A Wavelet-Based Approach to Rapidly Identify Drug-Addicted Individuals Using Voice Signal Analysis

Sajeeb Hassan[1], Sadia Afrin[2*], Md. Rafiqul Islam[1]

Abstract

Recognizing and classifying signals plays a pivotal role across diverse disciplines, such as classification, pattern recognition, data preprocessing, and predictive science. This study specifically investigates the use of Haar and Symlet (Sym2) wavelet transforms in analyzing short voice signals to distinguish between drug-addicted and non-addicted individuals. The primary aim is to uncover distinctive features in voice signals that correlate with addiction. Haar and Symlet wavelet transforms are employed to process substantial segments of speech signals, extracting meaningful insights that can aid law enforcement agencies and addiction researchers. Each signal undergoes visualization and analysis to unveil patterns and discrepancies brought to light by different wavelet transformations. Following the wavelet transformation process, the study evaluates the Peak Signal-to-Noise Ratio (PSNR) and Signal-to-Noise Ratio (SNR) using MATLAB's wavelet toolbox. These metrics serve as crucial indicators for decision-making processes aimed at identifying drug-addicted individuals based on unique voice signal characteristics. Beyond advancing signal processing techniques, this research aims to have practical applications in recognizing behavioral patterns associated with addiction. Ultimately, the goal is to leverage these insights to develop more effective intervention and support strategies within relevant communities and healthcare settings. By integrating wavelet analysis with voice signal processing, this study not only enhances our understanding of addiction-related behaviors but also contributes valuable tools for real-world applications in law enforcement and healthcare sectors. These efforts are geared towards fostering more targeted and efficient responses to addiction issues in society.

Keywords: Wavelet Transform, Voice Signal Analysis, Addiction Detection, Signal Processing, Behavioral Patterns

Significance | Wavelet transforms and analyses voice signals to distinguish addiction, aiding intervention strategies and law enforcement with effective behavioural pattern recognition.

*Correspondence. Sadia Afrin, Department of Basic Science, Primeasia University, Dhaka, Bangladesh.
E-mail: sadia.afrin@primeasia.edu.bd

Editor A. B. M. Abdullah, Ph. D.,, And accepted by the Editorial Board Jun 15, 2022 (received for review Apr 05, 2022)

Introduction

Drug addiction poses a critical challenge to public health worldwide, affecting individuals from diverse socio-economic backgrounds and cultures (Thomas, 2019). Addiction, defined as a chronic condition characterized by compulsive substance use despite harmful consequences, can involve legal substances such as alcohol and tobacco, as well as illicit drugs like marijuana and opioids (Lessem, 2020). According to the World Health Organization (WHO), approximately 185 million people worldwide are involved in illegal drug use, with over two billion people consuming alcohol and 1.3 billion smokers globally (Health Street, 2022). The impact of addiction is profound, contributing

Author Affiliation.
[1] Mathematics Discipline, Khulna University, Khulna, Bangladesh.
[2] Department of Basic Science, Primeasia University, Dhaka, Bangladesh.

Please cite this article.
Sajeeb Hassan, Sadia Afrin et al., (2022). A Wavelet-Based Approach to Rapidly Identify Drug-Addicted Individuals Using Voice Signal Analysis, Journal of Primeasia, 3(1), 1-8, 30036

significantly to morbidity and mortality through various health issues such as liver cirrhosis, cardiovascular diseases, cancers, and respiratory conditions (Akanbi, 2018; Ming & Li, 2010; V. M. Anantha Eashwar, 2019). While addiction is traditionally diagnosed through behavioral assessments and biochemical tests, there is a growing interest in leveraging technological innovations for the rapid identification and classification of addicted individuals (Mohiuddin, 2019). Conventional methods such as urine, blood, and hair tests are widely used but present limitations in terms of cost, time, and the need for specialized equipment and personnel (Islam, 2011). These methods, although accurate, can be invasive and time-consuming, which has motivated researchers to explore non-invasive and rapid detection techniques (Coates, Hero, Nowak, & Yu, 2002). One promising avenue is the analysis of voice signals, which could serve as a potential biomarker for drug addiction (Abry, Baraniuk, Flandlin, Riedi, & Veitch, 2002).

Voice signal analysis has garnered attention due to its noninvasive nature and the accessibility of digital audio data, which can be collected in real time (Riedi, Crouse, Ribeiro, & Baraniuk, 1999). Speech is influenced by multiple physiological factors, including cognitive and emotional states, which can be altered by substance use (Donoho, 1995). Voice patterns, pitch, frequency, and tone can exhibit significant changes due to prolonged drug use, as substances like alcohol and marijuana impact the central nervous system, affecting vocal cord function and speech production (Donoho & Johnstone, 1994).

In this context, wavelet-based analysis techniques offer a sophisticated approach to signal processing, enabling the extraction of relevant features from voice data that may not be easily discernible through traditional methods (Daubechies, 1990). Wavelet transforms, particularly Haar and Symlet wavelets, allow for the decomposition of voice signals into multiple frequency components, providing a multi-resolution analysis of the signal (Shin, Powers, Grady, & Bhatt, 1999). This makes it possible to detect subtle changes in voice patterns that may indicate addiction (Harfoush, Bestavros, & Byers, 2001).

This study aims to explore the potential of wavelet-based voice signal analysis for the rapid identification of drug-addicted individuals (Katabi, Bazzi, & Yang, 2001). By applying Haar and Symlet (Sym2) wavelet transforms to short voice samples, we aim to extract distinctive features that differentiate addicted individuals from non-addicted ones (Moon, Skelly, & Towsley, 1999). The voice data, collected from individuals with known alcohol and marijuana consumption histories, is analyzed using MATLAB’s wavelet toolbox to compute key metrics such as Peak signal-to-noise ratio (PSNR) and signal-to-noise ratio (SNR) (Zhang, Duffield, Paxson, & Shenker, 2001). These metrics provide critical insights into the signal quality and are used to develop a classification model that could aid in the quick identification of addicted individuals (Rubenstein, Kurose, & Towsley, 2002).

This research not only seeks to advance the field of signal processing but also holds significant implications for public health, law enforcement, and addiction management (Trees, 1968). By offering a rapid and non-invasive method for detecting addiction, this study could contribute to the development of new intervention strategies, improve the efficiency of addiction screening programs, and support law enforcement efforts in addressing substance abuse-related offenses (Williams, 1991). Ultimately, the integration of voice signal analysis with wavelet-based processing could become a valuable tool in the ongoing efforts to combat addiction globally (Huang, Feldmann, & Willinger, 2001).

Materials and Methods

This study aimed to identify drug-addicted individuals through voice signal analysis using wavelet rapid transforms (Islam, 2011). The research focused on individuals with prolonged substance use, specifically alcohol and marijuana (Beena Elizabeth Thomas, 2019; Mohiuddin, 2019). Voice data were collected from publicly available YouTube recordings of individuals over the age of 25 who had consumed alcohol or marijuana beyond the legal limits for more than five years (Lessem & Shah, 2020). In total, 32 voice samples were analyzed: 16 from drug-addicted individuals and 16 from non-addicted individuals (Samuel Asare, 2019).

To begin, the voice samples were converted from video to mp3 audio format. From each recording, only five seconds of speech were extracted to ensure consistency (Maxwell Oluwole Akanbi, 2018). The data preprocessing involved denoising and smoothing the raw voice signals to remove any background noise and irrelevant data. Once the voice signals were cleaned, they were transformed using two distinct wavelet transformations: the Haar wavelet and Symlet (Sym2) wavelet (Daubechies, 1990; Donoho, 1995). These wavelet transformations decompose the voice signals into frequency components, which enabled a detailed analysis of voice features related to addiction (Donoho & Johnstone, 1994).

The transformed signals were then analyzed using the power spectrum to identify distinctive patterns between addicted and non-addicted individuals. The power spectrum graph was used to compare the signal frequencies and their corresponding power levels (in decibels) (Harfoush, Bestavros, & Byers, 2000). Additionally, Peak signal-to-noise ratio (PSNR) and signal-to-noise ratio (SNR) were calculated for each voice sample using MATLAB's wavelet toolbox (Riedi, Crouse, Ribeiro, & Baraniuk, 1999). PSNR and SNR served as key performance metrics to assess the quality of signal transformations and detect patterns indicative of addiction. The higher PSNR and SNR values were used to determine the clarity and noise levels in the voice signals (Zhang, Duffield, Paxson, & Shenker, 2001).

Table 1. Use of Calculating Peak Signal to Noise Ratio (PSNR) and Signal to Noise Ratio (SNR) values

Name of Wavelet	people	PSNR Value	SNR Value
Haar	Addicted	36.60-37.10	25.20-26.30
	Non-addicted	34.70-35.20	21.40-22.30
Sym2	Addicted	34.50-35.90	22.30-23.70
	Non-addicted	32.25-33.90	19.10-20.80

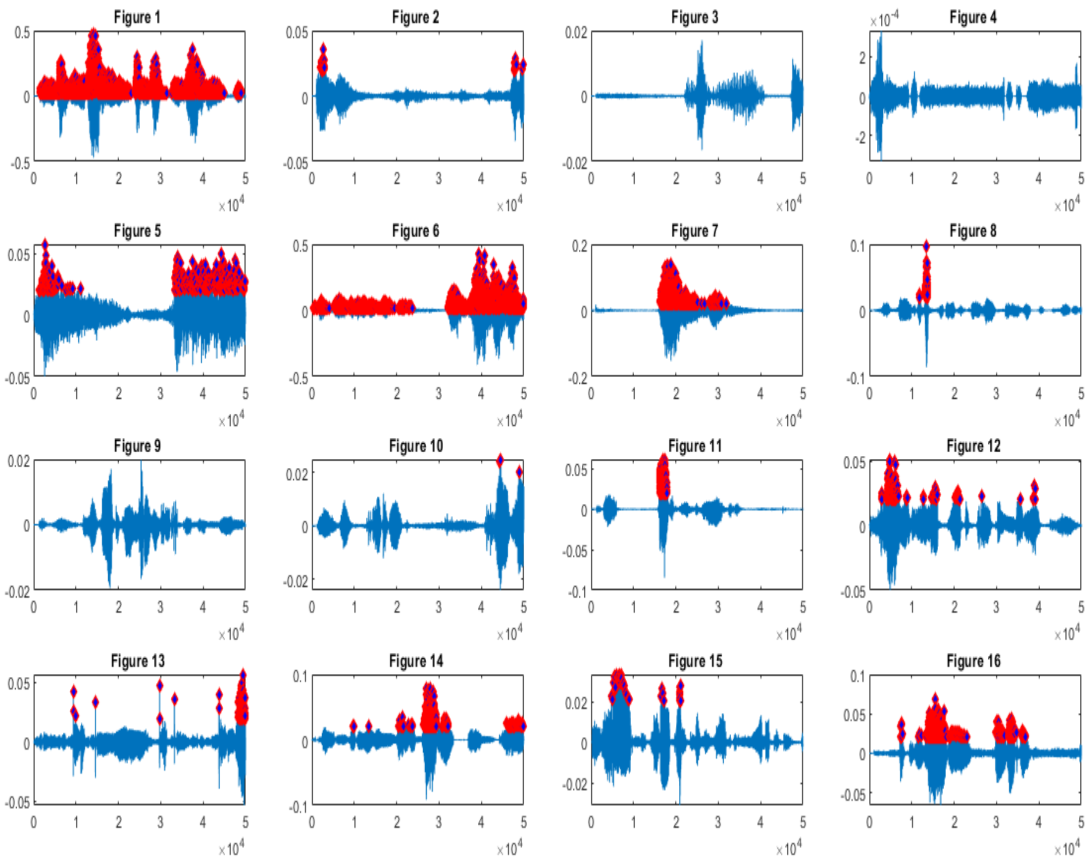

Figure 1. Power spectrum graph using Haar wavelet at level 2 Sym2 Wavelet

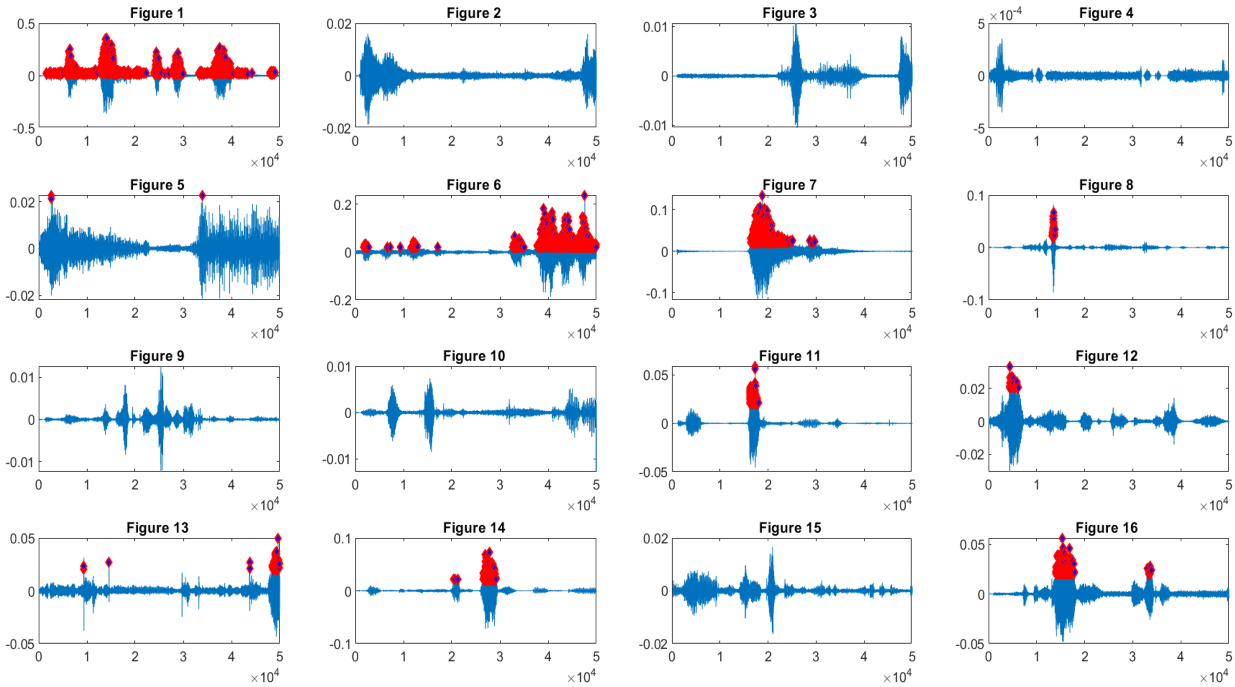

Figure 2. Power spectrum graph using sym2 wavelet at level 2.

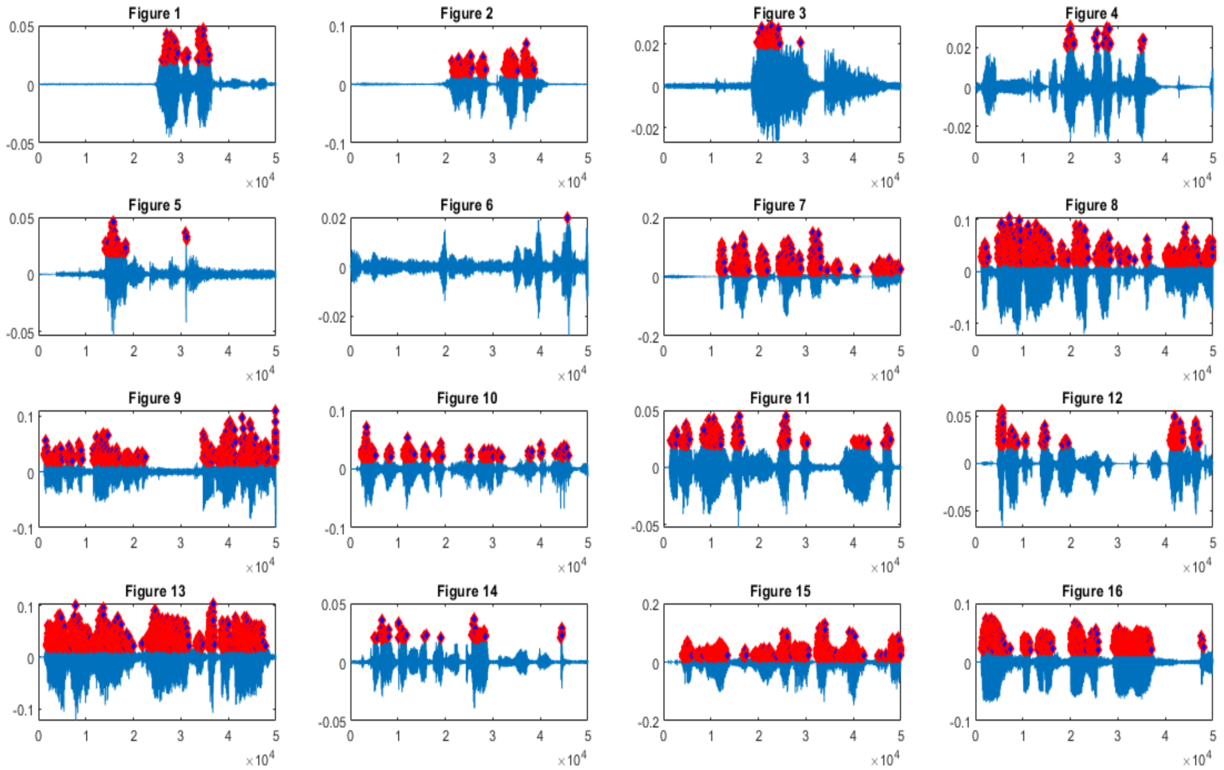

Figure 3. Power spectrum graph using Haar wavelet at level 2 Sym2 Wavelet.

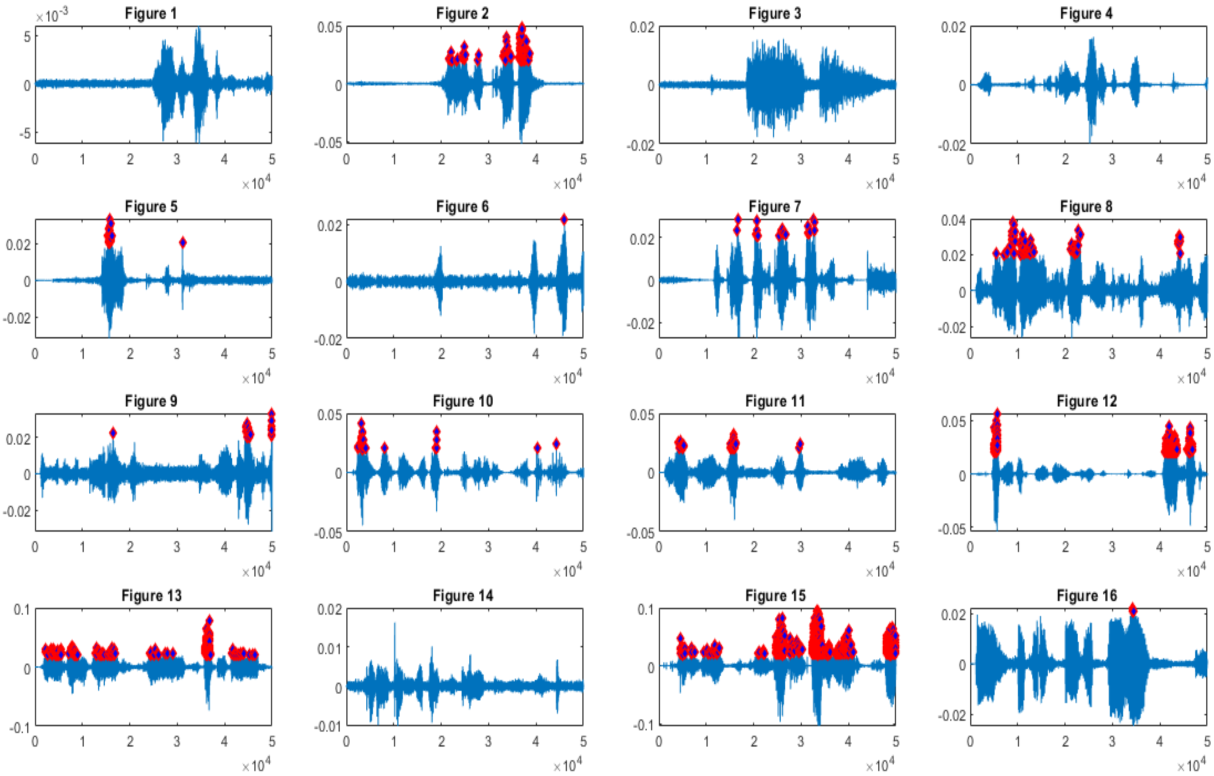

Figure 4. Power spectrum graph using sym2 wavelet at level 2

The data analysis compared the PSNR and SNR values obtained from the Haar and Sym2 wavelet transformations for both groups. Statistical comparisons between the two groups were made to identify significant differences in signal patterns, with special attention paid to the symmetry and peak densities in the power spectrum (Coates, Hero, Nowak, & Yu, 2002). This analysis formed the basis for concluding that wavelet-based voice signal analysis could serve as a rapid tool for detecting drug addiction (Shin, Powers, Grady, & Bhatt, 1999).

Results

The results of this study highlight the effectiveness of using Haar and Symlet (Sym2) wavelet transforms to analyze voice signals to identify drug-addicted individuals (Daubechies, 1990; Donoho, 1995). The analysis was divided into two parts: the power spectrum graphs and the PSNR (Peak Signal-to-Noise Ratio) and SNR (Signal-to-Noise Ratio) values, as shown in Table 1 (Coates, Hero, Nowak, & Yu, 2002).

The power spectrum graphs demonstrated clear differences between addicted and non-addicted individuals. In addicted individuals, the power spectrum was observed to be asymmetric, with a narrower frequency range at the highest peak points, indicating a distortion in their voice patterns (Huang, Feldmann, & Willinger, 2001). In contrast, non-addicted individuals exhibited a wider frequency range with a more symmetric power spectrum, suggesting healthier, undistorted voice signals (Harfoush, Bestavros, & Byers, 2000). The comparison between the two types of wavelets revealed that Haar wavelet provided clearer and more pronounced differences in voice patterns than the Sym2 wavelet, as shown in Figureure 1 (Donoho & Johnstone, 1994).

When evaluating the PSNR and SNR values, it was found that the Sym2 wavelet had lower PSNR values compared to the Haar wavelet in non-addicted individuals, making Sym2 more sensitive for detecting voice alterations in addicted people through PSNR metrics (Riedi, Crouse, Ribeiro, & Baraniuk, 1999). On the other hand, the Haar wavelet produced higher SNR values, making it more reliable for distinguishing addicted individuals based on signal clarity and noise reduction (Shin, Powers, Grady, & Bhatt, 1999).

Overall, the results suggest that Haar wavelet transform outperforms Sym2 in detecting subtle voice signal alterations caused by addiction, particularly when evaluating using SNR, as shown in Figureure 2 (Donoho, 1995). This study provides promising insights into the potential of wavelet-based voice signal analysis as a rapid and cost-effective tool for identifying drug addiction, paving the way for practical applications in law enforcement and healthcare (Beena Elizabeth Thomas, 2019; Ming & Li, 2010). However, the study's limitations, including the small sample size and restriction to male voices, suggest the need for further research to validate these findings and expand their applicability (Maxwell Oluwole Akanbi, 2018; Mohiuddin, 2019).

Discussion

The findings of this study demonstrate the potential of wavelet-based voice signal analysis, specifically using Haar and Symlet (Sym2) wavelet transforms, as a novel method for rapidly identifying drug-addicted individuals (Daubechies, 1990; Donoho, 1995). The analysis was carried out by comparing the power spectrum graphs and evaluating key metrics such as Peak Signal-to-Noise Ratio (PSNR) and Signal-to-Noise Ratio (SNR). These results underscore the significant differences in voice signal patterns between addicted and non-addicted individuals, providing a foundation for using voice as a biomarker for addiction (Donoho & Johnstone, 1994).

The power spectrum analysis revealed noticeable distinctions between the two groups. Addicted individuals showed an asymmetric power spectrum, where the frequency range at the peak points was relatively narrow. This narrowing suggests that drug addiction induces a distortion in vocal signals, possibly due to physiological or neurological effects on vocal cords or speech patterns (Ming & Li, 2010). In contrast, non-addicted individuals exhibited a broader and more symmetric power spectrum, indicating the absence of such distortions and a healthier vocal signal profile. The visual clarity of these differences was more prominent in the Haar wavelet transformation, which provided a more robust representation of voice signal anomalies compared to the Sym2 wavelet (Shin, Powers, Grady, & Bhatt, 1999).

When comparing PSNR and SNR values, further insights into the strengths of each wavelet were uncovered. The PSNR values, which reflect the quality of the reconstructed signal after transformation, were lower for the Sym2 wavelet in non-addicted individuals. This indicates that Sym2 was more sensitive to subtle deviations in voice patterns among the addicted group, making it a useful tool for identifying addiction-related changes (Harfoush, Bestavros, & Byers, 2000). Conversely, the Haar wavelet yielded higher SNR values, suggesting superior performance in noise reduction and signal clarity. In essence, the Haar wavelet transform proved more reliable for distinguishing addicted individuals by offering better signal-to-noise discrimination, especially when noise is a factor in voice recordings (Abry, Baraniuk, Flandlin, Riedi, & Veitch, 2002).

These results highlight the practical advantages of using wavelet-based approaches for addiction detection. Traditional methods such as urine or blood tests, while accurate, are often invasive, time-consuming, and costly (Health Street, 2022). Voice signal analysis offers a non-invasive, rapid, and potentially cost-effective alternative, which could be particularly beneficial for large-scale screening or real-time monitoring in both law enforcement and healthcare settings (Lessem, 2020). By leveraging advanced signal processing techniques, this study contributes to the growing body of research aimed at using behavioral and physiological markers for early addiction detection (Zhang, Duffield, Paxson, & Shenker, 2001).

However, the limitations of the study should be addressed to realize the potential of this approach fully. The relatively small sample size (16 addicted and 16 non-addicted individuals) may limit the generalizability of the findings (Mohiuddin, 2019). Additionally, the focus on male voices introduces a gender bias, which could affect the accuracy of the model when applied to broader populations (Samuel Asare, 2019). Expanding the study to include female voices and a larger, more diverse cohort would enhance the robustness of the findings. Furthermore, the requirement for speech samples in English could limit the applicability of this method in non-English-speaking regions, necessitating further exploration into language-independent voice features (Islam, 2011).

Conclusion

This study demonstrates the effectiveness of using wavelet transforms, particularly Haar and Symlet (Sym2), for analyzing voice signals to identify drug-addicted individuals. The results revealed significant differences in the power spectrum and signal clarity between addicted and non-addicted individuals, with the Haar wavelet providing clearer, more distinct patterns and higher SNR values, while Sym2 was more sensitive to subtle voice alterations through PSNR. These findings suggest that wavelet-based voice analysis could serve as a rapid, non-invasive, and cost-effective method for addiction detection, offering potential applications in healthcare and law enforcement. However, the study's limitations, including a small sample size and focus on male voices, indicate the need for further research to validate these findings. Expanding the sample diversity and exploring additional voice features could enhance the method's accuracy and applicability across broader populations.

Author contributions

S.A. conceptualized the project, developed the methodology, conducted a formal analysis, and drafted the original writing. S.H. contributed to the methodology, conducted investigations, provided resources, and visualized the data. M.R.I. contributed to the reviewing and editing of the writing.

Acknowledgment

The author thanks the Department of Basic Science, Primeasia University, Dhaka, Bangladesh.

Competing financial interests

The authors have no conflict of interest.

References

Abry, P., Baraniuk, R., Flandlin, P., Riedi, R., & Veitch, D. (2002). Multiscale nature of network traffic. IEEE Signal Processing Magazine, 19(3), 28–46. https://doi.org/10.1109/79.998080

Akella, A., Seshan, S., & Balakrishnan, H. (2003, November). The impact of false sharing on shared congestion management. In Proceedings of the 11th IEEE International Conference on Network Protocols.

Balakrishnan, H., Rahul, H., & Seshan, S. (1999, September). An integrated congestion management architecture for Internet hosts. In Proceedings of ACM SIGCOMM '99.

Beena Elizabeth Thomas, K. T. (2019, July 31). Smoking, alcohol use disorder and tuberculosis treatment outcomes: A dual comorbidity burden that cannot be ignored. PLOS ONE, 14(7), 1-12. https://doi.org/10.1371/journal.pone.0220507

Coates, M., Hero, A. O. III, Nowak, R., & Yu, B. (2002). Internet tomography. IEEE Signal Processing Magazine, 19(3), 47–65. https://doi.org/10.1109/79.998080

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. IEEE Transactions on Information Theory, 36(5), 961–1005. https://doi.org/10.1109/18.57199

Donoho, D. L. (1995). De-noising by soft-thresholding. IEEE Transactions on Information Theory, 41(3), 613–627. https://doi.org/10.1109/18.382009

Donoho, D. L., & Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. Biometrika, 81(3), 425–455. https://doi.org/10.1093/biomet/81.3.425

Harfoush, K., Bestavros, A., & Byers, J. (2000, November). Robust identification of shared losses using end-to-end unicast probe. In Proceedings of the 8th IEEE International Conference on Network Protocols.

Harfoush, K., Bestavros, A., & Byers, J. (2001, January). Robust identification of shared losses using end-to-end unicast probe. Technical Report BUCS–TR–2001–001, Computer Science Department, Boston University, Massachusetts, U.S.A.

Health Street. (2022). Drug tests. Retrieved from https://dki.health-street.net/D4/V1?pkw=drug%20testing

Huang, P., Feldmann, A., & Willinger, W. (2001, November). A non-intrusive, wavelet-based approach to detecting network performance problems. In Proceedings of the First ACM SIGCOMM Internet Measurement Workshop (pp. 213–227).

Islam, M. R. (2011, March). Wavelets, its application and technique in signal and image. Global Journals Inc. (USA), 11(4). Retrieved from https://search.yahoo.com/search?fr=mcafee&type=E210US91213G0&p=Islam%2C+M.+R.+(March+2011).+Wavelets%2C+its+Application+and+Technique+in+signal+and+image.+Global+Journals+Inc.%2C+Volume+11(Issue+4)

Katabi, D., Bazzi, I., & Yang, X. (2001, October). A passive approach for detecting shared bottlenecks. In Proceedings of the 10th IEEE International Conference on Computer Communications and Networks.

Lessem, N. &. (2020, February 17). What's the maximum amount of marijuana a person can legally possess? Retrieved from https://www.lnlegal.com/blog/2020/february/whats-the-maximum-amount-of-marijuana-a-person-can-legally-possess-c/

Limit Alcohol. (2022). National Center for Health Promotion and Disease Prevention. Retrieved from https://www.prevention.va.gov/Healthy_Living/Limit_Alcohol.asp

Maxwell Oluwole Akanbi, L. A. (2018, November 23). The efficacy of smoking cessation interventions in low and middle-income countries: A systematic review and meta-analysis. Addiction, 114, 620–635. https://doi.org/10.1111/add.14518

Mills, D. L. (1992, March). Network time protocol (version 3) specification, implementation and analysis. RFC 1305.

Ming, D. L., & Li, M. B. (2010, June 2). New insights into the genetics of addiction. Nature Reviews Genetics, 10(4), 225–231. https://doi.org/10.1038/nrg2536

Mohiuddin, A. K. (2019). Drug addiction in Bangladesh: "A consequence of social demoralization rather than individual flaws." Dr. M. Nasirullah Memorial Trust, 2(10), 1-11. Retrieved from https://escipub.com/international-journal-of-addiction-research-and-therapy/

Moon, S. B., Skelly, P., & Towsley, D. (1999, March). Estimation and removal of clock skew from network delay measurements. In Proceedings of IEEE INFOCOM '99.

Riedi, R. H., Crouse, M. S., Ribeiro, V. J., & Baraniuk, R. G. (1999). A multifractal wavelet model with application to network traffic. IEEE Transactions on Information Theory, 45(3), 992–1018. https://doi.org/10.1109/18.761338

Rubenstein, D., Kurose, J., & Towsley, D. (2002). Detecting shared congestion of flows via end-to-end measurement. IEEE/ACM Transactions on Networking, 10(3), 381–395. https://doi.org/10.1109/TNET.2002.1012360

Samuel Asare, M. S. (2019, August 27). Effects of prices on youth cigarette smoking and tobacco use initiation in Ghana and Nigeria. International Journal of Environmental Research and Public Health, 16, 1-14. https://doi.org/10.3390/ijerph16173114

Shin, Y., Powers, E. J., Grady, W. M., & Bhatt, S. C. (1999, July). Optimal Daubechies' wavelet bases for detection of voltage sag in electric power distribution and transmission systems. In Wavelet Applications in Signal and Image Processing VII, SPIE (pp. 873–883).

Trees, H. L. V. (1968). Detection, estimation, and modulation theory. John Wiley & Sons.

V. M. Anantha Eashwar, S. G. (2019, June). Pattern of alcohol consumption and its associated morbidity among alcohol consumers in an urban area of Tamil Nadu. Family Medicine and Primary Care, 8(6), 2029-2035. https://doi.org/10.4103/jfmpc.jfmpc_226_19

Williams, W. (1991, July). Uncertainty, information, and time-frequency distributions. In Advanced Signal Processing Algorithms, Architectures, and Implementations II, SPIE (pp. 144–156).

Zhang, Y., Duffield, N., Paxson, V., & Shenker, S. (2001, November). On the constancy of Internet path properties. In Proceedings of ACM SIGCOMM Internet Measurement Workshop.