



Application of AI-based Visual Analysis Technology in Vocational College Electronic Technology Teaching Evaluation

Sun Lifang*, Liu Yuan

Xinxiang Vocational and Technical College

Xinxiang, Henan, 453006, China

sunlifang@xxvtc.edu.cn

ABSTRACT

Using a sliding window combined with heterogeneous multi-column convolutional neural networks, we found that this new evaluation tool has higher accuracy than traditional evaluation tools and can provide more comprehensive evaluation information. This new evaluation tool can effectively improve teaching efficiency and provide more comprehensive evaluations. Advanced computer vision analysis technology can quickly collect and process valuable classroom information and convey it to teachers and peers in real-time, effectively motivating their participation and promoting their learning, ultimately achieving the best educational outcomes.

Keywords: Classroom Teaching Evaluation, Artificial Intelligence, Computer Vision Analysis Technology, And Formative Assessment

Received: 29 August 2024, Revised 5 December 2024, Accepted 22 December 2024

Copyright: with Authors

1. Introduction

In our country, the classroom plays a crucial role as a dominant activity. It not only helps us measure the effectiveness of our teaching but also provides valuable suggestions. We should actively participate in this activity and strive to promote the development of our educational philosophy and improve our teaching level [1]. By conducting classroom teaching evaluations in universities, we can better grasp teachers' achievements in teaching and their achievements at each stage. This will better guide teachers to achieve better results in teaching and push them to achieve greater achievements in education. It can be considered an essential reference for teachers to obtain better results [2]. We comprehensively assess teachers' performance, course arrangements, assignment quality, learning atmosphere, and student performance [3]. These assessments may come from stakeholders, such as administrators, teachers, students, or other relevant personnel. With the development of technology, classroom teaching evaluation has undergone significant changes [4]. Modern classroom teaching evaluation is

no longer limited to individual educators but involves students, educators, and researchers who jointly represent the effectiveness of classroom teaching in various forms. This diversified teaching model helps better reflect the learners' situation but also helps to measure the learners' learning outcomes [5] more accurately. Multiple forms of assessment are implemented, such as conducting check-ins before class and measuring the effectiveness of a class based on student participation, teacher's teaching abilities, and learning outcomes. At the same time, comprehensive assessments should also be carried out from the perspective of learner group participation, engagement, cognitive level, and other aspects [6].

2. Related Work

Xu, Xia, and other experts have conducted in-depth research on classroom teaching evaluation over the past 70 years, exploring current issues and delving into influencing factors to improve the teaching environment and promote a more scientific and reasonable curriculum setting for better outcomes. In addition, they have developed a comprehensive and targeted classroom teaching evaluation system that covers all courses, aiming to promote the current educational development [7]. Through the AHP (Analytic Hierarchy Process) technique, we have creatively improved the existing teacher evaluation method for online courses. Our goal is to address the shortcomings of current online courses and lay a solid foundation for better implementation of online education. Artificial intelligence technology is an emerging science that explores how computers can mimic human intelligence, and computer vision is one of its important research areas [8]. It aims to enable computers with visual abilities, like humans, to accurately recognise objects in the "seen" images and distinguish their different types. With the rapid advancement of technology, applying artificial intelligence and computer vision technology to classroom teaching evaluation has become an indispensable part of the present era [9]. Due to the advent of the information age, research on using computer vision information technology to improve classroom teaching quality is rapidly evolving [10]. Artificial intelligence technology studies how computers can simulate human intelligence, and computer vision is a research direction within artificial intelligence [11,12]. Computer vision technology focuses on enabling computers to simulate human vision, allowing computers to "see" objects in the real world and recognize their categories like humans. With the rapid development of new-generation information technology, applying artificial intelligence and computer vision technology to classroom teaching evaluation conforms to the trends and requirements of the times [13]. However, in practical applications, we still face many challenges, such as effectively using this method to obtain valid video data and construct an information-based and scientific classroom teaching evaluation system. Research on using computer vision technology for classroom teaching evaluation is still in its early stages [15,16]. Further research and exploration are needed on how to reasonably utilize the results of video data analysis using computer vision technology and how to establish a computer-based scientific classroom teaching evaluation model.

3. AI Vision Analysis Technology Based on Face Detection Algorithm

Computer vision technology is a discipline that aims to enable computers to "see" and attempt to understand the content they perceive, much like humans. Currently, mainstream tasks in computer vision include image classification, object recognition in images, object segmentation in images, predicting object positions in image sequences, and more [17, 18]. Image classification involves using computer vision algorithms to identify the main subject category of a given input image. Object detection consists of locating and categorizing objects in an image, using bounding boxes to enclose the detected objects and providing corresponding class labels.

Object recognition refers to the technology that recognizes specific objects in an image. Segmentation can be further divided into semantic segmentation and instance segmentation. Semantic segmentation involves segmenting objects with the same category at the pixel level, while instance segmentation separates each object at the pixel level. Object tracking predicts the position of a given target in a sequence of images, given its initial location in the image sequence. Within these tasks, computer vision technologies related to classroom teaching activities include face detection, face recognition, human pose estimation, facial expression recognition, and more.

3.1 Face Detection Algorithm

Through the analysis of input images, face detection algorithms can accurately locate the subjects' facial features and use edge detection techniques to precisely orient them, achieving fast and accurate detection of faces in specific scenarios. Advanced face detection algorithms allow us to accurately recognize facial information in images, including but not limited to feature extraction, model matching, and statistical analysis. Additionally, these technologies help us accurately locate learners in the classroom and record their attendance and departure times. Sum and Product are two methods available in the "operation" for selection. Sum refers to the selection of predicted probabilities through a sliding window. Sorting these probabilities in the order of $M = head+1, y' lot-c$ to obtain the final prediction result ($M = head+range-1, y' lot-c$ represents the sorted predicted probabilities), as shown in Formula (1).

$$y_{MCDNN}^i = \sum_{j=head}^M y^i - CNN_j \quad (1)$$

3.2 Face Recognition Technology

Face recognition is a high-precision technology that can extract individual identity information from images or videos. By pre-collecting and preprocessing these features, they can be compared with the images captured by the camera to determine the individual's true identity. Based on convolutional neural networks, this paper explores three different levels of applications in computer vision: object recognition, scene labeling, and scene recognition, as shown in Figure 1.

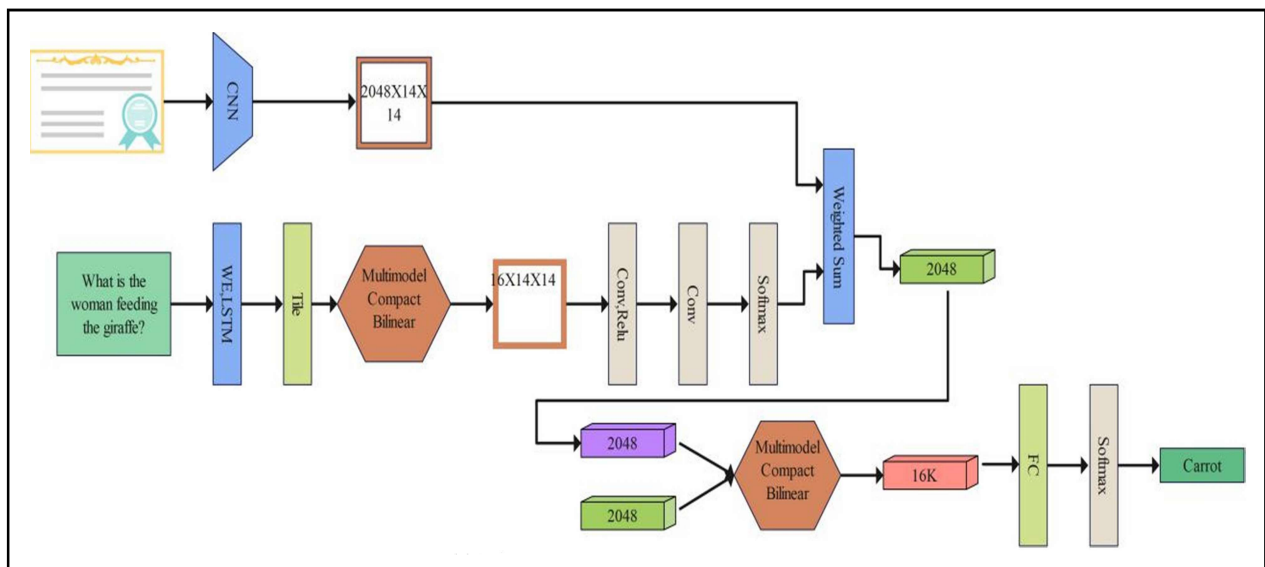


Figure 1. Workflow of Computer Vision Technology

Object recognition technology in computer vision has been proven to be very mature. It allows us to quickly and accurately obtain the required information from various datasets. It helps us perform better complex image processing tasks, such as image localization and qualitative and quantitative analysis. Due to its ability to extract independent information from many samples, object recognition technology is less affected by external environments, enabling more accurate object localization and tracking. This reduces human-generated noise and errors, providing more possibilities for feature and classification algorithm optimization and improvement. Product refers to the summation of possibilities obtained by using the sliding window and accumulating them, as shown in formula (2):

$$y_{MCDNN}^i = \prod_{j = \text{head}}^M y_{\text{sort-CNN}}^i \quad (2)$$

By applying the method proposed in this paper, we can obtain a series of confidences to enable the computer to accurately recognize objects in the image and determine their corresponding categories based on their features. With carefully designed algorithms, it is possible to effectively predict the object's category and compare it with benchmark labels in the dataset, thereby enhancing object recognition accuracy. Therefore, accuracy is an important metric for evaluating algorithm performance.

3.3 Introduction to Three Different Pose Estimation Techniques

Through human pose estimation, we can extract important features of the human body from an image and determine their positions. This allows us to recognize and accurately classify human actions, such as walking, standing, sitting, bowing, or raising hands. Through facial expression recognition, we can extract and study the facial expressions of individuals, categorizing them into different emotions and psychological states. For example, happiness, anger, sadness, fear, and disgust can all be accurately captured and described through facial expressions. In contrast to regular facial expressions, subtle facial expressions can better reflect an individual's feelings. By observing microexpressions, we can gain a more accurate understanding of an individual's inner emotions. Scene labeling is a technique that delves into the semantics of images. It not only helps us better understand objects in the image but also enables more accurate identification and labeling of all objects in the scene, thereby enhancing the functionality of the image. By considering various factors in the scene image, such as objects and environment, scene recognition can effectively determine the category of the scene, achieving a deeper semantic understanding of the image. Under the condition where head and range are both set to 1, $y_{CDNN} = y_{\text{sort-CNN}}$ in this model, and the model is based on arranging the values from the minimum to the n-ca., thus achieving the fusion of the sliding window, known as *minrule*, as shown in formula (3):

$$y_{MCDNN}^i = \min_{j \in [1, N]} y_{CN_j}^i \quad (3)$$

Changing the parameters of sliding window fusion can make it a more effective classifier fusion method, significantly improving the accuracy and efficiency of the model (*minrule*, *maxrule*, *averagerule*, and product rule, etc.).

4. Experimental Results and Analysis

4.1 Experimental Configuration

To test the reliability and generalization ability of our proposed method, we selected two widely used scene recognition datasets, MIT and SUN397. Compared with some common scene recognition algorithms, we found

that these algorithms have higher accuracy and generalization ability. 1) After careful selection, we finally chose three significant scales, highlighting the features of salient regions and effectively preserving the semantic information of scenes, thereby significantly improving scene recognition accuracy. 2) Through the recent Hybrid-CNN and Hybrid-CNN, we can better evaluate their accuracy and efficiency in dealing with complex problems. HIBCNN has five levels and three fully connected layers, while AlexNet has more complex levels. Their combination allows for better recognition of complex problems and better solutions. 3) Convolutional neural networks can extract useful information from multiple datasets, which may come from the model itself or external data sources. They can use bilinear interpolation to accurately transform the information of the model itself into a size of 227×227 , which meets the output requirements of HBCNN. 4) we can map image features to specific scene categories through multi-layer perceptrons. In this experiment, we processed the input and output layers and added a hidden layer.

4.2 Experimental Analysis

By dividing the 12-column convolutional neural network into four groups, we can explore the level of object recognition under different convolutional kernel sizes, thus improving the model's performance. These groups include CNN1, CNN2, CNN3, CNN4, CNN5, CNN6, CNN7, CNN8, CNN9, CNN10, CNN11, and CNN12. As shown in Figure 2, we achieved satisfactory results through experiments on the four groups of multi-column convolutional neural networks on the MNIST dataset.

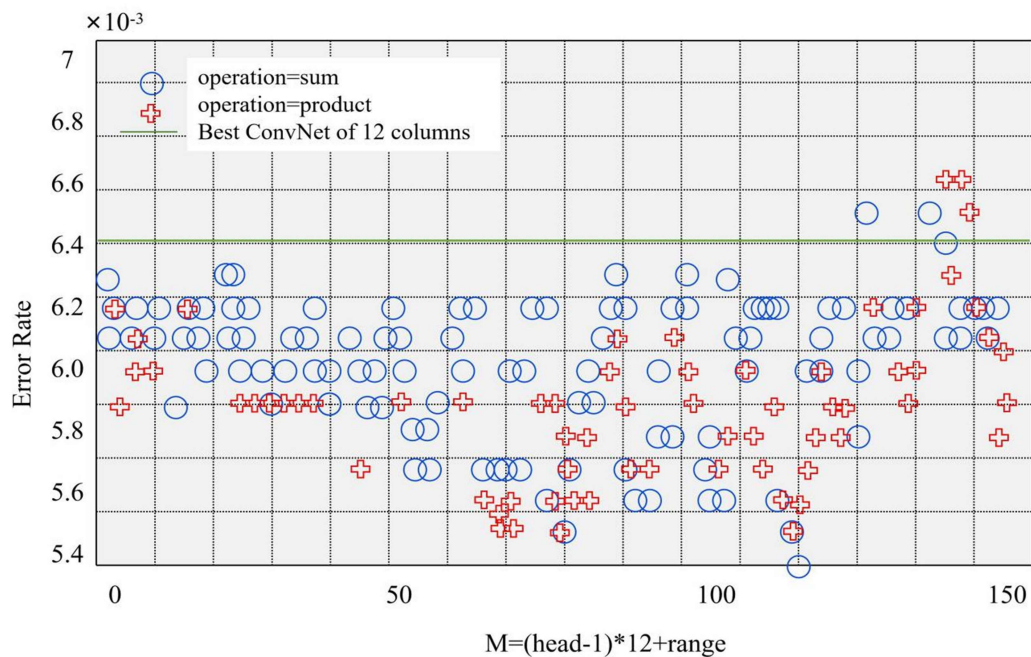


Figure 2. Results of the 12-column Convolutional Neural Network

In this experiment, we used a complex system consisting of 12 columns of convolutional neural networks with different structures, each forming an independent column. We applied the sliding window fusion technique to examine the MNIST dataset, and the specific results can be seen in Figure 2. We found that by setting $head=ceil(M/12)$ and $range=Mod((M-1), 12)+1$, where $m \in [1, 144]$, the performance of the 12-column convolutional neural network surpassed other models, with CNN10 achieving the best performance, reaching an accuracy of 0.65%. Additionally, when compared to other models, the accuracy of the 12-column convolutional neural

network was improved, gaining 0.54%. Through this experiment, we discovered that using the combination of sliding window fusion and heterogeneous multi-column convolutional neural networks significantly reduced the error rate to 16.92%, far superior to using only single-column convolution, and showed significant improvement in object detection, with advantages over previous research.

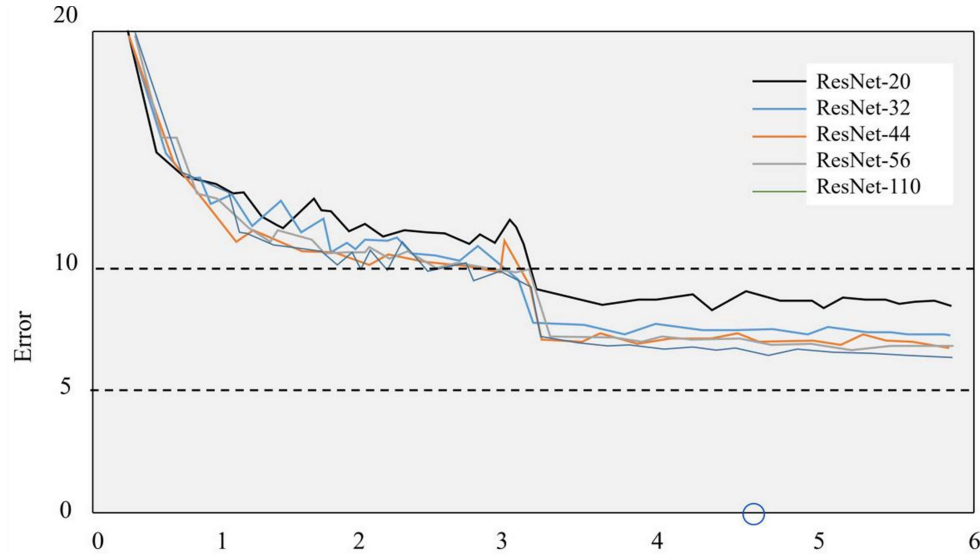


Figure 3. Error Rate Evaluation Results in MNIST

From Figure 3, it can be observed that both fusion techniques yield satisfactory results. However, the exhaustive sliding window fusion technique significantly enhances the overall performance, displaying a clear advantage. By employing the exhaustive sliding window fusion technique, the model's performance can be effectively improved, and its superiority surpasses that of traditional single fusion techniques. The experiments demonstrate the high efficiency of this technique, with five out of eight test groups showing outstanding performance, making it an exemplary fusion technique. Although training-based sliding window fusion techniques can enhance the model's performance without compromising its cost-effectiveness, this novel model exhibits higher network fusion capability compared to the exhaustive technique. Moreover, this new model boasts high reliability and can cater to various application scenarios with simplicity and lower costs.

5. Conclusion

Through in-depth research, we have found that leveraging computer vision analysis technology can effectively improve the efficiency of classroom teaching and provide abundant information, thereby expanding the scope of evaluation and enhancing work efficiency. We conducted a comprehensive study on three widely used object recognition datasets, MNIST, CIFAR-10, and Caltech-256, by employing the sliding window fusion technique in heterogeneous multi-column convolutional neural networks. Compared to traditional approaches, this network demonstrates a significant performance improvement and can achieve high accuracy. Advanced computer vision technology enables the rapid collection, organization, and analysis of various relevant course materials, which can be implemented to enhance course quality, stimulate teacher-student engagement, facilitate effective interactions, and ultimately achieve educational objectives. Additionally, leveraging artificial intelligence, big data, virtual reality, the Internet of Things, and large-scale integration can lead to more effective course management.

References

- [1] Yao, D. (2021). Research on the Application Analysis of Four-dimensional Teaching Method in Higher Vocational English Teaching Based on Big Data Analysis. *Journal of Physics: Conference Series*, 1744(3), 032059.
- [2] Yong, Q. L. (2021). Application Analysis of Artificial Intelligence in Library Network Security. *Journal of Physics: Conference Series*, 1744(3), 032024 (7pp).
- [3] Wei, H. (2021). Analysis on the Application of VR Technology in Practical Teaching of Law Education. *Journal of Physics: Conference Series*, 1744(3), 032089 (4pp).
- [4] Abbitt, J. T. (2011). Measuring technological pedagogical content knowledge in preservice teacher education: A review of current methods and instruments. *Journal of Research on Technology in Education*, 43(4), 281-300.
- [5] Xu, D. (2021). The Application and Analysis of Task-based Language Teaching in English Listening and Speaking Course for English Majors: A Case Study of Sichuan University of Culture and Arts. *Frontier of Higher Education*, 3(4), 10-13.
- [6] Alexander, W. R. J., Haug, A. A., Jaforullah, M. (2010). A two-stage double-bootstrap data envelopment analysis of efficiency differences of New Zealand secondary schools. *Journal of Productivity Analysis*, 34, 99-110.
- [7] Gao, L., Jiao, T., Feng, Q., et al. (2021). Application of artificial intelligence in diagnosis of osteoporosis using medical images: a systematic review and meta-analysis. *Osteoporosis International*, 32(7), 1279-1286.
- [8] Huang, J., Shen, G., Ren, X. (2021). Connotation Analysis and Paradigm Shift of Teaching Design under Artificial Intelligence Technology. *International Journal of Emerging Technologies in Learning (iJET)*, 16(5), 73.
- [9] Liu, H., Tan, W., Li, H., et al. (2021). Application of Artificial Intelligence Technology in the Teaching of Mechanical Education Courses in Universities. *Journal of Physics: Conference Series*, 1992(4), 042065-.
- [10] Yuan, Y. (2022). Quantitative analysis of Chinese classroom teaching activity under artificial intelligence background. *Education and Information Technologies*, 27(8), 11161-11177.
- [11] Simon, H. A. (1981). Studying Human Intelligence by Creating Artificial Intelligence: When considered as a physical symbol system, the human brain can be fruitfully studied by computer simulation of its processes. *American Scientist*, 69(3), 300-309.
- [12] Liang, G., Fan, W., Luo, H., et al. (2020). The emerging roles of artificial intelligence in cancer drug development and precision therapy. *Biomedicine & Pharmacotherapy*, 128, 110255.

- [13] Aggarwal, K., Mijwil, M. M., Al-Mistarehi, A. H., et al. (2022). Has the future started? The current growth of artificial intelligence, machine learning, and deep learning. *Iraqi Journal for Computer Science and Mathematics*, 3(1), 115-123.
- [14] Agatonovic-Kustrin, S., Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modelling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis*, 22(5), 717-727.
- [15] Luxton, D. D. (2014). Artificial intelligence in psychological practice: Current and future applications and implications. *Professional Psychology: Research and Practice*, 45(5), 332.
- [16] Hwang, G. J., Xie, H., Wah, B. W., et al. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, 1, 100001.
- [17] Komninos, N. (2006). The architecture of intelligent cities: Integrating human, collective and artificial intelligence to enhance knowledge and innovation. *2006 2nd IET International Conference on Intelligent Environments-IE 06*. IET, 1, 13-20.
- [18] Nishant, R., Kennedy, M., Corbett, J. (2020). Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. *International Journal of Information Management*, 53, 102104.