DE GRUYTER Journal of Intelligent Systems 2024; 33: 20230267

Review Article

Abbas Fadhil Jasim AL-Gburi*, Mohd Zakree Ahmad Nazri,
Mohd Ridzwan Bin Yaakub, and Zaid Abdi Alkareem Alyasseri*

A systematic review of symbiotic organisms search algorithm for data clustering and predictive analysis

https://doi.org/10.1515/jisys-2023-0267
received November 19, 2023; accepted December 30, 2023

Abstract: In recent years, the field of data analytics has witnessed a surge in innovative techniques to handle the ever-increasing volume and complexity of data. Among these, nature-inspired algorithms have gained significant attention due to their ability to efficiently mimic natural processes and solve intricate problems. One such algorithm, the symbiotic organisms search (SOS) Algorithm, has emerged as a promising approach for clustering and predictive analytics tasks, drawing inspiration from the symbiotic relationships observed in biological ecosystems. Metaheuristics such as the SOS have been frequently employed in clustering to discover suitable solutions for complicated issues. Despite the numerous research works on clustering and SOS-based predictive techniques, there have been minimal secondary investigations in the field. The aim of this study is to fill this gap by performing a systematic literature review (SLR) on SOS-based clustering models focusing on various aspects, including the adopted clustering approach, feature selection approach, and hybridized algorithms combining K-means algorithm with different SOS algorithms. This review aims to guide researchers to better understand the issues and challenges in this area. The study assesses the unique articles published in journals and conferences over the last ten years (2014–2023). After the abstract and full-text eligibility analysis, a limited number of articles were considered for this SLR. The findings show that various SOS methods were adapted as clustering and feature selection methods in which CSOS, discrete SOS, and multiagent SOS are mostly used for the clustering applications, and binary SOS, binary SOS with S-shaped transfer functions, and BSOSVT are used for feature selection problems. The findings also revealed that, of all the selected studies for this review, only a few studies specifically focused on hybridizing SOS with K-means algorithm for automatic data clustering application. Finally, the study analyzes the study gaps and the research prospects for SOS-based clustering methods.

Keywords: symbiotic organisms search algorithm, clustering, classification, feature selection, systematic review

*** Corresponding author: Abbas Fadhil Jasim AL-Gburi,** Faculty of Information Science & Technology, Universiti Kebangsaan Malaysia, Bangi, 43600 Selangor, Malaysia, e-mail: p104340@siswa.ukm.edu.my
*** Corresponding author: Zaid Abdi Alkareem Alyasseri,** Information Technology Research and Development Center (ITRDC), University of Kufa, Najaf, 54001, Iraq; College of Engineering, University of Warith Al-Anbiyaa, Karbala, 56001, Iraq, e-mail: zaid.alyasseri@uokufa.edu.iq
Mohd Zakree Ahmad Nazri: Faculty of Information Science & Technology, Universiti Kebangsaan Malaysia, Bangi, 43600 Selangor, Malaysia, e-mail: zakree@ukm.edu.my
Mohd Ridzwan Bin Yaakub: Faculty of Information Science & Technology, Universiti Kebangsaan Malaysia, Bangi, 43600 Selangor, Malaysia, e-mail: ridzwanyaakub@ukm.edu.my

1 Introduction

In recent times, data mining techniques have been prominently utilized for analyzing big data and decision-making in various areas, especially document clustering and classification [1], networking sensor problems [2], parameter tuning problems in artificial neural network (ANN) models [3], and maximization problem in renewable energy production [4,5].

Clustering is an important data mining task for grouping data points based on similarity [6]. Clustering methods require multiple parameters and are commonly used in big spaces. They exploit noisy, incomplete, and sampled data for this purpose; this significantly impacts their efficacy, which varies considerably across applications and data types. Numerous studies have been conducted, and a range of clustering strategies have been proposed, including Kernel Methods like support vector machine (SVM) [7], self-organizing maps [8], and the K-means [9]. This strategy was very effective but had drawbacks, such as initializations as well as speedy convergence to the local optimum. These limitations are addressed with an automated data clustering algorithm, known as metaheuristics, motivated by natural or physical events that occur [10]. The metaheuristic algorithm for optimization is a stochastic method for solving optimization issues; similar to a searching algorithm, it seeks the most optimal solution [11–13]. One of the methods proposed based on metaheuristics includes swarm intelligence algorithms such as ant colony optimization (ACO) [14,15] and particle swarm optimization (PSO) [16,17]. These algorithms were efficient and robust in handling various optimization issues. However, these algorithms frequently become trapped in local minima, are computationally demanding and complex, converge slowly, and are incompatible with classes of functions.

SOS is a stochastic metaheuristic method that uses randomization to determine a set of solutions, which has been proposed by Cheng and Prayogo [18]. This method was motivated by the shared behavior of organisms for survivability. This is called a symbiotic relationship when each organism relies interactively on each other in the ecosystem. According to [19,20], the SOS metaheuristic algorithm has made significant achievements in handling complex engineering applications and optimization problems for some reasons, such as: (1) the algorithms are easy to implement (2) do not require any tuning parameters to train, (3) they find better local optimal compared to other algorithms, and (4) they can be applied to solve various issues in different areas.

Subsequently, SOS algorithms have been prominently used, especially in engineering [21]. Despite their potential to increase retrieval recall and computational efficiency, only a few research in the literature have focused on clustering and classification, particularly the clustering issues and feature selection [21,22]. Prior studies focused on clustering and classification for optimal feature selection in brain–computer interfaces (BCIs) [23] and satellite image classification problems [24]. Among the literature, [25,26] were on track with text clustering based on the SOS algorithm.

In their empirical studies, Mohammadzadeh and Gharehchopogh [26] proposed an SOS-based method for feature extraction problems, while Cheng et al. [27] addressed text classification problems. Currently, no single review study has solely addressed methodologies for clustering and unsupervised feature selection. Although Gharehchopogh et al. [21] presented a solid review of the subject, the author did not focus explicitly on the clustering and classification approaches and included only a small number of publications that have been published concerning clustering and classification. Then, Abdullahi et al. [28] presented a comprehensive survey of SOS algorithms and their implementation. Regardless, no specific justification was provided based on the standard SLR approach for the selected articles to avoid biased reviews [29]. Hence, this research provides an SLR approach to enhance the efficacy of reviewing, summarizing, and analyzing the patterns observed in published articles. This research focuses on proferring the answers to the designed research questions (RQs) to determine the most widely used variations of the SOS-based method for both the feature selection and clustering problems, what are the reported hybridizations of the SOS algorithm with the K-means clustering algorithms in dealing with clustering problems, and importantly, what are the open issues and the prospect of SOS-based methods in the clustering and feature selection problems. Consequently, this study offers researchers and practitioners valuable insights and future directions. This study is conducted following the principles outlined by Kitchenham et al. [30]. These SLR contributions are as follows:

- Present an SLR of the existing SOS algorithms for clustering and feature selection tasks. This may be considered as a guide for other researchers in this domain to better understand the emergent trends and as well identify the issues that need to be resolved.
- Present quantitative analyses on state-of-the-arts hybridization of the SOS algorithms and K-means clustering methods that address the clustering issues, including the datasets and various performance evaluation metrics that can be utilized to evaluate the SOS algorithms performances.
- Present detailed explanations of the challenges and open issues in the field, and identify further issues that require urgent attention in clustering and feature selection problems.

The remainder of the study is outlined as follows: Section 2 discusses the study's context, including an explanation of the SOS algorithms and several applications of the SOS algorithms; Section 3 outlines the SLR method; Section 4 discusses the SLR results; and finally, Sections 5 and 6 describe the work limitations and conclusion, respectively.

2 Related work

Several works on the SOS algorithm have been published recently, and researchers have become increasingly interested in this field since its inception [18]. However, there were limited studies that provided a rigorous SLR to examine the implementation of the SOS algorithm in diverse scenarios. Specifically, Ezugwu and Prayogo [31] proposed an overview of SOS algorithms from its inception that included various applications using different variants and hybrid implementations. In addition, they recommended the algorithm be applied in extended applications as individual or hybrid. The exponential growth in research publications respective to the SOS algorithm acts as a significant signal of rapid expansion.

Subsequent work by Dokeroglu et al. [32] presented a review of metaheuristic algorithms based on a decade of application compared to classical algorithms. In this context, the metaheuristic algorithm was justified as efficient, prominently cited, interactive mechanisms between individuals, evolutionary operators, stagnation prevention methods, and parameter tuning/handling concepts. The study discusses significant challenges associated with metaheuristics and future directions. In addition, Abdullahi et al. [28] provided an in-depth examination of SOS advancements and applications focused on several existing symbiotic organism search methods. They comprehensively investigated relevant SOS development and its application in dealing with optimization challenges in various domains. Hybridization, discrete optimization, and limited and multiobjective optimization are some SOS algorithm study types.

A similar work by Gharehchopogh et al. [21] investigated the initial concept of symbiosis and SOS algorithms in parallel with its formulation and implementation. They illustrated their findings as a flowchart and pseudocode for each of the three processes involved in the algorithm. In addition, they discussed the variations of algorithms available. Subsequently, Hemeida et al. [33] reviewed basic theories and algorithms for optimization based on nature-inspired algorithms. Notably, they discuss variations of nature-inspired metaheuristic algorithms in terms of their concepts and mechanisms involved in an attempt to search for compatibility in real-world applications. They mainly focused on neural networks and feed-forward neural networks. Recently, Darvishpoor et al. [34] reviewed the application of heuristic and metaheuristic optimization algorithms in drones.

In addition to the SOS algorithm, there are many other machine learning algorithms that can be used for clustering and predictive problems. Some of the popular ones include K-means [35], fuzzy clustering algorithm [36], density-based clustering algorithm (DBSCAN) [37], hierarchical clustering [38], and Gaussian mixture models [39]. However, it was found that the SOS algorithm has several benefits over other clustering algorithms in different aspects. One of the benefits is that in all the reviewed articles, SOS algorithms have demonstrated improved performances and speedy convergences compared to other algorithms despite their simplicity and lesser nature of parameters [40]. Furthermore, unlike some clustering methods such as DBSCAN and K-means, the SOS algorithm can identify clusters of varying shapes and sizes, which do not require any tuning parameters to train [40].

Due to their robust performance and faster convergence, various studies employed the SOS algorithms in different applications to address specific problems. However, this research aims to provide a comprehensive survey of SOS algorithms related to clustering and classification processes. This is expected to fill the research gap and facilitate researchers with the necessary knowledge regarding SOS algorithms and their implementation using supervised and unsupervised feature selection problems. In addition, this review provided a microscopic perspective based on existing literature by dissecting the benefits and limitations of the current SOS methods regarding the feature selection and clustering problems. This will highlight current challenges and provide future recommendations.

3 Theoretical background

This section will discuss SOS algorithm variations and their application to clustering and feature selection.

3.1 Overview of the SOS algorithms

The SOS algorithms were pioneered by Cheng and Prayogo [18] and were inspired by the interactive performance of organisms to struggle for survival, known as symbiosis in biology. In order to find global optimal solutions to a particular objective function, this simple and efficient method guides a population of candidate solutions in iterative searches for possible optimum regions. It has, however, undergone various modifications that have strengthened and improved its adaptability to other problem spaces. This was initially built to handle the optimization issues in a continuous solution space. Since its evolution and adaptation, it has been more efficient in dealing with various problem spaces.

The major symbiotic relationship types found among organisms include mutualism, commensalism, and parasitism. Mutualism is a tripartite classification of interdependent relationships between two living creatures that mutually benefit from one another. One illustrative instance of such a phenomenon can be observed in the symbiotic relationship between Oxpecker birds and Rhinoceros. The Oxpecker species sustains itself by consuming insects and parasites on the Rhinoceros, an effective form of biological pest management. A commensalism symbiosis is a relationship in which one organism benefits completely from the other, while the other is unaffected by the interaction taking place instantaneously. An example of commensalism is the relationship between cattle egrets and insects that are disrupted by the cattle searching for food is an illustration of commensalism. However, when one creature derives benefits from another organism at the expense of the latter, this type of connection is known as parasitic symbiosis. Mosquitoes depend on human blood as a source of egg production. Figure 1 depicts the three types of symbiotic interactions that motivated the SOS algorithm.

Figure 1: Three types of symbiosis in the ecosystem [31].

The standard SOS algorithm is presented as Algorithm 1. It iteratively performs the steps beginning with the ecosystem initialization that can be denoted as $X = \{X_1, X_2, X_3, \ldots, X_{\text{ecosize}}\}$. The organisms' population size can be set from 25 (moderate), 50 (average), and 100 (large) [41]. Subsequently, the algorithm computes and compares respective objective function values to generate new organisms' positions with the best objective value selected as X_{best}. The algorithmic procedure involves iterative updating of the current optimal solution until the organism with the globally optimal answer is identified. In addition, the algorithm proceeds to assess alternative solution search spaces by engaging in both exploration and exploitation. Conversely, extending the execution period may also increase the computing expense of the algorithm. The subsequent sections, namely sections 2.2.1–2.2.3, comprehensively explain the steps involved in the SOS algorithm [31].

Algorithm 1 SOS algorithm

Input: Set ***ecosize***, create a population of organisms $\boldsymbol{X_i}$ $\boldsymbol{i = 1, 2, 3, \ldots,}$ ***ecosize***, initialize $\boldsymbol{X_i}$, Set stopping criteria.
Output: Optimal solution.
Identify the best organism $\boldsymbol{X_{\text{best}}}$
while the stopping criterion is not met do
  for $\boldsymbol{i = 1}$ to ***ecosize*** do
    Mutualism Phase
    $\boldsymbol{MV = \frac{X_i + X_j}{2}} \quad (\boldsymbol{j \neq i})$
    $\boldsymbol{B_{F1} = round(1 + r_1(0, 1))}$
    $\boldsymbol{B_{F2} = round(1 + r_2(0, 1))}$
    $\boldsymbol{X_i^* = X_i + r_1(0, 1)^*(X_{\text{best}} \;\; MV^*B_{F1})}$
    $\boldsymbol{X_j^* = X_j + r_2(0, 1)^*(X_{\text{best}} \;\; MV^*B_{F2})}$
    Evaluate $\boldsymbol{X_i^*}$
    if $\boldsymbol{X_i^*}$ then is better than $\boldsymbol{X_i}$
    $\boldsymbol{X_i \leftarrow X_i^*}$
    end if
    Evaluate $\boldsymbol{X_j^*}$
    if $\boldsymbol{X_j^*}$ then is better than $\boldsymbol{X_j}$
    $\boldsymbol{X_j \leftarrow X_j^*}$
    end if
    Commensalism Phase
    $\boldsymbol{X_i^* = X_i + r(1, 1)^*(X_{\text{best}} \;\; X_j)} \quad (\boldsymbol{j \neq i})$
    Evaluate $\boldsymbol{X_i^*}$
    if $\boldsymbol{X_i^*}$ then is better than $\boldsymbol{X_i}$
    $\boldsymbol{X_i \leftarrow X_i^*}$
    end if
    Parasitism Phase
    Create ***parasite_vector***
    Evaluate ***parasite_vector***
    if ***parasite_vector*** then is better than $\boldsymbol{X_j}$
    $\boldsymbol{X_j \leftarrow}$ ***parasite_vector***
    end if
    Identify the best organism $\boldsymbol{X_{\text{best}}}$
  end for
end while

3.1.1 Mutualism phase

In this phase, one organism X_i engages in interactions with another organism X_j, at random. It is important to note that X_i and X_j are distinct entities, where $X_i \neq X_j$. These interactions are established to form a mutually beneficial connection between the two organisms. The correlation between X_i and X_j aims to enhance the reciprocal survival rate of the two creatures within the environment. The candidate solutions $X_{i\text{new}}$ and $X_{j\text{new}}$ are derived using equations (1) and (2), respectively.

$$X_{i\text{new}} = X_i + \text{rand}(0,1) \times (X_{\text{best}} - X_{\text{mutual}}(B_{F1})), \tag{1}$$

$$X_{j\text{new}} = X_j + \text{rand}(0,1) \times (X_{\text{best}} - X_{\text{mutual}}(B_{F2})), \tag{2}$$

where X_{mutual} is represented in equation (3).

$$X_{\text{mutual}} = \frac{X_i + X_j}{2}, \tag{3}$$

$$B_{F1} = (1 + \text{round}(\text{rand}(0,1)), |\text{rand} \in [0,1]), \tag{4}$$

$$B_{F2} = (1 + \text{round}(\text{rand}(0,1)), |\text{rand} \in [0,1]). \tag{5}$$

The rand(0,1) function generates a vector of random numbers that follow a uniform distribution within the range of 0–1. The organism that demonstrates the highest objective or fitness function value to its level of adaptability within the ecosystem is denoted as X_{best} [42]. On the other hand, X_{mutual} implies the mutualistic characteristics demonstrated between the two organisms to advance the benefit of their survival. The values of the benefit factors B_{F1} and B_{F2} are selected by a random selection process, as specified by equations (4) and (5). These parameters indicate the degree of the advantage derived from the interaction with each organism. Then, the newly calculated fitness function value is represented as $f(X_{i\text{new}})$ and $f(X_{j\text{new}})$. They demonstrate superior performance compared to the earlier fitness functions, $f(X_i)$ and $f(X_j)$ [43]. Hence, the pair of equations (1) and (2) can be further transformed subsequently as follows:

$$X_{i\text{new}} = X_i + \text{rand}(0,1) \times (X_{\text{best}} - X_{\text{mutual}}(B_{F1})) \text{ if } f(X_{i\text{new}}) > f(X_i), \tag{6}$$

$$X_{j\text{new}} = X_j + \text{rand}(0,1) \times (X_{\text{best}} - X_{\text{mutual}}(B_{F2})) \text{ if } f(X_{j\text{new}}) > f(X_j). \tag{7}$$

3.1.2 Commensalism phase

In this phase, the organisms X_j are randomly chosen from the environment to interrelate with other organisms X_i. However, organisms X_i actively seek to maximize the advantages gained from the link with the organisms X_j, but X_j remain unaffected by this interaction, neither benefiting nor experiencing any negative consequences. In this form of interaction, the organism X_i occupies a favorable position relative to X_j, whereas the organism X_j does not experience any detrimental effects [44]. The emergence of a novel solution resulting from this symbiotic interaction is expressed in equation (8) [44].

$$X_{i\text{new}} = X_i + \text{rand}(-1,1) \times (X_{\text{best}} - X_j), \quad \text{if } f(X_{i\text{new}}) > f(X_i). \tag{8}$$

The $\text{rand}(-1,1)$ function generates a vector with random numbers that are uniformly distributed throughout the range of −1 to 1. Organism $X_{i\text{new}}$ can replace organism X_i if it exhibits a higher fitness value. In this context, the term $(X_{\text{best}} - X_j)$ represents the benefits offered by the organisms X_j to aid organisms X_i in maximizing their survival within the ecosystem population, with X_{best} referring to the most recent or updated organism.

3.1.3 Parasitism phase

A symbiotic relationship between two species, wherein one organism obtains exclusive benefits while causing harm to the other creature. An instance of parasitism can be observed in the symbiotic relationship involving three organisms: the Plasmodium parasite, the Anopheles mosquito, and the human host. In this form of symbiotic association, the human host has detrimental effects, yet the Anopheles mosquito, serving as the vector for the parasite, remains unaffected. Meanwhile, the Plasmodium parasite undergoes reproductive processes within the human organism [43]. Hence, to emulate the parasitic behaviors described above, organism X_i assume a role similar to the Anopheles mosquito by generating an artificial vector X_{parasite} within the solution search space. This is achieved by adjusting the randomly chosen dimension of organism X_i through a process of refinement [43]. Following this, a random selection is made from the ecosystem to identify the organism X_j, which then acts as the host for X_{parasite}. Subsequently, the X_{parasite} will endeavor to supplant X_j inside the ecosystem. If X_{parasite} demonstrates a superior level of fitness compared to X_j, then X_j will be substituted with X_{parasite}. According to Ezugwu et al. [44], X_j acquires immunity to X_{parasite}, leading to the eventual extinction of X_{parasite} within the ecosystem. This can be expressed as follows:

$$X_{\text{parasite}} = \text{rand}(0, 1) \times (\text{UB} - \text{LB}) + \text{LB}, \tag{9}$$

where LB (lower bound) and UB (upper bound) are the boundary limits that should be addressed.

It is important to note that modifications to either the mutualism phase, the commensalism phase, or both have accomplished most enhancements to the conventional SOS algorithm. Introducing an additional phase to the existing three phases is a phenomenon that occurs only in exceptionally unusual instances. This study comprehensively analyses several current advancements and hybridization techniques employed in SOS algorithms as discussed in the literature.

3.2 Variations of the SOS algorithm

This section presents variations of the recent advances of the SOS algorithms, which have been classified into three main categories: modified or enhanced SOS, hybrid SOS (HSOS), and multiobjective SOS. Each of these concepts will be elaborated further in the subsequent subsections.

3.2.1 Modified SOS (MSOS) algorithm

The SOS algorithm has undergone multiple revisions since its inception to offer an optimal and efficient solution for various optimization challenges. One of the examples is the MSOS algorithm, which was presented to improve the rate of convergence and the SOS algorithm performance [45]. An updated work by Chakraborty et al. [46] presented a novel variant of the SOS algorithm, known as nwSOS, to address higher dimensional optimization challenges in the context of segmenting COVID-19 chest X-ray images. Furthermore, Rodrigues et al. [47] introduced a novel approach known as the MSOS algorithm to address the challenges associated with workflow-scheduling problems. The suggested method involves selecting three phases of the SOS algorithm, where there is no pre-established symbiotic interaction among the population. Once solutions have been identified, a particular process is employed to allocate each solution to corresponding symbiotic relationship. Therefore, 20 samples of work scheduling problems have been used for evaluating the performance of the method. Subsequently, a comparative analysis was carried out between the proposed method and the original SOS method, revealing the superiority of the proposed model over a number of SOS versions. Abdullahi et al. [28] developed a novel approach called modified symbiotic organisms search with inertia weight to update the phases of SOS, aiming to enhance the solution quality.

Additionally, Secui [48] utilized a chaotic sequence created by the logistic map to boost the explorative abilities of SOS methods. A modification aimed at enhancing the advantageous aspects of SOS algorithms was discussed by Tejani et al. [49] referred to as symbiotic organisms search and artificial fish and bees algorithms. The selection of components was chosen adaptively in consideration of the organism optimization.

The work by Nama et al. [50] presents yet another attempt to enhance the standard SOS algorithm known as the improved symbiotic organism search algorithm by integrating random weighted reflective parameters and predations. Including a fourth symbiotic phase in the predation update category in the fundamental SOS framework was motivated by observing creatures within the ecosystem that frequently employ predation as a strategic approach. The author described that predation is a biological interaction that has a resemblance to the parasitism relationship, wherein a predator, actively seeking sustenance, consumes its prey, the organism being targeted. The commonality observed in both systems is a relationship in which one organism experiences harm while the other obtains advantages. However, it is important to note that the distinction between predation and parasitism is that the predator organism often kills and consumes its prey. However, not all parasites result in the death of their hosts.

Furthermore, Do et al. [51] designed a novel combination of deep ANN and an enhanced SOS algorithm. This approach was applied to the problem of material distribution optimization in functionally graded plates as a computational tool for the analysis of the ultimate component. Indeed, using deep ANNs enables the direct prediction of solutions through optimal mapping. In addition, a refined SOS (Sequential Optimisation Strategy) method has been employed to address two distinct challenges related to optimization buckling and unrestricted vibrations, each with different volume constraints.

3.2.2 Multiobjective SOS algorithms

The multiobjective optimization problem is commonly categorized as multiple criteria decision-making, where multiple objective functions are simultaneously optimized [52,53]. Major advances have been made in the SOS algorithms, for instance, the SOS method that was formulated to handle multiobjective problems (termed as multi-objective symbiotic organisms search [MOSOS]) [52]. This method is integrated with adaptive penalty functions to enable the effective handling of equality and inequality constraints associated with various problem types. The proposed approach demonstrates superior performance compared to other existing multi-objective optimization algorithms, including multi-objective cuckoo search with binary optimization, multi-objective particle swarm optimization, and non-dominated sorting genetic algorithm II (NSGA-II), as well as two gradient-based multiobjective algorithms, multi-objective genetic algorithm with elitism and multi-objective genetic programming. Ayala et al. [54] introduced an enhanced novel MOSOS approach that incorporates both nondominance and crowding distance criteria referred to as improved multi-objective symbiotic organisms search. It includes a normal (Gaussian) probability distribution function, which makes it perform better compared to other methods. In another study, Baysal et al. [23] investigated a multiobjective problem to define the tradeoff between the number of selected features and classification accuracy for the BCI system. This is called the nondominated sorting multiobjective symbiotic organism search (NSMOSOS) algorithm.

3.2.3 HSOS algorithms

As mentioned earlier, the fundamental SOS algorithm and its diverse adaptations have been used to effectively handle many problems relating to continuous and discrete optimizations. However, certain instances exist where these sets of algorithms have exhibited suboptimal performance or failed to achieve the desired answers. Consequently, researchers frequently find integrating multiple algorithms necessary to tackle this obstacle effectively [55,56]. Research has also indicated that hybrid algorithms were more likely to exhibit robustness and yield better performance than traditional techniques. In a recent study, Ikotun and Ezugwu [57] introduced a novel hybrid clustering approach of SOS and K-means. The findings derived from the comprehensive computational analysis demonstrate that the hybrid SOSK-means algorithm has enhanced efficacy in addressing automatic clustering tasks.

Furthermore, the SOS technique has been employed in hybrid image fusion techniques to achieve the highest possible level of image quality in the resulting fused images [58]. In another study, Rajah and Ezugwu [55] worked on various HSOS algorithms for automated clustering. The proposed approach was evaluated based on the quality of their solutions using the Davies–Bouldin Clustering Validity Index. In their recent study, Cheng et al. [25] introduced a novel approach named SOS-NN-LSTM to predict cash flow during the implementation of construction projects. Nama et al. [50] introduced a unique HSOS approach that combines the strengths of SOS and simple quadratic interpolation (SQI) to achieve a balanced implementation of the new algorithm in addressing the physical world and higher dimensional optimization challenges. Including the SQI in the new method has increased complexity compared to the standard SOS (second-order statistics) approach.

Wang et al. [59] have introduced a HSOS approach that integrates the global optimization capabilities of SOS with the local optimization capabilities of ACO to optimize assembly sequences. The computational complexity of the resultant method was heightened because of the integration of two optimization procedures. In their study, Ezugwu et al. [60] introduced two HSOS algorithms that integrate simulated annealing and genetic algorithm (GA) techniques. These algorithms were named SOSSA and self-organizing systems genetic algorithm [41].

In addition, Cheng et al. [61] worked on a novel artificial intelligence system known as self-organizing systems least squares support vector regression, which combines elements of artificial intelligence with the least squares support vector regression (SVR) to estimate the permanent deformation potential of asphalt pavement mixtures effectively. Zhang et al. [62] introduced an innovative data classification method that combines machine learning techniques with the SOS algorithm. In their recent study, Noureddine et al. [6] introduced a novel clustering technique known as self-organizing systems grey wolf optimization with tabu search based on the SOS approach. This hybrid approach integrates the grey wolf optimizer (GWO) and the Tabu Search algorithm as proposed by Aljarah et al. [63].

Recently, Goldanloo and Gharehchopogh [64] introduced two HSOS algorithms, namely the IOFA-SOS variant to enhance both the exploration and exploitation capabilities of the initial technique. This method aims to leverage the IOFA technique within the SOS algorithm, where the entire population is utilized to effectively identify solutions during the initial phases of implementation. As the algorithm progresses through each iteration, the number of solutions impacted within the population gradually decreases. The study involved conducting experiments on a set of 24 typical benchmark functions. The initial proposed approach demonstrated superior performance in smaller and medium dimensions while displaying moderately satisfactory performance in higher dimensions. On the other hand, the second proposed technique shows outstanding performance in increasing dimensions. The experimental findings demonstrated that the proposed methods yielded high-quality solutions in a reasonable computing time (Table 1).

4 Systematic literature review (SLR) methodology

The standard criteria used for this SLR are preferred reporting items for systematic reviews and meta-analyses (PRISMA) [72]. As depicted in Figure 2, the PRISMA framework offers a systematic and standardized approach for the identification, selection, and critical evaluation of relevant existing studies. In addition, it guides the process of selecting, identifying, and assessing research projects. A comprehensive description of the review procedure is provided in the subsequent subsection.

4.1 Data source and search strategies

This study uses eight different bibliographic digital databases for searching processes, which includes Scopus, Google Scholar, ScienceDirect, SpringerLink, IEEE Xplore Library, ACM Digital Library, PubMed, and Web of Science. To get the most up-to-date and thorough review, the time frame of 2014–2023 was considered. As depicted in Figure 3, the search string encompasses various combinations, including “Symbiotic Organisms Search,” “Symbiotic Organisms Search Algorithm” combined with “Clustering,” “Symbiotic Organisms Search Algorithm” combined with “Classification,” “Symbiotic Organism Search” combined with “K-Means Clustering,” “Symbiotic Organisms Search” combined with “Feature Selection,” and “Hybrid SOS-Based” combined with “Clustering” and “Feature Selection.”

Table 1: The benefits and limitations of the recent general applications of various clustering-based SOS methods

SOS variant	Problem description	Problem domain	Benefits	Limitations
Modified [47]	This study introduces a novel approach, the MSOS algorithm, which aims to tackle the challenges encountered in workflow-scheduling applications	Scheduling problem	The effectiveness of search (exploitation) and exploration accuracy can be improved by integrating an adaptive benefit factor and a modified parasitism vector	The algorithm was limited regarding neighborhood optima and a weaker convergence zone, which impeded the discovery of improved solutions
Modified [51]	The researchers introduced a novel integration of the deep ANN with an enhanced SOS method for addressing optimization problems	Construction engineering	The aim is to enhance computing efficiency by introducing a new methodology integrating deep neural network with an MSOS algorithm. The goal is to obtain the best possible material distribution	The architectural design of neural networks differs significantly from microprocessors, requiring emulation
Five new hybrid methods: SOS, SOSTLBO, self-organizing systems firefly algorithm (SOSFA), SOSPSO, and Self-organizing systems differential evolution (SOSDE) [55]	The authors discuss the challenges of automatic clustering analysis	Machine learning	Datasets are autonomously partitioned without prior knowledge of the number of clusters	Further optimization of the parameters is required to improve the robustness of the SOS algorithm in addressing intricate clustering situations
Hybrid [65]	A technique known as SOSK-means was developed to address the task of automatic grouping	Machine learning	The SOSK-means algorithm integrates the local exploitation capacity of the regular K-means algorithm with modified parameters and global exploration	The computational complexity of an algorithm lowers performance
Hybrid [66]	This study addresses the automatic clustering problem using a hybrid of K-means and SOS algorithms	Machine learning	The hybrid approach improved the effectiveness of addressing cluster problems	The algorithm is heavily reliant on the quality of the original solution, and there is a risk of the algorithm getting stuck in local optima
Hybrid [67]	The CSOS algorithm was created to facilitate the automatic production and merging of clusters, hence implementing a mechanism that enhances the efficiency of the process	—	The hybrid strategy employs the SOS as a global search metaheuristic to produce the optimal starting cluster centroids for the K-means algorithm	The algorithm's performance is influenced by difficulties such as local optima, premature convergence, and computational inefficiency
Modified [46]	The newly enhanced SOS method, specifically nwSOS, can address optimization challenges in more excellent dimensions	Machine learning; segmenting COVID-19 chest X-ray images	The method employed in this solution involves the utilization of a nonlinear methodology to determine the benefit factors, hence enabling the resolution of optimization issues in higher dimensions	The balance between exploitation and exploration is crucial for maximum performance
Hybrid [68]	Selection and parameter optimization for the classification of submerged targets using the SVM-SOS method	Machine learning; classification of underwater	This study focuses on utilizing the SOS method to optimize the selection of parameters for SVM, including the kernel type and kernel parameter	The SOS algorithm covers the parameter selection process and automatic kernel parameter selection to accomplish automatic SVM implementation

(Continued)

Table 1: *Continued*

SOS variant	Problem description	Problem domain	Benefits	Limitations
Multiobjective symbiotic [23]	NSMOSOS algorithm used to generate the optimal feature subset in BCI	Machine learning; motor imagery segmentation	The ability to generate the most effective subset of features in BCI	The uniform distribution of weights cannot ensure the uniform distribution of Pareto optimum solutions, despite this method's simplicity and ease of use
Five discrete SOS (DSOS) that are combinations of modified and hybrid methods [69]	This study aims to enhance the classification accuracy by simultaneously optimizing the feature subset and neighborhood size of the k-nearest neighbor model	Machine learning	Optimize the feature subset and neighborhood size of the k-nearest neighbor model simultaneously to increase classification accuracy	The current method fails to filter out redundant features and disregards defining the desired size of the feature subset, leading to a decrease in classification accuracy
Hybrid [70]	A novel approach is presented an accurate load forecasting based on a hybrid of the SVR technique and the SOSO method	Power system and plants	The SOS optimization kernel was proposed to identify the ideal parameters for SVR	The optimization techniques should be employed to optimize the selection of the most-relevant features while minimizing redundancy
Hybrid [71]	The SOS algorithm was used to optimize SVM for spyware classification	System security	Optimization of SVM using SOS algorithm for spyware feature selection	The generalizability of the technique to other datasets or types of spyware remains unclear due to the limited evaluation of model performance using only one dataset

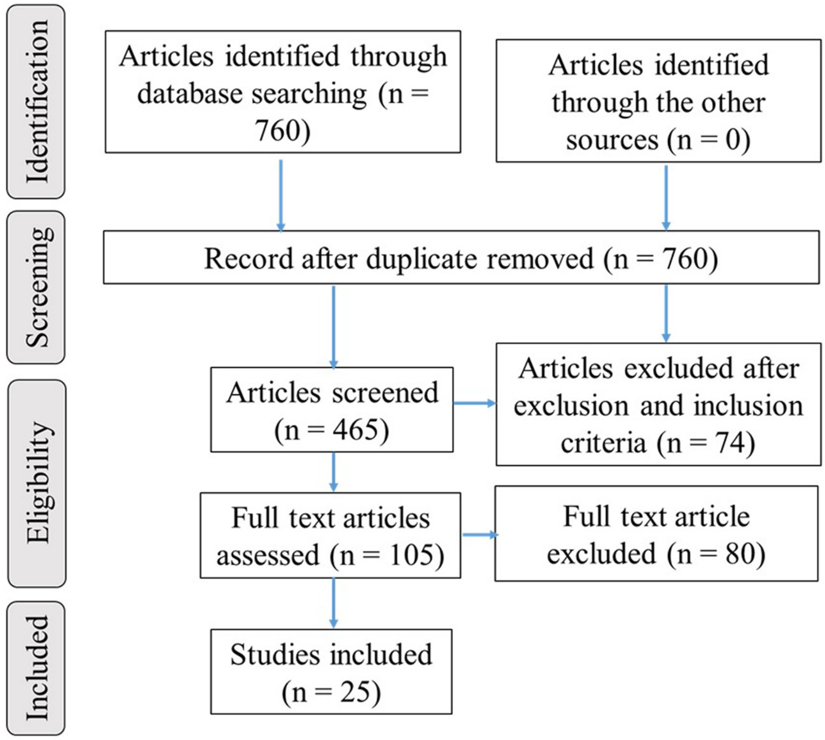

Figure 2: The SLR PRISMA method.

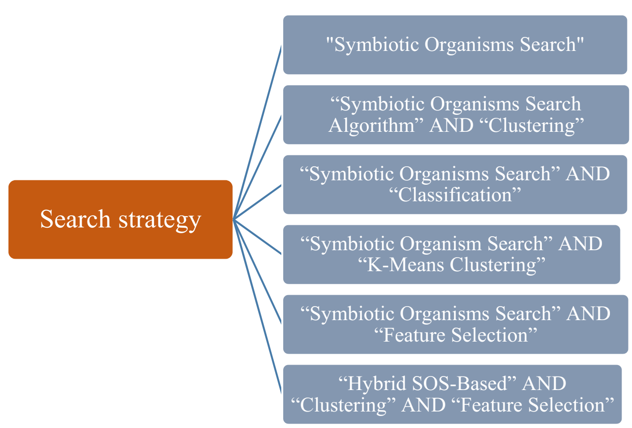

Figure 3: Search strategy.

4.2 Selection criteria

In conjunction with the automated search query, a manual search is performed to ensure a comprehensive identification of relevant articles and acquire a total of 760 studies. During the screening phase, 465 published articles were selected after eliminating duplicate and irrelevant studies. Upon reviewing the abstract, introduction, and title, a total of 295 published articles were excluded. The removal criteria were subsequently applied to the remaining 170 articles, resulting in the exclusion of 105 articles. A total of 65 research publications were later advanced to the subsequent step. After conducting a comprehensive assessment of all the articles during the eligibility phase, a total of 37 research publications were excluded. A total of 25 scholarly research publications were selected to conduct a comprehensive review of prediction techniques that utilize clustering techniques and symbiotic organisms search algorithms.

We included only articles that specifically addressed the topics of unsupervised clustering and feature selection. Through a rigorous selection process, we limited the number of publications to 25 articles, focusing exclusively on journal and conference articles. In addition to other SOS-based approaches for addressing clustering and feature selection challenges, our focus was directed on articles that specifically discuss hybrid methods that tackle both clustering and feature selection tasks. All these steps taken were based on Table 2 guidelines.

Table 2: Criteria for inclusion and exclusion

Inclusion	Exclusion
Studies that solely included experimental findings	Studies that excluded empirical findings were omitted
Studies that have been published in the year 2014–2023	Studies that were published earlier than 2014 were omitted
Studies solely related to SOS-based method	Studies related to other than metaheuristic algorithms were omitted
Articles discussing related unsupervised clustering and feature selection	Articles discussing other techniques for clustering and feature selection
Studies solely reported in English language	Studies not reported in English language
Only considered journals and conferences	Documents like books, thesis, and magazines are not considered

4.3 Quality assessment

The data analysis and the quality assessment of the evidence in an SLR are of equal importance. The presence of biases stemming from the chosen research methods has the potential impact on the results of a study that is inadequately executed, thus necessitating careful interpretation. Studies of this nature must either be explicitly rejected or, at the very least, clearly highlighted within SLR and carefully select appropriate criteria for evaluating the quality assessment and identifying any intrinsic biases present in any distinct study.

Standard quality checklist questions (SCQ) developed by Kitchenham and Charters [73] are utilized to guarantee the kind of the chosen publications for this study. For this reason, the studies that answered "yes" to at least seven questions that were selected, following the method proposed by Genc-Nayebi and Abran [74]. The quality assessment and data extraction will be conducted simultaneously to guarantee that the results make a substantial contribution to the review [30] (Table 3).

Table 3: Statistical Quality Control Questionnaire (SQCQ) [73]

No.	Quality questions
SCQ1	Is the report clear and coherent?
SCQ2	Is the aim of the study clearly specified?
SCQ3	Is data collection process designated properly?
SCQ4	Is the contexts of diversity been well investigated?
SCQ5	Are the research findings reliable?
SCQ6	Does the link between the data, interpretation, and conclusion exist?
SCQ7	Is there a clear method and process for experiments?
SCQ8	Is the research process adequately documented?
SCQ9	Are they important, if credible?
SCQ10	Could the findings of this research be replicated?

4.4 Data extraction and synthesis

Initially, we documented the necessary details, such as the article's title, year of publication, a list of authors, and publisher. Subsequently, we incorporated more data to facilitate the execution of the SLR model. This included the inclusion of SOS-based methods for unsupervised clustering and feature selection, hybrid methodologies, accuracy values, and evaluation metrics. The phase of data synthesis involves a thorough examination of the relevant and comparative findings obtained via the process of data extraction. These findings can then be presented to address the research issues at hand. Upon completion of data collection, the gathered data was subjected to analysis and subsequently visualized using a range of data visualization tools and techniques, including histograms, pie charts, and other relevant methods.

4.5 RQs

The selection of RQs holds significant importance in outlining the main goals and anticipated results of a study to fulfill the primary objective of our SLR. Hence, the RQs listed such the:

RQ1: What are the various SOS methods that have been adapted as clustering methods in the clustering application problems?

RQ2: What are the various SOS methods that have been adapted and utilized for feature selection problems?

RQ3: What are the reported hybridizations of the SOS algorithm with the K-means clustering algorithm that handled clustering problems?

RQ4: What is the rate of publication of the various SOS variants algorithms for clustering and feature selection?

RQ5: What are the challenges and prospects of SOS-based clustering algorithms?

5 Results

The findings described in Section 4 are provided. This section is subdivided into four different subsections: in Section 5.1, the analysis of the searched articles on various SOS methods adapted as clustering methods in the clustering application problems is shown; in Section 5.2, the various SOS methods adapted and utilized for feature selection problems are discussed; in Section 5.3, the reported hybridizations of the SOS algorithms with the K-means clustering algorithm that handled clustering problems are discussed; and in Section 5.4, the rate of publication of the various SOS variants algorithms for clustering and feature selection is discussed.

5.1 RQ1: What are the various SOS methods that have been adapted as clustering methods in the clustering application problems?

According to [23,66,69], the use of the SOS algorithm has been extensively utilized for classification, clustering, and unsupervised feature selection. Clustering is an unsupervised data analysis methodology employed to locate groups of items that exhibit homogeneity, determined by the values of their attributes. There are two main groups of clustering: (1) hierarchical clustering and (2) partitional clustering methods. The hierarchical clustering approach involves the iterative grouping of data objects in a hierarchical manner. Meanwhile, the partitional clustering technique is employed to produce a solitary partition of a dataset, to identify the inherent groupings within the data. This technique does not involve any hierarchical structure and relies on the utilization of a designated objective function. An example of this is K-means.

Furthermore, Yang and Sutrisno [67] suggested an automatic K-means clustering-based SOS (CSOS) to address the issues with automatic data clustering. In essence, the CSOS algorithm conducts a hybrid search approach by combining local and global search strategies inside a sub-ecosystem context that is established using the automated *k*-means clustering technique. The local search covers two distinct phases, namely mutualism and commensalism. Meanwhile, the global search in CSOS is characterized by the parasitism phase, when only the best solution inside a cluster can communicate with the best solutions in other clusters. Furthermore, Zhang et al. [62] introduced a data categorization methodology that combines a regularized extreme learning machine and an SOS algorithm. The newly developed method is referred to as SOS-RELM, which consists of two distinct phases of SVM, least squares SVMs, and backpropagation. Therefore, the SOS method is an optimization approach that is both efficient and effective. It optimizes the input weights, invisible biases, and regularization parameters in the second phase. In a separate work conducted by Liao and Kuo [69], they introduced DSOS algorithms for the optimization of feature subsets and neighborhood size.

In addition, Zhou et al. [66] introduced an approach for automatic data clustering utilizing the SOS algorithm. The k-means clustering algorithm's heavy dependence on the starting solution and ease of trapping in a local optimum have been discussed in this article. The SOS framework employed an automated k-means clustering technique to generate clusters. Within each cluster, the most optimal solutions engage in communication, integrating both local and global search strategies [65]. However, it leads to an increase in its computational cost.

Similarly, Acharya and Mishra [75] developed a multiagent SOS (MASOS) by integrating a multiagent system and self-adaptive benefit factors into the SOS algorithm. In this approach, each organism functions as an agent that engages in local interactions to identify the optimal solution. The execution of the SOS algorithm involved three distinct steps when an agent would select another agent from its nearby neighborhood. Regardless, the proposed work also suffered high computational complexity. Table 4 lists the adopted clustering approaches discussed.

Table 4: The clustering approach adopted

Authors	Adopted clustering approach
[67]	The initial number of clusters is set to be half of the ecosize, which is formed as sub-ecosystems. This is then subjected to optimization using CSOS
[55]	The optimization of the Bayesian information criterion (BIC) and Bayesian deviance information (BDI)
[65]	The Compact Separated Index and the Davies–Boulding Index are presented as optimization issues where the cuckoo virus infection serves as the objective function and must be minimized
[66]	The cluster is randomly initialized within the ecosystem to which the SOS algorithm is applied to optimize the clustering problem
[46]	To efficiently explore and exploit the search region, the benefit factors are determined using a nonlinear approach, and their weights are used
[68]	The determination of the initial number of clusters is based on the ecosize, which represents the sub-ecosystems formed by the self-organizing system (SOS) and subsequently optimized
[23]	To locate the ideal feature subset in the BCI data, an initial population is generated in the same way as using the standard SOS algorithm
[69]	The utilization of DSOS is initially observed in the optimization of cluster quantity for the hybrid approach between DSOS and discrete particle swarm optimization (DPSO), referred to as DSOSPSO
[70]	SVR is considered significant in a forecasting algorithm. The SOSO algorithm is incorporated in the SVR for optimization
[71]	SVM is optimized with SOS for the classification of spyware

5.2 RQ2: What are the various SOS methods that have been adapted and utilized for feature selection problems?

Feature selection aims to discern the pertinent characteristics and eliminate the extraneous features from the dataset to achieve advantages such as reduction in data dimensionality, improvement in the performance of classification algorithms, and facilitation of the learning process [23]. Nevertheless, it is notable that feature extraction methods could be categorized into two, namely filter-based methods and coating-based/wrapper methods [76]. Filter-based approaches are algorithmically agnostic and exhibit reasonably high computational efficiency. On the other hand, coating-based approaches employ classification algorithms as assessment criteria for optimal solutions.

It has been argued by researchers that coating-based approaches outperform filter-based methods due to their use of classification algorithms in the assessment criteria [77,78]. Based on this context, the authors [76] present three wrapper-based binary approaches using the SOS method to address feature selection problems. The first was binary SOS with S-shaped transfer functions (BSOSST), and the second was BSOSVT. These two coating-based methods are employed to develop the binary SOS (BSOS). In the third approach, enhanced elephant behavior cuckoo search optimization system has proven performance in terms of effectiveness and exploration capabilities.

In their study, Du et al. [79] introduced a novel approach known as the improved binary symbiotic organism search (IBSOS) that utilizes the wrapper method for addressing the issue of feature selection. Furthermore, the authors employed identical biological symbiosis tactics as those utilized in the continuous SOS method within the suggested IBSOS approach to maintain the balance between exploration and exploitation. Baysal et al. [23] presented a wrapper-based multiobjective algorithm, namely, the NSMOSOS algorithm, to produce the optimal feature subset in a BCI system. The investigation focused on evaluating the efficacy and resilience of the proposed algorithm in two datasets centered around motor imagery for feature selection, which achieved maximum accuracy results for both datasets used.

In their recent study, Mohammadzadeh and Gharehchopogh [26] introduced a novel feature selection method that utilizes the BSOS metaheuristic to enhance email spam detection using a feature selection technique to identify the most pertinent attributes from a vast array of input data. In a study conducted by Han et al. [80], it was proven that the BSOS algorithm demonstrates a propensity for identifying the minimal set of features across various datasets while concurrently attaining a notable level of accuracy in classification. Nonetheless, the BSOS approach exhibits limitations when dealing with datasets of low dimensionality and demonstrates reduced sensitivity in datasets with a high number of dimensions. In addition, Miao et al. [81] introduced an innovative approach for feature selection that utilizes the SOS method to enhance the precision and efficacy of sleep staging. This procedure involves categorizing sleep stages using physiological data.

Wrapper methods generally exhibit superior performance compared to filter methods. However, in the context of high-dimensional datasets like a microarray dataset, where the sample size is smaller than the dimension of the features, the wrapper method can become computationally burdensome. Therefore, researchers have devised hybrid approaches that integrate the usage of both filter and wrapper methods [69,82]. They experimented with a total of five novel DSOS algorithms [69]. These algorithms aimed to enhance the classification accuracy by simultaneously optimizing the feature subsets and neighborhood size of the k-nearest neighbor model. In this scenario, a two-step approach is utilized. Initially, a filter method is applied to eliminate a significant number

Table 5: List of the various SOS methods and feature selection approaches

S/N	SOS variant	Feature selection approach	References
1	Multiobjective SOS	Wrapper-based	[23]
2	Five DSOS that are combinations of modified and hybrid methods	Wrapper-based	[69]
3	Hybrid method	—	[70]
4	Hybrid method		[71]
5	Hybrid method	Wrapper/coating-based	[76]
6	Improved SOS	Wrapper-based	[79]
7	MSOS	Wrapper-based	[26]
8	MSOS	Wrapped-based	[80]
9	MSOS	Filter-based	[81]
10	Hybrid approach	Filter- and wrapper-based	[69]

of features. Subsequently, a wrapper method is implemented to choose the optimal subset from the remaining features. Table 5 summarizes various SOS and feature selection approaches.

5.3 RQ3: What are the reported hybridizations of the SOS algorithm with the K-means clustering method to address the clustering problems?

This section discusses the approaches that hybridize the SOS algorithm together with the K-means clustering for automatic data clustering. As shown in Table 6, the purpose of each hybridized method and the clustering method used in the related approaches are in the second and third columns, respectively. In contrast, the fourth, fifth, and sixth columns provide information on the datasets utilized to evaluate the algorithms, the approaches that were compared, and the criteria employed to assess their performance. Out of the 25 studies that were reviewed, it was found that only 12 K-means/SOS hybrid methods specifically focused on the matter of automatic data clustering.

Each of these has been further explained to ensure robust analysis, taking into consideration, the metrics used and the significance of the findings. For instance, Yang and Sutrisno [67] utilized a CSOS algorithm with K-means clustering method to build a hybridized method for automatic data clustering. The K-means approach is a partitioning clustering method that involves specifying the number of clusters to be created from a given data. This was aimed at overcoming the issues with the traditional clustering approach to address the clustering problems. The result shows that the hybrid algorithm improves the quality and performance of searching by combining the global and local searches from the data based on the successful execution numbers, and the mean computational time. Similarly, Rajah and Ezugwu [55] proposed five hybrid algorithms for automatic data clustering: SOS-based, SOSFA, SOSDE, SOSTLBO, and SOSPSO. These algorithms are inspired from the original SOS algorithm, which are evaluated using the average computational time. This approach overcomes the limitations of traditional clustering algorithms in handling automatic data clustering problems.

Furthermore, Ikotun and Ezugwu [65] developed a hybrid clustering method that integrates the search algorithm with K-means approach that utilizes the SOS algorithm as a global search method to generate the ultimate starting cluster centroids for the K-means method. Eleven datasets from the UCI machine learning repository and one artificial dataset are used to compare the performance of the proposed algorithm with the classical SOS, classical K-means, and other current hybrid clustering techniques. Compared to the traditional K-Means method, SOS-based clustering approach, and other hybrid clustering methods, the findings demonstrate that for solving automatic clustering, the hybrid SOSK-means performs better. In another study, Ikotun and Ezugwu [57] proposed an improved version of the SOSK-means hybrid algorithm for automatic data clustering. The proposed algorithm incorporates a random weighted reflection coefficient with a three-part mutualism phase to improve the efficiency of the hybrid method. A global threshold of the point-to-centroid distance distribution is utilized in this method to automatically determine the outliers and exclude them from the centroid update processes. This exclusion occurs when the new centroids are calculated in the K-means phase. Based on the research findings, the proposed hybrid method performs better than traditional K-means, SOS-based clustering, and other hybrid methods for automatic data clustering. Furthermore, to improve the accuracy of spyware classification and reduce the false positive rate (FPR), Gana et al. [71] proposed an improved SVM classifier using the grid-search optimization algorithm. This enhances the SVM classifier by optimizing its parameters and identifying the best features to select based on the SOS algorithm. Compared to a different clustering approach, the proposed method performed better when evaluated on spyware datasets, demonstrating increased accuracy and FPR.

In addition, Zare-Noghabi et al. [70] proposed a hybrid method for medium-term load forecasting. The proposed method combines the symbiotic organism search optimization (SOSO) and SVR techniques. SVR is the main forecasting algorithm in this method, and SOSO is integrated to improve SVR parameters. The effectiveness of this method compared to other methods is shown by evaluating it on the evolutionary united templar (EUNITE) competition dataset in terms of mean absolute percentage error. In comparison to alternative clustering methods, the proposed approach has demonstrated increased performance. Baysal et al. [23]

Table 6: The hybridization of K-means and SOS algorithms

SOS algorithm	Objective	Method for automatic clustering	Dataset used	Comparison to	Metrics used
CSOS + K-means [67]	A hybrid automatic clustering algorithm	Using half population size as the number of clusters	28 Benchmark function	GA, SOS, continuous ranked probability score, salp swarm algorithm with self-adaptive differential evolution (SaNSDE), GWO, and restart covariance matrix adaptation evolution strategy (rCMA-ES)	No. of successful execution, average computational time, and no. of evaluation
Five hybrid algorithms: SOS-based, SOSFA, SOSDE, SOSTLBO, SOSPSO [55]	Automatic clustering analysis problem	Optimization of BIC or the Davies–Boulding Index (BDI)		12 UCI machine learning repository datasets	Compare with variants of the proposed methods
SOSK-means [65]	A hybrid automatic clustering algorithm	Davies–Boulding index and the Compact Separated Index are presented as optimization problems with minimized objective functions		11 Datasets extracted from the UCI machine learning repository and one artificial dataset	SOSTLBO, SOSFA, SOSPSO, SOSDE, differential evolution (DE), differential chaotic particle swarm optimization, and grey cuckoo optimization with krill herd
Improved symbiotic organisms search with knowledge transfer (ISOSK)-means [57]	Automatic clustering problems	To calculate new centroids in the K-means phase for automatic outlier detection, global threshold of point-to-centroid distance distribution is used		24 Real-world datasets and 18 artificial datasets	DE, PSO, firefly algorithm, and invasive weed optimization
MSOS [47]	An MSOS algorithm used to address the clustering problems	The initial number of clusters is generated similar to the classic SOS algorithm		20 Benchmark data of the flow shop scheduling issue	Classical SOS
SVM-SOS [71]	Classification of spyware using (SOS) algorithm	SVM is optimized with SOS for the classification of spyware		Spyware datasets	SOS and SVM
SVR and SOSO [70]	A hybrid method for accurate load forecasting	SVR is used as the main component of the forecasting algorithm, embedded with SVR parameters for optimization		EUNITE competition dataset	Theta, Trigonometry smoothing, exponential smoothing, exponential smoothing state space model, autoregressive integrated moving average, and MLP neural network
Five hybrid algorithms: DSOS, distributed harmony search optimization system, DSOSPSO, APDHSOS, and AP2DHSOS [69]	Concurrent optimization of feature subsets and neighborhood size of the k-nearest algorithm	A hybrid of DSOS used for optimization and DPSO, known as DSOSPSO		16 Datasets	DACOR BABC DABC SADE PSOTVIW

(Continued)

Table 6: *Continued*

SOS algorithm	Objective	Method for automatic clustering	Dataset used	Comparison to	Metrics used
Multiobjective SOS [23]	NSMOSOS algorithm for optimization in feature subset for BCI	Generate initial population similar to classic SOS for optimization in BCI		Two motor imagery datasets	No feature selection, GA with aggressive mutation, relief algorithm, forward selection, classical SSO, SOS algorithm with limited no. of features, and weighted sum-based MOSOS
CSOS [66]	Solving automatic clustering problems based on SOS and K-means algorithms	Initialize the cluster randomly in the ecosystem, which the SOS optimizes for the clustering problem		10 Datasets extracted from the UCI machine learning repository	Flower pollination, PSO, differential evolution, multiverse optimizer, artificial bee colony, and cuckoo search
MSOS (nwSOS) [46]	The newly improved SOS method, specifically referred to as nwSOS, is employed to address optimization challenges in greater dimensions	The benefit factors employ a nonlinear methodology, wherein the weights of the benefit factors are used to investigate the search region efficiently		20 Standard benchmark functions with 100 and 500 dimensions	DE, PSO, SOS, teaching-learning-based optimization, SOS-ABF1&2, modified improved symbiotic organisms search, Improved symbiotic organisms search, SOS-ABF1, SOS-ABF2, Bayesian optimization algorithm, whale optimization algorithm, HSOS, salp swarm algorithm with improved symbiotic organisms search, enhanced self-organizing systems-based search algorithm, and cuckoo search optimization with social fabric-based individual
SOS-based [68]	Selection and parameter optimization for underwater target classification using the self-organizing system (SOS) method	Determining the initial cluster numbers based on the ecosize		28 Benchmark and 10 composition functions	GWO, GA, SaNSDE, SOS, rCMA-ES, and continuous ranked probability score optimization

proposed an NSMOSOS algorithm to investigate a multiobjective problem and found the balance between the classification accuracy and the number of selected features for the BCI system. The result demonstrated that the proposed method outperforms the classical SOS algorithm and other compared methods in terms of the number of successful runs, the number of features, and accuracy. In a recent study, decomposition-based multiobjective symbiotic organism search (MOSOS/D), using a novel optimizer for multiobjective problems, was proposed by Ganesh et al. [83]. The SOS algorithm is the basis of the proposed optimizer, which is considered a star-rising metaheuristic driven by the occurrence of symbioses among organisms. To assess the performance of the proposed approach, four different optimizers such as multi-objective evolutionary algorithm based on decomposition, NSGA-II, multi-objective marine predator slgorithm, and multi-objective equilibrium optimizer are compared using six commonly used performance metrics. The result shows that MOSOS/D is dominant to the other compared approaches. Similarly, Zhao and Liu [84] proposed an MSOS genetic multi-swarm symbiotic organism search (GMSOS) that uses good points set and memory mechanisms to improve the optimization ability and population diversity of the SOS method. The memory mechanism is used in three stages of the SOS algorithm, and the proposed method produces the initial population by using good point sets rather than a uniform distribution. By preventing simultaneous falls into local optima, these approaches aid in more efficiently balancing the search scope's exploration and exploitation. Compared to the other methods under consideration, results indicated that the suggested approach is dominant. Moreover, Chakraborty et al. [46] proposed a new MSOS algorithm (nwSOS), which is a modification of the original SOS algorithm that simulates the efforts of organisms to survive in the ecosystem. This method performs better on the controlled-source audio-magnetotellurics dataset with minimal tuning parameters as compared with other nature-inspired optimization methods. This method has shown promising results in terms of convergence analysis, statistical analysis, and complexity analysis.

The findings of hybridized algorithms are summarized in Table 6. The study encompasses a total of 20 distinct SOS variants hybridized with 12 different K-means automatic clustering algorithms.

5.4 RQ4: What is the rate of publication of the various SOS variants algorithms for clustering and feature selection?

Figure 4 illustrates the summarized search process for all the work completed aligned with the SLR, showing the number of publications from 2014 to 2023 that focus on the utilization of different variants of SOS algorithms for clustering and feature selection. The present study encompasses publications that explore the integration of clustering algorithms with different SOS methods, as discussed in the selected article. The graph depicts a declining growth pattern in the field of research during the initial phases, particularly from 2016 to 2018, wherein the number of published articles began to exhibit a notable decrease in 2017. The number

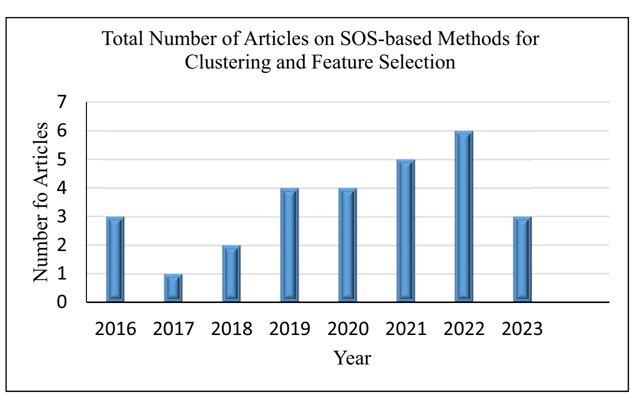

Figure 4: Rate of publications on SOS algorithms for clustering and feature selection.

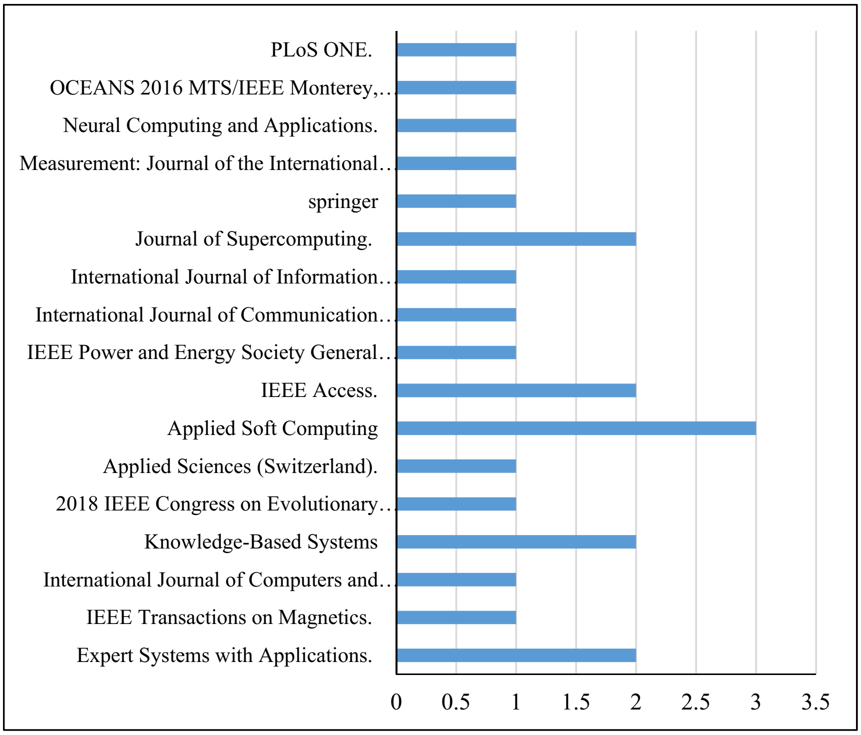

Figure 5: Number of articles per journal.

of articles under consideration for the study showed a predominantly upward trend after the year 2019. The years 2021 and 2022 had the most substantial publication output, with five and six articles, respectively, during the research period. This was succeeded by four papers each in 2019 and 2020. The analysis of a single article in the year 2017 indicates a decline in the rate of pertinent publications throughout that period. It is important to acknowledge that, owing to spatial constraints, Figure 5 displays only 70% of the reviewed articles along with their respective publication names. Figure 5 shows the various titles of the journals and the respective quantities of articles utilized in the investigation.

5.5 RQ5: What are the challenges and prospects of SOS-based clustering algorithms?

The symbiotic organism search (SOS) algorithms are fascinating metaheuristics inspired by the symbiotic interactions observed in ecosystems. An organism in nature develops various relationships to survive and thrive, which comprise mutualism, commensalism, and parasitism. This algorithm leverages these principles to better solve optimization problems. However, this section discusses some of the strengths and limitations of different approaches, as well as potential areas for future research. On this note, Chakraborty et al. [85] proposed a novel chaotic SOS (CSOS) optimization method to improve the efficiency of the traditional SOS optimization approach in multilevel image segmentation. This optimization ensures a strategic distance from premature convergence. The results indicate that the CSOS method has a higher convergence rate than the others; nevertheless, the main drawback is a little increase in temporal complexity while producing the chaotic sequence through the use of a nonlinear dynamical discrete map. Therefore, it was suggested that future research could focus on embedding chaos with additional developed metaheuristic algorithms to enhance the optimal solution. Similarly, Du et al. [79] proposed an IBSOS algorithm that uses a transfer function to binarize the continuous SOS algorithms and evaluate the influence of different transfer functions on the efficiency of the BSOS algorithm. This was aimed to resolve the challenges of the original BSOS algorithm, which has a premature convergence since it cannot directly solve the binary problem. In addition, the common drawbacks of the various SOS algorithms and their hybridizations are addressed by ISOSK-means, which improves the hybrid SOSK-means algorithm, in terms of stability, reliability, and efficiency [57]. It was shown that the proposed ISOSK-means algorithm has better performance than the original SOSK-means algorithm in terms of clustering accuracy, stability, and robustness. To further address some of the limitations of the existing SOS algorithms, such as the imbalance among the exploration and exploitation, as well as the over-exploration phenomenon, Chakraborty et al. [46] proposed an improved SOS algorithm in which the findings revealed that the proposed approach outperformed other compared algorithms when evaluated on 35 standard benchmark data and three engineering design challenges. Furthermore, a hybrid model was proposed by Ikotun and Ezugwu [65] to resolve the challenges of the individual SOS algorithms and the hybrid versions for a more reliable, stable, and effective performance. Although the initialization problem of the original K-means algorithm has been addressed by the hybrid SOSK-means algorithm, the amount of time needed for computing during the K-means phase is still proportional to the size of the dataset. As a result, better K-means variants can be integrated with SOS to shorten the time that the traditional K-means spend performing local searches. In addition, to further increase the performance of the hybrid algorithm, better SOS versions can be developed.

Having established the fact that the proposed SOS algorithms are efficient in handling automatic clustering problems, some insights into the challenges and advancements related to clustering using the SOS algorithm are discussed as follows:

5.5.1 Initial cluster centroids

The K-means clustering algorithms aim to divide given data into k pre-defined separate and nonoverlapping groups, each of which has a single group to which each data point belongs. However, given the sensitivity to the initial centroids of the cluster, the convergence probability into a local optimum, and the specification of the cluster number as an input parameter, the clustering performance is constrained. Therefore, one of the primary challenges in clustering algorithms, including SOS-based approaches, is determining the initial cluster centroids. Some researchers aimed to address this by combining SOS with the popular K-means algorithm [57,65,66]. The hybrid approach uses SOS as a global search metaheuristic to generate optimal initial cluster centroids for K-means. This helps improve K-means' performance by avoiding convergence into local optimal.

5.5.2 Parameter tuning

While knowing the number of clusters apriori is the most fundamental problem in cluster analysis, some of the metaheuristics algorithms, namely SOS-based algorithms, can be used to discover the number of clusters automatically. However, although the SOS-based algorithm is considered a parameter-free metaheuristic, some hybrid algorithms often require rigorous parameter tuning. Therefore, more researcher effort is needed to explore different mechanisms to fine-tune more parameters effectively for HSOS-based clustering methods. In addition, balancing exploration and exploitation is another important factor for achieving optimal results.

5.5.3 Solution quality and validity indices

Evaluating the quality of clustering solutions remains essential. Researchers often use validity indices (such as the Davies-Bouldin index) to assess the effectiveness of clustering algorithms. Future work should focus on enhancing solution quality assessment for SOS-based clustering approaches. It is essential to develop algorithmic methods to compare various SOS-based clustering methods based on multiple validity criteria, including internal, stability, and biological indicators [65].

5.5.4 Generalization and robustness

While SOSK-means shows promising results on specific datasets, its generalization across diverse data domains needs further investigation. Due to contamination introduced at various phases of measurement and processing, a dataset is frequently not pure, which necessitates data cleaning in data mining. The presence of noise and outliers in the data is another challenge. Therefore, researchers need to explore robustness under noisy data, varying cluster shapes, and different data distributions.

5.5.5 Scalability and efficiency

As with any metaheuristic, scalability becomes crucial when dealing with large datasets. One of the potential solutions would be to decrease the reliance of algorithms on user-dependent parameters, which can improve the effectiveness of clustering algorithms. Future research can also design improved algorithms that enhance the efficiency of SOS-based clustering methods, which will allow them to handle real-world applications effectively. In summary, the hybridization of SOS with K-means offers exciting possibilities for automatic clustering. Researchers can continue exploring novel adaptations, addressing challenges, and validating their approaches to various datasets.

6 Research limitations

This SLR identifies several strategies based on self-organizing maps (SOS) for clustering and feature selection. Our objective is to enhance the internal and external validity of our procedures by accomplishing the RQs. This section will address the limitations and obstacles that may arise concerning the validity of this argument.

- This SLR is exclusively focused on scholarly articles and conference papers that examine the topics of clustering, classification, or feature selection utilizing SOS algorithms. During the initial stages of the study, our search approach was employed to identify and subsequently exclude several research publications that were deemed irrelevant to this analysis and aligned with the criteria considered. Nevertheless, it is widely recognized that the inclusion of other sources, such as supplementary sourcebooks, should have provided a better quality review.
- Our research was restricted to materials written exclusively in the English language. The presence of potential publications in other languages in this field of research gives rise to linguistic bias. Fortunately, the documents collected were in the English language.
- Despite the study's focus on primary databases to review scholarly articles, it is plausible that other relevant studies available in additional digital libraries were not considered. To address this, we conducted a comparative analysis between the keywords and phrases used to search and a widely recognized compilation of scholarly research works. Nevertheless, it is possible to overlook specific synonyms while searching for keywords. Therefore, the SLR procedure has undergone revisions to ensure the exclusion of any essential terms.

7 Conclusion

This research work presents an SLR, which specifically adopts the PRISMA approach, to systematically review the existing research works on unsupervised clustering and feature selection problems based on the SOS-based methods. Specifically, the applicability of various techniques in SOS-based methods for clustering and classification is reviewed. We have also presented the hybridization of the K-means clustering algorithm with different SOS algorithms. The primary objective of each hybridization with the adopted clustering approach in the resultant hybridized algorithm was considered. The findings of the study revealed that various SOS methods were adapted as clustering and feature selection methods in which CSOS, DSOS, and MASOS were mostly used for the clustering applications, while BSOS, BSOSST, and BSOSVT were used for feature selection problems. The findings also revealed that, of all the selected studies for this review, only a few studies specifically focused on hybridizing SOS with the K-means algorithm for automatic data clustering application. In addition, the publication rate of research on SOS-based methods for clustering and feature selection has been presented. Five distinct RQs were created to meet the objectives of the study, and the answers to these questions were included in a comprehensive analysis of the many SOS approaches that incorporated the K-means clustering algorithm to produce hybrid methods.

To respond to our first question (RQ1), several articles that focused on various SOS algorithms for classification and clustering problems have been investigated. Adopted clustering approaches were also studied. In response to the second question (RQ2), different articles were selected among the reviewed articles, which were aimed primarily to resolve feature selection problems for clustering and classification purposes. These articles were few related to the overall articles considered for this study. In response to the third question (RQ3), various hybridizations of the SOS algorithms and K-means clustering algorithm that addressed clustering issues were also considered. The various approaches of reviewed hybridized algorithms for automatic clustering were also discussed in response to this RQ. The response to the fourth question (RQ4), provided a comprehensive investigation of the publication trends concerning SOS-based algorithms for clustering and feature selection problems in the last decade. The results of the analyzed publications show that, in the majority of the reviewed studies, there is a generally low rate of research publication that includes the hybridization of K-means with SOS algorithms. This suggests that this field of study needs a lot more attention, particularly to address automatic clustering issues. To achieve higher-quality clustering results, the study also shows that the current hybridized K-means algorithms with SOSs still need longer execution times when used for large dataset clustering.

Acknowledgments: We thank the anonymous reviewers for their comments and suggestions.

Funding information: This work was supported by the Malaysia Ministry of Higher Education and the Universiti Kebangsaan Malaysia under the FRGS/1/2022/ICT02/UKM/02/3 with machines and materials.

Author contributions: Conceptualization, all authors; methodology, all authors; writing – original draft preparation, Abbas Fadhil Jasim AL-Gburi; writing –review and editing, Mohd Zakree Ahmad Nazri, Mohd Ridzwan Bin Yaakub, Zaid Abdi Alkareem Alyasseri; funding acquisition, Mohd Zakree Ahmad Nazri; All authors have read and agreed to the published version of the manuscript.

Conflict of interest: Authors state no conflict of interest.

Data availability statement: Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

References

[1] Boushaki SI, Bendjeghaba O, Kamel N. Biomedical document clustering based on accelerated symbiotic organisms search algorithm. Int J Swarm Intell Res. 2021;12:169–85. doi: 10.4018/IJSIR.2021100109.
[2] Usery L. Problems of large spatial databases. Geogr Inf Sci Technol Body Knowl. 2019;2019. doi: 10.22224/gistbok/2019.2.13.
[3] Sabour S, Frosst N, Hinton GE. Dynamic routing between capsules. In: Advances in Neural Information Processing Systems. California, United States: Neural Information Processing Systems Foundation, Inc.; 2017.
[4] Kuriqi A, Pinheiro AN, Sordo-Ward A, Garrote L. Influence of hydrologically based environmental flow methods on flow alteration and energy production in a run-of-river hydropower plant. J Clean Prod. 2019;232:1028–42. doi: 10.1016/j.jclepro.2019.05.358.

[5] Kuriqi A, Pinheiro AN, Sordo-Ward A, Garrote L. Flow regime aspects in determining environmental flows and maximising energy production at run-of-river hydropower plants. Appl Energy. 2019;256:113980. doi: 10.1016/j.apenergy.2019.113980.
[6] Noureddine S, Zineeddine B, Toumi A, Betka A, Benharkat AN. A new predictive medical approach based on data mining and Symbiotic Organisms Search algorithm. Int J Comput Appl. 2022;44:465–79. doi: 10.1080/1206212X.2020.1809825.
[7] Gedam AG, Shikalpure SG. Direct kernel method for machine learning with support vector machine. In: 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies, ICICICT 2017; 2018.
[8] Brito Da Silva LE, Wunsch DC. An information-theoretic-cluster visualization for self-organizing maps. IEEE Trans Neural Netw Learn Syst. 2018;29:2595–613. doi: 10.1109/TNNLS.2017.2699674.
[9] Sinaga KP, Yang MS. Unsupervised K-means clustering algorithm. IEEE Access. 2020;8:80716–27. doi: 10.1109/ACCESS.2020.2988796.
[10] Ozdemir G, Karaboga N. A review on the cosine modulated filter bank studies using meta-heuristic optimization algorithms. Artif Intell Rev. 2019;52:1629–53.
[11] Das S, Mukhopadhyay A, Roy A, Abraham A, Panigrahi BK. Exploratory power of the harmony search algorithm: Analysis and improvements for global numerical optimization. IEEE Trans Syst Man Cybern Part B Cybern. 2011;41:89–106. doi: 10.1109/TSMCB.2010.2046035.
[12] Ding J, Wang Q, Zhang Q, Ye Q, Ma Y. A hybrid particle swarm optimization-cuckoo search algorithm and its engineering applications. Math Probl Eng. 2019;2019:1–12. doi: 10.1155/2019/5213759.
[13] Valian E, Tavakoli S, Mohanna S, Haghi A. Improved cuckoo search for reliability optimization problems. Comput Ind Eng. 2013;64:459–68. doi: 10.1016/j.cie.2012.07.011.
[14] Vanneschi L, Silva S. Particle swarm optimization. In: Natural Computing Series. Cham, Switzerland: Springer; 2023.
[15] Piotrowski AP, Napiorkowski JJ, Piotrowska AE. Population size in particle swarm optimization. Swarm Evol Comput. 2020;58:100718. doi: 10.1016/j.swevo.2020.100718.
[16] Neroni M. Ant colony optimization with warm-up. Algorithms. 2021;14:295. doi: 10.3390/a14100295.
[17] de Andoin MG, Echanobe J. Implementable hybrid quantum ant colony optimization algorithm. Quantum Mach Intell. 2022;4:12. doi: 10.1007/s42484-022-00065-1.
[18] Cheng MY, Prayogo D. Symbiotic organisms search: A new metaheuristic optimization algorithm. Comput Struct. 2014;139:98–112. doi: 10.1016/j.compstruc.2014.03.007.
[19] Kumar S, Tejani GG, Mirjalili S. Modified symbiotic organisms search for structural optimization. Eng Comput. 2019;35:1269–96. doi: 10.1007/s00366-018-0662-y.
[20] Mirjalili S, Lewis A. The whale optimization algorithm. Adv Eng Softw. 2016;95:51–67. doi: 10.1016/j.advengsoft.2016.01.008.
[21] Gharehchopogh FS, Shayanfar H, Gholizadeh H. A comprehensive survey on symbiotic organisms search algorithms. Artif Intell Rev. 2020;53:2265–312. doi: 10.1007/s10462-019-09733-4.
[22] Ganesh N, Shankar R, Čep R, Chakraborty S, Kalita K. Efficient feature selection using weighted superposition attraction optimization algorithm. Appl Sci. 2023;13:3223. doi: 10.3390/app13053223.
[23] Baysal YA, Ketenci S, Altas IH, Kayikcioglu T. Multi-objective symbiotic organism search algorithm for optimal feature selection in brain computer interfaces. Expert Syst Appl. 2021;165:113907. doi: 10.1016/j.eswa.2020.113907.
[24] Jaffel Z, Farah M. A symbiotic organisms search algorithm for feature selection in satellite image classification. In 2018 4th International Conference on Advanced Technologies for Signal and Image Processing, ATSIP 2018; 2018.
[25] Cheng MY, Cao MT, Herianto JG. Symbiotic organisms search-optimized deep learning technique for mapping construction cash flow considering complexity of project. Chaos Solitons Fractals. 2020;138:109869. doi: 10.1016/j.chaos.2020.109869.
[26] Mohammadzadeh H, Gharehchopogh FS. Feature selection with binary symbiotic organisms search algorithm for email spam detection. Int J Inf Technol Decis Mak. 2021;20:469–515. doi: 10.1142/S0219622020500546.
[27] Cheng MY, Kusoemo D, Gosno RA. Text mining-based construction site accident classification using hybrid supervised machine learning. Autom Constr. 2020;118:103265. doi: 10.1016/j.autcon.2020.103265.
[28] Abdullahi M, Ngadi MA, Dishing SI, Abdulhamid SM, Usman MJ. A survey of symbiotic organisms search algorithms and applications. Neural Comput Appl. 2020;32:547–66.
[29] Kitchenham B, Pretorius R, Budgen D, Pearl Brereton O, Turner M, Niazi M, et al. Systematic literature reviews in software engineering-A tertiary study. Inf Softw Technol. 2010;52:792–805.
[30] Kitchenham B, Pearl Brereton O, Budgen D, Turner M, Bailey J, Linkman S. Systematic literature reviews in software engineering – A systematic literature review. Inf Softw Technol. 2009;51:7–15. doi: 10.1016/j.infsof.2008.09.009.
[31] Ezugwu AE, Prayogo D. Symbiotic organisms search algorithm: theory, recent advances and applications. Expert Syst Appl. 2019;119:184–209.
[32] Dokeroglu T, Sevinc E, Kucukyilmaz T, Cosar A. A survey on new generation metaheuristic algorithms. Comput Ind Eng. 2019;137:106040. doi: 10.1016/j.cie.2019.106040.
[33] Hemeida AM, Hassan SA, Mohamed AAA, Alkhalaf S, Mahmoud MM, Senjyu T, et al. Nature-inspired algorithms for feed-forward neural network classifiers: A survey of one decade of research. Ain Shams Eng J. 2020;11:659–75.
[34] Darvishpoor S, Darvishpour A, Escarcega M, Hassanalian M. Nature-inspired algorithms from oceans to space: A comprehensive review of heuristic and meta-heuristic optimization algorithms and their potential applications in drones. Drones. 2023;7:427.
[35] Jain M, Kaur G, Saxena V. A K-Means clustering and SVM based hybrid concept drift detection technique for network anomaly detection. Expert Syst Appl. 2022;193. doi: 10.1016/j.eswa.2022.116510.

[36] Gerami Seresht N, Lourenzutti R, Fayek AR. A fuzzy clustering algorithm for developing predictive models in construction applications. Appl Soft Comput J. 2020;96:106679. doi: 10.1016/j.asoc.2020.106679.
[37] Bhattacharjee P, Mitra P. A survey of density based clustering algorithms. Front Comput Sci. 2021;15:151308.
[38] Li T, Rezaeipanah A, Tag El Din ESM. An ensemble agglomerative hierarchical clustering algorithm based on clusters clustering technique and the novel similarity measurement. J King Saud Univ – Comput Inf Sci. 2022;34:3828–42. doi: 10.1016/j.jksuci.2022.04.010.
[39] Weber CM, Ray D, Valverde AA, Clark JA, Sharma KS. Gaussian mixture model clustering algorithms for the analysis of high-precision mass measurements. Nucl Instrum Methods Phys Res Sect A Accel Spectrometers, Detect Assoc Equip. 2022;1027:166299. doi: 10.1016/j.nima.2021.166299.
[40] Cheng MY, Gosno RA. SOS 2.0: an evolutionary approach for SOS algorithm. Evol Intell. 2021;14:1965–83. doi: 10.1007/s12065-020-00476-8.
[41] Ezugwu AE, Adeleke OJ, Viriri S. Symbiotic organisms search algorithm for the unrelated parallel machines scheduling with sequence-dependent setup times. PLoS One. 2018;13:e0200030.
[42] Ezugwu AE, Adewumi AO. Soft sets based symbiotic organisms search algorithm for resource discovery in cloud computing environment. Futur Gener Comput Syst. 2017;76:33–50.
[43] Ezugwu AE-S, Adewumi AO. Discrete symbiotic organisms search algorithm for travelling salesman problem. Expert Syst Appl. 2017;87:70–8.
[44] Ezugwu AES, Adewumi AO, Frîncu ME. Simulated annealing based symbiotic organisms search optimization algorithm for traveling salesman problem. Expert Syst Appl. 2017;77:189–210. doi: 10.1016/j.eswa.2017.01.053.
[45] Banerjee S, Chattopadhyay S. Power optimization of three dimensional turbo code using a novel modified symbiotic organism search (MSOS) algorithm. Wirel Pers Commun. 2017;92:941–68. doi: 10.1007/s11277-016-3586-0.
[46] Chakraborty S, Nama S, Saha AK. An improved symbiotic organisms search algorithm for higher dimensional optimization problems. Knowl Syst. 2022;236:107779. doi: 10.1016/j.knosys.2021.107779.
[47] Rodrigues LR, Gomes JPP, Neto ARR, Souza AH. A modified symbiotic organisms search algorithm applied to flow shop scheduling problems. In 2018 IEEE Congress on Evolutionary Computation, CEC 2018 – Proceedings; 2018.
[48] Secui DC. A modified symbiotic organisms search algorithm for large scale economic dispatch problem with valve-point effects. Energy. 2016;113:366–84. doi: 10.1016/j.energy.2016.07.056.
[49] Tejani GG, Savsani VJ, Patel VK. Adaptive symbiotic organisms search (SOS) algorithm for structural design optimization. J Comput Des Eng. 2016;3:226–49. doi: 10.1016/j.jcde.2016.02.003.
[50] Nama S, Kumar Saha A, Ghosh S. A hybrid symbiosis organisms search algorithm and its application to real world problems. Memetic Comput. 2017;9:261–80. doi: 10.1007/s12293-016-0194-1.
[51] Do DTT, Lee D, Lee J. Material optimization of functionally graded plates using deep neural network and modified symbiotic organisms search for eigenvalue problems. Compos Part B Eng. 2019;159:300–26. doi: 10.1016/j.compositesb.2018.09.087.
[52] Panda A, Pani S. A symbiotic organisms search algorithm with adaptive penalty function to solve multi-objective constrained optimization problems. Appl Soft Comput J. 2016;46:344–60. doi: 10.1016/j.asoc.2016.04.030.
[53] Tiwari A, Pandit M. Bid based economic load dispatch using symbiotic organisms search algorithm. In: Proceedings of 2nd IEEE International Conference on Engineering and Technology, ICETECH 2016; 2016.
[54] Ayala HVH, Klein CE, Mariani VC, Coelho LDS. Multiobjective symbiotic search algorithm approaches for electromagnetic optimization. IEEE Trans Magn. 2017;53:1–4. doi: 10.1109/TMAG.2017.2665350.
[55] Rajah V, Ezugwu AE. Hybrid symbiotic organism search algorithms for automatic data clustering. In 2020 Conference on Information Communications Technology and Society, ICTAS 2020 – Proceedings; 2020.
[56] Agbaje MB, Ezugwu AE, Els R. Automatic data clustering using hybrid firefly particle swarm optimization algorithm. IEEE Access. 2019;7:184963–84. doi: 10.1109/ACCESS.2019.2960925.
[57] Ikotun AM, Ezugwu AE. Improved SOSK-means automatic clustering algorithm with a three-part mutualism phase and random weighted reflection coefficient for high-dimensional datasets. Appl Sci. 2022;12:13019. doi: 10.3390/app122413019.
[58] Serifoglu Yilmaz C, Yilmaz V, Güngör O. On the use of the SOS metaheuristic algorithm in hybrid image fusion methods to achieve optimum spectral fidelity. Int J Remote Sens. 2020;41:3993–4021. doi: 10.1080/01431161.2019.1711244.
[59] Wang Y, Geng C, Xu N. Assembly sequence optimization based on hybrid symbiotic organisms search and ant colony optimization. Soft Comput. 2021;25:1447–64. doi: 10.1007/s00500-020-05230-x.
[60] Ezugwu AE, Olusanya MO, Govender P. Mathematical model formulation and hybrid metaheuristic optimization approach for near-optimal blood assignment in a blood bank system. Expert Syst Appl. 2019;137:74–99. doi: 10.1016/j.eswa.2019.06.059.
[61] Cheng MY, Prayogo D, Wu YW. Prediction of permanent deformation in asphalt pavements using a novel symbiotic organisms search–least squares support vector regression. Neural Comput Appl. 2019;31:6261–73. doi: 10.1007/s00521-018-3426-0.
[62] Zhang B, Sun L, Yuan H, Lv J, Ma Z. An improved regularized extreme learning machine based on symbiotic organisms search. In: Proceedings of the 2016 IEEE 11th Conference on Industrial Electronics and Applications, ICIEA 2016; 2016.
[63] Aljarah I, Mafarja M, Heidari AA, Faris H, Mirjalili S. Clustering analysis using a novel locality-informed grey wolf-inspired clustering approach. Knowl Inf Syst. 2020;62:507–39. doi: 10.1007/s10115-019-01358-x.
[64] Goldanloo MJ, Gharehchopogh FS. A hybrid OBL-based firefly algorithm with symbiotic organisms search algorithm for solving continuous optimization problems. J Supercomput. 2022;78:3998–4031. doi: 10.1007/s11227-021-04015-9.

[65] Ikotun AM, Ezugwu AE. Boosting k-means clustering with symbiotic organisms search for automatic clustering problems. PLoS One. 2022;17:0272861. doi: 10.1371/journal.pone.0272861.
[66] Zhou Y, Wu H, Luo Q, Abdel-Baset M. Automatic data clustering using nature-inspired symbiotic organism search algorithm. Knowl Syst. 2019;163:546–57. doi: 10.1016/j.knosys.2018.09.013.
[67] Yang CL, Sutrisno H. A clustering-based symbiotic organisms search algorithm for high-dimensional optimization problems. Appl Soft Comput. 2020;97:106722. doi: 10.1016/j.asoc.2020.106722.
[68] Sherin BM, Supriya MH. SOS based selection and parameter optimization for underwater target classification. In: OCEANS 2016 MTS/IEEE Monterey, OCE 2016; 2016.
[69] Liao TW, Kuo RJ. Five discrete symbiotic organisms search algorithms for simultaneous optimization of feature subset and neighborhood size of KNN classification models. Appl Soft Comput J. 2018;64:581–95. doi: 10.1016/j.asoc.2017.12.039.
[70] Zare-Noghabi A, Shabanzadeh M, Sangrody H. Medium-term load forecasting using support vector regression, feature selection, and symbiotic organism search optimization. In: IEEE Power and Energy Society General Meeting; 2019.
[71] Gana NN, Abdulhamid SM, Misra S, Garg L, Ayeni F, Azeta A. Optimization of support vector machine for classification of spyware using symbiotic organism search for features selection. In: Lecture Notes in Networks and Systems; 2022.
[72] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ. 2021;372. doi: 10.1136/bmj.n71.
[73] Kitchenham B, Charters S. Guidelines for performing systematic literature reviews in software engineering. Engineering. 2007;65–1052. doi: 10.1145/1134285.1134500.
[74] Genc-Nayebi N, Abran A. A systematic literature review: Opinion mining studies from mobile app store user reviews. J Syst Softw. 2017;125:207–19.
[75] Acharya DS, Mishra SK. A multi-agent based symbiotic organisms search algorithm for tuning fractional order PID controller. Meas J Int Meas Confed. 2020;155:107559. doi: 10.1016/j.measurement.2020.107559.
[76] Mohmmadzadeh H, Gharehchopogh FS. An efficient binary chaotic symbiotic organisms search algorithm approaches for feature selection problems. J Supercomput. 2021;77:9102–44. doi: 10.1007/s11227-021-03626-6.
[77] Mohammadzadeh H, Gharehchopogh FS. A multi-agent system based for solving high-dimensional optimization problems: A case study on email spam detection. Int J Commun Syst. 2021;34. doi: 10.1002/dac.4670.
[78] Arora S, Anand P. Binary butterfly optimization approaches for feature selection. Expert Syst Appl. 2019;116:147–60. doi: 10.1016/j.eswa.2018.08.051.
[79] Du ZG, Pan JS, Chu SC, Chiu YJ. Improved binary symbiotic organism search algorithm with transfer functions for feature selection. IEEE Access. 2020;8:225730–44. doi: 10.1109/ACCESS.2020.3045043.
[80] Han C, Zhou G, Zhou Y. Binary symbiotic organism search algorithm for feature selection and analysis. IEEE Access. 2019;7:166833–59. doi: 10.1109/ACCESS.2019.2953800.
[81] Miao F, Yao L, Zhao X. Symbiotic organisms search algorithm using random walk and adaptive Cauchy mutation on the feature selection of sleep staging. Expert Syst Appl. 2021;176. doi: 10.1016/j.eswa.2021.114887.
[82] Apolloni J, Leguizamón G, Alba E. Two hybrid wrapper-filter feature selection algorithms applied to high-dimensional microarray experiments. Appl Soft Comput J. 2016;38:922–32. doi: 10.1016/j.asoc.2015.10.037.
[83] Ganesh N, Shankar R, Kalita K, Jangir P, Oliva D, Pérez-Cisneros M. A novel decomposition-based multi-objective symbiotic organism search optimization algorithm. Mathematics. 2023;11:1898. doi: 10.3390/math11081898.
[84] Zhao P, Liu S. An improved symbiotic organisms search algorithm with good point set and memory mechanism. J Supercomput. 2023;79:11170–97. doi: 10.1007/s11227-023-05083-9.
[85] Chakraborty F, Roy PK, Nandi D. A novel chaotic symbiotic organisms search optimization in multilevel image segmentation. Soft Comput. 2021;25:6973–98. doi: 10.1007/s00500-021-05611-w.