

A Review of Image Recognition Technology

Xingang Wang

Administration Center for the DTH Service in China, SAPPRFT, FuXingMengWai Street, XiChen District, Beijing, China, 100045

Abstract—With the development of the research of artificial intelligence, image recognition technology has become one of the hot research topics in the modern society, the research results have been widely applied to many fields, which has brought remarkable economic benefits and social benefits. This paper briefly describes the concept and basic process of image recognition technology, reviews the development history of image recognition technology and summarizes the application status of it. The main algorithms involved in the key steps, such as feature extraction and classification matching, are systematically summarized and analyzed in this paper. At last, the future direction of image recognition technology is prospected.

Keywords—image recognition; feature extraction; algorithm; development trend

I. INTRODUCTION

With the progress of science and technology and the further popularization of computer, human beings are no longer satisfied with sensing and processing external information only by their own organs. The subject of artificial intelligence has drawn wide attentions from industry, academia and media. It uses machines to simulate the human mind to deal with a large number of physical information, so as to achieve a certain subjective image recognition, natural language recognition and other image thinking work. The extensiveness of artificial intelligence makes the related application research reach deeply into many fields and achieved remarkable results, among which image recognition is one of the most representative technologies.

The ability of human eyes on image resolving is outstanding. To achieve an image recognition system to replace human eyes and brain has been a goal pursued by human beings for many years. Image recognition utilizes computers to simulate the process how human beings analyse and understand target images. Research shows that 60% of human perception information comes from the sense of vision, so the research of image recognition is of great significance[1]. In the recent 20 years of development, the objects of image recognition have gradually enriched from simple words, numbers to human faces, scenes and fine targets, etc. With the emergence of new technologies such as machine learning, excellent methods such as artificial neural networks and support vector machine classification have also injected new vitality into the image recognition technology. At present, the technology has shown great application value in numerous fields such as criminal investigation, biomedicine, astronomy and meteorology, financial banking, multimedia network communication, food testing, industry, agriculture and so on, and has extensive research prospects.

II. OVERVIEW OF IMAGE RECOGNITION

The principle of image recognition is to use computer technology and mathematics to preprocess the obtained target image information and extract the feature. Then the classifier classifies the image into the corresponding category, and finally the recognition result is gained from matching with the stored information. Generally, the process consists of the following steps: acquisition of image information, image preprocessing, image feature extraction and selection, classifier design and decision matching[2], as shown in Figure 1.

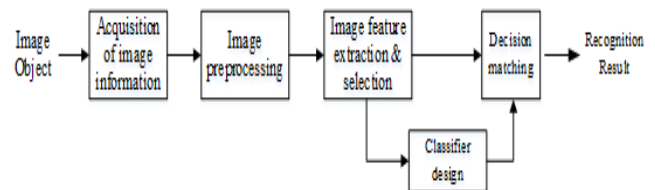


FIGURE I. BASIC STEPS OF IMAGE RECOGNITION

A. Acquisition of Image Information

Optical signals and analog signals are converted into electric information by means of cameras, image conversion cards, sensors and so on. Information can be expressed as a one-dimensional waveform, or two-dimensional text, images and so on, while a dynamic video is extracted into single frame images.

B. Image Preprocessing

Mainly refers to the transformation between images, which makes the information described by the image more concise and more unique, eliminating irrelevant features and non-uniformity, such as processing of the grayscale of color images, image smoothing, transformation, enhancement, filtering and so on. Sometimes in order to find the key research part in the image background, it is needed to perform some technical processing on the image such as edge detection and area connection, to achieve the effect of image segmentation.

C. Feature Extraction and Selection

In order to save resources, the elements that describe an object are usually extracted and selected according to some principles under the premise of meeting the requirements of classification precision and accuracy. Generally, this will be reflected by the reduction in the dimension of the feature vector or the simplification of the graph structure. This step is the core of the whole image recognition, and the key point is to ensure

the non-deformation of the image information and the uniqueness of the extracted feature vector.

D. Classifier Design.

The design of classifier is a training process. It establishes a sample image feature library through training and analysis based on certain rules, and determines the decision rules, according to which the lowest error rate of the classification can be achieved. In this step, attention should be paid to that the training sample set should be broad, and the classification performance of the training sample set should be ensured to be consistent with the sample to be identified.

E. Decision Matching

According to the differences discovered from the integration of image features, a reasonable classification algorithm is used to quantitatively analyze the extracted features, and the extracted feature sets will be put into a certain category of the stored feature category information for cognitive purpose. It is usually necessary to design a decision-making function based on the pattern matching theory.

III. THE DEVELOPMENT OF IMAGE RECOGNITION AND ITS APPLICATION STATUS

A. The Development of Image Recognition

Image recognition originated in the middle of the last century, has gone through three important stages. The objects for recognition have developed from simple symbols, texts to images and the present three-dimensional objects[3]. The first stage is the recognition of symbolic text, which appeared in 1950. As the technology advances, the current text recognition and complex symbol recognition already have very high accuracy, and many successful applications have emerged, including translation software. The second stage is the recognition and processing of digital image, which, not only in response to the digital trend of computer development, but also take advantage of the digital information features, such as large amount of storage, not easily distorted and so on. The third stage is the recognition of objects, which mainly refers to the perception of objects and environments in three-dimensional space. Based on digital image recognition, combined with artificial intelligence and 3D techniques, it is widely used in robots and other projects. It belongs to the advanced computer vision category. The rapid development of mobile Internet and the popularization of smart terminals have brought massive picture information, making the application range of image recognition further widened.

Major colleges, universities, enterprises and other research institutes attach great importance to the research of image recognition. IEEE and International Association for Pattern Recognition (IAPR) organized a special conference for academic exchange[4]. Intelligent algorithms are becoming more and more popular in the field of image recognition, such as retina recognition based on improved Gabor filter and scale invariant feature transform[5], SVM-based offline image target classification algorithm[6], convolution neural network based

face recognition method[7], deep learning based hand-painted sketch recognition[8].

At the same time, there are some successful commercial software put on the market. One of the most representative areas is OCR (Optical Character Recognition). OCR achieves text reading by computers. Originally it was widely used for information input of paper records, passport documents, invoices, bank statements, business cards, e-mail printing, or any suitable documents. With the proliferation of smart terminals and the Internet, OCR is now available for Internet-connected mobile device applications. It can extract text from the device's camera and return it to the device application for further processing (such as text-to-speech) or display[9]. A variety of commercial and open source OCR systems are suitable for most commonly used writing languages. Other representative products include Facebook's DeepFace and Google's Facenet, all for face recognition; Glow, a mobile photo gallery application developed by IQ Engine; as well as the Baidu scribbling notes, Baidu translation etc.

B. The Application Areas of Image Recognition

Image recognition technology has the advantages of transparency and intelligence. It is widely used in many fields such as criminal investigation, biomedicine, astronomy and meteorology, financial bank, multimedia network communication, industrial agriculture and so on [10].

1) *Remote sensing image recognition*: through processing and analyzing satellite remote sensing pictures and aerospace images, this technology can be used for geology and geophysics, weather forecasting, intelligent detection of agricultural diseases, environmental pollution detection, military target detection and tracking, guidance and so on.

2) *Public transport and police criminal investigation*: image recognition is widely used in this field, such as intelligent identification of fingerprints, handwriting and portraits at the scene of the crime, and debris restoration of important documents and pictures; it can also be used in transport aspect for navigation and license plate recognition, etc.; in public life aspect: fingerprint recognition and face recognition are widely used in access control to ensure public safety.

3) *Biomedicine*: image recognition is mainly used to improve clinical diagnosis and case studies, such as cancer cells, white blood cells, chromosomal examination, repair surgery control design and so on.

4) *Computer vision*: machine vision is the most important sense organs of intelligent robots. Intelligent robots can recognize 3D images through cameras, so as to realize product sorting and packaging, industrial workshop assembly line guidance, dangerous environment reconnaissance, and even postal services, hospitals and home services.

5) *Financial, communication and other fields*: today's commonly used bar codes and QR codes are inseparable from the image recognition technology, which greatly facilitate people's online transactions and social networking.

IV. THE MAIN METHODS OF IMAGE RECOGNITION

This section provides a brief introduction to the main algorithms involved in the key steps of image recognition.

A. Image Preprocessing

Image preprocessing involves many kinds of function blocks, and the method and the order of selecting each function have an important influence on the subsequent recognition results. Here are some of the main processing approaches.

1) *Grayscale processing*: grayscale processing is to convert a color image into a gray image to reduce the amount of computation of subsequent image processing[11]. Grayscale histogram is one of the most important applications, which shows the number of pixels of each gray level in the image, and can be used to extract the color features of the image. Digital image point operations to the histogram can achieve some different processing effects, for example, the use of piecewise linear transformation function for grayscale stretching can improve the contrast of the image in order to facilitate the observation and distinction.

2) *Image smoothing*: image smoothing is mainly used to highlight the main area of the image, reduce the noise interference of the high frequency part, and improve the image quality. The main methods include: mean filtering, median filtering, wavelet filtering, Gaussian low-pass filtering and statistical filtering, etc.[12] Image noise can be divided into many categories. Different types of noise can use different filtering methods. For example, median filtering is a non-linear image smoothing technique and is most suitable for filtering salt and pepper noise.

3) *Image segmentation*: threshold method, region growth method and edge detection method are the three most commonly used segmentation methods[13]. The threshold method performs parallel segmentation based on the gray value of different targets, the calculation is simple and fast, but the selection of the threshold has great effect on the result and is sensitive to noise, and only the gray value information is considered while the spatial information is ignored. The region growing method is a kind of serial region segmentation method, which divides pixels with similar property into one region. The result depends on the selection of the initial point, the designed principle of growth and termination, and it is more suitable for small region segmentation, but the disadvantage is its low efficiency. The edge detection method uses the discontinuity of image edge (such as mutations of brightness, color, texture) for segmentation. Usually the template operator is used to conduct derivative operation for detection. Commonly used differential operators include Prewitt operator, Roberts operator, Sobel operator, Laplace operator, Kirsch operator, Canny operator, etc.[14].

B. Feature Extraction and Selection

It is well-known that different images have different representative features. The proper selection of extraction methods has a great influence on the subsequent recognition

results. The four common image features and their typical methods are summarized below.

1) *The visual features include the edges, colors, textures, shapes, etc.* of the image, which have strong physical properties. The color feature extraction method is mainly based on the histogram feature extraction. The representative method of texture feature extraction is the Gray Level Co-occurrence Matrix (GLCM) proposed by Haralick et al. It can use the probability to reflect the comprehensive information of the direction of the image gray, the adjacent interval and the transformation amplitude[15]. The six texture features related to human visual perception proposed by Tamura is another representative method of texture feature extraction[16]. In addition to the first-order, second-order differential operators mentioned above, edge feature extraction methods also contain Hough transform detection[17], blur enhanced edge detection, image edge detection based on SVM, and so on.

2) *Statistical features mainly include gray histogram features, moment features and so on.* Commonly used methods for extracting moment features include the two-dimensional moment invariance theory based on planar geometry proposed by Hu[18], the discrete moment invariant algorithm and the entropy feature extraction method proposed by Pan et al.[19], and so on. The Zernike moment proposed by Miao improves the information redundancy of Hu's moment, and improves the recognition rate[20].

3) *The transform coefficient feature regards a series of mathematical transform coefficients as image features.* Common image transformation includes Fourier transform, wavelet transform, discrete cosine transform, wavelet packet transform. Laine et al. first proposed the feature extraction method based on wavelet transform in 1993, which provided a new idea for image feature extraction[21]. Pittner et al. proposed a unified wavelet transform framework[22]. In recent years, improved algorithms are emerging, such as Gabor wavelet transform feature extraction based on direction pre-classification, face recognition based on adaptive wavelet transform feature extraction[23] and so on.

4) *Algebraic features reflect the intrinsic properties of the image, which can be obtained by algebraic transformation or matrix factorization of the image storage matrix.* The most representative is the Singular Value Decomposition (SVD). In addition, there are Projection Pursuit (PP), Principal Component Analysis (PCA), Independent Component Analysis (ICA) and so on[24], which are all linear feature extraction methods. Among them, PCA can be used for dimensionality reduction of multidimensional data, but only the covariance matrix is considered, while the high-order statistical properties are ignored. The ICA takes into account the high-order statistical properties of the data and removes the correlation between the components. Fisher linear discriminant is a feature extraction method based on PCA dimensionality reduction, which takes into account the inter-class relations of the samples. Typical nonlinear feature extraction methods include Support Vector Machine (SVM), Kernel-based

Principal Component Analysis (KPCA), neural network and so on.

C. Category Recognition Methods

Category recognition methods are mostly related to pattern recognition algorithms and can be broadly divided into the following three categories.

1) *Statistical pattern recognition*: Such methods are based on mathematical theory of decision-making on the target image analysis and abstraction, to find out the law and the nature of the image to be identified. The key of this method lies in the extraction of image features and the design of decision functions. For feature extraction, PCA methods and so on are often used to reduce the dimension of the eigenvectors. The design of the decision function can take the principle of minimum distance, in addition to Bayesian statistical model, Markov statistical model, D-S evidence theory. This method is less effective for models with complex structures[25].

2) *Structure pattern recognition*: Structure pattern recognition is also known as syntax pattern recognition. The main principle is to split a complex image into several layers of sub-images, represented by a tree structure. A model is described by a set of primitive models located at the bottom of the tree and their syntactic structures. When recognizing an image, once its sub-image is identified, it can be analyzed and determined according to the established syntax, and will be classified into the category if it meets a certain grammar. This method has a strong anti-interference ability to image distortion, but how to extract the primitive model and construct a proper syntactic structure is the key issue[25].

3) *Pattern recognition based on soft computing*: In recent years, the development of new technologies has injected new vitality into pattern recognition. Recognition methods based on soft computing technology have drawn wide attention of the academic community due to their intelligence advantages. The main methods include fuzzy set theory, artificial neural network, genetic algorithm, clustering, rough set theory, SVM and so on. Fuzzy set theory is a local classification theory based on fuzzy mathematics, it introduced the concept of membership degree, and effectively resolved the problem of judgment conflict[26]. Artificial neural network is a parallel distributed network composed of a large number of neurons which simulates human nervous system, with the advantages of self-organization, self-adaptation and self-learning[27]. Support Vector Machine (SVM) is a classification technique proposed by Vapnik et al. in 1963, it has the global optimality and strict generalization ability, and requires less on quantity and quality of samples. It has high research potential[25].

V. TRENDS OF IMAGE RECOGNITION

Based on the development of image recognition technology and the status quo of research and application, this paper argues that the future development of image recognition technology will show the following trends:

A. Research on Algorithm Theory

Image recognition technology includes many steps, such as acquisition of image information, preprocessing, feature extraction, classification, recognition and so on, all of which are based on the continuous improvement and ingenious combination of various algorithms. To further improve the accuracy of image recognition, the research of algorithm theory is indispensable. At present, the algorithm design of feature extraction and recognition has high priority in many countries' research. Each algorithm itself has some limitations, so proposing of new algorithm, improving existing algorithms or putting forward more appropriate combination ways, will be the keys to a major breakthrough in future image recognition. In addition, the complementary combination of classical recognition algorithms and modern emerging technologies[28] is also one of the focuses of future research.

B. New Breakthroughs in Key Technologies

With the continuous development of computer technology and other related technologies, there are new possibilities and urgency of development in all aspects of image recognition. For example, in the field of image compression, the research focus has also gradually transformed from lossless compression with limited compression ratio to lossy compression. The traditional two-dimensional design has gradually developed into three-dimensional space, and three-dimensional reconstruction will play a more important role in the relic rehabilitation, medical imaging, three-dimensional maps and other fields. In addition, virtual reality, a hot technology in recent years, has provided new opportunities for digital identification technology.

C. High Integration with Other Technologies

Human perception of the outside world does not rely solely on the vision, so the high integration of image recognition technology and other technologies (such as speech recognition technology) can further improve the matching effect of artificial intelligence.

VI. CONCLUSIONS

With the development of technology, image recognition technology has achieved success in some extent. However, there are still many challenges for complex situations. For example, there are some limitations in image recognition algorithms, and there are no specific evaluation criteria, and there is still possibilities for the improvement in recognition accuracy. The coupling relationship between feature extraction and classification recognition algorithms needs to be further studied. However, we can believe that with the continuous improvement and innovation of technologies, image recognition technology will become more and more matured to provide people with more intelligent, automated, comprehensive and personalized services.

REFERENCES

- [1] Wang Chen, Zhang Tao, Lin Weimin, et al. Comprehensive Image Recognition and Its Application in Power Information Security [J]. Computer Technology and Development, 2012, 22 (4): 161-164.

- [2] Wang Mingxiao. Research of Image Recognition Algorithm and Its Implementation on Intelligent Terminal [D]. Beijing University of Posts and Telecommunications, 2010.
- [3] Li Longfei. Analysis of Computer Image Recognition Technology Development Status and Prospects [J]. Digital Communications World, 2016 (5): 29-29.
- [4] Li Wujun, Wang Chongjun, Zhang Wei, et al. Research Overview of Face Recognition [J]. Pattern Recognition and Artificial Intelligence, 2006, 19 (1): 58-66.
- [5] Meng X, Yin Y, Yang G, et al. Retinal Identification Based on an Improved Circular Gabor Filter and Scale Invariant Feature Transform [J]. Sensors, 2013, 13 (7): 9248-9266.
- [6] WANG Na, WAN Hong-lin, BAI Cheng-jie. Offline Image Target Classification Based on SVM [J] Application of Computers, 2016,25 (2): 208-211.
- [7] Zhang Y, Zhao D, Sun J, et al. Adaptive Convolutional Neural Network and Its Application in Face Recognition [J]. Neural Processing Letters, 2016, 43 (2): 389-399.
- [8] ZHAO Peng, WANG Fei, LIU Hui-ting, et al. Hand-drawn Sketch Recognition Based on Depth Learning [J]. Journal of Sichuan University (Engineering Science Edition), 2016,48 (3): 94-99.
- [9] Peng X, Cao H, Setlur S, et al. Multilingual OCR research and applications: an overview [C] // International Workshop on Multilingual Ocr. 2013: 1-8.
- [10] Chen Aixia, Du Youfu, Chen Zhongju. Application Status and Research Direction of Image Recognition [J]. Science and Technology of China, 2015 (16): 82-82.
- [11] R. C. Gonzalez, R. E. Woods. Digital Image Processing (Second Edition) [M]. Beijing: Publishing House of Electronics Industry, 2007.
- [12] Fan Lilan, Han Xiaowei, Zhang Guangyuan. Image Processing and Pattern Recognition [M] Beijing: Science Press, 2007.
- [13] Zhang Yong, Yuan Jiazheng, Liu Hongzhe. Overview of Image Segmentation Algorithms [C] // Annual Meeting of New Technologies and Applications in Networking. 2015.
- [14] Carpinteri A, Cornetti P, Kolwankar K M. Calculation of the tensile and flexural strength of disordered materials using fractional calculus [J]. Chaos Solitons & Fractals, 2004, 21 (3): 623-632.
- [15] R. M. Haralick, K. Shanmugam, I. Dinstein. Texture features for image classification [J]. IEEE Transactions on System, Man and Cybernetics, 1973, 8 (6): 610-621.
- [16] H. Tamura, S. Mori, T. Yamawaki. Texture features corresponding to visual perception [J]. IEEE Transactions on System, Man and Cybernetics, 1978, 8 (6): 460-473.
- [17] Ballard D. H. Generalizing the Hough transform to detect arbitrary shapes ☆ [J]. Pattern Recognition, 1981, 13 (2): 111-122.
- [18] Hu M K. Hu, M. K. .: Visual Pattern Recognition by Moment Invariants. IRE Transaction of Information Theory IT-8 [J]. Information Theory Ire Transactions on, 1962, 8 (2): 179-187.
- [19] Sun J D, Cui J T, Liu W G, et al. Spatial feature extraction and image retrieval based on entropy [J]. Systems Engineering & Electronics, 2006, 28 (6): 791-794.
- [20] Zhenjiang M. Zernike moment-based image shape analysis and its application [J]. Pattern Recognition Letters, 2000, 21 (2): 169-177.
- [21] A. Laine, J. Fan. Texture classification by wavelet packet signatures [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15 (11): 1186-1191.
- [22] S. Pittner, S. V. Kamarthi. Feature extraction from wavelet coefficients for pattern recognition tasks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21 (1): 83-88.
- [23] Sänket N J, Vyshak A V, Manikantan K, et al. Face recognition using adaptive filter wavelet transform based feature extraction [C] // International Conference on Science Engineering and Management Research. IEEE, 2015: 1-9.
- [24] Yan Weidong. Image Feature Extraction Methods [D]. Northwestern Polytechnical University, 2007.
- [25] Fan Huimin, Wang Hao. Overview of Pattern Recognition Methods [J]. Electronic Design Engineering, 2012, 20 (19): 48-51.
- [26] S. K. Pal. Fuzzy image processing and recognition: uncertainty handling and applications [J]. International Journal of Image Graphics, 2001, 1 (2): 169-195.
- [27] Alex Krizhevsky, Ilya Sutskever and G. Hinton. Imagenet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems 25. 2012. 1106-1114.
- [28] Shi J F, Sun B. Research on Image Recognition Based on Invariant Moment and SVM [C] // International Conference on Pervasive Computing Signal Processing & Applications. IEEE, 2010: 598-602.