Fusing Multiscale Texture and Residual Descriptors for Multilevel 2D Barcode Rebroadcasting Detection

Anselmo Ferreira  
Department of Information Engineering and Mathematics  
University of Siena  
Siena, Italy  
anzelmo.castelo@unisi.it

Changcheng Chen  
Shenzhen Key Laboratory of Media Security  
Shenzhen University  
Shenzhen, Guangdong, China  
cschen@szu.edu.cn

Mauro Barni  
Department of Information Engineering and Mathematics  
University of Siena  
Siena, Italy  
barni@di.unisi.it

Abstract—Nowadays, 2D barcodes have been widely used for advertisement, mobile payment, and product authentication. However, in applications related to product authentication, an authentic 2D barcode can be illegally copied and attached to a counterfeited product in such a way to bypass the authentication scheme. In this paper, we employ a proprietary 2D barcode pattern and use multimedia forensics methods to analyse the scanning and printing artefacts resulting from the copy (rebroadcasting) attack. A diverse and complementary feature set is proposed to quantify the barcode texture distortions introduced during the illegal copying process. The proposed features are composed of global and local descriptors, which characterize the multi-scale texture appearance and the points of interest distribution, respectively. The proposed descriptors are compared against some existing texture descriptors and deep learning-based approaches under various scenarios, such as cross-datasets and cross-size. Experimental results highlight the practicality of the proposed method in real-world settings.

Index Terms—2D Barcode; Copy Detection; Anti-Counterfeiting; Texture Descriptors.

I. INTRODUCTION

The widespread availability of high quality printers and scanners has made it cheaper and easier for a counterfeiter to make illegal copies of existing barcodes and attach them to counterfeited products. As a matter of fact, product counterfeiting is becoming an unprecedented problem for the global trade, as, for example, the 2016 report from Frontier Economics [4] predicted that, in 2022, 991 billion dollars will be traded in fake products, with an estimate of 4.2 to 5.2 million jobs lost because of piracy.

In order to tackle such an issue, 2D barcode authentication techniques proposed in the literature have been divided into active and passive categories. Active techniques [10], [11] are based on specific printing materials that could not be replicated by forgers and are commonly very expensive. On the other hand, passive-proof techniques [2], [12] exploit replication artifacts without using special imaging devices or barcode patterns, taking into account simple imaging devices such as smartphones. Notwithstanding, there is still an open field for other cheap passive anti-counterfeiting techniques, especially with regard to new types of 2D barcodes. Another important aspect to be dealt in such application is the high cost (financial and procedural) of building big datasets in order to make deep learning suitable training data. Therefore, approaches that could learn well from a small amount of training data are crucial to build efficient and cheap authentication approaches.

In this paper, we take such issues into consideration by proposing a two-stage multilevel 2D barcode image texture descriptor applied to a novel multilevel 2D barcode. Our approach takes into account a diverse, complementary and robust feature set composed of global and local features, with these features being designed to capture specific distortions suffered by the considered barcode in its counterfeited version. The global features are used to describe artifacts in image pixels, being calculated by the Multiscale Rotation Invariant Binary Gabor Patterns (MRI BGP), whereas the local features are calculated by a Bag of Multiscale Residual Words (BOMRW). A final descriptor fuses both features for a powerful authentication. We validate the effectiveness of our method on two datasets built with a large set of devices. Compared to other descriptors, the results of our extensive experiments showed promising results in challenging scenarios.

The remaining of this paper is organized as follows. Section II discusses the 2D barcode considered and our proposed authentication scheme for 2D Multilevel barcode copy detection. Section III shows experimental results on diverse scenarios. Finally, Section IV concludes the work and discusses future perspectives of research.

II. PROPOSED METHOD

Our solution is demonstrated with the proprietary generic multilevel 2D barcode proposed in [13]. Examples of genuine
and counterfeited multilevel 2D barcodes are shown in Figure 1. It can be seen from that figure that the additional scan-and-print operation in the illegal copying process yields specific textural distortions, which differentiate the genuine and counterfeited samples. However, as barcodes can be printed with different sizes and are acquired by different devices on varying angles and distances, the texture patterns can vary significantly even in the same class (i.e., genuine or counterfeited). For example, Figures 1(a) and (c) illustrate different textures within the same genuine or counterfeited barcodes, but rendered with different printing sizes. Furthermore, Figure 2 shows some patches from different genuine and counterfeited barcodes with the same area. It can be seen from Figure 2(b) and (d) that the counterfeited barcodes do not have a common halftone texture pattern. Worse still, the genuine patch in Figure 2(c) has some irregular halftones caused by acquisition noise and the printing process. Therefore, the halftone dots in Figure 2(a) are different from those in Figure 2(c). Such irregular halftoning patterns are not easy to capture with the existing textural descriptors. In this paper, we propose a diverse feature set in order to alleviate such a problem. The proposed features are summarized in the following subsections.

A. Global Feature Set: Multiscale Rotation Invariant Binary Gabor Patterns (MRIBGP)

The first proposed descriptor is based on [14]. The Binary Gabor Pattern (BGP) is calculated by convolving $J$ Gabor filters of different orientations with image patches. The outputs are the sums of filtering responses in a neighborhood and are concatenated in a response vector $r = \{r_j : j = 0, \ldots, J - 1\}$. By thresholding the values in $r$, a binary vector $b = \{b_j : j = 0, \ldots, J - 1\}$ is obtained. The BGP value is then calculated by converting the $J$-bits binary number in $b$ into decimal.

To achieve rotation invariance, the maximum $\text{BGP}$ is calculated after several circular bitwise right shifts in the binary vector to yield the Rotation Invariant $\text{BGP}$ ($\text{RIBGP}$). That is

$$\text{RIBGP} = \max\{\text{BGP}(\text{ROR}(b, j)) | j = 0, \ldots, J - 1\},$$  

where $\text{ROR}(b, j)$ is the circular bitwise right shift $j$ times on the $J$-bit number $b$.

By considering $J = 8$ (or 8-bit values), there are only 36 unique combinations of maximum $\text{RIBGP}$ values to be calculated, which will be the number of bins of a $\text{RIBGP}$ values histogram that is built to describe images textures. This approach considers $\gamma = 1.82$, three combinations of $\sigma$ and $\lambda$, and eight combinations of $\theta$ to perform texture classification. These parameters generate eight even-symmetric and eight odd-symmetric Gabor filters. Finally, a set of histograms are
Built for each resolution. Three combinations of $\sigma$ and $\lambda$ are considered, yielding three resolutions per filter symmetry. So, this leads to a $(3 + 3) \times 36 = 216$ dimensional feature vector.

In the proposed approach, we aim at analyzing the effect of the spatial frequency artifacts at multiple scales. Our descriptor, called the Multiscale Rotation Invariant Binary Gabor Patterns MRIBGP, applies the histogram of RIBGP values at different levels of the Gaussian Pyramidal decomposed image. In our approach, illustrated in Figure 3, the 2D barcode image firstly undergoes histogram equalization in order to mitigate the effects of different illumination conditions. Then, a Gaussian Pyramidal Decomposition is performed by downscaling and upsampling the barcode image and the 216-dimensional features are calculated for each scale. After a three-scale decomposition, a 648-dimensional feature vector is extracted and will describe both original and counterfeited barcodes.

**B. Local Feature Set: the Bag of Multiscale Residual Words (BOMRW)**

First devised for document classification applications, the Bag of Visual Words (BOVW) is a computer vision technique that describes image contents by occurrence counts in a dictionary of local image features. The BOVW approach was firstly presented in [3] and it works by detecting points of interest such as SIFT [8] in training images, which are then clustered by unsupervised classification algorithms to build a dictionary. The size of the visual vocabulary $VOC = \{v_1, v_2, v_3, ..., v_k\}$ is $k$, where $k$ is the number of centroids in the clustering process. A given image can then be represented by a set of descriptors defined as

$$I = \{d_1, d_2, d_3, ..., d_N\},$$

where $N$ is the total number of descriptors detected for a given image $I$. Each descriptor $d_i$, detected as an $N$-dimensional vector keypoint, is then mapped to a visual word $v_i$. This is done by finding the minimum Euclidean distance between $d_i$ and each $v_i$. That is

$$v(d_i) = \arg \min_{v \in VOC} \text{Dist}(v, d_i),$$

where $\text{Dist}(\cdot)$ evaluates the Euclidean distances between the input vectors.

For our problem, a 2D barcode can be acquired under various conditions with different rendering sizes, acquisition noises, etc. To mitigate these uncontrolled factors, we propose to characterize the points of interest distribution after diverse low-pass filterings of the input image, as described in Figure 4. The processing pipeline of our proposed Bag of Multiscale Residual Words (BOMRW) works by first applying histogram equalization to the input 2D barcode image. Then, successive Gaussian filterings of varying filter sizes ($3 \times 3$, $5 \times 5$ and $7 \times 7$ pixels) are applied to remove image noise and yield a residual multi-channel image. Afterwards, keypoints are extracted from the residual image and are clustered through a $k$-means procedure, where $k$ is set as the square root of the number of training samples. Finally, the distances between the cluster centroids and the detected keypoints are calculated to find correspondences. This procedure results in feature vectors (histograms) of visual words for the training images, which can be used to train any machine learning classifier. In the testing stage, the process is repeated, with keypoints being detected and their corresponding visual words found in the pre-trained $k$-means clusters. Then, a histogram of visual words is calculated and used as input to the machine learning classifier, which will be used for the BOMRW based barcode authentication. In our proposed approach, we consider the
using the BRISK descriptor [7] to extract keypoints, as it showed better robustness for this specific task than SIFT and SURF.

C. The Merged Feature Set: the Local and Global Multiscale Feature Set (LGMSF)

After both features are calculated, a final feature vector \( f^i \) is created by concatenating feature histograms from MRIBGP and BOMRW descriptors in a \((216 \times S) + \sqrt{|T_k|}\) dimensional vector, where \( S \) is the number of scales used in the MRIBGP descriptor and \( T_k \) is the number of keypoints detected in the training images by BOMRW. As a next step, we apply L1 normalization to these histograms. We chose L1 normalization as it minimizes the sum of the absolute differences between the target value and its mean. This procedure is of fundamental importance, as it reduces the differences of the range of features resulting from different descriptors. Specifically, the normalized vectors \( z_i \) can be written as

\[
z_i = \frac{f^i}{\sum_{j=1}^{D} |f_j|},
\]

where \( D = (216 \times S) + \sqrt{|T_k|} \) denotes the feature dimension.

This way, the proposed Local and Global Multiscale Feature Set (LGMSF) merges both the local and global features into one histogram. The resulting histogram will then contain richer information about the genuine and counterfeited barcodes, with descriptive information from edges and keypoints. Such features are then fed to a Support Vector Machines classifier with linear kernel. The LIBSVM library with five-fold cross-validation is adopted in our implementation [1] to find the best parameters for training the classifier. For reproducibility reasons, the source code of our proposed approach is available at GitHub

https://github.com/anselmoferreira/2d-barcode-authentication

III. EXPERIMENTS

In the first experiment, we compare the individual proposed methods against their original counterparts in order to understand how they evolve the original algorithms. For this experiment, we consider the Dataset I described in Table I created to generate genuine and counterfeited barcodes. In this dataset, 30 barcode images are taken under each devices combination, and this procedure leads to 300 genuine barcode images. To create the counterfeited samples, the printed genuine barcodes are first scanned at 600, 1200 and 2400 Points Per Inch (PPI) resolutions, respectively, and then are printed with the same paper of the genuine barcodes. Given a large amount of counterfeiting devices combinations, only some representative ones have been used to generate the counterfeited barcodes. This way, Dataset I includes 3775 counterfeited barcode images, totaling 3775 + 300 = 4075 images.

Figure 5 plots the clustered features of the proposed MRIBGP versus the original RIBGP, using the t-SNE visualization tool [9] after both descriptors are applied to dataset I. On the one hand, it can be seen that the RIBGP descriptor with only one scale leads to more confusion in the feature clustering. This is because the artifacts and halftones with different sizes have not been considered in the feature description. On the other hand, the advantage of the proposed multiscale approach is achieved by employing pyramidal decomposition in the input images, which transforms the halftones into multiple sizes and minimizes this way noisy effects present in both genuine and counterfeited captured samples.

Fig. 5. Clustered features from the proposed Multiscale Rotation Invariant Binary Gabor Patterns (left) versus those from the original Rotation Invariant Binary Gabor Patterns [14] (right). It can be seen that the proposed features cause less confusion, as the clusters from genuine barcodes are closer to each other and are farther from the counterfeited barcode clusters. This behavior shows that samples described by the proposed features are better separable by a classifier.

Table I. Devices used to generate 2.5 \( \times \) 2.5 cm\(^2\) area barcodes.

| Type       | Brand   | Model     | Resolution |
|------------|---------|-----------|------------|
| Smartphone | Google  | Nexus4    | 8MP        |
| Smartphone | Apple   | iPhone 8 Plus | 12MP     |
| Smartphone | Huawei  | P9        | 12MP       |
| Smartphone | Meizu   | Metal     | 13MP       |
| Smartphone | Nokia   | Lumia 930 | 20MP       |
| Printer    | HP      | Laserjet Pro M401dn | 600 and 1200dpi |
| Printer    | Toshiba | E-studio 2307 | 600dpi    |
| Scanner    | Epson   | Perfection V330 | 600, 1200 and 2400dpi |
| Scanner    | Microtek| D328K     | 600, 1200 and 2400dpi |
| Scanner    | BenQ    | K810      | 600, 1200 and 2400dpi |

1https://github.com/anselmoferreira/2d-barcode-authentication
The multi-scale transform and illumination correction from the proposed approach generate some clean samples to the descriptor, which leads to a better description of the counterfeited samples. Therefore, the proposed MRIBGP descriptor improves all the metrics of the original RIBGP approach [14]. Specifically, the genuine barcodes detection performance has been improved considerably, which is highlighted by the lowest mean false counterfeited detection (FPR) of this experiment. This supports our hypothesis that the rotation and multi-scale invariant descriptors can be crucial for the problem of multilevel 2D barcodes authentication, as the patterns are irregular due to the use of different devices and acquisition conditions. Similarly, the proposed bag of visual words approach (BOMRW) has the second best result of this experiment, being thus a better option for this problem than using the common approach with the SIFT descriptor (BOVW-SIFT), as the proposed residual image highlights better the distortions of irregular edges in counterfeited images. Finally, the effect of combining both descriptors in a final descriptor based on edges and points of interest in the proposed LGMFS leads to a better performance, especially because both fused descriptors take into the account varying halftones and edge behaviors, but considering different image structural features (e.g., pixels and keypoints). The LGMFS descriptor shows the best f-score of 0.97, 96.48% NACC and the best TPR of 98.58%. Therefore, our approach, even taking into account less data, showed a significantly better result than data-hungry approaches like CNNs.

We finish the validation experiments with an even more difficult scenario. In this experiment, the robustness of the proposed scheme under different printing area barcodes is investigated. To do that, we test the approaches on a 5 × 5 cm² area barcodes Dataset III described in Table IV. This new dataset contains 180 images, including 60 genuine and 120 counterfeited barcodes. Table V shows the performance of descriptors considered in this experiment.

From results shown in Table III, we find that starting discussing the CNNs performance in this more difficult scenario. These approaches showed better results only when trained from scratch and yield only reasonable authentication results, as their performance is highly dependent on the uniformity of training and testing sets and also the number of training data, which can be limited due to high costs of building such datasets. Such limitation results in an unacceptable false negative rate (i.e. percentage of undetected counterfeited samples) of almost 10% in all CNNs evaluated in this experiment.

Results in Table III also highlight the improvement of the proposed MRIBGP over the original RIBGP descriptor [14].
which happens in both experiments reported in this paper.

features complements the other to describe unstable patterns, such descriptors is particularly important when one set of

scenario. Finally, the combination of the proposed descriptors

BOVW-SIFT

its original counterpart

BOMRW

perfect NACC, improves the original

RIBGP

property of the proposed

MRIBGP

considering the NACC. Notwithstanding, the scale invariance

the best individual algorithm performance in 2.50% when

considered the NACC. Notwithstanding, the scale invariance

property of the proposed MRIBGP, which achieves an almost

perfect NACC, improves the original RIBGP descriptor by a

large margin. The proposed BOMRW descriptor also improves

its original counterpart BOVW-SIFT \[3\] in such a challenging

scenario. Finally, the combination of the proposed descriptors

(LGMFS) achieves 98.33% NACC and 0.99 f-measure. Fusing

such descriptors is particularly important when one set of

features complements the other to describe unstable patterns, which happens in both experiments reported in this paper.

IV. CONCLUSION

First, the experimental results show that shallower and simpler CNNs, such as the RESNET with 50 layers, do not

perform well in the cross-size evaluation although the training

and validation processes have been carried out properly, as

can be seen in Figure \[4\]. This can be explained by the fact

that the testing dataset has different halftone dots distributions

than the ones found in the training set, making it difficult for

such CNN to generalize to such a new dataset. We identified

that original barcode images with larger printing areas can be
easily classified as forgeries by such CNN when its

trained with patches from smaller size barcodes. A reason for

such a behavior is the fact that more halftone dots are found

in the barcode with a larger printing area, and these dots are

confused with replication artifacts (e.g., extra edges caused by
distortion) in the counterfeited barcodes from Dataset I (with

smaller barcode areas), which were used to train the CNNs.

However, DENSENNET with 121 layers \[6\] can handle that

problem as it consists of a large number of layers and dense

modules that better generalize to such a new testing dataset.

With these features in such a CNN, complex interdependencies

between different higher level features can be abstracted.

Finally, Table \[III\] also highlights the benefits of joining the

local and global descriptors also when considering the original

algorithms. Fusing the RIBGP and BOVW-SIFT improved

the best individual algorithm performance in 2.50% when

considering the NACC. Notwithstanding, the scale invariance

property of the proposed MRIBGP, which achieves an almost

perfect NACC, improves the original RIBGP descriptor by a

large margin. The proposed BOMRW descriptor also improves

its original counterpart BOVW-SIFT \[3\] in such a challenging

scenario. Finally, the combination of the proposed descriptors

(LGMFS) achieves 98.33% NACC and 0.99 f-measure. Fusing

such descriptors is particularly important when one set of

features complements the other to describe unstable patterns, which happens in both experiments reported in this paper.

REFERENCES

[1] C.-C. Chang and C.-J. Lin. LIBSVM: a Library for Support Vector

Machines. ACM Transactions on Intelligent Systems and Technology

(TIST), 2(3):27, 2011.

[2] C. Chen, M. Li, A. Ferreira, J. Huang, and R. Cai. A copy-proof

scheme based on the spectral and spatial barcoding channel models.

IEEE Transactions on Information Forensics and Security, 15:1056–

1071, 2020.

[3] G. Csurka, C. R. Dance, L. Fan, J. Willamowski, and C. Bray. Visual

categorization with bags of keypoints. In European Conference on

Computer Vision, pages 1–22, 2004.

[4] Frontier Economics. The economic impacts of counterfeiting and piracy.

Technical report, Frontier Economics, 2016.

[5] K. He, X. Zhang, S. Ren, and J. Sun. Deep Residual Learning for Image

Recognition. In Proceedings of the IEEE Conference on Computer

Vision and Pattern Recognition, pages 770–778, 2016.

[6] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. Densely

connected convolutional networks. In IEEE Conference on Computer

Vision and Pattern Recognition (CVPR), pages 2261–2269, 2017.

[7] S. Leutenegger, M. Chli, and Y. Siegwart. Brisk: Binary robust invariant

scalable keypoints. In IEEE International Conference on Computer

Vision, pages 2548–2555, 2011.

[8] D. Lowe. Distinctive image features from scale-invariant keypoints.

International Journal of Computer Vision, vol. 60(2):91–110, 2004.

[9] L. v. d. Maaten and G. Hinton. Visualizing data using t-sne. Journal of

machine learning research, 9(Nov):2579–2605, 2008.

[10] S. Ono, T. Maehara, and K. Minami. Coevolutionary design of a

watermark embedding scheme and an extraction algorithm for detecting

replicated two-dimensional barcodes. Applied Soft Computing, 46:991 –

1007, 2016.

[11] I. Tkachenko, W. Puech, C. Desruel, O. Strauss, J.-M. Gaudin, and

C. Guichard. Two-level QR code for Private Message Sharing and

Document Authentication. IEEE Transactions on Information Forensics

and Security, 11(3):571–583, 2016.

[12] S. Voloshynovskiy, T. Holotyak, and P. Bas. Physical Object Au-

thentication: Detection-theoretic Comparison of Natural and Artificial

Randomness. In IEEE International Conference on Acoustics, Speech

and Signal Processing (ICASSP), pages 2029–2033. IEEE, 2016.

[13] L. Zhang, C. Chen, and W. H. Mow. Accurate Modeling and Efficient

Estimation of the Print-Capture Channel With Application in Barcoding.

IEEE Transactions on Image Processing, 28(1):464–478, 2019.
[14] L. Zhang, Z. Zhou, and H. Li. Binary gabor pattern: An efficient and robust descriptor for texture classification. In *IEEE International Conference on Image Processing*, pages 81–84, Sept 2012.