



# Application of Artificial Intelligence Technology in Text Recognition and Detection Algorithms

Junxia Liang<sup>1</sup> and Yongjun Qi<sup>1,2</sup>(✉)

<sup>1</sup> Faculty of Megadata and Computing, Guangdong Baiyun University, Guangzhou 510450, Guangdong, China

qyj200702022@baiyunu.edu.cn

<sup>2</sup> School of Information and Communication Technology, Mongolian University of Science and Technology, Bayanzurkh district 13341, Ulaanbaatar, Mongolia

**Abstract.** With the development of the times, words play an indispensable role in daily life. Therefore, text recognition detection is very important. In recent years, although there have been many research achievements and progress in text recognition detection, there are still situations of ambiguity and distortion in text recognition. Therefore, it is necessary to further upgrade the text recognition detection algorithm. This paper studied the application of artificial intelligence (AI) technology in text recognition and detection algorithms, aiming to further improve the accuracy of text recognition and detection algorithms through AI technology. This article tested the improvement of text recognition detection accuracy using AI technology through experiments. The experimental data showed an improvement of at least 11% and at most 19%, indicating that AI technology can achieve good results in text recognition and detection algorithms.

**Keywords:** Text Recognition · Detection Algorithms · Artificial Intelligence · Algorithm Applications

## 1 Introduction

Text recognition is currently a popular research topic. Text recognition can scan documents and images, which can bring great convenience for people to obtain this article. Text recognition is constrained by the clarity and color difference of the image. Therefore, text recognition detection needs further improvement, and studying how to upgrade text recognition detection algorithms is of great significance.

Many scholars have conducted some research on text recognition. Scholar Chen X introduced an image database for handwritten text recognition research, which included digital images of approximately 5000 city names, 5000 state names, 10000 postal codes, and 50000 alphanumeric characters [1]. The architecture of the scene text recognition system proposed by Lin H had two unique characteristics. (1) It is end-to-end trainable, and most existing algorithm components are trained and tuned separately. (2) It naturally processes sequences of any length without involving character segmentation or horizontal scale normalization [2]. Petrova O believed that text recognition in natural scenes

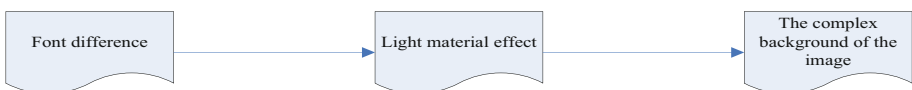
had always been an active research topic in the fields of computer vision and pattern recognition [3]. Although the research on text recognition is deep, there are still some shortcomings that need to be addressed.

The improvement of text recognition detection algorithms is now a very urgent problem to be solved. This article studied text recognition detection algorithms using AI technology. This article tested the user satisfaction of text recognition detection algorithms using AI technology through experiments, and found that the satisfaction was very good. The improvement of text recognition detection accuracy has been tested again, and it has been found that the accuracy improvement is good, which proves that AI technology can play a good role in text recognition detection algorithms.

## 2 Text Recognition and Detection Based on AI

### 2.1 Overview of Text Recognition

Optical character recognition, also known as text recognition, converts images of typed, handwritten, or printed text into machine encoded text electronically or mechanically, whether these characters come from scanned documents, document photos, or scene photos (such as text on signs and billboards in landscape photos) or subtitle text superimposed on images (such as from television broadcasts). Optical text recognition is a discipline that integrates pattern recognition, AI, and computer vision [4, 5]. Optical character recognition allows for easy viewing, searching, and recognition of text in images and labels. Before AI was applied to text recognition, people attempted to solve the machine learning problem of text recognition using traditional computer vision technology. These traditional methods focus on structured text in text recognition. In a standard background, using appropriate lines and standard fonts, most of the text is relatively dense and can be well recognized [6, 7]. The problems with text recognition are shown in Fig. 1:



**Fig. 1.** Problems with text recognition

(1) Significant differences in handwriting and font (such as handwriting and font). (2) Due to factors such as light and material, the same text can have significant differences in different scenes [8, 9]. (3) The complex background, noise, flash, multi word, geometric distortion, etc., of an image. It is in this context that text recognition based AI methods are widely studied and applied, and are widely applied in various industries. The use of AI optical character recognition technology can: (a) Reduce costs. (b) Accelerate workflow speed. (b) Automation of file routing and content processing. (d) Centralization and security of data. (e) Obtain the latest and most accurate information on improving services [10, 11].

Automatic text detection and reading in natural scenes is an important part, which can be used for several challenging tasks, such as image based machine translation,

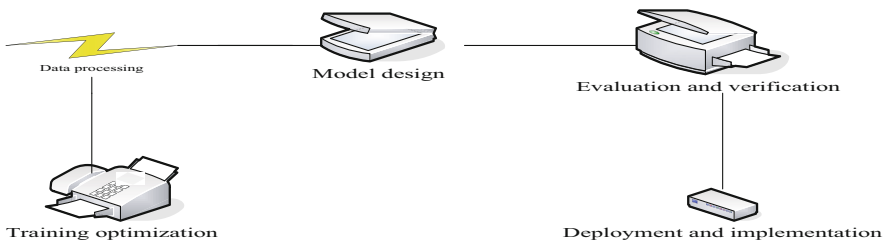
autonomous vehicle or picture, video indexing. In recent years, the task of detecting and recognizing text in natural scenes has received great attention from computer vision and document analysis. In addition, in other fields of AI, recent breakthroughs have made it possible to create better scene text detection and recognition systems.

## 2.2 Steps for Text Recognition

In general, text recognition in images involves two stages. One is to detect the text in the image, and the other is to use recognition technology to obtain information about the text. Text detection is a prerequisite for text recognition, while traditional text detection techniques mainly focus on document images. In these images, there is a significant difference between text and background, and background information does not cause significant interference with the text. In general, in these images, the text carrying important information is black, while the non important background is white. Therefore, it is easy to separate the required text itself from the image, and then recognize the separated text through text recognition technology to obtain text content information. However, in natural scenes, there are significant differences between text and background, such as excessive contrast between text and background, or exposure to light. In addition, in natural scene images, there are differences in text size, font, color, and other aspects. For example, large and small characters can appear simultaneously in images, and text can be non horizontal or even circular. A character can have many colors, which makes text recognition very difficult.

## 2.3 Overview of AI

AI, abbreviated as AI, belongs to a branch of computer science and is a discipline used to simulate, extend, expand, and learn the theories, methods, technologies, and application systems of human intelligence. AI products include intelligent systems, intelligent machines, robots, and so on [12, 13]. There are two ways to implement AI, one is the Engineering approach, which utilizes traditional programming techniques to present intelligent effects to the system, regardless of whether the methods used are those used by humans and animals. The second type is simulation, which needs to be executed in the same way as humans and other organisms [14, 15]. The process of using AI is shown in Fig. 2:



**Fig. 2.** Processes used by AI

## 2.4 Main Application Fields of AI

**Natural language processing:** In the application of AI, the most critical is natural language processing, which enables efficient interaction between people. Deep learning algorithms based on big data and parallel computing have made great progress in natural language processing [16, 17]. At present, the technologies in this field include language engineering, data processing, linguistics, etc. Among them, the most representative ones are customer service, robot chat, etc. In 2017, Google, Apple, Amazon, and other companies all released audio and home appliances controlled by sound [18, 19].

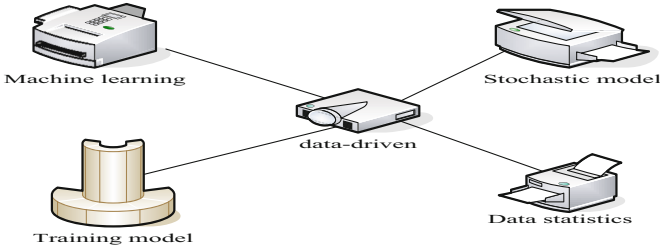
**In the field of computer vision:** AI is mainly applied in the field of computer vision, which imitates the human visual system and enables computers to determine the position, motion status, and recognize objects. In the process of application in the field of vision, there are a total of three stages: target detection, target recognition, and behavior recognition. At present, the most mature ones are facial recognition systems, pupil recognition systems, and fingerprint recognition systems. The most advanced machine vision technology can automatically extract the features of targets from a large amount of data and recognize them based on these features, greatly improving the accuracy of recognition [20].

## 2.5 Technical Support for AI

The technical support of AI includes algorithms: the so-called algorithms describe the strategies and mechanisms for solving problems in a systematic way, so that they can obtain the desired results in a limited time based on specific rules. In recent years, with the continuous development of new algorithms, the ability of machine learning has been greatly improved, especially in the increasingly mature situation of deep learning theory. Many companies provide advanced technology to the industry through cloud computing or open source methods, and package advanced algorithms into easy-to-use products, greatly promoting the development of AI technology.

**Big Data:** With the rapid development of mobile internet, massive amounts of data such as social media, mobile terminals, and inexpensive sensors are being quickly collected. With the continuous discovery of the value of data, various data management and analysis technologies are also constantly developing. In the field of AI, there are many machine learning algorithms, such as recognition of images, text, and speech, that require a large amount of data to be trained and continuously optimized. Nowadays, all of this is readily available, and big data has played a promoting role in the development of AI, laying a good foundation for its development.

**Computational power:** AI requires extremely high computational power. In the past, people's research on AI was often limited by the computing power of a single computer. In recent years, with the rapid development of cloud computing technology, its computing power has been greatly improved. Machine learning, especially deep learning, requires high computational resources, while cloud computing has a computing power of up to 1 trillion times per second. In addition, with the continuous progress of graphics processor technology, AI technology has also been greatly developed, and the use of multi core parallel methods can greatly improve the processing speed of AI. With the help of cloud computing technology, image processors can achieve large-scale and low-cost computing power. The key technologies in AI are shown in Fig. 3:



**Fig. 3.** Key technologies in AI

## 2.6 Algorithms in AI

**Genetic algorithm:** Genetic algorithm is an algorithm that imitates the law of survival of the fittest and survival of the fittest in nature and is developed to find the best solution. This algorithm starts from an initial population and restricts the chromosomes above the optimal solution gene. Each chromosome has a corresponding constraint characteristic value, which is manifested as the arrangement and combination of genes and determines the external characteristics of the individual. For example, the characteristics of black hair are determined by alleles on chromosomes, which control hair color. With the emergence of the first generation of groups, according to Darwin's theory of evolution, over time, previously unsuitable groups would continue to be eliminated and gradually produce more adaptable groups to new environments. In each generation, individuals who meet the conditions are selected as fathers, and the crossover and mutation of natural genes are simulated to generate new solutions. Finally, the best individual in the last group is decoded to become the near optimal solution of the problem.

**Ant colony algorithm:** Ant colony algorithm is a population optimization method that has its advantages and disadvantages compared to other methods. First of all, the ant colony algorithm introduces a positive feedback mechanism, so that it can continue to accelerate, so as to quickly get the optimization results. In this process, each individual can release its own pheromone and achieve communication between individuals. During the search process, multiple individuals can execute in parallel, greatly improving the computational power and execution efficiency of the algorithm. On this basis, a probability based search method is proposed to avoid errors caused by local minima during the search process.

## 2.7 Text Recognition Under AI

Convolutional based recurrent neural networks are widely used for text recognition in AI. This method first preprocesses the image for text recognition, and then proceeds in two steps: (1) Feature extraction is performed using convolutional neural network method. (2) Neural networks are an AI optical text recognition method that pre measures the position and value of text characters, enabling text recognition. Convolutional layers are widely used in image classification tasks due to their efficient feature extraction ability. Neural networks can be used to detect meaningful edges in images, thereby obtaining (at a higher level) shapes and composite images. For example, compared

to fully connected neural networks, convolutional neural networks can repeatedly use different pattern detection filters on images, thereby reducing the complexity of image recognition. Then, text recognition can be achieved by utilizing the relationships between characters. Circular networks have shown good performance in processing data with varying lengths. The most commonly used method is to use long-term short-term memory cells to solve asymptotic disappearance (asymptotic disappearance in a given weighting function, asymptotic disappearance).

## 2.8 Use of AI Technology

AI technology can grayize color images and process them accordingly. The process of turning an image into an image with only black and white colors is called binarization, which can be obtained through thresholding. Thresholding is the process from input image a to input image b. The calculation method for b is shown in formula 1:

$$b = \begin{cases} 1 & a \geq T \\ 0 & a < T \end{cases} \quad (1)$$

In formula 1, T represents the threshold. If the input image a is subjected to segmentation threshold processing, the inter class variance c obtained is shown in formula 2:

$$c = de^2 \quad (2)$$

In formula 2, d represents the proportion of foreground pixel count to the entire image, and e represents the proportion of background pixel count to the entire image. The maximum threshold f of variance c can be obtained by using the traversal method, and the calculation method of f is shown in formula 3:

$$f = \frac{\sqrt{b}}{c^2} \quad (3)$$

## 3 Simulation Experiment on the Application of AI Technology to Text Recognition Detection Algorithms

The use of AI in text recognition and detection is quite popular. This article conducted an experimental study on the improvement of text recognition and detection algorithms under AI technology, to verify whether the accuracy of text recognition and detection can be improved through AI algorithms. This article conducted an experimental survey on the satisfaction of 100 groups of users who used AI technology for text recognition and detection algorithms with the effectiveness of AI technology (out of 100 points). The satisfaction situation is shown in Fig. 4.

AI technology has better image processing capabilities and is expected to have good results in text recognition. From the experimental results in Fig. 4, it can be seen that users who use text recognition detection algorithms based on AI technology have a

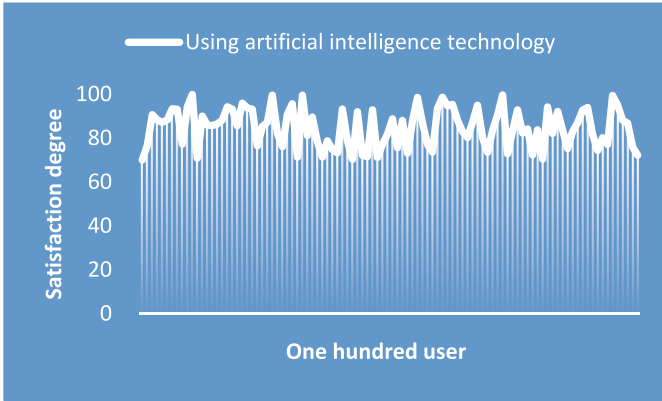


Fig. 4. Employee satisfaction of text recognition detection algorithm using AI technology

satisfaction score of 70 to 100 points with the algorithm. This indicates that users who use text recognition and detection algorithms based on AI technology are quite satisfied with the role of AI algorithms in text recognition and detection algorithms.

This article also tested the improvement of text recognition detection algorithm detection accuracy after using AI technology in five groups. The improvement of accuracy is shown in Fig. 5.

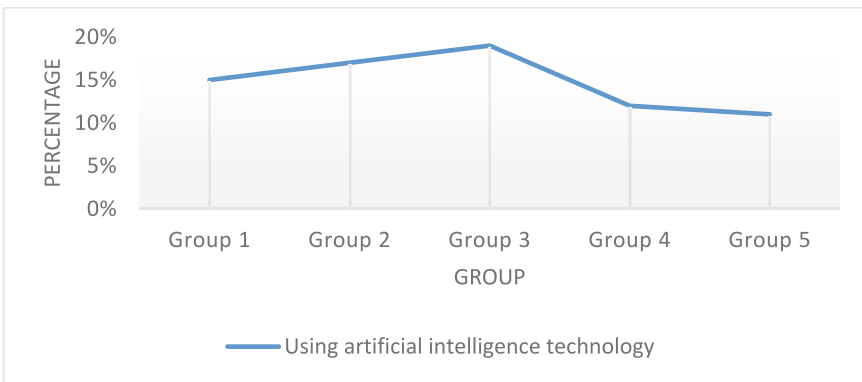


Fig. 5. The improvement of the accuracy of text recognition detection algorithm

From the experimental results in Fig. 5, it can be seen that the accuracy of the text recognition detection algorithm using AI technology has improved by a minimum of 11% and a maximum of 19%. From this experimental data, it can be seen that AI technology can improve the accuracy of text recognition detection algorithms, indicating that AI technology can achieve good results in text recognition detection algorithms.

## 4 Conclusions

Text recognition is a commonly used technology nowadays. Through text recognition, various types of text can be accurately read out, making it easier for people to obtain the text they need. When people need text that can be obtained by machines without manpower, it greatly reduces the workload of text processing. However, there are still many difficulties in text recognition. Because some texts have uneven or blurry areas, which affect the recognition effect, new technologies are needed to improve the quality of text recognition. This article studied the application of AI technology in text recognition detection algorithms, which aimed to improve the accuracy of text recognition detection algorithms through AI technology. This article tested the improvement of recognition accuracy of text recognition detection algorithms using AI technology through experiments, and found that the effect was good. Therefore, it has been proved that AI technology had a good effect in text recognition detection algorithms. Due to space limitations, there are still many shortcomings in the experiments conducted in this article, and further improvements are needed in the future. Finally, it is wished that the text recognition detection algorithm could be increasingly improved.

**Acknowledgements.** This work was supported by

*Key Research Project of Guangdong Baiyun College, No. 2022BYKYZ02.*

*Key Research Platform of Guangdong Province, no. 2022GCZX009*

Special project in key fields of colleges and universities in Guangdong province, No. 2020ZDZX3009.

## References

1. Chen, X., Jin, L., Zhu, Y.: Text recognition in the wild: a survey. *ACM Comput. Surv. (CSUR)* **54**(2), 1–35 (2021)
2. Lin, H., Yang, P., Zhang, F.: Review of scene text detection and recognition. *Arch. Comput. Methods Eng.* **27**(2), 433–454 (2020)
3. Petrova, O., Bulatov, K., Arlazarov, V.V.: Weighted combination of per-frame recognition results for text recognition in a video stream. *Компьютерная Оптика* **45**(1), 77–89 (2021)
4. Butt, H., Raza, M.R., Ramzan, M.J.: Attention-based CNN-RNN Arabic text recognition from natural scene images. *Forecasting* **3**(3), 520–540 (2021)
5. Bulatov, K., Razumnyi, N., Arlazarov, V.V.: On optimal stopping strategies for text recognition in a video stream as an application of a monotone sequential decision model. *Int. J. Doc. Anal. Recogn. (IJ DAR)* **22**(3), 303–314 (2019)
6. Francis, L.M., Sreenath, N.: Robust scene text recognition: using manifold regularized twin-support vector machine. *J. King Saud Univ.-Comput. Inf. Sci.* **34**(3), 589–604 (2022)
7. Sil, R., Roy, A., Dasmahapatra, M.: An intelligent approach for automated argument based legal text recognition and summarization using machine learning. *J. Intell. Fuzzy Syst.* **41**(5), 5457–5466 (2021)
8. Liao, M., Shi, B., Bai, X.: Textboxes++: a single-shot oriented scene text detector. *IEEE Trans. Image Process.* **27**(8), 3676–3690 (2018)
9. Kang, L., Riba, P., Rusinol, M.: Content and style aware generation of text-line images for handwriting recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**(12), 8846–8860 (2021)



10. Thompson, W.: Using handwritten text recognition (HTR) tools to transcribe historical multilingual lexica. *Scripta e-Scripta* **21**, 217–231 (2021)
11. Elleuch, M., Kherallah, M.: Boosting of deep convolutional architectures for Arabic handwriting recognition. *Int. J. Multimedia Data Eng. Manag. (IJMDEM)* **10**(4), 26–45 (2019)
12. Mintz, Y., Brodie, R.: Introduction to artificial intelligence in medicine. *Minim. Invasive Ther. Allied Technol.* **28**(2), 73–81 (2019)
13. Kaul, V., Enslin, S., Gross, S.A.: History of artificial intelligence in medicine. *Gastrointest. Endosc.* **92**(4), 807–812 (2020)
14. Gunning, D., Aha, D.: DARPA’s explainable artificial intelligence (XAI) program. *AI Mag.* **40**(2), 44–58 (2019)
15. Davenport, T.H., Ronanki, R.: Artificial intelligence for the real world. *Harv. Bus. Rev.* **96**(1), 108–116 (2018)
16. He, J., Baxter, S.L., Xu, J.: The practical implementation of artificial intelligence technologies in medicine. *Nat. Med.* **25**(1), 30–36 (2019)
17. Johnson, K.W., Torres Soto, J., Glicksberg, B.S.: Artificial intelligence in cardiology. *J. Am. Coll. Cardiol.* **71**(23), 2668–2679 (2018)
18. Haenlein, M., Kaplan, A.: A brief history of artificial intelligence: on the past, present, and future of artificial intelligence. *Calif. Manag. Rev.* **61**(4), 5–14 (2019)
19. Longoni, C., Bonezzi, A., Morewedge, C.K.: Resistance to medical artificial intelligence. *J. Cons. Res.* **46**(4), 629–650 (2019)
20. Hosny, A., Parmar, C., Quackenbush, J.: Artificial intelligence in radiology. *Nat. Rev. Cancer* **18**(8), 500–510 (2018)