Multimedia Network Data Fusion System Integrating SSA and Reinforcement Learning

Fangrui Li
School of Applied Technology, Zhumadian Preschool Education College, Zhumadian 463000, China

***Abstract*—To improve the performance and efficiency of multimedia network data fusion system, this study proposes an improved sparrow search algorithm on the ground of reinforcement learning algorithm and sparrow search algorithm, and improves the multimedia network data fusion model on the ground of this algorithm. A performance comparison experiment was conducted on the improved sparrow search algorithm, and it was found that the algorithm entered a convergence state after 380 iterations in a unimodal function. Its time consumption is lower than other comparison algorithms, and it has not fallen into the local optimal situation after 500 iterations in the multimodal benchmark function. Its performance is significantly superior to other comparison algorithms. Moreover, the study conducted relevant experiments on the multimedia network data fusion model and found that the F1 value output by the model was 0.37, with an accuracy of 92.4%, which is higher than other data fusion models. And the mean square error of this model reaches 0.52, and the processing time is 0.1 seconds, which is lower than other comparative data fusion models. The quality of output data and data processing efficiency of this model are better. The relevant outcomes demonstrate that the improved sparrow search algorithm possesses good global search and convergence performance. And the improved multimedia network data fusion model has better accuracy and efficiency, and has good practical application value. This study can provide reference and reference for multimedia network data fusion systems.**

Keywords*—*Sparrow search algorithm; reinforcement learning; multimedia network; data fusion; performance improvement

I. Introduction

As the boost of the Internet and mobile communication technology, the generation and transmission of multimedia data are showing an explosive growth trend. These data include various forms such as images, audio, and video, which have had a huge impact on people's lives and work. However, they also pose new challenges to the efficient and fast transmission of data [1]. The multimedia network data fusion model (DFM) can integrate and analyze data from different media and sources, helping people obtain more comprehensive information and reducing unnecessary data consumption [2]. However, data from different media have different characteristics and expressions, and the scale of multimedia network data is enormous, posing great challenges to the efficiency of processing and analysis [3]. Sparrow Search Algorithm (SSA), as a heuristic search algorithm, can optimize feature selection and fusion in multimedia data fusion. Reinforcement learning algorithms can achieve adaptive adjustment of SSA parameters and select the best fusion strategy in multimedia data fusion. Although various data fusion algorithms have been proposed, such as Particle Swarm Optimization (PSO), Bat Algorithm (BA), and Adaptive Fusion Spanning Tree (AFST), these methods still have limitations in terms of real-time performance, accuracy, and global search capability. Especially in the context of the rapid growth of multimedia data and the diversification of data types, existing technologies are unable to meet the growing demand for data processing. In addition, there are few attempts in existing research to combine sparrow search algorithm with reinforcement learning algorithm to improve data fusion performance [4-5]. Therefore, this research proposes the construction of a multimedia network DFM that integrates SSA and reinforcement learning, and constructs a data fusion system on the ground of this to improve the effectiveness and performance of multimedia network data fusion. The core issue of the research is how to improve the performance and efficiency of the multimedia network data fusion system. The purpose of the research is to develop an efficient multimedia network data fusion system to address the challenges of the current surge in data volume and diversification of data types. To achieve this goal, a method combining SSA and reinforcement learning algorithms was proposed, and an improved Sparrow Search Algorithm (ISSA) was created. The research aims to improve the performance and efficiency of the ISSA algorithm in processing and analyzing large-scale multimedia data by integrating it into the multimedia network data fusion model. This not only includes improving the accuracy of data fusion, but also involves optimizing data processing speed to ensure that valuable information can be quickly and effectively extracted from large amounts of complex data. The relevance of the research is reflected in the following aspects: Firstly, with the widespread application of multimedia data in various fields such as social media, online education, remote healthcare, etc., the accuracy and efficiency of data fusion directly affect the quality of decision-making and user experience. Secondly, with the promotion of the Internet of Things and 5G technology, the amount of data will further increase, which puts higher demands on data fusion technology. Finally, the advancement of data fusion technology is of great significance for promoting technological innovation and industrial development in related fields.

The study first provides a brief overview of the current status of multimedia network data fusion technology, then clearly points out the limitations of the existing technology, and proposes improvement solutions. Through empirical experiments, the study has demonstrated the superior performance of improved algorithms and models in processing multimedia network data, including higher accuracy, faster processing speed, and better global search capability. These achievements not only address the challenges faced by existing technologies, but also provide new directions for the development of multimedia network data fusion.

II. Related Work

As the boost of science and technology, optimization algorithms is widely applied in various fields. To select the optimal parameters for proton exchange membrane fuel cell stacks and improve their performance, Zhu et al. proposed an adaptive sparrow search algorithm and applied it to parameter identification. Then, the study conducted performance verification experiments on this method and found that the adaptive sparrow search algorithm can reduce the square error of the battery stack voltage, and the computational efficiency of this algorithm was better than other algorithms [6]. In response to the issue of insufficient accuracy in predicting rubber fatigue life, Wang et al. proposed an optimized rubber life prediction model on the ground of an ISSA to verify its effectiveness. The validation results found that the model has better prediction accuracy, convergence speed, and stability performance than traditional models [7]. To solve the problem of neglecting indicator weights in the dynamic priority scheduling algorithm of the power system, Meng et al. proposed an improved dynamic priority scheduling algorithm on the ground of improved reinforcement learning and verified its performance. The relevant results found that the improved algorithm has a faster learning speed compared to traditional algorithms, and it also achieved optimization of weight parameters, reducing scheduling costs [8]. To achieve observability and ease of operation of calibration parameters, Nobre et al. proposed a calibration assistance model on the ground of reinforcement learning and conducted empirical experiments on the model. The relevant results found that the model is applicable to positioning and drawing on multiple platforms, and has lower requirements for professional operations compared to traditional models [9].

The progress of science and technology has also brought about explosive growth in data, and there is currently an increasing amount of research on data fusion. For enhancing the accuracy and robustness of the integrated navigation system, Mai et al. presented a multi-source collaborative positioning system on the ground of an adaptive Kalman filter and conducted performance simulation experiments on the system. The relevant outcomes showcased that the navigation and positioning performance was markedly enhanced compared to traditional fonts [10]. To achieve the static gesture recognition function of Kinect camera depth maps, Sharma et al. proposed feature extraction and data fusion on the ground of static gesture datasets, constructed a gesture recognition model, and validated the model. The relevant outcomes showcased that the recognition accuracy of this model can reach 95.7%, which possesses the value in practical application [11]. For the difficulty in reconstructing the pseudo static displacement of the low-frequency part of the moving load, W et al. proposed a pseudo static displacement reconstruction model on the ground of acceleration and response data fusion to verify its effectiveness. The relevant results found that this model can improve the accuracy and accuracy of the reconstruction results compared to traditional models [12]. To address the occurrence of tilt events in the Linz Donawitz steelmaking converter, De et al. proposed a tilt event warning model on the ground of sensor data fusion technology to verify its effectiveness. The relevant results found that the model can improve alarm accuracy and improve the reliability of the anti collapse system [13].

In summary, research on data fusion is becoming increasingly mature. However, there is still limited research on improving the SSA using reinforcement learning algorithms and combining this optimization algorithm with a multimedia network DFM to construct a multimedia network DFM for SSA and reinforcement learning. Therefore, this study constructs a multimedia network DFM on the ground of SSA and reinforcement learning to efficiently handle data fusion problems in multimedia networks and improve the accuracy of data fusion.

III. Construction of a DFM Integrating SSA and Reinforcement Learning

To facilitate the normal transportation of multimedia network data faster and more efficiently, and to efficiently process multimedia network data, this study constructs the ISSA on the ground of SSA and reinforcement learning algorithm. Then it applies the algorithm to the multimedia network DFM to improve its data fusion quality and data processing efficiency.

A. SSA Algorithm on the Ground of Reinforcement Learning Optimization

At present, multimedia network DFMs often suffer from inaccurate information fusion and slow processing speed when dealing with large-scale unstructured or nonlinear data. This study proposes to integrate sparrow search algorithm with reinforcement learning algorithm to construct an ISSA on the ground of reinforcement learning algorithm, thereby improving the processing speed of multimedia network data models. SSA, as a swarm intelligence optimization algorithm on the ground of sparrow food search behavior, has good applicability in solving complex nonlinear optimization problems [14-16]. The process of SSA is shown in Fig. 1.

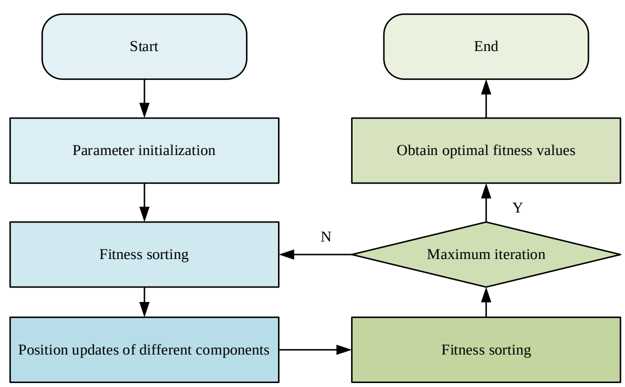


Fig. 1. Flow of the SSA.

As shown in Fig. 1, firstly, the SSA sets parameters like the size, maximum of iterations, and producer ratio of the sparrow population, and initializes the sparrow population. Subsequently, the SSA calculates the fitness values of all sparrows in the population and sorts their fitness values in order of size. The formula for calculating the fitness matrix of sparrows is shown in (1).

$$\begin{cases} \mathbf{F_X} = \left[\mathbf{f(X_1)}, \mathbf{f(X_2)}, \ldots, \mathbf{f(X_N)}\right] \\ \mathbf{f(X_i)} = \left[f(x_{i,1}), f(x_{i,2}), \ldots, f(x_{i,d})\right] \\ \mathbf{X} = \left[\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_N}\right], \mathbf{X_i} = \left[x_{i,1}, x_{i,2}, \ldots, x_{i,d}\right] (i = 1, 2, \ldots N) \end{cases} \tag{1}$$

In Eq. (1), x represents sparrow. $\mathbf{F_X}$ represents the vector matrix of sparrow fitness values. $\mathbf{X}$ represents the vector matrix of sparrows. $\mathbf{X_i}$ represents the row vector of sparrow position. $\mathbf{f(X_i)}$ represents the fitness value of sparrows. N serves as the population size of sparrows. d represents the dimension of the variable. The SSA sorts the fitness values on the ground of the fitness matrix to determine the individuals with the best and worst fitness values. Individuals with high fitness values become producers in the population, responsible for searching for areas and directions for the population. Their relevant formula is shown in Eq. (2).

$$X_{i,j}(t+1) = \begin{cases} X_{i,j}(t) \cdot \exp\left(\dfrac{-i}{\alpha \cdot T}\right), W < ST \\ X_{i,j}(t) + N_{rand} \cdot L, W \geq ST, L = 1 \times d \end{cases} \tag{2}$$

In Eq. (2), t serves as the quantity iterations. α serves as a random number, located between [0,1]. W serves as the warning value. ST serves as the safety value, located between [0.5,1]. N_{rand} represents a normally distributed random number. L represents the dimension matrix. The calculation formula for the accompanying position update is illustrated in Eq. (3).

$$X_{i,j}(t+1) = \begin{cases} N_{rand} \cdot \exp\left(\dfrac{X_{worst}(t) - X_{ij}(t)}{i^2}\right), i > \dfrac{N}{2} \\ X_{bestj}(t) + \left|X_{ij}(t) - X_{bestj}(t+1)\right| \cdot A^+ \cdot L, i \leq \dfrac{N}{2} \\ A^+ = A^T\left(AA^T\right)^{-1}, A = 1 \times d \end{cases} \tag{3}$$

In Eq. (3), X_{worst} represents the worst-case position. X_{bestj} represents the optimal position. A is a matrix, with an initial value typically of 1 or -1. A^T is the transposition of A. The relevant formula is shown in Eq. (4).

$$X_{i,j}(t+1) = \begin{cases} X_{bestj}(t) + \beta \cdot \left|X_{i,j}(t) - X_{bestj}(t)\right|, f_i > f_g \\ X_{i,j}(t) + K\left(\dfrac{\left|X_{i,j}(t) - X_{worstj}(t)\right|}{(f_i - f_w) + \varepsilon}\right), f_i = f_g, K \in [-1,1] \end{cases} \tag{4}$$

In Eq. (4), ε is a constant. $X_{bestj}(t)$ serves as the optimal position. β represents the step size control parameter, f_g serves as the best fitness. f_w represents the worst fit. After updating the relevant position, it counts and sorts the individual fitness values again until the iterations' maximum is achieved. According to the SSA calculation principle, the SSA mainly improves search performance through information exchange between individuals, and is prone to falling into local optima. The Q-learning algorithm, as a reinforcement learning algorithm, has good applicability, convergence, and generalization ability in discrete space problems. Applying Q-learning algorithm to SSA can improve the optimal search strategy of SSA, improve its convergence speed, and achieve adaptive parameter adjustment of SSA. The basic principle of the Q-learning algorithm is showcased in Fig. 2.

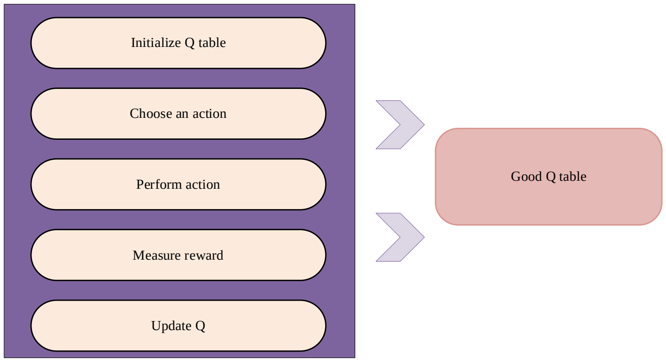


Fig. 2. Q. Basic principle of the learning algorithm.

Fig. 2 showcases that the Q-learning algorithm can use a value function to calculate the cumulative reward for adopting an action strategy in a certain state. The relevant calculation is shown in Eq. (5).

$$V_\pi(s) = E_\pi\left[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \ldots \middle| s_t = s\right] \tag{5}$$

In Eq. (5), R represents the reward value. γ represents the discount factor. V represents the state function. γ represents the state. π represents the action strategy. The reinforcement learning is for finding the optimal strategy to maximize the state function. The relevant calculation is demonstrated in Eq. (6).

$$V_{\pi^*}(s) = \max_\pi V_\pi(s) \tag{6}$$

In Eq. (6), π^* represents the optimal strategy. The Q-learning algorithm can define the state action value as a function $Q_\pi(s,a)$. It represents the accumulated discount reward obtained when adopting a strategy. The calculation formula is shown in Eq. (7).

$$Q_\pi(s,a) = E_\pi\left[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \ldots \middle| s_t = s, a = a_t\right] \tag{7}$$

In Eq. (7), a represents a specific action. At this point, the relevant calculation is showcased in Eq. (8).

$$V_{\pi^*}(s) = \max_\pi Q_{\pi^*}(s,a) \tag{8}$$

The update formula for the Q-value function is shown in Eq. (9).

$$Q(s,a) \leftarrow Q(s,a) + \partial(r + \gamma \max_{a'} Q(s',a') - Q(s,a)) \tag{9}$$

In Eq. (9), ∂ represents the Q learning rate. r represents the current reward. s' represents the state of the next decision. s' represents the action of the next decision. $\max_{a'} Q(s',a')$ serves as the maximum cumulative reward value after updating the status. In addition, this study also utilizes a random walk strategy for enhancing the SSA and avoiding local optima. The calculation formula for the random walk process is shown in Eq. (10).

$$\begin{cases} X(t_{\max}) = \left[0, cussum(2k(t_1)-1), \mathrm{K}\ , cussum(2k(t_n)-1)\right] \\ k(t) = \begin{cases} 1, rand > 0 \\ 0, rand < 0 \end{cases} \end{cases} \quad (10)$$

In Eq. (10), $t_{\max}$ serves as the maximum of iterations. $cussum$ serves as the cumulative sum. k serves as a random function. In addition, to ensure that the population walking area in the sparrow algorithm is within the feasible range, the calculation formula for sparrow position update is shown in Eq. (11).

$$X_i(t) = \frac{\left(X_i(t) - X_{\min}\right) * \left(X_{\max}'(t) - X_{\min}'(t)\right)}{(X_{\max} - X_{\min})} + X_{\min}'(t) \quad (11)$$

In Eq. (11), $X_{\max}$ serves as the maximum value. $X_{\min}$ serves as the minimum. $X_{\max}'(t)$ serves as the maximum as the iterations is t. $X_{\min}'(t)$ serves as the minimum as the iterations is t. This study utilizes Q-learning algorithm to improve the SSA, so that the SSA will decide the next action on the ground of the current state and reward signal. If an action leads to a better state and reward signal, it will repeat the action in the future when encountering similar states, otherwise avoid the action. The optimized SSA process is shown in Fig. 3.

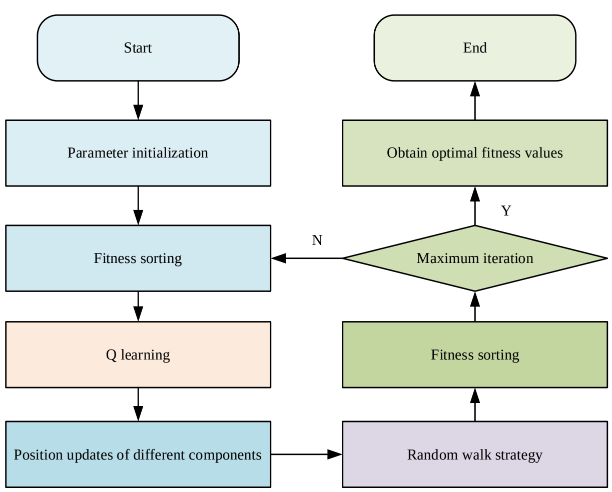


Fig. 3. The ISSA flow.

As shown in Fig. 3, this study first sets the parameters of the sparrow algorithm, defines the search direction as spatial actions, defines fitness as a reward function, and initializes the population and Q table. Subsequently, it enters the search process and calculates the fitness value of the sparrow population. During the search, the algorithm will select an action on the ground of the current state and Q table, and execute this action. The Q-learning algorithm updates the Q-table on the ground of the status and reward values after executing the action. When the Q value in the Q table converges, a random walk strategy is added for updating the global optimal position and improve its search ability. After obtaining the updated sparrow population position, it sorts the fitness values again. If a population with a higher fitness value appears after the update, it updates the optimal solution. Otherwise, it maintains its current optimal solution until reaching the maximum number of iterations.

B. Improved Multimedia Network DFM on the Ground of the ISSA

Data fusion technology is a technology that fuses or integrates data from different sources, types, and features, enabling comprehensive display of information to users. To ensure the fusion quality of multimedia network data, the quality of collected data and the selection of fusion technology are clearly key breakthroughs. Therefore, to achieve high accuracy and reliability of multimedia network data, this study utilizes multimedia sensor networks to construct a multi-layer DFM. Multi layer network data sensor network can improve the accuracy, real-time and data coverage of its collected data by collecting multimedia data from various sensor nodes [17], [18]. The basic structure of the relevant model is shown in Fig. 4.

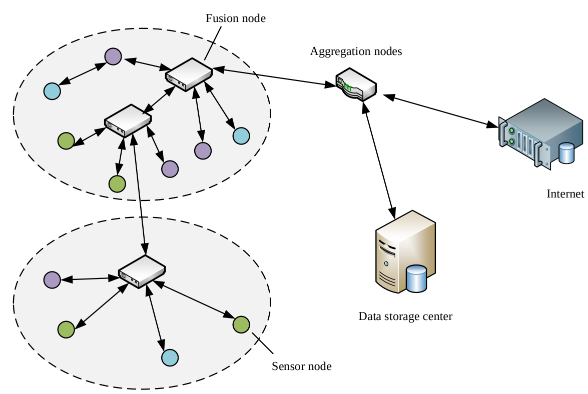


Fig. 4. Multi-layer DFM.

Fig. 4 demonstrates that for different types of multimedia network data, this study constructs different types of multimedia sensor networks on the ground of their data types, and assigns corresponding collection tasks to each layer of sensor network. Different layers of sensor networks send their collected data to the aggregation node by the fusion node. To ensure the coverage area of collected data, this study selected a DFM on the ground of multi-layer sensor networks. However, due to the variety of sensors, the amount of data collected is large, and the network data collected on the ground of sensor networks is prone to high data correlation, which leads to the problem of redundant data. Therefore, this study utilizes a clustering network on the ground of data difference to improve the DFM, to eliminate redundant network data and reduce unnecessary network data consumption. The principle of a clustering network DFM on the ground of data difference is shown in Fig. 5.

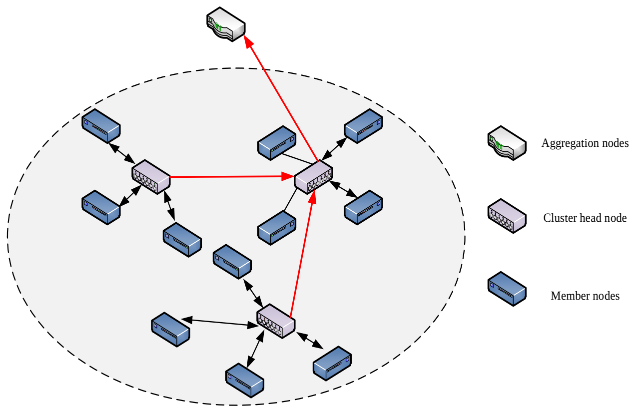


Fig. 5. Principle of the DFM on the ground of the topology of the cluster network.

Fig. 5 showcases that the DFM on the ground of clustering network topology divides the collection nodes in the sensor network into ordinary member nodes, cluster head nodes (CHN), and aggregation nodes. Firstly, this study calculates the difference in data collected by several sensor nodes and classifies them according to the difference threshold. The data difference matrix is shown in Eq. (12).

$$\delta_{nn} = \begin{bmatrix} 0, \delta_{12}, \delta_{13}, \mathrm{K}, \delta_{1n} \\ \delta_{21}, 0, \delta_{23}, \mathrm{K}, \delta_{2n} \\ \mathrm{L} \\ \delta_{n2}, \delta_{n2}, \delta_{n3}, \mathrm{K}, 0 \end{bmatrix} \tag{12}$$

In Eq. (12), the diagonals of the difference matrix are all 0. The difference between nodes not calculated is set to 1. Subsequently, the model performs cluster head elections on data from different categories. The selected cluster head can allocate its Time Division Multiple Access (TDMA) time slot to the ordinary member nodes within its corresponding cluster. Ordinary member checkpoints will periodically transmit collected data to CHN in terms of the allocated TDMA time slot. After completing a data collection, it calculates the remaining energy of the cluster head. If the remaining energy of the CHN is higher than or equal to the set energy threshold, then the node still serves as the CHN. Otherwise, it will reselect the CHN. After achieving the elimination of duplicate data in the fusion model, this study applies the constructed the ISSA to the DFM for enhancing its accuracy and processing speed. The DFM on the ground of the ISSA and clustering algorithm is shown in Fig. 6.

Fig. 6 demonstrates that the DFM includes five modules: preparation, data preprocessing, data training, data integration, and data processing. Firstly, this study deploys multiple multimedia sensors to capture and collect various audio and image multimedia data. Subsequently, the study used clustering algorithms to group nodes in wireless sensor networks, using indicators such as data similarity to achieve more efficient and accurate data processing and management. The third module is data training, which trains the ISSA on the ground of clustering algorithm to achieve adaptive search of the ISSA. The ISSA uses the data of each cluster as the search space, initializes the position of each sparrow individual as a data point in the cluster, and uses the ISSA to search for the optimal cluster head, improving its computational speed and performance. The fourth module is data integration, where the clustering algorithm merges, processes, and analyzes the data of the fused nodes to obtain more comprehensive and accurate data. Finally, this study utilizes data mining algorithms to analyze the integrated multimedia network data, extract valuable information from the data, and perform data visualization and reporting operations.

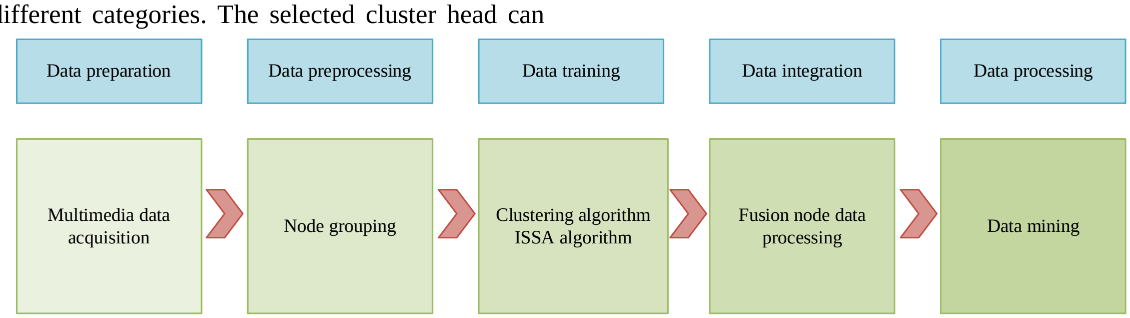


Fig. 6. DFM on the ground of the ISSA and clustering algorithm.

IV. Empirical Experiment on Multimedia Network DFM

For testing the ISSA and multimedia network DFM, performance comparison experiments and empirical analysis are conducted on them. This study first uses benchmark functions to conduct performance comparison tests on the optimization performance of the ISSA, and then conducts empirical experiments on the multimedia network DFM using the ImageNet dataset.

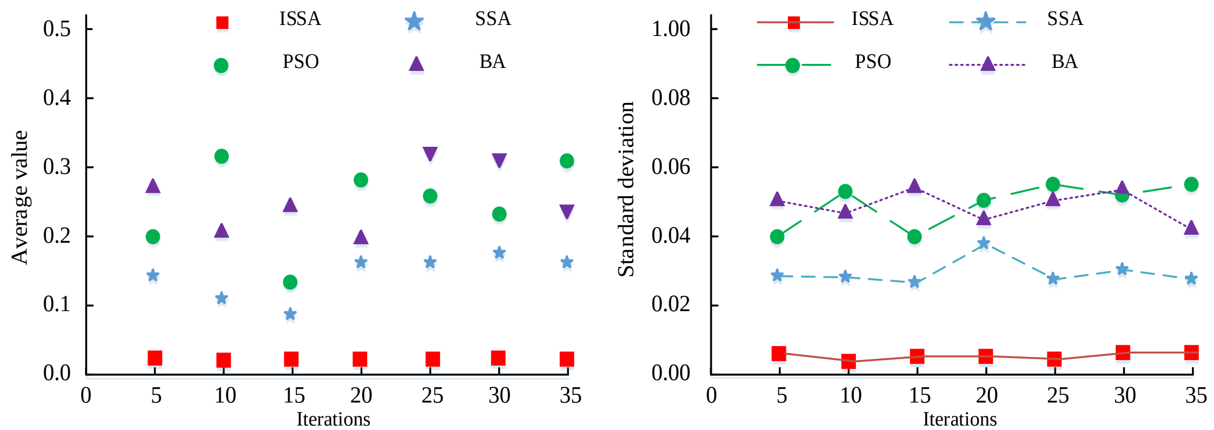
A. *Performance Verification Experiment of SSA Algorithm*

For testing the optimization of the ISSA, this study utilizes the PENALIZED benchmark function to verify its optimization performance, and compares its convergence performance with the single peak function Rosenbrock and the multi peak benchmark function Ackley. The minimum value of the PENALIZED benchmark function is 0. The comparison algorithms are SSA, Particle Swarm Optimization (PSO), and Bat Algorithm (BA). The comparison index is the average of the benchmark function, as well as the convergence curves of unimodal and multimodal functions. The experimental environment is Windows 10 and MATLAB 2018. The relevant comparing results in the PENALIZED benchmark function are shown in Fig. 7.

Fig. 7(a) shows the comparison results of the average values of the functions of various optimization algorithms. As shown in Fig. 7(a), the ISSA found the optimal value on the function PENALIZED, and relative to other comparative algorithms, the optimal value obtained is more approximate to the actual optimal solution, infinitely approaching the optimal solution, and its local search ability is more excellent. Fig. 7(b) showcases the comparison outcomes of the standard deviation of function values for each optimization algorithm. As shown in Fig. 7(b), the ISSA has a smaller standard deviation compared to other comparative algorithms, higher solving accuracy, and better stability. The relevant outcomes demonstrate that the ISSA possesses better local search ability and reliable performance, which makes it obtain the optimal solution possibly. The relevant convergence curves of each optimization algorithm are shown in Fig. 8.

Fig. 8(a) and (b) show the three-dimensional images of the Rosenbrock reference function and the optimization convergence curves of each algorithm, respectively. Fig. 8 indicates that the ISSA enters a convergence state when the number of iterations is 380, and its convergence performance is much higher than other comparison algorithms, and the convergence curve of the ISSA is steeper than other algorithms. This demonstrates that the ISSA can first search for the optimal value. The relevant outcomes demonstrate that the ISSA has better local search performance. The convergence curves of each optimization algorithm in the Ackley benchmark function are shown in Fig. 9.

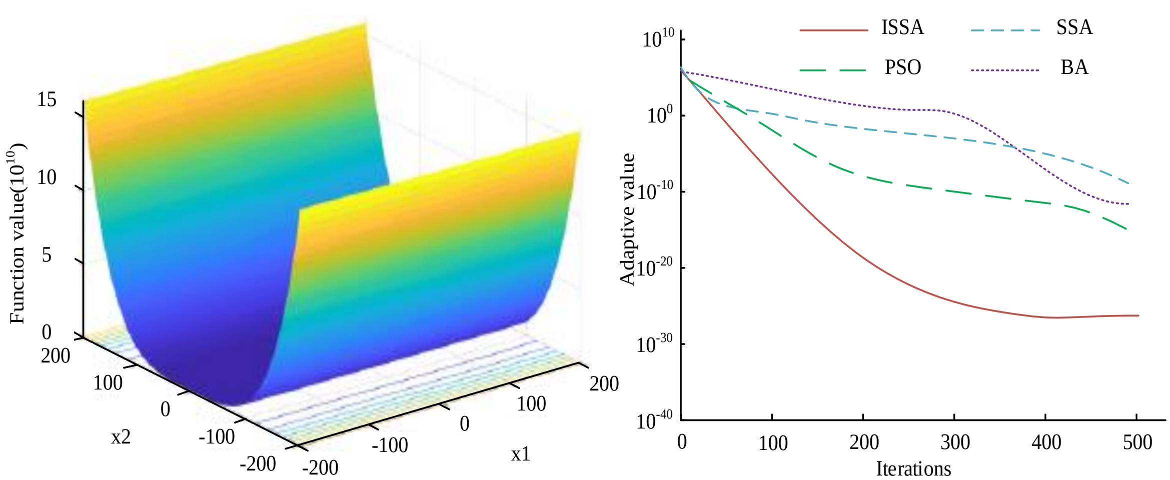
Fig. 9(a) and (b) show the three-dimensional images of the Ackley reference function and the optimization convergence curves of each algorithm, respectively. As shown in Fig. 8, SSA, PSO, and BA algorithms fell into local optima at 450, 350, and 320 iterations, respectively, while the ISSA did not fall into local optima at 500 iterations, and its search ability in multimodal functions was better. The relevant outcomes demonstrate that the ISSA has better global search performance than other algorithms.



(a) Mean value of benchmark function testing for each algorithm

(b) Standard deviation of benchmark function testing for each algorithm

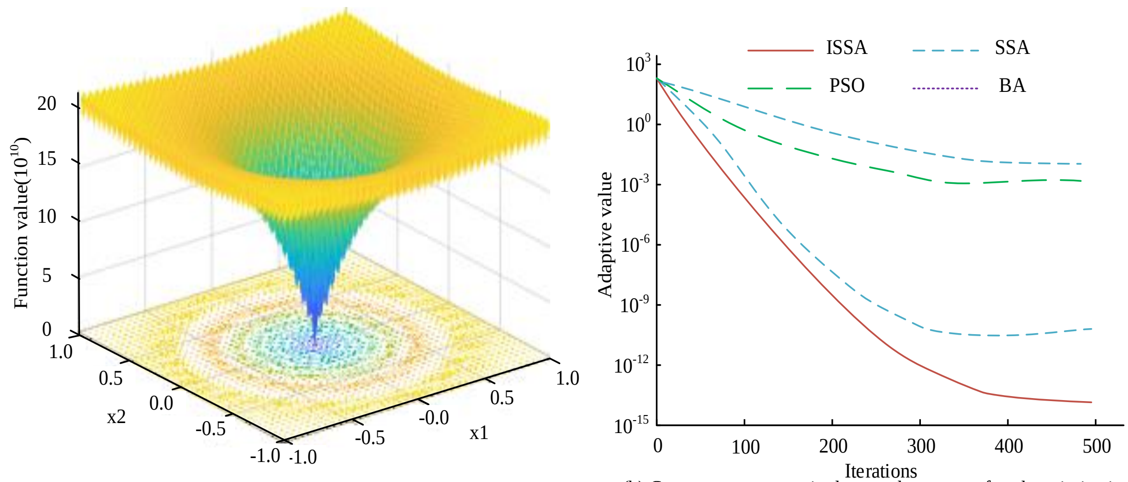
Fig. 7. Mean value and mean square error of the PENALIZED function.



(a) Rosenbrock function 3D stereogram

(b) Convergence curves in the search process of each optimization algorithm

Fig. 8. Optimal convergence curve of each optimization algorithm in the Rosenbrock benchmark function.



(a) Ackley function 3D stereogram

(b) Convergence curves in the search process of each optimization algorithm

Fig. 9. Optimal convergence curve of each optimization algorithm in the Ackley benchmark function.

B. Empirical Experiment on Multimedia Network DFM

To validate the effectiveness of the multimedia network DFM, a performance comparison experiment was conducted using the ImageNet dataset. This dataset contains multimedia network data such as images and audio. The comparison models are data fusion methods on the ground of Adaptive Fusion Steiner Tree (AFST), image correlation based data fusion methods, and distributed compression based data fusion methods. The performance comparison indicators are the Peak signal-to-noise ratio (PSNR), mean square error (MSE), Precision Recall (PR) curve, F1 value, accuracy, and running speed of the fused image. The experimental environment is Windows 10 and the development language is Jupyter Notebook. The PSNR and MSE comparison of each DFM are showcased in Fig. 10.

Fig. 10(a) shows the PSNR comparison results of various DFM. Fig. 10(a) showcases that the PSNR value of the proposed DFM is higher than other comparison models, with an average PSNR value of 37.5dB, which is 4.8dB higher than the ASFT DFM and has higher transmission image quality. Fig. 9(b) shows the comparison results of the MSE of each DFM, as shown in Fig. 10(b). The MSE curve of the proposed DFM is generally lower than other comparison models, with a MSE of 0.52, which is 0.19 smaller than the ASFT DFM. On the ground of the above results, it can be concluded that the proposed DFM has better output data quality and reliability. The PR curve and F1 value comparison are shown in Fig. 11.

Fig. 11(a) showcases the accuracy recall curves of each DFM, as shown in Fig. 11(a). The PR curve offline area of the proposed data fusion method is 0.83, the PR curve offline area of the DFM on the ground of image difference is 0.76, the PR curve offline area of the ASFT based DFM is 0.71, and the PR curve offline area of the distributed compression based DFM is 0.67. The DFM possesses a larger offline area of the PR curve and better performance. Fig. 11(b) showcases the comparison results of F1 values for various DFM. As shown in Fig. 11(b), the F1 value of the proposed data fusion method is 0.37, which is 0.05 higher than the F1 value of the ASFT based DFM, and exceeding the F1 value of other data models, resulting in better performance. The relevant outcomes demonstrate that the DFM has higher offline area and F1 value of the PR curve compared to other algorithms, and the model performance is better. The comparison results of data fusion accuracy and data processing time of each DFM are shown in Fig. 12.

Fig. 12(a) shows the accuracy comparison results of various DFMs. As shown in Fig. 12(a), the accuracy of the data fusion method is 92.4%, which is 13.6% higher than the image difference DFM, 10.1% higher than the distributed compressed DFM, and 7.5% higher than the ASFT DFM. The DFM possesses better data fusion accuracy. Fig. 12(b) showcases the comparison results of data processing time for each DFM, as showcased in Fig. 12(b). The proposed data fusion method has a processing time of 0.1 seconds, which is 0.08 seconds faster than the image difference DFM, 0.04 seconds faster than the distributed compression DFM, and 0.03 seconds faster than the ASFT DFM. Its data processing speed is faster. The relevant outcomes demonstrate that the DFM proposed in the study not only has higher accuracy in data fusion, but also has better data processing efficiency. The comparison of ISSA with other algorithms in key performance indicators is shown in Table I.

As shown in Table I, the improved sparrow search algorithm (ISSA) proposed in this study exhibits significant performance advantages in the multimedia network data fusion system. Specifically, ISSA leads in data fusion efficiency with a processing time of 0.10 seconds, compared to Particle Swarm Optimization (PSO)'s 0.35 seconds, Bat Algorithm (BA)'s 0.28 seconds, and Adaptive Fusion Spanning Tree (AFST)'s 0.42 seconds. In terms of accuracy, ISSA reached 92.43%, surpassing PSO's 85.67%, BA's 87.21%, and AFST's 90.15%. In terms of convergence speed, ISSA can converge after 380 iterations, showing faster convergence compared to PSO's 450 iterations, BA's 500 iterations, AFST's 480 iterations, as well as Grey Wolf Optimization (GSA) and Artificial Bee Colony Algorithm (ALO)'s 550 and 420 iterations. In addition, ISSA ranks first with a score of 8.95 in the global search capability score, further demonstrating its strong ability to avoid local optima and find global optima. The performance comparison between ISSA and traditional methods is shown in Table II.

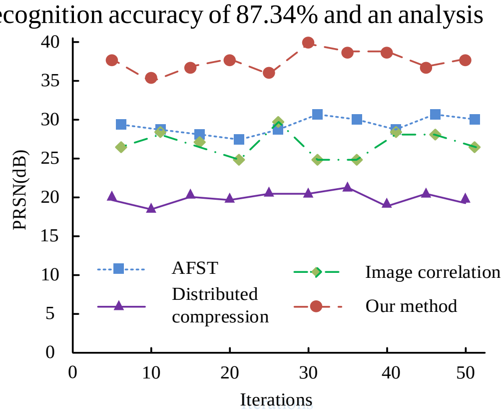
TABLE I. COMPARISON OF MULTIMEDIA NETWORK DATA FUSION MODELS

Model/Algorithm Name	Data Fusion Efficiency (s)	Accuracy (%)	Convergence Speed (Number of Iterations)	Global Search Capability Score (1-10)
ISSA (Proposed in this study)	0.10	92.43	380	8.95
PSO (Particle Swarm Optimization)	0.35	85.67	450	6.87
BA (Bat Algorithm)	0.28	87.21	500	7.12
AFST (Adaptive Fusion Steiner Tree)	0.42	90.15	480	7.56
GSA (Grey Wolf Optimizer)	0.56	88.53	550	5.98
ALO (Artificial Bee Colony Algorithm)	0.29	89.34	420	7.43

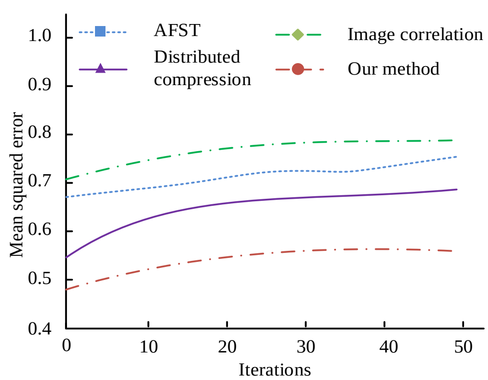
TABLE II. PERFORMANCE COMPARISON BETWEEN ISSA AND TRADITIONAL METHODS

Application Scenario	Task Description	Index	ISSA Performance Metrics	Traditional Method Performance Metrics	Performance Improvement
Video Surveillance Analysis	Real-time activity analysis in surveillance video to detect anomalies	Accuracy	94.56%	89.12%	5.44%
		Response Time	0.25 seconds	0.45 seconds	0.20 seconds
Social Media Content Filtering	Filtering harmful content on social media platforms	Accuracy	93.21%	88.47%	4.74%
		Processing Speed	150 posts/sec	90 posts/sec	60 posts/sec
Medical Image Analysis	Assisted diagnosis to identify abnormal areas in medical imagery	Accuracy	92.78%	87.34%	5.44%
		Analysis Time	0.30 seconds per image	0.60 seconds per image	0.30 seconds
Traffic Flow Management	Real-time traffic flow analysis for optimizing traffic signal control	Optimization Efficiency	95.23%	90.48%	4.75%
		Calculation Time	0.15 seconds per cycle	0.35 seconds per cycle	0.20 seconds per cycle
Environmental Monitoring	Monitoring environmental data to predict pollution events	Prediction Accuracy	91.56%	86.23%	5.33%
		Update Frequency	Every 5 minutes	Every 10 minutes	Doubled

As shown in Table II, ISSA achieved a detection accuracy of 94.56% in video surveillance analysis, with a response time of 0.25 seconds. Compared with the traditional method's accuracy of 89.12% and response time of 0.45 seconds, it shows a 5.44% improvement in accuracy and a 0.20 second reduction in response time. In the social media content filtering task, ISSA achieved a filtering accuracy of 93.21% and a processing speed of 150 posts per second. Compared to the traditional method's accuracy of 88.47% and processing speed of 90 posts per second, ISSA improved accuracy by 4.74% and processing speed by 60 posts per second. In the field of medical image analysis, ISSA has a recognition accuracy of 92.78% and an analysis time of 0.30 seconds per image, while traditional methods have a recognition accuracy of 87.34% and an analysis time of 0.60 seconds per image. ISSA has improved accuracy by 5.44% and reduced analysis time by 0.30 seconds. In traffic flow management, the optimization efficiency of ISSA is 95.23%, with a calculation time of 0.15 seconds per cycle. Compared with the traditional method's 90.48% optimization efficiency and 0.35 seconds calculation time, the efficiency has increased by 4.75%, and the calculation time has been reduced by 0.20 seconds per cycle. In environmental monitoring, the prediction accuracy of ISSA is 91.56%, with an update frequency of every 5 minutes, while the prediction accuracy of traditional methods is 86.23%, with an update frequency of every 10 minutes. ISSA has improved accuracy by 5.33% and doubled the update frequency.



(a) The PRSN results for each algorithm



(b) The SSIM results for each algorithm

Fig. 10. PSNR and MSE of each DFM.

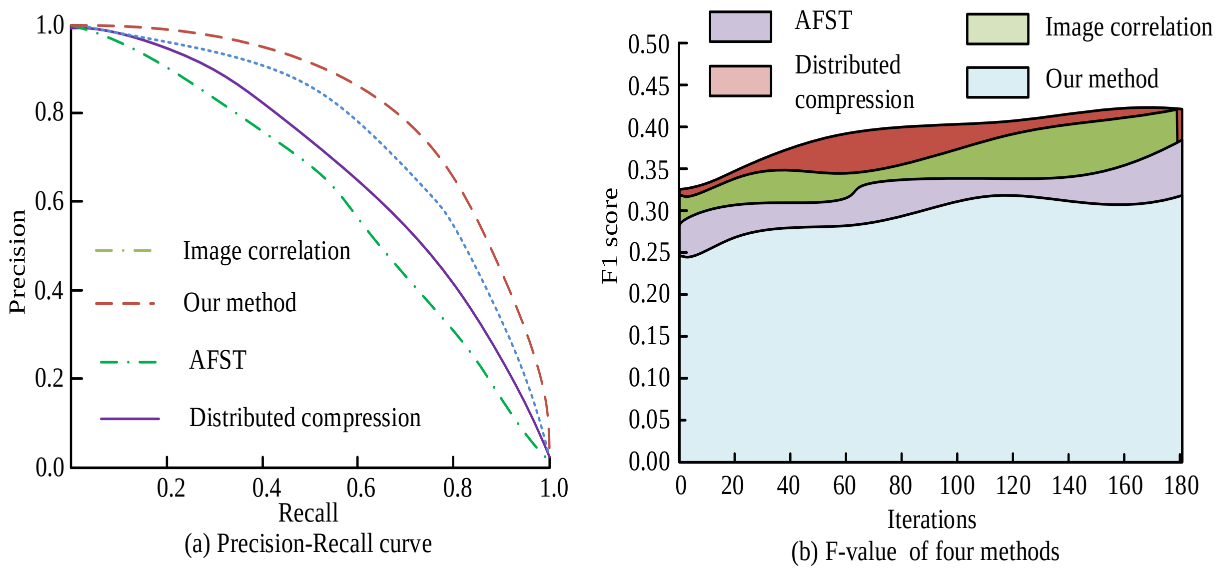


Fig. 11. Comparison results of the PR curves and F1 values of each DFM.

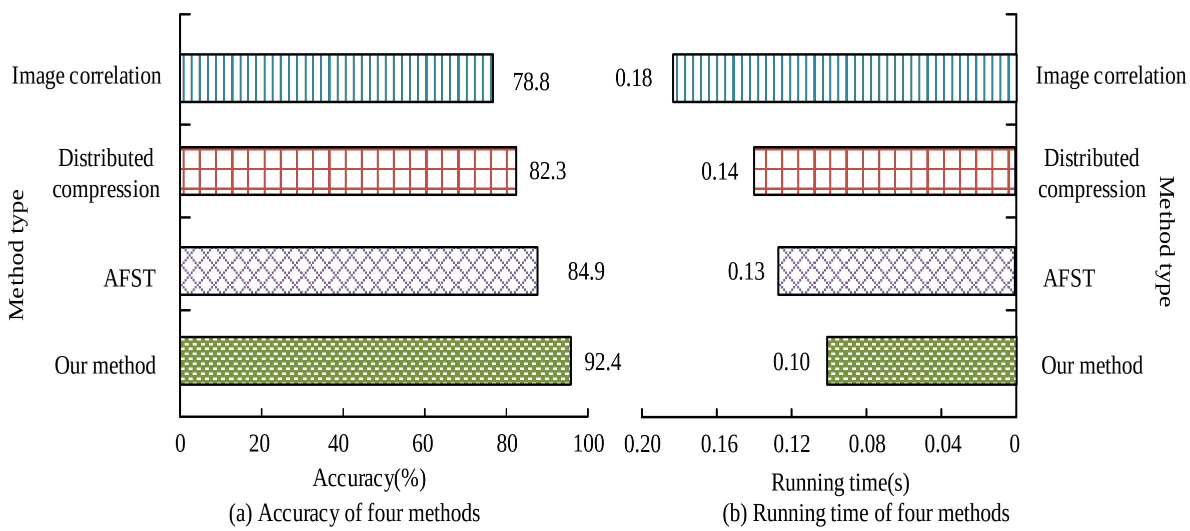


Fig. 12. Comparison results of data fusion accuracy and data processing time of each DFM.

V. CONCLUSION

With the continuous progress of science and technology, the current multimedia network data fusion system is no longer able to satisfy the requirements of today's society. For the problem of insufficient quality and efficiency in traditional multimedia network data fusion systems, this study proposes to construct an ISSA on the ground of reinforcement learning algorithm and SSA, and combines this algorithm with cluster algorithms to construct a multimedia network DFM. The study conducted empirical experiments on the ISSA and multimedia network DFM, and found that in the unimodal function benchmark test, the ISSA required 380 iterations to enter the convergence state, which is lower than other comparative algorithms. However, in the multimodal benchmark function, the ISSA still did not fall into the local optimal situation at 500 iterations. In addition, in the empirical experiment of the multimedia network DFM, it was found that the PRNS value of the output image data of the model was 37.5dB, with a MSE of 0.52. The area under the PR curve is 0.83, the F1 value is 0.37, the accuracy is 92.4%, and the processing time is 0.1s. Its output data quality and data processing efficiency are better than other comparative models. Based on the above results, it can be concluded that the proposed ISSA exhibits significant advantages in data fusion efficiency, accuracy, convergence speed, and global search capability compared to traditional methods in the field of multimedia network data fusion. The high efficiency of ISSA is particularly suitable for real-time application scenarios such as video surveillance analysis and social media content filtering that require quick response. Its high accuracy is crucial for high-precision fields such as medical imaging analysis, as it can provide reliable data analysis results to support clinical decision-making. In addition, ISSA has demonstrated rapid convergence and powerful global search capabilities in complex optimization problems such as traffic flow management and environmental monitoring that require quick adaptation to changes and finding optimal solutions. Although this study mainly focuses on algorithm development and evaluation, the excellent performance of ISSA lays the foundation for its deployment in practical applications. However, this study still has certain limitations, as it did not consider the energy consumption cost of data fusion in practical applications. Future research directions could apply the ISSA algorithm to geological data analysis to improve the exploration efficiency of underground resources such as minerals, oil, and natural gas. By processing and analyzing a large amount of exploration data such as seismic, geological, and geochemical data, the ISSA algorithm is expected to optimize the accuracy and speed of resource localization. Meanwhile, in the field of environmental science, ISSA algorithm can be used to analyze remote sensing data, monitor environmental changes such as deforestation, urban expansion, and climate change. Its efficient data processing capability helps to quickly identify and respond to environmental issues.

REFERENCES

[1] R. Huang, H. He, X. Zhao, Y. Wang, and M. Li, "Battery health-aware and naturalistic data-driven energy management for hybrid electric bus based on TD3 deep reinforcement learning algorithm, " Applied Energy, vol 321, no 1, pp. 1-15, 2022.

[2] Lee J H, Park J, Bennis M, Ko Y C. Integrating LEO Satellites and Multi-UAV Reinforcement Learning for Hybrid FSO/RF Non-Terrestrial Networks. IEEE Transactions on Vehicular Technology, 2023, 72(3):3647-3662.

[3] Abedin S F, Mahmood A, Tran N H, Han Z, Gidlund M. Elastic O-RAN Slicing for Industrial Monitoring and Control: A Distributed Matching Game and Deep Reinforcement Learning Approach. IEEE Transactions on Vehicular Technology, 2022, 71(10):10808-10822.

[4] Liu P, Zhou J, Lv J. Exploring the first-move balance point of Go-Moku based on reinforcement learning and Monte Carlo tree search. Knowledge-Based Systems, 2023, 261(15):1-11.

[5] Xu P, Wang B, Zhang Y, Wang B, Zhu H. Online topology-based voltage regulation: A computational performance enhanced algorithm based on deep reinforcement learning. IET Generation, Transmission & Distribution, 2022, 16(24):4879-4892.

[6] Zhu Y, Yousefi N. Optimal parameter identification of PEMFC stacks using Adaptive Sparrow Search Algorithm. International Journal of Hydrogen Energy, 2021, 46(14):9541-9552.

[7] Wang X, Liu J. Intelligent prediction of fatigue life of natural rubber considering strain ratio effect. Fatigue & Fracture of Engineering Materials and Structures, 2023, 46(5):1687-1703.

[8] Meng S, Zhu Q, Xia F. Research on parameter optimisation of dynamic priority scheduling algorithm based on improved reinforcement learning. IET Generation, Transmission & Distribution, 2020, 14(16):3171-3178.

[9] Nobre F, Heckman C. Learning to calibrate: Reinforcement learning for guided calibration of visual–inertial rigs. The International Journal of Robotics Research, 2019, 38(12-13):1388-1402.

[10] Mai Z, Xiong H, Yang G, Zhu W, He F, Bian R. Mobile target localization and tracking techniques in harsh environment utilizing adaptive multimodal data fusion. IET Communications, 2021, 15(5):736-747.

[11] Sharma P, Anand R S. Depth Data and Fusion of Feature Descriptors for Static Gesture Recognition. IET Image Processing, 2020, 14(5):909-920.

[12] W Y H, Liu P, H C C, Z B L, J Q B. Displacement reconstruction of beams subjected to moving load using data fusion of acceleration and strain response. Engineering structures, 2022, 268(1):1-13.

[13] De Menezes R P, Salarolli P F, Batista L G, Furtado H S, Cuadros M. A.S L. Slopping index for LD converters based on sound and image data fusion by fuzzy Kalman filter. Ironmaking & Steelmaking, 2022, 49(2):178-188.

[14] Tian W, Liao Z, Zhang Z, Wu H, Xin K. Flooding and Overflow Mitigation Using Deep Reinforcement Learning Based on Koopman Operator of Urban Drainage Systems. Water Resources Research, 2022, 58(7):1-29.

[15] Wu X, Chen H, Wang J. Adaptive Stock Trading Strategies with Deep Reinforcement Learning Methods. Information Sciences, 2020, 538:142-158.

[16] Xiao M, Xie W, Fang C, Wang S, Li Y, Liu S. Distribution line parameter estimation driven by probabilistic data fusion of D-PMU and AMI. IET generation, transmission & distribution, 2021, 15(20):2883-2892.

[17] Liu B, Zhan X, Gao Y. Investigation on the Commonality and Consistency among Data Fusion Algorithms with Unknown Cross-covariances and an Improved Algorithm. Advances in Space Research, 2021, 67(7):2044-2057.

[18] Fang Y, Luo B, Zhao T, He D, Jiang B, Liu Q. ST-SIGMA: Spatio-temporal semantics and interaction graph aggregation for multi-agent perception and trajectory forecasting. CAAI Transactions on Intelligence Technology, 2022, 7(4):744-757.