

Complex audio signal data compression and reconstruction: A benchmark data pre-processing approach for machine classification of chronic respiratory diseases

DIGITAL HEALTH
Volume 10: 1–8
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076241302234
journals.sagepub.com/home/dhj



Timothy Albiges , Zoheir Sabeur  and Banafshe Arbab-Zavar

Abstract

Objective: To develop and evaluate innovative methods for compressing and reconstructing complex audio signals from medical auscultation, while maintaining diagnostic integrity and reducing dimensionality for machine classification.

Methods: Using the ICBHI Respiratory Challenge 2017 Database, we assessed various compression frameworks, including discrete Fourier transform with peak detection, time-frequency transforms, dictionary learning and singular value decomposition. Reconstruction quality was evaluated using mean squared error (MSE). The study has been conducted at Bournemouth University from January 2023 to 2024.

Results: The multi-resolution wavelet transform (MRWT) framework demonstrated superior performance with the lowest average MSE score of 0.037. The proposed time-frequency framework with MRWT achieved 80% accuracy in distinguishing chronic obstructive pulmonary disease from healthy samples.

Conclusion: Our study advances signal processing in medical auscultation, while it offers insights into effective compression and reconstruction methods for preserving diagnostic information. The MRWT approach shows promising outcomes for balancing compression efficiency and reconstruction accuracy in complex audio signals.

Keywords

Compressed sensing, dictionary learning, signal compression, signal reconstruction, complex signals, diagnostic integrity

Submission date: 17 June 2024; Acceptance date: 7 November 2024

Introduction

The compression and reconstruction of intricate audio signals in medical auscultation pose significant challenges in signal analysis due to the complex nature of auscultation audio. These signals encounter various noises, including background lung, heart and digestive organ sounds, as well as other internal and external sounds. Additionally, the non-stationary nature of breathing, combined with noises, further complicates the audio signal structure notwithstanding its analysis. This study explores novel methodologies for preserving the diagnostic integrity of respiratory sounds while reducing their data footprint. It aims to reconstructing audio signals from their features

and establishing a robust mapping between compressed features and original audio signals.

The motivation for this research stems from multiple critical needs in medical auscultation data processing. These include:

Department of Computing and Informatics, Bournemouth University,
Bournemouth, UK

Corresponding author:

Zoheir Sabeur, Department of Computing and Informatics, Bournemouth University, Talbot Campus, Poole House, Room 322, Fern Barrow, BH12 5BB, Dorset, England, UK.

Email: zsabeur@bournemouth.ac.uk



1. The necessity for dimensional reduction to enhance the performance of machine learning algorithms which are used for respiratory conditions classification.
2. A growing need for reduced lung sound data sizes to enable machine computations on small remote devices and facilitating point-of-care intelligent diagnostics in the future.

Our problem formulation centres on developing a compression framework that significantly reduces data size while preserving critical features necessary for the accurate classification of respiratory conditions, such as chronic obstructive pulmonary disease (COPD). This framework aims to balance data compression and diagnostic integrity, develop highly efficient machine-learning models and expand the potential for remote health monitoring and diagnosis in the near future.

The compression and reconstruction of pulmonary audio signals whilst preserving diagnostic information presents a complex challenge in medical signal processing. A comprehensive review of the literature reveals several key approaches and insights which have contributed to advancements in this domain.

Skalicky et al.¹ focused on the detection of respiratory phases in breath sounds. Whilst they do not directly address compression, they provide valuable insights into the temporal characteristics of respiratory signals. Their work underscores the importance of preserving phase information in any compression scheme, as they are crucial for accurate diagnosis. This study highlights the need for compression methods to retain temporal features alongside spectral information.

Liu et al.² employed Convolutional Neural Networks (CNNs) for detecting adventitious respiratory sounds. Although their primary focus was on machine classification rather than compression, their work demonstrates the potential of adopting deep learning techniques in extracting salient features from respiratory audio signals. It also suggests that neural network-based approaches could be explored, while dimensionality reduction under compression schemes, could offer a balance between data reduction and preservation of diagnostic information.

Charleston-Villalobos et al.³ applied Empirical Mode Decomposition (EMD) to analyse crackle sounds. EMD decomposition of signals into intrinsic mode functions could be leveraged in compression schemes, particularly for capturing non-stationary and nonlinear aspects of respiratory sounds. This approach might be beneficial for preserving information about transient events like crackles, which are crucial for diagnosis.

İçer and Gençç⁴ focused on the classification and analysis of non-stationary characteristics of crackle and rhonchus lung adventitious sounds. Their work emphasises the importance of capturing time-varying features in respiratory audio and suggests that effective compression

methods must adapt to the non-stationary nature of these signals. This study reinforces the need for dynamic and adaptive compression techniques that preserve signal spectral characteristics such as their temporal evolution.

Li and Yi⁵ explored feature extraction of lung sounds based on bispectrum analysis. It captures the non-linear interactions between frequency components and could be valuable for developing compression methods. This approach might be beneficial for compressing signals with subtle harmonic structures that are diagnostically significant.

Aras and Gangal⁶ compared different features derived from Mel Frequency Cepstrum Coefficients (MFCCs) for lung sound classification. Their work is relevant to compression strategies as it highlights the effectiveness of cepstral analysis in capturing perceptually relevant aspects of audio signals. The incorporation of MFCC-like features in compression schemes could lead to dimensionality reduction that aligns well with human auditory perception and, by extension, clinical diagnostic practices.

Oletic and Bilas⁷ made significant strides in applying compressive sensing techniques to detect asthmatic wheezes from respiratory sound spectra. Their work addresses the challenge of compressing respiratory audio signals while maintaining diagnostic accuracy. It demonstrates the feasibility of detecting specific respiratory conditions from compressed spectra and provides guidance for future research in pulmonary audio compression.

Sakai et al.⁸ explored sparse representation-based extraction of pulmonary sound components from low-quality auscultation signals. Their approach is particularly relevant to the compression challenge, as sparse representations inherently provide a form of data reduction. The success of this method in extracting subtle pulmonary sound components suggests that sparsity-based compression techniques could be highly effective in preserving diagnostic information while significantly reducing data dimensionality.

In summary, the literature reveals diverse approaches to analysing and processing respiratory audio signals, as shown in Table 1. While many of these studies do not directly address compression, they offer valuable insights into the characteristics of pulmonary sounds that must be preserved for accurate diagnosis. The challenge lies in synthesising these various approaches – wavelet analysis, EMD, Bispectrum analysis, cepstral analysis, compressive sensing and sparse representations – into a coherent compression framework that can potentially and effectively reduce data dimensionality while retaining the complex and non-stationary nature of audio signals, for achieving diagnostic accuracies with performing machine learning classifiers.

Our research addresses critical gaps in the existing pulmonary auscultation signal processing literature. Firstly, more comprehensive compression techniques must be tailored explicitly for pulmonary auscultation signals. These unique characteristics distinguish them from other audio

signals that contain adventitious lung sounds like transient crackles and harmonic wheezes. Secondly, the exploration of advanced time-frequency transforms in conjunction with dictionary learning for medical audio compression has been limited despite the potential of these techniques to capture the complex temporal and spectral features of respiratory sounds. Thirdly, there needs to be more focus on preserving diagnostic integrity during compression, which is a crucial aspect for maintaining the clinical relevance of the compressed signals for diagnostics.

To address these gaps, our approach introduces several novel elements. We integrate multi-resolution wavelet transform (MRWT) with dictionary learning. The combined techniques lead to efficient signal compression whilst they preserve the salient features of the auscultation signals. Indeed, this integration allows for a more nuanced signal representation across different time scales and frequencies. Furthermore, we have developed a framework that carefully balances compression efficiency with reconstruction accuracy, ensuring that the compressed signals retain their diagnostic value. This balance is highly critical in medical applications where data fidelity directly impacts diagnostic outcomes. Finally, our study extends beyond signal processing by evaluating the impact of compression on subsequent machine-learning classification of respiratory conditions. This holistic approach assures that the compressed signals maintain their integrity and remain suitable for advanced diagnostic algorithms, while bridging the gap between signal processing and clinical application.

Methodology

This study, conducted at Bournemouth University from January 2023 to half of 2024, utilised the open-access ICBHI Respiratory Challenge 2017 Database,⁹ comprising 920 respiratory sound samples. Our research methodology encompassed two primary phases: Signal compression and reconstruction. The latter is performed with the aim of preserving diagnostic integrity whilst reducing data size.

The signal compression phase evaluated four distinct methods, each designed to address specific aspects of audio signal processing in medical auscultation. Firstly, we implemented the fast Fourier transform (FFT) with peak detection to assess whether dominant frequencies alone could sufficiently capture the essential characteristics of pulmonary audio for accurate reconstruction. This method served as a baseline, allowing us to evaluate the efficacy of more sophisticated approaches.

Secondly, and as part of the core of our investigation, we developed two time-frequency transform methods coupled with dictionary learning and singular value decomposition (SVD). These methods, which utilise short-time Fourier transform (STFT) and MRWT, respectively, were designed to capture both temporal and spectral features of the audio signals. The integration of dictionary learning and SVD

aimed to achieve efficient compression whilst preserving crucial diagnostic information.

Lastly, we examined the effectiveness of MRWT combined with SVD, while we omitted the dictionary learning step. This method was specifically included to evaluate the impact of dictionary learning on the overall compression and reconstruction process.

Signal compression

Time-frequency transforms with dictionary learning and SVD.

Our primary focus was on two time-frequency transform methods, the STFT (Method A) and the MRWT (Method B), which are integrated into a framework with dictionary learning and SVD. The time-frequency allows for the frequency to change over time.

The STFT, computed as:¹⁰

$$X(f, \tau) = \sum_{n=0}^{N-1} x(n) \cdot \omega(n - \tau R) \cdot e^{-i2\pi f n}$$

where ω is the window function, τ is the frame index, R is the hop length, $x(n)$ is the input signal and f is the frequency, providing a localised frequency analysis over time.

The MRWT, calculated as:¹¹

$$W_n(\tau, s) = \sum_{n=0}^{N-1} x_n \psi \left[\frac{n - \tau}{s} \right] dt$$

where s is the scaling factor, and τ is the translation function, it utilises a scaling window to capture both global trends and local nuances of the audio signal.^{11,12} ψ is the mother wavelet, in this case, the Complex Morlet Wavelet.

Dictionary learning, integral to STFT and MRWT methods, acquires sparse dictionaries that efficiently represent the signal characteristics. It is formulated as:¹³

$$\min_{D, \{\alpha_i\}} \sum_{i=1}^N \left(\frac{1}{2} \|x_i - D\alpha_i\|_2^2 + \|\lambda\alpha_i\|_1 \right)$$

where D is the dictionary matrix, and α_i are the coefficient vectors. The representation by these atoms efficiently captures the signal using orthogonal matching pursuit (OMP) with a small number of non-zero coefficients. This implies that only a few atoms are needed to represent the audio time-frequency matrix.¹⁴ In Equation 4, the objective function is minimised with respect to two main variables: the dictionary matrix D and the coefficient vectors α_i . The function is comprised of two terms: the reconstruction error term, and the sparsity-promoting regularisation term. The reconstruction error term, $\frac{1}{2} \|x_i - D\alpha_i\|_2^2$, measures the fidelity of the approximation obtained by representing the input signal x_i as a linear combination of atoms from the dictionary D . Here, $\|x_i - D\alpha_i\|_2^2$ denotes the squared L2 norm, quantifying the difference between the original

signal and its approximation. The second term, $\lambda\|\alpha_i\|_1$, encourages sparsity in the coefficient vectors α_i by penalising large coefficient magnitudes. The parameter λ controls the trade-off between fidelity and sparsity, influencing the degree of compression achieved during the dictionary learning. Overall, this equation encapsulates the essence of dictionary learning, wherein an efficient dictionary and sparse coefficients are jointly learned to represent the input signals optimally.

Finally, SVD is a mathematical technique that decomposes matrices, which can be non-square, into three main components: The principal component vectors and the singular values, as $A = UDV^T$.¹⁵ In this context, the matrices U and V^T capture distinct patterns within the input matrix. Specifically, the matrix U delineates patterns among the rows of the input matrix, while the matrix V^T encapsulates patterns among the columns. D are the singular values representing the strengths of the U and V^T column vectors.

DFT and peak detection. Method C uses the FFT to swiftly convert audio signals from the time domain to the frequency domain, revealing the spectral landscape.¹⁶ This transformation is formulated as follows:

$$X_k = \frac{1}{N} \sum_{n=0}^{N-1} x_n e^{-2\pi i k n / N}$$

where k represents the frequency index, following the FFT, peak detection identifies and selects dominant frequencies, aiming to capture the most salient features of the frequency domain whilst discarding less significant information.

MRWT with SVD. This method D, excluding dictionary learning, was implemented to isolate and evaluate the specific contributions of dictionary learning to the compression process. It follows the same MRWT procedure described above, followed directly by SVD.

By systematically comparing these methods, we aimed to determine the most effective approach for compressing pulmonary audio signals whilst maintaining their diagnostic integrity. This comprehensive methodology allowed us to evaluate the trade-offs between compression efficiency and signal fidelity, which is crucial for advancing the field of medical auscultation signal processing.

This study looks at four methods, with a proposed framework that consists of the following steps:

1. Signal preprocessing (normalisation and downsampling)
2. Time-frequency transform with STFT
Method A: TF transform STFT; or Method B: TF transform MRWT
3. Dictionary learning for sparse representation
4. Singular value decomposition for further dimensionality reduction
5. Signal reconstruction using inverse transforms

6. Evaluation using mean squared error (MSE) metric

Additionally, we tested two alternative approaches for comparison:

1. Method C: FFT with dominant peak detection (dominate frequency):
 - (a) Apply fast Fourier transform to the pre-processed signal
 - (b) Identify and select dominant frequency peaks
 - (c) Reconstruct the signal using only these dominant peaks
2. Method D: Proposed framework without dictionary learning (MRWT and SVD):
 - (a) Follow steps 1, 2, 4, 5 and 6 of the main framework, while omitting the dictionary learning step

These alternative approaches were evaluated alongside our main framework in order to assess each component impact on signal compression and reconstruction quality. The comparison allows us to quantify the benefits of including dictionary learning, while using more advanced time-frequency transforms over traditional FFT-based methods.

Signal reconstruction

The signal reconstruction in our proposed methodology meticulously restores compressed audio signals to their original form while preserving crucial diagnostic information. The reconstruction phase is vital for maintaining the integrity and, therefore, the important clinical relevance of the audio signals.

The inverse fast Fourier transform (iFFT) is initially applied to transform frequency-domain signals back into the time domain, effectively reconstructing the original audio signals. Furthermore, signals obtained from Time-Frequency Transforms undergo inverse transforms. These transformations convert time-frequency domain representations into time-domain signals whilst achieving the faithful reconstruction of the original signal.

Additionally, dictionary learning and SVD are employed, requiring matrix multiplication to reconstruct compressed signals. In dictionary learning, sparse dictionaries obtained during compression represent the original data via sparse linear combinations of learned dictionary atoms. Similarly, SVD components, including principal component vectors and singular values, reconstruct signals through matrix multiplication.

Finally, the reconstruction quality is assessed using the MSE criterion, which computes the differences between reconstructed and original audio signals. Lower MSE values indicate superior fidelity in preserving the original signal characteristics.

Table 1. Comparison of state-of-the-art respiratory sound analysis, highlighting their objectives and limitations.

Study	Objective	Limitations
Skalicky et al. ¹	Detection of respiratory phases in breath sounds	Limited to phase detection, not focused on compression
Liu et al. ²	Detection of adventitious respiratory sounds using CNN	Focused on classification, not on signal compression
Charleston-Villalobos et al. ³	Crackle sounds analysis by empirical mode decomposition	Limited to crackle analysis, not general compression
İçer and Gengeç ⁴	Classification and analysis of non-stationary characteristics of crackle and rhonchus lung sounds	Focused on specific sound types, not on compression
Li and Yi ⁵	Feature extraction of lung sounds based on bi-spectrum analysis	Limited to feature extraction, not compression
Aras and Gangal ⁶	Comparison of features derived from Mel Frequency Cepstrum Coefficients for lung sound classification	Focused on feature comparison, not on compression techniques
Oletic and Bilas ⁷	Asthmatic wheeze detection from compressed and reconstructed spectra	Limited to wheeze detection, not general compression
Sakai et al. ⁸	Extraction of pulmonary sound components	Focused on noise reduction, not comprehensive compression

Table 2. Signal reconstruction comparison to original audio.

Method	MSE score			
	Min	Mean	Median	Max
Method A: TF framework STFT	9.5252×10^{-4}	1.0178×10^{-1}	7.5069×10^{-2}	9.3763×10^{-1}
Method B: TF framework MRWT	1.1009×10^{-3}	3.6894×10^{-2}	2.9774×10^{-2}	2.3373×10^{-1}
Method C: dominate frequencies	5.9575×10^{-2}	1.9902×10^{-1}	1.8230×10^{-1}	7.8066×10^{-1}
Method D: MRWT and SVD	6.4706×10^{-1}	1.1391×10^{-1}	8.9350×10^{-2}	8.9497×10^{-1}

Results

Our findings reveal that the DFT coupled with peak detection failed to reconstruct the signal accurately. In contrast, the STFT demonstrated promising performance with an average MSE score of 0.199. Notably, the framework incorporating the MRWT achieved the lowest average MSE score of 0.037. The full results are shown in Table 2.

The observed outcomes can be attributed to several factors. Although widely used, the FFT may struggle to capture complex signal dynamics accurately, especially in the presence of noise and overlapping components. The

success of STFT can be credited to its ability to provide localised frequency analysis, which may be better suited for capturing the temporal variations in the original signal. However, the superior performance of MRWT could be attributed to its adaptive nature, allowing it to effectively capture both global trends and local fluctuations within the signal.

The box plot in Figure 1 visually illustrates the MSE scores for the four methods. The top and bottom of each box represent the upper and lower quartiles, showing the interquartile range where the middle 50% of the data lies. The red line inside each box indicates the median. The whiskers extend to the smallest and largest values

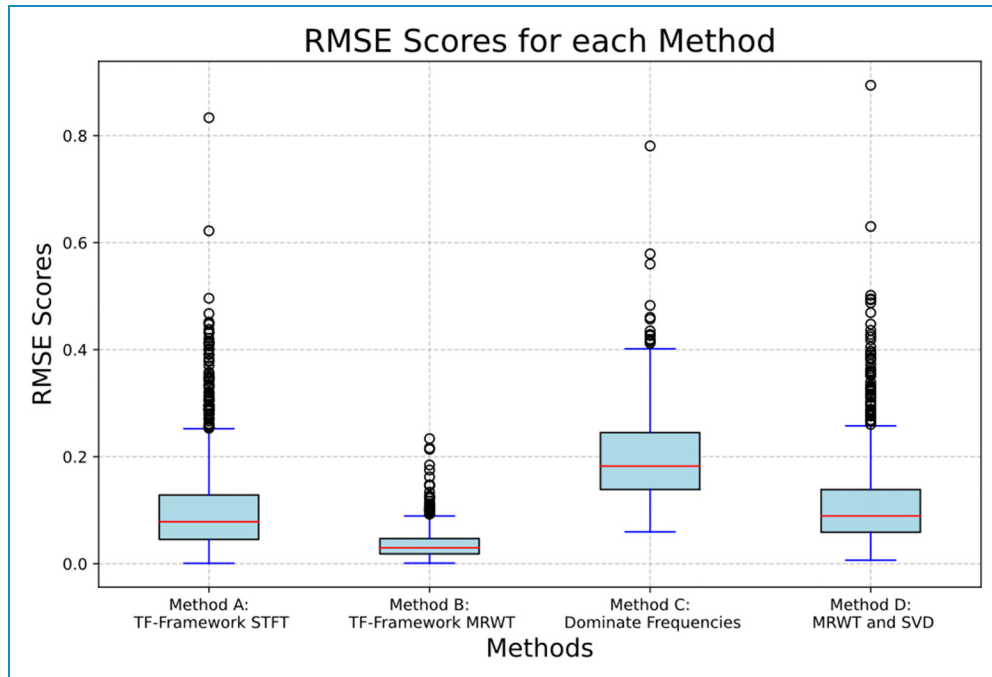


Figure 1. A box plot of the results of each method's MSE distribution.

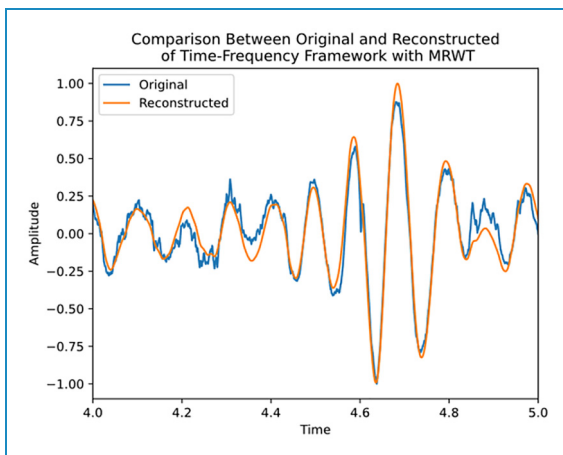


Figure 2. Signal comparison of original and reconstructed using TF framework with MRWT, method B.

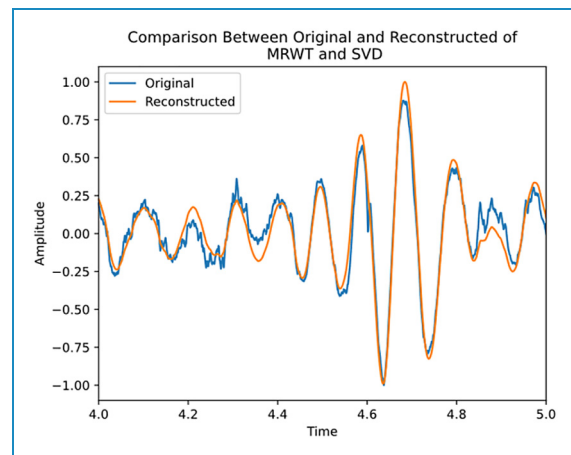


Figure 3. Signal comparison of original and reconstructed using MRWT and SVD, method D.

within 1.5 times the interquartile range from the quartiles. Outliers are shown as circles. This plot provides a visual summary of each method's distribution and variability of MSE scores.

The zoomed-in view presented in Figures 2 and 3 underscores the precision of MRWT in capturing subtle signal details, further validating its effectiveness in signal reconstruction in both the MRWT and SVD and the TF framework with MRWT. However, the FT framework with STFT suffered from phase alignment in the

rebuilding, highlighted in Figure 4, and that of the variance of amplitudes in Figure 5. The TF framework with MRWT, incorporating dictionary learning, offers the advantage of compressing data to a significantly smaller number of features, totalling 252, compared to a higher dimensionality of 448,118 feature points without dictionary learning. Furthermore, despite the reduction in feature space, the reconstruction accuracy is not only as good but slightly increased, as indicated by a higher MSE score.

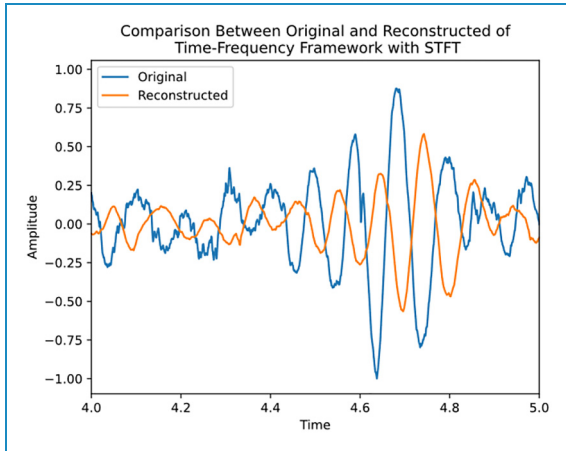


Figure 4. Signal comparison of original and reconstructed using TF framework with STFT, method A.

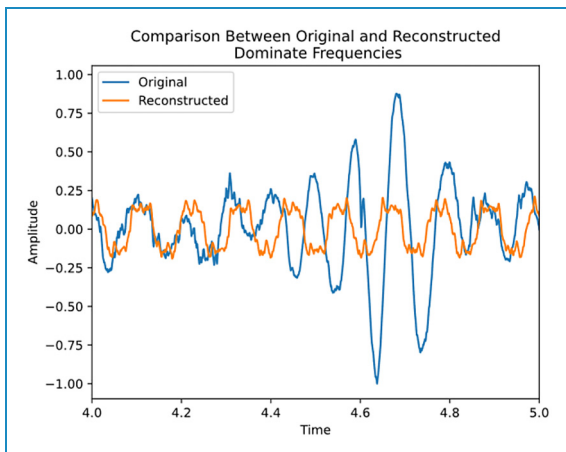


Figure 5. Signal comparison of original and reconstructed using dominate features, method C.

Discussion

The compression and reconstruction of complex audio signals in medical auscultation present significant challenges due to the intricate nature of these signals. Our research addresses this issue by exploring innovative methodologies that balance compression efficiency with the preservation of diagnostic integrity. The compression phase reduces signal complexities while retaining essential diagnostic insights, ensuring that the reconstructed audio closely aligns with the original data. We evaluated several compression frameworks, starting with the discrete Fourier transform (DFT) and peak detection, followed by the MRWT combined with SVD and finally, various Time-Frequency Transform methods, including STFT and MRWT with dictionary learning and SVD. The reconstruction process restores compressed audio signals to their original form, preserving critical diagnostic information through inverse transforms, matrix

multiplications and evaluations using the MSE metric. Our results indicated varying performance among the methods, with the time-frequency framework incorporating MRWT demonstrating superior performance through the lowest MSE scores. The adaptability and resilience of MRWT and dictionary learning in capturing global trends and local fluctuations contribute to their effectiveness.

Previous studies underscore the significance of signal compression methodologies in enhancing the diagnostic value of auscultation audio signals. Oletic and Bilas⁷ demonstrated the potential for accurately detecting asthmatic wheezing from respiratory sound spectra using compressive sensing techniques. This research highlights the feasibility of leveraging compressed feature requirements to efficiently process large volumes of audio data without compromising diagnostic accuracy. Furthermore, it underscores the importance of exploring innovative signal-processing techniques to extract clinically relevant information from noisy and complex audio signals containing various adventitious lung sounds. Similarly, Sakai et al.⁸ elucidated the implications of sparse representation-based extraction of pulmonary sound components from low-quality auscultation signals. Their findings suggest sparse representation methods offer a robust approach to extracting subtle pulmonary sound components, such as vesicular sounds and crackles, from noisy recordings. This underscores the potential for advanced signal processing techniques to improve the accuracy and efficiency of respiratory system diagnosis, paving the way for developing more sophisticated diagnostic tools in clinical practice.

Our research underscores the importance of selecting appropriate signal-processing techniques to balance compression efficiency with reconstruction accuracy. The MRWT time-frequency framework emerges as a promising approach for signal reconstruction in complex audio signals, warranting further exploration to enhance diagnostic accuracy and computational efficiency. This strategy enabled us to train machine learning classifiers for chronic respiratory conditions confidently. Using our proposed time-frequency framework with MRWT, we achieved an accuracy of 80% and an F1 score of 78.5% in distinguishing COPD from healthy samples. For the more complex task of differentiating COPD from healthy and pneumonia cases, we obtained an accuracy of 70% with an F1 score of 70%. These results demonstrate our compression method's effectiveness, which reduces data size while preserving diagnostically relevant information. Albiges et al.¹⁷ detailed the full classification results, validating our signal compression methodology's practical usefulness in distinguishing COPD from other respiratory conditions.

Conclusion

In conclusion, our study introduces novel methods for compressing and reconstructing auscultation audio signals, aiming to preserve diagnostic accuracy while reducing

storage, transmission needs and classification dimensionality. By capturing and reconstructing essential audio signal information using data from the ICBHI Respiratory Challenge 2017 Database, we aimed to differentiate COPD from other respiratory conditions effectively. Our findings underscore the importance of selecting signal-processing techniques that balance accuracy and computational efficiency. Further research is needed to explore broader clinical applications and refine classification algorithms to improve diagnostic and prognostic accuracy in medical auscultation.

Acknowledgements: The authors acknowledge the support of Bournemouth University, Department of Computing and Informatics throughout the deployment and conduct of this research work.

Contributorship: Timothy Albiges authored the paper and conducted the experiment. Professor Zoheir Sabeur and Dr Banafshe Arbab-Zavar interpreted the results and reviewed the manuscript.


Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Ethical approval: Ethical Considerations and Patient Consent This study utilised the ICBHI Respiratory Challenge 2017 Database (Rocha et al., 2019), an open-access dataset. The original data collection and publication were conducted in accordance with the Declaration of Helsinki, and the Ethics approved the protocol in the Research Committee of the School of Health Sciences and Technologies, Polytechnic Institute of Porto, Portugal. As per the dataset's documentation, all participants provided informed consent for their data to be used for research purposes. The database was anonymised before public release, ensuring patient privacy and confidentiality. Our study, which involves secondary analysis of this publicly available dataset, does not require additional patient consent. However, we have used the data responsibly and in compliance with the original ethical approval and consent of the dataset creators. Our research was conducted under the ethical approval granted by the Bournemouth University Ethical Board (reference number 40455), which covers using this public dataset for our specific research objectives.

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was conducted with no external funding. The lead author, Timothy Albiges, is a self-funded postgraduate PhD student at Department of Computing and Informatics, Bournemouth University.

Guarantor: The lead author, Timothy Albiges, serves as the guarantor for the content of this paper.

ORCID iDs: Timothy Albiges  <https://orcid.org/0000-0002-0202-8862>

Zoheir Sabeur  <https://orcid.org/0000-0003-4325-4871>

References

- Skalicky D, Koucky V, Hadraba D, et al. Detection of respiratory phases in a breath sound and their subsequent utilisation in a diagnosis. *Appl Sci* 2021; 11: 6535.
- Liu R, Cai S, Zhang K, et al. Detection of adventitious respiratory sounds based on convolutional neural network. In 2019 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), 2019; pp.298–303. 21–24 November 2019, Shanghai, China. IEEE
- Charleston-Villalobos S, González-Camarena R, Chi-Lem G, et al. Crackle sounds analysis by empirical mode decomposition. *IEEE Eng Med Biol* 2007; 26: 40–47.
- İçer S and Genç Ş. Classification and analysis of non-stationary characteristics of crackle and rhonchus lung adventitious sounds. *Digit Signal Process* 2014; 28: 18–27.
- Li S and Yi L. Feature Extraction of Lung Sounds Based on Bispectrum Analysis. In: *2010 Third Int Symposium Information Process*, 2010, pp.393–397.
- Aras S and Gangal A. Comparison of Different Features Derived from Mel Frequency Cepstrum Coefficients for Classification of Single Channel Lung Sounds. In: *2017 40th Int Conf Telecommun Signal Process Tsp*, 2017, pp.346–349.
- Oletic D and Bilas V. Asthmatic wheeze detection from compressively sensed respiratory sound Spectra. *IEEE J Biomed Health* 2018; 22: 1406–1414.
- Sakai T, Satomoto H, Kiyasu S, et al. Sparse representation-based extraction of pulmonary sound components from low-quality auscultation signals. In: *2012 IEEE Int Conf Acoust Speech Signal Process*, 2012, pp.509–512.
- Rocha BM, Filos D, Mendes L, et al. An open access database for the evaluation of respiratory sound classification algorithms. *Physiol Meas* 2019; 40: 035001.
- Beyerbach D and Nawab H. Principal components analysis of the short-time Fourier transform. In: *Proc ICASSP 91 1991 Int Conf Acoust Speech Signal Process*, vol.3, 1991, pp.1725–1728.
- Torrence C and Compo GP. A practical guide to wavelet analysis. *Bull Am Meteorol Soc* 1998; 79: 61–78.
- Du M, Chan FHY, Lam FK, et al. Crackle detection and classification based on matched wavelet analysis. In *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1997; 4: 1638–1641.
- Kreutz-Delgado K, Murray JF, Rao BD, et al. Dictionary learning algorithms for sparse representation. *Neural Comput* 2003; 15: 349–396.
- Mairal J, Bach F, Ponce J, et al. Online dictionary learning for sparse coding. In: *Proc 26th Annu Int Conf Mach Learn – ICML '09*, 2009, pp.689–696.
- Fu J, Yuan H, Zhao R, et al. Clustering K-SVD for sparse representation of images. *EURASIP J Adv Sig Process* 2019; 2019: 47.
- Cooley JW and Tukey JW. An algorithm for the machine calculation of complex Fourier series. *Math Comput* 1965; 19: 297–301.
- Albiges T, Sabeur Z and Arbar-Zavar B. Compressed sensing data with performing audio signal reconstruction for the intelligent classification of chronic respiratory diseases. *Sensors* 2023; 1439. doi: 10.3390/s23031439