www.cad-journal.net
Computer-Aided Design AND APPLICATIONS

Computer-Aided Industrial Product Design based on Image Enhancement Algorithm and Convolutional Neural Network

Kai Xiao[1] and Taile Ni[2]

[1]School of Communication, Guangxi Vocational Normal University, Nanning, Guangxi 530007, China, xikaii@163.com
[2]School of Literature, Journalism&Communication, Xihua University, Chengdu, Sichuan 610000, China, nitaile@mail.xhu.edu.cn

Corresponding author: Taile Ni, nitaile@mail.xhu.edu.cn

Abstract. In order to effectively realize the integration of conceptual design and detailed design, AI should be comprehensively used in product design to process the geometric feature information of products and realize the comprehensive design and high-quality production of new products. This article discusses the application of digital image processing technology based on computer vision and artificial intelligence (AI) in computer aided industrial design (CAID), and proposes an image enhancement algorithm for industrial design products based on convolutional neural network (CNN). The algorithm enhances the stability and anti-over-fitting performance of the network by compressing the model of CNN with image enhancement processing, and optimizes the design process and work efficiency of CAID. In the case of expanding network scale, this article adopts parallel verification method, which is more effective than single user verification method. The simulation shows that the model constructed in this article can quickly train to achieve the target accuracy of CAID image processing, and can support the feature extraction and modeling stage of CAID, which can not only optimize the design content, but also improve the design efficiency.

Keywords: Artificial Intelligence; Computer Vision; Convolutional Neural Network; Computer Aided Industrial Design
DOI: https://doi.org/10.14733/cadaps.2024.S3.92-106

1 INTRODUCTION

The arrival of the Industrial 4.0 era, represented by AI, has prompted AI to penetrate into many fields of people's lives and work. As an industrial design closely related to people's lives, the application of AI in industrial design has become the development trend of modern industrial design. Cao [1] conducted algorithm coefficient feature analysis for drone image stitching. It constructs an accurate image quality matching equation for the sensor. By combining the sensor deep learning index of convolutional networks, the edge difference index comparison of structural similarity was optimized. The combination of Image registration algorithm and CNN model can achieve more accurate and robust mosaic effect. The Image registration algorithm is used to register two images in the video to obtain the transformation matrix between them. Then, input the two registered images into the CNN model for feature extraction and stitching. Specifically, a deep learning-based image stitching scheme can be used. Apply the trained model to video image stitching tasks in practical scenarios to achieve more accurate and robust stitching results. In general, the combination of Image registration algorithm and CNN model is applied to video image mosaic, which can achieve more accurate and robust mosaic effect by learning image feature representation and mosaic algorithm, and can also adapt to image mosaic tasks in various scenes and conditions. Industrial design is not only the carrier of materialization of sci& tech, but also the carrier of commercialization of sci& tech. With the increasing complexity of new product design, the details and problems of the product need to be comprehensively considered in the stage of new product design, and further optimized design and innovative design are needed. Dhaya [2] analyzed the image accuracy resistance analysis of digital watermarks. It constructs a robust image Convolutional neural network Feature selection construction. A robust image watermarking scheme based on lightweight CNN is a deep learning method for image watermarking technology. This scheme improves the security and robustness of the watermark by using a lightweight CNN model to extract and embed watermark information. The lightweight CNN model refers to a neural network model with fewer parameters and computational complexity, which can effectively reduce the computational complexity and memory occupation of the model, making it suitable for resource constrained embedded systems and mobile devices. The security of this scheme mainly depends on the security of the watermark information and the robustness of the model. In order to improve the security of watermark information, some encryption techniques and privacy protection measures can be used, such as chaotic encryption, hash functions, etc., to protect the integrity and confidentiality of watermark information. Computer image processing technology is used to understand and analyze images, which usually refers to preprocessing, segmenting, detecting the characteristics of objects in images and describing them with some data structure. Further analysis also includes classification, recognition and understanding according to image content. Applying the powerful AI image processing method to CAID can not only develop a more efficient design process and guide the transformation of designers, design industry or design education, but also create a brand-new lifestyle for the public and promote social development when people generally pay attention to efficiency, experience and emotion.

A novel machine learning based hybrid CNN model for tumor recognition in medical image processing can significantly improve the accuracy and reliability of tumor recognition. This model combines the traditional Convolutional neural network (CNN) and the latest deep learning technology, including Recurrent neural network (RNN), attention mechanism and self-attention mechanism. This Mixture model can simultaneously use the spatial and temporal information of the image to improve the recognition accuracy. Dhiman et al. [3] first used CNN to extract features from medical images, and then input the extracted features into RNN to utilize the temporal information of the image. On the basis of RNN, attention and self-attention mechanisms can be introduced to further enhance the model's ability to recognize images. This new hybrid CNN model based on machine learning has broad application prospects in tumor recognition tasks in medical image processing, and can provide strong support for medical diagnosis and treatment. Gonthina et al. [4] conducted a structural safety evaluation of image quality. Use a deep CNN model to extract features related to concrete cracks from the preprocessed images. Classify images using a trained deep CNN model to determine if there are cracks in the image, and locate the cracks using bounding boxes or masks. Quantifying the located cracks can use various indicators, such as the length, width, depth, shape, etc. of the cracks. The evaluation of the results shows that using the test dataset to evaluate the accuracy and reliability of crack identification and quantification can compare the performance of different models to select the optimal model. Through the above steps, image processing technology can be used to identify and quantify concrete cracks based on deep CNN, thereby improving the safety and reliability of concrete structures. The traditional image classification is mainly manual classification, but with the rapid growth of the number of images, manual classification is costly and takes a long time, which is no longer suitable for the current image classification task. In addition, due to human subjectivity, it is difficult for people to classify images uniformly. The technical stage of image processing should mainly be the analysis and recognition of different features of the image, and at the same time, it is necessary to describe the relationship between the nature of the image and the modeling elements and colors more deeply, and find out the meaning and recognition focus of the image, so as to guide and plan the next specific action of image recognition. In order to effectively realize the integration of conceptual design and detailed design, AI should be comprehensively used in product design to process the geometric feature information of products and realize the comprehensive design and high-quality production of new products. Deep learning is widely used in classification tasks because of its unique structure and learning ability. Compared with traditional methods, it has stronger generalization ability and can handle more complex classification tasks. And CNN is a representative deep learning structure. CNN belongs to a kind of important research in the field of AI computing, and has achieved outstanding research results in some fields and been widely used. This article mainly discusses the application of AI in CAID, and proposes a CAID method based on computer vision and CNN.

Hou et al. [5] analyzed the detection structure and public data security analysis of road traffic images. The application of invasive sensing, image processing technology, and machine learning methods has significant advantages in road surface monitoring and analysis. These technologies can achieve long-term monitoring of road dynamic response, which will help timely detect road diseases and prevent road safety accidents. These technologies can automatically or semi-automatically detect and identify some typical types of road surface diseases, such as potholes, cracks, ruts, etc., greatly improving work efficiency and accuracy. In traditional road surface monitoring methods, this type of work usually requires manual labor, which is labor-intensive and inefficient. In addition, analysis work needs to be based on a large amount of monitoring data, and methods such as machine learning can process large-scale data and effectively mine information from the data. The key to the development of digital image information field is image classification technology, which is also the basis of the development of image retrieval system. Image classification can automatically classify massive images, and image retrieval can quickly and efficiently obtain effective information on this basis. The two complement each other and are the stage of continuous exploration in the field of AI [10]. The development of computer vision has made great progress with the improvement of computer hardware equipment and software technology. Efficient computing, fast data transmission, high-capacity storage and mobile network communication all bring convenience to the design of industrial products, but also break through the limitation of computer size [11]. The earliest graphic image processing was mainly an activity aimed at satisfying people's clarity of image quality and the best visual effect, and it was not further applied to production and life, which can be said to be a purely artistic requirement [12]. Therefore, in the stage of processing, it is mainly to modify the image with low resolution and poor image quality, so that the modified image can give people the best visual effect. This article discusses the application of digital image processing technology based on computer vision and AI in CAID, and makes the following innovations:

⊖ By analyzing the visual information, this article preprocesses and saves each frame image and arranges it into a sequence with a certain data structure type. By using the method of computer vision to process images, it detects, extracts and recognizes any frame image, which provides decisions for the next processing and analysis of CAID system.

⊜ Model compression of CNN for image enhancement processing enhances the stability and anti-over-fitting performance of the network, and optimizes the design process and work efficiency of CAID, thus realizing the application of AI in product design.

This article introduces the method of industrial product image feature fusion based on computer vision, and uses CNN in CAID image enhancement processing. Then the practical application effect of this method is verified on PASCAL-VOC data set. Finally, review the full text, summarize and analyze the research results, compare the experimental data and analyze the shortcomings, and make plans for the future research work.

2 RELATED WORK

Hu et al. [6] conducted a super-resolution low-light remote sensing image training architecture for image enhancement. Low light remote sensing image enhancement is an important issue, as low light conditions can affect the quality and recognition ability of the image. CNN based methods can effectively address this issue. Collect a dataset of low light remote sensing images, including images from different scenes, times, and angles. Preprocess image data, including unifying image size, adjusting brightness and contrast, etc. Build a CNN model, including convolutional layer, pooling layer, fully connected layer, and output layer. Train using preprocessed datasets and CNN models, adjust model parameters to improve model accuracy and generalization ability. Apply the trained model to practical low light remote sensing image enhancement problems, and perform image enhancement according to actual needs, such as improving image brightness and contrast, removing noise, etc. In summary, CNN based weak light remote sensing image enhancement methods can achieve image enhancement through deep learning technology, improving image quality and recognition ability. It should be noted that in practical applications, adjustments and optimizations need to be made based on specific problems to achieve better results. Hu et al. [7] processed the intensity distribution of image intensity information automatically by software. By constructing image methods based on deep learning, the magazine image model threshold of other extracted image data targets is obtained. The deep learning method based on the Mask R-CNN model can be used for image fluorescence processing to extract and recognize fluorescent regions in images. Firstly, the Mask R-CNN model extracts features through pre trained models, which can extract regions of interest and bounding boxes from the image. Then, these regions and bounding boxes are fed into a fully connected layer for classification to determine whether they contain fluorescent regions. If a region is classified as containing fluorescent regions, the model will generate a mask to separate the region from the background. This mask can be used for further analysis of fluorescence images, such as measuring fluorescence intensity, analyzing fluorescence morphology, etc. In addition, deep learning methods based on the Mask R-CNN model can also be used for tasks such as object detection and image segmentation. This model has strong feature extraction and classification capabilities and can be applied to many different image processing tasks. At present, many advanced image denoising methods are based on CNN, such as self-attention mechanism-based image denoising methods, residual learning based image denoising methods, and multi-scale processing based image denoising methods. These methods can achieve better denoising effects and performance by combining different denoising algorithms and CNN models. When using CNN for image denoising, many factors need to be considered, such as the quality of the dataset, the structure and parameter settings of the model, and the selection of denoising algorithms. Therefore, it is necessary to determine the optimal image denoising method through analysis and experimentation of different factors.

Ilesanmi and Ilesanmi [8] conducted image noise removal analysis. By considering the dataset analysis of neural networks, it constructed a formal data analysis for CNN image network denoising. Image denoising is an important issue in the field of image processing, which can be used to improve image quality and visual experience. Convolutional neural network (CNN) is a deep learning technology suitable for image denoising. It can achieve efficient image denoising by learning the image feature representation and denoising algorithm. When using CNN for image denoising, it is first necessary to collect a dataset containing both noisy and non-noisy images in order to train the CNN model. Then, convolutional and pooling layers can be used to extract features from the image, and denoising algorithms can be used to post-process the feature map to achieve image denoising. Jia et al. [9] improved the encoding performance of convolutional network image processing. It constructs a multi model filtering mechanism for image and video network discrimination. Content aware Convolutional neural network used for intra loop filtering in efficient video coding is an image processing technology based on deep learning, which can be used to improve the efficiency and performance of video coding. This neural network can predict the texture and details in images by learning image features and texture information in videos, and efficiently encode them accordingly. At the same time, the neural network can also perform adaptive filtering based on the content of the image to reduce encoded image distortion and noise. Specifically, the neural network includes a series of convolution layers and pooling layers to extract the feature representation of the image, and uses Recurrent neural network (RNN) to capture the temporal correlation of the image. Liu et al. [10] performed blur operation image processing for Gaussian image preprocessing. In iris recognition, blurry images are often encountered due to issues such as non-circular pupil shape, changes in light intensity and brightness, and user cooperation. To solve this problem, fuzzy image enhancement technology can be used. Among them, deep learning technology has great advantages in this regard. For example, deep learning facial database pre training and data enhancement methods can be used to enhance blurry iris images. Specifically, deep learning detection and segmentation methods for iris images can be trained by enhancing the data of blurred images and lighting changes. At the same time, to address the issues of non-circular and scale changes in the iris, deep learning semantic segmentation methods can also be attempted, by pre training public datasets and incorporating multi-scale schemes during training to perform iris segmentation. In summary, deep learning technology plays an important role in enhancing blurry images in iris recognition, and can improve the quality of iris images through deep learning methods, thereby improving recognition accuracy. Pu et al. [11] collected a large amount of image data on coal and gangue, including different angles, lighting, and scenes. Then, pre trained CNN models such as AlexNet, VGG, ResNet, etc. can be used for feature extraction and classification of images. In Transfer learning, you can use the pre trained CNN model as the basic network, then fix its parameters, and only allow its output layer to train to adapt to different classification tasks. This can improve the classification accuracy of the model while retaining the basic network feature extraction ability. In order to improve the accuracy and robustness of image recognition, data augmentation techniques such as rotation, translation, scaling, etc. can also be used to increase data diversity and improve the generalization ability of the model.

Sharma et al. [12] analyzed the training image data construction of Convolutional neural network. The deep learning CNN model using image segmentation for plant disease detection can be used to analyze and extract features of plant diseases, classify and recognize them. Using cross validation methods to train and test the model can evaluate its accuracy and reliability. Various indicators can be used to evaluate the performance of the model, such as accuracy, recall, F1 score, etc. Evaluate the model's resistance to noise and interference, as well as its adaptability to different experimental conditions. If the model can work stably under various conditions, it indicates that the model is relatively stable. Overall, deep learning CNN models using image segmentation for plant disease detection have high accuracy and reliability, and can successfully identify and classify various plant diseases. At the same time, the model also has good generalization ability and stability, and can adapt to different experimental conditions and environments. However, the running speed of this model may be slow and needs to be optimized during training and testing. The application of deep learning in the field of image detection has made significant progress. Srivastava et al. [13] extracts image features through layer-by-layer convolution and pooling, and uses Backpropagation for training. CNN has high accuracy and robustness in image detection, but requires a large amount of training data and computational resources. SVM is a classical machine learning algorithm, which separates different types of samples by finding Hyperplane. In image detection, SVM can be used for classification tasks, but its performance is limited for complex images and multi category tasks. In image detection, Random Forest can be used for Feature selection and classification tasks, but its performance for complex images and multi category tasks is limited. Overall, different deep learning image detection algorithms have their advantages and disadvantages, and the specific selection needs to be determined based on actual application scenarios and task requirements. CNN has high accuracy and robustness in image detection, but requires a large amount of training data and computing resources; The performance of SVM and Random Forest is limited in simple image and multi category tasks. Therefore, it is necessary to choose the most suitable algorithm based on actual needs. Zeng et al. [14] conducted high-resolution image adaptive 3D enhanced tone analysis. It constructs an image adaptive fusion image parameter color tone difference measurement. It automatically generates enhanced lookup tables suitable for different images and scenes by learning a large amount of image data, thereby achieving real-time enhancement of photos. The main idea of IA-TLUT is to divide a large amount of image data into different categories and learn a lookup table for each category. These lookup tables can be used to enhance new images with attributes similar to the learned category, such as brightness, contrast, color saturation, etc. Zhao et al. [15] analyzed the coupling relationship of low light level images. It constructs image enhancement strategies through learning and analyzing deep frameworks. Low light image enhancement is an important issue, as low light conditions can affect the quality and recognition ability of images. A unified depth framework is a method that can handle multiple image enhancement tasks in a unified manner, including low light image enhancement. The core idea of a unified depth framework is to treat multiple image enhancement tasks as the same problem, but with different parameters and weights. In this way, the same deep learning model can be used to handle different image enhancement tasks, thereby saving computational resources and improving efficiency. Train the model using datasets from multiple image enhancement tasks, and improve the accuracy and generalization ability of the model by adjusting model parameters and optimizing algorithms. Use a test dataset to test the performance of the model, evaluate its enhancement effect and recognition ability. Apply the trained model to actual image enhancement tasks, and perform image enhancement according to actual needs, such as improving image brightness and contrast, removing noise, etc.

3 METHODOLOGY

3.1 Visual Perception and Feature Fusion of Industrial Products

Since the industrial revolution, with the help of industry and market forces, the social influence and social penetration of design have been continuously enhanced, and the field of design has been continuously expanded. The purpose of industrial design is to fully understand the actual needs of users through information interaction, and then constantly improve the existing products. In order to effectively improve the quality of information interaction, we must do a good job in the dissemination and promotion of industrial products. Through the application of image recognition technology, relevant personnel can have an all-round understanding of the actual situation of users and the surrounding environment, and then innovate in product form design and content design. It uses technical knowledge and experience to visually evaluate the packaging, publicity and display of products.

The image processing of product design involved in this article is one of the research contents of computer vision, with computer as the core part and visual perception equipment to obtain continuous visual information. By analyzing the visual information, each frame of image is preprocessed and saved, and arranged into a sequence with a certain data structure type. By using the method of computer vision to process images, any frame of image is detected, extracted and recognized, which provides decisions for the next processing and analysis of CAID system, thus realizing the application of AI in product design.

CNN can be understood as a feedforward network, which belongs to an artificial neural network. In the network, each neuron is transmitted step by step from the input layer and finally reaches the output layer. This transmission is irreversible, that is, it only transmits in one direction, so it cannot form a cycle, that is, there is no periodicity. CNN is formed by the superposition of convolution layer of feature extraction and sampling layer of feature processing. Its function scope is to improve the recognition accuracy of deep neural network on the basis of unchanged tag data. The demand for its parameters is relatively large, and too many parameters will make the network over-fit the training data set. The CNN operation flow of product image processing is shown in Figure 1.

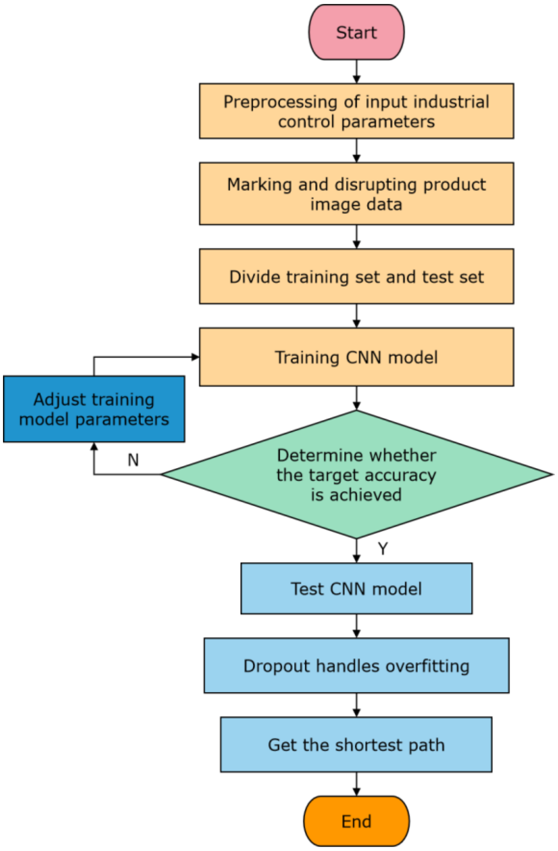

Figure 1: CNN operation flow of product image processing.

For the pooling operation, if the continuous range of images is taken as the operation position, the reddening application with convolution characteristics will be derived between the operation position and the nerve, which will make the pooling operation have the characteristics of translation invariance, resulting in the continuous output of the same classification results when the classifier and the same features are pooled. Normalization is a method to avoid the fitting problem with the increase of training neural network. Through the analysis of a large number of data, it can be seen that over-fitting generally occurs in a relatively small space, because the function value formed by the data changes greatly, resulting in a relatively large parameter value of the function and a relatively large absolute value of some functions with relatively small intervals.

Assuming that two cameras shoot a certain point $P(x_c, y_c, z_c)$ in the real scene at the same time, the imaging points of the P point in the left and right cameras are $P_1(u_1, v_1), P_2(u_2, v_2)$, and its P_1, P_2 ordinate is the same, but its abscissa is u_1, u_2, so the position deviation of the left and right image pairs in the horizontal direction is $u_1 - u_2$, which is recorded as parallax d. Then the sizes of u_1, u_2 and v_1 can be obtained from the triangle similarity principle:

$$\begin{cases} u_1 = f\dfrac{x^c}{z^c} \\ u_2 = f\dfrac{\left(x^c - b\right)}{z^c} \\ v_1 = v_2 = f\dfrac{y^c}{z^c} \end{cases} \tag{1}$$

After camera calibration and correction, the epipolar lines of the left and right images are collinear and parallel to the horizontal direction, that is to say, the left and right matching points have the same ordinate, so the search can be carried out along the one-dimensional scanning line in stereo matching. The relationship between matched pixel pairs is:

$$d(x, y) = x_L - x_R \tag{2}$$

Where x_L is the characteristic pixel in the left image, x_R is the corresponding pixel in the right image, and $d(x, y)$ is called parallax, that is, the offset of the characteristic point in the left and right image positions.

In the stage of product production, image recognition technology can not only shorten the product manufacturing time, improve the product yield, improve the production efficiency of factories and enterprises, but also improve the product quality. This provides a guarantee for designers in the production flow. The task of image classification depends on supervised learning to a great extent, that is, each sample has a corresponding label. Through the deep neural network, we can constantly learn the corresponding features of each tag and finally realize the classification. In this case, the capacity of data set and the quality of labels often play a decisive role in the performance of the model. The product image feature fusion mode is shown in Figure 2.

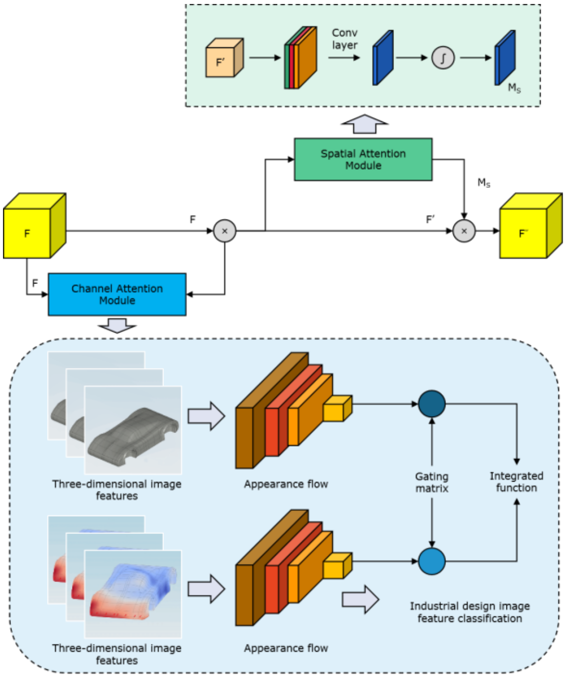

Figure 2: Feature fusion mode of product image.

In the stage of image analysis, by dividing it, different types of operations are produced, which are all one kind of image classification. On the one hand, the application of information interaction has greatly changed the technical means, procedures and methods of industrial design. In the image, the local domain will be used in other fields because of its own characteristics, so it is necessary to aggregate the characteristics of the local domain of the image.

Let X_i^k represent the sum of inputs of k layer neurons i, Y_i^k is the output, and the weights of $k-1$ layer neurons j to k layer neurons i are W_{ij}, then there is the following functional relationship:

$$Y_i^k = f\left(X_i^k\right) \tag{3}$$

$$X_i^k = \sum_{j=1}^{n+1} W_{ij} Y_j^{k-1} \tag{4}$$

Generally, f A is an asymmetric Sigmoid function:

$$f\left(x_i^k\right) = \frac{1}{1+\exp\left(-X_i^k\right)} \tag{5}$$

If the output layer is the m layer, the actual output of the i neuron in the output layer is Y_i^m. Let the corresponding human body signal be Y_i, and define the error function e as:

$$e = \frac{1}{2}\sum_i \left(Y_i^m - Y_i\right)^2 \tag{6}$$

Computer vision is to use machines instead of human eyes to measure and judge. In the application of vision technology, the camera converts the captured image into a graphic signal, which will be processed by a special image processor. The processing system will convert it into a digital signal according to the distribution of the image and the size, brightness and color of the object, and then the image processing system will analyze the target characteristics according to these signals. In the stage of image acquisition and transmission, the image quality is degraded due to factors such as environment, equipment and transmission mode. It can be seen from Figure 3 that due to the influence of illumination conditions and reflectivity, the obtained product image has the characteristics of low contrast, low signal-to-noise ratio and uneven illumination. This fact will further reduce the identification accuracy of existing algorithms.

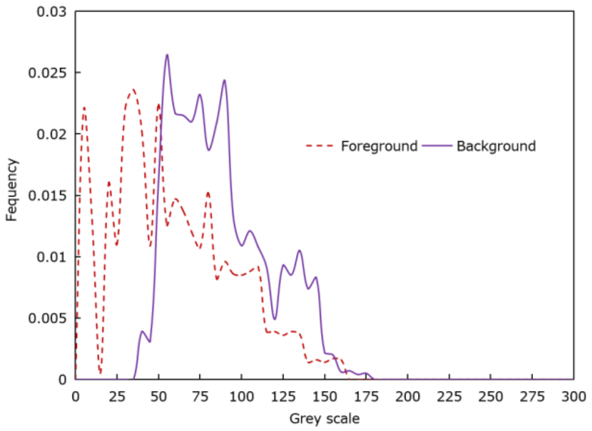

Figure 3: Gray distribution of foreground area and background area.

Over-fitting occurs because the trained neural network is relatively large, and there is not enough training data, or because the trained model is complex and the data interference is great, the importance of its accuracy is reduced. Once over-fitting occurs, its model will become more complicated with the training. In order to avoid this phenomenon, what needs to be done in the study is to increase the number of training samples.

3.2 CAID Image Enhancement Processing

CNN is usually over-parameterized. Although it is beneficial for the network to converge to the local optimal value of the loss function during training, in order to make the network model be used on devices that do not have strong computing power, such as mobile devices, redundant parameters need to be removed from the trained network, and the accuracy of the network cannot be affected. The convolution layer or weight of the network is compressed by the compression algorithm, and the forward propagation speed of CNN is accelerated by the new structural design, and the phenomena of over-fitting and gradient dispersion are suppressed. The principle of CAID image enhancement based on CNN is shown in Figure 4.

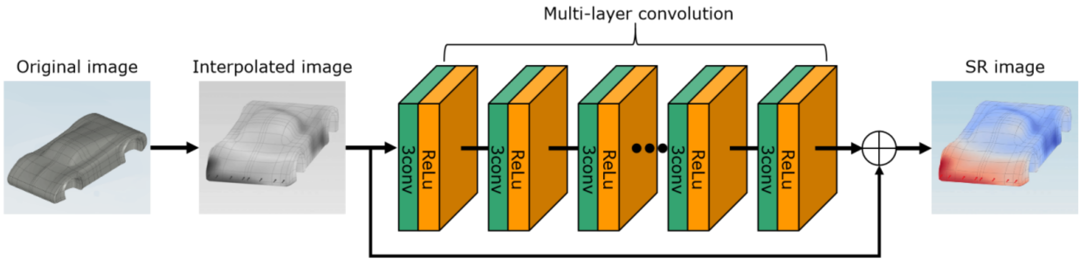

Figure 4: CAID image enhancement based on CNN.

The variable features of the target can be regarded as the new features of the target, and the cascade classifier generated by training at the current moment cannot detect and identify the new features of the target that will appear at the next moment. If the new features of the target at the next moment can be obtained in advance, then the training cascade classifier will be able to detect the target with changing features.

Given the parameter weight W and the deviation b, the probability that the transaction x belongs to the category i is:

$$P(Y=i|x,W,b)=softmax_i(Wx+b)=\frac{e^{W_ix+b_i}}{\sum_j e^{W_jx+b_j}} \tag{7}$$

Negative logarithmic likelihood function is used as the loss function in supervised fine tuning. Assuming that y^i is the true classification of input x^i, the loss function has the following form:

$$L(W,b,x)=-\sum_{i=1}^{N}\log\left(P\left(Y=y^i|x^i,W,b\right)\right) \tag{8}$$

For the two difference images obtained, set an appropriate threshold T, and a binary image $b_{(k-1,k)}(x,y), b_{(k,k+1)}(x,y)$ can be obtained, where the value is 0 when it is less than the threshold, and 1 otherwise. For each pixel (x, y), the two binary images are logically AND-ed to obtain a binary image $B_{(k-1,k)}(x, y)$:

$$B_{(k-1,k)}(x, y) = b_{(k-1,k)}(x, y) \otimes b_{(k,k+1)}(x, y) \quad (9)$$

The binary image $B_{(k-1,k)}(x, y)$ is morphologically filtered to further remove the isolated noise in the image and the "hole" in the target.

4 RESULT ANALYSIS AND DISCUSSION

The proposed method is experimentally studied by using PASCAL-VOC database. PASCAL-VOC database is divided into 20 classes, which is a public database widely used in the field of semantic segmentation. The algorithm simulation is based on Tensor Flow platform, and NVIDIA-2080Ti GPU graphics card is used in hardware equipment. The use of GPU is helpful to accelerate the training of neural network. During training, randomly select 150-300 images from each task, and then select another 150 images as inspection images. The number of training rounds is 500, the number of rounds per round is 50, and the number of batches is 20. Both the number of rounds and the number of cycles are super-parameters of the network and must be debugged manually. Batch size has specific requirements for video memory, and its settings are different for different devices. The environmental configuration parameter requirements of the system are listed in Table 1.

Project	*Edition*
Operating system	Windows 11
CPU	Intel(R) Core(TM) I7-13700H
Internal storage	32GB
Hard disc	1TB
GPU	RTX 2080Ti
Memory	11G
Framework	TensorFlow 2.5
Database administration	Navicat for SQLite
Compiler	Python 3.9

Table 1: Requirements for environmental configuration parameters of the system.

Among many low-level features of image processing, the color feature of the image can directly express the visual effect, followed by the shape feature and volume feature of the image. In the image histogram, all the color data are obtained after computer statistics, which clearly describes the color distribution state and quantity in the whole image, and at the same time, the basic tone and color distribution information of the image can be reflected from the histogram, which can provide a good reference for the next image processing.

The convolution layer failed to extract the main digital features of the samples, resulting in the extraction of secondary features, which eventually led to wrong classification. The solution to this problem is to send the sample to a deeper convolution layer and continue to extract features, that is, the convolution layer of this layer needs to continue the convolution operation to extract the main features of the sample. Figure 5 shows the error curve of CNN establishment process.

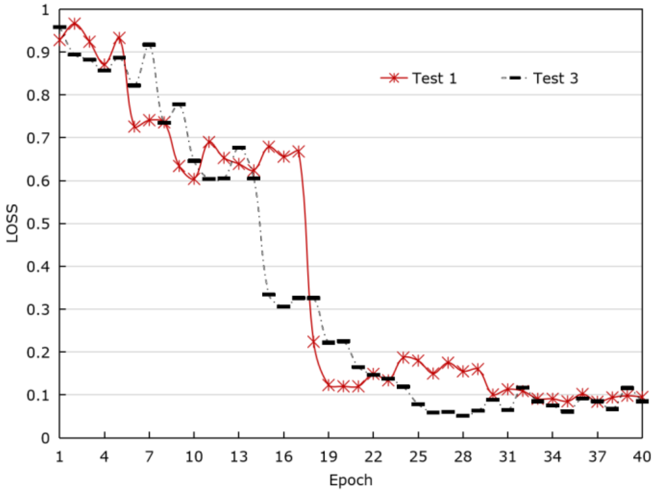

Figure 5: Training error curve.

The error rate obtained by this classifier in the test set is about 0.1. This method meets the accuracy requirements after 20 iterations. In this model, a pooling layer and a convolution layer are added, and a small-step and small-scale classifier is used to eliminate redundant features in the classifier, thus improving the classification accuracy of the classifier. Although the introduction of pooling layer will lose some information to a certain extent, the loss of information of a pool layer with a step size or smaller is limited, and its role in expanding the perception field and reducing over-fitting will make the features learned by the network more representative.

The neurons in the same feature map use the same filter bank, and the filter banks of different feature maps in the same layer are not the same. In array data, such as pictures, the values of local data are highly correlated, and such unique local features are easy to detect the local features in pictures and also the features of other regions, which is the ideological source of weight sharing. The function of pooling is to merge similar features into one, reduce the feature map and keep the main feature information. On this basis, 70% of the samples are training samples and 30% are testing samples. Training the CNN, the results are shown in Figure 6.

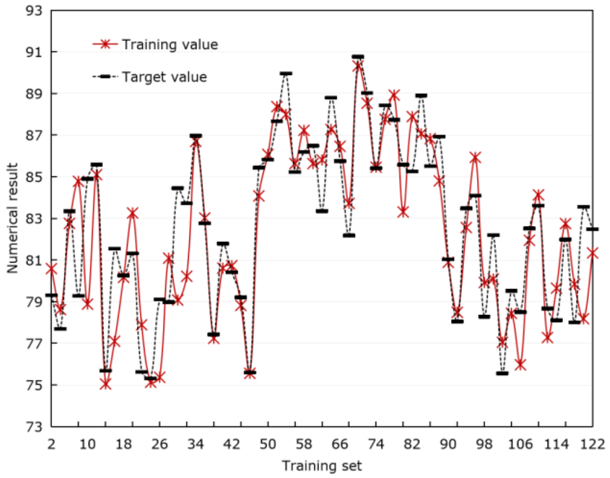

Figure 6: Results after network training.

From the training results shown in Figure 6, it can be seen that the training value is in good agreement with the target value. With the support of CAD technology, through the change of local parameters, the relevant parts of the structure have changed, and the scale has driven the graphics. There are many other kinds of activation functions in CNN, all of which have their own advantages and disadvantages. However, when selecting the activation function, the traditional activation function will not be used. At present, the widely used activation functions have certain applicability. The specific selection varies from person to person, but after the selection, we need to pay attention to the characteristics of the activation function itself, and pay more attention to the setting of relevant parameters in the training network.

It is a central problem for CNN to get the appropriate convolution kernel through training. The trained convolution kernel can be applied to other tasks after fine-tuning, which is due to its universal applicability and makes the well-trained CNN have good performance in other tasks. In Figure 7, the comparison results of image database updating algorithms are shown.

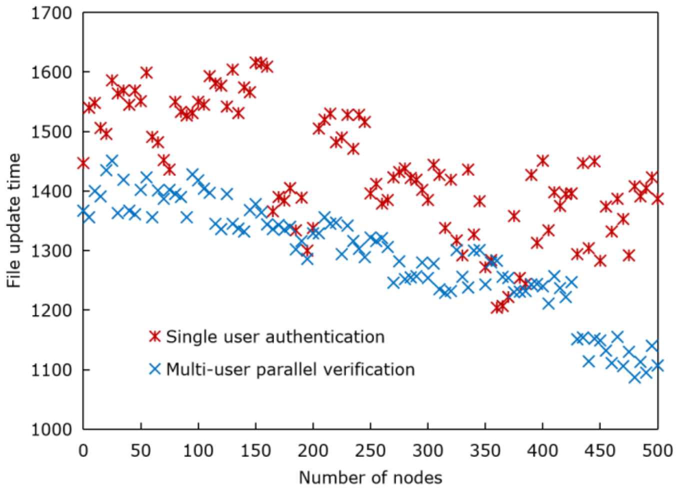

Figure 7: Performance comparison of image database updating algorithm.

Combined with the test results, in the case of expanding network scale, this article adopts parallel verification method, which is more effective than single user verification method. Applying AI to CAID can effectively convey visual culture, bring users a better experience and effectively guarantee the quality of industrial design. For products and users, the use of AI can enhance the design value of products, which not only makes products more scientific and reasonable in terms of shape, function, structure and materials, but also saves unnecessary functions and materials, and can greatly improve the overall beauty and social and cultural functions of products.

5 CONCLUSIONS

The arrival of the Industrial 4.0 era represented by AI has prompted AI to penetrate into many fields of people's lives and work. Efficient computing, fast data transmission, high-capacity storage and mobile network communication all bring convenience to the design of industrial products, but also break through the limitation of computer size. This article mainly discusses the application of AI in CAID, and proposes a CAID method based on computer vision and CNN. Simulation shows that the model constructed in this article can quickly train to achieve the target accuracy of CAID image processing, and can provide support for the feature extraction and modeling stage of CAID. Applying AI to CAID can effectively convey visual culture, bring users a better experience and effectively guarantee the quality of industrial design. The independent characteristics of various defects on the surface of industrial products need to be further studied and analyzed to extract more images of various defects. In order to improve the accuracy and comprehensiveness of defect identification, we can store different types of image feature information when creating defect types and feature databases, so as to make up for the shortcomings.

6 ACKNOWLEDGEMENT

This work was supported by 2023 Research-Based Improvement Project on the Basic Capabilities of Young and Middle-aged Teachers in Guangxi Universities: Visual Transformation and Activation of Museum Cultural Creative Products (No.: 2023KY0918).

Kai Xiao, https://orcid.org/0009-0004-7900-2881
Taile Ni, https://orcid.org/0000-0003-3215-3354

REFERENCES

[1] Cao, W.: Applying image registration algorithm combined with CNN model to video image stitching, The Journal of Supercomputing, 77(12), 2021, 13879-13896. https://doi.org/10.1007/s11227-021-03840-2
[2] Dhaya, R.: Light weight CNN based robust image watermarking scheme for security, Journal of information technology and digital world, 3(2), 2021, 118-132. https://doi.org/10.36548/jitdw.2021.2.005
[3] Dhiman, G.; Juneja, S.; Viriyasitavat, W.; Mohafez, H.; Hadizadeh, M.; Islam, M.-A.; Gulati, K.: A novel machine-learning-based hybrid CNN model for tumor identification in medical image processing, Sustainability, 14(3), 2022, 1447. https://doi.org/10.3390/su14031447
[4] Gonthina, M.; Chamata, R.; Duppalapudi, J.; Lute, V.: Deep CNN-based concrete cracks identification and quantification using image processing techniques, Asian Journal of Civil Engineering, 24(3), 2023, 727-740. https://doi.org/10.1007/s42107-022-00526-9
[5] Hou, Y.; Li, Q.; Zhang, C.; Lu, G.; Ye, Z.; Chen, Y.; Cao, D.: The state-of-the-art review on applications of intrusive sensing, image processing techniques, and machine learning methods in pavement monitoring and analysis, Engineering, 7(6), 2021, 845-856. https://doi.org/10.1016/j.eng.2020.07.030
[6] Hu, L.; Qin, M.; Zhang, F.; Du, Z.; Liu, R.: RSCNN: a CNN-based method to enhance low-light remote-sensing images, Remote Sensing, 13(1), 2020, 62. https://doi.org/10.3390/rs13010062
[7] Hu, Z.; Fang, W.; Gou, T.; Wu, W.; Hu, J.; Zhou, S.; Mu, Y.: A novel method based on a Mask R-CNN model for processing dPCR images, Analytical Methods, 11(27), 2019, 3410-3418. https://doi.org/10.1039/C9AY01005J
[8] Ilesanmi, A.-E.; Ilesanmi, T.-O.: Methods for image denoising using convolutional neural network: a review, Complex & Intelligent Systems, 7(5), 2021, 2179-2198. https://doi.org/10.1007/s40747-021-00428-4
[9] Jia, C.; Wang, S.; Zhang, X.; Wang, S.; Liu, J.; Pu, S.; Ma, S.: Content-aware convolutional neural network for in-loop filtering in high efficiency video coding, IEEE Transactions on Image Processing, 28(7), 2019, 3343-3356. https://doi.org/10.1109/TIP.2019.2896489
[10] Liu, M.; Zhou, Z.; Shang, P.; Xu, D.: Fuzzified image enhancement for deep learning in iris recognition, IEEE Transactions on Fuzzy Systems, 28(1), 2019, 92-99. https://doi.org/10.1109/TFUZZ.2019.2912576
[11] Pu, Y.; Apel, D.-B.; Szmigiel, A.; Chen, J.: Image recognition of coal and coal gangue using a convolutional neural network and transfer learning, Energies, 12(9), 2019, 1735. https://doi.org/10.3390/en12091735
[12] Sharma, P.; Berwal, Y.-P.-S.; Ghai, W.: Performance analysis of deep learning CNN models for disease detection in plants using image segmentation, Information Processing in Agriculture, 7(4), 2020, 566-574. https://doi.org/10.1016/j.inpa.2019.11.001

[13] Srivastava, S.; Divekar, A.-V.; Anilkumar, C.; Naik, I.; Kulkarni, V.; Pattabiraman, V.: Comparative analysis of deep learning image detection algorithms, Journal of Big data, 8(1), 2021, 1-27. https://doi.org/10.1186/s40537-021-00434-w

[14] Zeng, H.; Cai, J.; Li, L.; Cao, Z.; Zhang, L.: Learning image-adaptive 3d lookup tables for high performance photo enhancement in real-time, IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(4), 2020, 2058-2073. https://doi.org/10.1109/TPAMI.2020.3026740

[15] Zhao, Z.; Xiong, B.; Wang, L.; Ou, Q.; Yu, L.; Kuang, F.: RetinexDIP: A unified deep framework for low-light image enhancement, IEEE Transactions on Circuits and Systems for Video Technology, 32(3), 2021, 1076-1088. https://doi.org/10.1109/TCSVT.2021.3073371