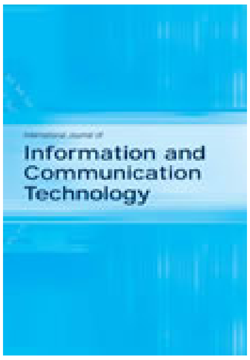


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

A machine classification learning model based on factor space mathematical theory in higher vocational education

Ruishuai Chai

DOI: 10.1504/IJICT.2025.10071985

Article History:

Received:	06 May 2025
Last revised:	23 May 2025
Accepted:	23 May 2025
Published online:	10 July 2025

A machine classification learning model based on factor space mathematical theory in higher vocational education

Ruishuai Chai

Henan Industry and Trade Vocational College,
Zhengzhou 450046, China
Email: zhang815474413@163.com

Abstract: To address teaching semantic gap issues in image sample learning for vocational education evaluation, this paper first applies factor domain theory to the teaching semantic embedding domain. Based on the relationships among semantics, it studies conjunction and reduction of factors, as well as the expansion and contraction of the factor domain. The enhanced factor space approach is then utilised in vocational education evaluation. Visual features are extracted using the residual network (ResNet101), and a generative adversarial network (GAN) is trained to produce more realistic picture characteristics. By combining teaching attributes and noise, the generator outputs picture characteristics which are then combined with class marks to train the categoriser, thereby completing the classification of teaching evaluation images. Experimental results reveal that the offered model achieves a classification accuracy of 92.5%, effectively helping to enhance the quality of higher vocational education.

Keywords: vocational education; factor space mathematical theory; machine classification learning; ResNet101 model; generative adversarial network; GAN.

Reference to this paper should be made as follows: Chai, R. (2025) 'A machine classification learning model based on factor space mathematical theory in higher vocational education', *Int. J. Information and Communication Technology*, Vol. 26, No. 25, pp.33–47.

Biographical notes: Ruishuai Chai studied in Henan Normal University from 2000 to 2004 and received his Bachelor's degree in 2004. From 2007 to 2009, he studied in Huazhong University of Science and Technology and received his Master's degree in 2009. Currently, he works in Henan Industry and Trade Vocational College. He has published ten papers. His research interests are mathematics education and student work.

1 Introduction

In today's era of rapid digitalisation and intelligent development, vocational education contributes significantly to the cultivation of high-quality technical talents. Therefore, innovation in teaching models and methods is crucial (Nadzinski et al., 2023). As one of the core technologies in the field of artificial intelligence (Ilić et al., 2021), machine learning has gradually emerged in higher vocational education, providing new ways to improve teaching effectiveness and optimise the teaching process. It can analyse and process large amounts of teaching data, uncover potential patterns and rules, and thereby achieve personalised teaching (Zhou and Liu, 2023). The mathematical theory of factor space takes factors as its core and, through the construction of factor space, systematically describes and analyses the attributes and characteristics of things, revealing the internal connections and causal relationships between them (Wang, 2018). The introduction of factor space mathematical theory into machine learning model construction in higher vocational education can more accurately extract key factors affecting classification, optimise the structure and parameters of classification models, improve the classification accuracy and generalisation ability of models (Yu, 2013), and enable them to better adapt to the complex data and diverse needs of higher vocational education.

In higher-level vocational studies, there is a significant lack of data on students' evaluations of teachers' teaching. How to classify teaching effectiveness using a small amount of image data has become a hot topic of research. Cai (2023) predicts the image's attributes and categorises it by selecting the nearest-matching attributes from the available set. Zhang et al. (2024) proposed a new attribute learning model by computing nearest attribute neighbours to automatically extract semantic information, and combined different language resources to mine attribute information, thereby completing the classification of educational evaluation images. Wang et al. (2022) investigated the correlation between teaching characteristics and student traits using mutual information analysis. The system implemented attribute prediction as a latent variable task, with a support vector machine (SVM) classifier fusing attribute and class information to optimise object recognition performance. Hou (2021) proposed a higher vocational education evaluation model based on SVM and decision trees, which uses SVM to mine the relationships between teaching attributes, thereby improving classification accuracy. Long and He (2023) combined a ranking function with the correlations between attributes calculated using a ranking algorithm, thereby improving the effectiveness of image classification.

Recent machine learning classification models primarily focus on directly learning the mapping among image visual characteristic domain and semantic domain. Hassan and Shamsudin (2019) map image characteristics to a semantic domain, afterwards compute the cosine similarity to identify the closest class embedding for label prediction. Ai and Feng (2022) projected visual features into the attribute embedding space, established a mapping between image representations and semantic concepts, and finally decoded them back to the original visual space to construct the categorisation framework. Wu (2024) drew inspiration from multi-task learning and transfer learning (Weiss et al., 2016) to design a model that jointly learns attribute labels and category labels. This architecture employs shared image feature representations for both attribute and category studying, with dual category labels co-located at the same hierarchical level for joint optimisation, forming an attribute-category parallel model, thereby improving the accuracy of image classification.

In machine classification learning models with fewer teaching evaluation image samples, both known and unknown categories require the establishment of an intermediary knowledge space that functions as a common representation domain. The unified knowledge space admits multiple representational forms-from lexical embeddings to linguistic descriptions and conceptual vectors. Consequently, this paper opts for factor domain to serve as the auxiliary knowledge domain. Chen and Song (2023) proposed the factor library theory, which classifies concepts based on attributes to extract useful information from massive educational evaluation data, providing a new approach to big data processing. Lu and Wang (2023), inspired by factor analysis, proposed a differential calculation algorithm based on determinism as the basic statistical measure, which improved the classification accuracy. Guleria and Sood (2023) used the concept of factor visibility to study classification algorithms and proposed a sweep-class chain algorithm to accelerate the classification speed and improve the accuracy of the algorithm.

Based on a comprehensive analysis of the aforementioned studies, it can be concluded that current vocational education teaching evaluation image sample learning suffers from the semantic gap problem. To address this issue, this article presents a machine-learning categorisation model grounded in the mathematical theory of factor space, specifically designed for vocational education teaching. First, the factor space algorithm is improved by applying factor domain to another semantic domain for teaching, to establish alignment between semantic abstractions and perceptual feature representations, a link is made among data characteristics and the data visually expressed in images. In light of the semantic relationships between factors, this study investigates the conjunction and reduction of factors, as well as the expansion and contraction of factor spaces. It then improves existing factor space algorithms and modifies them using machine learning-related algorithms to make the new algorithm more suitable for sparse teaching evaluation image classification technology. The improved factor space algorithm was then incorporated into the categorisation approach for evaluating higher vocational education. Visual features of images were extracted using ResNet101, and a generative adversarial network (GAN) was trained to produce more realistic picture characteristics. Picture characteristics were generated by merging teaching attributes and noise, and the classifier was trained using these characteristics paired with their class labels. Finally, the trained classifier receives test data as input features, which outputs class labels in the images, thus finalising the instructional assessment picture categorisation. The experimental outcome reveals that the categorisation accuracy of the proposed approach is enhanced by 5%–17.5% compared with the baseline model, demonstrating good classification accuracy.

2 Relevant technologies

2.1 Mathematical theory of factor space

Factor domain mathematical theory is a mathematical framework for describing and processing the relationships among factors in uncertain and complex systems (Li and Xu, 2001). In mathematical terms, factors constitute a particular class of mappings that transform objects into their phase representations, elucidating the essential relationship between objects and their factors. The state space mentioned in artificial intelligence is not a mathematical space but a system of variables. Each variable is a factor with a phase domain, and each factor's phase domain is a coordinate axis. Any object can be mapped to a point in the information space through a set of factors (Forni et al., 2015), as implied in Figure 1.

A thing is not related to any factor. The so-called relationship between thing u and factor f means that when talking about u in terms of f, there is a corresponding state $f(u)$. Let U be a set composed of some objects, V be a set composed of some factors, and for any $u \in U$, all factors related to u are in V. Then $(U, V]$ is called a left pairing. If a relationship R: $R(u, f) = 1 \Leftrightarrow u$ is defined between U and V that is related to f. In addition, let $D(f) \triangleq \{u \in U \mid R(u, f) = 1\}$ be the set of all possible states of the system. Then, we can define the following mappings f and $X(f)$ as the state spaces of f.

$$f : D(f) \to X(f),\ u \to f(u) \tag{1}$$

$$X(f) \triangleq \{f(u) | u \in D(f)\} \tag{2}$$

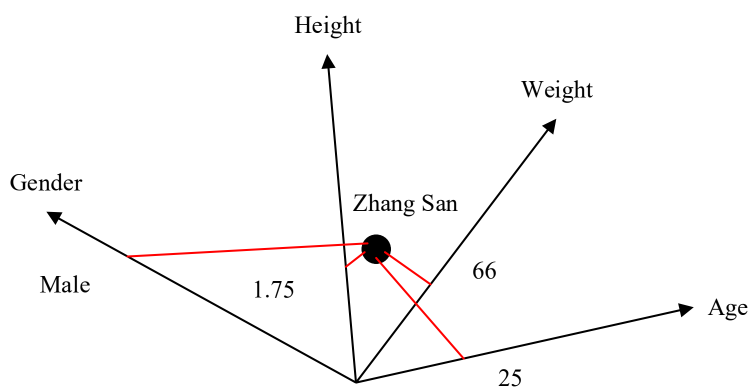
Given a left pairing $(U, V]$, take a factor family $F \subset V$, and call the set family $\{X(f)\}_{(f\in *)}$ a factor domain on U. If it meets the following axioms:

1 $F = F(\vee, \wedge, c, 1, 0)$ is a complete Boolean algebra

2 $X(0) = \{\varnothing\}$

3 for any $T \subset F$, if the factor family T is pairwise independent, then $\underset{f\in T}{\vee} f = \prod_{f\in T} f$.

Figure 1 Schematic diagram of factorisation (see online version for colours)



2.2 *Attribute classifier*

In the classification of images with fewer teaching evaluation samples, attributes need to be introduced as an intermediate bridge between known and unknown categories to infer the unknown categories. After constructing a knowledge sharing space and mapping relationships between attribute features and entity category labels, the key step in attribute-based image classification technology is to train attribute classifiers (Zhu et al., 2020). If the model needs to learn attribute classifiers, it is necessary to construct a mapping model between image low-level features and attributes. This model mainly consists of two parts: image feature extraction and attribute classifier training.

First, input the image and extract its features. The extracted features are visual features at the low-level semantic layer. To obtain more useful features, perform secondary feature extraction. However, traditional methods of extracting image features require a lot of manual intervention. Currently, the most widely used feature extraction method is deep learning. Deep learning methods mainly input images into neural networks, which automatically extract features from the images. The commonly used model is the convolutional neural network (CNN) (Li et al., 2021). After obtaining image features, it is essential to choose an appropriate classifier to classify the images and obtain category labels. Commonly used classifier models include Softmax classifiers.

3 Design and improvement of factor space algorithms

3.1 Improvements to factor space sparse representation algorithms

To address the semantic gap problem in sparse image sample learning, the factor space projection onto the semantic embedding space ensures consistency across semantic abstractions and primitive visual characteristics, constructing an explicit mapping between feature representations and image semantics. According to the semantic relationships between factors, this study investigates the conjunction and reduction of factors, as well as the expansion and contraction of factor spaces. Existing factor space algorithms are improved, and machine learning-related algorithms are applied to enhance the factor space algorithms, making the new algorithm more suitable for sparse image sample classification techniques.

For machine learning classification tasks, given a dataset where some data samples are critical and others may not be very useful, the selection of data samples is an important data preprocessing step. Select relevant samples from a given data set to decrease the difficulty of the learning task. Consider the data set as a matrix, where each row represents a data sample and each column represents the features of the sample. The process of selecting features from samples is the process of making features 'sparse', which can reduce the difficulty of learning tasks, reduce computation and storage requirements, and improve the interpretability of the learned model.

To simplify the task, the factors and attributes are 'thinned out'. The specific process of 'thinning out' is as follows.

1 For a given set of factors $F = \{f_1, f_2, \ldots, f_m\}$, each factor has corresponding attribute values $f_1 = \{a_{11}, a_{12}, \ldots, a_{1n}\}$ and $f_m = \{a_{m1}, a_{m2}, \ldots, a_{mn}\}$, forming a factor-attribute matrix. The expression for the learning dictionary is as follows.

$$\min_{B,\alpha_i} \sum_{i=1}^{m} \|f_i - B\alpha_i\|_2^2 + \lambda \sum_{i=1}^{m} \|\alpha_i\|_1 \quad (3)$$

where B is the dictionary matrix and α is the sparse representation of f.

2 Optimise equation (3) using variable alternating optimisation to learn sparse representations α and dictionary matrix B. First, fix dictionary B and find the corresponding sparse representation α for each factor f. The optimisation process is as follows.

$$\min_{\alpha_i} \|f_i - B\alpha_i\|_2^2 + \lambda \|\alpha_i\|_1 \quad (4)$$

3 Update dictionary B based on the sparse representation α obtained from equation (4), as shown in equation (5).

$$\min_{B} \|F - BA\|_F^2 \tag{5}$$

where $F = \{f_1, f_2, \ldots, f_m\}$, $A = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$, and $\|\bullet\|_F$ are the Frobenius norms of the matrices.

4 The singular value decomposition (SVD) algorithm was used to update the above equation column by column, and the optimised solution was obtained in equation (6).

$$\begin{aligned}\min_{b_i}\left\|D-\sum_{j=1}^{k} b_j\,\alpha^j\right\|_F^2 &= \min_{b_i}\left\|\left(F-\sum_{j\neq i}^{k} b_j\,\alpha^j\right)-b_i\alpha^i\right\|_F^2 \\ &= \min_{b_i}\left\|E_i-b_i\alpha^i\right\|_F^2\end{aligned} \tag{6}$$

where b_i is the i^{th} column of the dictionary matrix B, and α^i is the i^{th} row of the sparse matrix A.

When updating the i^{th} column of the dictionary matrix, the other columns in the matrix are fixed, i.e., $E_i = D - \sum_{j\neq i}^{k} b_j \alpha^j$ in equation (6) is fixed. If the model wants to minimise $\min_{b_i}\left\|E_i - b_i\alpha^i\right\|_F^2$, just perform SVD on E to obtain the orthogonal vector corresponding to the maximum value.

5 After obtaining the initial dictionary matrix B, repeat the above steps iteratively until the dictionary matrix B and the sparse representation α are finally obtained.

3.2 Machine classification learning based on improved sparse representation algorithms

In data sets commonly used for sparse sample learning, the number of entity categories is less than the number of attributes contained in the entities. In image classification processes with fewer images, it is necessary to describe the information of entities through effective key attributes to reduce storage space and computation time. To highlight key attribute information, the dictionary matrix composed of categories and attributes is sparsified.

In machine classification studying, for the goal of describing the class of entities with least attribute data, it is necessary to remove unnecessary redundant information, set the coefficients before non-critical attributes to 0, and set the coefficients before critical attributes to non-zero. The over-complete dictionary of the category – attribute is constructed as Φ, the matrix is represented as $B = [f^1, \ldots; f^m; \ldots; f^M]$, f^m is the atoms in the dictionary representing the attribute information, $\alpha = [\alpha_1; \ldots; \alpha_m; \ldots; \alpha_M]$ is the coefficient vector of the attribute property, it is hoped that the non-zero elements in α are as few as possible, and s is used to represent the target information of the attribute.

When solving for sparse solutions of α, a suboptimal approximation algorithm is used to solve for sparse solutions. After obtaining the sparse solutions, the global optimal solution is obtained using the basis tracking method of convex relaxation (Ma and Xu, 2021), as shown in equation (7), where $\| \alpha \|_0$ is the sparsity of α.

$$\min \| \alpha \|_0 \quad \text{s.t.} \quad \Phi \alpha = s \tag{7}$$

For the minimisation problem of the l_0 norm in formula (8), by replacing it with the l_1 norm, the following equation is obtained.

$$\min \| \alpha \|_1 \quad \text{s.t.} \quad \Phi \alpha = s \tag{8}$$

Let $A = [\Phi, -\Phi]$, $b = s$, $c = [1, 1]T$, $x = [u, v]^T$, and $\beta = u = u - v$ be the variables. Then, equation (8) can be converted into a standard programming problem, as shown in equation (9).

$$\min c^T x \quad \text{s.t.} \quad Ax = b, x \geq 0 \tag{9}$$

The impact of chief attribute message of the objective signal s can be characterised through the magnitudes of the coefficient matrix α, as shown below.

$$(a_m, z) = |\alpha_m| \tag{10}$$

After performing sparse representation on the attribute information of each entity class, the attributes belonging to the entity domains undergo enumeration and calculation to ascertain the weight of the common attribute in every domain. Use the association probability $p(a_m, rel)$ and non-association probability $p(a_m, norel)$ to compute the values, as shown below.

$$p(a_m, rel) = \frac{count\left(a_m = a_m^p \in s_p\right)}{count\left(a_m = a_m^p\right)} \quad 11)$$

$$p(a_m, norel) = \frac{count\left(a_m = a_m^p \notin s_p\right)}{count\left(a_m = a_m^p\right)} \tag{12}$$

where a_m^p is the parameter of a_m in domain s_p; $count(a_m = a_m^p)$ is the p^{th} sample containing attribute a_m in domain s_p; $count(a_m = a_m^p \notin s_p)$ is the number of samples in domain s_p that do not contain a_m. Subsequently, the weight $\omega_p^{a_m}$ is as follows.

$$\omega_p^{a_m} = \frac{p(a_m, rel)}{p(a_m, norel)} \tag{13}$$

Based on the degree of correlation between a_m and various fields, the weights of each identified field resulting from clustering among each known field instances for a_m are as follows.

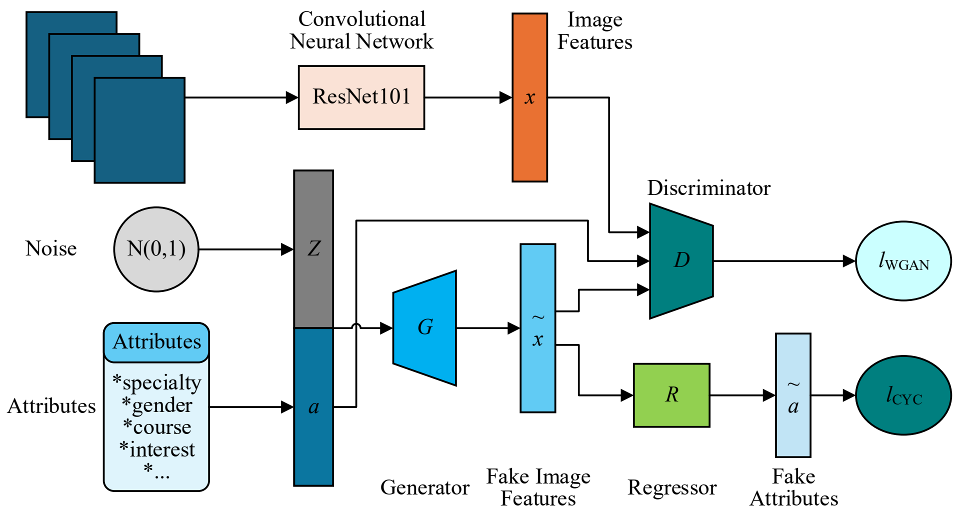
$$\hat{\omega}_p^{a_m} = \frac{\omega_p^{a_m}}{\Sigma_p \omega_p^{a_m}} \tag{14}$$

4 Classification of higher vocational education evaluation based on an improved factor space algorithm

4.1 The classification model for vocational education evaluation based on ResNet101 and GAN

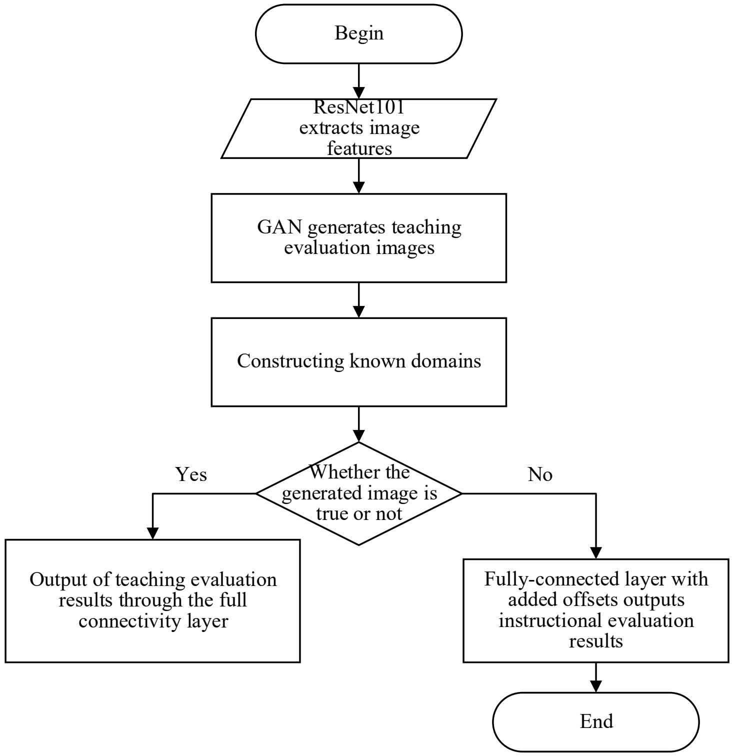
To address the issue of insufficient image evaluation datasets in existing vocational education evaluation systems, which leads to a lack of objectivity in evaluation results, this paper proposes a model combining ResNet101 (Zhang, 2022) and GAN (Goodfellow et al., 2020) to form the RGAN, as implied in Figure 2. The model is primarily composed of ResNet101, GAN, and a categoriser.

Figure 2 RGAN model (see online version for colours)



ResNet101 is adopted to capture picture characteristics. GAN's framework incorporates two adversarial networks: a generator network and a discriminator network. The generator synthesises artificial picture characteristics from arbitrary noise inputs, while the discriminator evaluates the authenticity of these produced pictures. During training, the two sub-networks compete with each other, enabling the model to generate more realistic features. Next, integrate test set attributes with noise vectors as generator input to synthesise output picture characteristics. Train the classifier using synthesised picture characteristics paired with their corresponding class labels from the test set. The test phase involves feeding images into ResNet101's residual architecture for feature embedding generation. The captured characteristics are fed into the categoriser to forecast teaching assessment classes for unseen images, yielding final categorisation outcome. The flow of higher vocational teaching evaluation classification method based on improved factor space is shown in Figure 3.

Figure 3 The flow of higher vocational teaching evaluation classification method



4.2 Classification model training based on improved factor space algorithm

The RGAN's procedure is primarily consist of the following phases: building a known field, feature extraction, network training, and testing.

Stage 1 Construct the known field. Compute the similarity among categories utilising the attribute message between teaching evaluation categories, perform hierarchical clustering on the teaching evaluation data, and construct the known domain of visible categories.

Stage 2 Characteristic extraction phase. Employ ResNet101 to pull out characteristics from the evaluation pictures. For the sake of boosting computational efficiency of network training. The network performs convolution and pooling operations on the images, and then outputs the characteristics of the pictures after passing by a fully connected level.

Stage 3 Network training phase. After preprocessing images in the dataset, they are input into the ResNet101 network. Once processed by a fully connected level that incorporates offset vectors, the attribute features of images are output. The mean square error operation is chosen as the loss operation for this training process, as demonstrated below.

$$loss : MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - y\_hat_i\right)^2 \tag{15}$$

where y_i is the real outcome, y_hat_i is the model forecasting outcome, and n is the number of teaching evaluation result categories: fail, pass, good, and excellent.

Then, train the GAN. The generator model is trained by extracting picture characteristics through ResNet101, combining the picture characteristics with class marks, passing them vy a fully connected level and selecting the ReLU function as the activation operation, that is a nonlinear operation, as shown below.

$$f(x) = \max(0, x) \tag{16}$$

After computation through a feed-forward level with an additive bias, the ReLU is chosen for the activation operation. A process of training the discriminator method entails capturing picture characteristics using ResNet101, combining the picture characteristics with class marks, then processed by a linear layer including a bias term. The activation operation chosen for the fully linked level is the leaky-Relu operation, as shown below, where $a \in (1, +\infty)$.

$$f(x) = \begin{cases} x & x \geq 0 \\ \dfrac{x}{a} & x < 0 \end{cases} \tag{17}$$

Subsequently, only the generated network is optimised to enhance the quality of its synthesised picture characteristics. The loss functions chosen for training are implied below.

$$\begin{aligned} \ell_{GAN(\theta_G,\theta_D)} = {} & E_{(x,a)\sim P^{x,a}}\left[D(x, a; \theta_D)\right] - E_{\tilde{x}\sim P_G^{x,a}}\left[D\left(\tilde{x}, a; \theta_D\right)\right] \\ & - \lambda E_{(\tilde{x},a)\sim P_\alpha^{x,a}}\left[\left(\left\|\nabla_{\hat{x}} D\left(\hat{x}, a; \theta_D\right)\right\|_2 - 1\right)^2\right] \end{aligned} \tag{18}$$

where D is the discriminated network, G is the generated network, θ is the random value, E is the expectation, a is picture's attribute message, x is the picture characteristic, $\tilde{x}$ is a image feature of the generated unknown object category, ∇ is the gradient computation, λ is the regularisation strength, $\hat{x} = \alpha x + (1-\alpha)x$, where $\alpha \sim U(0, 1)$.

Classifier training aims to combine the attribute features and noise obtained from the test set in Stage 1, feed them into the pre-trained generator, output the picture characteristics of unseen classes in the testing data and train the categoriser in conjunction with the class labels of unobservable classes. During the training process, the loss function employed is outlined below.

$$\ell_{CYC(\theta_R,\theta_G)} = E_{\alpha\sim P_S^a, z\sim N(0,I)}\left[\left\|a - R\left(G\left(a, z; \theta_G\right)\right); \theta_R\right\|_2^2\right] \tag{19}$$

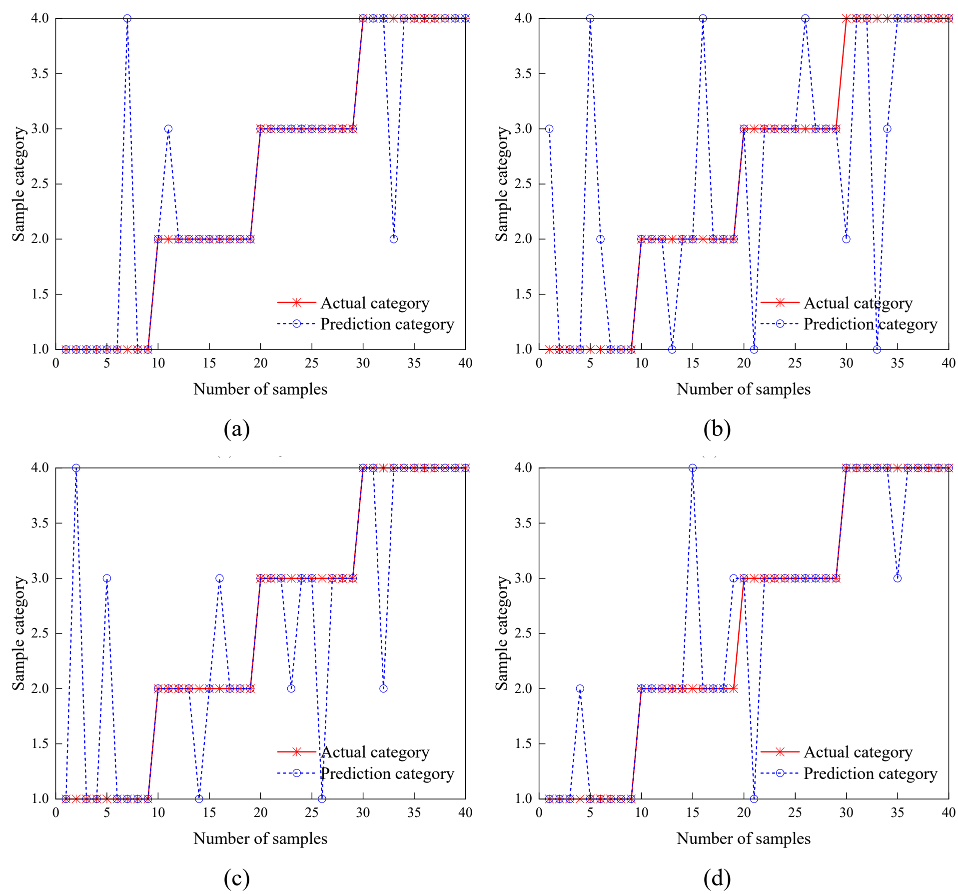
where R is the regressor, z is the arbitrary noise.

Stage 4 Testing stage. Utilise the features extracted from the test set images using ResNet101 in Stage 1, then input the captured picture characteristics into the trained categoriser. After passing through the categoriser, the class labels of the test pictures are output, thereby achieving the classification of vocational education evaluation.

5 Experimental results and analyses

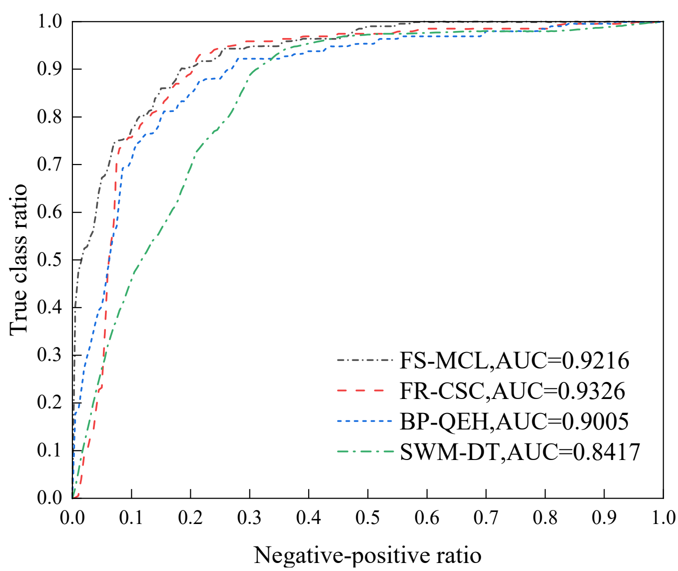
To validate the performance of the FS-MCL model designed in this paper, this study randomly selected relevant image data related to vocational education evaluation from the academic affairs management system of a vocational college and conducted simulation experiments. The initial data set consisted of 1,364 student classroom evaluation data items from eight semesters. This paper divides the teaching evaluation data samples into training sets and validation data sets in a ratio of 7:2:1. The graphics card model of the experimental server is Nvidia GTX 1080 Ti, the processor is Intel Core i5, the memory capacity is 8G, the operating system is Windows 10.

Figure 4 Comparison of classification results of teaching evaluation by different models, (a) FS-MCL (b) SVM-DT (c) BP-QEH (d) FR-CSC (see online version for colours)



The four categories of teaching evaluation results – unqualified, qualified, good, and excellent – are labelled as 1, 2, 3, and 4, respectively. The classification results of the three methods, FS-MCL and SVM-DT (Hou, 2021), BP-QEH (Ai and Feng, 2022), and FR-CSC (Guleria and Sood, 2023), are shown in Figure 4. When the number of image samples in the teaching evaluation was 40, SVM-DT, BP-QEH, FR-CSC, and FS-MCL correctly classified 30, 33, 35, and 37 samples, respectively, with classification accuracies of 75%, 82.5%, 87.5%, and 92.5%. The classification accuracy of FS-MCL improved by 17.5%, 10%, and 5% compared with SVM-DT, BP-QEH, and FR-CSC, respectively. SVM-DT considers the relationship between teaching attributes and student attributes and classifies teaching evaluation results using SVM, but the classification accuracy of SVM depends on the penalty parameter and kernel parameter. BP-QEH maps semantic spaces by reconstructing a small number of teaching evaluation images, but the reconstruction model may learn details unrelated to classification and ignore key discriminative features. FR-CSC uses the idea of explicit and implicit factors for classification research, but the division between explicit and implicit factors depends on prior knowledge. If the division is inappropriate, it will lead to a decline in model performance. FS-MCL applies factor space to semantic embedding domain, keeping the high-level semantic space consistent with the underlying picture characteristic domain, thereby greatly improving classification accuracy.

Figure 5 Comparison of ROC curves for different models (see online version for colours)



The ROC curves for different models are shown in Figure 5. The ROC curve for FS-MCL is the most convex, completely 'enveloping' SVM-DT and BP-QEH. That is, as the number of objects classified as positive increases, the number of misclassifications is less than that of the SVM-DT and BP-QEH algorithms. The ROC curve of FS-MCL does not completely envelop FR-CSC, but the AUC value is greater than that of the FR-CSC algorithm. The AUC value of FS-MCL is 0.9326, while that of FR-CSC is 0.9005. The classification performance of FS-MCL is superior to that of the FR-CSC algorithm. FS-MCL builds a direct link among data characteristics and the information expressed in pictures through factor space, studies the combination and reduction of factors based on semantic relationships, and the expansion and contraction of factor space. It also improves existing factor space algorithms and achieves good classification accuracy.

In addition to analysing the classification results, this paper also compares SVM-DT, BP-QEH, FR-CSC, and FS-MCL using precision, recall, F1, and training real-time rate (TR), as shown in Table 1. F1 is a composite measure of Precision and Recall (Hussain et al., 2019) that clearly reflects the classification performance of a model. The F1 score of FS-MCL is 93.4%, which is 6.7%-17.7% higher than the other three models. When comparing training real-time rates, FS-MCL's training real-time rate was 108 s/epoch, which was 372 s/epoch, 234 s/epoch, and 141 s/epoch lower than SVM-DT, BP-QEH, and FR-CSC, respectively. The FS-MCL model significantly improves the classification accuracy for test samples by incorporating residual networks to eliminate redundant attributes. This approach effectively enhances the correlation between attributes and target categories while reducing inter-attribute correlations. In addition, FS-MCL fully considers the correlations between attributes and between attributes and categories, thereby improving classification real-time performance.

Table 1 Comparison of classification performance indicators

Model	*Precision*	*Recall*	*F1*	*TR*
SVM-DT	74.3%	77.1%	75.7%	480s/epoch
BP-QEH	83.9%	82.6%	83.2%	342s/epoch
FR-CSC	85.2%	88.1%	86.7%	249s/epoch
FS-MCL	94.7%	92.1%	93.4%	108s/epoch

6 Conclusions

As the artificial intelligence technique rapidly growing, machine learning has been increasingly applied in the evaluation of higher vocational education. To address the current issue of the teaching semantic gap in evaluating image sample learning, this paper proposes a machine Categorisation studying model in light of factor domain mathematical theory for higher vocational education. The chief innovations of this model are summarised as bellow.

1 By applying factor domain theory to the instructional semantic embedding domain, this approach achieves consistency among high-level semantic space and low-level picture characteristic space, establishing direct connections between data features and image-conveyed information. The research investigates factor conjunction/reduction and factor space expansion/contraction based on semantic relationships.

2 Improve the factor space algorithm. Extend the theory of factor space based on machine learning algorithms, optimise the relationships between factors through clustering algorithms and sparse representation algorithms, and transfer them to the study of the relationships between categories and categories, attributes and categories in the image learning process of teaching evaluation, thereby improving the efficiency of image sample learning and the accuracy of classification.

3 The improved factor space algorithm is applied to a classification model for vocational education evaluation. Visual features are extracted from images using ResNet101, and a GAN is trained to produce more realistic picture characteristics. By combining attributes and noise, picture characteristics are output by the generator. The features and class marks are then combined to train the categoriser, thereby completing the classification of teaching evaluation images.

4 Experimental outcome on actual data sets show that the designed model gains a teaching evaluation classification accuracy rate of 92.5% and an AUC value of 0.9326, providing more reliable basis for teaching evaluation and effectively contributing to the improvement of vocational education quality.

Acknowledgements

This work is supported by the 2024 Henan Province Higher Education Teaching Reform Research and Practice Project named: Research and Practice on the 'Three Iterations, Four Dimensions, and Five Collaborations' Digital Intelligence Talent Training Model in Higher Vocational Colleges in Henan Province (No. 2024SJGLX0669).

Declarations

All authors declare that they have no conflicts of interest.

References

Ai, L. and Feng, T. (2022) 'Application of BP neural network in teaching quality evaluation of higher vocational education', *Mobile Information Systems*, Vol. 20, No. 11, pp.24–32.

Cai, Z. (2023) 'Optimizing vocational education design with big data and educational technology', *International Journal of Emerging Technologies in Learning*, Vol. 18, No. 24, pp.164–179.

Chen, C. and Song, P. (2023) 'Evaluation method of diversified teaching effect of higher vocational education based on knowledge map', *International Journal of Sustainable Development*, Vol. 26, Nos. 3–4, pp.318–328.

Forni, M., Hallin, M., Lippi, M. and Zaffaroni, P. (2015) 'Dynamic factor models with infinite-dimensional factor spaces: one-sided representations', *Journal of Econometrics*, Vol. 185, No. 2, pp.359–371.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2020) 'Generative adversarial networks', *Communications of the ACM*, Vol. 63, No. 11, pp.139–144.

Guleria, P. and Sood, M. (2023) 'Explainable AI and machine learning: performance evaluation and explainability of classifiers on educational data mining inspired career counseling', *Education and Information Technologies*, Vol. 28, No. 1, pp.1081–1116.

Hassan, S. and Shamsudin, M.F. (2019) 'Measuring the effect of service quality and corporate image on student satisfaction and loyalty in higher learning institutes of technical and vocational education and training', *International Journal of Engineering and Advanced Technology*, Vol. 8, No. 5, pp.533–538.

Hou, J. (2021) 'Online teaching quality evaluation model based on support vector machine and decision tree', *Journal of Intelligent & Fuzzy Systems*, Vol. 40, No. 2, pp.2193–2203.

Hussain, S., Muhsion, Z.F., Salal, Y.K., Theodorou, P., Kurtoğlu, F. and Hazarika, G. (2019) 'Prediction model on student performance based on internal assessment using deep learning', *International Journal of Emerging Technologies in Learning*, Vol. 14, No. 8, pp.14–20.

Ilić, M.P., Păun, D., Popović Šević, N., Hadžić, A. and Jianu, A. (2021) 'Needs and performance analysis for changes in higher education and implementation of artificial intelligence, machine learning, and extended reality', *Education Sciences*, Vol. 11, No. 10, p.568.

Li, H.X. and Xu, L.D. (2001) 'Feature space theory – a mathematical foundation for data mining', *Knowledge-Based Systems*, Vol. 14, Nos. 5–6, pp.253–257.

Li, Z., Liu, F., Yang, W., Peng, S. and Zhou, J. (2021) 'A survey of convolutional neural networks: analysis, applications, and prospects', *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 33, No. 12, pp.6999–7019.

Long, C. and He, J. (2023) 'Application of big data in entrepreneurship and innovation education for higher vocational teaching', *International Journal of Information Technology and Web Engineering*, Vol. 18, No. 1, pp.1–16.

Lu, Y. and Wang, T. (2023) 'Quality evaluation model of vocational education in China: a qualitative study based on grounded theory', *Education Sciences*, Vol. 13, No. 8, p.819.

Ma, W. and Xu, F. (2021) 'Study on computer vision target tracking algorithm based on sparse representation', *Journal of Real-Time Image Processing*, Vol. 18, No. 2, pp.407–418.

Nadzinski, G., Gerazov, B., Zlatinov, S., Kartalov, T., Dimitrovska, M.M., Gjoreski, H., Chavdarov, R., Kokolanski, Z., Atanasov, I. and Horstmann, J. (2023) 'Data science and machine learning teaching practices with focus on vocational education and training', *Informatics in Education*, Vol. 22, No. 4, pp.671–690.

Wang, P. (2018) 'Factor space-mathematical basis of mechanism based artificial intelligence theory', *CAAI Transactions on Intelligent Systems*, Vol. 13, No. 1, pp.37–54.

Wang, Z., Xu, T. and Wang, L. (2022) 'Teaching evaluation method based on fuzzy support vector machine algorithm', *Mobile Information Systems*, Vol. 22, No. 12, pp.15–23.

Weiss, K., Khoshgoftaar, T.M. and Wang, D. (2016) 'A survey of transfer learning', *Journal of Big Data*, Vol. 3, pp.1–40.

Wu, S. (2024) 'Application of multimedia technology to innovative vocational education on learning satisfaction in China', *Plos One*, Vol. 19, No. 2, p.e0298861.

Yu, Y. (2013) 'Research on higher tourism vocational education and management based on the factor analysis', *Research Journal of Applied Sciences, Engineering and Technology*, Vol. 5, No. 24, pp.5542–5549.

Zhang, Q. (2022) 'A novel ResNet101 model based on dense dilated convolution for image classification', *SN Applied Sciences*, Vol. 4, pp.1–13.

Zhang, Y., Sun, X. and Yu, J. (2024) 'transformative technologies in the evaluation of a vocational education system', *Journal of Web Engineering*, Vol. 23, No. 2, pp.275–298.

Zhou, Z. and Liu, Y. (2023) 'Political teaching application in high vocational care courses based on machine learning systems', *Soft Computing*, Vol. 27, No. 11, pp.7657–7666.

Zhu, Y., Min, W. and Jiang, S. (2020) 'Attribute-guided feature learning for few-shot image recognition', *IEEE Transactions on Multimedia*, Vol. 23, pp.1200–1209.