International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 6544-6553

https://doi.org/10.70517/ijhsa463562

The trend of applying data mining methods to analyze the combination of fine art and philosophy in artistic creation in a big data environment

Xin Liu[1,*]
[1] School of Plastic Arts, Dankook University, Yongin, 17113, Korea
Corresponding authors: (e-mail: liuxin1991929@163.com).

Abstract With the development of big data technology, data mining methods provide a new way to combine the study of fine art and philosophy. In this study, K-means clustering algorithm and support vector machine model are used to extract features and classify the dataset containing 2,713 samples of typical philosophical fine arts works in China and 3,671 fine arts works integrating philosophical ideas from abroad. The study shows that the performance of K-means clustering after feature extraction is significantly better than other algorithms, with a clustering accuracy of 89.16% and a profile coefficient of 40.4. In terms of philosophical-emotional feature mining, the Pearson's correlation coefficient of the support vector machine model in the utility prediction is 0.676, and the average absolute error is only 0.113, which is superior to the comparative models such as CNN and LSTM. The study classifies works of fine art into four major philosophical clusters: existentialism, metaphysics, social criticism and discernment, and perceptual experience, and predicts and analyzes the affective characteristics of each type of work. A survey of the number of participants in the International Society for the Philosophy of Works of Art 2020-2024 shows that there is a growing trend towards the integration of fine art and philosophy. The study provides data support and theoretical framework for deepening the combination of fine art creation and philosophical thought, which is of great significance in promoting the multidimensional development of art creation.

Index Terms Big data, data mining, K-means clustering, support vector machine, fine art works, philosophical emotions

I. Introduction

As artistic creations continue to change, the market value and size of the market for art works continues to rise. This is due to the fact that the demand for digital art and traditional art works has increased and become more demanding, and more and more art works focus on the deep meaning and connotation of the works in order to convey important ideas and messages, resulting in a gradual increase in the proportion of art creations with philosophical themes [1]-[4]. In the development of human civilization, art and philosophy have always been two inseparable fields, and the importance of the two to human society and individuals is self-evident, they can stimulate people's thinking and emotions, and enrich people's spiritual world [5], [6]. Art as a source of aesthetic experience, through the viewing and appreciation of art works, people can experience the value of beauty and be inspired at the level of sensibility, while philosophy is to think and explore the values from the perspective of rationality, and through the philosophical discourse and argumentation, people can search for the truth and wisdom, and the combination of the two can provide people with a comprehensive and multi-dimensional path of thinking, prompting people to form a more in-depth view of life and the world. The combination of the two can provide people with a comprehensive and multi-dimensional path of thinking, prompting people to form a more in-depth outlook on life and the world [7]-[10]. In addition, as a form of artistic expression, fine art is loved by people for its unique visual language and artistic expression. Through the use of color, lines, shapes and other elements, as well as the visual composition and expressive display, it conveys the creator's emotions and thoughts, and contains a variety of art forms such as painting, sculpture, photography, etc., each of which has its own unique characteristics and rules [11]-[13].

Nowadays, the digital resources of art continue to climb, art creation has entered a new stage, and the attention to injecting philosophy into fine art works is increasing [14], [15]. But the research on art creation in the integration of fine art and philosophy is still in the blank stage. And data mining can reveal implicit, previously unknown and potentially valuable information from a large amount of data in databases, providing technical support for revealing the trend of combining fine art and philosophy [16].

With the rapid development of information technology and computer technology, the application of big data analysis methods in the field of art is gradually deepened. As an important carrier for human beings to express their emotions and thoughts, artistic creation has long maintained a close connection with philosophical thought. However, the analysis of this connection often stays at the level of qualitative research and lacks effective quantitative analysis methods. Based on the big data environment, this study applies data mining technology to systematically analyze the embodiment of philosophical thoughts in fine art works, and explores the development trend of the combination of fine art and philosophy, which is of great academic value and practical significance. Existing studies show that the philosophical connotation of fine art works is multidimensional and complex, and traditional analysis methods are difficult to comprehensively capture its characteristics. Data mining is able to extract the implied philosophical features and emotional expression patterns from a large number of art works by combining unsupervised and supervised learning. In this study, K-means clustering algorithm and support vector machine model are chosen as the core technical tools for the cluster analysis of philosophical features and emotion recognition of art works, respectively. K-means algorithm is suitable for dealing with unlabeled dataset of art works due to its high efficiency and stability, and it calculates the sample similarity through the Euclidean distance to realize the automatic clustering of philosophical features. For the nonlinear classification problem, the support vector machine model with radial basis kernel function is used for the multi-classification recognition of philosophical sentiment of art works, and the complete binary tree SVM is constructed to realize the accurate classification of sentiment. In order to comprehensively evaluate the model performance, the study designs evaluation indexes including clustering accuracy and variance ratio criterion index, and compares and analyzes them with other mainstream algorithms. The data sources include the dataset of typical philosophical art works in China and the dataset of art works integrating philosophical ideas from abroad, and the validity of the analysis is ensured by data preprocessing. This study not only focuses on the philosophical clustering of art works, but also analyzes the emotional expression characteristics of each type of works to explore the influence of different philosophical ideas on art creation. Through the investigation of the number of participants in the International Society for Philosophy of Art Works and the degree of combination of art works and philosophical ideas, the study reveals the development trend of the combination of art and philosophy. This study aims to construct a framework for analyzing the characteristics of philosophical analysis of art works through data mining methods, to provide data support and theoretical guidance for the multidimensional development of art creation, to promote the in-depth integration of art and philosophy, and to inject new vitality into art creation.

II. Methodology

II. A.K-means algorithm

The K-means algorithm is described as shown in Fig. 1, where the cluster form center is generally defined as the mean of the points within the cluster for the specific operation.The K-means algorithm is an unsupervised learning algorithm because it can be trained on unlabeled datasets.

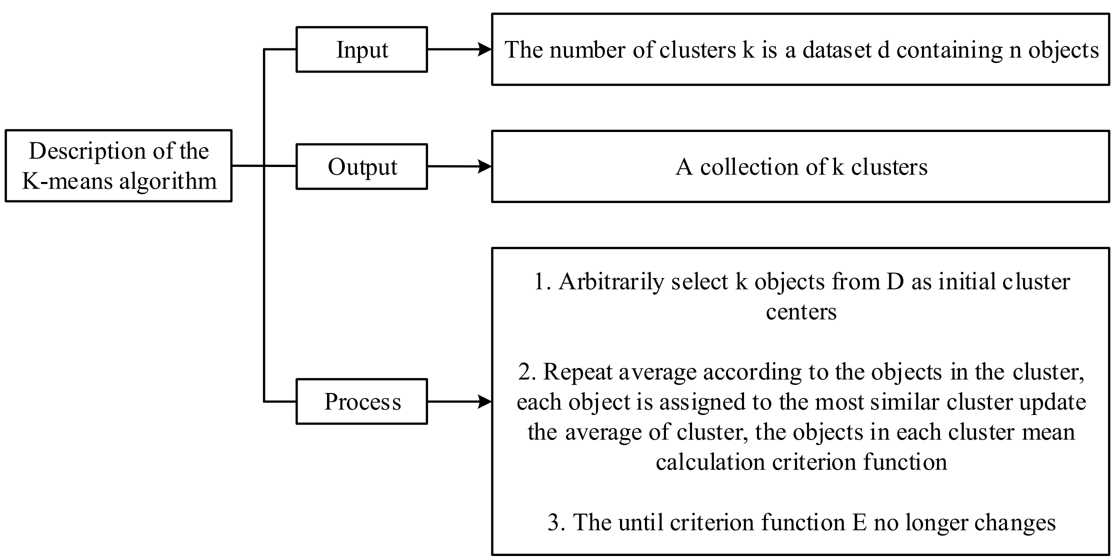

Figure 1: K-means algorithm

II. A. 1) Clustering distance approach selection

In the K-means algorithm [17], the plotted data set is divided into K clusters based on the size of the distance between the samples, so that the points in the clusters are as closely connected as possible, and the distance between the points in the clusters is as large as possible, and the available distance metrics are: based on the Euclidean distance, based on the Manhattan distance, and based on the Chebyshev distance.

In this paper, Euclidean distance is chosen, which is used to calculate the similarity or distance between data samples, and then determine the size of the probability that they belong to the same cluster. For any two points on the plane, the Euclidean distance d between them is given in Eq:

$$d_{XY} = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{1}$$

where x_i and y_i denote the coordinates of any point.

II. A. 2) Determination of K-value

When the basic information in a set of data tables is unknown, because the classification principle and the number of classifications will seriously affect the difference in the classification effect of the entire data table, so the K-means clustering algorithm on the choice of K value has: elbow method and contour coefficient method of two methods.

The measure used in the elbow method is SSE (Sum Of The Squared Errors), when K is smaller than the true number of clusters, as K increases, it will greatly improve the degree of aggregation between classes, SSE will fall dramatically, when K reaches the true number of clusters, as K increases, the degree of aggregation between classes will not be greatly improved, SSE will not fall very much, the K-SSE fold plot looks like an elbow.

When using the contour coefficient, the formula for calculating the contour coefficient is as follows, assuming first that the sample set has been divided into K clusters:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \tag{2}$$

where $a(i)$ is the average distance of the sample point $x(i)$ to other sample points of the same cluster as cohesion. $b(i)$ is the average distance of sample point $x(i)$ to all samples in the nearest cluster as the separation degree. argmin denotes the value of the variable when the function reaches the minimum value, p is a sample in any cluster C_k, then the nearest cluster is defined as follows:

$$C_j = \arg\min(C_k)\frac{1}{n}\sum_{p \in C_k} | p - x_i |^2 \tag{3}$$

After evaluating the distance of a point to a cluster by calculating the average distance between x_i and all the samples within that cluster, the closest cluster for each sample point is determined, and when the profile coefficients of all the sample points x_i have been computed, the average profile coefficient is the arithmetic mean of the profile coefficients computed for all the sample points x_i, where $-1 \le s(i) \le 1$. Obviously, according to the wheelhouse coefficient formula, the lower the cohesion, the higher the separation, the better the clustering effect, and the larger the average contour coefficient. Therefore, the clustering effect is best at the value of K corresponding to the time when the average wheelhouse coefficient reaches its maximum value.

II. A. 3) Specific calculation steps

The specific computational steps of the K-means clustering algorithm are as follows:

The first step is to first create a collection of n data samples containing Ω:

$$\Omega = \{x_i \mid x_i = (x_{i1}, x_{i2}, \cdots, x_{id}), i = 1, 2, \cdots, n\} \tag{4}$$

where $x_i = (x_{i1}, x_{i2}, \cdots, x_{id})$ is a d-dimensional vector denoting the d distinct attributes of the i th data, and n is the sample size. The clustering center is as follows:

$$C = \{c_j \mid c_j = (c_{j1}, c_{j2}, \cdots, c_{jd}), j = 1, 2, \cdots, k\} \tag{5}$$

where $c_j = (c_{j1}, c_{j2}, \cdots, c_{jd})$ is the centroid of the j th cluster class, each centroid c_j contains d different attributes and k is the number of cluster classes.

The second step is to choose an appropriate k-value. For the choice of K value there are generally two cases, the first is directly given the K value, and the second is to determine the K value that makes the clustering effect optimal through various model analyses.

The third step is to calculate the distance between the sample and each clustering center or the distance between two data x_i and c_j.

The fourth step is to recalculate the cluster centers for the same cluster class as c_j after classifying the sample data with the following expression:

$$c_{jl} = \frac{1}{N(\phi_j)} \sum_{x_i \in \phi_j} x_{il}, l = 1,2,\cdots.d; j = 1,2,\cdots,k \tag{6}$$

where $N(\phi_j)$ is the number of data in the same cluster class, $x_i = (x_{i1}, x_{i2}, \cdots, x_{id})$.

The fifth step is to determine whether the K-means run can be stopped, generally the clustering center is no longer changing or the clustering results converge, generally using the criterion function, which is usually defined in terms of the sum of squares of the errors, and theoretically the smaller its value is, the better the clustering effect is represented. The expression is as follows:

$$E = \sum_{j=3}^{k} \sum_{x_i \in \phi_j} dis(x_i, c_j) \tag{7}$$

where E is the sum of squared errors for all objects in the dataset, $dis(x_i, c_j)$ is the distance between the two datasets x_i and c_j, x_i is the point in the space denoting the given data object, c_j is the average of the centroids of the j th cluster class.

II. B. Expression of Philosophical Emotions in Art Creation

II. B. 1) Nonlinear Support Vector Machines

Since the classification of philosophical-emotional feature data of art works in this paper is a nonlinear problem, the nonlinear support vector machine (SVM) [18], [19] needs to be investigated. For nonlinear classification, the kernel function is utilized to map the data from the low-dimensional space to the high-dimensional space, which is turned into a linear solution. If the mapping to the high-dimensional space is a linearly indivisible problem, it is similar to the soft-interval SVM, which needs to introduce relaxation variables to solve the problem. When solving the optimal classification function, the kernel function is needed to replace the dot product operation, as in Eq:

$$K(\overrightarrow{x_i}, \overrightarrow{x_j}) = \Phi(\overrightarrow{x_i}) \cdot \Phi(\overrightarrow{x_j}) \tag{8}$$

At this point the optimization problem is transformed into a dyadic problem described as Eq:

$$\begin{cases} \max Q(\vec{\alpha}) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j K(\overrightarrow{x_i}, \overrightarrow{x_j}) \\ s.t. \quad \sum_{i=1}^{l} \alpha_i y_i = 0, \quad 0 \le \alpha_i \le C, \quad i = 1,2,\cdots,l \end{cases} \tag{9}$$

The optimal $\overrightarrow{\alpha^*}, \overrightarrow{\omega^*}$ and b^* are obtained, and the optimal classification function is finally solved as Eq:

$$\begin{aligned} f(x) &= \operatorname{sgn}(\overrightarrow{\omega^*} \Phi(\vec{x}) + b^*) \\ &= \operatorname{sgn}(\sum_{i=1}^{l} \alpha_i^* y_i \Phi(\vec{x}_i) \cdot \Phi(\vec{x}) + b^*) \\ &= \operatorname{sgn}(\sum_{i=1}^{l} \alpha_i^* y_i K(\overrightarrow{x_i}, \vec{x}) + b^*) \end{aligned} \tag{10}$$

There are four types of kernel functions in common use today:

(1) Linear kernel function:

$$K(\vec{x}, \overrightarrow{x_i}) = \vec{x} \cdot \overrightarrow{x_i} + c \tag{11}$$

(2) The number of polynomial kernel drawings:

$$K(\vec{x}, \overrightarrow{x_i}) = (s(\vec{x} \cdot \overrightarrow{x_i}) + c)^d \tag{12}$$

(3) RBF (radial basis) kernel function:

$$K(\vec{x}, \overrightarrow{x_i}) = \exp(-\frac{\| \vec{x} - \overrightarrow{x_i} \|^2}{2\sigma^2}) \tag{13}$$

(4) Sigmoid kernel function:

$$K(\vec{x}, \overrightarrow{x_i}) = \tanh(s(\vec{x} \cdot \overrightarrow{x_i}) + c) \tag{14}$$

In the above equation, the kernel parameters are s, c, d, σ, and their selection has a certain impact on the complexity of the classifier and the results of classification. In this paper, RBF is chosen to construct the SVM classifier for two main reasons: compared with the linear kernel function, RBF is able to realize the nonlinear mapping, while the linear kernel function cannot. Moreover, the linear kernel function is a special RBF, and the RBF contains all the functions of the linear kernel function.

In this study, according to the characteristics of the dataset where the number of samples far exceeds the number of feature dimensions, nonlinear SVM is selected and RBF is used as the kernel function of SVM classifier. In the SVM classification algorithm based on RBF, the penalty parameter C and the kernel parameter σ need to be determined to improve the accuracy of the model in recognizing the philosophical sentiment of art works.

II. B. 2) Sentiment Multi-Classification Model Based on Support Vector Machines

Currently, there are two commonly used SVM multi-class classification methods: the direct method and the indirect method. The direct method is to construct a multi-class classification mathematical model. The indirect method refers to the use of multiple SVMs to optimize the combination of multi-classification problems into multiple binary classification problems. The more typical methods include: one-to-one SVM, one-to-many SVM, complete binary tree SVM and so on. In this paper, complete binary tree SVM is chosen to realize the recognition of philosophical emotional features in art works.

In the complete binary tree SVM algorithm, for a $N(N \geq 2)$ class problem, $\log_2 N$ SVM classifiers need to be built. The algorithm first divides all the classes into two categories and builds the first binary classifier. Later, using the "direct method" to binary classify each of the two categories, there are three SVM classifiers to classify the samples into four categories. And so on, the four categories continue to be binary classified until the classification needs are met.

This multiclassification model is based on the emotional psychological model, first divided into two categories of vitality and calmness according to the energy of fine art works, and then divided into two categories of sadness and happiness along according to the stress axis of fine art works, and finally divided the emotional feature data into four categories. The multifaceted philosophical emotional categorization of fine art works is shown in Figure 2.

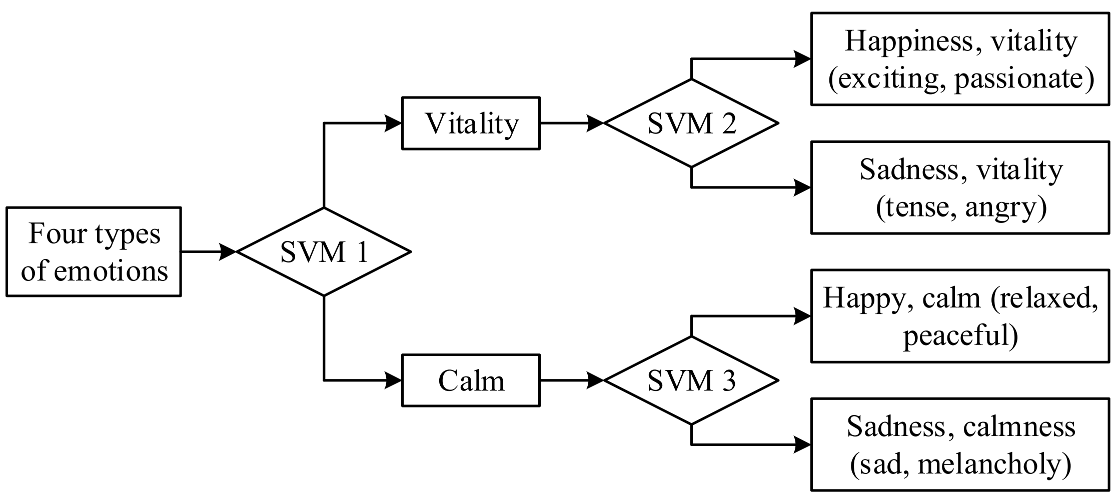

Figure 2: The classification of multi-philosophical feelings of art works

II. C. Experimental design

This section describes the hardware and software environments used in the experiments, the datasets and data processing data, as well as the training process and evaluation metrics.

II. C. 1) Environment setup

An Ubuntu 20.01 server with two NVIDIA RTX3090 (24G) GPUs was chosen as the environment for this experiment. The specific hardware configuration and software environment are as follows:

(1) Hardware Configuration
CPU: Intel(R) Xeon(R) CPU E5-2680 v4@ 2.40GHz
GPU: 4 NVIDIA GeForce RTX 3090
Single GPU video memory: 24G
Memory: 64G
Hard disk: 4TB
(2) Experimental platform software configuration
Operating system: Ubuntu 20.04

CUDA: 11.7
PyTorch: 11.3
Python: 3.8
IDE: Pycharm2024.1.2

II. C. 2) Data sets

In this paper, in addition to the dataset of typical philosophical art works in China, data mining experiments were additionally conducted using data of foreign art works integrating philosophical ideas in order to analyze the cluster class division and emotion recognition of art works combined with philosophical ideas.

The foreign datasets used in this paper are composed of 3,671 pieces originating from Princeton University Art Museum, Smithsonian Freer Gallery, Metropolitan Museum of Art, Harvard University Art Museums, etc., respectively. The dataset of Chinese typical philosophical art works contains 2713 art work samples, each sample is of 256x256 resolution, and some of them are obtained after cropping and rotating process.

II. C. 3) Data processing

The data processing part is standardized by cropping, rotating and compressing. Some of the images in the dataset need to be compressed to a 256×256 resolution version during processing, and the two datasets are input together for training, with a uniform division of 20% of the number as the test set.

III. Results and analysis

III. A. Philosophical Ideas Clustering Performance and Cluster Classification

In addition to the profile coefficients mentioned above, this study proposes two clustering performance metrics applicable to unsupervised clustering performance evaluation:

Clustering Accuracy (CA):

$$CA = \frac{\max_1 + \max_2 + \cdots + \max_i}{L} * 100 \tag{15}$$

In this case, the clustering accuracy was calculated based on the maximum length of the sequence of the most frequently occurring clustering labels in each section. This length is considered as the most representative clustering label in the section, which is considered to be a better match to the true category. Where CA represents the defined unsupervised clustering accuracy, max_i represents the maximum length of the most frequently occurring clustering labels in the i th section, and L denotes the length of the dataset, the clustering accuracy is obtained by dividing the sum of the maximum lengths in the i th section by the dataset and multiplying by 100.

Criterion index of variance ratio (CH)

CH is used to assess the tightness and separation of clusters based on the ratio of the degree of dispersion within a class to the degree of dispersion between classes, and a larger CH index indicates better clustering. The formula for its CH value is as follows:

$$CH = \frac{sum(S_B(c))}{sum(S_W(c))} * \frac{n-k}{k-1} \tag{16}$$

where n is the number of samples, k is the number of clusters, $S_W(c)$ is the average distance of the samples within the cluster, and $S_B(c)$ is the distance between the center of the cluster and the center of the entire data set.

In order to assess the superiority of K-means algorithm in classifying the data on the tendency of fine art works in art creation to express themselves at the philosophical level, the datasets before and after the feature screening are applied to do the clustering study respectively, and the classification effect is observed.

The experiments in this section run the common clustering algorithms Hierarchical Clustering (HCA), Spectral Clustering (Spec), Gaussian Mixture Clustering (GMM), and Unsupervised Extreme Learning Machine Clustering (US-ELM) for 10 times respectively. The number of clusters to be clustered is set to 4 and the resultant data of the run is taken as the average of the 10 runs expressed.

The experimental results of evaluating the clustering effect of different clustering algorithms before and after feature extraction are shown in Table 1. Comparing the clustering effect before and after feature extraction, in the dataset after picking features, all five clustering algorithms have a certain magnitude of improvement in the evaluation index. And through the data in the table, it is found that the K-means clustering algorithm of this paper obviously achieves better clustering results compared with other clustering algorithms in the dataset before feature picking and in the dataset after feature picking. Among them, the CA of K-means clustering algorithm reaches 89.16% on the data after feature extraction, and the SI and CH indices reach 40.4 and 2196, respectively, which demonstrates better clustering performance. In conclusion, the clustering algorithm selected in this paper can mine the multidimensional philosophical features in art works and can improve the clustering performance to a certain extent.

Table 1: The results of the selection of the characteristics were evaluated

Feature selection	Clustering algorithm	CA	SI	CH
Pre	K-means	82.55	37.2	1785
	HCA	76.02	30.1	969
	GMM	75.58	34.5	1442
	Spec	65.13	32.9	1164
	US-ELM	79.42	29.7	1598
Post	K-means	89.16	40.4	2196
	HCA	81.24	35.6	1414
	GMM	80.92	38.2	1858
	Spec	71.52	36.1	1560
	US-ELM	84.05	31.9	1931

Figure 3 shows the results of philosophical embodied clustering division based on K-means clustering algorithm. From the figure, it can be seen that the clustering delineation of the algorithm is better, with a clear hierarchy and no overlapping between the cluster classes. Among them, Cluster 1~Cluster 4 represent the embodiment of existentialist philosophy, metaphysical philosophy, social criticism and discursive philosophy, and perceptual experience philosophy in art works in turn. Taking Cluster 1 as an example, works in this category explore the meaning of individual existence and life, and express loneliness and anxiety through exaggerated art forms, such as Munch's The Scream.

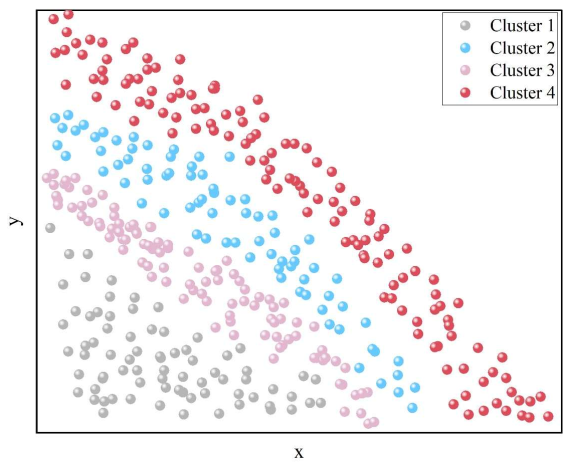

Figure 3: Philosophy embodies the clustering of clustering

III. B. Philosophical Emotional Expression Features Mining

In this section, Pearson's correlation coefficient (r) and mean absolute error (MAE) of potency and arousal are computed as evaluation metrics for the task of philosophical emotion recognition of artworks, respectively. And the performance of different models, including Convolutional Neural Network (CNN), Long Short-Term Memory Network (LSTM), Extreme Gradient Boosting (XGBoost), Random Forest (RF), and in this paper, Support Vector Machine (SVM) models, on the task of predicting emotional arousal and validity is compared.

The Pearson correlation coefficient, a statistical measure of the degree of linear correlation between two variables, and the high correlation indicates the effectiveness of the model in predicting the intensity of philosophical emotions in fine art works, is calculated as follows:

$$r=\frac{\sum(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum(x_i-\overline{x})^2\sum(y_i-\overline{y})^2}} \tag{17}$$

The Mean Absolute Error (MAE) is a measure of the difference between the predicted and actual values, with smaller values indicating that the model's predictions are closer to the true situation, and is calculated as follows:

$$MAE = \frac{\sum |y_i - \hat{y}_i|}{N} \tag{18}$$

where N is the number of music clips, y_i is the true value of the emotion of the music clips, and $\hat{y}_i$ is the predicted value of the emotion of the music clips. x_i and y_i are the ith pair of observations in the sample data, and $\bar{x}$ and $\bar{y}$ are the mean values of x and y, respectively.

Table 2 shows the performance of different models in philosophical sentiment recognition of art works. From the results in the table, the SVM model scores the highest Pearson correlation coefficient, 0.676 and 0.607 respectively, regardless of the Valance value or the Arsoual value. And in terms of the mean absolute error (MAE), the SVM model performs significantly better than the comparative models, with MAEs of 0.113 and 0.121 respectively. The experimental results confirm that the support vector machine's effectiveness in mining the philosophical emotional features of art works.

Table 2: The performance of different models in the analysis of the philosophy of art

Model	Valence		Arousal	
	r	MAE	r	MAE
CNN	0.614	0.184	0.574	0.179
LSTM	0.538	0.155	0.516	0.169
XGBoost	0.532	0.185	0.449	0.144
RF	0.629	0.128	0.589	0.149
SVM	0.676	0.113	0.607	0.121

Figure 4 shows the trend of predicting the main philosophical emotions of the art works based on the support vector machine model for each of the cluster categories of 4 above. Among them, the main philosophical emotions of works in the Existentialism category include loneliness, anxiety, fear, absurdity, and pain. Works in the Metaphysics category include the philosophical emotions of sublimity, awe, serenity, order, and eternity. Works in the social criticism and discursive categories are dominated by philosophical emotions such as anger, grief, irony, oppression, and resistance. Philosophical as well as perceptual-experiential works include philosophical emotions such as transience, immersion, peace, and meditation. From the figure, it can be visualized that the predicted data of this paper's model matches the actual data to a high degree, which proves the effectiveness of the support vector machine model in the task of recognizing philosophical emotions in art works.

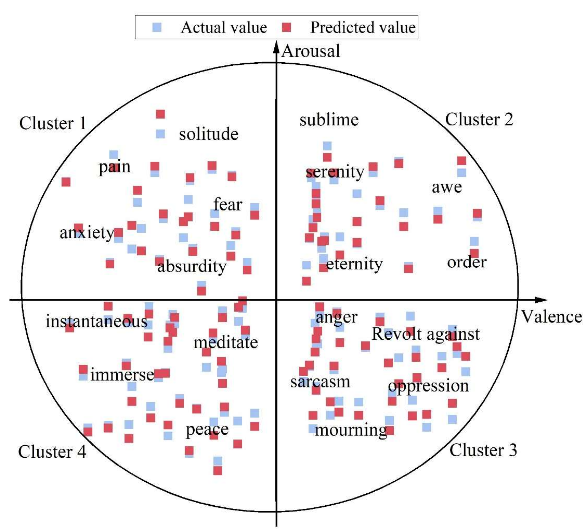

Figure 4: The trend of the main philosophical emotional prediction of art works

III. C. Development trend of the combination of art creation and philosophy

The International Society for the Philosophy of Works of Art is an international philosophical organization that aims to promote philosophical research on works of art and enhance philosophical exchanges. Since its establishment, it has attracted creators of works of art from many countries to participate in the discussion, creating an international academic organization and communication atmosphere with a pluralistic nature for the development of the combination of art and philosophy.

This section investigates the number of attendees from the five major participating member countries of the International Society for Philosophy of Fine Arts in the United States, the United Kingdom, France, Germany, and China during the period of 2020-2024, in order to analyze the development trend of the combination of fine arts creation and philosophy. Figure 5 shows the trend of the participants in the five countries. As can be seen from the figure, the number of participants in each country has been steadily increasing, with the number of participants from the United States, the United Kingdom, France, Germany, and China increasing from 58, 62, 56, 60, and 61 respectively in 2020 to 93, 81, 91, 87, and 91 respectively in 2024. The International Society for the Philosophy of Artwork translates the decentralized practice of combining philosophy and artwork into creative theory through academic seminars, providing a methodological framework for creators of artwork.

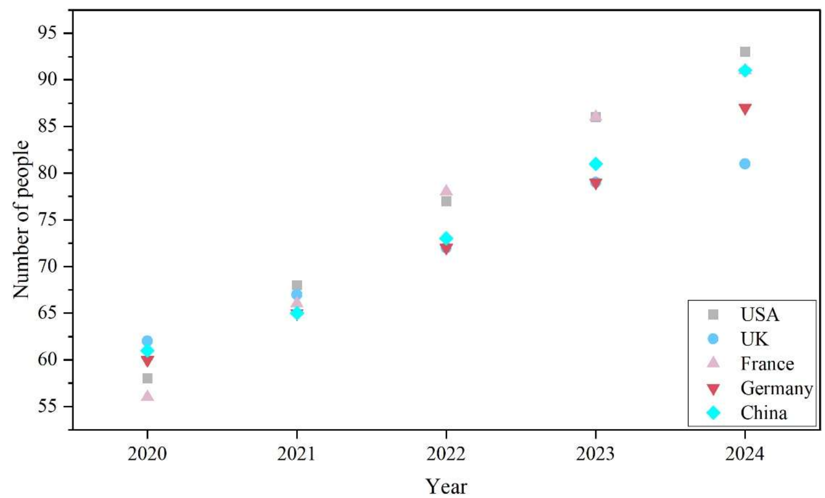

Figure 5: The change of participants

In addition, this section also analyzed the degree to which the 240 works in the art exhibition section of the exchange meeting were combined with philosophical thoughts, and evaluated from seven aspects: the depth of philosophical thoughts, the application of aesthetic theories, the fit between vision and philosophy, the appropriateness of theme expression, multi-dimensional philosophical expression, emotional expression, and public acceptance. The degree of combination was analyzed using the Richter five-point Scale, with scores from 1 to 5 representing "very good", "good", "average", "bad", and "very bad" in sequence. Table 3 presents the survey results on the degree of integration between artworks and philosophical thoughts. As can be seen from the table, among the 240 works, more than 65% of the works scored "average" or above in terms of the depth of philosophical thought, the application of aesthetic theory, the fit between vision and philosophy, the appropriateness of theme expression, multi-dimensional philosophical expression, emotional expression, and public acceptance. Only 13.3% to 31.7% of the works were rated as "bad" or "very bad". The vast majority of artworks have a good combination effect with philosophical thoughts. They not only reflect the pursuit of deep humanistic thinking in artistic creation, but also demonstrate the concrete expression of philosophical thoughts through visual means.

Table 3: The result of a combination of philosophical thought

Dimension	5	4	3	2	1
Philosophical depth	43	59	62	44	32
Aesthetic theory application	27	75	77	35	26
Visual and philosophical relevance	55	66	51	37	31
The topic expresses adaptability	36	84	65	41	14
Multidimensional philosophical expression	108	43	57	21	11
Emotional performance	73	60	55	28	24
Public acceptance	31	46	117	26	20

IV. Conclusion

Based on the big data environment, this study applies K-means clustering algorithm and support vector machine model to analyze the trend of the combination of fine art and philosophy in art creation, and the following conclusions are obtained:

The K-means clustering algorithm optimized by feature extraction performs well in the classification of philosophical features of fine art works, with a clustering accuracy of 89.16%, a profile coefficient of 40.4, and a variance ratio criterion index of 2196, which is significantly better than other clustering algorithms. The art works are successfully categorized into four philosophical clusters: existentialism, metaphysics, social criticism and discernment, and perceptual experience.

In the mining of philosophical emotional expression features, the support vector machine model significantly outperforms comparative models such as CNN and LSTM, and the Pearson correlation coefficient in arousal prediction reaches 0.607, with an average absolute error of only 0.121, which proves the model's effectiveness in the task of philosophical emotion recognition of art works.

By analyzing the data of the International Society for the Philosophy of Works of Art's participation from 2020 to 2024, it was found that the number of participants from various countries showed a steady upward trend, and the number of participants from China increased from 61 in 2020 to 91 in 2024, reflecting the increasing attention to the research on the combination of fine art and philosophy. Finally, the survey on the degree of integration of 240 works with philosophical thoughts shows that 108 works were rated as “very good” in the dimension of multi-dimensional philosophical expression, reflecting the pursuit of deep philosophical thinking in art creation.

Funding

This work was supported by the phased achievements of the major art project: "Research on innovative Methodology of Intelligent Design in the Information Age" (No.22ZD17).

References

[1] Bramantyo, T. (2021). Digital art and the future of traditional arts. Journal of Music Scholarship, 42(1), 96-110.
[2] Gryglewski, P., Ivashko, Y., Chernyshev, D., & Dmytrenko, A. (2020). ART AS A MESSAGE REALIZED THROUGH VARIOUS MEANS OF ARTISTIC EXPRESSION. Art Inquiry, (22).
[3] Al-Obaikan, K. H. (2023). Symbolic Connotations in the Works of the Artist Eugene Delacroix. Al-Academy, (110), 69-90.
[4] Logue, J. (2023). Classics in Western Philosophy of Art: Major Themes and Arguments. Essays in Philosophy, 24(1/2), 121-125.
[5] Irvin, S., & Dodd, J. (2017). In advance of the broken theory: philosophy and contemporary art. The Journal of Aesthetics and Art Criticism, 75(4), 375-386.
[6] Khwaja, A. (2024). Exploring the Intersection of Art and Philosophy: A Journey Through Conceptual and Visual Interpretations. Frontiers in Humanities and Social Research, 1(2), 48-69.
[7] McMahon, J. A. (2019). The ancient quarrel between art and philosophy in contemporary exhibitions of visual art. Curator: The Museum Journal, 62(1), 7-17.
[8] Gordon, R. E. (2020). The Philosophy of Freedom and the History of Art: An Interdisciplinary View. Philosophies, 5(3), 18.
[9] Berger, S. (2017). The art of philosophy: visual thinking in Europe from the late renaissance to the early enlightenment. In The Art of Philosophy. Princeton University Press.
[10] Fernández, M. A. (2025). Art as a Mirror of Cultural Values: A Philosophical Exploration of Aesthetic Expressions. Cultura: International Journal of Philosophy of Culture and Axiology, 22(3).
[11] Abdinabi, A., & Rayhona, T. (2024). Specific Characteristics Of The Foundations Of Artistic Creativity In The Types And Genres Of Fine Art. Web of Teachers: Inderscience Research, 2(3), 219-228.
[12] Urinbaev, I. (2024). THE CONCEPT AND MEANING OF ARTISTIC IMAGE IN FINE ARTS. Science and innovation, 3(C5), 54-61.
[13] Wango, K. (2024). Use of Elements and Application of Principles of Art and Design in the Creation of Pictorial Compositions among Fine Art Students at Kenyatta University, Nairobi, Kenya. East African Journal of Arts and Social Sciences, 7(1), 189-218.
[14] Hegel, G. W. F. (2017). The philosophy of fine art. In Aesthetics (pp. 496-500). Routledge.
[15] Winter, D. (2024). Aesthetic aspects of digital humanism: An aesthetic-philosophical analysis of whether ai can create art. Introduction to digital humanism, 211-224.
[16] Łydżba-Kopczyńska, B. I., & Szwabiński, J. (2021). Attribution markers and data mining in art authentication. Molecules, 27(1), 70.
[17] Duygu Selin Turan & Burak Ordin. (2025). The incremental SMOTE: A new approach based on the incremental k-means algorithm for solving imbalanced data set problem. Information Sciences,711,122103-122103.
[18] Francesca Maggioni & Andrea Spinelli. (2025). A novel robust optimization model for nonlinear Support Vector Machine. European Journal of Operational Research,322(1),237-253.
[19] Bingbing Yu,Bo Wang,Yi Li,Yuantong Zhang & Guohao Wang. (2024). Prediction of blast-hole utilization rate using structured nonlinear support vector machine combined with optimization algorithms. Applied Intelligence,54(19),9136-9157.