Deep learning-based automatic blood pressure measurement: evaluation of the effect of deep breathing, talking and arm movement

Pan, F., He, P., Chen, F., Pu, X., Zhao, Q. & Zheng, D.

DOI 10.1080/07853890.2019.1694170
ISSN 0785-3890
ESSN 1365-2060

Publisher: Taylor and Francis

Deep learning based automatic blood pressure measurement: Evaluation of the effect of deep breathing, talking and arm movement

Deep learning based automatic blood pressure measurement

Original Article

Fan Pan[1], Peiyu He[1], Fei Chen[2], Xiaobo Pu[3], Qijun Zhao[4], and Dingchang Zheng[5]

[1] College of Electronics and Information Engineering, Sichuan University, Chengdu, China

[2] Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, China

[3] Department of Cardiology, West China Hospital, Sichuan University, Chengdu, China

[4] College of Computer Science, Sichuan University, Chengdu, China

[5] Research Centre of Intelligent Healthcare, Faculty of Health and Life Science, Coventry University, Coventry, UK

Address correspondence to Peiyu He, College of Electronics and Information Engineering, Sichuan University, Chengdu 610064, China. Electronic mail: hpysbsy@163.com.

Address correspondence to Dingchang Zheng, Research Centre of Intelligent Healthcare, Faculty of Health and Life Science, Coventry University, CV1 5FB, UK. Electronic mail: dingchang.zheng@coventry.ac.uk.

Abstract:

Objectives: It is clinically important to evaluate the performance of a newly developed blood pressure (BP) measurement method under different measurement conditions. This study aims to evaluate the performance of using deep learning based method to measure BPs and BP change under non-resting conditions.

Materials and Methods: 40 healthy subjects were studied. Systolic and diastolic BPs (SBPs and DBPs) were measured under four conditions using deep learning and manual auscultatory method. The agreement between BPs determined by the two methods were analysed under different conditions. The performance of using deep learning based method to measure BP changes was finally evaluated.

Results: There were no significant BPs differences between two methods under all measurement conditions (all $P > 0.1$). SBP and DBP measured by deep learning method changed significantly in comparison with the resting condition: decreased by 2.3 and 4.2 mmHg with deeper breathing (both $P < 0.05$), increased by 3.6 and 6.4 mmHg with talking, and increased by 5.9 and 5.8 mmHg with arm movement (all $P < 0.05$). There were no significant differences in BP changes measured by two methods (all $P > 0.4$, except for SBP change with deeper breathing).

Conclusion: This study demonstrated that the deep learning method could achieve accurate BP measurement under both resting and non-resting conditions.

Key Messages:

- Accurate and reliable blood pressure measurement is clinically important. We evaluated the performance of our developed deep learning based blood pressure measurement method under resting and non-resting measurement conditions.
- The deep learning based method could achieve accurate BP measurement under both resting and non-resting measurement conditions.

Introduction

The gold standard for clinical blood pressure (BP) measurement is manual auscultatory method, which reads both systolic and diastolic BP (SBP and DBP) values with a stethoscope from a sphygmomanometer.[1] If clinical users are to be encouraged to use the manual measurement technique there is one challenging problem to be overcome. Users often find the manual identification of systole and diastole by a stethoscope difficult. So, expertise with the stethoscope is the most important aspect of the manual measurement, which requires training, skill, experience and good hearing.[2] This also results in the potential for inaccurate measurement due to small changes in the sounds heard, as well as loss in confidence.[3]

Deep learning technologies have been widely used in medicine including identifying moles from melanomas, diabetic retinopathy, and cardiovascular risk and breast lesion detection in mammograms.[4-7] Chen et al proposed a deep neural network (DNN) based classifier to recognize audible heart sound with more than 91% accuracy.[8] We have recently developed a deep learning based automatic auscultatory BP measurement method, where the convolutional neural network (CNN) was employed to identify the audible KorS sounds and a mapping algorithm was used to determine the automatic BP value.[9] In that paper, the accuracy of the deep learning based BP measurement has been evaluated under resting condition with reference manual auscultatory method. The overall measurement errors of the deep learning based method were 1.4±2.4 mmHg for SBP and 3.3±2.9 mmHg for DBP, suggesting that the deep learning based method is an effective technique to measure BPs.

It is well accepted that the BP measurement is highly affected by measurement conditions, including the back support, body and arm position.[10] Cushman et al has reported that, if the patient's back is not supported, the measured SBP and DBP increased by 5-15 and 6 mmHg, respectively.[11] Higher SBP and DBP (by 3-10 mmHg and 1-5 mmHg, respectively) were also observed with the measurements performed in the supine than the seated position.[12] Zheng et al has also quantified the effect of other different measurement conditions (deeper breathing, talking, arm and head movement) on auscultatory BP measurement. In comparison with the resting condition, there were significant manual auscultatory BPs differences measured from deeper breathing, talking and arm movement.[13] However, the performance of our recently proposed deep learning method has not been evaluated under non-resting measurement conditions.

The aim of this study is to evaluate the performance of using deep learning based method[9] to measure BPs and BP change under non-resting conditions (deeper breathing, talking, and arm movement) in comparison with the manual auscultatory method.

Materials and methods

Subjects

International Standards Organization (ISO) requires that the overall mean and standard deviation (SD) of the difference between a new BP measurement technique and the reference BP (from manual auscultatory method) should be ≤ 5 and 8 mmHg, respectively.[14] Sample size calculation was performed based on a paired t-test for mean difference to allow a mean 5mmHg BP difference to be detected with a typical 8mmHg SD of BP measurement. 21 subjects were therefore enough to achieve a confidence level of 95% and a statistical power of 80%. 40 healthy and normotensive subjects were enrolled in this study, of which 30 were male and 10 were female. The mean age of subjects was 43 years, with a range of age from 23 to 65 years. The investigation conformed with the principles in the Declaration of Helsinki. This experiment received ethical permission from the Newcastle & North Tyneside Research Ethics Committee, and all subjects gave their written informed consent to participate in the study. Table 1 summarizes the subject information, including age, sex, height, and arm circumference.

BP measurement protocol

The BP measurement experiment was performed by a trained operator in a quiet and temperature-controlled clinical measurement room. Subjects were asked to rest on a chair for at least 10 minutes before the measurements were taken. Manual SBP and DBP were measured from the left arm using traditional manual auscultatory method with a sphygmomanometer and a stethoscope. The whole procedure followed the guidelines recommended by the British Hypertension Society and American Heart Association.[1,15] The cuff pressure was linearly deflated from at a rate of 2-3 mmHg/s automatically.

For each subject, there were three repeated sessions. During each session, BP measurements were performed sequentially under four different measurement conditions (simply as ‘resting’, ‘talking’, ‘arm movement’ and ‘deeper breathing’). The sequence of these conditions was randomized between subjects. For the three non-resting conditions, subjects were asked to breathe deeply and regularly, to talk by counting numbers and to move the right arm (the arm without cuff) forward and backward. The conditions were designed to induce small typical effects in clinical practice, with deeper breathing at a level greater than normal breathing, but which could be easily and comfortably sustained by volunteers. In total, there were 12 BP measurements from each subject (three repeated sessions with four measurements for each). Between the repeated sessions there was a time interval of 3-4 minutes, and at least 1 minute between the four measurements within a session. During each BP measurement, the cuff pressure and stethoscope sound were simultaneously and digitally recorded to a data capture computer at a sample rate of 2000 Hz for offline deep learning analysis in the next step.

BP measurement using deep learning method

A deep learning method using convolutional neural networks (CNN) to identify the audible Korotkoff sound (KorS) has been developed in our previous publication.[9] As shown in Figure 1, after the audible KorS and non-audible KorS beats were identified by the trained CNN, the cuff pressures that corresponded to the first and last audible KorS beats were used to determine automatic SBP and DBP. To follow the guideline of manual auscultatory BP measurement, the following rule was applied in automatic BP determination: SBP was determined with at least two consecutive audible KorS beats identified by CNN, and DBP was determined at the beat at which all sounds disappear completely.[1]

Data and statistical analysis

There were 24 SBP values and 24 DBP values from each subject (from 2 measurement methods, 4 measurement conditions and 3 repeat sessions). The SPSS software package (SPSS Inc., Chicago, IL, USA) was employed to perform repeated measures analysis of variance to study the measurement repeatability. The value of $P < 0.05$ was considered a statistically significant difference. Figure 1 also shows the flow of data and statistical analysis. The overall mean and standard deviation (SD) for SBP and DBP were calculated across all the subjects separately for each measurement condition and method. The mean BP differences between deep learning and manual methods were also calculated, separately for each condition. All differences were paired values in each subject. In addition, Bland-Altman analysis was applied to investigate the agreement between BPs determined by two methods, respectively for each condition.

Analysis of variance (ANOVA) with post-hoc multiple comparisons was used to investigate the effect of conditions on measured BPs and the significance of BP changes with non-resting measurement conditions. Finally, the SBP and DBP changes (differences between non-resting and resting conditions) measured by the two methods were compared.

Results

BP agreement between deep learning and manual auscultatory method

Statistical analysis showed that there was no significant BP difference (for both SBP and DBP) between the 3 repeat measurement sessions (all $P > 0.05$). The means from the three repeats for each subject was then used to for the following analysis. The overall BP (mean ± SD) from four measurement conditions were given in Table 2, separately for the deep learning and manual auscultatory methods. There were no statistically significant differences between the deep learning method and manual auscultatory method under the four measurement conditions for both SBP and DBP (all $P > 0.1$). Figure 2 shows the

Bland-Altman plots of the SBP and DBP measured by deep learning and manual auscultatory method under the four measurement conditions.

Effect of measurement condition on BPs

Figure 3 demonstrates the overall mean + standard error of the mean (s.e.m.) of SBP and DBP under each measurement condition, separately for deep learning and manual methods. Table 3 gives the mean paired differences between the three non-resting conditions and resting condition.

From manual auscultatory method, the mean SBP and DBP measured with deeper breathing decreased significantly by 3.6 and 3.9 mmHg, respectively, in comparison with those for the resting condition ($P < 0.05$). Also, they increased significantly by 4.0 and 6.2 mmHg with talking, and increased by 6.0 and 6.0 mmHg with arm movement (all $P < 0.05$).

With the BPs determined by deep learning method, the mean SBP and DBP measured with deeper breathing decreased significantly by 2.3 and 4.2 mmHg, respectively, in comparison with those for the resting condition ($P < 0.05$), increased significantly by 3.6 and 6.4 mmHg with talking, and increased by 5.9 and 5.8 mmHg with arm movement (all $P < 0.05$).

Comparison of BP changes between the two measurement methods

As shown in Figure 4, there were no significant differences in BP changes measured by the deep learning and manual methods (all $P > 0.4$, except SBP measure with deeper breathing, where $P = 0.02$). This indicated that the small BP changes caused by non-resting conditions can be accurately measured by the deep learning method.

Discussion

This study demonstrated that when compared with the reference manual auscultatory method, there was no significant difference between deep learning method and manual method under four different measurement conditions. Although we have previous reported that deep learning method could measure BPs accurately under resting condition with the measurement error of 1.4±2.4 mmHg for SBP and 3.3±2.9 mmHg for DBP,[9] it is important to evaluate its measurement performance under non-resting condition. In this study, the deep learning method achieved less than 1 mmHg measurement error (all SD < 4 mmHg) under both resting and non-resting condition (deeper breathing, talking and arm movement condition). This level of accuracy was within the requirement of BP device validation from Association for the Advancement of Medical Instrumentation (AAMI) (average difference no greater than 5 mmHg and SD no greater than 8 mmHg).[16] This finding emphasizes the deep learning method could be used to achieve accurate measurement under both resting and non-resting conditions.

It was also observed that 2.0% of SBP differences and 2.7% of DBP differences were over 8 mmHg between deep learning and manual methods. Figure 5 gives an example that the DBP difference was over 18 mmHg. Due to the low signal amplitude, there were seven manually audible KorS beats that have not been successfully identified by our proposed deep learning method, indicating that the amplitude of KorS beats is an important factor influencing the accuracy of audible KorS identification. In future studies, more manually audible KorS beats with low amplitude are needed for training purpose, or additional signal enhancement algorithm is required to achieve better KorS identification.

Another finding from our study is that measurement conditions had significant influence on measured BPs. This study reconfirmed that talking, arm movement and deeper breathing must be avoided during BP measurement procedure to achieve an accurate and reliable BP measurement.[10,15,17,18] Serval studies have quantitatively reported the effect of different measurement conditions on BP measurement[13,19,20]. The results of the current study using deep learning method agreed with the findings from previous reports where BPs were obtained by manual auscultatory method or clinically validated BP devices. In principle, our deep learning BP measurement method was developed from the auscultatory method. The main difference is that CNN, rather than human ear, was used to identify audible and non-audible Korotkoff sounds. Therefore, the effect of measurement conditions on BP is expected to be similar with manual auscultatory method. More importantly, this study has demonstrated that there was no significant difference in BP changes (differences between non-resting and resting conditions) determined by the manual auscultatory and deep learning methods, providing quantitative evidence that the performance of deep learning BP measurement method is as well as manual auscultatory method under different conditions.

One limitation of this study is that 40 healthy subjects from 23-65 years old with normotension were studied. Although 40 subjects were statistically enough for a technology development study, a future clinical population study is recommended with a bigger sample size in large cohorts. Further studies should be focused on validating the generalizability of our conclusion to children, adolescent population and those older than 65, furthermore to a broader population with existing cardiovascular disease or co-morbidities such as obesity, diabetes, hypertension, hyperlipidemia, or peripheral vascular disease.

Conclusions

In summary, this study has demonstrated that deep learning BP measurement method can achieve accurate measurement under both resting and non-resting conditions.

Acknowledgement

The authors acknowledge the support from the volunteers for participating in the study.

Declaration of interest

The authors report no declarations of interest. This project funded by China Postdoctoral Science Foundation (reference number 2019M653409). The experiment was conducted with the support from the Engineering and Physical Sciences Research Council (EPSRC) Healthcare Partnership Award (reference number EP/I027270/1).

References

[1] Beevers G, Lip G Y, O'brien E. ABC of hypertension: Blood pressure measurement. Part II-conventional sphygmomanometry: technique of auscultatory blood pressure measurement[J]. BMJ, 2001, 322(7293): 1043-7.

[2] Zhang M, Zhang X, Chen F, et al. Effects of room environment and nursing experience on clinical blood pressure measurement: an observational study[J]. Blood Press Monit,

2017, 22(2): 79-85.

[3] Bank I, Vliegen H W, Bruschke A V. The 200th anniversary of the stethoscope: Can this low-tech device survive in the high-tech 21st century?[J]. Eur Heart J, 2016, 37(47): 3536-3543.

[4] Esteva A, Kuprel B, Novoa R A, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. Nature, 2017, 542(7639): 115-118.

[5] Gulshan V, Peng L, Coram M, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs[J]. JAMA, 2016, 316(22): 2402-2410.

[6] Poplin R, Varadarajan A V, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning[J]. Nat Biomed Eng, 2018, 2(3): 158-164.

[7] Kooi T, Litjens G, Van Ginneken B, et al. Large scale deep learning for computer aided detection of mammographic lesions[J]. Med Image Anal, 2017, 35: 303-312.

[8] Chen T E, Yang S I, Ho L T, et al. S1 and S2 Heart Sound Recognition Using Deep Neural Networks[J]. IEEE Trans Biomed Eng, 2017, 64(2): 372-380.

[9] Pan F, He P, Chen F, et al. A novel deep learning based automatic auscultatory method to measure blood pressure[J]. Int J Med Inform, 2019, 128: 71-78.

[10] Muntner P, Shimbo D, Carey R M, et al. Measurement of Blood Pressure in Humans: A Scientific Statement From the American Heart Association[J]. Hypertension, 2019, 73(5): e35-e66.

[11] Cushman W C, Cooper K M, Horne R A, et al. Effect of back support and stethoscope head on seated blood pressure determinations[J]. Am J Hypertens, 1990, 3(3): 240-1.

[12] Tolonen H, Koponen P, Naska A, et al. Challenges in standardization of blood pressure measurement at the population level[J]. BMC Med Res Methodol, 2015, 15: 33.

[13] Zheng D, Giovannini R, Murray A. Effect of respiration, talking and small body movements on blood pressure measurement[J]. Journal Of Human Hypertension, 2011, 26: 458.

[14] Graves J W, Quinn D. Noninvasive Sphygmomanometers—Part 2: Clinical Investigation of Automated Measurement Type. 81060–2, (ANSI/AAMI/ISO, 2013).

[15] Pickering T G, Hall J E, Appel L J, et al. Recommendations for blood pressure measurement in humans and experimental animals: Part 1: blood pressure measurement in humans: a statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Council on High Blood Pressure Research[J]. Hypertension, 2005, 45(1): 142-61.

[16] Stergiou G S, Alpert B, Mieke S, et al. A Universal Standard for the Validation of Blood Pressure Measuring Devices: Association for the Advancement of Medical Instrumentation/European Society of Hypertension/International Organization for Standardization (AAMI/ESH/ISO) Collaboration Statement[J]. Hypertension, 2018, 71(3): 368-374.

[17] O'brien E, Asmar R, Beilin L, et al. European Society of Hypertension recommendations for conventional, ambulatory and home blood pressure measurement[J]. J Hypertens, 2003, 21(5): 821-48.

[18] Williams B, Poulter N R, Brown M J, et al. Guidelines for management of hypertension: report of the fourth working party of the British Hypertension Society, 2004-BHS IV[J].

J Hum Hypertens, 2004, 18(3): 139-85.

[19] Le Pailleur C, Montgermont P, Feder J M, et al. Talking effect and "white coat" effect in hypertensive patients: physical effort or emotional content?[J]. Behav Med, 2001, 26(4): 149-57.

[20] Herakova N, Nwobodo N H N, Wang Y, et al. Effect of respiratory pattern on automated clinical blood pressure measurement: an observational study with normotensive subjects[J]. Clin Hypertens, 2017, 23: 15.

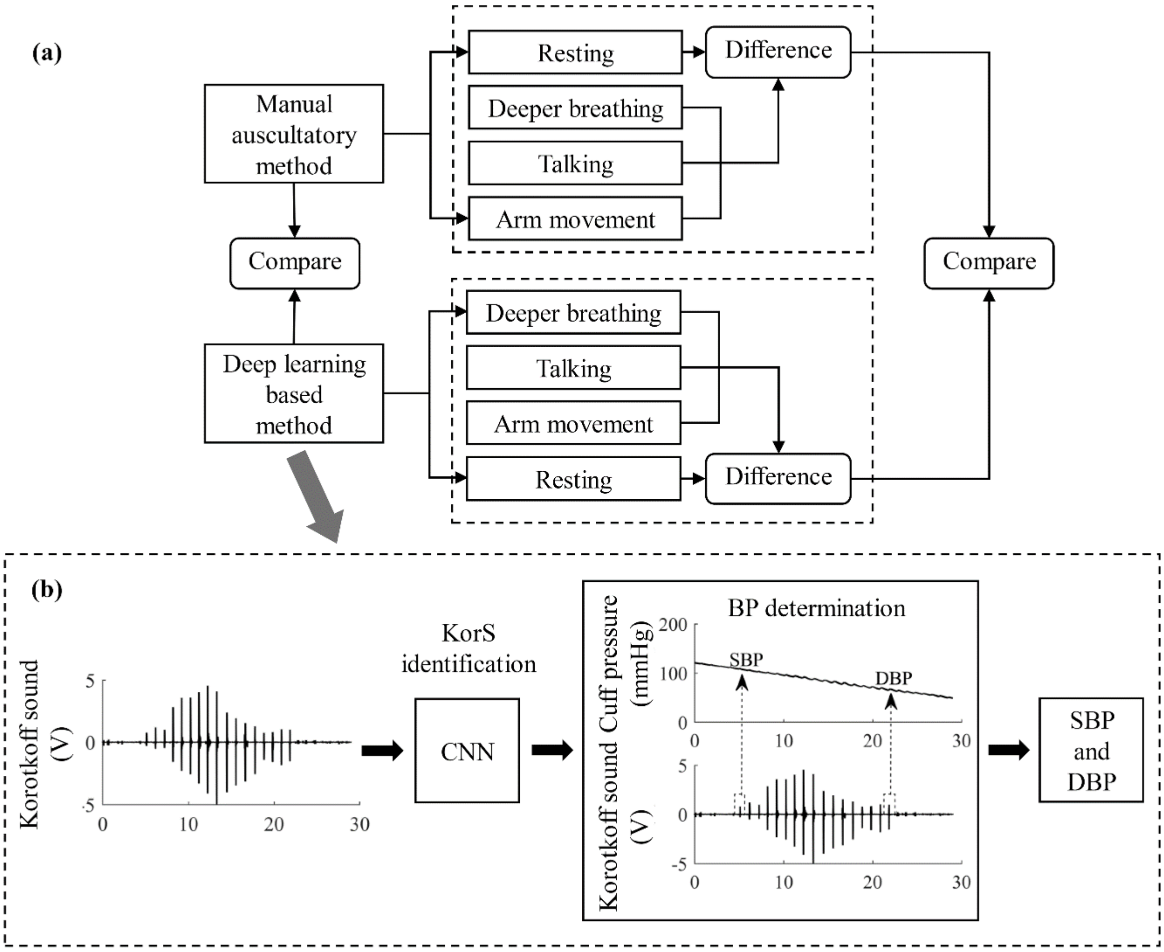

Figure 1. (a) Data and statistical analysis procedure diagram, and (b) flow diagram of BP measurement using deep learning method.

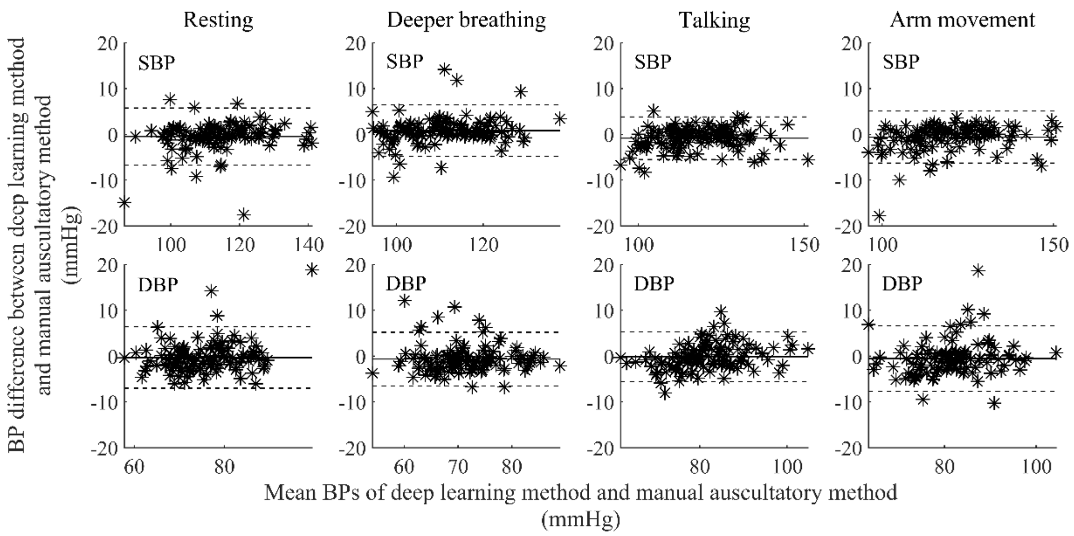

Figure 2. Bland-Altman plots of SBP and DBP from the deep learning method versus manual

auscultatory method. The limits of agreement (1.96 * SD of BP difference) are given using the dashed lines.

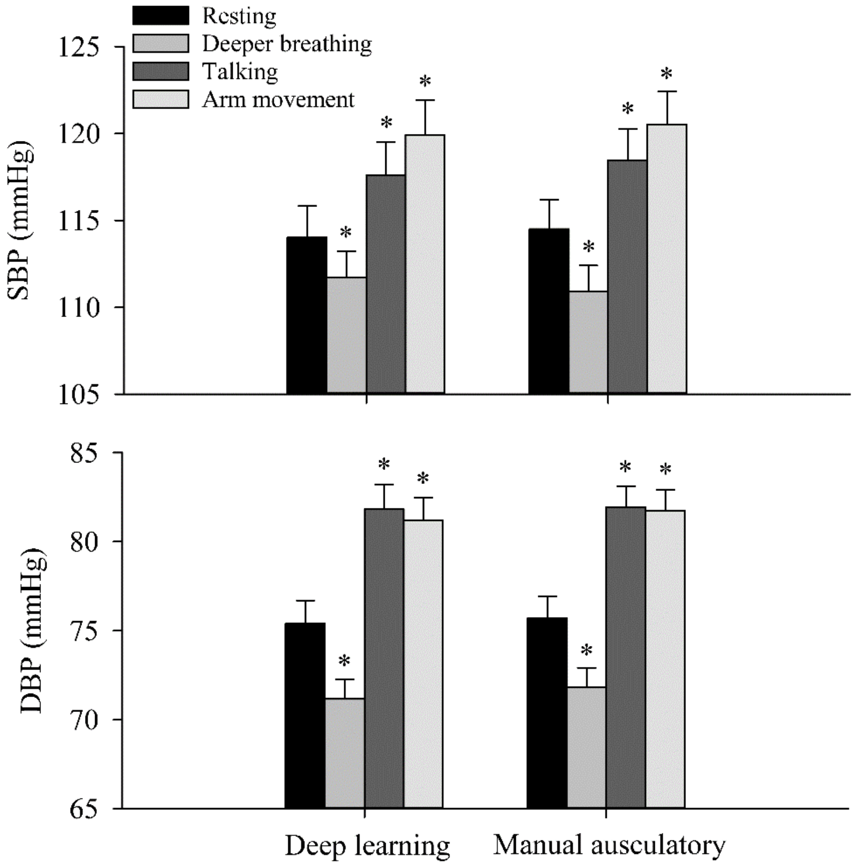

Figure 3. Overall mean + s.e.m. for SBP and DBP for both method under each measurement condition. *Significantly different in comparison with the resting condition.

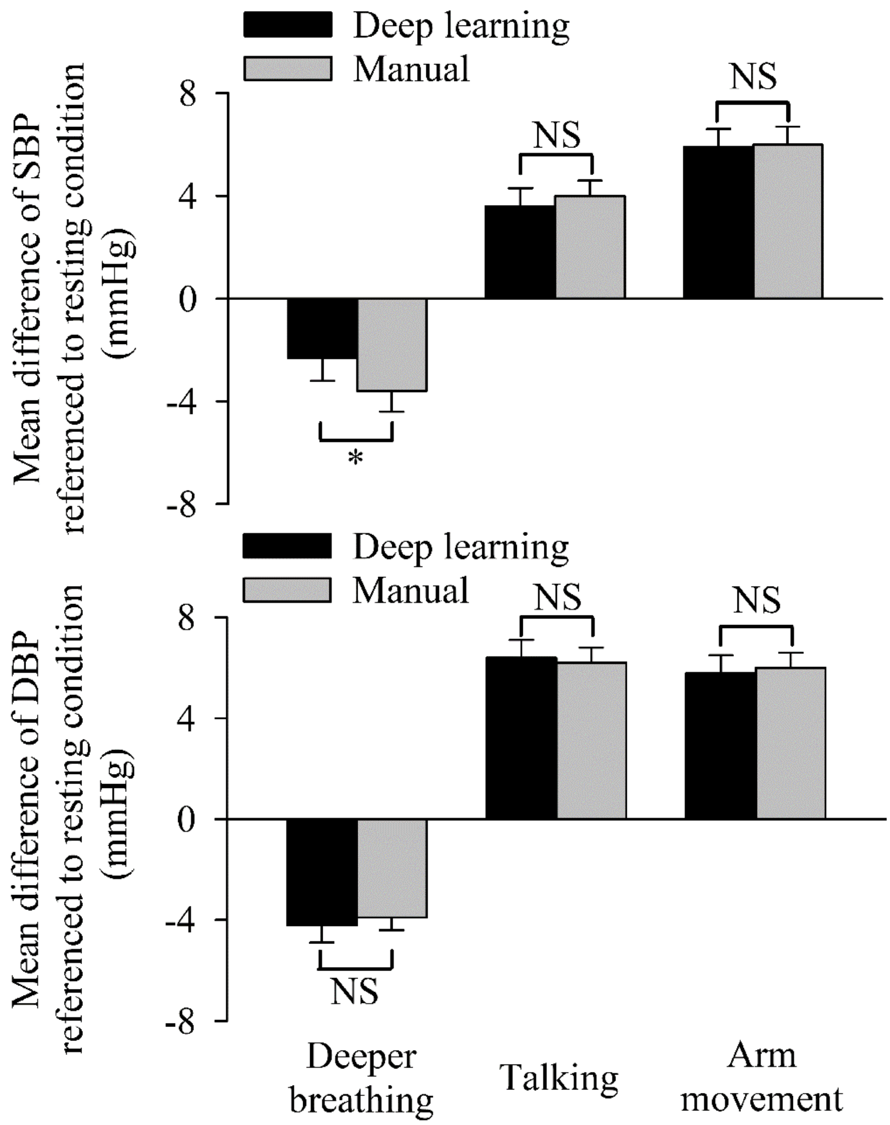

Figure 4. Comparison of BP changes (mean + s.e.m.) measured by the deep learning and manual methods. *Significant difference between comparisons ($P < 0.05$).

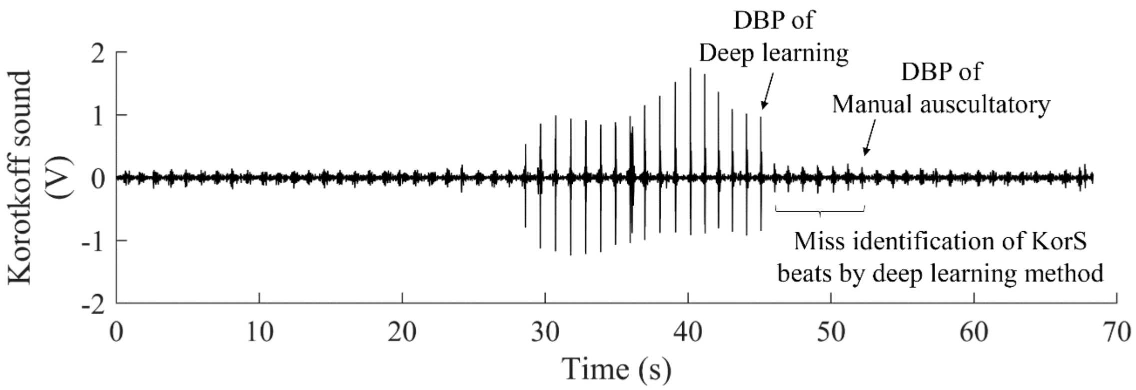

Figure 5. An example of incorrect identification of KorS beats by deep learning method, illustrating the large DBP difference between the deep learning and manual methods.

Table 1 General data information for the subjects studied.

Subject information				
No. subjects	40			
No. male	30			
No. Female	10			
	Min	Max	Mean	SD
Age (years)	23	65	43	12
Height (cm)	152	192	173	10
Weight (kg)	43	105	73	11
Arm circumference (cm)	24	39	28	2.7

Table 2. Over all mean ± SD of BP measured using manual auscultatory and deep learning method, and the BP differences between two methods from different measurement conditions (number of conditions = 4)

Condition	*Manual auscultatory method*		*Deep learning method*		*Mean differences of BP referenced to the manual method*	
	SBP (mmHg)	DBP (mmHg)	SBP (mmHg)	DBP (mmHg)	SBP (mmHg)	DBP (mmHg)
Resting	114.5±10.6	75.7±7.5	114.0±11.2	75.4±8.2	-0.5±3.2	-0.3±3.4
Deep breathing	110.9±9.2	71.8±7.0	111.7±9.7	71.2±6.9	0.8±2.9	-0.7±3.0
Talking	118.5±11.3	81.9±7.9	117.6±11.7	81.8±8.8	-0.9±2.4	-0.1±2.8
Arm movement	120.5±11.9	81.7±7.5	119.9±12.7	81.2±8.1	-0.6±2.9	-0.5±3.6

Table 3 Overall mean differences ±s.e.m. of BP difference measured by deep learning and manual methods, in comparison with the values from the resting condition

Condition	*Mean differences of BP referenced to the resting condition*			
	SBP (mmHg)		*DBP* (mmHg)	
	Deep learning method	*Manual auscultatory method*	*Deep learning method*	*Manual auscultatory method*
Deeper breathing	-2.3±0.9*	-3.6±0.8*	-4.2±0.7*	-3.9±0.5*
Talking	3.6±0.7*	4.0±0.6*	6.4±0.7*	6.2±0.6*
Arm movement	5.9±0.7*	6.0±0.7*	5.8±0.7*	6.0±0.6*

* Significantly different ($P < 0.05$) in comparison with the resting condition.