ENHANCEMENT OF CARDIAC AND RESPIRATORY SOUNDS FOR CELLPHONE REPRODUCTION BY MEANS OF DIGITAL SOUND PROCESSING METHODS

Maria Belen Echenique

School of Engineering
Pontificia Universidad Católica de Chile
Santiago, Chile
mbecheniquep@gmail.com

Eduardo J. Godoy

School of Medicine
Pontificia Universidad Católica de Chile
Santiago, Chile
ejgodoy@uc.cl

Rodrigo F. Cádiz

School of Engineering and Music Institute
Pontificia Universidad Católica de Chile
Millennium Institute for
Intelligent Healthcare Engineering - iHEALTH
Millennium Institute for Care Research - MICARE
Santiago, Chile
rcadiz@uc.cl

Marcelo E. Andia

School of Medicine
Pontificia Universidad Católica de Chile
Millennium Institute for
Intelligent Healthcare Engineering - iHEALTH
Santiago, Chile
meandia@uc.cl

ABSTRACT

Telemedicine has become increasingly popular due to its convenience and accessibility. However, one of its drawbacks is the difficulty of performing lung and heart auscultation remotely. As a solution, we propose smartphone-based tele-auscultation for capturing lung and heart sounds. However, one of the obstacles to this approach is that low-frequency sounds, such as heart sounds, cannot be accurately reproduced by a smartphone speaker, due to their frequency response. To address this issue, we suggest using pitch-shifted versions of these sounds, specifically designed for hearing through smartphones. We conducted an initial evaluation of these processed sounds by obtaining 10 heart sounds and 20 lung sounds from open-source databases, which were pitch-shifted using algorithms based on Paul's Stretch and SoundStretch libraries, respectively. These processed audios were validated by 40 final-year medical students using a web survey and conventional headphones and were compared against the original versions. The results showed that 72% and 80% of responses indicated that clinical information was preserved in the samples of respiratory and heart sounds, respectively. These findings suggest that pitch-shifted sounds could be potentially used in tele-auscultation devices such as smartphones. However, further research is needed regarding the recording and playback capabilities of smartphones.

1. INTRODUCTION

1.1. Telemedicine

Telemedicine is the use of telecommunication technology such as videoconferencing to provide health care to patients who are away from the point of service [1, 2]. It is already being applied to multiple clinical specialties including neurology, psychiatry, cardiology, surgery, ophthalmology, genetics, oncology, gastroenterology, geriatrics, and dermatology [3, 4, 5, 6, 7].

The use of telemedicine has various advantages for patients, as it is more accessible, it saves them time, and reduces travel expenses and medical costs. Indeed, some organizations such as the American Telemedicine Association are dedicated entirely to promoting the use of telemedicine as a way to transform health care [8]. There are several types of interactions in telemedicine. One of the most basic types involves communication through videoconference between clinician and patient. The obvious limitation is the lack of physical examination, which is essential to medicine. When physical examination is required, the use of platforms designed to integrate special medical devices (e.g. a digital stethoscope, oxymeter, etc.) and a facilitator that ensures they're used correctly are needed [9]. This, in turn, means a person seeking tele-healthcare needs to travel to a specialized center, increasing costs. With the COVID-19 pandemic, telemedicine became a critically essential service for infected and uninfected alike, reducing the spread of the virus and the exposure of healthcare professionals [10]. The pandemic has changed global health care and has driven a major shift towards the use of non-invasive technologies for monitoring patients remotely, including the use of mobile health technologies [11].

1.2. Smartphones

Smartphones are one of the technologies that have most rapidly spread around modern society, with studies estimating over 7 billion devices around the world [12]. The amount of adults who own a smartphone is also increasing exponentially. In 2008, less than 25% of the adults in the United Kingdom owned a smartphone, while in 2021, 88% of them did [13]. In addition, a study made in 2016 revealed that, in the majority of the European Union, for people aged 16 - 74 years, smartphones were the most used devices to access the internet [14].

In the last few years, smartphones have developed the ability to transfer high-quality text, sound, images and video that, until recently, only complex and expensive technical platforms could provide. Their use of mobile internet and Wi-Fi networks makes them able to transmit and receive data from almost everywhere [15]. For this reason, numerous algorithms have been developed in order to detect various medical conditions.

For example, microphones in smartphones have been used to detect coughs and sneezes [16], or to perform spirometry based on respiratory sounds [17]. These applications require the user to hold the smartphone at a certain distance and breathe, and then estimate the flow or find respiratory patterns using sound processing and data classification techniques. Other sensors, such as cameras, have been proposed for different noninvasive and incrementally cost-effective monitoring, treatment and diagnostic methods. Smartphone-based imaging has been proposed for monitoring vital signs, blood pressure, diagnosis of skin cancer and other skin lesions, and treatment of burns, image-guiding-surgery and venipuncture. [18].

There are multiple advantages of using smartphones as medical devices and today, there are more than 350 thousand medical applications developed for them and available to consumers [19]. The many sensors in recent models allow for various physiological parameters to be determined. This can be done through the use of gyroscopes, accelerometers, light sensors, dual cameras, microphones, pressure and temperature sensors, among others [18].

1.3. Auscultation

Auscultation has been a part of physical examination since the time of Hippocrates. Nowadays, a stethoscope is used to perform it. Heart sounds are generated by periodic closures of the heart valves, while lung breath sounds are created by the movement of air through the airways. The range of clinically important heart sounds is between 20 - 600 Hz, while those of the lung lie in the range of 100 - 1000 Hz. [20].

Auscultation is an operator-dependent procedure. Despite this, it can provide vital information very quickly, and without added cost. For instance, in the case of lung auscultation, there is evidence that its performance is very high for some diagnoses, such as wheezing, pneumonia and other bacterial diseases [21].

Given the above, one of the most promising devices for telemedicine would be mobile phones, which could help to replace the presence of clinical staff and solve limitations such as the current practice of performing pulmonary and cardiac auscultation during a physical examination in a teleconsultation. Since smartphones contain embedded microphones, they could then be used to record the sounds produced by the heart and lungs [22, 23], but there are some limitations such as the difference in microphones between different mobile phone models [24]. Another limitation is that the quality of mobile phone speakers does not allow for the correct reproduction of low frequency waves, which are a main component of heart sounds. However, there are several ways to improve these sounds in order to allow them to be played on a smartphone.

The aim of this study is to determine whether the use of two separate algorithms to improve lung and heart sounds, respectively, have an impact on the clinical information contained in them. To achieve this, two algorithms were proposed for enhancing lung and heart sounds separately and then they were tested with medical interns who were asked to answer if the original and processed audio samples contained the same information. To validate these algorithms we conducted an online survey where medical students compared two sounds, one unprocessed and the other processed with the corresponding pitch-shifting procedure, presented in random order, as discussed below.

2. METHODOLOGY

For the study, respiratory and cardiac audio samples from an online database were processed using two different techniques. The database, pre-processing, methodology used to process the recordings for each case, and validation methods are detailed in this section. A summary of the methodology applied is shown in figure 1.

2.1. Databases

Two open-source databases were used in the study: the first corresponds to cardiac sounds [25] and the second to respiratory sounds [26].

For cardiac sounds, a database of sounds collected from infants to young adults was used [25, 27], whose sounds

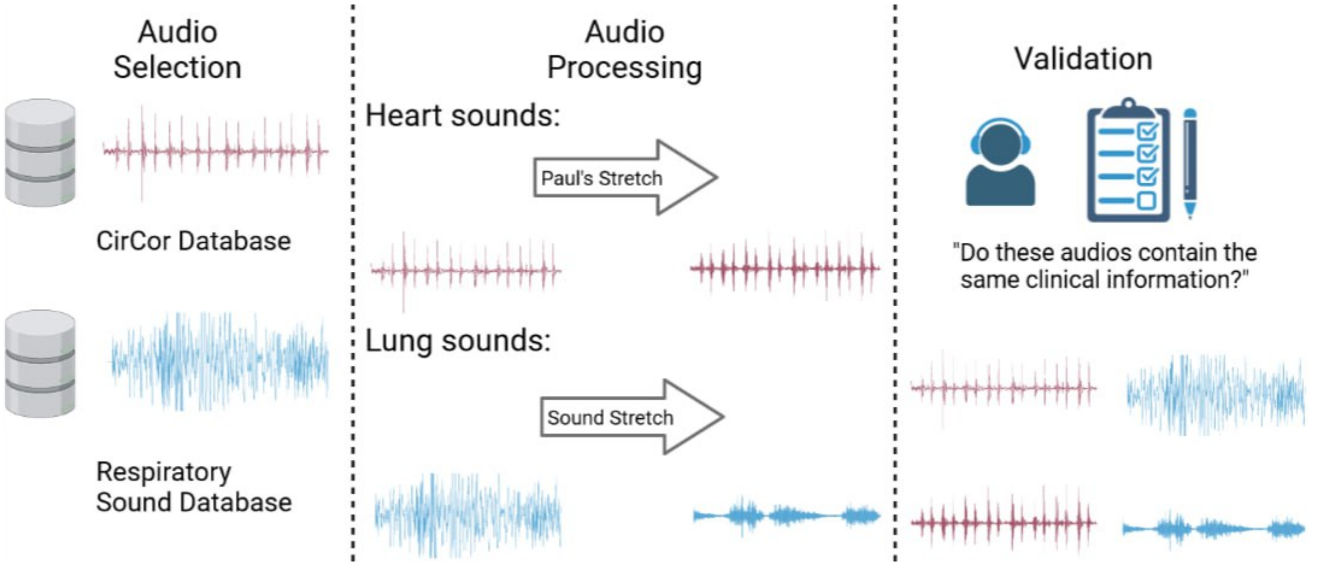

Figure 1: Summary of the methodology followed in this research. The first step consisted of selecting audio samples from two different databases, the second step was processing the audio samples and the third one consisted of validating the processed recordings with final-year medical students. (Created with BioRender.com)

were obtained using a 3M Littmann 3200 stethoscope. Ten audio samples were manually selected, five of which contained heart murmurs, and the other five were normal.

The second database contains normal and abnormal (or pathological) breath sounds [26, 28]. These audio samples were obtained from adults using four methods: an AKG C417L microphone, a 3M Littmann Classic II SE stethoscope, a 3M Litmmann 3200 stethoscope, and a WelchAllyn Meditron Master Elite stethoscope. Twenty of these audio samples were selected, six corresponded to normal sounds (three vesicular, three tracheal) and fourteen were abnormal, which included: seven sounds containing pure crackles, five containing pure wheezes, one containing crackles and wheezing, and one audio sample containing a rhonchus.

2.2. Dynamic range compressor

To enhance the audio samples of the database, the signal peaks were reduced by a dynamic range compressor. This is done in order to reduce the impact of eventual noises and high-volume peaks produced by sliding and friction when auscultating. The parameters of the implemented compressor are:

- Threshold: level in Decibels at which compression begins.
- Ratio: the fraction in which it is needed to compress the parts of the signal that exceed the threshold
- Attack and Release times: controls how fast the compressor acts
- Knee width: for the aim of this compressor, a soft knee was implemented to make the change in compression less noticeable. The Knee width controls how soft the change will be.
- Make-up gain: allows the input and output loudness to match

To implement the compressor, the process described in [29] was used and some parameters were automated with the procedure described in [30], in order to manually adjust as few inputs as possible. This can be done since most of the parameters are used in separate blocks.

Since the sound signals to be used are time-varying, it is desirable that these parameters are also time-varying, so the crest factor as a short-term signal measure method proposed in [30] was used to obtain their values. The maximum attack time ($T_{A_{max}}$) used was $80ms$ and the maximum release time ($T_{R_{max}}$) used was $1s$. The ratio was set to infinity as in [30] and the knee width at $1dB$. The missing parameter to be automated is the threshold, which was set to the average between the RMS value of the signal and the peak of the signal.

2.3. Filtering

This step was carried out differently for cardiac and respiratory audio samples. For cardiac ones, first, they were filtered with a band-pass second-order Butterworth filter with a low-cut-off frequency of 20Hz and a high-cut-off frequency of 600Hz, then they were normalized with regard to the maximal value. The respiratory audio samples were also filtered with a band-pass second-order Butterworth filter, but with low- and high- cut-off frequencies of 100 and 2000 Hz respectively. Then they were also normalized.

The cut-off frequencies were chosen according to the clinically important ranges described in 1.3.

2.4. Pitch shifting using Paul's Extreme Sound Stretch algorithm

The Paul's Extreme Sound Stretch algorithm is designed to greatly increase the duration of audio samples while maintaining a subjective quality of them [31]. The [31] algorithm consists of:

1. Segmentation of the recording into small overlapping windows. In this case, the window size chosen was 0.05 seconds, overlapping by 0.0125 seconds.
2. FFT is applied to each segment. The original phases are replaced by random phases.
3. IFFT is applied.
4. The signal is reconstructed by increasing the space between windows.
5. The sampling frequency is increased by the desired amount to increase the pitch of the recording. In this case, it was increased by double.

This algorithm was used for the enhancement of the cardiac sounds. It was not used for respiratory sounds because some specific characteristic sounds like crackles are too sharp and fast-transient sounds such as these are very sensitive to the change of spacing between windows, so the high-frequency content of the sound is affected.

2.5. Pitch shifting using the *SoundTouch* library

SoundTouch is an audio processing library open-source created by Olli Parviainen that allows to modify pitch, speed, and other parameters independent of each other [32]. To increase the pitch of an audio recording, the sample is first stretched (time-stretching) using the synchronized overlap-and-add method (SOLA) method [33], which is an algorithm used for modifying the rate of an audio sample without losing quality or affecting the pitch. Then, the audio is brought to its original length using sample rate transposing, which performs a cubic interpolation and applies an anti-alias filter, allowing to reduce the length of the sample affecting the pitch.

The SOLA algorithm consists of cutting the audio sample into 10-100ms sections and then joining them by repeating or deleting samples, depending on whether you want to lengthen or shorten the audio sample. To prevent discontinuity problems, the places where the sequences are joined are partially overlapped [34].

This algorithm was used for the enhancement of the respiratory sounds, because it adjusts better to sounds that are sharp and contain fast-transient sounds, such as crackles.

2.6. Validation of the clinical usefulness of the algorithm

The evaluation of the proposed algorithms was carried out by means of an online survey where the respondents were final-year medical students who volunteered to participate. The objective was to verify that the selected processing maintains the relevant clinical information of the lung and heart sounds included in the study. To carry this out, 30 sounds from the *open source* database detailed in section 2.1 were used, 20 of which were respiratory and 10 were cardiac. The survey was designed in *Survey Monkey* consisting of 30 questions. In each question, participants were presented with two audio samples, one of which was the unprocessed lung or heart audio sample, and the other one was the same audio sample processed with the corresponding pitch-shifting procedure. The questions were presented in random order and the participants were asked whether both audio samples contained the same clinical information, with the possible answers being “Yes”, “No” and “Other (specify)”. The last response was added in case the participants desired to incorporate supplementary information or had any issues with a specific recording. The instructions included using headphones to standardize listening and eliminate variability in the results due to this factor. The details about the characteristics of each audio sample are shown in the table 1.

3. RESULTS

Responses were collected from 40 evaluators, who on average felt that a 74% of the processed sounds do contain the same information as the original recording, while 24% of the audio samples did not. A total of 1% of the answers were “Other (specify)”. One of the samples was discarded because 10% of the respondents answered that it was too noisy to recognize the pathological sound.

For respiratory sounds, a 72% of the respondents considered that the processed sounds contained the same information as the original (tables 2 and 3). The audio sample with the worst score was the rhonchus, followed by the expiratory crackles, vesicular breath sounds and tracheal breath sounds. The sounds with the best scores were inspiratory crackles, wheezes, and also an audio sample with both these sounds present. Some comments that evaluators made in the “Other” answer were:

- Audio 102_1b1_Ar: in the processed audio sample, heart sounds are heard more faintly, which may affect the clinical decision if the disease has a heart origin.
- Audio 125_1b1_Tc: the processed audio sample is noticeably better
- Audio 135_2b1_Pl: a respondent was unable to identify the clinical finding in the original audio sample
- Audio 194_1b1_Lr: a respondent was unable to listen to breath sound in the original audio sample

ID	Type	Characteristic
2530_PV	Cardiac	Normal
39043_PV	Cardiac	Normal
40058_PV	Cardiac	Normal
50673_PV	Cardiac	Normal
68213_TV	Cardiac	Normal
9979_MV	Cardiac	Murmur
40840_PV	Cardiac	Murmur
49630_TV	Cardiac	Murmur
50238_MV	Cardiac	Murmur
73497_AV	Cardiac	Murmur
102_1b1_Ar	Respiratory	Vesicular breath sound
125_1b1_Tc	Respiratory	Tracheal sound
182_1b1_Tc	Respiratory	Tracheal sound
183_1b1_Pl	Respiratory	Vesicular breath sound
194_1b1_Lr	Respiratory	Vesicular breath sound
217_1b1_Tc	Respiratory	Tracheal sound
107_3p2_Pl	Respiratory	Expiratory crackle
107_3p2_Tc	Respiratory	Inspiratory wheeze
109_1b1_Ll	Respiratory	Inspiratory crackle
117_1b2_Tc	Respiratory	Expiratory wheeze
122_2b1_Al	Respiratory	Expiratory wheeze
122_2b1_Tc	Respiratory	Expiratory wheeze
122_2b1_Ar	Respiratory	Rhonchus
130_3p2_Pl	Respiratory	Expiratory crackle
130_2b4_Ll	Respiratory	Expiratory crackle
135_2b1_Pl	Respiratory	Expiratory crackle
158_1p3_Pr	Respiratory	Crackles, wheezing
169_1b1_Lr	Respiratory	Inspiratory crackle
201_1b1_Ar	Respiratory	Inspiratory wheeze
226_1b1_Ll	Respiratory	Expiratory crackle

Table 1: Audio samples' IDs (copied directly from the respective database), their type (cardiac or respiratory), and their respective characteristic(s). In the cardiac case, murmurs represent abnormal sounds. For the respiratory case, tracheal and vesicular breath sounds correspond to normal respiratory sounds, while crackles, wheezes, and rhonchus represent abnormal ones.

In the case of heart sounds (table 4) it can be observed that 80% of the respondents indicated that the sounds did contain the same information. This percentage is slightly lower for normal sounds (78%) than for murmurs (83%). The only relevant comment in the "Other" answer made from respondents was for recording 40058_PV and made note of important background noise in the newly processed audio sample, which made it difficult to interpret.

4. DISCUSSION

In the case of respiratory sounds, we found that abnormal sounds obtained a higher score than normal sounds. Most pathological sounds had a score above 75%, with the exception of expiratory crackles and rhonchus. In the case of the rhonchus sound, there was only one sample in the survey and it is non-representative, so the processing of more samples could show different results. In most cases, the algorithm employed did maintain the clinical information, and indeed, when participants commented on some recordings, they reported that the processed sample was much better and helped with certain subtleties in the original samples, even though the processed and original audio samples were presented randomly.

In the case of heart sounds, the results were similar in normal and abnormal groups. According to some of the comments made by participants, heart murmurs were made clearer with the algorithm.

Overall, the results showed that clinical information was maintained in more than 70% of the cases. However, there are some important limitations to address. In some audio samples, the processing introduced high-frequency sounds, which could have been interpreted as a difference in the clinical information presented in the them. This seemed to be particularly important in the sample that contained a rhonchus sound.

In addition, these audio samples were presented without an adequate clinical history, which is essential to diagnosis.

Respiratory Sound	No	Yes	Other
Abnormal	25.4%	73.6%	1.1%
Normal	32.0%	66.0%	2.0%
Total average	**27.1%**	**71.6%**	**1.3%**

Table 2: Answers to question "do these audio samples contain the same clinical information?" for respiratory sounds, separated into normal and abnormal.

Respiratory Sound	No	Yes	Other
Expiratory crackles	37.5%	60.5%	2.0%
Inspiratory crackle	15.0%	83.8%	1.3%
Crackles, wheezing	5.0%	95.0%	0.0%
Vesicular breath sound	32.5%	63.8%	3.8%
Tracheal breath sound	31.7%	67.5%	0.8%
Ronchi	50.0%	47.5%	2.5%
Expiratory wheezing	16.7%	83.3%	0.0%
Inspiratory wheezing	16.3%	83.8%	0.0%
Total average	**27.1%**	**71.6%**	**1.3%**

Table 3: Answers to question "do these audio samples contain the same clinical information?" for respiratory sounds separated by characteristic sound.

Heart sounds	No	Yes	Other
Normal	20.5%	78.0%	1.5%
Murmur	18.0%	82.0%	0.0%
Total average	**19.25%**	**80.0%**	**0.75%**

Table 4: Answers to question "do these audio samples contain the same clinical information?" for heart sounds. Murmurs represent the abnormal ones.

Furthermore, the participants did not use standardized equipment in order to listen to the sounds and, even though they were asked to wear headphones, the quality and perception of sound could have varied between them.

5. CONCLUSION AND FURTHER RESEARCH

This study shows that lung and heart sounds maintain their clinical information when using pitch shift algorithms based on the SoundStretch and PaulStretch methods, respectively. This would allow the use of a smartphone in a telemedicine setting for the playback of these sounds. Some studies report that capturing lung and heart sounds with a smartphone could be feasible, and thus, it could be employed as a recording device by the patient, and as a playback device for the clinician.

Further research is needed regarding the quality of lung and heart sounds captured by smartphones compared to those recorded by a digital stethoscope, and also whether the algorithms introduced in this paper can improve those sounds. Also, validation of the results with more experienced physicians would allow for a better comprehension of our findings.

6. ACKNOWLEDGMENTS

This work was funded by ANID – Millennium Science Initiative Program – ICN2021_004, ANID - Fondecyt grant #1161328 and supported by ANID - Millennium Science Initiative Program ICS2019_024.

7. REFERENCES

[1] R. G. Stovel, N. Gabarin, R. B. Cavalcanti, and H. Abrams, "Curricular needs for training telemedicine physicians: A scoping review," *Medical Teacher*, vol. 42, no. 11, pp. 1234 – 1242, 2020.

[2] J. Matusitz and G.-M. Breen, "Telemedicine: Its effects on health communication," *Health Communication*, vol. 21, no. 1, pp. 73–83, 2007.

[3] A. D. Cerbo, J. C. Morales-Medina, B. Palmieri, and T. Iannitti, "Narrative review of telemedicine consultation in medical practice," *Patient Preference and Adherence*, vol. 9, pp. 65–75, 2015.

[4] L. M. Hilty, P. Arnfalk, L. Erdmann, J. Goodman, M. Lehmann, and P. A. Wäger, "The relevance of information and communication technologies for environmental sustainability – a prospective simulation study," *Environmental Modelling Software*, vol. 21, no. 11, pp. 1618–1629, 2006.

[5] D. M. Hilty, A. Crawford, J. Teshima, S. Chan, N. Sunderji, P. M. Yellowlees, G. Kramer, P. O'neill, C. Fore, J. Luo, and S.-T. Li, "A framework for telepsychiatric training and e-health: Competency-based education, evaluation and implications," *International Review of Psychiatry*, vol. 27, no. 6, pp. 569–592, 2015.

[6] N. Sunderji, A. Crawford, and M. Jovanovic, "Telepsychiatry in graduate medical education: A narrative review," *Academic Psychiatry*, vol. 36, no. 1, pp. 55 – 62, 2015.

[7] N. Shaikh, C. U. Lehmann, P. H. Kaleida, and B. A. Cohen, "Efficacy and feasibility of teledermatology for paediatric medical education," *Journal of Telemedicine and Telecare*, vol. 14, no. 4, pp. 204–207, 2008.

[8] V. G. Chellaiyan, A. Y. Nirupama, and N. Taneja, "Telemedicine in india: Where do we stand?" *Journal of Family Medicine and Primary Care*, vol. 8, no. 6, pp. 1872–1876, 2019.

[9] J. Baker and A. Stanley, "Telemedicine technology: a review of services, equipment, and other aspects," *Current Allergy and Asthma Reports*, vol. 18, no. 11, p. 60, 2018.

[10] S. Ahmed, K. Sanghvi, and D. Yeo, "Telemedicine takes centre stage during covid-19 pandemic," *BMJ Innovations*, vol. 6, no. 4, pp. 252–254, 2020. [Online]. Available: https://innovations.bmj.com/content/6/4/252

[11] J. Ilowite, G. Lisker, and H. Greenberg, "Digital health technology and telemedicine-based hospital and home programs in pulmonary medicine during the covid-19 pandemic," *American Journal of Therapeutics*, vol. 28, no. 2, pp. e217–e223, 2021.

[12] I. Berlanga Fernández, V. E. Gozálvez Pérez, P. Renés Arellano, and J. I. Aguaded Gómez, "Diez años de smartphones : un análisis semiótico-comunicacional del impacto social de la telefonía móvil," 2018.

[13] Uswitch, "UK mobile phone statistics, 2023," 2023. [Online]. Available: https://www.uswitch.com/mobiles/studies/mobile-statistics/

[14] Eurostat. (2016, December) Internet use by individuals - almost 8 out of 10 internet users in the eu surfed via a mobile or smart phone in 2016… different patterns across member states in managing personal information. [Online]. Available: https://ec.europa.eu/commission/presscorner/detail/en/STAT_16_4477

[15] F. A. Allaert, L. n. Legrand, N. Abdoul Carime, and C. Quantin, "Will applications on smartphones allow a generalization of telemedicine?" *BMC Medical Informatics and Decision Making*, vol. 20, no. 1, p. 30, 2020.

[16] E. C. Larson, T. Lee, S. Liu, M. Rosenfeld, and S. N. Patel, “Accurate and privacy preserving cough sensing using a low-cost microphone,” in *Proceedings of the 13th International Conference on Ubiquitous Computing*. New York, NY, USA: Association for Computing Machinery, 2011, p. 375–384.

[17] E. C. Larson, M. Goel, G. Boriello, S. Heltshe, M. Rosenfeld, and S. N. Patel, “Spirosmart: Using a microphone to measure lung function on a mobile phone,” in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ser. UbiComp ’12, 2012, pp. 280–289.

[18] E. Agu, P. Pedersen, D. Strong, B. Tulu, Q. He, L. Wang, and Y. Li, “The smartphone as a medical device: Assessing enablers, benefits and challenges,” in *2013 IEEE International Workshop of Internet-of-Things Networking and Control (IoT-NC)*, 2013, pp. 48–52.

[19] E. Olsen, “Digital health apps balloon to more than 350,000 available on the market, according to iqvia report,” *El País*, 8 2021. [Online]. Available: https://www.mobihealthnews.com/news/digital-health-apps-balloon-more-350000-available-market-according-iqvia-report

[20] M. Abella, J. Formolo, and D. G. Penney, “Comparison of the acoustic properties of six popular stethoscopes,” *The Journal of the Acoustical Society of America*, vol. 91, no. 4, pp. 2224–2228, 1992.

[21] F. Bertrand Z., D. Segall K., I. Sánchez D., and P. Bertrand N., “La auscultación pulmonar en el siglo 21,” *Revista chilena de pediatría*, vol. 91, pp. 500 – 506, 2020.

[22] S.-H. Kang, B. Joe, Y. Yoon, G.-Y. Cho, I. Shin, and J.-W. Suh, “Cardiac auscultation using smartphones: Pilot study,” *JMIR Mhealth Uhealth*, vol. 6, no. 2, p. e49, 2018.

[23] H. Ferreira-Cardoso, C. Jácome, S. Silva, A. Amorim, M. T. Redondo, J. Fontoura-Matias, M. Vicente-Ferreira, P. Vieira-Marques, J. Valente, R. Almeida, J. A. Fonseca, and I. Azevedo, “Lung auscultation using the smartphone—feasibility study in real-world clinical practice,” *Sensors*, vol. 21, no. 14, 2021.

[24] H. Ferreira-Cardoso, “Pulmonary auscultation using mobile devices – feasibility study in respiratory disease,” Master’s thesis, Universidade Do Porto, Portugal, 2021.

[25] J. Oliveira, F. Renna, P. D. Costa, M. Nogueira, C. Oliveira, C. Ferreira, A. Jorge, S. Mattos, T. Hatem, T. Tavares, A. Elola, A. B. Rad, R. Sameni, G. D. Clifford, and M. T. Coimbra, “The circor digiscope phonocardiogram dataset (version 1.0.0).” Physionet, 2022.

[26] B. M. Rocha, D. Filos, L. Mendes, I. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jácome, A. Marques, R. P. Paiva, I. Chouvarda, P. Carvalho, and N. Maglaveras, “A respiratory sound database for the development of automated classification,” in *Precision Medicine Powered by pHealth and Connected Health*, N. Maglaveras, I. Chouvarda, and P. de Carvalho, Eds. Springer Singapore, 2018, pp. 33–37.

[27] J. Oliveira, F. Renna, P. D. Costa, M. Nogueira, C. Oliveira, C. Ferreira, A. Jorge, S. Mattos, T. Hatem, T. Tavares, A. Elola, A. B. Rad, R. Sameni, G. D. Clifford, and M. T. Coimbra, “The circor digiscope dataset: From murmur detection to murmur classification,” *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 6, pp. 2524–2535, 2022.

[28] B. M. Rocha, D. Filos, L. Mendes, I. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jácome, A. Marques, R. P. Paiva, I. Chouvarda, P. Carvalho, and N. Maglaveras, “Respiratory sound database: Use audio recordings to detect respiratory diseases.” Kaggle, 2018. [Online]. Available: https://www.kaggle.com/datasets/vbookshelf/respiratory-sound-database

[29] D. Giannoulis, M. Massberg, and J. Reiss, “Digital dynamic range compressor design—a tutorial and analysis,” *AES: Journal of the Audio Engineering Society*, vol. 60, no. 6, pp. 399–408, 2012.

[30] d. Giannoulis, m. massberg, and j. d. reiss, “Parameter automation in a dynamic range compressor,” *AES: Journal of the Audio Engineering Society*, vol. 61, no. 10, pp. 716–726, 2013.

[31] P. Nasca, “Paul’s extreme sound stretch,” 2012. [Online]. Available: https://hypermammut.sourceforge.net/paulstretch/

[32] O. Parviainen, “Soundtouch audio processing library,” 2022. [Online]. Available: https://hypermammut.sourceforge.net/paulstretch

[33] S. Roucos and A. Wilgus, “High quality time-scale modification for speech,” in *ICASSP ’85. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 10, 1985, pp. 493–496.

[34] O. Parviainen, “Time and pitch scaling in audio processing,” *Software Developer’s Journal*, vol. 4, 2006.