

# Legal Impact of Digital Information Technology on the Chain of Evidence in Criminal Cases

Liupeng Zhao<sup>1,✉</sup>

<sup>1</sup> *TC Beirne School of Law, The University of Queensland, Brisbane, Queensland, 4072, Australia*

## ABSTRACT

Criminal evidence serves as the foundation for criminal proceedings, with evidence used to ascertain the facts of cases being critical to achieving fairness and justice. This study explores the application of digital information technology in building a data resource base for criminal cases, formulating standard evidence guideline rules, and optimizing evidence verification procedures. A named entity recognition model based on the SVM-BiLSTM-CRF framework is proposed, coupled with an evidence relationship extraction model using the Transformer framework to improve evidence information extraction through sequential features and global feature capturing. Results show that the F1 value for entity recognition in criminal cases reaches 94.19%, and the evidence extraction model achieves an F1 value of 81.83% on the CAIL-A dataset. These results are utilized to construct evidence guidelines, helping case handlers increase case resolution rates to approximately 99%. The application of digital technology enhances evidence collection efficiency, accelerates case closures, and offers a pathway to improving judicial credibility.

*Keywords:* SVM-BiLSTM-CRF model, Transformer model, Entity relationship extraction, Named entity recognition, Criminal case

## 1. Introduction

Throughout the history of the development of judicial proof methods, mankind has successively experienced the era of "divine evidence" based on divine evidence, the era of "human evidence" based on interrogation and interrogation, and the era of "physical evidence" based on scientific proof [17, 13]. With the advent of the digital era, digital evidence has rapidly penetrated into a variety of judicial proof activities, is a powerful influence and promote the development and change of judicial

✉ Corresponding author.

*E-mail addresses:* [s997881751@163.com](mailto:s997881751@163.com) (Liupeng Zhao).

Received 01 September 2024; accepted 05 November 2024; published 31 December 2024.

DOI: [10.61091/jcmcc123-08](https://doi.org/10.61091/jcmcc123-08)

© 2024 The Author(s). Published by Combinatorial Press. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

proof method. We are entering the information to prove the main "digital evidence" era [19, 8].

Accurate definition of digital evidence is the starting point for the study of digital evidence, which is the main type of electronic evidence. At present, some domestic literature on electronic evidence, there are two representative views. The first point of view will be electronic evidence and computer evidence concept of equivalence, and defined as, in the computer or computer system operation process to record the contents of the case to prove the facts of the electromagnetic recordings, also known as computer evidence [4, 10]. The second point of view is a broad definition method, if from the perspective of evidence is too broad. By the electronic form of material transformed into subsidiary material that is derived (such as computer printouts), not included in the scope of electronic evidence, only as electronic evidence of a way of manifestation. This clearly points out that the object of electronic evidence is to prove the role of electronic data and information, the process of evidence is to obtain, analyze electronic data and information process [12, 6, 11].

At present, China's reference to digital evidence is not common, but the existence of electronic data and information in the form of two, analog and digital, and the two can be converted to each other [2, 18]. Therefore, electronic evidence can be divided into digital evidence and analog evidence. Analog signal has continuity, its content once the clip changes, easier to identify. The digital signal is discrete, its content is not easy to leave traces after the change, difficult to identify [20]. Different technical basis for digital evidence and analog evidence of the huge difference between. With the development of digital technology, digital form of information will occupy a dominant position in social life, digital evidence will be more and more important. Computer is the most important kind of digital equipment, and digital evidence includes computer evidence [14, 5].

After the revision of the criminal procedure law in 2012, in article 48 will be "electronic data" and "audio-visual materials" side by side, together as the eighth type of evidence to be stipulated, the independent legal status of electronic data can be established. After the electronic data become a kind of independent evidence, convenient for the judicial practice of its use, but due to the electronic data and other types of evidence of different classification standards, especially with the documentary evidence, audio-visual materials and physical evidence, there is a certain degree of cross overlap, in China's system of classification of evidence has caused a certain degree of confusion [16, 3].

The performance of digital information technology in criminal evidence review judgment is very bright, but confined to the technology and criminal justice and other aspects of the specificity of the limitations, technology in the criminal evidence chain review judgment in the application of the dilemma. In this paper, we start from the application of evidence chain in criminal cases empowered by digital information technology, relying on digital technology to establish a data resource base for criminal cases, optimize the evidence verification idiom in criminal cases, and reconstruct the evidence standard guideline rules. On this basis, the SVM model is used to extract the evidence feature vectors of criminal cases, which is used as the input of BiLSTM-CRF model to realize the recognition of named entities in criminal cases. Combining the recognition results with the evidence framework of criminal cases, the information extraction model of criminal case evidence entities based on the Transformer model is designed, and the validation analysis is carried out for the effectiveness of the above types of models.

## 2. Evidence in Criminal Cases Supported by Digital Information Technology

The development of digital information technology has had a broad and profound impact on social life, and the legal field is no exception. And with the reform of the criminal justice system, digital information technology has also begun to influence and apply to the examination and judgment of criminal evidence, which is a kind of deep innovation of "science and technology + law + reason". The examination and judgment of the evidence chain in criminal cases is a very important link in criminal proceedings, and plays a key role in the formation of the judgment of the case. Once the chain of evidence review judgment deviation, the factual determination will certainly also appear deviation. Therefore, not only to the individual evidence for evidence and proof of the review, but also to the whole case of evidence for a comprehensive review and judgment, and strive to make conclusive conclusions, otherwise, shall not be determined the facts of the case. Therefore, digital information technology as an emerging technology, the use of the chain of evidence in criminal cases in the examination and judgment, to the chain of evidence in criminal cases review judgment has brought new opportunities and challenges.

### *2.1. Evidence and digital repositories in criminal cases*

**2.1.1. Judgment of evidence in digital criminal cases.** Whether digital information technology can be applied in the field of judging the chain of evidence in criminal cases depends mainly on whether it is possible to depart from the standard of evidence to guide the norms. Digital information technology can only realize the intelligent operation of the evidence judgment on the basis of the evidence type is determined and standardized. If for different cases, the same kind of evidence standards and norms are not the same, then digital information technology will not have an established standard to learn and judge the same evidence of proof and proof content, so that it can not realize the chain of evidence in criminal cases of intelligent construction and review role. Therefore, for the combination of digital information technology and the chain of evidence in criminal cases, the introduction of a set of standard guidelines for evidence is a particularly important factor.

In judicial practice, we can see that for different cases, the collection of evidence and evidence citation standards and norms have a certain pattern. Different murder cases can be broadly categorized into the following types, namely, cases in which there are traces left at the scene, cases in which there are eyewitnesses at the scene, cases in which the criminals have confessed to the crime, and cases in which no one has confessed to the crime. And the evidence of these types of cases have similarity and regularity, if the evidence standards involved in these four types of cases can be standardized, fixed and standardized, then according to the standard guidelines of these evidence, can be roughly concluded whether the evidence of a particular case has reached the full effect. Commonality level, all the legal evidence needs to be able to form a complete and clear chain of evidence, and can be consistent with the facts identified, evidence and evidence, evidence and facts can match each other. From the level of differences in the structure of the evidence, the structure of the evidence of different cases are more or less a certain gap, this difference is mainly reflected in the form of evidence.

**2.1.2. Data repository for criminal cases.** Based on digital information technology and combining various types of technology to establish a criminal case data resource base, its purpose is to provide reliable data support for the creation of the evidence chain of criminal cases, and to assist

the case handlers with case-like information to better clarify the evidence of criminal cases, so as to obtain a complete evidence chain of criminal cases. The construction content of the criminal case data resource base is shown in Figure 1, which mainly contains a library of evidence standards, a database of issues, a library of case information, a library of case characteristics, a library of electronic files, a library of judicial documents, a library of cases, a library of judicial interpretations of laws and regulations, and a library of case-handling operational documents.

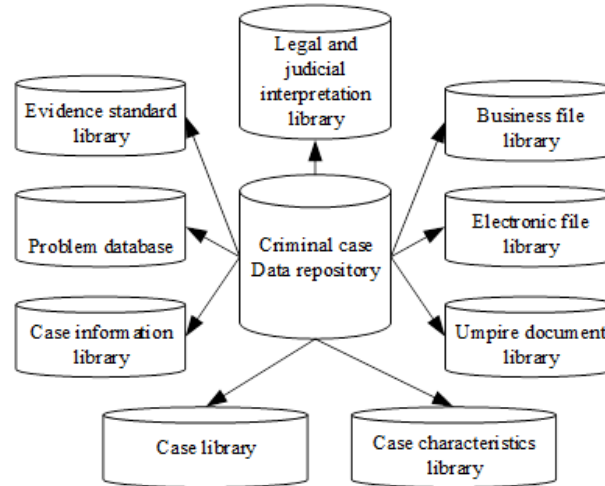


Fig. 1. Criminal case data repository construction

In the process of promoting the reform of the trial-centered litigation system, it has established a sound data resource base for criminal cases, integrating six sub-databases, including the evidence standard library, the electronic file library, the case library (including Supreme Court gazetted cases and guiding cases), the adjudicative documents library, the laws and regulations judicial interpretation library, and the case-handling operational documents library, and has shared the information resources of criminal cases with the public security and procuratorate, so as to provide the application of digital information technology with The department also shares information resources on criminal cases with the Public Security Bureau and the Procuratorate, providing strong information and data resource support and guarantee for the application of digital information technology.

## 2.2. Evidentiary standards and checking procedures in criminal cases

**2.2.1. Guidelines on standards of evidence in criminal cases.** In the process of applying digital information technology to the construction of the chain of evidence in criminal cases, it is necessary to characterize the standard of evidence in criminal cases and the standard of evidence in similar cases, with the aim of explaining what evidence needs to be collected to construct a complete chain of evidence in different types of cases through the standard of evidence, and enhancing the reliability of the chain of evidence in criminal cases. At this time, the importance of criminal case evidence standard guidelines are self-evident, and the construction process of evidence standard guidelines are as follows:

- (a) Categorize criminal cases according to the charges. For criminal cases involving common crimes in the case of its evidence standards as the focus of the construction, through the different types of cases to form a complete and closed chain of evidence required elements linked together to complete the construction of a unified standard of evidence.
- (b) On the basis of the classification of cases, according to the different characteristics of each stage

of the litigation to refine the development of the filing of the case, the end of investigation, prosecution, trial and other stages of the implementation of specific standards.

- (c) The formed evidence standard "structural transformation". Formed by the above two steps of the unified evidence standard material mostly belongs to the computer can not identify, perceive the unstructured text, and therefore the need for natural speech processing analysis technology will be transformed into a computer can handle the structured data. Through such a way of conversion, the evidence standard material can be efficiently run in the computer under the processing of the formation of data for different types of cases of uniform evidence standards.

The standards resulting from these three steps will be embedded in a digital information technology-supported framework for intelligent case handling, which will play an expected role in advancing the goals of the reform of the criminal procedure system and the realization of standard guidelines for criminal cases.

**2.2.2. Procedures for verification of evidence in criminal cases.** Based on the criminal case data resource base and the guidelines on evidence standards, it assists case handlers in realizing efficient sorting of evidence in criminal cases, and the intelligent assisted case handling framework can realize efficient verification of evidence in criminal cases. The evidence verification and analysis process relying on digital information technology has the advantages of objectivity and leaving traces throughout the entire process, strengthening the leading role of judicial personnel from the opposite direction, and practicing the goal of assisting the administration of justice through digital information technology.

The same litigation stage within the evidence verification program selection process shown in Figure 2. For the evidence calibration process of judgment bias problems, can be set in the operating system of the case officer's easy access to the program, giving the judicial personnel two types of procedural options to be resolved. First, the testing procedures between different litigation processes across, by the case officer to submit the relevant evidence materials and materials description, the system to make a mark after the direct transmission to the next litigation stage of the verification process. Second, the same litigation stage of the evidence verification process across. For the evidence based on ordinary verification process can not successfully enter the next verification stage of the evidence material, can give the case officer test program options, by the case officer discretionary treatment. If the case officer believes that the need to supplement the material, can be based on the system prompts additional collection of evidence materials. If the case officer with their own experience and judgment, that the case requires all the materials collected in the case, you can directly choose to enter the next stage.

### 3. Intelligent Extraction Modeling of Evidence in Criminal Cases

The review of evidence is the core of criminal case processing, directly determines the realization of judicial justice or not. The use of digital information technology for cracking criminal wrongdoing provides a new choice of path, the use of digital information technology can solve the problem of criminal cases in various regions of the application of evidence standards confusion, can be found in a timely manner in the qualification of evidence, the evidence of ability, forcing the standardization of evidence collection procedures, from the source to prevent the occurrence of wrongdoing. Combining

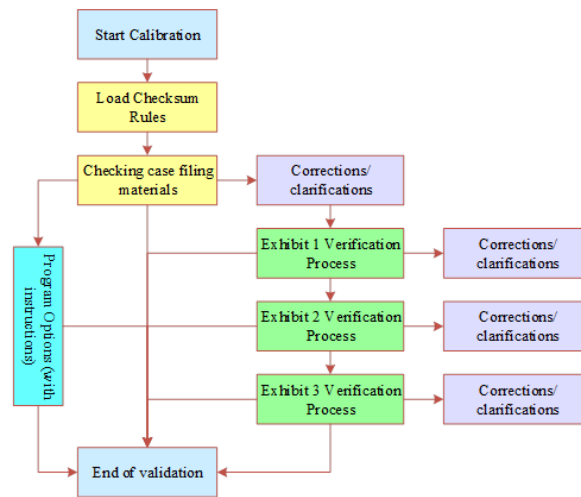


Fig. 2. The criminal case verification procedure

digital information technology with the review of evidence in criminal cases can also alleviate the pressure of "too many cases, too few people" that currently exists in the courts, and solve the problem of irregularities in case-handling procedures.

### 3.1. Named entity identification of evidence in criminal cases

**3.1.1. Support vector machines.** Based on the criminal case adjudication documents in the criminal case data repository, the named entity recognition of criminal cases helps the case handlers to better analyze the relevant chain of evidence of criminal cases. In order to better realize the named entity recognition of criminal cases, this paper screens criminal case adjudication documents in the input part.

Support Vector Machine (SVM) is a classical two-class classification algorithm, which finds a segmentation hyperplane with better robustness, so it is widely used on many tasks and shows strong advantages. Therefore this method is used in this chapter to incorporate features for model inputs [9].

Firstly, the content of the body of the adjudication document as a corpus is processed by using a word segmentation tool, and for the word embedding part we choose the TF-IDF method to get the word vector  $\{w_1, w_2 \dots w_n\}$ , where  $n$  is the feature dimension of the dataset. A sentence  $S_i$  of length  $m$  in  $S_i = \{w'_1, w'_2 \dots w'_m\}$ , processed to get the feature vector  $\{x'_1, x'_2 \dots x'_n\}$ . In this paper, we use the kernel function method, the kernel function implicitly maps the samples from the original feature space to a higher dimensional space, and solves the linear indivisibility problem in the original feature space. For example, in a transformed feature space  $\phi$ , the decision function of the support vector machine is:

$$\begin{aligned}
 f(x) &= \text{sgn}(w^T \phi(x) + b), \\
 &= \text{sgn}\left(\sum_{n=1}^N \lambda_n y^{(n)} K(x^{(n)}, x) + b\right),
 \end{aligned} \tag{1}$$

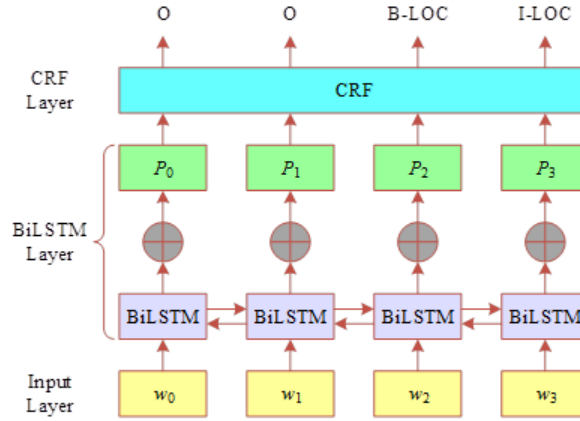
where  $K(x, z) = \phi(x)^T \phi(z)$  is the kernel function. The feature vectors are used as inputs to the model along with the label  $y_i$ .

The feature input part of the criminal case adjudication documents is first converted into corresponding word vectors by querying the word vector table, and then input into SVM for judgment. If it does not contain feature entities, all the words are labeled as O. Otherwise, the corresponding

character vector sequences are obtained by querying the character vector table, and these character vector sequences are input into BiLSTM for entity recognition. Finally the CRF module processes the output of BiLSTM to derive an optimal labeling sequence.

**3.1.2. BiLSTM+CRF modeling.** After using SVM to classify the vector features of the adjudication documents, this paper takes them as the input of the BiLSTM-CRF model, obtains the named entities of criminal cases through the BiLSTM model, and inputs them into the CRF layer in order to obtain the optimal sequences that conform to the characteristics of criminal cases, which helps the case handlers to better grasp the data related to the chain of evidence of a certain criminal case [7].

Figure 3 shows the structure of BiLSTM-CRF model, which is mainly divided into word vector input layer, BiLSTM layer and CRF layer. The main process of the experiment is to input the word vector sequence first, followed by feature extraction through the BiLSTM layer, so as to get the probability of each word on each label, and finally use the CRF layer to constrain the different combinations, and finally get the optimal label sequence. The main role of the CRF layer is to add constraints between the predicted labels, so as to make the number of invalid predicted labels greatly reduced, and the task of named entity recognition achieve better results.



**Fig. 3.** The model structure of the BiLSTM-CRF

The transfer matrix is one of the parameters of the BiLSTM-CRF model, which serves to correct, i.e., set constraints on, the firing scores output by the BiLSTM layer, where the value of each term is called the transfer score. The initial values of the transfer matrix are randomly generated and continuously updated with subsequent training to finally obtain a best result.

After the BiLSTM and CRF layers, the scoring for the final sequence consists of two parts. Namely:

$$score(X, y) = \sum_{i=1}^n P_{i, y_i} + \sum_{i=1}^{n+1} A_{y_{i-1}, y_i}, \quad (2)$$

where  $X$  is the sequence being scored and before the plus sign is the BiLSTM firing score, where  $P_{i, y_i}$  denotes the firing score of the  $y_i$ th label in the  $i$ rd word vector. After the plus sign is the transfer score of the CRF layer, where  $A_{y_{i-1}, y_i}$  denotes the score of the  $y_{i-1}$ th label to the  $y_i$ th label. Thus, although the firing score of a label is high, if the transfer score is low, it will pull down the final score, making the probability of that label much lower.

After calculating the score for each possible sequence, it is normalized using SoftMax. Then:

$$p(y|X) = \frac{e^{score(X,y)}}{\sum_{\tilde{y} \in Y(x)} e^{score(X,\tilde{y})}}, \quad (3)$$

where  $Y(x)$  represents all possible labeling sequences.

Maximizing the likelihood probability is used during model training, here the log-likelihood is used and the result is:

$$\log(p(y|X)) = \log\left(\frac{e^{score(X,y)}}{\sum_{\tilde{y} \in Y(x)} e^{score(X,\tilde{y})}}\right) = score(X,y) - \log\left(\sum_{\tilde{y} \in Y(x)} e^{score(X,\tilde{y})}\right). \quad (4)$$

Finally, the optimal labeled sequence is derived using Viterbi decoding in the prediction stage, which results in:

$$y^* = \arg \max_{\tilde{y} \in Y(x)} score(X, \tilde{y}). \quad (5)$$

The BiLSTM-CRF model combined with SVM is used in the task of named entity recognition in criminal cases, which fully integrates a little bit of feature engineering and deep learning, and is able to effectively analyze the contextual information of the adjudication documents of the criminal cases, which greatly improves the accuracy of the task of named entity recognition in criminal cases. Through the named entity recognition, it helps the case handler to accurately grasp the development of the relevant evidence chain of the criminal case, and also provides accurate recognition results for accurately extracting the key information in the criminal case.

### 3.2. Intelligent Evidence Extraction Model for Criminal Cases

**3.2.1. Evidentiary framework for criminal cases.** The creation of a framework for evidence in criminal cases is essentially a categorization of events in the text of a criminal case. By dividing sentences describing different events into different stages and sorting them, i.e., a *EF* contains  $n$  stages, i.e.,  $EF = \{S_i | 0 \leq i < n\}$ , the stages in the framework are sorted according to  $S_i$  from smallest to largest.

Based on the above method a framework of evidence for a criminal case can be obtained. Therefore, the core of the algorithm can be divided into two stages, one is to determine the stages in the evidence framework of a criminal case, and the other is to train the classifier to classify the sentences in the text of a particular criminal case decision into different stages.

A good evidence framework for criminal cases should fulfill the following four characteristics:

(a) Temporal consistency

The times in events in the same phase remain continuous and the timelines between phases have clear demarcations. Let  $S_1$  and  $S_2$  denote the two stages in *EF*,  $ES_1$  denote the set of  $\alpha$  in  $S_1$ ,  $ES_2$  denote the set of  $\alpha$  in  $S_2$ , and  $T$  denote the time of the event, then:

If  $\exists \alpha_1 \in ES_1, \exists \alpha_2 \in ES_2$ , such that  $\alpha_1.E.T < \alpha_2.E.T$ , then  $\forall \alpha_i \in ES_1, \forall \alpha_j \in ES_2$ , both satisfy  $\alpha_i.E.T < \alpha_j.E.T$ .

(b) Completeness

The stages in the event framework are to contain the activities of any process. Let  $S_\alpha$  denote the stage that contains activity  $\alpha$ , and  $P_k \in EF_k$ , then they are satisfied:

If  $\alpha \in P_k$ , then  $S_\alpha \in EF_k, |\{S | S \in EF_k\}| \geq |\{S_\alpha | \alpha \in P_k\}|$ .



## (c) Brevity

The event framework should be as concise as possible while satisfying completeness, since noise is often present in process texts. Let  $PT$  denote the set of all events in a particular process text,  $P_k \in PT$ ;  $EN_k$  denote the set of noise,  $EN_k \in PT$ , then they satisfy:

$$\alpha_i \in EN \text{ if } \alpha_i \in PT \text{ and } \alpha_i \notin SEQ_i, |\{\alpha | \alpha \in PT\}| = |\{\alpha | \alpha \in P_K\}| + |\alpha | \alpha \in EN|$$

## (d) Minimum availability

When establishing an evidence framework for criminal cases, the reason for establishing the framework should be fully considered, and the one with the largest granularity should be chosen as long as it can be used.

After determining the stages of the criminal case evidence framework, the classifier needs to be trained to categorize the long sentences in the text of a particular criminal case into different stages.

Classification targets include stages  $S_t$ ,  $id$  and noise. Let  $Y$  denote the target label of the classifier, then  $Y$  is satisfied:

$$Y = \{S_t \cdot id | S_t \in EF\} \cup \{Noise\}, \quad (6)$$

$$|Y| = |\{S_t \cdot id | S_t \in EF\}| + 1. \quad (7)$$

This step is essentially a simple text categorization task, so in this paper, TextCNN is chosen for criminal case text-based niang and after full connectivity and SoftMax outputs the results, the output of which is a probability distribution over different labels. The loss function is calculated as follows:

$$loss = - \sum_i (yi \log (yi') + (1 - yi) \log (1 - yi')), \quad (8)$$

where  $y'$  is the output after SoftMax.

**3.2.2. Extraction of relationships between evidentiary entities.** Based on the recognition results of the named entities of criminal cases and the classification results of the evidence framework obtained in the previous paper, this paper proposes a Transformer inter-entity relationship extraction model based on multiple embeddings, which is applied to the extraction of entity information of criminal cases, to further enhance the ability of the case handlers in the collection of evidence of criminal cases [1].

Figure 4 shows the structure of Transformer inter-entity relationship extraction model based on multiple embedding. It mainly contains four parts, i.e., the input part is the multiple embedding of character sequences and word sequences, and the encoding part uses the Transformer module to encode the output vectors of the input part. The decoding part splices the output vectors obtained in the encoding part with the embedding vectors of the word sequences and passes them to the Transformer module for decoding, and an improved attention mechanism is added after the Transformer module to fully utilize the global information of the feature vectors, and the output part uses CRF to obtain the optimal output sequences and complete the extraction of entities of evidence of criminal cases.

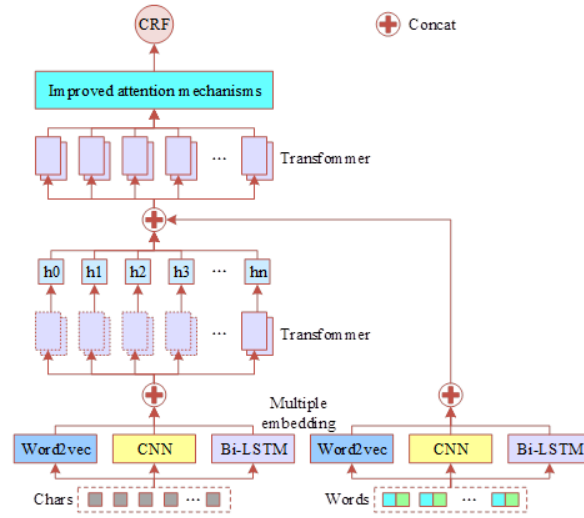


Fig. 4. Interphysical relationship extraction model structure

(a) Sequence feature fusion module based on multiple embedding

For the input character sequence  $X_i = [x_1, x_2, x_3, x_4, \dots, x_n]$ , after Word2vec embedding into  $h_i$  vectors, the embedding is for the basic embedding representation of the input sequence, many features of the input sequence can not be obtained, so the input sequence will be passed into the convolutional neural network for capturing local features. The specific formula is as follows:

$$hc_i = Dropout(X_i), \quad (9)$$

$$hc_i = CNN(hc_i), \quad (10)$$

$$hc_i = Relu(hc_i), \quad (11)$$

$$hc_i = Pool(hc_i), \quad (12)$$

$$hc_i = Linear(hc_i), \quad (13)$$

$$hc_i = Dropout(hc_i). \quad (14)$$

Convolutional neural networks can capture features at different scales of the input sequence due to the presence of convolutional kernels, i.e., they can capture the local information of the input sequence. In addition, since neither Word2vec nor convolutional neural networks take into account the inflectional order information of the input sequence, the input sequence is passed into the recurrent neural network. Then:

$$hw_1 = Dropout(X_i), \quad (15)$$

$$hw_i = BiLSTM(hw_i), \quad (16)$$

$$hw_i = Relh(hw_i), \quad (17)$$

$$hw_i = Pool(hw_i), \quad (18)$$

$$hw_i = Dropout(hw_i). \quad (19)$$

Bi-LSTM network, due to the bidirectionality of its structure, for the input character sequences and word sequences, can fully take into account their inflectional information, and the utilization of the information of the preceding and following texts is more adequate. The final input vectors are spliced with the embedding vectors of Word2vec, convolutional neural network and recurrent neural network, so that the input vectors can be obtained by fully integrating the multiple features of the input sequences, and the specific representation is as follows:

$$h_i = \text{Concat}(h_i, hc_i, hw_i). \quad (20)$$

(b) Global feature capture module based on attention mechanism

The attention mechanism is introduced in the Transformer module to give different weights to different sentences depending on their relevance to the relation, to reduce the influence of noise data and to improve the performance of the relation extraction model. The computational process of the sentence attention mechanism layer is as follows:

Calculate the attention score [15]. Assuming that the number of sentences in the sentence set bag of co-entity pairs is  $k$ , the match  $e_i$  between sentence feature vector  $s_i$  and relation label  $r$  is computed as:

$$e_i = s_i Ar. \quad (21)$$

The attention score is then calculated as:

$$\alpha_i = \frac{\exp(e_i)}{\sum_k \exp(e_k)}. \quad (22)$$

Calculate the feature vector of the entity pair  $g$ . The feature vector  $g$  of the entity pair is the vector that incorporates all the sentence features in the bag of the entity pair, which is essentially the weighted sum of all the sentence feature vectors. Assuming that the weights are each sentence's Attention Mechanism Score  $\alpha_i$ , the feature vector  $g$  for that entity pair is computed as:

$$g = \sum_i \alpha_i s_i. \quad (23)$$

In the process of calculating the attention score, it can be found that the calculation of the attention score relies on the relation label vector, when the Transformer model is trained, the relation labels corresponding to the bag are used, but when the Transformer model is tested, the bag has no relation labels. For this problem, the solution of Transformer model is to traverse all the relations, and then get the probability distribution of a certain relation for the relation, and select the maximum value as the probability of the relation.

#### 4. Application of Digital Information Technology in the Chain of Evidence in Criminal Cases

Evidence review judgment and determination in criminal justice activities in the core position, the relationship between the direction of the case, the focus on evidence review in line with the direction of trial-centered reform of the criminal procedure system. The application of digital information technology in the criminal field can help the case handler to accurately characterize the case, making

the direction of the evidence and the means of investigation more targeted to enhance the quality of the case. This chapter mainly focuses on the effectiveness of the previously proposed model of naming entity identification and relationship extraction of evidence in criminal cases, aiming to provide reference for case handlers to provide more diversified ways of evidence chain collection.

#### *4.1. Criminal Case Data Set Construction*

**4.1.1. Sources of data on criminal cases.** Because there is a lot of personal privacy information in criminal cases, and there are very limited publicly available datasets of Chinese adjudication documents for research, this paper uses the publicly available dataset of the "China Legal Research Cup" (CAIL2018). This dataset includes about 3 million criminal legal documents, involving more than 180 legal articles and 8 major categories of crimes, which is the first public, large-scale criminal case dataset in China. Due to the limitation of resources, this paper selects with 100,000 data from it to produce the CAIL-A experimental dataset, and divides it into training set and test set according to the ratio of 7:3 to ensure the accuracy and reliability of the model.

**4.1.2. Pre-processing of data on criminal cases.** As the CAIL-A experimental dataset contains a large number of criminal cases, its information is more cluttered contains a variety of noise, to this end, this paper preprocesses the dataset through the following steps.

- (a) Data clarity. Before using Jieba for word splitting, some irrelevant symbols are removed by regular expressions, including spaces and special symbols, year numbers and Arabic numerals, as well as foreign and garbled characters.
- (b) Chinese Segmentation. In this paper, the Jieba Segmentation Module that comes with Python is utilized to slice the text data of criminal cases and remove the words with a length lower than 2.
- (c) Construct dictionary. The corresponding word frequency is counted and a unique ID is assigned to each lexical element, and then the lexical element sequences of case descriptions and legal articles are converted into index sequences.
- (d) Label vectorization. In this paper, one-hot coding is used to transform non-numerical features by 0-1, as a way to enhance the correctness of entity identification and relationship extraction of evidence in criminal cases.

#### *4.2. Validation of physical identification of evidence in criminal cases*

**4.2.1. Comparative analysis of model energy efficiency.** In order to verify the effectiveness of naming recognition of evidence entities in criminal cases based on SVM-BiLSTM-CRF model, this paper carries out the validation analysis on the homemade CAIL-A dataset of criminal cases. The F1 value is chosen as the evaluation index, and with the same batch size and epoch settings and the same pre-trained word vectors, the SVM-BiLSTM-CRF model in this paper is compared with the two single-feature-layer models, LSTM-CRF and IDCNN-CRF, in terms of the energy efficiency based on the parameters recorded during the model training process. The F1 values are obtained in continuous iterations connected to form a change curve, and the change curves of the F1 values of the three models with the increase of the number of trainings are shown in Figure 5.

Starting from the first iteration, the SVM-BiLSTM-CRF model proposed in this paper achieves the best F1 value of 94.19% after 40 epoch iterations, while the two constructs based on IDCNN

and LSTM obtain the best F1 value after 72 and 90 epoch iterations, respectively, with the specific values of 92.49% and 91.52%, respectively. It can be seen that the F1 value of the SVM-BiLSTM-CRF model is 1.7% and 2.67% higher than the other two models, which indicates that the model designed in this paper has a better accuracy and convergence speed in performing entity recognition of the evidence feature information of criminal cases. This is due to the fact that this paper first uses SVM model to classify the input vector features of the criminal case, and then combined with the BiLSTM model can realize the effective extraction of the relevant feature information of the criminal case, and then in the CRF layer to get the optimal sequence of evidence of the criminal case.

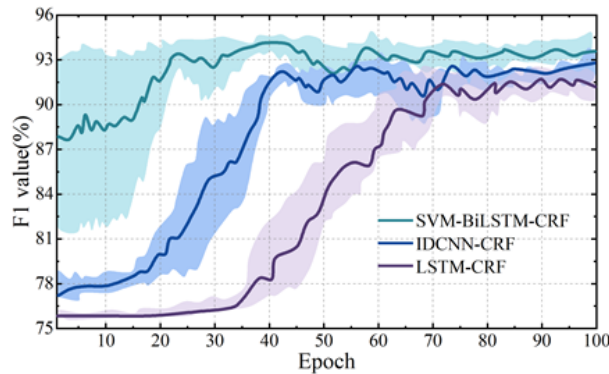


Fig. 5. F1 values follow the training chart

**4.2.2. Comparative analysis of entity identification.** Comparing the experimental result data in Table 1, compared with the one-way LSTM-CRF model, the F1 value of the SVM-BiLSTM-CRF model improves the recognition results on entities such as the name of the criminal, the location of the crime, the time of the crime, and the name of the relevant arresting authority by 6.68%, 10.88%, 11.10%, and 7.01%, respectively, and compared with the IDCNN-CRF model. Compared with the IDCNN-CRF model, the entity recognition accuracy of this paper's model increases by 4.88%, 5.36%, 1.79%, and 2.48%, respectively. The experimental results show that the SVM-BiLSTM-CRF model is more effective than the LSTM-CRF model and IDCNN-CRF model in recognizing the key entity information in the evidence chain of criminal cases, and it has a certain improvement in recognizing various types of named entities. Overall, it has a better effect on the recognition rate of named entities in criminal case adjudication documents, and confirms the superiority of this model in the task of recognizing named entities in criminal case adjudication documents, which can provide reliable recognition results for the case handlers to analyze different types of criminal cases and find out the relevant features of evidence entities from them. Moreover, under the same neural network training conditions, the SVM-BiLSTM-CRF model by combining with SVM makes the model in the training process, the loss value is reduced more, the convergence speed is faster, which effectively reduces the computational amount of the neural network and the model training time.

### 4.3. Verification of evidence information extraction in criminal cases

**4.3.1. Effectiveness of evidence information extraction.** The criminal case evidence information extraction model based on the criminal case evidence framework, combined with multiple embedded sequential feature fusion and global feature capture modules, is designed to obtain more case-related evidence information based on the named entity features of a criminal case after acquiring its features. In order to verify the effectiveness of the model, this paper additionally selects

Model	Type	ACC (%)	RE (%)	F1 (%)
LSTM-CRF	PER	83.72	81.08	83.18
	LOC	87.69	82.74	82.32
	TIME	88.63	80.67	81.43
	ORG	85.36	82.89	86.54
IDCNN-CRF	PER	87.48	84.92	85.25
	LOC	86.21	82.83	87.26
	TIME	90.42	83.91	86.19
	ORG	89.54	85.35	85.23
SVM-BiLSTM-CRF	PER	91.75	86.53	88.74
	LOC	90.83	85.48	91.28
	TIME	92.04	87.97	90.47
	ORG	91.76	88.46	92.61

**Table 1.** Entity identification results of various models

CJRE data as experimental data on the basis of CAIL-A dataset. CasRel, TPLinker, GPLinker, and READK are selected as the baseline models, and the extraction effect of the model’s criminal case evidence information is comparatively evaluated with CJRE and CAIL-A datasets in the same experimental environment. Table 2 Criminal case evidence information extraction results of different models.

By analyzing the results of the above comparison experiments, the following conclusions can be drawn:

- (a) The criminal case evidence information extraction model proposed in this paper is significantly better than the CasRel model, and the F1 value is improved by 2.94% and 11.23% on the two datasets respectively, due to the fact that CasRel’s relational ternary extraction is to recognize the head entity first, and then the head entity is used as an input, and another network is used to extract the tail entities and relations. If there is a deviation in the identification of the head entity, it will affect the results of the subsequent tail entities and relations, i.e., there may be a certain error propagation problem. In contrast, the criminal case evidence information extraction model proposed in this paper is a single-stage entity-relationship joint extraction model based on the criminal case evidence framework and named entity recognition results, which solves the cascading error problem of CasRel model and performs better.
- (b) Compared with other baseline models, the criminal case evidence information extraction model proposed in this paper achieves the best experimental results, with its F1 values of 62.06% and 81.83% on the two datasets. This indicates that the introduction of criminal case evidence framework and attention global feature capture mechanism in the relational extraction model can effectively improve the performance of criminal case evidence information extraction model.

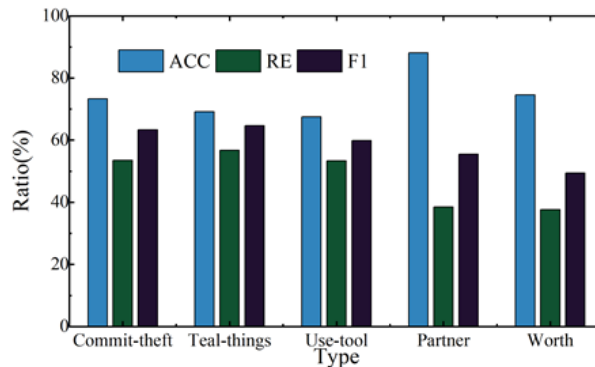
In order to further analyze the predictions of the model for extracting information from evidence in criminal cases, the model was used to extract information from various types of relationships in the CJRE dataset. There are five main types of relationships defined in the CJRE dataset, i.e., Commit-theft, Teal-things, Use-tool, Partner, and Worth. Figure 6 shows the effect of the model on the extraction of various types of relationships in the CJRE dataset.

As can be seen from the figure, the model’s prediction effect for the partnership relationship (Partner) and the value relationship (Worth) is biased, with F1 values of 55.45% and 49.38%, respectively,

Model	CJRE			CAIL-A		
	ACC (%)	RE (%)	F1 (%)	ACC (%)	RE (%)	F1 (%)
CasRel	60.24	59.42	60.29	79.37	74.64	73.57
TPLinker	60.79	58.35	60.45	80.18	76.35	78.32
GPLinker	61.48	58.61	60.78	78.45	78.28	79.06
READK	61.51	57.59	60.64	79.26	80.59	80.18
Ours	66.73	60.28	62.06	82.63	83.17	81.83

**Table 2.** Information extraction results for criminal cases

mainly due to the small amount of data, the model fails to adequately learn the distribution of the location of this type of entity, resulting in the frequent omission of these two relationships in the prediction results of a given sample. Overall, this paper based on the criminal case evidence framework and named entity identification results to establish the criminal case evidence entity information relationship extraction model has a good criminal case information extraction ability, can assist the case officer from the massive criminal case data to obtain the relevant entity category information needed to fill the criminal case evidence chain.



**Fig. 6.** CJRE relationship extraction

**4.3.2. Ablation experiment P-R curves.** Sequential feature module with multiple embedding is used to obtain the entity information of criminal cases, and global feature capture with attention mechanism is used to realize the extraction of entity information of criminal case evidence. For the effectiveness of the above different modules, this paper sets up multiple sets of experimental comparisons to prove that both sequence feature and global feature incorporation are helpful for entity information extraction effect enhancement. The details are as follows:

Experiment A takes the BiLSTM model as the baseline model, Experiment B adds the criminal case evidence framework on the basis of A, Experiment C adds the sequence feature module with multiple embeddings on the basis of B, and Experiment D adds the attention global feature capture module on the basis of C. Comparison experiments are conducted for the above four models, and Figure 7 shows the P-R curves of the four sets of experiments.

From the figure, it can be seen that when the recall is small ( $<0.05$  or so), the four groups of experiments perform about the same in terms of precision. However, as the recall rate increases, the model that has both the sequence feature fusion module with multiple embeddings and the attention global feature capture mechanism performs significantly better than the other three models. The

enhancement of the criminal case evidence framework for the model is visible to the naked eye when the recall rate is low, but as the recall rate increases, the enhancement starts to become less obvious. Comparing Experiment B and Experiment C, it is found that the external information provided by the multiple-embedded sequence feature fusion module improves the model's relational classification performance better than the evidential information provided by the criminal case evidence framework when the recall rate is large. The mean P@N obtained in experiment D was 0.759, which was 14.65% higher than that of experiment A. This further illustrates the effectiveness of the Evidence Framework for Criminal Cases, Sequential Feature Fusion with Multiple Embeddings, and the Global Feature Capture Module Optimized for Attention Mechanisms in enhancing the ability to extract information about the entities of evidence in criminal cases.

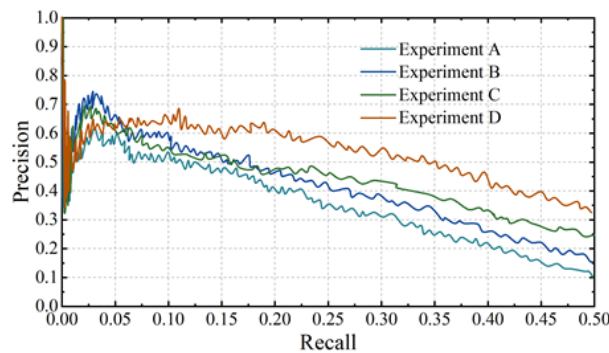


Fig. 7. Ablation experiment P-R curve

**4.3.3. Effectiveness of information extraction applications.** The criminal case data based on the criminal case named entity identification and evidence entity information extraction model can realize the depth construction of the evidence chain of the criminal case, which is an important goal of the construction and improvement of the evidence chain of the criminal case empowered by digital information technology. And give the extraction results to establish the criminal case evidence standards and rules, in order to help case officers in the massive data to obtain the evidence information that meets the needs.

The extracted entity information of criminal cases is applied to a city police cloud platform, and the corresponding evidence rules are formulated, taking intentional homicide as an example, and the evidence classification of intentional homicide is carried out through the evidence rules, so as to obtain the evidence classification of intentional homicide cases concluded by the city in 2020~2023, such as shown in Figure 8. Among them, the entity information represented by Y1~Y9 are witness testimony, physical evidence, appraisal opinions, investigation transcripts, defendant's confession, identification transcripts, audiovisual materials, electronic data, and victim's statement, respectively.

With the digital information technology development of criminal case entity information extraction, can be for different types of evidence in criminal cases for effective extraction, to help case officers better clarify the intentional homicide case involves nine different types of evidence information, in order to form a complete chain of evidence in criminal cases. On the whole, the chain of evidence in intentional homicide cases can be more than 90% complete for witness testimony (Y1), investigation transcripts (Y4) and defendants' confessions (Y5). This is due to the fact that all of the above evidence can be transformed into the form of documents, so that the information extraction model designed in this paper can be used to better extract the relevant information therein, providing reliable evidence support for the judge to decide the intentional homicide case. In the improvement



of litigation efficiency, it provides clear and explicit guidelines for evidence collection and review for judicial officers in handling different types of cases, and reduces the differences in case handling due to the differences in the individual case handling experience and subjective judgment of judicial officers, resulting in greater differences in case handling. The awareness of the high rules of evidence standards and strict procedural operation and management allows judicial officers to continuously strengthen their awareness of the standardization of criminal case handling in the process of handling cases, reduce the incidence of missing and defective evidence, improve the efficiency of criminal case handling, and enhance the credibility of the judiciary.

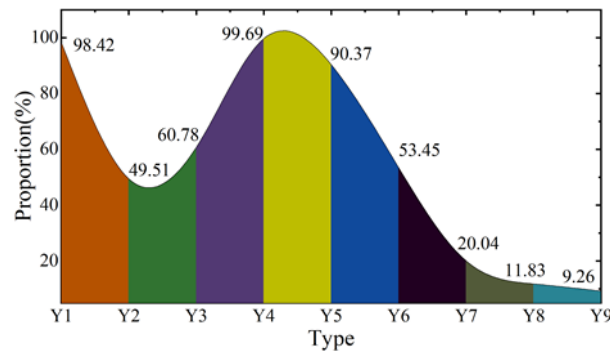


Fig. 8. The certificate of intentional homicide is classified

## 5. Conclusion

The article proposes a named entity recognition model for criminal cases based on SVM-BiLSTM-CRF, and establishes a criminal case evidence entity information extraction model based on Transformer model in combination with the criminal case evidence framework. The constructed named entity recognition model for criminal cases achieves the best F1 value of 94.19% after 40 epoch iterations, which is 1.7% and 2.67% higher than the IDCNN-CRF and LSTM-CRF models, respectively. The F1 values of the criminal case evidence extraction model on the CJFE and CAIL-A datasets are 62.06% and 81.83%, respectively, and the average P@N value can reach 0.759, which is 14.65% higher than that of the BiLSTM model. With the results of extracting information on the entity of evidence in criminal cases to establish the standards and rules of evidence in criminal cases, taking intentional homicide cases as an example, the construction of part of the chain of evidence in them can realize the completion rate of more than 90% of criminal cases. As a result, the use of digital information technology can assist case managers in mining the evidence information in the data of such cases, promote case managers to solve criminal cases faster, and contribute to the enhancement of judicial credibility.

**Conflict of interest:** The authors declare that they have no conflicts of interest.

## References

- [1] A novel multitask transformer deep learning architecture for joint classification and segmentation of horticulture plantations using very high-resolution satellite imagery. *Computers and Electronics in Agriculture*, 227:109540, 2024. <https://doi.org/10.1016/j.compag.2024.109540>.

- [2] Z. A. Baig, P. Szweczyk, C. Valli, P. Rabadia, P. Hannay, M. Chernyshev, M. Johnstone, P. Kerai, A. Ibrahim, K. Sansurooah, et al. Future challenges for smart cities: cyber-security and digital forensics. *Digital Investigation*, 22:3–13, 2017. <https://doi.org/10.1016/j.diin.2017.06.015>.
- [3] A. M. Balogun and T. Zuva. Criminal profiling in digital forensics: assumptions, challenges and probable solution:1–7, 2018. <https://doi.org/10.1109/ICONIC.2018.8601268>.
- [4] M. A. Biasiotti. A proposed electronic evidence exchange across the european union. *Digital Evidence and Electronic Signature Law Review*, 14:1, 2017.
- [5] L. Caviglione, S. Wendzel, and W. Mazurczyk. The future of digital forensics: challenges and the road ahead. *IEEE Security & Privacy*, 15(6):12–17, 2017. <https://doi.org/10.1109/MSP.2017.4251117>.
- [6] H. B. P. de Gallo, O. C. Medina, J. G. Dorado, J. A. Fleming, C. Párraga, E. R. Notario, and E. Sanchez. The application of blockchain to the chain of custody of judicial evidence. *Eureka*, 2023.
- [7] J. Duan, Z. Tan, M. Zhang, and H. Wang. New word detection using bilstm+ crf model with features. *IEICE TRANSACTIONS on Information and Systems*, 103(10):2228–2236, 2020. <https://doi.org/10.1587/transinf.2019EDP7330>.
- [8] D. Freed, J. Palmer, D. E. Minchala, K. Levy, T. Ristenpart, and N. Dell. Digital technologies and intimate partner violence: a qualitative analysis with multiple stakeholders. *Proceedings of the ACM on human-computer interaction*, 1(CSCW):1–22, 2017. <https://doi.org/10.1145/3134681>.
- [9] X. Gao, W. Bai, Q. Dang, S. Yang, and G. Zhang. Learnable self-supervised support vector machine based individual selection strategy for multimodal multi-objective optimization. *Information Sciences*, 690:121553, 2025. <https://doi.org/10.1016/j.ins.2024.121553>.
- [10] B. Karimov. Scientific and theoretical questions of the category of digital evidence. *Review of law sciences*, 2020.
- [11] A. Keane and P. McKeown. The modern law of evidence, 2022.
- [12] L. Klasén, N. Fock, and R. Forchheimer. The invisible evidence: digital forensics as key to solving crimes in the digital age. *Forensic science international*, 362:112133, 2024. <https://doi.org/10.1016/j.forsciint.2024.112133>.
- [13] Q. Le, O. Boydell, B. Mac Namee, and M. Scanlon. Deep learning at the shallow end: malware classification for non-domain experts. *Digital Investigation*, 26:S118–S126, 2018. <https://doi.org/10.1016/j.diin.2018.04.024>.
- [14] A. H. Lone and R. N. Mir. Forensic-chain: blockchain based digital forensics chain of custody with poc in hyperledger composer. *Digital investigation*, 28:44–55, 2019. <https://doi.org/10.1016/j.diin.2019.01.002>.
- [15] Z. Lou, S. Luo, X. Huang, and D. Wei. Conv-attention: a low computation attention calculation method for swin transformer. *Neural Processing Letters*, 56(2):71, 2024. <https://doi.org/10.1007/s11063-024-11483-6>.
- [16] F. E. Salamh, U. Karabiyik, M. K. Rogers, and E. T. Matson. A comparative uav forensic analysis: static and live digital evidence traceability challenges. *Drones*, 5(2):42, 2021. <https://doi.org/10.3390/drones5020042>.
- [17] M. S. M. B. Shah, S. Saleem, and R. Zulqarnain. Protecting digital evidence integrity and preserving chain of custody. *Journal of Digital Forensics, Security and Law*, 12(2):12, 2017. <https://doi.org/10.15394/jdfs1.2017.1478>.

- 
- [18] Z. Tian, M. Li, M. Qiu, Y. Sun, and S. Su. Block-def: a secure digital evidence framework using blockchain. *Information Sciences*, 491:151–165, 2019. <https://doi.org/10.1016/j.ins.2019.04.011>.
- [19] M. J. Vermeer, D. Woods, and B. A. Jackson. Identifying law enforcement needs for access to digital evidence in remote data centers, 2018.
- [20] I. Yaqoob, I. A. T. Hashem, A. Ahmed, S. A. Kazmi, and C. S. Hong. Internet of things forensics: recent advances, taxonomy, requirements, and open challenges. *Future Generation Computer Systems*, 92:265–275, 2019. <https://doi.org/10.1016/j.future.2018.09.058>.