Research on Automatic Recognition Technology of Library Books Based on Image Processing

Haiyan Xun
Library of Shandong Women's University, Jinan, Shandong, 250300, China
E-mail: xunhaiyan2023@126.com

Keywords: image processing, automatic recognition technology, image retrieval

Received:

The intelligence of computers is the future development direction. In today's society, the amount of information is increasing, which puts forward higher requirements for retrieval technology and automation levels. With the development and popularization of the Internet era, online shopping, study, and even work have become the norm in people's lives. However, a large amount of data is generated on the Internet every day, and how to obtain the information people need from this data is a key research problem today. It can be seen that the traditional global search or image search can no longer meet the amount of information people need, so the content-based search method will inevitably become a more popular database retrieval technology. In recent years, content-based image capture has become a research center in the field of image information removal. This article first studies image processing and capture technology. It provides a detailed overview of image segmentation technology and image segmentation models introduced in image processing, as well as three image segmentation methods. This article also introduces the basic principles and framework of the CBIR procurement system. The most important technology in the CBIR system is the analysis and description of image features, and several methods commonly used to express the content of feature images. Secondly, it analyzes the image edge detection algorithm in detail and finally introduces the functional division and workflow of the library book automatic recognition system. It also provides all the construction environments, processes, and algorithms to perform the basic functions of the system.

Povzetek: Predstavljena je tehnologije obdelave in zajemanja slik, podrobno še segmentacija slik in metode segmentacije.

1 Introduction

In modern scientific research and daily life, people usually prefer to obtain information through images. However, traditional manual identification has been unable to meet the needs of social development for the research on the automatic identification technology of artificial intelligence library books. With the development of various computer technologies and networks, the informatization development of various enterprises has become the current mainstream development direction. In this information age, people usually learn about social changes through website information. The library is the dissemination center of all kinds of information in colleges and universities [1-2]. With the popularization of information technology, major libraries have also established online libraries. Such online libraries can not only effectively manage the books in the library, but also facilitate people to read books. At the same time, the online library also speeds up the time for people to find the books they need. However, due to the large number of secretaries in the library, the order of books in the online library system is also entered according to the order of the books in the offline library. Therefore, there may be errors in the information of the book in the system due to the incorrect placement of the book. This has a very bad influence on the development of online libraries [3].

Image segmentation usually refers to the segmentation of digital images. The method of image segmentation is the process of dividing an image into several discontinuous regions. The basic idea is to classify and group images with similar feature values in digital images, and divide the images into regions with different important levels, thereby reducing the amount of data and information contained in each part of the image. The structure information is suitable for the later stage of image processing, and it is further processed to store information about the target structure while significantly reducing the amount of target data [4-5].

Image segmentation techniques can be divided into three categories. One is based on the extracted regional features or the morphological results obtained from the regional features. It uses thresholds, regional methods, and textures to classify image pixel groups; the second is to pay attention to the picture's border, which divides the margin to obtain the region according to the obtained operator. The last method is to use statistical features and

prior knowledge features for image segmentation. This method is to segment first, then highlight the objects, and then combine them into segmented regions according to the domain of knowledge [6].

According to the three categories of image segmentation, the most commonly used segmentation techniques are the threshold method, edge detection method, and region improvement method. In threshold segmentation, the maximum threshold and blur threshold are more suitable for the segmentation of specific types of images (vessels) [7-8]. Although that which separates the background from the target is not very obvious, the segmentation effect is not very good; the method of edge detection is to first detect the pixels along the edge of the image, and then stitch the end pixels together to form a segmented area. The edge detection method uses part of the window operation to find the edge information in the image. This method is suitable for images with a solid color background [9-10].Literature survey are discussed in Table 1.

Table 1: Related works

Ref. No	Methodologies	Summary of findings	Limitation
[11]	To facilitate effective deep-learning medical image processing, the study presents TorchIO, a Python package. Patch-based sampling, spatial metadata consideration, and data augmentation are highlighted as ways to overcome the difficulties associated with processing MRI and CT images	Patch-based sampling, augmentation, preprocessing, and loading are made easier by TorchIO. It supports composition, transform inversion, and simulation of MRI-specific artifacts. It interfaces with PyTorch and conventional medical image processing libraries. Open-science principles are encouraged by the availability of the source code, tutorials, and documentation.	The efficiency of TorchIO, which simplifies medical image processing, could differ depending on the application case. The modularity and interoperability of the library with various deep-learning frameworks for medical pictures should be taken into consideration by users.
[12]	To extract gear tooth profiles properly without depending on conventional meshing theory, the research presents a novel method called Engagement-Pixel Image Edge Tracking (EPIET). The process entails extracting meshing points, calibrating tool locus coordinates, and obtaining immediate contact images.	The method's efficiency in removing tooth profile edges is illustrated by a case study on a face gear. Its promise for computerized design of complicated conjugate curved surfaces is demonstrated by the results, which show feasible accuracy and stability when compared to standard meshing equations.	The paper discusses major error sources associated with the presented method. While it shows promise for gear profile extraction, further research is needed to explore its applicability to diverse gear types and potential limitations in complex scenarios.
[13]	Using a case study methodology, this paper examines how public libraries, namely the Seattle Public Library (SPL) system, curate and use available demographic data. As part of the study, use cases are created, a dashboard tool prototype is developed using available census data, and information needs are identified through interviews with SPL regional managers.	The study provides new perspectives on how open data might be used to help SPL regional managers find the information they need. The needs within two SPL regions are successfully addressed via a dashboard tool prototype. As a result of the findings, public libraries may find it easier to keep up with changing neighborhood demographics by developing reproducible data analytic techniques.	Limitations include the case study's specialization to the SPL system, which may restrict generalizability even though the paper provides insightful information. Furthermore, issues concerning the timeliness and accuracy of publicly available demographic data may affect how broadly applicable these approaches are in various public library environments.
[14]	The goal of the project is to increase the efficiency of library book borrowing by developing an automated book sorter that makes use of RFID	The device’s great precision and quick reaction time when processing book numbers and operating equipment are demonstrated through practical operation. Effective book classification	Even though the automated book sorter works well, there could be issues with managing different book sizes or fixing system

	technology and a single-chip microprocessor control system. The sorter uses an AC motor to power a conveyor belt, recognizes electronic tags using RFID and a microcontroller, and precisely places books in recycling bins.	is made possible by the combination of RFID technology and a single-chip microcomputer, which improves library operations and improves student learning outcomes.	errors. It might take constant improvement and adjustment to maximize performance in different library circumstances.
[15]	Using Particle Image Velocimetry (PIV) to assess multiphase flows, the study highlights the importance of appropriate image segmentation for precise phase dynamics separation. Triangular meshing and particle detection are used in the suggested approach to provide reliable phase separation and interface detection.	The new technique uses a 2D unstructured mesh for phase separation and interface detection and successfully identifies tracer particles by utilizing seeding density differentiation between phases. The effectiveness of the method is demonstrated by parametric analysis on synthetic images, and successful tracking of complicated interface evolution is revealed by experimental application in immiscible multiphase flow in porous media.	While there is promise in the method, more validation under various experimental conditions is required to verify its wider application, as potential constraints may develop in other flow scenarios.

1.1 Problem statement

There are problems with misplacing books since the current library book management systems are not effective in identifying books. By using image processing technologies, this research proposes an automatic identification technique to address the issue. Accurate character recognition, post-processing analysis, and bookcase detection are the system's main goals. Improvements in orderliness, real-time book tracking, and user-friendly reading environments are all part of the plan to improve library management.

2 Research on image segmentation technology and retrieval technology

2.1 Image segmentation technology

General model of image segmentation

According to the basic concept of image segmentation, the known image segmentation is the process of dividing an image into several regions. This segmented area is a combination of pixels with common numerical features. For example, different objects occupy different areas of the image, and the image and background objects occupy different areas of the image. The segmented regions must have sufficient homogeneity and connectivity to be segmented together[8].

Let F be the collection of all the image's pixels G, and the hypothesis about the uniformity of $\bigcup_{j=1}^{n} S_j = F$ is P (.)[3].

The mathematical explanation of the above situation is shown by formulas (1), (2), (3), (4):

$$\bigcup_{j=1}^{n} S_j = F \tag{1}$$

$$S_i \cap S_j = \varnothing, (i \neq j) \tag{2}$$

$$P(S_j) = \text{TRUE}, (\forall j) \tag{3}$$

$$P(S_i \cup S_j) = FALUS, (i \neq j) \tag{4}$$

Summary of image segmentation methods

(1) Threshold segmentation

Threshold segmentation can be divided into three methods: global threshold method, multi-threshold method, and adaptive threshold method. Among them, the global threshold method is easier to apply to complete images with obvious contrast between the target and the background, especially if the level of the gray background is fixed, the effect is much better.

The multi-threshold method is suitable for images with different target types and background areas, and different thresholds can achieve different targets. The main core meaning of multi-threshold is to set multiple thresholds so that for images with different target types and background areas, multiple grayscales in each target and background area in the image can be compared, which enables more accurate detection. The gray value of the pixel, and more accurately segment the image. The adaptive threshold will be adjusted according to the change in the gray background level, and the background from the target will also change. If the threshold is fixed, the target effect of this kind of image acquisition is not good, and the adaptive threshold method, that is, the gray threshold obtained under different image conditions is also different, and it may have a better segmentation effect [9].

The essence of image segmentation is to split a complete image into multiple small images, and the basis for splitting into small images is based on the feature attributes of each pixel in the image, the pixel points that are closer and have similar feature attributes are grouped into a small image, and then these pixel points are used as the center to expand outward to find similar pixel points, and if similar ones are found, they can be fused into the small image, and so on until the whole image is split.

The global threshold method is suitable for the complete image with obvious contrast between the target and the background, but it has certain disadvantages when it is used to segment the image where the contrast between the target and the background is not obvious and incomplete. The disadvantage of incomplete images is that some of the pixel point properties in the image are changed or lost, the histogram of these incomplete images is generally close to a single peak, and the values derived from traditional thresholding are generally not at the bottom of the valley in the histogram, so if the incomplete images are then segmented using traditional thresholding segmentation methods cannot get the correct threshold values and thus also cannot segment the images reasonably.

When the image threshold is segmented, the segmentation of the target and the background is based on comparing the gray pixel value of the position with the specified threshold. The classic method of selecting a threshold is the maximum between-class variance method, which is used alone to segment a single threshold. The principle is to maximize the difference between the two parts of the classroom by setting a threshold.

(2) Edge detection

The edge is one of the important components of the image, so edge detection must be performed when image segmentation. The edge of the image is generally between the target image and the background. If the edge information in the image can be obtained during image segmentation, this can improve the efficiency of image segmentation [10].

The gradient operators of the commonly used edge detection algorithms include the Robert operator, Sobel operator, Prewitt operator, Laplace operator, etc.

Roberts's edge detection operator has two pairs of detection operators, which are the detection operators in the vertical direction and the diagonal direction. These two pairs of detection operators detect the difference between two adjacent pixels in their respective directions at the same time. It can be seen that the Roberts edge detection operator can detect the edges of the image more accurately, but the Roberts edge detection operator is more sensitive to the noise generated in image segmentation, so it is not suitable for image segmentation with blurred edges and large segmentation noise [11].

The principle of the Roberts edge detection operator is to use the vertical and diagonal detection operators to detect the difference between two adjacent pixels in two directions in an image. The characteristics of the Roberts edge detection operator are as follows: Roberts is more sensitive to the noise generated by image segmentation, which makes this operator not suitable for image segmentation with fuzzy edges and more complicated segmentation. However, noise will inevitably occur in the image segmentation process, so there is a certain error in the edge of the image detected by this operator.

The biggest advantage of the Sobel operator is that it has a certain ability to suppress noise when performing edge detection. This ability enables the Sobel operator to help the image segmentation more smoothly, which makes the edges of the image segmentation clearer, and it can also effectively remove false edges. The traditional Sobel edge detection operator's edge positioning and noise smoothing are contradictory. To overcome this shortcoming, people use edge detection combined with template matching to effectively adjust this contradiction.

Sobel operator principle: In image processing, the difference between adjacent pixels or spaces is often used to represent image edge information. The principle of the Sobel operator is to use the corresponding image data convolution mode to perform weighted calculation and approximate calculation on discrete data. It uses the small convolution mode to perform horizontal edge detection and vertical border recognition.

The edge location of the Sobel edge detection operator is contradictory to noise smoothing. To overcome this shortcoming, people use edge detection combined with template matching to effectively adjust this contradiction. The first is to increase the inclusion of the Sobel operator relative to the edge direction. Due to the different edge directions, 6 templates with different edges moving 45 degrees clockwise can be added [12]. Secondly, the Sobel operator algorithm is improved, that is, the result of the S convolution operation on the 8 modes M1-M8: S1 = a1+2a8+a7-a3-2a4-a5. So Si (1 = <i = <8)), and finally S = max {Si} (1 = <i = <8).

The traditional calculation method of the Prewitt operator is: first do convolution, then compare the dimensions, and use the result of the convolution operation as the final

value. 7 convolution operations are unnecessary, resulting in a large amount of calculation. This amount of calculation is too disadvantageous for processing images quickly in real-time.

There are 7 extra convolution operations under the conventional calculation method of the Prewitt operator, which causes a large amount of calculation. This is very disadvantageous for fast real-time processing of images. In response to this problem, the researchers considered minimizing the convolution operation, thereby reducing the number of operations. If the comparison operation is performed first to maximize the convolution mode, and then the convolution is performed as expected, then if the convolution operation is performed only once, the amount of calculation will be greatly reduced, thereby improving the computational efficiency.

The Laplace operator is a second-order differential, which doubles the influence of noise, resulting in sharp edges in the image. Its advantages are isotropy, that is, rotation invariance and displacement invariance. It can be known from differential calculus that only linear combination operators composed of even-order derivatives and even-order odd-order derivatives must be isotropic.

Image Retrieval Related Technologies： CBIR technology can be traced back to 1992. This technology is mainly used to solve the problem of excessive newspaper image capture. CBIR is a selection method that directly uses image content as image information to request an image. The main techniques involved are: highlighting features, comparing similarities, and extracting matching results. For CBIR, there are many algorithms, and the basic extraction methods include extraction methods based on color features, texture features, shape features, and spatial features [13-14].

Based on the example of image query, the CBIR system framework can be divided into three modules: function search, function comparison, and result display. The system framework is shown in Figure 1 below.

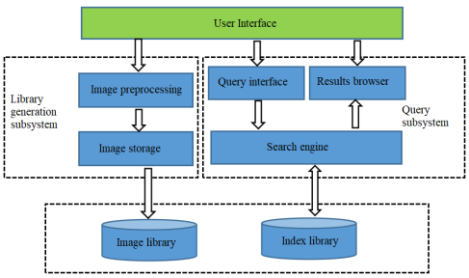

Figure 1: The system framework of the CBIR system

3 Image edge detection algorithm

Mathematical Background： The definition of the first derivative of the image grayscale in the digital picture is shown in formula (5), and the expression of the second derivative is shown in formula (6):

$$f'(x) = f(x+\Delta x) - f(x) = f(x+1) - f(x) \quad (5)$$

$$f''(x) = f'(x+1) - f'(x) = f(x+2) - 2f(x+1) + f(x) \quad (6)$$

Edge Model： When selecting the original features of an image, edges are important features in image analysis. They are located at the boundaries of different regions and represent significant local changes in image intensity. These changes appear as discontinuities in image intensity, disc discontinuities, or line discontinuities. As shown in Figure 2, phase discontinuity refers to the sudden change of image intensity from one initial value to another, and line discontinuity refers to the sudden change of image intensity from one initial value to another, and then restores to the initial value. In reality, the low-pass filtering function is introduced by the sensor, and the two discontinuities are shown as slope edges and roof edges [15].

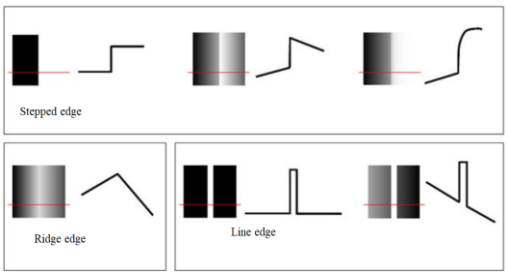

Figure 2: Edge model diagram

3.1 Canny edge detection operator

There are two main types of basic linear recognition methods for raster images. One is based on a grayscale equation containing additional conditions (such as gradient), and the other is based on the edge of the object. In most applications, the purpose of detecting a straight line is to determine the object in the image, and the object should be bounded at the boundary. A boundary is a special pixel that must meet a certain boundary when creating a target edge. Although it is directly based on grayscale and its transformation, it is usually not tested whether it meets the edge properties, and the observed lines may not be the actual edges that are needed. This is the direct reason why related algorithms are generally considered wrong. For this reason, it is more convenient to perform line detection based on edge analysis [16-17].

Among many edge detection algorithms, the canny algorithm generally uses a fixed threshold for edge detection. Of course, the Canny algorithm generally has two thresholds. When using this algorithm to segment the target image and the image with more background, its efficiency is not high. In addition, some local edges with slow changes in gray value will be lost, so this algorithm may also have the defect of false edges or edge loss. The canny operator is recognized as an operator with a low

error rate, accurate positioning, and strong noise-elimination ability. It can make more accurate decisions on edge pixels, and it is widely used in edge detection. The edge analysis algorithm used by this algorithm is executed as follows: let f (x, y) be the input gray image, and G (x, y) be the two-dimensional Gaussian function. Canny's edge detection algorithm first performs audio filtering and Gaussian filtering on the input image to find the best value between the edge positions. According to formula (7), a well-balanced and uniform image is obtained, in which the two-dimensional Gaussian function is expressed by formula (8):

$$F(x,y) = G(x,y) * f(x,y) \quad (7)$$

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (8)$$

Gaussian filtering is performed on the image, that is, the specified o, x, and y are taken to obtain the corresponding Gaussian kernel, which is convolved with the image [12]. On this basis, the partial derivative is used to calculate the gradient amplitude and gradient direction as shown in formulas (9) and (10):

$$M(x,y) = |\nabla F(x,y)| = \sqrt{\frac{\partial F}{\partial x}^2 + \frac{\partial F}{\partial y}^2} \quad (9)$$

$$n = \nabla F(x,y) / |\nabla F(x,y)| \quad (10)$$

Among them, the calculation of the partial derivative can be obtained by the first-order difference approximation. The gradient operator used by the canny operator is shown in formulas (11), (12), and (13):

$$C_x = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, C_y = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \quad (11)$$

$$\frac{\partial F}{\partial x} = \frac{1}{2}(F(x,y+1) - F(x,y) + F(x+1,y+1) - F(x+1,y)) \quad (12)$$

$$\frac{\partial F}{\partial y} = \frac{1}{2}(F(x+1,y) - F(x,y) + F(x+1,y+1) - F(x,y+1)) \quad (13)$$

Because canny edge detection requires a single edge response, the principle of non-maximum suppression (as shown in Figure 3) is introduced to refine the edges of the image.

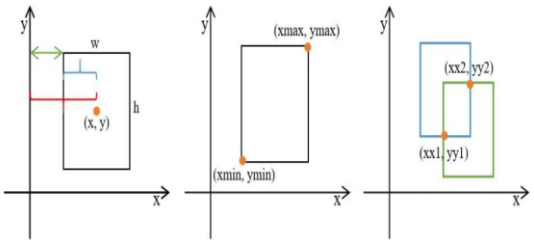

Figure 3: Non-maximum suppression

To obtain more accurate edge results, the dual-threshold method is used to filter noise below the lower threshold and maintain strong edges above the higher threshold. Then, after the cropping operation, the gaps between the strong corners are connected to obtain a convenient edge with a pixel angle. The double threshold method is as follows:

Set the upper limit to TH and the lower limit to TL. Usually, TH/TL is between 2-3. Using the two boundaries of the gradient magnitude image M(x, y), the strong edge is obtained according to formula (14), and the weak edge is obtained according to formula (15):

$$M_H(x,y) = M(x,y) \geq T_H \quad (14)$$

$$M_L(x,y) = M_H(x,y) - M(x,y) \geq T_L \quad (15)$$

The Canny algorithm generally has two thresholds, and its efficiency is not high when it is used to segment the target image and the image with more background. The core reason for this problem is that the traditional Canny algorithm uses a fixed threshold when detecting the edge of the image. This leads to the lack of flexibility and adaptability of the traditional Canny algorithm. Therefore, researchers have added a genetic algorithm to the traditional Canny algorithm. The Canny algorithm based on the genetic algorithm can adaptively generate a dynamic threshold according to the complexity of the image and the change of the gray value, thereby effectively solving the problem of low image segmentation efficiency.

The traditional Canny edge detection algorithm causes the problem of image segmentation due to the problem of fixed threshold, which may cause false edges or missing edges with slow changes in gray value. This is very unfavorable for image edge detection and acquisition. In response to this problem, researchers have proposed a method that combines statistical filters and gray-scale iteration to calculate the threshold.

This can not only suppress noise through statistical filters but also ensure that edge detection is not affected by noise; secondly, the threshold is determined by gray-scale iteration calculation, which avoids the problem of edges that are lost due to slow gray-scale changes.

The canny operator is an edge detection algorithm. Its specific steps are as follows: First, the image is smoothed by Gaussian filtering to reduce the noise of image segmentation. Secondly, use relevant calculation formulas to find the size and azimuth of the image gray gradient. Then the non-maximum value suppression is performed on the gradient amplitude obtained in the gradient direction, and the local maximum point is found. Finally, upper and lower thresholds are used to detect and connect edges.

3.2 Image edge tracking

After the image is divided into several areas, the computer will usually visualize and describe the set of pixels it contains suitably for further processing. To obtain the information of a region, it is usually represented by an outer boundary or a group of inner pixels. According to the linear characteristics of the grid, this paper uses Freeman's chain code as the edge description and introduces a fault-tolerant mechanism to track the straight line pattern, and then detect the straight line [18-19].

1. Freeman chain code

The serial code is used to indicate a boundary composed of straight-line segments connected in a specified direction and length. Usually based on 4 adjacent pixels or 8 adjacent pixels, each connection address uses a digital code, as shown in Figure 4, it is a Freeman series code. This article uses this structure to describe the subsequent results.

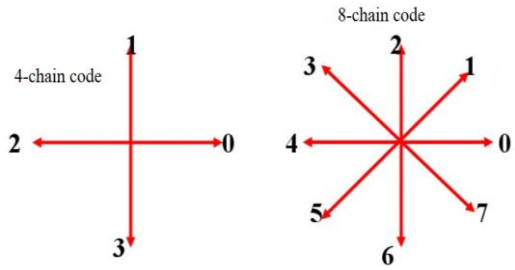

Figure 4: Freeman chain code

2. The characteristics of the grating line

To draw quickly, people have carried out extensive research on straight-line grids. Any straight-line segment can only be composed of adjacent straight-line segments with a specific pattern. The "pattern" described here refers to a short straight line that has the same length or is separated by only one pixel. Raster line pixels have strong directivity, see Figure 5.

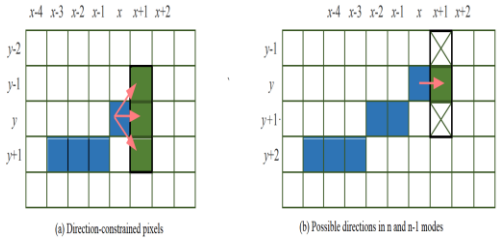

Figure 5: The characteristics and direction of the grating line

3. Fault tolerance mechanism

Under unfavorable conditions, if the linear direction and pattern of the grating are strictly used for pixel occlusion and tracking, many line segments will be damaged. These small line segments are not easy to connect due to the large gap. Therefore, a specific fault tolerance mechanism must be introduced in the subsequent process. As shown in Figure 6, in the constraint mode, the raster line tracking algorithm cannot track the entire target at one time. Therefore, the whole line can be realized through post-processing such as segmentation, capture, and polyline. However, this document introduces a fault-tolerant mechanism in the tracking process, which allows tracking in three directions on one side and accommodates a single outlier to achieve a more complete straight-line pattern.

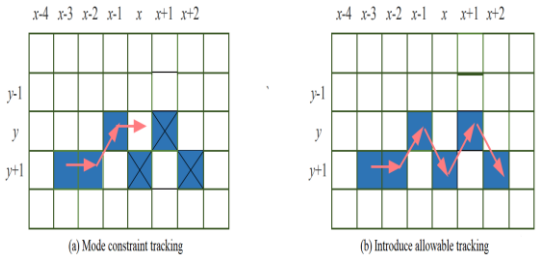

Figure 6: Straight-line tracking under non-ideal conditions

In addition, to achieve a complete line segment in one round of tracking and improve the efficiency of the algorithm, as shown in Figure 7, the line detection algorithm proposed in this paper starts to track the edges on both sides in six directions. Only two processes of edge extraction and straight-line tracking are required. These two processes do not require post-processing operations such as segmentation, capture, and connection in traditional edge point tracking algorithms.

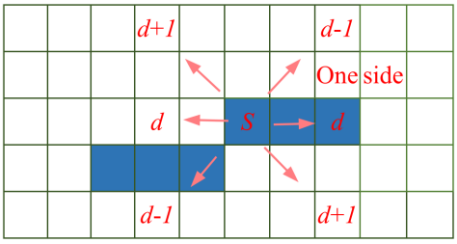

Figure 7: Tracking in six directions on both sides

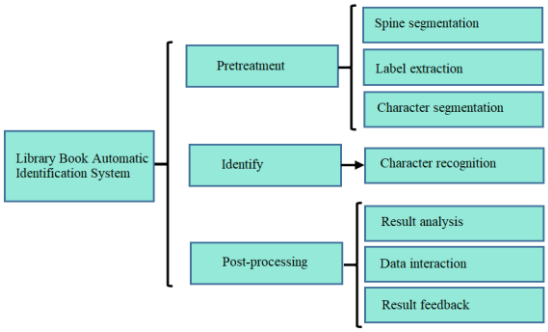

Figure 8: System function design

4. Result and discussion

4.1 Library book automatic identification system

System design scheme

1. Function description

As shown in Figure 8, the library book automatic recognition system is mainly composed of three modules: 1) Preprocessing module, which processes the image of the spine on the bookshelf from the outside. The main functions include trunk segmentation, label, telephone number, and character number extraction and segmentation;2) The recognition module is mainly used to recognize the character images in the preprocessing link, and then recognize the books on the shelf;3) Post-processing functions include result identification (clutter analysis), database interaction analysis, and result feedback. The pre-processing module is a server that processes the spine images collected by the administrator. The processing of the LFD proposed in the article is based on a line detection algorithm, which is used to detect the edge of the spine of a book; Extracting the phone number label from the label in the HSV space ROI is based on the characteristics of the label (such as aspect ratio, area, etc.); Phone number character segmentation, according to the projection method, it divides the phone number string in the ROI into a single character image sequence, which is used in the system recognition process.

In the recognition module, the server recognizes the character sequence of the phone number of each book by calling a pre-trained deep learning model (feedforward double convolutional network) and saves the recognition results for further processing.

The post-processing module analyzes and summarizes the past verification results. The library is an analysis of the verification results by interacting with the database management system. The system proposed in this paper currently only focuses on the recognition of books in the wrong order, which can improve the intelligent borrowing and reading functions. In addition, it is also important to provide feedback on the analysis results, and guide and assist the management staff in their work.

2. Development environment

In this work, some algorithms are written in the Microsoft Visual Studio 2013 environment using C++ language, which is efficient and stable. In addition, some functions in OpenCV are used for image processing, Caffe's deep learning architecture is used for book recognition, and MySQL is used for database management.

(1) Microsoft Visual Studio 2013

As an integrated solution, Microsoft Visual Studio 2013 is suitable for all groups, whether it is individuals or development teams of different sizes. It is currently the most popular integrated development environment for the Windows platform. As a development tool, it has fast application development, efficient teamwork, and innovative user experience. Excellent debugging functions are often used for visual programming.

(2) OpenCV

OpenCV is an open-source computer vision library, which covers the most advanced computer vision and cloud machine preparation sub-learning.OpenCV is not only compatible with interface languages such as C/C++, Python, and Matlab, but also compatible with multiple operating systems such as Windows, Linux, and Mac OS. Due to its high efficiency, openness, independence, and other advantages, it is very suitable for real-time image processing, so some functions of the OpenCV3.0 version are used in the development process.

(3) Caffe

Caffe (Conventional Architecture for Feature Extraction) is an expressive, innovative, efficient, and easy-to-extensible tool. It is highly appreciated by researchers in the design of deep learning network frameworks. Caffe focuses on machine vision, adopts C++/CUDA architecture, has command line, Python, and Matlab interfaces, and has a wealth of open-source models, demonstrations, and in-depth learning examples.

The most important feature is the ability to seamlessly switch between CPU and GPU.

(4) MySQL

MySql is a relational database, which uses the standard form of the SQL language. It is now one of the most popular relational data management systems. MySQL is small and fast. Since the code is open source, it also reduces the cost of ownership. The system is written in C and C++, creating a portable source code.

4.2 Detailed design of main functional modules

1. Preprocessing module

As mentioned above, the process is based on the edge analysis and direction-constrained line tracking algorithm proposed in this paper. The first algorithm attenuates the spine images collected by the image manager; On this basis, according to Canny's point of view, the boundary of the spine in the edge image is a straight line. Using the straight-line tracking algorithm proposed in this paper, the long straight-line segment is detected as a candidate, and the critical point is selected. As the candidate line segment for the last spine, the edge is used for segmentation; the initial image is segmented along the identified edge line of the vertebrae, and the rectangular frame around the book is removed. Because the book is tilted, the rectangular frame must be tilted to get the image of the spine of the book.

2. Identity Module

The recognition part only needs to call the network model trained through the C++ interface provided by Caffe. It is necessary to scale the single-character image sequence received by the preprocessing module to a specific size (28×28) network that satisfies the input conditions, and the average image is reduced (whitened) and then flows into the organization for further calculations to obtain the recognition result. Check the maximum probability in the softmaxlayer, and read the identification number of the maximum value, which is the result of input recognition.

3. Post-processing module

The post-processing module of the system is the process of analyzing, classifying, and returning the previously obtained verification results. This includes interaction with existing library databases. Due to unclearness and other reasons, part of the phone number is missing. This document uses regular expressions to perform fuzzy query processing on data to achieve higher accuracy. The specific implementation code is as follows:

```
for (int i=0; i<= book[j].length () +1; i= i+ book[j].length ()) {
str = book[j].insert(1," %");
}
mysql_query(con, "set names gbk");
query = "select booked from book where booked like "+str+" limit 1";
strcpy(q, query.c_str());
rt =mysql_real_query (con, q, strlen (q));
```

4.3 Test and analysis of automatic identification system of library books

Test Tools： In the testing process, to ensure the proper operation of the system and to provide a reliable basis for the analysis and processing of different data. The first step is to select the corresponding functional modules to be used. Then the sub-modules are connected as a whole according to the actual requirements and tested to verify their feasibility before starting the formal operation. This experiment mainly adopts two kinds of sensors, namely a reed switch and infrared sensor, to simulate the reader content and specific location information contained in the book image collected from the reader reading related information, as well as the identification and processing of the library internal staff through different channels in the library to the library shelves, paper, etc.; the automatic library book identification system based on image processing is designed according to different needs. Through this algorithm, the information of bookshelves and paper in the library is extracted effectively, so that the readers can get the satisfaction of their psychological feelings according to the requirements in the reading process. The experiment also analyzes the specific detection of book features and bibliographic contents by two sensor models, English RGB and SIFT, based on image pre-processing technology recognition. The experimental results show that the book recognition system based on image processing can effectively detect the contents involved in the reading process of readers. The algorithm can recognize the book information accurately and without errors.

System Performance： This paper designs an automatic library book recognition system based on image processing technology. By comparing various kinds of books and journals, it is selected among them which are suitable for the research and production use of this subject, the best economic applicability, and the conditions of the practical situation are generally applied to various fields. The algorithm has good performance in terms of hardware, and it can meet the laboratory real-time detection requirements. This can effectively solve the current problem of not being able to provide high-performance data transmission channels in a large-scale integrated environment; at the same time, the highly automated characteristics of image processing technology make the system highly flexible and adaptable to a variety of different scenarios and can be realized

according to different needs for various types of image information and retrieval results, etc., which makes the system can be well applied in a variety of application scenarios This enables the system to be used in a variety of application scenarios.

4.4 Compared with other current frameworks

The strength of our suggested method is its capacity to recognize the category of every book in a picture, regardless of how the volumes are stacked. Table 1 compares the current framework in use with other frameworks. Our technique Figure 11 achieves accurate categorization of book kinds, which goes beyond prior methods like the one described in Figure 9, which only focused on spine segmentation, and Figure 10, which expanded to spine segmentation and text extraction. Notably, current techniques were unable to identify the particular genre of books. They demonstrate the superiority of our suggested strategy over other current frameworks for book categorization through a comparative examination.

Table 1: Comparison of alternative frameworks currently in use

Methods	Book Spine Segmentation Accuracy (%)	Text Extraction Accuracy (%)	Book type categorization Accuracy (%)
Method 1 [20]	92	0	0
Method 2 [21]	95	86	0
Method 3 [22]	95	87	90
Proposed	97	94	95

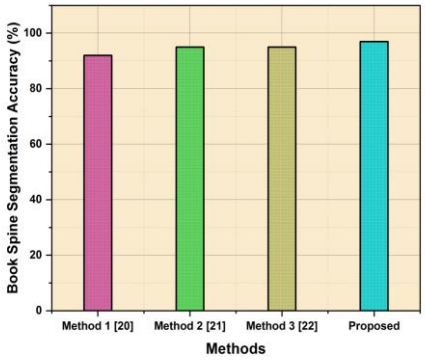

Figure 9: Accuracy of Book Spine Segmentation

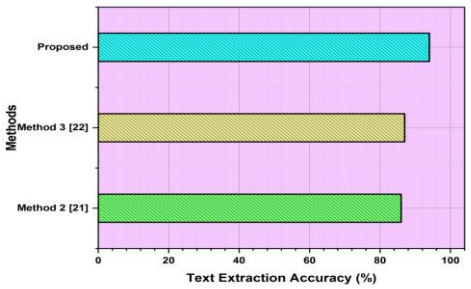

Figure 10: Accuracy of text extraction

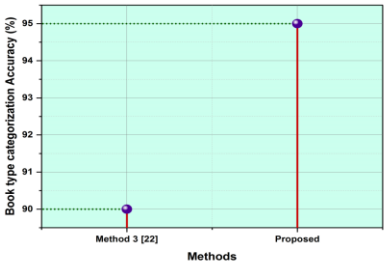

Figure 11: Accurate classification of book types

5 Discussion

The article would be greatly improved by include a thorough examination of the performance of the recommended system in a separate discussion section. To fully understand the distinctiveness and efficacy of the suggested approach, a comparison with analogous efforts is required. Gaining insight into the unique contributions of the proposed system requires an understanding of any observed performance differences. This improvement provides more than just a synopsis; it enables readers to explore the system's useful implications and applicability in a wider range of scenarios. This thorough explanation helps readers understand the system's importance and possible ramifications by clarifying its advantages and disadvantages in relation to current methods. In the end, this method provides a nuanced comprehension of the contributions made by the proposed system, enabling a better assessment of its advantages and disadvantages in practical situations.

6 Conclusions

Active library management uses modern management methods to provide borrowing and intelligent management, to adapt to changing times. Based on the technical status of the library, this research explores and analyzes the application of automatic identification technology in disordered books, and designs a specific system. Only the correct level can be guaranteed, the straight lines in the image can be developed faster, and

the noise reduction can be improved. The results of current experiments make this algorithm suitable for machine vision and image processing applications in terms of time, efficiency, and reliability. To build a fully automatic identification system, the construction and design of each link in the system must be completed. The modules in the currently designed automatic identification system include pre-processing, identification, and post-processing modules. Each module plays a different role in the entire system, but it is indispensable. This thesis researches the automatic library book recognition technology based on image processing, mainly by classifying different book labels to achieve the reader input reading mark on the shelf, and then get the desired information according to the corresponding characteristics. Through experiments and practical tests, certain results have been achieved. The experimental results show that the automatic book recognition system based on image processing is effective and efficient in label classification and information retrieval. Some of the limitations include the inability to segment incomplete images using conventional threshold approaches, the noise sensitivity of some edge detection algorithms, and possible inefficiencies in the fixed-threshold implementation of the Canny edge detection algorithm. Because it depends on particular sensors, the book recognition system could not be as flexible in different library settings.

Data availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest

Funding Statement

This study did not receive any funding in any form.

References

[1] Senkyire, I.B. and Liu, Z., 2021. Supervised and semi-supervised methods for abdominal organ segmentation: A review. *International Journal of Automation and Computing*, *18*(6), pp.887-914..

[2] Yi, P.A.N., Jin, L.I.U., Xu, T.I.A.N., Wei, L.A.N. and Rui, G.U.O., 2021. Hippocampal segmentation in brain mri images using machine learning methods: A survey. *Chinese Journal of Electronics*, *30*(5), pp.793-814..

[3] Fan, H., Sun, Y., Zhang, X., Zhang, C., Li, X. and Wang, Y., 2021. Magnetic-resonance image segmentation based on improved variable weight multi-resolution Markov random field in undecimated complex wavelet domain. *Chinese Physics B*, *30*(7), p.078703..

[4] Yi, P.A.N., Jin, L.I.U., Xu, T.I.A.N., Wei, L.A.N. and Rui, G.U.O., 2021. Hippocampal segmentation in brain mri images using machine learning methods: A survey. *Chinese Journal of Electronics*, *30*(5), pp.793-814..

[5] Zhang, Y. and Tian, Y., 2021. A new image segmentation method based on fractional-varying-order differential. *Journal of Beijing Institute of Technology*, *30*(3), pp.254-264..

[6] Li, Y., Blois, G., Kazemifar, F. and Christensen, K.T., 2021. A particle-based image segmentation method for phase separation and interface detection in PIV images of immiscible multiphase flow. *Measurement Science and Technology*, *32*(9), p.095208..

[7] Yongbin, Y.U., Chenyu, Y.A.N.G., Quanxin, D.E.N.G., Tashi, N., Shouyi, L. and Chen, Z., 2021. Memristive network-based genetic algorithm and its application to image edge detection. *Journal of Systems Engineering and Electronics*, *32*(5), pp.1062-1070..

[8] Liu, X. and Richardson, A.G., 2021. Edge deep learning for neural implants: a case study of seizure detection and prediction. *Journal of Neural Engineering*, *18*(4), p.046034..

[9] Gong, Y., Liu, Y. and Yin, C., 2021. A novel two-phase cycle algorithm for effective cyber intrusion detection in edge computing. *EURASIP Journal on Wireless Communications and Networking*, *2021*(1), pp.1-22..

[10] Xu, Z., Ji, X., Wang, M. and Sun, X., 2021, June. Edge detection algorithm of medical image based on Canny operator. In *Journal of Physics: Conference Series* (Vol. 1955, No. 1, p. 012080). IOP Publishing..

[11] Yu, T., 2021, February. Computer Application of Edge Detection Of Substrate Image Considering Fuzzy Microflora Optimization Algorithm. In *Journal of Physics: Conference Series* (Vol. 1792, No. 1, p. 012033). IOP Publishing..

[12] Gustavo Z. Felipe, Jacqueline N. Zanoni, Camila C. Sehaber-Sierakowski, Gleison D. P. Bossolani, Sara R. G. Souza, Franklin C. Flores, Luiz E. S. Oliveira, Rodolfo M. Pereira,Yandre M. G. Costa. Automatic chronic degenerative diseases identification using enteric nervous system images[J]. Neural Computing and Applications, 2021, 33(22):310-328.

[13] Bhardwaj, C., Jain, S. and Sood, M., 2021. Transfer learning based robust automatic detection system for diabetic retinopathy grading. *Neural Computing and Applications*, *33*(20), pp.13999-14019..

[14] Niu, Y., Fang, L., Sun, S., Jiang, J. and Chen, P., 2018. The Design of Book Sorter Base on Radio Frequency Identification. *Journal of Applied Science and Engineering Innovation*, *5*(1), pp.18-21..

[15] Ostler, K.R., Norlander, B. and Weber, N., 2021. Using open data to inform public library branch services. *Public Library Quarterly*, *40*(4), pp.365-377..

[16] Wang, X., Li, B. and Geng, Q., 2012, August. Runway detection and tracking for unmanned aerial vehicle based on an improved canny edge detection algorithm. In *2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics* (Vol. 2, pp. 149-152). IEEE..

[17] Chen, Q., Xiao, L. and Zhuang, S., 2012, May. A New Data Reduction Approach over the Stream Processor Architecture. In *2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum* (pp. 2300-2304). IEEE..

[18] Li, R., Wu, H., Liu, S., Rahman, M.A., Liu, S. and Kwok, N.M., 2018, April. Image edge tracking via ant colony optimization. In *Ninth International Conference on Graphic and Image Processing (ICGIP 2017)* (Vol. 10615, pp. 569-576). SPIE..

[19] Lu, J., Cai, Z., Yao, B. and Chen, B., 2020. A novel engagement-pixel image edge tracking method for extracting gear tooth profile edge. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, *234*(2), pp.405-416.):

[20] Jubair, M.I. and Banik, P., 2013, July. A technique to detect books from library bookshelf image. In 2013 IEEE 9th International Conference on Computational Cybernetics (ICCC) (pp. 359-363).

[21] Tabassum, N., Chowdhury, S., Hossen, M.K. and Mondal, S.U., 2017, February. An approach to recognize book title from multi-cell bookshelf images. In *2017 IEEE International Conference on Imaging, Vision & Pattern Recognition (icIVPR)* (pp. 1-6).

[22] Fatema, K., Ahmed, M.R. and Arefin, M.S., 2022. Developing a system for automatic detection of books. In *Second International Conference on Image Processing and Capsule Networks: ICIPCN 2021 2* (pp. 309-321). Springer International Publishing.