

HearBP: Hear Your Blood Pressure via In-ear Acoustic Sensing Based on Heart Sounds

Zhiyuan Zhao* Fan Li* Yadong Xie* Huanran Xie* Kerui Zhang* Li Zhang† Yu Wang‡

* School of Computer Science, Beijing Institute of Technology, Beijing, China.

† School of Mathematics, Hefei University of Technology, Hefei, China.

‡ Department of Computer and Information Sciences, Temple University, Philadelphia, PA, USA.

Email: {3120205519, fli, ydxie, huanranxie, 3220221001}@bit.edu.cn, lizhang@hfut.edu.cn, wangyu@temple.edu

Abstract—Continuous blood pressure (BP) monitoring using wearable devices has received increasing attention due to its importance in diagnosing diseases. However, existing methods mainly measure BP intermittently, involve some form of user effort, and suffer from insufficient accuracy due to sensor properties. In order to overcome these limitations, we study the BP measurement technology based on heart sounds, and find that the time interval between the first and second heart sounds (TIFS) of bone-conducted heart sounds collected in the binaural canal is closely related to BP. Motivated by this, we propose HearBP, a novel BP monitoring system that utilizes in-ear microphones to collect bone-conducted heart sounds in the binaural canal. We first design a noise removing method based on U-net autoencoder-decoder to separate clean heart sounds from background noises. Then, we design a feature extraction method based on Shannon energy and energy-entropy ratio to further mine the time domain and frequency domain features of heart sounds. In addition, combined with the principal component analysis algorithm, we achieve feature dimension reduction to extract the main features related to BP. Finally, we propose a network model based on dendritic neural regression to construct a mapping between the extracted features and BP. Extensive experiments with 41 participants show the average estimation error of 0.97mmHg and 1.61mmHg and the standard deviation error of 3.13mmHg and 3.56mmHg for diastolic pressure and systolic pressure, respectively. These errors are within the acceptable range specified by the FDA’s AAMI protocol.

I. INTRODUCTION

Blood pressure (BP) is a crucial indicator in the assessment of cardiovascular health and disease management, which can provide doctors with in-depth insights for diagnosis. Hypertension is a condition in which BP is continuously elevated and is one of the critical risk factors for cardiovascular disease [1]. Hypotension, in contrast, may indicate cardiac or renal disease, syncope, coma, or even shock [2]. In addition, once BP continues to rise without control, it can lead to many life-threatening diseases, such as myocardial infarction, cerebral infarction, or renal failure [3]. Therefore, accurate and continuous BP measurement is of great significance for people’s daily lives.

The “gold standard” is placing specific medical devices in the patient’s arterial blood vessels [4]. However, it is technically challenging to place medical devices such as arterial

catheters in the human body, especially in children, and it can also cause pain and even infection to patients [5]. As a safe and convenient alternative method, non-invasive BP measurement has received increasing attention. The commonly used clinical non-invasive methods mainly utilize BP sensors with inflatable cuffs to compress arterial blood vessels, and measure BP based on the special sounds (e.g., Korotkoff sound) of blood vessels under different pressures and various oscillation phenomena [6], [7]. However, cuff devices typically have a relatively single function and poor portability.

In recent years, the development of wearable devices has provided new ideas for BP measurement. A more convenient way can be achieved by embedding an inflatable device into a wristband device [8], [9] or an in-ear device [10], [11]. But these wearable BP monitors also need to compress blood vessels to cause comfort issues and cannot achieve continuous BP measurement. In order to achieve more comfortable BP measurements, many works utilize electrocardiography (ECG), photoplethysmography (PPG), and seismocardiography (SCG) to calculate pulse transit time (PTT) and pulse arrival time (PAT) [12]–[15]. However, these methods usually rely on users to perform special operations (for example, the user needs to continuously press the wristband device on the chest), which increases the user’s burden and is not convenient for long-term continuous measurement. Therefore, wearable devices that can measure BP without user effort have emerged. For example, BP measurement is accomplished by integrating accelerometers into compression shorts to obtain PTT [16]. However, this work is limited by sensors, which is only suitable for BP measurement when the body is stationary. In addition, reflected wave transit time (RWTT) and pulse time difference (PTD) can also be extracted from the PPG signal for continuous BP measurement [17], [18]. However, PPG sensors are usually affected by lighting conditions, sensor location, and skin color, thereby reducing the accuracy of BP measurement.

Motivated by the above limitations, we design and implement a convenient, comfortable, and continuous in-ear BP monitoring system, HearBP, through an acoustic sensing scheme based on heart sounds. The key inspiration for the design comes from the following aspects. Firstly, due to the complex vascular structure of the ear, it is a perfect location for sensing vital signs and heart-related biometrics [19]. Secondly, existing work has explored the possibility of using in-ear

Fan Li is the corresponding author. The work of Fan Li is partially supported by the National Natural Science Foundation of China (NSFC) under Grant No. 62372045 and No. 62072040. And the work of Li Zhang is partially supported by the NSFC under Grant No. 61972131.

microphones to capture heart sounds (a.k.a., phonocardiogram or PCG) [20]. Furthermore, we introduce the time interval between the peak points of first PCGs and second PCGs (TIFS), and explore the possibility of using PCGs for BP measurement. In recent years, acoustic sensing technology has made great progress [21], [22]. In-ear smart devices equipped with microphones and speakers are gradually becoming new sensing platforms, and their market size is maintaining rapid growth. According to a survey [23], the number of wireless earphones will reach 1.3 billion by 2028, and more than 60% of adults have been found to wear earphones for more than one hour per day [24]. Meanwhile, many types of commercial earphones are equipped with multiple microphones for noise reduction. In addition, Nokia Bell Labs releases eSense, an in-ear sensing platform, which further arouses the research community’s attention on the development and computation of in-ear sensing [25]. These principles motivate the basic idea of HearBP: use an in-ear microphone to capture bone-conducted PCGs, and then extract BP-related features from PCGs for BP measurement.

With the goal of developing an earphone-oriented continuous BP monitoring solution, we first identify and discuss several technical challenges that need to be addressed. First of all, the PCG is very weak and can easily be overwhelmed by strong background noises (e.g., earphone audio, bone-conducted noises generated by daily activities, and other body sounds). Therefore, it is necessary to separate the weak PCGs from strong noises without affecting the general use of earphones. Secondly, although TIFS features extracted from bone-conducted PCGs show a correlation with BP, extracting more unique and reliable BP-related features from PCGs is still an unexplored field. Last but not least, when using these features to establish a relationship mapping with BP, it is necessary to choose appropriate machine learning algorithms and modeling methods. This model needs to undergo rigorous training and validation to ensure the accuracy of BP measurement and the stability of the model.

In order to solve the above challenges, we first design a noise elimination method that utilizes a U-Net autoencoder-decoder-based deep learning architecture to map the log-mel spectrograms of the collected PCGs with noises to the log-mel spectrograms of the clean PCGs. Secondly, we find through experiments that relying solely on the extracted TIFS features cannot accurately measure BP. Therefore, we further mine more time domain and frequency domain features based on shannon energy and energy-entropy ratio. In addition, we also use principal component analysis (PCA) to reduce the dimension of these features to extract the main features related to BP, thereby reducing feature redundancy and computational overhead. Finally, we propose a network model based on dendritic neural regression (DNR) [26] to establish a mapping relationship between the above features and BP, which can ensure that the error of BP measurement meets the standards of the Association for the Advancement of Medical Instruments (AAMI) [27] and meets the grade A standards of the Britain Hypertension Society (BHS) [28]. In summary, we make the

following major contributions in HearBP:

- We propose a novel BP monitoring system, HearBP, which uses earphones to collect bone-conducted PCGs in the ear canal. To the best of our knowledge, we are the first to sense PCGs through earphones for BP measurement. HearBP is low-cost, comfortable, convenient, and supports long-term monitoring, which makes it suitable for widespread adoption.
- We propose an effective noise elimination method, which can reduce the influence of earphone background audio, bone-conducted noises generated by daily activities, and other body sounds.
- We fully mine the time and frequency domain features of PCGs based on shannon energy and energy-entropy ratio, and use PCA to reduce the dimension of features to obtain the main features related to BP. In addition, we build a mapping relationship between BP and PCG features based on DNR.
- We implement HearBP and evaluate it with 41 volunteers under various scenarios. The results show that the estimation errors of HearBP for diastolic blood pressure (DBP) and systolic blood pressure (SBP) are 0.97 ± 3.13 *mmHg* and 1.61 ± 3.56 *mmHg*, respectively.

II. RELATED WORK

In this section, we review 3 kinds of BP measurement systems related to HearBP, highlight their properties, and compare the pros and cons.

Professional equipment is widely used as a safe and non-invasive BP measurement method. The idea is to use a BP sensor with an inflatable cuff to compress the arterial blood vessels, and measure BP according to the special sound and various oscillation phenomena of the arterial blood vessels under different pressures, such as mercury or electronic sphygmomanometer [29]. But when the cuff is inflated, it can block blood flow, causing users to feel uncomfortable or even painful, and continuous measurement can lead to various risks such as tissue hypoxia and injury [30].

Wristband sensors make portable BP measurements possible in daily life. HUAWEI WATCH D [8] and eBP [10] can achieve accurate BP measurement in the wristband and ear-worn devices, respectively, but they still rely on inflatable cuffs to compress blood vessels and cause comfort issues. SeismoWatch [14] places the watch against the sternum and measures PTT using accelerometers and optical sensors to estimate BP. However, it generally relies on users performing specific operations, which is inconvenient for continuous BP measurement. In addition, Crisp-BP [17] and Stereo-BP [18] measure PPG signals through wristband PPG devices, and they further extract RWTT and PTD features to achieve convenient and continuous BP measurement. But PPG sensors are usually sensitive to location, lighting, and skin color information, which affects the accuracy of BP measurement.

Acoustic sensors have gradually emerged in the field of BP measurement due to the increasing maturity of acoustic technology. Some researchers attempt to calculate PTT by

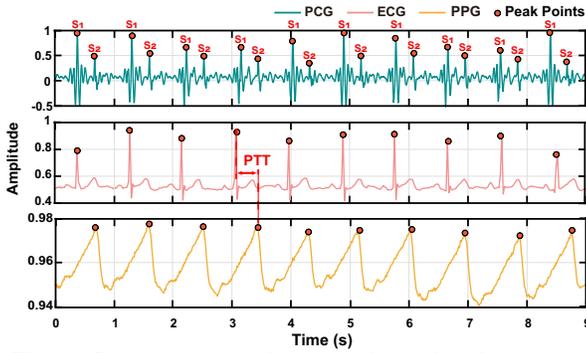


Fig. 1: Correspondence between three signals in time.

recording ECG (or PCG) and PPG using the built-in microphones of smartphones, respectively [31], [32]. But these methods usually require users to perform special operations to disturb them. Peng et al. [33] design and fabricate a flexible piezoelectric composite ultrasound sensor for continuous BP measurement by tracking the vessel wall through ultrasonic motion. However, the use of couplant gel between the flexible sensor and human skin during measurement is not practical for daily continuous BP monitoring.

Compared with existing work, HearBP is a novel BP measurement system that only adopts in-ear microphones in ear-worn devices to acquire PCGs. It is less affected by the environment, does not disturb users, and can continuously measure BP more comfortably and conveniently.

III. FEASIBILITY ANALYSIS

Thanks to the amplification enhancement caused by the occlusion effect [34], the in-ear microphone can capture the bone-conducted PCGs with improved SNR (signal-to-noise ratio). To verify the feasibility of using bone-conducted PCGs for BP measurement, we conduct a pilot study with 6 users.

A. Relationship between PCG, ECG and PPG in Time

PTT is linearly correlated with BP within a certain time range, and it can be calculated using the time interval between the R wave peak points of ECG and the corresponding peak points of the PPG. As the first PCGs (S_1) and second PCGs (S_2) are the two main components of PCG, we study the corresponding relationship among the peak points of these two components, the peak points of the R wave, and the peak points of PPG in time. We synchronously collect the PCG, ECG, and PPG of 6 users. Fig. 1 shows a user's PCG, ECG, and PPG. It can be seen that the peak points of S_1 and S_2 correspond significantly with the peak points of R-wave and PPG in time, respectively.

Going a step further, we verify this correspondence through experiments. We fit the time data corresponding to the peak points of S_1 and R waves of 6 users, as well as the peak points of S_2 and PPG, as shown in Fig. 2. We choose the goodness of fit R^2 as the evaluation indicator for the degree of fit. R^2 is between $[0, 1]$, and the closer it is to 1, the better the regression line fits the observed values, the stronger the linear relationship, and vice versa. By calculation, the R^2 of

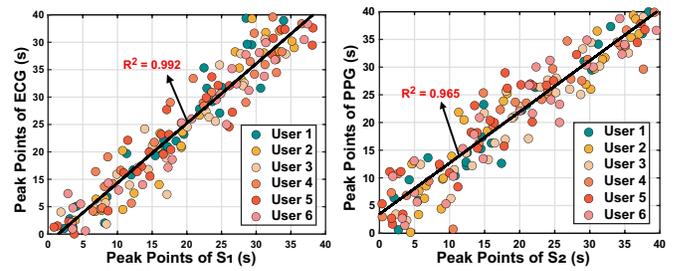


Fig. 2: Correspondence between peak points in time.

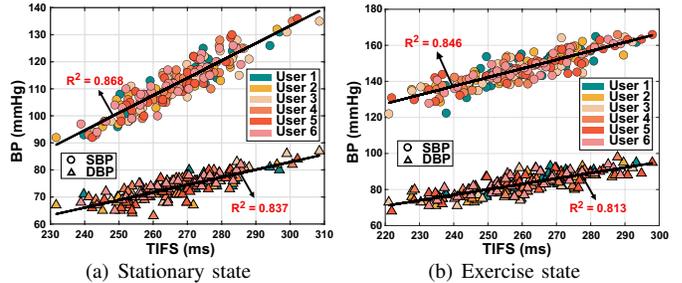


Fig. 3: Correspondence between BP and TIFS.

the time data corresponding to the peak point of S_1 and R wave is equal to 0.992, and the R^2 of the time data corresponding to the peak point of S_2 and PPG is equal to 0.965. Such results also confirm that the peak point of S_1 and R wave, and the peak point of S_2 and PPG have a strong correlation in time. PTT can be calculated using the time interval between the peak point of the R wave and the corresponding peak point of the PPG. Naturally, we can conclude that the time interval between the peak points of S_1 and S_2 (TIFS) may replace PTT. Next, we verify this conclusion through the PCG and BP readings of 6 users.

B. Relationship between TIFS and BP

We verify the correlation between TIFS and SBP as well as DBP in two states. In one case, 6 users are in a static state, and in the other case, they are asked to perform the same movement (30 jumping jacks). As shown in Fig. 3, the R^2 between TIFS and SBP as well as DBP is greater than 0.8 no matter what the case. These results also further confirm that TIFS is related to BP, which provides a research basis for using PCGs to monitor BP. Of course, it should be pointed out that through subsequent experiments, we demonstrate that BP measurement solely based on TIFS cannot meet current international standards for BP measurement (e.g., AAMI and BHS). Therefore, we explore the extraction of more BP-related features from PCGs to further improve the accuracy of BP measurement in Section IV-D.

IV. SYSTEM DESIGN

In this section, we first present the system overview of HearBP and then detail the behind techniques.

A. Overview of System Architecture

Fig. 4 shows the architecture of HearBP, which can be divided into two phases, i.e., offline phase and online phase.

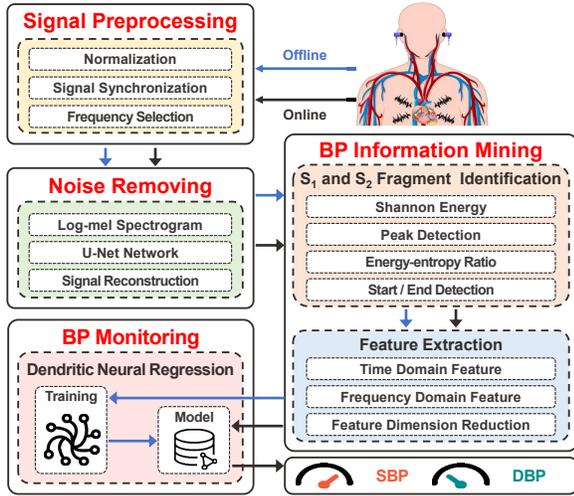


Fig. 4: System architecture of HearBP.

In the offline phase, HearBP aims to collect data from users and train BP monitoring regression models. The system uses earphones with in-ear microphones to continuously collect bone-conducted PCGs from the user’s left and right ears. The recorded sounds first undergo *Signal Preprocessing*, including normalization, signal synchronization, and frequency selection, to maintain stable sound volume, ensure better synchronization between left and right ear audio, and eliminate high-frequency noise interference. Then, HearBP performs *Noise Removing* to recover clean binaural bone-conducted PCGs. Specifically, we extract the log-mel spectrogram of sounds and ground truth PCGs, establish the mapping relationship between them by designing a U-net autoencoder-decoder network to remove noise interference, and reconstruct clean bone-conducted PCGs using spectrum inversion algorithm. After that, *BP Information Mining* module begins to explore the time domain and frequency domain features of PCGs based on the shannon energy peak point detection algorithm and the energy-entropy ratio endpoint detection algorithm, respectively. In addition, we also use PCA algorithm to reduce the dimension of these features to obtain the main features related to BP, which reduces the computational complexity while reducing feature redundancy. Finally, the extracted features are used to train the DNR neural network in *BP Monitoring*.

In the online stage, HearBP first records bone-conducted sounds through a pair of in-ear microphones. Then, through *Signal Preprocessing*, *Noise Removing*, and *BP Information Mining*, the features are sent to *BP Monitoring* module to achieve the purpose of measuring BP.

B. Signal Preprocessing

1) *Normalization*: When users wear earphones each time, the tightness and angle may vary slightly, resulting in unstable sound volume recorded by the in-ear microphone. Therefore, before performing other operations, the sound should be normalized first. We employ a loudness-based standardization by adjusting the average volume of the sound to the standard loudness ($-24dB$) recommended by ATSC [35].

2) *Signal Synchronization*: We use a pair of in-ear microphones to collect bone-conducted sounds, but because they conduct to the two ear canals through different paths, there is a time difference in the sounds. To ensure better synchronization between the sounds recorded by the two microphones, we calculate the cross-correlation between the pair of sounds. Then, we adjust the sounds based on the relative time delay indicated by the maximum cross-correlation value.

3) *Frequency Selection*: We use a low-pass filter to select the frequency band. Since the frequency range of bone-conducted PCG is generally below $160Hz$ [36], we use a low-pass filter with a cut-off frequency of $160Hz$ for bone-conducted sounds to eliminate high-frequency interference.

C. Noise Removing

Although we have eliminated the interference of high-frequency noises through *Frequency Selection*, there are still three significant sources of interference, including earphone background audio, bone-conducted sounds caused by daily activities (e.g., walking sounds), and other body sounds (e.g., lung sounds). Intuitively, we can extract clean PCGs through empirical mode decomposition and wavelet-based methods. However, these methods usually rely on the assumption of low aliasing of PCG and interference, which is not applicable in our case. Therefore, we design an innovative interference removal method based on U-net autoencoder-decoder network. Its core idea is to map the log-mel spectrogram of the preprocessed bone-conducted sounds to the log-mel spectrogram of the real PCGs (also after *Signal Preprocessing*).

1) *Log-mel Spectrogram*: First, we divide the bone-conducted sounds and ground truth PCGs after *Signal Preprocessing* into frames, with a frame length of 256 samples and a frame shift of 32 samples. Then, hamming window function is applied to each frame to highlight the short-term features of the audio. Furthermore, we employ FFT to compute log-mel spectrograms of bone-conducted sounds and real PCGs of each window. Here we use 1,024 FFT bins with zero padding and 64 Mel bins. The resulting log-mel spectrogram is a 64×64 matrix of audio of each window. The reason why the log-mel spectrogram is chosen instead of the mel spectrogram is that the former can compress the high-frequency part and provide more detailed information in the low-frequency region where the PCGs frequency is located. Another benefit of doing so is that it eliminates redundant information and reduces data dimensions, thereby improving the efficiency of data processing. It is worth noting that since bone-conducted sounds are captured through both ears, we compute log-mel spectrograms for each channel and stack them together to form a $64 \times 64 \times 2$ input. And the output is the log-mel spectrogram of single-channel PCGs. In addition, we use the maximum normalization method to standardize the log-mel spectrogram between $[0, 1]$ to facilitate network training.

2) *U-Net Network*: Fig. 5 shows the U-Net autoencoder-decoder structure for noise removal. In the autoencoder, the model consists of repeated blocks, each of which contains a 3×3 convolutional layer, a batch normalization layer, a ReLU

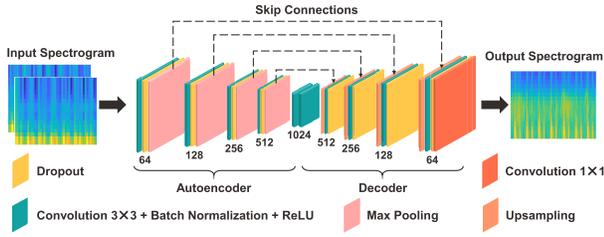


Fig. 5: Structure of the U-net autoencoder-decoder.

activation function layer, and a maximum pooling layer with a stride of 2 to downsample the data. In addition, we also add a dropout layer in each block to avoid overfitting. During each downsampling, the number of feature maps is doubled to enable the network to learn complex structures in the data. Whereas in the decoder, the data undergoes successive upsampling, and the number of feature maps is halved at each step. After each upsampling, the feature map is merged with the corresponding feature map from the autoencoder, and then undergoes convolution and batch normalization with operations similar to those in the autoencoder. In the last layer, a 1×1 convolution is used to map the final feature map into a single 64×64 output image.

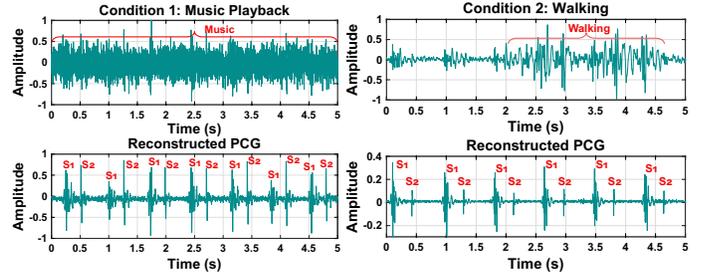
When training the network, the input is the spectrograms of bone-conducted sounds and their corresponding PCGs spectrograms. We use leave-one-out cross-validation for testing, that is, one user's data is used for testing and the other users' data is used for training the model. We empirically train the model for 150 epochs using the Adam optimizer with a learning rate of 0.001 and a batch size of 128. We choose the mean squared error $\sum (s_{ij} - \tilde{s}_{ij})^2 / TF$ as the loss function, where i and j denote the time and frequency indices, respectively. T and F denote the total number of bins in the time and frequency dimensions, respectively. This loss minimizes the distance between the spectrograms \tilde{s}_{ij} of ground truth PCGs and the spectrograms s_{ij} of bone-conducted sounds.

3) *Signal Reconstruction*: Our purpose is to convert the reconstructed clean log-mel spectrogram into a time domain waveform in order to mine more BP-related features contained in PCGs. To this end, we employ the Griffin-Lim algorithm [37] for spectrum inversion, with each converted waveform having a duration of 3s and an overlap of 2s. Then, the transformed waveforms are merged into a continuous time-series signal by averaging the overlapping regions.

Fig. 6 shows examples of PCGs extraction under two different conditions (e.g., music playback and walking), respectively. It can be seen that the impact of these noises is significantly reduced after applying our proposed method. We also conduct experiments in Section V-D1 to further demonstrate the effectiveness of the noise removal method.

D. BP Information Mining

After obtaining clean bone-conducted PCGs, the next step is to dig out more features related to BP to make up for the fact that the accuracy of BP measurement cannot meet the actual needs when only relying on TIFS. Although we currently



(a) Music playback

(b) Walking

Fig. 6: Examples of noise removing.

cannot fully understand the relationship between BP and PCGs, some existing works [38], [39] are able to demonstrate its association with symptoms such as cardiac abnormalities and pulmonary hypertension by analyzing the time domain and frequency domain features of PCGs. Therefore, similarly, we also start with the time domain and frequency domain features of PCGs to find the relationship between them and BP. The extraction of these features mainly relies on the two main components S_1 and S_2 of PCGs, so the first step is to recognize the fragments of S_1 and S_2 .

1) *S_1 and S_2 Fragment Identification*: S_1 and S_2 fragment identification mainly includes two parts, one is to identify the peak points of S_1 and S_2 , and the other is to identify their start and end times. We design a peak point detection algorithm based on shannon energy envelope and an endpoint detection algorithm based on energy-entropy ratio.

Shannon Energy Calculation: We first normalize clean bone-conducted PCGs to the range $[-1, 1]$ by dividing by their absolute maximum. Then the shannon energy is calculated by $E = -x^2 \log(x^2)$, where x is the normalized bone-conducted PCGs. Next, the bone-conducted PCGs are divided into frames, with a frame length of 25ms and a frame shift of 10ms. We average the shannon energy over a sliding time window of 25ms as $E_{ave} = \sum_1^n E/n$. n is the length of the window. In addition, E_{ave} is normalized by subtracting its average value $m(E_{ave})$ and dividing by its standard deviation $s(E_{ave})$ as $E_N = [E_{ave} - m(E_{ave})]/s(E_{ave})$.

Peak Detection: We apply two thresholds to E_N to identify potential peaks of S_1 and S_2 . We set the high threshold T_1 as a preset coefficient for the maximum amplitude average values to detect high amplitude peaks. The low threshold T_2 is set to half of the average shannon energy envelope to obtain low amplitude peaks that may be considered noise by the high threshold. According to the fact that the systolic period is shorter than the diastolic period, these peaks are assigned to S_1 or S_2 , respectively. Fig. 7 shows the process of peak point detection. The upper figure shows the shannon energy envelope of bone-conducted PCGs, and the lower figure shows the results obtained through the peak point detection algorithm. It can be seen that our proposed algorithm can effectively locate the peak positions of S_1 or S_2 .

Energy-entropy Ratio Calculation: The energy-entropy ratio is calculated from the energy and spectral entropy, so we introduce the definitions of these two quantities, respectively.

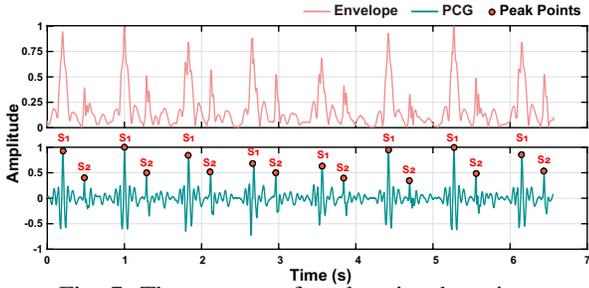


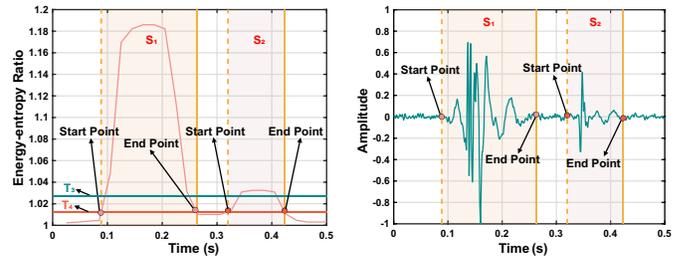
Fig. 7: The process of peak point detection.

Let $x_k(m)$ be the k -th frame of bone-conducted PCGs obtained after windowing and framing processing, and the frame length is n , then the energy of this frame can be defined as $LE_k = \lg(1 + \sum_{m=1}^n x_k^2(m)/a)$, where a represents a constant used to alleviate energy fluctuations. Then we perform FFT on $x_k(m)$, and the normalized spectral probability density function of the frequency component f_l of the l -th spectral lines can be defined as $p_k(l) = Y_k(l) / \sum_{r=0}^{N/2} Y_k(r)$, where $Y_k(l)$ is the energy spectrum of f_l after $x_k(m)$ undergoes FFT. N is the length of the FFT and is set to 1,024. The spectral entropy of each $x_k(m)$ can be defined as $H_k = -\sum_{l=0}^{N/2} p_k(l) \log p_k(l)$. Therefore, the energy entropy ratio of $x_k(m)$ can be expressed as $EEF_k = \sqrt{1 + |LE_k/H_k|}$. Fig. 8(a) shows the energy-entropy ratios corresponding to S_1 and S_2 . The results indicate that the energy-entropy ratio can effectively capture S_1 and S_2 .

Start/End Detection: Then, we employ the dual-threshold method to detect the start and end points of S_1 and S_2 . Specifically, we first set a high energy-entropy ratio threshold T_3 and the segment which is larger than T_3 can be considered to contain S_1 and S_2 . Then, a low threshold T_4 is set to find the start and end points of S_1 and S_2 . We search from the beginning of the segment to the left and find the first point that intersects with T_4 as the start point of S_1 and S_2 . Similarly, we search from the ending of the segment to the right to find the end points of S_1 and S_2 . Fig. 8(b) shows the results of using the energy-entropy ratio to detect the start and end points of S_1 and S_2 . It can be seen that the dual-threshold method accurately identifies the start and end points.

2) *Feature Extraction:* After obtaining the peak point positions, and start and end points of S_1 and S_2 , we extract their time domain and frequency domain features.

Time Domain Feature: In addition to TIFS, we first extract the absolute maximum amplitudes of S_1 and S_2 based on their peak position information. Then, the duration of S_1 (from the start point of S_1 to the end point of S_1), the duration of S_2 (from the start point of S_2 to the end point of S_2), the systolic time interval (from the start of the previous S_1 to the start of the next S_2), the diastolic time interval (from the start of the previous S_2 to the start of the next S_1), the kurtosis of S_1 and S_2 , the energy of S_1 and S_2 are extracted based on the start and end points information. In particular, the kurtosis is used to describe the degree of smoothness of a signal, which can be calculated as $K = \int_{-\infty}^{+\infty} [x(t) - \tilde{x}]^4 \rho(x) dx / \sigma^4$. Where



(a) Corresponding energy-entropy ratio

(b) Endpoint detection results

Fig. 8: The process of endpoint detection.

$x(t)$ is the instantaneous amplitude, \tilde{x} is the mean amplitude, $\rho(x)$ is the probability density, and σ is the standard deviation. When K is less than 4, it indicates a smoother signal, and when K is greater than 4, it indicates a steeper signal.

Frequency Domain Feature: Fragments of S_1 and S_2 can be obtained through the start and end points of S_1 and S_2 , respectively. We perform FFT on these fragments to obtain the spectrum, and then use 40 spectral values with $4Hz$ intervals as spectral features in the 0 to $160Hz$ frequency band and express them as $F_{S_1} = [f_4^{S_1}, f_8^{S_1}, \dots, f_{160}^{S_1}]_{1 \times 40}$, $F_{S_2} = [f_4^{S_2}, f_8^{S_2}, \dots, f_{160}^{S_2}]_{1 \times 40}$. Where $f_u^{s_i}$ is the spectral value of S_1 or S_2 at the frequency uHz .

In addition, considering that BP is also related to demographic factors [17], we also add the user's gender, age, height, weight, and body mass index information to the features. In the end, we extract a total of 96 dimensional features. However, whether these features are closely related to BP needs further analysis, and the high dimension of features is also a burden for network training. For these reasons, we next use PCA to reduce the dimension of the features.

Feature Dimension Reduction: Based on the variance distribution of features, PCA can retain the most important principal components and discard the unimportant principal components to achieve feature dimension reduction. We calculate the variance ratios of the extracted 96 dimensional features and sort them according to the magnitude of the variance ratios, as shown in Fig. 9. We also calculate the cumulative variance ratio of all features. It can be seen from Fig. 9 that TIFS does have the highest variance ratio, but the value is only 63.79%, which is also an important reason why it is impossible to rely solely on TIFS for accurate BP measurement. Through the cumulative variance ratio curve, we can find that the cumulative variance ratio of the first 15 principal components can reach 97.13%, and the other 14 time domain and frequency domain features except TIFS contribute 31.51% of the variance ratio, which also provides favorable conditions for the accurate measurement of BP. In addition, the growth rate of the cumulative variance ratio gradually slows down and becomes stable from the 16-th principal component, and a total of 81 principal components from the 16-th to the 96-th provide only a variance ratio of 2.87%. The main reason is that the eigenvalues of these features are almost zero, which means they do not provide additional information. Therefore, we can conclude that the majority of useful information is contained

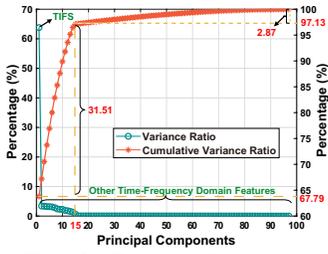


Fig. 9: Feature dimension reduction.

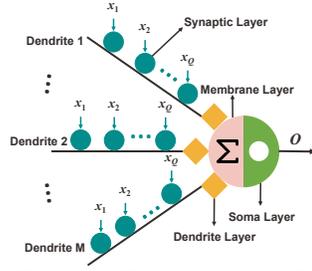


Fig. 10: The structure of DNR.

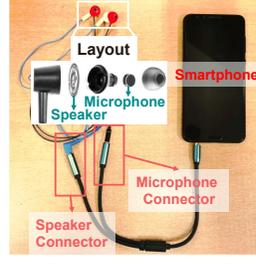


Fig. 11: HearBP prototype.

Methodology	Formula
Mean Error	$\frac{\sum_{i=1}^n (p_i - r_i)}{n}$
Standard Deviation	$\sqrt{\sum_{i=1}^n (p_i - r_i - ME)^2 / n}$
Goodness of Fit	$\frac{\sum_{i=1}^n (p_i - \bar{r})^2}{\sum_{i=1}^n (r_i - \bar{r})^2}$

Fig. 12: Evaluation methodology.

in the first 15 principal components, and discarding the latter 81 principal components does not result in a significant loss of information. If the first 15 principal components are retained, the feature dimension can be reduced from 96 to 15 to achieve approximately 6 : 1 compression, which can reduce the computational complexity while reducing feature redundancy. We also prove this conclusion in Section V-D2 to illustrate the effectiveness of using PCA for feature dimension reduction.

E. BP Monitoring

We design a network model to establish a mapping between extracted features and BP based on DNR. It consists of four layers: synaptic layer, dendritic layer, membrane layer, and soma layer, as shown in Fig. 10.

As the entrance to the model, synapses transmit electrical or chemical signals from other neurons. To simulate the process of signal conversion, the synaptic layer of the DNR dynamically forms the output of the i -th synapse located in the j -th dendritic branch using the equation $Z_j(x_i) = 1/[1 + e^{-k(\omega_{ij}x_i + \theta_{ij})}]$, where x_i represents the i -th input feature, and k is a positive constant. ω_{ij} and θ_{ij} are two types of parameters that optimization algorithms need to train, which determine the neural architecture and model performance of DNR. The dendritic layer processes the output of multiple synapses in the synaptic layer through a multiplication function. The output function of the j -th dendritic branch is represented as $V_j = \prod_{i=1}^Q Z_j(x_i)$, where Q denotes the number of dendritic branches. All signals from the dendritic layer in DNR are linearly summed in the membrane layer. The output definition of the membrane layer is as $D = \sum_{j=1}^M (\varepsilon_j * V_j)$, where ε_j is the strength of the j -th dendritic branch, and M is the number of dendritic branches. Finally, the soma layer fulfills the function of the soma, and neurons fire if the output of the membrane layer exceeds its threshold. This process is expressed using a sigmoid function $O = 1/[1 + e^{-k(D-\gamma)}]$ to calculate the final output of the entire model, where γ represents the threshold that changes from 0 to 1.

In addition, we utilize the AMSGrad to optimize the model structure of DNR. As a gradient-based optimization algorithm, AMSGrad is based on the minimum square error between the actual output O and the desired target T (SBP or DBP).

V. IMPLEMENTATION AND EVALUATION

In this section, we introduce the implementation details and provide the evaluation results.

A. Experiment Setup

1) *Implementation*: Although in-ear microphones are integrated into existing commercial earbuds, we cannot obtain sound data due to hardware limitations. Therefore, we develop a prototype system as shown in Fig. 11, including a speaker and a microphone located about 0.5cm in front of it, and the specific layout of the prototype is given in the white line box. In addition, we use an audio splitter cable to connect the earphones to the smartphone to receive the sound collected by the in-ear microphone. The device also has silicone earbud tips with user-specific sizes to ensure good ambient noise isolation.

2) *Data Collection*: We recruit 41 participants (24 males and 17 females, aged from 19 to 54) with no known health conditions relevant to the evaluation. We conduct experiments with four different conditions (e.g., music playback, walking and speaking, walking, and stationary state) in four environments (e.g., laboratory, dormitory, home, and playground). At the same time, trained experts use a digital stethoscope to record the ground truth of PCGs in the heart valve area. The ground truth is provided by an FDA-approved cuff BP measurement device Omron U30. BP measurements are alternated between the cuff device and HearBP with a measurement interval of 2 minutes. In addition, in order to expand the scope of HearBP for BP measurement, we require each user to have at least one exercise (30 jumping jacks) each time when collecting data. We conduct at least 5 data collections for each participant over a period of three months. Finally, we collect more than 1,012 minutes of binaural bone-conducted PCGs, of which 60% of the data are used for training and 40% for testing. All procedures are approved by the Institutional Review Board (IRB) at our institute.

3) *Evaluation Methodology*: We use three measurement methods shown in Fig. 12 to evaluate the accuracy, including mean error (ME), standard deviation of mean error (STD), and goodness of fit (R^2). Where r_i is the BP reference value from ground truth recordings. p_i is the predicted value of HearBP. \bar{r} is the mean value of reference BP. n is the number of samples.

B. Overall Performance

Fig. 13 shows the Bland-Altman diagram for the estimated SBP and DBP. The red line represents the mean error, and the pink line represents the limit of agreement (LOA, defined as $ME \pm 1.96 \times STD$). The mean errors of SBP and DBP are 1.61mmHg and 0.97mmHg, respectively, and the standard

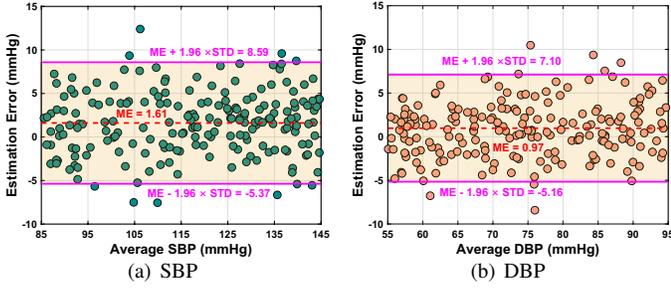


Fig. 13: Bland–Altman diagram of SBP and DBP.

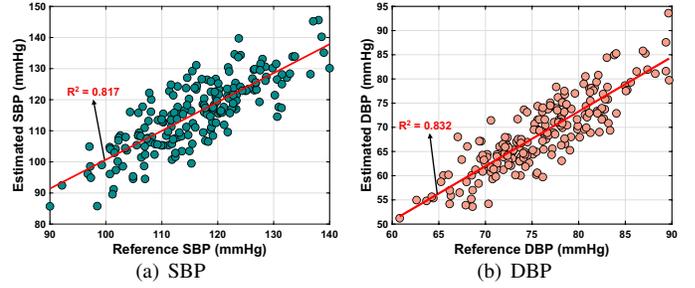


Fig. 14: Correlation diagram of SBP and DBP.

TABLE I: Comparison of HearBP with BHS standard.

Criteria		Cumulative Error Percentage		
		$\leq 5mmHg$	$\leq 10mmHg$	$\leq 15mmHg$
BHS	Grade A	60%	85%	95%
HearBP	SBP	69.29%	90.65%	98.37%
	DBP	73.56%	94.63%	99.29%

deviations of mean errors of DBP and SBP are $3.56mmHg$ and $3.13mmHg$, respectively. The results confirm that HearBP is accurate for users' BP monitoring and meets the AAMI standards ($ME \leq 5mmHg$ and $STD \leq 8mmHg$). In addition, more than 95% of the points are within the LOA of SBP and DBP, which proves highly acceptable measurement results. Meanwhile, Fig. 14 shows the correlation diagram between the estimated results and the reference values, and we can observe that points are clustered around the regression line. R^2 for BP and DBP are 0.817 and 0.832, respectively, which indicates a close correlation. TABLE I provides the BHS standards for BP measurement, and it can be found that both SBP and DBP meet the grade A standards. The above results all indicate that HearBP has high measurement accuracy for SBP and DBP.

C. Performance of Continuous Measurement

We ask 8 participants to record their binaural bone-conduction PCGs for 24 hours. The ground truth is measured every hour using a cuff BP measuring device. Fig. 15 shows the mean estimation errors for DBP and SBP for all 8 participants. The estimation error of DBP and SBP fluctuates below $1.65mmHg$ during awake time, which confirms the validity of HearBP. However, the estimation errors of DBP and SBP fluctuate significantly during sleeping time. The reasons are mainly reflected in two aspects. One reason is that most people tend to turn over while sleeping, which leads to loose wearing of the earphones or even slipping out of the ear canal, resulting in poor quality of collected PCGs. On the other hand, frequent interruptions during sleeping and insufficient sleep can affect participants' mood and heart rate, which adversely affect BP measurements.

D. Key Algorithm Performance

1) *Effectiveness of Noise Removing*: In order to investigate the effectiveness of the proposed noise removing method, we collect data for testing under the four conditions (e.g., music playback, walking and speaking, walking, and stationary state).

TABLE II: Results using different features.

Feature Type	TIFS	TIFS+TD	TIFS+FD	All	PCA
SBP ME/STD ($mmHg$)	8.23/13.66	4.96/7.02	5.59/8.86	1.47/3.03	1.61/3.56
DBP ME/STD ($mmHg$)	7.75/12.17	5.41/8.54	5.27/6.51	0.73/2.79	0.97/3.13

* TD and FD stand for time domain and frequency domain, respectively.

TABLE III: Comparison of DNR with other ML techniques.

Algorithm	GBR	DTR	RFR	SVR	DNR
SBP ME/STD ($mmHg$)	7.24/11.37	5.63/7.21	4.52/6.58	6.49/6.63	1.61/3.56
DBP ME/STD ($mmHg$)	8.71/13.38	4.91/8.89	5.13/6.64	3.26/7.59	0.97/3.13

Fig. 16 shows the estimation errors of BP under these four conditions without and with noise removing. If HeartBP is implemented without noise removing, the estimation errors of DBP and SBP can reach over $9.83mmHg$, which confirms that noise greatly reduces the accuracy of HearBP and makes it unable to meet the requirements of BP measurement in AAMI ($\leq 5mmHg$). After applying noise removing, the estimation errors drop below $2.57mmHg$. The good results not only verify the effectiveness of the noise removing method, but also show that HeartBP is a practical method that is expected to be applied to real daily usage.

2) *Effectiveness of Feature Extraction and Dimension Reduction*: We investigate the effectiveness of different types of features and PCA feature dimension reduction. In the case of five different features, we calculate the BP measurement results, as shown in TABLE II. It can be seen that relying solely on TIFS cannot provide BP measurement accuracy that meets the AAMI standard. Time domain features or frequency domain features can play a great role in improving the accuracy of BP measurement, but only one of them can only assist TIFS to achieve BP measurement accuracy close to the AAMI standard. Applying all features to BP measurement can obviously achieve satisfactory accuracy, and both SBP and DBP can well meet the accuracy requirements of AAMI. In addition, after PCA feature dimension reduction, the accuracy of BP measurement is almost no different from that of using all features. These results confirm the effectiveness of feature extraction and PCA feature dimension reduction, enabling HearBP to effectively meet the practical needs of BP

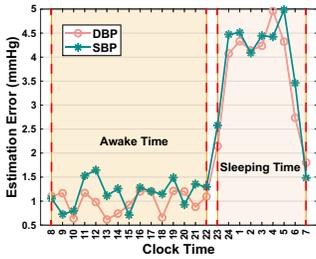


Fig. 15: 24-hour performance.

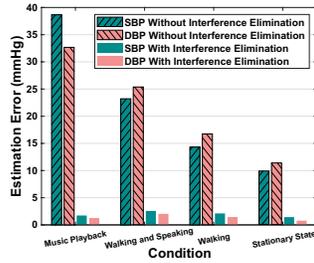


Fig. 16: Performance with & without noise removal.

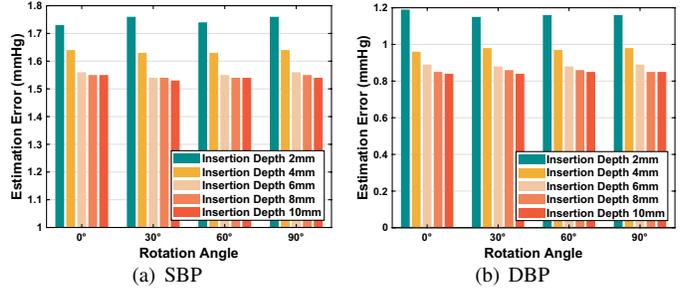


Fig. 17: Impact of earphone wearing status.

TABLE IV: Comparisons of Hear-BP with typical works on BP measurement via wearable devices and acoustic sensing.

Comparison Works	SeismoWatch [14]	eBP [10]	Crisp-BP [17]	Diehl et al. [11]	Ganti et al. [15]	Stereo-BP [18]	Peng et al. [33]	WIB [32]	HearBP
Feature Type	PTT	CFA	RWTT	None	PTT	PTD	None	PTT	TIFS+TFD
Participant	13	35	35	27	21	20	1	30	41
SBP ME/STD (mmHg)	N/A	1.80/7.20	1.67/7.31	1.03/7.19	4.75/2.29	3.97/3.09	0.76	N/A	1.61/3.56
DBP ME/STD (mmHg)	1.30/6.90	3.10/7.90	0.86/6.55	4.41/7.76	2.72/0.75	3.83/2.95	0.85	5.8	0.97/3.13

* CFA, and TFD stand for change of pulse amplitude, time-frequency domain, respectively.

measurement and remove redundant features.

3) *Effectiveness of DNR*: We compare DNR with 4 traditional machine learning (ML) methods, including gradient boosting regression (GBR), decision tree regression (DTR), random forest regression (RFR), and support vector regression (SVR). The results are shown in TABLE III. It can be seen that these regression methods cannot make DBP and SBP meet the standards required by AAMI, so these methods are not suitable for constructing the mapping relationship between the extracted features and BP. On the contrary, the use of DNR can ensure that the measurement accuracy of BP fully meets the standards required by AAMI, which reflects that DNR can effectively establish the mapping relationship between the extracted features and BP. These results are also important reasons why we chose DNR for BP measurement.

E. System Robustness: Impact of Earphone Wearing Status

Due to different wearing habits, the earphone insertion depth and rotation angle of the same participant may vary, resulting in the sound recorded by the in-ear microphone may vary. We require participants to collect data for experiments under different insertion depths (e.g., 2mm, 4mm, 6mm, 8mm, and 10mm) and rotation angles (e.g., 0°, 30°, 60°, and 90°). Fig. 17 shows the estimation errors of SBP and DBP quantified by the above two factors. Overall, HearBP obtains stable results in different wearing status. In addition, we notice that deeper insertion depths lead to lower estimation errors, which may be due to the deeper insertion of the earphones providing better ear canal sealing and stronger position fixation.

F. Comparison with Typical Related Works

TABLE IV shows the comparison results of HearBP with other typical works on BP measurement via wearable devices

and acoustic sensing, including the type of features used, the number of participants, and the estimation accuracy of BP (i.e., mean error/standard deviation for SBP and DBP). But these methods are all subject to certain conditions. SeismoWatch [14], [15], and WIB [32] require the use of additional sensors and user cooperation. The PPG sensors adopted by Crisp-BP [17] and Stereo-BP [18] are usually affected by lighting conditions, skin color, and location information. eBP [10] and [11] require a balloon to compress the artery, which reduces user comfort. [33] uses special materials between the sensor and the human skin when measuring BP. Compared with these works, HearBP can continuously measure BP in a more comfortable and convenient way without user effort, without being limited by various conditions mentioned above.

VI. CONCLUSION

In this paper, we design and implement HearBP, which utilizes in-ear microphones to collect bone-conducted heart sounds in the binaural canal for BP monitoring. We propose effective methods to filter out interference from earphone audio, daily activities, and other body sounds. We explore BP-related time domain and frequency domain features and reduce redundant information using Shannon energy, energy-entropy ratio, and principal component analysis. Based on these features, a network model based on dendritic neural regression is proposed to realize BP monitoring. Extensive experiments involving 41 participants demonstrate that HearBP can provide accurate BP monitoring, with estimation errors of $0.97 \pm 3.13 \text{ mmHg}$ and $1.61 \pm 3.56 \text{ mmHg}$ for DBP and SBP, respectively. These errors are within the acceptable range specified by the AAMI protocol.

REFERENCES

- [1] A. K. Cheung, T. I. Chang, W. C. Cushman, S. L. Furth, F. F. Hou, J. H. Ix, G. A. Knoll, P. Muntner, R. Pecoits-Filho, M. J. Sarnak *et al.*, “Executive summary of the kdigo 2021 clinical practice guideline for the management of blood pressure in chronic kidney disease,” *Kidney International*, vol. 99, no. 3, pp. 559–569, 2021.
- [2] MedicineNet, “Low Blood Pressure (Hypotension),” https://www.medicinenet.com/low_blood_pressure/article.htm, 2022.
- [3] Centers for Disease Control and Prevention, “Coronary Heart Disease, Myocardial Infarction, and Stroke — A Public Health Issue,” <https://www.cdc.gov/aging/publications/coronary-heart-disease-brief.html>, 2019.
- [4] S. Romagnoli, Z. Ricci, D. Quattrone, L. Tofani, O. Tujjar, G. Villa, S. M. Romano, and A. R. De Gaudio, “Accuracy of invasive arterial pressure monitoring in cardiovascular patients: an observational study,” *Critical Care*, vol. 18, no. 6, pp. 1–11, 2014.
- [5] S.-A. Cho, Y.-E. Jang, S.-H. Ji, E.-H. Kim, J.-H. Lee, H.-S. Kim, and J.-T. Kim, “Ultrasound-guided arterial catheterization,” *Anesthesia and Pain Medicine*, vol. 16, no. 2, pp. 119–132, 2021.
- [6] G. Van Montfrans, G. Van Der Hoeven, J. Karemaker, W. Wieling, and A. Dunning, “Accuracy of auscultatory blood pressure measurement with a long cuff,” *Br Med J (Clin Res Ed)*, vol. 295, no. 6594, pp. 354–355, 1987.
- [7] M. Ramsey, “Noninvasive automatic determination of mean arterial pressure,” *Medical and Biological Engineering and Computing*, vol. 17, pp. 11–18, 1979.
- [8] HUAWEI WATCH D, <https://consumer.huawei.com/en/wearables/watch-d/>, 2023.
- [9] HeartGuide, <https://omronhealthcare.com/products/heartguide-wearable-blood-pressure-monitor-bp8000m/>, 2023.
- [10] N. Bui, N. Pham, J. J. Barnitz, Z. Zou, P. Nguyen, H. Truong, T. Kim, N. Farrow, A. Nguyen, J. Xiao, R. Deterding, T. Dinh, T. Vu, “eBP: A wearable system for frequent and comfortable blood pressure monitoring from user’s ear,” in *The 25th annual international conference on mobile computing and networking*, 2019, pp. 1–17.
- [11] M. Diehl, J. Zeilfelder, C. Zimmermann, and W. Stork, “Evaluation of a new method for continuous blood pressure measurement in the ear,” in *2022 IEEE International Symposium on Medical Measurements and Applications*, 2022, pp. 1–6.
- [12] P. A. Obrist, K. C. Light, J. A. McCubbin, J. S. Hutcheson, and J. L. Hoffer, “Pulse transit time: Relationship to blood pressure and myocardial performance,” *Psychophysiology*, vol. 16, no. 3, pp. 292–301, 1979.
- [13] C. F. Wippermann, D. Schranz, and R. G. Huth, “Evaluation of the pulse wave arrival time as a marker for blood pressure changes in critically ill infants and children,” *Journal of Clinical Monitoring*, vol. 11, pp. 324–328, 1995.
- [14] A. M. Carek, J. Conant, A. Joshi, H. Kang, and O. T. Inan, “SeismoWatch: wearable cuffless blood pressure monitoring using pulse transit time,” *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, vol. 1, no. 3, pp. 1–16, 2017.
- [15] V. G. Ganti, A. M. Carek, B. N. Nevius, J. A. Heller, M. Etemadi, and O. T. Inan, “Wearable cuff-less blood pressure estimation at home via pulse transit time,” *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 6, pp. 1926–1937, 2020.
- [16] A. Carek and C. Holz, “Naptics: convenient and continuous blood pressure monitoring during sleep,” *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 3, pp. 1–22, 2018.
- [17] Y. Cao, H. Chen, F. Li, and Y. Wang, “Crisp-BP: Continuous wrist ppg-based blood pressure measurement,” *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*, 2021, pp. 378–391.
- [18] A. N. Balaji, A. Ferlini, F. Kawsar, and A. Montanari, “Stereo-BP: Non-invasive blood pressure sensing with earables,” *Proceedings of the 24th International Workshop on Mobile Computing Systems and Applications*, 2023, pp. 96–102.
- [19] A. Ferlini, A. Montanari, C. Min, H. Li, U. Sassi, and F. Kawsar, “In-ear ppg for vital signs,” *IEEE Pervasive Computing*, vol. 21, no. 1, pp. 65–74, 2021.
- [20] S. Nirjon, R. F. Dickerson, Q. Li, P. Asare, J. A. Stankovic, D. Hong, B. Zhang, X. Jiang, G. Shen, and F. Zhao, “Musicalheart: A hearty way of listening to music,” *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, 2012, pp. 43–56.
- [21] C. Cai, R. Zheng, and J. Luo, “Ubiquitous acoustic sensing on commodity iot devices: A survey,” *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 432–454, 2022.
- [22] C. Cai, H. Pu, P. Wang, Z. Chen, and J. Luo, “We hear your pace: Passive acoustic localization of multiple walking persons,” *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 2, pp. 1–24, 2021.
- [23] Statista Research Department, “Volume of headphones worldwide 2018–2028,” <https://www.statista.com/forecasts/1284259/sales-volume-headphones-market-worldwide>, 2023.
- [24] A. S. AlQahtani, A. N. Alshammari, E. M. Khalifah, A. A. Alnabri, H. A. Aldarwish, K. F. Alshammari, H. F. Alshammari, and A. M. Almudayni, “Awareness about the relation of noise induced hearing loss and use of headphones at hail region,” *Annals of Medicine and Surgery*, vol. 73, p. 103113, 2022.
- [25] F. Kawsar, C. Min, A. Mathur, A. Montanari, U. G. Acer, and M. Van den Broeck, “esense: Open earable platform for human sensing,” *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*, 2018, pp. 371–372.
- [26] J. Ji, M. Dong, Q. Lin, and K. C. Tan, “Forecasting wind speed time series via dendritic neural regression,” *IEEE Computational Intelligence Magazine*, vol. 16, no. 3, pp. 50–66, 2021.
- [27] G. S. Stergiou, B. Alpert, S. Mieke, R. Asmar, N. Atkins, S. Eckert, G. Frick, B. Friedman *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,” *Hypertension*, vol. 71, no. 3, pp. 368–374, 2018.
- [28] E. O’Brien, J. Petrie, W. Littler, M. De Swiet, P. L. Padfield, D. Altman, M. Bland, A. Coats, N. Atkins, “The british hypertension society protocol for the evaluation of blood pressure measuring devices,” *J hypertension*, vol. 11, no. Suppl 2, pp. S43–S62, 1993.
- [29] K. Asayama, T. Ohkubo, S. Hoshida, K. Kario, Y. Ohya, H. Rakugi, and S. Umemura, “From mercury sphygmomanometer to electric device on blood pressure measurement: correspondence of minamata convention on mercury,” *Hypertension Research*, vol. 39, no. 4, pp. 179–182, 2016.
- [30] J. Gravenstein, D. A. Paulus, J. Feldman, and G. McLaughlin, “Tissue hypoxia distal to a penaz finger blood pressure cuff,” *Journal of Clinical Monitoring*, vol. 1, pp. 120–125, 1985.
- [31] A. D. Junior, S. Murali, F. Rincon, and D. Atienza, “Estimation of blood pressure and pulse transit time using your smartphone,” in *2015 Euromicro Conference on Digital System Design*. IEEE, 2015, pp. 173–180.
- [32] J. Zhang, W. Yuan, Y. Chen, M. Li, J. Wang, and Q. Zhang, “WIB: Real-time, non-intrusive blood pressure detection using smartphones,” *ACM Transactions on Sensor Networks*, vol. 19, no. 4, pp. 1–27, 2023.
- [33] C. Peng, M. Chen, H. K. Sim, Y. Zhu, and X. Jiang, “Noninvasive and nonocclusive blood pressure monitoring via a flexible piezo-composite ultrasonic sensor,” *IEEE Sensors Journal*, vol. 21, no. 3, pp. 2642–2650, 2020.
- [34] K. Carillo, O. Doutres, and F. Sgard, “Theoretical investigation of the low frequency fundamental mechanism of the objective occlusion effect induced by bone-conducted stimulation,” *The Journal of the Acoustical Society of America*, vol. 147, no. 5, pp. 3476–3489, 2020.
- [35] Advanced Television Systems Committee, “Techniques for establishing and maintaining audio loudness for digital television,” <https://www.atsc.org/atsc-documents/type/1-0-standards/>, 2013.
- [36] P. Arnott, G. Pfeiffer, and M. Tavel, “Spectral analysis of heart sounds: relationships between some physical characteristics and frequency spectra of first and second heart sounds in normals and hypertensives,” *Journal of Biomedical Engineering*, vol. 6, no. 2, pp. 121–128, 1984.
- [37] N. Perraudin, P. Balazs, and P. L. Søndergaard, “A fast griffin-lim algorithm,” in *2013 IEEE workshop on applications of signal processing to audio and acoustics*, 2013, pp. 1–4.
- [38] M. Elgendi, P. Bobhate, S. Jain, J. Rutledge, J. Y. Coe, R. Zemp, D. Schuurmans, and I. Adatia, “Time-domain analysis of heart sound intensity in children with and without pulmonary artery hypertension: a pilot study using a digital stethoscope,” *Pulmonary Circulation*, vol. 4, no. 4, pp. 685–695, 2014.
- [39] W. Yang, J. Xu, J. Xiang, Z. Yan, H. Zhou, B. Wen, H. Kong, R. Zhu, and W. Li, “Diagnosis of cardiac abnormalities based on phonocardiogram using a novel fuzzy matching feature extraction method,” *BMC Medical Informatics and Decision Making*, vol. 22, no. 1, pp. 1–13, 2022.