layers F_1 and F_2 are 100 and 50, respectively. We applied the drop out [28] with a keep probability 0.1 to the RNN layers, and the decov [29] with a weight 0.1 to the FC layers. Both the drop out and the decov are known to have a regularization effect, which prevents from overfitting. In terms of the parameter initialization, the weight matrices of the FC layers are initialized using He initialization [30], and the biases are initialized as zero. The weight matrices of the RNN layers are initialized using Xavier initialization [31], and the initial bias value is one. We use Adam optimizer [32] with an initial learning rate 0.001 to train the model parameters, and the number of epochs is 60. For training phase, it computes a predicted blood pressure by feed-forward propagation; the RNN layers generate two vectors given a input, and the fully connected layers take the concatenation of the two vectors as an input and generate an output. It computes a cost (error) by comparing the predicted blood pressure and a true blood pressure, and all weights and bias values are updated via back propagation algorithm. For each epoch, the feed-forward and back propagation are conducted throughout all data with a mini-batch as a unit. In this paper, we set the size of mini-batch as 100.

Table 3 summarizes the mean and standard deviation of the absolute errors obtained from the three predictions. Small mean and standard deviation mean that it predicts the blood pressure accurately. Among the three predictions, the DBP prediction is the most accurate while the SBP prediction exhibits the worst results. The Figures 5–7 depict histograms of errors, where horizontal axis represents error bins; for example, a bin [1–2) represents the range $1 \leq e < 2$ where e indicates an error. The three figures seem to have a form of Gaussian distribution, and they generally follow the trend of the true blood pressures. For instance, in Figure 6, the peak of distribution is located around the interval [0–1), which implies that the predicted MBP values are mostly correct compared to the true MBP values. However, the shapes of three figures are a bit left skewed, so the overall mean is between 8.2 mmHg and 11.1 mmHg while the standard deviation is between 8.7 mmHg and 12.7 mmHg. Figure 8 shows Bland-Altman diagrams of the three blood pressures. The diagrams imply that errors tend to grow when the average of a predicted blood pressure and a true blood pressure is high. This can be interpreted that it is hard to correctly predict the true blood pressure when the average is abnormally high because such cases were barely seen in the data.

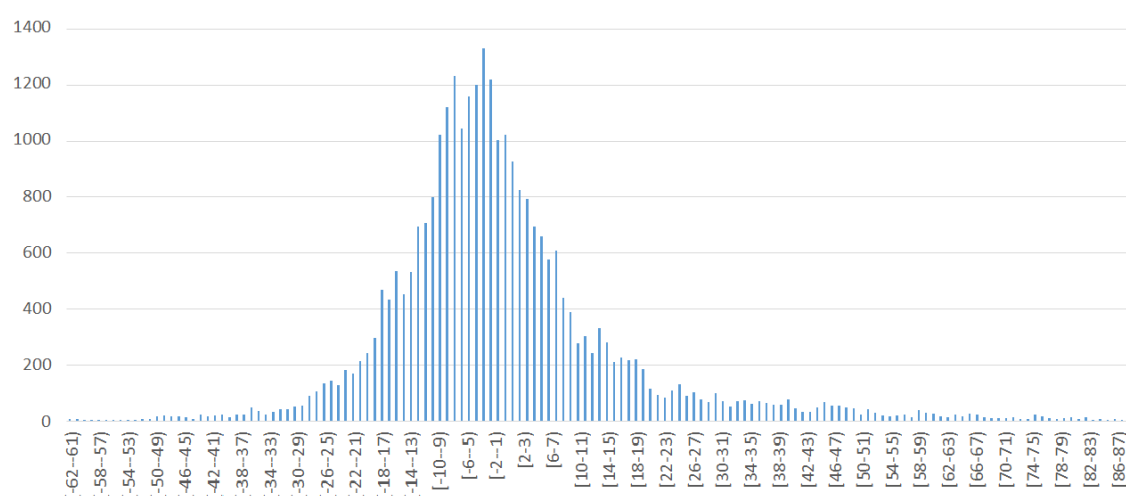


Figure 5. Error histogram of SBP prediction, where horizontal axis represents error bins and vertical axis is the number of associated data.

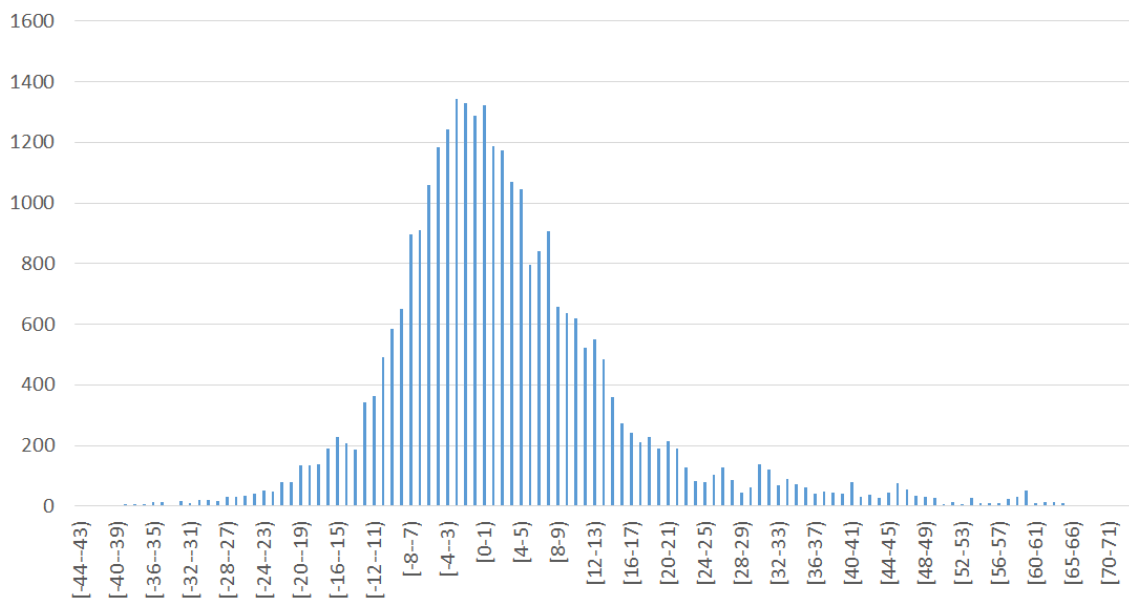


Figure 6. Error histogram of MBP prediction, where horizontal axis represents error bins and vertical axis is the number of associated data.

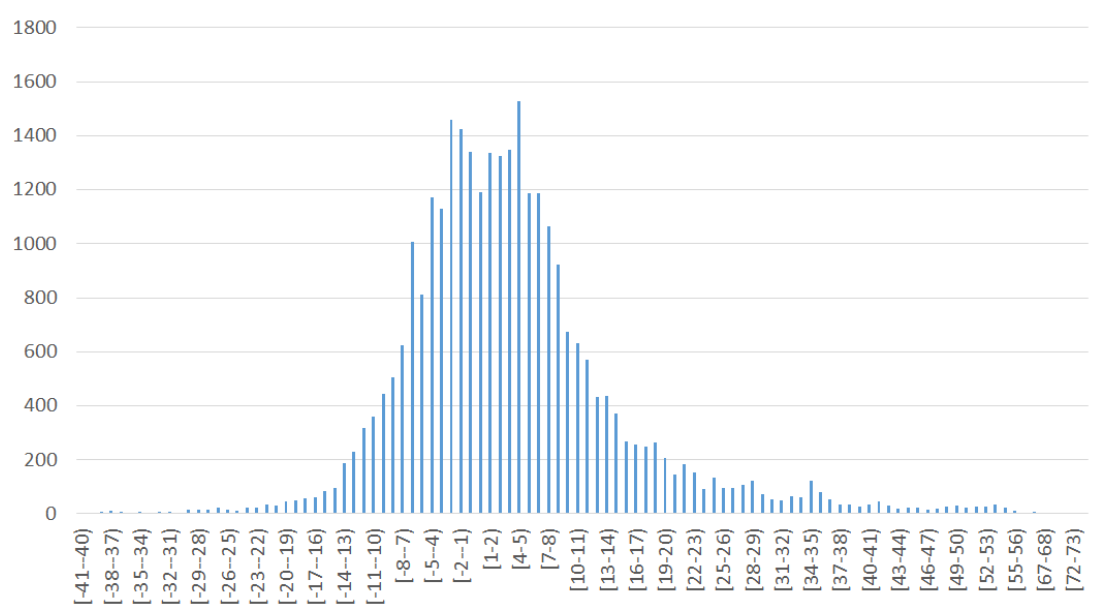


Figure 7. Error histogram of DBP prediction, where horizontal axis represents error bins and vertical axis is the number of associated data.

Table 3. Mean and standard deviation of absolute errors.

Experiment	Mean (mmHg)	Standard Deviation (mmHg)
SBP prediction	11.056	12.687
MBP prediction	9.253	9.835
DBP prediction	8.231	8.689

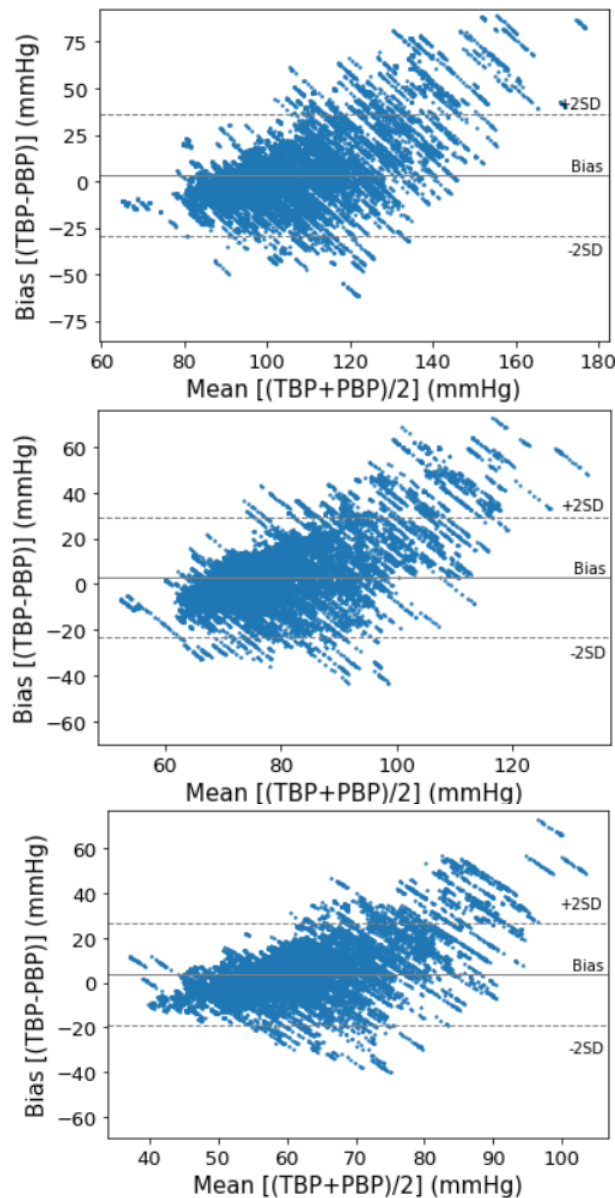


Figure 8. Bland-Altman diagrams of SBP prediction, MBP prediction, and DBP prediction, where TBP and PBP mean true blood pressures and predicted blood pressures, respectively.

4. Discussion

We investigated whether RNN could predict future blood pressure (e.g., 3 min ahead) during anesthesia-induction period. We found that our model could predict 3-min ahead blood pressure with absolute error around 10 mmHg for each SBP, DBP, and MBP. Although this error seems to be large for helping clinicians to use our model as decision support tool in the hemodynamic management during anesthesia for now, we suggest it is feasible for RNN to predict future blood pressure using only features those obtained from various anesthesia monitors, ventilator and drug infusion pump in relatively short periods.

We examine the plots of predicted blood pressure and true blood pressure. To do so, we trained the model with 90% of shuffled data, and the remaining data is used for examination. Figure 9 shows three plots of SBP prediction, where the two upper examples are relatively well predicted cases and the bottom example shows a poorly predicted case. Please note that the model gives its first prediction at the 120-th timestep because it sees the sequential data of three minutes (i.e., 60 timesteps) and predicts three minutes later (i.e., 60 timesteps). Because the SBP is sampled every minute, the plot of true SBP

looks like stairs. Generally speaking, the three figures in Figure 9 show that the model well predicts the trend of future blood pressure; it captures when the SBP will arise, keep or fall. Interestingly, as shown in the second figure of Figure 9, the predicted SBP fluctuates as the true SBP even though it predicts the SBP of three minutes after. On the other hand, in the bottom figure, the predicted SBP follows the trend of true SBP but there is a steady gap between them. We believe that such gap will be reduced if we collect more data to incorporate various patterns of blood pressure.

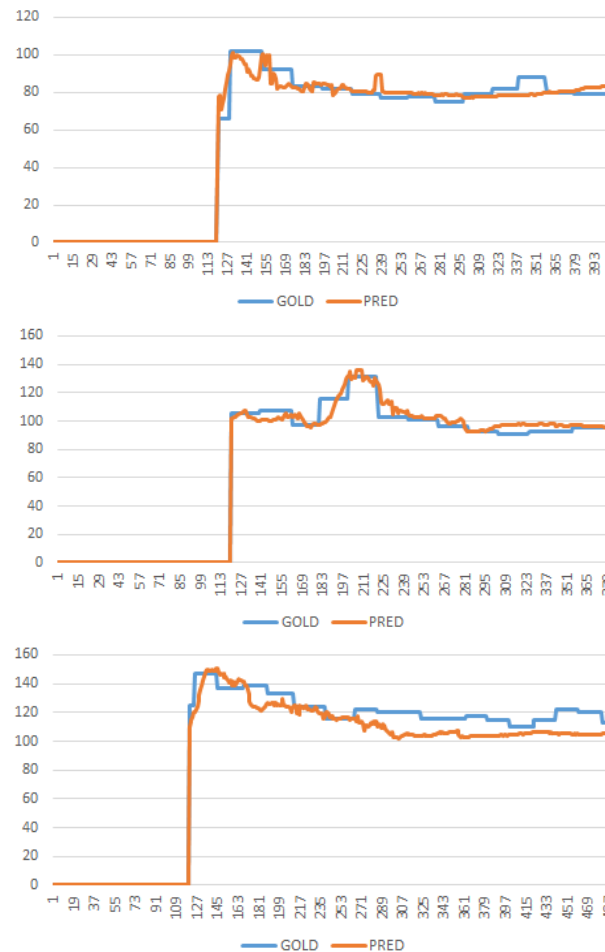


Figure 9. Plots of predicted SBP (orange color) and true SBP (blue color), where horizontal axis represents a timestep and vertical axis is the SBP (mmHg).

Among the hemodynamic changes occurring during surgery, hypotension is known to be frequent and has been reported to cause adverse outcomes after surgery [1]. Definition of intraoperative hypotension varies among investigators which ranges from MBP of 55 mmHg to 65 mmHg. In [33], it was revealed that MBP less than 60 mmHg for 11 to 20 min and MBP less than 55 mmHg for more than 10 min are associated with acute kidney injury. The mean absolute error of MBP predicted by our proposed model was 9 mmHg, which may not be helpful to clinicians in some critical situations. For example, if the actual MBP is 58 mmHg, then MBP predicted by our model may range from 49 mmHg to 67 mmHg. Such variation of the predicted MBP might cause two opposite ways of management. If the predicted MBP is 49 mmHg, one will explore possible causes for hypotension, whereas one just observes blood pressures and do nothing if the predicted MBP is 67 mmHg. Of course, there can be another case that the predicted MBP is helpful. Assuming that actual MBP is around 75 mmHg, and predicted MBP may range from 66 mmHg to 84 mmHg. This is generally not harmful to most surgical patients. The Association for the Advancement of Medical Instrumentation (AAMI) established standards for the validation of automatic arterial pressure monitoring. It was defined

as acceptable if error (e.g., mean absolute error) is not greater than 5 mmHg and standard deviation of errors is not greater than 8 mmHg for SAP and DAP [34]. In this regard, as the mean absolute errors of our model for SBP and MBP were 11 mmHg and 9 mmHg, respectively, which does not meet the AAMI standards. However, there is no consensus on the accuracy of clinically acceptable blood pressure because the AAMI standards are for the approving clinical validation of new automated blood pressure devices.

One may argue that there might be better parameter settings or better structure of the model. The training recipe and parameter setting used in this paper is obtained via a grid-searching. We varied the number of RNN layers and fully connected layers, and tried various dimensions. A part of the grid-searching result is shown in Table 4, where the relative change of MAE is computed using the best MAE 11.056 of Table 3; the relative change is $(\text{current MAE} - 11.056) / 11.056 \times 100$, so greater value means worse result. It seems that the bidirectional RNNs generally work better than the unidirectional RNNs. The FC dimensions represent F_1 and F_2 ; [100, 50] means F_1 and F_2 are 100 and 50, respectively, and [100] implies that it uses a single FC layer with $F_1 = 100$. It seems that using two FC layers is much better than using a single FC layer, and the regularization methods (e.g., drop out, decov) prevent the model from overfitting.

Table 4. Part of SBP prediction results with different settings, where FC means fully connected layers.

Bidirectional	RNN Dimensions	RNN Drop Out	FC Dimensions	FC Decov	MAE (Relative Change)
No	[10, 10]	No	[100]	No	12.803 ($\Delta 15.8\%$)
No	[15, 15]	No	[100]	No	12.715 ($\Delta 15.0\%$)
No	[20, 20]	No	[100]	No	12.996 ($\Delta 17.5\%$)
No	[15, 15]	No	[100, 50]	No	11.789 ($\Delta 6.6\%$)
Yes	[15, 15]	No	[100, 50]	No	11.770 ($\Delta 6.5\%$)
Yes	[15, 15]	No	[100, 50]	0.3	11.693 ($\Delta 5.8\%$)
Yes	[15, 15]	No	[100, 50]	0.1	11.537 ($\Delta 4.4\%$)
Yes	[15, 15]	0.25	[100, 50]	0.1	12.062 ($\Delta 9.1\%$)

This study aims at a real-time prediction of blood pressure, so one may ask ‘Does this model really work in real time?’ because our model has a quite complex structure (e.g., a composite model of RNN and fully connected layers). We found that the average elapsed time for prediction of a batch of 100 unseen data is about 26.56 milliseconds. As our model must give a prediction result every three seconds, it is definitely capable of the real-time prediction.

Although our model exhibits its potential as a real-time predictor of future blood pressure, there is a room for improvement, especially about the error. About the SBP prediction, its mean absolute error 11.056 indicates that we still have a lot to do. The main reason for this is that our data is obtained from only 102 operations, which is not much enough for incorporating diverse patterns of operations. Thus, this study can be a first step that proves the feasibility of the real-time prediction of future blood pressure. We believe that our model will achieve further improvement as we will keep collecting more data. Another minor limitation of our work is that it gives its first result after some timesteps (e.g., 120 timesteps), which can be addressed if we collect the vital signs before the induction of anesthesia.

5. Conclusions

In this study, we prepared and preprocessed the vital signs, and design a recurrent neural network for real-time prediction of future blood pressures. We demonstrated that the model has a potential to predict the future blood pressures by histograms of absolute errors, but also observed its limitation (e.g., mean and standard deviation of absolute errors). By the plots of predicted blood pressures, we showed that the model can foresee the trend of blood pressures. We also proved that our model works in real time by measuring average elapsed time for prediction. This study is not a final stop and not ready for deployment, but shows a feasibility of the RNN-based model for the prediction of future

blood pressures. We believe that this study will help to reduce facing emergent situations by warning to the medical team before it happens. For example, if our model reports that a future SBP will be low (e.g., 60 mmHg), then the medical team may inject arteriopressor to prevent potential hypotension. To improve the performance of the proposed model, we will keep collecting more data and finding better model structures. Furthermore, we will investigate other useful devices (or sensors) as well as a combination of clinical values (e.g., EMR) and the vital signs. We will also extend our study to develop a real-time system of an intraoperative prediction of future blood pressures.

Author Contributions: Conceptualization, Y.-S.J. and S.H.K.; Data curation, A.R.K.; Investigation, A.R.K.; Methodology, Y.-S.J.; Project administration, Y.-S.J.; Resources, W.J., S.L., S.J.L., M.L., Y.H.C. and B.S.K.; Software, W.J.; Validation, Y.-S.J., A.R.K., W.J., S.L., S.J.L., M.L., Y.H.C., B.S.K. and S.H.K.; Writing—original draft, Y.-S.J.; Writing—review & editing, A.R.K. and S.H.K.

Acknowledgments: This research was supported by the Soonchunhyang University Research Fund and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2018R1A6A3A01011337).

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

References

- Walsh, M.; Devereaux, P.J.; Garg, A.X.; Kurz, A.; Turan, A.; Rodseth, R.N.; Cywinski, J.; Thabane, L.; Sessler, D.I. Relationship between Intraoperative Mean Arterial Pressure and Clinical Outcomes after Noncardiac Surgery: Toward an Empirical Definition of Hypotension. *Anesthesiology* **2013**, *119*, 507–515. [[CrossRef](#)]
- Goldberg, M.E.; Larijani, G.E. Perioperative hypertension. *Pharmacotherapy* **1998**, *18*, 911–914. [[PubMed](#)]
- Kendale, S.; Kulkarni, P.; Rosenberg, A.; Wang, J. Supervised Machine-learning Predictive Analytics for Prediction of Postinduction Hypotension. *Anesthesiology* **2018**, *129*, 675–688. [[CrossRef](#)] [[PubMed](#)]
- Hatib, F.; Jian, Z.; Buddi, S.; Lee, C.; Settels, J.; Sibert, K.; Rinehart, J.; Cannesson, M. Machine-learning Algorithm to Predict Hypotension Based on High-fidelity Arterial Pressure Waveform Analysis. *Anesthesiology* **2018**, *129*, 663–674. [[CrossRef](#)] [[PubMed](#)]
- Drucker, H.; Burges, C.J.C.; Kaufman, L.; Smola, A.; Vapnik, V. Support vector regression machines. In Proceedings of the 9th International Conference on Neural Information Processing Systems, Denver, CO, USA, 3–5 December 1996; pp. 155–161.
- Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A Training Algorithm For Optimal Margin Classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, Pittsburgh, PA, USA, 27–29 July 1992; pp. 144–152.
- Henrique, B.M.; Sobreiro, V.A.; Kimura, H. Stock price prediction using support vector regression on daily and up to the minute prices. *J. Financ. Data Sci.* **2018**, *4*, 183–201. [[CrossRef](#)]
- Anandhi, V.; Chezian, R.M. Support Vector Regression to Forecast the Demand and Supply of Pulpwood. *Int. J. Future Comput. Commun.* **2013**, *2*, 266–269. [[CrossRef](#)]
- Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
- Liaw, A.; Wiener, M. Classification and Regression by randomForest. *R News* **2002**, *2*, 18–22.
- Hassoun, M.H. *Fundamentals of Artificial Neural Networks*; MIT Press: Cambridge, MA, USA, 1995.
- Hochreiter, S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *Int. J. Uncertain. Fuzziness Knowl. Based Syst.* **1998**, *6*, 107–116. [[CrossRef](#)]
- Nair, V.; Hinton, G.E. Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th International Conference on International Conference on Machine Learning, Haifa, Israel, 21–24 June 2010; pp. 807–814.
- He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–788.
- Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2261–2269.
- Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.

17. Hinton, G.E.; Osindero, S.; Teh, Y.W. A fast learning algorithm for deep belief nets. *Neural Comput.* **2006**, *18*, 1527–1554. [[CrossRef](#)] [[PubMed](#)]
18. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proceedings of the 3rd International Conference on Learning Representations, Scottsdale, AR, USA, 2–4 May 2015; pp. 1–14.
19. Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-Based Learning Applied to Document Recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [[CrossRef](#)]
20. Redmon, J.; Divvala, S.K.; Girshick, R.B.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 779–788.
21. Mollah, A.F.; Majumder, N.; Basu, S.; Nasipuri, M. Design of an Optical Character Recognition System for Camerabased Handheld Devices. *Int. J. Comput. Sci. Issues* **2011**, *8*, 283–289.
22. Medsker, L.; Jain, L.C. *Recurrent Neural Networks: Design and Applications*; CRC Press: Boca Raton, FL, USA, 1999.
23. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to Sequence Learning with Neural Networks. In Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; pp. 3104–3112.
24. Luong, T.; Pham, H.; Manning, C.D. Effective Approaches to Attention-based Neural Machine Translation. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Lisbon, Portugal, 17–21 September 2015; pp. 1412–1421.
25. Liu, Q.; Wu, S.; Wang, L.; Tan, T. Predicting the Next Location: A Recurrent Model with Spatial and Temporal Contexts. In Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, AR, USA, 12–17 February 2016; pp. 194–200.
26. Meftah, S.; Semmar, N. A Neural Network Model for Part-Of-Speech Tagging of Social Media Texts. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation, Miyazaki, Japan, 7–12 May 2018; pp. 2821–2828.
27. Cho, K.; van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 25–29 October 2014; pp. 1724–1734.
28. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
29. Cogswell, M.; Ahmed, F.; Girshick, R.B.; Zitnick, L.; Batra, D. Reducing Overfitting in Deep Networks by Decorrelating Representations. In Proceedings of the 4th International Conference on Learning Representations, San Juan, Puerto Rico, 2–4 May 2016; pp. 1–12.
30. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. In Proceedings of the 2015 IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1026–1034.
31. Glorot, X.; Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 13–15 May 2010; pp. 249–256.
32. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. In Proceedings of the 3rd International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015; pp. 1–15.
33. Sun, L.Y.; Wijeyesundera, D.N.; Tait, G.A.; Beattie, W.S. Association of Intraoperative Hypotension with Acute Kidney Injury after Elective Noncardiac Surgery. *Anesthesiology* **2015**, *123*, 515–523. [[CrossRef](#)] [[PubMed](#)]
34. *Non-Invasive Sphygmomanometers—Part 2: Clinical Validation of Automated Measurement Type*; Standard, International Organization for Standardization: Geneva, Switzerland, 2009.

