Research Article

Yutong Sun*

High-resolution image processing and entity recognition algorithm based on artificial intelligence

Abstract

Objective – With the popularity of high-resolution devices such as high-definition, ultra-high-definition televisions, and smartphones, the demand for high-resolution images is also increasing, which puts forward higher requirements for high-resolution image processing and entity recognition technology.

Method – This article introduced the research progress and application of high-resolution image processing and entity recognition algorithms from the perspective of artificial intelligence (AI). First, the important role of AI in high-resolution image processing and entity recognition was introduced, and then the applications of deep learning-based algorithms in high-resolution image grayscale equalization, denoising, and deblurring were introduced. Subsequently, the application of AI-based object detection and image segmentation algorithms in entity recognition was explored, and the superiority of AI-based high-resolution image processing and entity recognition algorithms was verified through training and testing. The accuracy of the model was verified through testing experiments. Finally, a summary and outlook were made on high-resolution image processing and entity recognition algorithms based on AI.

Result – After experimental testing, it was found that high-resolution image processing and entity recognition based on AI had higher efficiency, and the overall image recognition ability was improved by 29.6% compared to traditional image recognition models. The recognition speed and accuracy were also improved.

Conclusion – High-resolution image processing and element recognition algorithms based on AI enabled observers to see the detailed information in the image more clearly, thus improving the efficiency and accuracy of image analysis. Through continuous improvement of algorithm performance, real-time application, and expansion of cross-disciplinary applications, people can look forward to the development of more advanced and powerful image processing and entity recognition technologies, which will bring huge impetus to research and application in various fields.

Keywords: artificial intelligence, deep learning, convolutional neural network, image processing, entity recognition algorithms

1 Introduction

With the widespread application and popularization of digital image acquisition devices, the image data generated in people's daily lives have shown explosive growth. For example, image sharing on social media, the popularity of smartphones, and the widespread use of surveillance cameras have all provided a large amount of data sources for the development of image recognition. Second, the rapid development of artificial

* **Corresponding author: Yutong Sun,** School of Software Engineering, Jilin Technology College of Electronic Information, Jilin 132000, Jilin, China, e-mail: 17644228211@163.com

intelligence (AI) technology has played a crucial role in promoting image recognition. Specifically convolutional neural network (CNN) has made significant breakthroughs in the field of image recognition. In machine learning, pattern recognition, and image processing, feature extraction starts from the initial measurement data set and establishes derivative values (features) designed to provide information and non-redundancy, thereby promoting subsequent learning and generalization steps, and it can bring better interpretability in some cases. Through deep learning methods, features can be automatically learned and extracted from raw pixel data, achieving more accurate and efficient image recognition. However, with the development of AI, people have put forward higher requirements for image processing and recognition to achieve more efficient, accurate, and intelligent decision-making and applications. Multiple factors provide broad space and opportunities for the application and development of image recognition. With the further advancement and innovation of AI, image recognition can have a broader and far-reaching impact in various fields.

AI, as a hot research direction at present, has been integrated into people's daily lives in various ways, thus greatly improving people's living standards. Mintz and Brodie believed that radiological imaging, pathological sections, and patient electronic medical records were being evaluated through machine learning, which can aid in the diagnosis and treatment process of patients and enhance the capabilities of doctors. They also described the current situation of AI in medicine, its application in different disciplines, and future trends [1]. He et al. discussed some key practical issues of implementing AI in the existing clinical workflow, including data sharing and privacy, algorithm transparency, data standardization and interoperability across multiple platforms, as well as concerns about patient safety [2]. Challen et al. believed that it was necessary to assist clinical safety professionals in critically evaluating current medical AI research from a quality and safety perspective and to support the development of AI by emphasizing some clinical safety issues [3]. Buch et al. believed that machine learning could make AI-driven applications superior to dermatology in correctly classifying suspicious skin lesions [4]. Schwalbe and Wahl believed that AI-driven health interventions might improve health outcomes in low- and middle-income countries [5]. Ouyang et al. believed that AI brought new methods to improve teaching and learning in online higher education [6]. To overcome the difficulties associated with the automatic detection of oral illnesses, Mira et al. described an approach based on smartphone image diagnosis powered by a deep learning algorithm. The centered rule method of image capture was offered as a quick and easy way to get high-quality pictures of the mouth [7]. These scholars generally lean toward the medical field in their research and application of AI, with less involvement in image recognition.

High-quality images can provide more details and a better sensory experience, while image processing based on AI can significantly improve image quality and improve image details. Moen et al. believed that the latest advances in machine learning supported a series of algorithms with strong image recognition capabilities, thus providing great assistance for researchers' research [8]. Mohan and Poobal improved the crack detection level on concrete surfaces by analyzing and researching image processing techniques, targets, accuracy levels, error levels, and image datasets [9]. Minaee et al. studied the relationships, advantages, and challenges of deep learning-based image segmentation models and examined widely used datasets. They compared the performance and discussed promising research directions [10]. As a type of image recognition, scholar Phillips et al. believed that collaboration between humans and machines provided practical benefits for the accuracy of facial recognition in important applications. Only when humans and machines work together can optimal facial recognition be achieved [11]. Maier et al. believed that the development of medical image processing and deep learning was inseparable and pointed out the problems and challenges it would face in the future [12]. Integrating machine learning technologies into AI is at the forefront of the scientific and technological tools employed to combat the COVID-19 pandemic. Almotairi assessed different uses and deployments of modern technology for combating the COVID-19 pandemic at various levels, such as image processing, tracking of disease, prediction of outcomes, and computational medicine. The results prove that computerized tomography scans help to diagnose patients infected by COVID-19 [13]. According to Sahins study, fracture detection and classification are performed using various machine learning techniques using a dataset containing various bones (normal and fractured). Classifier results are presented comparatively as accuracy, training time, and testing time, and linear discriminant analysis (LDA) reaches the highest accuracy rate with 88.67% and 0.89 AUC. The proposed computer-aided diagnosis system will reduce the burden on

physicians by identifying fractures with high accuracy [14]. The above research indicates that image recognition based on AI is widely applied in various fields, but there are still issues with incomplete technology.

In order to improve the application of AI in image recognition, this article first discussed high-resolution image processing based on AI and introduced the important aspects of image preprocessing, as well as image segmentation and feature extraction. Second, the algorithm for entity recognition based on AI was discussed, and the application of CNN in image recognition and related algorithms was introduced. The application of image recognition was discussed, and high-resolution image processing and entity recognition algorithms based on AI were summarized and prospected.

2 High-resolution image processing based on AI

2.1 Image preprocessing

Due to factors such as environmental conditions, shooting equipment, and image resolution, the captured image may have some problems that interfere with the normal display of the image. Therefore, it is necessary to preprocess the image, such as grayscale equalization, denoising, and deblurring.

Grayscale equalization is an image processing technique used to enhance the contrast and brightness of an image. Before grayscale equalization, people need to first understand the image histogram. The image histogram represents the distribution of pixel grayscale levels in the image. If the histogram of the image is evenly distributed, it means that the contrast and brightness of the image are relatively balanced. However, if the histogram of the image is not evenly distributed, it may cause some areas to be too bright or too dark, affecting the visual effect of the image. Grayscale equalization enhances the contrast and brightness of the image by redistributing the pixels of the image, resulting in a uniform distribution of the image's histogram. The common grayscale equalization methods are shown in Table 1.

Table 1: Common grayscale equalization methods

Name	Method	Advantage
Global grayscale balance	Global grayscale equalization refers to reallocating the grayscale levels of the entire image, so that the probability of each grayscale level appearing in the image is equal.	Can enhance the contrast of images
Local grayscale equalization	Local grayscale equalization refers to dividing an image into several small regions and performing grayscale equalization on each small region.	Avoiding excessive brightness enhancement caused by global grayscale equalization
Adaptive histogram equalization	AHE is a method based on local grayscale equalization. It divides the image into several small regions, balances the gray level of each small region, and uses bilinear interpolation to combine the results.	It can be balanced based on local features of the image

Grayscale equalization is a simple and effective image enhancement technique that can enhance the contrast and brightness of images. In practical applications, it is necessary to select appropriate grayscale equalization methods based on the characteristics of the image.

Image denoising mainly aims to reduce the noise in the image as much as possible, preserve the details and texture of the image, and minimize image distortion and blur as much as possible [15]. Traditional image-denoising methods are mainly based on mathematical models, such as wavelet transform, total variation denoising, and so on. However, these methods often require precise parameter settings and are sensitive to factors such as noise type and signal-to-noise ratio, thus making it difficult to achieve ideal results. The initial learning rate setting range for network hyperparameters is as follows (Table 2).

Table 2: Setting range of initial learning rate for network hyperparameters

SGD	$[1 \times 10^{-2}, 1 \times 10^{-1}]$
Momentum	$[1 \times 10^{-3}, 1 \times 10^{-2}]$
Adagrad	$[1 \times 10^{-3}, 1 \times 10^{-2}]$
Adadelta	$[1 \times 10^{-2}, 1 \times 10^{-1}]$
RMSprop	$[1 \times 10^{-3}, 1 \times 10^{-2}]$
Adam	$[1 \times 10^{-3}, 1 \times 10^{-2}]$
Adamax	$[1 \times 10^{-3}, 1 \times 10^{-2}]$
Nadam	$[1 \times 10^{-3}, 1 \times 10^{-2}]$

These ranges usually refer to situations where training starts from scratch. If fine-tuned, the initial learning rate can be reduced by one to two orders of magnitude.

Image-denoising algorithms based on AI, especially algorithms based on deep learning, are more suitable for different types of noise and have better adaptability and robustness. The deep learning-based image-denoising algorithm is a method that utilizes deep learning models to process image noise [16]. The basic idea of this algorithm is to train a deep CNN to learn the features of noise and texture in images, so as to accurately remove noise. Compared with traditional rule-based or mathematical model-based methods, deep learning algorithms can more accurately handle various types of image noise without the need to manually select appropriate filters or set parameters. Image smoothing is an image-denoising algorithm based on deep learning, which aims to reduce the impact of noise in the image, making it clearer and easier for subsequent processing. Image smoothing is mainly achieved through filters. Common filters include mean filter, median filter, Gaussian filter, etc. The effect of image smoothing on image denoising is shown in Figure 1. Among them, a represents the effect before denoising, b and c represent the effect during the denoising process, and d represents the effect after denoising is completed. In comparison, image clarity has significantly improved.
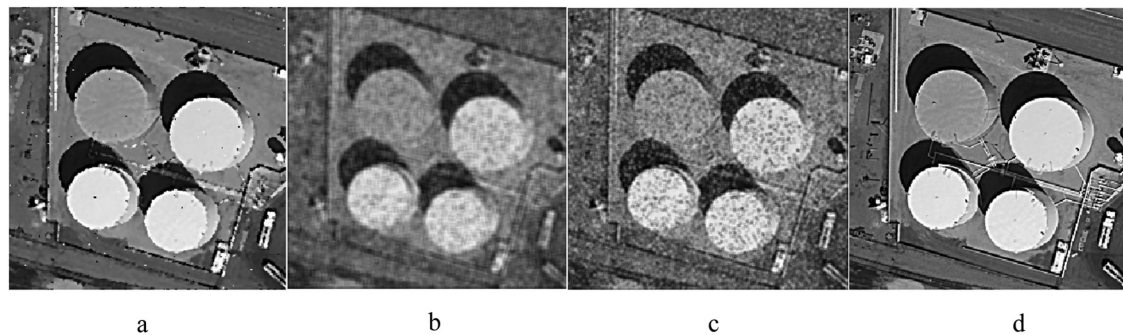

| a | b | c | d |

Figure 1: Image-denoising effect display diagram. (a) Display of the effect before image denoising. (b) Display of the effect during image denoising process. (c) Display of the effect during image denoising process. (d) Display of the effect after image denoising is completed.

Image blur is often caused by image motion or camera shake, which can have a significant impact on image quality and clarity. Therefore, image deblurring algorithms have important application value in image processing. In contrast, image deblurring is more difficult. Ambiguity is a type of information loss, and its types and degrees are diverse, so it is difficult to describe it with a simple mathematical model. The traditional image deblurring algorithm is mainly based on blind deconvolution, which estimates the fuzzy kernel to restore the original image. However, this method often requires accurate estimation of fuzzy kernels, and different models need to be designed for different types of fuzzy scenes, making it difficult to generalize. The image deblurring algorithm based on deep learning can automatically learn the fuzzy kernel and generate clear images [17]. In addition, deep learning algorithms can also combine various advanced technologies, such as super-resolution and multi-scale strategies, to further improve the effectiveness and robustness of image deblurring. The image deblurring effect based on deep learning is shown in Figure 2. Among them, a
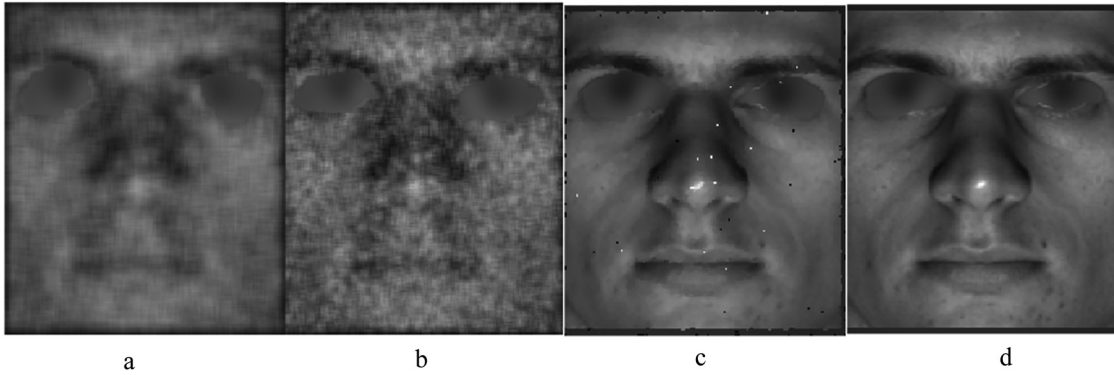
Figure 2: Display of image deblurring effect. (a) The effect of image before deblurring. (b) The effect of image deblurring process. (c) The effect of image deblurring process. (d) The effect of image deblurring after completion.

represents the effect before deblurring, b and c represent the effect during the deblurring process, and d represents the effect after deblurring is completed. It can be clearly seen that the blurred image becomes more detailed.

2.2 Image segmentation

High-resolution image segmentation refers to classifying high-resolution images at the pixel level and dividing them into several independent objects or regions that correspond to the real world. In recent years, with the development of deep-learning technology, image segmentation models based on deep learning have gradually become mainstream, which can effectively solve the problem of high-resolution image segmentation. DeepLab is a deep-learning image segmentation model based on hollow convolution. The model expands the receptive field through cavity convolution, which can effectively capture the context information in the image. DeepLab also introduces the idea of multi-scale information fusion, which can segment images at different scales and improve the accuracy of segmentation. The high-resolution image segmentation model based on deep learning has good segmentation performance and broad application prospects.

Common image segmentation algorithms include the threshold method, color clustering method, edge detection method, and region growth method. The threshold method is a basic and intuitive image segmentation method. It divides image pixels into different categories by setting a threshold, so that pixels with similar pixel values are grouped together. The calculation process of the threshold method is relatively simple and does not require complex mathematical operations or iterative processes, so it has high computational efficiency. The threshold method is suitable for application scenarios that require high real-time performance.

Although the threshold method performs well in terms of effectiveness, its adaptability is weak when dealing with environments with changes in lighting and complex backgrounds. The color clustering method can overcome the impact of environmental changes on segmentation results. However, compared to the threshold method, its calculation time is longer. Edge detection methods are relatively simple, have short computational time, and have good effectiveness. However, most algorithms are difficult to extract the complete closed edges of the target, which can easily lead to false edges. The principle of the region growth method is simple and intuitive, but it takes a long time to segment the entire image and is a recursive algorithm. In summary, different image segmentation algorithms have their own advantages and limitations. In practical applications, it is necessary to select appropriate algorithms or combine multiple algorithms based on specific needs and scenarios for image segmentation to obtain more accurate and effective segmentation results. Image segmentation is a very complex process that requires comprehensive consideration of multiple factors such as image features, algorithm performance, and segmentation results. Different segmentation methods have their own advantages and disadvantages. In specific applications, it is necessary to choose appropriate methods based on the actual situation.

2.3 Image feature extraction

Image feature extraction is an important task in the field of computer vision, with the goal of extracting representative and discriminative features from images for subsequent image analysis and recognition tasks. Image features are descriptions of image content, which can be local or global features. Common image feature representations include color histograms, texture features, edge features, shape features, etc. These features can be extracted and represented through mathematical models and algorithms. Image feature extraction methods based on deep learning automatically learn feature representations through neural networks, such as CNN, recurrent neural networks, etc. These methods can learn more advanced and abstract feature representations from the original image and have strong expressive ability and adaptability. Image feature extraction also needs to consider feature selection and dimension reduction. In large-scale image data, feature dimensions are usually very high, which may lead to computational complexity and redundancy. Therefore, feature selection and dimensionality reduction are commonly used techniques for selecting the most representative subset of features or reducing feature dimensions. Common methods include principal component analysis, LDA, and sparse encoding. Finally, in image recognition and retrieval tasks, feature matching and similarity measurement are required to measure the similarity or distance between features.

Image feature extraction plays a crucial role in many applications of computer vision, such as image classification, object detection, face recognition, and image retrieval. By accurately, robustly, and efficiently extracting image features, the performance of image analysis and recognition systems can be greatly improved.

3 AI-based entity recognition algorithm

3.1 Convolutional neural network

CNN, as a deep learning model, has achieved significant results in the field of image processing, such as image classification, object detection, image segmentation, and image generation. By automatically learning features in images and possessing the characteristics of translation invariance and local connectivity, they can process large-scale and complex image data and significantly improve image recognition efficiency [18,19].

The main structures of CNN include convolutional layers, pooling layers, nonlinear activation layers, and fully connected (FC) layers. Compared with multilayer perceptron, CNN has better cost performance and performance, especially in processing large-scale image data. The convolutional layer and pooling layer are the core components of CNN. The pooling layer reduces the computational burden by reducing the number of connections between convolutional layers and reduces the excessive sensitivity of convolutional layers to location. These design features enable CNN to ensure the invariance of input image pixels in displacement, scaling, and distortion to a certain extent [20,21].

The relevant algorithms for the basic model of CNN are as follows:

$$g(t) = \int x(a)z(t - a)\mathrm{d}a. \tag{1}$$

Among them, $x(a)$ and $z(a)$ represent convolution operations:

$$g(t) = (x \times z)(t). \tag{2}$$

Among them, the x function represents the first input parameter in the convolution operation, and the z function represents the second parameter. The previous formula stated that the convolution operation is on a continuous interval. If the convolution function is on a discrete interval, the convolution operation formula is as follows:

$$g(t) = (x \times z)(t) = \sum_{-\infty}^{+\infty} x(a)z(t - a). \tag{3}$$

If the input and convolution kernel are both two-dimensional functions, the formula is as follows:

$$g(m, n) = (X, Z)(m, n) = \sum_{m}\sum_{m} X(i, j)Z(m - i, n - j). \tag{4}$$

3.1.1 Convolutional layer

The convolutional layer of CNN can be divided into lower and higher layers. The convolutional layer contains multiple convolutional kernels. Each consists of a learnable weight matrix. This is usually an odd-size matrix. The convolutional layer performs convolution operations on the weight matrix of the input feature map and the convolutional kernel. The convolutional kernel performs convolution operations on the feature map through sliding windows to extract feature information from different positions. This operation captures the local patterns and structures of the image and achieves higher-level feature extraction through the combination of multiple convolutional layers. When processing feature maps of the same batch, the convolutional kernel parameters are fixed, and the shared parameter mechanism makes the operation simple and efficient. On large data sets, training is carried out to reduce the number of parameters and reduce the risk of overfitting. Shared parameters enable CNN to extract features from different regions of the image using the same convolutional kernel, capture local patterns and structures of the image, and improve algorithm efficiency and generalization ability.

3.1.2 Pooling layer

The pooling layer is usually located after the convolutional layer. The main purpose of using pooling layers is to perform downsampling and dimensionality reduction processing. The scale invariance, translation invariance, and rotation invariance of the input image are implemented. The robustness of the output feature map to individual neurons is enhanced, and the sensitivity to distortion and error is reduced. By pooling operations, important information can be preserved while reducing the spatial dimensions of the feature map, improving the computational efficiency of the network, and making it more robust to changes in input. The pooling layer usually adopts the method of maximum pooling or average pooling and obtains a more compact feature representation by statistically summarizing the input regions. This can effectively reduce the number of parameters, reduce the risk of overfitting, and improve the generalization ability of the model. The use of pooling layers has enabled CNN to achieve significant results in image processing and recognition tasks and has become an important component in the field of computer vision. Pooling layers can reduce the size of feature maps, thereby reducing computational complexity. Of course, convolutional layers with a stride greater than 1 can also reduce the size of feature maps, and the pooling layer can bring invariance to features such as translation and rotation.

3.1.3 Nonlinear activation layer

The linear activation function (such as the linear rectification function ReLU) is often used in deep neural networks. It has a simple computational form and efficient derivative properties, making the training of the network more stable and fast. However, the linear activation function has a disadvantage that it cannot capture the nonlinear relationship of input data, which limits the expression ability of the network. To overcome this problem, various nonlinear activation functions are introduced, such as Sigmoid, Tanh, and Leaky ReLU. These functions have nonlinear shapes and can better model complex data distributions and feature representations. As shown in Figure 3, the Sigmoid function maps the input to a continuous value between 0 and 1, while the Tanh function maps the input to a continuous value between −1 and 1. Leaky ReLU introduces a small slope in the negative value region, solving the problem of dead neurons in the ReLU function. The introduction of nonlinear activation layers enables CNN to learn more complex and abstract feature representations. They enhance the expressive power of the network and improve the fitting and generalization abilities of the model. At the same time, the selection of nonlinear activation function also needs to be adjusted according to specific tasks and network architecture to obtain the best performance and effect.
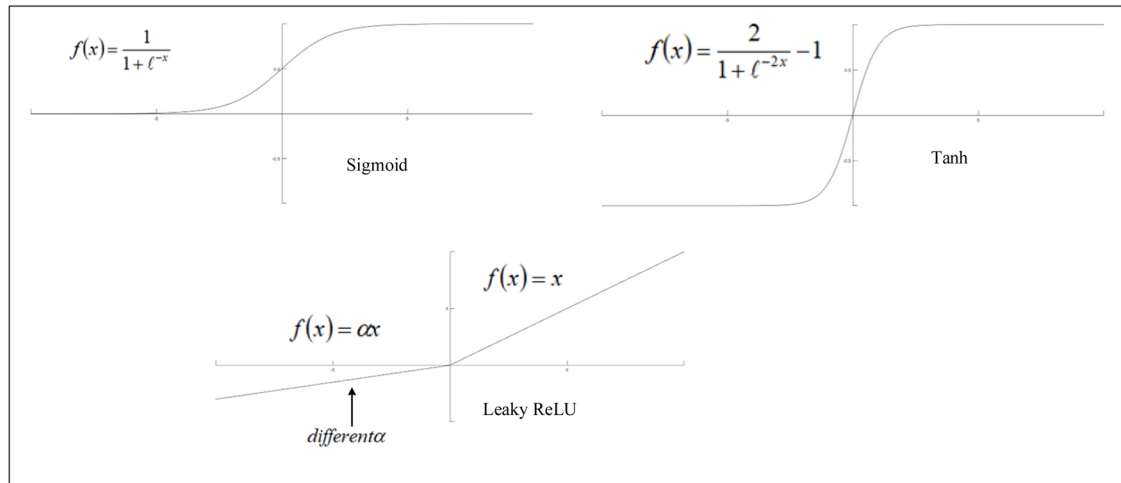
Figure 3: Various nonlinear activation functions.

3.1.4 FC layer

FC layer is a common type of layer in CNN, and its function is to transform the feature maps extracted from the convolutional layer into the final output result. It flattens the features of the previous layer into vectors through matrix multiplication and nonlinear transformation, multiplies them with the weight of each neuron, and calculates them with the offset term. The FC layer is used for the final classification operation in image recognition tasks, thus mapping the extracted features with category labels and outputting the final classification results. The FC layer can integrate and combine global information on extracted features, thus achieving higher levels of semantic understanding and classification capabilities. However, due to the large number of parameters, it is easy to lead to overfitting of the model, so attention should be paid to regularization and model structure design. In summary, the FC layer plays an important role in CNN, thus achieving the transformation and classification tasks from features to output results. It can discard some features that cannot help us predict correctly. All features can be retained, but the size of the parameters can be reduced. High terms can lead to overfitting, so if we can make the precision of high terms close to 0, we can fit well.

3.2 Algorithm for entity recognition based on CNN

Element recognition aims to detect and recognize various types of elements (such as text, edges, lines, corners, etc.) from the input image and convert them into computer-processable forms for subsequent image analysis and processing. The method based on AI learns a large amount of image data and compares the image results calculated by the network with the standard image results. The cost function is to minimize this error, and the cost function is as follows:

$$N(\theta) = \frac{1}{i}\sum_{m}^{i}\left(\frac{1}{2}\ \|h_{z,d}(x^{(m)}) - y^{(m)}\|^2\right). \tag{5}$$

By using the expected residual value and the estimated residual value calculated through convolutional networks to obtain its mean square error, the cost function of the entire network is obtained:

$$L(\Theta) = \frac{1}{2j}\sum_{1}^{j}\|R(y_m;\Theta) - (y_m - x_m)\|_F^2. \tag{6}$$

The input of each dimension is normalized, and the mean and variance are calculated in each mini-batch to replace the mean and variance of the entire training set. The input vector is manipulated. By translating and

scaling changes, the distribution of x becomes a standardized standard distribution within a fixed interval range. The change framework is as follows:

$$h = f\left(s \cdot \frac{x - \mu}{\sigma} + d\right). \tag{7}$$

Among them, μ is the translation parameter and σ is the scaling factor. Through these two parameters, scaling and translation transformations are carried out:

$$\hat{x} = \frac{x - \mu}{\sigma}. \tag{8}$$

The obtained data satisfy the standard distribution with a mean of 0 and a variance of 1. In formula (7), d represents the retranslation parameter, and s represents the rescale parameter. By introducing formula (8) into formula (9), the formula is obtained as follows:

$$y = s \cdot \hat{x} + d. \tag{9}$$

Part of the overall data is trained, and then the mean and variance of x_m neurons are calculated:

$$\mu_m = \frac{1}{i}\sum x_m, \tag{10}$$

$$\sigma_m = \sqrt{\frac{1}{i}\sum (x_m - \mu_m)^2 + \varepsilon}. \tag{11}$$

Among them, m represents the size of the training section, and ε represents a positive value infinitely close to 0.

4 Application of high-resolution image processing and entity recognition algorithms based on AI

In practical applications, high-resolution image processing and entity recognition algorithms based on AI have broad application prospects. Medical imaging analysis, such as medical imaging, contains a large amount of detailed information, which is of great significance for diagnosis and treatment. High-resolution image processing methods based on AI can perform denoising, enhancement, and super-resolution processing on medical images, improving image quality and diagnostic accuracy. At the same time, AI based entity recognition algorithms can also recognize various structural elements in medical images, thereby more accurately locating lesions and diagnosing diseases.

The application of high-resolution image processing and entity recognition algorithms based on AI needs to consider the performance of image recognition models. After testing, it can be found that image recognition models based on CNN have significant advantages in terms of speed and convergence. As shown in Figure 4, the model was tested 1,000 times. The x-axis represented the number of test training, and the y-axis represented the loss value of the loss function. It could be observed that as the number of training increases, the loss value continued to decrease, indicating that the training speed of image recognition models based on CNN was very fast.

Subsequently, recognition tests were conducted on the model to assess its denoising ability, deblurring ability, grayscale balance ability, image segmentation ability, image feature extraction ability, and overall image recognition ability.

As shown in Figure 5, six tests were conducted on the AI-based image recognition model to compare it with traditional image recognition models. It was found that the denoising ability was improved by 32.91%; the deblurring ability was improved by 28.83%; the grayscale balance ability was improved by 26.05%; the image segmentation ability was improved by 29.69%; the image feature extraction ability was 30.52%; the overall
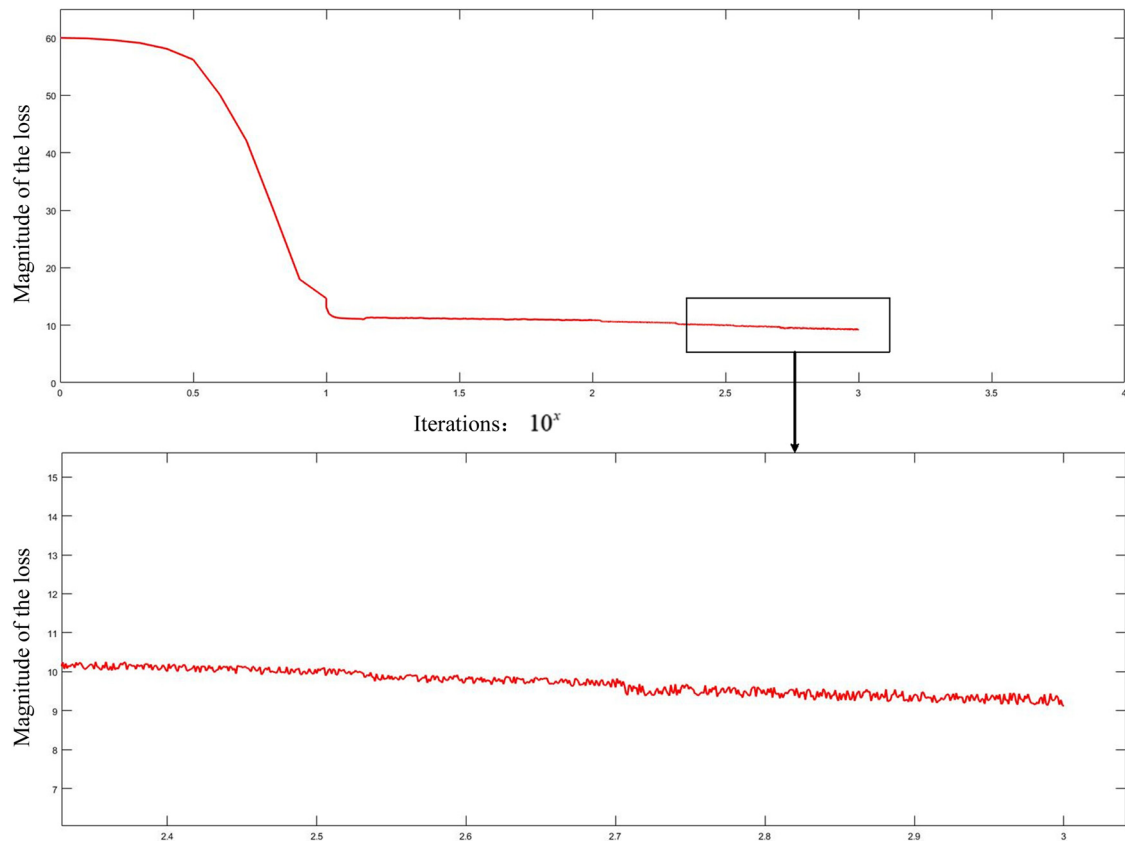
Figure 4: Loss function curve of image recognition model based on AI.
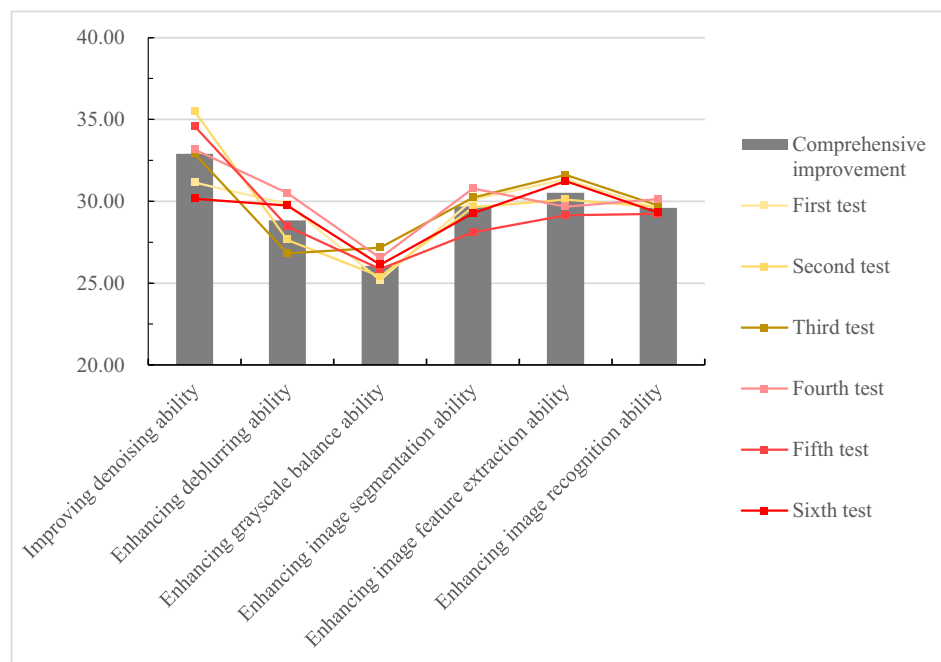

Figure 5: Improvement of image recognition model capability by AI.

recognition ability was improved by 29.6%. From Figure 5, it can be seen that AI has significantly improved the model's image recognition ability, including noise reduction ability, deburring ability, grayscale balance ability, image segmentation, image feature extraction ability, and overall recognition ability.

It could be seen that high-resolution image processing and entity recognition based on AI were more efficient.

Finally, the recognition accuracy of the model was tested, and four types of images, including portraits, buildings, cars, and plants, were selected for testing and divided into two groups: distant and close range. Each type of two groups was tested 200 times, and the recognition accuracy of 50 tests, 100 tests, 150 tests, and 200 tests were recorded, as shown in Table 3. Among them, the recognition accuracy of cars and plants in the vision group was the highest; in the close-up group, the recognition accuracy of portraits and cars was higher, while the success rate of plants and buildings was lower. Overall, the recognition rate for vehicles with more obvious features was higher, while the success rate for plant recognition with less obvious features was lower. The overall recognition accuracy of the model met the standard. For plant image recognition, the model's ability to extract plant image features could be improved through training, and the recognition accuracy could be improved by increasing the number of samples.

Table 3: Accuracy of image recognition model items

		Accuracy rate of 50 tests (%)	Accuracy rate of 100 tests (%)	Accuracy rate of 150 tests (%)	Accuracy rate of 200 tests (%)
Prospect	Portrait	94	95	96.7	97.5
	Build	92	94	96.7	97
	Vehicle	98	98	98	98.5
	Plant	90	93	95.3	96
Close shot	Portrait	98	99	99.3	99.5
	Build	96	98	98	98
	Vehicle	100	100	99.3	99.5
	Plant	98	99	98.6	98

5 Outlook and conclusions

With the continuous progress of AI and deep-learning technology, it can be foreseen that it has broad development prospects in the future. First, in terms of algorithms, with the continuous optimization of neural network structures and the improvement of deep learning models, the performance of high-resolution image processing and entity recognition algorithms would continue to improve. By introducing more complex network structures, more effective feature extraction methods, and more accurate classification algorithms, people can expect more accurate and fast image processing and entity recognition results. Next is the promotion of real-time applications. With the advancement of hardware technology and the improvement of computing power, high-resolution image processing, and entity recognition algorithms based on AI would be more suitable for real-time applications. For example, in fields such as autonomous driving, security monitoring, and medical image analysis, real-time image processing and element recognition can provide more timely and accurate results, thus providing strong support for decision-making and applications. Finally, there is the expansion of cross-domain applications. High-resolution image processing and entity recognition algorithms can not only be applied to traditional computer vision fields, but also be extended to other fields. For example, in the fields of agriculture, environmental monitoring, cultural heritage protection, etc., image processing and entity recognition technology can be used in image analysis, object detection, image reconstruction, etc., thus providing strong support for research and application in related fields. In summary, high-resolution image processing and entity recognition algorithms based on AI have broad development prospects. By continuously improving algorithm performance, achieving real-time applications, and expanding cross-domain applications, people can expect the development of more advanced and powerful image processing and entity recognition technologies, which would bring huge impetus to research and applications in various fields.

References

[1] Mintz Y, Brodie. R. Introduction to artificial intelligence in medicine. Minim Invasive Ther Allied Technol. 2019;28(2):73–81.
[2] He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Nat Med. 2019;25(1):30–6.
[3] Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. BMJ Qual Saf. 2019;28(3):231–7.
[4] Buch VH, Ahmed I, Maruthappu M. Artificial intelligence in medicine: current trends and future possibilities. Br J Gen Pract. 2018;68(668):143–4.
[5] Schwalbe N, Wahl B. Artificial intelligence and the future of global health. Lancet. 2020;395(10236):1579–86.
[6] Ouyang F, Zheng L, Jiao P. Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. Educ Inf Technol. 2022;27(6):7893–925.
[7] Mira ES, Sapri AM, Aljehanı RF, Jambı BS, Bashir T, El-Kenawy ES, et al. Early diagnosis of oral cancer using image processing and artificial intelligence. Fusion: Pract Appl. 2024;14(1):293–308.
[8] Moen E, Bannon D, Kudo T, Graf W, Covert M, Van Valen D. Deep learning for cellular image analysis. Nat Methods. 2019;16(12):1233–46.
[9] Mohan A, Poobal S. Crack detection using image processing: A critical review and analysis. Alex Eng J. 2018;57(2):787–98.
[10] Minaee S, Boykov Y, Porikli F, Plaza A, Kehtarnavaz N, Terzopoulos D. Image segmentation using deep learning: A survey. IEEE Trans Pattern Anal Mach Intell. 2021;44(7):3523–42.
[11] Phillips PJ, Yates AN, Hu Y, Hahn CA, Noyes E, Jackson K, et al. Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. Proc Natl Acad Sci. 2018;115(24):6171–6.
[12] Maier A, Syben C, Lasser T, Riess C. A gentle introduction to deep learning in medical image processing. Z Med Phys. 2019;29(2):86–101.
[13] Almotairi KH. Impact of artificial intelligence on COVID-19 pandemic: a survey of image processing, tracking of disease, prediction of outcomes, and computational medicine. Big Data Cognit Comput. 2023;7(1):11.
[14] Sahin ME. Image processing and machine learning-based bone fracture detection and classification using X-ray images. Int J Imaging Syst Technol. 2023;33(3):853–65.
[15] Fan L, Zhang F, Fan H, Zhang C. Brief review of image denoising techniquesVisual Computing for Industry. Biomed Art. 2019;2(1):1–12.
[16] Shi Q, Tang X, Yang T, Liu R, Zhang L. Hyperspectral image denoising using a 3-D attention denoising network. IEEE Trans Geosci Remote Sens. 2021;59(12):10348–63.
[17] Bendjillali RI, Beladgham M, Merit K, Taleb-Ahmed A. Illumination-robust face recognition based on deep convolutional neural networks architectures. Indones J Electr Eng Comput Sci. 2020;18(2):1015–27.
[18] Dhillon A, Verma GK. Convolutional neural network: a review of models, methodologies and applications to object detection. Prog Artif Intell. 2020;9(2):85–112.
[19] Jadhav SB, Udupi VR, Patil SB. Identification of plant diseases using convolutional neural networks. Int J Inf Technol. 2021;13(6):2461–70.
[20] Wang S. Feed-forward neural network optimized by hybridization of PSO and ABC for abnormal brain detection. Int J Imaging Syst Technol. 2015;25(2):153–64.
[21] Wang SH. Polarimetric synthetic aperture radar image segmentation by convolutional neural network using graphical processing units. J Real-Time Image Process. 2018;15:631–42.