Copy Detection Pattern-Based Authentication for Printed Documents with Multi-Dimensional Features

Pei Zhang, Weiming Zhang, Nenghai Yu
CAS Key Laboratory of Electro-magnetic Space Information
University of Science and Technology of China
Hefei, China
e-mail: zhangwm@ustc.edu.cn

***Abstract*—A copy detection pattern (CDP) is an effective method to distinguish counterfeits of printed documents. CDP is a digital image filled with pixels of random grey levels, which is embedded into digital documents and printed as the legitimate ones. A general counterfeit method is to scan the legitimate documents and reprint them. The printed documents will be scanned to proceed the authentication. CDP will undergo distortion during the print-scan operation and we will detect counterfeits by measuring the extent of the distortion. In this paper, we propose a novel CDP-based authentication method for printed documents. Firstly, we analyze some attack methods which make the existing CDP-based detection methods ineffective. Secondly, we propose new metrics for measuring the distortion of CDP and construct a classifier based on support vector domain description (SVDD). Experimental results show that the proposed authentication method significantly outperforms previous methods.**

Keywords-copy detection pattern; printed documents; counterfeits authentication

I. INTRODUCTION

It is admitted that printed documents play an important role in many aspects of social life. These printed documents include ID cards, passports, driving licenses, diplomas, contracts commodity packaging and so on. With printing and scanning devices developing greatly, making fake printed documents has become quite simple. For fakers, a straightforward counterfeiting method is scanning legal printed documents and re-printing them. Of course, fakers may use some image processing methods to tamper the scanned documents so as to make the re-printed documents look like the legal ones. To distinguish legal documents from fake documents, many anti-counterfeit methods have been proposed such as special printed materials (such as ink) [1], watermarks [2] and holograms [3]. However, these methods need corresponding devices to produce legal documents and proceed the authentication so that they cost high.

To seek for an effective and low-cost method to distinguish legal documents from fake documents, Yu et al. [4] acquire some facts about the print-scan operation. From the view of information theory, they regard the print-scan operation as a lossy channel, then information loss must happen in a printed and scanned document and the amount of information loss is positively correlated with times of the print-scan operation. In other words, the more times of printing and scanning a digital document undergo, the more information loss will happen. From the view of the printed document, there are more quality distortion appearing in a printed and scanned document that undergo several times of printing and scanning. Based on the theories mentioned above, there are methods proposed, such as 2D code based method [5], [6], digital forensic based method [7], [8].

The straightforward counterfeiting method is as follows. The legal digital documents possessed by the authorities are printed, the legal printed documents are produced, they will be scanned for the next authentication, we can call them the 1st print-scan documents (1st PSD). However, fakers will scan the legal printed documents and reprint them to produce fake ones. These fake printed documents will be also scanned for the same authentication. It is worth noting that these fake printed documents have already undergone the print-scan operation twice, so we can call them the 2nd print-scan documents (2nd PSD). Note that there exist differences between the 1st PSD and 2nd PSD. What we need to do is measuring the amount of information loss (i.e. the extent of quality distortion) and enabling discrimination of the legal printed documents and fake printed documents.

In fact, all the contents (texts or images) included in the printed document can serve as the indicator of the information loss resulting from the print-scan operation. However, we could intentionally design a special pattern with the highest sensitivity to the information loss. Copy detection patterns (CDP) proposed by Justin Picard [9] is the very solution based on the idea mentioned above. What is called CDP is a special pattern filled with pixels of random grey levels. The small size and arbitrary shape make it be easily embedded in digital documents and be printed with the documents together (see Fig. 1). The printed document is scanned and the amount of information loss, which is contained in the pixels of CDP, is measured by designing some features. Finally, according to certain criterions, a decision on the authenticity of the printed document is made.

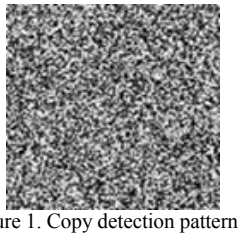

Figure 1. Copy detection pattern (CDP)

978-1-7281-0426-3/19/$31.00 ©2019 IEEE

Authentication methods based on CDP is an effective and low-cost method to distinguish legal documents from fake documents. What is more, CDP-based authentication method can be combined with other fake documents authentication methods (special printed materials, watermarks, holograms and so on) to improve authentication precision. Nevertheless, fakers can scan the legal printed document with a high-resolution scanner and use some image processing technology to restore information loss contained in CDP before reprinting it so that a perfect duplicate may be recognized as legal by the authentication method. The image processing technology can be viewed as the attack methods to authentication methods based on CDP. This is the very drawback for authentication methods based on CDP, however, previous research works do not consider it.

Thus, to build a complete printed documents authentication system based on CDP, we have two contributions in this paper. Firstly, we try a series of attack methods based on restoration of digital images, and experiment results show that some attack methods can make state of the art CDP-based authentication methods perform badly. Secondly, to resist these attacks, we improve the accuracy and robustness of the CDP method by developing some new features by which a classifier is trained to identify the information loss contained in a scanned CDP.

The paper is organized as follows. Details of CDP scheme as well as previous research works about CDP are introduced in Section II. Section III discusses some effective attack methods that increases error rates of current authentication methods based on CDP. In Section IV, we propose new features used to measure information loss contained in CDP. The details about the classifier are also discussed there. In Section V, experiment results show the availability of attack methods, new features and the classifier. Finally, we conclude the paper in Section VI.

II. Details of CDP Scheme and Related Works

This section is split into three sub-sections. In section A, we describe specific properties of CDP and discuss why CDP is suitable for detecting copies of the printed documents. We present the processes of the whole document authentication system and give a block diagram in section B. In section C, we focus on various metrics of information loss contained in CDP and review related research works.

A. Properties of CDP

Firstly, we must admit a basic principle called 'information loss principle': every time a digital document is printed and scanned, some information is lost about it. The principle reflects that the print-scan operation is a lossy channel. Based on this principle, all contents of the digital document can undergo information loss after the print-scan operation so that they can play the same role as CDP, discerning between true or false documents according to the amount of information loss contained in pixels.

Why we select CDP as the indicator? General contents of digital documents (texts or images) are not very sensitive to the print-scan channel so that the differences between the digital document and the corresponding version of the print-scan operation are not obvious. CDP is a kind of pattern that is designed intentionally for monitoring the effects of the print-scan operation on image information loss. It is an arbitrary shape (in this paper, we use square) filled with pixels of random grey levels. There are many methods to generate a CDP, the most convenient one is to use a prefabricated pseudo random number generator with a secret key. The secret key can be related to the document itself and anyone possessing the secret key can recreate the same CDP. The property that pixels contained in CDP are non-predictable enables maximum entropy. Information loss resulting from the print-scan operation is irreversible and CDP can be viewed as a kind of physical unclonable function. So creating an identical copy of CDP is impossible for fakers. Selecting CDP to protect the security of the document is reasonable.

B. Processes of Document Authentication System Based on CDP

Now we describe the processes of document authentication system based on CDP as illustrated in Fig. 2. We have three blocks in the light of truth or false of the document. The first block is the generation process of the legal document containing CDP, the second block is the counterfeit process of the document containing CDP and the third block is the authentication process of the document containing CDP.

Generation process of the documents containing CDP (Fig.2, block 1). Possessors of legal documents hold legal digital documents as well as secret keys matched with these documents. They input a secret key to a pseudo random number generator and produce a CDP. Then they embed the CDP into a legal digital document and print it as a legal printed document.

Counterfeit process of the documents containing CDP (Fig.2, block 2). Fakers can obtain legal printed documents containing CDP, of course, they must not obtain legal digital documents or the secret key. Then, they scan printed documents, translating them into digital ones. Before reprinting them, they maybe use some digital image processes to retrieve information loss that documents undergo during the print-scan operation. In fact, except for CDP, the rest of the document is easy to copy, so the main purpose of image restoration is to restore CDP. Finally, fakers produce copies of legal printed documents and they hope the copies can pass the next authentication.

Authentication process of the documents containing CDP (Fig. 2, block 3). Authentication is the core of the whole document authentication system. Firstly, printed documents are scanned, translating the printed version of the document into the digital version. Then the feature extractor extracts features of the scanned document according to a prefabricated feature extraction algorithm. Finally, the classifier takes a decision on the basis of features extracted.

C. Various Metrics of Information Loss Contained in CDP

As mentioned in this section A, CDP is a kind of physical unclonable function and copying CDP to produce an identical one is an impossible task. However, it is not

necessary for fakers to create a perfect copy of CDP. As for fakers, they only need to create CDP that can deceive the existing printed documents authentication methods. Therefore, except for the secret key used to generate CDP, the security and effectiveness of the CDP-based authentication methods depends on what kind of metrics (or called features) of information loss contained in CDP can be used and the decision method (we will talk about the decision method in Section IV). A rational metric should suitably reflect the information loss contained in CDP after undergoing the print-scan operation and avoid the interference of factors that are not related to the print-scan operation. Many metrics of information loss contained in CDP are proposed in related works, we will review some of them in the rest of the section.

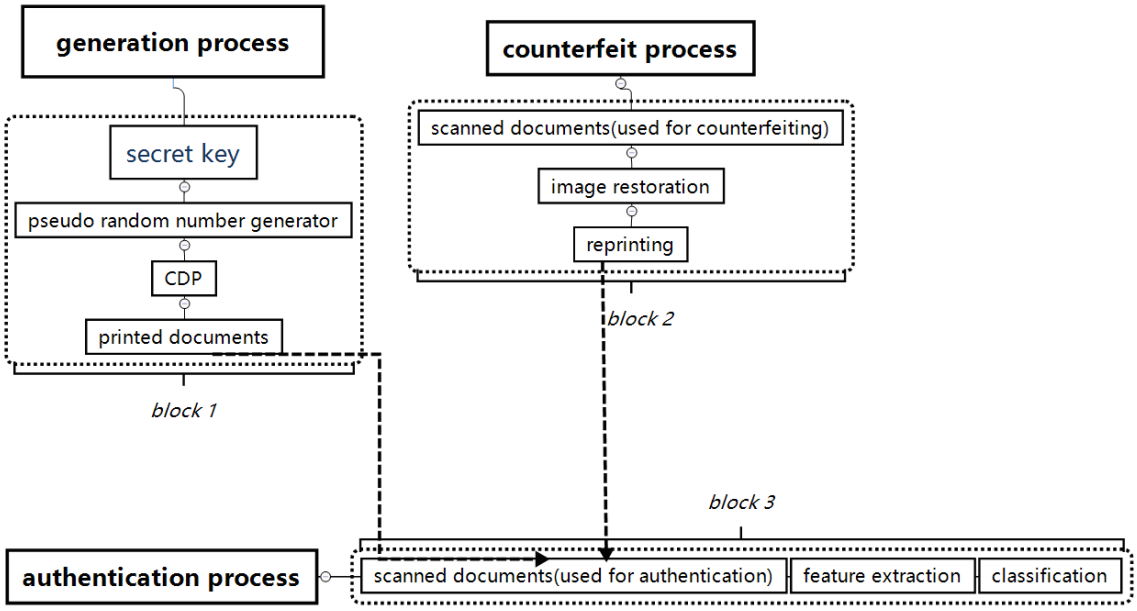

Figure 2. Processes of document authentication system based on CDP

In [10], there are four state of the art metrics proposed. The first one is called entropy metric. The amount of information is generally measured by entropy of a signal in information theory. Let x be a vector representing intensity values of pixels in a scanned CDP and n is the length of the vector x. The entropy of CDP is defined as follows:

$$F_{entropy}(CDP_x) = -\sum_{i=1}^{n} p(x_i)\log_2 p(x_i) \tag{1}$$

where, $p(x_i)=P_r(X=x_i)$ is the probability mass function of x.

The second one is called Fourier domain sharpness metric. Let $x_{i,j}$ represent the pixel values of a scanned CDP. M and N are the dimensions of the scanned CDP. The 2-dimensional Fourier transform formula is defined as follows:

$$X_{u,v} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} x_{m,n} e^{-i(2\pi mu/M)} e^{-i(2\pi nv/N)} \tag{2}$$

And the Fourier domain sharpness metric is defined as follows:

$$S_0(w) = \sum_{u=0+w}^{M-1-w}\sum_{v=0+w}^{N-1-w} |X_{u,v}| \tag{3}$$

$$S_1 = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} |X_{u,v}| \tag{4}$$

$$F_{fourier_w}(CDP_x) = \frac{S_0(w)}{S_1} \tag{5}$$

where, w is set as 7pi/8 empirically.

The third one is called wavelet domain sharpness metric. Firstly, the scanned CDP is decomposed by two-level wavelet transform. Then, the standard deviation of the second-level HH sub-band coefficient is selected as the metric. Note that the type of wavelet is 'Daubechies8' wavelet in the wavelet transform.

The fourth one is called prediction error metric. The metric is used on the basis of the idea that the intensity values of CDP pixels can be predicted from their neighbours. Let $x_{i,j}$ represent the pixel values of a scanned CDP. The prediction of $x_{i,j}$ is defined as follows:

$$x'_{i,j} = (x_{i,j-1} + x_{j-1,i}) - x_{i-1,j-1} \tag{6}$$

And the prediction error metric is defined as follows:

$$F_{prediction}(CDP_x) = \frac{1}{(M-1)(N-1)}\sum_{i=1}^{M-1}\sum_{j=1}^{N-1} |x_{i,j} - x'_{i,j}| \tag{7}$$

As mentioned above, these metrics can partly reflect the amount of information loss that CDP has undergone during the print-scan operation. We will show that there exist some attack methods that can interfere these metrics in next section.

III. Evaluation of CDP Metrics

It is impossible to perfectly copy a CDP because of the intense sensibility of CDP to the print-scan operation. From the view of fakers, they have no need of duplicating a CDP to the utmost. In fact, they only improve the quality of the scanned CDP to some extent so that the reprinted CDP can pass the next authentication. In other words, what they require to do is attacking the existing printed documents authentication methods with the help of the scanned CDP that obtains restoration and enhancement. We propose some attack methods to evaluate the performance of existing CDP authentication methods from the view of the fakers in this section.

A. Effects of Print-Scan Operation on CDP

What are called attack methods are equally restoration methods of the scanned CDP. Although the print-scan operation has high complexity and CDP is physical unclonable function, it is meaningful to try some image processing methods to offset certain amount of information loss because of incompleteness of existing authentication methods. To start with, we focus on some facts about the print-scan operation [4]. There are various distortions in the print-scan operation, noting that most of them are nonlinear. For simplicity, we divide them into three primary parts. The first one is called digital halftone. It converts a digital image into a corresponding binary image used for printing, leading to the appearance of the quantization noise. Digital halftone happens in printing process. The second one is called geometric transformations. Geometric transformations include cropping, rotation, scaling and so on, happening in both printing and scanning process. The third one is called blurring, resulting from the ink effect of printer and the optical imaging of scanner. It makes the scanned image look like less clear than the original digital image from the view of human.

B. Some Effective Attack Methods

Based on distortions that CDP has undergone during the print-scan operation, we find five restoration methods. The first one is called Wiener filtering [11]. Wiener filtering is a usual linear image restoration method. Firstly, Linear degradation model of the scanned CDP can be defined as follows:

$$g(x,y) = h(x,y) * f(x,y) + n(x,y) \tag{8}$$

where, $g(x,y)$ is the scanned CDP, $f(x,y)$ is the original digital CDP, $h(x,y)$ is called the degradation function (also called point spread function), $n(x,y)$ is the noise. Then, a statistical error function is defined as follows:

$$e^2 = E\left\{(f(x,y) - f'(x,y))^2\right\} \tag{9}$$

where, $f'(x,y)$ is the estimate of the original digital CDP, E is the expected value operator. Wiener filter find a $f'(x,y)$ that minimizes the function. In the frequency domain, we can easily find the solution to the problem as follows:

$$F'(u,v) = [\frac{1}{H(u,v)} \frac{|H(u.v)|^2}{|H(u,v)|^2 + K}] G(u,v) \tag{10}$$

where $G(u,v)$, $F'(u,v)$ and $H(u,v)$ are the Fourier transform of the scanned CDP, the estimate of the original digital CDP and the degradation function respectively, K is a constant number.

The second one is called constrained least squares filtering [12]. It is also a linear restoration method. The constrained least squares filtering defines a criterion function C as follows:

$$C = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\nabla^2 f(x,y)]^2 \tag{11}$$

Subject to the constrain

$$\| g(x,y) - h(x,y) f'(x,y) \|^2 = \| n(x,y) \|^2 \tag{12}$$

where ∇^2 is the Laplace operator, $\|w\|^2 \triangleq w^T w$ is the Euclidean vector norm. We need to find the minimum value of C. The solution to the constrained optimization problem is expressed in the frequency domain as follows:

$$F'(u,v) = [\frac{H^*(u,v)}{|H(u,v)|^2 + \gamma |P(u,v)|^2}] G(u,v) \tag{13}$$

where we need to adjust the value of γ to satisfy the constraint. $H^*(u,v)$ is the complex conjugate of $H(u,v)$. $P(u,v)$ is the Fourier transformation of the Laplacian operation template $p(x,y)$ defined as follows:

$$p(x,y) = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix} \tag{14}$$

The third one is a kind of iterative nonlinear restoration method based on the Lucy-Richardson algorithm [13]. The Lucy-Richardson algorithm adopts the method of iteration to restore the original image. The iteration expression is as follows:

$$f'_{k+1}(x,y) = f'_k(x,y) [h(-x,-y) * \frac{g(x,y)}{h(x,y) * f'_k(x,y)}] \tag{15}$$

where $f'_k(x,y)$ is the k time estimate of the original digital CDP.

In [4], a sharpen filter, which is designed for the restoration of images that undergo quality degradation because of the print-scan operation, is proposed. The template T of the filter is defined as follows:

$$T = \begin{pmatrix} 0 & -\alpha & 0 \\ -\alpha & 1+4\alpha & -\alpha \\ 0 & -\alpha & 0 \end{pmatrix} \tag{16}$$

The parameter α of the above filter requires to be adjusted to get a proper result. The sharpen filter is selected

as the fourth one. Besides, the last one is the smart sharpen filter that is built in the software of Photoshop.

Note that it is necessary to estimate the point spread function in the first three methods of restoration. Based on the relative analysis of the print-scan operation discussed in this Section A, although the whole print-scan operation composes of certain kinds of nonlinear filtering, the Gaussian low-pass blurring is proved suitable for simulating the effects of the print-scan operation on CDP [4]. In experiments showed in Section V, we select the Gaussian low-pass blurring as the point spread function.

IV. New Metric and Classifier

In this section, we propose new metric (or call it new feature) to measure the information loss contained in CDP after the print-scan operation. Compared to those single dimensional metrics mentioned above, new metric consists of high dimensional features so that it can reflect the effects of the print-scan operation on CDP more precisely. According to features extracted, we need to construct a classifier. In fact, the type of classifier depends on property of the metric. The classifier based on single dimensional metric generally uses the method of threshold to distinguish the positive samples (legitimate ones) from the negative samples (counterfeits) and the threshold is usually decided empirically. However, in the case of obtaining high dimensional features that describe the same object, we can train a classifier that makes the utmost of multiple features to take the place of the method of threshold. We are certain that such classifier can perform more accurately.

A. New Metric

The key point of distinguishing fake CDPs from true CDPs depends on two aspects: the metric for measuring information loss that CDPs undergo during the print-scan operation and the classification method. Further, a proper metric will serve as the precondition of the classifier. Inspired by the steganalysis methods [14], [15], we propose a new metric for measuring the information loss contained in CDP after the print-scan operation. This metric can originally take effect on detecting hidden secret messages in images. To some extent, we can assume that the print-scan operation adds its 'signature' to CDPs so it is reasonable to amend this metric to our work. Besides, this metric is based on JPEG quantized block discrete cosine transformation (DCT) coefficients so the CDPs must be converted into the format of JPEG [16]. The construction of the new metric consists of the following three steps.

The first step is the generation of JPEG 2-D array and mode 2-D array. Firstly, apply 8×8 block discrete cosine transformation (DCT) to the scanned CDP and generate a 2-D array. Note that the DCT coefficients in the 2-D array have been quantized with a JPEG quantization table and taken absolute values. From this, the generated 2-D array is called JPEG 2-D array. As for the generation of mode 2-D array, we assemble the DCT coefficients at the same position from all the 8×8 blocks in JPEG 2-D array to generate 63 mode 2-D arrays. For example, we select the DCT coefficients at the first row and the second column from all the 8×8 blocks in JPEG 2-D array to generate an array called mode 2 array. Note that we discard the direct current component so we obtain 63 mode 2-D arrays (from mode 2 to mode 64).

The second step is the generation of difference JPEG 2-D arrays and difference mode 2-D arrays. We denote the JPEG 2-D array by $F(u,v)$ $(u \in [1, S_u],\ v \in [1, S_v])$where S_u and S_v are the size of the JPEG 2-D array in horizontal direction and vertical direction respectively. The difference JPEG 2-D arrays are defined as follows:

$$F_h(u,v) = F(u,v) - F(u+1,v) \tag{17}$$

$$F_v(u,v) = F(u,v) - F(u,v+1) \tag{18}$$

$$F_d(u,v) = F(u,v) - F(u+1,v+1) \tag{19}$$

$$F_{md}(u,v) = F(u+1,v) - F(u,v+1) \tag{20}$$

where $u \in [1, S_u - 1], v \in [1, S_v - 1]$ and $F_h(u,v),\ F_v(u,v), F_d(u,v), F_{md}(u,v)$ denote the difference arrays in the horizontal, vertical, main diagonal, and minor diagonal directions respectively.

Similarly, we denote the mode 2-D array by $M^z(s,t)(s \in [0, S_s - 2], t \in [0, S_t - 2], z \in [2,64])$, and note that $S_s = S_u/8,\ S_t = S_v/8$. Then the difference mode 2-D arrays are defined as follows:

$$M_h^{(z)}(s,t) = M^{(z)}(s,t) - M^{(z)}(s+1,t) \tag{21}$$

$$M_v^{(z)}(s,t) = M^{(z)}(s,t) - M^{(z)}(s,t+1) \tag{22}$$

$$M_d^{(z)}(s,t) = M^{(z)}(s,t) - M^{(z)}(s+1,t+1) \tag{23}$$

$$M_m^{(z)}(s,t) = M^{(z)}(s+1,t) - M^{(z)}(s,t+1) \tag{24}$$

where $M_h^z(s,t), M_v^z(s,t), M_d^z(s,t), M_m^z(s,t)$ denote the horizontal, vertical, main diagonal, and minor diagonal difference mode 2-D array respectively.

The third step is using Markov random process to model these difference arrays, generating transition probability matrixes (TPMs). TPMs of all four directions of difference JPEG 2-D arrays are defined as follows:

$$p\{F(u+1,v) = n \mid F(u,v) = m\} = \frac{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1} \delta(F(u,v) = m, F(u+1,v) = n)}{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1} \delta(F(u,v) = m)} \tag{25}$$

$$p\{F(u,v+1) = n \mid F(u,v) = m\} = \frac{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1} \delta(F(u,v) = m, F(u,v+1) = n)}{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1} \delta(F(u,v) = m)} \tag{26}$$

$$p\{F(u+1,v+1)=n \mid F(u,v)=m\}=\frac{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1}\delta(F(u,v)=m,F(u+1,v+1)=n)}{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1}\delta(F(u,v)=m)} \quad (27)$$

$$p\{F(u,v+1)=n \mid F(u+1,v)=m\}=\frac{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1}\delta(F(u+1,v)=m,F(u,v+1)=n)}{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1}\delta(F(u+1,v)=m)} \quad (28)$$

where $m,n \in \{-T,-T+1,\ldots,0,\ldots,T\}$, T is a predefined threshold and

$$\delta(F(u,v)=m,F(u,v+1)=n)=\begin{cases}1, & F(u,v)=m,F(u,v+1)=n\\ 0, & otherwise\end{cases} \quad (29)$$

Similarly, TPMs of all four directions of difference mode 2-D arrays are defined as follows:

$$p\{M_h(s+1,t)=n \mid M_h(s,t)=m\}=\frac{\sum\delta(\begin{matrix}M_h^{(z)}(s,t)=m\\ M_h^{(z)}(s+1,t)=n\end{matrix})}{\sum\delta(M_h^{(z)}(s,t)=m)} \quad (30)$$

$$p\{M_v(s,t+1)=n \mid M_v(s,t)=m\}=\frac{\sum\delta(\begin{matrix}M_v^{(z)}(s,t)=m\\ M_v^{(z)}(s,t+1)=n\end{matrix})}{\sum\delta(M_v^{(z)}(s,t)=m)} \quad (31)$$

$$p\{M_d(s+1,t+1)=n \mid M_d(s,t)=m\}=\frac{\sum\delta(\begin{matrix}M_d^{(z)}(s,t)=m\\ M_d^{(z)}(s+1,t+1)=n\end{matrix})}{\sum\delta(M_d^{(z)}(s,t)=m)} \quad (32)$$

$$p\{M_m(s,t+1)=n \mid M_m(s+1,t)=m\}=\frac{\sum\delta(\begin{matrix}M_m^{(z)}(s+1,t)=m\\ M_m^{(z)}(s,t+1)=n\end{matrix})}{\sum\delta(M_m^{(z)}(s+1,t)=m)} \quad (33)$$

where $m, n \in \{-T_2, \ldots, T_2\}$, T_2 is a predefined threshold. And

$$\delta(A)=\begin{cases}1, & if \quad A \quad holds\\ 0, & otherwise\end{cases} \quad (34)$$

In this paper, we use four difference JPEG 2-D arrays and two difference mode 2-D arrays and set T and T_2 as 4 respectively. A 486-dimentional features will be generated after calculation.

B. Classification Method

With new metric constructed, we apply it to scanned CDPs, either legal ones or fake ones. To some extent, we can be certain that the amount of information loss contained in CDPs after undergoing the print-scan operation is reflected in features extracted. Now that we obtain high dimensional features and we can train a classifier replacing the previous method of threshold. The classifier only need to distinguish the legal CDPs (positive samples) from fake CDPs (negative samples), so it seems to be a two class classifier. As for positive samples, we can obtain legal CDPs as many as possible and insure they are produced in the same conditions including paper type, ink type, printer and scanner type. In other words, positive samples have high similarity. As for negative samples, there is a problem that finding all types of samples is impossible. Although we refer to a limited number of effective attack methods to produce fake CDPs from the view of fakers in previous section, we can't control behaviors of fakers. In other words, image restoration methods and high-resolution print-scan devices used by fakers will create various negative samples beyond our assumption. It is possible for fakers to produce some negative samples which possess similar statistical properties as positive samples in the same metric. The new metric proposed by us has better performance than metrics proposed in previous researches, but it is not perfect because there exists a risk of being attacked. Thus we resort to the powerful function of classifier on the ground of the new metric. We face the phenomenon that positives samples are sufficient but negative samples are incomplete. So a proper solution to the current problem is adopting support vector domain description (SVDD) [17], [18].

TABLE I. EFFECTIVENESS OF VARIOUS ATTACK METHODS

Feature Type	Attack Methods						
	Without Attack	Photoshop	Sharpness a=0.5	Sharpness a=1	Winer Filtering	Lucy-Richardson	Constrained Least Squares Filtering
Entropy [12]	1.47%	22.70%	28.23%	16.13%	11.73%	25.70%	10.40%
Fourier [12]	12.77%	39.87%	37.37%	29.07%	23.97%	36.83%	18.60%
Wavelet [12]	0.00%	32.47%	51.90%	38.47%	1.83%	6.47%	0.93%
Predict [12]	4.07%	26.77%	26.87%	18.63%	20.07%	39.00%	19.60%
486	0.00%	2.43%	2.47%	4.83%	2.80%	2.57%	2.23%

SVDD is a one-class classification method inspired by the support vector machine. Its purpose is to find a sphere with minimal radius covering all positive samples but regarding all other samples as outliers. Seeking for an optimal sphere is a constraint optimization problem. Firstly, we define a data set $\{x_i, i=1,\ldots,N\}$, x_i is the feature of positive samples and negative samples. The object function is defined as follows:

$$F(R,\xi_i)=R^2+C\sum_i\xi_i \quad (35)$$

Subject to the constraint

$$(x_i - a)^T (x_i - a) \le R^2 + \xi_i \quad \forall i, \xi_i \ge 0 \tag{36}$$

where R is the radius of the sphere, a is the center of the sphere, ξ_i is called slack variable, C is a parameter used for adjusting the weight between R^2 and $\sum_i \xi_i$. The function has to be minimized under the constraint. By constructing the Lagrangian, we get

$$L(R, a, \alpha_i, \xi_i) = R^2 + C\sum_i \xi_i - \sum_i \alpha_i \{R^2 + \xi_i - (x_i^2 - 2ax_i + a^2)\} - \sum_i \gamma_i \xi_i \tag{37}$$

Set the partial derivatives to 0 and note that Lagrange multipliers $\alpha_i \ge 0$ and $\gamma_i \ge 0$, we get

$$\sum_i \alpha_i = 1 \tag{38}$$

$$a = \frac{\sum_i \alpha_i x_i}{\sum_i \alpha_i} = \sum_i \alpha_i x_i \tag{39}$$

$$C - \alpha_i - \gamma_i = 0 \quad \forall i \tag{40}$$

Rewrite the object function, we get

$$L = \sum_i \alpha_i (x_i \cdot x_i) - \sum_{i,j} \alpha_i \alpha_j (x_i \cdot x_j) \tag{41}$$

Subject to the constraints

$$0 \le \alpha_i \le C, \quad \sum_i \alpha_i = 1 \tag{42}$$

A sample z will be classified as positive when it satisfies the follow condition

$$(z \cdot z) - 2\sum_i \alpha_i (z \cdot x_i) + \sum_{i,j} \alpha_i \alpha_j (x_i \cdot x_j) \le R^2 \tag{43}$$

V. Experiments

In this experiments, the type of printer we used is EPSON L380, the type of scanner we used is Canon LIDE 120 and the type of paper we used is photo-paper. We printed 150 CDPs (everyone is 100×100 pixels) generated from the same secret key with the resolution of 300 dpi as positive samples. Then from the view of fakers, we scanned all positive samples with the resolution of 600 dpi and applied five kinds of attack methods proposed in Section III to all positive samples respectively. So every attack method will produce 150 fake CDPs, adding up to 750 fake CDPs as negative samples. Then we reprinted these counterfeits with the resolution of 600 dpi rather than 300 dpi in order to make them have the same size as legal ones. Finally, we scanned all CDPs (including positive samples and negative samples) with the resolution of 300 dpi for the following authentication.

A. Evaluation of CDP Metrics

We propose five kinds of attack methods to evaluate the performance of existing CDP metrics in Section III. We respectively use positive samples and negative samples produced by every attack method to train a two-class classifier based on support vector machine (SVM). The effectiveness of different attack methods on different metrics is shown as TABLE I. We use error rate (the ratio of the sum of false positive samples and false negative samples to the sum of positive samples and negative samples) to show the effectiveness. The higher the error rate is, the better the attack method is, the less robust the metric is. Feature types of Entropy, Fourier, Wavelet and Predict showed in the first four rows are proposed in [10], the feature type of 486 showed in the last row is proposed in this paper. Experimental results show that the metric we proposed has low error rate in all different attack methods than those state of the art metrics proposed in [10].

B. One-class Classifier

When we are not sure that what kind of attack method used by fakers, we use positive samples and all negative samples to train a one-class classifier based on SVDD described in Section IV. The performance of the classifier is shown as TABLE II. Where FR means false alarm rate (the ratio of false positive samples to positive samples), FA means missing alarm rate (the ratio of false negative samples to negative samples) and PE means error rate. Moreover, in order to reduce complexity, we select the first 14 features with low error rate as a whole for training. From TABLE II, we can find 486-dimensional features have lower PE but higher FR than 14 selected features. Although FR is high, it is not important in our situation. The phenomenon that PE is high means that more legal CDPs are misclassified as fake CDPs. However, we can control the production of legal CDPs so that we can discard those CDPs which are regarded as fake CDPs by the classifier. We ensure that the FA is low, meaning that most counterfeits will not pass our classifier. So the 486-dimensional features can be used for training the classifier when a higher FR can be accepted and the 14-selected features can be used for training the classifier when low complexity can be needed.

TABLE II. Performance of the SVDD Classifier

Feature Type	Ratio Type		
	FR	FA	PE
Select14	6.67%	8.54%	8.48%
486	16.67%	6.85%	7.15%

VI. Conclusion

In this paper, we discussed the properties of CDP and various authentication methods based on CDP. Then we analyzed the deficiencies of previous metrics from the view of fakers. We used some image restoration methods as attack methods and found that they can be effective for restoring information loss contained in the scanned CDPs. It proved that previous authentication methods are vulnerable to being attacked. So we proposed new metric to measuring the distortion of CDP after the print-scan operation. Finally, considering that fakers will create various fake CDPs beyond our assumption, we trained a one-class classifier based on SVDD. our experimental results show that the metric and

classifier we proposed as a whole achieve significantly lower error rate and better robustness compared with previous methods. In our future work, we will combine CDP-based authentication method with other anti-counterfeit methods to construct a complete and efficient authentication system for printed documents.

ACKNOWLEDGMENT

This work was supported in part by the Natural Science Foundation of China under Grant U1636201 and 61572452.

REFERENCES

[1] B Song, H Wang, Y Zhong, B Chu, Y Su, "Fluorescent and magnetic anti-counterfeiting realized by biocompatible multifunctional silicon nanoshuttle-based security ink," Nanoscale, 2017, 10 (4) :1617-1621.

[2] Wen Zhang, Jie Meng, Conglong Ma, "Research progress of applying digital watermarking technology for printing," 2018 Chinese Control And Decision Conference (CCDC).

[3] R Rossi, G Ruffato, M Massari, F Romanato, "Novel computer generated holograms for high-security anti-counterfeiting applications," 19th Italian National Conference on Photonic Technologies 2017 :37 (4 .)-37 (4 .).

[4] L Yu, X Niu, S Sun, "Print-and-scan model and the watermarking countermeasure," Image & Vision Computing, 2005 , 23 (9) :807-814.

[5] J Ehlenbröker, V Lohweg, "System for simple coding, authentication and copy detection of printed documents," US Patent 9,900,463, 2018.

[6] I Tkachenko, W Puech, C Destruel, O Strauss, JM Gaudin, "Two-Level QR Code for Private Message Sharing and Document Authentication," IEEE Transactions on Information Forensics and Security 2015 , 11 (3) :571-583.

[7] H Jain, G Gupta, S Joshi, N Khanna, "Passive classification of source printer using text-line-level Geometric distortion signatures from scanned images of printed documents," arXiv preprint arXiv:1706.06651, 2017 - arxiv.org.

[8] CW Wong, M Wu, "Counterfeit Detection Based on Unclonable Feature of Paper Using Mobile Camera," IEEE Transactions on Information Forensics and Security 2017, PP (99) :1-1.

[9] Picard J, "Digital authentication with copy-detection patterns," Proc.SPIE, 2004, 5310, pp. 176–183.

[10] Hass B, Dink AE, "Copy detection pattern-based document protection for variable media," Iet Image Processing, 2012, 6 (8): 1102-1113.

[11] AC Yağan, MT Özgen, "A spectral graph wiener filter in graph fourier domain for improved image denoising," Signal & Information Processing, 2017: 450-454.

[12] R Li, W Zhan, Z Hao, "An improved constrained least-squares filter image restoration algorithm," Boletin Tecnico/technical Bulletin, 2017, 55 (1): 236-243.

[13] N Kumar, H Shukla, R Tripathi, "Image Restoration in Noisy Free Images Using Fuzzy Based Median Filtering and Adaptive Particle Swarm Optimization - Richardson-Lucy Algorithm," International Journal of Intelligent Engineering and Systems, 2017, 10 (4): 50-59.

[14] YQ Shi, C Chen, W Chen, "A Markov process based approach to effective attacking JPEG steganography," 8th Information Hiding Workshop, Old Town Alexandria, VA, USA (2006).

[15] C Chen, YQ Shi, "JPEG image steganalysis utilizing both intrablock and interblock correlations," IEEE International Symposium on Circuits & Systems, 2008: 3029-3032.

[16] Acharya T, Tsai P S, "JPEG - Still Image Compression Standard JPEG2000 Standard for Image Compression: Concepts, Algorithms and VLSI Architectures," 2005.

[17] DMJ Tax, RPW Duin, "Support vector domain description," Pattern Recognit Lett , 1999, 20 (11–13): 1191-1199.

[18] L Liu, Y Lu, C Suen, "An Image-Based Approach to Detection of Fake Coins," IEEE Transactions on Information Forensics & Security, 2017, PP (99) :1-1.

ICICN 2019

Pattern Recognition and Classification