IEEE Journal of Translational Engineering in Health and Medicine

Received 13 January 2024; revised 6 July 2024; accepted 22 July 2024.
Date of publication 29 July 2024; date of current version 8 August 2024.

Digital Object Identifier 10.1109/JTEHM.2024.3433448

A Development of a Sound Recognition-Based Cardiopulmonary Resuscitation Training System

DONG HYUN CHOI[1], YOON HA JOO[2,3], KI HONG KIM[3], JEONG HO PARK[3], HYUNJIN JOO[4], HYOUN-JOONG KONG[5,6], HYUNJU LEE[7], KYOUNG JUN SONG[8], AND SUNGWAN KIM[1], (Senior Member, IEEE)

[1]Department of Biomedical Engineering, Seoul National University College of Medicine, Jongno, Seoul 03080, South Korea
[2]Biomedical Research Institute, Seoul National University Hospital, Jongno, Seoul 03080, South Korea
[3]Department of Emergency Medicine, Seoul National University Hospital, Jongno, Seoul 03080, South Korea
[4]Innovative Medical Technology Research Institute, Seoul National University Hospital, Jongno, Seoul 03080, South Korea
[5]Department of Transdisciplinary Medicine, Innovative Medical Technology Research Institute, Seoul National University Hospital, Jongno, Seoul 03080, South Korea
[6]Department of Medicine, Seoul National University College of Medicine, Jongno, Seoul 03080, South Korea
[7]Laboratory of Emergency Medical Services, Biomedical Research Institute, Seoul National University Hospital, Jongno, Seoul 03080, South Korea
[8]Department of Emergency Medicine, Seoul Metropolitan Boramae Medical Center, Dongjak, Seoul 07061, South Korea

(Dong Hyun Choi and Yoon Ha Joo contributed equally to this work.) CORRESPONDING AUTHORS: K. H. KIM (emphysiciankkh@gmail.com) AND S. KIM (sungwan@snu.ac.kr)

This work was supported by the National Research Foundation (NRF) of South Korea Grant funded by the Korea Government Ministry of Science and ICT (MSIT) under Grant NRF-2021R1C1C1006638.

ABSTRACT The objective of this study was to develop a sound recognition-based cardiopulmonary resuscitation (CPR) training system that is accessible, cost-effective, easy-to-maintain and provides accurate CPR feedback. Beep-CPR, a novel device with accordion squeakers that emit high-pitched sounds during compression, was developed. The sounds emitted by Beep-CPR were recorded using a smartphone, segmented into 2-second audio fragments, and then transformed into spectrograms. A total of 6,065 spectrograms were generated from approximately 40 minutes of audio data, which were then randomly split into training, validation, and test datasets. Each spectrogram was matched with the depth, rate, and release velocity of the compression measured at the same time interval by the ZOLL X Series monitor/defibrillator. Deep learning models utilizing spectrograms as input were trained using transfer learning based on EfficientNet to predict the depth (Depth model), rate (Rate model), and release velocity (Recoil model) of compressions. Results: The mean absolute error (MAE) for the Depth model was 0.30 cm (95% confidence interval [CI]: 0.27–0.33). The MAE of the Rate model was 3.6/min (95% CI: 3.2–3.9). For the Recoil model, the MAE was 2.3 cm/s (95% CI: 2.1–2.5). External validation of the models demonstrated acceptable performance across multiple conditions, including the utilization of a newly-manufactured device, a fatigued device, and evaluation in an environment with altered spatial dimensions. We have developed a novel sound recognition-based CPR training system, that accurately measures compression quality during training. Significance: Beep-CPR is a cost-effective and easy-to-maintain solution that can improve the efficacy of CPR training by facilitating decentralized at-home training with performance feedback.

INDEX TERMS Cardiopulmonary arrest, sound recognition, deep learning, feedback communications.

I. INTRODUCTION

OUT-OF-HOSPITAL cardiac arrest (OHCA) poses a major public health burden, having a 1-year survival rate of only 7.7% [1]. Providing high-quality bystander cardiopulmonary resuscitation (CPR) is crucial for improving outcomes in patients with OHCA [2], [3]. Current guidelines recommend that CPR be performed as quickly as possible with chest compressions at a depth of 5–6 cm and a rate of 100–120/min, ensuring full recoil between compressions [4]. However, despite its critical importance for OHCA outcomes, the bystander CPR rate remains low [5].

Public CPR training is an effective approach to encourage laypersons to perform bystander CPR and to ensure its proper execution [3]. CPR training plays a crucial role in empowering individuals with the necessary skills and knowledge to provide CPR effectively [6]. Historically, CPR training has primarily involved group sessions conducted at certified training centers using expensive manikins [7], [8]. However, this approach has several limitations, including restricted accessibility due to participants' time constraints, the finite capacity of group sessions, and the inability to conduct group training during pandemic situations [9]. For instance, there were reports of group CPR training sessions being disrupted, postponed, or shortened during the COVID-19 pandemic [10], [11].

To address the challenges associated with group CPR training, researchers have proposed several low-cost CPR training devices that can be used at home [12], [13]. While these devices offer basic training, they lack quality measurement capabilities and cannot provide feedback on performance. One example is the Push Heart, a heart-shaped sponge-like device that costs around $20 and is easy to carry. However, this device does not provide feedback on CPR quality, which resulted in a lower proportion of adequate compression depth after training compared to training with feedback [13]. Some CPR training devices emit an audible "click" sound when compressed to a depth of more than 5 cm or when sufficient downward pressure is applied [14], [15]. However, they cannot differentiate chest compressions exceeding 6 cm in depth, which is associated with an increased risk of complications and decreased survival [16], [17], nor can they measure other aspects of compression quality. Additionally, they do not record a trainee's performance throughout a training session, making it impossible to evaluate training efficacy. Nonetheless, these devices have shown the potential of using sound to assess chest compression quality.

Recent advancements in deep learning have facilitated the analysis and classification of various sounds [18], [19]. Hence, the development of a CPR training device capable of generating distinct sound signals according to different compression patterns could enable the measurement of chest compression quality. Leveraging the widespread adoption of smartphones, the built-in microphone and processing capabilities of these devices can be utilized to capture the audio signals produced by the CPR training device [20]. These signals can then be processed using a deep learning model, and the results can be displayed on the smartphone's screen as visual feedback. Therefore, the objective of this study was to develop and validate a cost-effective and easy-to-maintain sound recognition-based CPR training system that can measure compression quality.

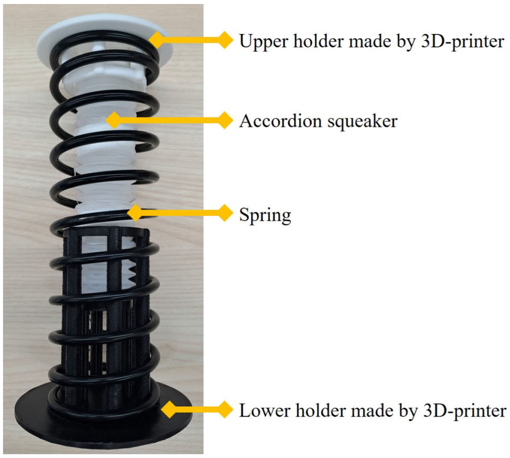

FIGURE 1. The Beep-CPR hardware.

II. METHODS

A. DEVELOPMENT of BEEP-CPR HARDWARE

We developed a novel CPR training device called 'Beep-CPR'. Beep-CPR is a compressible device that emits a 'beep' sound when compressed and released. It was constructed using a self-produced holder, a commercially available spring, and three interconnected accordion squeakers (Figure 1). Beep-CPR has a diameter of 10 cm, a length of 20.6 cm, a maximum compression depth of 8 cm, and a weight of 500 g.

The holder was designed using Inventor 2023 (Autodesk, Inc., CA, USA) and produced with a 3D printer called Cubicon Single Plus (HyVISION, Gyeonggido, Republic of Korea). It consists of an upper and a lower part, which secure the spring and accordion squeakers. The lower part holds the base of the spring to prevent it from falling, while the upper part is fixed on the spring and makes direct contact with the trainee's hands. This design allows force to be transmitted to the spring and accordion squeakers. Additionally, the lower part of the holder minimizes the horizontal bending of the spring and accordion squeakers, permitting only vertical movement.

For the spring, we used a commercially available SWP-A material spring with a free length of 20 cm, an outer diameter of 6.5 cm, a wire diameter of 0.55 cm, and a spring constant of 4.68 N/mm. The spring has an inner diameter of 5.4 cm and is wider than the accordion squeakers, covering them completely. With a maximum contraction of 14.1 cm, the desired maximum compression depth of 8 cm can be achieved.

Three accordion squeakers were connected together and placed inside the spring, secured to both the upper and lower parts of the holder. When the trainee places their hands on the upper part of the holder and presses it, the spring and accordion squeakers are compressed vertically. As a result, the accordion squeakers produce a distinct, loud, high-pitched 'beep' sound when they are compressed or released.

B. DATA ACQUISITION AND POST-PROCESSING

The ZOLL X Series monitor/defibrillator (ZOLL Medical Corporation, MA, USA), equipped with Real CPR Help technology, was used to acquire compression data [21]. The ZOLL CPR-D padz (ZOLL Medical Corporation, MA, USA), connected to the device was placed on the upper holder of the Beep-CPR.

The researchers performed compressions on Beep-CPR with the ZOLL CPR-D padz attached in a closed room (3.85 m × 3.35 m × 2.60 m). A Galaxy Note 20 smartphone (Samsung, Gyeonggido, Republic of Korea) was placed within 1 m of the Beep-CPR device and was used to record the 'beep' sounds produced during Beep-CPR compressions using its built-in microphone. To obtain compression data of diverse depths and rates, we followed a preplanned data collection schedule with different target depths (3–5, 5–6, and 6–8 cm) and rates (70–100, 100–120, and 120–150/min), which were determined according to a previous study (Supplementary Table 1) [22]. The compression data measured by the ZOLL X Series monitor/defibrillator was synchronized with the Beep-CPR sounds recorded by the smartphone using time-based synchronization. This synchronization was initiated by three consecutive compressions on the Beep-CPR at the start of each session, which served as a reference point. These initial compressions were followed by a brief pause. Subsequent compressions were then performed according to the data collection schedule. The investigators compressing the Beep-CPR observed the ZOLL X Series monitor/defibrillator, which displayed the depth and rate of the compressions, and attempted to adhere to the target depth and rate. Compression data that did not meet the targets were not discarded but were also utilized. We were not able to apply a similar process to acquire data with diverse compression release velocities because release velocity is not displayed on the monitor. After the experiment, data for each compression's depth, rate, release velocity, and timestamp were retrieved from the ZOLL X Series monitor/defibrillator. After the experiment, data on compression depth, rate, and recoil velocity, along with the timestamp for each compression, can be extracted via USB from the ZOLL X Series monitor/defibrillator. We also recorded the ambient sounds of the room without Beep-CPR compressions, which included talking, traffic, and music. These ambient sounds were used to train the model to differentiate noise from Beep-CPR compressions, preventing the model from making erroneous predictions when there were no compressions.

We extracted the raw audio files recorded during Beep-CPR compressions with a sampling rate of 10,000 Hz and segmented them into 2-second audio fragments using a sliding window of 0.4 seconds. Each 2-second audio fragment was then transformed into a spectrogram using the short-time Fourier transform method. Subsequently, the portion of the spectrogram corresponding to frequencies less than 512 Hz was removed, and the spectrogram was reshaped to (224, 224, 3) for (time, frequency, RGB channel). Since the frequency of the sound emitted by Beep-CPR was high-pitched around 1,000 Hz, this process aimed to remove the effect of low-frequency noise below 500 Hz, including speech [23]. Each spectrogram was matched with the depth, rate, and release velocity measured by the ZOLL X Series monitor/defibrillator for the latest compression that occurred during the time window of the corresponding spectrogram (Figure 2). If there was no compression matched to a spectrogram, then the spectrogram was labeled as 'no compression'.

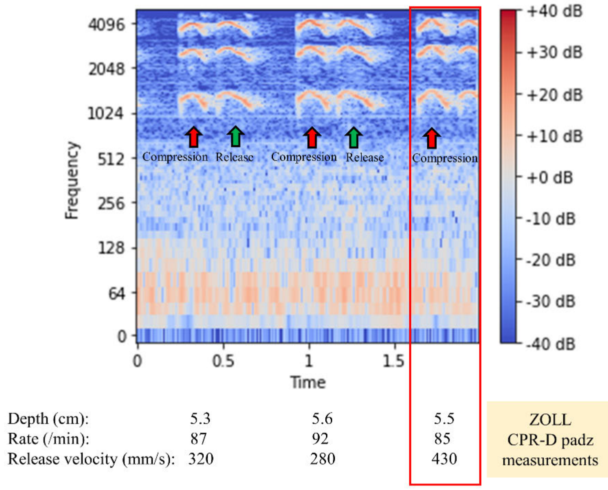

FIGURE 2. Spectrogram demonstrating chest compression quality.

C. DEVELOPMENT of DEEP LEARNING MODELS

A hierarchical structure was employed to classify spectrograms and assess the quality of Beep-CPR compressions. First, we developed a deep learning model (Beep model) to determine the presence of Beep-CPR compressions (compression vs. no compression) in a spectrogram. The Beep model used the architecture of "EfficientNetV2B0" except for the last layer, which was replaced by a global average pooling layer followed by a fully connected layer with a sigmoid activation function [24]. We applied transfer learning, in which the pretrained weights of "EfficientNetV2B0" were used and all its layers were frozen during the training of the Beep model. If a spectrogram was classified as "no compression," no further steps were taken. If a spectrogram was classified as "compression," we moved on to the second step to predict the depth, rate, and release velocity of the compression. Three separate deep learning regression models, one for predicting compression depth (Depth model), another for predicting compression rate (Rate model), and the last for predicting release velocity (Recoil model), were developed. To develop the Depth, Rate, and Recoil models, we performed fine-tuning based on "EfficientNetV2B0". The last layer was removed from the "EfficientNetV2B0" architecture, and the remaining layers were connected to a global average pooling layer, followed by two fully connected hidden layers with a rectified linear unit activation function and a fully connected layer with a linear activation function. Only the weights of the last 10 layers of the models were trained, while the weights of the remaining layers were frozen.

IEEE Journal of Translational Engineering in Health and Medicine

The total dataset of spectrograms was randomly split into training, validation, and test datasets at a ratio of 8:1:1. The deep learning models were trained using the training set with a learning rate of 0.001 for 50 epochs. Hyperparameter tuning was performed using the hyperband tuner in the validation set, with the number of nodes in the fully connected hidden layers of the models as the hyperparameters [25]. The hyperparameter tuning results for each model are shown in Supplementary Table 2.

D. CORRELATION BETWEEN SOUND FREQUENCY AND INSTANTANEOUS VELOCITY in BEEP-CPR COMPRESSIONS

The correlation between the frequency of the sound emitted by Beep-CPR and the instantaneous velocity during compressions was analyzed. The vertical acceleration of the Beep-CPR upper holder was measured using an accelerometer 805M1-0020 (Measurement Specialties Inc., Virginia, USA), and the velocity was calculated by integrating the measured acceleration. The sampling rate was 100 Hz. Sound and acceleration data were collected during 10 compressions at random depths and speeds. Pearson's correlation coefficient was used to evaluate the correlation between the instantaneous compression velocity and the fundamental frequency of the emitted sound.

E. STATISTICAL ANALYSIS

The performance of Beep, Depth, Rate, and Recoil models was evaluated on the test set. The Beep model's accuracy at classifying a spectrogram as 'compression' or 'no compression' was assessed. The Depth, Rate, and Recoil models were evaluated using the mean absolute error (MAE) and Bland-Altman plots. Intraclass correlation coefficient (ICC) using a two-way mixed model was also obtained to assess agreement between the actual and predicted values. The agreement was interpreted as slight, fair, moderate, substantial, and almost perfect if the ICC was in the range of 0.01–0.20, 0.21–0.40, 0.41–0.60, 0.61–0.80, and 0.81–0.99, respectively [26]. We categorized the regression output of the Depth model into 3 depth categories (<5, 5–6, >6 cm), the Rate model into 3 rate categories (<100, 100–120, >120/min), and the Recoil model into 2 categories (<40, >40 cm/s) [27], considering the target range for recommendation. Classification accuracies of the models were evaluated according to categorizations.

Continuous variables were reported in terms of means and standard deviations, while categorical variables were reported as numbers and proportions. Confidence intervals (CIs) were calculated by bootstrapping 1,000 samples. Model development and statistical analysis were performed using Python version 3.8 (Python Software Foundation, Wilmington, DE, USA) and TensorFlow version 2.10.0.

F. EXTERNAL VALIDATION of the BEEP-CPR SYSTEM

To assess the generalizability of the Beep-CPR system, four types of external validation were conducted. First, evaluation of a newly-manufactured device. Second, validation of a fatigued device after subjecting it to 10,000 compressions via a mechanical compressor. Third, evaluation using a spring with different stiffness (5.66 N/mm). Finally, testing in an environment with altered spatial dimensions (4.75 m × 4.55 m × 2.60 m). In each setting, 3 minutes of compression data were collected at various depths and rates. The models' performance was further assessed following linear calibration. This calibration was based on the mean difference between actual and predicted values for the first five compressions.

TABLE 1. Number of spectrograms in each dataset.

	Training	Validation	Test
Total	4,852 (100.0%)	607 (100.0%)	606 (100.0%)
Compression			
Yes	3,997 (82.4%)	496 (81.7%)	497 (82.0%)
No	855 (17.6%)	111 (18.3%)	109 (18.0%)
Depth			
<5 cm	1,867 (46.7%)	214 (43.1%)	238 (47.9%)
5–6 cm	1,201 (30.0%)	164 (33.1%)	148 (29.8%)
>6 cm	929 (23.2%)	118 (23.8%)	111 (22.3%)
Rate			
<100 /min	1,254 (31.4%)	148 (29.8%)	161 (32.4%)
100–120 /min	1,529 (38.3%)	198 (39.9%)	171 (34.4%)
>120 /min	1,214 (30.4%)	150 (30.2%)	165 (33.2%)
Release velocity			
<40 cm/s	1,720 (43.0%)	201 (40.5%)	220 (44.3%)
≥40 cm/s	2,277 (57.0%)	295 (59.5%)	277 (55.7%)

G. ETHICS STATEMENT

Institutional Review Board approval was not needed because this study did not involve any human subjects.

III. RESULTS

A dataset comprising 6,065 spectrograms was generated from approximately 40 minutes of audio data. The dataset was randomly split into training (N = 4,582), validation (N = 607), and test (N = 606) datasets. The distribution of compression depth, rate, and release velocity in each dataset can be found in Table 1. Although the preplanned schedule enabled the acquisition of a dataset with varying compression depth and rate, the dataset exhibited a slight imbalance due to the investigators' inability to consistently achieve the intended target depth and rate during compressions. The fundamental frequency of the sound emitted by Beep-CPR was linearly correlated ($R^2 = 0.86$) with the instantaneous compression velocity (Figure 3), providing evidence for the use of Beep-CPR sound spectrograms in predicting compression depth, rate, and recoil.

The Beep model showed 100% accuracy in discriminating "compression" and "no compression" spectrograms on the test dataset. The performances of the Depth, Rate, and Recoil models were assessed using the 497 "compression" spectrograms in the test dataset. The MAE of the Depth model was 0.30 cm (95% CI: 0.27–0.33). The predictions by the Depth

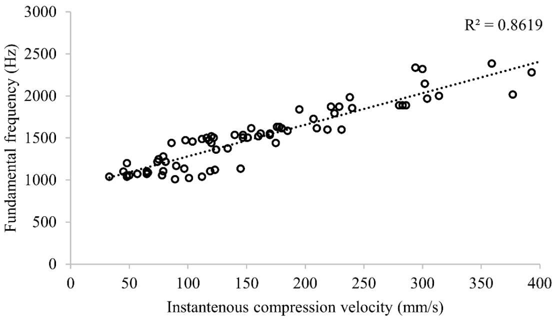

FIGURE 3. Correlation plot between fundamental frequency of sound and instantaneous compression velocity.

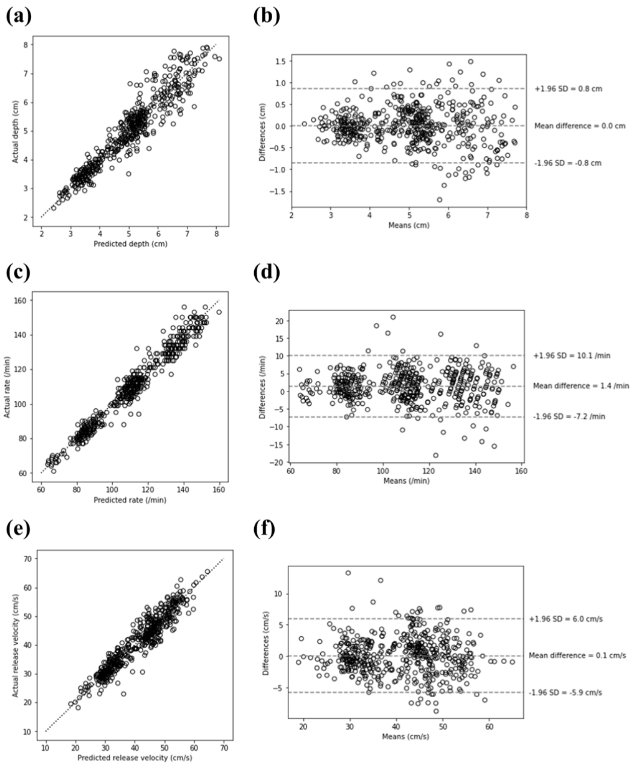

FIGURE 4. Correlation plots and Bland-Altman plots for predicted and actual values of compression depth, rate, and release velocity.

model and the actual compression depths showed an almost perfect (ICC = 0.95) agreement (Figure 4a and b). While the Depth model's classification accuracy was 81.7%, most of the misclassification occurred on the classification boundaries (Table 2). The MAE of the Rate model was 3.6/min (95% CI: 3.2–3.9). The Rate model predictions and the actual compression rates also showed an almost perfect (ICC = 0.98) agreement (Figure 4c, 4d). The classification accuracy of the Rate model was 95.6% (Table 3). For the Recoil model, the MAE was 2.3 cm/s (95% CI: 2.1–2.5) with an almost perfect (ICC = 0.95) agreement between predicted and actual release velocity (Figure 4e, 4f, and Table 4).

The Beep model accurately predicted 1,896 (98.9%) among the 1,917 spectrograms in the overall external

TABLE 2. Classification accuracy of the depth model.

Actual compression depth	Predicted compression depth		
	<5 cm	5–6 cm	>6 cm
<5 cm	201 (84.5%)	37 (15.5%)	0 (0%)
2.0–2.5 cm	1 (100%)	0 (0%)	0 (0%)
2.5–3.0 cm	16 (100%)	0 (0%)	0 (0%)
3.0–3.5 cm	59 (100%)	0 (0%)	0 (0%)
3.5–4.0 cm	57 (98.3%)	1 (1.7%)	0 (0%)
4.0–4.5 cm	37 (86.0%)	6 (14.0%)	0 (0%)
4.5–5.0 cm	31 (50.8%)	30 (49.2%)	0 (0%)
5–6 cm	18 (12.2%)	111 (75.0%)	19 (12.8%)
5.0–5.5 cm	15 (16.5%)	72 (79.1%)	4 (4.4%)
5.5–6.0 cm	3 (5.3%)	39 (68.4%)	15 (26.3%)
>6 cm	0 (0%)	17 (15.3%)	94 (84.7%)
6.0–6.5 cm	0 (0%)	7 (18.9%)	30 (81.1%)
6.5–7.0 cm	0 (0%)	9 (25.0%)	27 (75.0%)
7.0–7.5 cm	0 (0%)	1 (5.0%)	19 (95.0%)
7.5–8.0 cm	0 (0%)	0 (0%)	18 (100%)

TABLE 3. Classification accuracy of the rate model.

Actual compression rate	Predicted compression rate		
	<100 /min	100–120 /min	>120 /min
<100 /min	152 (94.4%)	9 (5.6%)	0 (0%)
60–70 /min	14 (100%)	0 (0%)	0 (0%)
70–80 /min	20 (100%)	0 (0%)	0 (0%)
80–90 /min	97 (99.0%)	1 (1.0%)	0 (0%)
90–100 /min	21 (72.4%)	8 (27.6%)	0 (0%)
100–120 /min	1 (0.6%)	165 (96.5%)	5 (2.9%)
100–110 /min	1 (1.2%)	79 (98.8%)	0 (0%)
110–120 /min	0 (0%)	86 (94.5%)	5 (5.5%)
>120 /min	0 (0%)	5 (3.0%)	160 (97.0%)
120–130 /min	0 (0%)	4 (7.4%)	50 (92.6%)
130–140 /min	0 (0%)	1 (1.6%)	61 (98.4%)
140–150 /min	0 (0%)	0 (0%)	33 (100%)
150–160 /min	0 (0%)	0 (0%)	16 (100%)

validation dataset (Supplementary Table 3). When validated using a newly-manufactured device, a fatigued device, a spring with different stiffness, and in an environment with altered spatial dimensions, the MAE of the Depth model was 0.71 cm (95% CI: 0.66–0.76), 0.81 cm (95% CI: 0.76–0.87), 0.96 cm (95% CI: 0.92–1.01) and 0.79 cm (95% CI: 0.74–0.85) respectively. Substantial agreement (ICC ranging from 0.67 to 0.72) was found between the predicted and actual depths in all settings except when validating with a spring of different stiffness (ICC = 0.14). The MAEs of the Rate model in the same order were 8.7/min (95% CI: 8.0–9.5), 11.8/min (95% CI: 11.2–12.4), 9.2/min (95% CI: 8.7–9.8) and 10.2/min (95% CI: 9.5–11.0). The predicted and actual rates demonstrated substantial to almost perfect agreement (ICC ranging from 0.69 to 0.83). The MAEs of the Recoil model in the same order were 8.9 cm/s (95% CI: 8.6–9.2), 14.3 cm/s (95% CI: 13.6–14.9), 6.4 cm/s (95% CI: 6.0–6.7) and 9.5 cm/s (95% CI: 9.1–9.8). The predicted and actual release velocities demonstrated substantial

IEEE Journal of Translational Engineering in Health and Medicine

TABLE 4. Classification accuracy of the recoil model.

Actual release velocity	Predicted release velocity	
	<40 cm/s	≥40 cm/s
<40 cm/s	205 (93.2%)	15 (6.8%)
10–20 cm/s	2 (100%)	0 (0%)
20–30 cm/s	62 (100%)	0 (0%)
30–40 cm/s	141 (90.4%)	15 (9.6%)
≥40 cm/s	11 (4.0%)	266 (96.0%)
40–50 cm/s	11 (6.0%)	173 (94.0%)
50–60 cm/s	0 (0%)	87 (100%)
60–70 cm/s	0 (0%)	6 (100%)

TABLE 5. Comparison of materials and costs required to measure the quality of CPR: beep-CPR versus traditional CPR training system.

Beep-CPR		Traditional CPR training system	
Materials	Cost (U.S. dollars)	Materials	Cost (U.S. dollars)
Accordion squeakers	1	Laser displacement sensor: VL53L0X V2	20
		Microprocessor: Raspberry Pi 4	35
		Battery	1
		Wires	1
Total cost	1	Total cost	52

CPR, cardiopulmonary resuscitation

(ICC ranging from 0.61 to 0.80) agreement (Supplementary Figures 1-4).

Calibration improved the performance of the models in external validations, except in the setting with a spring of different stiffness. The MAEs for the Depth, Rate, and Recoil models were reduced to 0.51–0.58 cm, 6.4–9.1/min, and 3.3–7.3 cm/s, respectively.

IV. DISCUSSION

In this study, we developed a sound recognition-based CPR training system named Beep-CPR. Beep-CPR assesses the quality of chest compressions by analyzing the sounds emitted from accordion squeakers inside the device. The audio data recorded with a smartphone is first converted into a spectrogram. Deep learning is then applied to this spectrogram to create algorithms that predict compression depth, rate, and release velocity. The predicted values from the Depth, Rate, and Recoil models exhibited almost perfect agreement with the actual values in the internal validation. External validation was conducted to assess the generalizability of the Beep-CPR system. The models demonstrated acceptable performance when externally validated using a newly-manufactured device, a fatigued device, and when evaluated in an environment with altered spatial dimensions. These results indicate that Beep-CPR accurately measures chest compression quality and has the potential to be used for CPR training in diverse settings.

Beep-CPR shows potential for widespread adoption. First. Beep-CPR is cost effective. Traditional CPR training systems use various electronic components to measure the quality of CPR: 1. Sensors for measuring depth, rate, and release velocity, 2. A microprocessor to converts sensor signals into numerical values and transmits them to the smart device, 3. A battery for power supply, 4. Wires connecting all the components. These electronic components make the system more expensive (Table 5). For example, the Resusci Anne QCPR manikin (Laerdal Medical, Stavanger, Norway), one of the most widely used CPR training manikins, can measure compression depth with less than 1 mm error but costs more than $2,000 and cannot measure depths greater than 6 cm [28]. However, Beep-CPR only uses inexpensive accordion squeakers to measure CPR quality (Table 5). Second, it is easy to maintain. This reduction in electronic components not only minimizes potential of failure but also simplifies the overall structure of the device. In the event that a component does fail, the system's design ensures that repairs are both straightforward and efficient. The parts are modular, meaning that individual components can be easily accessed and replaced without the need for specialized tools or extensive technical knowledge. This modularity also allows for quick identification of faulty components, reducing downtime and ensuring that the device can be returned to operational status rapidly. This innovative system holds the potential for decentralized CPR training, eliminating the need for expensive, difficult-to-repair manikins and overcoming limitations related to time and space.

Public CPR training plays a crucial role in elevating bystander CPR rates and enhancing the outcomes of patients with OHCA. Numerous studies have demonstrated that various types of CPR training effectively enhance the knowledge and willingness of laypersons to administer CPR [29], [30], [31]. One study showed that a 5% increase in participation in CPR training courses resulted in a significant 14% increase in the 30-day survival rate of OHCA patients [3]. Another study identified associations between county-level CPR training rates and improvements in good neurological recovery rates [32]. However, despite its paramount importance, barriers to CPR training, including issues of accessibility and financial constraints, can hinder the positive impact [33].

Several previous studies also developed low-cost CPR training devices, but their ability to provide accurate performance feedback was limited. The Push Heart, while inexpensive, lacks the capability to assess compression quality. As a result, CPR training using Push Heart resulted in a lower percentage of adequate compression depth when compared to training using Little Anne (1.5% vs. 5.5%) [13]. Another study employed toilet paper for CPR training and found it to be noninferior to traditional CPR training [12]. Similarly, a study utilizing a foam die and a plastic bag found no difference in CPR performance after 6 months when compared to conventional CPR training [34]. However, despite their cost-effectiveness and easy-to-maintain, these devices cannot measure CPR performance or provide feedback - a significant limitation given the known benefits of feedback in improving CPR skill acquisition and retention [35].

We employed sound recognition based on spectrograms and deep learning to effectively assess CPR quality. To analyze these sounds effectively, a method inspired by recent studies in audio signal processing was utilized. These studies have demonstrated the effectiveness of converting audio signals into spectrogram images to be used as input into convolutional neural network-based models [18], [19], [36]. This spectrogram-based approach offers three advantages. First, spectrograms effectively capture both the frequency distribution and temporal information of the audio signal. Second, the spectrogram can potentially predict multiple CPR quality indicators. Compression depth and release velocity can be estimated from the spectrogram, as these parameters relate to the instantaneous velocity integrated over time. Additionally, the compression rate can be determined by analyzing the time intervals between the high-pitched sounds emitted by the squeakers, which are visible in the spectrogram. Third, while deep learning models pretrained on audio data are relatively scarce, there is an abundance of pretrained models for image data [37]. By converting audio to spectrograms, these widely available image-based models can be leveraged for the audio analysis task.

External validation demonstrated acceptable model performance; however, an increase in prediction errors was observed. This degradation in accuracy is hypothesized to be attributed to subtle variations in the acoustic properties of the squeakers, potentially arising from manufacturing variability and material fatigue. As with many sensors that require calibration before use, the process can be adopted to mitigate these errors. Additionally, errors are expected to decrease as Beep-CPR manufacturing becomes industrialized and automated. While other aspects of external validation were acceptable, the Depth model notably underperformed when predicting compressions using a spring with different stiffness. Spectrogram analysis revealed substantial differences in the acoustic spectrum compared to the data from the original springs used for model development. These spectral differences can be attributed to changes in the compression wave patterns caused by variations in spring stiffness. The altered spring properties likely affected the dynamics of chest compressions, including the speed and force of compressions and decompressions. Consequently, these mechanical changes are assumed to affect the acoustic properties of the squeakers, resulting in sound signatures that diverged significantly from those in the original training dataset.

This study has several limitations that need to be addressed. First, sound was recorded using only a single smartphone model. Second, chest compressions were performed by only two researchers, limiting the variety of compression patterns reflected in the data. Third, the Beep-CPR is restricted to use with springs of varying stiffness. Since patients have various thorax resistances, CPR training using a single type of spring may be insufficient for trainees to learn how to perform high-quality CPR for diverse patients.

In our future research, we plan to develop a more robust Beep-CPR system by collecting diverse data from various training environments, diverse trainees, multiple smartphone models, springs with varying stiffness, and different levels of background noise. This comprehensive approach to data collection is expected to enhance the model's adaptability and performance across a wider spectrum of real-world conditions. A randomized trial is planned to evaluate the effectiveness, learner experience, and engagement of Beep-CPR training compared to traditional CPR training in improving trainees' CPR knowledge and skills.

V. CONCLUSION

We developed a novel sound recognition-based CPR training system capable of accurately measuring chest compression quality during training. The proposed system is cost-effective and easily repairable, and it has the potential to enhance the efficacy of CPR training by facilitating decentralized at-home training and providing feedback on CPR performance. By translating this engineering methodology into real-world use for both community and clinical settings, we look forward to improving outcomes for OHCA patients.

REFERENCES

[1] S. Yan et al., "The global survival rate among adult out-of-hospital cardiac arrest patients who received cardiopulmonary resuscitation: A systematic review and meta-analysis," *Crit. Care*, vol. 24, no. 1, p. 61, Feb. 2020, doi: 10.1186/s13054-020-2773-2.

[2] M. Talikowska, H. Tohira, and J. Finn, "Cardiopulmonary resuscitation quality and patient survival outcome in cardiac arrest: A systematic review and meta-analysis," *Resuscitation*, vol. 96, pp. 66–77, Nov. 2015, doi: 10.1016/j.resuscitation.2015.07.036.

[3] T. W. Jensen et al., "Training in basic life support and bystander-performed cardiopulmonary resuscitation and survival in out-of-hospital cardiac arrests in Denmark, 2005 to 2019," *JAMA Netw. Open*, vol. 6, no. 3, Mar. 2023, Art. no. e233338, doi: 10.1001/jamanetworkopen.2023.3338.

[4] A. R. Panchal et al., "Part 3: Adult basic and advanced life support: 2020 American heart association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care," *Circulation*, vol. 142, no. 2, pp. S366–S468, Oct. 2020, doi: 10.1161/cir.0000000000000916.

[5] R. Chocron et al., "Bystander cardiopulmonary resuscitation quality: Potential for improvements in cardiac arrest resuscitation," *J. Amer. Heart Assoc.*, vol. 10, no. 6, Mar. 2021, Art. no. e017930, doi: 10.1161/jaha.120.017930.

[6] I. Lund-Kordahl, M. Mathiassen, J. Melau, T. M. Olasveengen, K. Sunde, and K. Fredriksen, "Relationship between level of CPR training, self-reported skills, and actual manikin test performance—An observational study," *Int. J. Emergency Med.*, vol. 12, no. 1, p. 2, Jan. 2019, doi: 10.1186/s12245-018-0220-9.

[7] A. Cheng, A. Lockey, F. Bhanji, Y. Lin, E. A. Hunt, and E. Lang, "The use of high-fidelity manikins for advanced life support training—A systematic review and meta-analysis," *Resuscitation*, vol. 93, pp. 142–149, Aug. 2015, doi: 10.1016/j.resuscitation.2015.04.004.

[8] E. McCoy et al., "Randomized controlled trial of simulation vs. Standard training for teaching medical students high-quality cardiopulmonary resuscitation," *Western J. Emergency Med.*, vol. 20, no. 1, pp. 15–22, Dec. 2018, doi: 10.5811/westjem.2018.11.39040.

[9] C. J. Thorne et al., "The nationwide impact of COVID-19 on life support courses. A retrospective evaluation by resuscitation council U.K.," *Resuscitation Plus*, vol. 13, Mar. 2023, Art. no. 100366, doi: 10.1016/j.resplu.2023.100366.

[10] N. Rajaram, H. Krishna, R. Singh, and A. K. Narayan, "Impact of COVID-19 on basic life support training among medical students: An experimental study," *Cureus*, vol. 14, no. 4, Apr. 2022, Art. no. e23775, doi: 10.7759/cureus.23775.

IEEE Journal of Translational Engineering in Health and Medicine

[11] Q. N. N. Tran et al., "Efficacy of the continuous resuscitation training with the gap period due to the COVID-19 pandemic," *Int. Med. Educ.*, vol. 2, no. 3, pp. 188–197, Aug. 2023, doi: 10.3390/ime2030018.

[12] R. Ohle, M. Moskalyk, E. Boissonneault, A. Bilgasem, E. Tissot, and S. McIsaac, "Is a homemade cardiopulmonary resuscitation (CPR) trainer non-inferior to a commercially available CPR mannequin in teaching high-quality CPR? A non-inferiority randomized control trial," *Resuscitation Plus*, vol. 6, Jun. 2021, Art. no. 100134, doi: 10.1016/j.resplu.2021.100134.

[13] S. Tanaka et al., "Comparison of quality of chest compressions during training of laypersons using push heart and little anne manikins using blinded CPRcards," *Int. J. Emergency Med.*, vol. 10, no. 1, p. 20, Dec. 2017, doi: 10.1186/s12245-017-0147-6.

[14] M. A. Baubin, H. Gilly, A. Posch, A. Schinnerl, and G. A. Kroesen, "Compression characteristics of CPR manikins," *Resuscitation*, vol. 30, no. 2, pp. 117–126, Oct. 1995, doi: 10.1016/0300-9572(95)00874-s.

[15] B. Lynch, E. L. Einspruch, G. Nichol, L. B. Becker, T. P. Aufderheide, and A. Idris, "Effectiveness of a 30-min CPR self-instruction program for lay responders: A controlled randomized study," *Resuscitation*, vol. 67, no. 1, pp. 31–43, Oct. 2005, doi: 10.1016/j.resuscitation.2005.04.017.

[16] H. Hellevuo et al., "Deeper chest compression–more complications for cardiac arrest patients?" *Resuscitation*, vol. 84, no. 6, pp. 760–765, Jun. 2013, doi: 10.1016/j.resuscitation.2013.02.015.

[17] I. G. Stiell et al., "What is the optimal chest compression depth during out-of-hospital cardiac arrest resuscitation of adult patients?" *Circulation*, vol. 130, no. 22, pp. 1962–1970, Nov. 2014, doi: 10.1161/circulationaha.114.008671.

[18] D. Ciric, Z. Peric, J. Nikolic, and N. Vucic, "Audio signal mapping into spectrogram-based images for deep learning applications," in *Proc. 20th Int. Symp. INFOTEH-JAHORINA*, Mar. 2021, pp. 1–6.

[19] V. Franzoni, G. Biondi, and A. Milani, "Emotional sounds of crowds: Spectrogram-based analysis using deep learning," *Multimedia Tools Appl.*, vol. 79, nos. 47–48, pp. 36063–36075, Aug. 2020, doi: 10.1007/s11042-020-09428-x.

[20] M. Alkhodari and A. H. Khandoker, "Detection of COVID-19 in smartphone-based breathing recordings: A pre-screening deep learning tool," *PLoS One*, vol. 17, no. 1, Jan. 2022, Art. no. e0262448, doi: 10.1371/journal.pone.0262448.

[21] N. Segal, F. Laurent, L. Maman, P. Plaisance, and P. Augustin, "Accuracy of a feedback device for cardiopulmonary resuscitation on a dental chair," *Emergency Med. J.*, vol. 29, no. 11, pp. 890–893, Nov. 2012, doi: 10.1136/emermed-2011-200003.

[22] S. Duval et al., "Optimal combination of compression rate and depth during cardiopulmonary resuscitation for functionally favorable survival," *JAMA Cardiology*, vol. 4, no. 9, p. 900, Sep. 2019, doi: 10.1001/jamacardio.2019.2717.

[23] A. Palumbo et al., "A novel portable device for laryngeal pathologies analysis and classification," in *Advances in Biomedical Sensing, Measurements, Instrumentation and Systems*. Cham, Switzerland: Springer, 2010, pp. 335–352.

[24] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 6105–6114.

[25] L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar, "Hyperband: A novel bandit-based approach to hyperparameter optimization," 2016, *arXiv:1603.06560*.

[26] V. C. T. Sung, "Agreement in assessing optic discs with a digital stereoscopic optic disc camera (discam) and Heidelberg retina tomograph," *Brit. J. Ophthalmology*, vol. 86, no. 2, pp. 196–202, Feb. 2002, doi: 10.1136/bjo.86.2.196.

[27] A. Kovacs et al., "Chest compression release velocity: Association with survival and favorable neurologic outcome after out-of-hospital cardiac arrest," *Resuscitation*, vol. 92, pp. 107–114, Jul. 2015, doi: 10.1016/j.resuscitation.2015.04.026.

[28] S. G. Beesems and R. W. Koster, "Accurate feedback of chest compression depth on a manikin on a soft surface with correction for total body displacement," *Resuscitation*, vol. 85, no. 11, pp. 1439–1443, Nov. 2014, doi: 10.1016/j.resuscitation.2014.08.005.

[29] M. Riggs, R. Franklin, and L. Saylany, "Associations between cardiopulmonary resuscitation (CPR) knowledge, self-efficacy, training history and willingness to perform CPR and CPR psychomotor skills: A systematic review," *Resuscitation*, vol. 138, pp. 259–272, May 2019, doi: 10.1016/j.resuscitation.2019.03.019.

[30] K. Sipsma, B. A. Stubbs, and M. Plorde, "Training rates and willingness to perform CPR in king county, washington: A community survey," *Resuscitation*, vol. 82, no. 5, pp. 564–567, May 2011, doi: 10.1016/j.resuscitation.2010.12.007.

[31] J. Urban, H. Thode, E. Stapleton, and A. J. Singer, "Current knowledge of and willingness to perform hands-only CPR in laypersons," *Resuscitation*, vol. 84, no. 11, pp. 1574–1578, Nov. 2013, doi: 10.1016/j.resuscitation.2013.04.014.

[32] Y. S. Ro et al., "Association between county-level cardiopulmonary resuscitation training and changes in survival outcomes after out-of-hospital cardiac arrest over 5 years: A multilevel analysis," *Resuscitation*, vol. 139, pp. 291–298, Jun. 2019, doi: 10.1016/j.resuscitation.2019.01.012.

[33] C. Sasson et al., "Barriers and facilitators to learning and performing cardiopulmonary resuscitation in neighborhoods with low bystander cardiopulmonary resuscitation prevalence and high rates of cardiac arrest in columbus, OH," *Circulation, Cardiovascular Quality Outcomes*, vol. 6, no. 5, pp. 550–558, Sep. 2013, doi: 10.1161/circoutcomes.111.000097.

[34] V. Van Raemdonck, K. G. Monsieurs, D. Aerenhouts, and K. De Martelaer, "Teaching basic life support: A prospective randomized study on low-cost training strategies in secondary schools," *Eur. J. Emergency Med.*, vol. 21, no. 4, pp. 284–290, Aug. 2014, doi: 10.1097/mej.0000000000000071.

[35] J. Yeung, R. Meeks, D. Edelson, F. Gao, J. Soar, and G. D. Perkins, "The use of CPR feedback/prompt devices during training and CPR performance: A systematic review," *Resuscitation*, vol. 80, no. 7, pp. 743–751, Jul. 2009, doi: 10.1016/j.resuscitation.2009.04.012.

[36] P. M. Vincent, K. Srinivasan, and C.-Y. Chang, "Deep learning assisted neonatal cry classification via support vector machine models," *Frontiers Public Health*, vol. 9, Jun. 2021, Art. no. 670352, doi: 10.3389/fpubh.2021.670352.

[37] A. Stancic, V. Vyroubal, and V. Slijepcevic, "Classification efficiency of pre-trained deep CNN models on camera trap images," *J. Imag.*, vol. 8, no. 2, p. 20, Jan. 2022, doi: 10.3390/jimaging8020020.