WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology

ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 4, Number 3, Year 2024
DOI: https://doi.org/10.62051/ijcsit.v4n3.22
Journal homepage: https://wepub.org/index.php/IJCSIT/index

A Review of Digital Image Processing Techniques and Future Prospects

Chenxi Liu

College of Photonic and Electronic Engineering, Fujian Normal University, Fuzhou, 350007, China
Liuchenxi5842@outlook.com

ABSTRACT

With the rapid development of computer vision and image processing technology, digital image processing has become a hot spot of computer science research. Based on the research background of digital image processing technology, this paper comprehensively combs through the relevant theories, core technologies and industry applications. By analysing the academic papers and technical reports in recent years, the research results and development trends in the field of image processing are summarized. The steps of digital image processing such as preprocessing, enhancement, recovery, segmentation, representation description and recognition are discussed, and the advantages and disadvantages of the existing algorithms and their application effects in image quality improvement and information extraction are evaluated. For example, the performance of Wavelet transforms and deep learning in image denoising is compared, pointing out the advantages of deep learning in dealing with high noise images. The processing results were quantitatively analysed using international standard image libraries and PSNR and SSIM metrics, yielding reliable experimental results. The article also discusses the application of parallel computing in processing large-scale image data, and the improvement of computational efficiency by GPU acceleration and distributed architecture. Combined with new advances in artificial intelligence and machine learning, it looks at future trends in digital image processing, including the application of deep learning in image analysis, 3D reconstruction, and virtual and augmented reality. The article points out that with the support of a large amount of data and high-performance computing, digital image processing technology is expected to promote a new round of technological revolution, especially in the fields of intelligent manufacturing, telemedicine and automated driving, which have broad application prospects and important social significance.

KEYWORDS

Digital image processing; Deep learning; Image denoising; Parallel computing; Image recognition

1. INTRODUCTION

Digital image processing technology is an emerging discipline since the 1960s, which integrates the theories and methods of computer science, mathematics, physics and other disciplines. With the rapid development of computer technology and digital imaging equipment, digital image processing technology has been widely used in many fields, such as medical imaging, remote sensing image analysis, industrial automation detection. Especially in recent years, the artificial intelligence technology represented by deep learning has brought revolutionary changes to the field of image processing, significantly improving the accuracy and efficiency of image analysis.

As early as the beginning of the 20th century, people began to explore the digitisation of image signals. As early as around 1920, images were first transmitted from London to New York via submarine cables. 1950, MIT produced the first computer with a graphic display, the Cyclone I. By 1962, MIT graduate student Ivan Suzelan successfully developed the groundbreaking Drawing Tablet course, which was the first interactive drawing system in history, and from then on, computer and graphic processing technologies were closely linked. By 1964, the Jet Propulsion Laboratory in California, USA processed a large number of lunar photographs returned by the Voyager 7 spacecraft to correct various image distortions in the spacecraft's camera, and achieved remarkable results. This marks the digital image processing technology in the future of space technology plays a huge role. Since then, digital image processing technology has entered a stage of rapid development.

The basic principle of digital image processing is to convert continuous analogue image signals into discrete digital signals, and then carry out all kinds of processing and analysis of digital images in the computer. A digital image can be represented by a two-dimensional matrix of M×N, where each element of the matrix represents a pixel of the image. The value of a pixel reflects the grey scale or colour information of that point. By mathematically transforming and calculating this matrix, image enhancement, restoration, segmentation, feature extraction, compression and other processing can be achieved. Common image processing algorithms include histogram equalisation, Fourier transform, wavelet transform, edge detection operator, morphological operations and so on.

In 1964, Rosenfeld proposed a series of computer image processing algorithms in his book ‘Picture Processing by Computer’, which laid a theoretical foundation for the development of digital image processing. In the same year, the United States, Bell Labs developed the first image processing software package BEIL, to achieve a variety of image enhancement and analysis functions. 1970s, CT, MRI and other medical imaging equipment has greatly promoted the development of medical image processing technology, emergence of a large number of medical image segmentation, alignment, three-dimensional reconstruction and other algorithms.

NASA for the analysis of land resources satellite remote sensing image data, the establishment of the world's first remote sensing image processing system ERTS, creating a precedent for remote sensing image processing. Since then, remote sensing image processing technology has made great progress, in agriculture, forestry, geology, marine, military and other fields play an irreplaceable role. 1972, the International Telecommunication Union (ITU) formulated the world's first still image coding standard CCITT, the use of differential pulse-code modulation (DPCM) technology to achieve image compression, opened up the image compression and coding technology of the new era.

In the 1980s, with the progress of VLSI chip manufacturing technology, digital signal processing chip (DSP) began to be used in image processing systems, greatly improving the real-time performance of image processing. 1988, Mallat proposed a multi-resolution analysis of the wavelet transform theory, multi-scale analysis of the image and feature extraction provides a new tool. In the 1990s, image processing technology began to develop in the direction of intelligence and network, and in 1997, Long et al. proposed the introduction of neural networks into the image segmentation method, which created a precedent of intelligent image segmentation.

In recent years, artificial intelligence technology represented by deep learning has brought disruptive changes to the field of image processing. 2012, Krizhevsky et al. proposed a deep convolutional neural network AlexNet won the ImageNet image classification competition with an error rate of 15.3%, which is nearly 10 percentage points lower than the traditional method, and unveiled the beginning of the glory of deep learning in the field of computer vision. Deep Learning is a brilliant prelude in the field of computer vision. Since then, various deep learning models such as VGGNet, GoogLeNet, ResNet, etc., have been emerging, constantly refreshing the performance records of image classification, target detection, semantic segmentation and other tasks. In terms of medical image analysis, the U-Net network architecture proposed by Ronneberger et al. in 2015 has achieved fine segmentation of cells, organs, and lesion regions, and is regarded as a milestone in medical image segmentation.

Overall, digital image processing has experienced the development from theory to practice and from simple to intelligent. It has greatly expanded human's ability to acquire and perceive visual information, and has profoundly influenced and changed people's production and life style. Looking to the future, digital image processing has many urgent problems to overcome and a broad space for development, such as efficient storage and retrieval of massive image data, image understanding and content generation, three-dimensional scene reconstruction and understanding. With the new generation of artificial intelligence, big data, cloud computing, Internet of Things and other technologies, digital image processing will usher in a more brilliant tomorrow.

2. OVERVIEW OF DIGITAL IMAGE PROCESSING

2.1. Review of Development History

Digital image processing technology has made great progress since the 1960s. In the early days, image processing mainly used analogue circuits, processing speed is slow, low precision, single function. 1970s, with the improvement of computer performance, digital image processing gradually emerged. The landmark event is the United States in 1964, the Jet Propulsion Laboratory (JPL) use of computers to digitise lunar photographs, greatly improving the quality of the image [1]. Into the 80's, digital image processing theory continues to improve, a variety of new algorithms emerge. Such as Canny edge detection operator [2], Otsu image binarisation algorithm [3], wavelet transform compression coding, etc., greatly enriching the means of image processing.

Since the 1990s, with the rapid development of multimedia technology and the Internet, digital image processing has entered a new stage. Massive image data has laid the foundation for the application of deep learning and other artificial intelligence techniques in the image field. A series of milestone results continue to emerge, such as Krizhevsky et al [4] in the ILSVRC-2012 image classification competition using convolutional neural networks (CNN) to reduce the top-5 error rate to 15.3%, creating a new era of deep learning in the field of computer vision. Since then, deeper and more complex CNN networks such as VGGNet [5], GoogLeNet [6], and ResNet [7] have been introduced one after another, and the top-5 error rate has been constantly refreshed, and is now as low as about 2%. In terms of target detection, from the traditional HOG+SVM [8] to two-stage detectors such as RCNN series [9], and then to one-stage detectors YOLO [10], SSD [11], etc., the mAP indexes are constantly improved, and the processing speed reaches real-time. Semantic segmentation from FCN [12], SegNet [13] to DeepLab series [14], its accuracy and efficiency have made an order of magnitude leap.

In addition to excellent performance, deep learning has also greatly broadened the scope of image processing applications. From the initial OCR, face detection to image content understanding, image generation, image super-resolution and a series of advanced visual tasks, all show great application prospects. At present, image processing has been widely penetrated into industrial production, medical diagnosis, human-computer interaction, public security, automatic driving and other fields, resulting in significant social and economic benefits. With the continuous evolution of 5G, Internet of Things, AI chips and other new generation information technology, digital image processing will usher in a broader development space. At the same time, challenging issues such as interpretability, robustness, data efficiency, generalisation ability, etc. need to be solved, which provides a continuous impetus for the innovation of related theories and methods. It is foreseeable that digital image processing technology will play an increasingly important role in promoting the development of artificial intelligence and empowering thousands of industries.

2.2. Fundamentals and Applications

The basic principle of digital image processing is to convert an image into a digital signal and perform a series of processes to improve image quality, extract features or achieve other purposes. The core of the process is to convert a continuous analogue image into a discrete digital matrix by sampling and quantisation, with each matrix element corresponding to a pixel of the image. Common digital images include binary, grey scale and colour images, which use 1-bit, 8-bit and 24-bit binary numbers to represent the pixel values respectively. The digital image processing process usually includes image acquisition, pre-processing, segmentation, feature extraction, classification and recognition steps.

In the pre-processing stage, the commonly used methods are image filtering, histogram equalisation, geometric transformation, etc. For example, median filtering, histogram equalisation, geometric transformation, etc. are used. For example, median filtering can effectively remove pepper noise, Gaussian filtering can smooth the image, and histogram equalisation can enhance image contrast. Image segmentation is to divide the image into non-overlapping regions, and commonly used methods include threshold segmentation, edge detection, region growing, etc. Otsu thresholding method can adaptively determine the binarised threshold, and Canny operator can accurately extract the image edges. Based on the image segmentation results, the colour, texture, shape and other features of the image can be further extracted, and commonly used feature descriptors include colour histogram, local binary pattern (LBP), scale-invariant feature transform (SIFT) and so on.

In the recognition and classification stage, machine learning algorithms play an important role. Support vector machine (SVM), random forest and other traditional machine learning methods have achieved good results in handwriting recognition, face recognition and other tasks. In recent years, deep learning methods, especially convolutional neural networks (CNN), have made significant breakthroughs in the field of image recognition, and the error rate in the ImageNet Image Classification Challenge has been reduced from 28.2% in 2010 to 2.25% in 2017, even exceeding the level of human recognition. CNN extracts local features of the image through the convolutional layer, achieves feature invariance and dimensionality reduction, through the fully connected layer to achieve classification decisions, end-to-end training makes it perform well in complex image recognition tasks.

In addition to recognition and classification, digital image processing is widely used in many fields. In medical image analysis, image segmentation can accurately outline lesions, and texture analysis can quantify the inhomogeneity of lesions to help diagnose and stage diseases. In remote sensing image analysis, image fusion can improve the spatial resolution of the image, and change detection can find out the temporal dynamics of features. In industrial vision inspection, shape descriptors can be used to accurately locate defects, and surface illumination models can be used to detect surface scratches, skinning and other subtle defects. Digital image processing is also widely used in face recognition, fingerprint recognition, iris recognition and other biometric features, as well as automatic driving, security monitoring, medical imaging and many other fields.

In short, digital image processing through the digital image, the use of computer algorithms from the image to extract, analyse and identify information, greatly expanding the ability of human access to image information and use. With the development of computer hardware and software level, machine learning theory, and big data resources, digital image processing will play an increasingly large role in intelligent perception, human-computer interaction, scientific research and other aspects.

3. CORE TECHNOLOGY ANALYSIS

3.1. Image Enhancement and Restoration

Image enhancement and restoration is one of the core technologies in digital image processing, which aims to improve or restore image quality, making the processed image more suitable for specific applications. Commonly used image enhancement techniques include histogram equalisation (Histogram Equalization), gamma correction (Gamma Correction), contrast stretching (Contrast Stretching) and so on. Histogram equalisation enhances image contrast by redistributing the distribution of image grey levels to make the grey level distribution more uniform. Gamma correction adjusts the grey level of an image through a non-linear transformation, which can be used to correct the non-linear response of the display device and improve the brightness and contrast of the image. Contrast stretching is used to stretch the grey scale range of an image through a linear transformation to improve the contrast of an image.

Image restoration techniques are used to recover degraded images, common degradations include blur, noise, motion blur and so on. Wiener Filter is a classical image restoration algorithm, by minimizing the mean square error to design the optimal filter, the degraded image in the frequency domain of the inverse filter, while taking into account the impact of the noise power spectrum. Lucy-Richardson algorithm is based on the maximum likelihood estimation, using an iterative approach to the restoration of degraded images, in each iteration to estimate the product of the real image and the degraded function with the observed image. The Lucy-Richardson algorithm is based on maximum likelihood estimation and uses an iterative approach to recover the degraded image, estimating the ratio of the product of the real image and the degraded function to the observed image in each iteration until convergence. Blind Deconvolution (Blind Deconvolution) to deal with motion blur, in the case of unknown degradation function to restore the image, the common methods are maximum likelihood estimation and expectation maximisation algorithm.

The key factors in image enhancement and restoration include the evaluation criteria of the image, the use of prior knowledge and the computational efficiency of the algorithm. Commonly used objective evaluation metrics include peak signal-to-noise ratio (PSNR), structural similarity (SSIM), etc. However, subjective visual quality assessment is more important. A priori knowledge, such as degradation models and statistical properties of noise, can be used to guide the algorithm design and constrain the solution space. In addition, image restoration algorithms are usually computationally intensive and require a trade-off between restoration quality and computational efficiency.

In recent years, emerging theories such as sparse representation and low-rank decomposition provide new ideas for image enhancement and restoration. The dictionary learning method based on sparse representation can be used for image denoising and super-resolution reconstruction through adaptive learning of image sparse representation dictionary. Low-rank decomposition decomposes the image into low-rank and sparse components, the low-rank component corresponds to the image structure information, and the sparse component corresponds to the noise and texture details, and image enhancement can be achieved by processing the components separately.

In conclusion, as the core technology of digital image processing, image enhancement and restoration play an important role in improving the visual quality of images and recovering degraded images. Traditional methods such as histogram equalisation and Wiener filtering have been widely used, while emerging theories such as sparse representation and low-rank decomposition have injected new vitality into them. In the future, image enhancement and restoration techniques will develop in the direction of intelligence and adaptivity, and artificial intelligence methods such as deep learning are expected to make breakthroughs to meet the growing demand for image processing.

3.2. Feature Extraction and Identification

Image feature extraction is the process of extracting feature vectors from an image that can reflect the essential attributes of the image and distinguish different categories of images. For image detail features, Kim et al. proposed a multi-scale edge detection algorithm based on Haar wavelet transform, which achieves effective extraction of image detail features by extracting horizontal, vertical and diagonal edge features in different scales. For texture features, Ojala et al. proposed a texture description operator called Local Binary Pattern (LBP), by comparing the size relationship between a pixel and its neighbourhood pixels, encoding the comparison results into binary strings, and then counting the frequency of occurrence of these binary strings to form a description of the image texture features. Experiments show that the LBP operator has grey scale and rotation invariance, and can effectively distinguish different categories of texture images.

On the basis of feature extraction, the purpose of image recognition is to classify images according to these features. Support Vector Machine (SVM) is a commonly used method for image recognition. Its basic idea is to find an optimal classification hyperplane in the feature space, so that different categories of samples can be separated by the hyperplane. Chapelle et al. studied the impact of various kernel functions of SVM on image recognition performance, and found that SVM has the highest recognition accuracy when using RBF kernel function, and achieves 98.76% accuracy in MNIST handwritten digit recognition task. In addition, the rise of deep learning technology in recent years has brought new breakthroughs in image recognition, Krizhevsky et al. proposed a deep convolutional neural network (CNN) model AlexNet, which contains five convolutional layers and three fully-connected layers, and achieved a top-5 error rate of 15.3% on the ImageNet image classification task, which is far more than the traditional methods. Since then, Google, Microsoft and other technology companies and academic institutions have carried out research on CNN models, constantly refreshing the performance record of image recognition.

In general, feature extraction and recognition is the core technology in digital image processing, and its research depth and application expansion have greatly promoted the development of computer vision field. Traditional feature extraction methods such as Haar wavelet, LBP, etc. have become more and more mature in many years of practice, but how to automatically learn a more robust and discriminative image feature representation is still an urgent problem to be solved. In the image recognition task, the deep learning model shows amazing performance advantages, but for the recognition problem in small samples and unconstrained scenes, how to improve the generalisation ability of the deep model, how to use the a priori knowledge to guide the model training, and how to achieve the interpretability of the model, etc., are all research directions worthy of further exploration. In the future, with the continuous improvement of computer hardware performance and the accumulation of large-scale labelled data, the image feature extraction and recognition technology will be greatly developed, providing solid support for intelligent monitoring, automatic driving, human-computer interaction and other application fields.

4. COMMONLY USED PROCESSING ALGORITHMS

4.1. Airspace Processing Methods

Null domain processing is the most direct and basic method in digital image processing. Its basic idea is to operate directly on the pixel values of the image, so as to achieve the purpose of image enhancement, restoration, segmentation, etc. Common null domain processing methods include grey scale transformation, histogram correction, spatial filtering and so on. Common null domain processing methods include grey scale transformation, histogram correction, spatial filtering and so on.

Gray scale transformation is a simple but effective image enhancement method. By establishing a functional relationship between the grey value of the input image and the output image, the grey range of the image can be stretched or compressed to improve the contrast and brightness of the image. For example, the logarithmic transformation $s=c\log(1+r)$ can be used to broaden the dark details of the image, and the power-law transformation $s=cr\gamma$ can control the overall brightness of the image. Research has shown that $\gamma=0.6$ is the best for human eyes.

Histogram correction by redistributing the grey values of the image, so that the gray level histogram shows a specific shape, such as Gaussian distribution or uniform distribution. CLAHE algorithm divides the image into 8 × 8 small blocks, histogram equalization of each sub-block, and limit the maximum enhancement of the grey values (e.g., 40) to avoid over-enhancement of the noise.Experiments conducted by Menotti et al. show that CLAHE can be a good solution to the problem of noise enhancement of faces. CLAHE can reduce the error rate of face recognition by 8.1%.

Spatial filtering is an important means of eliminating image noise and extracting edge features by weighting the pixel values with the grey scale information in the pixel neighbourhood. Mean filtering uses the arithmetic mean of all pixels in the neighbourhood to replace the central pixel, which is effective in suppressing Gaussian noise, but causes blurring of edges. Weighted median filtering considers the distance from the neighbourhood pixels to the centre point and sets the weighting coefficient, which overcomes the shortcomings of traditional median filtering and retains more edge details while removing impulse noise.

Sobel, Prewitt, Roberts and other differential operators use the differential approximate gradient of the neighbourhood pixel values to extract the image edges. Canny edge detection further introduces the non-maximum value suppression and double-threshold connection to improve the accuracy of edge localization, which is recognized as the optimal edge detection algorithm.Arbeláez et al. experimentally proved that combining the local Canny edge with the global optimization model can obtain a better edge than the one obtained by human beings. Arbeláez et al. demonstrated that the combination of local Canny edge and global optimisation model can achieve higher than average human edge detection effect (ODS F-measure = 0.79).

Perona-Malik anisotropic diffusion filter is widely used for image denoising and enhancement in CT, MRI, PET and other medical images. The core is a nonlinear partial differential equation, and the diffusion coefficient is dynamically adjusted with the modulus of the gradient, which reduces the diffusion rate at larger gradients and suppresses the diffusion of the edges and details. Yoo et al. used the improved Perona-Malik filter to process clinical CT images, and the signal-to-noise ratio was increased from 22.9 to 35.4 dB, with a markedly improved sharpness.

In conclusion, the airspace processing method is intuitive, efficient, and the basis of digital image processing, and is widely used in noise reduction, enhancement, segmentation, and so on. With the improvement of computer performance, more advanced null domain algorithms keep emerging, such as adaptive filtering, morphological filtering, partial differential equations, etc., which greatly expand the application scope and performance of null domain processing and lay a solid foundation for image understanding and computer vision.

4.2. Frequency Domain Processing Methods

Fourier transform is one of the most widely used methods in digital image frequency domain processing. Due to the separability of Fourier transform, the two-dimensional DFT can be decomposed into two one-dimensional DFTs by first doing one-dimensional DFT on rows and then one-dimensional DFT on columns, which greatly reduces the computational complexity.

In the frequency domain, the low-frequency components of the image reflect the basic contour and grey-scale distribution information of the image, while the high-frequency components correspond to the detailed information of the image such as edges and textures. By filtering the image in the frequency domain, certain frequency components can be selectively enhanced or suppressed to achieve image enhancement, denoising, edge extraction and other operations. Commonly used frequency domain filters include low-pass filters, high-pass filters, band-pass filters and band-stop filters.

In addition to ideal filters, Butterworth filters, Gaussian filters are also widely used in frequency domain image processing.

In short, the frequency domain processing method by transforming the image from the spatial domain to the frequency domain, using the characteristics of the image in the frequency domain, the design of appropriate filters for image enhancement, recovery, segmentation and other processing, is one of the important means of digital image processing. Compared with the air-domain processing method, the frequency-domain method has the advantages of flexible design, efficient processing, and can achieve better results in many occasions. With the development of fast algorithms and the improvement of computational power, frequency domain processing methods will play an increasingly important role in the field of image processing.

5. CHALLENGES AND OPPORTUNITIES

5.1. Current Problems

Although digital image processing technology has made great progress, but still faces many challenges. The first and foremost is the image quality problem, noise, blur, distortion and other factors seriously affect the visual effect of the image and the integrity of the information. Taking medical imaging as an example, images collected by CT, MRI, X-ray and other imaging equipment are often subject to a certain degree of noise pollution, causing interference in the subsequent analysis and diagnosis. On the other hand, with the increasing resolution of imaging equipment, the huge amount of image data brings great pressure on storage, transmission and calculation. In the field of remote sensing, for example, the data volume of a scene of high-resolution satellite images can reach tens of gigabytes (GB), which brings a severe test to ground reception, processing and application.

In addition, target detection and recognition in complex scenes is always a difficult problem. Traditional feature extraction algorithms, such as SIFT, SURF, etc., are not robust enough under the influence of lighting changes, perspective changes, occlusion and other factors. Although deep learning alleviates this problem to a certain extent, it also suffers from defects such as sample dependence, large computation volume and poor interpretability. Taking face recognition as an example, under uncontrolled conditions such as pose, expression, illumination, etc., even with advanced CNN network, its recognition accuracy is hardly satisfactory [16]. For medical imaging, the morphological differences of tissues and organs, and the diversity of lesion sites bring great challenges to computer-aided diagnosis (CAD) systems.

Image security and privacy protection are also of increasing concern. On the one hand, image tampering, face-swapping, DeepFake and other technologies pose a threat to personal privacy and social order; on the other hand, the storage and transmission of massive image data in a networked environment also faces the risk of data leakage and illegal use. How to achieve intelligent image analysis and processing under the premise of ensuring user privacy is an urgent problem to be solved. The image analysis framework based on encrypted computing, federated learning and other privacy-preserving mechanisms is expected to become a future research hotspot [17].

Finally, cross-modal image fusion and understanding need to be studied in depth. In many application scenarios, it is often necessary to comprehensively analyse heterogeneous data from multiple sources, such as visible light images, infrared images, SAR images and so on. The images obtained by different imaging mechanisms have large differences in resolution, viewing angle, information content, etc. How to achieve accurate alignment and semantic fusion of different modal images is a challenging topic. Most of the current mainstream multimodal deep learning methods are based on shared representation learning, but the full use of cross-modal information and semantic alignment still need to be strengthened [18]. With the rise of new neural network structures such as Transformer, new breakthroughs in cross-modal image understanding are expected.

5.2. Future Trends

With the rapid development of artificial intelligence, big data, 5G and other new generation information technology, digital image processing technology will usher in a broader application prospect. Deep learning-based image processing methods are one of the important development directions in the future. Deep learning models such as convolutional neural networks (CNN), generative adversarial networks (GAN), etc., can be trained through the massive data, automatically extract and learn the image of multi-level, high-level semantic features, significantly improve the performance of image classification, target detection, semantic segmentation and other tasks, has been successfully applied to face recognition, automatic driving, medical image analysis and many other fields. For example, the AlexNet model proposed by Krizhevsky et al. achieved a breakthrough result of top-5 error rate of 15.3% on ImageNet image classification task, which opened a new era of deep learning in the field of computer vision. In the future, deep learning models will be optimised in the direction of deeper, wider and lighter weight to adapt to more complex and changing application scenarios.

Multimodal fusion is also an important research hotspot in digital image processing technology. With the explosive growth of heterogeneous data from multiple sources, single-modal image information is difficult to meet the growing demand for perception and understanding. By fusing images with text, audio and other modal data, it can characterise objective things more comprehensively and three-dimensionally, and provide multi-channel and multi-level information required for human-computer interaction. For example, the CLIP model developed by OpenAI has built a bridge connecting vision and language by pre-training on a large number of graphic and text pairs, and has achieved remarkable performance in tasks such as zero-sample learning and cross-modal retrieval. Multimodal fusion is expected to become the mainstream paradigm of future artificial intelligence, greatly expanding the application space of digital images in smart cities, smart homes, virtual reality and other fields.

At the same time, digital image processing technology is accelerating the migration to the mobile side, terminal side, with lower latency, higher security, stronger customisation to meet the explosive growth of application needs. Thanks to the continuous improvement of mobile computing power and the maturity of model compression, knowledge distillation and other lightweight technologies, it has become possible to run complex visual algorithms in real time on mobile phones, cameras and other resource-constrained devices. The launch of Huawei HiAI, Apple CoreML and other mobile AI engines has further driven the landing of a wide range of applications such as face unlocking, photo beauty, AR games and so on. In the future, in algorithms, chips, architectures and other levels, end-cloud collaboration, hardware and software will become the core paradigm of digital image processing to cope with the explosive growth of data volume and computation.

As the concepts of meta-universe and digital twin continue to heat up, the boundaries between the ‘real world’ and the ‘virtual world’ are becoming increasingly blurred. As an important bridge connecting the real and virtual worlds, digital image processing technology will play an indispensable role in building a lifelike digital world. Through accurate perception of the physical world, efficient reconstruction and in-depth understanding, digital image will help human beings to build a parallel digital space with the real world, realising full-fidelity perception, immersive experience, real-time interaction, and bringing about a new change in the way of life and production. In the near future, human beings will live in a meta-universe that seamlessly integrates the real and virtual worlds, and digital image processing, as the underlying support technology, will usher in a broader space for development.

6. CONCLUSION

Comprehensive research content of this paper can be concluded as follows: digital image processing technology since the 1960s after more than half a century of development, has become the most active in the field of computer science and artificial intelligence, one of the most widely used branches. It covers many aspects from image acquisition, compression and coding, enhancement and restoration, to target detection, feature extraction, pattern recognition and so on. Traditional spatial and frequency domain processing methods, such as histogram equalisation, spatial filtering, frequency domain filtering, etc., have laid the foundation for image quality improvement and information extraction. In recent years, with the rise of deep learning and other artificial intelligence technologies, new algorithms such as convolutional neural networks and generative adversarial networks have been widely used in semantic segmentation of images, target detection, face recognition, image synthesis and other tasks, which have significantly improved the performance of image understanding and generation.

However, current digital image processing still faces many challenges. First of all, the huge amount of heterogeneous image data from multiple sources brings great pressure on storage, computation and analysis, which requires the development of more efficient data organisation and management and parallel computing methods. Secondly, image processing in complex dynamic scenes, such as occlusion, motion blur, lighting changes, etc., puts forward higher requirements for algorithm robustness. Again, cross-modal image processing, such as infrared and visible image fusion, medical imaging and anatomical atlas alignment, etc., requires deep excavation of the intrinsic connection of different modal data. In addition, the interpretability and credibility of image processing need to be strengthened, and it is necessary to develop interpretable model design methods and evaluation systems. At the same time, image processing in mobile terminals, Internet of Things and other resource-constrained environment of lightweight implementation, as well as data security and privacy protection, are also urgent practical problems.

Looking into the future, digital image processing technology will continue to develop in the direction of intelligence, automation, real-time and lightweight. On the one hand, the new artificial intelligence technology represented by deep learning will further empower image processing and achieve higher-level semantic image understanding, knowledge representation and reasoning judgement. Unsupervised and self-supervised learning, continuous learning, reinforcement learning, graphical neural networks and other methods will be used more often. On the other hand, the algorithm architecture design will pay more attention to the computational efficiency, energy consumption and model size, so as to adapt to diversified application scenarios and hardware platforms. New neuron models such as SNN, PNN, etc. and new devices such as amnesia will bring innovative opportunities for image processing. At the same time, blockchain, federated learning, privacy computing and other technologies will ensure the security of data sharing in a distributed environment, AR/VR, automated driving, industrial quality inspection, security monitoring, biomedical and other fields will become an important direction for image processing technology to land.

To sum up, as one of the core support technologies of artificial intelligence, digital image processing has made great progress and plays an important role in various fields of national economy and social life, such as industry, agriculture, transport, medical treatment and aerospace. The evolution and integration of the new generation of information technology will certainly promote the image processing from the perception layer to the cognitive layer, from the single modality to the multi-modal fusion, from the centralised to the distributed span, opening a new era of intelligent image processing. We have reason to believe that the future of digital image processing science and industry is full of imagination and vitality.

REFERENCES

[1] Liu Jinxiu. Research on rapid measurement technology of particle size distribution of loose accumulator based on digital image processing technology [J]. 2014.

[2] Zuo Chao, Chen Qian. Computational Optical Imaging: Where to come from, Where to go, Where to go, Where to go? [J]. Infrared and Laser Engineering, 2022.

[3] Y Yu. Influence and Application of Digital Image Processing Technology on Oil Painting Creation in the Era of Big Data [D]. Lecture Notes in Electrical Engineering, 2022.

[4] Yang ZhuoLin. A Review of Research on the Application of Digital Image Processing Technology in Asphalt Pavement Technology [J]. Fujian Quality Management, 2019.

[5] S Jiang, H Zhang. Method for extracting crack area on structure surface based on digital image processing technology [D]. Journal of Physics Conference, 2020.

[6] Li Qiuping. A review of patented technologies for digital image restoration [J]. Big Science and Technology, 2019.

[7] PCD Silva, C Junior, J Huguenin, et al. Investigation of copper and zinc alloy surface exposed to corrosion environment by digital image processing [D]. Journal of Materials Research & Technology, 2023.

[8] Hu YT, Huang S, Liu N, et al. Progress of imaging histology in primary liver cancer [J]. 2021.

[9] JIANG Hao-Ran. Application and development of digital image processing [J]. Electronic World, 2020.

[10] Guo Yanyan. A review of image stitching technology [J]. Information Weekly, 2019.

[11] Xuan Kefan, Li Xiaopeng, Zhang Jiabao, et al. A review of organic materials to improve soil structure and its quantitative research methods [J]. Journal of Irrigation and Drainage, 2023.

[12] LIU Yufei, FAN Jiansheng, NIE Jianguo, et al. Review and prospect of digital image method for structural surface crack identification [J]. Journal of Civil Engineering, 2021.

[13] ZHANG Yu, YAN Jiayong, GU Liang. A review of the development of new laparoscopic technology [J]. Chinese Journal of Medical Devices, 2019.

[14] LIU Mei, GOLD, ZHANG Yaya, et al. Application and prospect of microscopic image processing technology [J]. Information Technology and Informatisation, 2022.

[15] Zhang Hongling. Detection and analysis of comprehensive performance of recycled concrete using digital technology [J]. Bonding, 2021.

[16] Ranjan R, Sankaranarayanan S, Bansal A, et al. Deep learning for understanding faces: Machines may be just as good, or better, than humans [J]. IEEE Signal Processing Magazine, 2018, 35(1): 66-83.

[17] Acar A, Aksu H, Uluagac A S, et al. A survey on homomorphic encryption schemes: Theory and implementation [J]. ACM Computing Surveys, 2018, 51(4): 1-35.

[18] Baltrusaitis T, Ahuja C, Morency L P. Multimodal machine learning: A survey and taxonomy [J]. IEEE transactions on pattern analysis and machine intelligence, 2019, 41(2): 423-443.