ISSN: 2664-9640

Test Paper Score Scanning Input and Analysis System Based on Computer Vision

Peixin Ren

Beijing Wuzi University, Beijing 101100, China;

renpeixin0924@icloud.com

Abstract

With the increasing number of students, the workload of teachers is also increasing. After the examination, the traditional manual input of a large number of examination results into the system in turn has become one of the burdens of teachers. Therefore, this paper intends to build a system that can automatically input the test results into the score system. This paper mainly uses the computer vision image detection technology in machine learning to enable the computer to identify the scores on the examination paper and the information of the examinees, and finally upload the relevant information to the system correctly. Compared with the traditional manual input, the implementation of this system can greatly reduce the burden and time cost of teachers.

Keywords

Computer vision, machine learning, image detection technology, score entry system, artificial intelligence.

1. Purpose of system design

The main purpose of this system design is to reduce the pressure of teachers in the university campus to log in the test scores to the system. The student number and score corresponding to the test paper are automatically written into the system through automatic identification. On many university campuses, because of the relatively small number of students, from an economic point of view, it is more appropriate to use human rather than answer sheets and card readers when scoring test papers. However, it is inefficient for teachers to input scores into the system manually in turn. Therefore, the method of automatically identifying student numbers and scores and transmitting them to the system is considered for optimization.

2. Meaning of system realization

After the implementation of the system, we can use the external shooting equipment to capture the corresponding position on the test paper, and send the captured pictures to the system, and finally we can automatically input the score of the test paper and the corresponding students into the transcript through the system. Compared with the traditional input method, the speed and the time cost of the automatic input system will be greatly reduced. And the relevant experiments show that in the same number of papers, compared with the automatic detection, the time and accuracy of automatic detection are more stable, while the time of manual detection will be longer and the accuracy will begin to be lower when the number of papers is gradually increased. When the detection time becomes longer, the human eye will become tired, so the efficiency and accuracy of artificial will be reduced. Therefore, it is necessary to realize the automatic entry of scores under the requirement of high efficiency, which has great practical application value and broad market prospects.[1]

ISSN: 2664-9640

3. System design

For the design of the system, we mainly use the python programming language of version 3.9 under the pytorch development environment and its third-party function library to complete. The difficulty of the whole system is how to get the information and scores of the examinees through the test paper pictures. For this difficulty, we use image detection technology in machine learning to complete it. The specific ideas of the whole system are as follows:

Acquire an image NEed to be recognized

We print out a score table on the test paper in advance (the size of each grid in the table is equally divided), the content of the table is the score of each question on the test paper (filled in by the teacher, each grid contains only one number, every two grids represent a score, here we default that the score of each question is between 0 and 99. The single digit is filled with 0 as a two-digit number) and the examinee's examination number (all the contents filled in can be handwritten, but no modification is allowed). By using the OpenCV third-party library in python to call our external camera device, we can intercept this part of the test paper to form an image and divide the intercepted image evenly to obtain a single digital image for recognition. It stipulates that the first few digits are the student number, and the following pictures are scored for each question[2]by digital splicing of every two images. Finally, the OS file management library in python is also used to put all the photos from each test paper into a folder.

Recognizing score image

How to identify the score in the image. We adopt the following two schemes:

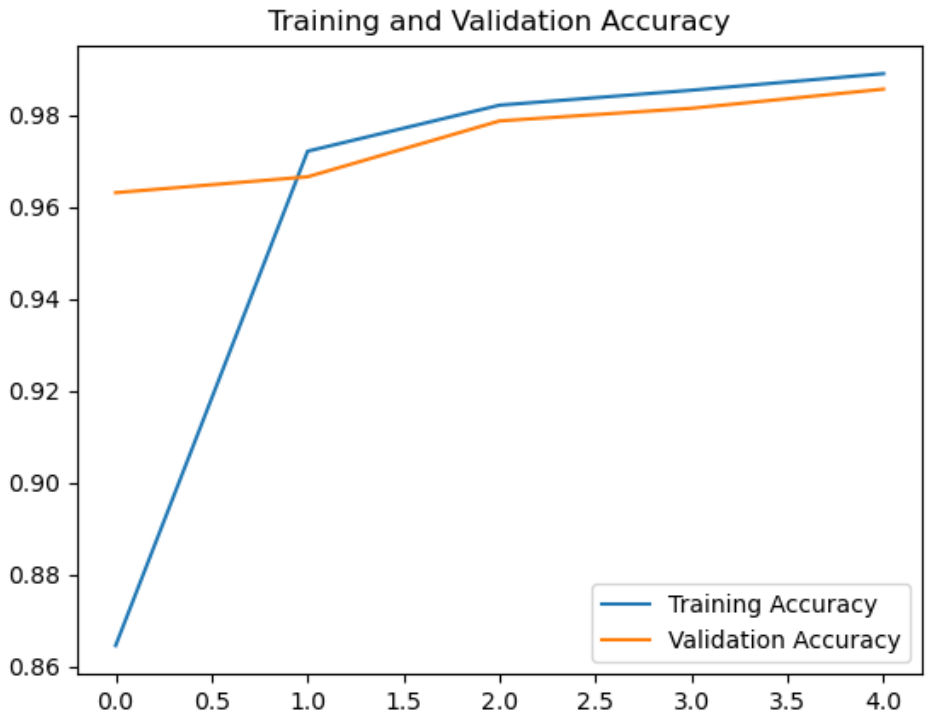
```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 28, 28, 32)        320

max_pooling2d (MaxPooling2D  (None, 14, 14, 32)        0
)

conv2d_1 (Conv2D)            (None, 14, 14, 64)        18496

max_pooling2d_1 (MaxPooling  (None, 7, 7, 64)          0
2D)

conv2d_2 (Conv2D)            (None, 7, 7, 64)          36928

max_pooling2d_2 (MaxPooling  (None, 3, 3, 64)          0
2D)

conv2d_3 (Conv2D)            (None, 3, 3, 32)          18464

max_pooling2d_3 (MaxPooling  (None, 1, 1, 32)          0
2D)

flatten (Flatten)            (None, 32)                0

dense (Dense)                (None, 128)               4224

dense_1 (Dense)              (None, 128)               16512

dense_2 (Dense)              (None, 64)                8256

dense_3 (Dense)              (None, 10)                650

=================================================================
Total params: 103,850
Trainable params: 103,850
Non-trainable params: 0
_________________________________________________________________
```

1. In the first scheme, we design a convolutional neural network based on the sequential model in the TensorFlow framework. The network structure is shown in the following figure:

We first use four convolution layers to extract the features of the image, and then use the fully connected network to classify the image from 0 to 9. As for how to train the network, the data

ISSN: 2664-9640

set we choose directly is the mnist handwritten digit data set. The data set is an image data set of handwritten digits, and the handwritten digits from a large number of different people are counted, wherein the training set has 60000 images, and the test set has 10000 images. So this data set is very close to the number of pictures we need to detect. However, since the image format of the handwritten data set is a handwritten digital image with gray values of 28 * 28 pixel points from 0 to 9, with white characters on a black background, the image pixel value is from 0 to 255, and the larger the point is, the whiter it is. This is different from the RGB image we passed in, so before passing it into the network, we need to change the format of the incoming image to a grayscale image, set a threshold according to it, and flip its pixel value according to the set threshold to make it similar to the image of the mnist data set. In order to achieve the best results, we can also convert the image, image scaling, rotation, filtering, binarization, edge extraction and other processing. Finally, the processed image file is sent to our algorithm model to identify the numbers in the image. We use the PIL third-party library in python to implement the modification of the image format. The following figure shows the accuracy trend of the training set and validation set for model training:



As shown in the figure above, our accuracy in training the model can finally reach 98%, so we can consider that the model is feasible. For the input of the network, we can directly input the picture folder, which can better achieve end-to-end and improve the efficiency of the whole system.

2. In the second scheme, we recognize the number of pictures by calculating the similarity between pictures. First of all, we write ten numbers from 0 to 9 on the answer sheet in advance, and then we will test the similarity between the pictures and the numbers written in advance. We think that which digital picture is more similar is the number. We also use python language to calculate the similarity, trying to calculate the cosine similarity (representing the picture as a vector, representing the similarity of two pictures by calculating the cosine distance between the vectors) and using SSIM (Structural Similarity Measure) to calculate the similarity of the picture (a full reference image quality evaluation index). Image similarity is measured from

ISSN: 2664-9640

three aspects of brightness, contrast and structure. In the process of practical application, we determine the final detection results by combining the size of the two similarities.

Obtain the results and enter them into the system

Using the first scheme, we trained the model to predict our incoming score image. We pass the score image folder of each test paper into the model. Because the model we designed is a multi-classification model, the final result is shown in the form of probability, so we need to determine which digit of 0-9 in the picture according to the probability predicted by the model. You can get the student's test number and score in turn according to the order. Finally, write the test number and the corresponding score into the excel form.

If we use the second scheme, we can directly get the recognition result by calculating the similarity of the picture, and then get the score of the student's test number through the splicing method, and after getting the student number, we will compare the student number with the student number in the database again, if there is a problem, we take the high similarity in the database as the correct student number. Finally, the test results are also written into excel tables through python.

Using the data analysis function of Excel, students' scores can be further analyzed to achieve the purpose that teachers can better understand students' scores. For this function, we also use the openpyxl library in python to detect the data and write it into the local Excel table.

Analysis of results

After writing the scores of each question on the test paper into the Excel table, we can directly use excel to match the student number with the score of each question in the test paper corresponding to the student number. We can not only understand the total score of the students' papers, but also understand the score of each question in the students' papers. For teachers, this is a great help to understand the students' performance. They can clearly see the students' shortcomings and train them.

Overall flow chart of the system

4. System analysis and improvement

The author completes this system function is mainly uses the pytorch development environment python language programming to complete, uses in the python multitudinous formidable third party library to be possible simple effectively to complete us to need the function, the external camera equipment then carries on the connection with the USB connection with python's OpenCV library to carry on the call. In the actual application, the whole system still has the shortcomings of low accuracy and long acquisition time. The detection algorithm in the system can be improved as follows:

1. In this system, we use the external camera to obtain the score area of the test paper by using ordinary screenshots, but we still need to obtain its location information manually. We can

ISSN: 2664-9640

actually print the form on the test paper with a specific label around it, and use computer infrared scanning to automatically get a screenshot of the form area. Or we can also use the machine learning method to directly obtain the test area we want by training the image depth network.

2. In this system, we use the foreign mnist handwritten digit training set to train the algorithm. Because of the format of the picture, we can't change the picture we detected into the format of the mnist data set completely. Even if the conversion is successful, the picture effect is not very good, and the picture may be distorted. Unable to achieve the image effect with the training set. And because it is a foreign statistical picture, handwriting habits may be a little different from the numbers we write, these reasons may lead to inaccurate recognition results. Direction of improvement.

3. In this system, we use a common convolutional neural network model to detect images, which can only detect a single target, but in fact, we need to detect multiple numbers in our test paper, so we can use the popular target detection algorithm in recent years to directly detect the whole test paper and find the target we want to learn. For example, if we can integrate the YOLO algorithm series into this system, we can directly obtain the student number of the examinee and the score of the test paper, and even achieve the goal of identifying the test paper in real time, so that we can greatly reduce our workload and time cost.

4. Because of the variety of people's handwriting habits and the different thickness of pen and ink, our test results are inaccurate. So if we want to improve the accuracy of the algorithm, it is necessary to improve the generalization of the algorithm. The simplest way is to add dropout layer and normalization layer to the neural network. Or we can also use some classic feature acquisition deep neural network models, with previous experience may be able to better improve the accuracy of our detection.Supported by: Beijing Wuzi University Students' Innovation and Entrepreneurship Program

Acknowledgements

Supported by: Beijing Wuzi University Students'Innovation and Entrepreneurship Program Training Program(2022103022).

References

[1] Biao Linxuan. Design of an Intelligent Test Paper Score Recognition System Based on Machine Vision [J]. Electronic Production , 2019(10):3.

[2] Kan Biao, Gong Kejian, Wang Ye, et al. The invention relates to a test paper and a method for rapidly inputting test paper scores based on a machine vision technology. :, CN112215192A[P]. 2021.