Received: July 8, 2023. Revised: September 22, 2023.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Integrating Grey Wolf Optimizer for Feature Selection in Birdsong Classification Using K-Nearest Neighbours Algorithm

Ricardus Anggi Pramunendar[1]* **Pulung Nurtantio Andono[1]** **Guruh Fajar Shidik[1]**
Rama Aria Megantara[1] **Dewi Pergiwati[1]** **Dwi Puji Prabowo[1]** **Yuslena Sari[2]**
Lim Way Soong[3]

[1]*Study Program in Informatics Engineering, Faculty of Computer Science*
Universitas Dian Nuswantoro, Semarang 50131, Indonesia
[2]*Department of Information Technology, Faculty of Engineering*
Universitas Lambung Mangkurat, Banjarmasin 70123, Indonesia
[3]*Center ojliobotics and Automation, Faculty of Engineering and Technology,*
Multimedia University, Malacca, Malaysia
* Corresponding author's Email: ricardus.anggi@dsn.dinus.ac.id

Abstract: This study aims to improve the classification accuracy of birdsongs by selecting the most pertinent features. This is important because birds are exceptional environmental regulators, but many species are endangered. The community can be assisted in distinguishing bird species and conserving the local environment if the classification is more precise. Nevertheless, because of disruptive noise and unfavorable qualities in the whispering of these bird species, feature selection focuses on enhancing performance accuracy. The use of the gray wolf optimizer (GWO) technique has been employed to identify the most optimum features from the data after outlier removal by the application of k-means clustering, reducing noise through YAMNet, and feature synthesis using gammatone cepstral coefficients (GFCC). This work utilizes the GWO algorithm to address the constraint management challenges associated with high-dimensional data in birdsong classification. The fitness functions used in this research are derived from the K-nearest neighbors (KNN) algorithm. The objective is to devise innovative ways for effectively managing constraints in the context of high-dimensional data. The number of features was reduced by more than 30.7% compared to the initial number of features and obtained an accuracy of 81.06%, as determined by the evaluation outcomes. This discovery improves precision by 4% and surpasses prior research. In summary, this work showcases the effectiveness of the optimization method, especially of GWO. It makes a valuable contribution to advancing a new workflow for analyzing high-dimensional data, specifically in enhancing the classification of birdsongs.

Keywords: Bird classification, YAMNet, GFCC, GWO, KNN.

1. Introduction

Birds significantly impact human life through stress recovery, attention restoration, and as a warning signal for dangerous weather changes and other situations. In addition, birds are among the most critical environmental monitoring indicators for environmental preservation and biodiversity assessment in developing ecological civilization. Birds may provide direct aid to humans by preserving ecological equilibrium, functioning as teaching and scientific research experimental materials, enhancing the naturalness of the environment, and monitoring the natural environment.

However, human actions and climate change can occasionally endanger birds and other fauna [1-3]. According to the world bank, a total of 160 bird species out of the 4,584 existing globally are now facing the imminent threat of extinction. Therefore, the accurate assessment of the ecosystem's condition and the sustainability of activities such as ecotourism and bird watching might be compromised, along with potential implications for the ecological equilibrium of the environment.

The government has implemented a range of conservation initiatives across several sectors. The prioritization of bird habitat preservation results in their protection. Furthermore, community members and scientists must collaborate to learn about various species' taxonomy, morphology, habitat, extinction risk, and conservation strategies [2, 4].

The research encountered significant difficulties in accurately identifying varied bird species, primarily due to a limited grasp of their distinguishing characteristics. These challenges were particularly apparent during the classification process in a natural forest environment, where the abundance of birdsongs and reliance on human auditory perception posed significant limitations [3].

There are substantial differences in the birdsongs for the different bird species. Hence, birdsongs may be used for classification research [1, 5]. Classification of birdsongs enables a clear comprehension of the current situation and the dynamic variations of the bird population. The protection and evaluation of ecosystems rely heavily on analyses and investigations of bird species diversity [6]. Several researchers have made attempts at this [1-3, 5-7]. They want information not just about birdsongs in their area but also about birdsongs from other places. This phenomenon may be attributed to the fact that every nation will inevitably encounter the issue of bird species becoming extinct [7]. Moreover, several research emphasize the capacity to classify birdsong due to their efficacy, non-invasive nature, and extensive coverage [6].

The classification of birdsongs typically encompasses four sequential stages: data acquisition, preliminary data manipulation, extraction of distinctive acoustic features, and the application of classification algorithms [2]. Parameter feature extraction and feature selection in classifiers are of utmost importance in classifying birdsong [2, 3]. Several studies have attempted to combine numerous characteristics to improve the accuracy of classification results, as compared to using a single feature [1, 8]. Nevertheless, many individuals use the feature selection technique to get the most desirable features [3, 6].

The current investigation suggests the use of the GFCC feature extraction technique. Although GTCC-based techniques have shown promising results in the past, it is essential to exercise caution when selecting feature variations produced from voice data. Previous research has shown that the particle swarm optimization (PSO) approach for feature selection has difficulties due to its inclination to quickly converge to the local optimum, as indicated by previous studies. The abovementioned technique discerns remarkable attributes not investigated in previous PSO studies [9]. In contrast, the GWO method can alleviate premature convergence, enhancing its probability of achieving a global optimum [10, 11]. The use of the KNN approach for the classification procedure was motivated by its effectiveness in previous studies [2, 3] for classifying birdsongs.

This paper presents the primary contributions: (1) The audio files are subjected to the GTCC extraction method, and the significant features are determined using the GWO algorithm. (2) The feature selection process aims to eliminate redundant and extraneous features, resulting in an optimal subset of features. (3) An investigation explores the correlation between a feature selection method and classification performance. (4) The parameters necessary for the classification of birdsongs are identified.

The structure of this article is organised in the following manner: Section 2nd provides an overview of relevant previous research that applies to the present study. The 3rd part delineates the experimental technique and model that we propose. Section 4 provides a comprehensive account of the outcomes and assessment of the conducted experiment, while section 5 presents the concluding remarks of the summary.

2. Previous research

A group of ornithologists researched to classify wild birds. In ornithology and conservation monitoring, there has been a growing interest in the automated classification of birdsong based on species. Their objective is to enhance the classification of birdsongs. The current discussion may be divided into two major domains: scientific research related to the classification of birdsong and preferred optimizer techniques classification of birdsong has been conducted by Andono et al. [2]. The study included the use of a total of 264 bird species, which were afterwards divided into 18 distinct experimental models. This approach allowed for the calculation of average performance across the various models. This work employs a hybrid approach that combines the GTCC and mel-frequency cepstral coefficients (MFCC) techniques and uses dimension reduction techniques. The findings indicate that the mean accuracy performance attains a value of 78.11%. Furthermore, Murugaiya et al. [12] used preprocessing approaches for optimum data quality. GTCC feature extraction was also utilized, leading to a performance accuracy of 89.5%.

Xie et al. combined three distinct feature extraction techniques to classify 16 bird species [6]. This research uses a support vector machine (SVM) for classification and provides a performance of 96.25%. This study utilizes transfer learning, and increased performance is obtained from a combination of features. Combination feature classification methods are also carried out by Raghuram et al. (2016) for birdsongs of 35 bird species using several combined features, such as pitch, energy, duration, MFCC, and tempo. This research succeeded in achieving an accuracy of 83.33% by utilizing the random forest (RF) classification method. However, RF and SVM need substantial memory when dealing with large feature datasets, affecting the overall execution time.

Pahuja and Kumar [8] took a different approach. They used the short-time fourier transform (STFT) spectrogram with multi-layer perceptron (MLP) classification to sort the songs of eight common Eurasian bird species with 96.1% accuracy. Supriya et al. [13] classify birdsongs using GMM based on the MFCC feature, and the results demonstrate that GMM is better than SVM. Both studies propose a new workflow by pre-processing bird signals for better performance. Likewise, Sukri et al. [14] and Murugaiya et al. [12] perform pre-processing to obtain a clear signal with increasing accuracy. Nevertheless, the current feature generates a diverse range of noise, hence potentially compromising its accuracy. More enhancements might be made to improve its performance.

Furthermore, several studies have been undertaken by Kahl et al. [15], Jancovic and Kokuer [16], Chandu et al. [17], and other academics, exploring various methodologies for the classification of birdsong using convolutional neural networks (CNNs). The researchers directed their attention to the implementation of their CNN. Despite its effectiveness, CNN requires substantial data storage and processing resources.

The main objective of the research is to determine which features have the most optimal features. It is well acknowledged that a limited body of research is dedicated to feature selection in the domain of birdsong classification. Several researchers have used the feature selection approach in their studies, including Ji et al. [18], Murugaiya et al. [1], and Andono et al. [3]. The study conducted by Ji et al. [18] used a selection feature based on the neighborhood component analysis (NCA) method in combination with the local binary pattern (LBP) feature approach. The results suggest a significant improvement in accuracy by more than 1%. Murugaiya et al. [1] use linear discriminant analysis (LDA) to determine the most suitable features and achieve a performance accuracy of 96.7%. The dataset consisted of 10 distinct species, each having 12 instances of birdsong data. Andono et al. [3] use a metaheuristic technique, PSO, to get optimal feature outcomes from a range of GTCC result characteristics. Based on the data congruent with the findings of [2], this study achieved a performance improvement of over 1% in accuracy.

From all the research that has been apparent above, it is proven that most researchers utilize different approaches to identify birdsong and employ diverse data sources with variable amounts of labels. Furthermore, there is a lack of consideration among researchers on the impact of specific characteristics on accurate performance. Consequently, this study aims to build upon prior research by identifying the optimal elements that contribute to achieving the highest level of performance.

This works with the GWO method for feature selection. The GWO method is known for preventing premature convergence and attaining a global optimum, improving the effectiveness of classifying different birdsong variations. The feature data used in this study is acquired by employing the GTCC approach, which has enhanced sound quality by eliminating preexisting noise. In this work, a technique known as dimensional reduction was used to mitigate the presence of residual noise.

3. Proposed research

Previous studies suggest that the efficacy of a classification system is influenced by the data, features, and classification algorithms used. Findings from the classification strategy show that essential features do not necessarily mean there will be an absence of accuracy. These characteristics can be employed depending on the context to present the correct vocal quality. However, these factors are impacted by the data source. This research introduces a unique approach to recognising birdsongs that integrates the inherent vocal characteristics (as illustrated in Fig.1).

3.1 Data collection and pre-processing

This study utilises data on birdsongs, which have distinct characteristics and are affected by their surroundings. The data used is from the cornell lab

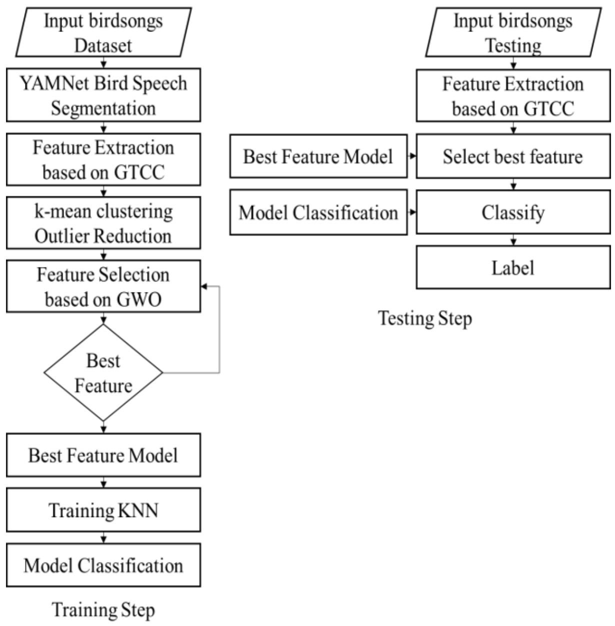

Figure. 1 Research proposed training and testing model

of ornithology center for conservation bioacoustics (CCB), which included 21,375 melodies from 264 distinct bird species (https://www.kaggle.com/competitions/birdsong-recognition/). This data has been utilised in prior studies [2, 3] and will also be utilised in this study. Our research involves two main phases: data preparation and classification of birdsongs based on their characteristics.

Initially, a data collection is compiled based on the attributes of files of the same type. The investigation entails extracting audio from the newly provided data set using time-coded annotations, which enables precise space validation. The files in the training set are then processed to generate new datasets with unique information.

Speech boundaries are identified in the audio stream, segmenting and grouping the noises. Birdsongs are prioritised, while other noises are disregarded. In this stage, The YAMNet AudioSet ontology is utilised to extract the portion of the audio stream classified as birdsongs [6, 19, 20].

3.2 GFCC features

The gammatone frequency cepstral coefficients (GFCC) are used to derive unique features [6]. The fundamental basis of the GFCC is the use of a set of gammatone filter banks, which are afterwards transformed into a cochleagram. This cochleagram serves as a time-frequency depiction of a given signal. The GFCC characteristics are then derived from this cochleagram.

3.2.1. Gammatone filter

Gammatone filters imitate the operation of the human auditory system. The centre frequency (f_c) of a gammatone filter is determined by Eq. (1), where 'a' is a constant that controls the gain value, φ is the initial phase of the filter, 'n' (less than four) defines the filter order, and 'b' is determined by Eq. (2) that associate with the equivalent rectangular bandwidth (ERB). A collection of gammatone filters with distinct center frequencies are represented as channels with distinct f_c. That corresponds to a representation like the FFT-based spectrogram.

$$g(t) = at^{n-1}e^{-2\pi bt}cos(2\pi f_c + \varphi) \quad (1)$$

$$b = 25.17\left(\frac{4.37f_c}{1000} + 1\right) \quad (2)$$

3.2.2. Windowing

The feature GFCC commission necessitates a window to encompass K in each frame and move each L po characterisesint. $x\big(t; f_c(m)\big)$ characterizes every frame, with the central frequency (f_c) in the m filter. The Cochleagram of each frame is computed as the average across the t window (3).

$$one\bar{x}(n;m) = \frac{1}{K}\sum_0^{K-1}\gamma|x\big(nL + i; f_c(m)\big) \quad (3)$$

Eq. (3) represents the magnitude of the complex number. It also represents the factor dependent on frequency. m denotes the number of filter bank channels with K values of 400, L values of 160, and m values of 32 for 16 kHz signals that generate 100fps.

3.2.3. DCT

Cepstral coefficients devoid of correlation were obtained using discrete continue transform (DCT). The range u value from 0 to 31 is equivalent to the MFCC operation (4). The GFCC method generates 39 features, each comprising 13 GFCC values and 26 GFCC deltas.

$$g(n;m) = \left(\frac{2}{m}\right)^{0.5}\sum_{i=0}^{m-1}\left\{\frac{1}{3}log(\bar{x})\,cos\left[\frac{\pi u}{2m(2i-1)}\right]\right\} \quad (4)$$

3.3 Outlier reduction based on K-means clustering

In some cases, duplicate data can occur. Recognising these feature sets as interconnected enables a reduction in the feature vector without a substantial loss of information. K-means clustering helps reduce duplicate data for a given label while retaining most pertinent information. K-means can reduce anomalies by combining generated data and selecting the most significant members. As data is produced each time a speech signal undergoes feature extraction, there is no data loss, and data aggregation ensures high-quality output. The k-means algorithm organises data into distinct clusters, each sharing and possessing distinctive characteristics. This method reduces the objective function by restricting within-cluster variation and increasing between-cluster variation. In this instance, the objective function is represented by Eq. (5), where 'f' represents the signal frequency and 'ce' represents the randomly determined centroid.

$$dist = \sqrt{(f - ce)^2} \tag{5}$$

3.4 Feature selection based on grey wolf optimization

Grey wolf optimizer (GWO) was employed to identify the optimal features [19]. In GWO, the wolf mimics the leadership role and intelligent hunting behavior seen in nature, such as exploring, encircling, and assaulting prey. The wolf inside the GWO is classified into discrete categories, each fulfilling a unique function. The first group, also known as the alpha, is the most powerful and serves as the decision-maker. The second group, the beta, serves as an advisor to the alpha, while the third group, the delta, also contributes. Optimisation is a function of alpha, beta, and delta. A fourth pack, Omega, is responsible for hunting down other canines. In this circumstance, GWO establishes a group at the start of the initial population and actively modifies the wolf position to achieve the optimal outcome. The following measures were taken to develop new competencies:

- Initialize the population value (*init_pop*), the maximum number of iterations in a single process (*max_iter*), and the random numbers *Xi* and *Yi*, which represent the initial position of the wolf. The GWO method will concentrate on three values, namely alpha, beta, and delta, whose respective positions correspond to the solutions *Xa, Xb*, and *Xd*. The remaining wolves, particularly omega, represent a potential solution.
- Initialize the three coefficients' vectors as $\vec{a}, \vec{A}$ and $\vec{C}$.
- Each wolf position (*X*) is used as a reference to select the corresponding features within their position range.
- Determine the value of $\vec{a}$ that decreases linearly using Eq. (6), where max_iter is the utmost number of iterations.

$$\vec{a} = 2 - 1 \times (2/max\_iter) \tag{6}$$

- Determine the value of $\vec{A}$ and $\vec{C}$ using Eqs. (7) and (8) where $\vec{r_1}$ and $\vec{r_2}$ are random vectors in [0,1].

$$\vec{A} = 2\vec{a}\,\vec{r_1} - \vec{a} \tag{7}$$

$$\vec{C} = 2\,\vec{r_2} \tag{8}$$

- Calculate the wolf's movement according to Eqs. (9) and (10) and update its position.

$$\vec{D} = |\,\vec{C}.\vec{X_p}(t) - \vec{X}(t)| \tag{9}$$

$$\vec{X}(t+1) = \vec{X_p}(t) - \vec{A}\,\vec{D} \tag{10}$$

where t represents the current iteration, $\vec{A}$ and $\vec{C}$ represent the coefficient vectors, $\vec{X_p}$ represents the prey position vector, and $\vec{X}$ represents the wolf position vector. The movement and new positions of the alpha, beta, and delta wolves can be calculated using Eqs. (11) to (17), given Eqs. (9) and (10).

$$\vec{D_\alpha} = \vec{C_1}\,.\vec{X_\alpha} - \vec{X} \tag{11}$$

$$\vec{D_\beta} = \vec{C_2}\,.\vec{X_\beta} - \vec{X} \tag{12}$$

$$\vec{D_\delta} = \vec{C_3}\,.\vec{X_\delta} - \vec{X} \tag{13}$$

$$\vec{X_1} = \vec{X_\alpha} - \vec{A_1}\,.\vec{D_\alpha} \tag{14}$$

$$\vec{X_2} = \vec{X_\beta} - \vec{A_2}\,.\vec{D_\beta} \tag{15}$$

$$Open\vec{X_3} = \vec{X_\delta} - \vec{A_3}\,.\vec{D_\delta} \tag{16}$$

$$X(t+1) = \frac{X_1 + X_2 + X_3}{3} \tag{17}$$

- Verify new solutions for the three coefficient vectors of $\vec{a}$, $\vec{A}$ and $\vec{C}$ and penalize with them if necessary.
- Determine the most current fitness values. Because the most recent values are greater than the previous ones, the positions of the wolves *Xa, Xb*, and *Xd* are updated accordingly.

- Compare the halting criteria to the value of *max_iter*.

3.5 K-nearest neighbors' classification (KNN)

The k-nearest neighbours (KNN) algorithm is a straightforward yet effective classification method that retains all extant data instances and classifies new ones based on similarity. When presented with new, unlabelled data, it identifies the k most similar instances in this set by comparing their features using labelled training datasets. KNN is renowned for its precision, robustness against outliers, and lack of data-specific assumptions. However, it may require substantial computational resources and memory, posing potential computational and storage difficulties.

3.6 Performance evaluation

Accuracy, a prevalent performance metric in classification models, quantifies the proportion of correctly classified data points. It is computed as the proportion of proper classifications relative to the total data (18). While useful, accuracy may not be the optimal metric for evaluating unbalanced datasets. However, accuracy is useful for evaluating data.

$$accuracy = t/n \times 100 \qquad (18)$$

3.7 Research design

The objective of the research was to demonstrate the significance of including both automatically obtained and manually selected variables to get a high level of classification accuracy. The approach was evaluated using data from the center for conservation bioacoustics (CCB) at the cornell lab of ornithology. Comparative analysis with previous research revealed a significant increase in classification accuracy.

The researchers used GFCC for feature extraction. The experiment used specific settings, including a Hamming window function with a frequency multiplier of 0.03. A sample frequency of 8000 Hz was used to measure time in samples and influence audio quality. The temporal parameters of frame length and hop duration play a crucial role in establishing the material characteristics and spacing for audio analysis. In audio processing, frequency and time resolution were manipulated by adjusting segment length, which consisted of 8192 samples, and segment duration, which was computed based on the segment length and the sampling frequency (*fs*). To comprehensively analyze the entire audio segment, the number of hops was determined by considering the segment's length, each frame's duration, and each hop's duration. The determination of sample-based frame length and spacing was achieved by calculating frame length, which was based on the time of the frame and the sampling rate (*fs*). Similarly, the calculation of hop length, which was based on the duration of the hop and the fs, also played a role in this determination. An FFT of 2048 samples was used for the spectrum analysis and frequency domain conversion of audio sources. The voice detection limitations were found using a window overlap of 0.02 times the frequency value. 625,381 data points were obtained by applying the GFCC and k-means clustering techniques on a dataset consisting of 21,375 birdsong recordings.

The GWO approach was used to ascertain the most crucial attributes using a population size and a maximum number of iterations set at 100. During the training phase, the researchers estimated the fitness function value for each feature selection approach using the KNN algorithm. The default specifications for the proposed method were determined by selecting the weights that demonstrated the highest performance.

The GWO and PSO algorithms were executed using identical parameters, including a population size and an iteration limit of 100. This research aimed to assess and contrast the efficacy of two optimization strategies. However, the assessment process involves not only comparing the two approaches but also considering various additional optimization strategies to get the ideal feature outcomes, such as genetic algorithm (GA) [19] and swarm magnetic optimiser (SMO) [20].

For performance assessment, the dataset was partitioned into five groups of similar size. Subsequently, five experiments were carried out to assess the performance of each partition. The training protocol and subsequent examination of classification performance yielded outcomes culminating in generating a convolution matrix during the performance assessment.

4. Research outcome and analysis

This section includes the fundamental processing processes, feature extraction, and classification.

4.1 Data gathering and preprocessing

The dataset includes 21,375 recordings of 264 species of birds. For signal integrity, the unique wavelengths of each sound necessitate the division of this massive data set into multiple sections.

Table 1. Feature data after reduction

No	gtcc1	gtcc5	gtcc10	gtcc15	.	gtcc39
1	3,25	-1.03	-6,00	-0.83		-2,96
2	1,63	-1.79	0,26	-0.33		-2,59
3	1,62	-2.39	0,65	-0.17		-2,71
4	1,56	-1.99	0,49	0.38		-2,60
5	1,66	-1.52	0,27	0.87		-1,81
6	1,63	-1.85	0,42	1.36		-2,16
7	1,77	-1.70	0,32	-0.15		-2,96
...	...	…	....	...		...
625.381	1,33	0,88	0,59	-0,29		0,23

Windowing replicates each sample frame from beginning to end to prevent audio discontinuities and ensure audio consistency. Not all data represent birdsongs, necessitating the classification of probable boundaries to isolate birdsongs (as depicted in Fig. 2). Investigations reveal that several data segments lack bird noises; these are subsequently eliminated. Finally, combining birdsong signals (as shown in Fig. 3) creates a comprehensive birdsong signal.

4.2 Feature engineering

The signal quality is better by taking traits from the signal data and turning them into numbers that machine learning programs can use. Over 21,375 birdsongs are captured and extracted using the GFCC method. The k-means clustering method reduces the amount of data by finding the most important overall cluster member for each grouped dataset. This makes the data become 625,831 records with 39 features (as shown in Table 1).

4.3 GWO-based feature selection

The results obtained from the submitted features have shown a more than 30.7% reduction in the overall number of features compared to the original set. When considering the qualities of a wolf colony, the selection of features necessitates the identification of the wolf's location to acquire the necessary traits. The position above is then used to determine the numerical worth of the chosen attributes (as seen in Table 2). The column labelled *x* represents the initial position of the wolf inside the alpha group. During the inquiry, a set of 27 features was selected, indicating an average position of 0.78 and a minimum value of 0.51. Based on the provided data, it can be determined that an overall of 12 features were not selected, with an average value of 0.41.

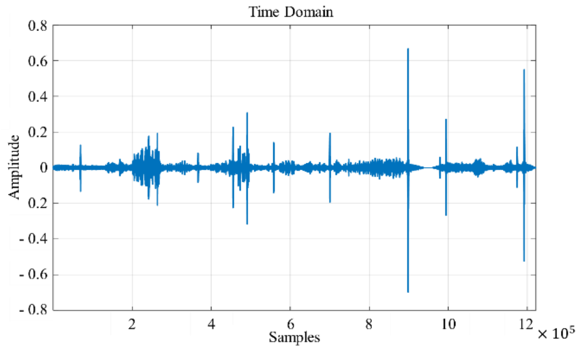

Figure. 2 Original bird audio

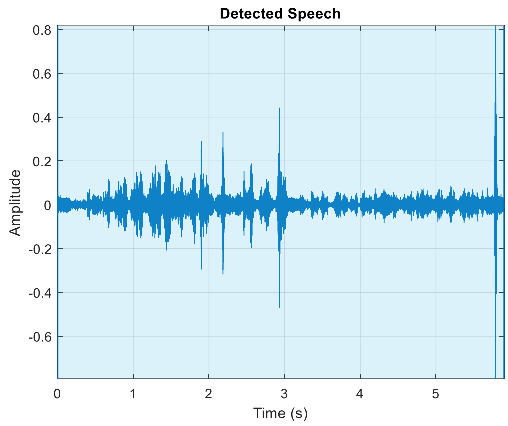

Figure. 3 Clear bird voice

Table 2. Selected features applied using GWO

F	*x*	*F*	*x*	*F*	*x*	*F*	*x*
1	0.85	10	0.80	22	0.67	32	0.87
2	0.57	11	0.93	25	0.90	34	0.99
3	0.64	13	0.91	26	0.79	35	0.59
4	0.99	16	0.66	27	0.52	36	0.85
5	0.87	17	0.56	28	0.82	37	0.93
6	0.91	20	0.70	29	0.62	38	0.85
8	0.66	21	0.70	31	0.79		

The optimal feature is determined using the GWO method's highest fitness scores. The K-nearest neighbours (KNN) method evaluates the chosen features with an accuracy metric. For each wolf position, the fitness value is calculated. The k parameter is assigned to 3 in the KNN method.

The evaluation result of several parameter combinations related to the number of wolves and the number of iterations has determined that the most favourable quantity of wolves is 100. The average level of accuracy achieved was 80.13% after conducting ten iterations (as seen in Fig. 4). The data demonstrates that there is a discernible range of effects when the quantity of wolves is increased, exhibiting a variability that ranges from

Received: July 8, 2023. Revised: September 22, 2023.

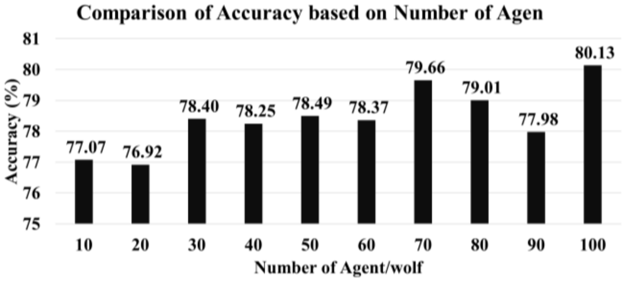

Figure. 4 Highest accuracy the number of wolves

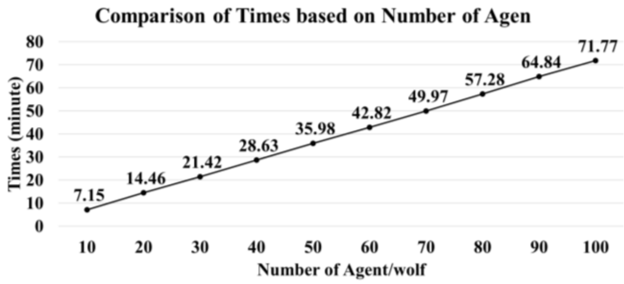

Figure. 5 Times computation the number of wolves

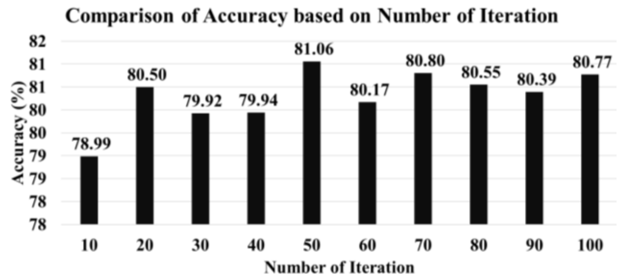

Figure. 6 Highest accuracy the number of wolves

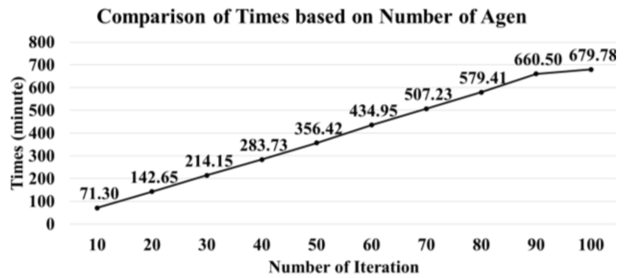

Figure. 7 Times computation the number of wolves

76.92% to 80.13%. The results of this study indicate a favourable association between the number of wolves and the likelihood of achieving a greater degree of accuracy. However, it is crucial to note that the augmentation of wolf quantity also impacts the time computation (as seen in Fig. 5).

The optimal number of wolves is used to determine the performance accuracy based on the number of iterations. The investigation revealed that the highest level of accuracy, achieving 81.06%, was completed during the 50th iteration. Furthermore, the performance results obtained ranged between 79 and 81 (Fig. 6), indicating a consistent pattern in the data. This observed consistency extends to subsequent iterations, showing a positive correlation between the number of iterations and the related increase in time (Fig. 7).

The analysis of variance (ANOVA) done on the experimental data shown in Figs. 4 and 6 indicates a statistically significant disparity. The obtained F-statistic value of 25.95 the result value is a statistically significant result. Furthermore, the very small P-value of 7.58E-05 provides evidence of a noticeable disparity in accuracy between the two groups. The observed discrepancy is not coincidental but somewhat influenced by the parameter settings of the GWO.

4.4 Classification accuracy

This study aims to increase classification accuracy by selecting appropriate features. The GWO method is used to find the best features for classifying sound patterns, then compared to earlier studies to determine how feature selection affects classification performance. In the classification analysis, the Proposed obtains an accuracy rate of 81.06 % for the KNN algorithm, which is the highest of any method. With an accuracy rate of 80.77%, the PSO method [3] also exhibits excellent performance. Alternatively, the procedure without feature selection [2] has a slightly reduced accuracy rate of 77.06%. Figure 8 also demonstrates that applying feature selection techniques such as GWO and PSO can enhance the accuracy of the KNN algorithm.

Fig. 9 illustrates the classification algorithm's (LDA, DT, KNN, NB, NN) efficacy using the GWO feature selection method. With an accuracy rate of 81.06%, the KNN algorithm has the maximum accuracy rate in the GWO method, followed by the DT algorithm with an accuracy rate of 73.01%. The NN algorithm also demonstrates solid performance with an accuracy rate of 76.51%. However, the LDA and NB algorithms have a lower accuracy rate of 44.03% and 36.2%, respectively. Based on these results, it can be concluded that the KNN algorithm is the most accurate classification algorithm for the GWO method. Using the GWO method to classify features, the DT and NN algorithms also demonstrate decent performance, whereas the LDA and NB algorithms perform less well.

4.5 Feature selection comparison

Previous experiments showed that the best performance was obtained from this proposal compared to previous research for selecting birdsong features. However, when compared to other optimization procedures used to get optimum

Received: July 8, 2023. Revised: September 22, 2023.

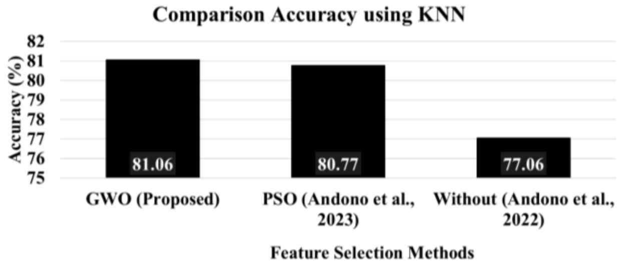

Figure. 8 Comparison accuracy using KNN

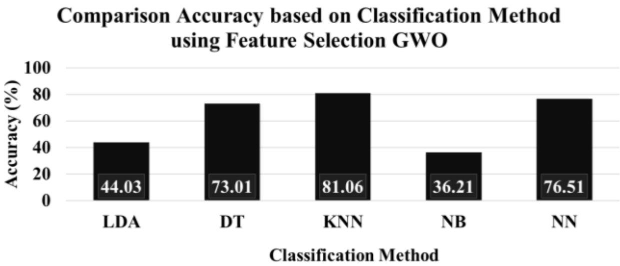

Figure. 9 Comparison accuracy Feature Selection GWO

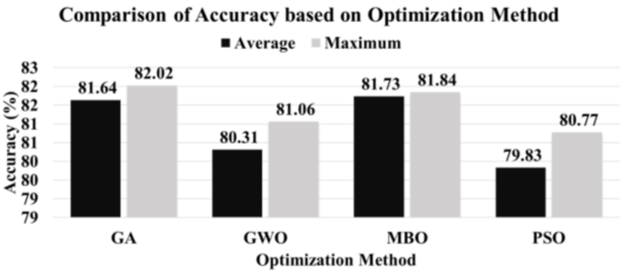

Figure. 10 Comparison accuracy using KNN

features, it becomes apparent that this proposal lacks robustness in terms of accuracy (Fig. 10). The results that were produced for the best features by each approach exhibit varying numerical values, which in turn impact both the computational time and performance outcomes. The number of features packed in the GA, MBO, GWO, and PSO algorithms are 36, 38, 27, and 32, respectively. The corresponding accuracy values are shown in Fig. 10.

The experiments included a comparative analysis of the optimum and maximal average results. The statistical analysis demonstrates that the P-value is amount 1.10E-08, suggesting a major disparity across the experimental groups of GA [19], GWO, MBO [20], and PSO [3] concerning the observed mean values. This discovery provides more evidence supporting the conclusions shown in Fig. 10. As a result, a significant discrepancy may be seen in the average values across these distinct groups.

Based on the observed results, it is apparent that a strategy enabling wolves to either approach or retreat from the superior reference could be an appropriate choice for improvement, compared to either approaching the ideal reference or avoiding the inferior one. MBO can explore alternative search spaces in instances when proximity to local optima is achieved by deviating from the reference point. Furthermore, it can facilitate acquiring the global ideal solution due to its dynamic nature since it actively moves rather than being static as a mere reference point. In contrast, the GA employs a strategy to sustain variety among individuals in the population, facilitating the more comprehensive exploration of the search area. Nevertheless, both approaches exhibit a drawback in terms of computing efficiency when evaluated under identical settings.

4.5 Discussion

This study examines the use of feature selection in classifying birdsongs. In its stages, feature extraction using GFCC yields 39 features and dimension reduction using k-means clustering, as established by prior research. This study suggests the use of GWO as a criterion for selection. Based on experiments, the number of features can be reduced by more than 30.7% compared to the original set. These features are derived based on the GWO method's highest fitness value in Table 2, which displays 27 selected features.

The accuracy of the selected features is determined using KNN with the parameter k set to 3. This investigation uses the number of wolves and the number of iterations as test parameters. At the 50th iteration, the accuracy attained a maximum of 80.21%, according to the investigation results. Even though the increase in accuracy tends to occur as the number of iterations increases, it turns out that there are still minor fluctuations after the performance apex is reached. With extremely low p-values, the results of the analysis indicate that there are statistically significant differences between the iteration groups. This shows that the average of at least one iteration group is significantly distinct from the standards of the other iteration groups. In addition, there is a significant disparity in terms of accuracy between the time groups. The classification comparison results, however, indicate that the performance of the proposed method is the best. Overall, this study integrates GWO-based feature processing and feature selection methods to enhance the classification quality of birdsongs. The results demonstrate that optimal feature selection can lead to superior classification performance.

5. Conclusions

The research conclusions indicate that using the GFCC technique for feature processing and implementing the k-means clustering technique for data reduction generated the construction of a dataset including 625,831 data. Each of the data in the dataset includes 39 unique features. The records above originate from 21,375 bird songs. The grey wolf optimization (GWO) approach effectively resulted in a notable decrease in the total number of features by 30.7%. This finally resulted in classifying and selecting 27 features based on their unique fitness values. At the 50th iteration, the evaluation of accuracy utilizing the KNN technique reveals that the accuracy reaches a maximum weight of 81.06%. Although the accuracy value cannot exceed other optimization methods, the analysis of variance (ANOVA) results indicated a significant difference in accuracy between the optimization methods. They yielded superior outcomes compared to further research investigations.

Based on the limitations, it is advisable to use alternative optimization methodologies for feature processing, such as diverse feature selection strategies and more expansive classification algorithms for future research. The main objective of this endeavor is to enhance the precision and dependability of birdsong classification while advancing our understanding of optimal feature selection techniques.

Conflicts of interest

According to the international journal of intelligent engineering and systems rules and my ethical duties as a researcher, I guarantee that this paper has not been published, copied, or sent to another journal before. I sent a full report to the international journal of intelligent engineering and systems, letting them know about any possible conflicts of interest that may come up due to this study, and I got all the writers' permission to do so.

Author contributions

The contributors to this article are Ricardus Anggi Pramunendar (RAP), Pulung Nurtantio Andono (PNA), Guruh Fajar Shidik (GFS), Rama Aria Megantara (RAM), Dewi Pergiwati (DP), Dwi Puji Prabowo (DPP), Yuslena Sari (YS), and Lim Way Soong (LWS), which everyone made the following contributions: Conceptualization, RAP and PNA; methodology, LWS and RAP; software, RAM, RAP and GFS; investigation, YS and DPP; formal analysis, RAP, and GFS; resources, DP, and DPP; data curation, DPS, DP, and DPP; writing—original draft preparation, GFS, PNA and RAP; writing—review and editing, GFS and YS; visualization, DP and DPP; supervision, RAP; project administration, DP and DPP; funding acquisition, RAP.

Acknowledgements

This research was conducted by a group of lecturers from Dian Nuswantoro University in Semarang, Indonesia; Lambung Mangkurat University in Banjarmasin, Indonesia; and Multimedia University in Malacca, Malaysia.

References

[1] R. Murugaiya, P. E. Abas, and L. C. D. Silva, "A novel approach of audio-based feature optimisation for bird classification", *Pertanika Journal of Science and Technology*, Vol. 29, No. 4, pp. 2383–2407, 2021.

[2] P. N. Andono, G. F. Shidik, D. P. Prabowo, D. Pergiwati, and R. A. Pramunendar, "Bird Voice Classification Based on Combination Feature Extraction and Reduction Dimension with the K-Nearest Neighbor", *International Journal of Intelligent Engineering and Systems*, Vol. 15, No. 1, pp. 262–272, 2022, doi: 10.22266/ijies2022.0228.24.

[3] P. N. Andono, G. F. Shidik, D. P. Prabowo, D. H. Yanuarsari, Y. Sari, and R. A. Pramunendar, "Feature Selection on Gammatone Cepstral Coefficients for Bird Voice Classification Using Particle Swarm Optimization", *International Journal of Intelligent Engineering and Systems*, Vol. 16, No. 1, pp. 254–264, 2023, doi: 10.22266/ijies2023.0228.23.

[4] N. J. Bennett et al., "Conservation social science: Understanding and integrating human dimensions to improve conservation", *Biological Conservation*, Vol. 205, pp. 93–108, 2017.

[5] S. A. Brooker, P. A. Stephens, M. J. Whittingham, and S. G. Willis, "Automated detection and classification of birdsong: An ensemble approach", *Ecological Indicators*, Vol. 117, No. June, 2020.

[6] S. Xie et al., "Multi-view features fusion for birdsong classification", *Ecological Informatics*, Vol. 72, No. November, p. 101893, 2022.

[7] J. Atanbori, W. Duan, J. Murray, K. Appiah, and P. Dickinson, "Automatic classification of flying bird species using computer vision techniques", *Pattern Recognition Letters*, Vol. 81, pp. 53–62, 2016.

[8] R. Pahuja and A. Kumar, "Sound-spectrogram based automatic bird species recognition using MLP classifier", *Applied Acoustics*, Vol. 180, p. 108077, 2021.

[9] F. Briggs, R. Raich, and X. Z. Fern, "Audio

classification of bird species: A statistical manifold approach", In: *Proc of IEEE International Conference on Data Minin*g, *ICDM*, pp. 51–60, 2009.

[10] S. M. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer", *Advances in Engineering Software*, Vol. 69, pp. 46–61, 2014.

[11] S. Gupta and K. Deep, "A novel Random Walk Grey Wolf Optimizer", *Swarm and Evolutionary Computation,* Vol. 44, No. November 2017, pp. 101–112, 2019.

[12] R. Murugaiya, P. E. Abas, and L. C. D. Silva, "Probability Enhanced Entropy (PEE) Novel Feature for Improved Bird Sound Classification", *Machine Intelligence Research*, Vol. 19, No. 1, pp. 52–62, 2022.

[13] Supriya, P. R, S. Bhat, Shivani, and S. S, "Classification of birds based on their sound patterns using GMM and SVM classifiers", *International Research Journal of Engineering and Technology,* Vol. 5, No. 4, pp. 4708–4711, 2018.

[14] M. M. M. Sukri, U. Fadlilah, S. Saon, A. K. Mahamad, M. M. Som, and A. Sidek, "Bird Sound Identification based on Artificial Neural Network", In: *Proc. of* 2020 *IEEE Student Conference on Research and Developmen*t, SCOReD 2020, No. September, pp. 342–345, 2020.

[15] S. Kahl et al., "Large-scale bird sound classification using convolutional neural networks", *CEUR Workshop Proceedings*, Vol. 1866, 2017.

[16] P. Jancovic and M. Kokuer, "Bird species recognition using unsupervised modelling of individual vocalisation elements", *IEEE/ACM Transactions on Audio Speech and Language Processing*, Vol. 27, No. 5, pp. 932–947, 2019.

[17] B. Chandu, A. Munikoti, K. S. Murthy, G. Murthy, and C. Nagaraj, "Automated Bird Species Identification using Audio Signal Processing and Neural Networks", In: *Proc. of 2020 International Conference on Artificial Intelligence and Signal Processing*, AISP 2020, 2020.

[18] X. Ji, K. Jiang, and J. Xie, "LBP-based bird sound classification using improved feature selection algorithm", *International Journal of Speech Technology,* Vol. 24, No. 4, pp. 1033–1045, 2021.

[19] P. I. Santosa and R. A. Pramunendar, "A Robust Feature Construction for Fish Classification Using Grey Wolf Optimizer", *Cybernetics and Information Technologies*, Vol. 22, No. 4, pp. 152–166, 2022.

[20] P. D. Kusuma and F. C. Hasibuan, "Swarm Magnetic Optimizer: A New Optimizer that Adopts Magnetic Behaviour", *International Journal of Intelligent Engineering and System*s, Vol. 16, No. 4, pp. 264–275, Aug. 2023, doi: 10.22266/ijies2023.0831.22.