Retraction

Retracted: Analysis of Ancient Literary Works Based on Intelligent Image Text Recognition

Mobile Information Systems

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Fan, "Analysis of Ancient Literary Works Based on Intelligent Image Text Recognition," *Mobile Information Systems*, vol. 2022, Article ID 1113856, 9 pages, 2022.

Hindawi

Research Article

Analysis of Ancient Literary Works Based on Intelligent Image Text Recognition

Jia Fan (ORCID)

School of Literature and Journalism, Xihua University, Chengdu, Sichuan 610039, China

Correspondence should be addressed to Jia Fan; 1701100116@xy.dlpu.edu.cn

The analysis of ancient literary works in the era of digital intelligence needs to keep pace with the times. In order to improve the analysis effect of ancient literary works, this paper combines the intelligent image text recognition algorithm to extract the features of ancient literary works and proposes an intelligent algorithm that can be used for the analysis of ancient literary works. Moreover, this paper combines the analysis needs of literary works to improve the algorithm. In order to verify the role of the intelligent image text recognition algorithm proposed in this paper in the analysis of ancient literary works, this paper scans a large number of pictures of ancient literary works in the library by scanning to construct the experimental database of this paper. Finally, this paper combines experimental research to verify the algorithm proposed in this paper. From the experimental results, it can be seen that the method proposed in this paper has a certain effect, and it can be used as a reference for the digital processing and digital preservation of subsequent literary works, and it can also be used as a reference for the management of digital libraries.

1. Introduction

Human beings have always evolved and developed along with image culture and have deep historical origins with it. In ancient times, the ancestors generally used the method of physical records to record daily events. Recording in kind means that the primitive ancestors recorded figures in kind or expressed their opinions and feelings. This kind of method of remembering things in the image of real objects existed in large numbers in the ancient times of many nations or countries. They have a certain impact on other methods of note-taking and the invention of later writing. This primitive calculation method in kind is still used by many ethnic groups (including some ethnic minorities in our country) that are still in a primitive state in today's world. For example, when the Li ethnic minority group in Hainan Island settled accounts, every few tens of straws were put in one section to remember that human beings have always evolved and developed along with image culture. In ancient times, the ancestors generally used the method of physical records to record daily events. Recording in kind means that the primitive ancestors recorded figures in kind

or expressed their opinions and feelings. This method of remembering things in the image of objects existed in large numbers in the ancient times of many nations or countries [1]. They have a certain impact on other methods of note-taking and the invention of later writing. This primitive calculation method in kind is still used by many ethnic groups (including some ethnic minorities in our country) that are still in a primitive state in today's world.

In ancient times, the level of social productivity was extremely low, and the living conditions of primitive people were very difficult. Moreover, the living environment is extremely dangerous, social life and spiritual life are also in a barbaric and ignorant state, there is little production experience, and the level of labor skills and knowledge is very low. In order to survive, people have to use primitive and crude production tools to fight against nature. In order to exchange ideas and convey information, language came into being [2]. However, language is fleeting, neither can it be preserved nor can it be transmitted to distant places. Moreover, it is impossible to rely solely on human brain memory. As a result, the original method of note-taking, namely, pictures, naturally appeared. Before social

production and social relations have developed to the point that people feel that they must use language to record things or transmit information, they can only use pictures directly to represent things, and they do not think of using them to record the names of things—words in the language. Over time, these have become customary. With the passage of time, more and more such pictures have played a role in mutual communication [3].

This paper combines intelligent image text recognition technology, proposes an intelligent algorithm that can be used for the analysis of ancient literary works, combines the needs of literary works analysis to improve the algorithm, and combines experimental research to verify the algorithm of this paper.

2. Related Work

In order to solve the problem of binarization of degraded document images, scholars have devoted themselves to research and proposed many methods. Among the traditional threshold calculation methods, the representative algorithms of the global threshold method include the simple iterative method, Otsu algorithm, and histogram peaks algorithm. [4]. The global threshold method determines a threshold according to the gray value of the image and then divides the image into two parts, the foreground and the background, according to this value. The method is simple to implement and fast in execution. However, when the background noise of the image to be processed is complicated, the use of a fixed threshold may lose the foreground information or retain a large amount of noise information, which is obviously not the most ideal method for binarization. Therefore, a local threshold method based on histogram was born, among which the more representative algorithms mainly include the Niblack algorithm [5], Sauvola algorithm [6], Wolf algorithm [7], and other methods. Some scholars have proposed binarization algorithms based on local contrast, such as the Bernsen algorithm [8], a contrast calculation method based on local maximum and minimum gray levels, also known as the LMM algorithm [9], Gatos algorithm [10], and BESE algorithm [11]. Compared with the global threshold method that selects a single threshold, the local threshold effect is more accurate. However, because the threshold adjustment of this method varies according to the size of the sliding window, there may also be phenomena such as misjudgment of the foreground and background. In addition to some methods based on histograms, traditional threshold processing algorithms also appear to select thresholds based on image features, such as threshold segmentation algorithms based on image texture features [12]. This method first uses the Otsu algorithm to iteratively extract candidate thresholds and then extracts the texture features associated with each candidate threshold from the run-length histogram. Finally, the optimal threshold was selected to maintain the ideal document texture characteristics. Some scholars also consider the image as a three-dimensional terrain and put forward a very interesting thresholding idea. On this basis, a water flow model is proposed, which extracts characters from the

background through a thresholding method. Compared with histogram-based methods, feature-based methods can usually get more robust binarization results because high-resolution image features can be used to classify foreground and background pixels [13]. This type of algorithm is more suitable for images with relatively single noise information. Since the noise of the degraded document image is more complicated and the noise information of most images has a low contrast with the foreground text, it is difficult to effectively remove the noise without losing the strokes of the foreground text.

Using a single traditional threshold segmentation method for binarization processing, it is difficult to show ideal results on complex degraded document images. Therefore, some scholars combine a variety of image processing techniques and make full use of certain characteristics of the image to perform binarization processing. The main methods include global and local combination methods [14] and the edge detection method used in literature [15], [16] The proposed methods include background estimation method, gradient normalization and saliency map method [17], and Laplace energy method [18], and so on. This type of algorithm is more adaptable and can achieve ideal results for general image binarization problems. However, due to the complexity and diversity of the types of degraded document images, in the current existing multithreshold fusion algorithm, even if a nonfixed threshold is used for processing, it is difficult to remove the background without losing the foreground text. Noise information is removed cleanly.

The method based on statistical learning refers to the application of the method in mathematical statistics to the binarization of the image; that is, the problem is transformed into a clustering or classification problem. Among them, the representative methods based on statistical learning mainly include the support vector machine algorithm [19], k-means algorithm [20], and fuzzy c-means algorithm [21]. Under the premise of a large number of data sets, using this type of method to process general types of images with a single noise information has an ideal effect. However, various types of degraded document images with complex noise causes the processed binarized image to be disconnected from the foreground text, hollow strokes, and even information loss. This is a problem that is not allowed for some precious archive files. If you want to get better binarization results, you need to perform secondary processing on this basis. In addition, the model built by this type of method depends on the data set. If the processed image is similar to the data set, the effect is better; otherwise, the effect is poor. Therefore, the generalization ability of this type of algorithm is poor.

3. Intelligent Ancient Literature Image Text Recognition Algorithm

The global threshold refers to the process of selecting a threshold for an image and applying it to the entire image. The algorithm mainly traverses each pixel of the image and compares it with the selected threshold and uses a formula to determine the category of the point. The global threshold

method is simple to implement and fast in execution. However, if the background noise of the document image is complex, the effect of the global threshold will be unsatisfactory. The following mainly introduces Otsu as a representative algorithm.

The Otsu algorithm is also called the Otsu method, which is a clustering algorithm. According to the characteristics of the gray value of the image, the algorithm obtains the best threshold when the variance between classes is the largest, and then divides the image into two categories: background and foreground. If there is a wrong classification of the categories, this will cause the variance between the two categories to become smaller. Therefore, we have to choose an appropriate threshold to maximize the variance between classes and minimize the probability of class misclassification.

We use L to represent the number of gray levels of the image, and use $\sum_{i=1}^{L} p_i$ to represent the total number of image pixels. Among them, n_i represents the number of pixels when the gray level is i. Relative to the histogram, the probability of gray level i is given as

$$p_i = \frac{n_i}{\sum_{i=1}^{L} n_i}, \sum_{i=1}^{L} p_i = 1. \tag{1}$$

If it is assumed that the best threshold value is $T = K, T \in (I, L)$, then we can use this value to divide the image into target and nontarget categories, which are represented by C_1 and C_2 respectively. Among them, the gray value range of class C_1 is $[1, K]$, and the gray value range of class C_2 is $[k + 1, L]$. Then, the probability that a pixel may be divided into the foreground $p_1(k)$ and the background probability $p_2(k)$ can be given as

$$p_1(k) = \sum_{i=1}^{K} p_i,$$
$$p_2(k) = \sum_{i=k+1}^{L} p_i = 1 - p_1(k). \tag{2}$$

Then, the average gray values of pixels classified into foreground C_1 and background C_2 are, respectively, given as follows:

$$g_1(k) = \frac{1}{p_1(k)} \sum_{i=1}^{K} i p_i,$$
$$g_2(k) = \frac{1}{p_1(k)} \sum_{i=k+1}^{L} i p_i. \tag{3}$$

The overall gray value of the image can be expressed as

$$g_{avg}(k) = \sum_{i=1}^{L} i p_i. \tag{4}$$

The between-class variance of foreground and background is

$$\sigma_{12}^2(k) = p_1(k) p_2(k) [g_1(k) - g_2(k)]^2. \tag{5}$$

It can be obtained from formula (7) that when the difference between g_1 and g_2 is larger, the variance $\sigma_{12}^2(k)$ between foreground and background is also larger, and vice versa. Therefore, the between-class variance can be used as a measure of the separability between the foreground and the background, which represents the difference between the two categories. The Otsu algorithm mainly uses the normalized histogram to select the segmentation threshold, so the operation is simple, relatively easy to implement, and it is widely used in the image processing neighborhood. The following will show the results of degraded document images processed by the Otsu algorithm. Among them, Figure 1 is the original picture, and Figure 2 is the effect picture after the Otsu algorithm is used.

From Figure 2, we can see that when there are smudges around the text, that is, when the contrast between the text and the background area is low, the foreground text area will be misclassified as the background area. In addition, the Otsu algorithm uses a uniform threshold to process the entire image, so some noise information is still retained, and the background area is not cleanly processed. It can be seen that the Otsu algorithm can get better results when the foreground and background present high contrast, but when the target area and the background area have low contrast or when dark blocks are formed due to ink pollution, the results of the algorithm are not ideal. Therefore, the Otsu algorithm is not suitable for degraded document images with complex backgrounds.

When the contrast between the foreground and the background in an image is not constant due to ink, lighting, etc., the effect of the global threshold method will become unsatisfactory. Therefore, the threshold segmentation method based on the local characteristics of the image should be adopted. The local threshold algorithm is an algorithm that takes into account the different characteristics of the local area of the image and selects different optimal thresholds for segmentation in different areas. The local threshold algorithm determines the local area of the image to be processed according to the size of the sliding window. Therefore, the size of the sliding window has a great influence on the effect of the local threshold method. Because the foreground text size, font, and stroke of the degraded document image vary, it is necessary to adjust the size of the sliding window according to the characteristics of different images in order to obtain the best binarization effect. The following mainly introduces four representative local threshold binarization algorithms: Niblack algorithm, Sauvola algorithm, Bernsen algorithm, and LMM algorithm.

The Niblack algorithm is a simple and effective local binarization algorithm. Its main idea is to take a pixel as the center and calculate the best threshold in the area according to the point in its neighborhood. If it is assumed that there is a pixel point p on an image, and its coordinates are (x, y), then the threshold T_p of this point is given as

$$T_p(x, y) = \mu_p(x, y) + k * \sigma_p(x, y). \tag{6}$$

Among them, $\mu_p(x, y)$ represents the average gray value of all pixels within the size of the sliding window around the
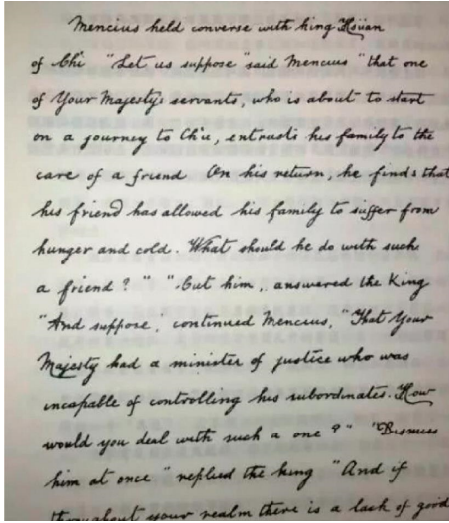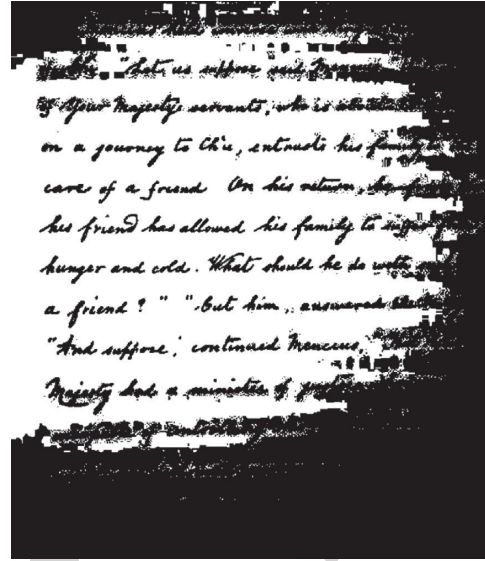
FIGURE 1: Original image.

FIGURE 2: The result after Otsu processing.
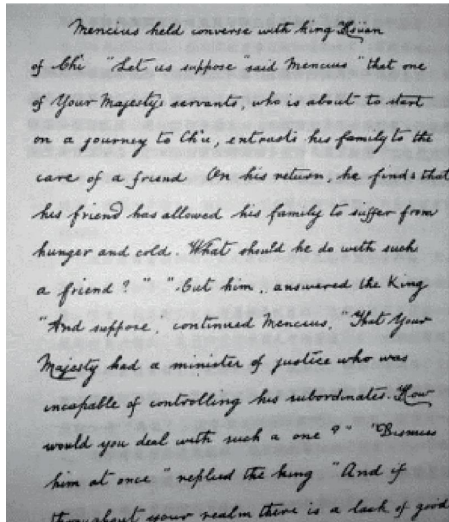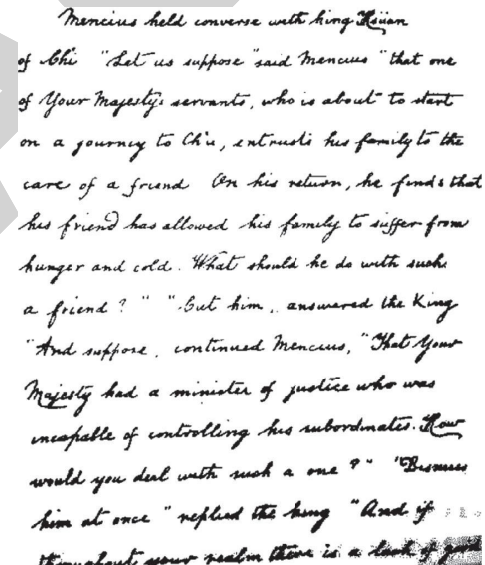

FIGURE 3: Original image.

FIGURE 4: Effect picture after processing with the Niblack algorithm.

pixel p, k is the coefficient for dynamically adjusting the threshold, and $\sigma_p(x, y)$ represents the standard deviation in the neighborhood. The following will show the result of the degraded document image processed by the Niblack algorithm. Among them, Figure 3 is the original picture, Figure 4 is the effect picture after using Niblack algorithm processing.

The Sauvola algorithm is a local threshold algorithm that uses the local mean as a benchmark and then fine-tunes the binarization effect according to the standard deviation. This algorithm can well solve the problem that the global threshold method cannot solve, that is, the binarization of the image with uneven illumination. The Sauvola algorithm is a local binarization algorithm for degraded document images, which is improved on the Niblack algorithm. If it is assumed that there is a pixel point P on the image, and its coordinates are (x, y), then the threshold T_p of this point is

$$T_p(x, y) = \mu_p(x, y) * \left[1 - \frac{K * \left(1 - \sigma_p(x, y)\right)}{R}\right]. \quad (7)$$

Among them, $C_p(x, y) = f_{\max}(x, y) - f_{\min}(x, y)$ represents the average gray value of all pixels within the size of the sliding window around the pixel p, k is a coefficient for dynamically adjusting the threshold, $\sigma_p(x, y)$ represents the standard deviation in the neighborhood, and R represents a constant. Generally, $R = 128$ and $k = 0.5$. When we choose the size of the sliding window to be $40 * 40$, the experimental results of the Sauvola algorithm are shown in the figure below. Among them, Figure 5 is the original picture, and
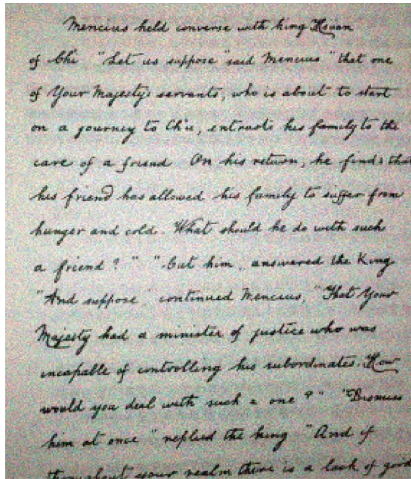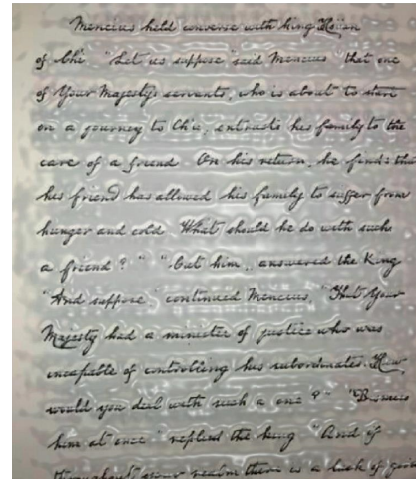
Figure 5: Original image.

Figure 6: Sauvola processed results.

Figure 6 is the effect picture after the Sauvola algorithm is used.

From the experimental results, it can be clearly seen that the Sauvola algorithm has a better improvement than the Niblack algorithm with the presence of large ink spots in the image. However, when the image contrast is low, a large number of noise points are generated around the font, as shown in Figure 6, or the foreground is misjudged as background pixels, resulting in the loss of foreground fonts.

The Bernsen algorithm is a binarization algorithm based on local contrast. The so-called local contrast refers to the assumption that there is a pixel P on the image. Its coordinates are (x, y) and the local contrast $C_p(x, y)$ of the pixel is given as

$$C_p(x, y) = f_{\max}(x, y) - f_{\min}(x, y). \tag{8}$$

Among them, $f_{\max}(x, y)$ represents the maximum gray value in the sliding window centered on the point P, and $f_{\min}(x, y)$ is the minimum gray value. The average gray value $T_p(x, y)$ in the neighborhood of this pixel is

$$T_p(x, y) = \frac{f_{\max}(x, y) - f_{\min}(x, y)}{2}. \tag{9}$$

First, we artificially set the value s, and $S = 15$ in the algorithm. For any pixel, the algorithm calculates the value of its local contrast. If the contrast value is greater than S, set the threshold of the current point as the average gray value in the neighborhood and perform threshold segmentation according to the threshold. If the value of contrast is less than the preset value S, indicating that the area is a background area, set the current gray value to 255. The experimental results of Bernsen algorithm are shown in the figure below. Among them, Figure 7 is the original picture, and Figure 8 is the effect picture after processing with Bernsen algorithm.

From the experimental results shown in Figure 8, we can see that the Bernsen algorithm can effectively solve problems such as traces. However, there are still a large number of misjudgment points, and too much noise is retained, and even part of the foreground area is misjudged as noise. In
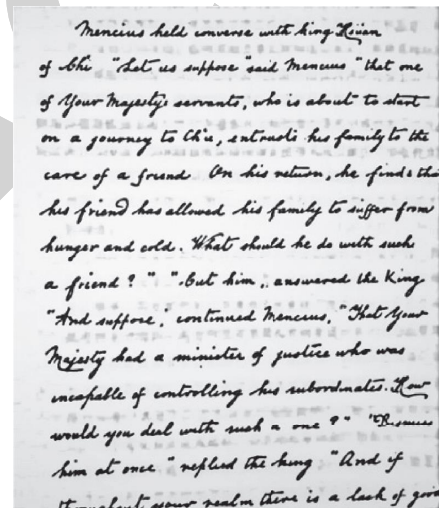

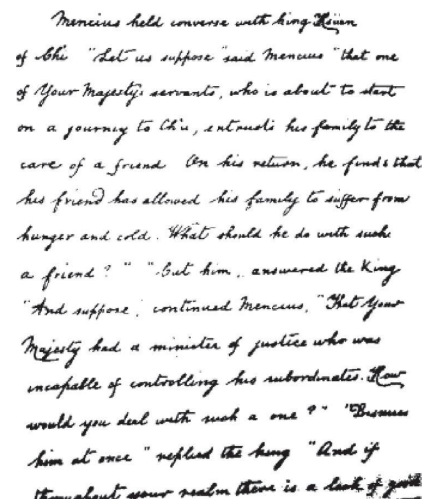
Figure 7: Original image.

Figure 8: Bernsen processing results.

addition, the algorithm takes too long to run, which is very time-consuming and not very practical when it needs to process large quantities of data.

The LMM method normalizes the contrast, which can avoid the influence of the appeal situation on the binarization effect. The definition of image contrast is as follows:

$$C_p(x, y) = \frac{f_{\max}(x, y) - f_{\min}(x, y)}{f_{\max}(x, y) + f_{\min}(x, y)}. \tag{10}$$

Among them, $f_{\max}(x, y)$ is the maximum gray value in the neighborhood of the sliding window centered on the pixel (x, y), and $f_{\min}(x, y)$ represents the minimum gray value. The denominator is the sum of the maximum gray value and the minimum gray value. This method is used to normalize the contrast of the image, which is also the difference in the definition of image contrast between the LMM algorithm and the Bernsen algorithm. The advantage of normalizing the contrast of the image is that no matter whether the image is dim or bright, it can get a similar contrast value. After that, the algorithm uses the Otsu method to extract high-contrast pixel values and then uses the following formula to perform threshold segmentation:

$$H(x, y) = \begin{cases} 0, & N_c > N_{\min} \& f_p \le \mu_p(x, y) + \dfrac{\sigma_p(x, y)}{2}. \\ \\ 1, & \text{others}. \end{cases} \tag{11}$$

The following parameters are for sliding windows. Among them, $N_{\min}$ is the minimum number of pixels in the sliding window, $C_p(x, y) = f_{\max}(x, y) - f_{\min}(x, y)$ is the average gray value of the pixels in the sliding window, $\sigma_p(x, y)$ is the standard deviation of the pixels in the sliding window, and N_c is the number of foreground pixels in the sliding window.

The support vector machine algorithm has good generalization ability for unknown samples, so the model can be applied to the fields of pattern recognition, regression estimation, probability density function estimation, time series forecasting, and so on. At the same time, it is also widely used in many fields such as handwritten digit recognition, text classification, image classification, and recognition in pattern recognition. The support vector algorithm is to find a hyperplane, through training samples, so that the sample points fall on both sides of the hyperplane, that is, divided into two categories. At the same time, the separation distance between the two sides to the hyperplane is maximized to distinguish the difference between the perceptron and the support vector machine algorithm. The problem of image binarization can actually be regarded as a pixel-level classification problem. Given a large amount of training data, the main method is to train a model for classification tasks. Among them, each pixel is assigned to a foreground or background label. Let us assume that we have N sample data as follows:

$$S = \{(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i)\}, i \in [1, N]. \tag{12}$$

Among them, x_i is the i-th instance, and y_i is the category of x_i, $y_i \in \{-1, +1\}$; that is, when binarization is performed, y_i may be the foreground or the background category. We assume that the hyperplane is $g(x)$: wx + b = o, and the distance d_i from the sample points on both sides to the hyperplane is

$$d_i = \frac{|g(x_i)|}{\|w\|}. \tag{13}$$

From this, a binary classification problem is transformed into

$$\min_{w, b} \frac{1}{2} \|w\|^2, \tag{14}$$

$$\text{s.t. } y_i(wgx_i + b) - 1 \ge 0, i = 1, 2, \dots, N.$$

The optimal solutions w^*, b^* are obtained by solving, and the hyperplane can be obtained from this as

$$w^* x + b^* = 0. \tag{15}$$

The classification decision function is

$$f(x) = \text{sign}(w^* x + b^*). \tag{16}$$

We directly use the support vector machine algorithm to process the degraded document image. Because the problems of degraded document images are more complex and diverse, the effect of directly performing linear classification binarization is not ideal. Therefore, Xiong et al. made improvements based on this algorithm and proposed a low-quality document image binarization based on support vector machines. In this algorithm, the support vector machine algorithm has been successfully applied to the processing of old degraded documents. The main method is to use the algorithm to preprocess the image, classify the image block, and achieve rough segmentation. After getting different types of image blocks, different methods are used to process, correct, and adjust details for different classification results. For an image, the algorithm divides it into three categories, but for some complex document images, the three categories may not be enough, and different follow-up threshold processing is required for different categories. Therefore, the generalization ability of the algorithm is not strong.

We assume that for the given n data sets $X = \{x_1, x_2, \dots, x_n\}$, the data set X is divided into C clusters. Among them, the center of each cluster is $m_j, 1 \le j \le C$, and the cluster loss function based on the membership function is defined as (17)

$$J_L = \sum_{j=1}^{C} \sum_{i=1}^{n} [\mu_j(x_i)]^b \|x_i - m_j\|^2. \tag{17}$$

Among them, i and j are the membership functions of the i-th sample corresponding to the j-th category, $\sum_{j=1}^{C} \mu_j(x_i) = 1$, and b represents the smoothing factor. Generally, $b = 2$.

We used J_L to obtain partial derivatives of m_j, μ_j, respectively and to make the result equal to 0. At this point, we get the following formula:

$$m_j = \frac{\sum_{i=1}^{n} \left[\mu_j(x_i)\right]^b g x_i}{\sum_{i=1}^{n} \left[\mu_j(x_i)\right]^b}, \mu_j(x_i) = \frac{\left\|x_i - m_j\right\|^{(2/b-1)}}{\sum_{s=1}^{C} \left\|x_i - m_j\right\|^{(2/b-1)}}. \tag{18}$$

Through the iterative solution, we get the optimal solution.

We use F-measure to directly evaluate the binarization process. This metric is suitable for two types of problems with unbalanced sample distribution. The F value is calculated as follows:

$$F = \frac{2 \times RC \times PR}{RC + PR}. \tag{19}$$

Among them, RC means recall rate, and PR means precision rate. The calculation methods are as follows:

$$RC = \frac{TP}{TP + FN},$$
$$pr = \frac{TP}{TP + Fp}. \tag{20}$$

Among them, TP is a case where it is judged to be a positive sample, but it is actually a positive sample. The meaning of FP is actually a negative sample, but it is classified as the number of positive samples. FN is actually a positive sample and is classified as a negative sample. The F value is a measure of the number of correct foreground pixel values in the generated result. The higher the F value, the higher the accuracy of the algorithm.

3.1. Peak Signal-to-Noise Ratio (PSNR). The calculation method of peak signal-to-noise ratio is as follows:

$$PSNR = 10 \log \left(\frac{C^2}{MSE}\right),$$
$$MSE = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} (I(x, y) - K(x, y))^2}{MN}. \tag{21}$$

Among them, M represents the length of the image, N represents the width of the image, $I(x, y)$ and $K(x, y)$, respectively, represent the original image and the true value image, and C represents the absolute error value between the foreground and the background. The peak signal-to-noise ratio was originally used in the communication field to represent the ratio between the maximum possible power of a signal and the noise power that affects the signal. We now apply it to the algorithm to measure the processing of degraded document images to calculate the similarity between the document image processed by the algorithm and the true value image. The larger the ratio, the better the performance of the algorithm, and the closer the binarization result is to the true value image.

The degree of distortion is inversely proportional to the distance between adjacent pixels. The distance on the diagonal is larger than the distance between the horizontal and vertical directions, so the distortion caused by changing the pixels on the diagonal is smaller than the horizontal or vertical direction. Therefore, the reciprocal distance distortion metric can be used as a criterion for judging the binarization result of degraded documents. The calculation method is as follows:

$$DRD = \frac{\sum_{k=1}^{S} DRD_k}{NUBN}. \tag{22}$$

Among them, NUBN represents the number of image blocks that are not completely black or not completely white in the Ground Truth image. DRD_k represents the weighted sum of distortion of the K-th flipped pixel, and the calculation method is as follows:

$$DRD_k = \sum_{i=-2}^{2} \sum_{j=2}^{2} \left|GT_k(i, j) - B_k(x, y)\right| g W_{NM}(i, j). \tag{23}$$

Among them, W_{NM} represents the weight matrix, $GT_k(i, j)$ represents the pixel value of the K-th pixel at the flip center pixel point (x, y) in the Ground Truth image, and $B_k(x, y)$ is the k-th pixel in the image to be measured at the pixel (x, y). The smaller the reciprocal distance distortion metric, the smaller the distortion and the better the algorithm performance.

The error rate metric is calculated as follows:

$$NRM = \frac{NR_{FN} + NR_{FP}}{2}. \tag{24}$$

Among them, NR_{FN} represents the false negative rate, and NR_{FP} represents the false positive rate. The calculation methods are as follows:

$$NR_{FN} = \frac{FN}{FN + TP},$$
$$NR_{FP} = \frac{FP}{FP + TN}. \tag{25}$$

Among them, TN refers to the situation where the actual result and the judgment result are both negative samples. The error rate measurement also compares the difference between the image processed by the algorithm and the Ground Truth image. However, the error rate measurement mainly focuses on the ratio of false matches. Therefore, the smaller the calculation result, the better the algorithm performance.

4. Analysis of Ancient Literary Works Based on Intelligent Image Text Recognition

In order to verify the role of the intelligent image text recognition algorithm proposed in this paper in the analysis of ancient literary works, a large number of pictures of ancient literary works are scanned in the library by scanning, and an experimental database is constructed. After that, this

TABLE 1: Statistical table of the analysis effect of ancient literary works based on intelligent image text recognition.

No	Text feature recognition	Literary analysis
1	93.56	76.82
2	92.23	76.28
3	95.73	76.40
4	96.73	79.06
5	93.32	86.39
6	96.94	77.57
7	94.06	74.32
8	96.58	76.42
9	94.88	84.43
10	94.06	80.55
11	97.78	78.35
12	92.85	80.04
13	97.14	78.04
14	93.67	79.98
15	97.66	87.95
16	95.76	74.34
17	97.81	85.82
18	96.27	86.41
19	93.14	80.78
20	93.90	78.80
21	94.70	86.67
22	97.36	88.34
23	95.88	84.65
24	97.18	86.21
25	97.29	75.04
26	93.35	81.64
27	95.07	79.78
28	94.65	85.94
29	92.42	85.69
30	97.38	74.48
31	92.90	86.89
32	93.15	82.49
33	96.48	80.40
34	92.11	80.31
35	95.45	86.53
36	93.70	77.92
37	95.83	77.72
38	93.68	83.26
39	94.16	84.45
40	95.45	83.66
41	92.38	85.04
42	97.43	86.89
43	93.53	86.54
44	95.27	81.42
45	96.30	76.33
46	97.60	78.75
47	93.32	85.46
48	94.69	81.35
49	92.33	84.80
50	95.67	86.98
51	94.64	84.85
52	92.82	78.34
53	93.93	86.57
54	92.53	85.14
55	95.07	88.14
56	97.81	81.00
57	94.27	81.57
58	96.15	85.73
59	97.96	81.52
60	92.52	80.27

paper randomly combines these database files to obtain multiple sets of image collections of ancient literary works. At the same time, this paper recognizes these images, counts the recognition accuracy and work analysis effect of the intelligent image text recognition algorithm proposed in this paper, and obtains the results shown in Table 1.

From the above research, it can be seen that the intelligent image-based text recognition method proposed in this paper can play an important role in the analysis of ancient literary works, and it can be used as a reference for the digital processing and digital preservation of subsequent literary works. At the same time, it can also be used as a reference for the management of digital libraries.

5. Conclusion

The form of literary works began to generalize, and the literary form showed a trend of diversification. As we all know, literature in the traditional sense is now fading from the sacred aura of the superstructure in the "image age," and its era as the darling of the art temple is gone forever. Moreover, it will gradually become a new industry in the commodity society, seeking survival and development like other industries. Facing the increasingly fierce competition and challenges, the pure and serious traditional literature began to decline. The real life of the novel and the creation and fiction are obviously out of balance, the spiritual sublimation and the ideological connotation are mostly poor, and the concern for human nature is gradually disappearing, and the quality is beginning to be rapidly vulgar. This paper combines intelligent image and text recognition technology to propose intelligent algorithms that can be used for the analysis of ancient literary works and improves the algorithm based on the needs of literary works analysis. Finally, this paper combines experimental research to verify the algorithm of this paper. From the experimental results, we can see that the method proposed in this paper has a certain effect.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] J. Zhao, X. Xu, H. Jiang, and Y. Ding, "The effectiveness of virtual reality-based technology on anatomy teaching: a meta-analysis of randomized controlled studies," *BMC Medical Education*, vol. 20, no. 1, pp. 127–136, 2020.

[2] S. J. Bennie, K. E. Ranaghan, H. Deeks et al., "Teaching enzyme catalysis using interactive molecular dynamics in virtual reality," *Journal of Chemical Education*, vol. 96, no. 11, pp. 2488–2496, 2019.

[3] S. F. M. Alfalah, J. F. M. Falah, T. Alfalah, M. Elfalah, N. Muhaidat, and O. Falah, "A comparative study between a virtual reality heart anatomy system and traditional medical

teaching modalities," *Virtual Reality*, vol. 23, no. 3, pp. 229–234, 2019.

[4] M. Reymus, A. Liebermann, and C. Diegritz, "Virtual reality: an effective tool for teaching root canal anatomy to undergraduate dental students–a preliminary study," *International Endodontic Journal*, vol. 53, no. 11, pp. 1581–1587, 2020.

[5] V. L. Dayarathna, S. Karam, R. Jaradat et al., "Assessment of the efficacy and effectiveness of virtual reality teaching module: a gender-based comparison," *International Journal of Engineering Education*, vol. 36, no. 6, pp. 1938–1955, 2020.

[6] O. Hernandez-Pozas and H. Carreon-Flores, "Teaching international business using virtual reality," *Journal of Teaching in International Business*, vol. 30, no. 2, pp. 196–212, 2019.

[7] V. Andrunyk, T. Shestakevych, and V. Pasichnyk, "The technology of augmented and virtual reality in teaching children with ASD," *Econtechmod: Scientific Journal*, vol. 4, no. 7, pp. 59–64, 2018.

[8] R. Mayne and H. Green, "Virtual reality for teaching and learning in crime scene investigation," *Science & Justice*, vol. 60, no. 5, pp. 466–472, 2020.

[9] C. Yang, S. Huan, and Y. Yang, "A practical teaching mode for colleges supported by artificial intelligence," *International Journal of Emerging Technologies in Learning (IJET)*, vol. 15, no. 17, pp. 195–206, 2020.

[10] K. Kim and Y. Park, "A development and application of the teaching and learning model of artificial intelligence education for elementary students," *Journal of The Korean Association of Information Education*, vol. 21, no. 1, pp. 137–147, 2017.

[11] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education–where are the educators?" *International Journal of Educational Technology in Higher Education*, vol. 16, no. 1, pp. 39–27, 2019.

[12] S. C. H. Yang, W. K. Vong, R. B. Sojitra, T. Folke, and P. Shafto, "Mitigating belief projection in explainable artificial intelligence via Bayesian teaching," *Scientific Reports*, vol. 11, no. 1, pp. 1–17, 2021.

[13] Y. Lee 서울영도초등학교, "An analysis of the influence of block-type programming language-based artificial intelligence education on the learner's attitude in artificial intelligence," *Journal of The Korean Association of Information Education*, vol. 23, no. 2, pp. 189–196, 2019.

[14] J. M. Alonso, "Teaching explainable artificial intelligence to high school students," *International Journal of Computational Intelligence Systems*, vol. 13, no. 1, pp. 974–987, 2020.

[15] R. Yang, "Artificial intelligence-based strategies for improving the teaching effect of art major courses in colleges," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 22, pp. 146–160, 2020.

[16] P. Ribeiro, H. Simões, and M. Ferreira, "Teaching artificial intelligence and logic programming in a competitive environment," *Informatics in Education*, vol. 8, no. 1, pp. 85–100, 2009.

[17] V. N. Ganesh, S. G. Acharya, S. Bhat, and S. V. Yashas, "Machine vision robot with real time sensing," *Journal of Advancements in Robotics*, vol. 1, no. 3, pp. 30–34, 2020.

[18] H. H. Chu and Z. Y. Wang, "A study on welding quality inspection system for shell-tube heat exchanger based on machine vision," *International Journal of Precision Engineering and Manufacturing*, vol. 18, no. 6, pp. 825–834, 2017.

[19] N. Noguchi, "Agricultural vehicle robot," *Journal of Robotics and Mechatronics*, vol. 30, no. 2, pp. 165–172, 2018.

[20] V. Villani, F. Pini, F. Leali, C. Secchi, and C. Fantuzzi, "Survey on human-robot interaction for robot programming in industrial applications," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 66–71, 2018.

[21] S. Huang, K. Shinya, N. Bergström, Y. Yamakawa, T. Yamazaki, and M. Ishikawa, "Dynamic compensation robot with a new high-speed vision system for flexible manufacturing," *International Journal of Advanced Manufacturing Technology*, vol. 95, no. 9-12, pp. 4523–4533, 2018.