Research on the Optimization of Intelligent Scoring of Physical and Chemical Experiment Operation Test in Junior Middle School

ZhiJie Lu[1](✉) and Juan Xu[2]

[1] ShangHai Municipal Educational Examinations Authority, Shanghai 200433, China
lzj@shmeea.edu.cn
[2] Shanghai Xiding Intelligent Technology Co., Ltd., Shanghai 200433, China

Abstract. In order to solve the problem of fairness, time and effort in large-scale on-site scoring of physical and chemical experiment operation test in junior high school, this paper proposes a method based on frame adjustment and particle swarm optimization clustering, and makes a comparative application in the Shanghai physical and chemical experiment operation test system in junior high school. The results show that the new optimization method has significantly improved the accuracy and evaluability of intelligent scoring of experimental operation test. The accuracy of intelligent scoring accuracy of designated physical experiments has been improved by 10–20%. With the improvement of subsequent class libraries and the iteration of algorithms, the effect of relevant models and algorithms is expected to be further improved. This method has practical theoretical guiding significance for the improvement of the existing large-scale scoring of the physical and chemical experiment operation test in junior high school, and also has certain reference value for the machine scoring research of other experimental operation tests.

Keywords: Experiment Operation Examination · Intelligent Scoring · Model and Algorithm Optimization

1 Foreword

In order to implement the requirements of the National Opinions on Strengthening and Improving Experimental Teaching in Primary and Secondary Schools [1], Shanghai formulated the Measures for the Implementation of the Shanghai Junior High School Academic Level Test[2] in April 2019, which specifies that comprehensive test subjects should be added to the junior high school academic level test, including physics, chemistry, interdisciplinary case analysis and physical and chemical experiment operation test, in which the physical and chemical experiment operation test should be tested in the laboratory on-site actual operation mode, The test is divided into sections and sections, with a full score of 15 points (including 10 points for the physical experiment operation test and 5 points for the chemical experiment operation test), and the test duration is 15 min for each section. In order to reduce the chance of the experiment operation test, the

C. F. Peng et al. (Eds.): EIMT 2023, AHSSEH 8, pp. 1151–1165, 2023.
https://doi.org/10.2991/978-94-6463-192-0_150

implementation plan specifies that each candidate needs to complete two consecutive physical and two chemical experiment operation tests. The total score of 15 points in the exam will be recorded into the total score of the candidates in the midterm examination.

Many scholars have also done some research on the experimental operation test. Cao Jingjing [3] mentioned in the physical and chemical experiment operation test of junior high school in Shanxi Province that the "24 out 1" method was used to solve the complexity of the test organization. Zhuo Min and Zhang Dongmei [4] believe that there are four deficiencies in the current way, including the low level of standardization construction of the experimental examination room, the need for further improvement of the test content, the need to strengthen the professional needs of the evaluators, and the need to further improve the rigor of the on-site evaluation process. At the same time, the time, personnel the financial investment and efficiency are compared. Three suggestions are put forward to strengthen the construction of standardized examination rooms, strengthen the construction of examination room monitoring system and improve equipment standards. However, the above research mainly summarized and analyzed the physical and chemical experiment operation test from the macro level, and did not propose effective solutions to the problems and difficulties in the core link of the scoring process.

Based on the above background, the author of this paper proposed an intelligent scoring model based on frame adjustment and particle swarm optimization algorithm optimization through comprehensive analysis and comparison of the four technical support companies participating in the Shanghai Physical and Chemical Experiment Operation Test, and compared it with the existing scoring application system. The results show that the accuracy and evaluability of intelligent scoring for the specified experiment have been significantly improved.

2 Basic Principles and Models of Intelligent Scoring in Physical and Chemical Experiment Operation Test of Junior High School

2.1 Basic Principles of Intelligent Scoring in Physical and Chemical Experiment Operation Test of Junior High School

Intelligent scoring is mainly used to shoot the whole physical and chemical experiment operation test process through two or more cameras installed in different positions on the experimental platform, form multi-angle video data, and judge the correctness of the operation of the evaluation points (scoring points) in the experiment process through special scoring principles and algorithms, and give corresponding scores. According to different experiments, the principle of intelligent scoring is slightly different. Generally, there are four basic principles and algorithms for comparison and judgment as shown in Table 1.

In the physical and chemical experiment operation test, because of the differences in scene, light, equipment, content, location and interference factors, it is still difficult to accurately judge the correctness of each operation or operation process of each candidate. As a technical supporter, the adoption of different models and algorithms must ensure sufficient accuracy, and also consider the research and development costs. Therefore,

Table 1. Principle and algorithm of intelligent scoring comparison and judgment in physical and chemical experiment operation test of junior high school

number	Judgment principle	Classification		Algorithm
1	Image recognition judgment	Single view	single-graph	R-CNN
		Multi-view diagram	multi-graph	SPP-N
				Fast R-CNN
				Faster R-CNN
				YOLO
				SSD
2	Action recognition and judgment	Single view	single-view	CNN
				CNN&RNN
				3D CNN
				Factor-ized spatio-temporal CNNs
				LSTM
				SNN
				DBN
				DTD&DNN
				P-CNN
		Multi-view	multi-view	MOCAP
		Depth and RGB data	RGB-D	RNN&LSTM
3	Action + image judgment	Single view + single view	S-S	Composition algorithm
		Single view + multiple views	S-M	
		Multi-view + single view	M-S	
		Multi-view + multi-view	M-M	
		RBG + Multigraph	R-M	
4	Auxiliary sensor	Single sensor	S-Sensor	Composition algorithm
		Multi-sensor	M-Sensor	Composition algorithm

there must be a balance between the scoring accuracy and the research and development costs recognized by the examination organization department.

Generally, the training process of intelligent scoring model algorithm needs to go through evaluation guide scanning, simulation question scanning, equipment list scanning, evaluation point scanning, scoring test, formal test training, and score processing and optimization. These seven links are a whole and also an optimization basic framework. Every link must be solid, so that it is possible to truly realize the intelligent scoring system that meets the requirements of unified examination. Most of the technical solution models of the existing technical support parties are trained based on the balance between cost and scoring accuracy, using all or part of the seven links.

2.2 Intelligent Scoring Model of Physical and Chemical Experiment Operation Test in Junior High School

Based on the seven-training links, the existing technical support parties have built their own intelligent scoring models and scoring engines according to their different schemes and patents [5–8]. Figure 1 shows the schematic diagram of intelligent scoring for the physical and chemical experiment operation test in junior high school. The scoring quality of the scoring engine mainly depends on the classification quality of the algorithm for the training set and the signature class library. The more comprehensive the training set is, the more accurate the algorithm is, the more distinct the three class libraries of the scoring engine are, and the higher the scoring quality and accuracy of the scoring engine are.

However, in the actual operation process, due to the difference in cost control, algorithm and training set, as well as the incomplete collection of environmental characteristics and equipment characteristics, the existing technical support companies still have the following obvious deficiencies in the signature class library trained through the limited training set:

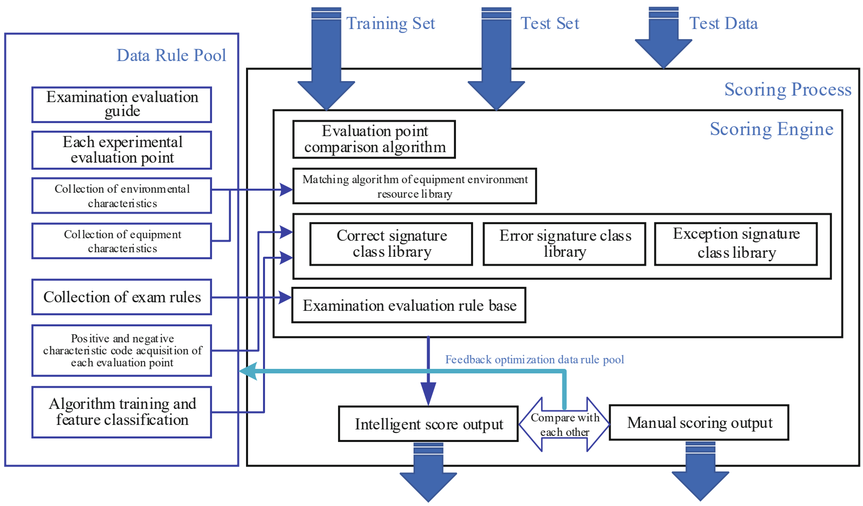


Fig. 1. Schematic diagram of intelligent scoring model for physical and chemical experiment operation test in junior high school

First, the analysis of each experimental evaluation point required by the evaluation guide is insufficient. Most of the evaluation points are based on the existing test paper analysis, rather than the full coverage of each experiment based on the analysis of the evaluation guide itself;

Second, due to the differences between the technical routes of the technical support parties, the information collected by the parties on the environment and equipment characteristics cannot be shared, and there is no unified standard, so there is no complete and fully covered resource base;

Third, various rules in the examination rule base may have certain conflicts among different rules in the rule base due to differences in different regions, simulation and training requirements;

Fourth, the data of the correct, incorrect and exceptional characteristic code class library of the same evaluation point is insufficient (most technical solutions only have the correct and other two class libraries based on cost considerations), resulting in low differentiation or accuracy of the scoring results.

3 Improvement of Intelligent Scoring Model and Algorithm

In view of the above shortcomings, combined with the large-scale scoring practice of the physical and chemical experiment operation test of junior high school in Shanghai in 2021, this paper proposes a particle swarm optimization algorithm based on frame adjustment and improved classification of test sets. There are two key points. First, the framework is adjusted and optimized based on the seven-training links. All the required experimental contents in the evaluation guide are split and refined to form a meta-evaluation point library, which is uniformly numbered. Each meta-evaluation point is allocated with corresponding equipment list resources, test evaluation rules, correct feature code library, error feature code library and exception feature code library. The second is to optimize the fast classification algorithm of the test set for the correct, error and exception signature class library. Because the score of the physical and chemical experiment operation test includes both the evaluation of the experimental results and the evaluation of the experimental process, and the evaluation of the process, the collection of the three feature library data and the collection of sequence information are relatively difficult. At the same time, in order to reduce the time of the scoring process, it is necessary to quickly classify the relevant videos, pictures or fused images in a certain link. After analyzing the original classification algorithm of technical support, we found that the existing algorithm is difficult to obtain accurate scoring support for the evaluation of experimental process and complex results. Therefore, there is an embarrassing situation that "the accuracy rate is more than 80%, but where is the inaccurate 20%? I don't know".

When optimizing the classification algorithm, the balance between classification accuracy, speed, computation and cost must be considered. Obviously, the amount of information related to the configuration resources of the meta-evaluation point will affect the calculation amount of the classification and comparison of the feature class library. If the meta-evaluation point feature class library is large, in principle, the accuracy of the scoring is very high, but the calculation amount of the scoring engine will also become very large, so how to find a balance between the calculation amount and the scoring accuracy and speed is also very important. How to obtain the best scoring accuracy and speed through the characteristic code data of small sample size at the meta-evaluation point, and how to submit the exceptional characteristic code data for manual judgment and processing as far as possible, so as to reduce the misjudgment of machine intelligent scoring as far as possible, so as to improve the accuracy of effective scoring.

After comprehensive balancing, this paper proposes to use the improved particle swarm optimization clustering algorithm (IPSOCA, Improved Particle Swarm Optimization Clustering Algorithm) to deal with the intelligent scoring of physical and chemical experiment operation tests, and to transform the scoring problem of test sets into the clustering problem of test sets on the feature class library. First, three initial populations (correct signature group, error signature group and exception signature group) are generated, and a group of particles are randomly initialized in the feasible space of the test set. Each particle is a feasible solution in the optimization space, and an objective function determines a fitness value for it. Each particle moves in the solution space, and its evolution direction and distance are determined according to its speed and position, The particles will follow the current optimal particle population search in the solution space. In each iteration, the particle will track two extreme values to update itself. One is the optimal solution found by the particle itself, that is, the current optimal solution, and the other is the optimal solution found by the entire population, that is, the global optimal solution.

3.1 Model Analysis and Algorithm Steps of Improved IPSOCA

In the improved particle swarm optimization based clustering algorithm (IPSOCA), the digital coding method based on the cluster center is adopted. For data n samples to be classified, the dimension is d, and each particle V_i is composed of k cluster centers, which can represent the numerical code with length $l = k \times d$. A particle to be classified can adopt the following formula (1):

$$V_i(t) = (\underbrace{a_{11}a_{12}\cdots a_{1d}}_{c_1}\cdots\underbrace{a_{i1}\cdots a_{id}}_{c_i}\cdots\underbrace{a_{k1}a_{k2}\cdots a_{kd}}_{c_k}) \tag{1}$$

Among them, $c_1, c_2, \cdots, c_k$ the corresponding numerical codes represent the coordinates of the cluster centers of each group in the sample space of the classification data, and the group can be randomly initialized using the results of five iterations of K-means iteration.

This paper uses CD kernel distance to measure the distance calculation and analysis of fitness function in particle swarm optimization algorithm. Assuming that there are d-dimensional data samples $X = \{x_1, x_2, \cdots, x_n\}$ to be classified in the input pattern space O^d, the idea of kernel function clustering is to use a nonlinear mapping relationship $\lambda : O^d \rightarrow Hx_i \rightarrow \lambda(x_i)$ to map the samples x_i in the input pattern space O^d to a high-dimensional classification feature space H, with the purpose of highlighting the differences in classification characteristics between different types of samples, so that the samples to be classified become linearly separable or approximately linearly separable in the feature space, Then cluster in this high-dimensional classification feature space, so as to realize the rapid classification of the test set in the physical and chemical experiment operation test in the three initial populations.

Among them, test set sample

$$x_i = [x_{i1}, x_{i2}, \cdots x_{id}]^{\mathrm{T}}, \lambda(x_i) = [\lambda_1(x_i), \lambda_2(x_i), \cdots, \lambda_{\mathrm{H}}(x_i)]^{\mathrm{T}}$$

Through mapping relationship, $x_i \cdot x_j$ Convert to $\lambda^T(x_i) \cdot \lambda\left(x_j\right)$. In the high-dimensional feature space H, the inner product kernel function can be defined as (2):

$$W(x_i, x_j) = \lambda^T(x_i) \cdot \lambda(x_j) \tag{2}$$

In the actual calculation, we use radial basis kernel function. Therefore $W(x_i, x_j) = 1$, in the feature space, the CD kernel function distance between the sample pattern x_i and x_j of the test set can be expressed as (3):

$$\begin{aligned} \left\|\lambda(x_i) - \lambda(x_j)\right\|^2 &= \lambda^T(x_i) \cdot \lambda(x_i) - 2 \cdot \lambda^T(x_i) \cdot \lambda(x_j) + \lambda^T(x_j) \cdot \lambda(x_j) \\ &= (\lambda(x_i) - \lambda(x_j))^T(\lambda(x_i) - \lambda(x_j)) \\ &= W(x_i, x_i) - 2 \cdot W(x_i, x_j) + W(x_j, x_j) \\ &= 2 \cdot (1 - W(x_i, x_j)) \end{aligned} \tag{3}$$

The *CD* core distance can be defined as (4):

$$\begin{aligned} CD(k) &= \frac{\frac{1}{k}\sum_{i=1}^{k}[\frac{1}{M_i}\sum_{X_i \in C_i} \max_{X_q \in C_i}\{d(x_i, x_q)\}]}{\frac{1}{k}\sum_{i=1}^{k}[\min_{j \in K, j \neq i}\{d(c_i, c_j)\}]} \\ &= \frac{\sum_{i=1}^{k}[\frac{1}{M_i}\sum_{X_i \in C_i} \max_{X_q \in C_i}\{d(x_i, x_q)\}]}{\sum_{i=1}^{k}[\min_{j \in K, j \neq i}\{d(c_i, c_j)\}]} \end{aligned} \tag{4}$$

where, $c_i = \frac{1}{M}\sum_{x_j \in C_i} x_j$ is the cluster center of the i cluster, and $d(x_i, x_j)$ is the Euclidean distance between the data points x_i and x_j of any two test sets,

To this end, we can map the data to be classified in the test set to the high-dimensional feature space H (three initial populations) by using the *CD* kernel function. The calculation of formula (4) can be further optimized as follows formula (5):

$$\begin{aligned} CD_{kernel}(k) &= \frac{\sum_{i=1}^{k}[\frac{1}{M_i}\sum_{X_i \in C_i} \max_{X_q \in C_i}\{\left\|\lambda(x_i) - \lambda(x_q)\right\|^2\}]}{\sum_{i=1}^{k}[\min_{j \in K, j \neq i}\{\left\|\lambda(c_i) - \lambda(c_j)\right\|\}]} \\ &= \frac{\sum_{i=1}^{k}[\frac{1}{M_i}\sum_{X_i \in C_i} \max_{X_q \in C_i}\{2 \cdot (1 - W(x_i, x_q))\}]}{\sum_{i=1}^{k}[\min_{j \in K, j \neq i}\{2 \cdot (1 - W(c_i, c_j))\}]} \end{aligned} \tag{5}$$

Through the test and analysis of the classification data of the actual test set, the *CD* core distance can effectively solve the problem that the distribution of the data points to be classified in the test set is quite different, but with the increase of the value of k and n, the calculation amount of the system will also significantly increase, and the classification speed will also decline. We take the *CD* kernel function as the fitness function, so the fitness function of the i particle V_i to be classified in the test set is:

$$f_{V_i} = \frac{1}{CD_{kernel}(k) + 1} \tag{6}$$

Formula (6) shows that if the fitness function is to be the maximum value, the *CD* kernel function is to be the minimum value, so that the data points to be classified can reach the optimal division, and the accuracy of classification and scoring can be achieved. To this end, the following algorithm steps can be used to optimize the original scoring algorithm. The clustering algorithm based on particle swarm optimization steps are as follows:

(1) Enter: n signature data set of data; Number of clusters; Maximum number of iterations; Stop threshold.
(2) $t = 0$, initialize the particle swarm to be classified, and generate each particle to be classified using five K-means iteration methods;
(3) The current optimal position and global optimal position of particles are generated according to the fitness value;
(4) Update the velocity and position of all particles according to the particle swarm velocity and position update formula to generate a new classified particle swarm;
(5) Calculate the fitness of the updated particle, update the current optimal position and global optimal position of the particle;
(6) $t = t + 1$, , If the change rate of clustering error is less than the stop threshold, or the upper limit of the number of iterations is reached, the program terminates and the global optimal position is output. Otherwise, skip to step (4).

3.2 Analysis of Application Effect of Improved IPSOCA Model Algorithm

In order to compare the application effect of intelligent scoring algorithm optimization in physical and chemical experiments, this paper selects the experiment of "Exploring the Principle of Bar Balance" in the physical experiment project of physical and chemical experiment operation test in junior high school for comparative analysis. The experiment includes seven evaluation points, which are relatively basic experiments in all physical experiments, and the relevant physical experiment test set also has good distinguishability.

3.2.1 The Existing Scoring Framework and Algorithm of the Technical Support Party

The existing intelligent scoring model and algorithm of technical support party A is mainly to correctly compare and classify the captured video image with the other two feature class library data, as shown in Fig. 2.

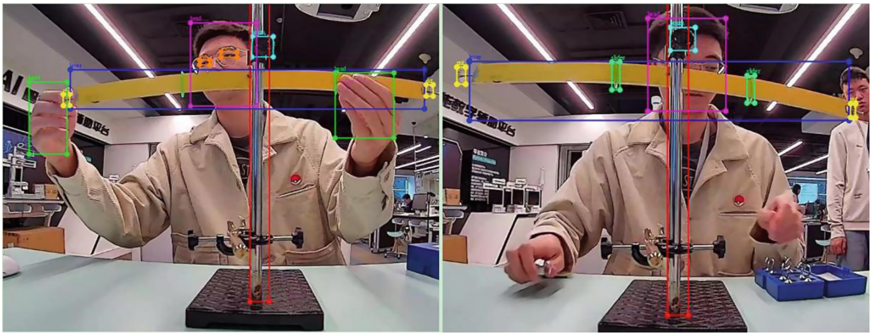


Fig. 2. Operation picture of the original scoring algorithm in exploring the principle of leverage balance [Photographed by the author]

The intelligent scoring model and algorithm of Company A is mainly based on the classification library of the correct operation of the experimental examination video stream fusion image capture to judge and obtain the scores of different evaluation points. This method has the characteristics of simple operation and rapid scoring. However, because there are only two types of feature class library: correct and other, the scoring accuracy is greatly affected. In fact, it is impossible to judge whether the data classified into the "other feature" class library is scored or not. However, for the completeness of the scoring, the data classified into the "other feature" class library is generally placed in the "no score" scoring result. This reduces the accuracy of scoring to a certain extent.

3.2.2 Adopt New Framework Adjustment and IPSOCA Model and Algorithm

In order to optimize the relevant content, this paper first constructs the meta-evaluation point library, equipment library, correct signature class library, error signature class library, exception signature class library and test rule library of the "Exploring the Principle of Bar Balance" experiment. At the same time, after adopting IPSOCA algorithm, we optimized the intelligent scoring framework and process, and further regulated and optimized the experimental scoring framework of "Exploring the Principle of Bar Balance" in strict accordance with seven training links and contents. The relevant steps are shown in Table 2.

According to the optimization links and steps, the basic contents of the meta-evaluation point library, the equipment library, the correct signature class library, the error signature class library, the exception signature class library and the examination rule library of the "Exploring the Principle of Bar Balance" experiment have been comprehensively updated, as shown in Table 3, because there are many examples involved, only some signature examples are listed in Table 3. After completing the framework adjustment and resource data information update, we can start to adjust the algorithm model. With the help of company A's scoring platform, we have built a new scoring engine and updated the scoring framework and scoring algorithm model.

Table 2. List of intelligent scoring optimization links of physical and chemical experiment operation test in junior high school based on IPSOCA algorithm optimization

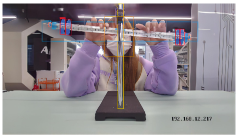
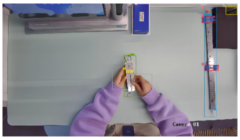
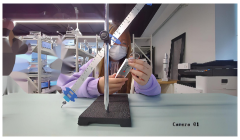
step	Optimization link	Content	Operating instructions	remarks
1	Relevant experimental requirements of scanning evaluation guide	Experiment name numbering	Separate number	Explore the principle of bar balance
		Experimental element evaluation point numbering	Subdivide to the smallest unit	
		Scanning and numbering of laboratory equipment list	Build a multi-view feature library	Form a laboratory equipment library of a single manufacturer
		Establish evaluation point resource information base	Form a single evaluation point resource library	
		Establish examination evaluation rule base	Form various examination rules	Regular expression of test fairness and justice rules
		Example of correct operation	Form correct signature class library	
		Typical error operation example	Form error signature class library	
		Establish abnormal feature information base	Form exception signature class library	Submit the basis for manual judgment and handling
2	Training simulation question scanning	Display format	Uniform standard format	
		content analysis	Single evaluation point resource pool matching	
		Answer setting	Unified input mode, etc.	Implement according to the unified examination interface specification

(*continued*)

Table 2. (*continued*)

step	Optimization link	Content	Operating instructions	remarks
		Evaluation point matching		
3	Examination equipment list scanning	Name range matches evaluation point		Adapt to equipment resource library
		Pattern recognition and matching of experimental equipment	Create multi-view signature	
		Scan special equipment	Update equipment resource library	
4	Scanning of evaluation points of training simulation questions	Evaluation point matching	Update meta-evaluation point library	
		Scan the correct signature class library	Update the correct signature class library	
		Scan error signature class library	Update error signature class library	
		Scan abnormal alarm information base	Update exception signature class library	
5	Examination video stream matching test	Comparison of evaluation point sequence integrity	Establish the scoring point sequence framework	If it is incomplete, go to rule judgment
		Single evaluation point clustering analysis based on new algorithm	Basic operation of intelligent scoring	Direct grading based on clustering results
		Proportion and accuracy of exceptions	There is a situation that the machine cannot recognize	Submit manual processing according to the results
		Judgment of examination rules	Judge whether there is any violation of discipline, cheating, etc.	It requires manual late judgment

Table 3. List of experimental scoring resources based on the new framework and IPSOCA model algorithm optimization [All the pictures involved were taken by the author]

No.	Optimization link	Content	Signature information or operation instructions	remarks
1	Evaluation point library	1. Support the middle point of the lever on the iron frame		
		Description of other evaluation points......	......	Reference Guide
2	Equipment warehouse	1. Spring dynamometer		Single manufacturer
		Description of other experimental equipment......	......	Single manufacturer
3	Correct signature library	1. Example 1 of correct signature code		Corresponding evaluation point, forming corresponding correct characteristic code 1
		Examples of other correct signatures......	......	......
4	Error signature library	1. Error signature example 1		Corresponding evaluation point, forming corresponding error characteristic code 1
		Examples of other error signatures......	......	......
5	Exception signature library	1. Exception signature example 1		Corresponding evaluation point, forming corresponding exception characteristic code 1
		Examples of other exception signatures......	......	......
6	Scoring and arbitration of examination evaluation	Sequencing and scoring of evaluation points	Set the correct order of evaluation points according to the requirements of simulation questions	Preset
		Examination rule base	Set rules based on fairness, justice and cheating	Preset
		Determination of exceptions	In case of any exception, the system prompts and submits it for manual processing	Preset
		Video stream cutting label and classification	Capture evaluation points, combine the order of evaluation points, and intelligently assign points	
		Statistical analysis of data	Test the same amount of video stream, compare the new and old algorithms and manual scoring data	

3.2.3 Analysis and Comparison of Relevant Scoring Data

After adopting the new adjustment framework and scoring optimization algorithm, this paper makes a comparative analysis of the scoring results of the experimental test video of "Exploring the Principle of Bar Balance" in the randomly selected 96 junior high school physical and chemical experimental operation tests, and obtains the data shown in Table 4.

From the data in Table 4, we can see that the scoring accuracy of the new framework and IPSOCA model algorithm is significantly higher than that of the existing intelligent scoring framework and algorithm of technology company A. In the comparison of the scoring data of this experiment, the scoring results of the new architecture and IPSOCA algorithm are also slightly higher than the results of the one-to-four teachers' on-site scoring. This result has an important stage significance for the machine scoring of large-scale physical and chemical experiment operation tests. It also shows that in some specific experiments, through the construction of a complete meta-evaluation point resource class

Table 4. Analysis of intelligent scoring data based on framework adjustment and IPSOCA algorithm optimization

No.	Scoring method	Explore the principle of bar balance(96 people operation video)						remarks
		Total score points	Scoring point sequence OK	Score of correct class library or evaluation point	No score for wrong class library or evaluation point	Exception class library	Examination rules punishment	
1	New architecture and IPSOCA algorithm	672	668	634	26	12	1	Exceptional data needs to be judged manually
2	Original scoring framework and scoring algorithm*	672	—	539	133	—	—	The original method only has correct and other two class libraries
3	1 to 4 teacher on-site scoring method	672	671	636	36	—	1	The grading teacher cannot be disturbed
4	1-to-1 multi-angle teacher review score	672	669	632	40	—	1	Can be compared as standard answer

* Note: The original scoring framework and scoring method refer to the existing intelligent scoring framework and algorithm of technology company A for comparative testing

library and the optimization of algorithms, It can realize the reliable automatic machine intelligent scoring of experimental operation test.

4 Conclusion and Prospect

Through analysis and comparison, this paper believes that the machine intelligent scoring of physical and chemical experimental operation test in junior high school is one of the effective ways to solve the problems of fairness, fairness and time-consuming and laborious in large-scale grading of experimental operation test at present. As long as a scientific scoring framework and efficient scoring algorithm are adopted, and an accurate, unified and standardized meta-evaluation point scoring resource information library is established, it is possible to make the accuracy of machine intelligence scoring reach or even exceed the results of manual on-site scoring through continuous optimization and iteration process.

Of course, the optimization of models and algorithms is a continuous iterative process, which requires the continuous improvement of training sets, the continuous standardization of test sets, and the continuous optimization of algorithms for different meta-evaluation points to continuously improve the accuracy and assignability of intelligent scoring. This paper will also continue to follow up the latest research progress of relevant content in the follow-up research, and make continuous efforts for the intelligent scoring of the physical and chemical experiment operation test in junior high school.

References

1. Ministry of Education Guiding Opinions of the Ministry of Education on Further Promoting the Reform of the Examination and Enrollment System in High School [EB/OL] (2016-09-19). http://www.moe.gov.cn/srcsite/A06/s3732/201609/t20160920_281610.html.
2. Shanghai Education Notice of Shanghai Municipal Education Commission on Printing and Distributing the Measures for the Implementation of Shanghai Junior High School Academic Level Examination [EB/OL] (2019-04-08). http://edu.sh.gov.cn/zcjd_area_3678/20200706/0015-xw_99986.html.
3. Cao Jingjing Give full play to the guiding role and comprehensively improve the level of experimental teaching and education in Shanxi Province – Summary of the physical and chemical experiment operation test for junior high school graduates in Shanxi Province [J] Shanxi Education, 2018 (10): 51–52
4. Zhuo Min, Zhang Dongmei Consideration on the construction of equipment related to the experimental operation test for junior high school physics and chemistry students in Liaoning Province [J] Education and Equipment Research, 2022 (9): 83–86
5. Liu Kai et al A specific action detection method for intelligent experiment evaluation: China, 114743270. A [P]. 2022-07-12
6. Liu Lifei et al A moving region detection method based on image segmentation for experimental evaluation: China, 114742858. A [P]. 2022-07-12
7. Never wait too late A motion recognition method and related device based on artificial intelligence: China, 110765967. A [P]. 2020-02-07
8. Sun Xiaohua et al Intelligent scoring system for experimental examination based on AI vision: China, 111915460. A [P]. 2020-11-10