

Article

Innovative Bacterial Colony Detection: Leveraging Multi-Feature Selection with the Improved Salp Swarm Algorithm

Ahmad Ihsan ^{1,*}, Khairul Muttaqin ¹, Rahmatul Fajri ², Mursyidah Mursyidah ³ and Islam Md Rizwanul Fattah ⁴

¹ Department of Informatics, Faculty of Engineering, Universitas Samudra, Langsa 24416, Aceh, Indonesia; khairulmuttaqin@unsam.ac.id

² Department of Chemistry, Faculty of Engineering, Universitas Samudra, Langsa 24416, Aceh, Indonesia; rahmatulfajri@unsam.ac.id

³ Department of Multimedia Engineering Technology, Politeknik Negeri Lhokseumawe, Kota Lhokseumawe 24301, Aceh, Indonesia; mursyidah@pnl.ac.id

⁴ Centre for Technology in Water and Wastewater, School of Civil and Environmental Engineering, Faculty of Engineering and IT, University of Technology Sydney, Ultimo, NSW 2007, Australia; islammdrizwanul.fattah@uts.edu.au

* Correspondence: ahmadihsan@unsam.ac.id

Abstract: In this paper, we introduce a new and advanced multi-feature selection method for bacterial classification that uses the salp swarm algorithm (SSA). We improve the SSA's performance by using opposition-based learning (OBL) and a local search algorithm (LSA). The proposed method has three main stages, which automate the categorization of bacteria based on their unique characteristics. The method uses a multi-feature selection approach augmented by an enhanced version of the SSA. The enhancements include using OBL to increase population diversity during the search process and LSA to address local optimization problems. The improved salp swarm algorithm (ISSA) is designed to optimize multi-feature selection by increasing the number of selected features and improving classification accuracy. We compare the ISSA's performance to that of several other algorithms on ten different test datasets. The results show that the ISSA outperforms the other algorithms in terms of classification accuracy on three datasets with 19 features, achieving an accuracy of 73.75%. Additionally, the ISSA excels at determining the optimal number of features and producing a better fit value, with a classification error rate of 0.249. Therefore, the ISSA method is expected to make a significant contribution to solving feature selection problems in bacterial analysis.

Keywords: bacterial colony; multi-feature selection; classification accuracy; improved salp swarm algorithm



Citation: Ihsan, A.; Muttaqin, K.; Fajri, R.; Mursyidah, M.; Fattah, I.M.R. Innovative Bacterial Colony Detection: Leveraging Multi-Feature Selection with the Improved Salp Swarm Algorithm. *J. Imaging* **2023**, *9*, 263. <https://doi.org/10.3390/jimaging9120263>

Academic Editor: Arslan Munir

Received: 26 September 2023

Revised: 20 November 2023

Accepted: 21 November 2023

Published: 28 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Microorganisms, also known as microbes, can pose a significant threat to human health and ecological systems. One of the most significant risks is their role in causing a variety of human diseases, including dysentery, tuberculosis, pneumonia, sepsis, typhus, diarrhea, and tetanus [1]. Furthermore, the negative effects of microorganisms also extend to agriculture, particularly the process of soil desalination [2]. One type of microbe that poses a threat is bacteria in water [3]. The World Health Organization (WHO) estimates that 2 billion people worldwide drink feces-contaminated water, which causes an estimated 525,000 deaths from diarrhea each year in children under five. According to the World Bank, poor water quality can limit economic growth, constrain human development, and disrupt food production [4].

Bacteria are microscopic organisms that are invisible to the naked eye. They are typically identified using a microscope, but this process can be time-consuming and laborious. Automatic identification of bacteria in water is a promising new technique that could

improve the speed and accuracy of bacterial identification. Traditional laboratory-based methods for identifying bacteria rely heavily on the expertise of laboratory technicians, which can lead to significant variations in results. Additionally, this laboratory-centric approach requires resource-intensive protocols that are expensive and time-consuming [5]. Therefore, an automated bacterial identification system is required that can leverage artificial intelligence (AI) to accurately and efficiently identify bacteria.

AI systems have become an attractive choice for bacterial identification [4]. One of the critical steps in AI-based bacterial identification is feature extraction, which involves identifying and extracting distinctive features of bacteria that can be used for classification. There are three general categories in feature extraction for the automatic classification of bacteria: first, the category of statistical approaches; second, machine-learning techniques such as artificial neural networks and support vector machines (SVMs); and third, geometric feature extraction from input images [6]. In practice, a combination of feature extraction techniques may be used to achieve the best results for bacterial classification.

Panicker et al. [7] proposed an approach for the automatic detection of tuberculosis bacteria from microscopic images of speckled sputum using geometric features, including color averaging of the extracted images. This study shows that color features have discriminatory properties and have a significant effect on improving image recognition. On the other hand, Luo et al. [8] used segmented images to extract the shape and size of microorganisms and then classify them based on their type. Ernest Bonah [9] introduced an approach to identifying food-related pathogens based on the dispersal patterns of bacterial colonies. They use a group of features that are analyzed using data mining algorithms to reduce feature dimensions, then adopt an SVM in the classification stage.

Convolutional neural networks (CNNs) have also become an interesting method for feature extraction and classification of bacterial images [10–14]. In addition to pre-processing approaches such as segmentation, CNNs also utilize data augmentation and transfer learning to improve the ability to classify and identify images. Despite having a high success rate in classification, this approach requires significant computational power and is often time-consuming. In addition, the success of CNNs in the classification task requires computer hardware with a high configuration, such as a graphics processing unit (GPU) or multi-CPU, to complete the task properly.

Zieliński et al. [15] employed the deep convolutional neural networks (CNNs) technique for the recognition of bacterial genera and species. Within this investigation, a CNN was utilized to extract distinctive attributes, while the classification phase involved support vector machines (SVMs). This approach is associated with limitations, including substantial resource consumption, a need for enhanced accuracy, and struggles in discerning-colored bacterial images. These complexities underscore the impetus for continued research in advancing more effective strategies for both feature extraction and classification.

In recent years, various types of orthogonal moments have been developed to describe and extract features from color images, including quaternion moments [16–20] and multi-channel moments [21]. Orthogonal functions with an order of fractions have been shown to have a better ability to represent the fine details of a given function compared to functions that have an order of integers. The interesting characteristics of fractional successive orthogonal moments have led the authors to develop Hu moment invariants and Zernike moments, which are used to extract fine features from colored images of bacteria.

This study applies feature selection techniques to reduce the number of extracted features and remove irrelevant ones [22–25]. One popular approach is the swarm intelligence algorithm [26]. The salp swarm algorithm (SSA) is a swarm intelligence algorithm that has been shown to be effective in identifying the most informative subset of features [27]. Several researchers have improved the SSA for feature selection. For example, Miodrag Zivkovic [28] developed an improved SSA that achieved better results on 21 benchmark datasets. Chaabane et al. [29] used SSA to optimize the selection of feature weights before training a machine-learning model for blind modulation identification. This study uses salp swarm optimization to minimize the misclassification rate in blind modulation identi-

fication. Shikai Wang et al. [30] developed a color image segmentation method using SSA to find the optimal threshold set in the multilevel thresholding process. This study adopts the salp movement behavior to find the optimal threshold set in the multilevel thresholding process. Xiaojun Xie et al. [31] developed an optimal feature selection combination strategy based on SSA for plant disease detection. While the SSA has the potential to select image features, it has some drawbacks, such as the difficulty of setting parameters and the risk of stagnation. In addition to SSA, there is another technique that can be used for feature selection, namely particle swarm optimization (PSO). The PSO algorithm has a simple number of features and fast convergence speed. However, PSO has several limitations, such as population diversity and local optima [32]. Furthermore, PSO does not effectively work for large-scale problems [33]. To address these issues, this study uses a feature selection technique based on the improved salp swarm algorithm (ISSA) [34]. The ISSA incorporates two significant enhancements over the SSA: it employs opposition-based learning (OBL) during SSA initialization to augment population diversity within the search space and incorporates a local search algorithm to enhance exploitation. The ISSA was chosen as a means to optimize the feature selection process to improve the quality of relevant features and classification accuracy.

This paper proposes an advanced multi-feature selection method for bacterial colonies, leveraging the integration of the improved salp swarm algorithm (ISSA), opposition-based learning (OBL), and a local search algorithm (LSA). The primary goal is to improve the accuracy and efficiency of bacterial identification using advanced AI techniques. The proposed method will be compared to other existing algorithms, such as the salp swarm algorithm (SSA) and particle swarm optimization (PSO), on 19 different bacterial features to assess its performance. Additionally, this study will highlight the ISSA algorithm's significance in improving feature selection and classification accuracy, addressing the current limitations of bacterial identification techniques. The scope also includes a comparative analysis of the ISSA algorithm's performance against other prevalent methods, highlighting its advantages and potential contributions to the field of microbial identification and classification.

2. Materials and Methods

This section provides a comprehensive overview of all the experimental procedures conducted in this study. Specifically, it details the microbiological methods used to acquire the image database, the pre-processing, segmentation, feature extraction, feature selection, and image classification techniques employed, and presents step-by-step flowcharts to illustrate each step.

2.1. Pre-Processing

Image pre-processing is a stage of image quality enhancement that precedes further analysis. The first step in pre-processing is to convert the image from RGB color space to grayscale. This step reduces the computational complexity of subsequent processing by using a single intensity channel. Additionally, the grayscale representation can facilitate a clearer interpretation of image features, which is essential for further analysis. This is performed using the National Television System Committee (NTSC) formula, namely $0.299 \times \text{red component} + 0.587 \times \text{green component} + 0.114 \times \text{blue component}$. The second step applies median filtering to reduce noise in the image without losing contrast. To overcome the possibility of uneven lighting, a top-hat transformation with a radius of 200 pixels is carried out and uses a single disc-shaped structural element. The third step, contrast adjustment, is the process of changing the intensity range of pixels in an image to make the image sharper, clearer, and more detailed so that information in the image can be more easily recognized and understood. In this step, the image is normalized using Equation (1); $p_{adjusted}$ is the replaced pixel value, p is the current pixel value, p_{min} is the minimum

pixel value, NC is the normalization coefficient, and MI is the maximum intensity value of the image.

$$\sum p_n \quad p_{adjusted} = \frac{p - p_{min}}{NC} \times MI \quad (1)$$

To segment the object from the background, a binary thresholding technique is applied. In this process, each pixel in the image is converted to a binary value (black or white) based on a predefined threshold value. Binary thresholding is used to segment the object from the background by converting each pixel in the image to a binary value (black or white) based on a predefined threshold. Pixels with an intensity above the threshold are considered white, while pixels with an intensity below the threshold are considered black. This stage isolates the foreground from the background, allowing for the isolation of objects of interest for further analysis. By converting each pixel to a binary value, this process effectively divides the image into distinct parts and extracts relevant information related to specific features of interest. The pre-processing stages are shown in Figure 1.

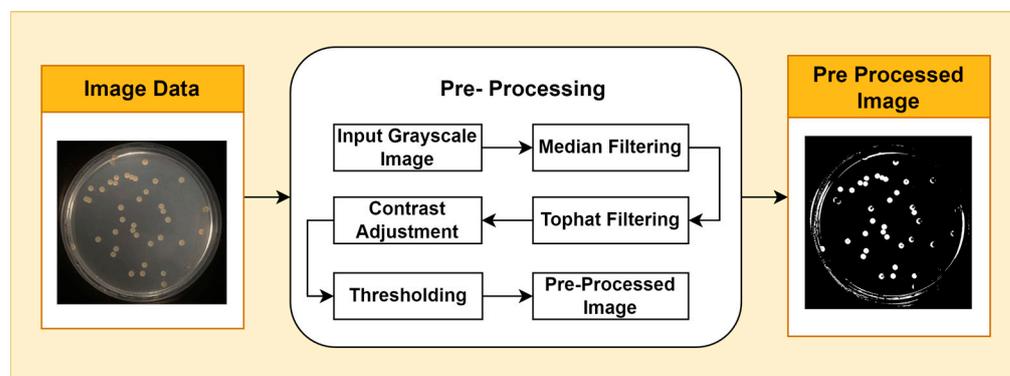


Figure 1. Pre-Processing Stages.

2.2. Segmentation

Image segmentation is the process of partitioning an image into multiple regions with semantically meaningful or visually distinct characteristics. It is a fundamental task in computer vision with a wide range of applications, such as object detection, tracking, and classification. One common approach to image segmentation is contour detection, which aims to identify and extract the boundaries that separate objects in the image. A popular contour detection algorithm is Canny’s edge detector, which uses a multi-stage process to identify edges with high probability and low localization error. Once the contours of bacterial cells have been detected, the Douglas–Peucker algorithm can be used to simplify them by approximating them with a sequence of line segments. The epsilon parameter controls the maximum distance allowed between a point on the original contour and the corresponding point on the approximated contour. A smaller epsilon value will result in a more detailed approximation, while a larger epsilon value will result in a more simplified approximation. The choice of epsilon value depends on the desired level of contour detail. Circularity analysis is a technique used to measure the extent to which an object or structure in an image approaches a circular shape. It is often used in biological image processing to identify and segment cells. A variety of circularity metrics can be used, such as the aspect ratio, circularity index, and Feret’s diameter. The choice of metric depends on the specific application. In this study, circularity analysis was performed using the circular Hough transform (Figure 2).

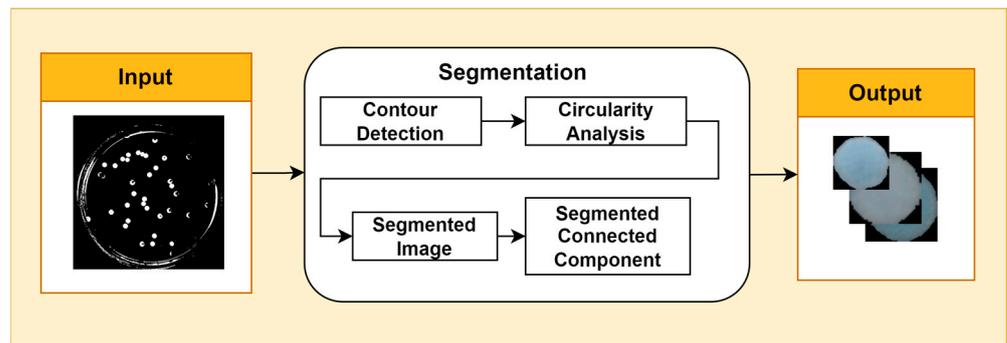


Figure 2. Segmentation Stage.

2.3. Feature Extraction

Feature extraction is an essential step in bacterial image analysis, as it allows us to identify, classify, and characterize different types of bacteria based on their visual features. There are many different types of features that can be extracted from bacterial images, including geometric features, color features, texture features, and moment features.

Geometric features provide information about the shape and size of bacterial cells, such as area, perimeter, eccentricity, circularity, roughness, convex area, and center of mass (centroid x and centroid y). Color features are measured using color metrics such as hue, saturation, and value (HSV). Texture features describe the visual patterns in an image, such as contrast, energy, homogeneity, dissimilarity, correlation, and opacity. Moment features are mathematical representations of the geometric properties of objects in an image.

In this study, we focus on two types of moment features: Hu moments and Zernike moments (Figure 3). Hu moments are invariant to translation, rotation, and scale, making them useful for feature extraction from bacterial images, which can vary in size and orientation. Zernike moments are more sensitive to small changes in image structure, making them useful for discriminating between different types of bacteria.

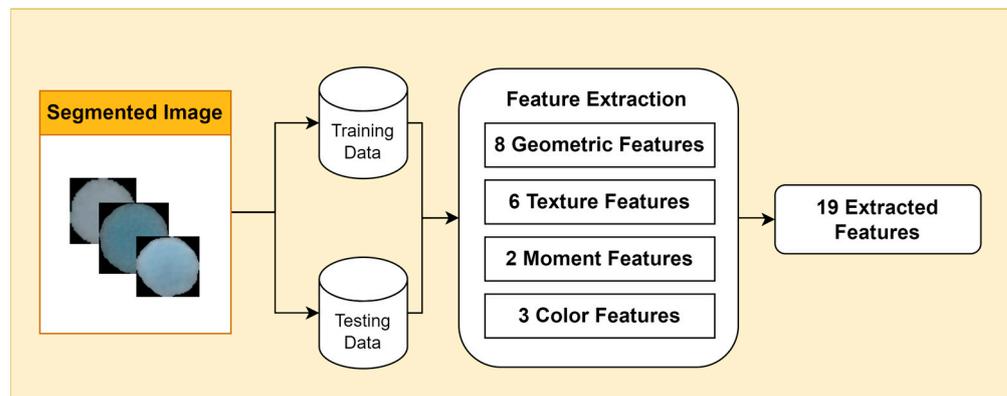


Figure 3. Feature Extraction Stages.

2.4. Feature Selection

2.4.1. Salp Swarm Algorithm

The salp swarm algorithm is a recently developed nature-inspired metaheuristic algorithm that mimics the particular swarming behavior of marine salps. Salps usually live in groups and often form shoals called salp chains. The first salp is referred to as the leader, while the others are followers. In the mathematical model of the salp, the position of the salp leader is updated using Equation (2).

$$x_j^1 = \begin{cases} F_j + c_1((ub_j - lb_j)c_2 + lb_j) & c_3 \geq 0.5 \\ F_j - c_1((ub_j - lb_j)c_2 + lb_j) & c_3 < 0.5 \end{cases} \quad (2)$$

where the value x_j^1 indicates the position of the leader in the salp chain, and F_j shows the position of the food source in the j dimension. Moreover, it represents the lower bound value of the j dimension. In b_j represents the upper bound value of j . Then, c_1 , c_2 , and c_3 represent random values; c_1 is the most important controlling parameter in SSA because it is responsible for maintaining the balance between exploration and exploitation in SSA. Equation (3) is used to calculate the value of parameter c .

$$c_1 = 2e^{-\left(\frac{t}{L}\right)^2} \tag{3}$$

In Equation (3), t represents the current iteration, and L represents the maximum number of iterations of the algorithm. Parameter values of c_2 and c_3 denote random values in the range $[0, 1]$. The positions of the followers in the salp chain are updated using Equation (4),

$$x_j^i = \frac{1}{2} \left(x_j^i + x_j^{i-1} \right) \tag{4}$$

where $i \geq 2$ and x_j^i is the position of i follower in dimension j search space. The SSA algorithm is shown in Algorithm 1.

Algorithm 1 Pseudocode of the SSA algorithm

```

initialize the salps' positions  $x_i$  ( $i = 1, 2, \dots, n$ )
while ( $t < \text{max iterations}$ )
    determine the fitness value of each salp
     $F = \text{best salp ((search-agent)}$ 
    Update the value of  $c$  parameter using Equation (2)
    for every salp ( $x_i$ )
        if ( $i == 1$ )
            Update leader position using Equation (1)
        Else
            Update follower position using Equation (3)
        end if
    end for
    reposition the salps that go out of the search space based on the lower and upper bounds of
    problem variables
     $T = t + 1$ 
end while
return  $F$ 

```

2.4.2. Opposition-Based Learning

OBL is an optimization technique used to improve the quality of the initial solution in a population by introducing variations to the solution. OBL operates by searching in two directions in the search space. Both directions involve the initial solution, as well as the opposite direction, which is represented by a solution that has the opposite properties. In the end, OBL chooses the best solutions from all the existing ones.

Opposite numbers, denoted as x , are defined as real numbers on the interval. The opposite number is \tilde{x} , denoted as follows and determined using Equation (5),

$$\tilde{x} = lb + ub - x \tag{5}$$

Equation (5) can be generalized to apply in a multidimensional search space. Therefore, Equations (6) and (7) will represent the position of each search agent and its opposite position in order to generalize this as follows:

$$x = [x_1, x_2, x_3, \dots, x_D] \tag{6}$$

$$\tilde{x} = [\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \dots, \tilde{x}_D] \tag{7}$$

The values of all the inner elements \tilde{x} will be determined using Equation (8):

$$\tilde{x}_j = lb_j + ub_j - x_j \quad j = 1, 2, 3, \dots, D \tag{8}$$

In the optimization based on opposed populations approach, the fitness function is represented as $f()$. Therefore, if the fitness value $f(\tilde{x})$ of the opposite solution is superior to $f(x)$, that of the initial solution x , then $x = \tilde{x}$; if not, $x = x$.

The integration of OBL (opposition-based learning) with SSA (salp swarm algorithm) involves the following steps: initializing the salp positions x as x_i and determining the opposite positions of the salp population OX as \tilde{x}_i , where $(i = 1, 2, \dots, n)$; subsequently, the n strongest salps selected from $\{X \cup OX\}$ will constitute the new initial population for SSA.

2.4.3. Local Search Algorithm

The developed local search algorithm (LSA) is shown in Algorithm 2. LSA will be called at the end of each iteration in SSA to improve the best solution. Initially, LSA stores the best solution value obtained from SSA in the last SSA iteration into the Temp variable. LSA performs a number of iterations to increase temperature. In each LSA iteration, the LSA randomly selects three features from the Temp. LSA rearranges the selected features based on their values. In addition, LSA will determine the fitness value of the new solution; if it is better than the initial fitness value, it will be set to Temp; otherwise, it will remain unchanged.

Algorithm 2 Pseudocode of the LSA algorithm

```

Temp = F (where F represents the current best solution at end of SSA's current iteration)
Lt = 1 (Lt is a variable used to store the current iteration of local search algorithm)
while (Lt < maximum number of local iterations)
    Randomly select three features from Temp
    if selected-feature == 1 (1 means the feature is selected and 0 means not selected)
        selected-feature = 0
    Else
        selected-feature = 1
    end if
    Calculate the fitness value of Temp
    if f(Temp) < f(F)
        F = Temp
    end if
    Lt = Lt + 1
end while
return F

```

2.4.4. Improved Salp Swarm Algorithm

ISSA is an improvement on the SSA algorithm. The stages of the proposed ISSA algorithm are shown in Figure 4. The steps of the ISSA algorithm include initializing SSA by generating the number of salps based on population size and selecting a feature subset randomly from the entire feature set. This is followed by the application of OBL to find the opposite solution from each initial solution and the reverse solution, which is calculated based on the KNN classification accuracy error, as shown in Figure 5. Additionally, SSA gives a t value to the best solution that OBL selects as having the lowest classification accuracy error. Depending on whether the salp is the leader or a follower in the salp chain, Equation (2) or (4) updates the position of each salp after that. Fitness evaluation is performed by determining the fitness values of all salps, and these values are updated if a better solution is found. The LSA algorithm is applied to the t best solution to find a better

solution, and the t value is updated if a better solution is found. The ISSA algorithm is repeated in iterations. ISSA returns the best solution that represents the best feature subset. In the testing phase, the features selected for the best solution are used to evaluate ISSA's performance on the test dataset.

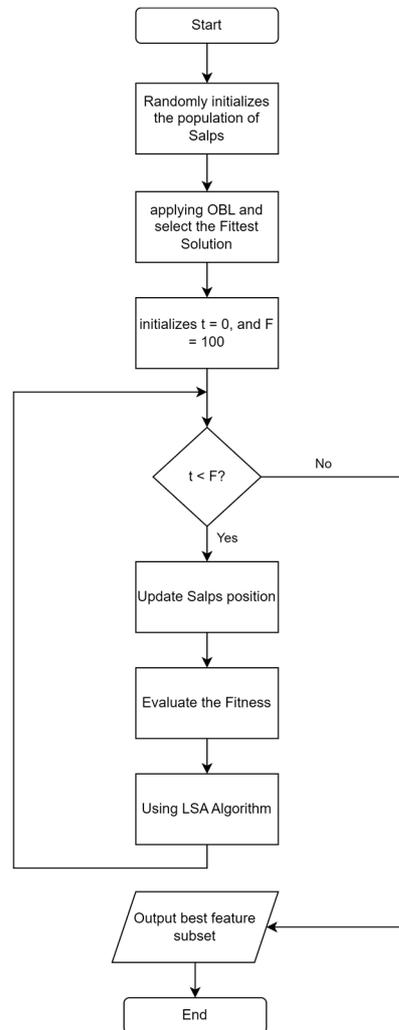


Figure 4. Flowchart ISSA algorithm based on OBL and LSA algorithms.

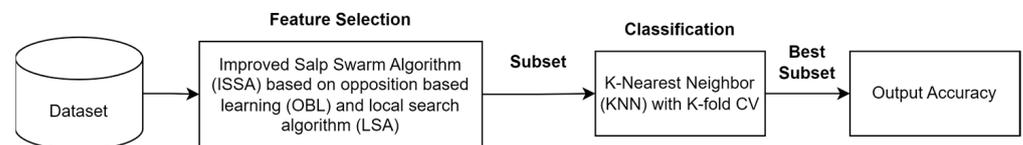


Figure 5. The Second Stage of Feature Selection.

The ISSA algorithm is compared to the standard SSA algorithm and the particle swarm optimization (PSO) algorithm. The population size utilized by all algorithms is 10, the maximum number of iterations for each optimization algorithm is 100, and the number of LSA iterations is fixed at 10. An indicator to evaluate the performance of the optimization algorithm is the classification accuracy using KNN classification.

3. Results and Discussion

Given the growing concern about the adverse impact of microorganisms on human health and ecological systems, this study aims to propose an advanced multi-feature selection method for bacterial colonies. This method incorporates the salp swarm algorithm

(SSA), opposition-based learning (OBL), and local search algorithm (LSA), with the primary objective of enhancing the accuracy and efficiency of bacterial identification through sophisticated artificial intelligence techniques. This study will conduct a comparison between the proposed method and other existing algorithms, such as the salp swarm algorithm (SSA) and particle swarm optimization (PSO), across 19 distinct bacterial features to evaluate its performance. Furthermore, this research will underscore the significance of the ISSA algorithm in refining feature selection and classification accuracy while addressing current constraints within bacterial identification techniques. The research scope also encompasses an analysis that compares the performance of the ISSA algorithm with other prevalent methods, emphasizing its benefits and potential contributions in the realm of microbial identification and classification.

3.1. Dataset

To evaluate and validate the performance of the ISSA algorithm compared to other algorithms, the bacterial benchmark dataset from the CINATE repository (Center for Innovation and Technological Support, ESB, UCP) was used [35]. Details of the datasets used can be seen in Table 1.

Table 1. Details of the datasets used.

No.	Dataset	Number of Features	Number of Samples
1	<i>E. coli</i> bacteria	19	427
2	<i>S. aureus</i> bacteria	19	371
3	<i>P. aeruginosa</i> bacteria	19	458

3.2. Parameter Setting

All experiments were conducted using a 10-fold cross-validation method. This validation method divides the data into 10 parts, with 70% of the data used for training and 30% used for testing. The experiment was repeated 10 times, and the results of the 10 trials were averaged to produce the performance indicators. Therefore, the results listed in Table 2 reflect the average value of 100 iterations in terms of accuracy, suitability, and selected features.

Table 2. Comparison between ISSA, PSO, and SSA methods based on average accuracy, average number of selected features, and average fit in 10 trials.

Data	Accuracy			Number of Selected Feature			Fitness		
	PSO	SSA	ISSA	PSO	SSA	ISSA	PSO	SSA	ISSA
Data Test 1	67.686	72.489	72.925	5	5	4	0.323	0.274	0.270
Data Test 2	68.995	69.432	74.672	6	9	6	0.310	0.307	0.254
Data Test 3	66.375	68.995	72.489	5	8	6	0.335	0.311	0.276
Data Test 4	63.755	68.122	78.602	5	5	5	0.361	0.318	0.214
Data Test 5	68.558	72.489	72.925	6	6	6	0.314	0.275	0.271
Data Test 6	68.995	70.305	72.925	9	8	5	0.311	0.298	0.270
Data Test 7	65.502	69.868	71.179	5	7	5	0.344	0.301	0.287
Data Test 8	70.742	71.179	73.799	5	9	6	0.292	0.290	0.262
Data Test 9	69.868	68.122	72.925	7	9	5	0.301	0.320	0.270
Data Test 10	69.868	71.179	75.109	7	8	6	0.301	0.289	0.249
Average	68.034	70.218	73.755	6	7	5	0.319	0.298	0.262

3.3. Results and Analysis

This section presents a comprehensive overview of the key findings and results of the conducted experiments. Specifically, this study compares the performance of SSA, PBO, and ISSA on bacterial colony images. The evaluation of ISSA's efficacy against other algorithms is based on three key metrics: classification accuracy, the number of selected features, and the match value. Table 2 summarizes the results of the comparative analysis of SSA, PSO, and ISSA. ISSA outperforms SSA and PSO in classification accuracy on three datasets with 19 features, achieving an accuracy of 73.75%. Additionally, ISSA selects a smaller number of features than SSA and PSO in all 10 trials. In addition to accuracy, ISSA also demonstrates superior performance in match value, with a significantly lower classification error rate of 0.2623 compared to SSA. These results highlight ISSA's significant improvement over the original SSA algorithm in terms of classification accuracy, feature selection, and match value.

Figure 6 demonstrates ISSA's superior exploration capability compared to other optimization algorithms. This is evident in ISSA's ability to access regions of the search space that are inaccessible to other algorithms, indicating its superior ability to maintain a diverse set of solutions. Additionally, Table 2 shows that ISSA consistently achieves better fit values than other algorithms, demonstrating its ability to avoid local optima. Finally, ISSA can select a smaller number of features than other optimization algorithms, further demonstrating its exploration capability.

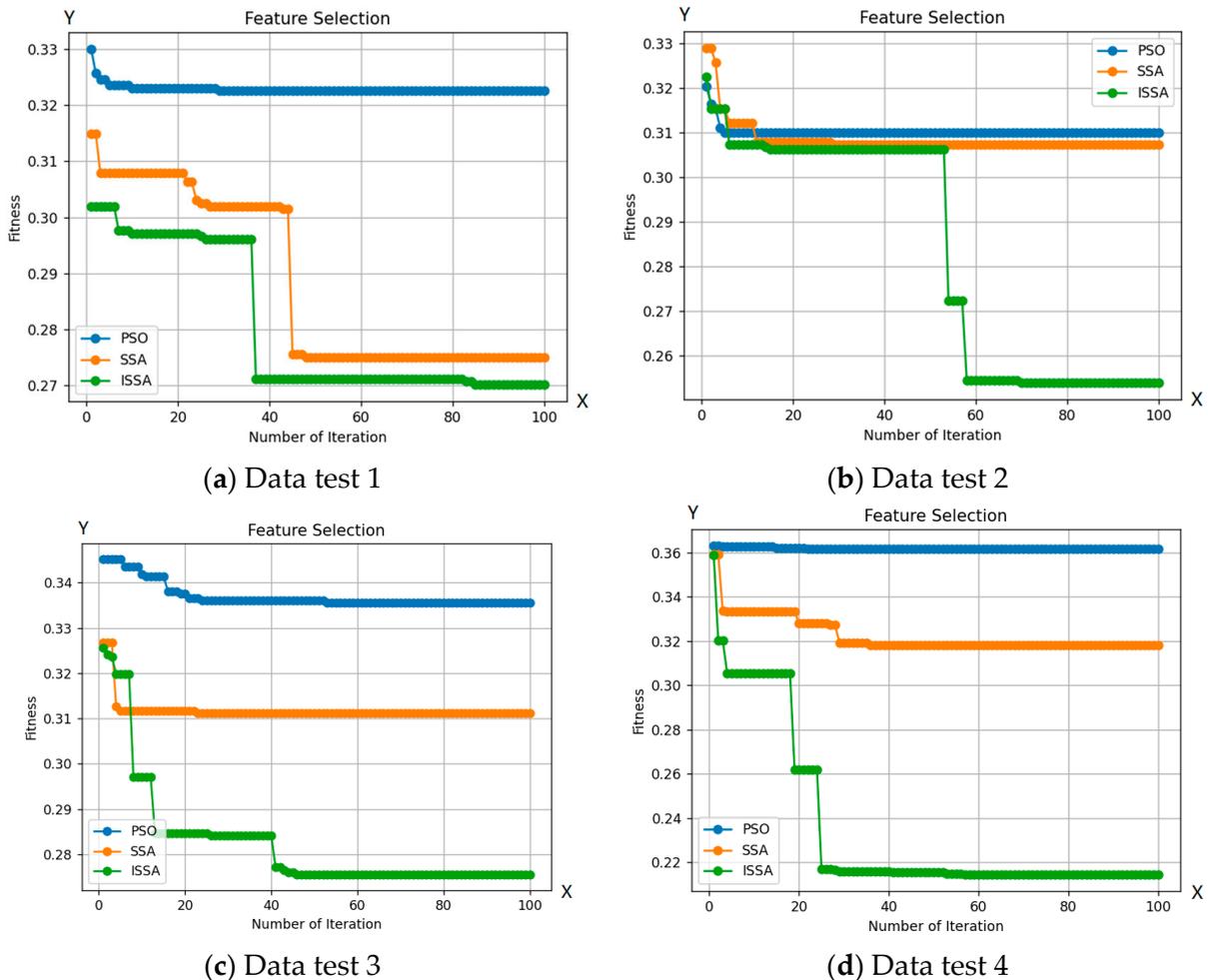
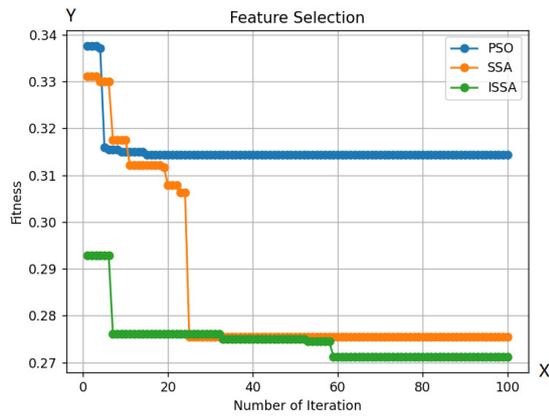
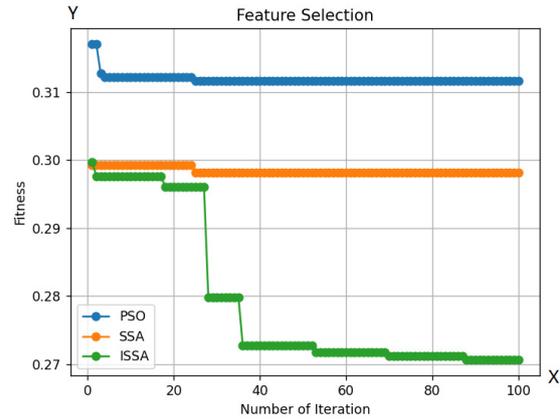


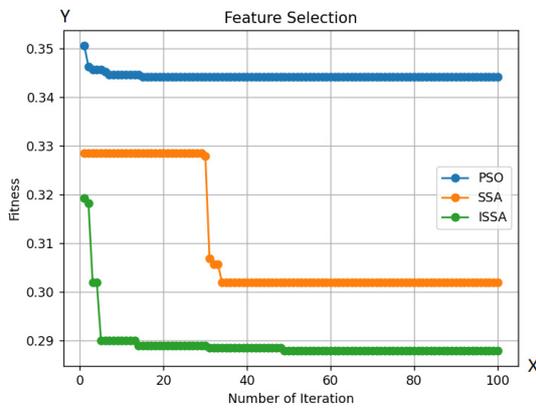
Figure 6. Cont.



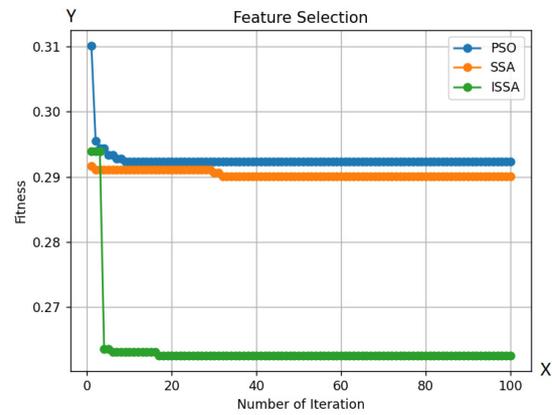
(e) Data test 5



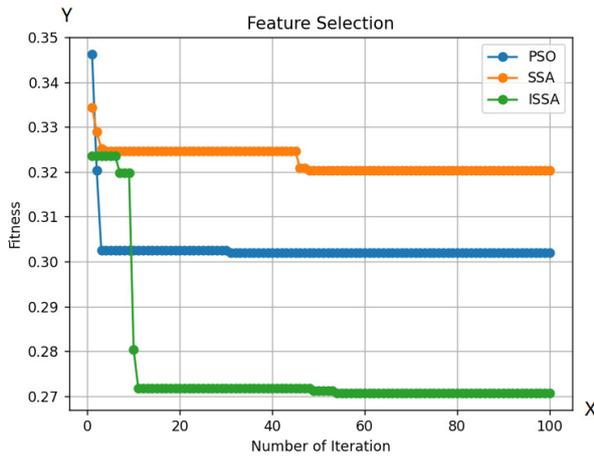
(f) Data test 6



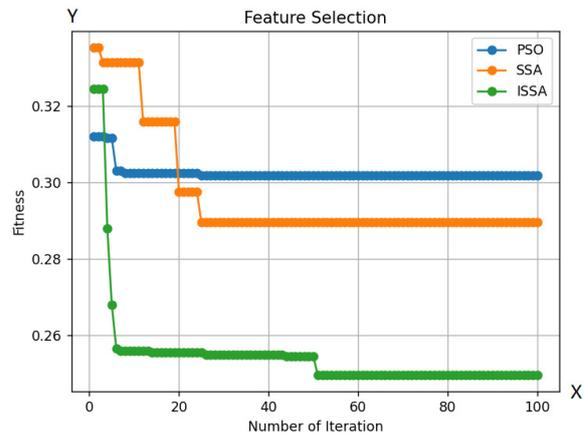
(g) Data test 7



(h) Data test 8



(i) Data test 9



(j) Data test 10

Figure 6. Comparison between the ISSA, PSO, and SSA methods with 10 test data.

Based on the number of features selected by ISSA, as seen in Table 2, ISSA demonstrates superior exploration capabilities compared to the SSA and PSO algorithms. This is further strengthened by the selection of a smaller number of features from the 10 included test data. Out of the 19 features tested on 10 test data, ISSA is able to choose 5 features to identify bacteria. The use of a more limited set of features can enhance bacterial classification. The exploration superiority of ISSA is also evident through its ability to select a smaller number of features, surpassing other optimization algorithms, as reflected in the accuracy results

obtained from the 10 test data. Additionally, ISSA shows a significantly lower classification error rate in 100 iterations compared to the SSA and PSO algorithms.

3.4. Comparison with Other Methods

In this section, the performance of the ISSA feature selection is proposed by considering the fitness value and classification error rate compared to other published methods. For instance, research conducted by Nie et al. [36] suggests the automation of the identification and classification of bacterial colony areas using a convolutional deep belief network (CDBN). The aim of this technique is to create a profound representation of small parts of images.

Through the training process, the front and back parts of the images are distinguished with a high level of accuracy using a support vector machine (SVM). Subsequently, a convolutional neural network (CNN) is employed to detect the presence of bacteria in those images. The accuracy of ISSA KNN with the K-fold cross-validation method achieves better results than those reported by Nie and his co-authors (Table 3).

Table 3. The comparison with other bacteria species recognition methods.

	CNN	SIFT + KNN	SIFT + SVM	ISSA + KNN
ACC (%)	0.6210	0.4967	0.5400	0.7375

4. Conclusions

This study presents a multi-feature selection technique, the improved salp swarm algorithm (ISSA), for accurate bacterial colony classification. We have significantly improved the ISSA approach's effectiveness and performance by integrating the salp swarm algorithm (SSA) with the opposition-based learning (OBL) strategy and a local search algorithm (LSA). Our research methodology is a streamlined three-stage process that enables the automated categorization of bacteria based on their distinctive features.

The ISSA technique has demonstrated outstanding performance through rigorous evaluation of a comprehensive bacterial colony feature dataset and comparative analysis with several other algorithms across ten distinct test datasets. Specifically, ISSA achieved a superior classification accuracy of 73.75% on three datasets with 19 features. Additionally, ISSA excelled at determining the optimal feature count, resulting in a significantly reduced classification error rate of 0.249. These numerical results demonstrate the ISSA technique's robustness and effectiveness in addressing the complex challenges of bacterial feature selection and classification.

While the ISSA technique has the potential to revolutionize bacterial colony classification, certain limitations must be considered. The ISSA algorithm's computational complexity and potential constraints in scaling the methodology for larger datasets highlight the need for further optimization and resource allocation. Additionally, practical implementation of ISSA in real-time applications may require additional fine-tuning to maximize its efficiency and applicability.

Overall, the ISSA technique represents a significant advancement in bacterial colony classification, offering improved accuracy and efficiency while addressing the computational complexities of the classification task. By acknowledging the limitations and leveraging the strengths of the proposed technique, this study contributes to the ongoing efforts to refine bacterial analysis methodologies for enhanced scientific and practical applications.

Author Contributions: Conceptualization, writing—original draft preparation, A.I. methodology, and writing—review and editing, K.M. experimental work, validation, visualization, and data curation, R.F. formal analysis and investigation. writing—review and editing, M.M. Supervision, I.M.R.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi, Republik Indonesia (Ministry of Education, Culture, Research, and Technology, Republic of Indone-

sia), which supported this activity through DIPA DRTPM 159/E5/PG.02.00.PL/2023 in the Penelitian Dasar Kompetitif Nasional (PDKN) funds scheme under Contract Number 054/E5/PG.02.00.PL/2023.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in FigShare at <https://doi.org/10.6084/m9.figshare.20109377.v2> (accessed on 15 August 2023), reference number [35].

Acknowledgments: Our deepest gratitude to all those who have helped successfully implement this research, especially Direktorat Riset, Teknologi, Pengabdian Kepada Masyarakat (DRTPM), and Lembaga Penelitian Pengabdian Kepada Masyarakat-Penjaminan Mutu (LPPM-PM) Universitas Samudra, which already invested trust in our capability to conduct this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Serwecińska, L. Antimicrobials and antibiotic-resistant bacteria. *Water* **2020**, *12*, 3313. [CrossRef]
2. Mokrani, S.; Nabti, E.H.; Cruz, C. Recent trends in microbial approaches for soil desalination. *Appl. Sci.* **2022**, *12*, 3586. [CrossRef]
3. Benami, M.; Gillor, O.; Gross, A. Potential microbial hazards from graywater reuse and associated matrices: A review. *Water Res.* **2016**, *106*, 183–195. [CrossRef] [PubMed]
4. Joy, C.; Sundar, G.N.; Narmadha, D. AI driven automatic detection of bacterial contamination in water: A review. In Proceedings of the 5th International Conference on Intelligent Computing and Control Systems (ICICCS 2021), Madurai, India, 6–8 May 2021; pp. 1281–1285. [CrossRef]
5. Nurliyana, M.R.; Sahdan, M.Z.; Wibowo, K.M.; Muslihati, A.; Saim, H.; Ahmad, S.A.; Sari, Y.; Mansor, Z. The detection method of *Escherichia coli* in water resources: A Review. *J. Phys. Conf. Ser.* **2018**, *995*, 012065. [CrossRef]
6. Elaziz, M.A.; Hosny, K.M.; Hemedan, A.A.; Darwish, M.M. Improved recognition of bacterial species using novel fractional-order orthogonal descriptors. *Appl. Soft Comput. J.* **2020**, *95*, 106504. [CrossRef]
7. Panicker, R.O.; Kalmady, K.S.; Rajan, J.; Sabu, M.K. Automatic detection of *Tuberculosis bacilli* from microscopic sputum smear images using deep learning methods. *Biocybern. Biomed. Eng.* **2018**, *38*, 691–699. [CrossRef]
8. Luo, J.; Ser, W.; Liu, A.; Yap, P.H.; Liedberg, B.; Rayatpisheh, S. Microorganism image classification with circle-based multi-region binarization and mutual-information-based feature selection. *Biomed. Eng. Adv.* **2021**, *2*, 100020. [CrossRef]
9. Bonah, E.; Huang, X.; Yi, R.; Aheto, J.H.; Osa, R.; Golly, M. Electronic nose classification and differentiation of bacterial foodborne pathogens based on support vector machine optimized with particle swarm optimization algorithm. *J. Food Process Eng.* **2019**, *42*, e13236. [CrossRef]
10. Abdullah, H.-C.K.; Ali, S.; Khan, Z.; Hussain, A.; Athar, A. Computer vision based deep learning approach for the microscopic images. *Water* **2022**, *22*, 2219. [CrossRef]
11. Wahid, M.F.; Ahmed, T.; Habib, M.A. Classification of microscopic images of bacteria using deep convolutional neural network. In Proceedings of the 2018 10th International Conference on Electrical and Computer Engineering (ICECE), Dhaka, Bangladesh, 20–22 December 2018; pp. 217–220. [CrossRef]
12. Talo, M. An automated deep learning approach for bacterial image classification. *arXiv* **2019**, arXiv:1912.08765. [CrossRef]
13. Li, S.; Song, W.; Fang, L.; Chen, Y.; Ghamisi, P.; Benediktsson, J.A. Deep learning for hyperspectral image classification: An overview. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6690–6709. [CrossRef]
14. Yanik, H.; Hilmi Kaloğlu, A.; Değirmenci, E. Detection of *Escherichia Coli* bacteria in water using deep learning. *Teh. Glas.* **2020**, *14*, 273–280. [CrossRef]
15. Zieliński, B.; Plichta, A.; Misztal, K.; Spurek, P.; Brzywczy-Włoch, M.; Ochońska, D. Deep learning approach to bacterial colony classification. *PLoS ONE* **2017**, *12*, e0184554. [CrossRef] [PubMed]
16. Singh, C.; Singh, J. Quaternion generalized Chebyshev-Fourier and pseudo-Jacobi-Fourier moments for color object recognition. *Opt. Laser Technol.* **2018**, *106*, 234–250. [CrossRef]
17. Wang, C.; Wang, X.; Li, Y.; Xia, Z.; Zhang, C. Quaternion polar harmonic Fourier moments for color images. *Inf. Sci.* **2018**, *450*, 141–156. [CrossRef]
18. Chen, B.; Shu, H.; Coatrieux, G.; Chen, G.; Sun, X.; Coatrieux, J.L. Color Image Analysis by Quaternion-Type Moments. *J. Math. Imaging Vis.* **2015**, *51*, 124–144. [CrossRef]
19. Huang, C.; Li, J.; Gao, G. Review of quaternion-based color image processing methods. *Mathematics* **2023**, *11*, 2056. [CrossRef]
20. He, B.; Liu, J.; Yang, T.; Xiao, B.; Peng, Y. Quaternion fractional-order color orthogonal moment-based image representation and recognition. *Eurasip J. Image Video Process.* **2021**, *2021*, 17. [CrossRef]
21. Hosny, K.M.; Darwish, M.M.; Eltoukhy, M.M. Novel multi-channel fractional-order radial harmonic fourier moments for color image analysis. *IEEE Access* **2020**, *8*, 40732–40743. [CrossRef]
22. Tariq Al-Rayes, H.; Tariq Ibrahim, H.; Jalil Mazher, W.; Ucan, O.N.; Bayat, O. Feature selection using salp swarm algorithm for real biomedical datasets. *IJCSNS Int. J. Comput. Sci. Netw. Secur.* **2017**, *17*, 13–20.

23. Elaziz, M.A.; Moemen, Y.S.; Hassanien, A.E.; Xiong, S. Toxicity risks evaluation of unknown FDA biotransformed drugs based on a multi-objective feature selection approach. *Appl. Soft Comput.* **2020**, *97*, 105509. [[CrossRef](#)]
24. Abd El Aziz, M.; Hassanien, A.E. An improved social spider optimization algorithm based on rough sets for solving minimum number attribute reduction problem. *Neural Comput. Appl.* **2018**, *30*, 2441–2452. [[CrossRef](#)]
25. Ewees, A.A.; El Aziz, M.A.; Hassanien, A.E. Chaotic multi-verse optimizer-based feature selection. *Neural Comput. Appl.* **2019**, *31*, 991–1006. [[CrossRef](#)]
26. Brezočnik, L.; Fister, I.; Podgorelec, V. Swarm intelligence algorithms for feature selection: A review. *Appl. Sci.* **2018**, *8*, 1521. [[CrossRef](#)]
27. Mirjalili, S.; Gandomi, A.H.; Mirjalili, S.Z.; Saremi, S.; Faris, H.; Mirjalili, S.M. Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. *Adv. Eng. Softw.* **2017**, *114*, 163–191. [[CrossRef](#)]
28. Zivkovic, M.; Stoean, C.; Chhabra, A.; Budimirovic, N.; Petrovic, A.; Bacanin, N. Novel improved salp swarm algorithm: An application for feature selection. *Sensors* **2022**, *22*, 1711. [[CrossRef](#)]
29. Chaabane, S.B.; Belazi, A.; Kharbech, S.; Bouallegue, A. Improved salp swarm optimization algorithm: Application in feature weighting for blind modulation identification. *Electronics* **2021**, *10*, 2002. [[CrossRef](#)]
30. Wang, S.; Jia, H.; Peng, X. Modified salp swarm algorithm based multilevel thresholding for color image segmentation. *Math. Biosci. Eng.* **2020**, *17*, 700–724. [[CrossRef](#)]
31. Xie, X.; Xia, F.; Wu, Y.; Liu, S.; Yan, K.; Xu, H.; Ji, Z. A novel feature selection strategy based on salp swarm algorithm for plant disease detection. *Plant Phenomics* **2023**, *5*, 0039. [[CrossRef](#)]
32. Qiu, C. A novel multi-swarm particle swarm optimization for feature selection. *Genet. Program. Evolvable Mach.* **2019**, *20*, 503–529. [[CrossRef](#)]
33. Gu, S.; Cheng, R.; Jin, Y. Feature selection for high-dimensional classification using a competitive swarm optimizer. *Soft Comput.* **2018**, *22*, 811–822. [[CrossRef](#)]
34. Tubishat, M.; Idris, N.; Shuib, L.; Abushariah, M.A.M.; Mirjalili, S. Improved salp swarm algorithm based on opposition based learning and novel local search algorithm for feature selection. *Expert Syst. Appl.* **2020**, *145*, 113122. [[CrossRef](#)]
35. Rodrigues, F.K.T.; Miguel, P.; Luis, J. *Petri Dishes Digital Images Dataset of E. coli, S. aureus and P. aeruginosa*; Centre of Biotechnology and Fine Chemistry, Catholic University of Portugal: Lisboa, Portugal, 2022. Available online: https://figshare.com/articles/dataset/Dataset_bioengineering_17489364/20109377/2 (accessed on 17 July 2023).
36. Nie, D.; Shank, E.A.; Jojic, V. A deep framework for bacterial image segmentation and classification. In Proceedings of the BCB '15: Proceedings of the 6th ACM Conference on Bioinformatics, Computational Biology and Health Informatics, Atlanta, GA, USA, 9–12 September 2015; pp. 306–314. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.