Research on Artificial Intelligence Machine Learning Character Recognition Method Based on Feature Fusion

Wen Huang¹*, Haibiao Zhang ²

¹ School of Cyberspace Security, Chengdu University of Information Technology, Chengdu 610225, China
² School of Mathematics and Physics, Southwest Forestry University, Kunming, 650224, China

*Corresponding author e-mail: willfreelyhw@gmail.com

Abstract. With the continuous development of artificial intelligence technology, people's research on it has become more in-depth and machine learning has become a new research core issue in artificial intelligence. Before artificial intelligence technology was mature, it was a very difficult task to realize automatic character recognition and input. At present, human beings have made some progress in intelligent robots, voice recognition and network search. The text recognition method based on machine learning is of great significance to the promotion of information technology. Text recognition based on machine learning, although it has formed relatively rich theoretical support in principle and skill, but still faces various factors during actual practice, making its recognition effect difficult to be reflected. Based on the feature fusion perspective, this paper combines the background and main identification methods of text recognition in artificial intelligence machine learning to analyze the purpose, which lays a foundation for future specific applications.

1. Introduction
Before artificial intelligence technology was mature, it was a very difficult task to realize automatic character recognition and input. The development of science and technology and Internet technology has provided the basis for people to move towards intelligence, while machine learning is a very important component in the field of artificial intelligence, but its application still needs further expansion [1]. As an important component of artificial intelligence, machine learning has been gradually applied to various fields in recent years. Its intelligent characteristics make the substantial effects achieved in its application process extremely obvious [2]. Traditional character recognition methods are based on the intuitive morphological characteristics of characters. Through statistical analysis of the morphological differences between characters, a group of approximately optimal statistical parameters that can represent the differences of characters are found to screen and recognize characters, thus achieving the purpose of computer character recognition and automatic entry and preservation [3]. The development of science and technology and Internet technology has provided the basis for people to move towards intelligence, while machine learning is a very important component in the field of artificial intelligence, but its application still needs further expansion [4]. How to quickly and accurately
record these text messages into various electronic devices such as computers has become an urgent problem to be solved.

Text recognition based on machine learning, although it has formed relatively rich theoretical support in principle and skill, but still faces various factors during practical practice, making its recognition effect difficult to be reflected [5]. At this stage, humans have made some progress in intelligent robots, speech recognition, and network search [6]. Deep question-and-answer technology and autonomous driving technologies, which have received extensive attention, are based on artificial intelligence machine learning [7]. Before the artificial intelligence technology is still mature, it is a very difficult task to realize the automatic recognition and entry of words [8]. The traditional method of text recognition is based on the visual morphological features of the text. Through the statistical analysis of the morphological differences between the characters, a set of approximate optimal statistical parameters representing the difference of the characters is used to screen and identify the characters. Thereby achieving computer text recognition and automatic entry for preservation purposes [9]. Nowadays, robots, network search and speech recognition in artificial intelligence have made some progress, but text recognition has not been effectively applied, and more in-depth and comprehensive research is needed [10]. Based on the feature fusion perspective, this paper combines the background and main identification methods of text recognition in artificial intelligence machine learning to analyze the purpose, which lays a foundation for future specific applications.

2. Construction and Implementation of Text Recognition System Classifier

Automatic text recognition and entry is a very difficult but urgent problem. The traditional character recognition methods that we use in our life are usually based on their basic shapes, and the differences between characters are counted and analyzed. Then find an optimal set of statistical parameters that can represent the differences between characters to realize the screening and recognition of characters, and further realize the recognition and automatic entry of characters. Spatial distribution feature and stroke density feature are feature extraction methods based on ignoring a large number of character details and textures, which can produce good class spacing for characters with relatively simple structure. The core content of machine learning is how to construct the classifier and how to train the classifier to correctly classify a large number of unknown and known data. The reasoning process, as the intellectual essence of the corresponding machine learning, combined with the professional setting of the reasoning strategy, enables the relationship among the learner, data and knowledge adjustment of the character recognition system to be fully reflected intuitively. In the text recognition system based on machine learning method, the introduction of classifier is an important performance different from traditional text recognition methods. This enables researchers to focus more on how to obtain better discernable feature vector extraction methods instead of caring about the internal patterns of data.

The change of machine learning mode has led to the rapid growth of the generated Internet data at a speed of one million times, and with the increase of the degree of interconnection, the format and form of interactive data generated by online teaching are also constantly updated. The research on the reliability of network topology requires a comprehensive, complete and reasonable analysis from the two aspects of network devices and identification of load capacity. The algorithm flow is shown in Figure 1.
In general, the form of knowledge representation is determined by the algorithm selection of the machine learning system itself, and the learner of the same structure can also be applied to different fields. The law between the data can be learned from the classifier by providing the training data to the classifier. Such a character recognition system will have the advantages of strong adaptability and easy upgrade and extension. The learner input must be directly related to the corresponding system environment. The corresponding learner inducts and transforms the environmental data according to the internal learning algorithm, and updates the new information into the knowledge base operation. Machine learning algorithms have well-defined symbolic representations and algorithmic principles, with the ability to transform data into knowledge representation arrays [11]. Learning process is a complex intelligent activity closely connected with reasoning process. Machine learning can be classified according to learning strategy, knowledge description or application field. Symbolic representation and principle division of machine learning algorithm are carried out to clarify its characteristics of transforming data into knowledge for array representation. During this period, knowledge representation is directly related to classifier type and structure. Since the computer technology came into being, people have optimized the relevant methods, but the final result of character recognition is not satisfactory. For example, for some very simple English recognition, there are still many problems, which can not get a good recognition rate. For other languages with larger character sets, this method has a lower recognition rate.

3. Main Methods of Character Recognition in Artificial Intelligence Machine Learning

3.1. Analysis of Recognition Classifier

In order to analyze the method of character recognition in artificial intelligence machine learning, it is necessary to study the classifier of character recognition, and BP neural network human classifier is one of many classifier of character recognition. Conventional character recognition is to find out a group of similar statistical parameters that can represent the differences of characters by making corresponding statistical analysis on the morphological differences between characters according to the intuitive morphological characteristics of characters, so as to carry out comprehensive screening and recognition on them. In the process of cross-validation, the character recognition system will judge some characters,
and the ratio between the number of wrong characters and the total number of words is the error self-checking rate [12]. Text recognition of world scenes refers to the recognition of images with text meaning from images containing a large number of natural backgrounds. It is the most difficult object to recognize in text recognition systems. During the actual application, it is still difficult to obtain good recognition accuracy when it is applied to English with fewer characters. This phenomenon is mainly caused by the types of characters, the dithering of characters’ own pixels, and the long recognition time. Due to the relatively closed and centralized management of the data center network, it is drawn by different performance requirements in the actual deployment process, showing the coexistence of various network forms. In line with the premise of cointegration test. Check whether there is a long-term equilibrium relationship between relevant variables. The inspection results are shown in Table 1.

| Characteristic value | Trace estimation | Critical value |
|----------------------|------------------|----------------|
| 0.852                | 1117             | 139.7          |
| 0.725                | 956              | 141.2          |
| 0.953                | 1272             | 159.6          |
| 0.359                | 956              | 188.7          |

A single classifier is distinguished on the basis of the characteristics of each group, that is to say, if all characters are to be effectively recognized during the operation of the system, the classifier application process must be able to contain all the above characters. In the process of control input, all neuron activities can be effectively input, thus reducing waste among neurons and making the training speed faster.

3.2. Analysis of Character Recognition Performance

The character recognition system will make corresponding judgment on some characters in the process of cross-validation, and will also detect the wrong number of characters. The ratio between the wrong number of characters and the total number of characters is also the final error self-checking rate. Considering the symmetric demand of big data, when the network capacity is relatively large relative to the bandwidth demand, the symmetric traffic demand will be tree-shaped. When the traffic demand is lower than, the calculated total cost increases linearly, because the routes have not changed for different traffic demands. They are all tree-like routes, and only the reserved bandwidth of each link is changed. Corresponding parameters of different identification items are shown in Table 2. The network cost of character recognition is shown in Figure 2.

| Identification type | 1    | 2    | 3    | 4    | 5    |
|---------------------|------|------|------|------|------|
| TDR                 | 24.007 | 33.386 | 25.842 | 17.556 | 24.028 |
| Single Domain       | 35.318 | 34.277 | 22.875 | 21.374 | 25.315 |
| Hybrid              | 20.509 | 42.606 | 14.614 | 24.824 | 30.504 |
When setting up a text recognition method based on machine learning, the characteristics of classifier construction of machine learning itself should be made clear, and the efficiency of correctly analyzing and classifying a large number of unknown and known sight distances by aggregate classifiers should be ensured, so as to ensure that the text recognition system based on machine learning can be fully formed [13]. Even though artificial intelligence is superior to human beings in terms of computing ability, they are unlikely to have the kind of independent consciousness peculiar to human beings. When recognizing characters, the group features represented by different character images must be accepted by the whole system, so the classifier needs to have the common features possessed by each group. Handwritten recognition represents clear text meaning in each input area. However, printed text images need to be segmented by some means to obtain regions with clear text meaning. For the different results in the classifier, the features with the most differences are counted and the data represented by these features are filtered, and the final recognition result is the one with the most votes.

4. Conclusion
The emergence of artificial intelligence has brought about great changes to our life and production methods. In addition to the development of science and technology and network technology, people have successfully entered the information age. Human civilization has entered the information age, and all aspects of people's daily work and life are affected by information technology. Among many intelligent products, artificial intelligence machine learning has received extensive attention. With the continuous development of artificial intelligence technology, people's research on it has become more in-depth and machine learning has become a new research core issue in artificial intelligence. Among many intelligent products, artificial intelligence machine learning has received extensive attention, which is the core technology of automatic driving technology, and has important research value and broad development space. How to build a character recognition system with the characteristics of automatic layout analysis, strong fault tolerance, high recognition rate, error self-learning and self-correction, and easy expansion is the research goal of character recognition automation. Artificial intelligence machine has attracted people's attention due to its unique characteristics among many intelligent products, which is also an indispensable key technology in automatic driving technology, and its research has significant practical significance.
Reference

[1] Bartsch G, Mitra A P, Mitra S A, et al. Using Artificial Intelligence and Machine-Learning Algorithms with Gene Expression Profiling to Predict recurrence of Nonmuscle-Invasive Urothelial Carcinoma of the Bladder at Initial Presentation. [J]. Journal of Urology, 2016, 195(2):493-498.

[2] Norman M. Artificial Intelligence& Machine Learning in the Earth Sciences [J]. ACTA GEOLOGICA SINICA(English edition), 2019, 93(z1):48-51.

[3] Das S, Dey A, Pal A, et al. Applications of Artificial Intelligence in Machine Learning: Review and Prospect [J]. International Journal of Computer Applications, 2015, 115(9):31-41.

[4] Foulquier N, Redou P, Gal C L, et al. Pathogenesis-based treatments in primary Sjogren's syndrome using artificial intelligence and advanced machine learning techniques: a systematic literature review [J]. Human Vaccines and Immunotherapeutics, 2018, 14(3):1-18.

[5] Sinha Y, Jain P, Kasliwal N. Comparative Study of Preprocessing and Classification Methods in Character Recognition of Natural Scene Images [J]. Advances in Intelligent Systems and Computing, 2016, 390:119-129.

[6] Amin A. Prototyping structural description using an inductive learning program [J]. Systems Man & Cybernetics Part C Applications & Reviews IEEE Transactions on, 2015, 15(12):1103-1123.

[7] Hussain R, Gao H, Shaikh R A. Segmentation of connected characters in text-based CAPTCHAs for intelligent character recognition [J]. Multimedia Tools and Applications, 2016, 76(2):1-15.

[8] Price S, Flach P A. Computational support for academic peer review: A perspective from artificial intelligence [J]. Communications of the ACM, 2017, 60(3):70-79.

[9] Zhang X Y, Yin F, Zhang Y M, et al. Drawing and Recognizing Chinese Characters with Recurrent Neural Network [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, PP(99):849-862.

[10] Hausknecht M, Li W K, Mauk M, et al. Machine Learning Capabilities of a Simulated Cerebellum [J]. IEEE Transactions on Neural Networks and Learning Systems, 2016, 99(3):1-13.

[11] Goodfellow I, Mcdaniel P, Papernot N. Making machine learning robust against adversarial inputs [J]. Communications of the ACM, 2018, 61(7):56-66.

[12] Hoang N D, Pham A D. Hybrid Artificial Intelligence Approach Based on Metaheuristic and Machine Learning for Slope Stability Assessment: A Multinational Data Analysis [J]. Expert Systems with Applications, 2015, Accepted, In Press:60-68.

[13] William H. S. Antônio, Silva M D, Miani R S, et al. A Proposal of an Animal Detection System Using Machine Learning [J]. Applied Artificial Intelligence, 2019:1-14.