Research and Application of Digit Recognition Based on K-Nearest Neighbor Classifier

Bolin Lu

Daqing No.1 middle school, Daqing, Heilongjiang, 163256, China
agreyry52971@student.napavalley.edu

Abstract. The electrical or mechanical translation of pictures of typed, handwritten, or printed text into machine-encoded text is known as optical character recognition (OCR). Pattern recognition, artificial intelligence, and computer vision are all areas of study in OCR. Early iterations operated on one typeface at a time and required training with photos of each character. modern systems that bring high precision to many fonts have now become a natural choice, supporting various image file formats for input. However, little research has been done on programs that specifically recognize numbers. The goal of this article is to perform digit recognition based on the KNN (k-nearest neighbor classifier). For digit recognition based on the KNN classifier, one can use the KNN algorithm to classify each character in an image. The image is first preprocessed and then each digit is segmented. The features of each digit are then extracted and used as input to the KNN classifier. The dataset comes from the MNIST (Modified National Institute of Standards and Technology database) dataset.
Keywords: optical recognition, number, image

1 Introduction

Object ID is a computer technology for processing computer images and image processing, which identifies a specific target object (person, building, car, etc.) through digital images and videos. It is one of the most useful and commonly used computer image reflection and image imaging techniques. It is used in video surveillance and image fax systems The history of the object description file can be divided into two cycles Around 2014, before 2014, most object recognition technologies were built on These features are the background for practicing Classical Support (SDM) and distinguishing objects. which achieved real-time face detection, and deformable part models (DPMs), which modeled objects as flexible configurations of parts. After 2014, deep learning-based methods became dominant in object detection. Convolutional neural networks (CNNs) are utilized in these techniques to learn features. and classifiers end-to-end from data. Some of the milestone detectors in this period are R-CNN (Regions with CNN features) , which used region proposals to localize objects, YOLO (You Only Look Once), which used a single network to predict bounding boxes and class probabilities, and Faster R-CNN, which integrated region proposal network (RPN) and R-CNN into a unified framework. OCR has many real-world applications across different industries. The

© The Author(s) 2024
B. H. Ahmad (ed.), *Proceedings of the 2023 International Conference on Data Science, Advanced Algorithm and Intelligent Computing (DAI 2023)*, Advances in Intelligent Systems Research 180,
https://doi.org/10.2991/978-94-6463-370-2_5

history of OCR can be found throughout the late 19th and early 20th centuries when inventors such as Charles R. Carey, Edmund Fournier d'Albe, and Gustav Tauschek created a device that scans text and generates sound and signals for visually impaired individuals. In the 1930s and 1940s, Emanuel Goldberg and Vannevar Bush invented a machine that was able to read text and convert it into telegraph code or microfilm. In the 1950s and 1960s, IBM（International Business Machines Corporation）, RCA (Royal College of Art), Cinch, Vannevar Bush, and others launched the first commercial OCR system. In the 1970s and 1980s, OCR technology improved with the use of scanners, optical sensors, and digital computers. Ray Kurzweil developed the first Omni font OCR system that can recognize all text documents, regardless of font and style. He also manufactured a reading machine for people who are blind that uses a text-to-speech synthesizer. In the 1990s and early 21st century, OCR technology became more widespread and accessible through online services such as ADOBE Acrobat, and Web OCR. Due to the development of artificial intelligence, machine learning, and computer vision, OCR technology can also more accurately and strongly recognize complex texts such as handwritten documents, natural scenes, or multilingual texts.

Digit recognition is part of OCR. The advantage of using the KNN （k-nearest neighbor classifier) is that the algorithm is easy to understand, the accuracy is high for this problem, and the time complexity is low compared to most other algorithms.

Algorithms focusing on number recognition can take less time to process each digit than the universal character recognition algorithms, leading to less processing time for large data inputs.

2 Related research

2.1 A Survey of Character Recognition Technology

The technique of digitizing handwritten or printed text is known as character recognition. The main steps are preprocessing, segmentation, feature extraction, classification, and post-processing [1]. There are challenges and applications of character recognition in various domains, such as document analysis, biometrics, education, and entertainment [2]. There are many methods for character recognition, such as character recognition based on the concept of fuzzy sets, which are derived from the pixel values of the character image using fuzzy membership functions [3], and character recognition based on neural networks and genetic algorithms, which uses a feedforward neural network with one hidden layer to classify the character images, and a genetic algorithm to optimize the network weights and architecture [4]. an approach for sentiment analysis on Twitter data using hybrid feature selection and machine learning techniques. it proposes a feature selection method called the hybrid feature selection method (HFSM) that combines the advantages of both filter and wrapper methods [5]. The proposed method is then used to select the most relevant features from the Twitter data. The OCR can recognize handwritten documents such as invoices, payment drafts, bills, etc. And convert them into electric form [6].

2.2 Application of K Nearest Neighbor Classifier in Image Recognition

The KNN can be used in object detection for medical blob detection, based on the blurred blobs that resulted from a sequence of photos, a scale-space tree is built. The function in the building of a scale space tree is played by features and geographical data. Following the blob extraction procedure, users construct ground truth pictures by classifying the blobs that were identified inside the image. To illustrate its effectiveness, the KNN method is used inside the same process to separate the image's blob into several classes [7]. And can be used in social Image Segmentation. The proposed approach follows an allocating identical surrounding pixels to the same area iteratively and vice versa [8].

2.3 Research Status of Digital Recognition Based on k-Nearest Neighbors

The applications of nearest-neighbor classifier in OCR are attributed to the ease of defining the N value. By digitization, pre-processing, feature extraction, classification, and classification, the system could recognize the handwritten character [9]. It can be used as the confirmation algorithm, unlike other OCR systems, this system is based on the angles of the digits rather than on pixels. The flexibility to function in various lighting and exposure situations, the insensitivity to potential digit rotations, and the capacity to deduce and employ heuristics for character identification are some of the benefits of their suggested system [10].

3 Method

3.1 The principle and workflow of k nearest neighbor classifier

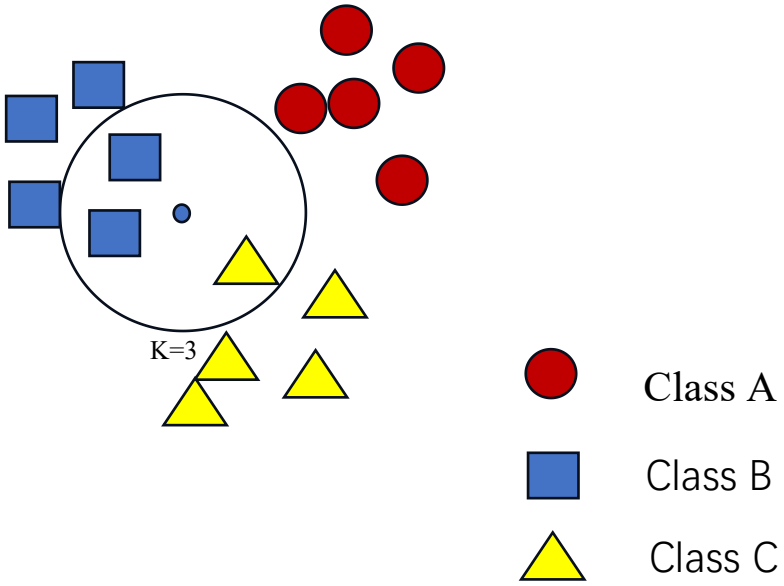

Fig. 1. The figure shows the principle of KNN algorithms (Picture credit :Original)

The KNN method involves 4 steps. The first step is to calculate the distance between each sample point in the training sample and the test sample (common distance measures such as Manhattan distance, Euclidean distance, Minkowski distance, and Chebyshev distance etc.), The second step is to sort all the distance values. The third step is to select the first k samples with the smallest distance. The last step is to vote according to the labels of the k samples to get the final classification category.

For example, as shown in Figure 1 when k=3, the sample found 3 nearest data, including 2 class B and 1 class C. So according to the number of each class included, the sample is considered as class B.

3.2 Data preprocessing and feature extraction

For the data, the dataset was made of a 28*28 grayscale image, which has a scale of 0-255 representing each pixel's brightness, and the image has a label showing the number it represents. The KNN algorithm could automatically extract the feature of each number.

4 Experiment

4.1 Dataset introduction and preparation

A sizable library of handwritten numbers called the MNIST database (Modified National Institute of Standards and Technology database), is frequently used to train and test different image-processing algorithms. The developer determined that the training data set was obtained from an NIST (National Institute of Standards and Technology) employee at the U.S. Census Bureau and that the test data set was obtained from an American high school student, which was not appropriate for the machine learning experiment. MNIST database is also frequently used in testing and training in the field of machine learning. The training set is made of 250 distinct people's handwritten digits, of whom 50% are students from high school and 50% are the people who work in the Census Bureau. The test set also includes handwritten digital data in the same proportion.

To deal with arrays, we utilize the NumPy (Numerical Python) Python module, which is one of the most used libraries in Python. Matrix operations, the Fourier transform, and functions for working with linear algebra are also included. NumPy was created by American data scientist and businessman Travis Oliphant in the year 2005. Because it is an open-source project, you can use it without any restrictions. The abbreviation for numerical Python is NumPy. The Python package A wide range of supervised and unsupervised machine learning methods are available through sklearn (Scikit-learn). It is built on SciPy, one of the most widely used libraries for machine learning. Matplotlib is an excellent Python visualization library to process 2D array displays. To deal with the bigger SciPy stack, a multi-platform data visualization toolkit named Matplotlib was developed and is based on NumPy arrays. In the year 2002, John Hunter uttered the remark for the first time.

4.2 Experimental setup and parameter selection

The n value is being tested and compared to get the most accurate value in the range from 1 to 15. Cpu (Central Processing Unit) : Intel(R) Core (TM) 11th Gen i7-11800H @ 2.30GHz 2.30 GHz

4.3 Experimental results and conclusion

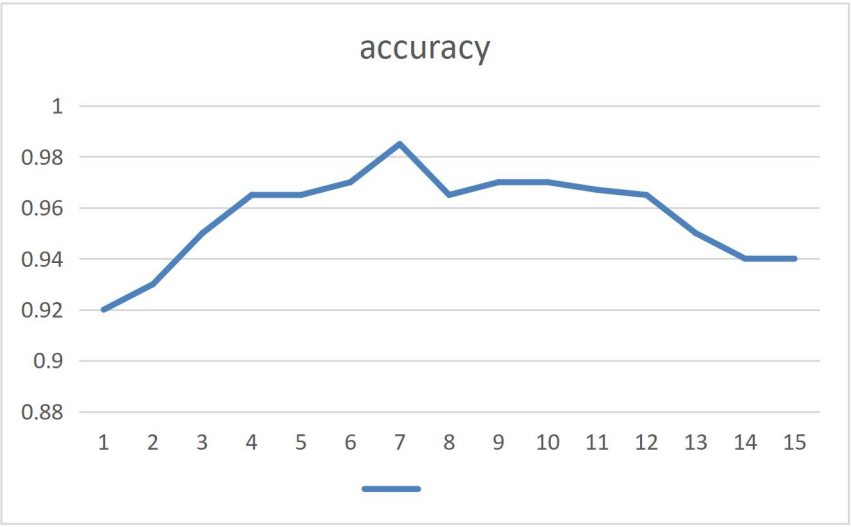

Fig. 2. The figure shows the accuracy with different n value (Picture credit :Original)

The figure shown above is the accuracy of different n values, the accuracy is calculated by testing 5 times and getting the average accuracy of each n value. When n value is below 7, the accuracy tends to increase with the increase of n, and the accuracy is highest when n value is 7. When the n value is above 7, the value tends to decrease and stay stable at about 0.94.

5 Conclusion

In the research, we apply the KNN method to the program use the MNIST dataset as a train set and test set, and let the program be able to recognize 28*28 handwritten images. Based on the sklearn library in Python, the program could get the accuracy of recognition from different n values and determined the best n value for number recognition, which is 7 and the accuracy is 0.986.

The limitation of the program is that when the model is applied, the program is calculated by CPU instead of GPU (Graphics Processing Unit), which slows down the program a lot. The program could only recognize the 28*28 image, so it could not apply to real-life circumstances. The program also does not perform well in an environment with noise. In the future, the program should be able to use GPU to calculate, and convert other types of images into 28*28 grayscale images, and obtain more accuracy in both environments with noise and environments without noise.

For application, number recognition could be used in circumstances with a lot of data to process, such as detection and statistics of a vehicle's license plate, banks for processing checks, and post offices for scheduling letters, floating population registration, or recognizing the handwritten letters.

References

1. Govindan V K, Shivaprasad A P. Character recognition—a review[J]. Pattern recognition, 23(7): 671-683 (1990).
2. Mithe, R., Indalkar, S., & Divekar, N. Optical character recognition. International journal of recent technology and engineering (IJRTE), 2(1), 72-75 (2013).
3. Rossant, F., & Bloch, I. A fuzzy model for optical recognition of musical scores. Fuzzy sets and systems, 141(2), 165-201 (2004).
4. Trier, Ø. D., Jain, A. K., & Taxt, T. Feature extraction methods for character recognition-a survey. *Pattern recognition*, *29*(4), 641-662 (1996).
5. Singh, A., Bacchuwar, K., & Bhasin, A. A survey of OCR applications. International Journal of Machine Learning and Computing, 2(3), 314 (2012).
6. Stoliński, S., & Bieniecki, W. Application of OCR systems to processing and digitization of paper documents. Information Systems in Management VIII, 102 (2011).
7. Inthajak, K., Duanggate, C., Uyyanonvara, B., Makhanov, S. S., & Barman, S. Medical image blob detection with feature stability and KNN classification. In 2011 Eighth International Joint Conference on Computer Science and Software Engineering (JCSSE) (pp. 128-131). IEEE (2011).
8. Wazarkar, S., Keshavamurthy, B. N., & Hussain, A. Region-based segmentation of social images using soft KNN algorithm. Procedia Computer Science, 125, 93-98 (2018).
9. Kumar, M., Jindal, M. K., & Sharma, R. K. k-nearest neighbor based offline handwritten Gurmukhi character recognition. In 2011 international conference on image information processing (pp. 1-4). IEEE (2011).
10. Matei, O., Pop, P. C., & Vălean, H. Optical character recognition in real environments using neural networks and k-nearest neighbor. Applied intelligence, 39, 739-748 (2013).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.