Electronics, Communications and Networks
A.J. Tallón-Ballesteros et al. (Eds.)

doi:10.3233/FAIA231261

Overview and Application of Sparrow Search Algorithm in Deep Learning

Hongjun WANG, Teng FEI[1] , Lanxue LIU and Kangle WU
Institute of Information Engineering, Tianjin University of Commerce, Tianjin, China

Abstract. A brand-new swarm intelligence optimization algorithm called the Sparrow Search Algorithm (SSA) was put forth in 2020. By simulating the foraging process of sparrows, SSA efficiently solves optimization problems, exhibiting advantages such as faster convergence and excellent optimization abilities. This paper introduces the basic principles of the Sparrow Search Algorithm, analyzes its existing issues, summarizes improvements made to existing Sparrow algorithms, and then discusses the application of SSA in deep learning.

Keywords. Sparrow Search Algorithm, Algorithm Enhancement, Deep Learning, Algorithm Application

1. Introduction

Since Professor Holland proposed the Genetic Algorithm based on Darwin's theory of evolution and natural selection mechanisms in 1975, numerous scholars have drawn inspiration from various biological populations and physical phenomena, resulting in the creation of various swarm intelligence optimization algorithms, including the Grey Wolf Optimization Algorithm[1], Particle Swarm Algorithm[2], Harris Hawk Algorithm[3], and Whale Optimization Algorithm[4].

The Sparrow Search Algorithm (SSA) [5] is an emerging swarm intelligence algorithm inspired by the feeding and dodging habits of sparrows. SSA's unique design imparts it with exceptional optimization capabilities and rapid convergence speed. Compared to other algorithms, SSA demonstrates higher intelligence and adaptability, enabling it to tackle complex optimization problems. Notably, SSA excels in handling nonlinear problems, characterized by its rapid convergence and strong optimization prowess. It has found widespread application in deep learning.

2. Basic Principles of Sparrow Algorithm

The SSA algorithm is a novel swarm intelligence optimization algorithm that mimics the foraging and counter-attacking behavior of sparrows. The population is split up into distinct roles: finders, joiners, and scouts. Finders are responsible for locating food sources and determining the direction and area for foraging. Joiners utilize the details

[1] Corresponding Author, Tengfei, Professer, Tianjin University of Commerce, Tianjin, China; E-mail: feiteng@tjcu.edu.cn.

that finders provide them to get food. Scouts, randomly distributed within the population, release danger signals when they perceive a threat, prompting the entire population to counter-attack. The roles of finders and joiners can interchange, but their ratios remain constant.

The formula for updating the position of finders after t iterations is as follows:

$$X_i^{t+1} = \begin{cases} X_i^t \cdot \exp(-\frac{i}{a \cdot T}) & R_2 < ST \\ X_i^t + Q \cdot L & R_2 \geq ST \end{cases} \tag{1}$$

Here, t represents the present iteration number, X_i^t indicates the location of the i-th sparrow following t iterations, Q is a random number with a normal distribution in the interval $[0,1]$, T is the largest number of iterations, and a is a random value in the range of $(0,1)$, L is the unit matrix, R_2 is a random value in the range of (0,1), and the protection threshold is typically set between 0.5 and 1. When $R_2 < ST$, it shows that predators have not been spotted by the population, enabling finders to explore widely. When $R_2 \geq ST$, it indicates that a scouting sparrow has detected a predator, prompting an immediate release of the danger signal, causing other sparrows to fly to different positions for foraging.

Joiners are individuals with lower fitness values. They observe finders obtaining food and compete for it. Joiners that fail to obtain food will fly to areas where food is abundant. The following formula is used to update the joiners' position:

$$X_i^{t+1} = \begin{cases} Q \cdot \exp(-\frac{X_{worst} - X_i^t}{i^2}) & i > n/2 \\ X_p^t + \left| X_i^t - X_p^t \right| \cdot A^+ \cdot L & otherwise \end{cases} \tag{2}$$

Here, the current population's worst location is depicted by X_{worst}, and the best location among finders is depicted by X_p^t, A is a matrix with a dimension of $1 \times D$, every element was given a random value of 1 or -1. in addition $A^+ = A^T (AA^T)^{-1}$,When $i > n/2$, it suggests that the i-th joiner did not get any food, is starving, and must fly to a different area to forage. Otherwise, it indicates that the joiner will choose a location at random that is close to the current finder's optimal foraging location.

Scouts typically account for $10\% - 20\%$ of the sparrow population, their formula for updating positions is as follows:

$$X_i^{t+1} = \begin{cases} X_i^t + \beta \cdot \left| X_i^t - X_{best}^t \right| & f_i > f_g \\ X_i^t + K \cdot \left(\frac{\left| X_i^t - X_{worst}^t \right|}{(f_i - f_w) + \varepsilon} \right) & f_i = f_g \end{cases} \tag{3}$$

Here, X_{best}^t represents the globally best position, β is the step size adjustment coefficient, which is a random number following a normal distribution with mean 0 and variance 1, K is a random number in the range of [-1,1], and ε is a tinier constant that prevents division by zero. The global highest and lowest fitness values of the current sparrow population are denoted by f_g and f_w, respectively, whereas f_i denotes the fitness value of the i-th sparrow. When $f_i > f_g$, it suggests that the sparrow is at the edge of the population as a whole and is open to attack. When $f_i = f_g$, it indicates that the sparrow has realized that predators are a concern and that it has to surround other sparrows to stay safe from attacks.

3. Improvements to Sparrow Search Algorithm

The optimization capability of the Sparrow Search Algorithm (SSA) stands out within swarm intelligence algorithms and has garnered widespread attention. However, the SSA still exhibits certain shortcomings:

(1) Sensitivity to Initial Population: The search capability is heavily influenced by the initial population in the SSA. The use of a Random function for population initialization often leads to an overabundance of individuals at the edges of the search space, resulting in insufficient population diversity and algorithm stability. Consequently, the current SSA population initialization method may not fully meet the demands of solving complex optimization problems.

(2) Imbalance between Global Search and Local Exploitation: Both the global search capability and local exploitation capability of the SSA are closely tied to the foragers' hunting abilities. The foragers' hunting prowess determines the amount of food the sparrow population acquires, ultimately affecting the quality of the algorithm's solutions. The setting of a safety threshold plays a pivotal role in determining the foragers' search range, thus influencing the balance between global search and local exploitation capabilities.

(3) Lack of Mutation Mechanism: The absence of a mutation mechanism can lead to a deficiency in population diversity. During the search process, the Sparrow Algorithm may gradually exhibit population clustering, which results in decreased diversity in the later iterations. This leads to a reduction in the algorithm's convergence accuracy, making it susceptible to getting stuck in local optima.

3.1. Population Initialization

The quality of the population during the initialization stage directly determines the algorithm's optimization performance. The Sparrow Search Algorithm typically employs random initialization of the population when solving optimization problems, which can lead to an uneven distribution of the population and consequently reduce diversity.

Li et al.[6] utilized the Tent chaotic mapping for population initialization, thereby increasing population diversity, this enhancement improved the likelihood of obtaining optimal solutions in the early optimization iterations and accelerated convergence. Liu et al.[7] applied the Circle chaotic mapping to the population initialization of the Sparrow Algorithm, enhancing the variety of initial solutions. On the basis of several population initialization strategies, Wu et al.[8] presented an enhanced cyclic chaotic mapping theory, they produced an initial population with more diversity and randomness by fusing quantum computing with quantum gate mutation mechanisms.

3.2. Multiple Mechanism Improvements

Zhang et al.[9] introduced the Discrete Sparrow Search Algorithm, which first performs roulette selection after population initialization. They then introduced a sequence-based decoding method to update sparrow positions. Finally, they employed a global disturbance mechanism combining Gaussian mutation and swap operators to balance exploration and extraction abilities. These strategies improved solution quality and accelerated convergence. Hong et al.[10] proposed an improved Sparrow Search Algorithm (ISSA), which introduced a strategy focusing on forager concentration to update forager positions, enhancing global search capabilities. They also put forward vector enclosure models and direction selection strategies to update joiner positions, thus enhancing local search capabilities. Ma et al.[11] introduced an Enhanced Multi-Strategy Sparrow Search Algorithm (EMSSA) based on three strategies. Firstly, they proposed an adaptive tent chaotic theory to endow the initial population with greater diversity and randomness. Then, they constructed a weighted cosine method founded on growth functions to prevent settling into local optima. Finally, they designed a similar disturbance function and introduced triangular similarity theory, enhancing search capabilities. Li et al.[12] put forward five strategies to improve the Sparrow Search Algorithm, namely Improved Sine Mapping, Sine-Cosine Algorithm, Elite Opposition-Based Learning, Levy Flight, and Gaussian Mutation, to enhance SSA performance.

3.3. Integration with Other Algorithmic Improvement Strategies

Fan et al.[13] proposed a Hybrid Sparrow Search Algorithm that combines the Sparrow Search Algorithm with Particle Swarm Optimization(PSO). This method leverages the local optimal solutions from the SSA and the search efficiency of PSO to achieve global optimization. It demonstrates strong global search capabilities, avoiding the generation of local optima, and exhibits excellent search performance in both low- and high-dimensional spaces. Khaleel et al.[14] combined the SSA with Differential Evolution (DE) algorithm, expanding the search efficiency of the Sparrow Algorithm. Ren et al.[15] integrated the SSA with Firefly Algorithm, using firefly disturbance to update all sparrows to the best sparrow. This enhancement increased the global search capabilities of the Sparrow Algorithm.

4. Application of Sparrow Algorithm in Deep Learning

Deep learning is extensively used in many domains, including speech recognition and computer vision. The hyperparameters of these algorithms greatly affect how well they work. Determining the hyperparameters of deep neural networks and other complicated machine learning models can be difficult. Furthermore, local optima are frequently reached by the hyperparameter optimization algorithms now in use. The main focus of this paper is on using the Sparrow Search Algorithm to choose deep learning hyperparameters.

The unique incrementally generated Randomly Connected Network (SCN) model, which has a supervised mechanism, shows great advantages in tackling massive quantities of data regression and classification issues. Nonetheless, how specific network parameters are allocated and chosen has an impact on SCN's correctness. A Randomly Connected Network based on the Chaotic Sparrow Search Algorithm (CSSA) was introduced by Zhang et al.[16]. The Chaotic Sparrow Search Algorithm (CSSA) was the first algorithm they created. Because regularization parameters and the ratio factor of biases and weights affect SCN performance, CSSA was used to automatically generate better SCN parameters. According to experimental findings, CSSA-SCN functioned more effectively and practically than SCN. A Fast Randomly Connected Network (FSCN) built on an improved Sparrow Search Algorithm (ISSA) was proposed by Wu et al.[17]. They set the amount of hidden layer nodes immediately instead of removing the original SCN's steady increase in hidden layer nodes. They created the input weights and biases of the nodes in the hidden layer using the ISSA. Additionally, they also kept the supervisory mechanism in place to evaluate the biases and weights of each hidden layer node. In order to optimize specific parameters in FSCN and improve classification performance, nodes that failed to meet the requirements of the supervisory mechanism were regenerated using ISSA.

For medium- to long-term load forecasting, the Support Vector Machine (SVM)-based regression model is a common method. Traditional Support Vector Machines can be difficult to figure out the hyperparameters of, which results in subpar prediction performance. An Improved Sparrow Search Algorithm (ISSA) was proposed by Li et al.[18] to solve the problem of hyperparameter selection in the SVM model. ISSA-SVM greatly increased prediction accuracy when compared to the original Support Vector Machine, BP Neural Network, and Multivariate Linear Regression approaches.

For Deep Belief Networks (DBN) parameter optimization, Gai et al.[19] presented a gear fault severity detection method based on the Sparrow Search Algorithm (SSA). First, they labeled gear fault signals according to varying degrees of severity in order to start training the first DBN. Subsequently, they introduced SSA to optimize both the learning rate and batch size of the initial DBN, avoiding interference caused by subjective parameter selection. The Deep Belief Network that was tuned exhibited improved accuracy, stability, and feature extraction capabilities.

Xiong et al.[20] used an adaptive T-distribution mutation operator in conjunction with an opposition-based learning technique to optimize the SSA. The Deep Extreme Learning Machine's (DELM) input layer weights and bias settings were optimized using this method. As a result, they developed a thorough ISSA-DELM network model for astute quantitative evaluation of flaws in pipelines carrying natural gas. This model effectively mitigated the influence of random input weights and biases, resulting in improved quantitative prediction performance.

The Learning Vector Quantization (LVQ) neural network's starting weights were optimized by Zhang et al.[21] using the Sparrow Search Algorithm, the training convergence of the LVQ neural network was faster after it was optimized.

An innovative Long Short-Term Memory (LSTM) network optimized with the Improved Sparrow Search Algorithm (ISSA) was presented by Liu et al.[22]. Lithium battery Remaining Useful Life (RUL) was estimated using this network. Because the LSTM hyperparameters had a direct effect on the success of predictions, they were the first to be chosen for optimization. Through ISSA-assisted LSTM hyper parameters optimization, they were able to accurately predict RUL, higher accuracy and robustness were shown by the suggested LSTM network.

A novel Sparrow Search Algorithm and Convolutional Neural Network (SSA-CNN) algorithm for leak detection in oil pipelines was presented by Li et al.[23]. Initially, the suggested SSA-CNN technique transformed time series input data into a two-dimensional matrix and contrasted various pooling circumstances and convolution kernel sizes, the CNN parameters were then optimized using the SSA method. This technique not only beat conventional machine learning techniques, but it significantly enhanced CNN's classification power.

5. Conclusion

The Sparrow Search Algorithm, as a guided intelligent optimization method, has been widely applied to various optimization issues. However, it still has limitations such as premature convergence and suboptimal solution accuracy, which restrict its application scope. This paper provided an overview of the application of the Sparrow Search Algorithm in deep learning research, highlighting its advantages in hyperparameter selection, it offers a novel solution to the problem of hyperparameter optimization in deep learning models.

Acknowledgments

This work is support by Tianjin Natural Science Foundation (No. 20JCYBJC00320, College Students' innovation and Entrepreneurship Project (No. 2022SKYZ313) and Innovation and Entrepreneurship Project for College Students (202310069092).

References

[1] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," Advances in Engineering Software, vol. 69, pp. 46-61, Mar 2014, doi: 10.1016/j.advengsoft.2013.12.007.
[2] D. S. Wang, D. P. Tan, and L. Liu, "Particle swarm optimization algorithm: an overview," Soft Computing, vol. 22, no. 2, pp. 387-408, Jan 2018, doi: 10.1007/s00500-016-2474-6.
[3] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. L. Chen, "Harris hawks optimization: Algorithm and applications," FUTURE GENERATION COMPUTER SYSTEMS-THE INTERNATIONAL JOURNAL OF ESCIENCE, vol. 97, pp. 849-872, AUG 2019, doi: 10.1016/j.future.2019.02.028.
[4] S. Mirjalili and A. Lewis, "The Whale Optimization Algorithm," ADVANCES IN ENGINEERING SOFTWARE, vol. 95, pp. 51-67, MAY 2016, doi: 10.1016/j.advengsoft.2016.01.008.
[5] J. Xue and B. Shen, "A novel swarm intelligence optimization approach: sparrow search algorithm," Systems Science & Control Engineering, vol. 8, no. 1, pp. 22-34, 2020, doi: 10.1080/21642583.2019.1708830.
[6] Z. Li, J. Guo, X. Gao, X. Yang, and Y.-L. He, "A multi-strategy improved sparrow search algorithm of large-scale refrigeration system: Optimal loading distribution of chillers," Applied Energy, vol. 349, 2023, doi: 10.1016/j.apenergy.2023.121623.
[7] J. H. Liu and Z. H. Wang, "A Hybrid Sparrow Search Algorithm Based on Constructing Similarity," Ieee Access, vol. 9, pp. 117581-117595, 2021, doi: 10.1109/access.2021.3106269.
[8] R. Wu et al., "An improved sparrow search algorithm based on quantum computations and multi-strategy enhancement," Expert Systems with Applications, vol. 215, 2023, doi: 10.1016/j.eswa.2022.119421.
[9] Z. Zhang and Y. Han, "Discrete sparrow search algorithm for symmetric traveling salesman problem," Applied Soft Computing, vol. 118, 2022, doi: 10.1016/j.asoc.2022.108469.
[10] J. Hong, B. Shen, J. Xue, and A. Pan, "A vector-encirclement-model-based sparrow search algorithm for engineering optimization and numerical optimization problems," Applied Soft Computing, vol. 131, 2022, doi: 10.1016/j.asoc.2022.109777.
[11] J. Ma, Z. Hao, and W. Sun, "Enhancing sparrow search algorithm via multi-strategies for continuous optimization problems," Information Processing & Management, vol. 59, no. 2, 2022, doi: 10.1016/j.ipm.2021.102854.
[12] J. Li, J. Chen, and J. Shi, "Evaluation of new sparrow search algorithms with sequential fusion of improvement strategies," Computers & Industrial Engineering, vol. 182, 2023, doi: 10.1016/j.cie.2023.109425.
[13] Y. Fan, Y. Zhang, B. Guo, X. Luo, Q. Peng, and Z. Jin, "A Hybrid Sparrow Search Algorithm of the Hyperparameter Optimization in Deep Learning," Mathematics, vol. 10, no. 16, 2022, doi: 10.3390/math10163019.

[14] M. I. Khaleel, "Efficient job scheduling paradigm based on hybrid sparrow search algorithm and differential evolution optimization for heterogeneous cloud computing platforms," Internet of Things, vol. 22, 2023, doi: 10.1016/j.iot.2023.100697.
[15] X. Y. Ren, S. Chen, K. Y. Wang, and J. Tan, "Design and application of improved sparrow search algorithm based on sine cosine and firefly perturbation," Mathematical Biosciences and Engineering, vol. 19, no. 11, pp. 11422-11452, 2022, doi: 10.3934/mbe.2022533.
[16] C. Zhang and S. Ding, "A stochastic configuration network based on chaotic sparrow search algorithm," Knowledge-Based Systems, vol. 220, 2021, doi: 10.1016/j.knosys.2021.106924.
[17] H. Wu, A. Zhang, Y. Han, J. Nan, and K. Li, "Fast stochastic configuration network based on an improved sparrow search algorithm for fire flame recognition," Knowledge-Based Systems, vol. 245, 2022, doi: 10.1016/j.knosys.2022.108626.
[18] J. Li, Y. Lei, and S. Yang, "Mid-long term load forecasting model based on support vector machine optimized by improved sparrow search algorithm," Energy Reports, vol. 8, pp. 491-497, 2022, doi: 10.1016/j.egyr.2022.02.188.
[19] J. Gai, K. Zhong, X. Du, K. Yan, and J. Shen, "Detection of gear fault severity based on parameter-optimized deep belief network using sparrow search algorithm," Measurement, vol. 185, 2021, doi: 10.1016/j.measurement.2021.110079.
[20] J. Xiong, W. Liang, X. Liang, and J. Yao, "Intelligent quantification of natural gas pipeline defects using improved sparrow search algorithm and deep extreme learning machine," Chemical Engineering Research and Design, vol. 183, pp. 567-579, 2022, doi: 10.1016/j.cherd.2022.06.001.
[21] K. Zhang, Z. Chen, L. Yang, and Y. Liang, "Principal component analysis (PCA) based sparrow search algorithm (SSA) for optimal learning vector quantized (LVQ) neural network for mechanical fault diagnosis of high voltage circuit breakers," Energy Reports, vol. 9, pp. 954-962, 2023, doi: 10.1016/j.egyr.2022.11.118.
[22] Y. Liu, J. Sun, Y. Shang, X. Zhang, S. Ren, and D. Wang, "A novel remaining useful life prediction method for lithium-ion battery based on long short-term memory network optimized by improved sparrow search algorithm," Journal of Energy Storage, vol. 61, 2023, doi: 10.1016/j.est.2023.106645.
[23] Q. Li, Y. Shi, R. Lin, W. Qiao, and W. Ba, "A novel oil pipeline leakage detection method based on the sparrow search algorithm and CNN," Measurement, vol. 204, 2022, doi: 10.1016/j.measurement.2022.112122.