





Article

Multi-Time-Scale Features for Accurate Respiratory Sound Classification

Alfonso Monaco ¹, Nicola Amoroso ^{1,2,*}, Loredana Bellantuono ³, Ester Pantaleo ³,
Sabina Tangaro ^{1,4} and Roberto Bellotti ^{1,3}

¹ Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Bari, 70126 Bari, Italy; alfonso.monaco@ba.infn.it (A.M.); Sonia.Tangaro@ba.infn.it (S.T.); roberto.bellotti@ba.infn.it (R.B.)

² Dipartimento di Farmacia - Scienze del Farmaco, Università di Bari, 70126 Bari, Italy

³ Dipartimento Interateneo di Fisica, Università di Bari, 70126 Bari, Italy; loredana.bellantuono@ba.infn.it (L.B.); esterpantaleo@gmail.com (E.P.)

⁴ Dipartimento di Scienze del Suolo, della Pianta e degli Alimenti, Università di Bari, 70126 Bari, Italy

* Correspondence: nicola.amoroso@uniba.it

Received: 29 October 2020; Accepted: 25 November 2020; Published: 1 December 2020



Featured Application: The automated classification of respiratory sound has gained increasing attention in recent years and has been the subject of a growing number of international scientific challenges for the development of accurate classification algorithms to support clinical practice. The COVID-19 pandemic has highlighted an urgent need for such developments. In this work, an accurate algorithm for the classification of respiratory sounds—specifically, crackles, wheezes or a combination of them—is presented.

Abstract: The COVID-19 pandemic has amplified the urgency of the developments in computer-assisted medicine and, in particular, the need for automated tools supporting the clinical diagnosis and assessment of respiratory symptoms. This need was already clear to the scientific community, which launched an international challenge in 2017 at the International Conference on Biomedical Health Informatics (ICBHI) for the implementation of accurate algorithms for the classification of respiratory sound. In this work, we present a framework for respiratory sound classification based on two different kinds of features: (i) short-term features which summarize sound properties on a time scale of tenths of a second and (ii) long-term features which assess sounds properties on a time scale of seconds. Using the publicly available dataset provided by ICBHI, we cross-validated the classification performance of a neural network model over 6895 respiratory cycles and 126 subjects. The proposed model reached an accuracy of $85\% \pm 3\%$ and an precision of $80\% \pm 8\%$, which compare well with the body of literature. The robustness of the predictions was assessed by comparison with state-of-the-art machine learning tools, such as the support vector machine, Random Forest and deep neural networks. The model presented here is therefore suitable for large-scale applications and for adoption in clinical practice. Finally, an interesting observation is that both short-term and long-term features are necessary for accurate classification, which could be the subject of future studies related to its clinical interpretation.

Keywords: respiratory sound classification; machine learning; feature engineering; multi-time-scale analysis; Random Forest; deep learning; COVID-19 remote diagnostics

1. Introduction

Respiratory diseases are the third leading cause of death worldwide, accounting for an estimated 3 million deaths each year [1], and their burden is set to increase, especially in 2020 due to the severe acute respiratory syndrome coronavirus 2 epidemic caused by coronavirus disease 2019 (COVID-19) [2].

While the most critical manifestations of COVID-19 include respiratory symptoms and pneumonia of varying severity [3], there are clinically confirmed cases of patients with respiratory symptoms whose chest computed tomography (CT) did not reveal signs of pneumonia [4]. The situation is even more intriguing when considering asymptomatic subjects as they are known to ease the diffusion of the virus but their identification without instrumental examination is extremely challenging [5].

In this context, the development of accurate diagnostic decision support systems for the early detection of respiratory diseases to monitor patient conditions and assess the severity of symptoms is of paramount importance. The role played by artificial intelligence in this field has been thoroughly explored [6–9]; the identification of useful respiratory sounds, such as crackles or wheezes, allows the detection of abnormal conditions and therefore timely diagnosis. It is worth noting that crackles can be discontinuous and therefore detectable in a limited amount of time by “short-range” features, while wheezes have a prolonged duration and therefore are more suitably characterized by “long-term” features covering an extended time range [10].

Traditionally, sound analyses are based on time–frequency characterizations, such as Fourier transforms (FT) or wavelet transforms [11–15]. More recently, other approaches have shown promising results. For example, cepstral features have been successfully adopted for lung sound classification [16]. A multi-time-scale approach has been adopted based on the principal component analysis of Fourier transforms of signals [17]. Another proposed strategy is the empirical mode decomposition method [18–20], which exploits instantaneous frequencies and points toward a local high-dimensional representation. In some sense, these approaches can be considered precursors of the most recent deep learning approaches [21–24]. However, the high dimensionality of these strategies impairs their statistical robustness; more importantly, deep learning provides models which can be difficult to clinically interpret. In this work, we propose a joint set of multi-time-scale features, where “multi-time-scale” denotes the fact that these features are intrinsically designed to capture sound properties at different time scales, specifically the presence of crackles and wheezes; these features are used to feed a supervised learning framework [25] to accurately detect sound anomalies and gain further insights about the discriminating features between healthy and pathological conditions. The proposed system represents a very promising framework also in the field of telemedicine, because it could become the operational nucleus of a remote diagnostic system that would allow users to examine the respiratory conditions of patients, identifying possible problematic situations that require urgent intervention in real time.

2. Materials and Methods

In this work, we present a novel classification framework for respiratory sounds, specifically aimed at detecting the presence of significant sounds during the respiratory cycle (see Figure 1).

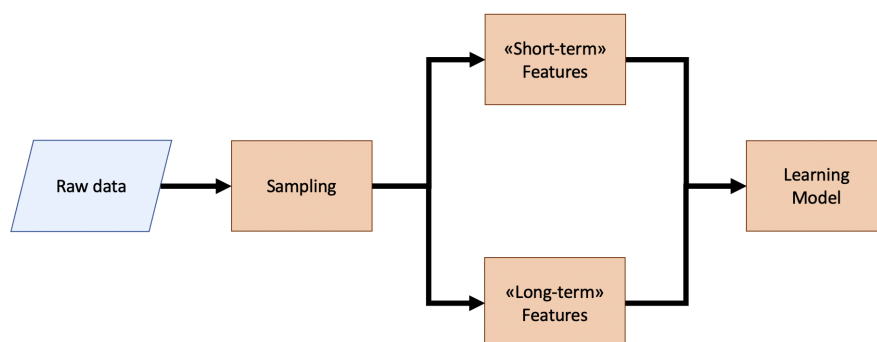


Figure 1. Flowchart of the proposed methodology. “Short-term” and “long-term” features are combined to detect significant respiratory sounds, such as crackles and wheezes.

The goal is the development of a diagnostic decision support system for the discrimination of healthy controls from patients with respiratory symptoms. The proposed approach consists of three

main steps: (i) data standardization, (ii) multi-time-scale feature extraction and (iii) classification. A detailed description of these steps is provided in the following sections.

2.1. The ICBHI Dataset

The ICBHI Scientific Challenge was launched in 2017 to provide a fair comparison of several respiratory sound classification algorithms [26]. One of the goals of the challenge was the creation of a common large and open dataset for respiratory analyses. The database was collected by two collaborating research teams from Portugal and Greece. The data collection required several years; the final dataset consists of 920 labeled audio tracks from 126 distinct participants and is currently the largest annotated, publicly available dataset. The sounds were collected from six different positions (left/right anterior, posterior and lateral) as illustrated in Figure 2.

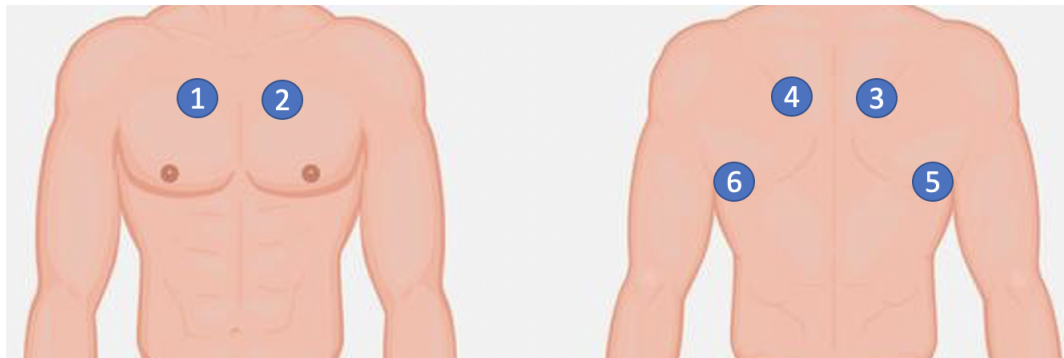


Figure 2. Locations from which respiratory sounds were collected: right anterior (1), left anterior (2), right posterior (3), left posterior (4), right lateral (5) and left lateral (6).

The available tracks have two different sampling rates: 44.1 kHz with 24 bits for sampling and 4 kHz with 16 bits for sampling. Metadata, which are not used here, also include sex, age, body-mass index for adults and height and weight for children. We segmented the audio tracks into respiratory cycles in order to increase the sample size; for each cycle, an annotation reporting the presence of significant sounds was available. Of course, we took into account this aspect during cross-validation analyses; respiratory cycles from the same patient were not split between training and validation to prevent a possible overfitting bias. The final dataset consisted of 6895 annotated respiratory cycles, including 321 cycles from healthy patients (HC) and 6574 cycles from patients with respiratory symptoms (RS).

2.2. Multi-Time-Scale Feature Extraction

We applied a multi-level feature extraction approach, where we progressed from a fine-scale representation to a global one through progressive generalizations [27]. The first level of analysis was the short-term level: we initially divided the signal into windows of 0.25 s (called short-term windows), and for each window, we computed 33 physical quantities, called short-term features. For each feature f , the short-term analysis produced a time series of values $F = (f_1, f_2, \dots, f_L)$; we extracted higher time-scale properties of the signal from the distribution of these values. The duration of the time windows on which short-term features were computed was set to 0.25 s as a compromise between two opposite needs. On the one hand, as the respiratory cycle in rest conditions lasts around two seconds for an inhalation and three seconds for an exhalation, the duration of the signal needed not to be excessive in order to effectively map the two phases of the respiratory cycle in a detailed way. Moreover, the shorter the time window, the higher the number of samples in the distribution of short-term features, and therefore the more meaningful the statistical indicators built from these distributions and used in the long-term analysis. On the other hand, a lower limit on the duration of the short-term time frames is essential to observe meaningful and informative features on the

respiratory cycle, since time windows much shorter than the typical duration of the process would provide information that would be affected by noise and difficult to interpret.

The second level of analysis was the long-term level and consisted of the calculation of 10 statistical moments associated with the time series of each of the 33 short-term features. At the end of the extraction process, each track was characterized by a total of 330 features. Figure 3 displays a pictorial representation of the feature extraction procedure.

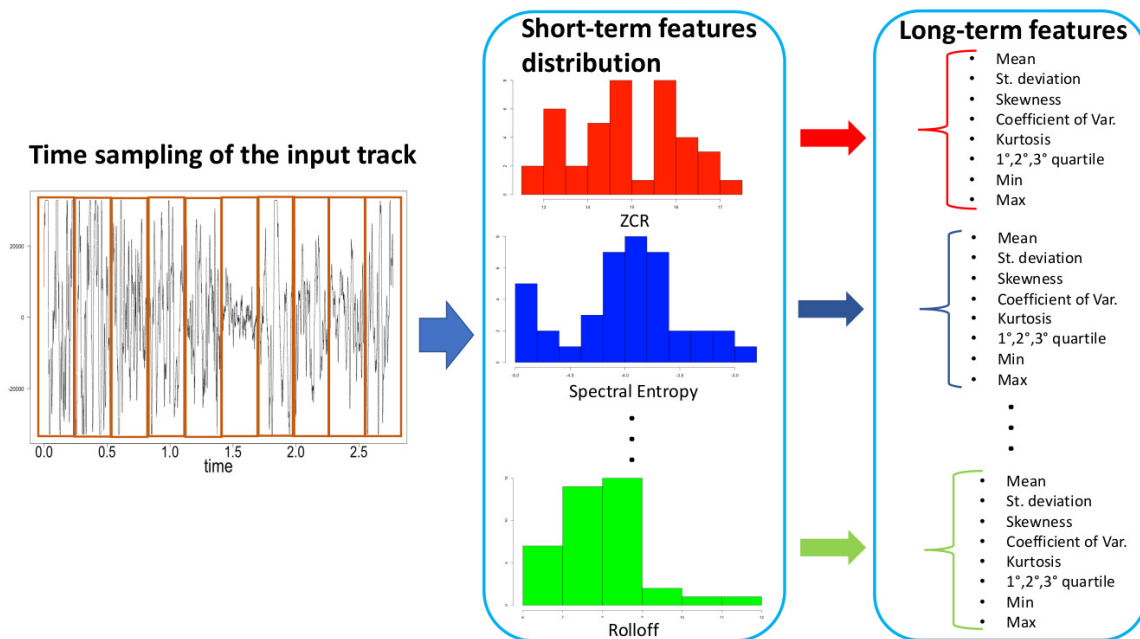


Figure 3. Pictorial representation of the feature extraction procedure. After dividing the input track into windows of 0.25 s, we estimated 33 short-term features. For each window and from the distribution of each of these 33 features, we calculated 10 statistical moments.

2.2.1. Short-Term Features

For the short-term analysis, we considered features in the time domain, features in the frequency domain, cepstral features and chroma for a total of 33 short-term features. We divided the input track into short-term windows and defined $x_i(n)$ with $n = 1, \dots, N$ being the sequence of sound intensities contained in the i -th window. In the time domain, we considered the zero crossing rate ZCR_i [28]:

$$ZCR_i = \frac{1}{2N} \sum_{n=1}^N |sgn[x_i(n)] - sgn[x_i(n - 1)]| \tag{1}$$

with

$$sgn(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases} \tag{2}$$

and the entropy of energy H_i [29]:

$$H_i = - \sum_{j=1}^K e_j \log_2 e_j \tag{3}$$

where e_j is the ratio between the energy of the j -th interval of K frames in which we divided the i -th short-term window and the energy E_i of the i -th short-term window, defined as [30]:

$$E_i = \sum_{n=1}^N |x_i(n)|^2. \tag{4}$$

In the frequency domain, we computed the spectral centroid C_i [31]:

$$C_i = \frac{\sum_{n=1}^{N/2} nX_i(n)}{\sum_{n=1}^{N/2} X_i(n)} \tag{5}$$

and its spectral spread S_i [32]:

$$S_i = \sqrt{\frac{\sum_{n=1}^{N/2} (n - C_i)^2 X_i(n)}{\sum_{n=1}^{N/2} X_i(n)}} \tag{6}$$

where $X_i(n)$, with $n = 1, \dots, N$ being the Fourier coefficients obtained by applying the discrete Fourier transform (DFT) on the i -th short-term window. In addition, we evaluated the spectral entropy SH_i [33]:

$$SH_i = - \sum_{j=1}^K n_j \log_2 n_j \tag{7}$$

with

$$n_j = \frac{E_j}{\sum_{j=1}^K E_j} \tag{8}$$

E_j ($j = 1, \dots, K$) represents the energy estimated on one of the K bins in which we divided the window. Furthermore, we computed the spectral flux Fl [34] as a measure of the spectral variation between two consecutive short-term windows $i - 1$ and i :

$$Fl(i, i - 1) = \sum_{n=1}^{N/2} (EN_i(n) - EN_{i-1}(n))^2 \tag{9}$$

where

$$EN_i(n) = \frac{X_i(n)}{\sum_{n=1}^{N/2} X_i(n)}. \tag{10}$$

The other considered features that we computed were the spectral roll-off [35], Mel-frequency cepstrum coefficients (MFCCs) [36,37] and the chroma vector [38].

The spectral roll-off R is defined as:

$$R = C \sum_{n=1}^{N/2} X_i(n) \tag{11}$$

where C is a threshold frequency below which most (typically 90%) of the spectral amplitude distribution is concentrated.

MFCCs were derived from a representation of the spectrum in which frequency bands were evenly distributed with respect to the Mel scale. Frequencies f_{Mel} in the Mel scale were related to frequencies in Hz f_{Hz} by the relation

$$f_{mel} = 1127 \cdot \ln \left(\frac{f_{Hz}}{700} + 1 \right). \quad (12)$$

MFCCs were computed through the following steps:

1. Calculate the DFT of the signal in the short-term window;
2. Identify M equally spaced frequencies on the Mel scale and build a bank of triangular spectral filters F_j with $j = 1, \dots, M$ centered on each corresponding M frequency in Hz;
3. Evaluate the spectral output powers O_j of each filter F_j ;
4. Estimate MFCCs as

$$c_m = \sum_{j=1}^M \log O_j \cos \left[m \left(j - \frac{1}{2} \right) \frac{\pi}{M} \right] \quad (13)$$

with $m = 1, \dots, M$. In the present work, we considered the first 13 MFCCs because they were deemed to contain sufficient discriminatory information in order to perform various classification tasks [27].

The chroma vector is a 12–element representation of the spectral energy and is calculated by grouping the DFT coefficients of the short-term window into 12 frequency classes related to semitone spacing. For each class q in $1, \dots, 12$, the q -th chroma element v_q is defined as the following ratio:

$$v_q = \sum_{k \in S_q} \frac{X_i(k)}{N_q} \quad (14)$$

where S_q is the subset of frequencies belonging to class q , and N_q is the number of elements in S_q . The last implemented feature is the standard deviation of the 12 components of the chroma vector.

2.2.2. Long-Term Features

From the time distributions of the 33 short-term features, we computed the following 10 statistical moments: the mean, standard deviation, coefficient of variation, skewness, kurtosis, first, second and third quartile, minimum and maximum [39].

2.3. Classification and Performance Assessment

To evaluate the informative content of the designed multi-time-scale features and thus to assess to which extent the performance depended on the feature representation or the classification models, we compared the performance of several classification methods. We used two state of the art classifiers: Random Forest (RF) [40] and the Support Vector Machine (SVM) [41]. Additionally, we explored the use of both an artificial neural network [42] and a fully-connected deep neural network [43].

2.3.1. Learning Models

RF is an ensemble of classification trees built with the bootstrapping of the training data-set. Through an iterative process during the construction of the trees, at each node, a subset of features is randomly selected which implies that the trees of the forest are weakly correlated to each other. In general, RF classifiers are easy to tune, very robust against overfitting and are particularly suitable when the number of features in the model exceeds the number of observations. In our analysis, we implemented a standard configuration in which each forest is grown with 1000 trees and $m = f/3$, with f being the number of features and m the number of features sampled to grow each leaf within a tree. An important property of Random Forest classifiers is that they can estimate the importance

of each feature during the training phase of the model. The algorithm can evaluate how much each feature decreases the impurity of a tree. In RF, the impurity decrease due to each variable is obtained from the average on all trees. Node impurity is measured by the Gini index [44].

SVM is a machine learning algorithm that employs mathematical functions, called kernels, to represent data in a new hyperspace that simplifies the representation of complicated patterns present in the data. Suppose it is desired to separate data belonging to two clusters; SVM finds the functional equation to separate the two clusters. When considering more variables, the separation line becomes a plane. By further increasing the variables, the separation becomes a hyperplane, obtained from a subset of points of the two classes, called support vectors. In general, the SVM algorithm finds a hyperplane that separates data into two classes, maximizing the separation. We implemented a default configuration with a linear kernel.

Artificial neural networks (ANNs) are computational networks inspired by the human nervous system that can learn from known examples and generalize to unknown cases. In this work, we used multilayer perceptron networks (MLPs) [45], the most commonly used ANNs, which utilize back-propagation for supervised learning. MLPs are composed of three neural levels: input, hidden and output layers. MLPs starts by feeding a features array to the input layer. The network then passes the input to the next hidden layers through connections, called dendrites; connection weights inhibit or amplify the signal, and neurons add up the input signals and transform them into output signals through an activation function. Our MLP model was composed of two hidden layers with 50 and 15 neurons, respectively, and used the sigmoid as an activation function.

The use of deep learning techniques has seen an exponential increase in the last decade; the reason for this is mainly due to the increasing availability of computing infrastructures that allow the learning of very expensive models from the computational perspective and to the increasing availability of infrastructures for data storage. Deep neural networks (DNNs) expand the architecture of ANNs: they are composed by a hierarchical architecture with many layers that constitute a non-linear information processing unit. The multiple levels of abstraction provide deep neural networks with a huge advantage in complex pattern recognition problems, adding information and analysis to each intermediate level to provide reliable output. The potential and capabilities of deep learning were unthinkable until a few years ago, even if its real advantage over ANNs, especially when the number of input cases is not very large, has been the subject of in-depth analysis and study. In this work, we used a DNN model composed of three hidden layers with 150 neurons in each layer.

2.3.2. Cross-Validation, Balancing and Performance Metrics

The examined dataset is particularly imbalanced in favor of patients; therefore, we adopted a three-fold classification framework to establish enough control in the validation sets. Moreover, it is essential that the numbers of patients and controls analyzed by a classification algorithm are balanced; otherwise, the algorithm could learn to discriminate one class well at the expense of the other classes. Among the simplest methods to balance a dataset are random over-sampling and random under-sampling strategies [46,47].

We performed the random under-sampling of the RS class during training for the machine learning models, while we performed the over-sampling of the HC class for the deep learning algorithm, which usually obtains higher performances with larger sample sizes. The balancing strategies were nested into the cross-validation procedure. Finally, respiratory cycles were randomly split into training and validation sets stratified over the subjects' identification codes; in this way, the presence of respiratory cycles from the same patient in both training and validation did not occur. We repeated the cross-validation procedure 500 times.

As a measure of performance, we evaluated accuracy, namely the rate of correct classifications, defined as follows:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (15)$$

where TP , TN , FP , FN represent true positives, true negatives, false positives and false negatives, respectively. In addition to accuracy, we also used precision:

$$Prec = \frac{TP}{TP + FP}; \quad (16)$$

classification error of the class HC:

$$EHC = \frac{FP}{TN + FP}; \quad (17)$$

and classification error of class RS:

$$ERS = \frac{FN}{TP + FN}. \quad (18)$$

All the data processing and statistical analyses were performed in Python version 3.7 (<https://www.python.org/downloads/release/python-370/>) and R version 3.6.1 (<https://www.r-project.org/>).

2.3.3. Feature Importance Procedure

To evaluate the robustness of the implemented model with respect to the used features, we applied a feature importance procedure. Initially, we estimated a feature importance ranking through the RF algorithm and the three-fold cross-validation procedure; namely, for each cross-validation cycle, we assigned a weight to each feature according to its importance evaluated by the Gini index, thus obtaining a partial ranking. We obtained an overall ranking by repeating the procedure 500 times averaging over all repetitions.

As an alternative approach, we applied a backward feature selection strategy. First, we considered a model exploiting the informative content of all the available 330 features; then, we removed the least important feature, assessed the classification performance and iterated the procedure until only four features were left. To avoid a double dipping bias, these feature selection analyses were performed within a nested cross-validation framework.

3. Results

3.1. Classification Performances

We evaluated the accuracy and precision of the HC vs. RS classification task and classification errors EHC and ERS for healthy controls and respiratory symptoms, respectively. The performances obtained by means of the implemented machine and deep learning algorithms are shown in Figure 4. An overview of the classification performances is summarized in Table 1.

Table 1. Summary table of the classification performance measures of the four implemented machine and deep learning models. Accuracy, precision and classification errors are reported with the respective standard deviations. According to a Kruskal–Wallis test, the differences within each column are statistically significant (p -value < 1%) MLP: multi-layer perceptron; SVM: support vector machine; DNN: deep neural network.

Learning Models	Accuracy	Precision	Error HC	Error RS
Random Forest	0.84 ± 0.04	0.84 ± 0.10	0.17 ± 0.10	0.16 ± 0.04
MLP	0.85 ± 0.03	0.80 ± 0.08	0.20 ± 0.08	0.14 ± 0.03
SVM	0.81 ± 0.03	0.79 ± 0.07	0.21 ± 0.07	0.19 ± 0.03
DNN	0.82 ± 0.02	0.87 ± 0.02	0.13 ± 0.02	0.23 ± 0.02

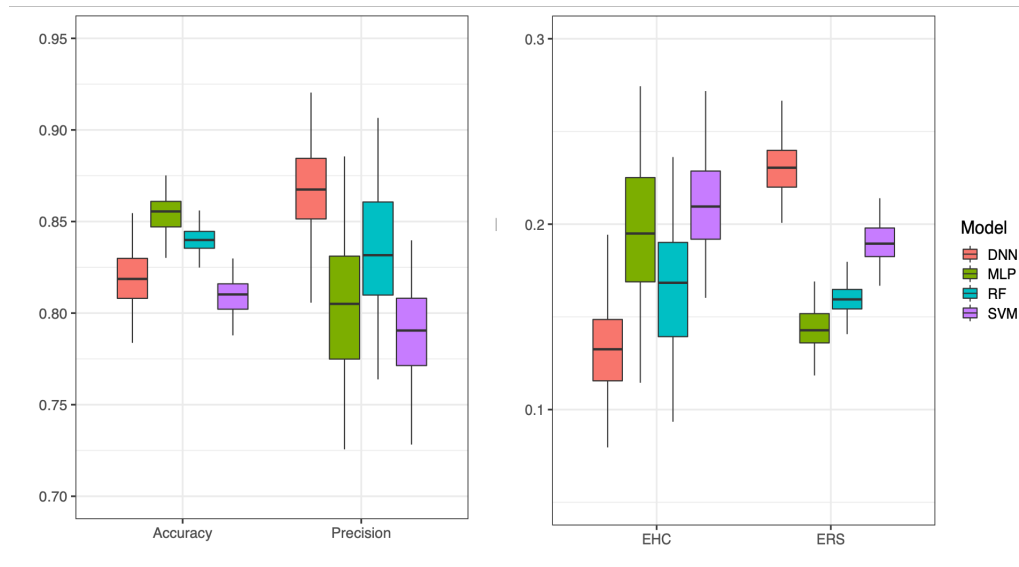


Figure 4. Box plots of healthy patients (HC) vs. patients with respiratory syndromes (RS) classification performance measures for the implemented machine and deep learning models. Distributions are obtained by means of a three-fold cross validation procedure repeated 500 times.

According to a Kruskal–Wallis test [48], the four methodologies are significantly different despite these differences being quite comparable with the inherent uncertainties for each metric. We also made pairwise comparisons of their predictions (see Figure 5 and Figure A1 in the Appendix A).

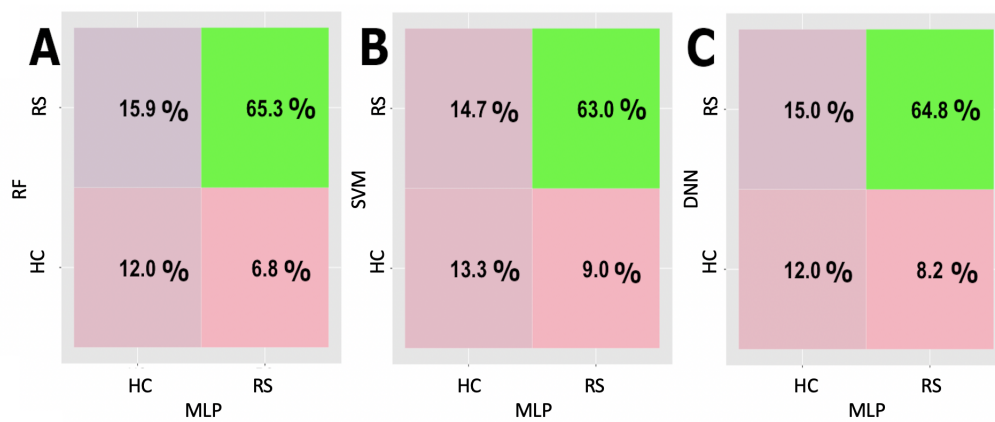


Figure 5. Contingency tables obtained by comparing MLP predictions with the output of Random Forest (RF) (panel A), SVM (panel B) and DNN (panel C) averaged over 500 rounds of 3-fold cross validation.

Overall, MLP had the best performance; thus, we only report its comparisons in the form of contingency tables. The remaining comparisons can be found in the Appendix A. It is worth noting that, in all three cases, the agreement between the classification models exceeds 76%.

3.2. Feature Importance

In the previous section, we showed that there was no significant difference among the four different classification algorithms adopted. Here, we investigate the most important features for classification and evaluate their importance. First of all, we investigated how precision and classification errors varied with the number of features used in the training (see Figure 6).

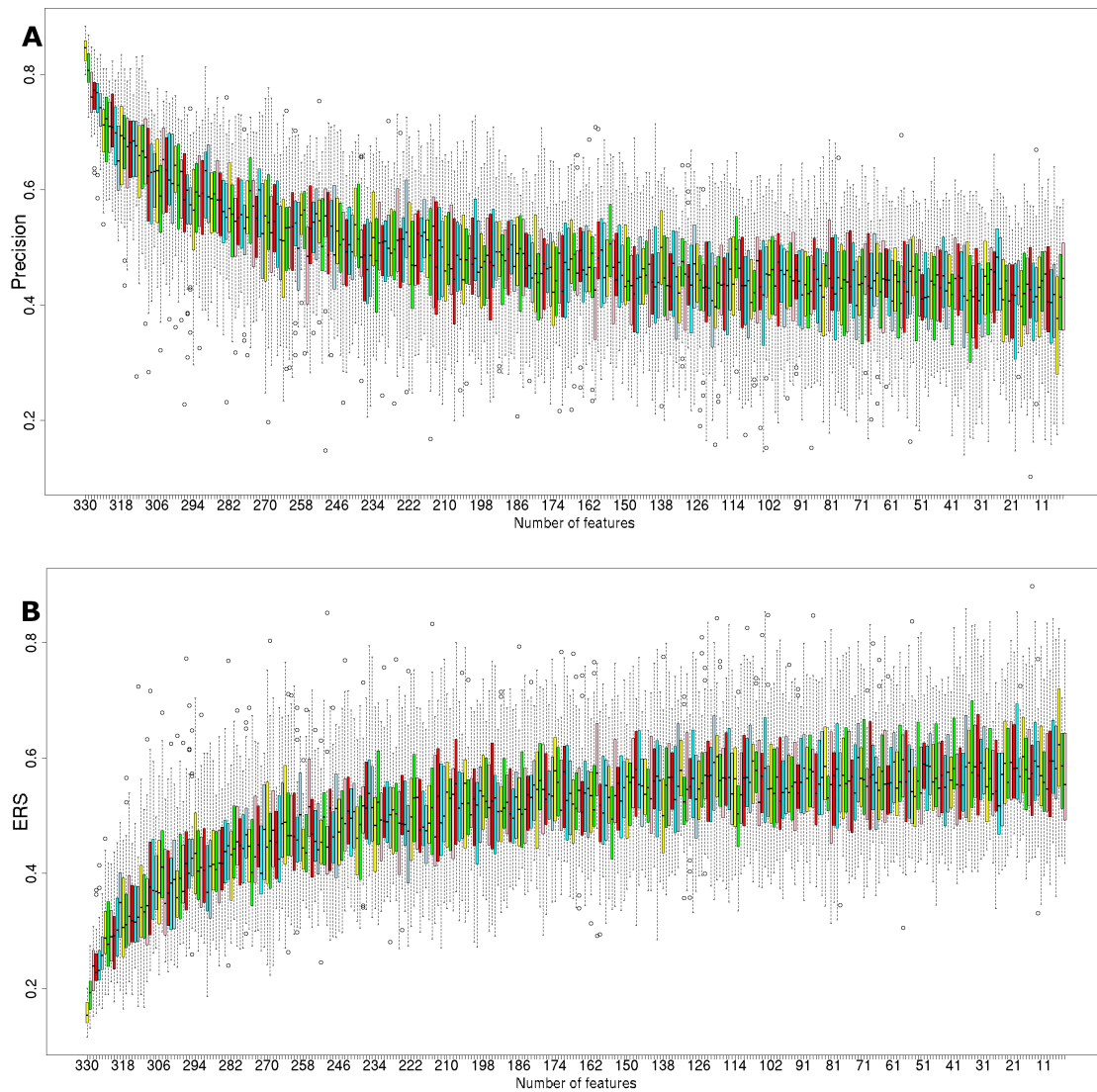


Figure 6. Precision (panel A) and RS classification error (panel B) obtained with the procedure described in Section 2.3.2. The number of features used to train the RF model ranges from 330 to 4.

By applying the procedure described in Section 2.3.3, we observed that classification performance worsens as the number of features used to train the model decreases. Moreover, we investigated feature importance in relation to the feature category (see Figure 7).

The mean decrease of the Gini index reaches a plateau at about 50 features. These 50 features were also categorized by type: 52% of the top 50 features (26 features) were the chrome vector, 14% were MFCCs and roll-off features and 10% were the ZCR and entropy.

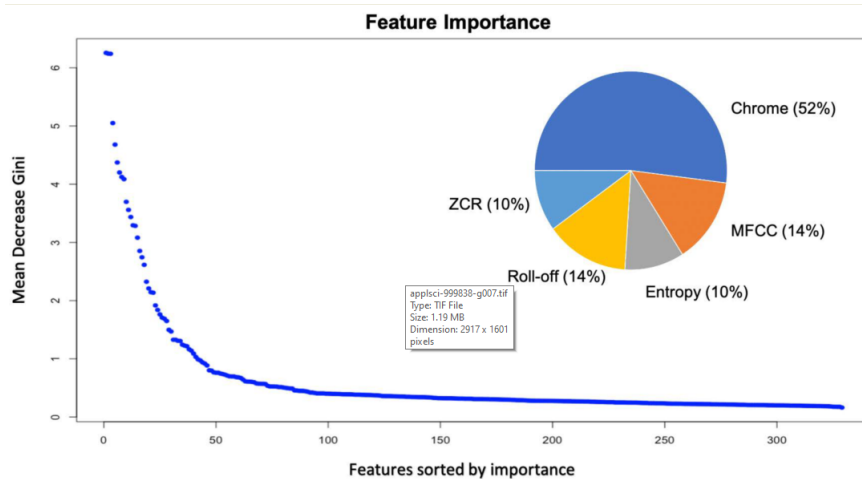


Figure 7. The feature importance in terms of the mean decrease of the Gini index and type composition of the 50 most important features.

4. Discussion

The classification of significant sounds has gained increasing importance in recent years. Several strategies have been proposed, and machine learning strategies account for a huge body of literature [49–57]. The plethora of different approaches and strategies address two distinct issues: on one hand, a robust classification for diagnostic purposes; on the other hand, “interpretability”—i.e., the design of specific features for the differential diagnosis of different pathologies. In fact, there is a consolidated consensus about the accuracy of machine learning approaches for sound analyses. Although it is difficult to compare results from different studies, state-of-the-art classification performances reach or exceed 80% accuracy, which compares well with the results obtained by our framework (81% ~ 85%). The use of different data or the adoption of specific study designs such as different cross-validation strategies (if any) makes comparison difficult in general. Table 2 presents an overview of recent peer-reviewed published studies using ICBHI data. These studies only occasionally deal with the classification of significant sounds, and their performance measures cannot be directly compared.

Table 2. Classification performances reported by other studies on the ICBHI data.

Study	Performance	Sensitivity	Specificity
Minami et al. [23]	0.81 (harmonic score)	0.54	0.41
Wu et al. [58]	0.88 (accuracy)	-	0.91
Ma et al. [59]	0.52 (performance score)	-	-
Yang et al. [60]	0.55 (average score)	-	-

The need for an objective comparison between different classification algorithms has led in recent years to the spread of international challenges, especially for machine learning applications, a common trait of which was the use of a shared framework for all the participants: a unique dataset for training and a blind test set [61–65]. The data investigated in this work were collected on the occasion of the previously mentioned international ICBHI challenge. Of the 18 different algorithms submitted to the challenge, only two reached the final stage and were presented at the ICBHI 2017. The first was an approach exploiting resonance-based decomposition [66]; the second was a method based on the application of hidden Markov models in combination with Gaussian mixture models [67]. These algorithms were evaluated on the basis of accuracy to detect wheezes, crackles or their simultaneous presence, thus resulting in a four-class classification problem; the test accuracy reported for both algorithms did not exceed 50%, which is far below the performances reported in the literature. A possible explanation for

this would be that, despite the large sample size, the collected data included some examples which were extremely difficult to classify—a possibility also mentioned by the organizers of the challenge.

The information content provided by the proposed features is encouraging not only for the high level of accuracy obtained but also for its robustness, which we evaluated by comparing several supervised learning frameworks. In fact, we observed that, for all pairwise comparisons, the agreement between the classification models exceeded 76%. It is well-known that, for each classification task, it is not possible to determine a priori which is the best classifier, and the impact of a specific classifier on classification performance can be substantial [68]. Nevertheless, we observed that the performances of four different approaches (RF, SVM, MLP and a DNN) differ by a few percentage points. The best-performing methods were the MLP in terms of accuracy and the DNN in terms of precision.

In recent years, deep learning strategies have experienced an exponential growth. Among the many available strategies that might be worth exploring in this setting, two deserve a mention: residual networks (ResNet) and long short-term memory (LSTM) networks [69–72]. These approaches should be considered especially for further studies that are aimed more at the deep learning domain.

It is also well known that, in general, learning algorithms require a sufficient number of training examples for each class and that unbalanced classes can weaken the learning process [73,74]; besides, the strength of DNN is related to the available sample size—the more the better. Accordingly, we adopted a three-fold cross-validation and different sampling strategies to allow the best operational conditions for each classifier. The use of three-fold instead of a more common five-fold or 10-fold cross-validation was motivated by the exiguous number of HC cases and ensured that the test set contained a representative number of HC examples. As concerns the sample strategies, we used under-sampling with standard machine learning algorithms and over-sampling for DNN. In fact, we observed that the use of under-sampling with DNN resulted in a significant performance deterioration.

Finally, we investigated which features were best at characterizing the presence of significant sounds. We observed that a relatively small amount of features (~50) was sufficient for an accurate classification. Besides, our findings demonstrated that, in the examined case variations, roughly this number of features results in negligible performance differences (see Figure 6); this is a relevant aspect, considering that using different feature importance thresholds can significantly affect the classification performance. Finally, by grouping these top rank features by type, we observed that the main contribution was given by the chroma vector. The chroma vector is a 12-dimensional representation of the spectral energy [75]. In general, this descriptor is suitable for music-speech applications when the signal is heavily affected by noise [27,76]; our findings would suggest that they can also be effectively used in the context of significant sound recognition.

5. Conclusions

In this work, we presented a multi-time-scale machine learning framework for the classification of respiratory sounds including crackles and wheezes. The proposed framework can accurately distinguish healthy controls from patients whose respiratory cycles present some significant sounds. Besides, we observed that the informative power of the proposed features is only slightly affected by the classifier choice; with four different classifiers (RF, SVM, MLP and DNN), we obtained accuracy values ranging from 85% for MLP to 81% for SVM. The best performing features among the 330 adopted were the chroma vector components. In this work, we addressed the binary classification problem for HC versus RS; in fact, we ran our analyses at the “patient” level. Future studies could address the recognition of significant sounds at the respiratory cycle level; of course, this problem poses some major difficulties, because it is a multi-class classification task. Our analysis presents the typical limitations of a feature-based learning approach. In fact, in a feature engineering process, a priori hypotheses are made that might imply that significant aspects of the signal are neglected. More general deep learning approaches such as LSTM and ResNet might be able to improve the classification

performance, although this would require further investigation. Nevertheless, the results presented here are promising and deserve further investigation.

Author Contributions: Conceptualization, N.A. and A.M.; methodology, N.A., A.M. and L.B.; software, A.M. and L.B.; formal analysis, A.M.; writing—original draft preparation, N.A. and A.M.; writing—review and editing, all the authors; visualization, N.A. and A.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors would like to acknowledge the IT resources made available by ReCaS, a project funded by the MIUR (Italian Ministry for Education, University and Research) in the “PON Ricerca e Competitività 2007–2013-Azione I-Interventi di rafforzamento strutturale” PONa3_00052, Avviso 254/Ric, University of Bari.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

In the manuscript, we focused on MLP because it was the best performing method and showed how it performed similarly to RF, SVM and DNN. Here, we show the remaining comparisons.

Even in this case, the agreement between the models is around 76%.

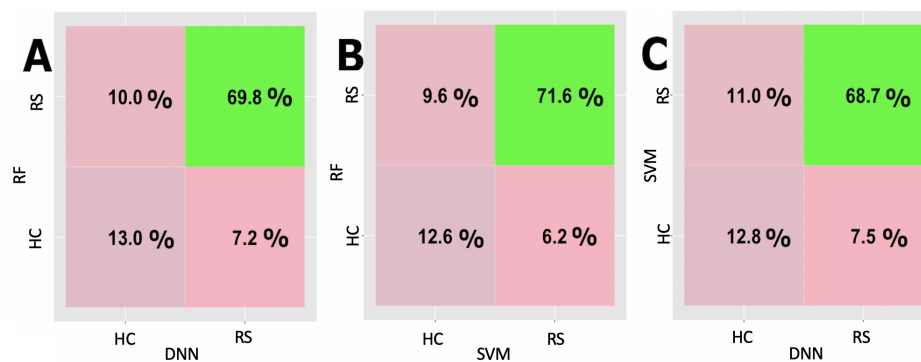


Figure A1. Contingency tables comparing RF and DNN predictions (panel A), RF and SVM predictions (panel B) and DNN and SVM predictions (panel C) averaged over 500 rounds of 3-fold cross validation.

References

- European Respiratory Society. The Global Impact of Respiratory Disease. In *Forum of International Respiratory Societies*, 2nd ed.; European Respiratory Society: Sheffield, UK, 2017.
- Williams, S.; Sheikh, A.; Campbell, H.; Fitch, N.; Griffiths, C.; Heyderman, R.S.; Jordan, R.E.; Katikireddi, S.V.; Tsiligianni, I.; Obasi, A. Respiratory research funding is inadequate, inequitable, and a missed opportunity. *Lancet Respir. Med.* **2020**, *8*, e67–e68. [[CrossRef](#)]
- Lai, C.C.; Liu, Y.H.; Wang, C.Y.; Wang, Y.H.; Hsueh, S.C.; Yen, M.Y.; Ko, W.C.; Hsueh, P.R. Asymptomatic carrier state, acute respiratory disease, and pneumonia due to severe acute respiratory syndrome coronavirus 2 (SARSCoV-2): Facts and myths. *J. Microbiol. Immunol. Infect.* **2020**, *53*, 404–412. [[CrossRef](#)] [[PubMed](#)]
- Guan, W.J.; Ni, Z.Y.; Hu, Y.; Liang, W.H.; Ou, C.Q.; He, J.X.; Liu, L.; Shan, H.; Lei, C.L.; Hui, D.S.; et al. Clinical characteristics of 2019 novel coronavirus infection in China. *medRxiv* **2020**. [[CrossRef](#)]
- Bai, Y.; Yao, L.; Wei, T.; Tian, F.; Jin, D.Y.; Chen, L.; Wang, M. Presumed asymptomatic carrier transmission of COVID-19. *JAMA* **2020**, *323*, 1406–1407. [[CrossRef](#)] [[PubMed](#)]
- Reichert, S.; Gass, R.; Brandt, C.; Andrès, E. Analysis of respiratory sounds: State of the art. *Clin. Med. Circ. Respir. Pulm. Med.* **2008**, *2*, CCRPM-S530. [[CrossRef](#)] [[PubMed](#)]
- Palaniappan, R.; Sundaraj, K.; Sundaraj, S. Artificial intelligence techniques used in respiratory sound analysis—a systematic review. *Biomed. Eng./Biomed. Tech.* **2014**, *59*, 7–18. [[CrossRef](#)]
- Pramono, R.X.A.; Bowyer, S.; Rodriguez-Villegas, E. Automatic adventitious respiratory sound analysis: A systematic review. *PLoS ONE* **2017**, *12*, e0177926. [[CrossRef](#)]
- Mekov, E.; Miravittles, M.; Petkov, R. Artificial intelligence and machine learning in respiratory medicine. *Expert Rev. Respir. Med.* **2020**, *14*, 559–564. [[CrossRef](#)]

10. Sovijarvi, A. Characteristics of breath sounds and adventitious respiratory sounds. *Eur. Respir. Rev.* **2000**, *10*, 591–596.
11. Homs-Corbera, A.; Fiz, J.A.; Morera, J.; Jané, R. Time-frequency detection and analysis of wheezes during forced exhalation. *IEEE Trans. Biomed. Eng.* **2004**, *51*, 182–186. [[CrossRef](#)]
12. Kandaswamy, A.; Kumar, C.S.; Ramanathan, R.P.; Jayaraman, S.; Malmurugan, N. Neural classification of lung sounds using wavelet coefficients. *Comput. Biol. Med.* **2004**, *34*, 523–537. [[CrossRef](#)]
13. Cnockaert, L.; Migeotte, P.F.; Daubigny, L.; Prisk, G.K.; Grenez, F.; Sá, R.C. A method for the analysis of respiratory sinus arrhythmia using continuous wavelet transforms. *IEEE Trans. Biomed. Eng.* **2008**, *55*, 1640–1642. [[CrossRef](#)] [[PubMed](#)]
14. Sello, S.; Strambi, S.k.; De Michele, G.; Ambrosino, N. Respiratory sound analysis in healthy and pathological subjects: A wavelet approach. *Biomed. Signal Process. Control.* **2008**, *3*, 181–191. [[CrossRef](#)]
15. Jin, F.; Krishnan, S.; Sattar, F. Adventitious sounds identification and extraction using temporal–spectral dominance-based features. *IEEE Trans. Biomed. Eng.* **2011**, *58*, 3078–3087. [[PubMed](#)]
16. Sengupta, N.; Sahidullah, M.; Saha, G. Lung sound classification using cepstral-based statistical features. *Comput. Biol. Med.* **2016**, *75*, 118–129. [[CrossRef](#)] [[PubMed](#)]
17. Xie, S.; Jin, F.; Krishnan, S.; Sattar, F. Signal feature extraction by multi-scale PCA and its application to respiratory sound classification. *Med. Biol. Eng. Comput.* **2012**, *50*, 759–768. [[CrossRef](#)]
18. Charleston-Villalobos, S.; González-Camarena, R.; Chi-Lem, G.; Aljama-Corrales, T. Crackle sounds analysis by empirical mode decomposition. *IEEE Eng. Med. Biol. Mag.* **2007**, *26*, 40. [[CrossRef](#)]
19. Lozano, M.; Fiz, J.A.; Jané, R. Estimation of instantaneous frequency from empirical mode decomposition on respiratory sounds analysis. In Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 3–7 July 2013; pp. 981–984.
20. Lozano, M.; Fiz, J.A.; Jané, R. Automatic differentiation of normal and continuous adventitious respiratory sounds using ensemble empirical mode decomposition and instantaneous frequency. *IEEE J. Biomed. Health Inform.* **2015**, *20*, 486–497. [[CrossRef](#)]
21. Perna, D. Convolutional neural networks learning from respiratory data. In Proceedings of the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Madrid, Spain, 3–6 December 2018; pp. 2109–2113.
22. Liu, R.; Cai, S.; Zhang, K.; Hu, N. Detection of Adventitious Respiratory Sounds based on Convolutional Neural Network. In Proceedings of the 2019 IEEE International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), Shanghai, China, 21–24 November 2019; pp. 298–303.
23. Minami, K.; Lu, H.; Kim, H.; Mabu, S.; Hirano, Y.; Kido, S. Automatic classification of large-scale respiratory sound dataset based on convolutional neural network. In Proceedings of the 2019 IEEE 19th International Conference on Control, Automation and Systems (ICCAS), Jeju, Korea, 15–18 October 2019; pp. 804–807.
24. Acharya, J.; Basu, A. Deep Neural Network for Respiratory Sound Classification in Wearable Devices Enabled by Patient Specific Model Tuning. *IEEE Trans. Biomed. Circuits Syst.* **2020**, *14*, 535–544. [[CrossRef](#)]
25. Duda, R.O.; Hart, P.E.; Stork, D.G. *Pattern Classification and Scene Analysis*; Wiley: New York, NY, USA, 1973; Volume 3.
26. Rocha, B.M.; Filos, D.; Mendes, L.; Serbes, G.; Ulukaya, S.; Kahya, Y.P.; Jakovljevic, N.; Turukalo, T.L.; Vogiatzis, I.M.; Perantoni, E.; et al. An open access database for the evaluation of respiratory sound classification algorithms. *Physiol. Meas.* **2019**, *40*, 035001. [[CrossRef](#)]
27. Pikrakis, A.; Giannakopoulos, T.; Theodoridis, S. A speech/music discriminator of radio recordings based on dynamic programming and bayesian networks. *IEEE Trans. Multimed.* **2008**, *10*, 846–857. [[CrossRef](#)]
28. Bachu, R.; Kopparthi, S.; Adapa, B.; Barkana, B. Separation of voiced and unvoiced using zero crossing rate and energy of the speech signal. In Proceedings of the American Society for Engineering Education (ASEE) Zone Conference Proceedings, Pittsburgh, PA, USA, 22–25 June 2008; pp. 1–7.
29. Rizal, A.; Hidayat, R.; Nugroho, H.A. Entropy measurement as features extraction in automatic lung sound classification. In Proceedings of the 2017 IEEE International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC), Yogyakarta, Indonesia, 26–28 September 2017; pp. 93–97.
30. Crocker, M.J. *Handbook of Acoustics*; John Wiley & Sons: Hoboken, NJ, USA, 1998.
31. Schubert, E.; Wolfe, J.; Tarnopolsky, A. Spectral centroid and timbre in complex, multiple instrumental textures. In Proceedings of the International Conference on Music Perception and Cognition, Evanston, IL, USA, 3–7 August 2004; pp. 112–116.

32. Lazaro, A.; Sarno, R.; Andre, R.J.; Mahardika, M.N. Music tempo classification using audio spectrum centroid, audio spectrum flatness, and audio spectrum spread based on MPEG-7 audio features. In Proceedings of the 2017 IEEE 3rd International Conference on Science in Information Technology (ICSITech), Bandung, Indonesia, 25–26 October 2017; pp. 41–46.
33. Misra, H.; Ikbali, S.; Bourlard, H.; Hermansky, H. Spectral entropy based feature for robust ASR. In Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, QC, Canada, 17–21 May 2004; Volume 1, p. I-193.
34. Sadjadi, S.O.; Hansen, J.H. Unsupervised speech activity detection using voicing measures and perceptual spectral flux. *IEEE Signal Process. Lett.* **2013**, *20*, 197–200. [[CrossRef](#)]
35. Kos, M.; Kačič, Z.; Vlaj, D. Acoustic classification and segmentation using modified spectral roll-off and variance-based features. *Digit. Signal Process.* **2013**, *23*, 659–674. [[CrossRef](#)]
36. Logan, B. Mel frequency cepstral coefficients for music modeling. *ISMIR* **2000**, *270*, 1–11.
37. Molau, S.; Pitz, M.; Schluter, R.; Ney, H. Computing mel-frequency cepstral coefficients on the power spectrum. In Proceedings of the 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 01CH37221), Salt Lake City, UT, USA, 7–11 May 2001; Volume 1, pp. 73–76.
38. Müller, M.; Kurth, F.; Clausen, M. Audio Matching via Chroma-Based Statistical Features. *ISMIR* **2005**, *2005*, 6.
39. Heng, R.; Nor, M.J.M. Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition. *Appl. Acoust.* **1998**, *53*, 211–226. [[CrossRef](#)]
40. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
41. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
42. Yegnanarayana, B. *Artificial Neural Networks*; PHI Learning Pvt. Ltd.: New Delhi, India, 2009.
43. Goodfellow, I.; Bengio, Y.; Courville, A.; Bengio, Y. *Deep Learning*; MIT Press: Cambridge, UK, 2016; Volume 1.
44. Han, H.; Guo, X.; Yu, H. Variable selection using mean decrease accuracy and mean decrease gini based on random forest. In Proceedings of the 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 26–28 August 2016; pp. 219–224.
45. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: New York, NY, USA, 2009.
46. Han, H.; Wang, W.Y.; Mao, B.H. Borderline-SMOTE: A new over-sampling method in imbalanced data sets learning. In *International Conference on Intelligent Computing*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 878–887.
47. Maglietta, R.; Amoroso, N.; Boccardi, M.; Bruno, S.; Chincarini, A.; Frisoni, G.B.; Inglese, P.; Redolfi, A.; Tangaro, S.; Tateo, A.; et al. Automated hippocampal segmentation in 3D MRI using random undersampling with boosting algorithm. *Pattern Anal. Appl.* **2016**, *19*, 579–591. [[CrossRef](#)]
48. Kruskal, W.H.; Wallis, W.A. Use of Ranks in One-Criterion Variance Analysis. *J. Am. Stat. Assoc.* **1952**, *47*, 583–621. [[CrossRef](#)]
49. Pesu, L.; Ademovic, E.; Pesquet, J.C.; Helisto, P. Wavelet packet based respiratory sound classification. In Proceedings of the IEEE Third International Symposium on Time-Frequency and Time-Scale Analysis (TFTS-96), Paris, France, 18–21 June 1996; pp. 377–380.
50. Güler, E.Ç.; Sankur, B.; Kahya, Y.P.; Raudys, S. Two-stage classification of respiratory sound patterns. *Comput. Biol. Med.* **2005**, *35*, 67–83. [[CrossRef](#)]
51. Riella, R.; Nohama, P.; Maia, J. Method for automatic detection of wheezing in lung sounds. *Braz. J. Med. Biol. Res.* **2009**, *42*, 674–684. [[CrossRef](#)] [[PubMed](#)]
52. Mayorga, P.; Druzgalski, C.; Morelos, R.; Gonzalez, O.; Vidales, J. Acoustics based assessment of respiratory diseases using GMM classification. In Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina, 31 August–4 September 2010; pp. 6312–6316.
53. Emmanouilidou, D.; Patil, K.; West, J.; Elhilali, M. A multiresolution analysis for detection of abnormal lung sounds. In Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012; pp. 3139–3142.
54. Palaniappan, R.; Sundaraj, K. Respiratory sound classification using cepstral features and support vector machine. In Proceedings of the 2013 IEEE Recent Advances in Intelligent Computational Systems (RAICS), Trivandrum, India, 19–21 December 2013; pp. 132–136.

55. Sen, I.; Saraclar, M.; Kahya, Y.P. A comparison of SVM and GMM-based classifier configurations for diagnostic classification of pulmonary sounds. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 1768–1776. [[CrossRef](#)] [[PubMed](#)]
56. Chambres, G.; Hanna, P.; Desainte-Catherine, M. Automatic detection of patient with respiratory diseases using lung sound analysis. In Proceedings of the 2018 IEEE International Conference on Content-Based Multimedia Indexing (CBMI), La Rochelle, France, 4–6 September 2018; pp. 1–6.
57. Yadav, A.; Dutta, M.K.; Prinosil, J. Machine Learning Based Automatic Classification of Respiratory Signals using Wavelet Transform. In Proceedings of the 2020 IEEE 43rd International Conference on Telecommunications and Signal Processing (TSP), Milan, Italy, 7–9 July 2020; pp. 545–549.
58. Wu, L.; Li, L. Investigating into segmentation methods for diagnosis of respiratory diseases using adventitious respiratory sounds. In Proceedings of the 2020 IEEE 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020; pp. 768–771.
59. Ma, Y.; Xu, X.; Li, Y. LungRN+ NL: An Improved Adventitious Lung Sound Classification Using non-local block ResNet Neural Network with Mixup Data Augmentation. In Proceedings of the Interspeech 2020, Shanghai, China, 25–29 October 2020; pp. 2902–2906.
60. Yang, Z.; Liu, S.; Song, M.; Parada-Cabaleiro, E.; Schuller12, B.W. Adventitious Respiratory Classification using Attentive Residual Neural Networks. In Proceedings of the Interspeech 2020, Shanghai, China, 25–29 October 2020; pp. 2912–2916.
61. Pelletier, B.; Hickey, G.M.; Bothi, K.L.; Mude, A. Linking rural livelihood resilience and food security: An international challenge. *Food Secur.* **2016**, *8*, 469–476. [[CrossRef](#)]
62. Kristan, M.; Leonardis, A.; Matas, J.; Felsberg, M.; Pflugfelder, R.; Cehovin Zajc, L.; Vojir, T.; Hager, G.; Lukežić, A.; Eldesokey, A.; et al. The visual object tracking vot2017 challenge results. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Venice, Italy, 22–29 October 2017; pp. 1949–1972.
63. Amoroso, N.; Diacono, D.; Fanizzi, A.; La Rocca, M.; Monaco, A.; Lombardi, A.; Guaragnella, C.; Bellotti, R.; Tangaro, S.; Initiative, A.D.N. Deep learning reveals Alzheimer’s disease onset in MCI subjects: results from an international challenge. *J. Neurosci. Methods* **2018**, *302*, 3–9. [[CrossRef](#)]
64. Choobdar, S.; Ahsen, M.E.; Crawford, J.; Tomasoni, M.; Fang, T.; Lamparter, D.; Lin, J.; Hescott, B.; Hu, X.; Mercer, J.; et al. Assessment of network module identification across complex diseases. *Nat. Methods* **2019**, *16*, 843–852. [[CrossRef](#)] [[PubMed](#)]
65. Honma, M.; Kitazawa, A.; Cayley, A.; Williams, R.V.; Barber, C.; Hanser, T.; Saiakhov, R.; Chakravarti, S.; Myatt, G.J.; Cross, K.P.; et al. Improvement of quantitative structure–activity relationship (QSAR) tools for predicting Ames mutagenicity: Outcomes of the Ames/QSAR International Challenge Project. *Mutagenesis* **2019**, *34*, 3–16. [[CrossRef](#)] [[PubMed](#)]
66. Selesnick, I.W. Wavelet transform with tunable Q-factor. *IEEE Trans. Signal Process.* **2011**, *59*, 3560–3575. [[CrossRef](#)]
67. Jakovljević, N.; Lončar-Turukalo, T. Hidden markov model based respiratory sound classification. In *International Conference on Biomedical and Health Informatics*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 39–43.
68. Ho, Y.C.; Pepyne, D.L. Simple explanation of the no-free-lunch theorem and its implications. *J. Opt. Theory Appl.* **2002**, *115*, 549–570. [[CrossRef](#)]
69. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
70. Ma, Y.; Xu, X.; Yu, Q.; Zhang, Y.; Li, Y.; Zhao, J.; Wang, G. LungBRN: A Smart Digital Stethoscope for Detecting Respiratory Disease Using bi-ResNet Deep Learning Algorithm. In Proceedings of the 2019 IEEE Biomedical Circuits and Systems Conference (BioCAS), Nara, Japan, 17–19 October 2019; pp. 1–4.
71. Székely, É.; Henter, G.E.; Gustafson, J. Casting to corpus: Segmenting and selecting spontaneous dialogue for TTS with a CNN-LSTM speaker-dependent breath detecto. In Proceedings of the ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; pp. 6925–6929.

72. Ardakani, A.A.; Kanafi, A.R.; Acharya, U.R.; Khadem, N.; Mohammadi, A. Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks. *Comput. Biol. Med.* **2020**, *121*, 103795. [[CrossRef](#)]
73. Mazurowski, M.A.; Habas, P.A.; Zurada, J.M.; Lo, J.Y.; Baker, J.A.; Tourassi, G.D. Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. *Neural Netw.* **2008**, *21*, 427–436. [[CrossRef](#)] [[PubMed](#)]
74. Cilli, R.; Monaco, A.; Amoroso, N.; Tateo, A.; Tangaro, S.; Bellotti, R. Machine learning for cloud detection of globally distributed Sentinel-2 images. *Remote Sens.* **2020**, *12*, 2355. [[CrossRef](#)]
75. Pirkakis, A.; Giannakopoulos, T.; Theodoridis, S. *A Computationally Efficient Speech/Music Discriminator for Radio Recordings*; ISMIR: Victoria, BC, Canada, 2006; pp. 107–110.
76. Hirsch, H.G.; Pearce, D. The Aurora experimental framework for the performance evaluation of speech recognition systems under noisy conditions. In Proceedings of the ASR2000-Automatic Speech Recognition: Challenges for the New Millenium ISCA Tutorial and Research Workshop (ITRW), Paris, France, 18–20 September 2000.

Sample Availability: Data used in this work are open access.

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).