International Journal for Housing Science and Its Applications
Publish August 4, 2025. Volume 46, Issue 3 Pages 1950-1960

https://doi.org/10.70517/ijhsa463154

Research on the Influence of Western Oil Painting on Chinese Painting Based on Association Rule Mining Algorithm in the Perspective of Cross-cultural Art Communication

Anxiang Fan[1,*]
[1] Hubei Institute of Fine Arts, Academy of Fine Arts, Wuhan, Hubei, 430205, China
Corresponding authors: (e-mail: 15629597327@163.com).

Abstract In today's deepening globalization, arts and crafts teaching is facing unprecedented challenges and opportunities. Cross-cultural art exchange is not only a trend, but also a necessity. The confidence and support of association rules affect the quality of mining, in order to improve the mining efficiency, this paper proposes an algorithm based on improved gravitational search mixed with particle swarm (GSA-PSO) for association rule mining. When mining association rules, support, confidence and boost are selected as the evaluation criteria of association rules and the optimal rules are selected using Pareto method. Through simulation tests on nine datasets of different sizes, the experiments prove that the algorithm proposed in this paper has a significant advantage over the other five algorithms in terms of the quality of the rules obtained. Using the dataset of Chinese and Western painting images for analysis, based on the Sankey diagram composed of the association relationship between Western-Chinese paintings, the adaptive threshold mining results of Western-Chinese painting types such as {spatial perspective-artistic conception} are strong association rules with the maximum enhancement greater than 1, but the enhancement for the fixed threshold {300m, 1h} is less than 1.

Index Terms association rule algorithm, particle swarm algorithm, data mining; painting influence

I. Introduction

In the wave of globalization, the world art has also realized multicultural exchanges, which not only promotes the generation of different art schools, but also promotes the continuous innovation and development of traditional art forms [1], [2]. Oil painting, as a type of art with a long history, has a long history of development in China, and each of its distinctive schools of oil painting has a positive reference effect on the development of Chinese painting [3]-[5]. Due to the relationship between the different regional cultures of China and the West, Chinese paintings and Western oil paintings have their own characteristics in terms of aesthetic orientation, painting materials and modeling methods [6], [7]. What embodies the cultural psychology and aesthetic consciousness of the Chinese nation is the Chinese painting's pictorial painting, while the product of the Western cultural context is the oil painting [8].

The pictorial painting of Chinese painting and the abstract painting of western oil painting are the same expression techniques, they have the same characteristics to a certain extent, and both pictorial painting and modern western abstract painting embody the transcendent, free and easy-going spirit of painters [9]-[11]. At the same time, Chinese painting and Impressionism share similar beauty, both are abstract ways to express the inner world of the author [12], [13]. Nowadays, with the mutual penetration and continuous exchange between Eastern and Western cultures, these two different ways of painting have their mutual integration, which is worth exploring. Therefore, Chinese artists, when facing the new opportunities for the development of oil painting, actively explore the influence of Western oil painting on Chinese painting, which points out the direction for the development of Chinese painting and promotes the progress of Chinese painting art creation [14]-[16].

Particle swarm optimization algorithm is combined with gravitational search algorithm to propose an algorithm based on improved gravitational search mixed with particle swarm for association rule mining, which is good at avoiding falling into local optimal solutions and improving the search performance of the algorithm. In order to comprehensively evaluate the performance of this algorithm mining, support, confidence and enhancement are selected as evaluation criteria, and the number of association rules mined by different algorithms on different datasets with data coverage is recorded. Case studies were conducted using Chinese and Western painting datasets, combining the painting features color, line, visual, spatial and other factors in the art of painting, and using the method of this paper, a network capable of describing the correlation mechanism between Western oil paintings and Chinese paintings was constructed.

II. Association rule data mining based on improved particle swarm optimization algorithm

II. A. Association Rule Mining

II. A. 1) Basic concepts of association rules

Some related concepts and definitions of association rules are as follows:

Items and item sets: assume that the set of attributes appearing in a transaction database D is $I=\{I_1,I_2,\ldots,I_i,\ldots,I_k\}$, and that each attribute I_i is referred to as an item, and that the collection of I consisting of the data items is the data item set. If there are k items in an item set, it is called $k-$ item set.

Transactions and transaction datasets: in a given transaction database D, the set of data items denoted by $T=\{t_1,t_2,\ldots,t_j,\ldots,t_d\}$ is called a transaction T, where t_j in each transaction is a itemset I a subset of $t_j \in I$. Each transaction has a unique identifier TID corresponding to it one-to-one, and all transactions as a whole constitute the entire transaction dataset, $D=\{T_1,T_2,\ldots,T_m\}$.

Association rule: an item set to item set implication is used to represent the relationship between the items in the data set, which is an association rule [17]. For an association rule $A \to B, A$, both A and B denote itemsets and must satisfy the following conditions $\{A,B \mid A\in I, B\in I, A\cap B=\varnothing\}$, and A and B are referred to as the association rule preterms and postterms, respectively.

II. A. 2) Evaluation methods for correlation rules

In association rules, there are various ways to evaluate their quality.

Support: $Sup(A\to B)$ denotes the percentage of occurrences of the itemsets A, B together in the transaction database D, which is defined as shown in Equation (1):

$$Sup(A\to B)=\frac{P(A\cup B)}{|D|} \tag{1}$$

where $P(A\cup B)$ denotes the number of times A and B appear in the database at the same time, and $|D|$ denotes the number of entries of all transactions in the current database.

Confidence: $\mathrm{Conf}(A\to B)$ denotes the probability of the occurrence of the itemset B if the itemset A already exists, defined as shown in equation (2):

$$Conf(A\to B)=\frac{P(A\cup B)}{P(A)} \tag{2}$$

where $P(A\cup B)$ denotes the number of times A and B appear in the database at the same time, and $P(A)$ denotes the number of times A appears in the database.

In addition to these traditional support-confidence evaluation methods, other scholars have proposed some new evaluation metrics based on them, which are more diversified to evaluate the association rules.

Lift: $Lift(A\to B)$ is a metric for evaluating the validity of a prediction model, and is the correlation coefficient, defined as shown in equation (3):

$$Lift(A\to B)=\frac{Conf(A\to B)}{Sup(B)} \tag{3}$$

Bounded by 1, when $Lift>1$, it means that the association rule $A\to B$ is positively correlated, and A promotes B; when $Lift<1$, it means that the association rule is negatively correlated, and A inhibits B; and $Lift=1$, it means that the antecedents and successors of the association rule are independent of each other.

Certainty Factor: abbreviated as CF, is used to de-evaluate the probability of the occurrence of the successor in the case of the occurrence of the prior in the association rule, with the value range of $[-1,1]$, defined as shown in Eq. (4):

$$CF=\begin{cases}\dfrac{Conf(A\to B)-Sup(B)}{1-Sup(B)}, Conf(A\to B)>Sup(B)\\ \dfrac{Conf(A\to B)-Sup(B)}{Sup(B)}, Conf(A\to B)<Sup(B)\\ 0, Conf(A\to B)=Sup(B)\end{cases} \tag{4}$$

$CF > 0$, the association rule $A \to B$ is positively correlated, and the closer it is to 1, the higher the confidence level of the rule is; $CF < 0$, it is negatively correlated; and $CF = 0$, it means that the precedence and the successor of the association rule are independent of each other.

The enhancement and certainty can be used in correcting some misleading rules in the strong association rules and avoiding to produce some numerically correct but practically unrelated association rules.

The number of terms NU of an association rule can also be used as an evaluation criterion for simplicity, defined as shown in Equation (5):

$$NU(A \to B) = NUM(A) + NUM(B) \tag{5}$$

where $NUM(A)$ and $NUM(B)$ denote the number of terms in the antecedent and the consequent, respectively, and a smaller number of terms means that the rule is more concise and easier to understand.

II. B. Multi-objective optimization

II. B. 1) Multi-objective optimization problems

In general, a multi-objective optimization problem can be expressed as finding a $\vec{x} = [x_1, x_2, \ldots, x_n]^T$ vector that satisfies the following m inequalities and optimizes k objective functions, with an expression as shown in Eq. (6):

$$\begin{aligned} &Manimize/Maximize\, f(x) = [f_1(x), f_2(x), \ldots, f_k(x)] \\ &subjecttoc(\vec{x}) = [c_1(\vec{x}), c_2(\vec{x}), \ldots, c_m(\vec{x})] \geq 0 \end{aligned} \tag{6}$$

These conditions form a feasible domain R^d containing the desired feasible solution, d denotes the dimension of the feasible domain, and $\vec{x}$ denotes an optimized solution in the feasible domain.

Let $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ be the two feasible solutions of a multi-objective optimization problem if x and y satisfy the conditions as shown in Equation (7):

$$\begin{aligned} f_i(x) \leq f_i(y), \forall i = \{1, 2, \cdots, m\} \\ f_i(x) < f_i(y), \exists i = \{1, 2, \cdots, m\} \end{aligned} \tag{7}$$

Then one can say that x Pareto dominates y, denoted $x \prec y$.

If a feasible solution $x^* = (x_1, x_2, \ldots, x_n)$ in the optimization problem satisfies equation (8):

$$\neg \exists x : x \prec x^* \tag{8}$$

Then, x^* can be said to be a Pareto optimal solution.

In a multi-objective optimization problem there are multiple Pareto optimal solutions, and these optimal solutions mailed together form a Pareto front, also known as a non-dominated solution set.

II. B. 2) Multi-objective optimization methods

In solving multi-objective optimization problems, the multi-objective optimization problem is often transformed into a single-objective problem to be solved, from which the already mature single-objective optimization can be used to solve multi-objective problems. Among them, the more traditional and commonly used methods can be divided into the following categories:

1. Linear weighting method: this method is based on the importance of each optimization objective to give different weighting coefficients, the multi-objective problem is transformed into a linear combination of optimization of single-objective optimization problem, weighting method as shown in Equation (9):

$$\begin{aligned} &Manimize/Maximize\, f(x) = \sum_{i=1}^{K} \omega_i * f_i(x) \\ &\omega_1 + \omega_2 + \ldots + \omega_K = 1, \omega_i > 0 \end{aligned} \tag{9}$$

Where ω_i is the weight coefficient.

2. Constraint method: this method is to use one of the many optimization objectives as the solution of the single-objective optimization problem, and the rest of the objectives as the constraints of this optimization problem, and the constraint method is shown in Eq. (10):

$$\begin{gathered} Manimize/Maximize\ f(x) = f_k(x), k \in [1, K] \\ f_i(x) \le \varepsilon_i, i \in [1, K] \cup i \ne k \end{gathered} \tag{10}$$

where ε_i is a constraint parameter, by setting different values for ε_i, different solutions can be obtained.

3. Objective planning method: this method is to set up an expectation value for each objective in the multi-objective, and to find the nearest optimal solution according to the setting of this expectation value [18]. According to the difference of this expectation value method, it can be divided into: weighted objective planning method, maximum and minimum objective planning method and so on.

II. C.Association rule mining based on GSA-PSO

II. C. 1) Encoding of GSA-PSO algorithm in association rules

PSO optimization algorithm is mainly for the continuous space function optimization of the problem of search operations, in the face of the many discrete optimization problems in practice, the proposed discrete binary particle swarm optimization algorithm, which the position of each particle in the algorithm with the traditional discrete method of binary coding in the form of representation. In the data mining preprocessing stage of traditional association rules, the data is binary encoded with 0s and 1s in the data transactionalization process, thus indicating whether the number of transactions occurs or not. The solution space of the association rule mining domain is the entire database, which belongs to the discrete domain, so in this paper we choose to use the Hollond method of encoding. That is, a particle represents an association rule, and if there are N transactions in the transaction set, the position of the particle is represented by an N-dimensional array, where each item has 2 parts and each part has two possible values, 0 and 1. The first part, if 1, indicates whether the transaction appears in the rule or not, and vice versa. The second part indicates whether the transaction is the first or the second of the rule, if 1 is the former and if the value of the item is the latter the rule is 0. For example, in the position of a particle, if the value of Ii is 00 or 01, it means that the item does not exist in this rule. If the value of Ii is 10, it means that the item is a posterior, and if the value is 11, it means that the ith item is a prior. The original scanning speed of the database has been greatly improved, and the efficiency of support and confidence calculation has also been improved a lot.

II. C. 2) Construction of the fitness function

The specific definitions of support and confidence are as follows:

Definition 1: $S(A \to B)$ denotes the percentage of transactions of A and B in the whole database, expressed mathematically:

$$\begin{aligned} & S(A \to B) = P(A \cup B) \\ & = \frac{\text{Number of transactions containing } A \text{ and } B}{\text{Number of transactions contained in the entire database}} \end{aligned} \tag{11}$$

Definition 2: $C(A \to B)$ denotes the percentage of all transactions that contain A that also contain B, expressed mathematically:

$$\begin{aligned} & C(A \to B) = P(A/B) = \frac{P(A \cup B)}{P(A)} \\ & = \frac{\text{Number of transactions containing } A \text{ and } B}{\text{Number of transactions in the entire database containing } A} \end{aligned} \tag{12}$$

So the confidence and support are utilized to construct the fitness function by combining the support with the confidence, multiplying it with the influence factors a and b, respectively, and then summing the results, and the final sum obtained is used as the fitness value.

For particle x, the fitness function is defined as follows:

$$F(x) = aS(x) + bC(x) \tag{13}$$

where a, b are the parameters of support and confidence in the fitness function satisfying $0 \le a \le 1$, $0 \le b \le 1$, and $a + b = 1, S(x)$, with $S(x)$ and $C(x)$ denoting the support and confidence, respectively. When a is 0 it means that the rule contains only confidence, and the algorithm may only produce similar rules with high confidence but low support, which is meaningless in practice. Similarly, when b is 0, it means that the rule is expressed only in terms of support, so some rules with low support but high confidence may be missed, which are often very valuable.

In the particle swarm algorithm, each particle has its own velocity, position and adaptation value. Let the particle swarm include M particles, the velocity of the particle at the moment of t is v_{id}, the position is $pbest$, the best position of the individual is $pbest_{id}$, and the best position of the group is $gbest_{id}$. The formula for updating the v_{id} of a particle with $pbest$ is expressed as:

$$v_{id}(t+1)=v_{id}(t)+c1r1\left(pbest_{id}-x_{id}\right)+c2r2\left(gbest_{id}-x_{id}\right) \tag{14}$$

$$x_{id}(t+1)=x_{id}(t)+v_{id}(t+1) \tag{15}$$

The optimized particle swarm update formula is:

$$v_{id}(t+1)=r1v_{id}(t)+c1r2a_{id}(t)+c2r3\left(gbest_{id}-x_{id}\right) \tag{16}$$

$$x_{id}(t+1)=x_{id}(t)+v_{id}(t+1) \tag{17}$$

The binary particle swarm algorithm has allowed the position update formula to be improved to conform to the binary nature. The specific formulas are as follows:

$$v_{id}(t+1)=r1v_{id}(t)+c1r2a_{id}(t)+c2r3\left(gbest_{id}-x_{id}\right) \tag{18}$$

$$x_{id}=\begin{cases}1, r4<sig\left(v_{id}(t+1)\right)\\ 0, r4\geq sig\left(v_{id}(t+1)\right)\end{cases} \tag{19}$$

where, $i=1,2\cdots N$, denotes the population size, $sig(x)=1\big/\left(1+e^{-x}\right)$, $d=1,2\cdots D$, denotes the particle dimension, $r1$, $r2$, $r3$, $r4$ are random numbers between $(0,1)$, $c1$, $c2$ are the learning factors, sig the range of values of the function is $(0,1)$, and since on the interval $[-10,10]$ the value of the sig function takes values within $(0,1)$, the independent variable x, i.e., the range of the $v_{id}(t+1)$ is on the $[-10,10]$, and so the velocities $v_{\max}$ of the particles are all less than 10. The velocity $v_{id}(t+1)$ determines the maximum distance that the particle can move in an iteration, and the larger the velocity, the closer the algorithm is to a local search near the current optimal solution. Moreover, $v_{\max}$ cannot exceed the limited range of the particle. Too large a $v_{\max}$ and the particle may cross the optimal solution, thus missing the opportunity of localized search. And too small will cause the global search ability of the particle will be reduced, making it fall into the local optimum.

II. C. 3) GSA-PSO based mining association rule process

The specific implementation steps for extracting rules with the GSA-PSO algorithm are as follows:

Input: total number of particle swarms n, learning factors $c1, c2$ maximum number of iterations D, influence factors a, b, minimum support and minimum confidence. Output: optimal association rules.

1: Initialization. Analyze the real problem given by modeling and encode the particles to obtain the support Sup and confidence $conf$. Individual best position of particles in the particle swarm is $pbest$ and the population best position is $gbest$.

2: Calculate the fitness value Fitness for each particle in the current population.

3: Start updating the optimal value of each particle, both individually and as a whole, and update the individual extreme value of the particle if the extreme fitness of the particle is less than the current fitness. The opposite is the same.

4: Record and update the velocity and position of each particle according to steps 1 and 2 so that new populations are created continuously.

5: Determine whether the algorithm meets the termination conditions, if so, perform 6, otherwise perform 2.

6: Extract the association rules.

II. D.Experimental results and analysis

II. D. 1) Data set selection

In order to better compare the performance of GSA-PSO with several other algorithms under different sized datasets, nine different real datasets from UCI repositories are selected in this paper. These datasets cover different engineering domains and have different number of transactions and dimensions.According to the different sizes, the datasets in this paper are divided into three different types Iris, Flag, and Cancer are small-scale datasets, German, Car, and CMC are medium-scale datasets, and Abalone, Wine, and Mushroom are large-scale datasets.

II. D. 2) Experimental results and analysis

In this section, the GSA-PSO algorithm is tested with SSA, WOA, PSO, ALO, and SMA algorithms, on nine datasets, and each algorithm is run five times, and the average value is used to analyze the results. The application of GSA-PSO on association rule mining will be analyzed from multiple perspectives of support, confidence, and lift, as well as coverage, number of rules, and execution time. The results of the small-scale dataset tests are shown in Tables 1 and 2.

It can be observed that on smaller scale datasets, such as on the Iris dataset, the GSA-PSO algorithm does not show significant superiority with other algorithms. On these datasets, it achieves only a moderate level in terms of three different objective function values. The relatively low problem complexity of the smaller datasets leads to insignificant performance differences between the six algorithms.The SMA algorithm falls into local optimal solutions early on and does not fully explore the entire search space, which makes it less efficient in utilizing the dataset. GSA-PSO, on the other hand, performs relatively evenly on all three datasets and maintains its objective function, although it obtains the most association rules on the Flag dataset.

Table 1: Small data set fitness function values

Algorithm	Iris			Flag		
	Sup	Conf	Lift	Sup	Conf	Lift
GSA-PSO	0.08	0.8	42.31	0.12	0.78	82.38
SSA	0.08	0.74	21.1	0.15	0.73	47
ALO	0.08	0.79	30.18	0.14	0.77	66.64
PSO	0.08	0.75	39.61	0.17	0.71	71.07
WOA	0.06	0.75	63.83	0.11	0.76	219.75
SMA	0.04	0.72	26.16	0.06	0.8	59.63

Table 2: Rule number on small data sets

Algorithm	Iris		Flag	
	Mean	Std	Mean	Std
GSA-PSO	7.9	2.79	35.55	6.93
SSA	7.9	2.48	23.7	2.02
ALO	7.1	2.13	22.15	1.13
PSO	10.1	4.17	13.8	4.91
WOA	8.3	3.23	11.56	2.88
SMA	34.1	47.07	29.17	17.82

Table 3: Medium scale data set fitness function value

Algorithm	German			CMC		
	Sup	Conf	Lift	Sup	Conf	Lift
GSA-PSO	0.28	0.84	78.99	0.16	0.9	224.1
SSA	0.25	0.82	71.98	0.15	0.83	89.8
ALO	0.24	0.76	35.55	0.17	0.88	107.45
PSO	0.22	0.7	40.36	0.07	0.77	198.14
WOA	0.24	0.77	185.13	0.13	0.79	198.36
SMA	0.18	0.75	81.88	0.13	0.82	507.45

The results of the tests on the medium-sized datasets are shown in Tables 3 and 4. While on the two medium datasets, GSA-PSO obtains the highest number of rules on German and CMC data at the same time most of the three objective functions achieve the highest values.SMA's lift on CMC dataset is almost tens of times that of other algorithms, but its support is very low, almost half of that of the other algorithms, which implies that it obtains the rules most likely to be meaningless, because the there are many more solutions with high confidence and high lift relative to those with high support. On this basis WOA and GSA-PSO perform better on both datasets with all three objective function values remaining at a more balanced level. As can be seen from the table, with the gradual increase in the size of the dataset, the advantages of GSA-PSO are revealed, compared to other algorithms, GSA-PSO maintains a careful development while fully exploring the data.

Table 4: Rule number of medium scale data set

Algorithm	German		CMC	
	Mean	Std	Mean	Std
GSA-PSO	41.5	7.5	29.3	5.72
SSA	31.5	6.66	25.5	4.7
ALO	40.9	9.79	27.3	3.55
PSO	30.1	8.69	22.9	8.83
WOA	27.5	7.82	28.1	6.17
SMA	26.3	7.04	18.7	5.24

The results of the large-scale dataset tests are shown in Tables 5 and 6. Finally, on the 2 large-scale datasets, GSA-PSO obtains the highest support and lift on Abalone and Mushroom, and leads the other algorithms by several orders of magnitude, and only the confidence level is slightly lower than the other algorithms on Abalone and Wine datasets. Compared to SSA, GSA-PSO leads almost across the board, both in terms of the number of rules and the objective function value, which implies that GSA-PSO is better than SSA in terms of both exploration and exploitation. Both ALO and PSO algorithms perform mediocrely on the nine datasets, and although they do not fall into extremes as easily as SMA, their exploration and exploitation capabilities are only of average caliber. So when analyzed in combination with the nine datasets of different sizes, overall GSA-PSO performs better, especially on the larger datasets showing more obvious advantages.

Table 5: Large data set fitness function values

Algorithm	Abalone			Mushroom		
	Sup	Conf	Lift	Sup	Conf	Lift
GSA-PSO	0.1	0.91	1031.66	0.43	0.89	68.78
SSA	0.09	0.89	938.33	0.39	0.85	51.06
ALO	0.08	0.82	767.24	0.39	0.87	68.28
PSO	0.06	0.85	483.55	0.35	0.81	37.46
WOA	0.07	0.97	451.08	0.42	0.84	55.31
SMA	0.05	0.98	266.64	0.31	0.72	29.28

Table 6: The number of rules on large data sets

Algorithm	Abalone		Mushroom	
	Mean	Std	Mean	Std
GSA-PSO	29.7	9.13	35.7	4.51
SSA	25.9	10.67	37.1	3.5
ALO	26.9	5.52	35.7	6.72
PSO	26.9	7.76	30.7	4.37
WOA	37.3	21.93	32.1	5.36
SMA	48.9	38.66	16.5	6.85

In order to further verify that the improvement of PSO algorithm in this paper is effective, the six algorithms will be analyzed in the following in terms of data coverage and time overhead, as shown in Fig. 1-Fig. 6. From the figure, it can be found that relative to the gap between the objective function value and the number of rules of the six algorithms on different datasets, except for PSO and SMA which are in a relative disadvantage, the gap between the algorithms in terms of the coverage on each dataset is not very obvious, almost at the same level. The coverage of GSA-PSO is almost the same relative to that of SSA, and the two algorithms have a high utilization rate of the dataset. From the coverage rate, it can be seen that the Pareto frontier solution sets obtained by several algorithms have excellent utilization of the dataset, almost all of them have more than one hundred percent, but combined with the analysis of the number of rules mined and the value of the objective function, GSA-PSO not only has a wide coverage of the dataset, but also carries out a more detailed development, so that the quality of the rules obtained is better, while maintaining a high coverage rate.

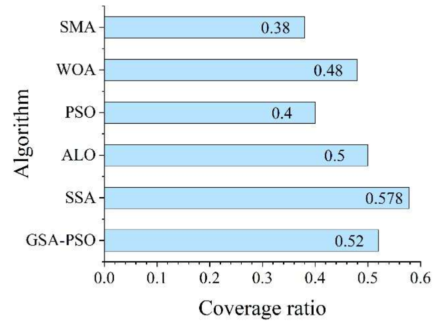

Figure 1: Iris Dataset Coverage

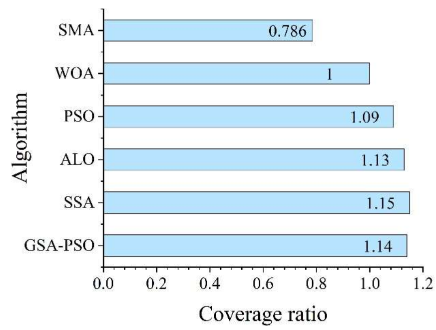

Figure 2: Flag Dataset Coverage

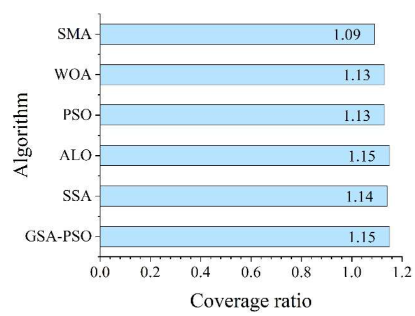

Figure 3: German Dataset Coverage

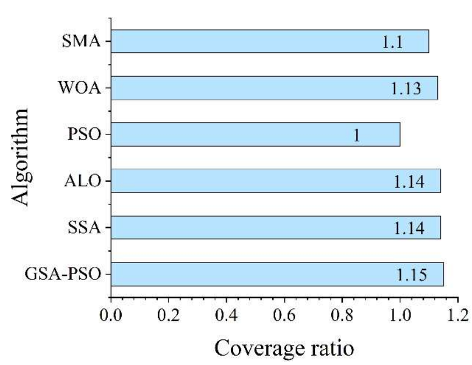

Figure 4: CMC Dataset Coverage

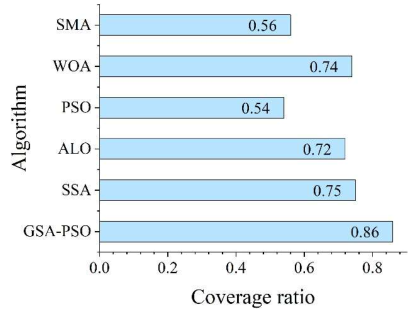

Figure 5: Abalone Dataset Coverage

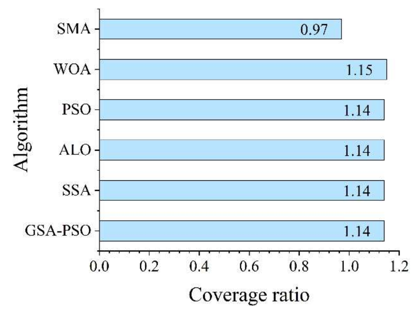

Figure 6 Mushroom Dataset Coverag

III. Mining the Relevance of Western Oil Painting and Chinese Painting

III. A. Sample data sets

A simulated sample dataset containing 2,321 samples was randomly generated through survey research and reference to Western and Chinese painting art.

III. A. 1) Constructing a sample dataset for painting feature simulation

Taking image data as the experimental object, image data including western oil painting and Chinese painting samples were collected, and nine categories of data characteristic of painting samples were extracted from the data, including line, color, composition, technique, visual elements, style, emotion, light and shadow, and color level.

III. A. 2) Data pre-processing

In this simulation experiment, data cleaning, data reduction, data conversion and other preprocessing operations are performed on the original data first. The data preprocessing operations include data cleaning, data reduction, and data conversion. After data processing, the processed database ends up with 1358 entries.

III. B. Analysis of the Influence of Western Oil Painting on Chinese Painting

The number of iterations of the algorithm is set to 20, the initial population size is set to 20, the minimum confidence level of the GSA-PSO algorithm is set to 10%, and the minimum support level is 1%. The autocorrelation rules are mined one by one for the characteristics of Chinese and Western painting art in the data. Take “color”, “line”, “vision”, “style” as an example, the mining results are shown in Table 7. The results are shown in Table 7.

A total of 12 strong association rules are obtained from association rule mining, including 4 spatial perspective, 4 art concepts and 4 humanism. In order to facilitate the display of association rule mining results with adaptive threshold selection, this paper takes art type and painting type as nodes, and elevation as the width of the connecting line, eliminates the weak association relationships with elevation less than one, and constructs the Sankey diagram, as shown in Fig. 7.

In order to verify the effectiveness of the algorithm, with reference to existing studies, the same data with thresholds of {300m, 1h} and {500m, 2h} are selected for spatio-temporal association rule mining with fixed spatial and temporal thresholds, and the results of spatio-temporal association rule mining with adaptive thresholds are analyzed in comparison with those of spatio-temporal association rule mining with adaptive thresholds, and the results are shown in Fig. 8. As can be seen from the figure, for the identified Western-Chinese painting types, the enhancement degree of adaptive threshold selection is improved compared to the fixed threshold in all cases. The adaptive threshold mining results for Western-Chinese painting types such as {spatial perspective-artistic concept} are strong association rules, with the maximum enhancement greater than 1, but the enhancement of fixed threshold {300m, 1h} is less than 1. This indicates that the traditional spatial-temporal association rule mining may not reach the maximum enhancement due to the subjective determination of a single threshold, which may miss some important rules not easily detected in the dataset. The rules are not easy to find but extremely important in the dataset. Therefore, compared with the traditional spatio-temporal association rule mining, the adaptive threshold selection association rule mining, by determining the most important temporal and spatial thresholds, results in a larger number of mining results, more comprehensive, and does not miss some important but hidden association rules in the transaction set, which verifies the effectiveness of the GSA-PSO algorithm.

Table 7: The adaptive threshold association rules mining results

Type	Art class	Optimum time threshold(h)	Optimum space threshold(m)	Maximum ascension
Space perspective	Color feature	5	1320	0.8149
	Line sketch	5	2080	0.8555
	Visual element	7	3543	0.8419
	Space image	3	732	1.9133
Art reading	Color feature	2	284	5.2591
	Line sketch	4	1610	1.4499
	Visual element	8	1516	1.1131
	Space image	3	2593	0.8719
Humanistic spirit	Color feature	3	595	7.4972
	Line sketch	6	1375	1.0763
	Visual element	7	1735	1.3414
	Space image	4	2560	0.8832

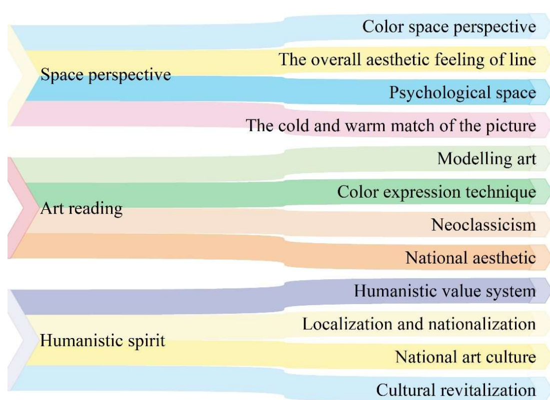

Figure 7: The relationship between art and painting

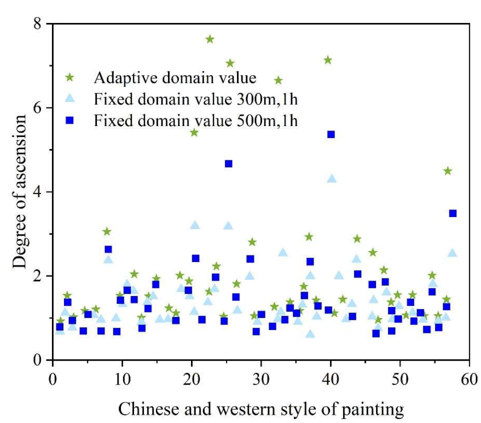

Figure 8: Improvement analysis

IV. Conclusion

In this paper, the GSA-PSO algorithm is used for association rule extraction, and through the application of this algorithm in mining the influence of western oil painting on Chinese painting, the simulation analysis is carried out with the data set of Chinese and western art paintings as the research object, and the support, confidence and enhancement are selected as the evaluation criteria of association rules to compare the performance of association rules with other algorithms, and the performance is compared with that of the traditional spatio-temporal association rule mining algorithms. Comparative analysis is carried out to verify the effectiveness of the algorithm. According to the results of this paper, the GSA-PSO algorithm has excellent performance advantages compared with other algorithms, and the results of adaptive threshold mining of {spatial perspective-artistic concepts} between Western and Chinese paintings are strong association rules, and the maximum degree of enhancement is more than one.

References

[1] He, M., Barkeshli, M., & Toroghi, R. M. (2024). CULTURAL DIPLOMACY IN ARTS EDUCATION AND OIL PAINTING: A SYSTEMATIC REVIEW OF EAST-WEST FUSION ARTISTIC EXPRESSIONS. Arts Educa, 41.

[2] Wu, S. (2019). Integration of China and the West: Diversified Expression of Contemporary Chinese Local Oil Painting. Open Access Library Journal, 6(10), 1.

[3] Yang, G. (2021). Inheritance and development of national elements in contemporary Chinese, American and Russian oil paintings. Linguistics and Culture Review, 1-15.

[4] Xinyuan, S. (2025). Exploring the Inheritance of Aesthetic Thinking in Chinese Contemporary Oil Painting. International Journal of Advanced Culture Technology, 13(1), 396-402.

[5] Zixuan, G. (2022). A Study of the Ethnicity of Chinese Oil Painting. Frontiers in Art Research, 4(15).

[6] Fan, L. (2024). The Expressionist Characteristics of Chinese Local Oil Painting Integrating the Freehand Brushwork of Traditional Chinese Painting. Art and Society, 3(1), 81-88.

[7] An, Z., & Nie, D. (2023). Impact Of Chinese Cultural Elements On Spiritual Expression In Oil Painting Imagery Landscape: An Exploratory Study In Arts Education. Arts Educa, 37.

[8] Feng, J. (2023). The Spirit of Freehand Brushwork in Chinese Oil Painting. Pacific International Journal, 6(1), 56-60.

[9] Yang, G. (2021). The imagery and abstraction trend of Chinese contemporary oil painting. Linguistics and Culture Review, 454-471.

[10] Xuan, Q. (2025). A Study of Imparities between Chinese Painting and Western Painting from Perspectives of Arts and Cultures. Journal of Literature and Arts Research, 2(1), 141-146.

[11] Liu, W. (2021). Analysis on the Collision and Fusion of Eastern and Western Paintings in the Context of Globalization. Thought, 7(8).

[12] Fan, Z. B., Zhu, Y. X., Marković, S., & Zhang, K. (2023). A comparative study of oil paintings and Chinese ink paintings on composition. The Visual Computer, 39(4), 1323-1334.

[13] Li, L. (2023). Multiple expression forms and their aesthetic trend of modern Chinese image oil painting. Frontiers in Art Research, 5(18).

[14] Sun, S., & Wang, K. (2023). An Introduction to the Influence of Neo-expressionism on Contemporary Chinese Oil Painting. Journal of Humanities, Arts and Social Science, 7(8).

[15] Hou, T., & Moyao, W. G. (2023). The Influence of Western Art Style on Oil Painting Education in Chinese Contemporary Art Colleges and Universities. The Educational Review, USA, 7(10), 1610-1614.

[16] Qu, L., & Liu, H. (2025). Research on the Development of Chinese Painting Under the Influence of Western Art Thought in the 1980s. Humanities and Social Science Research, 8(1), p104-p104.

[17] Haixing Wang,Longtao Guo,Hong Yin,Yuefeng Huang & Shimeng Li. (2025). Association Mining for Operation and Maintenance Safety Risks of EMUs Based on Unstructured Event Data. Urban Rail Transit,(prepublish),1-12.

[18] Davide La Torre & Hatem Masri. (2025). Multiple objective programming and goal programming: making better decisions with artificial intelligence and business analytics. Annals of Operations Research,346(2),1-2.