

Research Article

Preschool Education Interactive System Based on Smart Sensor Image Recognition

Xin Chen  and Guowen Jin

College of Sports, Suzhou University, Suzhou, 234000 Anhui, China

Correspondence should be addressed to Xin Chen; xinchen@lpubatangas.edu.ph

Received 16 December 2021; Revised 19 January 2022; Accepted 10 February 2022; Published 3 March 2022

Academic Editor: Deepak Kumar Jain

Copyright © 2022 Xin Chen and Guowen Jin. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The purpose of this research is to use the image recognition technology of smart sensors to establish a preschool education system for preschool children, which can realize interactive image-based learning and education for preschool children. This paper uses CMOS image sensor (IV2-6500CA of KEYENCE) to design the front-end image acquisition hardware system and establishes an image recognition system that can realize most of the recognition functions based on BP neural network algorithm and ImageNet, MS COCO, MNIST, and the Chars74K dataset and integrate the image recognition system into the preschool education interactive system. The image recognition system has a high accuracy rate, with an overall accuracy rate of 85.16%. Compared with the traditional preschool education system, it has a higher recognition rate, better teaching efficiency, and interactivity. It can recognize most of the objects that children touch and have a good interactive education effect.

1. Introduction

Online education is a new type of education and education mode under the integration of education and the Internet. This mode is mainly oriented to intelligent interaction and services. Traditional home children's literacy is mainly taught to children by teachers with the help of reference books, large-character cards, teaching aid maps, and other tools. This model is relatively dull and not interactive. With the popularization of electronic products, some teaching electronic devices emerge in an endless stream, so that children can easily learn knowledge in the process of entertaining and entertaining. The existing preschool education system still has many technical deficiencies, most of which are relatively backward. In image recognition, the two-dimensional code recognition technology is more used, and there is no database intelligent processing support. It is not

possible to shoot and identify real objects at will, and it is impossible to solve the questions that children encounter in daily exploration in real-time, accurately, and effectively. When encountering new and unknown things, they can only ask their parents for help. Therefore, the development of children's learning interactive systems with better performance and stronger functions is particularly important.

In recent years, because CMOS is superior to CDD in performance, researches on various applications of CMOS have emerged one after another. Researcher Assefa et al. introduced a small waveguide integrated germanium-on-insulator (GOI) photodetector with a capacitance of 10 ± 2 fF and an operating rate of 40 Gbps. During the source-drain implantation activation annealing process, the monolithic integration of thin single crystal germanium in the front-end CMOS stack is realized through rapid melt growth [1]. Bourdel et al. use ST Microelectronics CMOS 0.13 m

process to design a fully integrated ultra-wideband (UWB) pulse generator suitable for the Federal Communications Commission (FCC) 3.1-10.6 GHz frequency band. This generator is reserved for medium-speed applications and realizes pulses for on-off key control (OOK) modulation, pulse position modulation, or pulse interval modulation [2]. To improve the energy efficiency of CMOS, Bol has significantly improved the robustness and timing closure of ultralow voltage circuits in 65/45 nm CMOS by reducing the noise margin. This process not only satisfies a reasonable manufacturing output but also reduces energy efficiency in the case of long-term standby [3]. Although researchers have made a lot of progress in the improvement of CMOS energy efficiency, the research on improving the imaging quality of CMOS sensors still needs to be further improved. In power-constrained visual image recognition tasks based on IoT devices, the energy constraints due to the end of Dennard scaling limit the performance of neural network (NN) algorithms on popular digital platforms and cannot meet the energy efficiency requirements of embedded AI applications. To this end, Chen et al. proposed a CMOS image sensor based on the convolution kernel read-out method of mixed signal domain near-field processing structure. In this method, the visual data is collected from the smart CIS, which can realize the maximum kernel readout and the minimum sliding step of the convolution operation [4]. Image recognition technology has applications in all walks of life. For example, Jin et al. designed an intelligent vegetable cultivation device based on APP control, Internet communication, and image recognition technology. The device has functions such as remote control, precise seeding, quantitative dosing of liquid materials, and weed identification. The device mainly includes a tilling execution part, an image processing part, an STM32 microcontroller, and an application program that sends commands and controls the device to perform the corresponding tilling work. This research improves the intelligent and intensive level of vegetable cultivation, reduces the waste of production resources, and realizes the quantitative transportation of liquid materials [5]. Most image recognition is based on static images, and the research on dynamic image recognition is more difficult in the field of human motion recognition. Kim and Yoon proposed an effective feature extraction method in the existing human motion recognition technology, called adaptive local binary pattern (ALBP), for applications based on depth images. Compared with the traditional local binary mode (LBP), this algorithm can not only extract the textureless shape information in the depth image but also has a constant distance in the distance image [6]. Elngar and Kayed also apply face recognition in image recognition technology to vehicle intelligence. To prevent crimes such as vehicle theft, he proposed a development plan for vehicle protection and alarm systems based on Internet of Things technology and biometric authentication, which uses Pi cameras and PIR sensors. If the developed system detects an unauthorized person in the vehicle, the system will send its image and the location of the vehicle when it was stolen or damaged to the owner or the police via the Internet. Tests show that the accuracy of the system is 98.2% [7]. It can be seen from

the related work listed that, although the performance of CMOS smart sensors for image recognition is superior to that of CDD sensors, there are still some minor shortcomings, for example, when processing more complex images, it is easy to overload and heat, and the image quality is not as good as CDD. However, when it is applied to the preschool education system studied in this article, these small shortcomings do not affect the results.

This article creatively applies the intelligent sensor image recognition to the preschool education interactive system. Taking advantage of the CMOS image sensor's low energy consumption and fast readout speed, combined with deep learning algorithms, a dedicated interactive learning database composed of multiple image databases suitable for preschool children is established. Through the computer to recognize and analyze the image transmitted by the sensor, the system can help children carry out preschool education including the recognition of numbers, letters, and static objects. It is technically superior to the existing preschool education systems on the market, can achieve the purpose of learning in daily life, and help children better understand the surrounding world [8].

2. A Method for Establishing an Interactive System for Preschool Education Based on Smart Sensor Image Recognition Technology

2.1. Image Technology

2.1.1. The Basic Framework of Image Recognition. In early childhood education, images are more important. What we need to use are the recognition of words, letters, and numbers, as well as the recognition of some objects, including animals, plants, furniture, and toys [9].

Image recognition is a technology that analyzes the input image through the medium of a computer to distinguish it. The image recognition process is shown in Figure 1.

The basic framework of the image recognition system is shown in Figure 2, which is mainly composed of four parts: image enhancement, image segmentation, image description, and image classification. It combines the image recognition process, processes, and transmits images through hardware and software systems, and completes the process of recognition through feature extraction and classification and retrieval of the database.

2.1.2. Grayscale of Color Pictures. The pictures obtained by taking pictures are generally color pictures in RGB format. To recognize the text on the picture, the picture needs to be grayed out [10].

Each color in the RGB color space is composed of three components: R (red), G (green), and B (blue), each component is divided into 255 values, so the value range of each pixel point is 255 cubic. When the three components of R, G, and B take the same value (i.e., $R = G = B$), it is a grayscale image. Generally speaking, grayscale binarization is the first step in image processing. There are three main methods of image gray-scale processing: maximum value method, average value method, and weighted average method:

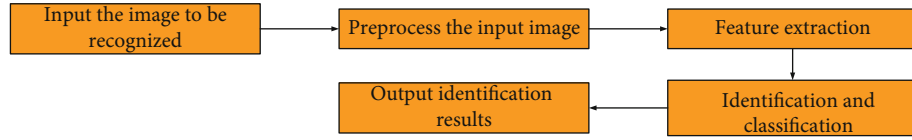


FIGURE 1: Image recognition process.

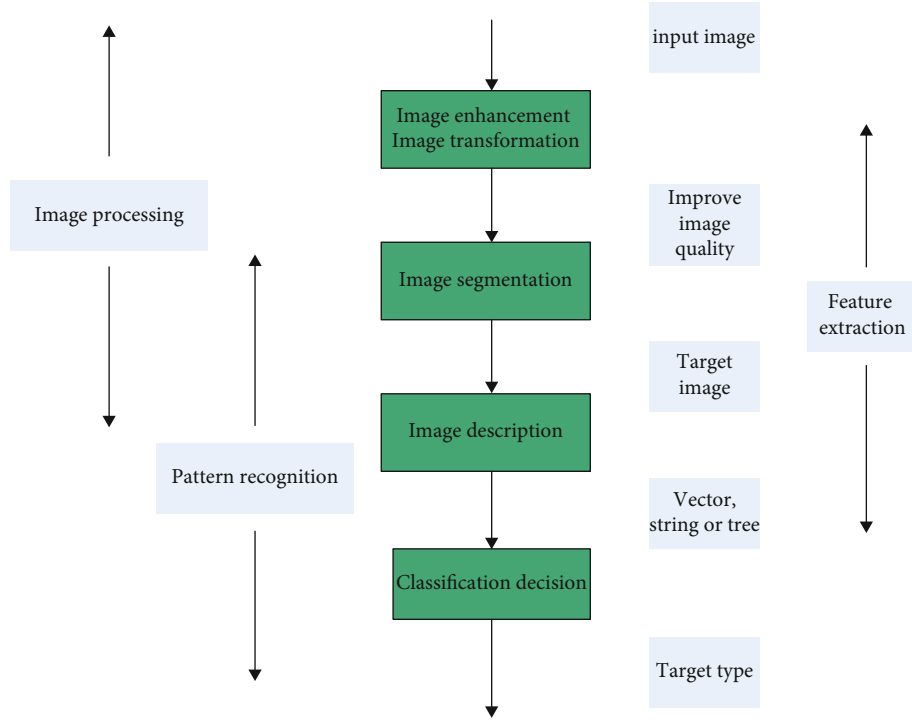


FIGURE 2: The basic frame diagram of the image recognition system.

- (1) The maximum method can be expressed by the following equation

$$F(i, j) = \max (R(i, j), G(i, j), B(i, j)). \quad (1)$$

- (2) The average value method can be expressed by the following equation

$$F(i, j) = \frac{(R(i, j) + G(i, j) + B(i, j))}{3}. \quad (2)$$

- (3) The weighted average method can be expressed by the following equation

$$F(i, j) = k * R(i, j) + m * G(i, j) + n * B(i, j). \quad (3)$$

In the three formulas (1), (2), and (3), (i, j) represents the coordinates of a certain pixel, and $F(i, j)$ represents the gray value of a certain pixel. In most cases, the grayscale effect of the weighted average method is better.

2.1.3. Quantification of Color Similarity. Image feature refers to the effective information that can characterize the mode of an image. Generally, after extracting the information, the information should be quantized into one or a group of (vector) data [11]. As a main feature of an image, color features are relatively more convenient to extract and quantify. However, in the identification and classification of patterns, the similarity of the two colors is often compared and used as the basis for discrimination and classification. Therefore, the choice of a good color similarity quantification algorithm is essential for image recognition. The color similarity quantification algorithm not only has a great influence on the result of image recognition but also has a great influence on the complexity and execution efficiency of the entire recognition algorithm. Here are some commonly used color similarity quantification algorithms:

(1) Minkowski distance

$$D(P, Q) = \left(\sum_i |H_i(P) - H_i(Q)|^R \right)^{1/R}. \quad (4)$$

If the feature vectors representing colors are mutually opposed, and these color feature vectors are equally important for pattern recognition, then, Minkowski distance can be used to quantify the similarity of two colors.

(2) Quadratic distance

$$D(P, Q) = \sqrt{(H_P - H_Q)A(H_P - H_Q)}. \quad (5)$$

Because the vectors that characterize color features generally have some correlation, there is a quadratic distance to quantify color similarity. In the above formula, A is a symmetric matrix that represents the correlation between feature vectors.

(3) Mahalanobis distance

$$D(P, Q) = \sqrt{(H_P - H_Q)C^{-1}(H_P - H_Q)}. \quad (6)$$

The principle of Mahalanobis distance is similar to quadratic distance, and the similarity between color feature vectors is also considered. It just replaces the matrix A with the negative first power of the covariance matrix C , which is composed of the covariance value C_i between each eigenvector. In the design of the pattern recognition algorithm, the appropriate color similarity quantification method should be selected according to the characteristic pattern of the image to be recognized.

2.1.4. Image Sensor. Smart sensors are widely used in home appliances, automotive industry, aerospace, machinery, chemistry, pharmaceuticals, and other fields. With the rise of emerging industries such as the Internet and mobile Internet, smart sensors including smart agriculture, smart transportation, health care, smart clothing, and other fields have a wide range of applications [12].

(1) CCD image smart sensor

CCD is called charge-coupled device. The inside of the CCD is composed of many photosensitive pixels. Each pixel is a photodiode, which detects the charge generated on the pixel. The amount of signal charge generated is directly proportional to the intensity of the incident light and the exposure time [13]. The working principle of CCD is shown in Figure 3.

(2) CMOS image smart sensor

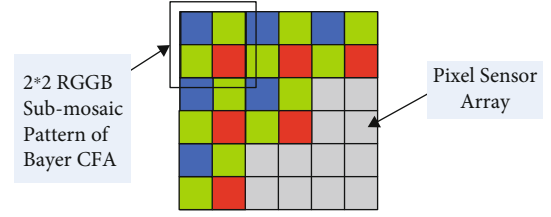


FIGURE 3: CCD working principle diagram.

CMOS, the scientific name, is complementary metal-oxide-semiconductor. The manufacturing technology of CMOS is the same as that of general computer chips. It is mainly composed of semiconductors composed of silicon and germanium. N (negatively charged) and P (positively charged) semiconductors coexist in CMOS [14]. Figure 4 shows the working principle. If external light illuminates the pixel array of the CMOS sensor, a photoelectric effect will be generated, and the corresponding charges will be generated between the pixel units. The row selection logic unit selects the corresponding row pixel unit, and the image signal level in the row pixel unit is sent to the corresponding analog signal processing unit via the signal bus of the corresponding column, converts the digital signal into a digital signal analog converter, and finally outputs it through the output interface.

2.1.5. Image Database. For image recognition, a detailed database is required to extract the image features obtained and match them with the image features in the database to achieve the purpose of recognition. Due to the larger, multi-dimensional, and diverse image data, the larger the database, the higher the accuracy of image recognition, which is essential for image recognition [15].

The database information table is shown in Table 1.

The four databases in Table 1 are different and have their own advantages. To achieve a more comprehensive picture recognition function, the four databases are selected to be combined to achieve complementary effects.

The four typical image data sets will be introduced below:

(1) ImageNet

ImageNet is a computer vision system recognition project and is currently the world's largest image recognition database. It was established by computer scientists at Stanford in the United States to simulate human recognition systems (as shown in Figure 5). It can recognize objects from pictures, and it currently contains 14197122 images, which is the largest known image database [16].

(2) MS COCO (Microsoft Common Objects in Context)

COCO is a data set that mainly targets scene interpretation and performs position calibration by accurately segmenting targets. This data set is mainly intercepted from complex daily scenes. It originated from the Microsoft COCO dataset which was funded and annotated by

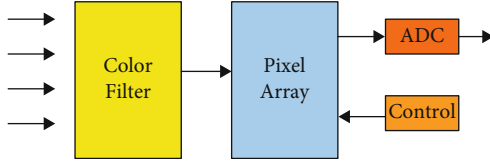


FIGURE 4: Schematic diagram of the working principle of CMOS.

Microsoft in 2014. The image includes 91 types of targets, 328,000 images, and 2,500,000 labels, and the number of individuals in the entire data set exceeds 1.5 million [17]. The way COCO recognizes objects is shown in Figure 6.

It can be seen from Figure 6 that when multiple objects in the picture are recognized, COCO first classifies the target objects, then marks the targets, classifies the same targets, and finally separates the marked objects from the surrounding background.

- (3) MNIST (Mixed National Institute of Standards and Technology)

MNIST is a large handwritten digit database, which is widely used for training and testing in the field of machine learning, compiled by a scholar from New York University. MNIST contains 60,000 training sets and 10,000 test sets. Each image has been scaled, normalized, and digitally centered. The fixed size is $28 * 28$ (as shown in Figure 7) [18].

- (4) The Chars74K dataset

The Chars74K dataset is very classic in the data set. It is a character recognition data set, which mainly includes English characters and Kannada characters. Because this data set has a total of 74000 images, it is also called Chars74K [19].

In English, Latin letters (excluding accents) and Hindu Arabic numerals are used. For simplicity, we call it the "English" character set. The data set includes 64 categories (0-9, A-Z, A-Z), 7705 characters obtained from natural images, 3410 hand-drawn characters using a tablet PC, and 62992 characters synthesized from computer fonts, which contains as shown in Figure 8.

2.2. Image Recognition and Classification Based on BP Neural Network. The BP neural network has three or more multilayer neural networks, and each layer is composed of several neurons. As a widely used neural network learning algorithm, it is more mature in all aspects [20].

The action function of BP neural network usually adopts S-type function (Sigmoid function):

$$y = f(WX + \theta) = \frac{1}{1 + e^{-(WX + \theta)}}. \quad (7)$$

The representation methods of threshold function and linear function are as follows:

Threshold function:

$$y = f(WX + \theta) = \begin{cases} 1 & WX + \theta \geq 0 \\ 0 & WX + \theta < 0 \end{cases}. \quad (8)$$

Linear function:

$$y = f(WX + \theta) = WX + \theta. \quad (9)$$

BP network training algorithm.

The steps of the BP algorithm can be summarized as follows:

- (1) Initialization
- (2) Making the following calculations for each sample:

Forward calculation: supposing the input of the j -th unit in n iterations is $x_i(n)$, the output is $y_i(n)$, and the net input s_j of the unit is

$$s_j(n) = \sum_{i=0}^P w_{ij}(n)x_i(n), \quad (10)$$

where P is the number of inputs added to unit j , and w_{ij} is the connection weight from unit i to unit j in the previous layer, then, there is an output:

$$y_j(n) = f_j(s_j(n)). \quad (11)$$

Among them, $f_j(s)$ is the action function. If the action function of unit j is the sigmoid function, then

$$y_j(n) = \frac{1}{1 + \exp(-s_j(n))}, \quad (12)$$

$$y_j'(n) = \frac{-\exp(-s_j(n))}{(1 + \exp(-s_j(n)))^2} = y_j(n)(1 - y_j(n)). \quad (13)$$

Reverse calculation: supposing the expected output is $d_j(n)$, then the error signal $e_j(n)$ is $e_j(n) = d_j(n) - y_j(n)$, and the total square error $\xi(n)$ at the output is

$$\xi(n) = \frac{1}{2} \sum_{j=1}^q e_j^2(n). \quad (14)$$

The mean value ξ_{AV} of the squared error is

$$\xi_{AV} = \frac{1}{N} \sum_{n=1}^N \xi(n). \quad (15)$$

The correction amount of weight w_{ij} is

$$\Delta w_{ij}(n) = -\eta \frac{\partial \xi(n)}{\partial w_{ij}} = -\eta \delta_j(n) x_i(n). \quad (16)$$

TABLE 1: Four database information tables.

Database name	Identify object	Number of image classes	Volume
ImageNet	Multiple images	Not classified	14197122 images
MS COCO	Scene understanding	91	>1500000 objects
MNIST	Handwritten numeral	1	60000 training sets, 10000 test sets
The Chars74K dataset	Character recognition	64	74107 characters



FIGURE 5: ImageNet's image recognition method.

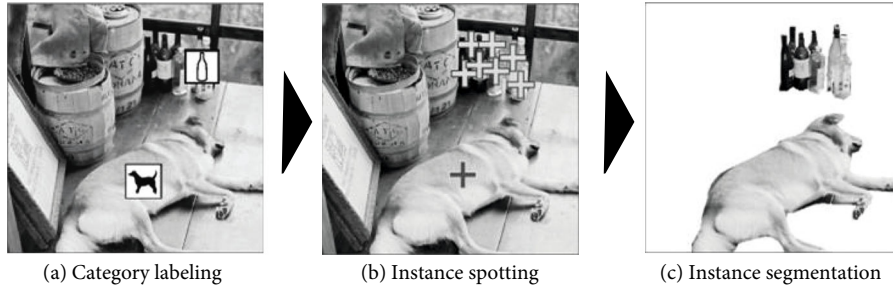


FIGURE 6: Image recognition method of MS COCO.

The minus sign indicates that the correction amount falls in the direction of the gradient, where $\delta_j(n) = e_j f'_j(s_j(n))$ is called the local gradient. When the unit j is an output unit, there are

$$\delta_j(n) = (d_j(n) - y_j(n)) f'_j(s_j(n)). \quad (17)$$

When unit j is an implicit unit, there are

$$\delta_j(n) = f'_j(s_j(n)) \sum w_{jk}(n) \delta_k(n). \quad (18)$$

Modify the weight as follows:

$$w_{ij}(n+1) = w_{ij}(n) + \eta \delta_j(n) x_i(n). \quad (19)$$

(3) $n = n + 1$, input a new sample until ξ_{AV} reaches the predetermined requirement

3. Preschool Education Experiment Based on Image Recognition

3.1. System Development. This article is mainly based on the Python language, using related technologies and methods to develop a learning interactive system that can recognize pictures. The structure of the preschool education system:

The system includes a mobile intelligent terminal and a service management center. The mobile smart terminal communicates with the service management center through the mobile network. The service management center includes an image recognition server, data processing server, and storage server [21]. The processing server is, respectively, connected to an image recognition server and a storage server for communication. The image recognition server includes an image recognition module, and the data processing server includes a classification retrieval module, an analysis processing module, a content output module, and a content interaction module. The mobile smart terminal is in communication connection with the image recognition server through a mobile network, and the mobile smart terminal is used to collect external image information and send the collected image information to the image recognition server. The image recognition module in the image recognition server is used to perform image recognition on the received image information and form image recognition information, and at the same time send the image recognition information to the data processing server. The data processing server retrieves, analyzes, and processes the image recognition information to form a data set, feeds the data set back to the mobile smart terminal, and saves the data set in the storage server [22]. The classification retrieval module is connected with the image recognition module, the analysis processing module is, respectively, connected with the classification retrieval module and the content output module, and the content output module is connected

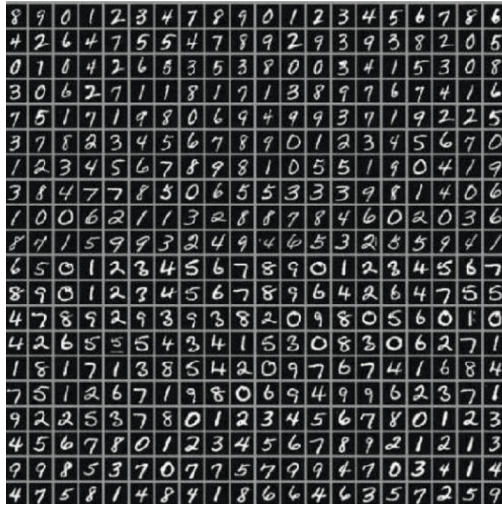


FIGURE 7: Some handwritten digits in MNIST.

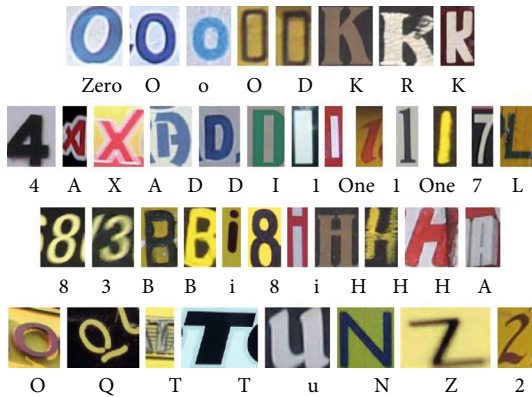


FIGURE 8: Part of the characters contained in The Chars74K dataset.

with the mobile intelligent terminal. The content interaction module is, respectively, connected with the analysis processing module and the mobile intelligent terminal.

The network structure diagram of this system is shown in Figure 9.

Its network structure is also relatively simple, which mainly includes the user side and the server side, and the user side includes PCs, tablets, and mobile phones. The general flow of the Web interaction of this system is that the user sends a request to the server. After the server receives the request, the system performs various operations, including preprocessing, identification, and search of picture pictures, and then returns the search results to the user.

3.2. System Test

3.2.1. *System Response Time Test.* In this test, the response time refers to the time from the end of the action of taking a picture to the time the system returns the information obtained by the picture recognition. The main factors affecting the response time of this system are the speed of image preprocessing and the speed of image recognition. The main influence on the image preprocessing time is the preprocess-

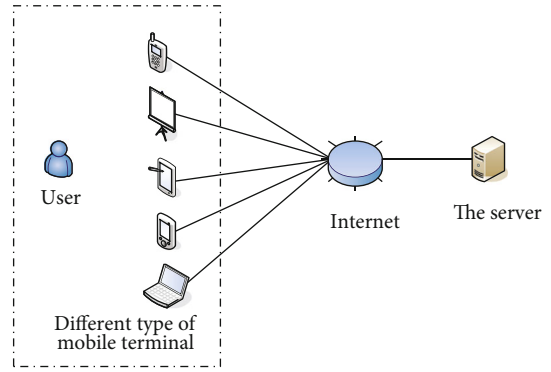


FIGURE 9: Network structure diagram of the preschool education system.

ing method; the main influence on the image recognition speed is the size of the database. This test selects 100 pictures taken in real scenes for testing, including 20 numbers, 20 letters, and 60 different objects (including animals, plants, toys, and furniture). There is only one target in the 100 pictures. The test results are shown in Table 2.

3.2.2. *System Stress Test.* System stress testing refers to testing the system’s ability to handle multiple requests at the same time under concurrent conditions, including the upper limit of the number of requests that the test system can accept and the number of concurrent systems within a reasonable response time. The pressure test results of this system are shown in Table 3.

From the test results, as the number of concurrent increases, the number of requests per second and the average response time of the page will also increase.

3.2.3. *Recognition Rate Test.* The samples used in this test selected 200 different characters and objects taken in real scenes, including 30 handwritten digits, 40 uppercase, and lowercase letters; 30 printed numbers in different fonts, 40 uppercase and lowercase letters; and 60 different objects. There is only one target in the 100 pictures, and the recognition effect is shown in Table 4.

From the test results, the recognition rate of handwritten digits is 83.3%, and the recognition rate of digits in different fonts printed is 93.3%, the recognition rate of handwritten letters is 80.0%, and the recognition rate of printed letters in different fonts is 92.5%, the recognition rate of different objects is 76.7%, and the total recognition rate is 85.16%. The return action occurs after the relevant information is retrieved and is generally presented on the screen of the mobile terminal in the form of text.

3.3. System Application

3.3.1. *Letter and Number Recognition Application Results.* After completing the system test, it is necessary to conduct application experiments on the system. Application experiment refers to combining the system with mobile devices to evaluate the educational results of preschool children using the system in a certain period and to set up a control to compare the teaching results of the experimental group

TABLE 2: System response time test.

Picture type test item	Pretreatment	Character recognition	Total
Number	191 ms	4380 ms	4571 ms
Letter	194 ms	4776 ms	4970 ms
Object	208 ms	5829 ms	6037 ms

TABLE 3: System stress test results.

Concurrent number	Request time per second	Average response time
25	21.27	3279 ms
50	47.26	3682 ms
100	89.18	4358 ms
200	165.45	5347 ms

with the control group. In this experiment, preschool children are divided into 4 groups according to their age, each group of 3 children of the same age, namely, the 3-year-old group, the 4-year-old group, the 5-year-old group, and the 6-year-old group using a preschool education machine equipped with an image recognition education interactive system. As well as the control 3-year-old group, 4-year-old group, 5-year-old group, and 6-year-old group using traditional preschool education machines, a one-month experiment was conducted on these 8 groups. Using the preschool education machine for one hour a day to learn letters and one hour to learn numbers. After one month, evaluate the learning results according to the number of letters and numbers (numbers include numbers above 10). The results are shown in Figure 10.

It can be seen from Figure 10 that the 3-6-year old children in the experimental group that used this image recognition interactive system knew more letters and numbers one month later than the 3-6 year old children who used the traditional preschool education machine. This difference is not large enough at the age of 3, but with age, the difference will become larger. And children's learning of letters is more affected by the educational interactive system. As the age increases, the better learning interactive system is more reflected.

3.3.2. Recognition Application Results of Different Objects when Single Target and Double Target. Sometimes a picture may have more than one identifiable target. Here, we will compare the learning situation of children with a picture with a target recognition object and a picture with two target recognitions, investigate children's knowledge and learning of different objects, and when there are two objects in a picture, whether children's knowledge and learning of objects will be affected, whether it will help or not be conducive to learning and memory. The samples used in the experiment are divided into single-target group and double-target group. In the single-target group, each picture has only one target. There are 10 pictures in each of the four categories of different animals, plants, toys, and furniture, for a total of 40 pic-

tures. There are two goals for each picture in the dual goal group, four types of goals are mixed and matched, and a total of 20 pictures. The target image involved in the picture is consistent with the single target group image, but the target of two pictures originally belonging to the single target group is moved to the same picture through PS processing. The subjects of the experiment were children aged 4, 5, and 6 years old. There are 3 children of the same age in each group, for a total of 9 children. The experiment lasted for three weeks and studied for one hour a day. After two weeks, the test was conducted, and the same pictures were used to see how much the children knew. The experimental results are shown in Figure 11.

It can be seen from Figure 11 that the degree of difficulty in understanding different objects is different. For children aged 4-6, it is more difficult to understand furniture than animals, plants, and toys. When faced with dual-target pictures, interactive learning becomes more difficult, which may be because the system is more difficult to recognize dual-target pictures than single-target pictures, and it is easy to be interfered with and misidentified. It may also be because children are easily confused when facing dual goals, and the specific reasons need to be explored by further experiments.

4. Discussion

This article aims to realize a system that can use image recognition technology for preschool education. The system includes (1) CMOS intelligent image sensor; (2) BP neural network algorithm; (3) ImageNet, MS COCO, MNIST, and the Chars74K dataset database. The main functions of the system are (1) preprocess the photos, (2) transmit the pre-processed photos to the network, (3) use algorithms in the database to identify the target on the photo, and (4) return the identified results to the mobile terminal. In general, the system can achieve the expected purpose and can conduct learning interaction based on image recognition [23].

To further understand the system, a series of experiments were carried out, including the test and application of the system. The testing of the system includes response time, pressure, and recognition rate testing. The result of the response time test shows that the system is better at character recognition, and the average recognition time of objects is longer than that of characters. The pressure test shows that the system is relatively stable and can withstand greater pressure, but when the pressure is too high, the processing time of the system will also increase, and it will not increase uniformly, but the greater the pressure, the greater the increase in the time spent. The recognition rate test result is 85.16%; the recognition rate of printed characters is 92.5% for letters and 93.3% for numbers; the recognition rate of handwritten characters is 80.0% for letters and 83.3% for numbers; the recognition rate for different objects is 76.7%. It can be seen that the system has a good recognition function and a high overall recognition rate. It is particularly good at recognizing printed characters. The recognition rate of handwritten characters is also higher,

TABLE 4: System recognition rate test results.

Picture type	Quantity	Number of recognition errors	Recognition accuracy	Total recognition rate
Handwritten numeral	30	5	83.3%	85.16%
Print numbers	30	2	93.3%	
Handwritten letter	40	8	80.0%	
Print letters	40	3	92.5%	
Different objects	60	14	76.7%	

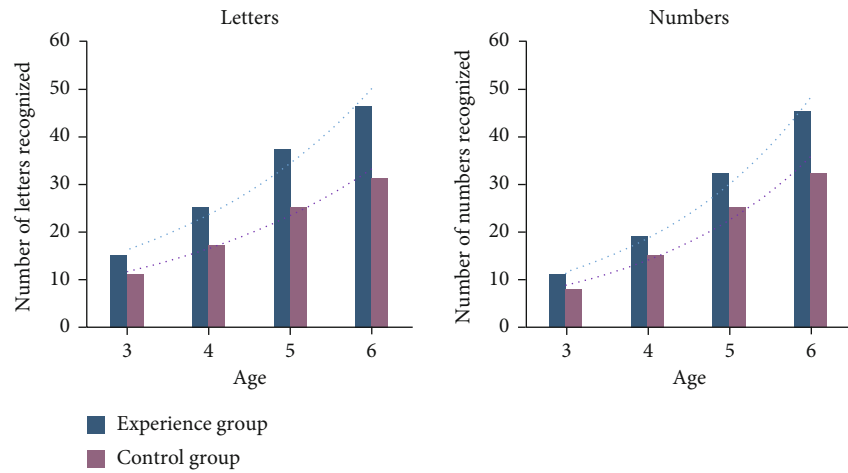


FIGURE 10: System application results of number and letter learning.

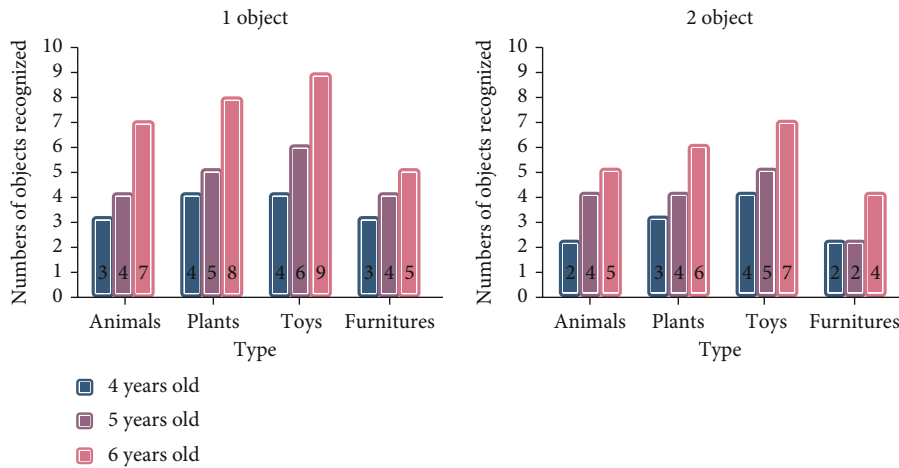


FIGURE 11: Recognition application results of different objects for single target and dual target.

which is better than the recognition of objects, which may be due to the complex shapes of objects.

The application experiment of the system is to investigate the performance of the system in practical application, so the system is implanted in the preschool education machine, and the preschool education machine is used to realize the function of the system [24]. Comparing the results of children using different kindergarten education machines to learn letters and numbers, it can be found that the kindergarten education machine implanted with the image recognition system can open the gap with the old kindergarten education machine in

letter teaching than number teaching. And for children aged 3-6, the new image recognition preschool education interactive system's teaching performance is clearly better than the old preschool education system. For the experiment of single-image single-target and single-image dual-target object recognition teaching, not only ages are grouped, but the types of objects are also distinguished. The experimental results prove that simultaneously identifying and learning multiple objects is not conducive to interactive teaching.

Although the experimental considerations in this article are more comprehensive, the sample size of the experiment

is not large enough, the experimental results are not enough to have strong representativeness, and individual problems still need further experiments. The system developed in this article still has some shortcomings, and further tests and improvements are needed in the usability of children, to achieve the purpose of easy and simple operation for children in daily life. In addition, the system can be further developed in terms of function plus audio-based auxiliary teaching, which is more conducive to children's understanding and memory and achieves the purpose of audiovisual teaching.

The system designed in this research has the following advantages and beneficial effects: the system uses the camera on the smart terminal to capture natural images and uses intelligent image recognition and database intelligent analysis technology to quickly and accurately identify physical images taken by children. Through the feedback of the photographed image information, it is uploaded to the intelligent database for statistical analysis to form the corresponding preschool education knowledge points, and then the corresponding preschool education knowledge points are presented on the screen from the child's perspective and aesthetics for interactive learning by the child. At the same time, it can also design corresponding puzzle games, cartoons, nursery rhymes, children's stories, etc. according to the children's age and preferences. Let children naturally complete the process from recognition to knowledge in relaxed interactive learning, to gain comprehensive cognition and learning, and at the same time, learning has a certain degree of fun.

5. Conclusions

The image recognition preschool education interactive system based on smart sensors designed in this research utilizes the advantages of CMOS sensors that can amplify signals to achieve better image processing and better recognition. In addition, the system combines multiple databases for image recognition, which effectively improves the recognition rate, expands the recognition range, and better meets the needs of children's learning. The overall recognition rate is as high as 85.16%, which is more effective than traditional learning systems, and improves the backwardness of the preschool education system. It provides better interactive education for preschool children, improves learning efficiency, truly realizes "learner-centered," and promotes students' cognitive development.

Data Availability

The data underlying the results presented in the study are available within the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by (1) the key research project of humanities and social sciences of colleges and universities in Anhui Province is "Research on the Construction of the Monitoring system of Campus Football influencing factors under the New situation of speeding up the Construction of Sports Power", Project number: SK2018A0469; and (2) Excellent academic and technical backbone (2016XJGG03).

References

- [1] S. Assefa, F. Xia, S. W. Bedell et al., "CMOS-integrated high-speed MSM germanium waveguide photodetector," *Optics Express*, vol. 18, no. 5, pp. 4986–4999, 2016.
- [2] S. Bourdel, Y. Bachelet, J. Gaubert et al., "A 9-pJ/pulse 1.42-Vpp OOK CMOS UWB pulse generator for the 3.1–10.6-GHz FCC band," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 1, pp. 65–73, 2018.
- [3] D. Bol, "Robust and energy-efficient ultra-low-voltage circuit design under timing constraints in 65/45 nm CMOS," *Journal of Low Power Electronics & Applications*, vol. 1, no. 1, pp. 1–19, 2019.
- [4] Z. Chen, H. Zhu, E. Ren et al., "Processing near sensor architecture in mixed-signal domain with CMOS image sensor of convolutional-kernel-readout method," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, no. 9, pp. 1–12, 2019.
- [5] J. Xin, L. Mingyong, Z. Kaixuan, J. Jiangtao, M. Hao, and Q. Zhaomei, "Development of vegetable intelligent farming device based on mobile APP," *Cluster Computing*, vol. 22, Supplement 4, pp. 8847–8857, 2019.
- [6] J. Kim and C. Yoon, "Three-dimensional head tracking using adaptive local binary pattern in depth images," *International Journal of Fuzzy Logic & Intelligent Systems*, vol. 16, no. 2, pp. 131–139, 2016.
- [7] A. A. Elngar and M. Kayed, "Vehicle security systems using face recognition based on internet of things," *Open Computer Science*, vol. 10, no. 1, pp. 17–29, 2020.
- [8] M. de Lourdes Cró and A. M. C. Pinho, "Resilience and psychomotricity," *International Journal of Privacy and Health Information Management*, vol. 4, no. 2, pp. 38–52, 2016.
- [9] R. G. Babu, P. Karthika, and G. Manikandan, "Polynomial equation based localization and recognition intelligent vehicles axis using wireless sensor in MANET," *Procedia Computer Science*, vol. 167, no. 3, pp. 1281–1290, 2020.
- [10] Y. Haibin, "Application analysis of new internet multimedia technology in optimizing the ideological and political education system of college students," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 5557343, 12 pages, 2021.
- [11] Z. Tu, B. Duan, Z. Wang, and X. Xu, "Bidirectional sensing of user preferences and application changes for dynamic mobile app recommendations," *Neural Computing and Applications*, vol. 33, no. 16, pp. 9791–9803, 2021.
- [12] H. Wei, A. Sehgal, and N. Kehtarnavaz, "A deep learning-based smartphone app for real-time detection of retinal abnormalities in fundus images," *Real-Time Image Processing and Deep Learning*, 2019.

- [13] M. Abdolmaleky, M. Naseri, J. Batle, A. Farouk, and L. H. Gong, "Red-green-blue multi-channel quantum representation of digital images," *Optik*, vol. 128, pp. 121–132, 2017.
- [14] X. Peng, L. Wang, X. Wang, and Y. Qiao, "Bag of visual words and fusion methods for action recognition: comprehensive study and good practice," *Computer Vision & Image Understanding*, vol. 150, no. C, pp. 109–125, 2016.
- [15] J. Lin, J. Li, and J. Chen, "An analysis of English classroom behavior by intelligent image recognition in IoT," *International Journal of System Assurance Engineering and Management*, vol. 24, no. 9, pp. 1–9, 2021.
- [16] W. Zeng, Z. Lin, Y. Chen et al., "Fruits and vegetables image recognition in intelligent refrigerator based on convolutional neural network," *Microcomputer & Its Applications*, vol. 87, no. 7, pp. 24–29, 2017.
- [17] T. Shen, C. Gao, and D. Xu, "The analysis of intelligent real-time image recognition technology based on mobile edge computing and deep learning," *Journal of Real-Time Image Processing*, vol. 14, no. 2, pp. 157–161, 2020.
- [18] H. Zhang, Q. Zheng, B. Dong, and B. Feng, "A financial ticket image intelligent recognition system based on deep learning," *Knowledge-Based Systems*, vol. 64, no. 10, pp. 106–255, 2021.
- [19] H. Chen, Y. Lu, and L. Tu, "Fault identification of gearbox degradation with optimized wavelet neural network," *Shock and Vibration*, vol. 20, no. 2, 262 pages, 2013.
- [20] Y. Liu, Y. Liu, S. Zhong, and S. Wu, "Implicit visual learning," *ACM Transactions on Intelligent Systems and Technology*, vol. 8, no. 2, pp. 1–31.24, 2017.
- [21] M. Adil, H. Song, J. Ali et al., "EnhancedAODV: a robust three phase priority-based traffic load balancing scheme for internet of things," *IEEE Internet of Things Journal*, 2021.
- [22] Q. Wang, H. L. Chen, W. J. Yan, Y. H. Chen, and X. Fu, "Face recognition and micro-expression recognition based on discriminant tensor subspace analysis plus extreme learning machine," *Neural Processing Letters*, vol. 39, no. 1, pp. 25–43, 2021.
- [23] H. Hu, Z. Gao, Y. Sheng, C. Zhang, and R. Zheng, "Traffic density recognition based on image global texture feature," *International Journal of Intelligent Transportation Systems Research*, vol. 17, no. 3, pp. 171–180, 2019.
- [24] S. Banerjee and A. Chatterjee, "Image set based ear recognition using novel dictionary learning and classification scheme," *The International Journal of Intelligent Real-Time Automation*, vol. 55, no. 3, pp. 37–46, 2016.