Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org

Volume 26, Issues 3, pp. 1293–1299, DOI 10.12694/scpe.v26i3.4209

ELECTRONIC INFORMATION IMAGE PROCESSING BASED ON CONVOLUTIONAL NEURAL NETWORKS

HONGMING PAN*

Abstract. In order to improve the accuracy and efficiency of mechanical part recognition, the author proposes a research on electronic information image processing based on convolutional neural networks. The author first performs saturation based grayscale processing on the image; Then, the binary image is obtained through significance enhancement, binarization using the Maximum Between Class Variance (OTSU) method, and morphological closure operation; Extract the part area using an improved seed filling method; Finally, the parts are identified by combining Scale Invariant Feature Transform (SIFT) features of image keypoints with Convolutional Neural Network (CNN) models. The experimental results show that the accuracy of the part recognition algorithm can reach 98.84%, and the recognition speed is about 5fps. Through experimental comparison and analysis, it has been proven that this method is fast and effective, with high accuracy and good robustness.

Key words: Part recognition, Image saturation, Seed filling method, Scale invariant feature transformation, Convolutional neural network

1. Introduction. With the popularization and development of science and technology, mobile phones, tablets, and computers have become necessities for people, whether it is work or life [1]. The emergence, development, and enrichment of technology have brought a new world to human civilization, in which electronic information provides convenient conditions for people to communicate in time and space, especially image processing technology, which is very widespread in daily life. At present, smartphones are extremely common, and users have a great demand for images. From the formation, acquisition, transmission, and reception of images, every step is inseparable [2-3]. However, in each stage, the image is more or less contaminated by noise, resulting in users not being able to achieve the expected image effect. However, directly optimizing or removing noise can also affect the accuracy of the image, so advanced noise removal techniques play a crucial role in the efficient use of images [4].

The use and transmission of image information is very common in today's era of electronic products, but the image information required by users is highly susceptible to external signal interference, so image denoising technology needs to be continuously improved [5]. The author described image denoising technology and superpixel generation algorithm, and explained the definitions, principles, and operational processes of these two aspects from a professional perspective, in addition, the author also analyzed the advantages and disadvantages of image denoising and superpixel generation algorithms, continued to innovate and develop based on the advantages, and continuously improved and optimized based on the disadvantages. The most important use of image processing is image recognition, which essentially involves identifying useful information in the image to be recognized through a trained network. Image recognition technology is widely studied due to its important application value and bright application prospects. Currently, image recognition is widely used in industries such as military, agriculture, and social life. In daily life, image recognition technology is constantly used: for example, in the popular smart home, character recognition is a relatively advanced image recognition. In addition, retinal scanning, fingerprint scanning, and other access control systems are also the same. Hospital clinical medical instruments use image recognition to judge and analyze the condition, and these applications have important practical significance [6].

2. Literature Review. With the progress of the times, the era of food and clothing has long passed. We now pay more attention to improving our quality of life, and science and technology are developing unprecedentedly. Whether it is dynamic publishing on social media or publishing articles on various forums, images

*Chongqing Industry & Trade Polytechnic, Chongqing, 408000, China (Corresponding author, HongmingPan8@163.com)

can add more colors and attract readers' attention. It can be seen that the use of images has penetrated into various fields; People in today's era are not only limited to dry and uninteresting written expressions, but also want to express their inner emotions and communicate with each other through images. Therefore, in the world of image transmission, image processing technology is particularly important. Li, Q. et al. proposed a novel and efficient neural convolutional network called MFU. The experimental results on various image denoising datasets (SIDD, DND, and synthetic Gaussian noise datasets) show that our MFU can produce comparable visual quality and accuracy results using state-of-the-art methods [7]. Vo, H. T. and others proposed a method based on deep learning network architecture for image classification of communication systems between autonomous vehicle [8]. Jun, M. et al. proposed a novel end-to-end dual stream convolutional neural network for single image dehazing. The network model consists of spatial information feature flow and high-level semantic feature flow. The spatial information feature flow preserves the detailed information of the dehazing image, while the advanced semantic feature flow extracts multi-scale structural features of the dehazing image. Designed a spatial information assistance module and placed it between feature flows. The experimental results show that the model proposed by the author can restore the desired visual effect without foggy images, and has good generalization performance in real haze scenes [9].

Artificial neural network technology has been successfully applied in multiple fields such as speech recognition, natural language processing, and computer vision. Convolutional neural networks are generated by simulating the visual nervous system and are mainly applied in the fields of natural language processing and computer vision. Convolutional neural networks are the intelligent crystallization of researchers continuously studying the features of advanced motor neurons. By referring to quantum theory and sharing feature parameters in a broad dimension, they significantly reduce the proportion of model storage. In the traditional sense, convolutional neural networks mainly extract some feature points first, and then represent the image through mathematical statistical models, which are often used to solve classification problems. The author applied convolutional neural networks to information image processing, and the final experiment proved that convolutional neural networks have very good performance in processing large-scale image datasets.

3. Method.

3.1. Convolutional Neural Network Structure. The most crucial aspect of convolutional neural networks in data processing is convolutional computation. The image data it processes is usually stored in BMP format, which is a format specified by the global non dynamic image storage organization. Its characteristic is to save the image to its original size without compression, so it requires a hardware system with large storage space. Therefore, after screening, the author selected the GE system, which is an advanced version of CPU and has fast pixel scanning ability, making it very suitable for image processing. And it has the ability to process multi format image conversion, and can transmit multi-dimensional images in multiple layers in parallel, laying a solid foundation for electronic information image processing [10].

3.2. Image Grayscale and Image Enhancement. Using high saturation red as the experimental background for image acquisition can improve the contrast between the parts and the background. The concept of image saturation comes from the HSV (Hue Saturation Value) color space, which is composed of hue h, saturation s, and brightness v. Due to the fact that images are generally represented by RGB color space components r, g, and b, conversion is required. Assuming $I_{max} = max\{r, g, b\}$ and $I_{min} = min\{r, g, b\}$, the relationship between the components s, v in the HSV color space and the RGB color space components r, g, and b is shown in equations 3.1 and 3.2:

$$s = \begin{cases} 0 \quad I_{min} \neq 0 \\ \frac{I_{max} - I_{min}}{T_{max}} \quad I_{max} \neq 0 \\ v = I_{max} \end{cases} \tag{3.1}$$

When there is reflection on the surface of the part, the color in that area appears bright white. The saturation value s in the reflection area of the part will decrease, and the contrast with the background will increase; When shadows appear in the background, the brightness of the area decreases, the saturation increases, and the contrast with the part also increases, as shown in the shaded part of the part. Therefore, extracting the image saturation layer as a grayscale image can effectively suppress the effects of reflection and shadows in the original RGB image. Next, the Luminance Contrast (LC) algorithm based on global contrast is adopted to further improve the contrast of the image [11]. The calculation method for the saliency value SalS (Ik) of pixel Ik in image I is shown in equation 3.3

$$SaIS(I_k) = \sum_{\forall I_i \in l} ||I_k - I_i|| \tag{3.2}$$

The range of values for I_i in the formula is [0.255], and $|| \cdot ||$ represents the distance between grayscale values.

3.3. Image binarization. The purpose of image binarization is to distinguish between parts and background pixels in the image. Firstly, the maximum inter class variance (OTSU) method is used for image binarization. Then, the morphological gradient operator is used to perform closure operations on the image, which can fill some small cracks and holes left by the binarization of the image, and fill the small broken curves into a whole [12].

3.4. Part area extraction. Using seed filling method to search for continuous white pixels in binary images can obtain the part regions and their boundary positions in the binary image. However, this algorithm takes a long time in the seed iteration process and has poor real-time performance. The author proposes a method based on image size transformation and adjusting boundary positions to improve the computational efficiency of seed filling method. The time required for seed filling method is closely related to the image size. Setting a size transformation parameter n and using bilinear interpolation to transform the image size, theoretically reducing the length and width of the image by n times before performing seed filling method operation, the time required will be shortened by 2n times, but it will bring significant operational errors. In order to reduce errors, the original image size is first reduced by n times to improve the computational speed of the seed filling method for obtaining the boundary position. Then, the reduced boundary position coordinates are enlarged by n times as the initial boundary position of the original size binary image. The coordinates of the pixels in the upper left and lower right corners of the initial boundary are (x_1, y_1), and (x_2, y_2), respectively, next, adjust the initial boundary in the original binary diagram to obtain the accurate part area boundary. Set the coordinates of the pixels in the upper left and lower right corners of the adjusted boundary to (x'_1, y'_1), (x'_2, y'_2), respectively.

The initial boundary position generally has a certain degree of error and needs to be adjusted. Each boundary has three positional relationships with the part: Intersection, separation, and tangent. The tangent state boundary is the precise boundary, and by traversing the pixel information of boundaries a, b, c, and d, the positional relationship between the boundary and the part area can be determined. If the boundary intersects with the part, it moves towards the outer direction. If it is apart, it moves towards the center of the image. Iterative operation is performed, and the change in position relationship is used as the termination condition to adjust the boundary to a tangent state, represented by (x'_1, y'_1), (x'_2, y'_2). In order to facilitate the training of subsequent CNN models, a rectangular image, namely the part area image, is extracted from the original RGB image based on the coordinates of boundary pixel points (x'_1, y'_1), (x'_2, y'_2). The short edges of the part area image are filled with red pixels to form a square, and then bilinear interpolation is used to unify the image size [13,14].

3.5. Overall Algorithm Structure. The part recognition algorithm proposed by the author first requires part area extraction, and then recognizes the extracted part area images, which can simplify the model and greatly reduce the workload of model training. The process of part recognition algorithm is shown in Figure 3.1.

3.6. CNN Based on Key Point Features of Parts. The training and testing of CNN based on part key point features are shown in Figure 3.2. Firstly, the SIFT algorithm is used to extract visual vocabulary vectors, namely feature vectors, from the image. These vectors represent locally invariant key point features in the image; Next, we use the Bag of Words (BoW) model to process the data, gather all feature vectors together, use K-means clustering algorithm to merge visual vocabulary with similar meanings, and construct a dictionary

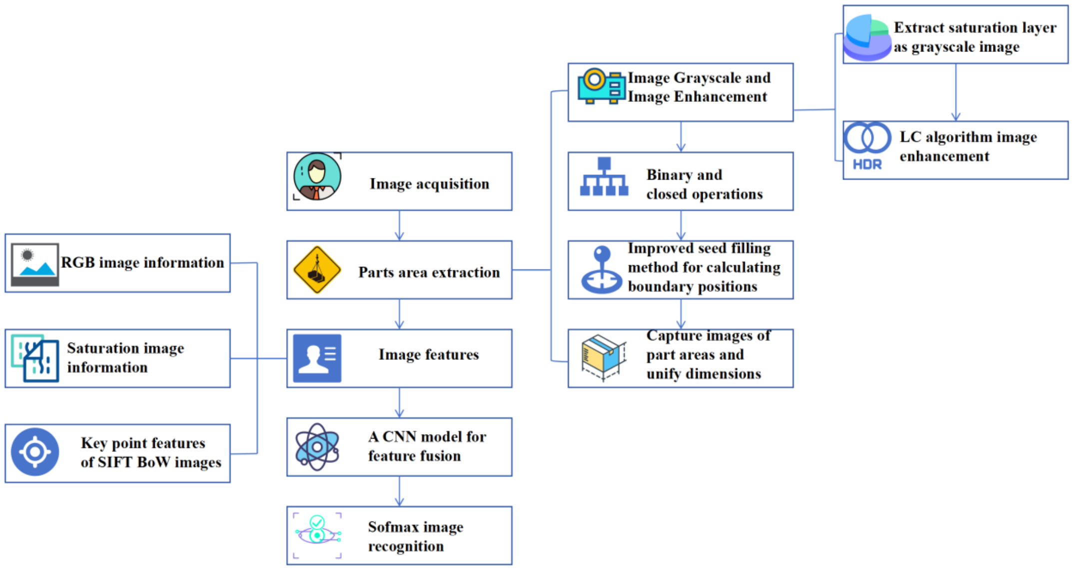

Fig. 3.1: Process of Part Recognition Algorithm

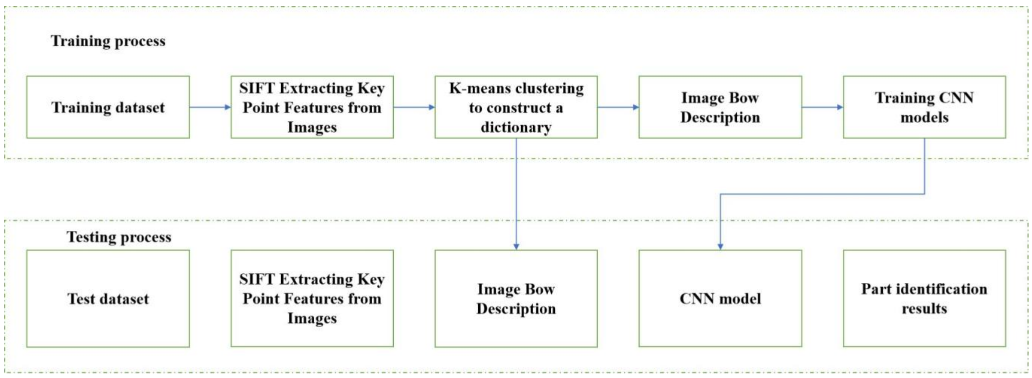

Fig. 3.2: Training and Testing of CNN Based on Key Point Features of Parts

containing K words, by counting the number of times each word in the dictionary appears in the image, the image can be represented as a K-dimensional feature vector, which is the BoW description of the image; Finally, the vector is used as input for a CNN based on part keypoints, and convolution and normalization operations are performed on it to obtain the feature vector of part keypoints [15].

3.7. CNN model for feature fusion. After experimental comparison, the VGG16 model was selected as the feature extraction network. The input was a 64 * 64 * 3 RGB image and a 64 * 64 * 1 saturation image stacked to form a 64 * 64 * 4 matrix, which underwent five double convolution operations and four max pooling operations. Due to the important role of the geometric shape of parts in part recognition, its importance is higher than features such as color and surface texture. By stacking high contrast saturation images with RGB images as inputs to the CNN model, the geometric shape features of the parts are enhanced during network training.

The SIFT-BOW model can extract local keypoint features of part images, which are feature vectors with a size of $1 \times 1 \times K$. Firstly, K 1xK convolution operations and normalization operations are performed on them, and then enter the fully connected layer to obtain a 128 dimensional feature vector of part keypoints. The

Table 4.1: Part Area Extraction Time

number	Part Name	Time/s for different size transformation parameters n					
		n=1	n=2	n=4	n=6	n=8	n=10
01	Gear shaft	1.045	0.388	0.213	0.175	0.164	0.162
03	bearing	1.001	0.354	0.182	0.153	0.141	0.134
04	Driven shaft	1.076	0.386	0.222	0.182	0.164	0.158
05	bolt	1.003	0.347	0.187	0.153	0.148	-
07	Gearbox cover	1.535	0.621	0.383	0.336	0.326	0.316
08	Gearbox seat	1.46	0.538	0.276	0.233	0.213	0.206
09	Big gear	1.205	0.436	0.243	0.193	0.182	0.177
12	Safety valve box	1.624	0.437	0.241	0.202	0.190	0.183
15	Belt pulley	1.113	0.392	0.210	0.176	0.163	0.160

key points of the part are concatenated with the feature vectors of the part shape to obtain a 256 dimensional feature vector. After passing through a fully connected layer, its size is reduced to $1 \times 1 \times c$. Finally, the recognition result is obtained using the Softmax algorithm. By concatenating local keypoint feature vectors at the end of the network, the feature loss generated during downsampling can be compensated for, the feature extraction ability can be enhanced, and the recognition accuracy can be improved [16].

3.8. Construction of experimental platform. Experimental environment: Windows 10, 64 bit operating system, memory size AMD Ryzen, GPU GTX-1650, and commonly used environments such as Python 3.6, Tensorflow 2.0, and Opencv were built. Place the tested part on a red background stage, use a strip LED light source for illumination, and keep the phone and computer in the same network environment using an IP camera for shooting and data transmission. In the experiment, the camera and desktop were kept horizontal, and the camera was about 36cm higher than the desktop. Collect samples under ambient light of different brightness levels in the experiment; The parts are arranged in a disorderly manner but not stacked on top of each other [17,18].

4. Results and Discussion. The dataset images for the experiment include 19 types of actual captured plunger pumps, gearboxes, gear oil pumps, safety valves, and ball valve components. 80 samples were collected for each type of component as initial samples. In order to increase data diversity, the images were expanded through horizontal and vertical flipping, as well as random changes in contrast and brightness, to improve the robustness of the model. A total of 1280 samples were obtained for each type of component. Train the model using a ratio of 7:3 between the training and testing sets. The original image size is 480 x 640, and the sample image size after extracting the part area is uniformly 64 x 64. Perform image grayscale, binarization, and morphological processing on the collected images in sequence, and then extract the part area. The core of the extraction algorithm is the size transformation parameter n. The author conducted experimental comparisons on the unchanged size of n=1 and the improved algorithm with n=2, 4, 6, 8, and 10, respectively, to explore the impact of different size transformation parameters n on the efficiency of part region extraction. The extraction time for some part areas collected in the experiment is shown in Table 4.1.

As shown in Table 4.1, directly extracting the part area from the original image, that is n=1, takes more than 1 second, which is inefficient. By changing the size of the image, the computational efficiency can be significantly improved. As n continues to increase, the improvement in computational efficiency gradually decreases. This is because image preprocessing takes time independent of the n value before size transformation. At the same time, in order to explore the impact of different size transformation parameters n on the quality of part region extraction, boundary position errors were calculated for some parts with n=2, 4, 6, 8, and 10 pairs, respectively. The results are shown in Table 4.2.

According to Table 4.2, when n=2, n=4, and n=6, the boundary position error pixels of each part are all lower than 3 pixels, indicating high accuracy; When n=8 and n=10, the boundary position error of the bolt significantly increases or even misses, due to the excessive reduction in image size causing loss of pixels in small parts. Therefore, considering the time required for extracting part regions in Table 4.3, the size transformation

Table 4.2: Boundary position errors for different parameters n

number	Part Name	Time/s for different size transformation parameters n				
		n=2	n=4	n=6	n=8	n=10
01	Gear shaft	1.24	1.24	1.24	1.3	1.34
03	bearing	0	0	0	0	0
04	Driven shaft	0.92	0.92	0.92	0.92	0.92
05	Snail inspection	1.6	1.6	2.6	14.8	-
07	Gearbox cover	1.04	1.04	1.04	1.04	1.04
08	Gearbox seat	0.42	0.42	0.42	0.42	0.42
15	Belt pulley	0.1	0.1	0.1	0.1	0.1

Table 4.3: Main parameters of CNN model

Parameter Name	Method/Value	Parameter Name	Method/Value
Number of images loaded each time	32	Training cycle	8
optimization algorithm	Adam	Single cycle time/min	8
Convolutional kernel	[3 × 3]	Trainable parameters/piece	2,290,724
Cluster center k	100	Number/type of sample categories	19

Table 4.4: Recognition Results of Different Algorithms

method	accuracy
PCA dimensionality reduction SVM	82.52%
SIFT-SVM	92.85%
SIFT-CNN	96.15%
CNN (direct extraction of image features)	97.80%
Feature fusion CNN model	98.84%

parameter n=4 is the optimal parameter for the part region extraction algorithm [19].

The selection of n value needs to consider the size of the parts. If identifying smaller parts, such as bolts, nuts, etc., the n value should not be too large. A value between 2 and 6 can maintain high recognition accuracy. If the size of the identified parts is large, such as large gears, safety valve bodies, etc., taking an n value between 6 and 10 can bring high efficiency and ensure recognition accuracy. The training time of this algorithm is only 8 minutes, which requires several hours of training compared to large object detection models. It also avoids the annotation work on the dataset. The training time of this algorithm is short and practical. After experimental comparison, the CNN model adopts the following parameters, as shown in Table 4.3.

In order to verify the progressiveness and effectiveness of the feature fusion CNN model, it is compared with a single model and existing recognition methods, and the results are shown in Table 4.4.

According to Table 4.4, the accuracy of the model using only Principal Component Analysis (PCA) and SIFT for feature extraction is not high, while the CNN model is good at extracting features from images and generalization. For CNN models that do not contain SIFT features, the accuracy of directly extracting features from RGB images is 97.80%. On this basis, the feature fusion CNN model superimposes saturation images on the input layer, enhancing the shape features of parts in network training. Combined with the key point feature vectors concatenated at the end of the network, the accuracy of part image recognition reaches 98.84%. Due to the unified size of 64×64 during the part extraction process, the CNN model takes about 2ms to predict, and the overall recognition algorithm speed is about 5fps, including the time for part area extraction.

5. Conclusion. The author proposes a research on electronic information image processing based on convolutional neural networks. In order to address the issues of easy reflection, unevenness, and uneven lighting on the surface of mechanical parts, as well as the interference caused by shadows and mixed other parts on recognition, a seed filling method is proposed to extract part regions through image saturation transformation and size transformation. This method can quickly and accurately extract part regions from images. Dividing the task of part image recognition into two steps: part region extraction and image recognition is beneficial for improving the scalability of the algorithm. Only by retraining a lightweight CNN classification model, new parts can be added for recognition, avoiding tedious part position annotation and complex boundary regression operations, which can reduce the computational power requirements of computing equipment. The author proposes a feature fusion CNN model that overlays RGB images with saturation images and inputs them into the CNN model. At the end of the CNN, feature vectors based on key points in the part image are concatenated and then passed down through the network. This can compensate for the loss of local information in the image during convolutional pooling, thereby further improving the accuracy of the recognition algorithm.

REFERENCES

[1] Yanshan, L. I., Zhou, L., Fan, X. U., & Chen, S. (2022). Ogsrn:optical-guided super-resolution network for sar image, 35(5), 204-219.

[2] He, H., Yang, P. F., Zhang, P. F., Li, G., & Zhang, T. C. (2023). Single-photon source with sub-mhz linewidth for cesium-based quantum information processing, 18(6), 61303.

[3] Zhang, Y. L., Feng, X. L., Chang, X. L., & Tie, L. M. (2023). Impacts of canopy structure on the sub-canopy solar radiation under a deciduous forest based on fisheye photographs, 15(3), 150-160.

[4] Wang, Z., Li, L., Xing, L., Wang, J., Sun, K., & Ma, H. (2023). Information purification network for remote sensing image super-resolution. Tsinghua Science and Technology, 28(2), 310-321.

[5] Liu, X., Nicolau, J. L., Law, R., & Li, C. (2023). Applying image recognition techniques to visual information mining in hospitality and tourism. International Journal of Contemporary Hospitality Management, 35(6), 2005-2016.

[6] Zhang, D., Zhao, L., Duanqing, X. U., & Dongming, L. U. (2022). Dual-constraint burst image denoising method. Frontiers of Information Technology & Electronic Engineering, 23(2), 220-233.

[7] Guo, X., He, P., Lv, X., Ren, X., Li, Y., & Liu, Y., et al. (2023). Material decomposition of spectral ct images via attention-based global convolutional generative adversarial network, 34(3), 143-153.

[8] Vo, H. T., Nguyen, N. L., Ngoc, D. N., Do, T. H., & Pham, Q. D. (2023). Binary Image Classification Using Convolutional Neural Network forV2V Communication Systems, 28(4), 743-753.

[9] Jun, M., Yuanyuan, L., Huahua, L., & You, M. (2022). Single-image dehazing based on two-stream convolutional neural network. Journal of Artificial Intelligence Technology (English), 76(9), 3859-3876.

[10] Yang, W., Huang, Z., & Zhu, W. (2023). A first-order rician denoising and deblurring model. Inverse Problems and Imaging, 17(6), 1139-1164.

[11] Chen, Q. (2022). Research on 3d mfl testing of wire rope based on empirical wavelet transform and srcnn. Journal of Vibroengineering, 24(4), 14.

[12] Singh, B. K., Nair, N., Falgun, P. A., & Jain, P. (2022). Quantitative evaluation of denoising techniques of lung computed tomography images: an experimental investigation. International Journal of Biomedical Engineering and Technology, 38(2), 151.

[13] Wen, Y., Sun, J., & Guo, Z. (2022). A new anisotropic fourth-order diffusion equation model based on image features for image denoising. Inverse Problems and Imaging, 16(4), 895-924.

[14] Joshi, S., Singla, N., Ahuja, S., Saini, S. K., Thakur, N., & Jindal, P., et al. (2022). Denoising of computed tomography using bilateral median based autoencoder network. International Journal of Imaging Systems and Technology, 32(3), 935-955.

[15] Huapeng Tang, Danyang, Q., Mengying, Y., Jiaqiang, Y., & Gengxin, Z. (2023). Research on color image matching method based on feature point compensation in dark light environment, 29(1), 78-86.

[16] Liu, X., Nicolau, J. L., Law, R., & Li, C. (2023). Applying image recognition techniques to visual information mining in hospitality and tourism. International Journal of Contemporary Hospitality Management, 35(6), 2005-2016.

[17] ZHOU Qian-fei,WANG Hui-fang,QING Guang-wei,HU Jing-bo. (2022). Detection of crane structural connecting bolts falling off based on uav image recognition. Manufacturing Automation, 44(12), 20-22.

[18] Liu, B., Li, J., Yang, X., Chen, F., Zhang, Y., & Li, H., et al. (2023). Diagnosis of primary clear cell carcinoma of the liver based on faster region-based convolutional neural network. Chinese Medical Journal, 136(22), 2706-2711.

[19] Zhao, S., Mao, G., Xiong, B., Huang, S., & Lin, J. (2023). Spatio-temporal graph mining model based on graph wavelet convolutional neural network. Computer Engineering, 49(7), 85-93.

Edited by: Bradha Madhavan
Special issue on: High-performance Computing Algorithms for Material Sciences
Received: Jun 28, 2024
Accepted: Aug 6, 2024