computers

MDPI

Article

Enhancing Counterfeit Detection with Multi-Features on Secure 2D Grayscale Codes

Bimo Sunarfri Hantono [1], Syukron Abu Ishaq Alfarozi [1], Azkario Rizky Pratama [1,*], Ahmad Ataka Awwalur Rizqi [1], I Wayan Mustika [1], Mardhani Riasetiawan [2] and Anna Maria Sri Asih [3]

[1] Department of Electrical and Information Engineering, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia; bhe@ugm.ac.id (B.S.H.); syukron.abu@ugm.ac.id (S.A.I.A.); ahmad.ataka.ar@ugm.ac.id (A.A.A.R.); mustika@gmail.com (I.W.M.)
[2] Department of Computer Science and Electronics, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia; mardhani@ugm.ac.id
[3] Department of Mechanical and Industrial Engineering, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia; amsriasih@ugm.ac.id
* Correspondence: azkario@ugm.ac.id

Abstract: Counterfeit products have become a pervasive problem in the global marketplace, necessitating effective strategies to protect both consumers and brands. This study examines the role of cybersecurity in addressing counterfeiting issues, specifically focusing on a multi-level grayscale watermark-based authentication system. The system comprises a generator responsible for creating a secure 2D code, and an authenticator designed to extract watermark information and verify product authenticity. To authenticate the secure 2D code, we propose various features, including the analysis of the spatial domain, frequency domain, and grayscale watermark distribution. Furthermore, we emphasize the importance of selecting appropriate interpolation methods to enhance counterfeit detection. Our proposed approach demonstrates remarkable performance, achieving precision, recall, and specificities surpassing 84.8%, 83.33%, and 84.5%, respectively, across different datasets.

Keywords: cybersecurity; authenticity; integrity; copy-detection pattern; counterfeit prevention; digital watermarking

check for updates

Citation: Hantono, B.S.; Alfarozi, S.A.I.; Pratama, A.R.; Rizqi, A.A.A.; Mustika, I.W.; Riasetiawan, M.; Asih, A.M.S. Enhancing Counterfeit Detection with Multi-Features on Secure 2D Grayscale Codes. *Computers* **2023**, *12*, 183. https://doi.org/10.3390/computers12090183

Academic Editors: Paolo Bellavista, Ömer Aslan and Refik Samet

Received: 20 May 2023
Revised: 31 August 2023
Accepted: 5 September 2023
Published: 14 September 2023

1. Introduction

Recent progress in scanning and printing technology, while beneficial to many aspects of human life, have led to various security challenges. These technologies have inadvertently facilitated the replication of codes, making them less expensive and easier to counterfeit and harder to detect, ultimately undermining efforts to ensure product authenticity. This challenge has directly contributed to financial losses. It has been reported that 5% to 8% of global sales of branded products experience huge losses due to duplication (https://www.oecd.org/gov/risk/counterfeits-and-the-uk-economy.pdf (accessed on 1 August 2022)). Various technologies have been developed to detect counterfeits. Past works have focused on a range of solutions, including UV ink, holograms, and optical variable devices (OVDs). UV ink, applied to a product, remains invisible until it is exposed to ultraviolet light [1]. Holograms, available in different designs and integrated into products, serve as unique identifiers to verify their authenticity [2]. Likewise, OVDs, which exhibit altered appearances based on factors like rotation and tilting, have been extensively employed [3]. However, these technologies are typically costly and require intricate construction processes.

Various products have employed 2D codes (such as QR codes and bar codes) embedded in products containing either public data or a specific URL linked to a product's serial number. This solution is popular mostly due to its simplicity and affordability. However, these advantages come at a price: 2D codes are generally easier to duplicate or replicate.

Researchers are actively investigating methods to develop a secure anti-counterfeit 2D code. Several security measures have been employed in these efforts, such as encryption, which involves converting plain text into an unreadable coded format that can only be deciphered with a decryption key. Watermarking is another technique used, where an invisible code is added to the image, making it challenging for counterfeiters to replicate the code accurately. Additionally, digital signatures are utilized, employing mathematical algorithms to verify the code's authenticity and ensure it has not been tampered with. These measures aim to enhance the security and integrity of the 2D codes in order to combat counterfeiting.

The aim of this paper is to investigate copy-detection patterns (CDPs) attached to the 2D code [4]. This CDP is generated using a combination of raw data embedded in the 2D code and a secret code stored in a server, ensuring the integrity of the CDP itself. Each time the 2D code is scanned and reprinted, some information within the CDP is lost, taking advantage of the image quality degradation. Despite great interest in the CDP technique, previous researchers utilized black-and-white elements (tiny black-and-white squares) in the CDP, which limited the sensitivity of the binary CDP to image quality degradation. This made it easier for counterfeiters to reproduce and counterfeit the code. The present study focuses on investigating the use of grayscale gradation elements in CDP images to address this limitation. The main contributions of this paper are as follows:

- We provide a comprehensive examination of grayscale gradation in CDPs and utilize the distribution of grayscale histograms as a feature to assess the image quality degradation;
- We investigate additional features in spatial and frequency domains using correlation and distance metrics;
- We generate six datasets representing diverse environmental settings;
- Finally, the performance of the proposed approach is compared to a recent and highly relevant paper, specifically the work in [5].

The structure of the paper is outlined as follows. In Section 2, we present a comprehensive review of previous studies focused on enhancing authenticity, with a particular emphasis on techniques such as the copy detection pattern (CDP) and watermarking. The experimental setup is described in Section 3, followed by the presentation of the experiment's results and a subsequent discussion in Section 4. Finally, a summary of our work is provided in Section 5.

2. Related Work

Several works aimed to improve the confidentiality and integrity of 2D codes. One example involved visual cryptography, as reported in [6]. In this work, an image was divided into sub-images (transparencies). The sub-images did not contain any information, thus keeping confidentially. Another approach was proposed in [7]. This approach makes use of the distance and angle variable between a QR code and its scanner to improve safety. Information from the code can only be retrieved once these variables match certain pre-defined values. The authors of [8] performed a simple symmetric encryption. The key used for the encryption process was attached to the QR code.

In the case of a QR code, the authenticity measure of the code is more important than confidentiality, as the code is designed to provide public information, such as a public URL or a unique serial number of a product. To this end, some researchers have proposed techniques to detect duplication. For instance, some have developed methods to detect counterfeiting attacks where a copied QR code is affixed to a fake item, as if it is the original one [9]. Examples of these works include CDP (copy detection pattern) methods, as reported in [4], or a more recent version combining CDP and QR codes [10,11]. Another line of work focused on employing the watermarking technique [12]. The principle of this technique is that the print-and-scan process creates pixel divergence and geometric distortion.

Various aspects of the CDP technique have been explored. For example, evaluation features have been used to measure the efficiency of restoring algorithms. Recently, Zhang et al. [13] proposed new features, denoted as the '486 feature-type', extracted from a CDP image. This showed superior performance compared to existing features, such as, entropy, wavelet domain sharpness, prediction error metric, Fourier domain sharpness, and the features outlined in [14]. The evaluation and comparison of these features were conducted using five restoration algorithms, including constrained least squares filtering, Wiener filtering, the filter method, the Lucy–Richardson algorithm, and the smart sharpen filter in Photoshop. These algorithms act as an "attack" to the CDP by making the fraudulent CDP indistinguishable from the authentic one. Each algorithm was evaluated using each of the five metrics, resulting in an error rate for each attack–metric pair. The 486 feature-type exhibited the lowest error rate across all attack methods, except for constrained least squares filtering and Wiener filtering. The authors developed a novel classifier based on the support vector domain description (SVDD) and evaluated its performance using the 486 features. The performance was evaluated based on three aspects: the false positive to positive sample ratio (FR), the false negative to negative sample ratio (FA), and the error rate (PE). The proposed SVDD classifier attained an FR of 16.67%, an FA of 6.85%, and a PE of 7.15% when using all 486 features. When using only 15 selected features, the classifier attained an FA of 8.54%, an FR of 6.67%, and a PE of 8.48%.

In another recent work, Chaban et al. compared two estimation methods, i.e., a binarization method based on Otsu's adaptive thresholding and a template estimation, which was developed using the LDA algorithm [15]. The Hamming distance for binary images (HAMMING), structural similarity index (SSIM), Jaccard index (JACCARD), and normalized cross-correlation (CORR) similarity metrics were used to assess the effectiveness of those procedures. The test was carried out across a variety of code densities in order to more accurately reflect the method's overall performance. A dataset's density was used with lower and higher entropy values. The actual dataset was produced using two sets of print-and-scan equipment. The proposed estimation achieved probabilities of 6.17% and 7.57% for the dataset with the maximum density tested (50%) using the Hamming distance. The results were less than the two baseline approaches, i.e., the LDA method's 15.24% and 16.34% and the Otsu method's 18.13% and 20.01%. In order to attain the desired performance, the authors evaluated the efficacy of the proposed approach through metric pairing. In this experiment, the HAMMING and SSIM techniques emerged as the top performers, registering the lowest miss and FA rates at 5.05% and 6.88%, respectively.

Khermaza et al. collected the CDP dataset, containing a digital template, original P&S, and fake CDP [5]. An experiment with several different types of data choices was performed. The data choices included the printed unique CDP (consisting of 5000 original CDPs, their templates, and 10,000 copies) and the printed CDP per batch (consisting of 2500 original CDPs, their templates, and 10,000 copies). The authors utilized a normalized cross-correlation score as a feature and neural networks as classifiers with various architectures. The original images were split into patches measuring 13 × 13. Two architectures were proposed. The first one employed a fully connected layer consisting of two, three, and four hidden layers, respectively, where the size of each layer was set to be the input size (169). The second architecture employed a bottleneck DNN (BN DNN) model using two fully connected hidden layers of 128 and 64 dimensions, respectively, at the encoder and decoder parts, and leveraged a latent representation of size 32. These architectures were based on an earlier work by Taran et al. [16]. The training parameters were set at 25 epochs with a batch size of 128. The activation function, the loss function, and the optimizer employed were ReLu, MSE (mean squared error), and Adam [17], with a learning rate of 10^{-3}, respectively. The best-performing method, having the lowest BER percentage, was shown to be the BN DNN method. It produced a mean BER of 23.27% for unique estimation attacks and 18.47% for batch estimation attacks.

Taran et al. regenerated original digital codes from printed ones by utilizing deep neural networks [16]. The neural network models learned digital codes and printed codes in the training phase. The models then generated estimated binary codes via a simple thresholding method. The estimated code was then printed and scanned on the corresponding equipment. To determine the authenticity of the code, the Pearson correlation and normalized Hamming distance were utilized. The findings demonstrated that the integration of modern machine learning technologies in the system made it more challenging for defenders to detect fakes. Taran et al., in their recent publication, examined the effectiveness of QR code authentication in real-life conditions using mobile phones [18]. To simulate real-life conditions, they printed the digital codes and captured images of the printed codes using a mobile phone with automatic settings. The phone was positioned parallel to the printed code at a distance of 11 cm. To create counterfeit copies, two standard copy machines were used on white and gray papers. The fake images were acquired in the same environment and with the same settings. The authors then employed machine learning techniques to authenticate the original codes and the counterfeits. The findings indicate that the classifier demonstrated a high level of accuracy when distinguishing between the original codes and the counterfeits, with a classification error of approximately 0.28%. Further, the classifier was capable of distinguishing fakes from different copy machines, but the study encountered a limitation in effectively distinguishing between counterfeit codes printed on white paper and those printed on gray paper.

The summarized works are presented in Table 1. The authors' primary objective was to differentiate between original and fake codes generated by a simple duplication or by predicting original templates through image processing or machine learning approaches. To differentiate counterfeit codes, a common thresholding approach was employed to normalized correlation measures like the Pearson correlation. In this research, additional features were introduced in both the spatial and frequency domains. Moreover, by exploring grayscale gradation in watermark images, changes in grayscale histogram distribution were utilized as indicators of counterfeit codes. Multiple machine learning techniques were implemented and validated using six distinct datasets collected from diverse sources. Lastly, a comparison was made between the proposed technique and the approach suggested by Khermaza et al. [5].

Table 1. Summary of related works.

Year **Authors**	**2023** **This Paper**	**2023** **Taran et al. [18]**	**2021** **Picard et al. [10]**	**2021** **Khermaza et al. [5]**	**2021** **Alfarozi and Pratama [11]**	**2019** **Zhang et al. [13]**	**2019** **Taran et al. [16]**
Binary/grayscale codes	grayscale CDP	binary	binary	binary	binary	binary	binary
Features for counterfeit detection	Spatial and frequency domain using a correlation, distance metrics, and histogram distribution	Pearson correlation and Hamming distance	Correlation coefficient	Normalized cross-correlation	Auto-generated features	486 features, including entropy, wavelet, and Fourier domain sharpness	Normalized hamming distance, Pearson correlation
Approach	Gradient boosting classifier	Support vector machine	Thresholding	Thresholding	CNN	Support vector domain description	Thresholding
Dataset	6 datasets, total of about 1200 images	Indigo mobile dataset, 300 template images, 1200 copy fake images	Not available	CDP dataset (27,500 images)	CDP dataset by [5]	150 template images, 750 fake copy images	PGC dataset, 384 ori images
Evaluation metrics	Precision, recall, and specificity	Percentage of classification error	Not available	BER	BER	False alarm rate, missing alarm rate, and mean error rate	ROC curves
Aims	Improving 2D secure QR codes and authentication method performance	Authentication QR codes in real-life conditions using mobile phones	Protecting copy attacks by integrating CDP into QR codes	Presenting a publicly available dataset and evaluating estimation attacks using neural network models as in [16]	Improving authentication method using CNN	Improving authentication performance	Generating reconstructed codes using deep neural networks to test the clonability attacks of QR codes.

3. Methodology

In this section, we discuss the proposed 2D secure code approach to prevent counterfeiting. The system consists of a SQR generator and authenticator, as shown in Figure 1.

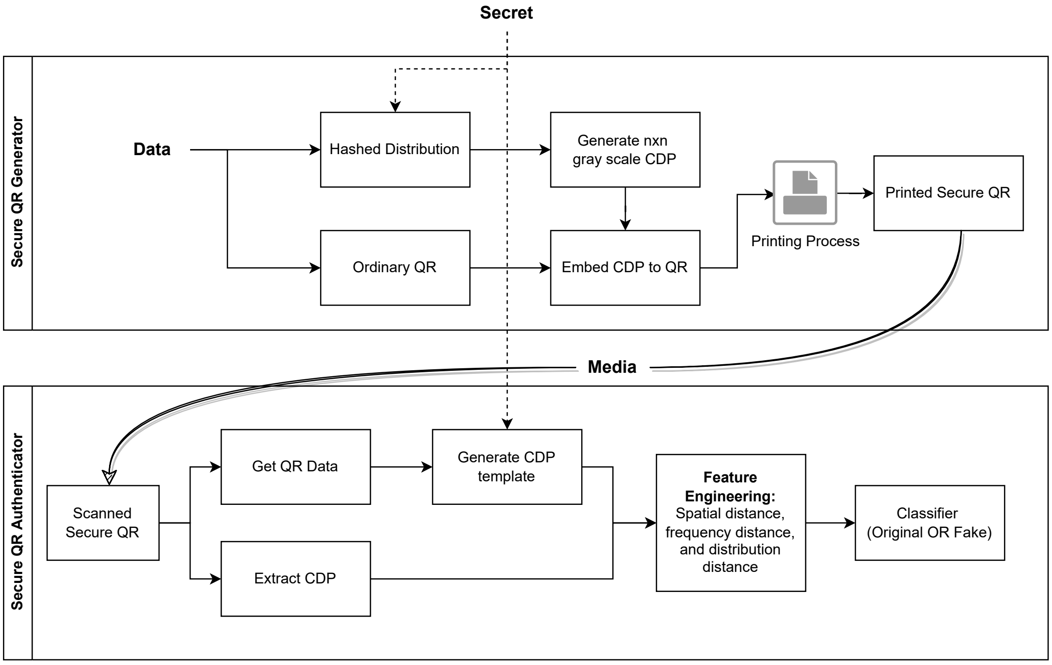

Figure 1. Print and scan process of SQR.

3.1. Secure QR Generator

The proposed secure QR (SQR) code is based on a CDP. The CDP is a pseudo-random digital image that contains high-frequency information, embedded in the QR code for counterfeit detection.

To construct CDP, we initially computed grayscale distribution based on QR data and a secret. While the QR data are public data stored in the 2D code, the secret is a private key stored in a central server. Thus, we generate a multinomial distribution $P = [p_1, p_2, \dots p_L]$ and its gray-level values $V = [v_1, v_2, \dots, v_L]$, where L is the quantization number determined by a hashing function. We opted for an $L = 8$ level grayscale quantization because a two-level intensity is deemed insufficiently secure against estimation attacks, as reported in [5]. We generate $n \times n$ CDP, such that

$$I_{i,j} = choice(V_{i,j}; dist = P) \tag{1}$$

where $I_{i,j}$ is the intensity level of the i-th row and j-th column. A static random number generator (RNG) in (1) is used to ensure that the unique generation of the hashing function results in the same CDP every time it is generated for a given set of d and s values, which corresponds to the data and secret respectively. Finally, the generated CDP is embedded into the SQR code and is ready to attach to a product. In our case, the CDP is generated with a size of 80 × 80 and a symbol size 1 × 1, printed on 0.14 mm for each symbol.

3.2. Dataset

We collected six datasets, each possessing unique characteristics based on the used P&S devices, such as the camera type, printer type, and lighting conditions. We invited six volunteers to print and capture our generated QR codes as original sets. We then asked them to reprint the sets and recapture them to create fake copies. These datasets were subsequently loaded into our detector. The summary of our dataset is shown in Table 2.

Table 2. Collected dataset for validation purposes.

Dataset	Printer	Camera	Scanned QR Resolution	Number of Original	Number of Copy
1	Fuji Xerox Versant	Realme GT Neo 3T	2025 × 2025	120	119
2	Canon C650	Fujifilm X-A7	3334 × 3334	200	155
3	Canon C650	Samsung A33 5G	1272 × 1286	200	200
4	Canon C650	iPhone XR	3024 × 3024	200	200
5	Canon C650	Realme GT Master ed.	1998 × 1998	200	199
6	Canon C650	POCO X3 NFC	2725 × 2725	200	227

The datasets were collected under various environmental settings, resulting in variations in quality and style due to different devices and lighting conditions. The SQR data for the original and copied images in each dataset are presented in Table 3. Dataset 1, dataset 3, and dataset 6 demonstrate a higher degree of consistency in terms of the lighting quality between the original and copy SQR images, where the intensity differences are smaller. Moreover, dataset 1 and dataset 6 exhibit higher levels of lighting compared to the other datasets, with the mean intensity of the whitespace border approaching a value of 255.

Table 3. Sample data of the original and copied secure 2D codes of 6 environment settings.

Class	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6
Original						
Copy						
Intensity difference *	0.2	13	4.9	18.0	12.6	0.0

* The absolute mean difference intensity, $I \in (0, 255)$, of the whitespace border of the QR code between the original and copy.

To assess the performance of the proposed method, 10-fold cross-validation is performed on each dataset. Finally, all datasets are combined into a single dataset and evaluated using the same procedure as the individual sets of data.

3.3. *Secure QR Authenticator*

Given a product with SQR, one may authenticate it by scanning the code. As the sizes of the scans may be different, scaling and resizing might be needed. The authenticator then extracts CDP and generates the CDP template based on the QR data (d) and secret key (s). This approach enhances the challenge for potential attackers attempting to forge counterfeit CDPs, as the template will differ for every unit item or product even with the same type of product, contingent on the QR data (i.e., serial number of the unit). The authenticator then extracts several features and classifies the code to show the originality states.

3.3.1. Scaling Methods

The scanned CDP and template are resized to the same dimensions so that the features can be measured. The original resolution of the CDP template is 80 × 80 pixels. This template is resized to be four times bigger, reaching 320 × 320 pixels. As the resolution of the scanned QR code (and the scanned CDP, correspondingly) might be different, we resize the scanned CDP to the same size as the scaled CDP template. The resized factors are different depending on the scanner or camera resolution.

Furthermore, since we encompass eight distinct gray intensity levels, it is necessary to estimate or generate intermediate gray-level values by considering the existing values of adjacent pixels when we scale the image. To achieve this, we employ both linear and non-linear interpolation techniques. In the former, we estimate the gray value between two given data points by projecting the values equidistantly between them, while in the latter, we interpolate the gray value between two known data points by assuming that the interpolation follows a polynomial function of degree 3.

3.3.2. Watermark localization

The watermark is placed in the center of a QR code. Thus, we need to extract the watermark from the image using the localization technique. In practice, the SQR image is not always aligned horizontally or vertically, it could be subject to rotation or affine transformations due to camera variations.

A four-point homography transformation is employed to align the QR image, ensuring it maintains a perfect square dimension. The homography transformation is described as the following equation,

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix}$$

where (x', y', w') represents the transformed coordinates and (x, y, w) represents the original coordinates. The homography matrix is denoted by $\mathbf{H}$ and has the following form:

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}$$

To compute the homography matrix, we need to solve a system of equations using the four corresponding points in the original and transformed coordinate systems [19]. After the transformation, the extraction of the watermark is straightforward due to its central positioning within the image, with its dimensions being proportionate to the QR code's size, as depicted in Figure 2. We utilize the OpenCV library to implement this transformation with its default interpolation technique, i.e., linear interpolation.

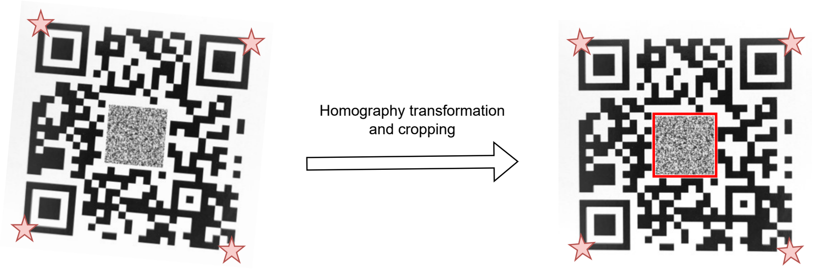

Figure 2. Localization and cropping method using four corners of a QR image. Stars represent the four-point homography transformation, while the red square denote watermark.

3.3.3. Feature Engineering

We utilize three feature representations of the CDP images: the spatial domain, frequency domain, and the grayscale distribution. An image can be represented in both the spatial and frequency domains. The spatial domain refers to the actual pixel values of an image, where each pixel is assigned a specific intensity value. On the other hand, the frequency domain represents the image in terms of its frequency components. This representation transforms the image from the spatial domain to the frequency domain using techniques such as the Fourier transform. In our case, discrete cosine transform (DCT) is used. The result involves a set of coefficients, describing the image's frequencies. In the frequency domain, low-frequency components represent the image's general structure and large-scale features, while high-frequency components represent fine details and small-scale structures. Analyzing an image in the frequency domain makes it possible to identify patterns and features that are not easily noticeable in the spatial domain representation. Specifically, the high frequency pertains to the existence of swiftly changing or fluctuating pixel values within the image. Within the DCT, high-frequency elements symbolize intricate details like distinct or sharp edges, textures, or swift pixel-intensity transitions in the image. On the other hand, lower frequencies depict the image's general structure or gradual variations. As shown by Schulze et al. [20], photocopied images and the first printing show different representations in the frequency domain. Such a representation offers superior performance in detecting copied documents compared to spatial features.

The grayscale distribution of an image refers to the distribution of gray-level values or pixel intensities in an image. In a grayscale image, each pixel is assigned an intensity value ranging from 0 (black) to 255 (white), with intermediate values representing shades of gray. The grayscale distribution of an image describes the distribution of these intensity values across the entire image.

The grayscale distribution can be visualized using a histogram, which is a graph that shows the frequency of occurrence of each intensity value in the image. The histogram of a grayscale image typically has a bell-shaped curve that reflects the most common intensity values in the image. The shape of the histogram can reveal information about the image's contrast, brightness, and overall tonality. When a document is copied, the details will be degraded and the images will be visually blurred. As a result, the grayscale distribution will also be affected. Lim et al. [21] illustrated the difference in the histogram distribution between blurred and sharpened images, demonstrating that the histogram can detect texture loss in an image.

This study incorporates multiple distance metrics to assess the similarity between extracted features from a sample and a reference template for each feature representation (i.e., histogram, and spatial and frequency domains). Four distance metrics, namely Euclidean distance, Cosine distance, Canberra distance, and correlation distance, are utilized. In addition, normalized cross-correlation is also employed, as described in [22]. The classifier is then supplied with a total of 13 distance metrics, which encompass the 12 different distance metrics derived from 3 domains, along with the normalized cross-correlation.

3.3.4. Machine Learning Authenticator

The gradient boosting classifier is a widely adopted and powerful machine learning algorithm that excels at handling complex and non-linear problems, such as a classifier [23]. By aggregating multiple weak models into a single strong model, it has the ability to effectively deal with multiple features and the intricate relationships among them, resulting in superior authentication performance. Additionally, the algorithm can manage noisy or irrelevant features, reducing the likelihood of overfitting and guaranteeing good generalization performance. These attributes make gradient boosting an ideal candidate for use as an authenticator in this study. Furthermore, our preliminary experiments on the CDP dataset demonstrated that gradient boosting is more stable than logistic regression and support vector machines (SVMs) in terms of performance.

In this study, we use the gradient boosting classifier with 100 estimators and a maximum depth of 1, creating an ensemble of 100 decision trees, where each tree has a maximum depth of 1. It starts by training the first tree on the data and updating weights based on prediction errors. Then, subsequent trees are trained to focus on the previously misclassified examples. The final prediction is made by aggregating the predictions of all the trees.

This approach gradually learns complex patterns by combining multiple simple trees. The input consists of 13 features and the output classifies the CDP as either original or fake.

3.4. Metrics

The machine learning models are evaluated in terms of five metrics, namely, accuracy, precision, recall, F1 score, and specificity. These metrics are used to calculate the model performance based on 10-fold cross-validation results. Specifically, the model is trained on nine folds and tested on the remaining fold. This process is repeated 10 times, each time using a different fold as the testing set, and the results are averaged to obtain the final evaluation of the model.

3.4.1. Accuracy

The accuracy measures the number of predictions of original and fake QR codes that are correct as a percentage of the total number of predictions that are made.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

3.4.2. Precision

The precision in detecting the original QR code is defined as the ratio of true positives (real original QR codes correctly identified) to the total number of QR codes classified as original. The $Precision$ is defined as

$$Precision = \frac{TP}{TP + FP}$$

3.4.3. Recall

The recall—in regard to detecting the original QR code—is defined as the ratio of true positives (original QR codes correctly identified) to the total number of actual original QR codes, regardless of whether they were detected as original or fake. $Recall$ is defined as

$$Recall = \frac{TP}{TP + FN}$$

3.4.4. F_1 Score

The F_1 score is the harmonic mean of precision and recall with an equal weight.

$$F_1 = 2 \cdot \frac{recall \cdot precision}{recall + precision}$$

3.4.5. Specificity

The specificity in detecting a fake QR code is defined as the ratio of the fake/copied QR code that is identified correctly (true negative) over the whole fake/copied QR codes, detected as fake or original QR codes. The $Specificity$ is defined as

$$Specificity = \frac{TN}{TN + FP}$$

4. Results and Discussion

We use the approach proposed by Khermaza et al. as a baseline for comparison in counterfeit detection [5]. Thus, we present our proposed approach on the same datasets to assess its improvements.

4.1. Baseline

We adopt the single-feature cross-correlation technique introduced in [5,22]. Additionally, we investigate the effects of interpolation methods on the authentication performance.

Table 4 shows the experimental results, bold typeface indicates the a particular interpolation method is better. The 'all' dataset row represents the results of the combined six datasets. In the baseline scenario, neither the linear nor third-degree polynomial interpolation shows significant enhancement in performance. Both methods yield a mere 2% difference, suggesting that neither scaling method demonstrates superiority when considering a single feature.

Table 4. Counterfeit detection evaluation based on a single feature proposed by [5], with different scaling methods (i.e., linear interpolation and third-order polynomial).

Dataset	Accuracy		Precision		Recall		F1 score		Specificity	
	Linear	**Poly3**	**Linear**	**Poly3**	**Linear**	**Poly3**	**Linear**	**Poly3**	**Linear**	**Poly3**
1	**49.76%**	48.91%	48.97%	**50.10%**	50.00%	50.00%	48.64%	**49.62%**	**49.32%**	47.80%
2	56.40%	**57.81%**	58.89%	**60.62%**	**74.50%**	73.50%	65.37%	**65.90%**	32.92%	**37.46%**
3	51.75%	**54.25%**	53.07%	**54.70%**	46.50%	**51.00%**	48.70%	**52.61%**	57.00%	**57.50%**
4	**54.75%**	48.25%	**54.40%**	48.97%	**65.00%**	49.50%	**58.66%**	48.77%	44.50%	**47.00%**
5	53.14%	**55.38%**	52.73%	**55.87%**	**57.50%**	55.50%	**54.83%**	54.79%	48.76%	**55.34%**
6	**53.33%**	52.45%	49.24%	**49.62%**	**49.00%**	45.50%	**48.66%**	46.56%	57.17%	**58.64%**
all	**54.28%**	52.30%	**53.58%**	52.07%	70.09%	**70.71%**	**60.63%**	59.87%	**38.18%**	33.55%

When considering counterfeit detection using a single feature, we can expect an accuracy range between 49% and 57% and an F1 score range between 48% and 65%. The dataset with the poorest performance is dataset 1, while dataset 2 exhibits the best performance in this setup. These discrepancies can primarily be attributed to the quality of the replicated data. As outlined in Section 3.2, dataset 1 is collected under consistent lighting conditions, resulting in minimal visual variations between the original and copied SQR images (i.e., an intensity difference of 0.2). In contrast, the original and copied images in dataset 2 possess some degree of visual distinctiveness (i.e., an intensity difference of 13.0), resulting in improved classification outcomes. A similar pattern is observed in dataset 4, where the original and copied images exhibit visual distinctiveness (i.e., an intensity difference of 18.0). The variation in visual characteristics likely influences the differences in performance when detecting counterfeits in this dataset

Upon closer examination of dataset 2, it is apparent that the specificity is below 38%, whereas the F1 score reaches 65.90%. This outcome indicates a deficiency in the capability to detect counterfeit QR codes within the dataset. One potential explanation for this is that the single-feature cross-correlation approach is unable to differentiate between copied and fake QR codes, even when dealing with a dataset containing substantial intensity differences between the original and copied codes, as seen in dataset 2. Consequently, there is a notable occurrence of false positive detections (identifying false original QR codes) and a low detection rate for true negatives.

4.2. Multiple Features

The incorporation of multiple features has been proven to significantly improve the performance of authentication systems, as illustrated in Table 5. Using the linear interpolation scaling method, the accuracy of the authentication system varies between 83% and 99%. Notably, when dealing with visually distinct data with intensity differences of more than 10, such as datasets 2, 4, and 5, the system exhibits excellent detection of original and copied QR codes, achieving F-measure scores of up to 99%. For the other three datasets, the system achieves a minimum F-measure of 83%. These positive results are also evident in the specificity metric, demonstrating the effectiveness of the proposed features in identifying original QR codes and minimizing the false detection of original QR codes.

Although the individual sets of data demonstrate high levels of accuracy, F-measure, and specificity, the same measures in the combined dataset (i.e., 'all' datasets) decline to 77.88%, 78.05%, and 75.55%, respectively, due to the presence of diverse variations and styles across the datasets. Consequently, developing a model specifically tailored to a particular dataset or a specific P&S (print and scan) device yields better performance compared to a model trained on multiple datasets simultaneously. However, in certain cases, such as in supply chain management, it may be challenging to build separate anti-counterfeit models for each dataset or device, as multiple individuals are involved in the business process.

In addition to the proposal of incorporating more features, it is noteworthy that the linear interpolation method consistently outperforms the third-order polynomial interpolation scaling method across all evaluation metrics. This finding could be attributed to the fact that the higher-order polynomial interpolation method alters the curvature of high-frequency data, resulting in a smoother signal compared to the linear interpolation method. Further research is necessary to fully understand these implications and explore the potential for optimizing the performance of authentication systems through advanced interpolation methods.

Table 5. Counterfeit detection evaluation based on multiple features with different scaling methods (i.e., linear interpolation and third-order polynomial).

Dataset	Accuracy		Precision		Recall		F1 score		Specificity	
	Linear	**Poly3**	**Linear**	**Poly3**	**Linear**	**Poly3**	**Linear**	**Poly3**	**Linear**	**Poly3**
1	**84.15%**	81.25%	**84.84%**	84.00%	**83.33%**	77.50%	**83.17%**	79.06%	84.92%	**85.00%**
2	**98.04%**	92.44%	**98.52%**	94.75%	**98.00%**	92.00%	**98.23%**	92.77%	**98.08%**	92.83%
3	**84.00%**	77.25%	**85.72%**	79.87%	**83.50%**	75.00%	**83.89%**	76.62%	**84.50%**	79.50%
4	**99.00%**	97.50%	**99.09%**	98.50%	**99.00%**	96.50%	**99.01%**	97.46%	**99.00%**	98.50%
5	**94.49%**	86.98%	**94.39%**	87.44%	**95.00%**	87.00%	**94.52%**	87.06%	**94.00%**	86.97%
6	86.70%	**87.86%**	85.83%	**85.98%**	87.00%	**90.50%**	85.85%	**87.66%**	**86.40%**	85.53%
all	**77.88%**	68.92%	**78.69%**	69.47%	**80.18%**	70.80%	**78.05%**	67.54%	**75.55%**	67.00%

5. Conclusions

Our study examined the effectiveness of different approaches in detecting counterfeit items by utilizing both single and multiple features. When focusing on a single feature, we observed accuracy ranging from 49% to 57% and F1 scores ranging from 48% to 65%. However, when incorporating multiple features, we observed a substantial improvement in the performance of the authentication system, with accuracy ranging from 84% to 99% and F1 scores ranging from 83% to 99%. These enhancements can be attributed to the effectiveness of the proposed features, specifically the utilization of grayscale distribution and image frequency, which significantly contribute to the classification process. It is worth noting that selecting appropriate interpolation methods in counterfeit detection systems may improve the performance of counterfeit detection. In our cases, linear interpolation mostly outperforms the third-order polynomial interpolation.

The classification performance in counterfeit detection is significantly influenced by the quality of the original and copied dataset. When the original and fake images are more similar, the classification process becomes more challenging. Dataset 2 consistently demonstrates the best performance, likely due to the visual distinctiveness between the original and copied images. Conversely, dataset 1 demonstrates the poorest performance, attributed to consistent lighting conditions that result in minimal visual variations.

While developing a model specifically tailored to a particular dataset yields superior performance, it may not always be practical in operational scenarios where images are captured by multiple individuals under various conditions. Our work demonstrates that a combined dataset ('all' datasets) consisting of approximately 2000 QR images achieves an F-measure of 78% and a specificity of 75.55%.

In summary, our study highlights the importance of dataset quality, the impact of interpolation methods, and the effectiveness of multiple features in the spatial domain, frequency domain, and grayscale distribution. These findings contribute to the understanding and improvement of counterfeit detection systems. In future works, we will explore more advanced testing methods, including the use of post-processing techniques like visual enhancement or machine learning to generate fake images after reprinting; moreover, we will assess whether modifications to the proposed secure QR codes are necessary.

Author Contributions: Conceptualization, B.S.H.; methodology, S.A.I.A. and A.R.P.; software, S.A.I.A.; validation A.A.A.R.; investigation, A.R.P. and A.A.A.R.; writing, A.R.P., S.A.I.A. and A.A.A.R.; data curation, M.R.; resources, A.M.S.A. and I.W.M.; writing—review and editing, B.S.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Lembaga Pengelola Dana Pendidikan (LPDP)/ the Indonesia Endowment Fund for Education Riset Inovatif Produktif (RISPRO) invitation scheme, contract number PRJ-8/LPDP/LPDP.4/2022.

Data Availability Statement: Not applicable.

Acknowledgments: The authors acknowledge the collaboration with PT. Telkom Indonesia regarding the collection of secure 2D data and in conducting R&D in SCM. The authors thank I Ketut Agung Enriko, Hedi Krishna, and Muhammad Andy Zaky for their suggestions and advice about the system architecture.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

URL	uniform resource locator
OVD	optical variable device
UV	ultraviolet
CDP	copy detection pattern
DCT	discrete cosine transform
P&S	print and scan
BER	bit error rate
QR code	quick response code
SQR	secure QR
ReLU	rectified linear unit
SVM	support vector machine

References

1. Muthamma, K.; Sunil, D.; Shetty, P. Luminophoric organic molecules for anticounterfeit printing ink applications: An up-to-date review. *Mater. Today Chem.* **2020**, *18*, 100361. [CrossRef]
2. Ruffato, G.; Rossi, R.; Massari, M.; Mafakheri, E.; Capaldo, P.; Romanato, F. Design, fabrication and characterization of computer generated holograms for anti-counterfeiting applications using OAM beams as light decoders. *Sci. Rep.* **2017**, *7*, 18011. [CrossRef] [PubMed]
3. Jung, C.; Kim, G.; Jeong, M.; Jang, J.; Dong, Z.; Badloe, T.; Yang, J.K.; Rho, J. Metasurface-driven optically variable devices. *Chem. Rev.* **2021**, *121*, 13013–13050. [CrossRef] [PubMed]
4. Picard, J. Digital authentication with copy-detection patterns. In *Optical Security and Counterfeit Deterrence Techniques V*; van Renesse, R.L., Ed.; SPIE: Bellingham, WA, USA, 2004; Volume 5310, pp. 176–183. [CrossRef]
5. Khermaza, E.; Tkachenko, I.; Picard, J. Can copy detection patterns be copied? evaluating the performance of attacks and highlighting the role of the detector. In Proceedings of the 2021 IEEE International Workshop on Information Forensics and Security (WIFS), Montpellier, France, 7–10 December 2021; pp. 1–6.
6. Moretto, T. Secure 2D Barcodes Based on Visual Cryptography. Master's Thesis, Università Ca'Foscari Venezia, Venice, Italy, 2018. Available online: http://dspace.unive.it/bitstream/handle/10579/13391/844996-1224864.pdf (accessed on 1 June 2023).
7. Pan, H.; Chen, Y.C.; Yang, L.; Xue, G.; You, C.W.; Ji, X. mqrcode: Secure qr code using nonlinearity of spatial frequency in light. In Proceedings of the 25th Annual International Conference on Mobile Computing and Networking, Los Cabos, Mexico, 21–25 October 2019; pp. 1–18.
8. Thamer, S.K.; Ameen, B.N. A new method for ciphering a message using QR code. *Comput. Sci. Eng* **2016**, *6*, 19–24.
9. Krombholz, K.; Frühwirt, P.; Kieseberg, P.; Kapsalis, I.; Huber, M.; Weippl, E. QR code security: A survey of attacks and challenges for usable security. In Proceedings of the Human Aspects of Information Security, Privacy, and Trust: Second International Conference, HAS 2014, Held as Part of HCI International 2014, Heraklion, Greece, 22–27 June 2014; pp. 79–90.

10. Picard, J.; Landry, P.; Bolay, M. Counterfeit Detection with QR Codes. In Proceedings of the 21st ACM Symposium on Document Engineering, New York, NY, USA, 24–27 August 2021. [CrossRef]
11. Alfarozi, S.A.I.; Pratama, A.R. CNN-Based Model for Copy Detection Pattern Estimation and Authentication. *J. Nas. Tek. Elektro Dan Teknol. Inf.* **2023**, *12*, 44–49. [CrossRef]
12. Shi, D.; Wang, Q.; Liang, C. Digital Watermarking Algorithm for Print-and-Scan Process Used for Printed Matter Anti-counterfeit. In Proceedings of the 2008 Congress on Image and Signal Processing, Sanya, China, 27–30 May 2008; Volume 5, pp. 697–701. [CrossRef]
13. Zhang, P.; Zhang, W.; Yu, N. Copy detection pattern-based authentication for printed documents with multi-dimensional features. In Proceedings of the 2019 7th International Conference on Information, Communication and Networks, ICICN 2019, Macau, China, 24–26 April 2019; pp. 150–157. [CrossRef]
14. Dirik, A.E.; Haas, B. Copy detection pattern-based document protection for variable media. *IET Image Process.* **2012**, *6*, 1102–1113. [CrossRef]
15. Chaban, R.; Taran, O.; Tutt, J.; Holotyak, T.; Bonev, S.; Voloshynovskiy, S. Machine learning attack on copy detection patterns: Are 1 × 1 patterns cloneable? In Proceedings of the 2021 IEEE International Workshop on Information Forensics and Security, WIFS 2021, Montpellier, France, 7–10 December 2021. [CrossRef]
16. Taran, O.; Bonev, S.; Voloshynovskiy, S. Clonability of Anti-counterfeiting Printable Graphical Codes: A Machine Learning Approach. In Proceedings of the ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing, Brighton, UK, 12–17 May 2019; pp. 2482–2486. [CrossRef]
17. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
18. Taran, O.; Tutt, J.; Holotyak, T.; Chaban, R.; Bonev, S.; Voloshynovskiy, S. Mobile authentication of copy detection patterns. *EURASIP J. Inf. Secur.* **2023**, *2023*, 4. [CrossRef] [PubMed]
19. Hartley, R.; Zisserman, A. *Multiple View Geometry in Computer Vision*; Cambridge University Press: Cambridge, UK, 2003.
20. Schulze, C.; Schreyer, M.; Stahl, A.; Breuel, T. Using DCT features for printing technique and copy detection. In Proceedings of the Advances in Digital Forensics V: Fifth IFIP WG 11.9 International Conference on Digital Forensics, Orlando, FL, USA, 26–28 January 2009; pp. 95–106.
21. Lim, H.; Yu, S.; Park, K.; Seo, D.; Paik, J. Texture-Aware Deblurring for Remote Sensing Images Using L0-Based Deblurring and L2- Based Fusion. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 3094–3108. [CrossRef]
22. Briechle, K.; Hanebeck, U.D. Template matching using fast normalized cross correlation. In Proceedings of the Optical Pattern Recognition XII, Orlando, FL, USA, 19 April 2001; Volume 4387, pp. 95–102.
23. Hastie, T.; Tibshirani, R.; Friedman, J.H.; Friedman, J.H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: Berlin/Heidelberg, Germany, 2009; Volume 2.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.