IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Traitement du Signal
Vol. 42, No. 2, April, 2025, pp. 675-687
Journal homepage: http://iieta.org/journals/ts

Application of Artificial Intelligence Image Recognition and Blockchain Technology in Enhancing the Security of Cross-Border Financial Transactions

Hongmei Zhang

School of Finance, Tongling University, Tongling 244000, China

Corresponding Author Email: 207543@tlu.edu.cn

https://doi.org/10.18280/ts.420207

Received: 20 November 2024
Revised: 17 March 2025
Accepted: 30 March 2025
Available online: 30 April 2025

Keywords:
cross-border financial transactions, AI image recognition, blockchain technology, identity fraud, transaction authenticity, security

ABSTRACT

In the context of globalization and the rapid development of international trade, the demand for cross-border financial transactions has been steadily increasing, while also facing security risks such as identity fraud and transaction authenticity. The application of artificial intelligence (AI) image recognition technology in the financial sector has demonstrated its potential in preventing forged documents, while blockchain technology, with its immutability and decentralized nature, excels in ensuring the authenticity and transparency of transaction data. Although existing research has made progress in the separate applications of AI image recognition and blockchain technology, there remain gaps and deficiencies in their integrated use. This paper aims to enhance the security and reliability of cross-border financial transactions by combining AI image recognition with blockchain technology. The research is divided into two main parts: First, a method based on AI image recognition for detecting forged documents to address identity fraud; second, the integration of blockchain technology to optimize the process of confirming the authenticity of cross-border financial transactions. Through these innovative technological integrations, this paper offers a comprehensive and efficient security solution for cross-border financial transactions.

1. INTRODUCTION

With the acceleration of globalization and the growing prosperity of international trade, the demand for cross-border financial transactions has been steadily increasing [1-4]. However, as the transaction volume rises, related security risks have become increasingly prominent, particularly issues of identity fraud and transaction authenticity. In recent years, AI technology, especially image recognition technology, has shown great potential in the financial sector [5, 6]. At the same time, blockchain technology, with its immutability and decentralized characteristics [7], has performed excellently in data security and transaction transparency [8]. Therefore, the integration of AI image recognition and blockchain technology provides a new solution for the security of cross-border financial transactions.

The significance of related research lies in the innovative integration of technologies to improve the security and reliability of cross-border financial transactions. Specifically, the use of AI image recognition technology can effectively identify and prevent forged documents [9], reducing the occurrence of identity fraud. Blockchain technology, on the other hand, ensures the authenticity and immutability of transaction data, thereby enhancing transaction transparency and trustworthiness [10-14]. By combining these two technologies, a more secure and efficient cross-border financial transaction system can be built.

Currently, AI image recognition technology is widely applied in the financial sector, especially in financial security, where it is used for facial recognition, behavior recognition, and more. These applications significantly enhance security protection and service quality. Blockchain technology, with its characteristics of decentralization and immutability, has deepened its use in scenarios such as cross-border payments and credit evaluation, improving transaction transparency and security. The trend of integrating blockchain and AI is becoming increasingly prominent, with the market for blockchain + AI expected to grow significantly, driving innovation in financial services, such as optimizing smart contract trading strategies and improving risk control efficiency. Despite challenges like data privacy and algorithm fairness, as technology develops and regulations progress, their application in the financial sector will become more profound. This will drive financial services toward greater intelligence, efficiency, and security, continuously empowering innovation and process optimization within the financial industry, while bringing new development opportunities and transformative momentum to the financial market.

Although existing research has made certain achievements in the separate applications of AI image recognition and blockchain technology [15-19], there are still many shortcomings in their integrated application. For example, Kao et al. [20] studied the use of AI image recognition technology for identity verification. Although it achieved significant improvements in accuracy, it still has vulnerabilities when dealing with complex forgery techniques. Smith and Dhillon [21] explored the application of blockchain in financial transactions, but due to the lack of deep integration with other technologies, it could not fully address the dual challenges of identity verification and transaction authenticity. Therefore, existing research still has considerable room for improvement in terms of integrated application and practical operability.

Identity fraud is a major issue in cross-border financial transactions, with counterfeit documents widely used in illicit financial activities. Traditional identity verification methods often have vulnerabilities and are unable to effectively identify high-quality counterfeit documents. To address this challenge, this paper proposes a counterfeit document recognition method based on AI image recognition technology, enhancing feature extraction and recognition accuracy of document images through a multi-region attention network. Regarding the confirmation of transaction authenticity, the decentralized nature of blockchain technology provides a secure and transparent solution for cross-border financial transactions, effectively preventing information tampering and forgery, and ensuring the rights of both parties involved in the transaction. The main research content of this paper is divided into two parts: First, regarding the issue of identity fraud in cross-border financial transactions, a method for detecting forged documents based on AI image recognition technology is studied and designed. Second, blockchain technology is integrated to optimize the process of confirming the authenticity of cross-border financial transactions. Through the combination of these two parts, this research aims to build a comprehensive and efficient security system, enhancing the overall security and trustworthiness of cross-border financial transactions. The value of the research lies in providing operational technical guidance and solutions for practical financial transactions, thereby promoting the healthy development of cross-border financial transactions.

2. FORGED DOCUMENT RECOGNITION FOR IDENTITY FRAUD IN CROSS-BORDER FINANCIAL TRANSACTIONS

Identity verification is crucial in cross-border financial transactions. To identify counterfeit documents, this paper proposes a new method based on AI's "image recognition" technology. Simply put, this method works like an intelligent system that can learn and determine whether a document image has been tampered with. Specifically, we use computer algorithms to recognize detailed features in the image, distinguishing authenticity in a way similar to the human eye. Additionally, we use blockchain technology, which acts as an "immutable ledger" to ensure that data in the transaction process cannot be tampered with, while protecting the privacy of both parties involved. Throughout the process, we also utilize "smart contracts," an automated protocol that ensures the security and validity of the transaction without requiring third-party intervention.

In cross-border financial transactions, identity fraud is a serious security issue, and the use of forged documents is one of the most common methods. To effectively recognize forged documents, this paper proposes a solution based on a multi-region attention network. This method utilizes a strategy that combines local texture features and high-level semantic features, aiming to enhance the accuracy and robustness of forged document recognition. The method extracts preliminary features of the document image through the backbone network and uses a local region segmentation-like approach to generate multiple attention maps, which focus on different potential discriminative regions in the document image. This approach ensures that we not only focus on the overall structure of the document image but also carefully observe the subtle differences in local regions, thus improving sensitivity to forged documents.

In practical applications, to avoid overfitting caused by the high-dimensional semantic representation formed by combining fine texture-level artifacts with other information in the deeper layers of the network, we retain the higher resolution of the shallow feature maps from the backbone network. Additionally, high-frequency components and more convolution operations are used to enhance the expression ability of shallow texture features. These operations ensure that when extracting the detailed features of document images, we can more accurately capture the subtle differences in forged documents. At the same time, the method can independently pool these texture features within each discriminative region into feature vectors for each local area to avoid interference from irrelevant features in the background, thus improving the precision of feature extraction. Finally, the method aggregates low-level texture features and high-level semantic features to form a multi-scale representation of the entire image for forged image recognition. This multi-scale representation can comprehensively consider both the local details and the overall structure of the document image, thereby effectively improving the accuracy of forged document recognition.

Figure 1 shows the framework of the forged document recognition method for identity fraud in cross-border financial transactions. The three key components of this method are specifically described at the model level as follows:

(1) Use the attention module to generate attention maps for multiple regions from the middle layers of the network. These attention maps can achieve a balance between different receptive field sizes and attention map resolutions, allowing for better capture of different details and features in the document image. When applied to the scenario of identity fraud in cross-border financial transactions, this module can more effectively focus on key areas of the document, such as the portrait, text, signature, and other important information. Through deep mining and detailed analysis of these features, the accuracy of forged document recognition can be significantly improved. By using multi-region attention mechanisms, it avoids the information loss caused by a single receptive field and ensures that subtle forgery traces are accurately captured at high resolution.

(2) Use high-pass filtering and densely connected convolutional layers as the texture enhancement module to extract and enhance the shallow texture information of the network. The purpose of this module design is to capture the high-frequency texture features of the document image at the shallow stage, avoiding dilution or coverage of these features during deeper processing. In the cross-border financial transaction scenario, forged documents typically have many small differences in details, such as print quality, paper texture, and subtle forgery traces. By using high-pass filtering and dense convolutions, these key detail features can be better extracted and enhanced.

(3) Replace traditional global average pooling with bilinear attention pooling. For shallow texture feature maps, bilinear attention pooling is used to obtain the global texture features

of each region, then the attention maps of different regions are summed to obtain a global attention map, which is bilinearly attention-pooled with the deep semantic features of the network, ultimately obtaining the global features. The integration process ensures that the model can capture both the local details of the document image and understand its overall structure and semantics, thus maintaining efficient recognition performance even when faced with complex and varied forgery techniques.

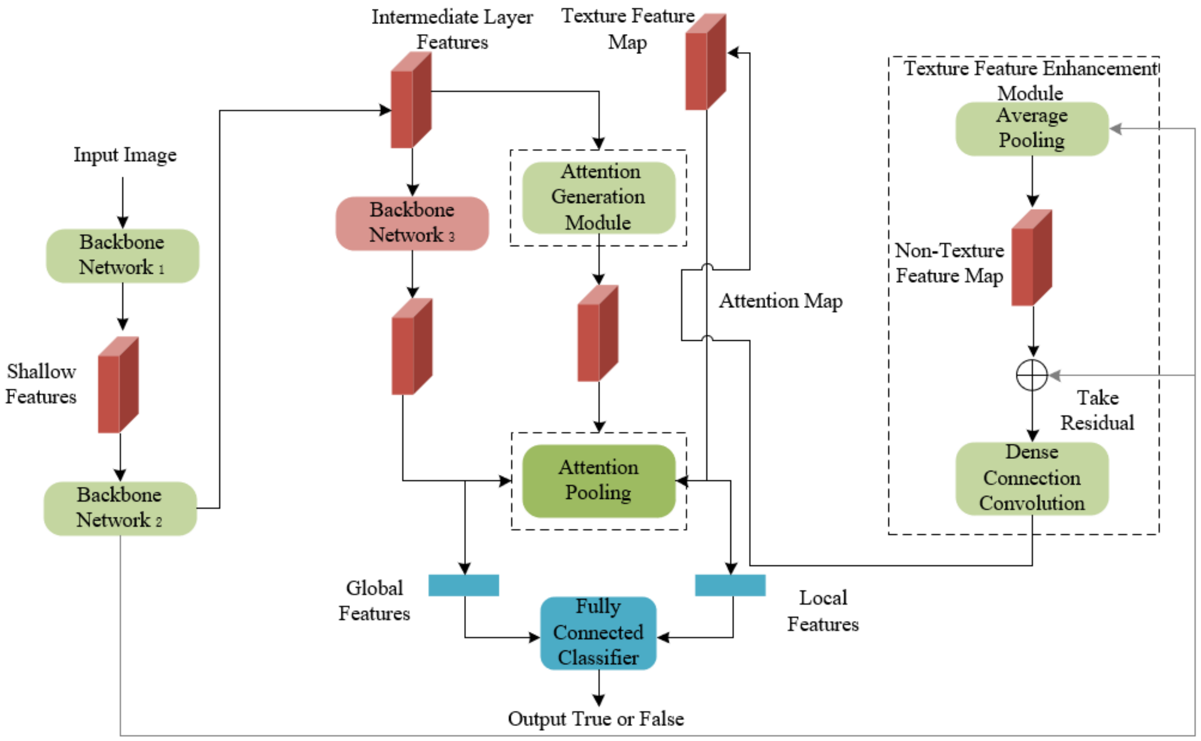

Figure 1. Framework of forged document recognition for identity fraud in cross-border financial transactions

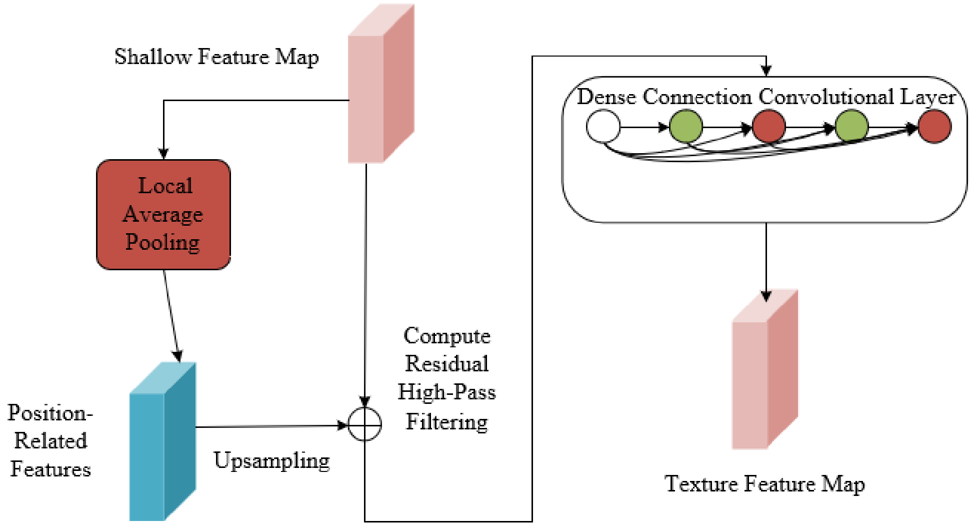
Shallow Feature Map
Local Average Pooling
Position-Related Features
Upsampling
Compute Residual High-Pass Filtering
Dense Connection Convolutional Layer
Texture Feature Map

Figure 2. Architecture of the texture feature enhancement module

Specifically, in forged document recognition for identity fraud in cross-border financial transactions, the principle of the multi-region attention map is to perform fine analysis and feature extraction on different regions of the document image to identify potential forgery traces. That is, when a real or forged document image U is input, the framework first processes the image through the backbone network d to extract the intermediate feature map $d_s(U)$. These feature maps contain rich information about the document image. To further refine the feature extraction process, the multi-region attention module is introduced to generate multiple attention maps X. These attention maps are achieved through a 1×1 convolution layer, BN batch normalization layer, and ReLU nonlinear activation function. The output dimension L of the 1×1 convolution layer represents the number of attention regions. By processing with BN and ReLU activation functions, the attention maps are ensured to be positive within the attended regions and zero in other areas.

In the scenario of identity fraud in cross-border financial transactions, the generation process of these multi-region attention maps can effectively capture key details in the document image. For example, facial images, text, signatures, and other important identification areas in the document, which are often the easiest targets for forgers. Through the multi-region attention module, multiple different attention maps can be generated, with each attention map X_j corresponding to a specific discriminative region. This means that each attention map focuses on a specific part of the document image, allowing for finer extraction and analysis of the features in that part.

The design of the texture feature enhancement module aims to retain and strengthen the texture information in the document image, thereby more effectively discovering forgery traces in local regions. This module first applies local average pooling on the feature map *TM* in the shallow network to downsample and obtain the pooled feature map *F*. The purpose of this process is to extract the low-pass components *F*, which mainly contain spatially correlated information of the image. This information is typically smoother and contains fewer high-frequency details. By prioritizing the extraction of these low-frequency components, the system can better understand the overall structure and layout of the image. Figure 2 shows the architecture of the texture feature enhancement module.

However, to accurately recognize subtle differences in forged documents, relying solely on low-frequency information is not sufficient. Therefore, the texture feature enhancement module further upsamples the pooled feature map *F* back to the original resolution of *TM* and calculates the residual between the two, which represents the high-frequency components of the shallow feature map, as shown in the following equation. The high-frequency components contain the details and texture information in the image, which are critical for identifying forgery traces. In cross-border financial transactions, forged documents often exhibit obvious flaws in details and textures, such as blurred facial image edges and subtle differences in text. By enhancing these high-frequency texture features, the system can more sensitively capture these forgery traces, thereby improving recognition accuracy and reliability.

$$S_{TM_s} = d_{TMs}\left(U\right) - F \tag{1}$$

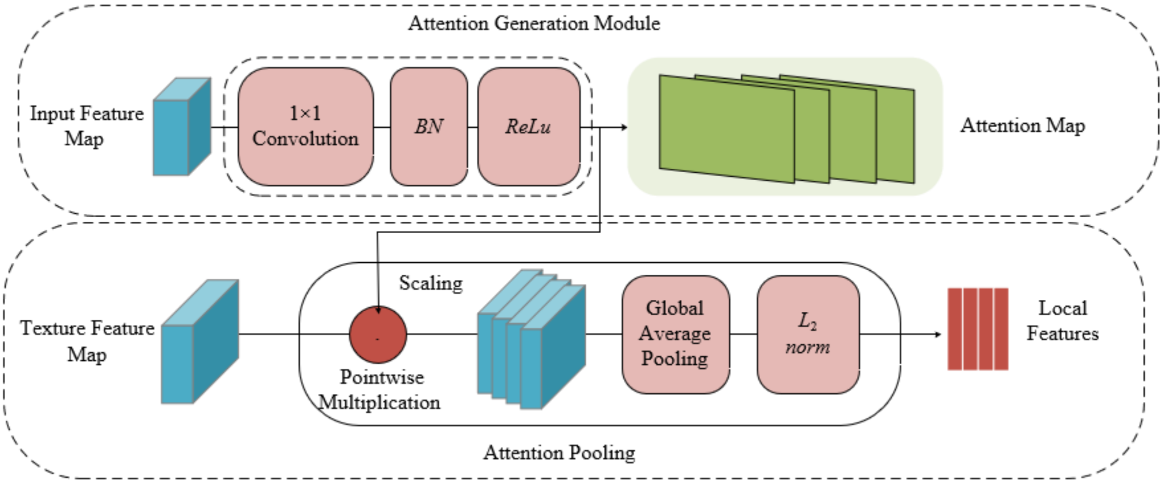

Figure 3. Attention generation module and bilinear attention pooling architecture

The introduction of bilinear attention pooling is intended to more effectively extract both local and global features, thereby accurately identifying forgery traces in the document. The core of this module lies in how to use the attention mechanism to focus on important areas in the image and combine shallow texture features with deep semantic features for efficient forged document recognition. Specifically, bilinear attention pooling first adjusts the attention map to the same resolution as the shallow texture feature map using bilinear interpolation. Then, these adjusted attention maps are used to extract local texture features. This ensures that the details of each local region are fully captured and utilized, especially when dealing with subtle differences in document images, making it easier to identify forgery traces. Figure 3 shows the attention generation module and bilinear attention pooling architecture.

In the application scenario of cross-border financial transactions, the authenticity recognition of documents needs to focus on the combination of details and overall features. By adding the attention maps from different regions, the system can generate a global attention map that represents the integrated information of all important areas in the image. Then, this global attention map is processed with bilinear attention pooling with the deep semantic feature map from the network to obtain the global features. This ensures that the system can not only focus on subtle local differences but also understand the overall structure and semantic information of the image.

To further improve the accuracy of recognition, this method introduces normalized average pooling when processing some texture feature maps D_j to address the issue of norm differences that may arise when using traditional global average pooling on different discriminative regions.

$$n_j = \frac{\sum_{l=0}^{G_t-1}\sum_{v=0}^{Q_t-1} D_{j,l,v}}{\left\|\sum_{l=0}^{G_t-1}\sum_{v=0}^{Q_t-1} D_{j,l,v}\right\|_2} \tag{2}$$

Through normalization, the system can treat the features of each region more fairly, avoiding the impact of attention map intensity on the feature vector norms, thereby capturing the texture information of each region more accurately. When processing deep features, by adding all the attention maps to obtain the global attention map X_{SUM}, and performing bilinear attention pooling with the feature map from the last layer of the network, the system can extract the global deep features *H*.

Finally, these global features are combined with the previously extracted local features *O* and input into the classifier for classification. This comprehensive processing flow ensures that the system can consider both local details and global information, improving the accuracy and reliability of forged document recognition.

In cross-border financial transactions, document forgery may involve various techniques, such as altering text information, splicing photos, and forging watermarks. Therefore, the recognition system must be able to focus on different regions of the document image and capture all possible forgery traces. To ensure that each attention map focuses on different and fixed semantic regions, this study proposes a region independence loss function to ensure the diversity and stability of attention maps. The construction of the region independence loss function mainly consists of two parts: intra-class loss and inter-class loss. The purpose of the intra-class loss is to ensure that each attention region always focuses on a fixed semantic position in different document images, for example, *X*1 always focuses on the eye region, and *X*2 always focuses on the lip region. This stability ensures that the system can reliably detect forgery traces in specific regions when facing different documents. The inter-class loss aims to ensure that different attention regions focus on different semantic positions, avoiding redundancy in information when multiple attention maps concentrate on the same discriminative region. Specifically, by performing bilinear attention pooling between the position-related feature map *F* extracted from the texture enhancement module and each attention region *X*, a semantic position vector *N* is obtained. Then, based on the center loss, the region independence loss ensures that the semantic position vectors of specific regions in different document images stay near a feature center *z*, achieving intra-class stability. The inter-class loss ensures that the distance between different attention regions is increased, making them focus on different semantic positions. In this way, the region independence loss function effectively addresses the degradation problem of multi-region attention networks, ensuring that the system can comprehensively and stably capture various forgery traces in forged document recognition for cross-border financial transactions. The expression for the region independence loss function is:

$$\begin{aligned} loss_{EUJM} &= \sum_{u=1}^{Y}\sum_{k=1}^{L} MAX\left(\left\|N_k^u - z_k^s\right\|_2 - l_{IN}\left(b_u\right), 0\right) \\ &+ \sum_{j,m\in(L,L),\, j\neq m} MAX\left(L_{OUT} - \left\|z_j^s - z_m^s\right\|_2, 0\right) \end{aligned} \tag{3}$$

The first part of *Loss*$_{RUM}$ is the intra-class loss. Here, the batch size is represented by *Y*, the number of attention maps is *L*, the maximum distance between each semantic position vector and the corresponding semantic position feature center is l_{IN}, and the minimum distance between each feature center is l_{OUT}. The feature center update rate is *β*, and *th* represents the stop gradient backpropagation. The expression for the feature center *z* is:

$$z^s = (1-\beta)th\left(z^{s-1}\right) + \beta\left(\frac{1}{Y}\sum_{u=1}^{Y} N^u\right) \tag{4}$$

The set weight is represented by *η*, and the loss function in this method includes both region independence loss and cross-entropy loss:

$$loss = loss_{ZR} + \eta * loss_{EUM} \tag{5}$$

In cross-border financial transactions, the accuracy of identity verification is critical, and forged document recognition is a key element in ensuring transaction security. The multi-region attention network proposed in this paper for this application scenario, by introducing an attention-guided data augmentation mechanism, aims to solve the problem of over-expansion of a single attention region, thus improving the comprehensiveness and robustness of forged document recognition. The core of this mechanism lies in the use of two strategies, hard erasure and soft erasure, to force the network to explore different discriminative regions during the training process. Specifically, the hard erasure strategy covers high-attention regions, forcing the network to look for discriminative features in other areas, while the soft erasure strategy applies Gaussian blur to make high-attention areas blurry, encouraging the network to extract information from other regions. This mechanism not only prevents the attention from concentrating on a single region but also promotes the exploration and division of various discriminative regions by the attention blocks, effectively improving the diversity and accuracy of forged document recognition. The expression for hard erasure is:

$$U'(a,b) = \begin{cases} 0, IF \;\; X_j^*(a,b) > 0.5 \\ U(a,b), otherwise \end{cases} \tag{6}$$

The expression for soft erasure is:

$$\begin{aligned} U_f(a,b) &= \frac{1}{Q}\sum_{u=-j}^{j}\sum_{k=-j}^{j} U(a+u, b+k) \\ &\cdot \frac{1}{2\tau\delta^2}\exp\left(-\frac{u^2+k^2}{2\delta^2}\right) \end{aligned} \tag{7}$$

$$U' = U \times \left(1 - X_j^*\right) + U_f \times X_j^* \tag{8}$$

During the training process, the enhanced images obtained through the data augmentation mechanism are input into the network for a second forward pass to obtain the cross-entropy loss. This loss is then combined with the cross-entropy loss obtained from the network's forward propagation and the region independence loss to compute the total loss function.

$$loss = loss_{ZR} + loss_{ZR}^{'} + \eta loss_{EUM} \tag{9}$$

3. BLOCKCHAIN-BASED OPTIMIZATION FOR CROSS-BORDER FINANCIAL TRANSACTION AUTHENTICITY CONFIRMATION

Figure 4 shows the cross-border financial transaction authenticity confirmation model. In blockchain-based cross-border financial transactions, in order to ensure the authenticity and security of the transaction, this paper designs transaction protocols that are suitable for both the pre-transaction stage and the formal transaction stage. The design of the pre-transaction stage mainly includes using zero-knowledge proof smart contracts to confirm transaction information and access permissions. Zero-knowledge proof technology allows both parties in a transaction to prove the authenticity and validity of their transaction data without disclosing specific transaction information. This method not only ensures the privacy of the participants' information but also prevents malicious tampering or leakage of information. In addition, the application of smart contracts ensures the automatic execution of transaction rules and conditions, avoiding the risks and uncertainties caused by human intervention. Through this design, the protocol in the pre-transaction stage can effectively solve the problems of information asymmetry and trust, laying a solid foundation for the smooth execution of the formal transaction.

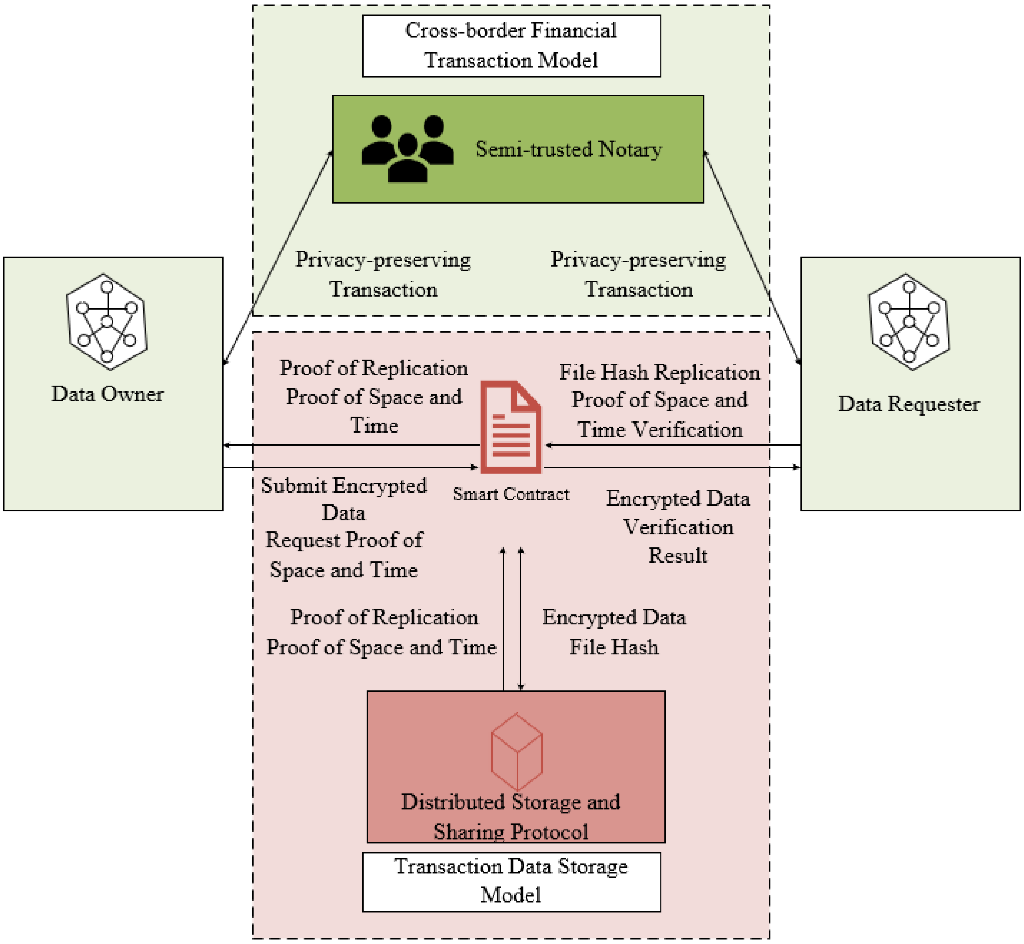
Cross-border Financial Transaction Model
Semi-trusted Notary
Privacy-preserving Transaction
Privacy-preserving Transaction
Data Owner
Data Requester
Proof of Replication Proof of Space and Time
File Hash Replication Proof of Space and Time Verification
Submit Encrypted Data Request Proof of Space and Time
Smart Contract
Encrypted Data Verification Result
Proof of Replication Proof of Space and Time
Encrypted Data File Hash
Distributed Storage and Sharing Protocol
Transaction Data Storage Model

Figure 4. Cross-border financial transaction authenticity confirmation model

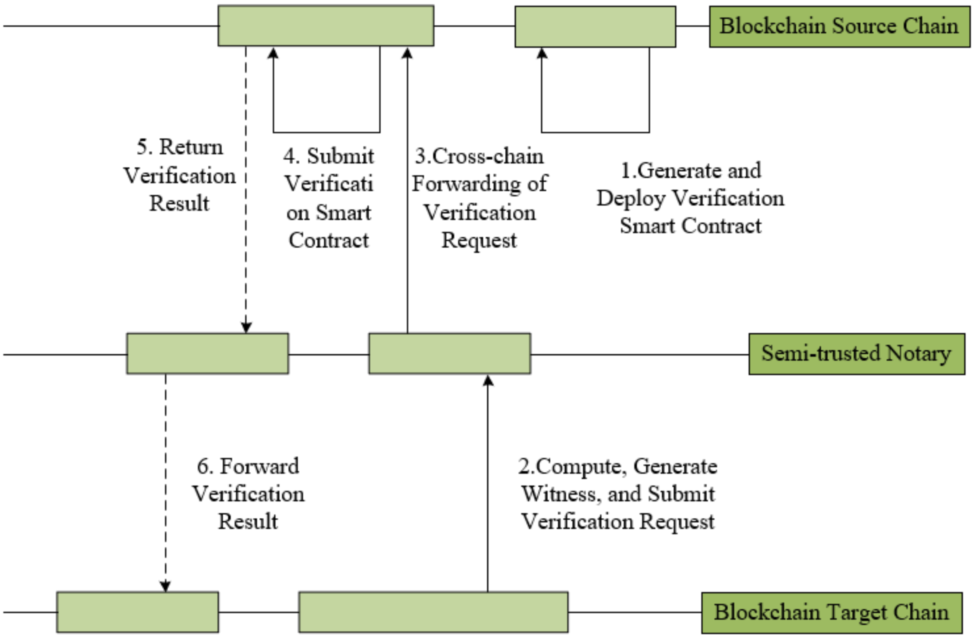
Blockchain Source Chain
5. Return Verification Result
4. Submit Verificati on Smart Contract
3.Cross-chain Forwarding of Verification Request
1.Generate and Deploy Verification Smart Contract
Semi-trusted Notary
6. Forward Verification Result
2.Compute, Generate Witness, and Submit Verification Request
Blockchain Target Chain

Figure 5. Sequence diagram of the cross-border financial pre-transaction scenario

In the formal transaction stage, this paper proposes completing the actual transaction operation with the assistance of a notary and using proxy re-encryption technology to implement a "decentralized" design for the notary. The core purpose of this design is to ensure transaction security while further enhancing the privacy and decentralization features of the transaction. The notary is responsible for verifying and recording the legitimacy of the cross-chain transaction, ensuring that the execution of the transaction complies with the conditions set in the pre-transaction stage. However, to prevent the notary from abusing their power or becoming a single point of failure, proxy re-encryption technology is introduced. This technology allows the transaction data to be securely re-encrypted and transferred without exposing the original data, thus decentralizing the role of the notary and dispersing potential risks.

The pre-transaction stage requires both parties to confirm the relevant information of the transaction. The following scenario occurs: The data requester *E* needs a specific type of data and then deploys a smart contract on the blockchain to complete cross-chain verification. The data owner *L* wishes to prove that they own the required data to the data requester *E* without revealing the data contents and hopes to obtain benefits through this data. *L* will calculate and construct a solution for the zero-knowledge proof based on the data they own, so that even without exposing the data itself, they can provide sufficient evidence to prove the validity of the data. In the pre-transaction stage, *L* notifies the notary to perform the cross-chain smart contract verification operation. If the verification by the zero-knowledge proof smart contract is successful, the smart contract will notify *E*, and the notary will notify *L* to confirm the transaction details, thus completing the pre-transaction stage. During this process, the notary cannot access any specific information about the data because the zero-knowledge proof only provides proof of the existence and validity of the data without disclosing its contents. This ensures that the data requester *E* can confirm that the data owner *L* indeed possesses the required data while protecting the privacy of the data owner *L*. Figure 5 presents the sequence diagram of the cross-border financial pre-transaction scenario. The specific steps are as follows:

Step 1: The data requester is usually one of the financial institutions or parties involved in the transaction. Based on the required data (such as data fields a_1,a_2,...,a_v), the requester uses the *DSL* language to write a *ZoKrates* program. *ZoKrates* is a programming framework used to generate zero-knowledge proofs. This program converts the input parameters into a series of mathematical constraints, ultimately generating a constraint system $R1CS$ is a core concept in zero-knowledge proofs that transforms program logic into a set of polynomial equations, which is convenient for subsequent proof and verification.

$$\begin{aligned}&def\ main\left(a_1,a_2,...,a_v\right)\\&\rightarrow R1CS=main\left(a_1,a_2,...,a_v\right)\end{aligned}\tag{10}$$

Step 2: Next, the data requester compiles the *ZoKrates* program into flat code for further processing. The compiled program is represented as an arithmetic circuit, followed by a trusted setup, which generates the proof key and verification key n_j. Trusted setup is a critical step in the zero-knowledge proof system, ensuring the system's security and validity. After completing the trusted setup, the *export-verifier* step is executed to generate a verification smart contract, which is then deployed on the blockchain. This smart contract can perform on-chain verification on the blockchain, ensuring that the verification process is public, transparent, and tamper-proof.

$$compile\left(R1CS\right)\rightarrow oj,nj\tag{11}$$

$$\exp ort\_verifier\left(nj,R1CS\right)\rightarrow TZ\tag{12}$$

Step 3: The data owner needs to prove that they own data that meets the requirements of the data requester. The data owner first executes the *compute-witness* step based on their data, generating a witness, which is part of the proof computation. Then, the *generate-proof* operation is executed to produce the zero-knowledge proof and the public reference string, which together constitute the data owner's proof materials. The data owner sends the proof *o* and the reference string *e* to the verification smart contract on the blockchain. The smart contract verifies the submitted proof, and if the verification passes, it confirms that the data owner indeed possesses the required data. The verification smart contract notifies both parties and the notary through the blockchain, confirming the authenticity of the data and preparing for the formal transaction.

$$\begin{aligned}&compute\_witness\left(R1CS,\left\{a_1,a_2,...,a_v\right\}\right)\\&\rightarrow witness\end{aligned}\tag{13}$$

$$generate\_proof\left(R1CS,witness\right)\rightarrow o,e\tag{14}$$

$$verify\_proof\left(nj,o,e\right)\rightarrow\langle TURE,FALSE\rangle\tag{15}$$

The specific steps for the formal cross-border financial transaction stage are as follows:

Step 1: Before the formal transaction stage begins, all parties need to complete the preparation work for the pre-transaction stage, including identity verification and transaction pre-audit. The two parties, the data owner *L* and the data requester *E*, publish their respective public keys OJ_L and OJ_E on the blockchain. As a distributed ledger, the blockchain ensures the authenticity and immutability of these public keys. When the transaction formally begins, the data owner *L*, after confirming the transaction, will use their public key OJ_L to encrypt the data *F* to be traded, generating the ciphertext Z_{OJL}. This encryption process ensures the privacy of the data; even if the data is intercepted during transmission, unauthorized third parties cannot obtain the plaintext content of the data.

$$Enc\left(OJ_L,F\right)\rightarrow Z_{OJ_L}\tag{16}$$

Step 2: Next, the data owner *L* needs to hand over the encrypted data to the notary. To ensure that the data requester *E* can decrypt the data, *L* will execute a key conversion algorithm, converting their key into a conversion key $EJ_{L\rightarrow E}$ that *E* can decrypt, and send this conversion key to the notary. At this point, the conversion key $EJ_{L\rightarrow E}$ is encrypted with the notary's public key to ensure that only the notary can decrypt and access it. Once the notary decrypts and obtains $EJ_{L\rightarrow E}$, they execute the ciphertext re-encryption algorithm *ReEncrypt*, converting the ciphertext Z_{OJL} into an intermediate ciphertext

Z_{OJL} that can be decrypted by the data requester E, and then send the Z_{OJL} to the data requester RE. This process ensures the secure transmission of the data, while leveraging the immutability and traceability of the blockchain to guarantee the credibility of the transaction.

$$\mathrm{Re}\,KeyGen\left(TJ_L,OJ_E\right)\rightarrow EJ_{L\rightarrow E} \tag{17}$$

Step 3: In this process, zero-knowledge proof technology is widely applied to various stages of the transaction, ensuring the authenticity of the transaction and privacy protection. During the ciphertext re-encryption and transmission process, the notary uses zero-knowledge proof to demonstrate the legitimacy and correctness of their actions to the blockchain network without disclosing any information related to the plaintext F. Specifically, the notary can generate a zero-knowledge proof to prove that they correctly executed the re-encryption operation without revealing the content of the ciphertext. Other nodes in the blockchain network verify this zero-knowledge proof to confirm the legality of the transaction, thereby ensuring the authenticity and reliability of the transaction.

$$REENC\left(EJ_{L\rightarrow E},Z_{OJ_L}\right)\rightarrow Z_{OJ_E} \tag{18}$$

Step 4: After receiving the intermediate ciphertext Z_{OJL}, the data requester E decrypts it using their private key TJ_E with the decryption algorithm *Decrypt* to obtain the plaintext F. The entire cross-border data transaction process is successfully completed while ensuring the privacy and security of the data. With the help of the blockchain, every step of the transaction process is recorded and verified, and any attempt to tamper with the transaction data is detected and prevented in real time.

$$Dec\left(C_{PK_R},SK_R\right)\rightarrow M \tag{19}$$

$$DE\left(Z_{OJ_E},TJ_E\right)\rightarrow L \tag{20}$$

4. EXPERIMENTAL RESULTS AND ANALYSIS

To validate the effectiveness of the counterfeit document recognition method, this paper uses a publicly available document image dataset, which includes both real and counterfeit document images from multiple countries. During testing, we assume that the quality of the document images is good and that the quality of the counterfeit documents is similar to that of the real documents. For the blockchain component, we simulated a smart contract environment based on the Ethereum platform, assuming that both transaction parties possess a certain level of technical expertise and are capable of interacting using zero-knowledge proofs.

From the data in Table 1, it can be seen that different auxiliary loss functions and data augmentation strategies significantly impact the training and testing results of the counterfeit document recognition method. Without any data augmentation strategy, the proposed loss function achieves an accuracy of 96.87% on the training set and 64.25% on the test set, which is notably higher than the cases with no auxiliary loss function and *AM Softmax*. After introducing the hard erasure data augmentation strategy, the test set results for *AM Softmax* and the proposed loss function are 62.36% and 63.54%, respectively, indicating that hard erasure has a certain negative impact on the recognition performance. However, when using the soft erasure data augmentation strategy, the test set accuracies for *AM Softmax* and the proposed loss function increase to 65.59% and 66.98%, respectively, further proving that the proposed loss function performs best when combined with the soft erasure strategy.

Table 1. Comparison of different auxiliary loss functions and data augmentation strategies

Loss Function Type	Data Augmentation Type	Training Set	Test Set
None	None	95.26	63.21
AM Softmax	None	95.64	63.58
Proposed Loss Function	None	96.87	64.25
AM Softmax	Hard Erasure	95.32	62.36
Proposed Loss Function	Hard Erasure	96.48	63.54
AM Softmax	Soft Erasure	95.21	65.59
Proposed Loss Function	Soft Erasure	96.38	66.98

Table 2. Performance comparison of different methods on the training set with different image compression types

Method	High Compression		Low Compression	
	ACC	*AUC*	*ACC*	*AUC*
Copy-Paste Detection	54.23	-	71.26	
Double JPEG Compression	57.26	-	77.69	-
GhostNet + Attention Mechanism	71.32	-	82.31	-
WGAN-GP	-	62.31	-	86.23
CycleGAN	85.36	88.26	94.23	95.64
Patch-GAN	78.52	91.36	92.31	93.21
StyleGAN Tampering Detection	86.51	91.56	-	-
CRNN+CNN	91.23	92.68	96.58	97.26
YOLOv8-OCR	-	85.63	-	97.51
CLIP+Prompt	85.69	87.54	95.31	98.36
Proposed Method	87.23	91.23	96.58	98.65

From the data in Table 2, it can be seen that different methods show significant differences in performance on high-compression and low-compression image training sets. On the high-compression image training set, the proposed method achieves an accuracy (ACC) of 87.23% and Area Under the Curve (AUC) of 91.23%, showing superior performance compared to other methods, second only to CRNN+CNN (91.23% ACC, 92.68% AUC). On the low-compression image training set, the proposed method achieves an accuracy and AUC of 96.58% and 98.65%, respectively, which is on par with CRNN+CNN, and slightly higher than CRNN+CNN in terms of AUC (97.26%). Overall, the proposed method consistently maintains high recognition accuracy and AUC values, especially on the low-compression image training set, where its AUC exceeds that of most methods.

The experimental results indicate significant performance differences across various methods for counterfeit document recognition. Under high compression, methods like CRNN-CNN and the proposed method performed excellently; under low compression, the proposed method, along with CLIP-Prompt, showed strong results. The proposed method maintained high ACC and AUC in both high and low compression scenarios, demonstrating good accuracy and robustness in counterfeit document recognition across different compression conditions. This aligns with the paper's design of "enhancing recognition accuracy by combining local texture features with high-level semantic features."

In blockchain transactions, accurate and efficient counterfeit document recognition can quickly verify the authenticity of the transaction parties, reducing validation time and costs. With its high accuracy, the proposed method can effectively filter counterfeit documents, improving the efficiency of identity verification in the pre-transaction phase of blockchain transactions. This lays a reliable foundation for subsequent blockchain transaction processes, helping optimize the authenticity confirmation process in cross-border financial transactions, reducing the risk of identity fraud, and ensuring the security and efficiency of blockchain transactions.

Table 3. Generalization performance comparison of different methods on different types of datasets

Method	**Standardized Public Dataset**	**Synthetic Generated Dataset**	**Industry Business Simulation Dataset**
Copy-Paste Detection	71.2	51.2	52.36
Double JPEG Compression	83.6	75.6	53.48
GhostNet + Attention Mechanism	82.5	74.1	52.31
WGAN-GP	98.6	84.3	64.52
CycleGAN	75.4	53.2	53.26
Patch-GAN	95.2	62.5	56.87
StyleGAN Tampering Detection	92.1	83.6	63.21
CRNN+CNN	-	-	64.59
YOLOv8-OCR	98.6	87.9	63.25
CLIP+Prompt	98.9	91.2	66.98
Proposed Method	98.7(+0.1)	97.8(+5.1)	77.62(+11.21)

Table 4. Performance comparison of the proposed method and baseline models with different backbone networks

Method	**High Compression**		**Low Compression**	
	ACC	*AUC*	*ACC*	*AUC*
ShuffleNetV2	85.32	88.96	94.21	95.36
Proposed Method (*ShuffleNetV2*)	85.62(+0.09)	86.32(-2.15)	95.33(+0.63)	97.56(+2.56)
ResNet	85.36	87.26	95.21	98.21
Proposed Method (*ResNet*)	87.26(+2.13)	91.23(+2.16)	96.58(+0.96)	98.26(+0.12)
DenseNet	86.23	88.65	95.21	98.47
Proposed Method (*DenseNet*)	88.36(+2.14)	92.36(+1.78)	96.39(+0.95)	98.36(+0.22)
RegNet	91.23	91.58	96.35	98.74
Proposed Method (*RegNet*)	91.23(+1.14)	92.36(+1.17)	97.58(+1.21)	98.36(+0.42)

From the data in Table 3, it can be seen that different methods show significant differences in generalization performance on the three types of datasets. On the standardized public dataset, the proposed method demonstrates a very high accuracy of 98.7%, second only to *CLIP+Prompt* (98.9%). On the synthetic generated dataset, the proposed method achieves an accuracy of 97.8%, significantly higher than other methods, especially compared to *CLIP+Prompt* (91.2%), where the proposed method improves by 5.1 percentage points. On the industry business simulation dataset, the proposed method also performs excellently with an accuracy of 77.62%, significantly leading other methods, where *CLIP+Prompt* and *YOLOv8-OCR* have accuracies of 66.98% and 63.25%, respectively. Overall, the proposed method demonstrates outstanding generalization ability on all datasets, especially on the more challenging synthetic generated dataset and industry business simulation dataset, where it stands out even more.

The experimental results demonstrate that the proposed method excels in generalization performance for counterfeit document recognition, particularly in the industry business simulation dataset, where the result of 77.62 (±11.21) indicates that it can maintain high accuracy even in complex real-world scenarios, with the fluctuation range remaining within an acceptable range. Compared to other methods, such as CLIP-Prompt (66.98) and StyleGAN tampering detection (63.21), the proposed method is more robust.

In blockchain transactions, generalization performance directly impacts the efficiency and security of identity verification. In the pre-transaction phase, zero-knowledge proof smart contracts rely on accurate document recognition to verify the identities of the transaction parties. The high generalization capability of the proposed method allows for rapid filtering of counterfeit documents, reducing validation time and the risk of misjudgment, thereby laying a reliable foundation for subsequent transaction processes. In the formal transaction phase, the combination of reliable identity verification with proxy re-encryption technology can further enhance transaction authenticity and the degree of decentralization, reducing fraud risks and ensuring efficient and secure cross-border financial transactions.

In summary, the proposed method significantly improves the accuracy and efficiency of counterfeit document recognition in blockchain transactions through outstanding generalization performance, optimizing the process of confirming transaction authenticity.

Table 4 shows the performance comparison between the proposed method and baseline models with different backbone networks under high and low compression scenarios. Using ShuffleNetV2 as the backbone network, the proposed method achieves an ACC of 85.62% and an AUC of 86.32% under high compression, which is an increase of 0.09 percentage points in ACC but a decrease of 2.15 percentage points in AUC compared to the baseline model. In the low compression case, the proposed method achieves an ACC of 95.33% and an AUC of 97.56%, which represents an improvement of 0.63 and 2.56 percentage points, respectively. Using ResNet as the backbone network, the proposed method shows an increase of 2.13 percentage points in ACC and 2.16 percentage points in AUC under high compression, and an increase of 0.96 percentage points in ACC and 0.12 percentage points in AUC under low compression. Using DenseNet and RegNet as backbone networks, the proposed method shows various improvements in ACC and AUC across all cases, especially under high compression. The ACC and AUC for DenseNet improve by 2.14 and 1.78 percentage points, respectively, while RegNet shows improvements of 1.14 and 1.17 percentage points. The experimental results indicate that the proposed method significantly improves performance across different backbone networks, especially under high compression data, with the most noticeable improvements in ResNet and DenseNet. This demonstrates that the proposed method can maintain high recognition accuracy and AUC values even in high compression data environments, showcasing its robustness and adaptability across different compression levels. Combining blockchain technology for transaction authenticity verification can further enhance the security of cross-border financial transactions by improving the reliability and efficiency of identity verification and reducing the risk of financial fraud. Overall, the experimental data across different backbone networks strongly supports the feasibility and effectiveness of applying the proposed method to cross-border financial transactions, providing reliable technical assurance for improving transaction security.

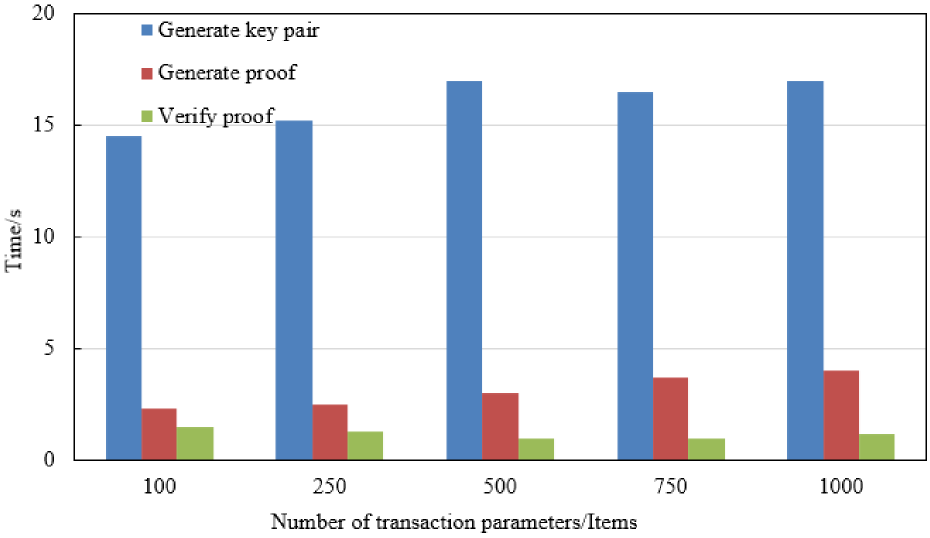

Figure 6. Average time for key pair generation, proof generation, and proof verification in the pre-transaction phase

Figure 6 shows the average time required for key pair generation, proof generation, and proof verification under different numbers of transaction parameters. Specifically, when the number of transaction parameters is 100, the time for key pair generation, proof generation, and proof verification is 14.5 seconds, 2.3 seconds, and 1.5 seconds, respectively. When the number of transaction parameters is 250, these times are 15.2 seconds, 2.5 seconds, and 1.3 seconds. With 500 parameters, the times are 17 seconds, 3 seconds, and 1 second. For 750 parameters, the times are 16.5 seconds, 3.7 seconds, and 1 second. With 1000 parameters, they are 17 seconds, 4 seconds, and 1.2 seconds. It can be observed that as the number of transaction parameters increases, the time for proof generation gradually increases, while the time for key pair generation and proof verification shows relatively little change. In fact, the proof verification time is even the shortest at 1 second for 500 and 750 parameters. The experimental data indicates that the blockchain-based cross-border financial transaction authenticity verification process shows stability and efficiency across different numbers of transaction parameters. Particularly, the proof verification time remains stable and is not significantly affected by the number of transaction parameters, consistently staying around 1 second. This demonstrates that the method offers high real-time performance and processing efficiency when handling a large number of transactions. The key pair generation time fluctuates slightly with the increase in transaction parameters but remains relatively stable, ranging from 14.5 to 17 seconds, indicating that its computational complexity is stable. The proof generation time increases as the number of transaction parameters grows, suggesting that this step is more sensitive to the number of parameters. However, the increase is within an acceptable range, providing a good balance between performance and computational cost.

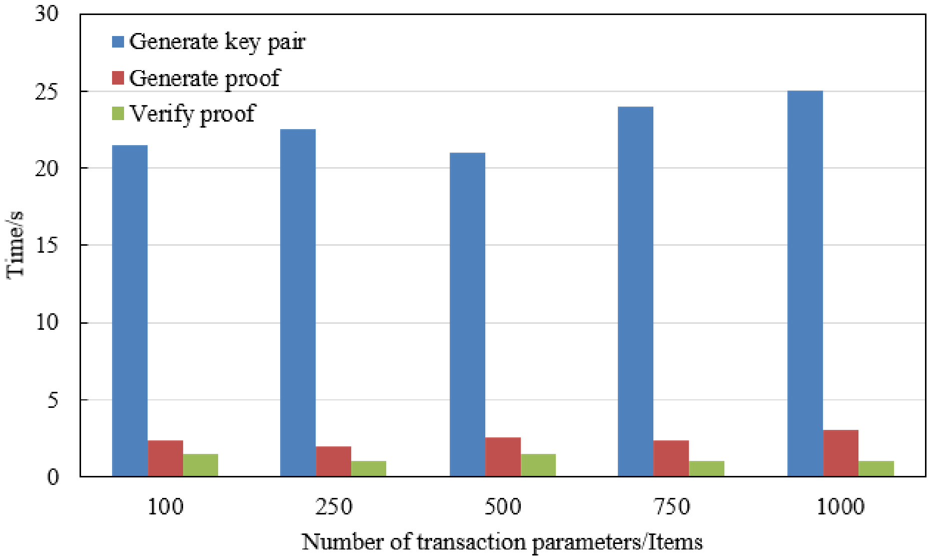

Figure 7. Average time for key pair generation, proof generation, and proof verification in the official transaction phase

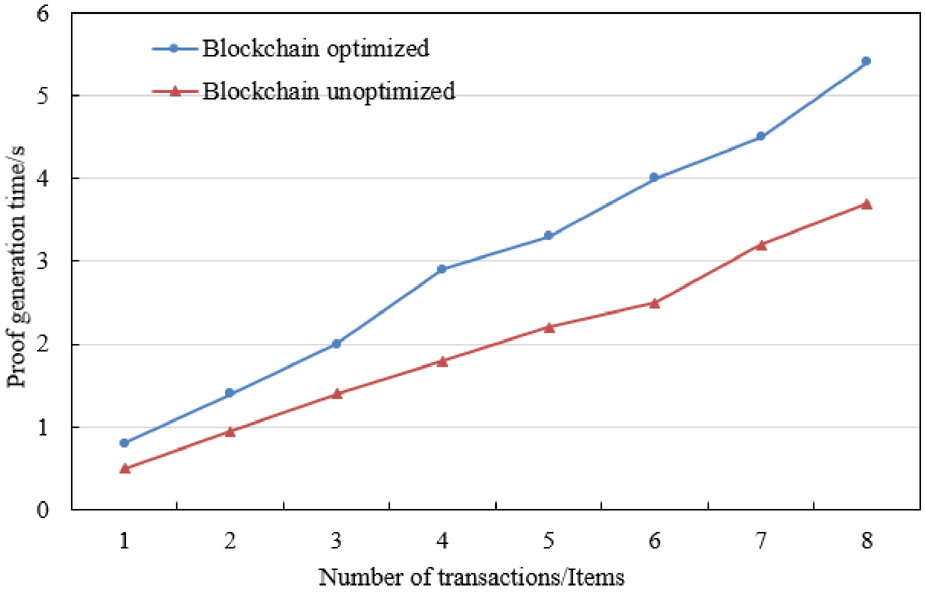

Figure 8. Proof generation time under different transaction numbers

Figure 7 shows the performance of the blockchain-based cross-border financial transaction authenticity verification method under different numbers of transaction parameters in the official transaction phase. Specifically, when the number of transaction parameters is 100, the time for key pair generation, proof generation, and proof verification is 21.5 seconds, 2.3 seconds, and 1.5 seconds, respectively. When the number of transaction parameters is 250, these times are 22.5 seconds, 2 seconds, and 1 second. With 500 parameters, the times are 21 seconds, 2.5 seconds, and 1.5 seconds. For 750 parameters, the times are 24 seconds, 2.3 seconds, and 1 second. With 1000 parameters, they are 25 seconds, 3 seconds, and 1 second. It can be observed that the key pair generation time generally increases with the number of transaction parameters. The time for proof generation shows little variation across different parameter numbers, and the proof verification time remains stable between 1 and 1.5 seconds, demonstrating strong stability. The experimental data suggests that the key pair generation time increases as the number of transaction parameters grows, indicating that the computational overhead for key pair generation rises with more parameters. However, this increase is within a reasonable range, and it does not significantly affect the overall performance of the system. The proof generation time fluctuates slightly as the number of parameters increases but mostly remains between 2 and 3 seconds, indicating that the computational complexity of this step is manageable. The proof verification time is extremely stable, remaining between 1 and 1.5 seconds, showing that the method offers high real-time performance and reliability even when processing large volumes of transactions.

Figure 8 shows the time required for cross-border financial transaction authenticity verification, both before and after blockchain optimization, when processing different numbers of transactions. Specifically, when the number of transactions is 1, the time before and after optimization is 0.5 seconds and 0.8 seconds, respectively. For 2 transactions, the times are 0.95 seconds and 1.4 seconds; for 3 transactions, the times are 1.4 seconds and 2 seconds; for 4 transactions, the times are 1.8 seconds and 2.9 seconds; for 5 transactions, the times are 2.2

seconds and 3.3 seconds; for 6 transactions, the times are 2.5 seconds and 4 seconds; for 7 transactions, the times are 3.2 seconds and 4.5 seconds; and for 8 transactions, the times are 3.7 seconds and 5.4 seconds. It can be observed that as the number of transactions increases, the time after blockchain optimization gradually increases, but the overall time is still kept within a reasonable range. The experimental data suggests that although the blockchain-optimized cross-border financial transaction authenticity verification method requires more time to process multiple transactions than before optimization, this increase is within an acceptable range, and it reflects the advantages of blockchain technology. Specifically, although the time after optimization increases significantly with the number of transactions, from 0.8 seconds for 1 transaction to 5.4 seconds for 8 transactions, the growth is more linear and stable compared to the pre-optimization scheme. This indicates that blockchain technology provides more reliable data verification and higher security when handling complex transactions.

Although the proposed integration of AI-based image recognition and blockchain technology has achieved notable results in enhancing the security of cross-border financial transactions, several limitations remain.

First, AI image recognition methods may struggle to accurately identify low-quality images or sophisticated malicious forgeries, indicating a need to further improve the robustness of the recognition network. Second, the high computational and storage overhead of blockchain technology can become a performance bottleneck when handling large-scale transactions, especially in real-time scenarios. Optimizing blockchain protocols to enhance transaction processing efficiency remains a critical research direction.

Finally, while blockchain offers decentralization advantages, transaction operations that involve notary assistance still rely on a certain level of trust mechanisms. Further exploration is needed to eliminate centralized trust points entirely and achieve fully trustless systems.

5. CONCLUSION

This study explored the application of combining AI image recognition with blockchain technology in cross-border financial transactions to enhance the security and reliability of transactions. The research is divided into two main parts: first, a forged document recognition method based on AI image recognition technology was designed to address identity fraud in cross-border financial transactions; second, blockchain technology was used to optimize the authenticity verification process of cross-border financial transactions. The experimental results show that the blockchain-optimized method, while requiring more time to process multiple transactions, increases in a more linear and stable manner compared to the pre-optimization solution, and the overall time remains within a reasonable range. By integrating AI image recognition technology, the accuracy of identity verification and the security of the entire transaction process had been effectively improved in real-world transaction scenarios, significantly reducing fraud risks.

The research has significant value in improving the security of cross-border financial transactions. By integrating AI image recognition technology, forged documents can be effectively detected, preventing identity fraud. Through the introduction of blockchain technology, the authenticity and immutability of transaction data are enhanced, ensuring the transparency and security of the transaction process.

Although the proposed integration of AI image recognition and blockchain technology has achieved significant results in enhancing the security of cross-border financial transactions, the study has also identified several challenges. These include the method's dependency on image quality and the efficiency issues faced by blockchain systems when handling large-scale transactions.

Future research can focus on optimizing image recognition algorithms to improve robustness against low-quality images, as well as exploring more efficient blockchain protocols to maintain high transaction speeds in real-time scenarios. Furthermore, with the rise of decentralized finance applications, applying the proposed approach to decentralized trading platforms represents a promising direction for further investigation.

ACKNOWLEDGMENT

Anhui Province University Science Research Project: Research on the Mechanism and Effect of Digital Finance Promoting High Quality Economic Development in Anhui Province under the New Development Pattern (Project No: 2022AH051730).

REFERENCES

[1] Sekgoka, C.P., Yadavalli, V.S.S., Adetunji, O. (2022). Privacy-preserving data mining of cross-border financial flows. Cogent Engineering, 9(1): 2046680. https://doi.org/10.1080/23311916.2022.2046680

[2] Jääskeläinen, M., Maula, M. (2014). Do networks of financial intermediaries help reduce local bias? Evidence from cross-border venture capital exits. Journal of Business Venturing, 29(5): 704-721. https://doi.org/10.1016/j.jbusvent.2013.09.001

[3] Weitzel, U., Kling, G., Gerritsen, D. (2014). Testing the fire-sale FDI hypothesis for the European financial crisis. Journal of International Money and Finance, 49: 211-234. https://doi.org/10.1016/j.jimonfin.2014.03.011

[4] Rao-Nicholson, R., Ayton, J.S. (2016). Euphoria in financial markets: How Indian companies generate value in their cross-border acquisitions. Research in International Business and Finance, 38: 494-508. https://doi.org/10.1016/j.ribaf.2016.07.022

[5] Farah, N., Souici, L., Sellami, M. (2005). Decision fusion and contextual information for Arabic words recognition for computing and informatics. Computing and Informatics, 24(5): 463-479.

[6] Sayallar, C., Sayar, A., Babalık, N. (2023). An OCR engine for printed receipt images using deep learning techniques. International Journal of Advanced Computer Science and Applications, 14(2): 833-840. https://doi.org/10.14569/ijacsa.2023.0140295

[7] Kuo, T.T., Kim, H.E., Ohno-Machado, L. (2017). Blockchain distributed ledger technologies for biomedical and health care applications. Journal of the American Medical Informatics Association, 24(6): 1211-1220. https://doi.org/10.1093/jamia/ocx068

[8] Morkunas, V.J., Paschen, J., Boon, E. (2019). How blockchain technologies impact your business model.

Business Horizons, 62(3): 295-306. https://doi.org/10.1016/j.bushor.2019.01.009

[9] Tyagi, P., Agarwal, K., Jaiswal, G., Sharma, A., Rani, R. (2024). Forged document detection and writer identification through unsupervised deep learning approach. Multimedia Tools and Applications, 83(6): 18459-18478. https://doi.org/10.1007/s11042-023-16146-7

[10] Bhandari, M., Tiwari, G., Dhakal, M. (2025). Enhancing transparency and accountability in sustainable finance through blockchain technology: A systematic review of the literature. Journal of Intelligent Management Decision, 4(1): 23-43. https://doi.org/10.56578/jimd040102

[11] Lakkakula, P., Bullock, D., Wilson, W. (2020). Blockchain technology in international commodity trading. Journal of Private Enterprise, 35(2): 23-46. https://doi.org/10.22004/ag.econ.303613

[12] Fu, W.S., Du, J.Q., Zhang, Y., Wang, Z.Q. (2024). A blockchain cross-chain solution based on relays. International Journal of Knowledge and Innovation Studies, 2(2): 70-80. https://doi.org/10.56578/ijkis020202

[13] Siddik, M.N.A., Kabiraj, S., Hosen, M.E., Miah, M.F. (2021). Blockchain technology and facilitation of international trade: An empirical analysis. FIIB Business Review, 10(3): 232-241. https://doi.org/10.1177/2319714520968297

[14] Kazan, G., Kocamış, T.U. (2023). Assessing the impact of blockchain technology on internal controls within the COSO framework. Journal of Corporate Governance, Insurance, and Risk Management, 10(1): 86-95. https://doi.org/10.56578/jcgirm100110

[15] Böhmecke-Schwafert, M. (2024). The role of blockchain for trade in global value chains: A systematic literature review and guidance for future research. Telecommunications Policy, 48(9): 102835. https://doi.org/10.1016/j.telpol.2024.102835

[16] Kumar, N., Goel, V., Ranjan, R., Hassan, M.M., Pandey, T.K., Dwivedi, A. (2024). A novel IoT-blockchain methodology to augment conviction in electronic health records management. Traitement du Signal, 41(5): 2279-2286. https://doi.org/10.18280/ts.410505

[17] Kowalski, M., Lee, Z.W., Chan, T.K. (2021). Blockchain technology and trust relationships in trade finance. Technological Forecasting and Social Change, 166: 120641. https://doi.org/10.1016/j.techfore.2021.120641

[18] Grzesiak, J., Górski, K., Kasperczak, A. (2024). Application of image recognition algorithms in unmanned aerial vehicles. Przegląd Elektrotechniczny, 100(10): 224-227. https://doi.org/10.15199/48.2024.10.46

[19] Dashkevich, N., Counsell, S., Destefanis, G. (2024). Blockchain financial statements: Innovating financial reporting, accounting, and liquidity management. Future Internet, 16(7): 244. https://doi.org/10.3390/fi16070244

[20] Kao, J.H., Chen, Y.H., Chuang, J.H. (2005). Identity verification by relative 3D structure using multiple facial images. Pattern Recognition Letters, 26(9): 1292-1303. https://doi.org/10.1016/j.patrec.2004.11.008

[21] Smith, K.J., Dhillon, G. (2020). Assessing blockchain potential for improving the cybersecurity of financial transactions. Managerial Finance, 46(6): 833-848. https://doi.org/10.1108/MF-06-2019-0314