Article

Automatic Robust Crackle Detection and Localization Approach Using AR-Based Spectral Estimation and Support Vector Machine

Loredana Daria Mang [1,*], Julio José Carabias-Orti [1], Francisco Jesús Canadas-Quesada [1], Juan de la Torre-Cruz [1], Antonio Muñoz-Montoro [1], Pablo Revuelta-Sanz [2] and Eilas Fernandez Combarro [2]

[1] Department of Telecommunication Engineering, University of Jaen, Campus Cientifico-Tecnologico de Linares, Avda. de la Universidad, s/n, 23700 Linares, Jaén, Spain; carabias@ujaen.es (J.J.C.-O.); fcanadas@ujaen.es (F.J.C.-O.); jtorre@ujaen.es (J.l.d.T.-C.); jmontoro@ujaen.es (A.M.-M.)

[2] Department of Computer Science, University of Oviedo, Campus de Gijón s/n, 33204 Gijón, Spain; revueltapablo@uniovi.es (P.R.-S.); efernandezca@uniovi.es (E.F.C.)

* Correspondence: lmang@ujaen.es

Abstract: Auscultation primarily relies upon the acoustic expertise of individual doctors in identifying, through the use of a stethoscope, the presence of abnormal sounds such as crackles because the recognition of these sound patterns has critical importance in the context of early detection and diagnosis of respiratory pathologies. In this paper, we propose a novel method combining autoregressive (AR)-based spectral features and a support vector machine (SVM) classifier to detect the presence of crackle events and their temporal location within the input signal. A preprocessing stage is performed to discard information out of the band of interest and define the segments for short-time signal analysis. The AR parameters are estimated for each segment to be classified by means of support vector machine (SVM) classifier into crackles and normal lung sounds using a set of synthetic crackle waveforms that have been modeled to train the classifier. A dataset composed of simulated and real coarse and fine crackles sound signals was created with several signal-to-noise (SNR) ratios to evaluate the robustness of the proposed method. Each simulated and real signal was mixed with noise that shows the same spectral energy distribution as typically found in breath noise from a healthy subject. This study makes a significant contribution by achieving competitive results. The proposed method yields values ranging from 80% in the lowest signal-to-noise ratio scenario to a perfect 100% in the highest signal-to-noise ratio scenario. Notably, these results surpass those of other methods presented by a margin of at least 15%. The combination of an autoregressive (AR) model with a support vector machine (SVM) classifier offers an effective solution for detecting the presented events. This approach exhibits enhanced robustness against variations in the signal-to-noise ratio that the input signals may encounter.

Keywords: respiratory; crackle; detection; AR model; SVM; accuracy; sensitivity; precision

1. Introduction

An increasing amount of economic and technological resources are being invested in accelerating the early detection of respiratory diseases (for example, pneumonia) as it is essential to providing effective medical treatment that minimizes risks to the health of the patient and optimizes healthcare costs at the same time. According to the World Health Organization (WHO), pneumonia is the leading cause of death in children, with one million deaths in children under 5 years of age in 2017, which is equivalent to 15% of all deaths in this population worldwide [1]. Although there are more sophisticated clinical tests (for example, blood tests and chest X-rays) to diagnose most lung diseases than the auscultation process [2], auscultation is the most widely used tool in health centers in low-income countries and even in most rural areas of high-income countries when a

first respiratory system examination is performed mainly due to its low cost and non-invasiveness. However, auscultation depends largely on the acoustic training of each pulmonologist to recognize, through the stethoscope, the presence of abnormal sounds that are often associated with various respiratory pathologies, as occurs between crackles and specific lung diseases such as congestive heart failure and pneumonia [3–5]. In addition, it is important to note that crackles may not be heard by the doctor due to their short duration, low rate of occurrence, and often low amplitude, which makes them difficult to be detected by the human auditory system [6].

Respiratory sounds can be categorized into normal and abnormal sounds according to the COmputerized Respiratory Sound Analysis (CORSA) guidelines [7]. Normal respiratory sounds (RSs) heard in healthy lungs are represented by a broadband spectrum locating most of the energy between 60–1000 Hz [8]. One of these abnormal or adventitious respiratory sounds are called crackles sounds (CSs), which are superimposed onto the RS, generated by unhealthy lungs as a consequence of pulmonary disorders. Crackles can be defined as discontinuous, short, explosive, and non-musical sounds with a spectral energy located between 100 Hz and 2000 Hz [9] as shown in Figure 1. Specifically, crackles can be classified as either coarse or fine. Coarse crackle sounds (CCs) are characterized by a temporal duration $\leq$20 ms and can be heard in the early inspiratory and the expiratory phases, showing a low pitch located around 350 Hz. One of the causes of these is the presence of air bubbles in large bronchi [10]. Fine crackle sounds (FCSs) are characterized by a temporal duration of $\leq$5 ms, active in the late inspiratory phase with a high pitch located around 650 Hz. The cause of FCSs is the abrupt opening of small airways [10]. In the literature [11–13], it is common characterize both CCSs and FCSs using temporal waveforms by means of two parameters: (i) t_{IDW} (initial deflection width), which indicates the temporal duration between the beginning of the crackle and the first deflection; and (ii) t_{2CD} (two-cycle deflection), which represents the time duration from the beginning of the crackle to the time point where two cycles are completed. Specifically, (i) FCSs show on average $t_{IDW} = 0.5$ ms and $t_{2CD} = 3.3$ ms whereas CCSs are represented on average by $t_{IDW} = 1.0$ ms and $t_{2CD} = 5.1$ ms according to Hoevers et al. [11]; (ii) FCSs show on average $t_{IDW} = 0.9$ ms and $t_{2CD} = 6.0$ ms whereas CCSs are represented on average by $t_{IDW} = 1.25$ ms and $t_{2CD} = 9.50$ ms according to Cohen et al. [12]; (iii) FCSs show on average $t_{IDW} = 0.7$ ms and $t_{2CD} = 5$ ms whereas CCSs are represented on average by $t_{IDW} = 1.5$ ms and $t_{2CD} = 10$ ms according to the American Thoracic Society (ATS) [13].
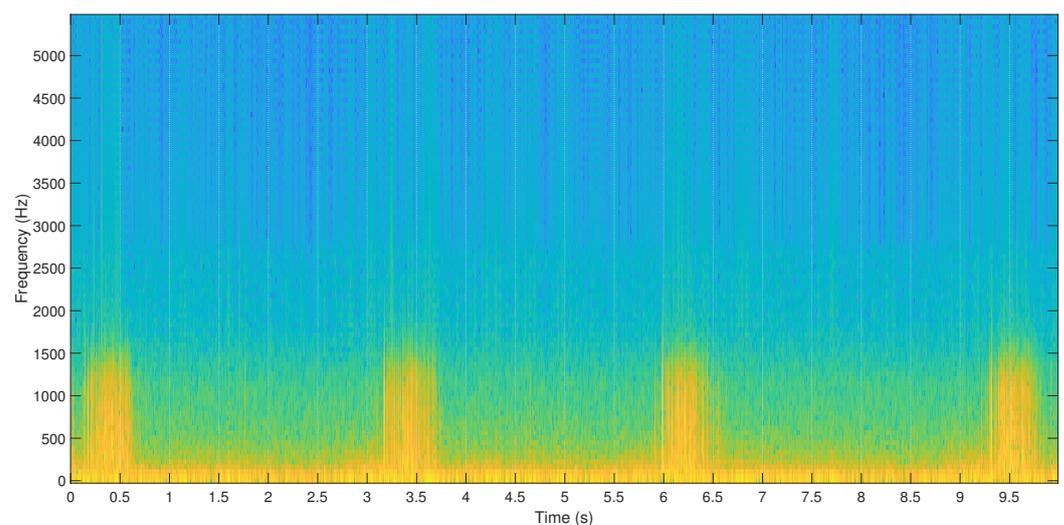

Figure 1. Magnitude spectrogram of a respiratory sound signal with a duration of ten seconds. The crackle sounds often tend to show vertical lines due to their short duration and explosive nature. It can be observed that most of the crackles are located in the intervals [0–0.7], [3.1–3.6], [6–6.4] and [9.4–9.7] seconds.

For many years, signal processing and machine learning approaches have been combined for event detection and classification tasks using spectro-temporal features [14–16]. Specifically for the task of crackle sound detection, several approaches have been proposed based on spectrogram analysis [17,18], autoregressive (AR) models [19,20], wavelet transform [21–24], fractal dimension filtering [25–28], entropy [29,30], empirical mode decomposition (EMD) [31], fuzzy systems [32], Gaussian mixture models (GMM) [33], logistic regression [34], support vector machines (SVM) [35–37], independent component analysis (ICA) [38], multi-perceptron networks (MPNs) [39], non-negative matrix factorization (NMF) [40], convolutional neural networks (CNNs) [41,42], recurrent neural networks (RNNs) [43,44] and hybrid neural networks [45,46]. Serbes et al. [23] detected crackle sounds testing different windows and wavelets for time-scale analysis. The extracted features were fed into SVM and k-nearest neighbor (kNN) models, but frequency bands containing no crackle information were removed using a dual-tree complex wavelet to reduce redundancy. Li and Hong [35] developed a crackle detection method using preprocessing, feature extraction (ratio between frequency of limbic signal, standard deviation of the limbic signal time, and the smoothing time of the limbic signal) and classification based on SVM. In [43], an event detection approach for crackle detection, based on RNNs using spectral features, was proposed. The approach exploited spectral information and temporal dependencies of the lung sounds, showing robustness regarding the contamination of the lung sound recordings with noise, bowel, and heart sounds. Garcia et al. [38] proposed to evaluate the efficiency of three popular ICA algorithms (FastICA, information maximization, and temporal decorrelation source separation) to determine the best method to extract crackle sounds. Information maximization achieved the best results, determining the presence of crackles in the independent components by means of their kurtosis and skewness, whereas the type of crackle was indicated by their characterization via the spectrogram of those components. In [42], CNNs were used to recognize lung sounds (including coarse and fine crackles) by means of spectrograms, Mel frequency cepstrum coefficient (MFCC) features and local binary pattern (LBP) features. Results demonstrated that the performance depends on learning parameters, batch size, and of iterations, band moreover, CNNs can replace conventional classifiers using fully connected layers to train the previous features. In [24], a classification system was proposed based on wavelets, genetic algorithms, and SVMs in order to discriminate between the presence or absence of crackles caused by pneumonia. Recently, Pal and Barney [28] proposed a crackle separation and detection technique combining an iterative envelope mean (IEM) method with the established fractal dimension (FD) technique [47]. The IEM method estimated the non-stationary and stationary parts of the lung sound signal and then the FD technique was applied to the estimated non-stationary output of the IEM method for further removing elements related to normal breath sounds. Authors indicated that IEM-FD filtering showed a high rate of crackle detectability when several signal-to-noise ratios (SNRs) were evaluated.

Crackle detection is a complex task due to factors such as: (a) nonstationarity of lung and crackle sounds, (b) low-magnitude relation between crackles and lung sounds (SNR), (c) crackle overlapping, (d) crackle waveform distortion by lung sounds, and (e) difficulty to establish the time domain parameters such as t_{IDW}, t_{2CD}, and t_{LDW}. The autoregressive (AR) model has proved useful for processing stationary random process, for example, for spectral analysis even with short time segments of information. In fact, respiratory sounds can be represented as a response of a system formed by lung parenchyma and chest wall to a white noise input which models the respiratory sounds source; hence, the system may be modeled as an all-pole filter.

Several AR-based models have been presented in the literature. In [48], AR-based features were used to feed two classifiers, a k-nearest neighbour (k-NN) classifier and a quadratic classifier, in order to discriminate between pathological and healthy patients. In [49], the authors combined sixth-order AR coefficients, wavelet coefficients, and crackle parameters as input for a k-NN and single layer perceptron. In a multichannel scenario (i.e., using several microphones), multivariate AR version called vector AR (VAR) has been

used in [50] as input for SVM and Gaussian mixture model (GMM) classifiers. Moreover, in [51], the authors proposed a crackle detection system using a threshold over the coefficients of a time-varying AR (TVAR) model showing superior results than using waveform analysis.

There are studies that have evaluated the detection of abnormal lung sounds from a physical perspective. In [52], it was asserted that determining the origin of abnormal sounds can provide valuable insights, potentially narrowing down the range of possible pathological causes and pinpointing the specific affected lung region. This information can prove invaluable in the context of diagnosis and treatment planning. To address this matter, [53] introduced an algorithm capable of effectively imaging sound sources, a capability demonstrated through computer simulations and experiments involving life-sized gelatin models of the human thorax. Their findings suggest that meaningful spatial information can be extracted from recordings employing as few as 16 microphones. Recently, in [54], a discussion of these types of studies covering auscultatory methods was presented. Alternatively, in this paper we aim to detect and localize the temporal occurrence of adventitious sounds on an input single-channel respiratory signal. In fact, the information about the temporal location will serve to focus a doctor's attention on the regions of the signal where crackles are present.

In this paper, we investigate the effect of the autoregressive (AR)-based frequency features that characterizes the spectral envelope of a breathing signal. In particular, we propose the utilization of complex valued poles derived from the AR model as inputs for a support vector machine (SVM) classifier using the radial basis function (RBF) kernel. The proposed approach has been compared with other state-of-the-art approaches including the iterative envelope mean-fractal dimension (IEM-FD) and the time-varying autoregressive (TVAR) methods. Finally, the proposed AR-based features are combined with a state-of-the-art convolutional neural network (CNN) architecture for crackle detection [55]. The outcomes of this comparison revealed notable enhancements in the performance metrics associated with the proposed methodology for characterizing and detecting crackles events. In fact, when dealing with limited data quantities, SVM presents a preferable choice whereas deep learning models typically entail a substantial number of tunable weights (free variables) that necessitate calibration with data. Actually, when the quantity of these weights exceeds or approximates the number of available training examples, deep models tend to essentially "memorize" the data, which can lead to overfitting. On the contrary, when a suitable kernel function is applied, SVM provides a robust, efficient, easily interpretable solution and is less prone to overfitting in classification problems.

The paper is organized as follows. Section 2 reviews the dataset used in this work and details the proposed method. Section 3 describes the metrics, setup, and the state-of-the-art methods in order to discuss the experimental results. Finally, conclusions and future work are presented in Section 4.

2. Materials and Methods

2.1. Dataset

In order to assess the crackle detection performance of the proposed method in a biomedical sound scenario, the dataset ψ was created using the same available software provided by the authors [28,56]. Specifically, the dataset ψ is composed of 2520 signals: (i) simulated signals related to FCSs and CCSs by modifying both parameters t_{IDW} and t_{2CD} according to [11,12,57]; (ii) real signals related to FCSs and CCSs with different t_{IDW} and t_{2CD} extracted from lung sound recordings. Specifically, real FCSs were extracted from a patient with idiopathic pulmonary fibrosis (IPF), and real CCSs were selected from a patient with bronchiectasis (BE) [56]. Each simulated and real signal has been mixed with noise N_R that shows the same spectral energy distribution as typically found in breath noise from a healthy subject measured over the lung bases on the right-hand side of the back as occurs in [56]. Moreover, several signal-to-noise ratios (SNRs) ranged from −10 to 10 dB in steps of 1 dB have been created to evaluate the robustness of the proposed

method detecting crackles. For each SNR, 15 simulated signals of every scenario have been evaluated, considering the effect of random variations of the local SNR around any given crackle. In this manner, a set of 315 signals have been generated considering all the SNRs for each type of simulated or real crackle signal. Although the dataset ψ is detailed in Table 1, more details can be found in [28,56].

Table 1. Dataset ψ. FCS: (i) ATS: t_{IDW} = 0.7 ms & t_{2CD} = 5 ms, (ii) Hoevers: t_{IDW} = 0.5 ms & t_{2CD} = 3.3 ms [45], (iii) Cohen: t_{IDW} = 0.9 ms & t_{2CD} = 6 ms. CCS: (i) ATS: t_{IDW} = 1.5 ms & t_{2CD} = 10 ms, (ii) Hoevers: t_{IDW} = 1 ms & t_{2CD} = 5.1 ms, (iii) Cohen: t_{IDW} = 1.25 ms & t_{2CD} = 9.5 ms. K_C: number of crackles per signal. $NOTS$: number of signals per SNR. N_S: number of signals generated taking into account all SNRs evaluated.

Scenario	Type	Model	K_C	$NOTS$	Noise	Diagnosis	SNR	N_S
Simulated	FCS [57]	ATS	10	15	N_R	-	[−10 dB, 10 dB]	315
	FCS [11]	Hoevers	10	15	N_R	-	[−10 dB, 10 dB]	315
	FCS [12]	Cohen	10	15	N_R	-	[−10 dB, 10 dB]	315
	CCS [57]	ATS	10	15	N_R	-	[−10 dB, 10 dB]	315
	CCS [11]	Hoevers	10	15	N_R	-	[−10 dB, 10 dB]	315
	CCS [12]	Cohen	10	15	N_R	-	[−10 dB, 10 dB]	315
Real	FCS [28,56]	-	10	15	N_R	IPF	[−10 dB, 10 dB]	315
	CCS [28,56]	-	10	15	N_R	BE	[−10 dB, 10 dB]	315

2.2. Modeling of Simulated Crackle Sounds

In order to train the classifier, a set of synthetic crackle sounds have been generated. Similar to [58], a crackle waveform $y(t)$ is created (Equations (1)–(4)) assuming: (i) the crackle has two cycles represented by the parameter t_{2CD}, (ii) the location where the first cycle of the waveform $y(t)$ equals zero is explicitly indicated by the parameter t_{IDW}, and (iii) most of the power in $y(t)$ is concentrated near the beginning of the waveform. To time shift most of the power to the beginning of the waveform, a modulating function $m(t)$ is generated. An example of simulated crackles in the time domain is shown in Figure 2.
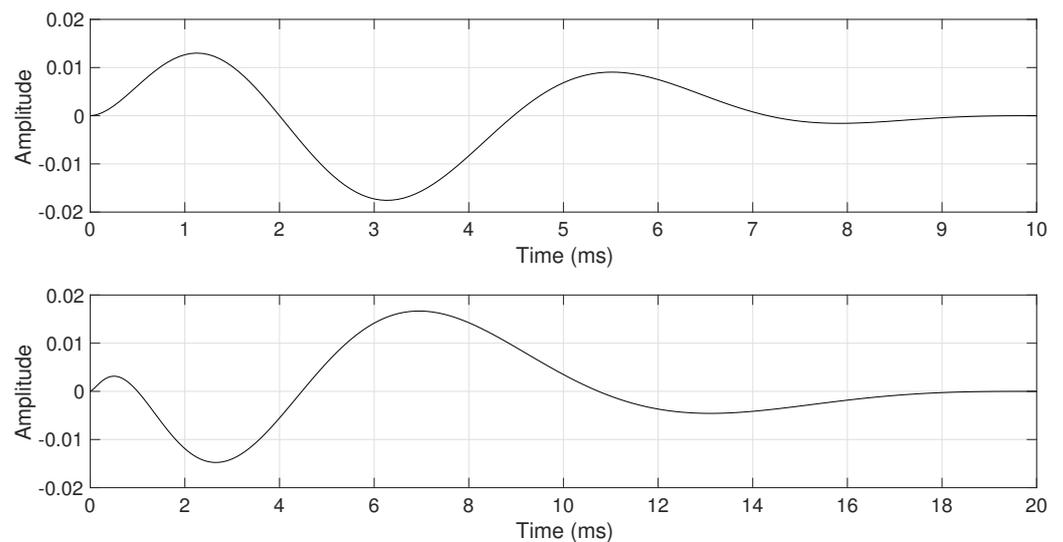

Figure 2. Two simulated crackles, normalized in energy, are modelled: (t_{IDW}, t_{2CD}) = (2 ms, 10 ms) in the top plot and (t_{IDW}, t_{2CD}) = (1 ms, 20 ms) in the bottom plot.

$$t_0 = \frac{t_{IDW}}{t_{2CD}} \tag{1}$$

$$y_0(t) = sin(4\pi t^\alpha), \alpha = \frac{log_{10}(0.25)}{log_{10}(t_0)} \tag{2}$$

$$m(t) = \frac{1}{2}\left(1 + cos\left(2\pi\left(t^{\frac{1}{2}} - \frac{1}{2}\right)\right)\right) \tag{3}$$

$$y(t) = y_0(t)m(t) \tag{4}$$

In this work, a set of $K = 187$ crackle waveforms was created based on a conservative strategy using all combinations between both parameters t_{IDW} and t_{2CD} [11,12,57]. Specifically, $t_{IDW} \in [0.5, 1.5]$ ms with a step size of 0.1 ms and $t_{2CD} \in [3.3, 20]$ ms with a step size of 1 ms. Finally, each signal $y(t)$ is normalized in energy, that is, $\sum_t y^2(t) = 1.0$. The magnitude Fourier transform (spectral pattern) of the previous set of K crackle waveforms is shown in Figure 3.
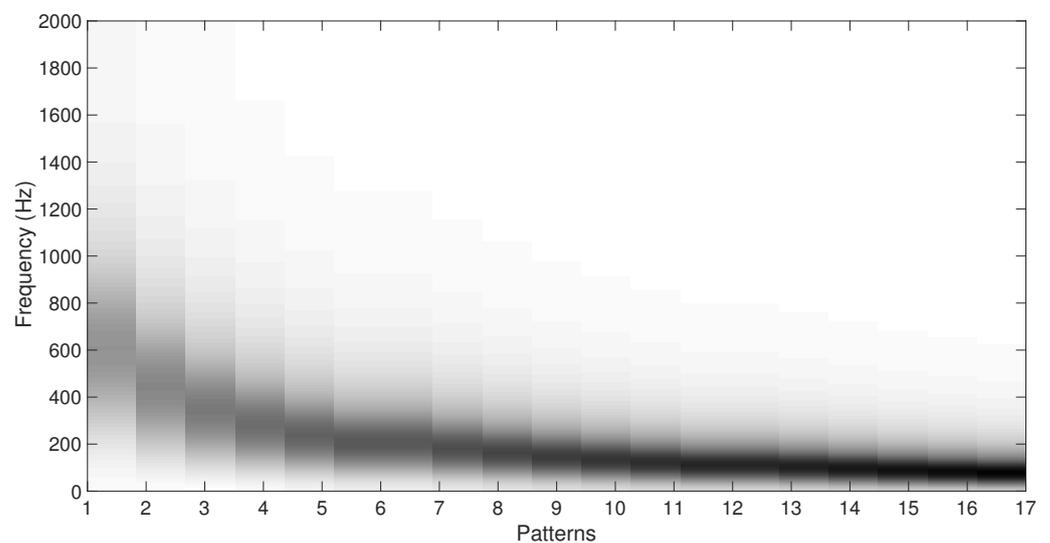

Figure 3. Magnitude spectrogram of the first eighteen spectral patterns combining the parameters t_{IDW} and t_{2CD} as previously mentioned. Higher energy is indicated by darker colour.

2.3. Proposed Method

In Figure 4, a block diagram of the structure of our proposed method is shown. First, a preprocessing stage is performed to discard information out of the band of interest and define the segments for the short-time signal analysis. Then, we window the initial signal into frames (parameters used are explained in the Setup section). We calculate the 14 complex envelope coefficients for each frame by using our AR model. Finally, we use these data to feed the support vector machine system.

2.3.1. Preprocessing

For the task of crackle detection, the signals are low-pass filtered and resampled to 4 kHz. The window length for the short-time analysis was chosen to be 256 samples (64 ms) with a hop size of 16 ms. This interval was a good compromise between parameter accuracy and stationary requirements. Finally, the sliding segments were weighted by a Hamming window to reduce the spectral leakage.
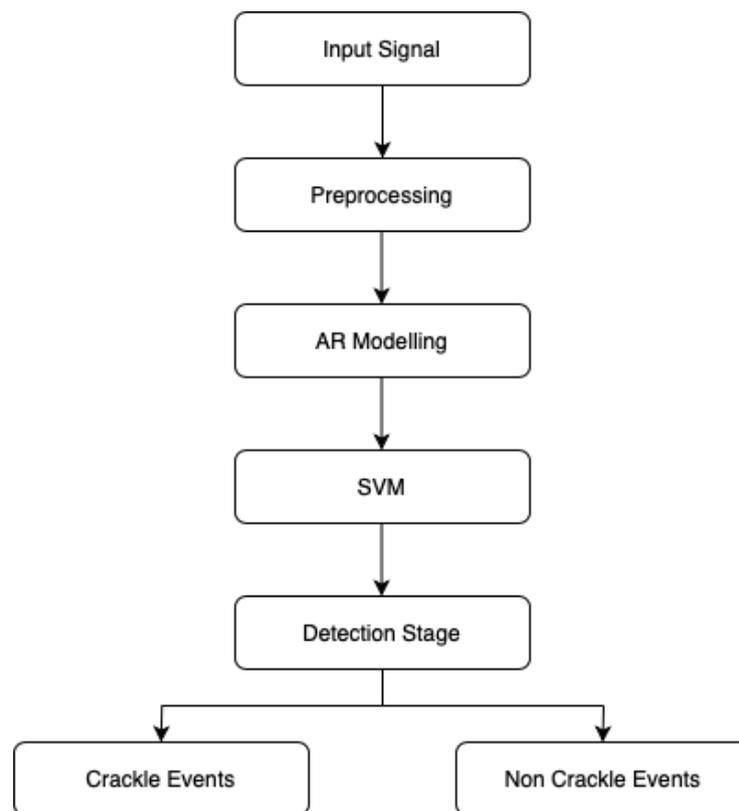
Figure 4. The block diagram of the proposed method.

2.3.2. Autoregressive (AR) Parameter Estimation

In this paper, AR parameters are estimated for each segment of the respiratory sound as

$$x_s(n) = -\sum_{k=1}^{K} a_s(k)x_s(n-k) + e_s(n),$$ (5)

where $s = 1, \ldots, S$ denotes the segment index, $a_s(k)$ is the k-th AR coefficient with K denoting the model order. $e_s(n)$ is IID Gaussian noise with zero mean and variance σ^2. Levinson–Durbin recursion was used in order to estimate the AR parameters. This recursion also gives the total mean square prediction error of a k-th order predictor. In this work, sixth-order AR model coefficients are extracted from respiratory sounds to form the feature set for the classifier as was shown in [48,49,59].

The above equation can be solved by using the z-transform. This allows the equation to be written as

$$X_s(z)\left(1 + \sum_{k=1}^{K} a_s(k)z^{-k}\right) = E_s(z)$$ (6)

and the transfer function $H_s(z)$ can be expressed using a rational format

$$H_s(z) = z^K \frac{1}{\prod_{k=1}^{K}(z - p_s(k))}$$ (7)

where $p_s(k)$ are the poles describing the estimated spectral envelope of the input segment. This approach has been widely used to characterize and synthesize the speech formants [60].

As shown in Figure 5, the frequency response of the all-pole AR model is depicted for two healthy and unhealthy excerpts. These features gave us the aim to proceed into the creation of a classification system between these two classes (healthy and unhealthy patients).
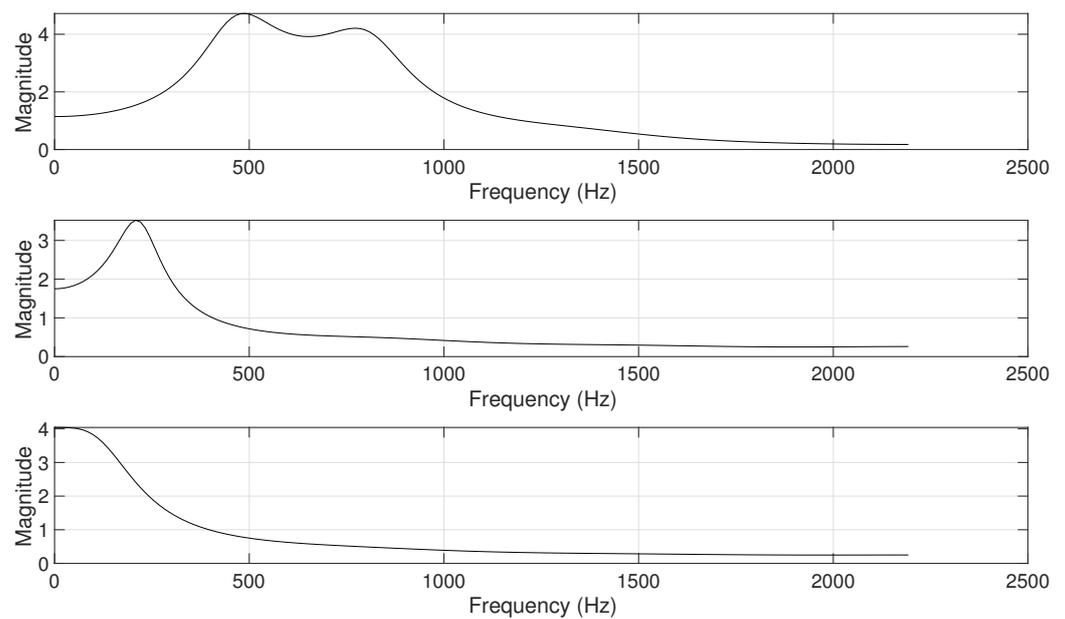
Figure 5. Frequency response for the all-pole AR model using 6-th coefficients modelling the breathing sound excerpt for a healthy patient (top plot), an unhealthy frame with coarse crackle (middle plot), and an unhealthy frame with fine crackle (bottom plot).

2.3.3. SVM Classifier

SVM is a supervised learning method which is often applied to classification or regression. It is a method for obtaining the optimal boundary of two sets in a vector space independently from the probabilistic distributions of training vectors in the sets. A support vector machine is used here to classify crackles and normal lung sounds. The RBF kernel function is used in this work to achieve an optimal result. In addition, the parameters of the SVM are optimized to obtain a better performance, which will be introduced in the third section.

Support vector machines (SVMs) are the famous and widely utilized supervised classifiers developed by [61]. A support vector machine (SVM) is a powerful and versatile machine learning algorithm used primarily for classification and regression tasks. SVMs are particularly effective in scenarios where the data are not linearly separable, meaning that the classes cannot be separated by a single straight line or hyperplane. SVMs achieve this by transforming the input data into a higher-dimensional space where separation becomes possible through the use of a hyperplane.

In a classification context, an SVM attempts to find the hyperplane that best separates different classes of data points while maximizing the margin between the two classes. The margin is the distance between the hyperplane and the nearest data points from each class. The SVM seeks to find the hyperplane that not only separates the classes but also maximizes this margin, which theoretically reduces the risk of overfitting and enhances the model's generalization capabilities.

The data points that are closest to the hyperplane and have the most influence on determining its position are referred to as "support vectors". These are the data points located at the edges of the margin or those that might be misclassified if the margin were to shift. The SVM focuses on these support vectors, making it a robust algorithm even in the presence of outliers or noisy data.

SVMs offer flexibility in dealing with different types of data through the use of various kernel functions, which can implicitly map the data into higher-dimensional spaces. This allows SVMs to handle non-linear relationships between variables.

3. Evaluation

3.1. Metrics

Three metrics are proposed to detect the presence or absence of crackle events: accuracy (*Acc*), sensitivity (S_e) and precision (P_r). All of them are calculated analyzing each signal within the dataset ψ by matching the original data and the estimated data provided by the proposed method. The metric *Acc* represents the ability to correctly detect the presence or absence of the crackles events when they are active or inactive in the signal, S_e represents the ability to detect the number of missed crackle events within the dataset, and P_r represents the ability to detect frames within a signal with crackles when no crackles events are active. These three mentioned metrics are calculated as shown in Equations (8)–(10). The parameter t_p indicates true positives, that is, number of active crackle events within the dataset ψ correctly detected as crackle. The parameter f_p represents false positives or false alarm events, that is, the number of crackles events inactive within the dataset ψ incorrectly detected as crackles. The parameter f_n indicates false negatives or missed events, that is, the number of crackles events active within the dataset ψ incorrectly detected as no crackle.

$$Acc = \frac{t_p}{t_p + f_p + f_n} \tag{8}$$

$$S_e = \frac{t_p}{t_p + f_n} \tag{9}$$

$$P_r = \frac{t_p}{t_p + f_p} \tag{10}$$

3.2. Setup

We conducted preliminary experiments that showed that the following parameters, used in this paper, provide the best trade-off between crackle detection performance and the computational cost: sampling rate f_s = 4410 Hz, window size N = 18 samples (approximately 4.1 ms), hop size of four samples (approximately 0.9 ms), and a number of FFT points equal to twice the window size in samples.

3.3. State-of-the-Art Methods for Comparison

Two recent and relevant state-of-the-art crackle detection methods IEM-FD [28] and TVAR [51] have been implemented in order to evaluate the performance of crackle detection of the proposed method.

3.4. Results and Discussion

In this section, we detail a systematic comparison of the results section obtained from the implemented methods: the IEM-FD, TVAR, and the proposed method. All of the separation techniques were implemented using the MATLAB (R2020b) programming language. The metrics employed to evaluate the implemented methods are described in Section 3.1. For evaluation, we compared the ground-truth and estimated temporal location and length of each crackle using a tolerance window of 0.6 ms.

The overall results for the implemented methods are shown in Figure 6. The boxplot displays the distribution of data based on a five-number summary ("minimum", first quartile (Q1), median, third quartile (Q3), and "maximum"). It details information about the atypical values and the location of them within the entire range of possible values. At the same time, it indicates whether the data are symmetrical, how tightly the data are grouped, and how the data are skewed.

Figure 6 shows these results in terms of accuracy *Acc*, sensitivity S_e and precision P_r of the IEM-FD method on the left side, TVAR method on the center and the proposed method on the right side. Additionally, Figure 7 displays the accuracy results of the compared methods as a function of the SNR value (between −10 and 10 dB).

In general terms, the values obtained from the proposed method are significantly higher in terms of accuracy *Acc* and precision P_r among the compared methods. Moreover, our method also outperforms the IEM-FD and the TVAR in terms of sensitivity (S_e).
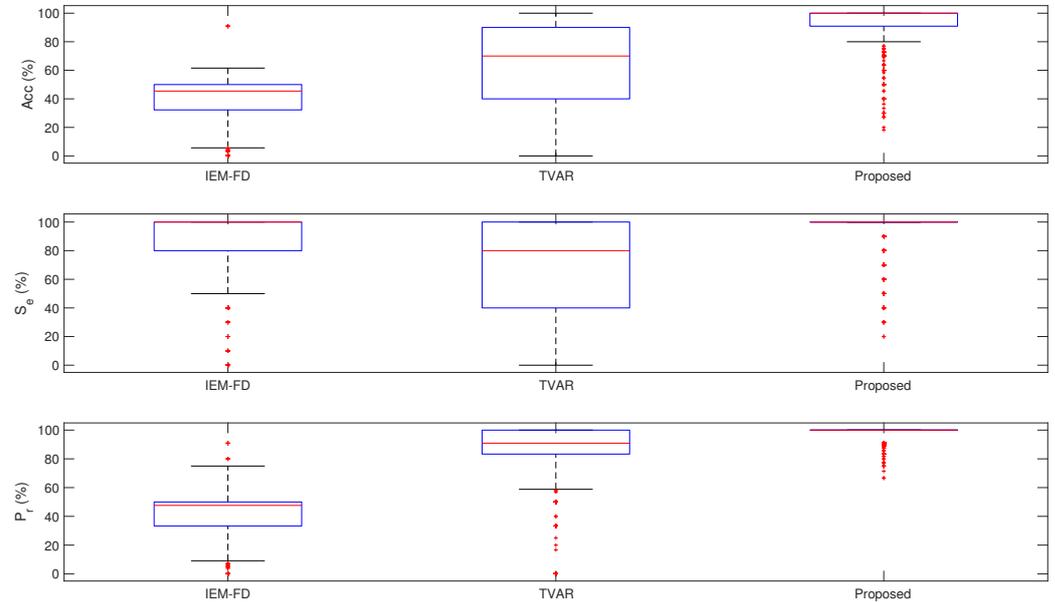

Figure 6. Accuracy, sensitivity and precision average results evaluating all scenarios and SNRs in the dataset ψ by IEM-FD [28], TVAR [51], and the proposed method.
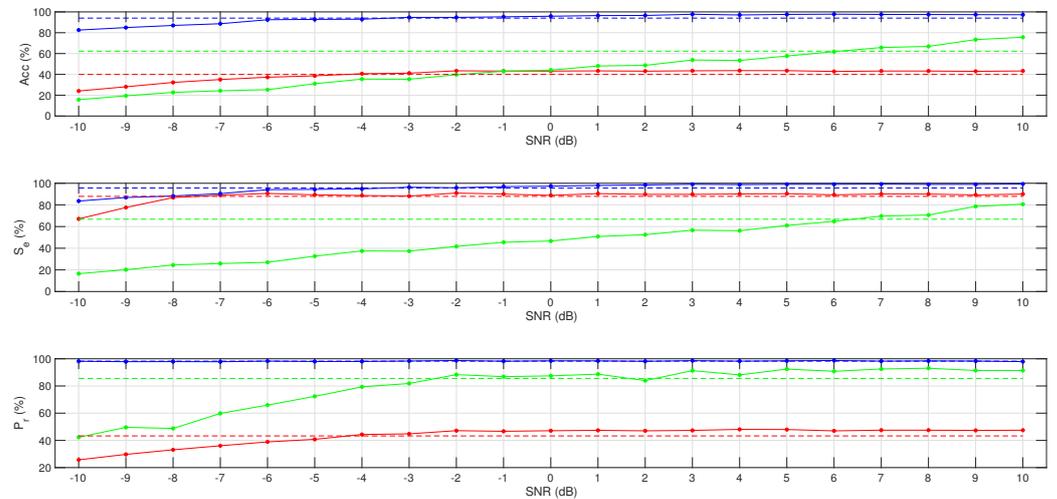

Figure 7. Accuracy, sensitivity, and precision average results evaluating all scenarios for each SNR in the database ψ by IEM-FD [28] (red color), TVAR [51] (green color), and the proposed method (blue color), where the dashed lines represent the mean value for each metric and method.

Regarding the boxplot accuracy values of the IEM-FD method, the median is 45.45%, the first quartile (Q1) is 32.25%, the second quartile (Q2) is 50%, the maximum is 61.54%, and the minimum is 5.66%. In fact, the performance of this method is limited by the number of false positives (i.e., normal events detected as crackles) as be seen in the precision values in Figure 6.

In the case of the TVAR method, the following measurements are found: the median is 70%, the first quartile (Q1) is 90%, the second quartile (Q2) is 40%, the maximum is 100%, and the minimum is 0%. This variability in the results suggests that the method is too dependent on the conditions of the input signals, as can be seen in Figure 7. In fact, the

TVAR method provides the worst results among the compared methods when dealing with low SNR values.

The proposed method shows the minimum standard deviation among the compared methods. In particular, the median is 100%, the first quartile (Q1) is 90.90%, the second quartile (Q2) is 100%, the maximum is 100%, and the minimum is 80%. Additionally, as can be seen in Figure 7, the results are robust independently of the input signal SNR conditions.

Now, focusing on a specific type of crackles, Figure 8 presents the Acc, S_e, and P_e results as a function of type of crackle (coarse (right side of the figure) or fine (left side of the figure)) and SNR value. As can be seen, the performance of the proposed method is clearly superior to the other compared methods in both cases (coarse and fine crackle detection). Comparing both sides of the figure, the results reveal that the proposed method underperforms when dealing with coarse crackles under low SNR scenarios. Similar behavior is observed in the case of the TVAR method although, as previously commented, the results are limited by the number of false positives, which provokes a clear underperformance in terms of S_e. Interestingly, the results seem more stable in the case of IEM-FD when comparing the performance of fine vs. coarse crackle detection. In any case, it is worth mentioning that the results are worse than the compared approaches and specially limited by the number of false negatives (i.e., crackles events detected as normal), which provokes a clear underperformance in terms of P_r. Note that the Acc can be seen as general metric accounting for both S_e and P_r values.
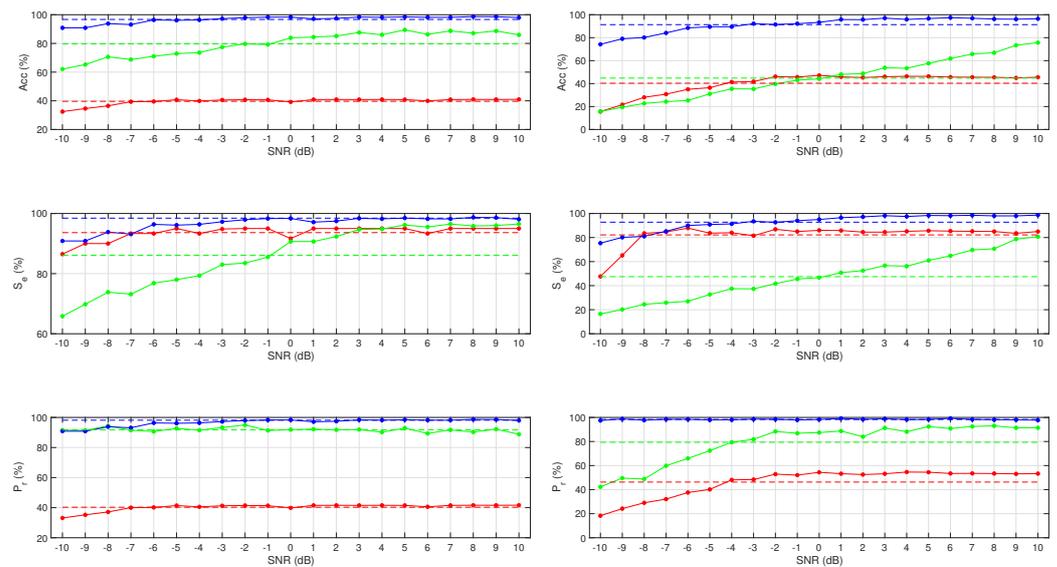

Figure 8. Accuracy, sensitivity, and precision average results evaluating all scenarios for each type (fine crackles on the left side and coarse crackles on the right side) of crackles and SNRs from database ψ by IEM-FD [28] (red color), TVAR [51] (green color), and the proposed method (blue color), where the dashed lines represent the mean value for each metric and method.

Table 2 shows the mean values in terms of accuracy Acc, sensitivity S_e and precision P_r for the IEM-FD method, TVAR method, and the proposed method for test samples ψ at SNR from -10 dB and until $+10$ dB for real and simulated crackles. In contrast to the previous results shown in this document, here we make a distinction by type of crackle within the test sample: ATS, Hoevers, Cohen, IPF, or BE, labeled here as simulated scenarios 1 to 6 and real scenarios 1 and 2. From these results, we can highlight the fact that our proposed method has widely higher percentage values in comparison with the IEM-FD and TVAR methods. It is important to note that all the methods have a clear underperformance in real scenario 2. That is, when dealing with real sounds with coarse crackles (i.e., bronchitis) superposed to simulated breathing. This suggests that the difference between simulated

and real fine crackles allows a suitable modeling valid for both cases whereas in the case of coarse crackles, the parameters may vary between simulated and real cases.

Table 2. Detailed results in terms of accuracy, sensitivity, and precision (mean values per crackle type) K_C: number of crackles per signal. *NOTS*: number of signals per SNR. N_S: number of signals generated taking into account all SNRs evaluated.

Scenario	Type	K_C	*NOTS*	Noise	Diagnosis	SNR	Accuracy [Acc]			Sensitivity [S_e]			Precision [P_r]		
							IEM-FD	TVAR	Proposed	IEM-FD	TVAR	Proposed	IEM-FD	TVAR	Proposed
Simulated 1	FCS [57]	10	15	N_R	-	[−10 dB, 10 dB]	46.94%	69.23%	98.36%	97.46%	73.02%	99.94%	46.94%	93.33%	98.42%
Simulated 2	FCS [11]	10	15	N_R	-	[−10 dB, 10 dB]	33.21%	81.36%	95.85%	100%	92.60%	97.81%	33.21%	87.27%	97.93%
Simulated 3	FCS [12]	10	15	N_R	-	[−10 dB, 10 dB]	47.72%	76.21%	98.03%	97.14%	81.56%	99.94%	47.72%	91.98%	98.08%
Simulated 4	CCS [57]	10	15	N_R	-	[−10 dB, 10 dB]	41.86%	46.00%	96.57%	93.17%	48.54%	98.25%	42.15%	82.64%	98.27%
Simulated 5	CCS [11]	10	15	N_R	-	[−10 dB, 10 dB]	47.63%	71.86%	98.63%	97.46%	76.73%	99.90%	47.63%	91.68%	98.71%
Simulated 6	CCS [12]	10	15	N_R	-	[−10 dB, 10 dB]	41.12%	38.68%	95.47%	91.49%	40.76%	96.86%	41.46%	76.70%	98.54%
Real 1	FCS [28,56]	10	15	N_R	IPF	[−10 dB, 10 dB]	30.70%	92.04%	94.59%	79.96%	97.30%	96.10%	33.26%	94.39%	98.32%
Real 2	CCS [28,56]	10	15	N_R	BE	[−10 dB, 10 dB]	30.64%	22.70%	74.78%	46.54%	23.87%	76.03%	53.86%	66.33%	97.72%

Once we have evaluated the robustness of the proposed method against different type of signals and SNR conditions and compared the results with classical approaches, we aim to also compare the results with more recent techniques based on deep learning. In fact, deep learning approaches have recently been widely investigated for the task of adventitious sound detection [41,62–68].

The investigation carried out by [55] highlighted the usage of convolutional neural networks (CNNs) as state-of-the-art solutions across various research domains. Leveraging the CNN architecture as detailed in [55], we subjected the same scenarios as previously outlined to testing, subsequently conducting a comparative analysis between the outcomes generated by the support vector machine (SVM) classifier and the CNN classifier. The comparative results, presented in Table 3, delineate the performance contrast in terms of accuracy (Acc), sensitivity (S_e), and precision (P_r) across the eight scenarios encompassing our proposed method. As can be observed in Table 3, both classifiers (SVM and CNN) achieve excellent results (>90%) in all metrics except in the real coarse crackle scenario. In any case, a slight improvement can be observed using the SVM-based classifier in terms of accuracy (Acc) and sensitivity (S_e). Conversely, CNN performs slightly better in terms of precision (P_r). This suggests that the proposed system is somewhat more robust in determining the occurrence of crackle events, which is very interesting from a clinical perspective. However, it is worth noting that SVM is simpler to train, has fewer parameters, and therefore is more robust to overfitting. Additionally, its parameters are much more interpretable than those of CNN.

It is pertinent to underscore the notable performance decline observed in the real-world scenarios examined in this study. As evident in Table 3, these scenarios exhibit discernible decreases across all three metrics. These findings shed light on the potential limitations inherent in the proposed methodology. In the context of real patient scenarios, it is imperative to recognize that the properties of the input signals may undergo fluctuations, resulting in consequential reductions in performance outcomes. As articulated in the preceding Section 4, forthcoming research endeavors will be geared towards data characteristic extraction and the modeling of time–frequency behaviors. These efforts are envisioned to foster improved generalization regarding the authentic behaviors exhibited by such sound signals in real-world contexts.

Table 3. Comparison of the results of the proposed method as input of an SVM and CNN in terms of accuracy, sensitivity, and precision (mean values per crackle type) K_C: number of crackles per signal. *NOTS*: number of signals per SNR. N_S: number of signals generated considering all SNRs evaluated.

Scenario	Type	K_C	*NOTS*	Noise	Diagnosis	SNR	Accuracy (%) [*Acc*]		Sensitivity (%) [S_e]		Precision (%) [P_r]	
							SVM	CNN	SVM	CNN	SVM	CNN
Simulated 1	FCS [57]	10	15	N_R	-	[−10 dB, 10 dB]	98.36	97.61	99.94	97.94	98.42	99.49
Simulated 2	FCS [11]	10	15	N_R	-	[−10 dB, 10 dB]	95.85	97.93	97.81	98.28	97.93	99.53
Simulated 3	FCS [12]	10	15	N_R	-	[−10 dB, 10 dB]	98.03	97.44	99.94	97.72	98.08	99.50
Simulated 4	CCS [57]	10	15	N_R	-	[−10 dB, 10 dB]	96.57	92.98	98.25	92.73	98.27	99.60
Simulated 5	CCS [11]	10	15	N_R	-	[−10 dB, 10 dB]	98.63	98.25	99.90	98.54	98.71	99.55
Simulated 6	CCS [12]	10	15	N_R	-	[−10 dB, 10 dB]	95.47	94.34	96.86	94.21	98.54	99.58
Real 1	FCS [28,56]	10	15	N_R	IPF	[−10 dB, 10 dB]	94.58	91.42	96.10	93.11	98.32	99.34
Real 2	CCS [28,56]	10	15	N_R	BE	[−10 dB, 10 dB]	74.78	72.34	76.03	74.00	97.72	98.12

4. Conclusions and Future Work

In this paper, a new crackle event detection method based on the combination of an autoregressive model, and a support vector machine classification model is proposed. We conclude that the proposed method is a suitable model for detection of the existence or non-existence of crackle events within a dataset with a significant high success rate.

The proposed method has achieved highly competitive results in the detection of crackle events in spite of the environmental factors, such as type of crackle (fine or coarse) or very low signal-to-noise ratio.

Future work will focus on combining recurrent and convolutional neural networks approaches using different time–frequency representations in order to develop novel criteria to determine the most reliable and discriminant feature map in terms of the abnormal respiratory sound to be detected.

Author Contributions: Conceptualization, L.D.M., J.J.C.-O. and F.J.C.-Q.; methodology, L.D.M., J.J.C.-O and F.J.C.-Q.; software, L.D.M., J.d.l.T.-C., A.M.-M. and P.R.-S.; validation, L.D.M. and P.R.-S.; formal analysis, L.D.M., J.J.C.-O., F.J.C.-Q. and E.F.C.; investigation, L.D.M., J.J.C.-O., F.J.C.-Q. and E.F.C.; resources, L.D.M., J.J.C.-O. and E.F.C.; data curation, L.D.M., J.J.C.-O., J.d.l.T.-C., A.M.-M. and P.R.-S.; writing—original draft preparation, L.D.M., J.J.C.-O. and F.J.C.-Q.; writing—review and editing, L.D.M., J.J.C.-O., F.J.C.-Q., J.d.l.T.-C., A.M.-M., P.R.-S. and E.F.C.; visualization, L.D.M., J.d.l.T.-C. and A.M.-M.; supervision, J.J.C.-O. and F.J.C.-Q.; project administration, F.J.C.-Q.; funding acquisition, F.J.C.-Q. and E.F.C. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: https://github.com/loredanadariamang/Applied_sciences_crackle_detection_localization.git (accessed on 21 September 2023).

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

References

1. World Health Organization. Pneumonia. 2017. Available online: https://www.who.int/health-topics/pneumonia#tab=tab_1 (accessed on 23 September 2023).
2. Pneumonia—Diagnosis and treatment—Mayo Clinic. 2021. Available online: https://www.mayoclinic.org/diseases-conditions/pneumonia/diagnosis-treatment/drc-20354210 (accessed on 23 September 2023).
3. Ponte, D.F.; Moraes, R.; Hizume, D.C.; Alencar, A.M. Characterization of crackles from patients with fibrosis, heart failure and pneumonia. *Med. Eng. Phys.* **2013**, *35*, 448–456. [CrossRef] [PubMed]
4. İçer, S.; Gengeç, Ş. Classification and analysis of non-stationary characteristics of crackle and rhonchus lung adventitious sounds. *Digit. Signal Process.* **2014**, *28*, 18–27. [CrossRef]
5. Pancaldi, F.; Sebastiani, M.; Cassone, G.; Luppi, F.; Cerri, S.; Della Casa, G.; Manfredi, A. Analysis of pulmonary sounds for the diagnosis of interstitial lung diseases secondary to rheumatoid arthritis. *Comput. Biol. Med.* **2018**, *96*, 91–97. [CrossRef] [PubMed]
6. Reyes, B.A.; Olvera-Montes, N.; Charleston-Villalobos, S.; González-Camarena, R.; Mejía-Ávila, M.; Aljama-Corrales, T. A smartphone-based system for automated bedside detection of crackle sounds in diffuse interstitial pneumonia patients. *Sensors* **2018**, *18*, 3813. [CrossRef] [PubMed]
7. Sovijarvi, A.; Dalmasso, F.; Vanderschoot, J.; Malmberg, L.; Righini, G.; Stoneman, S. Definition of terms for applications of respiratory sounds. *Eur. Respir. Rev.* **2000**, *10*, 597–610.
8. Salazar, A.J.; Alvarado, C.; Lozano, F.E. System of heart and lung sounds separation for store-and-forward telemedicine applications. In *Revista Facultad de Ingeniería Universidad de Antioquia*; 2012; pp. 175–181. Available online: https://revistas.udea.edu.co/index.php/ingenieria/issue/view/1223 (accessed on 21 September 2023).
9. Sovijarvi, A. Characteristics of breath sounds and adventitious respiratory sounds. *Eur. Respir. Rev.* **2000**, *10*, 591–596.
10. Pramono, R.X.A.; Bowyer, S.; Rodriguez-Villegas, E. Automatic adventitious respiratory sound analysis: A systematic review. *PLoS ONE* **2017**, *12*, e0177926. [CrossRef]
11. Hoevers, J.; Loudon, R.G. Measuring crackles. *Chest* **1990**, *98*, 1240–1243. [CrossRef]
12. Cohen, A. Signal processing methods for upper airway and pulmonary dysfunction diagnosis. *IEEE Eng. Med. Biol. Mag.* **1990**, *9*, 72–75. [CrossRef]
13. Speranza, C.G.; Moraes, R. Instantaneous frequency based index to characterize respiratory crackles. *Comput. Biol. Med.* **2018**, *102*, 21–29. [CrossRef]
14. Chan, T.K.; Chin, C.S. A Comprehensive Review of Polyphonic Sound Event Detection. *IEEE Access* **2020**, *8*, 103339–103373. [CrossRef]
15. Radad, M. Application of single-frequency time-space filtering technique for seismic ground roll and random noise attenuation. *J. Earth Space Phys.* **2018**, *44*, 41–51.
16. Hadiloo, S.; Radad, M.; Mirzaei, S.; Foomezhi, M. Seismic facies analysis by ANFIS and fuzzy clustering methods to extract channel patterns. In Proceedings of the 79th EAGE Conference and Exhibition 2017, Paris, France, 12–15 June 2017; Volume 2017, pp. 1–5.
17. Kaisia, T.; Sovijärvi, A.; Piirilä, P.; Rajala, H.; Haltsonen, S.; Rosqvist, T. Validated method for automatic detection of lung sound crackles. *Med. Biol. Eng. Comput.* **1991**, *29*, 517–521. [CrossRef] [PubMed]
18. Zhang, K.; Wang, X.; Han, F.; Zhao, H. The detection of crackles based on mathematical morphology in spectrogram analysis. *Technol. Health Care* **2015**, *23*, S489–S494. [CrossRef] [PubMed]
19. Hadjileontiadis, L.; Panas, S. Nonlinear separation of crackles and squawks from vesicular sounds using third-order statistics. In Proceedings of the 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Amsterdam, The Netherlands, 31 October–3 November 1996; Volume 5, pp. 2217–2219.
20. Charleston-Villalobos, S.; Martinez-Hernandez, G.; Gonzalez-Camarena, R.; Chi-Lem, G.; Carrillo, J.G.; Aljama-Corrales, T. Assessment of multichannel lung sounds parameterization for two-class classification in interstitial lung disease patients. *Comput. Biol. Med.* **2011**, *41*, 473–482. [CrossRef]
21. Hadjileontiadis, L.J.; Panas, S.M. Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter. *IEEE Trans. Biomed. Eng.* **1997**, *44*, 1269–1281. [CrossRef] [PubMed]
22. Lu, X.; Bahoura, M. An integrated automated system for crackles extraction and classification. *Biomed. Signal Process. Control* **2008**, *3*, 244–254. [CrossRef]
23. Serbes, G.; Sakar, C.O.; Kahya, Y.P.; Aydin, N. Pulmonary crackle detection using time–frequency and time–scale analysis. *Digit. Signal Process.* **2013**, *23*, 1012–1021. [CrossRef]
24. Stasiakiewicz, P.; Dobrowolski, A.P.; Targowski, T.; Gałązka-Świderek, N.; Sadura-Sieklucka, T.; Majka, K.; Skoczylas, A.; Lejkowski, W.; Olszewski, R. Automatic classification of normal and sick patients with crackles using wavelet packet decomposition and support vector machine. *Biomed. Signal Process. Control* **2021**, *67*, 102521. [CrossRef]
25. Hadjileontiadis, L.J. Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding-Part I: Methodology. *IEEE Trans. Biomed. Eng.* **2005**, *52*, 1143–1148. [CrossRef]
26. Hadjileontiadis, L.J. Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding-Part II: Application results. *IEEE Trans. Biomed. Eng.* **2005**, *52*, 1050–1064. [CrossRef] [PubMed]
27. Pinho, C.; Oliveira, A.; Jácome, C.; Rodrigues, J.; Marques, A. Automatic crackle detection algorithm based on fractal dimension and box filtering. *Procedia Comput. Sci.* **2015**, *64*, 705–712. [CrossRef]

28. Pal, R.; Barney, A. Iterative envelope mean fractal dimension filter for the separation of crackles from normal breath sounds. *Biomed. Signal Process. Control* **2021**, *66*, 102454. [CrossRef]

29. Liu, X.; Ser, W.; Zhang, J.; Goh, D.Y.T. Detection of adventitious lung sounds using entropy features and a 2-D threshold setting. In Proceedings of the 2015 10th International Conference on Information, Communications and Signal Processing (ICICS), Cairns, Australia, 2–4 December 2015; pp. 1–5.

30. Rizal, A.; Hidayat, R.; Nugroho, H.A. Pulmonary crackle feature extraction using tsallis entropy for automatic lung sound classification. In Proceedings of the 2016 1st International Conference on Biomedical Engineering (IBIOMED), Yogyakarta, Indonesia, 5–6 October 2016; pp. 1–4.

31. Hadjileontiadis, L.J. Empirical mode decomposition and fractal dimension filter. *IEEE Eng. Med. Biol. Mag.* **2007**, *26*, 30.

32. Mastorocostas, P.A.; Theocharis, J.B. A dynamic fuzzy neural filter for separation of discontinuous adventitious sounds from vesicular sounds. *Comput. Biol. Med.* **2007**, *37*, 60–69. [CrossRef] [PubMed]

33. Maruf, S.O.; Azhar, M.U.; Khawaja, S.G.; Akram, M.U. Crackle separation and classification from normal Respiratory sounds using Gaussian Mixture Model. In Proceedings of the 2015 IEEE 10th International Conference on Industrial and Information Systems (ICIIS), Peradeniya, Sri Lanka, 18–20 December 2015; pp. 267–271.

34. Mendes, L.; Vogiatzis, I.M.; Perantoni, E.; Kaimakamis, E.; Chouvarda, I.; Maglaveras, N.; Henriques, J.; Carvalho, P.; Paiva, R.P. Detection of crackle events using a multi-feature approach. In Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 3679–3683.

35. Li, J.; Hong, Y. Crackles detection method based on time-frequency features analysis and SVM. In Proceedings of the 2016 IEEE 13th International Conference on Signal Processing (ICSP), Chengdu, China, 6–10 November 2016; pp. 1412–1416.

36. Grønnesby, M.; Solis, J.C.A.; Holsbø, E.; Melbye, H.; Bongo, L.A. Feature extraction for machine learning based crackle detection in lung sounds from a health survey. *arXiv* **2017**, arXiv:1706.00005.

37. Pramudita, B.A.; Istiqomah.; Rizal, A. Crackle detection in lung sound using statistical feature of variogram. In *AIP Conference Proceedings*; AIP Publishing LLC: Melville, NY, USA, 2020; Volume 2296, p. 020014.

38. García, M.R.; Villalobos, S.C.; Villa, N.C.; González, A.J.; Camarena, R.G.; Corrales, T.A. Automated extraction of fine and coarse crackles by independent component analysis. *Health Technol.* **2020**, *10*, 459–463. [CrossRef]

39. Liu, Y.X.; Yang, Y.; Chen, Y.H. Lung sound classification based on Hilbert-Huang transform features and multilayer perceptron network. In Proceedings of the 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Kuala Lumpur, Malaysia, 12–15 December 2017; pp. 765–768.

40. Hong, K.J.; Essid, S.; Ser, W.; Foo, D.G. A robust audio classification system for detecting pulmonary edema. *Biomed. Signal Process. Control* **2018**, *46*, 94–103. [CrossRef]

41. Bardou, D.; Zhang, K.; Ahmad, S.M. Lung sounds classification using convolutional neural networks. *Artif. Intell. Med.* **2018**, *88*, 58–69. [CrossRef]

42. Nguyen, T.; Pernkopf, F. Lung sound classification using snapshot ensemble of convolutional neural networks. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020; pp. 760–763.

43. Messner, E.; Fediuk, M.; Swatek, P.; Scheidl, S.; Smolle-Juttner, F.M.; Olschewski, H.; Pernkopf, F. Crackle and breathing phase detection in lung sounds with deep bidirectional gated recurrent neural networks. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 3–7 July 2018; pp. 356–359.

44. Perna, D.; Tagarelli, A. Deep auscultation: Predicting respiratory anomalies and diseases via recurrent neural networks. In Proceedings of the 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), Cordoba, Spain, 5–7 June 2019; pp. 50–55.

45. 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]

46. Messner, E.; Fediuk, M.; Swatek, P.; Scheidl, S.; Smolle-Jüttner, F.M.; Olschewski, H.; Pernkopf, F. Multi-channel lung sound classification with convolutional recurrent neural networks. *Comput. Biol. Med.* **2020**, *122*, 103831. [CrossRef] [PubMed]

47. Hadjileontiadis, L.J.; Rekanos, I.T. Detection of explosive lung and bowel sounds by means of fractal dimension. *IEEE Signal Process. Lett.* **2003**, *10*, 311–314. [CrossRef]

48. Sankur, B.; Kahya, Y.P.; Güler, E.Ç.; Engin, T. Comparison of AR-based algorithms for respiratory sounds classification. *Comput. Biol. Med.* **1994**, *24*, 67–76. [CrossRef] [PubMed]

49. Kahya, Y.P.; Yeginer, M.; Bilgic, B. Classifying respiratory sounds with different feature sets. In Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, USA, 30 August–3 September 2006; pp. 2856–2859.

50. 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]

51. Dorantes-Mendez, G.; Charleston-Villalobos, S.; Gonzalez-Camarena, R.; Chi-Lem, G.; Carrillo, J.; Aljama-Corrales, T. Crackles detection using a time-variant autoregressive model. In Proceedings of the 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, BC, Canada, 20–24 August 2008; pp. 1894–1897.

52. Henry, B.; Royston, T.J. Localization of adventitious respiratory sounds. *J. Acoust. Soc. Am.* **2018**, *143*, 1297–1307. [CrossRef] [PubMed]

53. Kompis, M.; Pasterkamp, H.; Wodicka, G.R. Acoustic imaging of the human chest. *Chest* **2001**, *120*, 1309–1321. [CrossRef] [PubMed]

54. Rao, A.; Huynh, E.; Royston, T.J.; Kornblith, A.; Roy, S. Acoustic methods for pulmonary diagnosis. *IEEE Rev. Biomed. Eng.* **2018**, *12*, 221–239. [CrossRef]

55. Rocha, B.M.; Pessoa, D.; Marques, A.; Carvalho, P.; Paiva, R.P. Automatic classification of adventitious respiratory sounds: A (un) solved problem? *Sensors* **2020**, *21*, 57. [CrossRef]

56. Pal, R.; Barney, A. A dataset for systematic testing of crackle separation techniques. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 4690–4693.

57. Charbonneau, G. Basic techniques for respiratory sound analysis. *Eur. Respir. Rev.* **2000**, *10*, 625–635.

58. Kiyokawa, H.; Greenberg, M.; Shirota, K.; Pasterkamp, H. Auditory detection of simulated crackles in breath sounds. *Chest* **2001**, *119*, 1886–1892. [CrossRef]

59. Earis, J.; Cheetham, B. Current methods used for computerized respiratory sound analysis. *Eur. Respir. Rev.* **2000**, *10*, 586–590.

60. Benesty, J.; Sondhi, M.M.; Huang, Y. (Eds.) *Springer Handbook of Speech Processing*; Springer: Berlin, Germany, 2008; Volume 1.

61. Vapnik, V.; Guyon, I.; Hastie, T. Support Vector Machines.

62. Aykanat, M.; Kılıç, Ö.; Kurt, B.; Saryal, S. Classification of lung sounds using convolutional neural networks. *Eurasip J. Image Video Process.* **2017**, *2017*, 1–9. [CrossRef]

63. Kochetov, K.; Putin, E.; Balashov, M.; Filchenkov, A.; Shalyto, A. Noise masking recurrent neural network for respiratory sound classification. In Proceedings of the International Conference on Artificial Neural Networks, Rhodes, Greece, 4–7 October 2018; pp. 208–217.

64. Liu, R.; Cai, S.; Zhang, K.; Hu, N. Detection of adventitious respiratory sounds based on convolutional neural network. In Proceedings of the 2019 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), Bangkok, Thailand, 21–24 October 2019; pp. 298–303.

65. 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 19th International Conference on Control, Automation and Systems (ICCAS), Jeju, Republic of Korea, 15–18 October 2019; pp. 804–807.

66. 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.

67. Ngo, D.; Pham, L.; Nguyen, A.; Phan, B.; Tran, K.; Nguyen, T. Deep learning framework applied for predicting anomaly of respiratory sounds. In Proceedings of the 2021 International Symposium on Electrical and Electronics Engineering (ISEE), Ho Chi Minh City, Vietnam, 15–16 April 2021; pp. 42–47.

68. Demir, F.; Ismael, A.M.; Sengur, A. Classification of lung sounds with CNN model using parallel pooling structure. *IEEE Access* **2020**, *8*, 105376–105383. [CrossRef]