

Image Processing Applications in Smart Contracts: Automated Financial Transaction Verification



Meiling Lu¹, Yingxia Yang^{2*}

¹ School of Business Administration, Guangzhou Civil Aviation College, Guangzhou 510403, China

² The Affiliated Guangdong Second Provincial General Hospital of Jinan University, Guangzhou 510403, China

Corresponding Author Email: yyxgoodluck@163.com

Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.410422>

ABSTRACT

Received: 22 March 2024

Revised: 14 July 2024

Accepted: 26 July 2024

Available online: 31 August 2024

Keywords:

smart contracts, image processing, financial transaction verification, automation, table detection, text extraction

With the widespread adoption of smart contracts in automated financial transactions, the accurate and efficient processing of image data related to financial transactions has become a critical challenge. The successful execution of smart contracts relies on the precise verification of transaction voucher images, yet existing image processing technologies still face limitations in dealing with background complexity, noise interference, and text extraction accuracy. To address these issues, this study proposes a comprehensive image processing approach aimed at enhancing the automation of financial transaction verification. The research focuses on four key areas: separation of table lines and text regions in images, application of Sauvola local adaptive binarization, table detection and reconstruction, and text extraction and fracture restoration techniques. Through these efforts, the study aims to provide more efficient and reliable technical support for financial transaction verification in smart contracts, thereby advancing the development of smart contract technologies.

1. INTRODUCTION

With the rapid development of digital technology, the application of smart contracts in automated financial transactions has become increasingly widespread [1-3]. Smart contracts automatically execute contract terms through predefined programming logic, significantly improving transaction efficiency and transparency [4]. However, the effectiveness and security of smart contracts largely depend on accurate data input and verification, especially when processing images related to financial transactions [5-8]. Since financial transaction vouchers are often stored and transmitted in image form, the automated and precise verification of these image data has become a key challenge in the application of smart contracts.

In existing research, image processing technology has been widely applied in various fields, but its application in smart contracts is still in the exploratory stage [9-12]. For financial transaction verification, the application of image processing can effectively enhance the automation level of smart contracts, reduce human intervention, and lower error rates [13, 14]. This not only helps optimize financial transaction processes but also promotes the wider adoption and application of blockchain technology and smart contracts. Therefore, the study of the application of image processing technology in smart contracts has significant theoretical and practical value.

Although some studies have proposed application schemes for image processing in financial transactions, most methods face certain shortcomings in practical application [15-17]. Existing methods still encounter considerable challenges in

handling complex image backgrounds, eliminating noise interference, and accurately extracting text regions, particularly when dealing with large-scale data processing, where efficiency and accuracy issues are especially prominent [18, 19]. In addition, the accuracy of table detection and reconstruction needs to be improved, and text extraction and fracture restoration techniques in images require further optimization.

This paper proposes a comprehensive image processing method to address the above issues, with a focus on automated financial transaction verification, covering the following four main research contents: First, the study investigates the separation technology of table lines and text regions in images; second, it explores the application of the Sauvola local adaptive binarization method in financial transaction images; third, it develops a high-precision table detection and reconstruction algorithm; finally, it studies text extraction and fracture restoration techniques in images. Through these studies, this paper aims to provide more reliable and efficient technical support for automated financial transaction verification in smart contracts, which holds important academic and practical value.

2. SEPARATION OF TABLE LINES AND TEXT REGIONS IN IMAGES FOR AUTOMATED FINANCIAL TRANSACTION VERIFICATION

As a digital tool for the automated execution of contract terms, smart contracts rely on accurate data input and verification, particularly in scenarios involving financial

forms and receipt images. Figure 1 provides a binarization illustration of financial forms and receipt images where table and text regions are not separated. Financial forms and receipt images are common data carriers in financial transactions within smart contracts, and the information in these images is directly related to the accuracy and legality of the transactions. Therefore, accurately detecting, analyzing, and reconstructing tables from images is an important foundation for ensuring the accuracy of financial transaction verification. Traditional manual processing methods are not only inefficient but also prone to human errors, which cannot meet the high requirements of modern smart contracts for real-time processing and accuracy. To address this, this paper proposes a series of image processing techniques aimed at enhancing the automated verification capabilities of financial forms and receipt images.

Traditional binarization algorithms often struggle to handle the specific structures and complex backgrounds in these financial forms and receipt images. To address this, this paper improves traditional binarization algorithms to better meet the processing needs of financial forms and receipt images. Specifically, the approach first applies a local adaptive binarization method to binarize the image, which can adaptively adjust the binarization threshold based on the brightness variations in different regions of the image, thereby better preserving important details in the image. After binarization processing, a complete Canny edge image is obtained, and connected component analysis is performed on the Canny edge binary image. By analyzing the connected component information obtained, further merging and filtering operations are carried out on these connected components. Specifically, the width and height of the connected components are statistically analyzed to identify three distinct regions: small connected components such as broken strokes or small lines, large connected components such as table lines or long lines, and connected components similar in size to text. Then, by setting rules, large and small connected components are filtered out, retaining only connected components similar in size to text, which corresponds to the text portions of the original document. Finally, the remaining connected components are further merged, combining the strokes within the text into a complete connected component. Through this improved binarization processing method, this paper can more effectively separate the text regions in financial forms and receipt images while preserving the table structure, providing a more accurate data foundation for subsequent table detection and reconstruction, text extraction, and fracture restoration.



Figure 1. Binarization illustration of financial forms and receipt images where table and text regions are not separated

Since table document images often contain a large number of table lines and long straight lines, failing to filter these elements before binarization processing will have a significant negative impact on subsequent image processing steps. Specifically, the connected components formed by table lines and long straight lines in the image usually have a large width and height. If these connected components are not filtered out, the calculated average width and height of the connected components will deviate from the actual text size. This deviation will directly affect the subsequent connected component merging process, causing multiple independent text characters to be incorrectly merged into a whole, thus treating several characters as a single unit during local area binarization. Additionally, if the entire table and its internal text are merged into one connected component, the subsequent local area binarization will lose its effectiveness and become a global binarization process for the entire table region. This will not only destroy the details of the image but also severely affect the accuracy of text region extraction, ultimately having an adverse impact on the automated verification of smart contracts. Therefore, before the binarization processing of financial forms and receipt images, this paper prioritizes the filtering of table lines and long straight lines. By removing these interfering elements, this paper ensures that the average width and height of the connected components are closer to the actual text size, thus enabling accurate processing of each independent text character in subsequent local area binarization. The expressions for filtering out tables and long straight lines are as follows:

$$\delta_v^2 = \left(\sum (q_\pi - q_\delta)^2 \right) / v \quad (1)$$

$$|q_\pi - q_\delta| < 3\delta_v \quad (2)$$

$$\delta_g^2 = \left(\sum (g_m - g_\delta)^2 \right) / v \quad (3)$$

$$|g_\pi - g_\delta| < 3\delta_g \quad (4)$$

$$T_{zz1} \cap T_{zz2} \neq \emptyset \quad (5)$$

$$|Q_{zz1+zz2} / \text{MAX}(q_{zz1}, q_{zz2}) - 1| < \pi_g \quad (6)$$

$$|G_{zz1+zz2} / \text{MAX}(g_{zz1}, g_{zz2}) - 1| < \pi_g \quad (7)$$

In the process of automated financial transaction verification, processing financial forms and receipt images requires balancing efficiency and accuracy. This paper provides two significant advantages in meeting this demand by extracting connected components from the Canny edge image. First, the merging of connected components in the Canny edge image can be equivalent to the merging of connected domains in the original image. This means that in subsequent local area binarization processing, the processing effect obtained using the connected components from the Canny edge image is consistent with that obtained by directly using the connected domains in the original image. This method not only ensures the accurate separation of table lines and text regions in the image but also preserves the key structural information in the image, improving the accuracy of

automated financial transaction verification. Second, extracting connected components from the Canny edge image has a lower time cost. Since the Canny edge image is already the edge representation of the connected domains in the original image, detecting each point during processing can quickly obtain the corresponding connected components. This method significantly reduces computation time, improving image processing efficiency while ensuring processing effectiveness. Figure 2 shows a comparison of the binarization before and after the separation of table and text regions in financial forms and receipt images.



Figure 2. Comparison of binarization before and after separation of table and text regions in financial forms and receipt images

Furthermore, by marking and counting the connected components in the Canny edge image, this paper proposes an efficient method for separating table lines and text regions in financial forms and receipt images. First, a preliminary statistical analysis of the width and height of all connected components is performed, identifying three distinct regions: small connected components, large connected components, and connected components similar in size to text. Through this statistical analysis, rules are set to filter out connected components that do not match the text size, thereby retaining only the connected components close to the average text size. This step ensures that the connected components in the subsequent processing are mainly concentrated in the text regions, avoiding interference from non-text regions such as table lines. After completing the preliminary filtering, the remaining connected components are further merged, combining the strokes of each character into a complete connected component. The result of this operation is that the final local area binarization is only performed within each individual text region, ensuring the accuracy of the binarization processing.

3. SAUVOLA LOCAL ADAPTIVE BINARIZATION FOR AUTOMATED FINANCIAL TRANSACTION VERIFICATION IMAGES

In the process of automated financial transaction verification, when processing financial forms and receipt images, complex situations such as glare and changes in shooting angles are often encountered. To effectively address these image characteristics, this paper employs a binarization method that combines Canny edge detection with Sauvola local adaptive binarization. First, Canny edge detection is applied to the original image to extract the edge information of the image. This step effectively identifies the main structures in the image, such as the edges of text and table lines. Subsequently, connected component analysis is performed on the extracted Canny edge image to identify connected regions related to text and tables. Based on this, the Sauvola local adaptive binarization method is further applied. The Sauvola algorithm dynamically calculates the mean grayscale value and standard deviation of the neighboring region for each pixel to determine the binarization threshold

for the local area.

The Sauvola method is an improvement based on the Niblack algorithm. The Niblack method traverses each pixel in the financial forms and receipt images, with each pixel corresponding to a window. Suppose the mean grayscale value of the window pixels is denoted by l , the variance of the window pixels' grayscale is denoted by t , and the constant is denoted by j . The following formula gives the value calculation for the window center point:

$$S = l + j \cdot t \tag{8}$$

Suppose the dynamic range of the standard deviation is denoted by E , the current pixel coordinates are (a,b) , the area centered on this point is $e \times e$, and the grayscale value at (a,b) is denoted by $h(a,b)$. In the Sauvola algorithm, the pixel threshold at the window center point can be obtained by the following formula:

$$S(a,b) = l(a,b) + \left\{ 1 + j \cdot \left[t(a,b)/E - 1 \right] \right\} \tag{9}$$

The Sauvola algorithm can be represented by the following formulas:

$$l(a,b) = 1/e^2 \times \sum \sum h(a,b) \tag{10}$$

$$t(a,b) = \left[1/e^2 \times \sum \sum (h(a,b) - l(a,b))^2 \right]^{1/2} \tag{11}$$

4. TABLE DETECTION AND RECONSTRUCTION FOR AUTOMATED FINANCIAL TRANSACTION VERIFICATION IMAGES

In the process of automated financial transaction verification, accurately identifying and reconstructing horizontal lines in financial tables is crucial for ensuring the correct parsing of table data. This paper implements accurate detection and reconstruction of table lines through the BAG table horizontal search method, thereby enhancing the accuracy of automated verification. Figures 3 and 4 respectively show the schematic diagram of the BAG image and the BAG line detection.

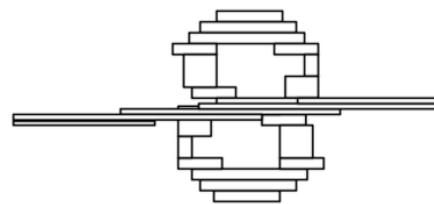


Figure 3. Schematic diagram of BAG

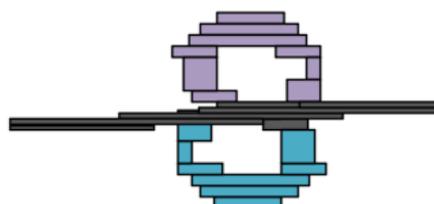


Figure 4. Schematic diagram of BAG line detection

This paper extracts the width of text strokes from the connected components and uses it as an approximation for the line width in the processing. This approximate processing method effectively meets the needs of subsequent line processing. After obtaining the average height of the lines, a BAG run block is selected as a candidate region that satisfies the width-to-height ratio greater than 10 and the height less than twice the approximate height of the line. This selection ensures that the chosen run block has certain linear characteristics, which is beneficial for subsequent line detection.

$$Y_i/Y_g > \pi_m \quad (12)$$

After finding the BAG run blocks that meet the above requirements, this paper conducts an adjacent run block extension search above and below these run blocks. This extension search confirms the continuity of the lines, avoiding omissions of broken or discontinuous lines, thereby ensuring the integrity of the lines. The extension must meet the following formulas:

$$(Y1_{HO} - Y2_{aT}) \cdot (Y2_{HO} - Y1_{aT}) = 0 \quad (13)$$

$$MAX(Y1_{aL} - Y2_{aR}, Y2_{aL} - Y1_{aR}) < 0 \quad (14)$$

This paper performs a traversal search of each BAG run block, ultimately obtaining a possible set of lines in the table. It should be noted that this set of lines not only includes complete lines in the table but may also include some broken line segments and even some lines that are not table lines. This means that the initially obtained set of lines needs further processing to eliminate interference and identify the true table lines.

$$TY_q = MAX(Y1_{aR}, Y2_{aR}) - MIN(Y1_{aL}, Y2_{aL}) \quad (15)$$

$$TY_g = MAX(Y1_{sT}, Y2_{bT}) - MIN(Y1_{bB}, Y2_{bB}) \quad (16)$$

$$TY_q/TY_g > \pi_m \quad (17)$$

After obtaining the set of lines that includes horizontal lines, this paper further merges adjacent short lines that clearly belong to the same line. Through this merging process, the impact of broken lines can be effectively reduced, connecting multiple short lines into complete horizontal lines in the table, thereby improving the accuracy of table reconstruction.

$$|MIN(Y1_{aW} - Y2_{aL}, Y2_{aW} - Y1_{aL})| < Q_{STR} \quad (18)$$

Finally, when the angle of inclination of the lines is relatively small, this paper introduces the analysis of horizontal spacing error. By analyzing the inclination angle of the lines in the BAG run block, it is determined whether it meets a certain tolerance, and merging is performed according to the requirements shown in the following formula. This ensures that even in the presence of slight inclinations, the horizontal lines of the table can still be correctly identified and reconstructed. Figure 5 shows the method for determining the inclination angle of the lines. If the inclination angle of the line

is determined by any two points among F1, F2, F3, and F4, then X and Y need to satisfy the following formula to allow merging.

$$|MIN(Y1_{aW} - Y2_{aL}, Y2_{aW} - Y1_{aL})| < Q_v \quad (19)$$

$$\varphi_v = \arctan((b_1 - b_2)/(a_1 - a_2)) \quad (20)$$

$$\delta_\varphi^2 = \left(\sum(\varphi_v - (\sum\varphi_v)/v)^2 \right) / v < \pi_{\delta^2} \quad (21)$$

$$MAX(Y1_{aL} - Y2_{aR}, Y2_{aL} - Y1_{aR}) / \cos((\sum\varphi_v)/v) < Q_\delta \quad (22)$$

Filtering all merged lines yields the set of table lines in financial forms and receipt images.

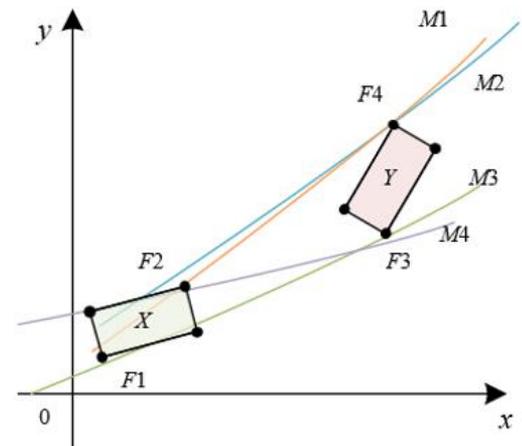


Figure 5. Method for determining the inclination angle of lines

Similar to the detection of horizontal lines, the detection of vertical lines also requires retrieval and merging from the BAG run block set. However, since the BAG run blocks usually present horizontal features in table images, the vertical search poses some unique challenges and steps. First, a BAG run block is selected as a base block from the BAG run block set. This base block typically has obvious horizontal features, and its width can well reflect the horizontal characteristics of the BAG run block. Subsequently, based on this base block, this paper begins searching upwards and downwards to find other run blocks that may connect to form a vertical line. To avoid interference from horizontal features, special attention must be paid during vertical retrieval to whether these candidates run blocks meet certain vertical feature requirements. Furthermore, this paper introduces the concept of "BAG run block vertical height," which refers to the total height of all run blocks connected to the base block within the width range of the base run block. This vertical height can more comprehensively describe the vertical characteristics of the run block, compensating for the shortcomings of relying solely on the run block height for detection. During the retrieval process, only run blocks that meet certain vertical height characteristics are selected as potential vertical line candidates, thus avoiding misjudgment and noise interference.

After completing the retrieval of the upper and lower run blocks, this paper merges the run blocks that meet the conditions. This merging process must ensure that the selected

run blocks can form continuous vertical line segments while avoiding merging errors caused by the horizontal features of the run blocks. In practice, this paper evaluates the vertical height and connectivity of the run blocks to ensure that the merged vertical line set has a high degree of accuracy. After obtaining the preliminary set of vertical lines, this paper further filters and optimizes the vertical lines. Considering that some vertical lines may be broken or discontinuous, this paper analyzes the inclination angle and overall connectivity of the vertical lines, merging short segments that may belong to the same vertical line to ensure that the final set of vertical lines is complete and coherent. By verifying and correcting the vertical line set, this paper further enhances the accuracy of table reconstruction.

In the process of automated financial transaction verification, the detection and merging of vertical lines in table and receipt images are key steps to ensure the correct identification of table structures. However, since vertical and horizontal lines in the image may share some common run blocks, these common run blocks usually have a larger width, exceeding the typical width range of vertical lines, thus posing challenges to vertical line detection. To solve this problem, this paper proposes an effective processing solution. In the vertical line detection and merging process, a base run block is first selected from the BAG run block set as the starting point of the vertical line, and then other run blocks that may connect are searched upwards or downwards. During this process, if a run block is found to be part of both the vertical line and the horizontal line, its width will often be significantly larger than the typical width of the vertical line. If this run block is directly merged into the vertical line, it may destroy the accuracy of vertical line detection, leading to misjudgment. Therefore, this paper proposes a strategy in which, during vertical line detection, if a run block with a width significantly greater than the current vertical line width is encountered, that run block is skipped and not merged into the currently constructed vertical line set.

Specifically, when performing vertical line merging, the algorithm checks the width of each run block to be merged. If the width of the run block is much larger than the width of the vertical line, this indicates that it may be part of a horizontal line or a wide common area block rather than a typical feature of a vertical line. In this case, the algorithm ignores the run block and continues searching in the current direction for other run blocks that meet the characteristics of a vertical line. By skipping these overly wide run blocks, the algorithm effectively avoids misjudgments caused by common blocks, thereby improving the accuracy of vertical line detection. The judgment conditions are given by the following formulas:

$$Y_G/Y_i > \pi_g \quad (23)$$

$$N_{12} \neq 0 \quad (24)$$

In the process of automated financial transaction verification, for the detection and merging of vertical lines in financial forms and receipt images, this paper proposes a vertical search and merging strategy based on BAG run blocks. Unlike horizontal line detection, vertical line detection cannot simply rely on the horizontal characteristics of the run blocks but needs to focus on their vertical characteristics. This is because BAG run blocks usually present horizontal characteristics, with their width reflecting horizontal attributes, while the merging of vertical lines needs to be

evaluated and processed based on vertical characteristics. During the vertical line detection and merging process, when a run block has more than one adjacent run block, it cannot be merged simply by selecting the run block with the largest width. This differs from the logic of horizontal line merging. In horizontal line detection, due to the obvious horizontal characteristics of BAG run blocks, merging can be done by simply selecting the run block with the largest width. However, in vertical line detection, the width of BAG run blocks does not accurately reflect the characteristics of vertical lines. Therefore, relying solely on horizontal width to judge vertical line merging will result in inaccurate detection outcomes. To effectively perform vertical line merging, this paper introduces the vertical characteristics of BAG run blocks as the basis for merging. When detecting vertical line merging, the algorithm analyzes the vertical characteristics of adjacent run blocks. These vertical characteristics refer to the height of the run blocks and their connectivity with other run blocks, that is, the extension of the run blocks in the vertical direction. By analyzing these vertical characteristics, the algorithm can better identify which run blocks belong to the vertical line structure.

Furthermore, this paper proposes a comprehensive strategy that considers the merging of multiple adjacent run blocks. During vertical line detection and merging, when encountering multiple adjacent run blocks, the algorithm does not simply select the run block with the largest width but instead evaluates the vertical height of all adjacent run blocks. Through this evaluation, the algorithm can identify which run blocks have continuity in the vertical direction and merge them into the vertical line. This method avoids the limitations of relying solely on width for merging, ensuring the accuracy and completeness of vertical line merging. Finally, to enhance the precision of vertical line merging, this paper also introduces further analysis of the connectivity of run blocks during the merging process. By calculating the vertical connectivity and height characteristics of adjacent run blocks, the algorithm can more accurately determine which run blocks should be merged into the same vertical line, thus avoiding misjudgment and incorrect merging. The above retrieval and merging method needs to satisfy the following formula:

$$TY_g/TY_i = \pi_g \quad (25)$$

After traversing all the BAG run blocks, the vertical line set in financial forms and receipt images is obtained. After completing the search for horizontal and vertical lines, the initial horizontal and vertical line sets obtained are merely the preliminary outline of the table structure. The preliminary results of the horizontal and vertical line sets may contain broken or redundant line segments due to image noise, line breaks, or other interference factors. If the correction and filtering are done separately for the horizontal and vertical line sets, it may not achieve the desired results. To achieve a complete reconstruction of the table, further correction and filtering of these line sets are necessary. Therefore, this paper proposes a holistic table reconstruction method, which involves analyzing and merging the structural characteristics of the table by combining the horizontal and vertical line sets to achieve more efficient and accurate line correction and filtering.

After obtaining the initial horizontal and vertical line sets, the system records the position information of each line. This position information includes not only the start and end

coordinates of the lines but also their relative position in the image. Using this information, the system can reverse reconstruct the table structure. Reverse table reconstruction means reconnecting the line segments that may belong to the same table line but were separated due to breaks or detection errors based on the recorded line position information. This process effectively repairs broken line segments in the table, making the entire table structure more complete. After the table structure is reverse reconstructed, the system further analyzes the merged line segments to identify and delete redundant or extraneous lines. For example, some lines may have been incorrectly detected as table lines due to noise or errors, and the system can identify and eliminate these erroneous lines by analyzing the overall structure of the table. Additionally, the system smooths the line segments of the table, ensuring smooth connections and avoiding unnecessary misalignment or overlap.

5. TEXT EXTRACTION AND FRACTURE RESTORATION IN AUTOMATED FINANCIAL TRANSACTION VERIFICATION IMAGES

In financial forms and receipt images, the integrity of textual information is crucial for accurate Optical Character Recognition (OCR), especially for Chinese characters, which are numerous, complex in strokes, and significantly vary in handwriting styles. If any part of the textual information is lost during image processing, it will lead to recognition errors and may even prevent the OCR system from correctly identifying the text. In practice, table lines often intersect with text strokes. For example, a table line may be tangent to a text stroke or intersect with an inclined text stroke. These situations can easily lead to the loss of foreground text information or stroke fractures when removing table lines. The fractured text strokes can severely affect the accuracy of OCR recognition, making it impossible to correctly identify the text, particularly in the case of Chinese characters, which have complex structures and high interconnectivity. To solve this problem, this paper proposes a BAG algorithm-based processing method that can automatically restore fractured text strokes by relying on the connectivity information of run blocks after removing table lines, ensuring the integrity of textual information. Figure 6 shows a schematic diagram of the connectivity information analysis of run blocks.

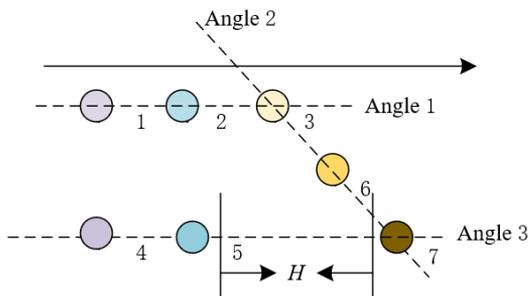


Figure 6. Schematic diagram of the connectivity information analysis of run blocks

Suppose that after removing the table lines, a text stroke T is fractured, with the fractured stroke divided into two parts, T1 and T2. The run block X in T1 is adjacent to the intersecting line, and the run block Y in T2 is adjacent to the intersecting line. According to the Euclidean distance calculation, X and Y

are the closest run blocks in T1 and T2, respectively, which means that if the stroke can be reconstructed between X and Y, the fractured text stroke T can be repaired. The BAG algorithm records the positions and adjacency relationships of run blocks, allowing the identification of the connection vector between X and Y. Using this position vector, the BAG run blocks located between X and Y in the original image can be identified, and the corresponding blank areas of these run blocks can be re-labeled as foreground pixels, thereby completing the fractured stroke and re-establishing the connection between X and Y.

To further improve the accuracy of the restoration, this paper introduces an improved processing method for the BAG run blocks at the fracture points of text strokes. Specifically, the improved method evenly distributes the BAG center points at the fracture points between the adjacent run blocks. This approach ensures that the restored strokes have better smoothness and consistency. The BAG run blocks between the fractured strokes can be obtained through the following formulas:

$$a_s = a_X + m \cdot (a_Y - a_X) / v \tag{26}$$

$$b_s = b_Y + m \cdot (b_X - b_Y) / v \tag{27}$$

$$g_s = (b_X - y_Y) / v \tag{28}$$

$$q_s = q_X + m \cdot (q_Y - q_X) / v \tag{29}$$

By following the above steps, all the text in financial forms and receipt images can be restored, ready for subsequent text information extraction.

6. EXPERIMENTAL RESULTS AND ANALYSIS

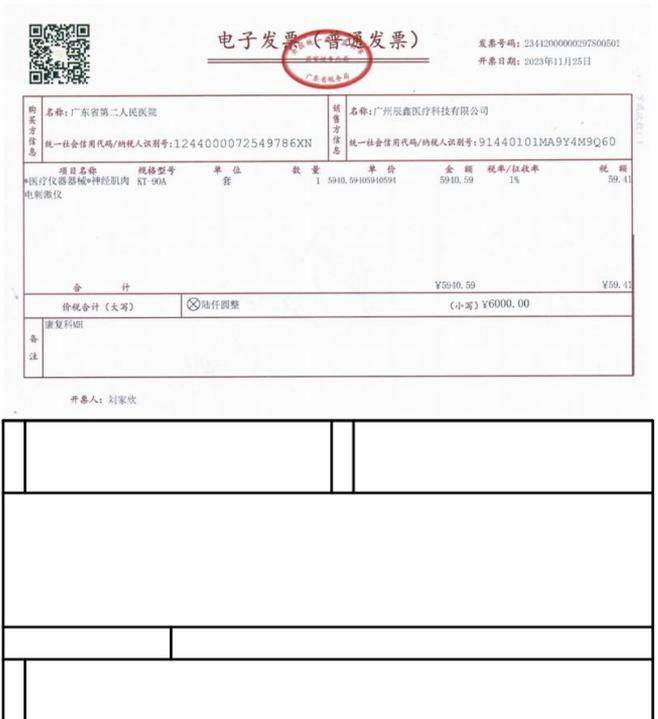


Figure 7. Example of financial forms and receipt images and their table extraction

This paper conducted systematic experiments on financial forms and receipt images, and Figures 7 and 8 demonstrate the effects of table extraction and text extraction. In the table extraction experiment, this paper applied a technique that combines Canny edge detection with Sauvola local adaptive binarization, testing it on various types of financial forms and receipt images. The results show that this method can effectively distinguish table lines from text regions and performs well in handling noise and complex backgrounds. Particularly in terms of edge detection accuracy and binarization adaptability, this method ensures accurate extraction of table structures under various conditions, significantly improving the robustness and reliability of table detection algorithms. A series of comparative experiments indicate that the improved table detection algorithm shows significant improvements in precision and recall, validating the practical application value of this method in financial document processing.

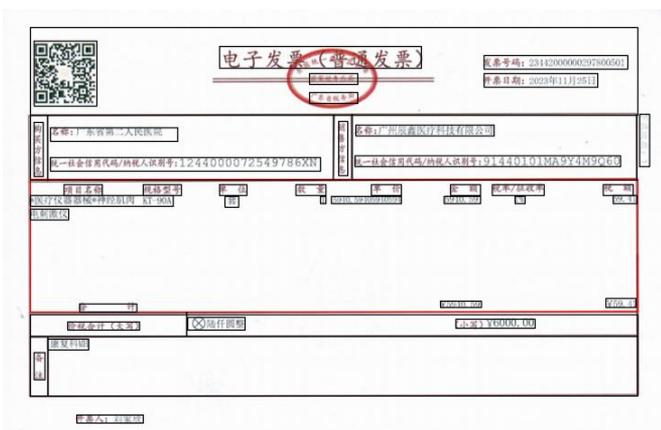


Figure 8. Example of text extraction from financial forms and receipt images

In further text extraction experiments, we combined table detection with text region separation techniques to extract text from images and restore fractured characters. The experimental results show that by repairing fractured characters, the extracted text exhibits high consistency in visual effect and recognition accuracy. Especially for complex receipt images, applying the Sauvola local adaptive binarization method significantly reduced character noise and fracture phenomena, achieving high-precision text extraction.

The analysis indicates that the proposed text extraction technique is not only capable of handling conventional financial forms but also adapts well to complex receipt images, providing reliable technical support for subsequent automated financial transaction verification.

Table 1. Impact of different numbers of local processing units on model performance

Number of Units	Training Set		Test Set	
	F	FPS	F	FPS
4	84.2	22.4	85.6	26.3
6	87.3	18.6	86.3	25.4
8	88.9	12.5	88.8	21.6
10	90.2	11.4	89.5	20.5
12	90.6	10.6	89.6	18.9

Table 1 shows the performance of the model on the training and test sets with different numbers of local processing units. As the number of processing units increases, the model's F-measure (F value) gradually improves, indicating that the model has higher accuracy when dealing with increased complexity. For example, when the number of units increases from 4 to 12, the F value on the training set increases from 84.2 to 90.6, while the F value on the test set increases from 85.6 to 89.6. However, with the increase in the number of units, the frames per second (FPS) decreases significantly, indicating a slowdown in processing speed. For instance, the FPS on the training set drops from 22.4 to 10.6, and on the test set, it drops from 26.3 to 18.9. This shows that while increasing the number of processing units improves the model's accuracy, it also increases the computational overhead. Analyzing these results, the following conclusions can be drawn: Increasing the number of processing units helps to enhance model performance, especially in automated financial transaction verification tasks, where this improvement is particularly significant. However, this improvement comes at the cost of reduced processing speed. In practical applications, a balance must be struck between accuracy and processing speed, choosing an appropriate number of processing units to meet the specific needs of the application scenario. For scenarios requiring high accuracy, a larger number of processing units can be selected; for scenarios with high real-time requirements, the number of processing units may need to be reduced to ensure system responsiveness.

Table 2. Ablation study

Canny	Sauvola	BAG	Training Set				Test Set			
			P	R	F	FPS	P	R	F	FPS
×	×	×	86.3	75.2	80.2	28.5	82.3	74.5	78.2	36.3
√	×	×	88.6	79.8	83.4	27.6	84.2	75.2	81	35.6
×	×	√	89.2	77.5	83.6	23	86.5	80.6	82.3	32.1
√	×	√	91.2	80.5	85.4	20.2	88.9	83.4	85.6	25.6
×	√	×	90.5	81.2	85	23.5	87.5	82.9	84.5	31
√	√	×	91.6	84.6	88.9	20.4	88	86.5	87.6	24.3
√	√	√	92.6	87.9	89.5	12.5	90.2	87.6	88.9	22.3

Table 2 shows the model's performance on the training and test sets using different combinations of image processing techniques, namely Canny edge detection, Sauvola local adaptive binarization method, and BAG technique. The experimental results show that using any single technique improves the model's performance. For example, using Canny

edge detection alone, the F value increases from 80.2 to 83.4; using the BAG technique alone, the F value increases from 80.2 to 83.6. Furthermore, when using both Canny edge detection and the BAG technique together, the F value further increases to 85.4. Notably, when all three techniques are combined, the model achieves the best performance, with F

values of 89.5 and 88.9 on the training and test sets, respectively. However, this improvement comes at the cost of processing speed (FPS), which drops from 28.5 to 12.5. From these data, it can be concluded that the combination of Canny edge detection, Sauvola local adaptive binarization method, and BAG technique significantly improves model accuracy, particularly in complex financial image processing. However, as more techniques are introduced, the processing speed decreases significantly, indicating an increase in computational load. Therefore, in practical applications, it is necessary to find a balance in the combination of these techniques based on the accuracy requirements and real-time needs of the task. For scenarios with extremely high accuracy requirements, the combination of all three techniques is undoubtedly the best choice.

Table 3 compares the performance of different methods for table line and text region separation on an invoice image set. The results show that the proposed method outperforms other methods across all metrics. Specifically, the proposed method achieves precision (P), effectiveness (E), and F values of 90.2,

87.2, and 88.9, respectively, significantly better than other methods. For example, the F value of adaptive projection analysis is only 74.5, while the F value of morphological operations, though reaching 85.6, has a processing speed (FPS) of only 4, far lower than the 22.3 FPS of the proposed method. Although the Hough transform and connected component analysis are close in F value, no FPS data are provided. Overall, the proposed method shows the best performance in both accuracy and separation effectiveness, while also maintaining a high level of processing speed. Analyzing these results, it can be concluded that the proposed method has significant advantages in table line and text region separation tasks, especially when processing complex invoice images. It can provide high accuracy while maintaining fast processing speed. In contrast, other methods either lack accuracy or have significant shortcomings in processing speed. Therefore, the proposed method is more practical for real-world applications, especially in financial transaction verification scenarios that require both accuracy and speed, making it the optimal choice.

Table 3. Performance of table line and text region separation on invoice image set

Method	<i>P</i>	<i>E</i>	<i>F</i>	<i>FPS</i>
Adaptive projection analysis	78.5	70.3	74.5	16
Morphological operations	85.6	85.6	85.6	4
Hough transform	85.2	80.5	82.2	-
Connected component analysis	87.6	82.3	84.3	-
Multi-task learning	82.6	74.6	78.9	21
Wavelet transform	83.5	80.2	82.3	16
The proposed method	90.2	87.2	88.9	22.3

Table 4. Performance of table detection and reconstruction on invoice image set

Method	<i>P</i>	<i>E</i>	<i>F</i>	<i>FPS</i>
Bounding box detection	74.5	52.1	60.2	7.2
Parallel line detection	83.2	73.2	78.5	13.2
Perspective correction	80.2	73.5	76.2	7.8
Local feature-based	79.5	78	78.4	-
Character alignment detection	86.6	84.2	85.6	1.5
Shape analysis-based	88.4	85.6	87.4	-
Texture analysis-based	91.2	83.1	87.2	-
RANSAC algorithm	83.6	81.2	82.3	25.6
Multi-model RANSAC	85.2	84.2	84.5	3
Weighted distance transform	91.2	83.6	87.6	11
SURF	90.5	83.9	87.2	10
The proposed method	92.5	87.2	89.6	12.5

Table 5. Performance of text extraction and fracture restoration on invoice image set

Method	<i>P</i>	<i>E</i>	<i>F</i>	<i>FPS</i>
OCR	78	71	73	1.2
Adaptive threshold segmentation	81	67	73	-
Local Fourier transform analysis	88	72	78	10
Complex wavelet transform	87	78	82	-
Laplacian pyramid	82	73.5	77.5	3
Convolutional autoencoder	88.5	78.5	82.6	8.5
Image matching	84.2	81.8	82.4	-
Adaptive region growing	84.6	83.2	84.5	30.2
Deep energy minimization	91.2	79.5	84.6	31
Multi-level graph cut algorithm	88.6	82.6	85.6	-
Multi-objective genetic algorithm	90.2	82.5	86.5	-
The proposed method	92.3	83.4	87	34.5

Table 4 compares the performance of different methods for table detection and reconstruction on an invoice image set. The results show that the proposed method achieved excellent

performance across all metrics. Specifically, the precision (P) of the proposed method reached 92.5, effectiveness (E) was 87.2, and the F value (F) was 89.6, all higher than other

methods. For example, the F value of the shape analysis-based method was 87.4, which, although close to the proposed method, was slightly inferior in precision. Additionally, although the weighted distance transform and SURF methods achieved F values of 87.6 and 87.2, respectively, they were still inferior to the proposed method in terms of precision and effectiveness. Notably, the proposed method maintained a high level of processing speed (FPS) at 12.5, much higher than the 1.5 FPS of character alignment detection and the 3 FPS of multi-model RANSAC, demonstrating good overall performance. From these data, it can be concluded that the proposed method excels in table detection and reconstruction tasks, especially when processing complex invoice images, outperforming other methods in multiple performance metrics. This indicates that the proposed method not only significantly improves detection and reconstruction accuracy but also maintains a high level of processing efficiency, which is particularly important for applications such as financial transaction verification that require real-time processing. In contrast, other methods exhibit deficiencies in precision, effectiveness, or processing speed, making it difficult to fully meet the demands of practical applications.

Table 5 presents a comparison of the performance of different methods in text extraction and fracture restoration tasks on an invoice image set. The proposed method performed excellently in terms of precision (P), effectiveness (E), and F value (F), achieving 92.3, 83.4, and 87, respectively, which are higher than other methods. For instance, the F value of the multi-objective genetic algorithm was 86.5, close to the proposed method, but slightly inferior in precision and effectiveness. The local Fourier transform analysis and convolutional autoencoder achieved F values of 78 and 82.6, respectively, performing well but still not as good as the proposed method. Additionally, the proposed method also excelled in processing speed (FPS), achieving 34.5, far surpassing other methods, such as OCR's 1.2 FPS and Laplacian pyramid's 3 FPS. From these data, it can be concluded that the proposed method has significant advantages in text extraction and fracture restoration tasks, especially when processing invoice images, where it can significantly improve processing speed while maintaining high precision and effectiveness. This is crucial for application scenarios that require efficient processing of large volumes of financial transaction images. In contrast, although other methods perform well in some metrics, they often have shortcomings in precision, effectiveness, or processing speed, making it difficult to meet the dual requirements of efficiency and accuracy in practical applications.

7. CONCLUSION

The comprehensive image processing method proposed in this paper aims to solve key issues in automated financial transaction verification by significantly enhancing the efficiency and accuracy of financial image processing through four main research areas. First, the paper studied the separation of table lines and text regions in images, developing algorithms capable of accurately separating table and text regions under complex backgrounds. Second, the application of the Sauvola local adaptive binarization method in financial transaction images was explored, improving the binarization effect of images. Third, a high-precision table detection and reconstruction algorithm was developed, significantly

enhancing the detection and reconstruction capabilities of table structures. Finally, the study focused on text extraction and fracture restoration techniques in images, effectively improving the accuracy and coherence of text extraction. The experimental results show that the proposed method outperforms existing methods across multiple performance metrics, particularly in precision, effectiveness, and processing speed, achieving significant progress.

The research in this paper has important practical value, especially in financial table processing, invoice image recognition, and other fields, where it can significantly improve the efficiency and accuracy of automated financial transaction verification systems. However, this paper also has certain limitations, such as the potential impact on model performance in extremely complex backgrounds or under severe noise interference. Additionally, due to the combination of multiple techniques, the computational complexity and hardware requirements of the proposed method are relatively high, which may limit its application in resource-constrained environments. Future research directions could focus on the following aspects: first, optimizing the computational efficiency of the model to reduce its dependency on hardware; second, further enhancing the robustness of the model to improve its adaptability in complex and dynamic environments; third, exploring the integration of the proposed method with other intelligent technologies, such as deep learning or reinforcement learning, to further enhance the automation and intelligence of image processing. With these improvements, the proposed method is expected to be promoted and applied in more practical scenarios.

FUNDINGS

This paper was supported by Guangdong Province Philosophy and Social Sciences 13th Five Year Plan Project (Grant No: GD20CSH09); Guangzhou Civil Aviation College 2024 Research Backbone Project (Grant No: 24X4306); 2023 Campus level Quality Engineering Education and Teaching Projects of Guangzhou Civil Aviation College (Grant No: JG202309) and Guangzhou Civil Aviation College 2023 College level College Students Innovation and Entrepreneurship Project (Grant No: CY23017 and CY23011).

REFERENCE

- [1] Park, J., Jeong, S., Yeom, K. (2023). Smart contract broker: Improving smart contract reusability in a blockchain environment. *Sensors*, 23(13): 6149. <https://doi.org/10.3390/s23136149>
- [2] Capocasale, V., Perboli, G. (2022). Standardizing smart contracts. *IEEE Access*, 10: 91203-91212. <https://doi.org/10.1109/ACCESS.2022.3202550>
- [3] Chen, E., Qin, B., Zhu, Y., Song, W., Wang, S., Chu, C.C.W., Yau, S.S. (2021). SPESC-Translator: Towards automatically smart legal contract conversion for blockchain-based auction services. *IEEE Transactions on Services Computing*, 15(5): 3061-3076. <https://doi.org/10.1109/TSC.2021.3077291>
- [4] Filatova, N. (2020). Smart contracts from the contract law perspective: Outlining new regulative strategies. *International Journal of Law and Information*

- Technology, 28(3): 217-242. <https://doi.org/10.1093/ijlit/aaaa015>
- [5] Ullah, F., Al-Turjman, F. (2023). A conceptual framework for blockchain smart contract adoption to manage real estate deals in smart cities. *Neural Computing and Applications*, 35(7): 5033-5054. <https://doi.org/10.1007/s00521-021-05800-6>
- [6] Ge, X. (2021). Smart payment contract mechanism based on blockchain smart contract mechanism. *Scientific Programming*, 2021(1): 3988070. <https://doi.org/10.1155/2021/3988070>
- [7] Guo, H., Chen, Y., Chen, X., Huang, Y., Zheng, Z. (2024). Smart contract code repair recommendation based on reinforcement learning and multi-metric optimization. *ACM Transactions on Software Engineering and Methodology*, 33(4): 106. <https://doi.org/10.1145/3637229>
- [8] Ning, J., Yu, S. (2021). Barcode location in financial statement system based on the partial differential equation image recognition algorithm. *Advances in Mathematical Physics*, 2021(1): 9177159. <https://doi.org/10.1155/2021/9177159>
- [9] Xiaohong, C., Yi, S., Zhaowen, L., Imran, M., Keping, Y. (2023). Web-based practical privacy-preserving distributed image storage for financial services in cloud computing. *World Wide Web*, 26(3): 1223-1241. <https://doi.org/10.1007/s11280-022-01090-7>
- [10] Zhang, H., Dong, B., Zheng, Q., Feng, B., Xu, B., Wu, H. (2022). All-content text recognition method for financial ticket images. *Multimedia Tools and Applications*, 81(20): 28327-28346. <https://doi.org/10.1007/s11042-022-12741-2>
- [11] Wang, W., Lu, M., Dai, X., Jiang, P. (2024). Binary wavelet transform-based financial text image authentication algorithm. *Automatic Control and Computer Sciences*, 58(3): 326-335. <https://doi.org/10.3103/S0146411624700202>
- [12] Chiu, C.H., Tsai, Y.C., Chen, H.L. (2021). Long text to image converter for financial reports. *International Journal of Data Mining, Modelling and Management*, 13(3): 211-230. <https://doi.org/10.1504/IJDM.2021.118019>
- [13] Pan, S., Wei, J., Hu, S. (2020). A novel image encryption algorithm based on hybrid chaotic mapping and intelligent learning in financial security system. *Multimedia Tools and Applications*, 79(13): 9163-9176. <https://doi.org/10.1007/s11042-018-7144-5>
- [14] Tian, M.W., Yan, S.R., Tian, X.X., Liu, J.A. (2019). Research on image recognition method of bank financing bill based on binary tree decision. *Journal of Visual Communication and Image Representation*, 60: 123-128. <https://doi.org/10.1016/j.jvcir.2018.12.016>
- [15] Wang, J. (2021). Application of wavelet transform image processing technology in financial stock analysis. *Journal of Intelligent & Fuzzy Systems*, 40(2): 2017-2027. <http://doi.org/10.3233/JIFS-189204>
- [16] Cho, S., Moon, J., Bae, J., Kang, J., Lee, S. (2023). A framework for understanding unstructured financial documents using RPA and multimodal approach. *Electronics*, 12(4): 939. <https://doi.org/10.3390/electronics12040939>
- [17] Kankam-Kwarteng, C., Donkor, G.N.A., Osei, F., Amofah, O. (2024). Do corporate social responsibility and corporate image influence performance of the financial sector? *Journal of Financial Services Marketing*, 29(2): 306-317. <https://doi.org/10.1057/s41264-023-00208-w>
- [18] Bouichou, S.I., Wang, L., Zulfiqar, S. (2022). How corporate social responsibility boosts corporate financial and non-financial performance: The moderating role of ethical leadership. *Frontiers in Psychology*, 13: 871334. <https://doi.org/10.3389/fpsyg.2022.871334>
- [19] Ali, H. Y., Danish, R. Q., Asrar-ul-Haq, M. (2020). How corporate social responsibility boosts firm financial performance: The mediating role of corporate image and customer satisfaction. *Corporate Social Responsibility and Environmental Management*, 27(1): 166-177. <https://doi.org/10.1002/csr.1781>