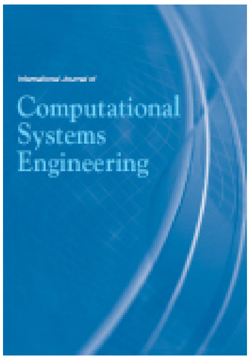


International Journal of Computational Systems Engineering

ISSN online: 2046-3405 - ISSN print: 2046-3391
https://www.inderscience.com/ijcsyse

A transfer learning-based model for assessing university students' innovation and entrepreneurship

Yiping Zheng

DOI: 10.1504/IJCSYSE.2022.10055142

Article History:

Received:	28 November 2022
Last revised:	26 December 2022
Accepted:	01 March 2023
Published online:	19 March 2024

A transfer learning-based model for assessing university students' innovation and entrepreneurship

Yiping Zheng

Cooperative School of International Education,
Tianjin University of Commerce,
Tianjin, 300134, China
Email: tucfiuzyp@tjcu.edu.cn

Abstract: In the current trend of innovation and entrepreneurship, the number of students who start their own innovation and entrepreneurship is increasing. In order to enable university students to assess their own innovation and entrepreneurship ability and avoid greater risks, the study starts with a convolutional neural network (CNN) based on migration learning, which is used to establish an assessment model of innovation and entrepreneurship ability. The model is based on migration learning and adversarial networks to strengthen the learning ability and incorporate game theory, with similar adversarial learning of classifiers and generators in the migration, so that the model has stronger learning ability and faster computing speed, and avoids problems such as over-convergence and excessive degrees of freedom. A variational self-encoder is used on the encoder to further improve the accuracy and precision of the input data recognition by compressing the information bottleneck.

Keywords: transfer learning; convolutional neural network; CNN; evaluation model; innovation and entrepreneurship.

Reference to this paper should be made as follows: Zheng, Y. (2024) 'A transfer learning-based model for assessing university students' innovation and entrepreneurship', *Int. J. Computational Systems Engineering*, Vol. 8, Nos. 1/2, pp.66–74.

Biographical notes: Yiping Zheng graduated from the Tianjin University of Commerce with a Bachelor's in Hotel Management in 2014. In 2018, she graduated from the Tianjin University of Commerce, majoring in Ideological and Political Education, with a Master's in Law. In 2022, she began her Doctoral study at the School of Marxism of Shanghai Jiao Tong University and is also a teacher at Tianjin University of Commerce. Her interests include Marxist theory, ideological and political education, college students' innovation and entrepreneurship.

1 Introduction

In the context of global economicisation, entrepreneurship is prevalent, and with today's ever-evolving technology and economic models, university students are increasingly innovative and entrepreneurial (Yuan and Hu, 2021). However, many university students are unprepared in terms of their abilities due to their lack of experience and qualifications, resulting in the eventual failure of their ventures, which results in greater losses. Innovation and entrepreneurship among university students also harbour economic development, and if too many university students fail to start their own businesses due to a lack of preparation and knowledge of their own entrepreneurial abilities, this also has a significant negative impact on economic development (Wang, 2020). To address this phenomenon, there is a need to create an assessment model to evaluate university users so that they can have a clearer perception and orientation of their own innovation and entrepreneurial abilities. Zhang and Xi (2021) combined the fuzzy analytic hierarchy process (AHP) and the BP neural network (GA-BPNN) optimised based on genetic algorithm (GA) to build an evaluation model of college students' innovation and entrepreneurship thinking ability to evaluate college students' innovation and entrepreneurship thinking ability. During the experiment, the validity of the model was verified. Zhou and Zhou (2022) based on the analysis of the current research status of innovation and entrepreneurship education quality evaluation, selected 21 indicators to build the evaluation index system of innovation and entrepreneurship education quality of college students. Subsequently, based on the evaluation index system, an evaluation model based on extenics was constructed and an empirical study was conducted. Based on the grounded theory, Xu (2022) built an evaluation system containing 29 indicators, and thus proposed a three-dimensional VPR evaluation model for innovation and entrepreneurship education. However, the current evaluation methods of innovation and entrepreneurship have some problems, such as low efficiency and poor accuracy. In view of this, the study started with convolutional neural networks (CNNs), which are suitable for evaluation models as they are suitable for all types of feature extraction (Chandra et al., 2019). To enhance its performance, migration learning in machine learning is performed on top of the CNN and its learning capability is further enhanced using the idea of adversarial in migration learning. The study expects that the adversarial CNN based on migration learning can enhance the accuracy and learning speed based on the advantages of CNN and strengthen the absorption of input data in order to be able to meet the requirements of accurate assessment, and apply the model to the assessment of the innovation and entrepreneurship ability of university students and play a role for university students' innovation and entrepreneurship. This model enriches the evaluation of innovation and entrepreneurship education, provides a theoretical reference for its practice, and is of great significance for solving the employment problem of college students and the development of market economy.

2 Related work

Many studies have been conducted at home and abroad on migration learning and related model evaluation research. He et al. (2022) proposed a migration learning method based on a deep multi-signal fusion adversarial model in order to solve the problem that fault diagnosis of axial piston pumps is difficult to complete. The method designed a multi-signal fusion module and embedded the residual network into the shared feature generation module to obtain rich feature information. The results showed that the method was accurate up to 98.5%. Sahinovic and Geisler (2022) combined density flooding theory simulations and migratory learning with elemental embedded neural networks to explore the formation energy and lattice parameters of vacancy layers in ABXn infinite layers and chalcogenide nitrides, oxides and fluorides, and the prediction accuracy was significantly improved. Liao et al. (2021) to achieve real-time intelligent manufacturing process monitoring, acoustic emission signals were converted into time-frequency and image formats for process classification based on the spindle speed, feed rate and depth of cut of the operation. The developed deep neural network for image classification was applied using a migration learning approach, and simulation results showed excellent classification accuracy of 95.58%. Wang et al.'s (2021) team, in order to be able to accurately classify diseases caused by COVID-19, used migration learning to allow the machine to perform chest CT image feature extraction and recognition to distinguish between different types of pneumonia. The experimental results showed that the combined accuracy of this method reached over 97.5% with high accuracy. Wang et al. (2022) proposed a fault diagnosis method combining deep learning and migration learning to enable accurate diagnosis of system-level faults in nuclear power plants by introducing an effective migration learning method that adequately reduces the differences in feature distribution between different power levels By introducing an effective migration learning method, the differences in feature distribution between different power levels are sufficiently reduced to achieve fault diagnosis at the target power level. The results show that the method is significantly effective in data discrepancy reduction.

There are also numerous studies on CNN for feature extraction. Han et al. (2020) used CNN in deep learning to extract typical features from images to accomplish multi-target recognition of remote sensing big data in complex scenes, and the results of the study showed that the algorithm has an accuracy of more than 80%, which is advantageous in target recognition. Alves et al.'s (2021) team investigated the effect of data enhancement techniques to improve the performance of CNNs by classifying anomalies between 11 different classes in the PV module from thermographic images in an unbalanced dataset. The study yielded a test accuracy estimate of 92.5% for the improved CNN through a cross-validation approach, which is an improvement over the conventional method. Ma et al. (2019) designed a new deep spanning CNN structure for fault diagnosis based on lightweight modelling requirements and techniques. Experimental results show that the algorithm outperforms existing traditional algorithms in terms of accuracy, storage space, computational complexity, noise immunity, and transmission performance. Foroughi et al. (2017) introduced a joint projection based on pairwise graph constraints and a low-order dictionary in order to avoid the adverse effects of illumination, occlusion, and disparity in the sample set on the classification system learning methods to ensure strong robustness while reducing redundancy in the execution structure. Gao et al. (2021) proposed a novel fault diagnosis method based on data self-generation and deep CNN, where the model can directly transform data into digital images and can improve system accuracy. Simulation results demonstrated that the method can effectively increase diagnostic information and help improve performance compared to traditional CNN. Wang et al. (2018) designed an automatic sorting system based on a deep learning method and applied region-based CNNs to locate and identify different types of images of industrial object models, and simulation experiments showed that the system can achieve efficient automated sorting of mechanical parts with robustness of the network. Ganesan et al.'s (2020) team took a CNN model in exploring the satellite fault problem, subjected it to training on satellite telemetry data, explored various processing schemes in the time and frequency domains to process the data input to the CNN, and the experimental results showed that it obtained high-quality classification results.

As can be seen from the above studies, there are numerous studies on migration learning and CNNs, and many of them have been applied to feature recognition for evaluation. It is feasible to combine migration learning with CNN for the creation of evaluation models.

3 Improved algorithms for adversarial CNN based on migration learning and their applications

3.1 Building a CNN based on migration learning

Transfer learning is a kind of machine learning. It refers to that machines learn new knowledge through existing knowledge to achieve the effect of drawing inferences from one instance. Transfer learning is suitable for expansion operations, and transfer learning can be used for operations with similarity or extensibility. The evaluation model is a typical classification model, which has a high degree of agreement with transfer learning. Therefore, it is feasible to apply the basic ideas and methods of transfer learning to the evaluation model of college students' innovation and entrepreneurship ability.

Among the evaluation models, perceptual machines are one of the commonly used ones. The input of a perceptual machine represents a multidimensional vector of the corresponding data, while the output is the discriminant of the corresponding data (Masykuri et al., 2021). A perceptual machine is a linear model for performing a binary category output that finds a high-dimensional spatial hyperplane of the input data, as shown in equation (1).

$$wx + b = 0 \tag{1}$$

In equation (1), w is the normal vector of the hyperplane in the high-dimensional space of the input data, and b is the intercept of the high-dimensional space of the input data. The classification method of the perceptron is to classify the found hyperplanes into two categories, i.e., $f(x) < 0$ and $f(x) > 0$, in order to achieve the classification effect. The perceptron learns the values of w and b to perform the classification discriminations. This is expressed as a symbolic function, as shown in equation (2).

$$sign(x) = \begin{cases} +1, & x \geq 0 \\ -1, & x < 0 \end{cases} \tag{2}$$

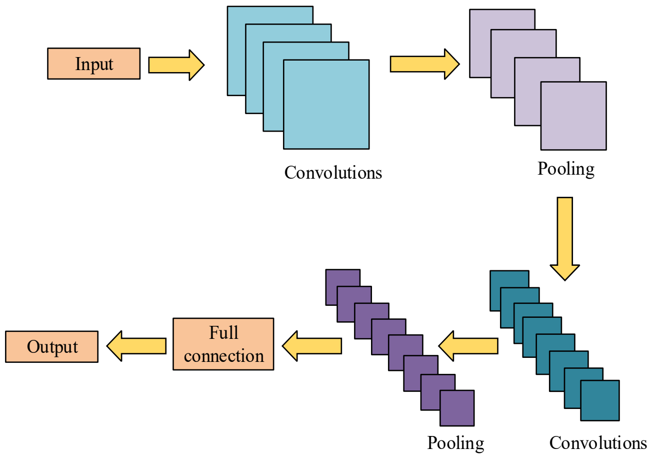
In order to determine the two parameters of the hyperplane, i.e., w and b, the loss function needs to be defined and minimised as much as possible, and the function needs to be continuously derivable, so the loss function needs to be chosen to minimise the sum of the total distances from the classified points in high-dimensional space to the binary hyperplane in high-dimensional space, as shown in equation (3). In equation (3) $L(w, b)$ represents the loss function of the parameters w and b, (x_i, y_i) are the coordinates of any point in the hyperplane, x_i and y_i are the corresponding coordinate values, and M represents the set of classification points.

$$L(w, b) = -\sum_{x_i \in M} y_i \left(wx_i + b \right) \tag{3}$$

The size of the loss function value reflects the classification accuracy of the perceptron, with smaller values representing fewer errors and more accurate classification.

In current machine learning, transfer learning, a large class of learning methods, including artificial neural network (ANN), is widely used. One of the basic learning models used in ANNs is the CNN, which consists of a convolution operator, a convolution feature, an activation function, a convolution kernel, a local sense, a convolution layer and a pooling layer, and is structured as a convolution layer, a pooling layer and a fully connected layer. Each convolutional layer consists of a different number of convolutional units, the parameters of which are derived from the loss function by back propagation. The pooling layer is used to pool the high-dimensional vectors obtained from the multi-layer convolution, i.e., to partition the high-dimensional vectors and reduce the dimensionality by taking the maximum, minimum and average values in each region. The main function of the fully connected layer is to sum all local features nonlinearly to form a global feature and use the global feature for classification (Deng et al., 2021). The basic structure of CNN is shown in Figure 1.

Figure 1 Basic structure diagram of CNN (see online version for colours)



After the output of the convolutional layer is fed to the activation layer, the activation function performs a nonlinear mapping of the convolutional layer, which allows the convolutional layer to extract more abstract features and thus enhance the functionality of the CNN (Kh et al., 2022). The activation function is usually a Sigmoid function, which is shown in equation (4).

$$h_\theta(l) = \frac{1}{1 + e^{-\theta l}} \tag{4}$$

In equation (1), θ is the mapping of the Sigmoid function and l indicates the number of layers corresponding to the value of the function. The essence of the convolution operation is the process of extracting valid features from the initial feature map through the convolution layers. Let the initial features of each convolutional layer with input be x_j, then the convolution operation is shown in equation (5). In equation (5), $f(x)$ is the activation function, M_j denotes the set of initial features, i is the matching result, k_{ij} the input of the i^{th} initial feature and the convolution kernel of the output result of the j^{th} initial feature of the convolution layer.

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} \cdot k_{ij}^l + c_j\right) \tag{5}$$

The Sigmoid operation resulted in the ω feature map f_ω as shown in equation (6). In equation (6), x is the input value, W is the weight and b represents the bias.

$$f_\omega = sigm\left(W^\omega x + b^\omega\right) \tag{6}$$

When the weights and update values of all neurons on the l layer need to be solved, the sensitivity magnitude at all nodes needs to be solved first, and the value is noted as θ. The sum of the sensitivity values defined by interest from the connectivity layer l to $l + 1$ is θ_j^{l+1}, multiplied by the corresponding weights W, and then the activation function is obtained by taking the inverse of $f(u^l)$, as shown in equation (7). In equation (7) u is the input value of the neurons in the l layer.

$$\theta_j^l = \theta_j^{l+1} W_j^{l+1} \cdot f\left(u^l\right) \tag{7}$$

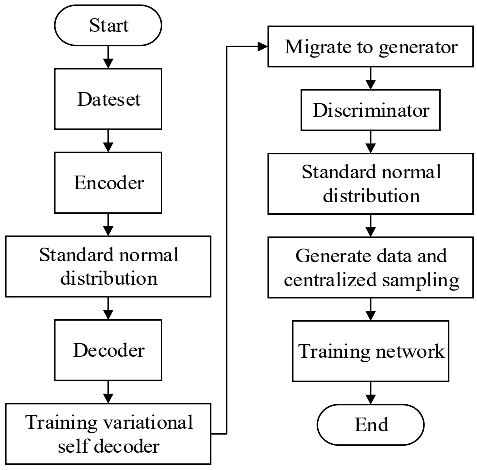
CNN based on transfer learning (TL-CNN) is a neural network that uses transfer learning as the theoretical basis for convolutional operations. There are three main types of transfer learning, namely, direct transfer learning, inductive transfer learning and unsupervised transfer learning. Direct transfer learning is mainly used when the original task and the target task are of the same type and the target and source domains are clearly linked. When using direct transfer learning, a sufficiently large number of unlabelled samples from the target domain and labelled samples from the source domain are required. Inductive transfer learning is mainly used when the original task and the target task type are different and there are no additional requirements for the target and source domains. Inductive learning migration is a knowledge-based summary of task features and can be divided into instance knowledge migration, feature knowledge migration, parameter knowledge migration and related knowledge migration. Unsupervised migration is mainly applicable when the original task and the learning target task type are different and there are no labelled samples in the target and source domains (Frid-Adar et al., 2018). Taking the nature of the three types of transfer learning together, the study mainly used unsupervised transfer learning.

Under unsupervised migration learning, the TL-CNN is based on the Bayesian formulation shown in equation (8). In equation (8) x is the input variable, y denotes the target variable, $p(y| x)$ denotes the class post-test probability, and $p(y| x)$ denotes the class conditional probability.

$$p(y \mid x) = \frac{p(x \mid y)p(x)}{p(x)} \tag{8}$$

TL-CNN performs deep learning by selecting the variational self-encoder as the source model in the first step, after which the model and model parameters are saved and the decoder is used as the starting point of the network. The whole network is adjusted so that the input of the generated model satisfies the standard normal distribution, and then iterated continuously. The whole process of TL-CNN is shown in Figure 2.

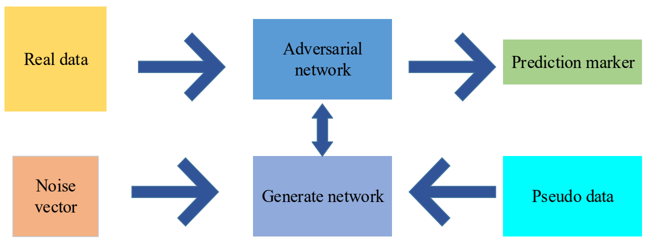
Figure 2 Basic flow chart of TL-CNN



3.2 *Improved algorithm based on countermeasure network and its application in evaluation model*

In machine learning models, the concept of adversarial networks is introduced in order to obtain more accurate and more credible data. As TL-CNN may have problems such as difficult model convergence, long model training time and uncontrollable model freedom, the research adopts an adversarial network based on TL-CNN to form a CNN based on transfer The aim is to improve the algorithm to compensate for these shortcomings.

Figure 3 Basic structure diagram of adversarial network (see online version for colours)



Adversarial networks are a theory based on game theory that learns the distribution of source data by combining a generative model with a discriminant model in an iterative training approach (Fu et al., 2021). In the network, the dichotomiser uses a discriminant model to determine whether the input data belongs to real or generated data, and to back-propagate the gradient on this basis. The generated data of the generative model is thus constantly approximated to the real data. The main advantages of adversarial networks are that they do not require pre-modelling and are simple in structure, and that only discriminators and generators are required for the entire network. The whole adversarial process is a continuous game between the discriminator and the generator, with the end point of the training being to find a Nash equilibrium between the two. The basic structure of the adversarial network is shown in Figure 3.

In Figure 3, the generator and discriminator are fitted with a CNN for the function. Let the fitting function of the discriminator be D, the fitting function of the generator be, the input data of the generator be Gr, and r be a low-dimensional random variable. The input to the discriminator has the real data as and the generated data as $xG(r)$. The goal of the generator is to make the generated data and the real data as identical in distribution as possible, so the generator and the discriminator will play each other during training and eventually reach a Nash equilibrium. Throughout the algorithmic process of the TLA-CNN, a small batch of stochastic gradient algorithms is first generated. The training number of the discriminator is, the small batch kn samples are selected from the data of the random noise distribution and the generative distribution, the first n samples are r_n and x_n respectively, and the random gradient rise to update the discriminator parameters is calculated as shown in equation (9).

$$k_D = \theta_D \frac{1}{n} \sum_{i=1}^{n} \left[\log D(x_i) + \log\left(1 - D\left(G(r_i)\right)\right) \right] \tag{9}$$

In equation (9), i represents the order of all samples in the corresponding set, is the x_i, i^{th} in the set of generated distribution data, and r_i is the i^{th} in the random distribution data. The parameters of the updated generator are then calculated using stochastic gradient descent, as shown in equation (10).

$$k_G = \theta_G \frac{1}{n} \sum_{i=1}^{n} \log\left[(1 - D\left(G(r_i)\right) \right] \tag{10}$$

The whole process of generating discriminators and generators, and updating parameters as described above, is shown in Figure 4.

The loss function of the TLA-CNN is then a maximum-minimum optimisation problem. For the generator G, the desired generating distribution is $p_r(r)$, i.e., satisfying a random noise distribution. The classifier D can directly measure the difference between $p_r(r)$ and $p(x)$. The objective function is therefore shown in equation (11).

$$\begin{aligned} \min_G \max_D f(D, G) &= E_{x \sim p(x)}[\log D(x)] \\ &\quad + E_{r \sim p(r)}[\log(1 - D(G(r)))] \end{aligned} \tag{11}$$

This optimisation will split equation (11) and optimise D and G in turn, as shown in equations (12) and (13) respectively.

$$\begin{aligned} \max_D f(D, G) &= E_{x \sim p(x)}[\log D(x)] \\ &\quad + E_{r \sim p(r)}[\log(1 - D(G(r)))] \end{aligned} \tag{12}$$

$$\min_G f(D, G) = E_{r \sim p(r)}[\log(1 - D(G(r)))] \tag{13}$$

In the TLA-CNN self-encoder, a variational self-encoder is used. Variational self-encoders utilise CNNs for representation learning, enabling knowledge compression of the original input so that it can be represented in a neural network architecture. This architecture is highly applicable if there is a strong correlation between the input features in the data and is able to pass the bottleneck of the network when forcing the input. The key to variational self-encoders is the bottleneck, and the information bottleneck can be minimised by minimising the reconstruction error. When the information bottleneck does not exist, the gap between the original input and the reconstruction will not exist and the whole network will only learn to remember the input values. The basic structure of a variational self-encoder is shown in Figure 5.

Figure 4 Basic flow chart of TLA-CNN

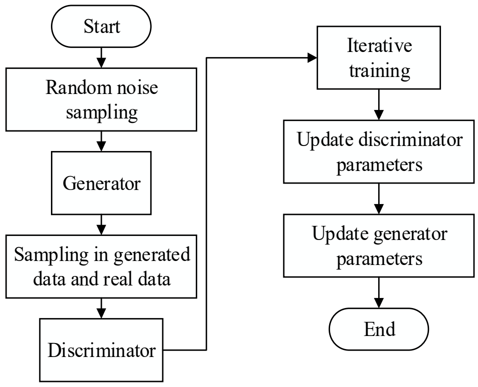
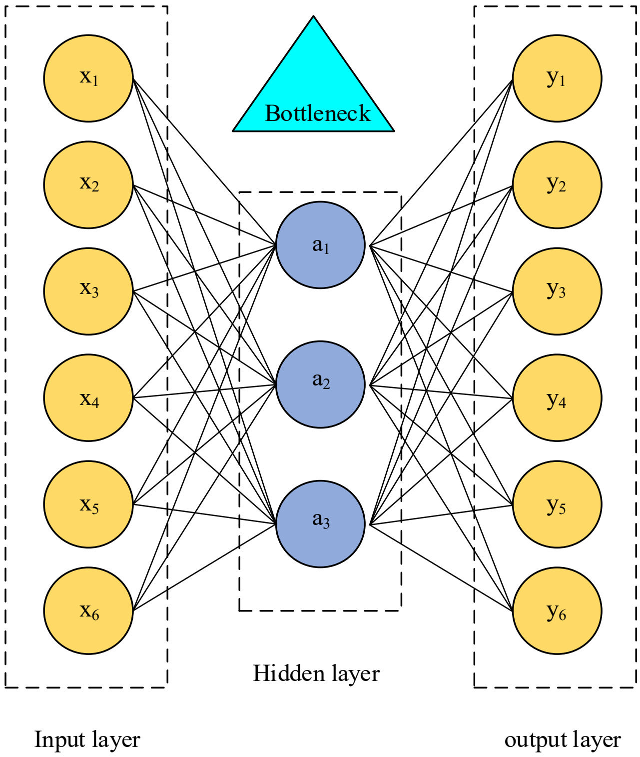


Figure 5 Basic structure diagram of variational self encoder (see online version for colours)



Due to the presence of bottlenecks, the amount of information in the input is suppressed, thus forcing the entire network to compress and learn from the input data. The variational self-encoder is sensitive to the input as a way to accurately build the reconstruction, but the sensitivity is not excessive and when excessive it tends to lead to over-simplification of the model, resulting in over-fitting of the training data, etc. To derive the loss function for the variational self-encoder, the maximum likelihood method was first used to estimate the learning parameters θ, and using the learning parameters θ, a lower bound on the likelihood function was derived as shown in equation (14).

$$\log L = \log p(x) \int_r q(r|x) dr \tag{14}$$

In equation (14), L represents the number of lower bound layers of the likelihood function and q represents the class reconstruction probability. The lower bound reaches its maximum value when L reaches its minimum. To arrive at the value of $q(r|\ x)$, a suitable distribution needs to be chosen, i.e., one of the Bernoulli distribution or the normal distribution. It is known from arithmetic that the normal distribution is more suitable for deriving the loss function and the results are simpler. The loss function of the variational self-encoder is shown in equation (15). In equation (15), μ is the function of the neural network with respect to the mean, k is a function constant based on a fixed variance under a normal distribution, and t represents the number of iterations.

$$\begin{aligned} fLoss = \min_{\theta} \theta \|x - f(r)\|^2 + \frac{1}{2}\Big[\Big(tr\sum(x)\Big) + \mu(x)^t \mu(x)\Big] \\ -k - \log \det \sum(x) \end{aligned} \tag{15}$$

The entire TLA-CNN utilises multiple features of the university student user to identify and extract as many features as possible through convolution, and enhances the learning of the features using migration learning and adversarial, resulting in a taxonomy-based assessment model in a machine learning manner. The final classification of the machine will assess whether each university user is able to directly engage in innovative entrepreneurship based on a combination of features, differentiated by 0 and 1, with 0 indicating no and 1 indicating yes. To verify the effectiveness of the improved model, the sample data needs to be pre-processed first, i.e., standardised normalisation. Afterwards, 75% of the sample data is used as the training set for the model to perform machine learning for training, and the remaining 25% of the sample data is used as the test set to test the model for multiple performance results. In order to make the result data more comparable, the experiments were tested on both CNN and RCNN algorithms in the same way to obtain comparative results of the improved algorithms.

4 Performance tests and analysis of results under multiple model comparisons

The hardware environment for the performance trials was an I7-8750 processor; 16G RAM, 2T hard disk, and the programming environment was PYTHON. The performance trials of the TLA-CNN model were conducted in two main directions, firstly, the analytical capabilities and metrics of various aspects of the TLA-CNN model, and secondly, the practical application effects of the model in the assessment of university students' innovation and entrepreneurship. To determine the training effectiveness of the model, the performance of the model in terms of accuracy and loss values in the training ensemble test set was recorded, as shown in Figure 6.

Figure 6 depicts the accuracy of the model in Figure 6(a) and the loss value of the model in Figure 6(b). As can be seen from the accuracy rates, the training set accuracy eventually stabilises at 100%, while the model's performance in the test set is closer to that of the training set, reaching a maximum of 94.63%. After the model was trained, the performance of the model was evaluated using a full range of neural network evaluation metrics, and the results are shown in Table 1.

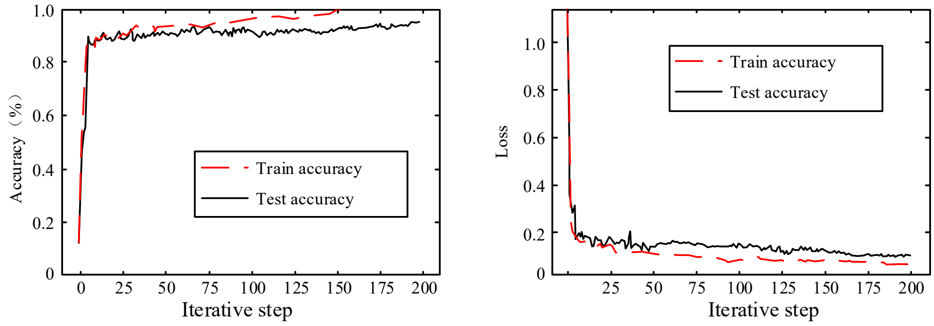
Table 1 Evaluation results of TLA-CNN model indicators

Algorithm/item	*Accuracy (%)*	*Precision (%)*	*Recall (%)*	*F1*
TLA-CNN	92.92	84.13	88.67	0.79
CNN	81.08	76.19	73.33	0.61
RCNN	87.64	78.47	84.67	0.73

The evaluation metrics in Table 1 include accuracy, precision, recall and F1 value, and the traditional CNN model and RCNN model were used to compare and analyse with the TLA-CNN model. The evaluation results show that the TLA-CNN model has the best performance in all four metrics. In terms of accuracy, the TLA-CNN achieves 92.92, while the RCNN and traditional CNN are 87.64 and 81.08 respectively. In addition, the recall rate of the TLA-CNN model reaches 88.67%, which is 4% higher than that of the traditional CNN. Combining the various data in Table 1, it can be seen that the TLA-CNN model has an all-round advantage over the traditional model. In addition to several metrics such as accuracy, the model was also tested for receiver operating characteristic curve (ROC) and the results are shown in Figure 7.

In addition to the TLA-CNN model itself, the ROC data of the traditional CNN and the RCNN model are also used here as the basis for comparative analysis. Looking at the images of the ROC curves, it can be seen that the ROC curve of the TLA-CNN model lies in the upper left of the other two models overall, furthest from the pure chance line, and the value of the AUC is also significantly larger than the other models, indicating that the TLA-CNN shows better judgement than the other models. In practice, the level of error of the neural network model can also have an impact on its application effectiveness, so the error metrics of the TLA-CNN model need to be measured and evaluated, and the results are shown in Figure 8.

Figure 6 Accuracy and loss value data of TLA-CNN model (see online version for colours)



(a) (b)

Figure 7 ROC analysis results of TLA-CNN model (see online version for colours)

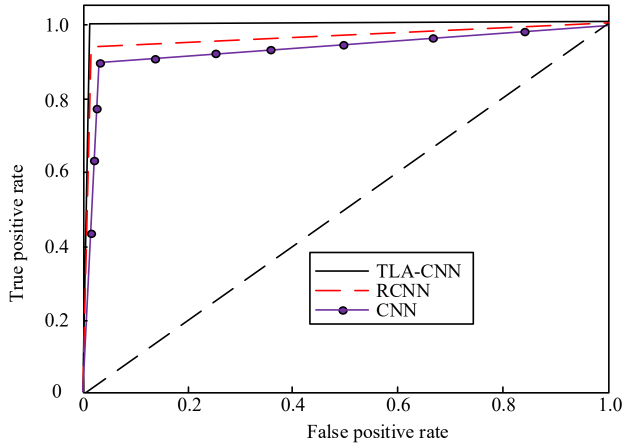
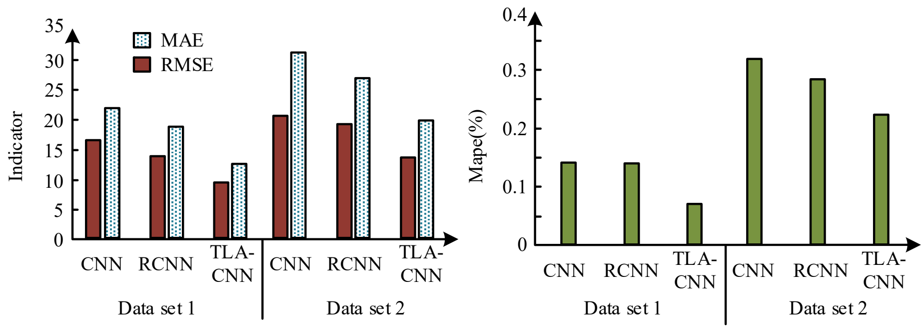
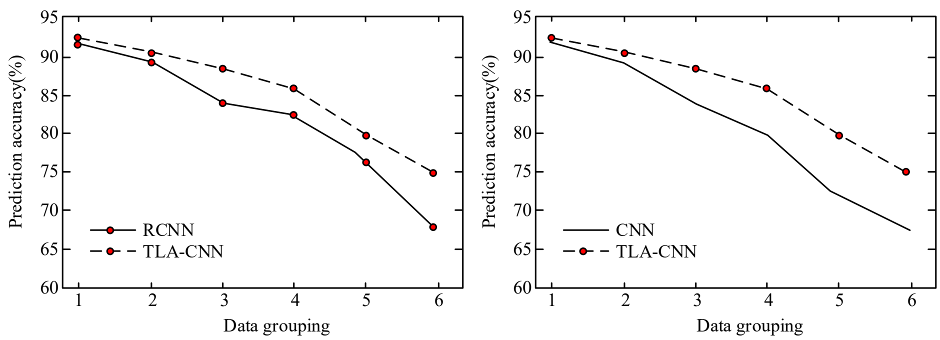


Figure 8 Error index evaluation results of TLA-CNN model (see online version for colours)



(a) (b)

Figure 9 Experimental results of TLA-CNN model's innovation and entrepreneurship ability evaluation (see online version for colours)



(a) (b)

The main error metrics used are mean absolute error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE). Figure 8(a) depicts the MAE and RMSE errors of the model under different data sets, and Figure 8(b) depicts the variation of MAPE under unused data sets. The three error metrics showed consistency in the test, i.e., the error metrics of TLA-CNN, RCNN and traditional CNNs fluctuated widely between different datasets, but when compared within the same dataset, it can be seen that TLA-CNN showed the least error profile under all three metrics of MAE, RMSE and MAPE. The TLA-CNN model showed only 0.71 MAPE in dataset 1 The MAPE of the TLA-CNN model in dataset 1 is only 0.71%, while the MAPE of both CNN and RCNN in this dataset is around 1.4, which shows that the TLA-CNN model also has a greater advantage over traditional algorithms in terms of error. After completing the evaluation of the performance of the neural network model, it is necessary to evaluate it from the perspective of practical application. For this purpose, a total of six datasets of student innovation and entrepreneurship information of different difficulty levels were searched for, and the TLA-CNN model and other comparison models were used to evaluate the students' innovation and entrepreneurship ability, and the evaluation results were compared with the expert evaluation results of the innovation and entrepreneurship teachers to observe the compliance. The comparison results are shown in Figure 9.

Figure 9(a) shows the results of TLA-CNN versus RCNN, and Figure 9(b) shows the results of TLA-CNN versus traditional CNN. It can be seen that all three models show a trend of decreasing accuracy as the difficulty and complexity of the dataset increases, but the accuracy of the TLA-CNN model is consistently higher than that of the other two models, and the highest evaluation accuracy in practical use is 92.67%.

5 Conclusions

With the development and transformation of society, as well as changes in the employment situation and employment pressure, the ability to innovate and start a business has become an important ability that affects the career development and comprehensive quality of university students. From the social level, the university student group has a high level of knowledge and creative ability, and is also the main group of people who meet the requirements of national entrepreneurship planning. Therefore, it is very important for universities to cultivate and assess the innovative and entrepreneurial abilities of university students. In view of this, this study constructs an assessment model for the innovation and entrepreneurship ability of university students based on transfer learning using an improved TLA-CNN model. The experimental results on the model showed that the accuracy of the training set of the model reached up to 100%, while the accuracy of the test set also reached 94.63%, showing a more reliable level of accuracy. In terms of error, the model achieves a minimum MAPE of 0.71%, which is significantly lower than that of traditional CNNs. In terms of practical application results, the TLA-CNN model achieves a maximum evaluation accuracy of 92.67% in actual use, and its accuracy rate is consistently higher than that of the traditional convolutional neural network model. The TLA-CNN model based on migration learning for the assessment of university students' innovation and entrepreneurship has proved to have better performance, but there is still room for improvement. The TLA-CNN model has a more obvious decline in assessment ability when facing data sets with higher complexity, so the resistance of the neural network model to data complexity should be improved for this problem in the next research.

Acknowledgements

The research is supported by: University-Industry Collaborative Education Program, Construction of the Practice System for Finance and Economics Students under the background of international education, No. 202102456044.

References

Alves, R., Gadd, J., Marra, E.G. and Lemos, R.P. (2021) 'Automatic fault classification in photovoltaic modules using convolutional neural networks', *Renewable Energy*, Vol. 179, No. 4, pp.502–516.

Chandra, B.S., Saatry, S. and Jana, S. (2019) 'Robust heartbeat detection from multimodal data via CNN-based generalizable information fusion', *IEEE Transactions on Biomedical Engineering*, Vol. 66, No. 3, pp.710–717.

Deng, H., Zhang, W.X. and Liang, Z.F. (2021) 'Application of BP neural network and convolutional neural network (CNN) in bearing fault diagnosis', *IOP Conference Series: Materials Science and Engineering*, Vol. 1043, No. 4, pp.42026–42034.

Foroughi, H., Ray, N. and Zhang, H. (2017) 'Object classification with joint projection and low-rank dictionary learning', *IEEE Transactions on Image Processing*, Vol. 27, No. 2, pp.806–821.

Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J. and Greenspan H. (2018) 'GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification', *Neurocomputing*, Vol. 321, No. 10, pp.321–331.

Fu, J., Li, W., Du, J. and Xu, L. (2021) 'DSAGAN: a generative adversarial network based on dual-stream attention mechanism for anatomical and functional image fusion', *Information Sciences*, Vol. 576, No. 9, pp.484–506.

Ganesan, M., Lavanya, R. and Devi, M.N. (2020) 'Fault detection in satellite power system using convolutional neural network', *Telecommunication Systems*, Vol. 4, No. 2, pp.101–107.

Gao, J., Han, H., Ren, Z. and Fam, Y. (2021) 'Fault diagnosis for building chillers based on data self-production and deep convolutional neural network', *Journal of Building Engineering*, Vol. 34, No. 4, pp.102–113.

Han, X., Jiang, T., Zhao, Z. and Lei, Z. (2020) 'Research on remote sensing image target recognition based on deep convolution neural network', *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 34, No. 5, pp.1789–1828.

He, Y., Tang, H., Ren, Y. and Kumar, A. (2022) 'A deep multi-signal fusion adversarial model-based transfer learning and residual network for axial piston pump fault diagnosis', *Measurement*, Vol. 192, No. 31, pp.110889–110902.

Kh, A., Yw, A., Wl, B. and Lw, C. (2022) 'CNN-BiLSTM enabled prediction on molten pool width for thin-walled part fabrication using laser directed energy deposition', *Journal of Manufacturing Processes*, Vol. 2022, No. 78, pp.32–45.

Liao, Y., Ragai, I., Huang, Z. and Kerner, S. (2021) 'Manufacturing process monitoring using time-frequency representation and transfer learning of deep neural networks', *Journal of Manufacturing Processes*, Vol. 68, No. 8, pp.231–248.

Ma, S., Cai, W., Liu, W., Shang, Z. and Liu, G. (2019) 'A lighted deep convolutional neural network based fault diagnosis of rotating machinery', *Sensors (Basel, Switzerland)*, Vol. 19, No. 10, pp.2381–2394.

Masykuri, W.S., Khatizah, E. and Bukhari, F. (2021) 'The application of perceptron method in predicting student graduation based on several identified key factors', *Journal of Physics Conference Series*, Vol. 1796, No. 1, pp.12060–12067.

Sahinovic, A. and Geisler, B. (2022) 'Quantifying transfer learning synergies in infinite-layer and perovskite nitrides, oxides, and fluorides', *Journal of Physics: Condensed Matter*, Vol. 34, No. 21, pp.214003–214014.

Wang, S.H., Nayak, D.R., Guttery, D.S., Zhang, X. and Zhang, Y.D. (2021) 'COVID-19 classification by CCSHNet with deep fusion using transfer learning and discriminant correlation analysis', *Information Fusion*, Vol. 68, No. 8, pp.131–148.

Wang, T., Yao, Y., Chen, Y., Zhang, M., Tao, F. and Snoussi, H. (2018) 'Auto-sorting system toward smart factory based on deep learning for image segmentation', *IEEE Sensors Journal*, Vol. 18, No. 20, pp.8493–8501.

Wang, X.K. (2020) 'College students' innovation and entrepreneurship resources recommendation based on collaborative filtering and recommendation technology', *Journal of Physics Conference Series*, Vol. 1533, No. 1, pp.22013–22019.

Wang, Z., Xia, H., Zhang, J., Annor-Nyarko, M., Zhu, S., Jiang, Y. and Yin, W. (2022) 'A deep transfer learning method for system-level fault diagnosis of nuclear power plants under different power levels', *Annals of Nuclear Energy*, Vol. 166, No. 2, pp.108771–108779.

Xu, X. (2022) 'The structural VPR model for the evaluation of innovation and entrepreneurship education', *Chinese Education & Society*, Vol. 55, Nos. 4–5, pp.307–321.

Yuan, F. and Hu, D. (2021) 'Research on the application of big data technology in college students' innovation and entrepreneurship guidance service', *Journal of Physics: Conference Series*, Vol. 1883, No. 1, pp.12171–12182.

Zhang, F. and Xi, L. (2021) 'An evaluation model for the innovation and entrepreneurship thinking ability of college students based on neural network', *International Journal of Emerging Technologies in Learning (iJET)*, Vol. 16, No. 2, pp.188–204.

Zhou, Y. and Zhou, H. (2022) 'Research on the quality evaluation of innovation and entrepreneurship education of college students based on extenics', *Procedia Computer Science*, Vol. 199, pp.605–612.