Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2025 SCPE. Volume 26, Issues 3, pp. 1386–1394, DOI 10.12694/scpe.v26i3.3901

CONSTRUCTION OF INFORMATION MANAGEMENT MODEL FOR COLLEGE STUDENTS BASED ON DEEP LEARNING ALGORITHMS AND DATA COLLECTION

LIN ZHU*

Abstract. In order to improve the efficiency of college student management and provide effective tools and assistants for college student management staff, the author proposes the construction of an information technology model for college student management based on deep learning algorithms and data collection. Firstly, use the FasterR-CNN model to detect the heads of personnel in the laboratory, Then, based on the output results of model detection, use the IoU algorithm to filter out duplicate detected targets, Finally, a coordinate based positioning method is used to determine whether there are people on each workbench in the laboratory, and the corresponding data is stored in the database. The main functions of this system include: (1) Real-time video monitoring and remote management of the laboratory, (2) Timed automatic photography detection and data collection provide data support for quantitative management in the laboratory, (3) Query and visualization of data on changes in laboratory personnel. The experimental findings demonstrate that our proposed model excels with an F1 Score exceeding 91%, showcasing robust generalization across detection confidence levels ranging from 50% to 99%. Notably, at a detection confidence of 96%, our model achieves its peak performance with an impressive F1 Score of 95.7%. This underscores the model's exceptional detection capabilities. Leveraging Faster R-CNN and IoU optimization, our laboratory personnel statistics and management system offer real-time personnel tracking and remote management functionalities tailored for office environments.

Key words: Convolutional neural network, Object detection, College student management, Personnel statistics, Management informatization

1. Introduction. The swift advancement of information technology has posed unparalleled challenges to university student management, yet it has also presented vast opportunities for enhancing student management practices. Especially regarding the precision and pertinence of student management work, it is currently the biggest challenge faced by traditional student management personnel. The development of information technology has brought good news to solve this problem [1]. It is precisely with such questions that analysis and research can bring innovative results to the model of student management in universities. Finally, with the strong utilization of information technology advantages, we will promote the healthy and stable development of student management in universities [2]. It must be acknowledged that the advent of the information age has brought tremendous convenience to people's lives and work. To some extent, the integration of information technology has undeniably enhanced the efficiency of our daily lives and professional endeavors. For university staff involved in student management, leveraging information technology has provided firsthand experience of its convenience and rapidity. Traditionally, communication and interaction between teachers and students primarily occurred through face-to-face interactions in the realm of student management. With the application of information technology, real-time video communication has become a reality [3]. Whether it's issues related to student learning or accommodation, many problems can be solved as quickly as possible through information-based communication methods. Especially in case of unexpected situations, the application of information technology highlights its advantages in terms of timeliness [4].

As the representative of information technology, the popularity of the Internet and mobile Internet has also had an increasingly profound impact on contemporary college students. Especially in the development of thinking, the Internet and mobile Internet have brought many new ideas and mysteries to college students, making those engaged in college student management clearly feel how backward the efficiency of traditional student management is [5]. Therefore, personnel engaged in the management of college students can learn and improve their knowledge and skills in student management through the help of the Internet and mobile Internet, a huge resource pool. Therefore, the application of information technology will greatly improve the efficiency of

*Student Affairs Section, Zibo Vocational Institute, Zibo, Shandong, 255300, China (lindazhuzhu1981@163.com)

university student management and become a powerful tool and assistant for university student management staff [6].

2. Literature Review. There is relatively little specialized research on big data in student management in universities, and more research is focused on the overall management of universities, mainly in the following two aspects. On the one hand, starting from the strategic thinking of empowering university management with big data, we aim to enhance the understanding of big data in university management. Mok, K. et al. have introduced a novel approach to signal management tailored for the comprehensive processing demands of smart cities. This method integrates deep learning and simulation techniques to address the challenge of handling large-scale data volumes effectively. Initially, the system undergoes offline training using deep learning networks based on computer vision, enabling the detection of various vehicle types. Training data, amassed from diverse city or country settings, facilitates a one-time training process for the deep networks. Subsequently, for each intersection requiring traffic flow prediction, a minimal dataset is gathered to construct a computer simulation model for localized traffic flow estimation. Finally, an adaptive traffic light management algorithm emerges through the fusion of deep learning-driven traffic monitoring systems and optimized simulation outcomes. This versatile approach offers seamless applicability across different intersections, requiring only minimal traffic data collection for each new intersection [7]. Yang, H. et al. have explored the integration of data mining technology within college student education management information systems, offering insights and strategies for enhanced student information management. Their work delves into research methodologies for employing data mining techniques within these systems, particularly focusing on educational data. They detail the implementation of the K-means and fuzzy C-means clustering algorithms, culminating in the development of a robust data mining system. Performance evaluation reveals an average algorithm processing time of 1.92 seconds within the system, ensuring a seamless user experience for education administrators [8]. Fan, J. et al. have primarily focused on leveraging data mining technology to advance university information management systems (IMS). Their investigation delves into the process of data mining (DM) and its application, particularly in association rule mining, within the realm of university IMS development. Their findings reveal a comprehensive course coverage, with specific parameters set for transaction support count and confidence level. Notably, the research highlights a tendency among students to select multiple courses in conjunction with one another. The integration of DM theory within university informatization initiatives is poised to significantly enhance the data analysis capabilities of management personnel, consequently elevating their proficiency in administration [9].

For the common office scenario of fixed indoor personnel and fixed workstations, the author takes ordinary university laboratories as an example and proposes an indoor personnel statistics method based on Faster R-CNN and Intersection over Union (IoU) optimization. This method utilizes deep neural networks to extract head features from images, resulting in higher detection accuracy; In addition, a coordinate based positioning method has been proposed, which can accurately determine whether there are people on each workbench in the laboratory. Finally, a laboratory personnel statistics and management system was designed and developed using the trained detection model, which effectively achieved remote, automated, and intelligent management of the laboratory. The experimental results indicate that the system can be applied to common indoor office scenarios.

3. Research Methods. Fast R-CNN is a rapid target detection method based on regional convolutional network (Region-based Convolutional Network). As an improvement of the R-CNN model, Fast R-CNN improves detection speed, but like R-CNN, it uses Selective Search (SS) method to extract candidate target regions (Proposals) from images. Therefore, there are still problems such as cumbersome detection steps, high time and memory consumption. FasterR-CNN introduces a Region Proposal Network (RPN) in the model to extract candidate target regions, achieving feature sharing in convolutional layers and greatly improving the generation speed of candidate target regions [10]. The Faster R-CNN network structure mainly consists of RPN and FastR-CNN detectors, where the input of RPN is image features extracted through a series of convolutions.

3.1. Feature extraction network. To illustrate, conventional deep neural networks like AlexNet, VGGNet, and GoogLeNet have the capacity to enhance image feature extraction by augmenting the number of network layers. However, when deep networks reach a point of convergence, elevating the number of layers may lead to a phenomenon termed "degradation," wherein the network's detection accuracy plateaus or diminishes.

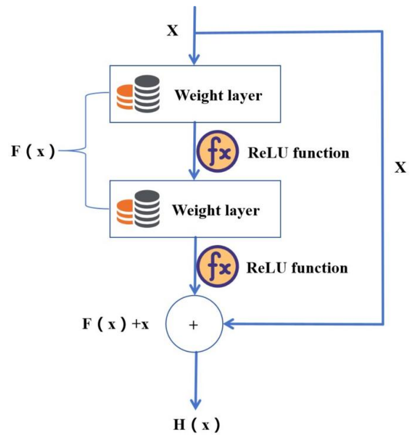


Fig. 3.1: Unit structure of residual network

Residual Neural Network (ResNet) can effectively solve the phenomenon of network degradation and has better image feature learning ability. Therefore, the author chose the residual network as the feature extraction network for Faster R-CNN [11].

The unit structure of the residual network is shown in Figure 3.1. Assuming that the original mapping output of the network units is $H(x)$, that is $H(x) = F(x) + x$. then $F(x) = H(x) - x$. Therefore, each convolutional output of the deep network will become a fitting residual. It can be simply understood that residual networks add some ”cross layer connections” in traditional deep convolutional networks (x in Figure 3.1), when the training error increases with the depth of the network, the residual network will skip certain convolutional layers and directly input the original data into the subsequent convolutional layers, which not only ensures the integrity of data transmission but also relatively reduces the training error and reduces the difficulty of deep network training [12].

3.2. Regional recommendation network. Traditional candidate target region extraction methods suffer from time-consuming issues, such as the sliding window and image pyramid used in Adaboost, and the SS used in R-CNN and Fast R-CNN. The RPN used by Faster R-CNN embeds the extraction of candidate target regions into the network and improves the generation speed of candidate target regions by sharing convolutional layer feature parameters. The author combines the actual pixel size of the target area and in order to obtain multi-scale detection boxes, RPN uses a 3x3 convolutional kernel to slide on the feature map output by the feature extraction network, and maps the region corresponding to the center of the convolutional kernel back to the original input image, generating a total of 12 anchors with 4 scales $\{16^{0.5}, 16, 16^{1.5}, 16^2\}$ and 3 aspect ratios {05,1,2}. Therefore, there are 12 suggested regions corresponding to the center of the convolution kernel in each sliding window. RPN is a fully convolutional network that inputs the original image convolutional feature map output by the feature extraction network. The suggested region corresponding to each anchor point is convolved through an intermediate layer to output a 512 dimensional feature vector, which is then fed into the classification layer and position regression layer, respectively. Among them, the classification layer outputs the classification information of the target in the corresponding anchor point, including the confidence level of the background and the confidence level of the target category; The position regression layer outputs the position information of the target in the anchor point, including the center point coordinates, length, and height of the target area. Finally, using the Non Maximum Suppression (NMS) algorithm, the candidate target regions are filtered based on the classification and position information of all anchor points, resulting in 2000 high-quality target candidate regions [13,14].

3.3. Fast R-CNN detection network. Once the Region Proposal Network (RPN) generates candidate regions of interest, they undergo further refinement through the Fast R-CNN detector for accurate classification and coordinate regression. To address the diverse sizes of these candidate regions, they are directed to the Region

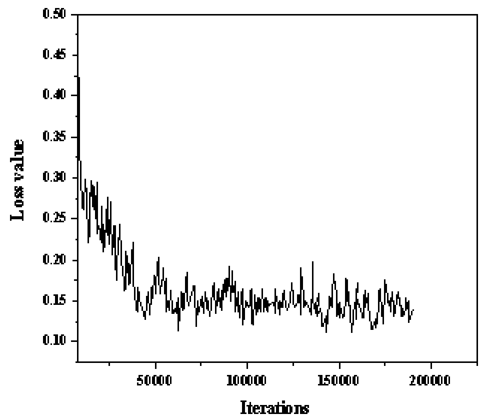


Fig. 4.1: Loss values during training process

of Interest (RoI) pooling layer, where they are resized uniformly for streamlined processing. The RoI pooling layer combines feature maps and target candidate regions for coordinate mapping, outputting a fixed size target candidate region. Subsequently, these target candidate regions are sent to the Fast R-CNN detector for training, obtaining the final detection results including classification information and coordinate information [15].

4. Result analysis.

4.1. Data generation and training. The author's experimental data was collected using a monocular camera located at the top of the laboratory. The image captured by the top camera shows a significant amount of occlusion in various parts of the human body. Therefore, the human head is selected as the detection target to determine the number and distribution of personnel in the experiment. A total of about 6000 original images were collected, and after flipping and symmetry, the dataset was expanded to about 24000 images. The image size is uniformly 1 510 x 860 pixels, and the number of people in each image ranges from 1 to 10. Randomly divide the dataset into training and testing sets in a ratio of 10:1 [16]. The author's experimental environment is Windows 10, GeForce GTX 1080Ti, and the network model is implemented using the mainstream deep learning framework TensorFlow. In the overall model architecture, ResNet101 serves as the feature extraction backbone. The training process employs a batch size of 4, initializing the learning rate at 0.0003. After 40,000 iterations, the learning rate decreases to 0.00003, followed by a further reduction to 0.000003 after 80,000 iterations, concluding with a total of 200,000 iterations [17].

4.2. Result Analysis. Following the training phase, the Faster R-CNN model attained an impressive mean average precision of 98.49% when evaluated on the test dataset. The training progression, as depicted in Figure 4.1, illustrates the loss curve over the course of training.

Figure 4.1 illustrates that the model's loss value stabilized around 0.15 after 180,000 iterations, indicating convergence. The exceptional detection performance of the model can be attributed to:

1. The scene background in the laboratory is single, with low personnel mobility, less personnel and background changes, and more prominent image features;
2. There are many training data samples, and the training set contains more than 20000 images, totaling about 70000 labeled human head samples;
3. For targets of different scales, a total of 12 anchor points with 4 scales and 3 aspect ratios are used, which can effectively detect targets of different scales;
4. The model utilizes RPN to generate high-quality target candidate regions, providing high-quality training data for subsequent Fast R-CNN networks.

Table 4.1: Four scenarios for model detection

Situation	Detected as the target	Detected that it is not the target
Actually, it's the goal	TP (really)	FN (False No)
Actually, it's not the target	FP (False)	TN (True or False)

In order to further study the generalization ability of the model, that is, its detection performance in actual scenes, 105 images were collected from the images captured by the camera as an incremental test set to test the detection performance of the model under different confidence levels. The most commonly used evaluation indicators for detection models are accuracy and recall. There are four situations when calling the model for detection: 1) Actually, it is the target, and detection considers it as the target; 2) In fact, it is the target, and detection considers it not the target; 3) In fact, it is not a target, and detection considers it as a target; 4) Actually, it is not the target, and the detection assumes it is not the target [18].

The four possible scenarios for the model to detect targets are shown in Table 4.1.

Therefore, the definitions of precision and recall can be given as follows: 4.1,4.2:

$$P = TP/(TP + FP) \tag{4.1}$$

$$R = TP/(TP + FN) \tag{4.2}$$

Among them, P is the model accuracy, and R is the model recall. Accuracy refers to how much of the detection results provided by the model are correct, while recall refers to how many actually correct targets have been detected. These two indicators usually have a trade-off. In order to comprehensively consider these two indicators, a new evaluation indicator is introduced, which is the weighted harmonic mean F-Score of accuracy and recall. The following equation 4.3:

$$F - Score = (1 + \beta^2) * \frac{P * R}{\beta^2 * P + R} \tag{4.3}$$

Among them, β for harmonic parameters, when $\beta > 1$, accuracy is considered more important, the author believes that recall and accuracy are equally important, that is, taking β=1. Therefore, the author's weighted harmonic mean is F1 Score, as follows:

$$F1 - Score = \frac{2 * P * R}{P + R} \tag{4.4}$$

The incremental test set consists of 105 images, with a total of 445 actual targets. Test the model using an incremental test set and calculate the F1 Score at different detection confidence levels, as shown in Figure 4.2.

From Figure 4.2, it can be seen that the model has a high F1 Score with a detection confidence level of 50% to 99%, and the F1 Score is greater than 91%, indicating that the model has strong generalization ability. At the same time, when the detection confidence is 96%, F1 Score reaches the highest level, reaching 95.7%, indicating that the model has the best detection performance at this time, that is, when the detection confidence is 96%, the model's generalization ability is the best.

During the incremental testing process, the detection time was 28.89 seconds, with an average detection speed of 275.1 ms per image, which is much lower than the detection speed in the middle, indicating that although the model implemented by the author has high detection accuracy, the detection speed is difficult to achieve the goal of real-time detection of video streams. In the field of object detection, there are usually three methods that can improve detection speed:

1. Reduce the size of the input image. This method is suitable for scenes with small monitoring areas. The author's input image size is 1 510 x 860 pixels, which only covers the experimental monitoring area. Therefore, this method is not applicable.

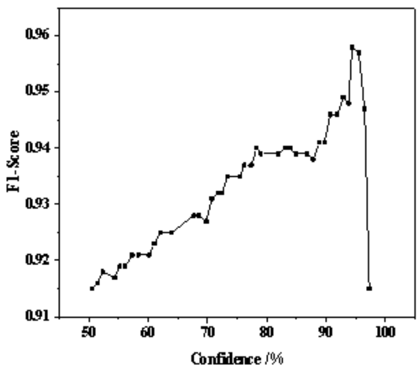


Fig. 4.2: Under different confidence levels

2. Using hardware devices with more powerful computing performance. This method requires a significant amount of capital and is therefore more suitable for the industrial sector.
3. 3) Utilize a more streamlined detection model based on network structure. This method usually sacrifices detection accuracy to a certain extent, such as SSD, making it more suitable for application areas that do not require high detection accuracy and high detection speed.

In the laboratory, personnel are mostly sitting and working, with low mobility. It is unnecessary to blindly pursue detection speed in this scenario, so the author did not use the three methods mentioned above. Instead, the system was designed to automatically detect every minute. Therefore, the detection speed of the model meets the design requirements of the system. For common office scenarios, a design that checks every minute is sufficient to provide reliable data for scientific management by managers [19].

4.3. IoU optimization. The accuracy and generalization ability of the model implemented by the author are quite outstanding, and the detection results are relatively good. The detection model detects a human head into two. In response to this situation, the author utilized the IoU algorithm for further optimization. IoU refers to the overlap rate of two detection boxes with overlapping areas, which is the ratio of the intersection and union between these two detection boxes. As shown in Figure 4.3, there is an overlapping area between Box A and Box B. Among them, S (A), S (B), and S (C) represent the areas of boxes A, B, and C, respectively.

After using the Faster R-CNN object detection model to detect the target image, all detection boxes containing position information output from the detection are input into the IoU algorithm. Among them, number is the number of targets detected by the model, which is the number of personnel in the laboratory output by the model detection. The final output N is the actual number of people in the laboratory filtered by IoU.

4.4. Personnel positioning. In response to the characteristics of low personnel mobility, single environment, and relatively fixed personnel positions in the laboratory, the author proposes a coordinate positioning method to determine whether there are people on each workbench and store the corresponding data in a database, providing reliable data support for scientific management of the laboratory. Divide the monitoring area into 12 rectangular areas in advance, representing the workstations within each area. Firstly, the center of mass of the human head is determined using the position information of the personnel target detected by the model, and then discrimination is carried out one by one. If the center of mass falls in which area, it is considered that there is a person on the workbench in that area. It can be considered that there are people on the workbench in areas 2, 3, 4, 6, and 8 [20].

4.5. System Implementation and Display. The system is developed using the open-source web development framework Django, and has two main functional modules: the system's historical data query and

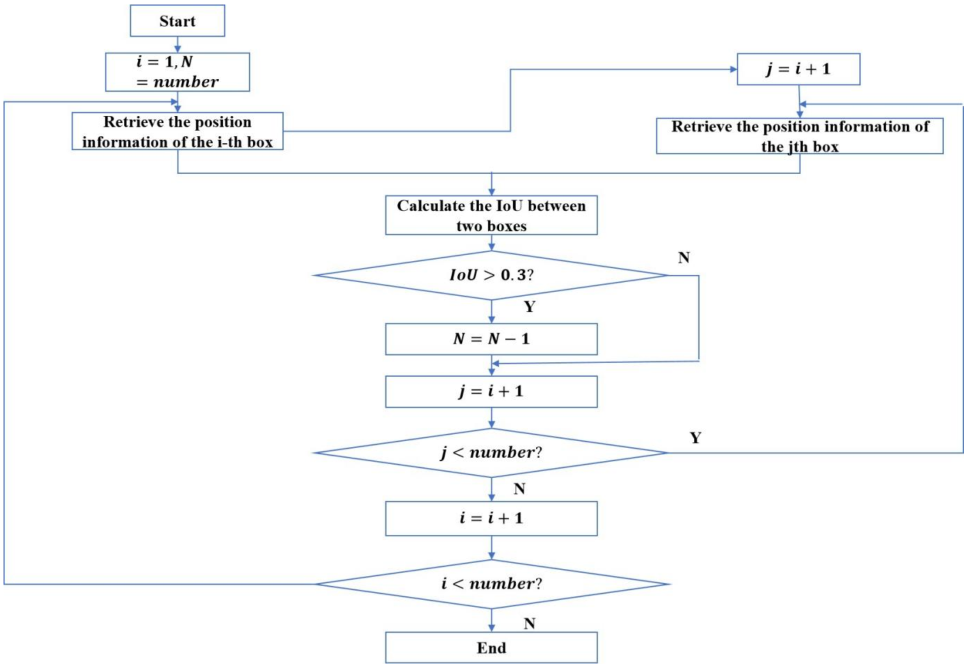
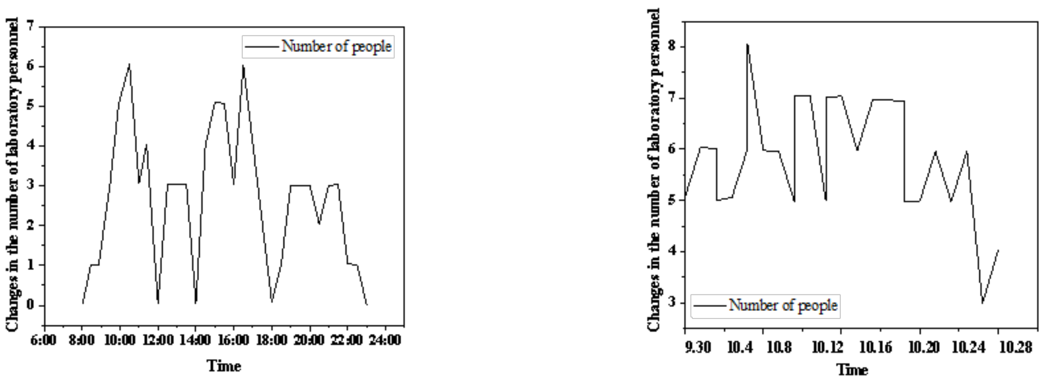


Fig. 4.3: IoU Algorithm Process



(a) Changes in daily population

(b) Changes in the number of people over a period of time

Fig. 4.4: Changes in the number of laboratory personnel

display module, and the real-time video monitoring and detection module. The system is developed based on B/S mode and has remote management function. Authorized users can access the system within the campus network by logging in through a PC browser. From 6:30 to 23:30 every day, the system server automatically captures a laboratory monitoring image every minute and calls the detection model to detect it; Then locate the personnel based on the test results to determine if each workbench is manned; Finally, store the corresponding data in the database for laboratory administrators to query.

Figure 4.4 shows the query page for daily changes in the number of people in the laboratory and changes in the number of people over a period of time. This system has been running stably in the laboratory for six months, and its promotion and application value has been verified [21].

5. Conclusion. In response to the characteristics of fixed personnel and fixed workstations in common office scenarios, taking ordinary university laboratories as an example, the author proposes an indoor personnel statistics method based on Faster R-CNN and IoU optimization. The experimental results show that the proposed method has good detection accuracy. Then, based on the detection results, a coordinate positioning method is used to determine whether each indoor workbench is occupied. Finally, a laboratory personnel statistics and management system was developed using the Django framework, achieving remote, automatic, and intelligent management of the laboratory. However, the system developed by the author has the problem of personnel positioning relying on the premise that personnel positions are relatively fixed. When personnel positions move, the system cannot make accurate judgments. Therefore, further research on target tracking algorithms between video frames will be carried out, while ensuring detection accuracy. By drawing the movement trajectory of personnel, dynamic positioning of personnel will be achieved.

6. Acknowledgement. The research is supported by: This work was supported by the Natural Science Foundation of Province, China (Grants No.2019JJ40097), by the Research Foundation of Hunan University of Science and Engineering, China (Grant No.17XKY068) and by the construct program of applied characteristic discipline in Hunan University of Science and Engineering.

REFERENCES

[1] Qu, J. (2021). Research on mobile learning in a teaching information service system based on a big data driven environment. Education and Information Technologies, 26(5), 6183-6201.
[2] Li, X., Liu, H., Wang, W., Zheng, Y., Lv, H., & Lv, Z. (2022). Big data analysis of the internet of things in the digital twins of smart city based on deep learning. Future Generation Computer Systems, 128, 167-177.
[3] Bi, H., Liu, J., & Kato, N. (2021). Deep learning-based privacy preservation and data analytics for IoT enabled healthcare. IEEE Transactions on Industrial Informatics, 18(7), 4798-4807.
[4] Kor, M., Yitmen, I., & Alizadehsalehi, S. (2023). An investigation for integration of deep learning and digital twins towards Construction 4.0. Smart and Sustainable Built Environment, 12(3), 461-487.
[5] Jasti, V. D. P., Zamani, A. S., Arumugam, K., Naved, M., Pallathadka, H., Sammy, F., ... & Kaliyaperumal, K. (2022). Computational technique based on machine learning and image processing for medical image analysis of breast cancer diagnosis. Security and communication networks, 2022, 1-7.
[6] Hu, J. (2021). Teaching evaluation system by use of machine learning and artificial intelligence methods. International Journal of Emerging Technologies in Learning (iJET), 16(5), 87-101.
[7] Mok, K. , & Zhang, L. . (2024). Adaptive traffic signal management method combining deep learning and simulation. Multimedia tools and applications(5), 83, 639-653.
[8] Yang, H. , & Zhang, W. . (2022). Data mining in college student education management information system. Int. J. Embed. Syst., 15, 279-287.
[9] Fan, J. , Zhang, M. , Sharma, A. , & Kukkar, A. . (2022). Data mining applications in university information management system development. Journal of Intelligent Systems, 31(1), 207-220.
[10] Dhillon, A., & Verma, G. K. (2020). Convolutional neural network: a review of models, methodologies and applications to object detection. Progress in Artificial Intelligence, 9(2), 85-112.
[11] Lindsay, G. W. (2021). Convolutional neural networks as a model of the visual system: Past, present, and future. Journal of cognitive neuroscience, 33(10), 2017-2031.
[12] Tian, Y. (2020). Artificial intelligence image recognition method based on convolutional neural network algorithm. Ieee Access, 8, 125731-125744.
[13] Zhou, D. X. (2020). Theory of deep convolutional neural networks: Downsampling. Neural Networks, 124, 319-327.
[14] Raghu, M., Unterthiner, T., Kornblith, S., Zhang, C., & Dosovitskiy, A. (2021). Do vision transformers see like convolutional neural networks?. Advances in neural information processing systems, 34, 12116-12128.
[15] Ilesanmi, A. E., & Ilesanmi, T. O. (2021). Methods for image denoising using convolutional neural network: a review. Complex & Intelligent Systems, 7(5), 2179-2198.
[16] Wu, X., Sahoo, D., & Hoi, S. C. (2020). Recent advances in deep learning for object detection. Neurocomputing, 396, 39-64.
[17] Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J. (2023). Object detection in 20 years: A survey. Proceedings of the IEEE, 111(3), 257-276.
[18] Oksuz, K., Cam, B. C., Kalkan, S., & Akbas, E. (2020). Imbalance problems in object detection: A review. IEEE transactions on pattern analysis and machine intelligence, 43(10), 3388-3415.
[19] Dhillon, A., & Verma, G. K. (2020). Convolutional neural network: a review of models, methodologies and applications to object detection. Progress in Artificial Intelligence, 9(2), 85-112.
[20] Diwan, T., Anirudh, G., & Tembhurne, J. V. (2023). Object detection using YOLO: Challenges, architectural successors, datasets and applications. multimedia Tools and Applications, 82(6), 9243-9275.

[21] Pal, S. K., Pramanik, A., Maiti, J., & Mitra, P. (2021). Deep learning in multi-object detection and tracking: state of the art. Applied Intelligence, 51, 6400-6429.

Edited by: Hailong Li
Special issue on: Deep Learning in Healthcare
Received: May 11, 2024
Accepted: Jun 9, 2024