CAUCHY – Jurnal Matematika Murni dan Aplikasi
Volume 11 (2) (2026), Pages 2366–2390
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Classification of Panthera Species Based on Vocalization Using Higuchi Fractal Dimension and K-Nearest Neighbor

Zharfanudin Afif and Dwi Juniati*

Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Surabaya, Indonesia

Abstract

The genus *Panthera*, comprising the lion, jaguar, leopard, tiger, and snow leopard, exhibits complex vocalization characteristics as biological communication mechanisms in its natural environment. However, manual acoustic identification is limited by environmental variability, individual sound variations, and recording noise, necessitating an automated approach. This study applies the Higuchi Fractal Dimension (HFD) for feature extraction and the K-Nearest Neighbor (KNN) algorithm for classification to recognize vocalizations of five extant *Panthera* species. The dataset consists of 100 recordings (20 per species) from the Tierstimmenarchiv repository, preprocessed by noise reduction, two-second segment selection, mono conversion, resampling, and amplitude normalization. An empirical fundamental frequency (F_0) verification showed a range of 53 to 367 Hz, confirming the adequacy of the DWT decomposition depth. The signals were decomposed using DWT (Daubechies 4, 7 levels), and HFD was computed on the resulting coefficients with $K_{\max}$ of 10, 16, 28, 32, 50, and 60. Performance was evaluated using repeated nested stratified cross-validation (5-fold outer loop repeated 10 times; 5-fold inner loop for selecting $K_{\max}$ and k), together with an ablation study of DWT approximation (cA) and detail (cD) coefficients and a non-DWT baseline computed directly from the preprocessed signal. The highest mean accuracy, 70.80% (SD = 8.77%), was obtained by combining all detail coefficients with the three highest-level approximation coefficients, and was stable across randomly sampled seeds (SD = 1.06%). The non-DWT baseline reached only 56.40%. These results indicate that HFD computed on DWT sub-band provides discriminative features for classifying *Panthera* vocalizations, although repeated individuals across folds cannot be ruled out.

Keywords: Bioacoustic Signals; Genus *Panthera*; Higuchi Fractal Dimension; K-Nearest Neighbor; Sound Classification.

1. Introduction

The genus *Panthera* comprises apex carnivorous predators commonly known as big cats [1], which include several iconic species such as the lion (*Panthera leo*), jaguar (*Panthera onca*), leopard (*Panthera pardus*), tiger (*Panthera tigris*), and snow leopard (*Panthera uncia*) [2]. The aforementioned species exhibit complex vocal characteristics as a form of biological communication within their natural environments. The ability to automatically identify each species' vocalization type represents a crucial aspect of animal sound signal processing, particularly in supporting technology-based wildlife monitoring. However, manual sound identification possesses inherent limitations because acoustic signals are influenced by environmental conditions, variations in individual vocal characteristics, and recording noise interferences.

*Corresponding author. E-mail: dwijuniati@unesa.ac.id

Received: August 1, 2026 Revised: October 4, 2026 Accepted: October 13, 2026
DOI: https://doi.org/10.18860/cauchy.v11i2.45890

Animal vocal signals exhibit complex and non-stationary waveform patterns, rendering conventional analysis methods often inadequate for optimally capturing the microcharacteristics of a signal. One approach capable of quantifying the complexity of a time series is the Higuchi Fractal Dimension (HFD). This method extracts geometric characteristics from a signal without requiring phase-space reconstruction, making it suitable for bioacoustic signal analysis. Several studies have shown that HFD performs effectively in animal sound classification. The application of HFD to domestic cat vocalization classification successfully distinguished various sound conditions with an accuracy of 87.5% [3]. Furthermore, parameter optimization of HFD in monkey sound classification achieved an accuracy of 90.38% [4], while its implementation on other bioacoustic subjects, such as seals [5], dolphins [6], birds [7], and crickets [8], also demonstrated competitive performance.

In the classification stage, most HFD-based studies on bioacoustic signals employ the K-Nearest Neighbor (KNN) algorithm as the classifier [3]-[8]. KNN is a distance-based classification algorithm that determines the class of a sample based on a set of nearest neighbors in the feature space. Its structural simplicity, capability to handle multi-class data, and ease of implementation have led to its widespread adoption in various sound signal classification studies. In the HFD-based approach, the fractal dimension values extracted from the audio signals serve as numerical features, which are subsequently used to determine characteristic proximity across classes using KNN.

Nevertheless, the classification performance of the KNN algorithm is strongly influenced by data distribution and the choice of k, necessitating systematic parameter optimization. Despite the success of HFD-KNN across various bioacoustic domains [3]–[8], its application to the genus *Panthera* remains unexplored, as prior studies have focused primarily on image-based classification [2]. Interspecific differences in laryngeal anatomy and roaring adaptations [9] suggest distinct vocal complexities suitable for HFD quantification. This study addresses these gaps by evaluating HFD-KNN on five *Panthera* species using systematic parameter tuning and repeated validation. Prior to detailing the methodology, a brief overview of the morphological, behavioral, and acoustic profiles of the five target species is presented below (Figures 1–5):

1. Lion (*Panthera leo*). The body exhibits a uniform color without spot patterns; it is yellowish-brown, white on the abdomen and inner sides of the legs, with black on the backs of the ears. In terms of general size and appearance, the Asian and African subspecies are similar; the most distinct physical difference between them is the presence of a longitudinal skin fold along the belly in the Asian subspecies. Variations among African subspecies have been recorded both between regions and within the same geographical area based on size, coat color, and mane development and coloration, with individuals from the southern part of the continent averaging larger sizes than those from the eastern region [10].



Fig. 1: *Panthera leo.* Photo by Andreas Vonlanthen on Unsplash[1], used under the Unsplash License.

2. Jaguar (*Panthera onca*). The jaguar (*Panthera onca*) is the largest felid species in the Neotropical realm, yet ironically remains the least studied member of the genus *Panthera*. Historically, its geographic distribution extended from the southwestern United States to central Argentina; however, it has now been extirpated from approximately 55% of its historical range, with a global conservation status of Near Threatened, and several populations are already classified as endangered or critically endangered [11].



Fig. 2: *Panthera onca*. Photo by Mike van den Bos on Unsplash[2], used under the Unsplash License.

3. Leopard (*Panthera pardus*). The leopard is the smallest among the species in the genus Panthera. As a solitary and adaptable species, it possesses the broadest geographical range among all wild felids, inhabiting rainforests, mountains, semi-arid environments, and suburban areas across sub-Saharan Africa, the Middle East, and South Asia to the Russian Far East. Despite its wide distribution, the leopard is listed as "Near Threatened" by the International Union for Conservation of Nature and Natural Resources (IUCN), with several Asian subspecies categorized as endangered. The leopard primarily preys on small to medium-sized ungulates, but maintains a varied diet including fish, reptiles, birds, and small mammals [12].



Fig. 3: *Panthera pardus*. Photo by Jane Stroebel on Unsplash[3], used under the Unsplash License.

[1] https://unsplash.com/id/foto/seekor-singa-jantan-yang-megah-berjalan-melintasi-sabana-berumput-2OME-KI5IYU

[2] https://unsplash.com/photos/k5brbv8T3yE

[3] https://unsplash.com/photos/NzHZ-KjWQzI

4. Tiger (*Panthera tigris*). The tiger is the largest species in the family Felidae. Its base coat coloration ranges from reddish orange to ochraceous orange and is characterized by predominantly vertical dark stripes, while the underparts are generally white or whitish. The skull is large, exceeding 255 mm in length and typically ranging from 285 to 360 mm. The total body length of adults generally ranges from 2200 to 3000 mm, with females typically being smaller than males [13].



Fig. 4: *Panthera tigris*. Photo by Waldemar Brandt on Unsplash[4], used under the Unsplash License.

5. Snow Leopard (*Panthera uncia*). The snow leopard (*Panthera uncia*) is native to the mountains of Central and South Asia, where it inhabits elevations between 3000 and 4500 m above sea level. This species is listed as endangered by the IUCN, and its population is reported to be continually declining.



Fig. 5: *Panthera uncia*. Photo by Mathieu Odin on Unsplash[5], used under the Unsplash License.

The biological characteristics described above provide the contextual basis for examining whether the vocal signals of these five *Panthera* species contain distinguishable acoustic patterns. To investigate these differences systematically, the following section describes the dataset, audio preprocessing procedures, feature extraction using DWT and HFD, KNN-based classification, and the evaluation framework employed in this study.

[4] https://unsplash.com/photos/zQsCISvBY2E
[5] https://unsplash.com/photos/YU8GTlre5Zk

2. Methods

The genus *Panthera* comprises apex predators, namely the lion, jaguar, leopard, tiger, and snow leopard, whose vocalizations vary in acoustic pattern across species. Since these sound signals are complex, non-stationary, and environment-sensitive, computational methods are needed to extract features that distinguish species reliably.

This study uses an audio signal processing and machine learning approach to classify Panthera species from their vocal signals, treating each recording as a time series of amplitude over time. Recordings were sourced from the Tierstimmenarchiv and selected for clear, continuous roar vocalizations of sufficient quality. Preprocessing standardized each clip (6 dB noise reduction, 2-second duration, mono channel, 96,000 Hz sampling rate, amplitude normalization) before feature extraction.

Feature extraction combined Discrete Wavelet Transform (DWT), with mother wavelet selection to find the optimal basis function, and Higuchi Fractal Dimension (HFD) to quantify signal complexity without phase-space reconstruction. The resulting HFD values served as numerical features capturing temporal differences between species. These features were then classified using K-Nearest Neighbor (KNN), chosen for its simplicity, multi-class capability, and prevalence in HFD-based bioacoustic classification studies.

Overall, the workflow proceeds from raw audio to species identification through: sound selection, preprocessing, DWT-based feature extraction with wavelet selection, HFD complexity extraction, KNN classification, and performance evaluation via accuracy and confusion matrix.

2.1. Audio Dataset and Data Selection

The audio data used in this study are secondary data obtained from the Animal Sound Archive (Tierstimmenarchiv) of the Museum für Naturkunde Berlin, in (.wav) format [14].

The audio data used in this study comprise five extant species within the genus Panthera: the lion (*Panthera leo*), jaguar (*Panthera onca*), leopard (*Panthera pardus*), tiger (*Panthera tigris*), and snow leopard (*Panthera uncia*). The selection of these taxa aligns with prior classifications of the genus *Panthera* [2]; however, the previous work distinguished male and female lions separately (e.g., lion and lioness), this study consolidates them under their singular taxonomic classification, *Panthera leo* alongside the other four primary species.

The initial stage of data collection was conducted by searching and downloading audio recordings based on the species labels available in the repository. Each recording was then manually inspected to ensure the suitability of the acoustic characteristics required for this study. The selection process was carried out by choosing recordings that exhibited clear, continuous roaring vocalizations and possessed adequate signal quality for further analysis.

Manual sound selection was conducted because the available bioacoustic recordings exhibited considerable variation in characteristics, such as differences in duration, vocalization types, sound intensity, and the presence of non-informative audio segments. Through this preliminary selection, the curated dataset is expected to provide a more consistent acoustic representation across species, thereby improving the quality of feature extraction and classification. A total of 20 audio samples were collected for each species, yielding 100 Panthera sound samples across the five distinct species.

Following the selection process, each selected audio recording was prepared for the audio preprocessing stage. The curated data were then further processed through noise level adjustment, segment selection to a fixed duration, mono conversion, sampling rate standardization, and amplitude normalization prior to feature extraction.

Recording-level metadata such as individual animal identity were not consistently available in the source repository; therefore, potential data leakage from repeated individuals across folds cannot be fully ruled out. This is acknowledged as a limitation of the current study.

2.2. Audio Pre-processing

The audio preprocessing stage is conducted to standardize the signal characteristics prior to feature extraction. This stage comprises noise adjustment, 2-second segment selection, mono conversion, sampling rate standardization, and audio normalization.

2.2.1. Filtering

Filtering is a signal processing operation that separates undesired components from noise [15]. Mathematically, the filtering process for discrete signals can be expressed by the convolution in Eq. (1) [16]:

$$y_n = \sum_{k=-m}^{m} W_k \cdot x_{n+k} \tag{1}$$

where y_n is the output signal after the filtering process, x_n is the input signal, W_k is the filter coefficient or weight, m is the number of samples in the filter window, and k is the summation index.

Noise reduction was performed using Audacity software (v3.7.8) with its default Noise Reduction parameters: a noise reduction level of 6 dB, a sensitivity of 6.00, and frequency smoothing of 3 bands. The noise profile was obtained from a brief non-vocal segment preceding the roar in each recording using the *Get Noise Profile* step, followed by application of the reduction to the full recording using the *Reduce* step.

2.2.2. Signal Segmentation

Each audio recording was standardized to a duration of 2-second so that all samples contain the same amount of data, i.e., $N = f_s \times T$ samples for sampling rate f_s and duration T. Segment selection was performed manually in Audacity. Each recording was opened, played back, and inspected in the waveform view, and a 2-second region containing the clearest roar was selected and cut. No quantitative selection criterion was applied. Because the selection was made by a single annotator, it may introduce subjective selection bias. The selected 2-second segments are provided in the Supplementary Material.

2.2.3. Stereo to Mono Conversion

Audio recordings with two channels consist of a left and right channel. The mono signal is obtained by averaging both channels (Eq. (2)).

$$x_{mono}(n) = \frac{x_L(n) + x_R(n)}{2} \tag{2}$$

where $x_{mono}(n)$ is the mono signal, $x_L(n)$ is the left channel, and $x_R(n)$ is the right channel.

Raw bioacoustic recordings often exhibit heterogeneous channel configurations, consisting of both mono and stereo formats. To maximize analysis consistency and ensure optimal feature extraction, all stereo audio files were converted to mono format. This standardization eliminates spatial channel redundancy, enabling the subsequent signal processing stages to focus strictly on a single, unified acoustic time series representation.

2.2.4. Sampling Rate Standardization

Differences in sampling rate change the number of samples per second and, through Eq. (4), the frequency band covered by each DWT sub-band. The most frequent sampling rate in the collected recordings was 96,000 Hz, which was therefore adopted as the common rate. In Audacity (v3.7.8) the Project Rate was set to 96,000 Hz, and recordings with a different original rate (including some at 48,000 Hz) were automatically resampled to this rate by Audacity's built-in sample-rate converter, using the high-quality conversion settings Sample Rate Converter = Best Quality (Slowest) and Dither = Shaped. The original sampling rate of each recording was not

logged during this step; recordings below 96,000 Hz were therefore upsampled, which adds no spectral content above their original Nyquist frequency. The relationship between the number of samples and the sampling frequency is given in Eq. (3):

$$f_s = \frac{N}{T} \tag{3}$$

where f_s is the sampling frequency, N is the number of samples, and T is the signal duration.

2.2.5. Audio Normalization

Normalization is a signal processing procedure aimed at standardizing the amplitude scale of audio signals within a specified range, thereby minimizing amplitude variations caused by differing recording conditions. The maximum amplitude value of each signal is adjusted by scaling the amplitude interval to the range of -1 to 1 [3], ensuring that all data share a uniform scale and that amplitude level discrepancies do not distort the signal characteristics. The normalization formula is defined as follows [17]:

$$X_{norm}(t) = \frac{x(t)}{max|x(t)|}$$

where x_t is the amplitude value of the original signal, and $max|x(t)|$ is the maximum amplitude value of the original signal.

2.3. Feature Extraction

2.3.1. Fundamental Frequency Verification

This stage aims to identify the highest fundamental frequency values across the 100 audio samples analyzed. The F_0 estimation was determined through frequency spectrum analysis by capturing the first prominent peak from each recording, thereby establishing the maximum fundamental frequency profile for each vocalization.

To ensure that the analyzed signals were free from acoustic anomalies, the reported fundamental frequency range of the lion, spanning 40 to 200 Hz [9], served as the foundational baseline. This reference was adopted due to the lack of dedicated comparative studies establishing standardized, quantitative F_0 ranges for each individual species within the genus *Panthera*.

2.3.2. Discrete Wavelet Transform

The feature extraction stage is conducted using the Discrete Wavelet Transform (DWT) combined with the Higuchi Fractal Dimension (HFD) to obtain characteristic information from each audio signal. The pairing of DWT decomposition prior to HFD computation is methodologically motivated rather than arbitrary. Raw bioacoustic signals are broadband and non-stationary, containing overlapping components across the fundamental frequency, harmonics, and non-informative high-frequency noise. Computing HFD directly on such a broadband signal risks conflating complexity contributions from biologically irrelevant frequency regions with the species-discriminative vocal patterns concentrated in lower-frequency bands. To address this, the process involves multi-level wavelet decomposition from level 1 to level 7 using the Daubechies 4 (db4) mother wavelet which has been shown to perform best in capturing bioacoustic signal characteristics [4]. By decomposing the signal into successive approximation sub-band via DWT, HFD can instead be computed at each frequency scale independently, yielding a multi-resolution complexity profile that captures species-specific characteristics across the physiologically relevant frequency range, as confirmed by the empirical F_0 verification. This staged approach is also consistent with prior HFD-based bioacoustic classification studies [3]-[8], which similarly combine DWT and HFD for feature extraction.

The selection of the 7 decomposition levels is not arbitrary. The upper frequency limit of each level is given by Eq. (4):

$$f_{maxj} = \frac{f_s}{2^{j+1}} \tag{4}$$

where f_s is the sampling rate (Hz), j is the DWT decomposition level, and f_{maxj} is the upper frequency limit of the approximation coefficients at jth level.

Accordingly, the frequency band distribution for DWT is presented in Table 1:

Table 1: DWT Frequency Sub-band Distribution

DWT Level	Component	Frequency Range (Hz)	Description
0	A0 (Original Signal)	0-48000	Signal Before Decomposition
1	A1	0-24000	Approximation level 1
2	A2	0-12000	Approximation level 2
3	A3	0-6000	Approximation level 3
4	A4	0-3000	Approximation level 4
5	A5	0-1500	Approximation level 5
6	A6	0-750	Approximation level 6
7	A7	0-375	Approximation level 7

DWT is performed by transforming the discrete signal into wavelet coefficients through successive filtering stages using a low-pass filter and a high-pass filter [18]. In general, the scale parameter a and translation parameter b are expressed in discrete form as $a = 2^m$ and $b = n2^m$, where m denotes the scale parameter and n denotes the translation parameter [19]. Based on these parameters, the DWT function is defined in Eq. (5):

$$DWT(a,b) = \frac{1}{\sqrt{|2^m|}} \sum_{t=-\infty}^{\infty} \psi(\frac{t - n2^m}{2^m}) f(t) \tag{5}$$

where ψ is the mother wavelet, 2^m is the scale (dilation) parameter, $n2^m$ is the translation parameter, and $f(t)$ is the original signal in the time domain.

2.4. Higuchi Fractal Dimension

A fractal dimension is an index for characterizing fractal pattern or sets by quantifying their complexity as a ratio of the change in detail to the change in scale [20]. Different from topological dimension, the fractal dimension can take non-integer values. For instance, a curve with a fractal dimension close to 1 looks like an ordinary line, whereas a curve with a fractal dimension close to 2 resembles a string of threads that fills the surface [21]. The Higuchi Fractal Dimension (HFD), originally introduced by Higuchi [22], is one of the algorithms used to calculate fractal dimension in the form of time series data. In many cases, Higuchi's fractal dimension is more appropriate than the spectral exponent to analyze an irregular time series. The notation and computational steps presented in this section follow the adaptation used by Juniati et al. [23], which is consistent with the original formulation.

Given a one-dimensional time series $X = x[1], x[2], ..., x[N]$. The algorithm for calculating the Higuchi fractal dimension is described as follows [23].

1. Set a new time series of k, X_m^k defined as follows:

$$X_m^k = x[m], x[m+k], x[m+2k], ..., x[m + int(\frac{N-m}{k})k]$$

where k and m are integers, k indicates discrete time intervals, while m = 1,2,...,k, indicates the value of the initial time [23].

2. The length of each new time series can be defined as follows:

$$L(m,k) = \frac{(\sum_{i=1}^{int(\frac{N-m}{k})} |x[m+ik] - x[m+(i-1).k]|)\frac{N-1}{int(\frac{N-m}{k})k}}{k}$$

where N is the length of the original time series X, and $|x[m+ik] - x[m+(i-1)k]| = h_i$. Thus $L(m,k)$ is the normalized number of segment length h^i indicate different distance values on the coordinate pair of points as far as k, starting on samples to m,$x[m]$ with $m = 1, 2, ..., k$ [23].

3. The length of the curve for the time interval k obtained by dividing all subseries $L(m,k)$ with k. For $m = 1, 2, ..., k$

$$L(k) = \frac{\sum_{m=1}^{k} L(m,k)}{k}$$

4. The Higuchi fractal dimension is then defined as the slope of the regression line fitted to the pairs $(\ln(1/k), \ln(L(k)))$ for $k = 1, 2, \ldots, k_{\max}$, since $L(k)$ is expected to follow a power-law relationship with k:

$$L(k) \propto k^{-D}$$

Taking the natural logarithm of both sides gives a linear relationship:

$$\ln(L(k)) = -D\ln(k) + c = D\ln\left(\frac{1}{k}\right) + c$$

where c is a constant (intercept). The slope D is estimated using the least-squares method over the $k_{\max}$ pairs of $(x_k, y_k) = \left(\ln\left(\frac{1}{k}\right), \ln(L(k))\right)$:

$$D = \frac{k_{\max}\sum_{k=1}^{k_{\max}} x_k y_k - \sum_{k=1}^{k_{\max}} x_k \sum_{k=1}^{k_{\max}} y_k}{k_{\max}\sum_{k=1}^{k_{\max}} x_k^2 - \left(\sum_{k=1}^{k_{\max}} x_k\right)^2}$$

where D is the Higuchi fractal dimension, obtained as the slope of the best-fit line over all $k = 1, \ldots, k_{\max}$ [23].

2.5. Classification

K-Nearest Neighbor (KNN) is a classification method that assigns a test sample to the class most common among its nearest training samples [24]. The Euclidean distance between two points is defined as follows:

$$D(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

where D is the distance between a test point x and a training point y. The KNN algorithm proceeds as follows:

1. Determine the number of nearest neighbors k.
2. Calculate the Euclidean distance between the test sample and every training sample.
3. Sort the training samples by increasing distance.
4. Select the k nearest training samples.
5. Assign the test sample to the majority class among these k neighbors.

In this study, odd values of k (1, 3, 5, 7, and 9) were selected to avoid tied votes among neighbors [25].

2.6. Accuracy

Accuracy is one of the performance indicators used to measure the precision of a classification model in predicting data classes. To evaluate the model's performance, a confusion matrix can be utilized, which is an evaluation method that compares the model's predictions against the true labels in the test data [26].

From the confusion matrix, the accuracy value is derived as a metric to assess the classification model's performance. For a multi-class classification problem involving C classes, accuracy is defined as the proportion of correctly classified instances relative to the total number of instances, which can be mathematically formulated as follows:

$$Accuracy = \frac{\sum_{i=1}^{C} TP_i}{N} \times 100\%$$

where TP_i is the number of correctly classified instances for class i, C is the total number of classes (in this study, $C = 5$), and N is the total number of instances in the test set. This is equivalent to the sum of the diagonal entries of the confusion matrix divided by the total number of instances.

For each individual class, precision, recall, and F1-score are computed using a one-vs-rest approach, treating class i as the positive class and all remaining classes as negative:

$$Precision_i = \frac{TP_i}{TP_i + FP_i}$$

$$Recall_i = \frac{TP_i}{TP_i + FN_i}$$

$$F1\text{-}Score_i = 2 \times \frac{Precision_i \times Recall_i}{Precision_i + Recall_i}$$

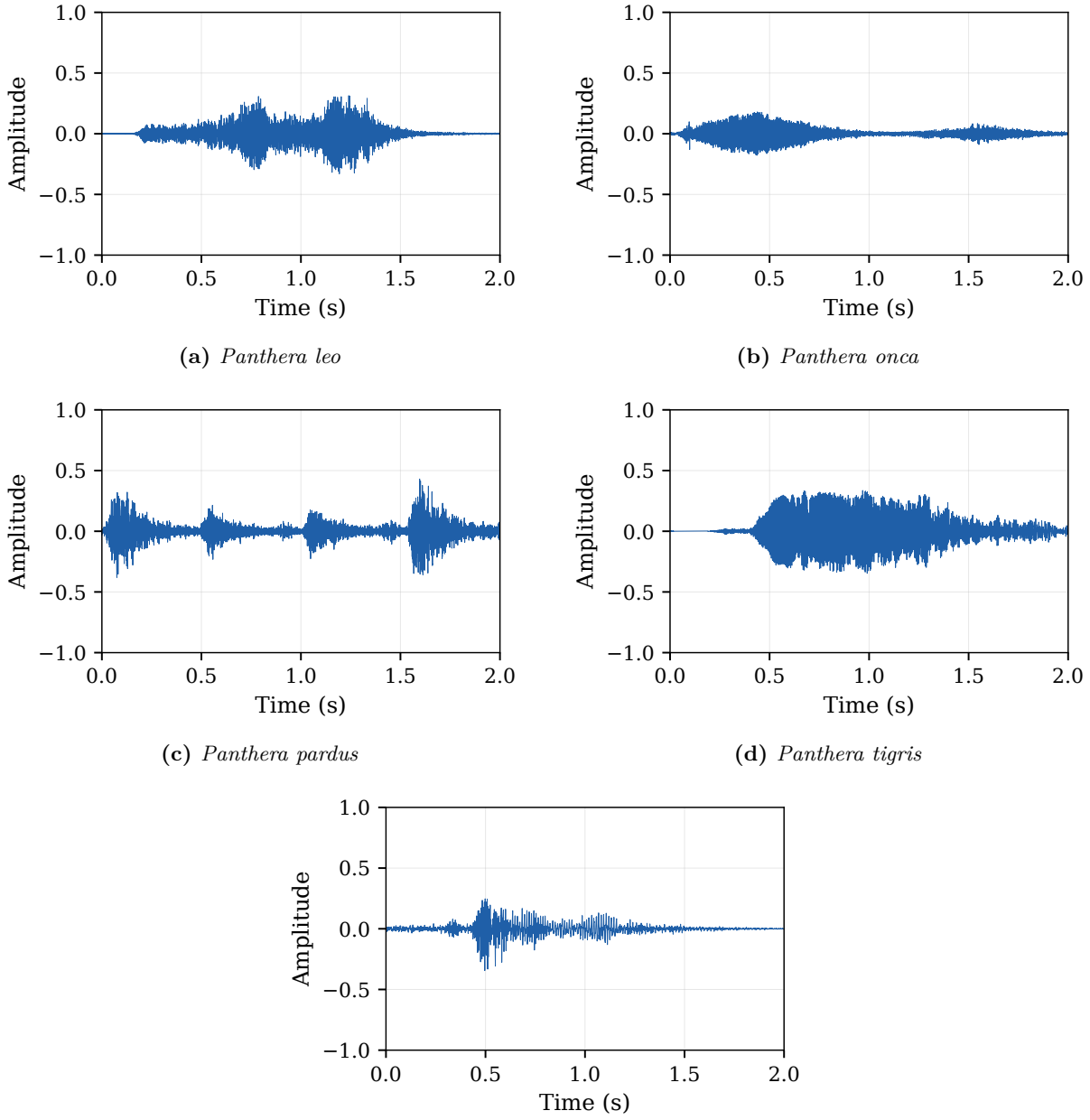
where TP_i is the number of instances correctly predicted as class i, FP_i is the number of instances incorrectly predicted as class i, and FN_i is the number of instances belonging to class i that were actually class i but incorrectly predicted as a different class.

3. Results and Discussion

This section elaborates in detail the stages of classifying species within the genus Panthera based on their acoustic vocalizations. The classification process in this study is conducted by computing fractal dimensions using the Higuchi method to extract the characteristic features of the audio signals. The obtained fractal dimension values are subsequently utilized as features for the data clustering and classification process using the K-Nearest Neighbor (KNN) algorithm. Through these stages, the proposed system is expected to effectively identify and distinguish species of the genus Panthera with higher accuracy based on their distinct vocal patterns.

3.1. Data Selection Process

The data utilized in this study are secondary animal vocalization data obtained from the Tierstimmenarchiv repository. The selected dataset consists of 100 audio samples from the genus *Panthera*, evenly distributed across five species: 20 lion vocalizations (*Panthera leo*), 20 jaguar vocalizations (*Panthera onca*), 20 leopard vocalizations (*Panthera pardus*), 20 tiger vocalizations (*Panthera tigris*), and 20 snow leopard vocalizations (*Panthera uncia*). The recordings were manually selected from over 300 available audio files in the repository. The selection was specifically restricted to audio files containing continuous roaring vocalizations to facilitate the continuity of subsequent analytical stages in the study. Figure 6 shows one example of the selected 2-second segment for each species before amplitude normalization.



(a) *Panthera leo*

(b) *Panthera onca*

(c) *Panthera pardus*

(d) *Panthera tigris*

(e) *Panthera uncia*

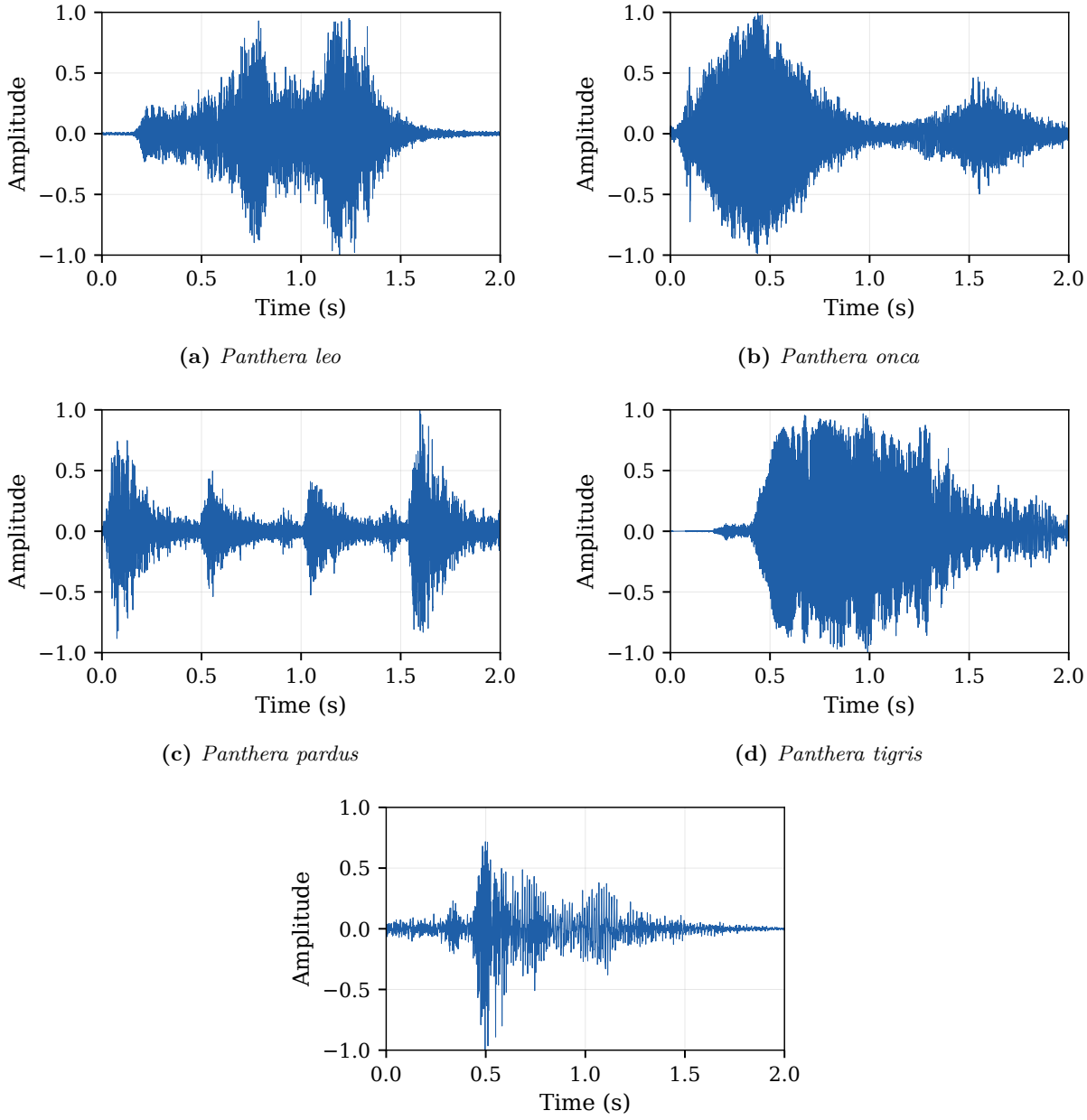
Fig. 6: Waveforms of the selected 2-second roar segments before amplitude normalization: (a) *Panthera leo*; (b) *Panthera onca*; (c) *Panthera pardus*; (d) *Panthera tigris*; (e) *Panthera uncia*.

3.2. Audio Signal Preprocessing

The acoustic recordings exhibit varying durations and recording parameters. All preprocessing steps except amplitude normalization were carried out in Audacity (v3.7.8) in the following order: noise reduction (Section 2.2.1), manual selection of a 2-second roar segment (Section 2.2.2), conversion from stereo to mono (Section 2.2.3), and resampling to 96,000 Hz (Section 2.2.4). Amplitude normalization to the range -1 to 1 was then applied in Python using Visual Studio Code (v1.137.0). Figure 7 shows the same segments as Figure 6 after amplitude normalization.

3.3. Feature Extraction Process

In the feature extraction stage, frequency spectrum analysis was performed on all 100 preprocessed audio samples to determine the fundamental frequency (F_0) and empirically justify the chosen DWT decomposition level. This analysis was conducted using the Frequency Analysis feature in



(a) *Panthera leo*

(b) *Panthera onca*

(c) *Panthera pardus*

(d) *Panthera tigris*

(e) *Panthera uncia*

Fig. 7: The segments in Figure 6 after amplitude normalization to the range −1 to 1.

Audacity, employing the Enhanced Autocorrelation algorithm, a Hann window function, and a window size of 8192. The Hann window and the 8192 size were selected to provide clearer waveform visualization and finer harmonic resolution, although the window size remains adaptable and can be scaled larger depending on the researcher's analytical preferences. For each sample, the fundamental frequency was identified from the first prominent peak in the autocorrelation spectrum representing the lowest frequency component of the harmonic series and recorded directly from the Peak indicator.

The analysis revealed that the fundamental frequency across the five *Panthera* species ranged from 53 Hz to 367 Hz. Specifically, the observed F_0 ranges for each species extracted and analyzed using Altair AI Studio (formerly RapidMiner Studio) software (v2026.1.1) were 151-328 Hz for *Panthera leo*, 139-342 Hz for *Panthera onca*, 146-299 Hz for *Panthera pardus*, 174-367 Hz for *Panthera tigris*, and 53-291 Hz for *Panthera uncia*. The highest fundamental frequency in the entire dataset was 367 Hz (observed in a *Panthera tigris* sample), while the lowest was recorded

from a *Panthera uncia* sample.

The observed maximum value of 367 Hz critically validates the decomposition depth selected for the Discrete Wavelet Transform (DWT). As indicated in Table 1, level 7 decomposition yields an approximation sub-band of 0–375 Hz. Because the empirical maximum F_0 (367 Hz) falls safely below this 375 Hz upper boundary, level 7 is confirmed to adequately retain the complete fundamental frequency content across all five species without spectral truncation.

Following this spectral verification, the feature extraction process was advanced using DWT across 7 decomposition levels with the Daubechies 4 (db4) mother wavelet. Below are the audio signal representations resulting from decomposition levels 1 through 7, analyzed and computed using Visual Studio Code (VS Code). Figures 8–12 show the resulting coefficients for *Panthera leo*, *Panthera onca*, *Panthera pardus*, *Panthera tigris*, and *Panthera uncia*, respectively:

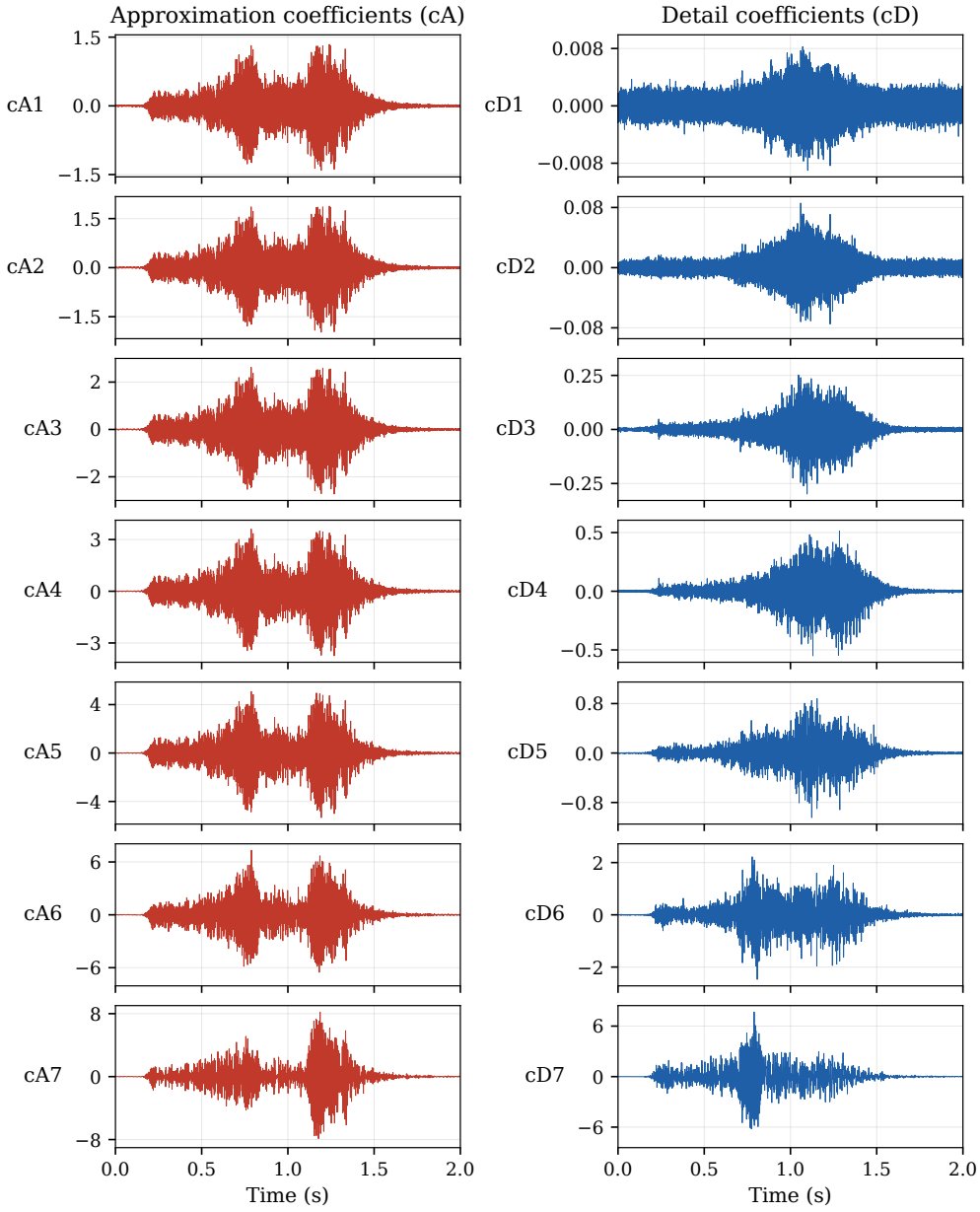


Fig. 8: DWT approximation (cA_1–cA_7) and detail (cD_1–cD_7) coefficients of a *Panthera leo* roar (db4, 7 levels).

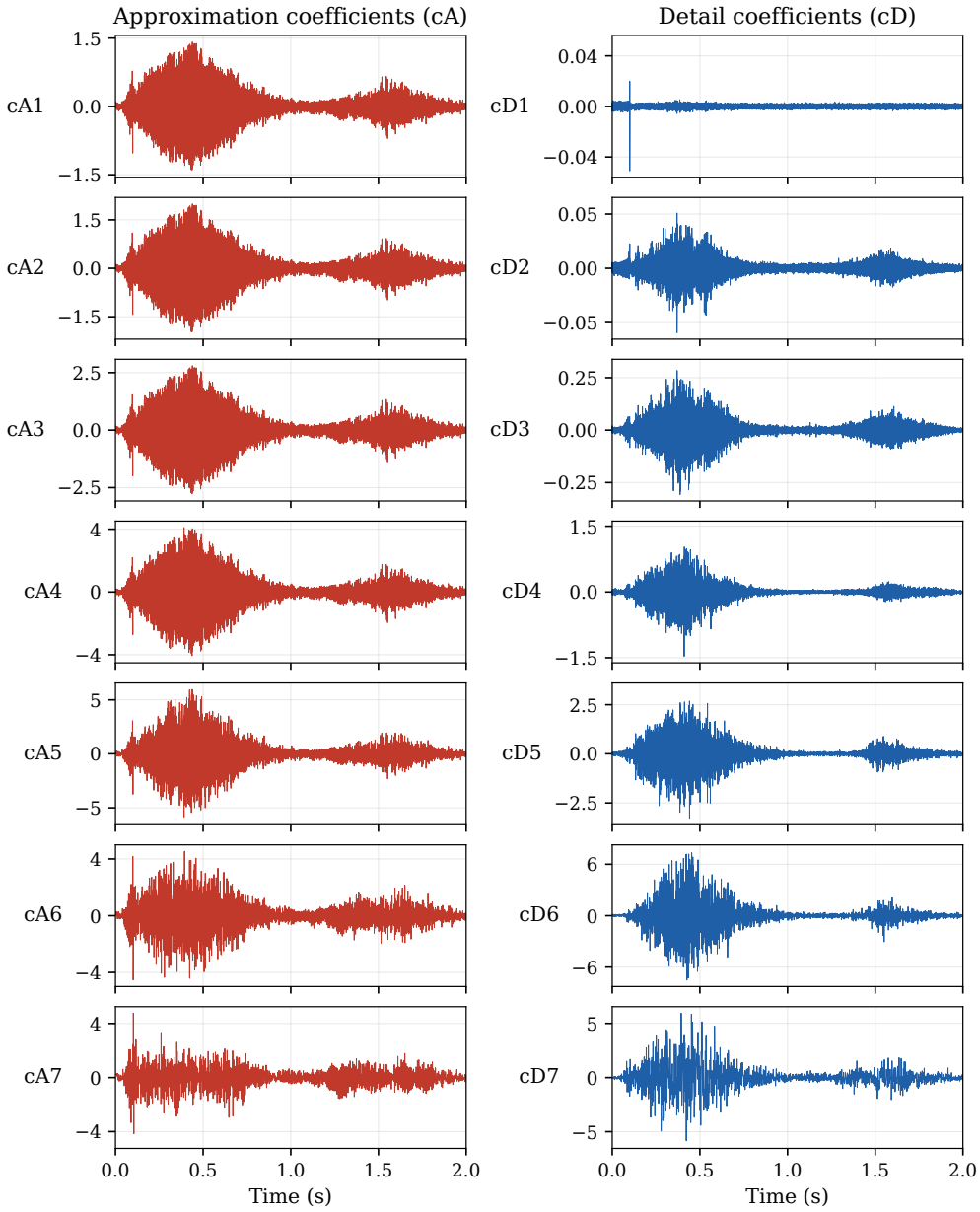


Fig. 9: DWT approximation (cA_1–cA_7) and detail (cD_1–cD_7) coefficients of a *Panthera onca* roar (db4, 7 levels).

3.4. Higuchi Method

After obtaining the approximation and detail coefficients from the DWT decomposition encompassing both the low-pass filter coefficients (cA_1 to cA_7) and high-pass filter coefficients (cD_1 to cD_7) the subsequent stage is computing the fractal dimensions using the Higuchi method. The calculation process begins by defining the $K_{\max}$ value as the primary parameter. Based on established literature references, $K_{\max}$ is evaluated comprehensively across a broader range rather than limited to a few specific values. In this study, $K_{\max}$ values of 10, 16, 28, 32, 50, and 60 are utilized. This set of candidate $K_{\max}$ values was informed by prior HFD-based bioacoustic classification studies within this research line: $K_{\max}$ values of 16, 28, 32, 50, and 60 were previously employed by Juniati et al. [23] in the classification of heart sound recordings, while $K_{\max} = 10$ was adopted from a related study on infant cry classification [27]. Rather than adopting a single value from either study, this study retains the combined set of candidates and determines the optimal $K_{\max}$ empirically for Panthera vocalizations through the nested

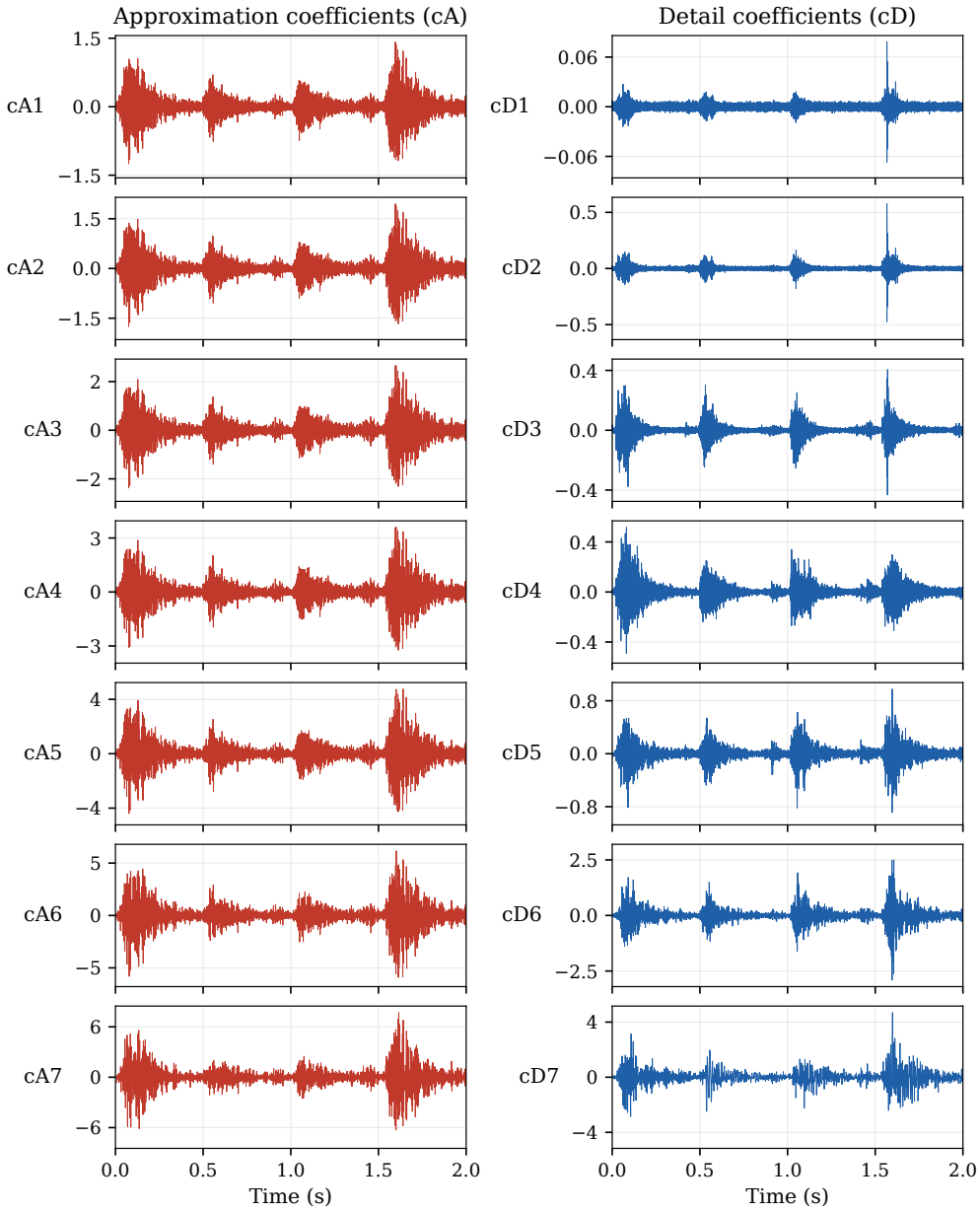


Fig. 10: DWT approximation (cA_1–cA_7) and detail (cD_1–cD_7) coefficients of a *Panthera pardus* roar (db4, 7 levels).

cross-validation procedure described in Section 3.5. The candidate k values for KNN (1, 3, 5, 7, 9) similarly follow common practice in this line of research to avoid tied votes among neighbors [25], with the final k likewise selected empirically rather than fixed in advance. All computations and algorithmic executions of the Higuchi fractal dimension were carried out using Visual Studio Code (VS Code). Tables 2–13 list the HFD values of one example recording per species for each $K_{\max}$ and coefficient type, and Table 14 lists the non-DWT values.

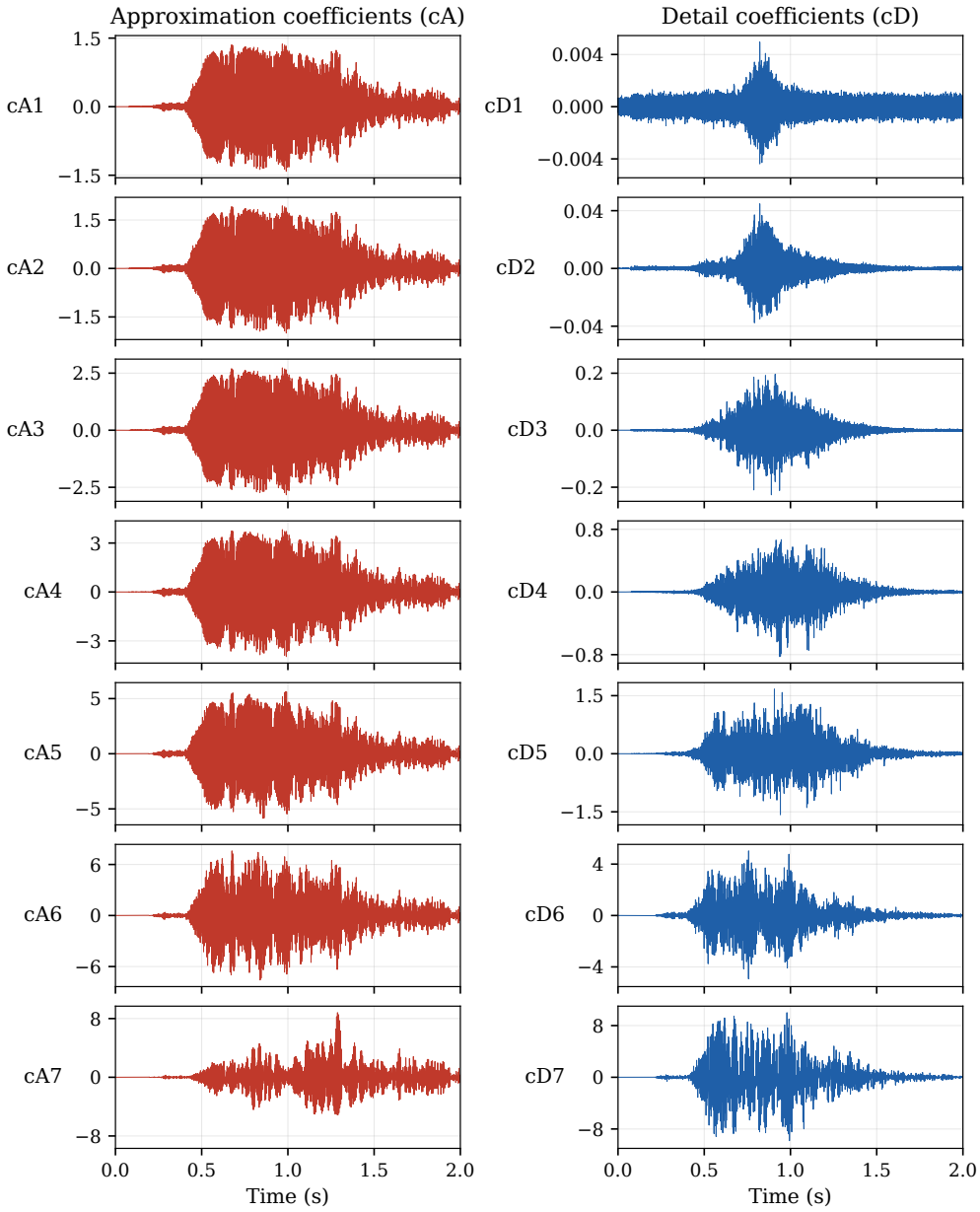


Fig. 11: DWT approximation (cA_1–cA_7) and detail (cD_1–cD_7) coefficients of a *Panthera tigris* roar (db4, 7 levels).

Table 2: Higuchi Fractal Dimension Calculation with $K_{\max}$ 10 (Low-Pass Filter Coefficient (cA))

i-th Data	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Label
1	1.090	1.141	1.203	1.278	1.438	1.765	1.901	PL
21	1.121	1.295	1.801	2.000	2.043	1.999	1.949	PO
41	1.035	1.097	1.274	1.518	1.766	1.961	2.050	PP
61	1.167	1.318	1.576	1.807	2.043	2.076	2.040	PT
81	1.026	1.069	1.150	1.318	1.654	1.970	2.074	PU

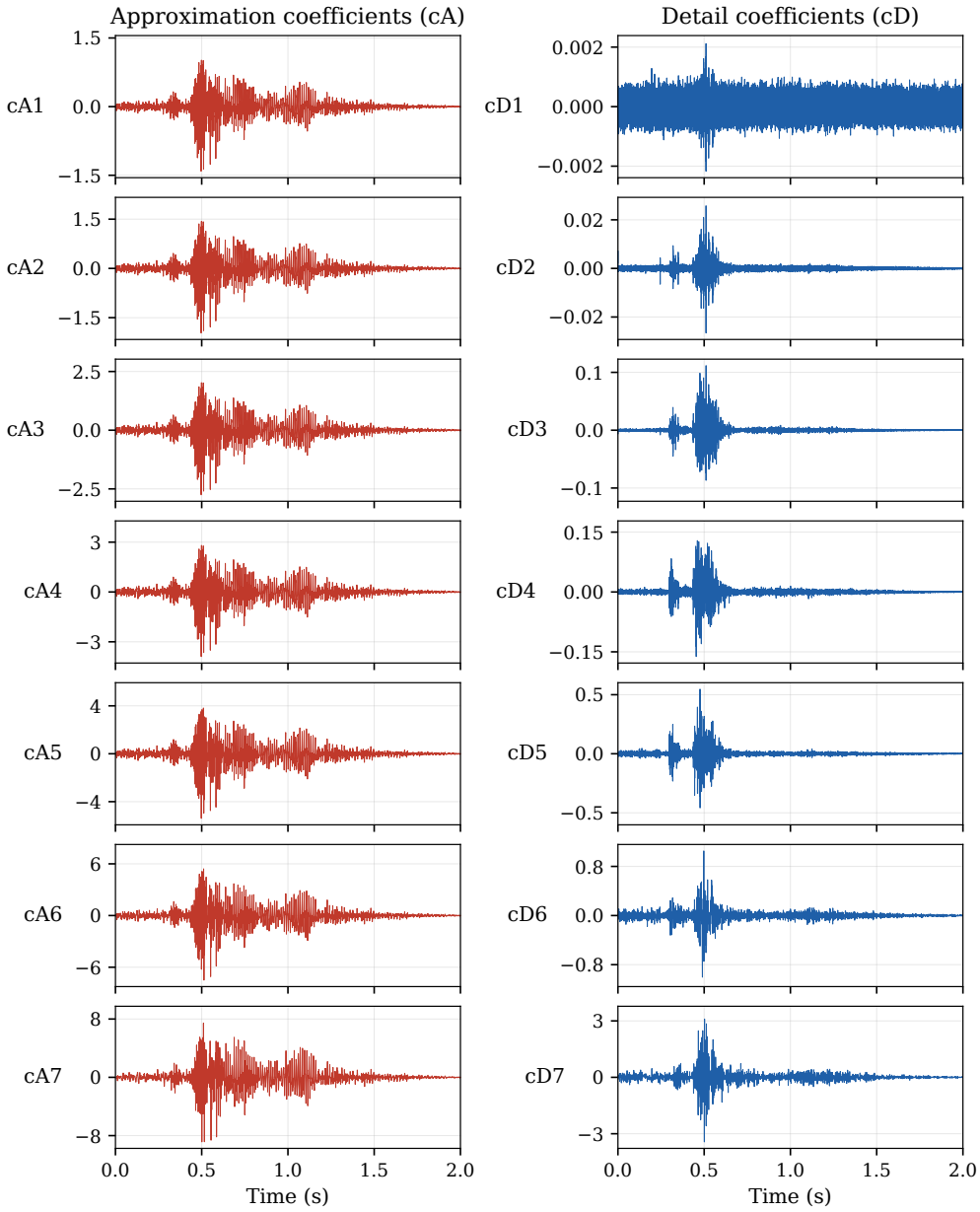


Fig. 12: DWT approximation (cA_1–cA_7) and detail (cD_1–cD_7) coefficients of a *Panthera uncia* roar (db4, 7 levels).

Table 3: Higuchi Fractal Dimension Calculation with $K_{\max}$ 10 (High-Pass Filter Coefficient (cD))

i-th Data	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Label
1	1.999	2.054	2.041	2.039	2.029	2.045	2.030	PL
21	2.053	2.063	2.040	2.044	2.038	1.962	2.003	PO
41	2.019	2.028	2.003	2.034	2.043	2.038	1.952	PP
61	1.890	2.050	2.049	2.050	2.039	2.051	2.044	PT
81	1.846	2.050	2.047	2.074	2.059	2.053	2.049	PU

Table 4: Higuchi Fractal Dimension Calculation with $K_{\max}$ 16 (Low-Pass Filter Coefficient (cA))

i-th Data	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Label
1	1.115	1.169	1.241	1.350	1.652	1.866	2.006	PL
21	1.188	1.561	1.848	1.997	2.043	1.985	1.967	PO
41	1.061	1.172	1.420	1.607	1.895	1.994	2.035	PP
61	1.235	1.436	1.699	1.980	2.030	2.043	2.018	PT
81	1.046	1.104	1.226	1.451	1.868	1.980	2.045	PU

Table 5: Higuchi Fractal Dimension Calculation with $K_{\max}$ 16 (High-Pass Filter Coefficient (cD))

i-th Data	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Label
1	1.994	2.032	2.020	2.021	2.014	2.016	2.001	PL
21	2.022	2.031	2.016	2.023	2.042	1.992	2.047	PO
41	2.013	2.001	2.007	2.017	2.028	2.021	1.938	PP
61	1.926	2.031	2.031	2.028	2.011	2.019	2.016	PT
81	1.897	2.029	2.034	2.035	2.032	2.026	2.035	PU

Table 6: Higuchi Fractal Dimension Calculation with $K_{\max}$ 28 (Low-Pass Filter Coefficient (cA))

i-th Data	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Label
1	1.146	1.214	1.306	1.521	1.743	1.907	1.991	PL
21	1.409	1.750	1.916	2.013	2.023	1.992	1.985	PO
41	1.125	1.331	1.495	1.803	1.962	1.996	2.010	PP
61	1.350	1.611	1.878	1.985	2.018	2.015	1.997	PT
81	1.081	1.175	1.358	1.734	1.917	2.007	2.021	PU

Table 7: Higuchi Fractal Dimension Calculation with $K_{\max}$ 28 (High-Pass Filter Coefficient (cD))

i-th Data	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Label
1	1.994	2.015	2.008	2.008	2.004	2.006	1.985	PL
21	2.009	2.005	1.993	2.025	2.022	2.006	2.010	PO
41	1.993	1.992	1.993	2.012	2.015	2.002	1.931	PP
61	1.954	2.017	2.015	2.013	2.002	2.011	1.995	PT
81	1.935	2.014	2.018	2.017	2.015	2.016	2.018	PU

Table 8: Higuchi Fractal Dimension Calculation with $K_{\max}$ 32 (Low-Pass Filter Coefficient (cA))

i-th Data	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Label
1	1.156	1.225	1.329	1.593	1.805	1.951	1.998	PL
21	1.492	1.768	1.926	2.009	2.012	1.990	1.991	PO
41	1.150	1.369	1.543	1.826	1.956	1.997	2.004	PP
61	1.389	1.641	1.910	1.984	2.013	2.011	1.995	PT
81	1.092	1.199	1.402	1.783	1.916	2.001	2.018	PU

Table 9: Higuchi Fractal Dimension Calculation with $K_{\max}$ 32 (High-Pass Filter Coefficient (cD))

i-th Data	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Label
1	1.993	2.014	2.006	2.007	2.002	2.004	1.987	PL
21	2.005	1.998	1.992	2.020	2.016	2.000	2.006	PO
41	2.001	1.993	1.996	2.010	2.012	1.998	1.929	PP
61	1.959	2.015	2.012	2.010	2.001	2.009	1.992	PT
81	1.942	2.013	2.016	2.014	2.014	2.014	2.015	PU

Table 10: Higuchi Fractal Dimension Calculation with $K_{\max}$ 50 (Low-Pass Filter Coefficient (cA))

i-th Data	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Label
1	1.192	1.272	1.442	1.671	1.861	1.950	1.997	PL
21	1.676	1.848	1.946	1.997	2.004	1.991	1.992	PO
41	1.264	1.430	1.715	1.893	1.974	1.992	1.987	PP
61	1.533	1.768	1.919	1.985	2.005	1.998	1.985	PT
81	1.142	1.298	1.624	1.845	1.945	1.996	2.005	PU

Table 11: Higuchi Fractal Dimension Calculation with $K_{\max}$ 50 (High-Pass Filter Coefficient (cD))

i-th Data	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Label
1	1.994	2.007	2.000	2.000	1.997	1.997	1.983	PL
21	1.997	1.982	2.001	2.011	2.008	2.001	1.989	PO
41	2.003	1.987	1.999	2.005	2.004	1.987	1.931	PP
61	1.972	2.008	2.005	2.002	1.996	2.000	1.978	PT
81	1.961	2.009	2.010	2.008	2.007	2.006	2.005	PU

Table 12: Higuchi Fractal Dimension Calculation with $K_{\max}$ 60 (Low-Pass Filter Coefficient (cA))

i-th Data	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Label
1	1.207	1.300	1.524	1.729	1.884	1.958	1.987	PL
21	1.704	1.870	1.970	2.000	2.002	1.991	1.990	PO
41	1.321	1.478	1.760	1.917	1.970	1.987	1.979	PP
61	1.582	1.846	1.945	1.988	2.004	1.994	1.977	PT
81	1.171	1.348	1.704	1.864	1.960	1.999	2.002	PU

Table 13: Higuchi Fractal Dimension Calculation with $K_{\max}$ 60 (High-Pass Filter Coefficient (cD))

i-th Data	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Label
1	1.994	2.005	1.998	1.998	1.994	1.995	1.977	PL
21	1.994	1.978	2.007	2.011	2.004	1.997	1.982	PO
41	2.003	1.987	1.998	2.004	2.001	1.983	1.924	PP
61	1.976	2.005	2.002	2.000	1.996	1.996	1.970	PT
81	1.967	2.008	2.008	2.006	2.006	2.005	2.001	PU

Table 14: Higuchi Fractal Dimension Calculation Directly from the Preprocessed Signal without DWT (Raw HFD) for Each $K_{\max}$

i-th Data	$K_{\max}$ 10	$K_{\max}$ 16	$K_{\max}$ 28	$K_{\max}$ 32	$K_{\max}$ 50	$K_{\max}$ 60	Label
1	1.049	1.068	1.097	1.104	1.129	1.141	PL
21	1.072	1.092	1.143	1.167	1.312	1.410	PO
41	1.021	1.027	1.047	1.056	1.099	1.126	PP
61	1.069	1.115	1.188	1.208	1.293	1.341	PT
81	1.009	1.016	1.034	1.040	1.066	1.079	PU

where PL denotes *Panthera leo*, PO denotes *Panthera onca*, PP denotes *Panthera pardus*, PT denotes *Panthera tigris*, and PU denotes *Panthera uncia*.

3.5. Classification Process

This study adopts a repeated nested stratified cross-validation design that separates hyperparameter selection from performance estimation. The outer loop (5-fold, repeated 10 times, random

seed = 873) estimates generalization performance on held-out data. Within each outer training fold, an inner loop (5-fold, random seed = 681) selects the $K_{\max}$ and k combination (from $K_{\max} \in \{10, 16, 28, 32, 50, 60\}$ and $k \in \{1, 3, 5, 7, 9\}$) that maximizes mean validation accuracy, using only the outer training partition; the outer test partition is never accessed during this selection step. The selected configuration is then refit on the full outer training partition and evaluated once on the corresponding outer test partition. This procedure yields 50 outer-fold accuracy estimates in total, from which the mean, standard deviation, and 95% confidence interval are reported. Both random seeds were fixed prior to inspecting the results.

In addition to this validation design, an ablation study was conducted to compare five feature groups. Four groups are derived from the DWT approximation (cA) and detail (cD) coefficients, in order to assess which coefficients contribute most to classification performance. The fifth group is a non-DWT baseline (raw HFD), in which the Higuchi fractal dimension is computed directly from the preprocessed 2-second signal without any wavelet decomposition, using the same $K_{\max}$ candidates (Table 14); this group therefore consists of a single feature per recording and isolates the contribution of the DWT decomposition. The entire classification workflow was implemented in Python (v3.12.7) using the scikit-learn library (v1.9.0), and the resulting accuracies are reported in Table 15:

Table 15: Repeated Nested Cross-Validation Results Across Feature Groups (50 outer-folds; outer seed = 873, inner seed = 681)

Feature Group	#Features	Mean Acc. (%)	Std (%)	95% CI (%)	Chosen $K_{\max}$ / k (mode)
cD_1–cD_7 + cA_5–cA_7	10	**70.80**	8.77	±2.43	10 / 1
cD_1–cD_7	7	67.10	10.65	±2.95	28 / 1
cA_1–cA_7 + cD_1–cD_7	14	61.70	10.86	±3.01	16 / 1
cA_1–cA_7	7	61.50	11.12	±3.08	16 / 1
Raw HFD (without DWT)	1	56.40	10.10	±2.80	32 / 1

As shown in Table 15, using repeated nested cross-validation, the highest mean accuracy of 70.80% (SD = 8.77%, 95% CI ±2.43%) was obtained using the feature group combining all detail coefficients (cD_1–cD_7) with the three highest-level approximation coefficients (cA_5–cA_7), with $K_{\max} = 10$ and $k = 1$ most frequently selected across the 50 outer-folds. The ablation results further show that detail coefficients alone outperform approximation coefficients alone (67.10% vs. 61.50%), while combining all approximation and detail coefficients together does not improve performance (61.70%), suggesting that lower-level approximation coefficients introduce redundant or noisy information rather than discriminative signal.

The non-DWT baseline reached a mean accuracy of 56.40% (SD = 10.10%, 95% CI ±2.80%), which is 14.40 percentage points below the best DWT-based feature group and also below every DWT-based group, including the approximation coefficients alone (61.50%). This gap exceeds the fold-to-fold uncertainty reflected in the reported confidence intervals (approximately 2.4–2.8 percentage points), indicating that computing the fractal dimension on the multi-level DWT sub-band provides discriminative information that a single HFD value of the full-band signal does not. It should be noted that this comparison contrasts a single scalar feature with multi-dimensional, multi-resolution feature vectors, so the observed gain reflects both the frequency scale decomposition and the larger feature dimensionality.

To confirm that the reported performance was not an artifact of a particular random seed choice, the nested cross-validation procedure for the best performing feature group was additionally repeated across 20 randomly sampled seed pairs (5 outer repeats each), drawn from a fixed master seed independent of the resulting accuracy. This yielded a mean of means accuracy of 69.89% with a standard deviation of only 1.06% across seed pairs, indicating that the reported result is stable with respect to the choice of random seed rather than dependent on a favorable configuration.

The detailed classification results corresponding to the best performing feature group, pooled

across all 50 outer-folds ($n = 1000$ predictions), are given in Table 16:

Table 16: Precision, Recall, F1-Score (pooled across 50 outer-folds, best-performing feature group)

Precision	Recall	F1-Score	Class
64.21%	61.00%	62.56%	PL
69.13%	79.50%	73.95%	PO
78.57%	82.50%	80.49%	PP
66.15%	63.50%	64.80%	PT
75.84%	67.50%	71.43%	PU
70.78%	70.80%	70.65%	Macro Avg

The class abbreviations used in Table 16 are as follows: PL denotes *Panthera leo*, PO denotes *Panthera onca*, PP denotes *Panthera pardus*, PT denotes *Panthera tigris*, and PU denotes *Panthera uncia*.

Based on the pooled confusion matrix (Table 17), the classification performance for each class can be evaluated as follows: for the *Panthera leo* class, a precision of 64.21% was achieved, as out of 190 samples predicted as this class across all folds, 122 were correct while 68 were misclassified from other species (16 from *Panthera onca*, 9 from *Panthera pardus*, 25 from *Panthera tigris*, and 18 from *Panthera uncia*), and a recall of 61.00% was obtained since out of 200 actual *Panthera leo* samples, 122 were correctly identified while 78 were misclassified (12 as *Panthera onca*, 10 as *Panthera pardus*, 41 as *Panthera tigris*, and 15 as *Panthera uncia*), resulting in an F1-score of 62.56%. For the *Panthera onca* class, the precision reached 69.13%, as out of 230 samples predicted as this class, 159 were correct while 71 were misclassified from other species (12 from *Panthera leo*, 24 from *Panthera tigris*, and 35 from *Panthera uncia*), and the recall reached 79.50% as 159 of the 200 actual samples were correctly identified while 41 were misclassified (16 as *Panthera leo*, 6 as *Panthera pardus*, 9 as *Panthera tigris*, and 10 as *Panthera uncia*), yielding an F1-score of 73.95%. For the *Panthera pardus* class, the precision reached 78.57%, as out of 210 samples predicted as this class, 165 were correct while 45 were misclassified from other species (10 from *Panthera leo*, 6 from *Panthera onca*, 22 from *Panthera tigris*, and 7 from *Panthera uncia*), while the recall reached the highest value among all classes at 82.50%, since only 35 of the 200 actual samples were misclassified (9 as *Panthera leo*, 10 as *Panthera tigris*, and 16 as *Panthera uncia*), producing an F1-score of 80.49%.

For the *Panthera tigris* class, the precision reached 66.15%, as out of 192 samples predicted as this class, 127 were correct while 65 were misclassified from other species (41 from *Panthera leo*, 9 from *Panthera onca*, 10 from *Panthera pardus*, and 5 from *Panthera uncia*), while the recall was the lowest among all classes at 63.50%, since 73 of the 200 actual samples were misclassified (25 as *Panthera leo*, 24 as *Panthera onca*, 22 as *Panthera pardus*, and 2 as *Panthera uncia*), yielding an F1-score of 64.80%. Lastly, for the *Panthera uncia* class, the precision reached 75.84%, as out of 178 samples predicted as this class, 135 were correct while 43 were misclassified from other species (15 from *Panthera leo*, 10 from *Panthera onca*, 16 from *Panthera pardus*, and 2 from *Panthera tigris*), while the recall reached 67.50% since 65 of the 200 actual samples were misclassified (18 as *Panthera leo*, 35 as *Panthera onca*, 7 as *Panthera pardus*, and 5 as *Panthera tigris*), resulting in an F1-score of 71.43%. Overall, these results produced a macro averaged precision of 70.78%, recall of 70.80%, and F1-score of 70.65%, consistent with the overall model accuracy of 70.80%.

Table 17: Pooled Confusion Matrix for the Best-Performing Feature Group ($n = 1000$ predictions across 50 outer-folds)

True \ Predicted	PL	PO	PP	PT	PU
PL	122	12	10	41	15
PO	16	159	6	9	10
PP	9	0	165	10	16
PT	25	24	22	127	2
PU	18	35	7	5	135

The confusion matrix reveals that the most frequent misclassifications occur between *Panthera leo* and *Panthera tigris*, and between *Panthera onca* and *Panthera uncia*. This pattern is broadly consistent with prior findings on laryngeal morphology, which suggest overlapping vocal complexity among closely related roaring species [9].

Although the nested cross-validation procedure separates hyperparameter selection from performance evaluation, it does not resolve the potential for data leakage arising from repeated individual animals appearing across folds, as individual-level metadata were not consistently available in the source repository. This remains an acknowledged limitation, and the reported accuracy should be interpreted as an upper bound under this constraint. Furthermore, the 2-second segments were selected manually by a single annotator, which may introduce selection bias that cannot be fully excluded. In addition, the comparison in this study is limited to HFD-based features; benchmarking against other conventional acoustic feature sets was beyond the scope of this work and is left for future research.

4. Conclusion

This study aims to apply the Higuchi Fractal Dimension (HFD) method for feature extraction and the K-Nearest Neighbor (KNN) algorithm for classification to recognize the vocal signals of five extant species within the genus *Panthera* (*Panthera leo*, *Panthera onca*, *Panthera pardus*, *Panthera tigris*, and *Panthera uncia*). Based on experimental results from 100 audio recordings processed through preprocessing, feature extraction, and Higuchi fractal dimension computation, classification performance is influenced by the $K_{\max}$ parameter in HFD, the k value in KNN, and the choice of feature subset. Using a repeated nested cross-validation design that separates hyperparameter selection from performance evaluation, the highest mean classification accuracy achieved was 70.80% (SD = 8.77%, 95% CI ±2.43%), obtained using detail coefficients (cD_1–cD_7) combined with the three highest-level approximation coefficients (cA_5–cA_7), with $K_{\max} = 10$ and $k = 1$ most frequently selected across 50 outer-folds. This result was stable with respect to random seed choice, with a mean accuracy of 69.89% and a standard deviation of 1.06% across 20 randomly sampled seed pairs.

The ablation study indicated that detail coefficients carry more discriminative information than approximation coefficients (67.10% vs. 61.50%), whereas combining all coefficients did not improve performance (61.70%), suggesting that lower-level approximation coefficients add redundant or noisy information. A non-DWT baseline, in which HFD was computed directly from the preprocessed signal, reached only 56.40%, which is 14.40 percentage points below the best DWT-based feature group. This indicates that the multi-level DWT decomposition contributes discriminative information beyond a single full-band HFD value. A notable methodological step in this study is the empirical verification of fundamental frequency prior to DWT decomposition, confirming that the observed F_0 range (53–367 Hz) falls within the 0–375 Hz sub-band of level-7 DWT decomposition, which provides a data-driven justification for the chosen decomposition depth.

Ultimately, these results indicate that HFD computed on DWT sub-band provides discriminative features that enable a KNN classifier to distinguish the vocalizations of species within the genus *Panthera* with moderate accuracy. However, because individual-level metadata were

unavailable, repeated individuals across folds cannot be ruled out, and the comparison was limited to HFD-based features; the reported accuracy should therefore be interpreted with caution.

CRediT Authorship Contribution Statement

Zharfanudin Afif: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Writing-Original Draft, Visualization, Editing. **Dwi Juniati**: Conceptualization, Methodology, Validation, Writing-Review & Editing, Visualization, Supervision.

Declaration of Competing Interest

The authors declare no competing interest.

Ethics Approval and Consent to Participate

Ethics approval was not required for this study.

Funding and Acknowledgments

This research received no external funding.

Data and Code Availability

The audio recordings analyzed in this study are publicly available in the Tierstimmenarchiv repository [14]. The Supplementary Material contains the list of recording identifiers for the 100 selected samples, the selected 2-second segments, and the Python source code for amplitude normalization, feature extraction, and classification. The remaining preprocessing steps were performed in Audacity as described in Section 2.2.

References

[1] M.D. Bjordal. "Why big cats are at high risk of extinction due to their exceptional predatory abilities. What conservation strategies are needed?" In: *USURJ: University of Saskatchewan Undergraduate Research Journal* 2.2 (2016). DOI: 10.32396/usurj.v2i2.122.

[2] Waeisul Bismi, Deny Novianti, and Muhammad Qomaruddin. "Analisis Perbandingan Klasifikasi Citra Genus Panthera dengan Pendekatan Deep Learning Model MobileNet". In: *Jurnal Informatika dan Rekayasa Perangkat Lunak* 6.1 (Mar. 2024), pp. 1–9.

[3] Lilis Mardiana and Dwi Juniati. "Penerapan Dimensi Fraktal Higuchi dan K-Nearest Neighbor (KNN) dalam Klasifikasi Kondisi Kucing Berdasarkan Suara". In: *MATHunesa: Jurnal Ilmiah Matematika* 11.2 (2023), pp. 82–91. DOI: 10.26740/mathunesa.v11n2.p82-91.

[4] Fika Yulistina and Dwi Juniati. "Penerapan Dimensi Fraktal Higuchi dalam Klasifikasi Jenis Monyet Berdasarkan Suara dengan Metode K-Nearest Neighbor (K-NN)". In: *MATHunesa: Jurnal Ilmiah Matematika* 12.1 (2024), pp. 110–120. DOI: 10.26740/mathunesa.v12n1.p110-120.

[5] Marisa Damayanti and Dwi Juniati. "Klasifikasi Jenis Anjing Laut Berdasarkan Suara dengan Metode Higuchi dan K-Nearest Neighbor". In: *MATHunesa: Jurnal Ilmiah Matematika* 11.2 (2023), pp. 102–111. DOI: 10.26740/mathunesa.v11n2.p102-111.

[6] Yolanda Gandes Pamela and Dwi Juniati. "Klasifikasi Jenis Delphinidae (Lumba-Lumba) dengan Dimensi Fraktal Menggunakan Metode Higuchi dan KNN (K-Nearest Neighbor)". In: *MATHunesa: Jurnal Ilmiah Matematika* 9.1 (2021), pp. 204–211. DOI: 10.26740/mathunesa.v9n1.p204-211.

[7] Putri Anis Nurul Yania and Dwi Juniati. "Penerapan Dimensi Fraktal Higuchi dan K-Nearest Neighbor dalam Klasifikasi Passeriformes (Burung Pengicau) Berdasarkan Suara". In: *MATHunesa: Jurnal Ilmiah Matematika* 11.2 (2023), pp. 92–101. DOI: 10.26740/mathunesa.v11n2.p92-101.

[8] Adin Nunasikhah and Dwi Juniati. "Klasifikasi Jenis Jangkrik Berdasarkan Suara Menggunakan Dimensi Fraktal Metode Higuchi dan K-Nearest Neighbor (KNN)". In: *MATHunesa: Jurnal Ilmiah Matematika* 10.1 (2022), pp. 199–207. DOI: 10.26740/mathunesa.v10n1.p199-207.

[9] Sarah A. Klemuk, Tobias Riede, Edward J. Walsh, and Ingo R. Titze. "Adapted to Roar: Functional Morphology of Tiger and Lion Vocal Folds". In: *PLOS ONE* 6.11 (Nov. 2011), e27029. DOI: 10.1371/journal.pone.0027029.

[10] Sarah K. Haas, Virginia Hayssen, and Paul R. Krausman. "Panthera leo". In: *Mammalian Species* 2005.762 (July 2005), pp. 1–11. DOI: 10.1644/1545-1410(2005)762[0001:PL]2.0.CO;2.

[11] Gerardo Ceballos et al. "Beyond words: From jaguar population trends to conservation and public policy in Mexico". In: *PLOS ONE* 16.10 (Oct. 2021), e0255555. DOI: 10.1371/journal.pone.0255555.

[12] Andrew B. Stein and Virginia Hayssen. "Panthera pardus (Carnivora: Felidae)". In: *Mammalian Species* 45.900 (June 2013), pp. 30–48. DOI: 10.1644/900.1.

[13] Vratislav Mazák. "Panthera tigris". In: *Mammalian Species* 152 (May 1981), pp. 1–8. DOI: 10.2307/3504004.

[14] Museum für Naturkunde Berlin. *Animal Sound Archive*. Occurrence dataset. Accessed via GBIF.org, August 2026. 2013. DOI: 10.15468/0bpalr.

[15] Satria Gunawan, Edi Suhardi Rahman, and Mardhatillah. "Metode Penentuan Perbaikan Noise Pada Data Musik Menggunakan Algoritma Least Mean Square". In: *Journal of Digital Technology and Computer Science* 1.1 (2023), pp. 48–56. DOI: 10.61220/digitech.v1i1.20235.

[16] Katsushi Shirahata, Shuhei Yoshimoto, Takeo Tsuchihara, and Satoshi Ishida. "Digital Filters to Eliminate or Separate Tidal Components in Groundwater Observation Time-Series Data". In: *JARQ* 50.3 (2016), pp. 241–252. DOI: 10.6090/jarq.50.241.

[17] Inung Wijayanto and Reni Dwifebrianti. "Jenis Tipe Jangkauan Suara pada Pria dan Wanita Menggunakan Metoda Mel-Frequency Cepstral Coefficient dan Jaringan Syaraf Tiruan Backpropagation". In: *Konferensi Nasional Sistem dan Informatika (KNSI)*. Institut Teknologi Telkom. Bandung, Indonesia, 2013.

[18] Esraa M. Shalby, Almoataz Y. Abdelaziz, Eman S. Ahmed, and Basem Abd-Elhamed Rashad. "A Comprehensive guide to Selecting Suitable Wavelet Decomposition Level and Functions in Discrete Wavelet Transform for Fault Detection in Distribution Networks". In: *Scientific Reports* 15 (Jan. 2025). DOI: 10.1038/s41598-024-82025-2.

[19] Aditya Mahmud Faza, Cepy Slamet, and Dian Nursantika. "Analisis Kinerja Kompresi Citra Digital dengan Komparasi DWT, DCT dan Hybrid (DWT-DCT)". In: *JOIN (Jurnal Online Informatika)* 1.1 (June 2016), pp. 1–5. DOI: 10.15575/join.v1i1.3.

[20] B.B. Mandelbrot. *The Fractal Geometry of Nature*. Einaudi paperbacks. Henry Holt and Company, 1983. https://books.google.co.id/books?id=0R2LkE3N7-oC.

[21] D. Juniati and I K. Budayasa. *Geometri Fractal dan Aplikasinya*. University Press, Unesa, 2016.

[22] Tomoyuki Higuchi. "Approach to an irregular time series on the basis of the fractal theory". In: *Physica D: Nonlinear Phenomena* 31.2 (1988), pp. 277–283. DOI: 10.1016/0167-2789(88)90081-4.

[23] D. Juniati, C. Khotimah, D. E. K. Wardani, and K. Budayasa. "Fractal Dimension to Classify the Heart Sound Recordings with KNN and Fuzzy c-Mean Clustering Methods". In: *Journal of Physics: Conference Series* 953.1 (2018), p. 012202. DOI: 10.1088/1742-6596/953/1/012202.

[24] Diah Wahyu Safitri and Dwi Juniati. "Classification of Diabetic Retinopathy Using Fractal Dimension Analysis of Eye Fundus Image". In: *AIP Conference Proceedings* 1867.1 (Aug. 2017), p. 020011. DOI: 10.1063/1.4994414.

[25] Felix Anthony, Hafiz Irsyad, and Muhammad Ezar Al Rivan. "KNN dan Gabor Filter Serta Wiener Filter untuk Mendiagnosis Penyakit Pneumonia Citra X-Ray pada Paru-Paru". In: *Jurnal Algoritme* 1.2 (Apr. 2021), pp. 147–155. DOI: 10.35957/algoritme.v1i2.893.

[26] Dewi Novita Sari, Muh. Arif Rahman, and Randy Cahya Wihandika. "Uji Parameter dan Arsitektur Convolutional Neural Network untuk Mendeteksi Citra Wajah Bermasker". In: *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)* 9.7 (Dec. 2022), pp. 1707–1714. DOI: 10.25126/jtiik.202296776.

[27] D. Widhyanti and D. Juniati. "Classification of Baby Cry Sound Using Higuchi's Fractal Dimension with K-Nearest Neighbor and Support Vector Machine". In: *Journal of Physics: Conference Series* 1747.1 (2021), p. 012014. DOI: 10.1088/1742-6596/1747/1/012014.