COMPARATIVE ANALYSIS OF DIFFERENT WAVELET BASED METHODS FOR DENOISING OF BIRD AUDIO RECORDINGS

Abstract

Bird audio signals are acoustic signals produced by birds for various purposes, including communication, territorial defense, courtship, and species recognition. These signals are rich in information and exhibit diverse patterns, structures, and frequency characteristics. Usually, bird audio recordings contain background noise captured by recording equipment, which reduces signal to noise ratio. As signal to noise ratio is a critical factor in determining the clarity of the bird vocalizations, it is necessary to remove this noise before actual signal processing. The main aim of this research work is to implement different wavelet based denoising methods for removal of noise from bird song recordings. Our proposed method improved signal to noise ratio after denoising while preserving prominent phrases in bird vocalizations.

Keywords: bird audio recordings; wavelet; wavelet signal denoiser; signal analyser; spectogram

Authors

Prof. Sushant. S. Khedgikar
Department of Electronics & Computer Engineering
P.E.S. College of Engineering
Aurangabad, India.

Dr. Minakshi R. Rajput
University Department of Information and Communication Technology
MGM University
Aurangabad, India.

I. INTRODUCTION

A bird audio signal is a natural signal in time domain. When we listen to it, we want to enhance region of interest from and suppress what we do not want to listen. We want to characterize the system that does so. Many bird species are being driven to extinction. There are thousands of bird species in the world, each with its own unique characteristics and habitats. Few example species are Perching Birds, Birds of Prey, Waterbirds, Seabirds and Songbirds [1].

Birds play a vital role in maintaining ecological balance and biodiversity. They control pest populations, disperse seeds, and contribute to pollination, which is essential for healthy ecosystems. Birds are often considered indicators of environmental health. Changes in bird populations can serve as early warning signs of habitat degradation, pollution, or climate change, alerting us to potential risks to other species and ecosystems. Many bird species attract birdwatchers and ecotourism, contributing to local economies. Protecting bird habitats can create sustainable income opportunities and promote conservation-oriented development. Birds have cultural significance in many societies and hold aesthetic value for people worldwide. Their songs, colors, and behaviors inspire art, literature, and cultural traditions [2]. In conclusion, protecting bird species is crucial for the preservation of biodiversity, ecosystem balance, and the overall well-being of our planet. By taking conservation measures and raising awareness, we can ensure that future generations can enjoy the beauty and benefits of these remarkable creatures.

Many bird species are common, but it is difficult for humans to identify a particular species. Ornithology includes automatic identification of bird calls from continuous environmental records. But most of the times, these recordings are noisy. Audio signal processing is must for identifying a particular bird. There is a need to find the frequencies that form an audio. We need to segment bird recording in time domain, also need to find out the part during which prominent note is heard and other parts on time axis during which different notes are heard. These notes are also varying in length. One note may be longer, another could be shorter. In addition to this, there is need to analyze the signal in frequency domain which involves finding the frequency contents in the bird song and locations at which transforms are prominent. Audio signal generally comprises of many sine waves which come together to form an audio piece. Analyzing bird songs involves both time and frequency domain analysis. But the problem with time-frequency analysis is that when we want to focus more on time domain, frequency region is narrowed and vice a versa and thus it’s not possible to get both time and frequency information at the same time.

Denoising bird audio songs typically involves a combination of techniques to reduce unwanted noise while preserving the quality and integrity of the bird vocalizations. Some common approaches used for denoising bird audio songs:

1. **Spectral Subtraction:** Spectral subtraction is a widely used technique for noise reduction. It estimates the noise spectrum by analyzing a noise-only portion of the recording and then subtracts the estimated noise spectrum from the entire signal. The resulting signal should ideally contain mostly the bird vocalizations with reduced noise [3],[4].

2. **Wavelet Denoising:** Wavelet-based denoising can be effective for removing noise from bird audio songs. It involves decomposing the signal using a wavelet transform, identifying and thresholding noise coefficients, and reconstructing the denoised signal [5] [6].

3. **Adaptive Filtering:** Adaptive filtering methods aim to adaptively estimate the noise characteristics and attenuate the noise components in the audio signal. Techniques such as the Wiener filter or the Least Mean Squares (LMS) algorithm are commonly used for adaptive noise reduction [7].

4. **Statistical Modeling:** Statistical modeling approaches can be employed to model the statistical properties of the bird vocalizations and the noise. By estimating the parameters of the models, it becomes possible to separate the bird songs from the noise components. Hidden Markov Models (HMMs) and Gaussian Mixture Models (GMMs) are commonly used for this purpose.

5. **Machine Learning-based Denoising:** Machine learning techniques can be used to train denoising models that learn the mapping between noisy bird songs and clean bird songs. These models can then be applied to denoise new audio recordings. Deep learning approaches such as Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs) have shown promising results in this area [8-10].

6. **Spatial Filtering:** Spatial filtering techniques can be useful when working with multi-microphone recordings. By utilizing the spatial information captured by the microphone array, it is possible to enhance the bird songs and suppress noise from specific directions or locations [11].

II. RESEARCH METHODOLOGY

In this research we tried to process bird audio recordings using wavelet based techniques and compared those techniques.

1. **Database used for Processing of birds' Audio Recordings:** The data set contains bird sound recordings, a specific subset gathered from the Xeno Canto collection to form a balanced data set across 88 species and their 264 recordings commonly heard in the United Kingdom [12]. For experimentation purpose we used xc25119 audio song recording of bird Redwing of iliacus species. Platform used for experimentations is MATLAB R2020b. Tools used are signal analyzer and wavelet signal denoiser app from signal and communications toolbox.

2. **Wavelet based Denoising:** Wavelets are mathematical function that cut data in terms of different frequency components basically sine waves. Wavelet transform has infinite set of possible basis functions. Wavelet analysis provides easy access to information that may not be provided by time-frequency analysis in Fourier analysis. Power spectra shows localization in frequency, representing how much power is concentrated in specific frequency interval. Sparseness in wavelet transform results into various applications like data compression, feature detection in images and noise removal etc. When we decompose a wave using wavelet, averaging filters and other filters are used that produce

details. If these details are small, they are omitted without affecting the main features within a signal. Thresholding is done to set all coefficients to zero which are less than a particular threshold. These coefficients are used in an inverse wavelet transformation to reconstruct the original signal [13].

We used wavelet denoiser app in MATLAB to denoise bird audio. Wavelet denoiser app helps us to visualize and denoise time series data. It easily adjust default parameters and apply different denoising techniques. Steps followed to denoise the signal using wavelet are explained in figure 1.

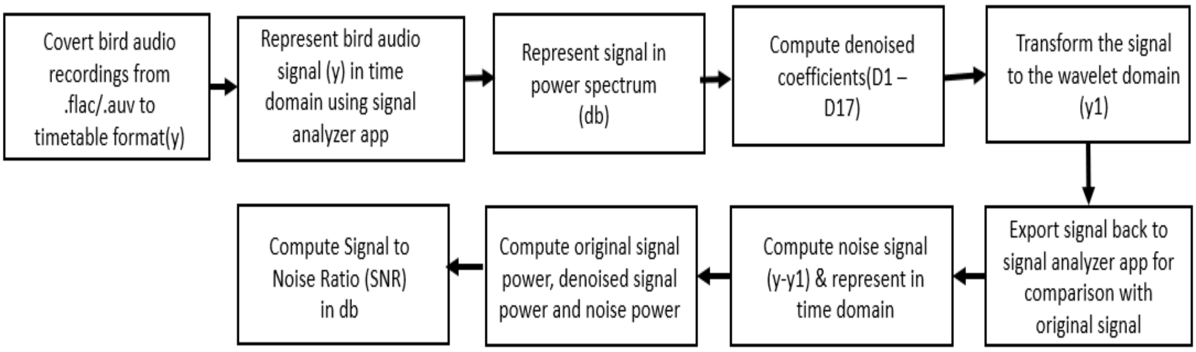

Figure 1: Block Diagram Representing Process Flow for Wavelet based Denoising of Bird Signal

Firstly, the original signal is represented in time domain along with power spectrum. Then it is transformed to the wavelet domain using parameters Wavelet: sym4, Level: 17, Denoising Method: Minimax, Threshold Rule: Soft, Noise: Level Independent. Then denoised signal is exported back to signal analyzer app for comparison with original signal.

xc25119 audio song recording represents Redwing bird vocalization. It is represented in time domain using signal analyzer app as shown in figure 2. There are three prominent phrases (actual bird vocal sound) with pauses in between. These phrases lie in frequency range 2 to 4 KHz as observed from power spectrum which is shown below the signal representation.

Power spectrum of the signal shown in figure 2 represents Power Spectral Density (PSD) representing the power distribution of the signal with respect to frequency. It indicates how much power is present at different frequencies. Signal components have higher power in the relevant frequency bands (2 -4 KHz), while noise components are more evenly distributed or concentrated at certain frequencies.

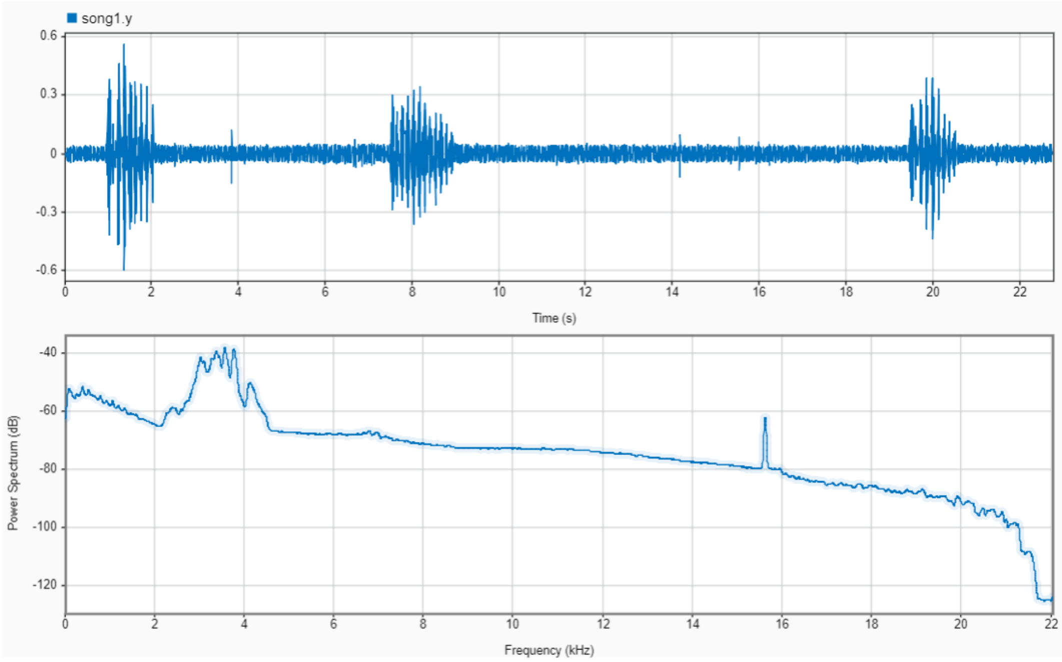

Figure 2: Original bird sound recording as time domain signal with its power spectrum

Figure 3 shows the spectrogram of the signal indicating the desired bird audio signal and the noise. As can be seen from the figure, all the three bird sounds are within 2-4 KHz with stronger amplitudes indicated by orange colors. The blue color indication with low amplitudes is of noise.

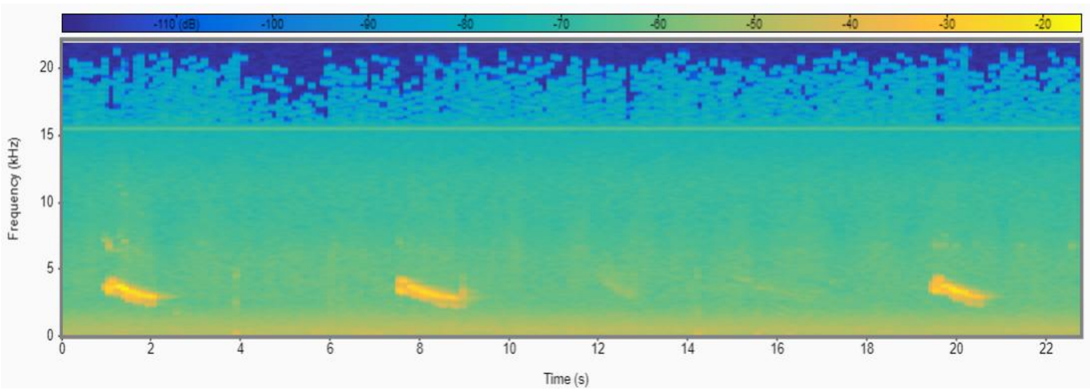

Figure 3: Spectrogram of the Original Signal

After denoising the signal using wavelet denoiser, higher signal strength is obtained only at desired frequencies ensuring noise reduction to some extent. Figure 4 shows the denoised signal along with original signal.

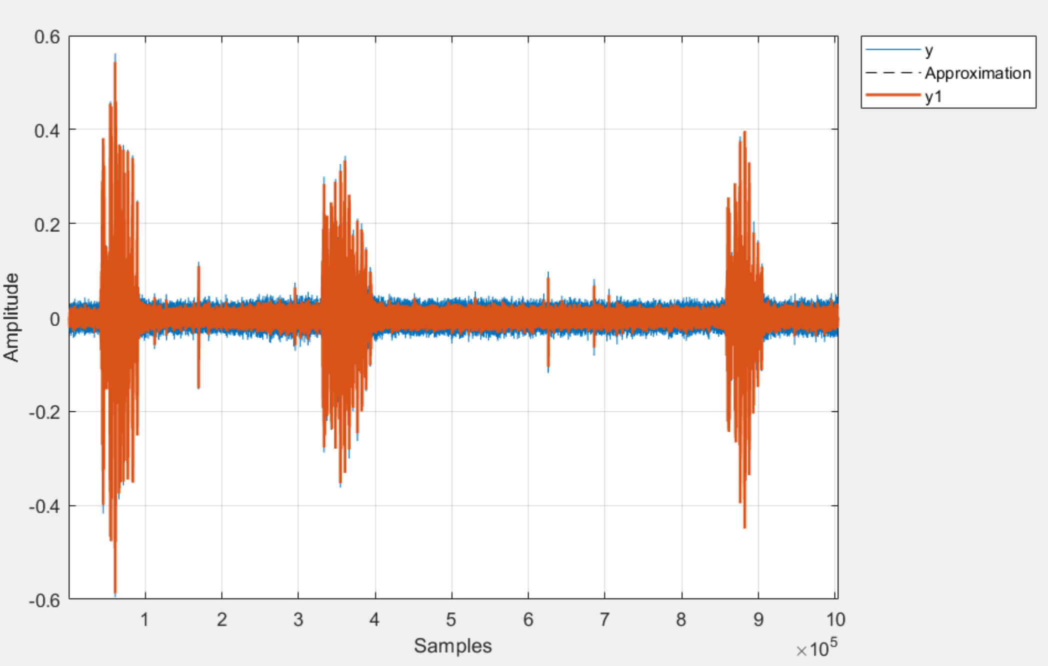

Figure 4: Denoised signal (y1) with original signal (y)

Figure 5 represents filtered denoised coefficients obtained from a wavelet denoising algorithm. As it can be seen, the low-frequency coefficients (D1 to D4) contain the essential signal information while the high-frequency coefficients usually represent the noise. The denoising process starts with decomposing the signal using a wavelet transform to obtain the wavelet coefficients at various scales. After applying thresholding to the high-frequency coefficients, the denoised coefficients are reconstructed to obtain the cleaned signal.

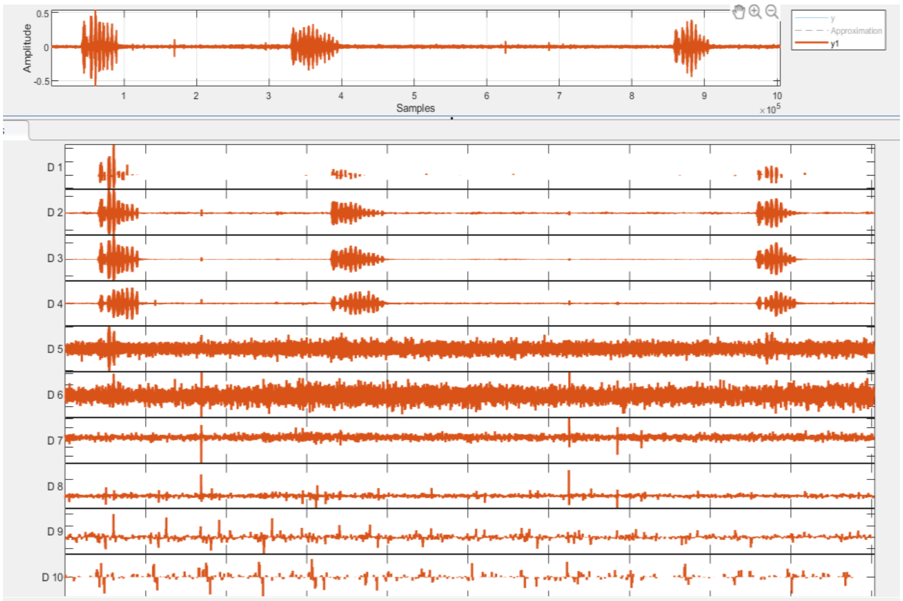

Figure 5: Denoised signal with 17 denoised coefficients

Denoising is often a trade-off between reducing noise and preserving signal details. If the threshold is too high, it may remove some relevant signal information along with the noise, leading to signal distortion. On the other hand, if the threshold is too low, some noise may remain in the denoised coefficients.

Noise signal represented in figure 6 is obtained by subtracting denoised signal (y1) from original signal (y). Figure indicates the frequency concentration of noise signal which has power distributed across a wide range of frequencies, it may affect the desired signal at its main frequency components.

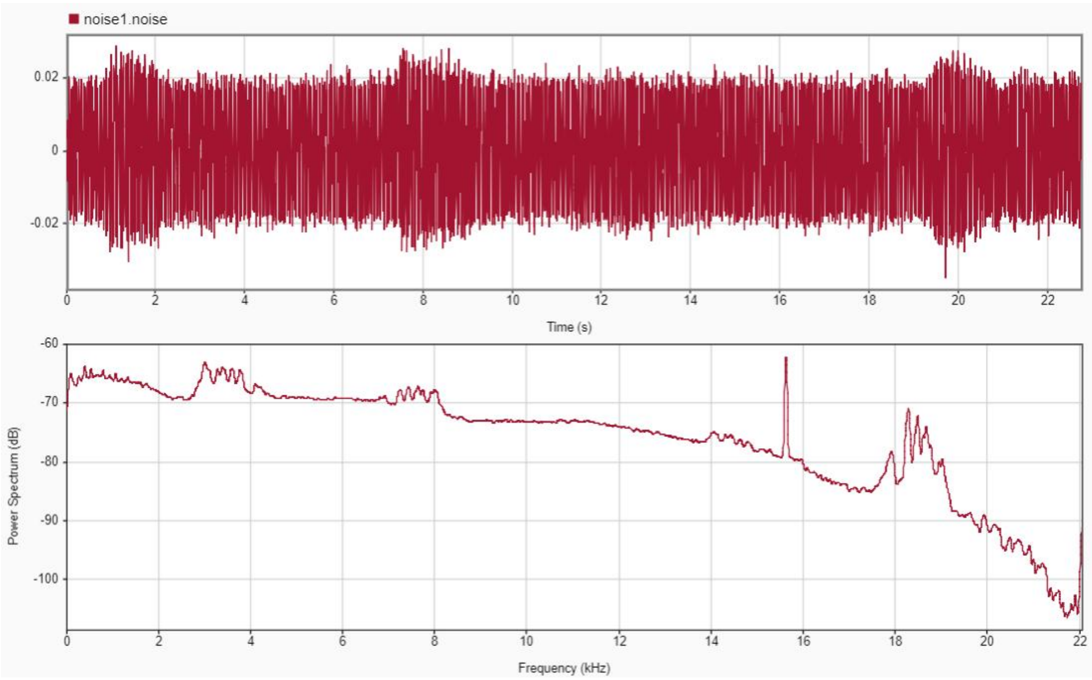

Figure 6: Noise signal representing power distribution across wide range of frequencies

Figure 7 shows all three signals together original signal(y), denoised signal(y1) and noise signal (y-y1) and their power spectra.

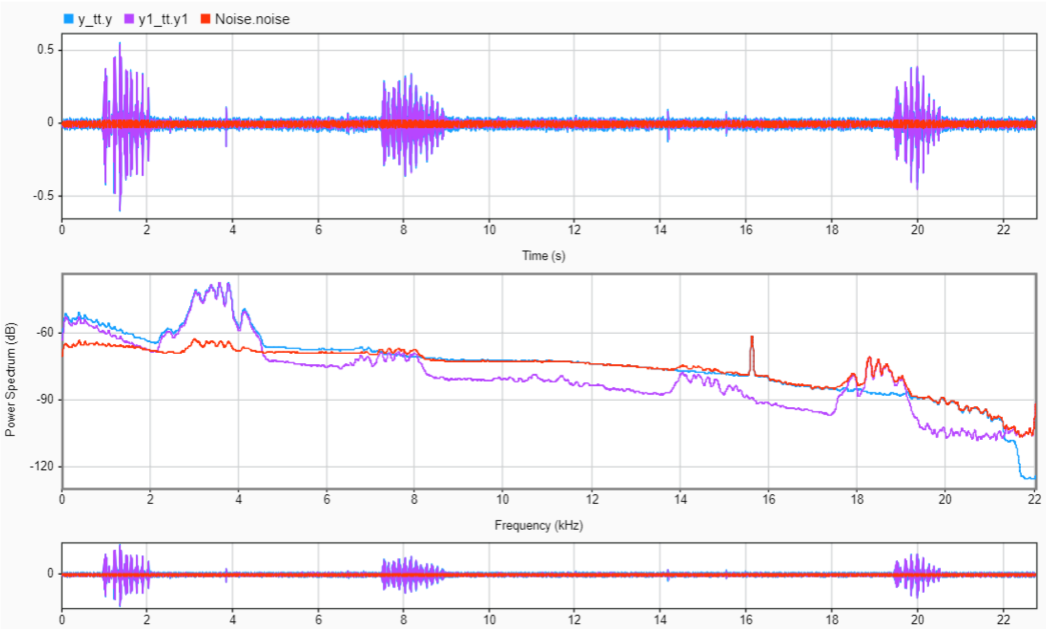

Figure 7: Original signal(y), denoised signal(y1) and noise signal (y-y1) and their power spectra.

III. COMPARATIVE ANALYSIS OF DIFFERENT WAVELET METHODS

Comparing the signal-to-noise ratio (SNR) of different denoising methods is a common way to assess their effectiveness in reducing noise from signals. A higher SNR value indicates a better denoising performance, as it means that the denoised signal has a higher proportion of signal power relative to the noise power. Table 1 compares different wavelet methods and parameters used for denoising the bird audio recording. It also compares signal to noise ratio (SNR in db) for different methods.

We first computed the power of the original signal (y) and then calculated the power of the noise signal as the difference between the original and denoised signals. Finally, we computed the SNR in decibels using the formula.

$$SNR = 10 * \log_{10} (\text{power of denoised signal / power of noise signal})$$

As observed in table some methods may preserve signal details better, while others may be more efficient in removing certain types of noise. In addition to SNR, we also listened to the denoised bird audio signals and examined denoised images to see if any artifacts or distortion are introduced. It is observed that the denoised bird recordings having higher SNR value which indicates a cleaner and more distinguishable bird vocal quality compared to the original signal.

Table 1: Different Wavelet Methods used for Denoising along with Signal to Noise Ratios (SNRs) obtained

Wavelet DenoisingMethod	Power of denoised signal	Noisepower	SNR(db)
Wavelet:symlet Level: 17Denoising Method: BayesThreshold Rule: Median Noise: Level Independent	1657.889611	6.369295039	24.19927
Wavelet: symletLevel: 17Denoising Method: Minimax Threshold Rule: Soft Noise: Level Independent	1471.104722	44.05326786	15.8004
Wavelet: fkLevel:17 Denoising Method: BayesThreshold Rule: Median Noise: Level Independent	1645.900626	10.02026444	22.23139
Wavelet: daubenchies Level: 17 Denoising Method: BayesThreshold Rule: Median Noise: Level Independent	1641.182598	11.85433247	21.50141
Wavelet: biorthogonal Level: 17 Denoising Method: Bayes Threshold Rule: Median Noise: Level Independent	1641.182598	11.85433247	21.50141

IV. CONCLUSIONS AND FUTURE SCOPE

Bird species identification from audio recordings is an interesting and challenging task. The quality and variety of these recordings will significantly impact the accuracy of your model. Background noise present in these recordings interferes the identification process. To preprocess bird audio recordings, removal of noise and other artifacts is necessary. The main aim of this research is to implement wavelet denoising algorithm to effectively reduce noise contents in bird audio recordings. This research explored various wavelet denoising methods and evaluated based on signal to noise ratio of denoised signals. Prominent notes in bird vocalizations are clearly heard after denoising which also proves effectiveness of wavelets in denoising audio signals.

This research will further be extended for identification of particular bird species. Machine learning model can be trained to extract features from spectrogram images obtained from denoised signals.

REFERENCES

[1] Website: https://animalcorner.org/birds/
[2] Norris, Ken; Pain, Deborah J. (2002). Conserving Bird Biodiversity (General Principles and their Application) || Why conserve bird diversity?, 10.1017/CBO9780511606304(2), 20–33. doi:10.1017/CBO9780511606304.003
[3] A. Patti and G. A. Williamson, ''Methods for classification of nocturnal migratory bird vocalizations using Pseudo Wigner-Ville transform, ''in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), May 2013, pp. 758–762.
[4] Z. Chen, "Noise Reduction of Bird Calls Based on a Combination of Spectral Subtraction, Wiener Filtering, and Kalman Filtering," 2022 IEEE 5th International Conference on Automation, Electronics and Electrical Engineering (AUTEEE), Shenyang, China, 2022, pp. 512-517, doi: 10.1109/AUTEEE56487.2022.9994282.
[5] Priyadarshani, Nirosha & Marsland, Stephen & Castro, Isabel & Punchihewa, A.. (2016). Birdsong Denoising Using Wavelets. PLOS ONE. 11. e0146790. 10.1371/journal.pone.0146790.
[6] Arja, Selin & Turunen, Jari & Tanttu, Juha. (2007). Wavelets in Recognition of Bird Sounds. EURASIP Journal on Advances in Signal Processing. 2007. 10.1155/2007/51806.
[7] Zhang, Xiaoxia & Li, Ying. (2015). Adaptive energy detection for bird sound detection in complex environments. Neurocomputing. 155. 10.1016/j.neucom.2014.12.042.
[8] Á. Incze, H. -B. Jancsó, Z. Szilágyi, A. Farkas and C. Sulyok, "Bird Sound Recognition Using a Convolutional Neural Network," 2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY), Subotica, Serbia, 2018, pp. 000295-000300, doi: 10.1109/SISY.2018.8524677.
[9] Kahl, Stefan & Wilhelm-Stein, Thomas & Hussein, Hussein & Klinck, Holger & Kowerko, Danny & Ritter, Marc & Eibl, Maximilian. (2017). Large-Scale Bird Sound Classification using Convolutional Neural Networks.
[10] T. Grill and J. Schlüter, "Two convolutional neural networks for bird detection in audio signals," 2017 25th European Signal Processing Conference (EUSIPCO), Kos, Greece, 2017, pp. 1764-1768, doi: 10.23919/EUSIPCO.2017.8081512.
[11] M. Dhanushree, R. Priyadharsini, T. Sree Sharmila, Acoustic image denoising using various spatial filtering techniques, International Journal of Information Technology , 11(4), 2019, pp- 659-665.
[12] website: https://www.kaggle.com/datasets/imoore/xenocanto-bird-recordings-dataset
[13] NPTEL course on "Fundamentals of Wavelets, Filter Banks and Time frequency Analysis, coordinated by IIT Bombay" https://archive.nptel.ac.in/courses/108/101/108101093/#

