

Automatic detection of patient with respiratory diseases using lung sound analysis

Gaëtan CHAMBRES

Univ. Bordeaux, LaBRI,
Bordeaux INP, CNRS, UMR 5800
Talence, FRANCE

Pierre HANNA

Univ. Bordeaux, LaBRI
Bordeaux INP, CNRS, UMR 5800
Talence, FRANCE

Myriam DESAINTE-CATHERINE

Univ. Bordeaux, LaBRI
Bordeaux INP, CNRS, UMR 5800
Talence, FRANCE

Abstract—In modern medicine, every cardiac assessment or respiratory check-up includes an audio auscultation during which one the medical specialist listens to sounds from the patient body with different tools (stethoscope, sonography). This shows how important sound analysis is for heart and lungs disease detection. During the ICBHI 2017 challenge, a database of 920 records acquired from 126 subject, was used to find a method that predicted if a respiratory cycle contains, or not, adventitious sounds like crackles, wheezes or both of them. The team which submits the best results reached around 50% of correct detection. Using a machine learning approach with a boosted decisional tree model and more audio features leads to the same results. A new approach consists in creating a new model at the patient level, which is able to decide if a patient sounds sick or not, by taking as input the predicted results of the first classification model. This new model permits to reach 85% of good predictions and could be used as a tool for helping doctors to make better diagnosis.

Index Terms—Lung, pulmonary disease, respiratory sound, audio features, classification, machine learning, patient level, diagnosis, pathology

I. INTRODUCTION

In modern medicine, diagnostics are more often related to visual or audio information. Medical information can be acquired from different methods, and it is mainly represented to the medical specialist as images or sounds. Being able to correctly identify health problem regarding this information requires experience and time, but according to the World Health Organization (WHO) statistics [1], 45% of the WHO Member States report to have less than 1 physician per 1000 population, the WHO ratio recommendation. With such bad statistics, taking into account that a diagnosis implies to study individually every patient during a non compressible time and that health specialists are already overbooked, conditions are not optimum for them, and mistakes can be made. This is why finding new ways to help doctors to save time is a priority. Automatic and reliable tools can help to diagnosis more people, or it can help specialists to make less mistakes due to the work overload. Computer science can bring this new tools.

Since several decades, computer science improvements constantly ameliorate the ability to analyze automatically media data in real time. Indeed the diagnosis-help tools must be able

to process image and/or audio information. Computer science could help nursing staff or doctors by proposing faster and reliable tools for diagnosis [2]. It could give to the patient customizable tools for medical monitoring.

Every cardiac assessment or respiratory check-up includes an audio auscultation in which one the medical specialist listens to sounds from the patient body with different tools (stethoscope, sonography). This shows how important sound analysis is for heart and lungs disease detection. As heart disease remains the leading cause of death globally [3], there are many studies about heart sound identification [4]. A review [5] published in 2012 about signal processing techniques for heart sound analysis conclude that "the ability to correctly acquire, analyze and interpret heart sound signals for improved clinical diagnostic processes has become a priority". Since then, there are lots of improvements, for processing records taken in noisy environments [6] for example. Furthermore, new kinds of methods drastically improve the domain, as machine learning and deep learning [7]. These approaches contribute a lot to computer vision [8], [9] or audio analysis [10], [11]. By extension, there are several applications on medical studies and heart sound analysis [12], [13].

Analysis of respiratory sounds is less studied, but there are some existing works [14]. A review [15] published in 2007, highlights that previous researches permit to identify markers, such as crackles or wheezes. As said before, with the apparition of machine and deep learning, the performances of sound detection and classification are drastically improved [7] and studies about lung sound analysis take advantage of this progression [16] [17]. With identified markers, lung sound analysis can be transformed into a classification problem [18], a kind of problem that machine learning performs well to solve. This seems to be a relevant approach, but this kind of analysis more often focuses on the properties of the sound record, and not always on the patient level. Indeed, a doctor will listen several expirations and several inspirations, at several places on the patient body before taking a decision. This article raises the problem of using classification results based on sound content (micro-level) to take a decision at the patient level (macro-level), suggesting it the patient sounds sick or healthy.

The main contribution of this work is to suggest a new approach in audio classification. In some cases, here for lung pathologies, machine learning for audio classification based on sound content is not the best solution, or at least not alone. In this study, a machine learning approach is presented and performs as well as the previous state of the art. Using this classification model and extrapolating the results to take a decision on the patient level leads to better results.

II. LUNG SOUND ANALYSIS

A. Adventitious respiratory sounds

According to Reichert and al [15], an adventitious sound relates to *"additional respiratory sounds superimposed on normal breath sounds. It can be continuous (like wheezes) or discontinuous (such as crackles). Some of them (like squawks) have both characteristics. The presence of such sounds usually indicates pulmonary disorders."* Wheeze and crackle are the main categories of adventitious sounds. Crackles, which can reflect a pathology on lung tissue or airways, are discontinuous and explosive sounds. They are characterized by their specific waveform, their duration, and their location in the respiratory cycle. In opposition, a wheeze is a continuous sound with a musical character. It is characterized by periodic waveforms with a dominant frequency usually over 100 Hz and with a duration over 100 ms. When wheezes are observed, they usually indicate airways obstruction. A diagnosis-help tool could be seen as a classification problem, with a prediction model based on algorithm that detects this kind of markers. To be able to create a classification model, an annotated database with sufficient subjects and information is needed.

B. ICBHI 2017 Challenge - The database

The 2017 International Conference on Biomedical Health Informatics (ICBHI) organized a scientific challenge about automatic identification of crackles and wheezes in respiratory sounds. For this challenge, a specific database [19] was created. This database was made by collecting records of respiratory sounds from different research teams with different recording tools.

The work discussed in this paper is based on this database because it is diversified and fully annotated. For a given patient, one or several records are provided with one annotation file by record. Table I is an example of the content of an annotation file for a given record.

According to this table I, for each cycle contained in the given record, their position in time is given, and information about containing adventitious sound or not is given as a binary score. In this example, *cycle 1* starts at 1.018 second, ends at 3.411 second and only contains crackles. *Cycle 5* starts at 10.923 second, ends at 13.292, and only contains wheezes. According to the binary score, *cycle 3* contains both crackles and wheezes, and no respiratory cycle of this record sounds normal, with no adventitious sound (binary score would be 0 for crackles and 0 for wheezes).

Table II indicates the repartition of patients, records and the different type of cycles for the train and test set. This table

TABLE I
CONTENT OF THE ANNOTATION FILE FOR THE RECORD
107_2B4_AL_MC_AKGC417L [19]

Cycles	Start	End	Crackles	Wheezes
cycle 1	1.018	3.411	1	0
cycle 2	3.411	5.827	1	0
cycle 3	5.827	8.339	1	1
cycle 4	8.339	10.923	1	0
cycle 5	10.923	13.292	0	1
cycle 6	13.292	16.018	1	0
cycle 7	16.018	18.482	1	0
cycle 8	18.482	19.542	1	0

shows one limitation of the database. Crackles and wheezes repartition is not well balanced between the train and test sets. Another important information is the fact that there is a majority of normal cycles in both sets. Despite the unbalance, it was decided to keep this train-test separation because it was the split used during the ICBHI Challenge [19]. Having the same split permits to compare the obtained results to the state of the art.

TABLE II
REPARTITION OF THE DATA IN THE ICBHI CHALLENGE DATABASE [19]

Database	Train set	Test set	Total
Nb of patients	79	49	128
Nb of records	539	381	920
Nb of cycles with crackles	1247	617	1864
Nb of cycles with wheezes	533	353	886
Nb of cycles with both	366	140	506
Nb of normal cycles	2152	1490	3642
Nb total of cycles	4298	2600	6898

C. ICBHI 2017 Challenge - Submitted results

The challenge was divided in two parts, train and test. During the training phase, research teams accessed the public part of the database (the train set) to test their methods. During the test phase, research teams accessed the private part of the database (the test set) to be able to compute the accuracy Acc of their method. This last score was the final one in the evaluation of the method. It is computed by making the mean of sensitivity Se and specificity Sp for the predictions. As a reminder, these are the used formulas :

$$Se = \frac{(TP_C + TP_W + TP_B)}{(TP_C + FN_C) + (TP_W + FN_W) + (TP_B + FN_B)} \quad (1)$$

$$Sp = \frac{TP_N}{TP_N + FN_N} \quad (2)$$

$$Acc = \left(\frac{Se + Sp}{2} \right) * 100 \quad (3)$$

where TP_i stands for *TruePositive* in class i and FN_i for *FalseNegative* in class i . TP_i corresponds to the number of correctly classed elements in class i and $TP_i + FN_i$ the number total of elements that should be in class i . The different classes i are C for crackles, W for wheezes, B for both and N for normal.

It was decided to choose two of the results proposed during the ICBHI Challenge as the main comparison for this work, because these research teams proposed two different methods to classify respiratory cycles into the four categories of adventitious sound described before. As the idea is to propose a new way of thinking using machine learning, a focus is made on similar method of classification. The table III summarizes the results of both methods.

TABLE III
SUMMARY OF THE STATE OF THE ART RESULTS

Method	Se	Sp	Acc	Score
HMM	0.4232	0.5669	49.50%	39.37%
NLSp	0.4890	0.7780	49.98%	49.86%

For the first method, *HMM* stands for Hidden Markov Model classification. After preprocessing records to reduce noise and the heart beats, Mel-Frequency Cepstral Coefficients (MFCC) are used as inputs in a hidden Markov model. The model permits to compute the probability for a respiratory cycle to contain crackles or wheezes. The best score during train phase is $Se = 0.4232$, $Sp = 0.5669$ and $Acc = 49.50\%$, and the best score during test phase is $Score = 39.37\%$.

For the second method, *NLSp* stands for Non-Linear Spectral features extraction method. After preprocessing records to reduce the sample rate from 44100Hz to 4000Hz, and minimize noise effect with a Butterworth band pass filter, a non-linear resonance based wavelet decomposition is applied. This decomposition gives three channels (high,low,residue) and features are computed in each channel with a short time Fourier transform. Finally, the classification prediction is computed by a Support Vector Machine (SVM). The best score during the test phase is $Score = 49.86\%$ while the method gave better results on the train phase. To be able to compare, these are the closest results obtained during the train phase : $Se = 0.4890$, $Sp = 0.7780$ and $Acc = 49.98\%$.

III. PROPOSED METHOD

As the goal is to demonstrate the importance of adopting a patient-level point of view for the classification, the proposed method is divided in two parts. The first part is about the classification of the respiratory cycles depending of the adventitious sounds and the second part is about extrapolate the classification results to consider the patient classification. These parts are respectively named the micro-level part and the macro-level part.

A. Micro-level part

The main goal of this part is to classify individually every respiratory cycles depending if adventitious sounds are detected or not. For that, all records are taken one by one, and for each records, features are extracted on the signal window containing every cycles. The classification of each cycle is computed with a boosted decisional tree, which gives, according to the features, the probability to belong to every class. The higher probability is kept as the decision.

1) *Features computation*: Features were computed thanks to the *Essentia* library [22]. The automatic *Freesound* features extractor is used, which is designed for batch computations on large sound collections and is used by *Freesound*¹. It is configured from a profile specifying the parameters to use for the features computation : in this case, a frame size of 2048, a hop size of 1024, a Blackmann-Harris window and finally, the time window corresponding to the respiratory cycle. A complete description of computed features can be found on the *Essentia* website². The different categories are the *lowlevel features* (spectral information like barkbands, energybands, melbands, mfcc, etc), the *rythm features* (beats loudness, bpm, etc), the *SFX features* (harmonicity and inharmonicity information) and the *tonal features* (chords strength, tuning frequency, etc).

To maximize the chance to correctly represent the audio signal, statistical values are computed from the features: standard deviation, variance, minimum and maximum, median, mean, means of first and second derivatives and variances of first and second derivatives.

After several classification tests based on the features, it appears that not all of them were relevant. A focus was done on *lowlevel features* because they provide spectral information, which is useful to identify wheezes, and information about punctual events, which is useful to identify crackles.

2) *The classification model*: The classification system is based on a boosted decisional tree algorithm. Each leaf of the tree will be a class, and the branches will be the features conjunction to reach the different leaves. Two approach with different kind of classification model were tested, a monaclass one and a multiclass one.

The monaclass approach is a succession of two monaclass model, where each tree have two leaves: *crackles* and *no crackles* for the first one, then *wheezes* and *no wheezes* for the second one. They respectively gives the probability for every respiratory cycles to contain crackles and then, the probability for every cycles to contain wheezes. The final decision is taken according to the diagram on figure 1 where the *YES* and *NO* ways are decided according to the highest probability.

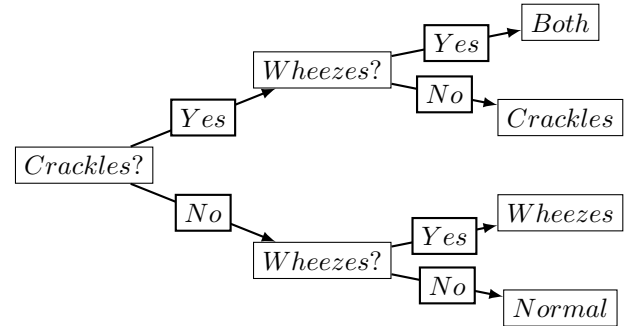


Fig. 1. Monoclass model : Decision diagram for final prediction

¹<https://freesound.org/>

²http://essentia.upf.edu/documentation/freesound_extractor.html

The second kind of model is a multiclass model, where a unique tree has 4 leaves for the 4 different classes: *normal*, *crackles*, *wheezes* or *both*. The final decision is taken by keeping the highest prediction of the four, one for each possible class.

3) *Results*: The system takes all the features as input and makes features conjunction, represented as branches of the tree, to reach the leaves. For a given respiratory cycle, by making the conjunction of the features, probability to belong to each class is given. The system is a boosted tree model. That means when the model is trained, several iterations are done to re-train the cycles which were not correctly classified. This permits to be more robust and precise. Parameters of the boosted decisional tree model were empirically set: the max depth of trees is 3, the maximum of boosting iterations is 100 and the learning rate is 0.1.

In the state of the art, the main features used are the mel-frequency cepstral coefficient (MFCC). To have a reliable comparison, a first bench of tests was done only using the 13 first mel-frequency cepstral coefficients as input features. In a search of improved precision, new results are computed using the *lowlevel* features. This two bench of tests were done for both monoclase and multiclass model. All the results are summarized in table IV, where *Mono/mfcc* and *Mono/lowlevel* indicate predictions made with the monoclase model, respectively with MFCC and lowlevel features as input. In the same way, *Multi/mfcc* and *Multi/lowlevel* indicate predictions made with the multiclass model, respectively with MFCC and lowlevel features as input.

TABLE IV
SUMMARY OF THE PREDICTION RESULTS

Model	Se	Sp	Acc
Mono/mfcc	0.2216	0.7604	49.10%
Mono/lowlevel	0.3505	0.5819	46.62%
Multi/mfcc	0.2081	0.7805	49.43%
Multi/lowlevel	0.3550	0.6255	49.02%

Although the accuracy is comparable to the state of the art because of the computation formula, sensitivity and specificity are not balanced. For the monoclase model with only MFCC features for example, normal cycles are correctly classed 76% of the time, but cycles with adventitious sounds are correctly classed only 22% of the time. Taking a look at the different confusion matrices permits to explain visually this problem. Table V refers to the confusion matrix of *Mono/mfcc*, table VI to *Mono/lowlevel*, table VII to *Multi/mfcc* and table VIII to *Multi/lowlevel*. In each table, C_i corresponds to the number of correctly predicted cycles for the class i , T_i corresponds to the number of cycle that should be predicted to be in class i and i corresponds to the different classes: n for normal, c for crackles, w for wheezes and b for both crackles and wheezes.

TABLE V
CONFUSION MATRIX OF THE PREDICTION OBTAINED WITH THE MONOCLASS MODEL AND MFCC FEATURES AS INPUT.

	Cn	Cc	Cw	Cb	Total
Tn	1133	224	111	22	1490
Tc	357	221	23	16	617
Tw	246	71	25	11	353
Tb	105	25	10	0	140
Total	1841	541	169	49	2600

TABLE VI
CONFUSION MATRIX OF THE PREDICTION OBTAINED WITH THE MONOCLASS MODEL AND LOWLEVEL FEATURES AS INPUT.

	Cn	Cc	Cw	Cb	Total
Tn	867	250	228	145	1490
Tc	298	274	22	23	617
Tw	156	43	97	57	353
Tb	65	29	28	18	140
Total	1386	596	375	243	2600

TABLE VII
CONFUSION MATRIX OF THE PREDICTION OBTAINED WITH THE MULTICLASS MODEL AND MFCC FEATURES AS INPUT.

	Cn	Cc	Cw	Cb	Total
Tn	1163	229	85	13	1490
Tc	374	211	31	1	617
Tw	265	63	20	5	353
Tb	103	29	8	0	140
Total	1905	532	144	19	2600

TABLE VIII
CONFUSION MATRIX OF THE PREDICTION OBTAINED WITH THE MULTICLASS MODEL AND LOWLEVEL FEATURES AS INPUT.

	Cn	Cc	Cw	Cb	Total
Tn	932	200	307	51	1490
Tc	283	256	33	45	617
Tw	165	38	121	29	353
Tb	68	35	20	17	140
Total	1448	529	481	142	2600

These confusion matrices permit to show several information. The models using MFCCs are better at classifying normal cycles. For the monoclase model, the precision of normal cycles correctly predicted is increased from 58% with lowlevel features (table VI) to 77% with MFCCs (table V). It is the same for the multiclass model, with a precision increased from 62% (table VIII) to 78% (table VII). On the other hand, crackles and wheezes are more correctly predicted when using the lowlevel features. It seems normal as lowlevel features catch more information about different audio events, for the crackles in particular which are less linked to spectral features. Finally, these matrices permit to see that the multiclass model is more efficient for wheezes detection and less precise for crackles detection. The precision for wheezes detection is increased from 27% (table VI) to 34% (table VIII) but the precision for crackles detection is decreased from 44% (table VIII) to 41% (table VI).

To have more readable results, table IX summarizes the results discussed from the confusion matrices, where each line of the table IX corresponds to a confusion matrix. The following values are provided in percent: the specificity Sp and the sensitivity Se_i for the three classes with c for crackle, w for wheeze and b for both. Acc stands for accuracy, with the same definition to table IV.

TABLE IX
SUMMARY OF THE CORRECTLY PREDICTED CYCLES FOR EACH CLASS IN PERCENT AND TOTAL ACCURACY

Model	Sp	Se_c	Se_w	Se_b	Acc
Mono/mfcc	76	36	7	0	49.10
Mono/lowlevel	58	44	27	13	46.62
Multi/mfcc	78	34	6	0	49.43
Multi/Lowlevel	63	41	34	12	49.02

B. Macro-level part

As the main idea of studying lung diseases with an automatic model is to propose a model which could help the lung specialist by saving time, the goal of the macro-level part is to suggest a "diagnostic" taking into account the totality of the predictions previously computed. Here, "diagnostic" means is the patient sounds sick or healthy. As a doctor not only listens one time the lung of his patient, but several times at different area of the body, we compute the different kind of cycle ratio, predicted for all the cycles of one patient. Depending on what kind of cycles appears the most, a decision is taken.

Has seen before, the classification models in the micro-level part give the probability for each respiratory cycles to be in each class depending the adventitious sounds. These results are quite good, but the hope was that when we look all the decisions for all the cycles of a patient, a trend appears. For example, for a patient with lot of crackles in his respiration, not all the cycles have to contain crackles. It could appear that one or more of them are normal. This is what looking at a trend is interesting. If this trend appears to be symptomatic, it is possible to suggest that the patient is sick. To identify the trend, the results are packed for all the cycles of the same patient. This leads to have a list of class for each patient, corresponding to the classes of the different cycles. After generating this list for every patient we want to classify, it remains to count how many cycles are normal, how many contains crackles, etc. The ratios are naturally obtained by:

$$Ratio_i = \frac{Cycles_i}{TC} \quad (4)$$

where $Cycles_i$ is the number of cycles predicted to be in class i and TC is the total of cycles analyzed for the current patient. Taking as goal to identify healthy patient, the decision should be taken depending the following question : *Is there a majority of normal cycles ?* and the two answers will be *if yes, the patient should be healthy* and *if no, the patient should be sick*. But, as explained before, and as confirmed by taking a deep look at the database, symptoms are non-systematical. It seems that crackles and wheezes do not appear on every

cycles of a sick patient and a lot of them have a majority of normal respiratory cycles, even the sick ones. This is why the decisive question is a bit transformed as follows: *Does the patient **only** have normal cycles ?* This is possible thanks to the problematic: it is more important to not miss a sick patient, thinking he is healthy, than to miss an healthy patient, thinking he is sick.

Keeping in mind that the classification results were not perfect, a threshold is used on the ratio. Two different thresholds have been empirically tested : $threshold = 0$ means that a patient is considered healthy if there is no cycle predicted with adventitious sounds at all (every cycles are predicted as normal), and $threshold = 0.1$ means that a patient is considered sick if less than 10% of the total of cycles are predicted as having adventitious sound. These 10% were empirically set to test the contribution of the threshold, but best results were obtained with $threshold = 0$. This proves that the respiratory cycles classification on the micro-level part is sufficiently strong to be used as input in the macro-level classification. The final results of the number of classified patient are visible in table X.

According to this table, 85% of the patients in the test dataset were correctly classified. It is important to keep in mind that the total number of patient is limited. For a more robust conclusion, the method should be extended to a bigger database.

TABLE X
CONFUSION MATRIX OF THE PREDICTION OF THE PATIENT HEALTH STATE

	Sick	Healty	Total
Sick	29	5	34
Healty	2	11	13
Total	31	16	47

IV. CONCLUSION

A. Discussion of the results

Looking at the sound content on respiratory sound seems to be difficult for a classification. Our classification cannot do better than the state of the art, and it seems difficult to have more than 50% of accuracy on the test set of this database. This is certainly due to the variable quality of the records, with sometimes a lot of noise and interferences. The unbalanced data could also be a cause of bad predictions, because it affects the model training. However, the new approach proposed permits to use these results to take a decision at the patient level. As a specialist will listen several cycles at different places on the patient body, we proposed to not only focus on the audio content for each respiratory cycle, but to predict a trend for each patient. This macro-level approach gives good results with an accuracy up to 85 %. It is possible to consider this model could be extended as automatic guide for diagnosis, and could help professionals by saving time or for educational purpose.

B. Perspectives

Like it was done in previous studies, the records could be preprocessed to avoid noise and interference. This could give better result for the respiratory cycles classification, but it is necessary to prove that the symptom information is not modified by the filtering step.

In macro-level part, some problems appear with the database. For example, some of the patient were annotated as having URTI for Upper Respiratory Tract Infection, but after taking a look at their records annotations it seems that no cycle present an adventitious sound : all recorded cycles are normal, with no crackle nor wheeze. With our system, we are not able to consider this kind of patient as sick. If we do not consider this type of subjects, the score could be increased up to *accuracy* = 90%.

A new approach for decision on patient level was discussed about establishing a correlation between the presence of crackles, wheezes or both and a particular pathology. Unfortunately, for 3 of the 7 pathologies contained in the dataset, Asthma, Lower Respiratory Tract Infections (URTI) and Pneumonia, there is no subject in the test set. Trying to find a correlation for these pathologies implies to consider a new repartition of the database, and lose the reliability of comparing the results to the state of the art.

The utility of features is to represents the content of the audio signal. Another idea is to compute features on the respiratory record and train a classification model to identify directly the pathology. Crackles and wheezes are ones of the main adventitious sounds, but they are not the only ones [15], so it could be interesting to not restrict the information to them. These could help to make the difference between patient with URTI and healthy patient. For the 3 other pathologies (Common Obstructive Pulmonary Disease, Bronchiolitis and Bronchiectasis), finding a profile by training the model must be a very complex problem, because of the unbalanced and limited database. Correctly identify the features reliable to detect adventitious sounds will help in the process of establishing a new and normalized database.

ACKNOWLEDGMENT

This work was done during a Master internship at LaBRI, supported by Univ. Bordeaux. We are thankful to our colleague Gyorgy Kurtag who provided expertise that greatly assisted the research. Finally, a very special thanks goes to Annick Mersier and the whole SCRIME³ team for providing all that was necessary for the smooth running of the internship.

REFERENCES

- [1] World Health Organization. Density of physicians. http://www.who.int/gho/health_workforce/physicians_density/en/, 2017. [Online; accessed 18-May-2018].
- [2] Klaus Schoeffmann, Bernd Mnzer, Michael Riegler, and Pl Halvorsen. 2017. Medical Multimedia Information Systems (MMIS). In Proceedings of the 2017 ACM on Multimedia Conference (MM '17). ACM, New York, NY, USA, 1957-1958.
- [3] World Health Organization. Cardiovascular diseases (cvds).<http://who.int/mediacentre/factsheets/fs317/en/>, 2017. [Online; accessed 16-May-2018].
- [4] X. Yang, F. Yang, L. Gobeawan, S.Y. Yeo, S. Leng, L. Zhong and Y. Su, A multi-modal classifier for heart sound recordings, in Computing in Cardiology, vol. 43, 2017.
- [5] B. Emmanuel. A review of signal processing techniques for heart sound analysis in clinical diagnosis. Journal of medical engineering and technology. vol 36. pages 303-7. 2012.
- [6] M. Abdollahpur, A. Ghaffari, S. Ghiasi and M. Mollakazemi, Detection of pathological heart sounds, in Physiological Measurement, vol.38, num. 8, pp 1616-1630, 2017.
- [7] X. Du, Y. Cai, S. Wang and L. Zhang, "Overview of deep learning," 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), Wuhan, 2016, pp. 159-164.
- [8] Q. Wu, Y. Liu, Q. Li, S. Jin and F. Li, "The application of deep learning in computer vision," 2017 Chinese Automation Congress (CAC), Jinan, 2017, pp. 6522-6527.
- [9] E. Nishani and B. io, "Computer vision approaches based on deep learning and neural networks: Deep neural networks for video analysis of human pose estimation," 2017 6th Mediterranean Conference on Embedded Computing (MECO), Bar, 2017, pp. 1-4.
- [10] Y. Petetin, C. Laroche and A. Mayoue, "Deep neural networks for audio scene recognition," 2015 23rd European Signal Processing Conference (EUSIPCO), Nice, 2015, pp. 125-129.
- [11] J. Li, W. Dai, F. Metze, S. Qu and S. Das, "A comparison of Deep Learning methods for environmental sound detection," 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, 2017, pp. 126-130.
- [12] G.D. Clifford, C. Liu, B. Moody, J. Millet, S. Schmidt and Q. Li, I. Silva and R.G. Mark, Recent advances in heart sound analysis, in Physiological Measurement, vol.38, num. 8, pp E10-E25, 2017.
- [13] V. Maknickas and A. Maknickas, Recognition of normalabnormal phonocardiographic signals using deep convolutional neural networks and mel-frequency spectral coefficients, in Physiological Measurement, vol. 38, num. 8, pp 1671-1684, 2017.
- [14] H. Polat and I. Güler, A simple computer-based measurement and analysis system of pulmonary auscultation sounds, in Journal of medical systems, vol. 28, num. 6, pp.665-672, 2004.
- [15] S. Reichert and R. Gass and E. Andr s, Analyse des sons auscultatoires pulmonaires (Analysis of respiratory sounds), in ITBM-RBM, vol. 28, pp 169-180, 2007.
- [16] G. Chang, and Y. Lai, Performance evaluation and enhancement of lung sound recognition system in two real noisy environments, in Computer Methods and Programs in Biomedicine, vol. 97, num. 2, pp 141-150, 2010.
- [17] A. Abbas and A. Fahim, An Automated Computerized Auscultation and Diagnostic System for Pulmonary Diseases; in Journal of Medical Systems, vol. 34, num. 6, pp 1149-1155, 2010.
- [18] C. Chen, W. Huang, T. Tan, C. Chang and Y. Chang, Using K-Nearest Neighbor Classification to Diagnose Abnormal Lung Sounds, in Sensors, vol. 15, num. 6, pp 13132-13158, 2015.
- [19] B. M. Rocha, D. Filos, L. Mendes, I. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jcome, A. Marques, R. P. Paiva, I. Chouvarda, P. Carvalho and N. Maglaveras, Respiratory Sound Database for the Development of Automated Classification, In Precision Medicine Powered by pHealth and Connected Health, 2017, pp 33-37
- [20] Jakovljevi N., Lonar-Turukalo T. (2018) Hidden Markov Model Based Respiratory Sound Classification. In: Maglaveras N., Chouvarda I., de Carvalho P. (eds) Precision Medicine Powered by pHealth and Connected Health. IFMBE Proceedings, vol 66. Springer, Singapore
- [21] Serbes G., Ulukaya S., Kahya Y.P. (2018) An Automated Lung Sound Preprocessing and Classification System Based OnSpectral Analysis Methods. In: Maglaveras N., Chouvarda I., de Carvalho P. (eds) Precision Medicine Powered by pHealth and Connected Health. IFMBE Proceedings, vol 66. Springer, Singapore
- [22] Bogdanov, D., Wack N., Gmez E., Gulati S., Herrera P., Mayor O., et al. (2013). ESSENTIA: an Audio Analysis Library for Music Information Retrieval. International Society for Music Information Retrieval Conference (ISMIR'13). 493-498.

³<https://scrime.u-bordeaux.fr/>